Compressed Video Streaming in Cooperative Wireless Networks with Interfering Transmissions

Antonios Argyriou, *Member, IEEE*

***Abstract*—In this paper, we present a video streaming system for wireless networks that employs utility optimization of pre-compressed video at the application layer, together with a novel cooperative wireless physical layer (PHY) that allows interfering transmissions. Our system model considers multiple and independent unicast or multicast transmissions between network nodes while a number of them serve as relays. For this new PHY the average transmission rate that each sender-destination pair can achieve is estimated first. Next, we show that the utility optimization problem can be simplified due to the features of the proposed PHY. Subsequently, we devise a utility optimization algorithm that is executed independently at each sender and it derives the optimal video packets that should be sent in the wireless link given the calculated rate constraint. Simulation results demonstrate that for unicast video streaming scenarios and with the utility optimization framework, allowing packets to interfere is a better choice than employing a typical cooperative PHY. For the case of multicast video delivery the performance of the proposed scheme is significantly improved for good channel conditions while the improvement is minimized in the low channel SNR regime when compared to the unicast scenario.**

***Index Terms*—Video streaming, wireless networks, interference, cooperative systems, physical layer network coding.**

I. Introduction

VIDEO transmission through streaming in wireless networks is still one of the most challenging problems of information transmission [1]. Contrary to elastic video distribution mechanisms (e.g. video download) that have seen huge benefits from the increases in network capacity, video streaming in wireless communication systems still suffers from several problems: Channel errors are difficult to be corrected in real-time in ways that are not observable by the user (e.g. with error concealment [1]). Channel coding mechanisms with forward error correction (FEC) or automatic repeat request (ARQ) introduce significant overheads that cannot solve entirely the problem of error resiliency due to the dynamic nature of the wireless channels. Furthermore, asymmetry of multi-hop wireless networks exacerbates the problem of video transmission through such a path [2]. Bandwidth fluctuations in the wireless case create very frequently the well known video stuttering problems requiring thus bigger playback buffers and larger playback delay.

Another important parameter to consider is the inherent characteristic of the video data. Video streams are characterized by different rate-distortion (RD) characteristics which means that data units do not have the same importance for the video decoding process [3]. Therefore, existing solutions have to adapt not only to the channel conditions but also to the transmitted video stream. Naturally, the RD tradeoff is part of source coding and transmission system. In imperfect channels joint source-channel coding should be exercised [4], [5]. However, when the source description is already coded, the options regarding the real-time actions that can occur during transmission are limited [6].

There is a need to radically address these problems and we believe that this can be accomplished when the technical progress at the lower layers of the protocol stack are carefully taken into account. One of the most promising approaches for achieving considerably better throughput and lowering delay at the lower layers of the protocol stack, is to allow packet transmissions to interfere for the advantage of all the involved users. Interference exploitation has recently attracted considerable attention from the wireless communication and information theory communities so as to increase the network capacity [7], [8], [9]. Physical layer network coding (PLNC) is another term frequently used for this technique. Allowing packet transmissions to interfere means that wireless signals are transmitted concurrently, reducing thus the total duration of the transmissions. However, a subset of the interfering signals must be known at the final destinations (a-priori information) in order to be able to decode the desired packet. By overhearing signals and with the help of a relay more transmissions can occur per time unit [10].

However, for the particular class of packet-based video communication systems the idea of allowing packets to interfere might not be always beneficial. The reason is that with video communications when a certain packet is transmitted the importance of this packet might be completely different from subsequent or previous packets as we explained earlier. A failed attempt to remove interference and decode a PHY packet might result in the inability to deliver an important video packet to the application. This failure might not be a problem for elastic data transfer in wireless networks since the performance is averaged over several channel realizations and every packet erasure is equally important. Therefore, in the case of video payload the problem that has to be solved is to identify the conditions for allowing specific packet transmissions to interfere given a certain RD characterization

Manuscript received November 19, 2011; revised May 22, 2012; accepted June 28, 2012. The associate editor coordinating the review of this paper and approving it for publication was B. Sikdar.

The material in this paper was presented in part at the Packet Video Workshop, Munich, Germany, May 2012.

Antonios Argyriou is with The Department of Computer and Communications Engineering, University of Thessaly, Volos, 38221, Greece.

Antonios Argyriou would like to acknowledge the support from the European Commission through the Marie Curie Intra-European Fellowship WINIE-273041 and the STREP project CONECT (FP7ICT257616).

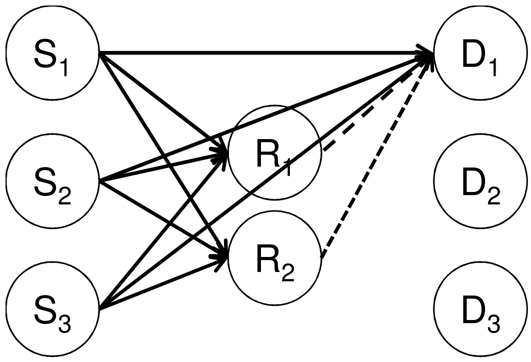


Fig. 1. The basic network scenario considered in this paper is cooperative unicast video transmission with the help of relay nodes R_1 and R_2. Different line styles indicate transmission in a different time slot. In the case of multicast transmissions all the destinations are decoding all the video streams.

of the transmitted bitstream and playback delay requirement of the streaming application. For the multicast case we expect that the optimal conditions for interfering transmissions change.

The rest of this paper is organized as follows. First, in Section II we present the related works. The system model that is used in this paper is presented in Section III. Subsequently, the cooperative PHY that incorporates interfering transmissions is analyzed in Section IV, while the impact of the application layer FEC on the packet error rate and the bandwidth overhead are evaluated in Section V. Section VI presents the utility optimization framework for video transmission and explains how it is coupled with the previously described PHY. In Section VII we present comprehensive simulation results for the proposed system configurations. Finally, Section VIII presents our conclusions and provides some possible directions for future work.

II. Related Works

Our work brings together techniques from RD-optimized video transmission and the novel idea of cooperation through interfering transmissions. We review the most closely related works to this paper next. The work presented by Katti *et al.* in [8] introduced the term analog network coding (ANC) which is another term for referring to concurrent interfering transmissions or PLNC. In that work the receiver decodes independently the overheard and relayed version of the superimposed signals leading to higher number of packet failures. The concept of PLNC in topologies that use one relay was also thoroughly studied from other researchers in [11], [12], [10] with similar results. Despite the flurry of activities in PLNC-based schemes, we are not aware any work that considers the exploitation of interfering transmissions for improving wireless video streaming.

On the contrary, wireless cooperative transmission of video streams has been studied in the literature more extensively. The case of layered encoded video was studied in [13], [14]. Video multicast with a more advanced cooperative scheme was considered in [15]. Distributed space-time codes were employed in that work in order to improve the decoding of the PHY when multiple receivers are involved. More recently, packet-based (or digital) network coding was also studied as a technique that can improve video transmission. The case of algebraic network coding and video transmission was studied in [16]. In that work the authors employ linear network codes for mixing video packets before transmission to a group of senders. In [17] the authors presented a rate-distortion optimized and network-coding-based, cooperative peer-to-peer packet repair solution for the multi-stream WWAN video broadcast. Even though in the above work the RD optimization is the objective, a more complex packet-level algebraic network coding solution was employed that means no exploitation of concurrent interfering transmissions. Also in [18] the authors studied the use of broadcasting from multiple stations for wireless video transmission but without a systematic cooperative protocol for allowing and exploiting interfering transmissions.

III. System Model and Overview

In this paper, we consider a wireless ad hoc network where all nodes can send data and also be potential relays. In this network structure all nodes can hear each other. The proposed cooperative PHY protocol optimizes the cooperative transmission for a single hop, i.e. within the defined network. Fig. 1 presents a representative network that is used for explaining several aspects of this work. Here we introduce the term *communication phase* as the basic time frame for the protocol analysis in this paper. The communication phase consists of a number of time slots. The proposed PHY works as follows. During the first slot of the communication phase a number of N network nodes broadcast their packets independently and concurrently. During the next $N-1$ slots of the communication phase the broadcasted and interfered packets are amplified and forwarded at the PHY by $N-1$ relays sequentially. The relay transmission order can be random but it should be decided in advance of a communication phase. Note that we focus on the system design when a group of N nodes has been decided that are the ones who are transmitting video and $N-1$ nodes that are relays. In [19] we investigated the topic of relay selection for such a scheme where interfering transmissions are allowed from two nodes. In this work we do not delve in such issues since relay selection, although it affects the performance of any cooperative scheme, is addressed at the MAC layer with special protocols.

Regarding the lower layer aspects of our system, we assume full overlap between the transmitted packets. This is possible since transmissions use time division multiple access (TDMA) with pre-defined slot boundaries. At the receivers the packets at the PHY are decoded, after the final relay forwards the respective interfered packet, with an MMSE-based joint symbol-level decoder that we describe in the next section. In case of unicast transmission the result is that destination D_1 in Fig. 1 recovers and stores only the packet originated from S_1. However, if all nodes are multicasting to all the three destinations in Fig. 1 then all the decoded packets at the PHY are useful and are stored by each one of the destinations. The optimality of interfering and decoding two packets for maximum throughput was studied and established in [8], [10], [20]. Therefore, the result of these works was that the cooperative system should allow the interference of two and not more packets for achieving throughput optimality.

An overview of the proposed video transmission system functionality is presented next. We consider the transmission of H.264 pre-compressed video where each video unit corresponds to a single I, P, or B frame [4]. A more efficient scalable video coded (SVC) bitstream could also be used with our scheme in order to provide more graceful quality variations. Nevertheless, the proposed scheme is orthogonal to the video coding standard. If the video description is available, the first task of the video transmission system is to pass the video units through an application-layer FEC encoder in order to create the packet to be actually transmitted. In the proposed system independent senders follow the previous process, and broadcast their respective packets regardless if they have only one destination (unicast) or multiple destinations (multicast). These packets interfere both at the potential relays and also at the intended destinations.

Given the above system model and general system design, the questions we try to answer in this paper are the following: 1) What is the optimal number of interfering senders when the video quality is the optimization objective instead of throughput? 2) Is it possible to maximize the utility of the transmitted video sequences in a decentralized fashion given that interfering transmissions are taking place? 3) Finally, are there any benefits of the cooperative interference scheme for video multicast?

IV. Physical Layer with Cooperative Collisions

In this section we describe the physical layer (PHY) that supports cooperative interfering transmissions. We denote with N the number of senders where each one of them transmits a video stream. The senders transmit by using pre-defined TDMA slots. The transmitted signals collide at M relay nodes. Fig. 1 presents a sample network topology with three senders and two relays. The relay nodes forward the interfered packets. Due to the broadcast nature of the wireless channel the concurrently transmitted packets will be available to every node that does not transmit but it overhears the channel. This basic observation is exploited by the protocol named *amplify and forward of over-the-air superimposed transmissions* (AFOST). For this protocol we present an example based on the topology depicted in Fig. 1. Fig. 2 presents the corresponding protocol behavior in the time domain. In this example there is one broadcast phase from all the N=3 senders and two forwarding phases from all the M=2 relays. Because of spatial diversity, different versions of the broadcasted signals are received at different network nodes including the relays. During the forwarding phase, each participating relay broadcasts the locally received interfered signals after it applies the appropriate power scaling. With the proposed scheme it must be $M \geq N-1$. The reason is that if a number of nodes transmit concurrently, there is a need for at least the same number of forwarding phases so that N linear equations are collected and the PHY decoder can then solve for the N unknown and concurrently transmitted symbols [21]. Note that the clear benefit when compared with the cooperative protocol that orthogonalizes transmissions is that the relays forward a composite signal that is useful for more than one receivers. Another observation is that these $N-1$ forwarding phases do not have to be executed from different relays even though in our analysis we make this assumption. It is possible that one relay repeatedly retransmits the same signal over independent channel realizations exploiting thus diversity over time [21]. These aspects will become clearer when we present the protocol operation with a mathematical notation next.

A. Description of AFOST

Let x_n denote the PHY symbol that is transmitted from sender S_n. The number of PHY symbols L_n denotes all the symbols of the packet that were transmitted during the broadcast phase. Let also the channel transfer function between the sender S_n and the relay R_m be h_{S_n,R_m}, and the channel between sender S_n and destination D_k be h_{S_n,D_k}. By using this notation, let us proceed with the detailed description of the transmitted and received signals for the protocol we sketched in the previous paragraph. We may write the received signals at a relay R_m as

$$y_{R_m} = \sum_{n=1}^{N} \sqrt{P} h_{S_n,R_m} x_n + w_{R_m}, \quad \forall m \in \mathcal{M}, \tag{1}$$

where P is the transmission power at each sender, and $w_{R_m} \sim \mathcal{CN}(0, \sigma)$ denotes the AWGN at the relay R_m. Similarly, the received signals at destination D_k during the broadcast phase can be written as

$$y_{D_k} = \sum_{n=1}^{N} \sqrt{P} h_{S_n,D_k} x_n + w_{D_k}. \tag{2}$$

For the one broadcast phase there are multiple forwarding phases, i.e. their precise number is $N-1$. In each of the forwarding phases a relay R_m broadcasts the received signals given in (1) by applying a power amplification factor g_m so as to maintain the power constraint [22]. The power gains are given as

$$g_m = \sqrt{\frac{P}{P\sum_{n=1}^{N} \gamma_{S_n,R_m} + \sigma^2}}, \tag{3}$$

where $\gamma_i = |h_i|^2$. Subsequently, the relays forward the amplified signals. In the forwarding phase from R_m, the received signal at D_k can now be written as

$$\begin{aligned} y_{D_k,R_m} &= g_m h_{R_m,D_k} \sum_{n=1}^{N} \sqrt{P} h_{S_n,R_m} x_n \\ &+ g_m h_{R_m,D_k} w_{R_m} + w_{D_k}. \end{aligned} \tag{4}$$

Note that (4) corresponds to the received signal from one relay. A point should also be made about the spectral efficiency of our protocol that is not compromised. The reason is that the number of concurrently transmitted packets is N (equal to the number of senders) while this is accomplished in a total of N transmission phases (1 direct plus $N-1$ forwarding). This translates to a spectral efficiency of 1 regardless of the number of relays.

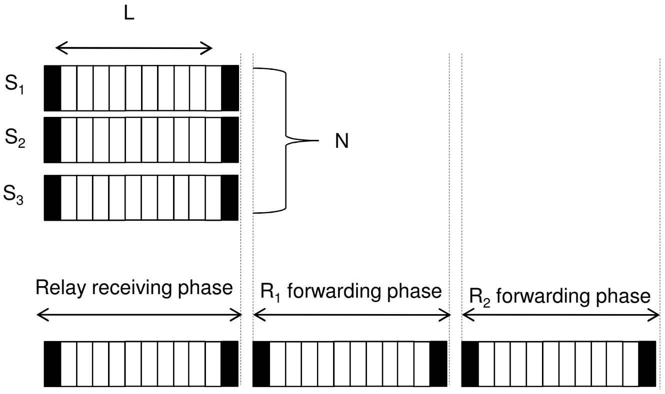


Fig. 2. The AFOST cooperative protocol for the simple topology with three senders. The dark-shaded blocks indicate the symbols that belong to the preambles and postambles of a packet and are used for channel estimation, while the non-shaded blocks are the symbols of the information packet.

Based on the above analysis, we write in vector form the received signal for destination D_k as follows

$$\mathbf{y}_{D_k} = \sqrt{P} \cdot \mathbf{G} \cdot \mathbf{H}_{D_k} \cdot \mathbf{x} + \mathbf{w}_{D_k}. \tag{5}$$

The $N \times N$ channel matrix for the cooperative system we introduced is:

$$\mathbf{H}_{D_k} = \begin{bmatrix} h_{S_1,D_k} & .. & h_{S_N,D_k} \\ h_{S_1,R_1}h_{R_1,D_k} & ... & h_{S_N,R_1}h_{R_1,D_k} \\ ... & ... & ... \\ h_{S_1,R_M}h_{R_M,D_k} & ... & h_{S_N,R_M}h_{R_M,D_k} \end{bmatrix}$$

The array $\mathbf{G}$ corresponds to the power gains of all the relays and $\mathbf{w}_{D_k}$ is the noise vector that includes the broadcast and each forwarding phase. All these arrays consist of complex numbers and are given below:

$$\begin{aligned} \mathbf{w}_{D_k} &= \begin{bmatrix} w_{D_k} & g_1 h_{R_1,D_k} w_{R_1} & ... & g_M h_{R_M,D_k} w_{R_M} \end{bmatrix}^T \\ \mathbf{x} &= \begin{bmatrix} x_1 & x_2 & ... & x_N \end{bmatrix}^T \\ \mathbf{y}_{D_k} &= \begin{bmatrix} y_{D_k} & y_{D_k,R_1} & ... & y_{D_k,R_M} \end{bmatrix}^T \end{aligned}$$

Depending on how may senders transmit concurrently the dimension of the matrices change accordingly.

B. PHY Decoding Algorithm

We now describe the PHY layer detection algorithm executed at each destination node. For a multi-user system the optimal detector is a minimum mean square error with successive interference cancellation (MMSE-SIC) receiver [21]. If the Hermitian of $\mathbf{H}$ is $\mathbf{H}^H$, then the pseudo-inverse channel matrix $\mathbf{H}^\dagger = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H$ is used as follows: The MMSE approach tries to find a coefficient matrix $\mathbf{Q}$ that minimizes the MMSE criterion. We have that $\mathbf{Q}^\dagger = (\mathbf{H}^H\mathbf{H} + \sigma^2\mathbf{I})^{-1}\mathbf{H}^H$. The $n-$th bit stream transmitted to destination node D_k is extracted with the help of the pseudo-inverse channel matrix $\mathbf{Q}^\dagger$ as follows. The signal is multiplied by the (estimated $\widetilde{\mathbf{Q}}^\dagger$) pseudo-inverse:

$$\hat{\mathbf{y}}_{D_k} = \widetilde{\mathbf{Q}}^\dagger_{D_k,n}\mathbf{y}_{D_k} = \widetilde{\mathbf{Q}}^\dagger_{D_k,n}\mathbf{H}_{D_k,n}\mathbf{x}_k + \widetilde{\mathbf{Q}}^\dagger_{D_k,n}\mathbf{w}_k \tag{6}$$

where $\widetilde{\mathbf{Q}}^\dagger_{D_k,n}$ indicates *all the rows in the pseudo-inverse channel matrix $\widetilde{\mathbf{Q}}^\dagger_{D_k}$ minus the n-th row*, while $\mathbf{H}_{D_k,n}$ symbolizes the $n-$th column for $\mathbf{H}_{D_k}$. Due to the whitening operation the variance of the noise matrix is $E[\widetilde{\mathbf{Q}}^\dagger_{D_k,n}\mathbf{w}_k|\mathbf{H}_{D_k,n}] = \Big(\sum_{m=1}^{M} g_m^2 |h_{R_m,D_k}|^2 + 1\Big) I_N$. In addition, in our case we apply an MMSE-SIC receiver where the power of the received signals in $\hat{\mathbf{y}}_{D_k}$ is ordered from higher to lower power. If we denote by $\bar{\mathbf{y}}$ the ordered version of the received signals contained in $\hat{\mathbf{y}}_{D_k}$ from higher to lower power[1], then we can apply the ordered SIC (OSIC) approach for detecting first the symbols that were received with the higher power. The destination uses MMSE equalization and estimates the higher power symbol (first in array $\bar{\mathbf{y}}$) x_l as

$$\hat{\mathbf{x}}_{l,D_k} = \widetilde{\mathbf{Q}}^\dagger_{D_k,1}\bar{\mathbf{y}}_{D_k,1} \tag{7}$$

It is important to note that $\hat{\mathbf{x}}_{l,D_k}$ indicates the estimate of symbols from a sender S_n but at node D_k.

C. Sum Rate

For the MMSE/OSIC receiver that we adopted only the ergodic capacity or the average achievable rate can be practically calculated. Essentially it is the rate after averaging over a significant number of channel realizations. For the link $S_n \rightarrow D_k$ we have that the average achievable rate will be:

$$\widetilde{R}^{afost}_{D_k,S_n} = \mathbb{E}\Big[\log_2(1+\frac{P}{\sum_{m=1}^{M} g_m^2|h_{R_m,D_k}|^2+1}\widetilde{\mathbf{Q}}^\dagger_{D_k,n}\mathbf{H}_{D_k,n})\Big] \tag{8}$$

This rate estimate is communicated from node D_k to S_n after it is estimated locally, since the destination node has the channel state information (CSI). Note that since the network is distributed and formed in an ad-hoc fashion, the transmission power P is not amenable to an optimization step since that would require a central control point that would perform this task.

D. Baseline Cooperative System

Here, we present a brief description of a typical cooperative PHY that employs amplify-and-forward (AF) [22], [23] that serves as the baseline system for comparison and we refer to it as COOP. In the general case of cooperative systems, the transmitter may select to use cooperative transmission when a desired rate is not met with a direct transmission. However, without loosing generality we assume that with this cooperative protocol the optimal mode is always selected whether it is cooperative or direct transmission. Now consider that the channel bandwidth is W, the transmitter power P, additive white Gaussian noise (AWGN) with zero mean and variance σ^2, and $\tilde{h}_i$ is the estimate for channel i. If we assume Rayleigh block fading channels where the attenuation is considered constant throughout the transmission of a single frame then the estimated rate of the Direct transmission mode is:

$$\widetilde{R}^{dir}_{D_k,S_n} = W \cdot log_2(1 + \frac{P|\tilde{h}_{S_n,D_k}|^2}{\sigma^2}). \tag{9}$$

[1]This ordering is easily accomplished through the channel estimation process.

On the other hand, the estimated rate of the cooperative transmission protocol COOP that occurs in two orthogonal time slots will be [22]:

$$\begin{aligned}\widetilde{R}^{coop}_{D_k,S_n} &= \frac{W}{2}\cdot min\{log_2(1+\frac{P|\tilde{h}_{S_n,R_m}|^2}{\sigma^2}), \\ & log_2(1+\frac{P|\tilde{h}_{S_n,D_k}|^2}{\sigma^2}\frac{P|\tilde{h}_{S_n,R_m}|^2|\tilde{h}_{R_m,D_k}|^2g^2}{\sigma^2(1+|\tilde{h}_{R_m,D_k}|^2g^2)})\}\end{aligned} \quad (10)$$

Similar conditions are used by state-of-art cooperative protocols [23].

V. Rate Estimation with FEC

With the proposed streaming system source packets are sent to the application-layer FEC Reed Solomon (RS) encoder. The RS encoder generates $I-J$ additional packets for J input source packets. Therefore, the overhead that is added with an $RS(I,J)$ code is $I-J$ packets or $J/I\%$. FEC is applied across the source packets so that each generated transport packet contains parts of both the source payload and the parity bits. This packetization strategy is crucial for efficient wireless transmission since most software implementations discard link-layer frames that have even single-bit errors that cannot be corrected. In such a case, applying FEC along each source packet would mean that in the case of a link-layer frame loss, the complete source packet would be lost, and RS decoding would thus fail. Subsequently, the I transport packets are sent to the PHY for wireless transmission. Link-layer retransmissions are not used in this work, since this approach would complicate the analytical models. Once the packets are transmitted at the physical layer, they pass through the application-layer RS decoder at the receiver. The probability of RS decoding failure at the receiver is given by

$$\rho = 1-\sum_{l=0}^{I-J}\binom{I}{l}p_w^l(1-p_w)^{I-l}, \quad (11)$$

where p_w is the wireless packet error rate (PER) before FEC error recovery. If RS decoding fails, the video unit cannot be decoded. The decoder starts to decode video packets and display them, after an initial *startup delay* that is configured by the application. Note here that the selected RS code is the same for all packets. Due to the additional complexity of an unequal error protection (UEP) scheme, we do not include it here although performance benefits are expected.

A. Effective Throughput

For calculating the effective throughput at the sender we utilize the rate estimates given in (8), (9), (10). We also assume that the transmission of acknowledgements on the reverse path is considered error-free, which is something that can be easily achieved by applying strong error correcting codes for the short ACK messages. Therefore, if the video payload consists of $data$ bytes, and the combined protocol overheads is hdr bytes, then the effective throughput is given by:

$$T = \frac{data}{data+hdr} * \frac{J}{I} * \widetilde{R} * (1-\rho) \quad (12)$$

In our system, we are more concerned with the raw application-layer data rate that can be achieved between a pair of nodes. To obtain this quantity, packet losses in (12) have to be ignored, which makes the maximum possible application data rate:

$$T_{max} = \frac{d}{d+hdr} * \frac{J}{I} * \widetilde{R} \quad (13)$$

Actually T_{max} corresponds to the effective throughput for a channel SNR $\to \infty$, where packet errors are non existent.

VI. Utility Optimization

The task of a multimedia communication system is to maximize the reconstruction quality of the media presentation at the receiver for the given channel conditions. One important detail is that video quality is usually measured through the end-to-end degradation/distortion of the reconstructed signal. Furthermore, distortion is related to the data rate allocated to the media presentation, a relationship that is captured through the RD function of the presentation. In this section we attempt to optimize video transmission with a utility-based framework that uses the proposed cooperative protocol that allows interference.

A. Utility Function

We formulate our optimization problem as a utility maximization. Different utility functions can be employed by the senders. In our case, the utility function is defined as the reduction of the reconstruction distortion of the video presentation, i.e.,

$$u(r_i) = \sum_i \Delta D(i) \quad \text{with} \quad \sum_i \Delta R(i) \leq T_{n,k}, \quad (14)$$

where the RD information associated with packet i consists of its size $\Delta R(i)$ in bytes and the importance of the packet for the overall reconstruction quality of the video presentation denoted as $\Delta D(i)$ [24], [25]. In practice, $\Delta D(i)$ is the total increase in the mean square error (MSE) distortion that will affect the video stream if the packet is not delivered to the client by its prescribed deadline [6]. It is important to note at this point that the value of the MSE distortion in $\Delta D(i)$ includes both the distortion that is added when packet i is lost and also the packets that have a decoding dependency with i (For example the ΔD for an I frame includes the ΔD of the P and B frames that depend on it). In this way the utility formulation considers also the possible drift that might occur due to the loss of particular packets/video frames. Now, in order to compute the utility $u(r_i)$ in (14) we previously label the video packets comprising the presentation in terms of importance using the procedure from [26]. Therefore, the index i in the summations in (14) enumerates the most important video packets in the presentation up to a data rate of r_i. In other words, $u(r_i)$ corresponds to the cumulative utility of the most important packets up to the rate point r_i.

However, the actual utility of the received packets must account for the lost video units. This should be done such that the overall utility $U_{n,k}(j)$ of the GOP j that belongs to the video stream that is transmitted over the wireless link

$S_n \to D_k$ is maximized. According to the above, the overall utility is defined as

$$U_{n,k}(j) = u[r_i(1-\rho)], \tag{15}$$

which is the utility of all the packets of GOP j up to rate point $r_i(1-\rho)$. This rate point corresponds to the packets that are not lost.

B. Optimization for COOP

After the rate $\widetilde{R}_{n,k}$ and the effective throughput $T_{n,k}$ at the physical layer have been calculated for the sender-receiver pair (n,k), its value needs to be used for the utility optimization step. Let us denote with $c_{n,l}$ the TDMA slot allocation vector that indicates that the n-th sender transmits in the l-th slot out of the N maximum. Then the utility optimization problem for the COOP protocol is defined as:

$$\max U_{n,k}(j) \quad \text{s.t.} \quad \left\{ \begin{array}{c} r_i \leq max(T_{n,k}^{dir}, T_{n,k}^{coop}) \\ \sum_{n=1}^{N}\sum_{l=1}^{N} c_{n,l} = N \\ c_{n,l} \in \{0,1\} \\ U_{n,k} \in u(r_i) \end{array} \right. \tag{16}$$

In the above, in the first constraint the best out of the direct or cooperative transmission modes is selected by COOP based on the rate estimate. The second and third constraints ensure that all the allocated slots to the N transmitting nodes is equal to their number. The last constraint means that the maximized utility should consist of a valid RD point that includes video packets up to packet i. This is necessary since there is a finite number of available rate points. The optimal solution to the above problem is out of the scope of this paper and heuristic solutions to it can be found in [27], [28] and in references therein. Naturally, due to the NP-completeness of the problem we resort here to a heuristic solution that does not require a centralized controller [28]. More specifically, we allow the nodes to share equally the slots after every communication phase, and then they select locally the optimal transmission mode which is either direct or cooperative transmission.

C. Optimization for AFOST

We use the formulation of the optimization problem given in (16) and we adapt it given that the AFOST protocol is used. First, when AFOST is used during every transmission slot, only the residual rate constraint $T_{n,k}^{afost}$ is needed. Second, every sender transmits in each of the N TDMA slots in a complete communication phase. Therefore, the second and third conditions can be eliminated from (16). Using the notation introduced previously we can write the simpler optimization problem as

$$\max U_{n,k}(j) \quad \text{s.t.} \quad \left\{ \begin{array}{c} r_i \leq T_{n,k}^{afost} \\ U_{n,k} \in u(r_i) \end{array} \right. \tag{17}$$

It is evident here the simpler problem formulation. The decision to couple the proposed PHY with the utility optimization pays off at three levels: First, it enables a simpler solution algorithm with a limited number of constraints, second it requires no central coordination for the slot allocation, and third it removes the non-linear constraint that dictates the maximum achieved rate based on the used PHY.

We proceed here by solving the optimization problem in (17). For this problem, we can apply Lagrange duality [29] to the first constraint in (17) to produce the following partial Lagrangian

$$L_n(\lambda_n, r_{n,k}) \quad = \quad U_{n,k} - \lambda_{n,k} \cdot (r_{n,k} - T_{n,k}), \tag{18}$$

where $\lambda_{n,k} > 0$ is the Lagrange multiplier for link $S_n \to D_k$. Similarly, $r_{n,k}$ is current instantaneous rate allocation. $L_n(\lambda_n, r_{n,k})$ represents the individual Lagrangian for link $S_n \to D_k$.

Now, (17) represents a concave optimization problem with linear constraints for the rate region as provided by the link rate constraint in (17). Then, each Lagrange multiplier expresses the price of each selected rate allocation for the outgoing link at node S_n. It is known that if λ_n^* is the optimal solution for the dual problem, then the corresponding $r^*(\lambda_{n,k}^*)$ is the solution to the primal problem defined in (17).

It can be shown that the following two equations represent a solution for the primal-dual optimization problems [29]. First, node S_n computes the optimal rate allocation on link $S_n \to D_k$ using

$$r_{n,k}^* = \arg\max_{r_{n,k}} \Big\{ U_{n,k} - \lambda_{n,k} r_{n,k} \Big\}. \tag{19}$$

Then, given $r_{n,k}^*$ we employ a sub-gradient method [30] to update the value of $\lambda_{n,k}$ as follows

$$\lambda_{n,k}(t+1) = \max\big\{0, \lambda_{n,k}(t) + \beta\Big(r_{n,k}^* - T_{n,k}\Big)\big\}. \tag{20}$$

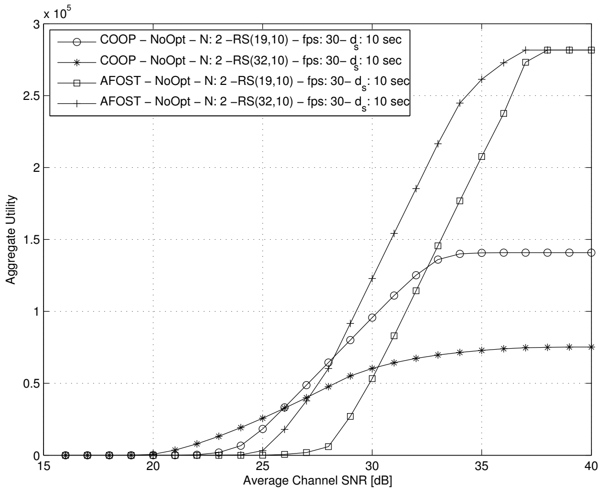
In the above equation β is a small constant that is appropriately selected. Sub-gradient adaptation methods such as (20) are typically used in optimization problems involving Lagrange relaxation. Lastly, (19) and (20) are consecutively applied every time node S_n performs rate allocation on its outgoing link. Thus, the rate allocation algorithm presented in this section calculates the Lagrangian multiplier for the transmitted bitstream at each sender separately.

D. Stream Adaptation

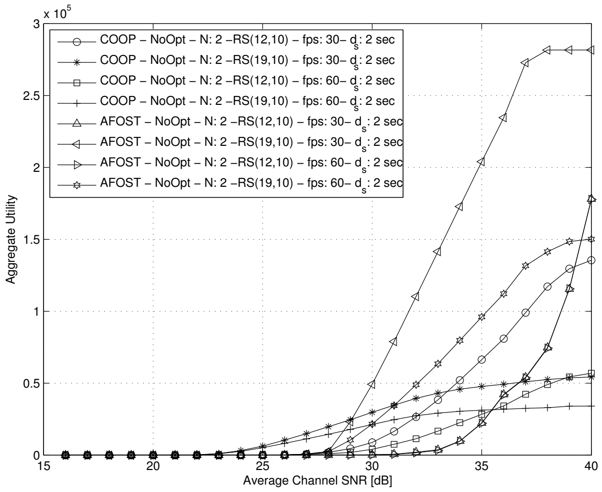
Now, as shown in [26] (17) can be efficiently solved using the rate-distortion characterization of the video packets comprising a flow. In particular, let $\Delta D(i_n)/\Delta R(i_n)$ be the utility gradient of packet i_n, i.e., packet i from flow n. Then, the practical solution of (17) comprises of filtering the video packets sent over the link $S_n \to D_k$ according to the following rule

$$\begin{array}{ll} \text{Send packet } i_n & : \ \Delta D(i_n)/\Delta R(i_n) > \lambda_{n,k}, \\ \text{Do not send packet } i_n & : \ \text{otherwise.} \end{array} \tag{21}$$

In essence, (21) allows the transmission over the wireless channel only the most important packets such that the overall utility of the video streams is maximized while at the same time the resulting transmission data rate does not exceed the maximum achieved by the underlying PHY.



(a) d_s=10 seconds and 30 fps



(b) d_s=2 seconds and 30/60 fps

Fig. 3. Average utility vs. the channel SNR for 2 concurrently transmitting senders and without enabling utility optimization.

VII. Performance Evaluation

In this section, a comprehensive evaluation of the proposed algorithms comprising our framework through simulations is presented. Both the PHY and video streaming algorithms were implemented in Matlab. The number of nodes tested is kept small since the simulator operates at the PHY symbol level (not packet-level) requiring thus significant amount of execution time. At the communication layer a typical cooperative system is implemented and it employs cooperative orthogonal amplify and forward at the PHY without collisions and it is named COOP. With the COOP scheme the receiver combines with maximum ratio combining (MRC) the directly received and forwarded signal to achieve lower BER [22]. The proposed PHY with cooperative collisions is named AFOST in the figures. Regarding the lower layer parameters we assume a channel bandwidth of W=20 MHz, while the same Rayleigh fading path loss model was used for all the channels. The assumptions in this case include a frequency-flat fading wireless link that remains invariant per transmitted PHY frame,

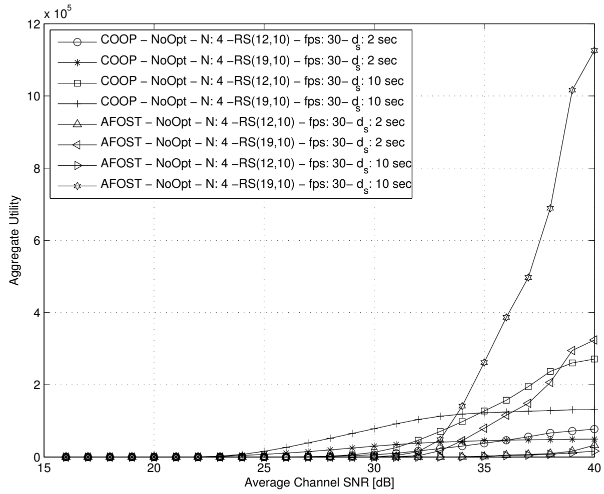


Fig. 4. Average utility vs. the channel SNR for 4 concurrently transmitting senders and without enabling utility optimization. d_s=2 and 10 seconds while the frame rate is 30 fps.

but may vary between simulated frames. The noise over the wireless spectrum is additive white Gaussian noise (AWGN) with the variance of the noise to be 10^{-9} at every node/link. For ensuring fairness for the comparison of the two PHYs, the average channel SNR was assumed to be the same for all the links but it varied independently during each channel realization.

For the video part of the simulation, the performance of two video streaming systems that use and do not use utility optimization is examined. They are named Opt and NoOpt respectively. The utility optimization was exercised for the duration of 10 GOPs. The video content consists of the CIF sequences MOTHER & DAUGHTER and FOREMAN that were compressed using an H.264 codec [31] at the rates of 203 kbps and 328 kbps, respectively. Both sequences were tested in independent experiments. A number of 300 frames of each sequence were encoded using the following frame-type pattern IBBBP..., i.e., there are three B frames between every two P frames. The GOP size was set to 32 frames. Also, the startup/playback delay d_s of the video presentation at every node is set according to the experiment. In all the figures, the results correspond to the average utility enjoyed by a destination for the duration of the sequence.

A. Results for Multiple Unicast Streams without Utility Optimization

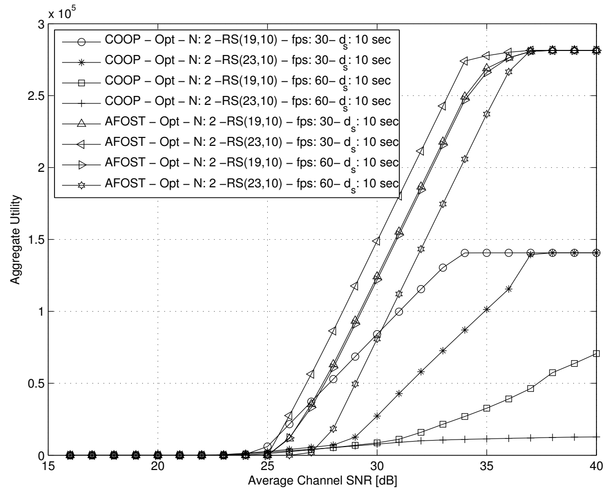
First, we present simulation results for the system where no utility optimization is applied, two senders concurrently transmit (N=2), while packets are transmitted according to the presentation order. Fig. 3(a,b) presents the related results for the COOP and AFOST systems. For a relatively high playback delay of d_s=10 sec in Fig. 3(a), we observe that AFOST performs considerably better than COOP. Different application layer FEC code rate is needed for each system in order to achieve the best performance while other tested code rates provide worse performance. In this case AFOST compensates with the higher code rate of $RS(32, 10)$ the slightly increased BER. If the playback delay is even higher then the performance of both systems would converge. For d_s=2 sec and frame rate of 30fps shown in Fig. 3(b), the performance of the COOP system with orthogonal cooperative transmissions is significantly inferior to AFOST. In the high SNR regime, where the BER is reduced for both systems, the only choice is to reduce the FEC rate but this is not enough since the data rate of the communication link must be high in order to compensate for the short d_s. Only the $RS(12, 10)$ code can provide some improvement for COOP. Therefore, AFOST can support more effectively this higher data rate and low-delay requirement of multiple video streaming senders. Also note that when the frame rate is increased to 60fps in Fig. 3(b), both systems suffer due to the increased need for bandwidth but AFOST still performs considerably better. Another interesting result is that for the poor channel conditions, the COOP system is still able to provide some meaningful aggregate utility contrary to AFOST. The reason is that this system can reduce the BER and it can provide at least some goodput to the application.

The results for four nodes (N=4) can be seen in Fig. 4. We can observe here that generally the COOP system behaves considerably worse than AFOST that achieves the best performance for a FEC code rate $RS(19, 10)$. The situation is deteriorated for both systems when the startup delay d_s is shorter by nearly an order of magnitude and equal to 2 sec. Still, the results are better for AFOST over COOP.

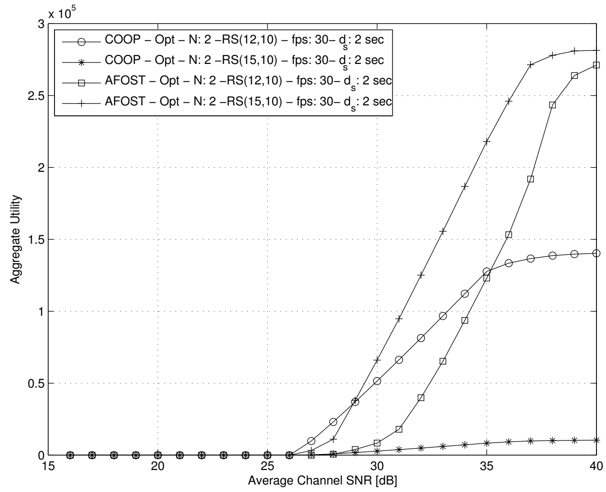
B. *Results for Multiple Unicast Streams and Utility Optimization*

Fig. 5 presents results for the case that the utility optimization framework we developed in Section VI was enabled. Again, two concurrent senders are tested. With the proposed utility optimization framework, the performance of the COOP transmission mode is good only for the case of a more relaxed startup delay requirement of d_s=10 and 30 fps as Fig. 5(a) indicates. AFOST is superior on all tested cases. In the case of d_s=2, the COOP system with $RS(12, 10)$ is better than AFOST for the same code rate, while the other FEC coding rate options under-perform significantly. But for a different FEC rate AFOST again outperforms COOP. From these first results, and even for two senders, two conclusions can already be made. First is that allocating the transmission rate and prioritizing important video packets improves considerably the video quality for AFOST, and second that utility optimization is crucial even if concurrent interfering transmissions with AFOST are not enabled and a standard orthogonal cooperative PHY is used.

Very interesting results are obtained for N=4 and are shown in Fig. 6(a,b). Although, COOP performs relatively good (but still inferior to AFOST) for a high d_s=10, for d_s=2 it is nearly impossible to compete. The reason is that the required transmission rate is very high for transmitting the most important and high utility packets on-time. The most important video units are larger in number of bits (I and P frames) and so they require more bandwidth. So achieving higher throughput at the PHY of the communication stack is more critical as more nodes share the medium. At the same time the aggregate utility is reaching significantly higher absolute values of nearly 8×10^5 in the high SNR regime. This behavior is attributed to the fact that the video units of highest importance are sent for all the four interfering senders. Furthermore, when we compare these results with the NoOpt case, we see that Opt outperforms NoOpt for the case of four interfering nodes. Therefore, the option to allow collisions to occur between more than two nodes makes sense when video transmission is jointly employed with a RD utility optimization framework. Alternatively, when the COOP transmission mode is selected, it is not so critical to employ Opt.



(a) d_s=10 seconds and 30/60 fps

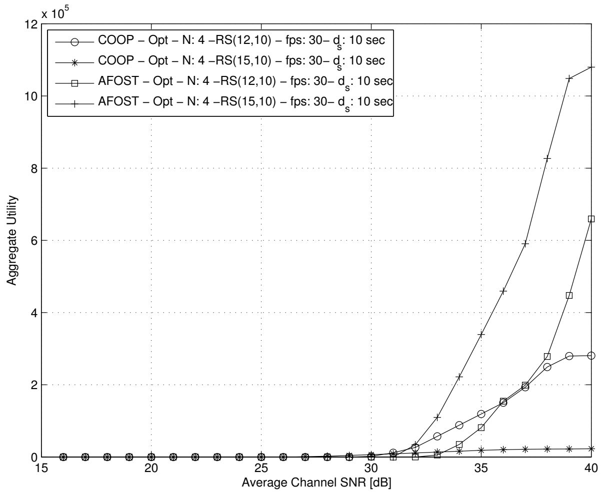


(b) d_s=2 seconds and 30 fps

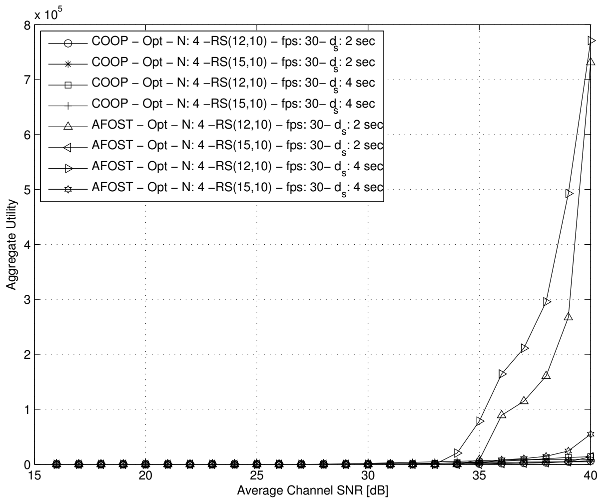
Fig. 5. Average utility vs. the channel SNR for 2 concurrently transmitting senders and enabled utility optimization.

C. *Results for Multiple Multicast Streams*

For the case of multicast transmission we consider that each one of the senders transmits an independent video stream to all the other destinations and not just one. More specifically in each experiment half of the senders transmit MOTHER & DAUGHTER and half FOREMAN. This means for example that when the utility is presented in the following figures each

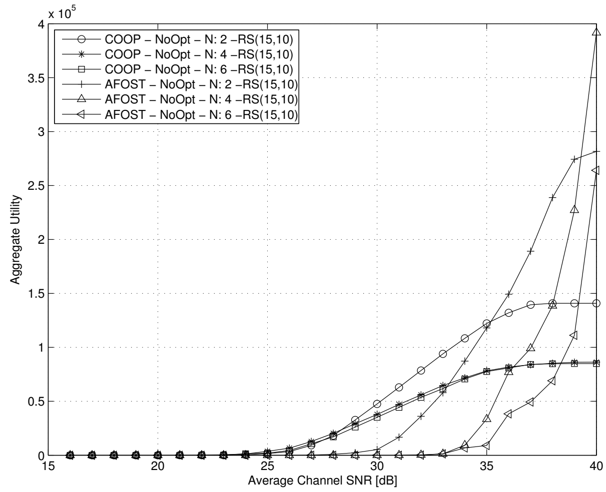


(a) d_s=10 seconds and 30 fps

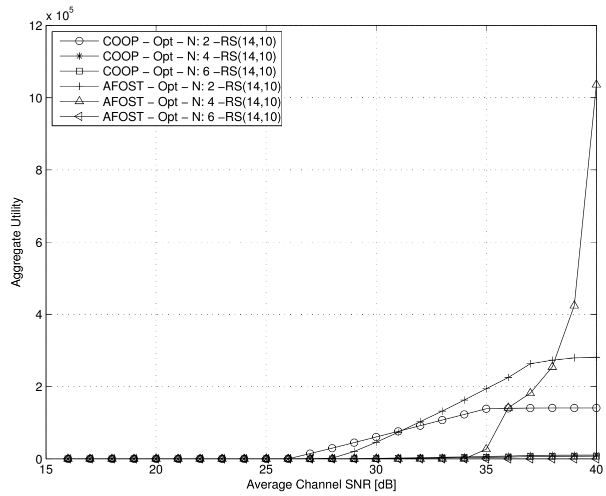


(b) d_s=2 and 4 seconds and 30 fps

Fig. 6. Average utility vs. the channel SNR for 4 concurrently transmitting senders and enabled utility optimization.



(a) Utility optimization disabled



(b) Utility optimization enabled

Fig. 7. Average utility vs. the channel SNR for multicast transmission of 2, 4, and 6 concurrently transmitting senders, d_s=4 seconds and 30 fps.

destination collectively measures and averages the utility of all the flows that are incoming towards it. The benefit we expect to come is from the fact that no packet that is decoded at the PHY is wasted since all the packets are locally useful for video playback.

Results are shown for 2, 4, and 6 senders and the NoOpt system in Fig. 7(a), while results for the Opt system are shown in Fig. 7(b). We notice in both figures that the rate of increase of the average utility of AFOST over COOP is higher as the channel quality is improved when compared with the previous experiments. This is the case for both Opt and NoOpt. The reason is that when the channel is poor the number of video packets that are not delivered to the application is increased while in the case that the PHY performs better then more PHY packets are successfully decoded simultaneously at all destinations and so more video units are played back. This result clarifies the main point we want to come across from the multicast experiment. That is, the combination of the proposed PHY and the utility optimization framework is more important to be used when the channel is good and not crucial when the channel is worse when compared to the unicast scenario.

VIII. Conclusions

In this paper we presented an RD-based utility optimization framework for video transmission in a wireless cooperative network that allows interfering transmissions to occur as part of the normal system operation. The first benefit from the adoption of the new PHY was a simpler formulation of the utility optimization problem. The second direct benefit was that when multiple senders transmit concurrently, a higher number of transmitted packets per unit of time can be recovered at the PHY. The performance results showed that even though the interference of only two wireless packet transmissions is optimal for increasing throughput, the utility of multiple transmitted video sequences can be increased even if more than two senders transmit concurrently. The later result was also shown to be possible even if the utility optimization framework is disabled and only the cooperative protocol that allows interference is used. If a delay constraint is present, the previous result is emphasized even more since the concurrent transmission naturally expedites the transmission of a higher number of packets on-time. The results for multicast video delivery demonstrated the need for channel adaptive scheme since the performance of the proposed framework is significantly improved for good channel conditions but when the channel is poor interfering transmissions is not the best choice. In our future plans we intend to focus first on the development of such a channel adaptive medium access control (MAC) protocol that reaps the full benefits of the proposed scheme depending on the channel conditions, and second on the combination of the proposed approach with more advanced error protection algorithms for the video streams like UEP.

REFERENCES

[1] M. van der Schaar and P. Chou, *Multimedia over IP and Wireless Networks*. Academic Press, 2007.

[2] H. Luo, A. Argyriou, D. Wu, and S. Ci, "Joint source coding and network-supported channel coding for video streaming in wireless multi-hop networks," *IEEE Transactions on Multimedia*, vol. 11, no. 7, pp. 1362–1372, November 2009.

[3] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Processing Magazine*, vol. 15, no. 4, pp. 74–90, November 1998.

[4] M.-T. Sun and A. R. Reibman, *Compressed video over networks*. Marcel Dekker Inc., September 2001.

[5] G. Cheung and A. Zakhor, "Bit allocation for joint source/channel coding of scalable video," *IEEE Transactions on Image Processing*, vol. 9, no. 3, pp. 340–356, March 2000.

[6] P. A. Chou and Z. Miao, "Rate-distortion optimized streaming of packetized media," *Microsoft Research Technical Report MSR-TR-2001-35*, 2001.

[7] R. Ahlswede, N. Cai, S.-Y. R. Li, and R. W. Yeung, "Network information flow," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1204–1216, July 2000.

[8] S. Katti, S. Gollakota, and D. Katabi, "Embracing Wireless Interference: Analog Network Coding," in *SIGCOMM*, 2007.

[9] S. Zhang, S. C. Liew, and P. P. Lam, "Hot topic: Physical-layer network coding," in *MobiCom*, 2006.

[10] A. Argyriou and A. Pandharipande, "Cooperative protocol for analog network coding in distributed wireless networks," *IEEE Transactions on Wireless Communications*, vol. 8, no. 10, pp. 2014–2023, October 2010.

[11] P. Larsson, N. Johansson, and K.-E. Sunell, "Coded bi–directional relaying," in *5th Scandinavian WS on Ad Hoc Networks*, May 2005.

[12] P. Popovski and H. Yomo, "The anti-packets can increase the achievable throughput of a wireless multi-hop network," in *IEEE International Conference on Communication (ICC)*, June 2006.

[13] O. Alay, T. Korakis, Y. Wang, E. Erkip, and S. Panwar, "Layered wireless video multicast using omni-directional relays," in *IEEE ICASSP*, 2008.

[14] ——, "Layered wireless video multicast using relays," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 8, pp. 1095 – 1109, August 2010.

[15] O. Alay, Z. Guo, Y. Wang, E. Erkip, and S. Panwar, "Enhanced parity packet transmission for video multicast using r-dstc," in *IEEE PIMRC*, 2010.

[16] H.Seferoglu and A. Markopoulou, "Opportunistic network coding for video streaming over wireless," in *Packet Video Workshop*, Lausanne, Switzerland, November 2007.

[17] X. Liu, G. Cheung, and C.-N. Chuah, "Structured network coding and cooperative wireless ad-hoc peer-to-peer repair for wwan video broadcast," *IEEE Transactions on Multimedia*, vol. 11, no. 4, pp. 730–741, June 2009.

[18] Y. Wu, P. A. Chou, and S.-Y. Kung, "Information exchange in wireless networks with network coding and physical-layer broadcast," in *39th Annual Conference on Information Sciences and Systems (CISS)*, Baltimore, MD, March 2005.

[19] A. Argyriou, "Coordinating interfering transmissions in cooperative wireless lans," *IEEE Transactions on Wireless Communications*, vol. 10, no. 11, pp. 3804–3812, November 2011.

[20] A. Argyriou and A. Pandharipande, "Cooperative protocol for analog network coding in wireless networks," in *IEEE International Conference on Communications (ICC)*, May 2010.

[21] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[22] J. N. Laneman, D. Tse, and G. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," *IEEE Transactions on Information Theory*, vol. 50, no. 12, pp. 3062–3080, December 2004.

[23] P. Liu, Z. Tao, S. Narayanan, T. Korakis, and S. Panwar, "CoopMAC: A cooperative MAC for wireless LANs," *IEEE Journal on Selected Areas in Communications, Special Issue on Cooperative Communications and Networking*, vol. 25, no. 2, February 2007.

[24] J. Chakareski, J. Apostolopoulos, S. Wee, W.-T. Tan, and B. Girod, "Rate-distortion hint tracks for adaptive video streaming," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 10, pp. 1257–1269, October 2005.

[25] J. Chakareski and P. Frossard, "Rate-distortion optimized distributed packet scheduling of multiple video streams over shared communication resources," *IEEE Transactions on Multimedia*, vol. 8, no. 2, pp. 207–218, April 2006.

[26] ——, "Distributed collaboration for enhanced sender-driven video streaming," *IEEE Transactions on Multimedia*, vol. 10, no. 5, pp. 858–870, August 2008.

[27] J. Huang, Z. Han, M. Chiang, and H. Poor, "Auction-based resource allocation for cooperative communications," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 7, pp. 1226 –1237, September 2008.

[28] J. Huang, Z. Li, M. Chiang, and A. K. Katsaggelos, "Joint source adaptation and resource allocation for multi-user wireless video streaming," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 5, pp. 582 – 595, May 2008.

[29] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[30] D. Bertsekas, *Network Optimization: Continuous and Discrete Models*. Athena Scientific, 1998.

[31] "ITU-T Recommendation H.264, Advanced video coding for generic audiovisual services," May 2003.

PLACE PHOTO HERE

Antonios Argyriou received the Diploma in electrical and computer engineering from Democritus University of Thrace, Greece, in 2001, and the M.S. and Ph.D. degrees in electrical and computer engineering as a Fulbright scholar from the Georgia Institute of Technology, Atlanta, USA, in 2003 and 2005, respectively.

Currently, he is a tenure-track faculty member at the Department of Computer and Communications Engineering, University of Thessaly, Greece. From 2007 until 2010 he was a Senior Research Scientist at Philips Research, Eindhoven, The Netherlands. From 2004 until 2005, he was a Senior Engineer at Soft.Networks, Atlanta, GA. Dr. Argyriou currently serves in the editorial board of the *Journal of Communications*. He has also served as guest editor for the *IEEE Transactions on Multimedia* Special Issue on Quality-Driven Cross-Layer Design, and he was also a lead guest editor for the *Journal of Communications*, Special Issue on Network Coding and Applications. Dr. Argyriou serves in the TPC of several international conferences and workshops in the area of communications, networking, and signal processing. His current research interests are in the areas of communication systems and computer networks. He is a member of IEEE.