This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the IEEE ICC 2011 proceedings

On the Effect of Cooperative Relaying on the Performance of Video Streaming Applications in Cognitive Radio Networks

Zhangyu Guan†‡, Lei Ding‡, Tommaso Melodia‡, and Dongfeng Yuan†
‡Department of Electrical Engineering, University at Buffalo, The State University of New York, Buffalo, NY, 14260
†School of Information Science and Engineering, Shandong University, China, 250100
Email: zhangyuguan@sdu.edu.cn, {leiding, tmelodia}@buffalo.edu, dfyuan@sdu.edu.cn

***Abstract*—The problem of optimal resource allocation to stream multimedia content in *cognitive ad hoc networks with cooperative relays* is addressed in this paper. Cooperative transmission is a promising technique to increase the capacity of wireless links by exploiting spatial diversity without multiple antennas at each node. However, mainstream research in this field focuses on optimizing physical layer performance measures, with little consideration for application-specific and network-wide performance measures. In this paper, the problem of joint video encoding rate control, power control, relay selection and channel assignment is formulated as a mixed-integer nonlinear problem (MINLP), and a solution algorithm based on a combination of the branch and bound framework and convex relaxation techniques is then proposed. The proposed solution jointly allocates channel, power, video encoding rate, and relay nodes to secondary users to maximize the network-wide video quality. Performance evaluation results show that cognitive networks with cooperative relaying can provide considerably higher video quality (in terms of the average peak signal-to-noise ratio (PSNR)) than traditional networking technologies that do not rely on cooperation and dynamic spectrum allocation.**

***Index Terms*—Cooperative transmission, cognitive radio networks, spectrum allocation, video streaming, cross-layer design.**

I. Introduction

The need to wirelessly share high-quality multimedia content is driving the need for ever-increasing wireless transport capacity, which is however limited by the scarcity of the available radio spectrum. Cognitive radio [1], [2] is a promising technology to increase the utilization efficiency of the existing radio spectrum. A key challenge in cognitive radio networks is the design of efficient spectrum sharing algorithms to enable wireless devices to opportunistically access portions of the spectrum as they become available. Consequently, techniques for dynamic spectrum access have received significant attention [3]–[6]. However, no previous research has concentrated on developing a cross-layer optimization framework to analyze the performance of multimedia streaming applications over cognitive radios, which is the subject of this paper.

Within the context of dynamic spectrum access technologies, we additionally consider techniques to leverage the *spatial diversity* that characterizes the wireless channel. Spatial diversity is traditionally exploited by using multiple transceiver antennas to effectively cope with channel fading. However, equipping a mobile device with multiple antennas may not be practical, since the minimum required separation between the antennas is dictated by the operating radio wavelength. The concept of *cooperative communications* has been hence proposed to achieve spatial diversity without requiring multiple transceiver antennas on the same node [7]–[9]. In cooperative communications, in their *virtual multiple-input single-output* (VMISO) variant, each node is equipped with a single antenna, and relies on the antennas of neighboring devices to achieve spatial diversity. There is a vast and growing literature on information and communication theory problems [10] in cooperative communications. However, the common theme of most research in this field is to optimize physical layer performance measures (i.e., bit error rate and link outage probability) from a broad system perspective, without considering in much detail how cooperation interacts with higher layers of the protocol stack to improve network performance measures. For example, [11], [12] investigate the achievable rates and diversity gains of given cooperative schemes focusing on a single source and destination pair. Some initial promising work on networking aspects of cooperative communications includes studies on medium access control protocols to leverage cooperation [9], [13], cooperative routing [14]–[17], optimal network-wide relay selection [18], and optimal stochastic control [19]. However, none of these works considers the impact of cooperation on the end-to-end video quality performance. Moreover, most of these works assume that the network is interference-free, which clearly constitutes an idealized version of the problem. Dynamic spectrum management with cooperative devices for multi-user multimedia traffic transmission is a substantially unexplored area.

For this reason, in this paper we introduce a cross-layer optimization framework to study the problem of joint control of the video encoding rate at the application layer, relay selection, channel assignment, and power control at the link and physical layers, to maximize the sum peak signal-to-noise ratio (PSNR) of multiple concurrent video sessions. We first formulate the problem of joint channel allocation, rate control, relay selection, and power control for video streaming in cognitive ad hoc networks. The problem turns out to be a non-convex and combinatorial optimization problem, for which we develop a new iterative solution algorithm that solves convex relaxations of the original problem within the framework of branch and bound. The proposed algorithm searches for the optimal solution iteratively. At each iteration, we relax the original non-convex problem to a series of convex problems, which are solved using polynomial-complexity interior-point algorithms [20], [21]. By applying the proposed algorithm, we show that *cooperative relaying can play an important role by significantly improving the network-wide video quality* of concurrent and competing video streaming applications. We additionally show that close-to-optimal performance can be achieved through joint spectrum management and cooperative relaying only, i.e., without relying on transmission power control.

The rest of the paper is organized as follows. In Section II, we describe the communication system model and formulate the cross-layer problem. The proposed solution algorithm is discussed in detail in Section III. In Section IV, we provide extensive performance evaluation of the performance limits achievable by the system under consideration by applying the proposed algorithm to application scenarios of interest. Finally, Section V concludes the paper.

II. System Model and Problem Formulation

We consider N_p primary users (PUs) and N_s secondary users (SUs). A portion of the spectrum is divided into N_f frequency channels. SUs are allowed to dynamically access portions of the spectrum without causing "harmful" interference to the PUs. Each SU compresses a video sequence at a given rate and then transmits the compressed video packets either using a direct link or through cooperative relaying. If a cooperative relay is used, the SU selects one from a number N_r of potential relay nodes. The video packets are enqueued at the source node whenever the video encoding rate is greater than the underlying capacity of the direct or cooperative link, and are dropped if they are not received before a pre-determined play-out deadline.

The objective of the problem is to maximize the sum video quality (expressed in terms of peak-signal-to-noise-ratio, PSNR) of all SUs by jointly optimizing the video encoding rate, relay selection, spectrum assignment and transmission power. Let $\mathcal{N}_p$, $\mathcal{N}_s$, $\mathcal{N}_r$ and $\mathcal{N}_f$ represent the sets of all PUs, SUs, potential relay nodes and available frequency channels, respectively. The PSNR of SU $n_s \in \mathcal{N}_s$, denoted as Q_{n_s}, is given by $Q_{n_s} = 10\log_{10}(D_{cst}/D_{n_s})$, where D_{n_s} represents the end-to-end video distortion measured in mean square error (MSE) and D_{cst} represents the worst-case MSE, which is assumed to be a constant equal to 255^2. The total video distortion D_{n_s} consists of *compression distortion* and *transmission distortion*, denoted as $D^c_{n_s}$ and $D^t_{n_s}$, respectively. The compression distortion $D^c_{n_s}$ can be modeled as a function of the video encoding rate R_{n_s} as

$$D^c_{n_s} = D^0_{n_s} + \frac{\theta_{n_s}}{R_{n_s} - R^0_{n_s}},\ D^0_{n_s} > 0,\ R_{n_s} > R^0_{n_s}, \quad (1)$$

where $D^0_{n_s}$, $R^0_{n_s}$ and θ_{n_s} are video-dependent parameters that can be measured offline or estimated in real time. $D^t_{n_s}$ is modeled as $D^t_{n_s} = k_{n_s} P^{los}_{n_s}$, where k_{n_s} represents the sensitivity of the video sequence to packet loss and $P^{los}_{n_s}$ represents the end-to-end packet loss rate, which can be caused by transmission errors and/or violations of the play-out deadline caused by queueing delay. If we represent the corresponding packet loss rate as $P^{err}_{n_s}$ and $P^{dly}_{n_s}$, respectively, then $P^{los}_{n_s} = P^{err}_{n_s} + P^{dly}_{n_s}$. As in [22], as a first approximation we model the packet queue as a $M/M/1$ queue. Then, the average queueing delay T_{n_s} can be expressed as [23] $T_{n_s} = L_{n_s}/(C_{n_s} - R_{n_s})$, where L_{n_s} is the average video packet length and C_{n_s} represents the capacity of the underlying link. The probability that the queueing delay exceeds the play-out deadline $T^0_{n_s}$ can then be expressed as $P^{dly}_{n_s} = e^{-T^0_{n_s}/(C_{n_s} - R_{n_s})}$.

Denote the vector of relay selection variables as $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_{n_s})_{n_s=1}^{N_s}$ with $\boldsymbol{\alpha}_{n_s} = (\alpha^{n_r}_{n_s})_{n_r=1}^{N_r}$, where $\alpha^{n_r}_{n_s} = 1$ if relay node n_r is selected by SU n_s and $\alpha^{n_r}_{n_s} = 0$ otherwise. Assume that each SU selects only the best relay and each relay node can be selected by at most one SU. Then, we have

$$\alpha^{n_r}_{n_s} \in \{0,\ 1\},\ \forall n_s \in \mathcal{N}_s,\ n_r \in \mathcal{N}_r \quad (2)$$

$$\sum_{n_r=1}^{N_r} \alpha^{n_r}_{n_s} \leq 1,\ \forall n_s \in \mathcal{N}_s, \quad (3)$$

$$\sum_{n_s=1}^{N_s} \alpha^{n_r}_{n_s} \leq 1,\ \forall n_r \in \mathcal{N}_r. \quad (4)$$

The capacity available to SU n_s can be expressed as

$$C_{n_s} = \left(1 - \sum_{n_r=1}^{N_r} \alpha^{n_r}_{n_s}\right) C^{dir}_{n_s} + \sum_{n_r=1}^{N_r} \alpha^{n_r}_{n_s} C^{cop}_{n_s n_r}, \quad (5)$$

where $C^{dir}_{n_s}$ represents the achievable capacity if SU n_s uses the direct link only and $C^{cop}_{n_s n_r}$ represents the capacity if SU n_s uses relay node n_r for cooperative relaying.

We consider a cooperative transmission scheme executed over two time-slots. At each time-slot, the SU selects a frequency for transmission. If a direct link is used, the same frequency is used in both time-slots. Denote the vector of frequency assignment variables for time-slot $t \in \{1,2\}$ as $\boldsymbol{\beta}_t = (\boldsymbol{\beta}_{t,n_s})_{n_s=1}^{N_s}$ with $\boldsymbol{\beta}_{t,n_s} = (\beta^{n_f}_{t,n_s})_{n_f=1}^{N_f}$, where $\beta^{n_f}_{t,n_s} = 1$ if frequency n_f is used by SU n_s for transmission in time-slot t and $\beta^{n_f}_{t,n_s} = 0$ otherwise. Then, we have

$$\beta^{n_f}_{t,n_s} \in \{0,1\},\ \forall t \in \{1,2\},\ n_s \in \mathcal{N}_s,\ n_f \in \mathcal{N}_f, \quad (6)$$

$$\sum_{n_f=1}^{N_f} \beta^{n_f}_{t,n_s} \leq 1,\ \forall t \in \{1,2\},\ n_s \in \mathcal{N}_s. \quad (7)$$

We assume that the decode-and-forward (DF) [8] strategy is employed at each relay node for cooperative relaying. Other forwarding strategies, e.g., amplify-and-forward (AF), will be addressed in our future work. Then, the direct and cooperative link capacity $C^{dir}_{n_s}$ and $C^{cop}_{n_s n_r}$ can be expressed as

$$C^{dir}_{n_s} = B\log_2\left(1 + SINR^{s2d}_{n_s}\right), \quad (8)$$

$$C^{cop}_{n_s n_r} = \min(C^{s2r}_{n_s n_r},\ C^{sr2d}_{n_s n_r})/2, \quad (9)$$

where $C^{s2r}_{n_s n_r}$ represents the link capacity for SU n_s from its source node to relay node n_r, and $C^{sr2d}_{n_s n_r}$ represents the capacity achieved by decoding the two signals received from source and relay nodes through maximal ratio combining [8]. $C^{s2r}_{n_s n_r}$ and $C^{sr2d}_{n_s n_r}$ can be expressed as

$$C^{s2r}_{n_s n_r} = B\log_2\left(1 + SINR^{s2r}_{n_s n_r}\right), \quad (10)$$

$$C^{sr2d}_{n_s n_r} = B\log_2\left(1 + SINR^{s2d}_{n_s} + SINR^{r2d}_{n_s n_r}\right). \quad (11)$$

In (8), (10) and (11), $SINR_{n_s}^{s2d}, SINR_{n_s n_r}^{s2r}$ and $SINR_{n_s n_r}^{r2d}$ represent the signal-to-noise-plus-ratio (SINR) for links from the source to destination node, source to relay node, relay to destination node, respectively, and can be expressed as

$$SINR_{n_s}^{s2d} = \sum_{n_f=1}^{N_f} \beta_{1,n_s}^{n_f} G_{n_s n_s}^{s2d} P_{n_s} / (\delta_{n_s n_f}^2 + I_{n_s}^{n_f}), \quad (12)$$

$$SINR_{n_s n_r}^{s2r} = \sum_{n_f=1}^{N_f} \beta_{1,n_s}^{n_f} G_{n_s n_r}^{s2r} P_{n_s} / (\delta_{n_r n_f}^2 + I_{n_r}^{n_f}), \quad (13)$$

$$SINR_{n_s n_r}^{r2d} = \sum_{n_f=1}^{N_f} \beta_{2,n_s}^{n_f} G_{n_s n_r}^{r2d} P_{n_r} / (\delta_{n_s n_f}^2 + I_{n_s}^{n_f}), \quad (14)$$

where $G_{n_s n_s}^{s2d}$, $G_{n_s n_r}^{s2r}$ and $G_{n_s n_r}^{r2d}$ represent the channel gain of link from the source node of SU n_s to its destination node, from the source node of SU n_s to relay node n_r, and from relay node n_r to the destination node of SU n_s, respectively. P_{n_s} and P_{n_r} represent the transmission power of source node of SU n_s and relay node n_r, respectively. $I_{n_s}^{n_f}$ and $I_{n_r}^{n_f}$ represent the interference measured at the destination node of SU n_s and relay node n_r on frequency n_f, respectively.

To avoid the restricting assumption of strict global synchronization, we approximate the interference by averaging over the two time-slots of a cooperative transmission scheme[1]. Then, $I_{n_s}^{n_f}$ can be expressed as

$$I_{n_s}^{n_f} = \sum_{m_s \in \mathcal{N}_s}^{m_s \neq n_s} \Bigg[\left(1 - \frac{1}{2} \sum_{n_r=1}^{N_r} \alpha_{m_s}^{n_r}\right) \beta_{1,m_s}^{n_f} G_{m_s n_s}^{s2d} P_{m_s} + \frac{1}{2} \sum_{n_r=1}^{N_r} \alpha_{m_s}^{n_r} \beta_{2,m_s}^{n_f} G_{n_s n_r}^{r2d} P_{n_r} \Bigg], \quad (15)$$

and $I_{n_r}^{n_f}$ can be expressed similarly.

In cognitive networks the interference-temperature constraints for PUs should be satisfied, since the SUs dynamically access a portion of spectrum assigned to the PUs. Use $I_{n_p n_f}^{max}$ to represent the maximum interference tolerable at destination node of the PU $n_p \in \mathcal{N}_p$ if frequency n_f is used. Then the interference caused by the SUs to PU n_p on frequency n_f, denoted as $I_{n_p}^{n_f}$, cannot exceed $I_{n_p n_f}^{max}$, i.e.,

$$I_{n_p}^{n_f} \leq I_{n_p n_f}^{max}, \ n_p \in \mathcal{N}_p, \ n_f \in \mathcal{N}_{f,n_p}, \quad (16)$$

where $\mathcal{N}_{f,n_p}$ represents the set of frequency channels used by PU n_p and the $I_{n_p}^{n_f}$ can be expressed similarly to $I_{n_s}^{n_f}$ in (15).

Denote the vectors of power allocation variables with $\boldsymbol{P}_s = (P_{n_s})_{n_s=1}^{N_s}$ and $\boldsymbol{P}_r = (P_{n_r})_{n_r=1}^{N_r}$, the vector of video encoding rate as $\boldsymbol{R} = (R_{n_s})_{n_s=1}^{N_s}$. Then, based on the above discussion the problem can formulated as

$$Given: \quad I_{n_p n_f}^{max}, \ n_f \in \mathcal{N}_{f,n_p}, \ n_p \in \mathcal{N}_p \quad (17)$$

$$Find: \quad \boldsymbol{R}, \ \boldsymbol{\alpha}, \ \boldsymbol{\beta}_1, \ \boldsymbol{\beta}_2, \ \boldsymbol{P}_s, \ \boldsymbol{P}_r \quad (18)$$

$$Maximize: \quad \sum_{n_s \in \mathcal{N}_s} Q_{n_s} \quad (19)$$

$$Subject\ to: \quad R_{n_s} > R_{n_s}^0, \ n_s \in \mathcal{N}_s \quad (20)$$

$$T_{n_s} \leq T_{n_s}^0, \ n_s \in \mathcal{N}_s \quad (21)$$

$$P_{n_s}^{min} \leq P_{n_s} \leq P_{n_s}^{max}, \ n_s \in \mathcal{N}_s \quad (22)$$

$$P_{n_r}^{min} \leq P_{n_r} \leq P_{n_r}^{max}, \ n_r \in \mathcal{N}_r \quad (23)$$

$$(2), (3), (4), (6), (7), (16). \quad (24)$$

III. Proposed Solution Algorithm

The problem formulated in Section II is a mixed integer and nonlinear problem (MINLP). In this section, we propose a solution algorithm based on a combination of branch-and-bound framework [24] and convex relaxation to the MINLP. It can be proven that the algorithm always converges to a solution as close to the optimum as we wish and that in practice has much lower computation complexity than exhaustive search.

If we denote Q^* as the optimal objective function in (19), and $\epsilon \in [0, 1]$ as the optimality precision, then the algorithm iteratively searches for a solution with objective function Q that satisfies $Q \geq \epsilon Q^*$. Let Λ_0 represent the domain set of the MINLP, which includes all possible combinations of transmission strategies, i.e., $\Lambda_0 = \{(\boldsymbol{r}, \boldsymbol{\alpha}, \boldsymbol{\beta_1}, \boldsymbol{\beta_2}, \boldsymbol{P_s}, \boldsymbol{P_r})\}$. Iteratively, the algorithm partitions Λ_0 into a set $\boldsymbol{\Lambda}$ of sub-domains, and at each iteration the algorithm selects a sub-domain $\Lambda_n \in \boldsymbol{\Lambda}$ and calculates a local upper bound and a local lower bound on the original objective function. Denote the upper bound as $up(\Lambda_n)$ and the lower bound as $lr(\Lambda_n)$, respectively. The algorithm also maintains a global upper bound $UP(\Lambda_0)$ and a global lower bound $LR(\Lambda_0)$ through the iterations and updates them as $UP(\Lambda_0) = \max_{\Lambda_n \in \boldsymbol{\Lambda}} up(\Lambda_n)$ and $LR(\Lambda_0) = \max_{\Lambda_n \in \boldsymbol{\Lambda}} lr(\Lambda_n)$. Then, the sub-domain with the highest local upper bound is selected for further partition. As sub-domains are partitioned, their measures become smaller and each sub-domain tends to contain a fixed transmission strategy. Correspondingly, the gap between $UP(\Lambda_0)$ and $LR(\Lambda_0)$ tends to zero. The iteration stops if $LR(\Lambda_0) \geq \epsilon UR(\Lambda_0)$ can be satisfied and set $Q = LR(\Lambda_0)$. Since the optimal objective function Q^* is guaranteed to be located between $LR(\Lambda_0)$ and $UP(\Lambda_0)$, we have $Q \geq \epsilon Q^*$. To speedup convergence of the iteration, if sub-domain $\Lambda_n \in \boldsymbol{\Lambda}$ satisfies that $up(\Lambda_n) < LR(\Lambda_0)$, it means that the optimal solution is certainly not located in Λ_n and therefore Λ_n will be removed from $\boldsymbol{\Lambda}$. Next, we discuss how to design $up(\Lambda_n)$ and $lr(\Lambda_n)$ and how to partition a selected sub-domain.

Convex Relaxation: The local upper bound $up(\Lambda_n)$ is obtained by relaxing the original MINLP to a convex problem so that it can be solved with polynomial computational time. We first introduce the following lemma.

***Lemma** 1:* The utility function Q_{n_s} is a concave function of video encoding rate R_{n_s} and underlying link capacity C_{n_s}.

[1]We have experimentally validated this assumption, and verified its negligible impact on the overall network performance - a more detailed discussion is omitted because of space limits.

We omit the proof of Lemma 1 because of space limits. We also observe that Q_{n_s} is an increasing function of C_{n_s}. Based on Lemma 1 and on this observation, we only need to relax C_{n_s} to be a concave function of the relay selection variables $\boldsymbol{\alpha}$, channel assignment variables $\boldsymbol{\beta_1}$ and $\boldsymbol{\beta_2}$, and power allocation variables $\boldsymbol{P}_s$ and $\boldsymbol{P}_r$. To this end, we first relax $\boldsymbol{\alpha}, \boldsymbol{\beta_1}$ and $\boldsymbol{\beta_2}$ to be real, then (2) and (6) can be rewritten as $\alpha_{n_s,n_r}^{min} \le \alpha_{n_s}^{n_r} \le \alpha_{n_s,n_r}^{max}$, $\forall n_s \in \mathcal{N}_s$, $\forall n_r \in \mathcal{N}_r$ and $\beta_{t,n_s}^{n_f,max} \le \beta_{t,n_s}^{n_f} \le \beta_{t,n_s}^{n_f,max}$, $\forall t \in \{1,2\}$, $n_s \in \mathcal{N}_s$, $n_f \in \mathcal{N}_f$, where α_{n_s,n_r}^{min} and $\beta_{t,n_s}^{n_f,min}$ are initialized to 0, and α_{n_s,n_r}^{max} and $\beta_{t,n_s}^{n_f,max}$ are initialized to 1. Then, the original problem can be relaxed considering the following three cases.

Case 1: The overall link capacity C_{n_s} formulated in (5)-(15) can be relaxed in different ways. We use a simple relaxation method based on "best-case of signal and interference". Given a selected sub-domain, if not all relay selection variables are fixed, i.e., $\alpha_{n_s,n_r}^{min} = \alpha_{n_s,n_r}^{max}$ does not hold for all $n_s \in \mathcal{N}_s$ and $n_r \in \mathcal{N}_r$, then the interference $I_{n_s}^{n_f}$ in (15) can be relaxed to $\hat{I}_{n_s}^{n_f}$ as follows

$$\hat{I}_{n_s}^{n_f} = \sum_{m_s \in \mathcal{N}_s}^{m_s \neq n_s} \left[\left(1 - \frac{1}{2} \max_{n_r \in \mathcal{N}_r} \alpha_{m_s,n_r}^{max}\right) \beta_{1,m_s}^{n_f,min} G_{m_s n_s}^{s2d} P_{m_s}^{min} + \frac{1}{2} \sum_{n_r=1}^{N_r} \alpha_{m_s,n_r}^{min} \beta_{2,m_s}^{n_f,min} G_{n_s n_r}^{r2d} P_{n_r}^{min} \right], \quad (25)$$

and $I_{n_r}^{n_f}$ and $I_{n_p}^{n_f}$ can be relaxed similarly. Then, the SINR in (12) can be relaxed to $\widehat{SINR}_{n_s}^{s2d} = \max_{n_f \in \mathcal{N}_f} \beta_{1,n_s}^{n_f,max} G_{n_s n_s}^{s2d} P_{n_s}^{max} / (\delta_{n_s n_f}^2 + \widehat{I}_{n_s}^{n_f})$, and $SINR_{n_s n_r}^{s2r}$ in (13) and $SINR_{n_s n_r}^{r2d}$ in (14) can be relaxed similarly. Then, capacities in (8), (9), (10) and (11) are constant, and hence the overall link capacity C_{n_s} in (5) is an affine (and therefore convex) function of $\boldsymbol{\alpha}$ only.

Case 2: Given a sub-domain, if $\alpha_{n_s,n_r}^{min} = \alpha_{n_s,n_r}^{max}$ hold true for all $n_s \in \mathcal{N}_s$ and $n_r \in \mathcal{N}_r$, then the SINR in (12) can be relaxed as $\widetilde{SINR}_{n_s}^{s2d} = \sum_{n_f=1}^{N_f} \beta_{1,n_s}^{n_f} G_{n_s n_s}^{s2d} P_{n_s}^{max} / (\delta_{n_s n_f}^2 + \widehat{I}_{n_s}^{n_f})$, and $SINR_{n_s n_r}^{s2r}$ in (13) and $SINR_{n_s n_r}^{r2d}$ in (14) can be relaxed similarly. Then, it can be proven that capacities in (8), (10) and (11) are concave functions of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$. Since the component-wise minimum of two concave functions is a concave function, the capacity in (9) and hence C_{n_s} in (5) are also concave.

Case 3: Given a sub-domain, if all relay selection variables and channel assignment variables are fixed, we only need to relax the SINR in (12) as $\overline{SINR}_{n_s}^{s2d} = \sum_{n_f=1}^{N_f} \beta_{1,n_s}^{n_f,max} G_{n_s n_s}^{s2d} P_{n_s} / (\delta_{n_s n_f}^2 + \widehat{I}_{n_s}^{n_f})$, and relax $SINR_{n_s n_r}^{s2r}$ in (13) and $SINR_{n_s n_r}^{r2d}$ in (14) similarly. Then, we can prove that the overall link capacity C_{n_s} in (5) is a concave function of $\boldsymbol{P}_s$ and $\boldsymbol{P}_r$.

Based on the above relaxations and considering that all constraints in (20)-(24) are either affine or convex, the relaxed problem is convex and can be solved efficiently with polynomial computational time.

Local Search: The local lower bound $lr(\Lambda_n)$ is obtained through a local search for a feasible solution based on the upper bound obtained as a solution of the relaxed problem. In case 1, since the link capacity C_{n_s} is a linear function of $\boldsymbol{\alpha}$, the relay selection solution is integer and hence feasible. For channel assignment, an arbitrary feasible $\boldsymbol{\beta}_t$, $t \in \{1,2\}$ can be used and we allocate the subchannels sequentially among all SUs. Maximum transmission power is used for all the source nodes of SUs and the selected relay nodes. Then, we can calculate a feasible value of the overall link capacity C_{n_s} for each SU and further we can calculate a feasible value as a lower bound on the objective function in (19) by solving a convex problem. In case 2, use $\boldsymbol{\beta}_t^* = (\beta_{t,n_s}^{n_f*})_{n_s \in \mathcal{N}_s}^{n_f \in \mathcal{N}_f}$ to represent a solution of the relaxed problem. A feasible channel assignment solution $\widehat{\boldsymbol{\beta}_t^*} = (\widehat{\beta_{t,n_s}^{n_f*}})_{n_s \in \mathcal{N}_s}^{n_f \in \mathcal{N}_f}$ can be obtained by performing $\widehat{\beta_{t,n_s}^{n_f*}} = round(\beta_{t,n_s}^{n_f*})$, $\forall n_s \in \mathcal{N}_s$, $n_f \in \mathcal{N}_f$. The resulting channel assignment variables are certainly feasible. Hence, we can calculate a feasible capacity for each SU and then calculate a feasible sum utility function. Finally in case 3, the solution of power allocation is certainly feasible and hence a feasible sum utility can be calculated.

Domain partition: The set $\boldsymbol{\Lambda}$ of sub-domains is initialized to $\{\Lambda_0\}$, and the domain Λ_0 is then partitioned into two sub-domains. When partitioning Λ_0, we consider the physical meaning of the variables. Note that in Λ_0 the relay selection variables are certainly not fixed. Hence the problem is relaxed using case 1. Denote the solution of relay selection of the relaxed problem as $\boldsymbol{\alpha}^* = (\alpha_{n_r}^{n_s})_{n_r \in \mathcal{N}_r}^{n_s \in \mathcal{N}_s}$. Since a larger value of $\alpha_{n_r}^{n_s,*}$ means that the relay n_r is more likely to be selected by the SU n_s, we partition Λ_0 by fixing the value of $\alpha_{n_r}^{n_s}$ that has the highest value, i.e., fixing the value of $\alpha_{n_r^*}^{n_s^*}$ such that $(n_r^*, n_s^*) = \arg\max_{n_r \in \mathcal{N}_r, n_s \in \mathcal{N}_s} \alpha_{n_r}^{n_s,*}$, and setting $\alpha_{n_r^*,n_s}^{max} = 0$ for all $n_s \in \mathcal{N}_s$, $\alpha_{n_r,n_s^*}^{max} = 0$ for all $n_r \in \mathcal{N}_r$ and $\alpha_{n_r^*,n_s^*}^{min} = 1$. When all relay selection variables are fixed in a sub-domain, the latter is further partitioned by fixing the channel assignment variables $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$ similarly to fixing the relay selection variables $\boldsymbol{\alpha}$.

Finally, a sub-domain with both relay selection and channel assignment variables fixed is further partitioned by partitioning the power allocation variables. The power allocation variable that has the highest power range is selected and partitioned from the middle point of the range. To this end, $P_{n_s^*}$ and $P_{n_r^*}$ are selected such that

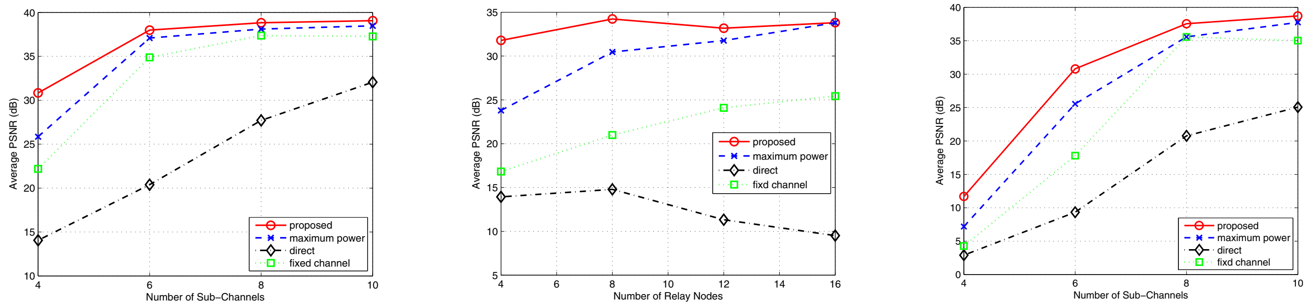
$$n_s^* = \max_{n_s \in \mathcal{N}_s} P_{n_s}^{max} - P_{n_s}^{min}, \quad (26)$$

$$n_r^* = \max_{n_r \in \mathcal{N}_r} (P_{n_r}^{max} - P_{n_r}^{min}) \cdot \max_{n_s \in \mathcal{N}_s} \alpha_{n_r,n_s}^{max}. \quad (27)$$

If $P_{n_s^*}^{max} - P_{n_s^*}^{min} \ge P_{n_r^*}^{max} - P_{n_r^*}^{min}$, then $P_{n_s^*}$ is partitioned from $(P_{n_s^*}^{max} + P_{n_s^*}^{min})/2$, otherwise $P_{n_r^*}$ is partitioned similarly. In (26) and (27) we only consider the source nodes of SUs and the selected relay nodes while transmission power of relay nodes that are not selected are not partitioned.

IV. Numerical Results

In this section, we present numerical results illustrating the performance of the system under consideration through the proposed optimal solution algorithm. Our most important finding is that *cooperative relaying can significantly improve the*



(a) (b) (c)

Fig. 2: Average PSNR achieved by the four strategies in case of (a) one PU, eight relay nodes and varying number of sub-channels, (b) two PUs, six subchannels and varying number of relay nodes, and (c) two PUs, eight relay nodes and varying number of sub-channels.

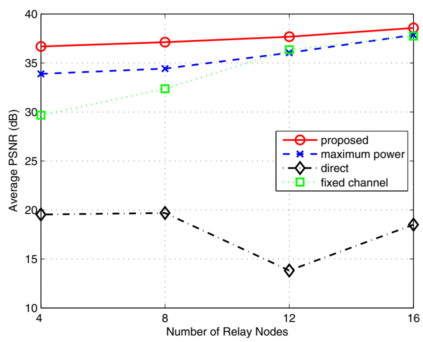


Fig. 1: Average PSNR achieved by the four strategies in case of one PU, six subchannels and varying number of relay nodes.

video quality since the constraints of interference-temperature for the PUs and minimum video encoding rate for the SUs are easier to satisfy by using cooperative relaying. We additionally show that the benefits of power control are limited - close-to-optimal performance can be achieved through joint optimal spectrum allocation and selection of cooperative relays, by transmitting at maximum transmission power.

We consider a terrain of $1000 \text{ m} \times 1000 \text{ m}$. The number of PUs, SUs and relay nodes are set to $N_p = 1,\ N_s = 8$ and $N_r = 4, 8, 12, 16$, respectively. Nodes are randomly deployed following a uniform distribution. The minimum and maximum transmission power of each node are set to 500 mW and 1000 mW, respectively. The average channel gain is assumed to be a function of the distance between two nodes, i.e., $G_{mn} = d^{-\gamma}(m,n)$, where $d(m,n)$ represents the distance in meters between node m and n, γ represents the channel path loss and is set to 4. The average AWGN power at each SU destination node on each sub-channel is set to 10^{-7} mW, and the maximum interference allowed at the PU is set to 10^{-8} mW. The number of sub-channels is set to $4, 6$ and 8, respectively, and the bandwidth of each sub-channel is set to 1.25 MHz. Two video sequences are considered, with parameters $D_1^0 = 0.38,\ \theta_1 = 2537$ kbps, $R_1^0 = 18.3$ kbps, $P_1^{err} = 0.01,\ k_1 = 750,\ T_1^0 = 350$ ms, $L_1 = 3040$ bits for sequence 1 and $D_2^0 = 0,\ \theta_2 = 858$ kbps, $R_2^0 = 0.67$ kbps, $P_2^{err} = 0,\ k_2 = 30,\ T_2^0 = 350$ ms, $L_2 = 3040$ bits for sequence 2. Four SUs transmit sequence 1, while the others transmit sequence 2. The optimality precision in the proposed algorithm is set to $\epsilon = 95\%$ and the performance of the resource allocation strategy obtained by applying the proposed algorithm (referred to simply as "proposed") is evaluated against three alternative resource allocation strategies: maximum transmission power ("maximum power"), direct transmission only ("direct") and fixed channel assignment ("fixed channel"). All results presented in this section are obtained by averaging over 30 independent simulations.

In Fig. 1, we report the average PNSR achieved by the four strategies for the case of one PU, six sub-channels, and varying number of relay nodes. The "proposed" strategy shows the best performance while "direct" results in the lowest average PSNR. In the case of four relay nodes, an improvement of over 15 dB can be achieved and further improvement can be observed with increasing number of relay nodes. We also observe that in all cases except for "direct", the average PSNR increases as the number of relay nodes increases. In case of "fixed channel", cooperative relaying is used to avoid "harmful" interference among SUs and PUs. Around 7 dB of improvement in average PSNR can be achieved with 12 relay nodes compared to 4 relay nodes. Since in "maximum power" interference can be avoided both through sub-channel assignment and by using cooperative relaying, the average PSNR increases more moderately as the number of relay nodes increases compared to "fixed channel".

Results for the average PSNR with varying number of sub-channels are given in Fig. 2(a), with one PU and eight relay nodes. The best performance is achieved by the "proposed" strategy. We observe that in all cases the average PSNR increases as the number of sub-channels increases. Especially, the average PSNR in the "direct" case keeps increasing considerably, since the "direct" strategy does not use cooperative relaying and the sub-channel assignment is mainly used for interference avoidance. When there are only four sub-channels, which is a low number compared to the SU number, significant improvement in the average PSNR can be made by the "proposed" strategy compared to the other three strategies. As the number of sub-channels increases from 4 to 10, the performance gain is reduced, e.g., from around 15 dB to around 5 dB compared to "direct". Moreover, in case of 10 sub-channels, "maximum power" and "fixed channel" can achieve a performance close to the "proposed". This is because with a sufficient number of sub-channels interference can be avoided through frequency diversity by using different sub-channels. Therefore, the beneficial effect of cooperative relaying is reduced.

In Fig. 2(b), the average PSNR is plotted for the case of two PUs, 6 sub-channels and varying number of relay nodes. The best performance is still achieved by the "proposed" strategy. However, we find that the average PSNR does not necessarily increase with the number of relay nodes. More sub-channels are occupied by PUs and consequently the interference constraints for the PUs are more likely to be violated resulting in an infeasible problem. In our simulations, if a problem is infeasible, the resulting average PSNR is set to 0. Additionally, the node locations in each simulation are generated randomly. Therefore, the curve corresponding to the "direct" case is not necessarily a strictly straight line with averaging results over a limited number of simulations. Similar result was observed in Fig. 1.

The average PSNR is shown in Fig. 2(c) with two PUs, eight relay nodes and varying number of sub-channels. Compared to the 1 PU case in Fig. 2(a), the average PSNR with two PUs is much lower. When there are only four sub-channels, the average PSNR is below 15 dB, which corresponds to unacceptable video quality. As the number of sub-channels increases, the average PSNR increases but at a decreasing rate except for the "direct" case. Moreover, even if there are sufficient number of sub-channels, cooperative relaying is still desirable for interference avoidance, since the "direct" strategy can only achieve an average PSNR of 25 dB, which corresponds to low video quality.

V. Conclusions

We studied the problem of video streaming in cognitive and cooperative wireless networks. We first formulated the problem of joint video encoding rate control, power control, relay selection and channel assignment as a MINLP, and then proposed a solution algorithm based on a combination of the branch and bound framework and convex relaxation techniques. Performance results showed that in cognitive networks where spectrum resources are not sufficient, cooperative relaying can play an important role in enabling interference avoidance and hence improving the quality of service of SUs.

Acknowledgement

This research was carried out while Zhangyu Guan was a visiting scholar at the State University of New York at Buffalo. The work of Zhangyu Guan and Dongfeng Yuan was supported by the NSFC (No. 60832008, No. 61071122).

References

[1] J. Mitola, *Cognitive Radio*. Licentiate thesis, Royal Inst. of Technol., KTH, September 1999.

[2] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt Generation/Dynamic Spectrum Access/Cognitive Radio Wireless Networks: A Survey," *Computer Networks Journal(Elsevier)*, vol. 50, pp. 2127–2159, September 2006.

[3] L. Cao and H. Zheng, "SPARTA: Stable and Efficient Spectrum Access in Next Generation Dynamic Spectrum Networks," in *Proc. of IEEE Intl. Conf. on Computer Communications (INFOCOM)*, (Phoenix, AZ, USA), pp. 870–878, April 2008.

[4] K. Liu and Q. Zhao, "A Restless Bandit Formulation of Opportunistic Access: Indexablity and Index Policy," in *Proc. of the 5th IEEE Conf. on Sensor, Mesh and Ad Hoc Communications and Networks (SECON) Workshops*, (Crowne Plaza, CA, USA), June 2008.

[5] L. Ding, T. Melodia, S. Batalama, J. Matyjas, and M. Medley, "Cross-layer Routing and Dynamic Spectrum Allocation in Cognitive Radio Ad Hoc Networks," *IEEE Transactions on Vehicular Technology*, vol. 59, pp. 1969–1979, May 2010.

[6] L. Ding, T. Melodia, S. N. Batalama, and J. D. Matyjas, "Distributed Routing, Relay Selection, and Spectrum Allocation in Cognitive and Cooperative Ad Hoc Networks," in *Proc. IEEE SECON*, (Boston, USA), June 2010.

[7] A. Sendonaris, E. Erkip, and B. Aazhang, "User Cooperation Diversity-Part I:System Description," *IEEE Trans. on Communications*, vol. 51, pp. 1927–1938, Nov. 2003.

[8] J. N. Laneman, D. N. C. Tse, and G. W. Wornell, "Cooperative Diversity in Wireless Networks: Efficient Protocols and Outage Behavior," *IEEE Trans. on Information Theory*, vol. 50, pp. 3062–3080, Dec. 2004.

[9] G. Jakllari, S. V. Krishnamurthy, M. Faloutsos, P. V. Krishnamurthy, and O. Ercetin, "A Cross-Layer Framework for Exploiting Virtual MISO Links in Mobile Ad Hoc Networks," *IEEE Trans. on Mobile Computing*, vol. 6, pp. 579–594, Jun. 2007.

[10] B. Aazhang, R. Blum, J. Laneman, K. Liu, W. Su, and A. Wittneben, "Special Issue on Cooperative Communications and Networking," *IEEE Journal. on Selected Areas in Communications*, vol. 25, Feb. 2007.

[11] A. Host-Madsen and J. Zhang, "Capacity Bounds and Power Allocation for Wireless Relay Channels," *IEEE Trans. on Information Theory*, vol. 51, pp. 2020–2040, Jun. 2005.

[12] G. Kramer, M. Gastpar, and P. Gupta, "Cooperative Strategies and Capacity Theorems for Relay Networks," *IEEE Trans. on Information Theory*, vol. 51, pp. 3037–3063, Sept. 2005.

[13] P. Liu, Z. Tao, S. Narayanan, T. Korakis, and S. S. Panwar, "CoopMAC: A Cooperative MAC for Wireless LANs," *IEEE Journal on Selecteld Areas in Communications*, vol. 25, pp. 340–354, Feb. 2007.

[14] S. Lakshmanan and R. Sivakumar, "Diversity Routing for Multi-hop Wireless Networks with Cooperative Transmissions," in *Proc. of IEEE Communications Society Conf. on Sensor, Mesh and Ad hoc Communications and Networks (SECON)*, (Rome, Italy), Jun. 2009.

[15] M. Kurth, A. Zubow, and J.-P. Redlich, "Cooperative Opportunistic Routing Using Transmit Diversity in Wireless Mesh Networks," in *Proc. of IEEE Intl. Conf. on Computer Communications (INFOCOM)*, Apr. 2008.

[16] J. Zhang and Q. Zhang, "Cooperative Routing in Multi-Source Multi-Destination Multi-hop Wireless Networks," in *Proc. of IEEE Intl. Conf. on Computer Communications (INFOCOM)*, Apr. 2008.

[17] A. Khandani, J. Abounadi, E. Modiano, and L. Zheng, "Cooperative Routing in Static Wireless Networks," *IEEE Trans. on Communications*, Nov. 2007.

[18] Y. Shi, S. Sharma, Y. T. Hou, and S. Kompella, "Optimal relay assignment for cooperative communications," in *Proc. ACM Intern. Symp. on Mobile Ad Hoc Networking and Computing (MobiHoc)*, (Hong Kong SAR, China), pp. 3–12, May 2002.

[19] E. M. Yeh and R. A. Berry, "Throughput Optimal Control of Cooperative Relay Networks," *IEEE Trans. on Information Theory*, vol. 53, pp. 3827–3833, Oct. 2007.

[20] I. E. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*. Philadelphia, PA, USA: SIAM, 1994.

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

[22] X. Zhu, E. Setton, and B. Girod, "Congestion-Distortion Optimized Video Transmission Over Ad Hoc Networks," *EURASIP Signal Processing: Image Communication*, pp. 773–783, Sept. 2005.

[23] D. Bertsekas and R. Gallager, *Data Networks*. USA: Prentice Hall, 2000.

[24] S. Boyd and J. Mattingley, "Branch and Bound Methods," *Notes for EE364b, Stanford University*, Mar. 2007.