QoS Driven Multi-user Video Streaming in Cellular CRNs The Case of Multiple Channel Access

Zhifeng He and Shiwen Mao
Department of Electrical and Computer Engineering
Auburn University, Auburn, AL 36849-5201 USA
Email: zzh0008@tigermail.auburn.edu, smao@ieee.org

***Abstract*—We tackle the challenging problem of streaming multi-user videos over the downlink of a cellular Cognitive Radio Network (CRN), where each Cognitive User (CU) can sense and access multiple channels at a time. Spectrum sensing, channel assignment, and power allocation strategies are jointly optimized to maximize the Quality of Service (QoS) for the CUs. We show that the formulated Mixed Integer NonLinear Programming (MINLP) problem can be decomposed into two sub-problems without sacrificing optimality: SP1 for the optimal spectrum sensing strategy, and SP2 for optimal channel assignment and power allocation. We show that SP1 can be optimally solved and then develop a Column Generation (CG) based algorithm to solve SP2 iteratively in a distributed manner. We also develop a heuristic algorithm for spectrum sensing with greatly reduced requirement on CU hardware, but with a highly competitive sensing performance. We analyze the proposed algorithms with respect to complexity and derive a performance upper bound. The proposed algorithms are validated with extensive simulations.**

***Keywords*-Quality of Service; cognitive radio networks; column generation; multi-user video streaming; optimization.**

I. Introduction

The Cognitive Radio (CR) technology has been widely recognized as an effective solution for efficient and flexible access to the radio spectrum. In a CR network (CRN), cognitive users (CU) detect primary user (PU) activities on licensed channels, and access the channels that are sensed idle opportunistically and unobtrusively [1]. Bandwidth-demanding and elastic mobile services, such as wireless video, will benefit enormously from this new wireless networking paradigm [2].

Although with great potential, the problem of video over CRNs brings about a whole level of technical challenges, particularly due to the extra dimension of dynamics on channel availability and the uncertainty from spectrum sensing and access. The manifold design trade-offs, multifarious network dynamics, limited network resources and, on the other hand, video's stringent QoS constraints, necessitate a holistic cross-layer design approach to "squeeze" the most out of the CRN. Usually such cross-layer design results in a tremendously complex global optimization problem, where all the layers (i.e., the PHY, MAC, network, and application layers) and all the users (i.e., PUs and CUs) are tightly coupled [3]. A *separation principle* that helps to decouple the design of spectrum sensing, access, and application QoS provisioning would be crucial for making the problem manageable [4].

In this paper, we tackle the problem of downlink multi-user video streaming in a cellular CRN, where each CU receives a video stream from the Cognitive Base Station (CBS). We consider the case that each CU is able to sense (with multiple sensors [5]) and access (with channel bonding/aggregation [6]) multiple channels. We adopt the H.264 Scalable Video Coding (SVC) (Quality Scalability) model from [7], and jointly design spectrum sensing, channel access, and power control for maximizing the QoS of all the CUs. There are two tightly coupled parts in this problem: the spectrum sensing problem (SP1) to determine which CU to sense which channel; and the channel assignment and power allocation problem (SP2) to allocate channels and transmit power to the CUs.

The formulated problem turns out to be a Mixed Integer NonLinear Programming (MINLP) problem, which is NP-hard in general. However, as in [4], where a *separation principle* is established to decouple the design of sensing strategy from that of sensor and access policy, we show that our problem can also be decoupled into two relatively easier sub-problems with a *decomposition principle* and develop an effective Column Generation (CG) based solution algorithm [8].

The major contributions made in this paper include:

1) A holistic *problem formulation* that jointly optimizes the spectrum sensing, channel assignment, and power allocation strategies for maximizing CU QoS.
2) A *decomposition principle* to decouple the original problem into a sensing strategy optimization problem SP1 and a resource allocation problem SP2, without sacrificing optimality, and *effective algorithms* to solve SP1 and SP2, respectively.
3) A *heuristic sensing scheme* that is much less demanding on CU hardware than the optimal sensing strategy, but can achieve highly competitive sensing performance.
4) An *upper bound* for the performance of the CG-based distributed algorithm and an analysis of complexity.
5) *Simulation validation* to demonstrate the superior per-

978-1-4799-6036-1/14 $31.00 © 2014 IEEE
DOI 10.1109/MASS.2014.111

formance of our proposed algorithms in terms of sensing performance and the QoS achieved by CUs.

The remainder of this paper is organized as follows. Section II reviews related work. The system model and problem formulation are presented in Section III, while the decomposition principle and the two sub-problems SP1 and SP2 are presented in Section IV. The CG-based distributed algorithm to solve SP2 is developed in Section V and analyzed in Section VI. The performance validation is provided in Section VII, and Section VIII concludes the paper.

II. Related Work

CR research has been largely focused on the aspects of spectrum sensing and dynamic spectrum access. In [9], the authors study the sensing-throughput tradeoff problem that optimizes the spectrum sensing time so that the CU's throughput can be maximized with restricted interference to the PUs. Unlike [9], the protocol proposed in [4] also considers the problem of which channel to sense, in addition to sensing parameters and access strategy optimization. Moreover, the design of sensing strategy is independent to sensing parameters design and the access strategy, as specified in a *principle of separation* [4]. These works focus on the optimization of sensing parameters only, and there is no collaboration between CUs. Considering the fact that different CUs may have different spectrum sensing performance, the authors in [10] propose an algorithm where groups of CUs are formed for cooperative sensing, aiming to find the best grouping scheme to discover most idle channels. Furthermore, the problem of sensing parameter optimization in addition to optimal sensor selection is addressed in [11], in order to achieve a trade-off between detection performance and sensing overhead.

Recently, cross layer design for video streaming over CRNs has attracted considerable interest. An auction game model is proposed in [12] to solve the problem of spectrum allocation in delay-sensitive content-aware multimedia delivering. Channel/path selection for multi-user video streaming is formulated as an MINLP problem in [3] to maximize the received video quality while restricting collisions with PUs. Packet scheduling is studied in [6] in which spectrum sensing at the PHY is integrated with packet scheduling at the MAC layer to improve delay-QoS provisioning over CRNs. The authors also analyze the throughput and delay performance with a Markov chain model and the $M/G_Y/1$ queuing model. Beyond these, other cross layer factors such as Fine Grained Scalability (FGS) coding, error control, and modulation, are jointly considered in [2] to achieve the maximum QoS for CUs in a cellular CRN. Interestingly, cross layer optimization of streaming videos over a CR link can also be modeled as a POMDP (Partially Observable Markov Decision Process) as in [13], in which intra refreshing rate, a video codec parameter, along with spectrum sensing and access strategies are jointly designed.

This paper is motivated by these interesting prior works, and is mainly focused on the joint design of spectrum sensing and resource allocation strategies for streaming multi-user videos over the downlink of a cellular CRN, where a novel decomposition principle is developed. The problem is not well addressed in prior work but is essential for supporting the demand of large bandwidth for video applications and enhance the QoS of CUs.

III. System Model and Problem Statement

A. System Model

We consider a primary network operating on N_1 licensed orthogonal channels, while each channel j has bandwidth B_j. A CR network is co-located with the primary network, consisting of a CBS and M CUs. The CUs sense the PU activities on the licensed channels and access the channels in an opportunistic manner. As in prior work [5], we first assume that each CU is equipped with N_1 sensors so that it can sense all the channels simultaneously. This assumption is relaxed in Section IV-B, where each CU can only sense a few channels at a time. The CBS determines the status of the licensed channels based on the sensing results reported from the CUs.

We consider the scenario of downlink multi-user video streaming, where the CBS transmits different video streams to the CUs using the channels sensed idle. Once the channel states are estimated, the CBS and CUs determine the allocation of the idle channels, and the CBS selects a power level k, $k = 1, \cdots, K$, for the video transmission to a CU on each allocated channel. We assume CU i can receive from the CBS simultaneously on at most C_i channels: when $C_i = 1$, it is the traditional single channel access mode; when $C_i > 1$, it is the case when channel bonding/aggregation is used [6]. Without loss of generality, we assume time is divided into a series of non-overlapping Group of Pictures (GOP) windows, each consisting of $\mathbb{T}$ time slots.

B. Problem Statement

Due to multipath fading and shadowing, different CUs usually experience different SNRs when detecting a PU signal, and thus may have different sensing performance. It is important to choose a suitable set of CUs to sense a licensed channel [10]. To achieve diversity gain, cooperative sensing is usually used to improve the detection performance by fusing the sensing results from multiple CUs, where a certain fusion rule is used to combine the CU sensing results. In this paper, we consider cooperative sensing with the OR fusion rule: if any of the CUs reports the presence of a PU signal on a channel, the CBS will determine that the channel is busy; otherwise, the channel is considered to be idle.

We use an $M \times N_1$ matrix $\mathbf{X}$ to represent the assignment of sensing tasks, where each element x_{ij} is defined as

$$x_{ij} = \begin{cases} 1, & \text{CU } i \text{ is assigned to sense channel } j \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

For energy detection, the probability of detection of PU signal on channel j by CU i, $P_{d_{ij}}$, and the probability of false alarm on channel j by CU i, $P_{f_{ij}}$, can be expressed as [9]

$$P_{d_{ij}} = \frac{1}{2}\text{erfc}\left(\left(\frac{\lambda_{ij}}{n_0 B_j} - \gamma_{ij} - 1\right)\sqrt{\frac{\kappa}{2\left(2\gamma_{ij}+1\right)}}\right) \quad (2)$$

$$P_{f_{ij}} = \frac{1}{2}\text{erfc}\left(\sqrt{2\gamma_{ij}+1}\ \text{erfc}^{-1}\left(2P_{d_{ij}}\right) + \sqrt{\frac{\kappa}{2}}\gamma_{ij}\right), \quad (3)$$

where n_0 is the power density of Additive White Gaussian Noise (AWGN) that is i.i.d. at each CU, γ_{ij} is the SNR of PU signal on channel j at CU i's side, λ_{ij} is the energy detection threshold on channel j by CU i, and κ is the number of samples collected. For cooperative sensing with the OR fusion rule, the probability of detection of PU signal on channel j, P_{d_j}, and the probability of false alarm on channel j, P_{f_j}, can be written as

$$P_{d_j} = 1 - \prod_{i=1}^{M}(1 - P_{d_{ij}})^{x_{ij}} \quad (4)$$

$$P_{f_j} = 1 - \prod_{i=1}^{M}(1 - P_{f_{ij}})^{x_{ij}}. \quad (5)$$

To provide a graceful protection to PUs, we set $P_{d_j} = P_{req}$, where P_{req} is the maximum interference from the CU system that can be tolerated by the PU system, and $\sum_{i=1}^{M} x_{ij} = \Lambda_j$, for all $j = 1, 2, \cdots, N_1$, where Λ_j is the minimum number of CUs to sense channel j. If we set $P_{d_{ij}} = \bar{P}_{d_j} = 1 - (1 - P_{req})^{1/\Lambda_j}$, it follows (4) that $P_{d_j} = P_{req}$.

Let r_j denote the real state of channel j: $r_j = 0$ when channel j is idle, and $r_j = 1$ otherwise. Also let s_j be the cooperative sensing result on channel j: $s_j = 0$ if the channel is determined idle, and $s_j = 1$ otherwise. We have

$$P(s_j{=}0) = P(r_j{=}0)(1 - P_{f_j}) + P(r_j{=}1)(1 - P_{d_j}) \quad (6)$$

$$P(s_j{=}1) = P(r_j{=}0)P_{f_j} + P(r_j{=}1)P_{d_j}. \quad (7)$$

The cooperative sensing results on the N_1 channels can be represented as $\vec{\mathbf{S}} = \{s_j, j = 1, \cdots, N_1\}$. There are 2^{N_1} possible outcomes for $\vec{\mathbf{S}}$, and let $\vec{\mathbf{S}}_h$ be the h-th outcome, $0 \le h \le 2^{N_1} - 1$. To determine the j-th element in $\vec{\mathbf{S}}_h$, let $s_j = \Gamma_j(h), j = 1, 2, \cdots, N_1$, denote the relationship between $\vec{\mathbf{S}}_h$ and s_j. Assuming independent channel states, the probability of getting outcome $\vec{\mathbf{S}}_h$ can be written as

$$\begin{aligned} & P(\vec{\mathbf{S}} = \vec{\mathbf{S}}_h) \\ = & \prod_{j=1}^{N_1} P(s_j = \Gamma_j(h)) \\ = & \prod_{j=1}^{N_1} \left[(1 - \Gamma_j(h))P(s_j = 0) + \Gamma_j(h)P(s_j = 1)\right]. \end{aligned} \quad (8)$$

We adopt the QoS model for H.264 SVC (Quality Scalability) from [7] as

$$\rho_i = \alpha_i + \beta_i \cdot R_i, \quad (9)$$

where ρ_i is the Y-PSNR (Peak Signal-to-Noise Ratio) of the received video at CU i, α_i and β_i are constants dependent on the content type of the video sequence, and R_i is the effective data rate of the video sequence. According to conditional expectation, the expected overall QoS can be derived as

$$\begin{aligned} & \mathrm{E}\left(\sum_{i=1}^{M} \rho_i\right) \\ = & \sum_{i=1}^{M} \sum_{h=0}^{2^{N_1}-1} \mathrm{E}(\rho_i | \vec{\mathbf{S}} = \vec{\mathbf{S}}_h) P(\vec{\mathbf{S}} = \vec{\mathbf{S}}_h) \\ = & \sum_{h=0}^{2^{N_1}-1} \sum_{i=1}^{M} \mathrm{E}(\rho_i | \vec{\mathbf{S}} = \vec{\mathbf{S}}_h) P(\vec{\mathbf{S}} = \vec{\mathbf{S}}_h). \end{aligned} \quad (10)$$

For $\vec{\mathbf{S}} = \vec{\mathbf{S}}_h$, let $\Phi_h = \{j : \Gamma_j(h) = 0, j = 1, 2, \cdots, N_1\}$ be the set of channels sensed idle. Let $\mathbf{Y}_h = \left[y_{ijk}^h\right], 1 \le i \le M$, $j \in \Phi_h$, $1 \le k \le K$, $0 \le h \le 2^{N_1 - 1}$ be the channel assignment and power allocation matrix, where $0 \le y_{ijk}^h \le 1$ is the amount of time that CBS transmits to CU i with a power level of k on channel j in a time slot, when the sensing outcome is $\vec{\mathbf{S}}_h$. The channel assignment and power allocation strategy can be expressed as $\mathbf{Y} = [\mathbf{Y}_0, \mathbf{Y}_1, \cdots \mathbf{Y}_{2^{N_1 - 1}}]$.

Putting it all together, it follows that

$$\begin{aligned} & \mathrm{E}(\rho_i | \vec{\mathbf{S}} = \vec{\mathbf{S}}_h) \\ = & \alpha_i + \beta_i \cdot \mathrm{E}(R_i | \vec{\mathbf{S}} = \vec{\mathbf{S}}_h) \\ = & \alpha_i + \beta_i \cdot \mathrm{E}\left(\sum_{j \in \Phi_h} \sum_{k=1}^{K} R_{ijk} \cdot y_{ijk}^h | \vec{\mathbf{S}} = \vec{\mathbf{S}}_h\right) \\ = & \alpha_i + \beta_i \sum_{j \in \Phi_h} \sum_{k=1}^{K} \left(P_{00}^j R_{00}^{ijk} + P_{10}^j R_{10}^{ijk}\right) \cdot y_{ijk}^h \end{aligned} \quad (11)$$

where G_k is the power of level k, d_{ij} is the channel gain between the CBS and CU i on channel j, and

$$\begin{aligned} P_{00}^j &= P(r_j = 0 | s_j = 0) = \\ & \frac{(1 - P_{f_j})P(r_j = 0)}{(1 - P_{f_j})P(r_j = 0) + (1 - P_{d_j})P(r_j = 1)} \\ P_{10}^j &= P(r_j = 1 | s_j = 0) = 1 - P_{00}^j \\ R_{00}^{ijk} &= B_j \log_2(1 + G_k d_{ij} / (n_0 B_j)) \\ R_{10}^{ijk} &= B_j \log_2(1 + G_k d_{ij} / (n_0 B_j (1 + \gamma_{ij}))). \end{aligned}$$

Define $w_{ijk} = \alpha_i + \beta_i B_j (P_{00}^j R_{00}^{ijk} + P_{10}^j R_{10}^{ijk})$. The master problem of maximizing the total expected QoS, denoted as

P0, can be formulated as follows.

$$\underline{\textbf{P0}} : \max : \sum_{h=0}^{2^{N_1}-1} \sum_{i=1}^{M} \sum_{j \in \Phi_h} \sum_{k=1}^{K} w_{ijk} \cdot y_{ijk}^h \cdot P(\vec{\mathbf{S}} = \vec{\mathbf{S}}_h) \tag{12}$$

$$\text{s.t.} \sum_{j \in \Phi_h} \sum_{k=1}^{K} y_{ijk}^h \leq C_i, \forall\, i, h \tag{13}$$

$$\sum_{i=1}^{M} \sum_{k=1}^{K} y_{ijk}^h \leq 1, \forall\, j, h \tag{14}$$

$$\sum_{i=1}^{M} \sum_{j \in \Phi_h} \sum_{k=1}^{K} y_{ijk}^h \cdot G_k \leq G_{total}, \forall\, h \tag{15}$$

$$\sum_{i=1}^{M} x_{ij} = \Lambda_j, \forall\, j \tag{16}$$

$$x_{ij} = \{0, 1\}, \forall\, i, j \tag{17}$$

$$y_{ijk}^h \begin{cases} \in [0,1], & \text{if } G_k d_{ij}/(n_0 B_j) \geq \bar{\gamma} \\ = 0, & \text{otherwise,} \end{cases} \quad \forall\, i, j, k, h. \tag{18}$$

In each time slot, constraint (13) enforces that CR i can access at most C_i channels simultaneously; constraint (14) enforces that the transmission time on each channel is within 1 time slot; constraint (15) enforces that the total transmission power of the CBS must not exceed the peak power limit G_{total}; constraint (16) enforces that there are Λ_j CUs to sense each channel j; and constraint (18) enforces that the necessary condition for the CBS to transmit to CU i on channel j with power level k is that the resulting SNR must be greater than a predefined threshold $\bar{\gamma}$ such that CU i can successfully decode the received video.

Note that constraint (18) indicates $y_{ijk}^h \geq 0$. Combined with constraint (14), it follows that each $0 \leq y_{ijk}^h \leq 1$. Therefore, constraint (18) can be rewritten as

$$y_{ijk}^h \begin{cases} \geq 0, & \text{if } G_k d_{ij}/(n_0 B_j) \geq \bar{\gamma} \\ = 0, & \text{otherwise,} \end{cases} \quad \forall\, i, j, k, h. \tag{19}$$

The upper bound of 1 on the y_{ijk}^h's is thus removed and the problem can be solved more efficiently, since usually LP solvers solve an LP without upper bounds faster than LPs with upper bounds.

IV. Problem Decomposition

A. Optimal Sensing Strategy for Problem P0

Problem P0 is an MINLP, which is NP-hard. However, we observe that the optimal sensing strategy can be obtained by solving a relatively easier problem as follows. We first introduce Lemma 1 as a basis for later analysis.

Lemma 1. *The objective value of P0 is a decreasing function of P_{f_j}, for all $j = 1, 2, \cdots, N_1$.*

Proof: Define $F(\mathbf{X}, \mathbf{Y}_h) = \sum_{i=1}^{M} \sum_{j \in \Phi_h} \sum_{k=1}^{K} w_{ijk} \cdot y_{ijk}^h \cdot P(\vec{\mathbf{S}} = \vec{\mathbf{S}}_h)$ and $f(\mathbf{X}, \mathbf{Y}) = \sum_{h=0}^{2^{N_1}-1} F(\mathbf{X}, \mathbf{Y}_h)$. The partial derivative of $F(\mathbf{X}, \mathbf{Y}_h)$ with respect to P_{f_j} is

$$\frac{\partial F(\mathbf{X}, \mathbf{Y}_h)}{\partial P_{f_j}} = -y_{ijk}^h \cdot P(\vec{\mathbf{S}} = \vec{\mathbf{S}}_h) \cdot \sum_{i=1}^{M} \sum_{k=1}^{K} \Big(\beta_i \cdot R_{00}^{ijk} \cdot ((1-P_{req}) P(r_j = 0) P(r_j = 1) + (1-P_{f_j}) P(r_j = 0)^2) + \alpha_i P(r_j = 0)(R_{00}^{ijk} P_{00}^j + R_{10}^{ijk} P_{10}^j)/B_j \Big) \leq 0.$$

It follows that $\frac{\partial f(\mathbf{X},\mathbf{Y})}{\partial P_{f_j}} = \sum_{h=0}^{2^{N_1}-1} \frac{\partial F(\mathbf{X},\mathbf{Y}_h)}{\partial P_{f_j}} \leq 0$. ■

Theorem 1. *The optimal spectrum sensing strategy to problem P0 can be obtained by solving the following problem SP1.*

$$\underline{\textbf{SP1}} : \forall\, j = 1, 2, \cdots, N_1$$

$$\min : P_{f_j} = 1 - \prod_{i=1}^{M} (1 - P_{f_{ij}})^{x_{ij}} \tag{20}$$

$$\text{s.t.} \sum_{i=1}^{M} x_{ij} = \Lambda_j. \tag{21}$$

Proof: Let the optimal solution to the original problem P0 be $(\mathbf{X}', \mathbf{Y}')$ and the solution to SP1 be $\mathbf{X}^*$. Since $\mathbf{X}'$ is optimal to the maximization problem P0, we have $f(\mathbf{X}^*, \mathbf{Y}') \leq \mathbf{f}(\mathbf{X}', \mathbf{Y}')$. On the other hand, since $\mathbf{X}^*$ is optimal to the minimization problem SP1, it follows Lemma 1 that $f(\mathbf{X}^*, \mathbf{Y}') \geq \mathbf{f}(\mathbf{X}', \mathbf{Y}')$. Therefore we conclude that $f(\mathbf{X}^*, \mathbf{Y}') = \mathbf{f}(\mathbf{X}', \mathbf{Y}')$ and $\mathbf{X}^*$ is optimal to problem P0.

After obtaining $\mathbf{X}^*$, we substitute it into problem P0 to solve for $\mathbf{Y}^*$. The $\mathbf{Y}^*$ obtained this way is also optimal to P0, i.e., we have $f(\mathbf{X}^*, \mathbf{Y}^*) \geq \mathbf{f}(\mathbf{X}^*, \mathbf{Y}) \geq \mathbf{f}(\mathbf{X}, \mathbf{Y})$, for all $\mathbf{X}, \mathbf{Y}$. The proof is completed. ■

Problem SP1 can be rewritten as the following problem SP1a,

$$\underline{\textbf{SP1a}} : \forall\, j = 1, 2, \cdots, N_1,$$

$$\max : \sum_{i=1}^{M} x_{ij} \cdot \log_2(1 - P_{f_{ij}}) \tag{22}$$

$$\text{s.t.} \sum_{i=1}^{M} x_{ij} = \Lambda_j, \tag{23}$$

which can be solved easily with an LP solver.

B. Heuristic Spectrum Sensing Algorithm

In the optimal sensing strategy, a CU may be required to sense many or even all the channels, which may be time consuming and hard-ware demanding. In this section, we introduce a heuristic algorithm where each CU only needs to sense a part of all the channels at a time slot thus the CU sensing overhead and hard-ware requirement is reduced.

Assume that CU i can sense at most $\Theta_i \ll N_1$ channels in a time slot. The idea is to sort the N_1 channels according to

Algorithm 1: Heuristic Spectrum Sensing Algorithm

```
1  Sort the N_1 channels in descending order of P(r_j = 0) and
   let the sorted channel set be Ξ ;
2  for j = 1 : N_1 do
3      Let j' = Ξ(j) ;
4      Solve problem SP1b and denote the solution as Θ_{j'} ;
5      if Θ_{j'} = ∅ then
6          Channel j' is determined to be busy ;
7      end
8      if x_{ij'} = 1 then
9          Θ_i = Θ_i − 1 ;
10     end
11 end
```

$P(r_j = 0)$, for all $j = 1, 2, \ldots, N_1$, in the descending order, and then minimize P_{f_j}, for all $j = 1, \ldots, N_1$, sequentially. The heuristic spectrum sensing algorithm is presented in Algorithm 1. In Line 4, the following problem SP1b is solved.

$$\underline{\textbf{SP1b}} : \min : P_{f_{j'}} = 1 - \prod_{i=1}^{M} (1 - P_{f_{ij'}})^{x_{ij'}} \tag{24}$$

$$\text{s.t.} \sum_{i=1}^{M} x_{ij'} = \Lambda_{j'} \tag{25}$$

$$x_{ij'} \leq \Theta_i. \tag{26}$$

In Lines 5∼7, if there is no feasible solution to problem SP1b, there is not a sufficient number of CUs to sense channel j', and we conservatively assume that channel j' is busy to avoid collision with PUs. Each time if CU i is assigned to sense a channel, Θ_i is decreased by 1 as in Lines 8∼10. When Θ_i reaches 0, constraint (26) enforces that CU i senses Θ_i channels at most.

In Section VII, we will show that the performance of Algorithm 1 is very close to that of the optimal sensing strategy in terms of both sensing performance and the expected overall QoS, even when $\Theta_i \ll N_1$.

C. Optimal Chanel Assignment and Power Allocation Solution

After obtaining $\mathbf{X}^*$, cooperative sensing is conducted and the CBS determines the set of available channels based on sensing results. From now on, we omit the subscript (or superscript) h in all the symbols, since the cooperative sensing results on the N_1 channels is already determined. Denote the number of channels sensed idle as N_2, and re-index the N_2 idle channels as $1, 2, \cdots, N_2$. Then the remaining channel assignment and power allocation problem SP2 can be written as follows.

$$\underline{\textbf{SP2}} : \max : \sum_{i=1}^{M} \sum_{j=1}^{N_2} \sum_{k=1}^{K} w_{ijk} \cdot y_{ijk} \tag{27}$$

$$\text{s.t.} \sum_{j=1}^{N_2} \sum_{k=1}^{K} y_{ijk} \leq C_i, \forall\, i \tag{28}$$

$$\sum_{i=1}^{M} \sum_{k=1}^{K} y_{ijk} \leq 1, \forall\, j \tag{29}$$

$$\sum_{i=1}^{M} \sum_{j=1}^{N_2} \sum_{k=1}^{K} y_{ijk} \cdot G_k \leq G_{total} \tag{30}$$

Constraint (19).

In practice, there may be a large number of CUs and licensed channels, and the CBS also has a great flexibility to choose the power level for transmission on a channel. Therefore, the constraint matrix of SP2 could be huge and it may be hard to solve with an LP solver due to its size. In the next section, we propose to use the Column Generation (CG) method [8] to solve SP2 and derive a decentralized algorithm for better scalability. With the proposed CG method, the CBS and CUs solve different sub-problems, thus alleviating the computational burden on the CBS.

V. Column Generation Method to Solve SP2

A. Dantzig-Wolfe Decomposition

We first reformulate problem SP2 from the standard form into a *disaggregated formulation* by applying Dantzig-Wolfe decomposition of LP problems [14].

For $i = 1, 2, \cdots, M$, let $\overline{\Theta}_i = \left\{\chi_i^1, \chi_i^2, \cdots, \chi_i^{Q_i}\right\}$ denote the set of feasible channel assignment and power allocation schemes to CU i. Then

$$\chi_i^q = \left\{y_{ijk}^q, j = 1, 2, \cdots, N_2, k = 1, 2, \cdots, K\right\},$$

for $q = 1, 2, \cdots, Q_i$, is a feasible scheme satisfying all the constraints, where $y_{ijk}^q = 1$ if the CBS transmits to CU i on channel j at power level k, and $y_{ijk}^q = 0$ otherwise. Thus, the feasible schemes are indeed the extreme points of the the feasible region of SP2, which is the key for Dantzig-Wolfe decomposition [14].

Introduce a variable $0 \leq z_i^q \leq 1$ to denote the amount of time the CBS transmits using feasible scheme χ_i^q within a time slot. Let the "utility" gained by using χ_i^q for CU i as

$$\varpi_i^q = \sum_{j=1}^{N_2} \sum_{k=1}^{K} w_{ijk} \cdot y_{ijk}^q.$$

Then SP2 can be represented in a *set-partition form*, termed the Master Problem (MP), as

$$\underline{\textbf{MP}}: \max: \sum_{i=1}^{M}\sum_{q=1}^{Q_i} \varpi_i^q \cdot z_i^q \tag{31}$$

$$\text{s.t.} \sum_{q=1}^{Q_i} z_i^q \leq 1, \forall\, i, \tag{32}$$

$$\sum_{i=1}^{M}\sum_{q=1}^{Q_i}\left(\sum_{k=1}^{K} y_{ijk}^q\right) z_i^q \leq 1, \forall\, j, \tag{33}$$

$$\sum_{i=1}^{M}\sum_{q=1}^{Q_i}\left(\sum_{j=1}^{N_2}\sum_{k=1}^{K} y_{ijk}^q \cdot G_k\right) z_i^q \leq G_{total}, \tag{34}$$

$$0 \leq z_i^q \leq 1, \forall\, i, q. \tag{35}$$

Constraint (32) ensures that $0 \leq y_{ijk} = \sum_{q=1}^{Q_i} y_{ijk}^q z_i^q \leq 1$, for all i, j, k; constraints (33) and (34) correspond to constraints (29) and (30), respectively; and constraints (28) and (19) are specified in the INitialization Problem (INP) and Pricing Problem (PP) defined next in Section V-B. For convenience of our later discussion, the problem containing a subset of the columns and cost coefficients (variables) of the MP is called Restricted MP (RMP).

B. Design of the Column Generation Method

Obviously, it is infeasible to solve the MP directly due to the exponential number of columns. However, usually most of the variables in the optimal solution to the MP are equal to zero, with only a small number of positive-valued variables. The MP solution can be re-optimized iteratively by finding the variables having the potential to improve the objective value at each iteration. This is done by iteratively solving the PP, which examines whether there exists a variable with a negative (in the case of a minimization problem) or positive (in the case of a maximization problem) reduced cost, and then generates the corresponding column to add it to the RMP.

The RMP contains only a small subset of all the feasible columns and variables of the MP and thus can be solved quickly. The simplex multiplier obtained from the RMP will be passed to the PP to identify a new column to enter the RMP again, until there is no variables whose reduced cost is negative (in the case of a minimization problem) or positive (in the case of a maximization problem). Thus an optimal feasible solution to the MP is found. The purpose of RMP is to generate the simplex multiplier for solving the PP.

The CG based Distributed Optimization Algorithm (CDOA) includes the following six steps.

Algorithm 2: CG Based Distributed Optimization Algorithm

Step 1: CU i solves the following i-th INP and reports its solution to the CBS, $i = 1, 2, \cdots, M$.

$$\underline{\textbf{INP}}: \max: \sum_{j=1}^{N_2}\sum_{k=1}^{K} w_{ijk} \cdot y_{ijk} \tag{36}$$

$$\text{s.t.} \sum_{j=1}^{N_2}\sum_{k=1}^{K} y_{ijk} \leq C_i, \tag{37}$$

$$\sum_{k=1}^{K} y_{ijk} \leq 1, \forall\, j, \tag{38}$$

$$y_{ijk} \begin{cases} \in \{0,1\}, & \text{if } G_k d_{ij}/(n_0 B_j) \geq \bar{\gamma} \\ = 0, & \text{otherwise}, \end{cases} \quad \forall\, j, k. \tag{39}$$

Each of the M solutions generates a feasible column of the MP. The CBS uses the M feasible columns and the corresponding cost coefficients to initiate the RMP, which has the same formulation with the MP, but with $Q_i = 1$, for all $i = 1, 2, \cdots, M$.

Step 2: The CBS solves the RMP, from which a vector of simplex multiplier $\Omega^T = (\nu^T, \mu^T, \varphi)$ is obtained, where $(\cdot)^T$ denotes the transpose of a vector, ν^T is a $1 \times M$ vector with the i-th entry ν_i corresponding to the i-th constraint in the RMP, μ^T is a $1 \times N_2$ vector with the j-th entry μ_j corresponding to the $(M+j)$-th constraint in the RMP, and φ is the simplex multiplier corresponding to the last constraint in the RMP. The objective value of the RMP is a *lower bound* to the MP.

Step 3: The CBS broadcasts Ω^T to all CUs and assigns CU i to solve the following i-th PP, to find the column and the corresponding variable with the most positive reduced cost [8] to enter the RMP to improve the objective value of the MP.

$$\underline{\textbf{PP}}: \max: \Delta_i = \sum_{j=1}^{N_2}\sum_{k=1}^{K} (w_{ijk} - \mu^j - \varphi G_k) \cdot y_{ijk} - \nu_i \tag{40}$$

$$\text{s.t.} \sum_{j=1}^{N_2}\sum_{k=1}^{K} y_{ijk} \leq C_i, \tag{41}$$

$$\sum_{k=1}^{K} y_{ijk} \leq 1, \forall\, j, \tag{42}$$

Constraint (39).

Step 4: Each CU i decides when to report its optimal solution to the i-th PP to the CBS according to a delay

$$\tau_i = \xi(\Delta_i), \tag{43}$$

where $\xi(\cdot)$ denotes a decreasing function of Δ_i. Define an index $a = \arg\max_{i=1,\cdots,M}\{\Delta_i\}$. In case that $\Delta_a > 0$, then the current optimal solution to the RMP is not optimal to the MP and CU a sends its solution in the earliest time τ_a (since it has the maximum value Δ_a). Other CUs overhearing CU

a's message will not send their respective messages. In case that $\Delta_a \leq 0$, the current optimal solution to the RMP is also optimal to the MP, and no CU sends message to CBS.

Step 5: The CBS verifies the optimality of the current solution: if nothing is received from the CUs after a predefined period of time, the CBS concludes that $\Delta_a \leq 0$ and thus the CG method is terminated; otherwise, go to Step 6.

Step 6: For index $a = \arg\max_{i=1,\cdots,M}\{\Delta_i\}$, let $Q_a = Q_a + 1$ and generate the column

$$H_a^{Q_a} = \left[e_a, \sum_{k=1}^{K} y_{a1k}^{Q_a}, \cdots, \sum_{k=1}^{K} y_{aN_2k}^{Q_a}, \sum_{j=1}^{N_2}\sum_{k=1}^{K} y_{ajk}^{Q_a} \cdot G_k \right]^T \quad (44)$$

with the solution to the a-th PP derived in Step 3, where e_a is a $1 \times M$ unit vector with the a-th entry being 1. Add the column and the corresponding variable $z_a^{Q_a}$ to the RMP and go to Step 2.

VI. Upper Bound and Complexity

A. Upper Bound for the MP

In the following, we derive an upper bound for the optimal objective value of the MP in each iteration of the CG method.

Theorem 2. *At each iteration, let Ω^T be the simplex multiplier vector of the RMP; $\Delta_a = \max_{i=1,\cdots,M}\{\Delta_i\}$ be the most positive reduced cost obtained from the PPs; $\vec{b}$ be a $(M+N_2+1)\times 1$ column vector with the i-th entry being the value of right hand side of the i-th constraint of the RMP, $i = 1, 2, \cdots, (M+N_2+1)$; $\vec{g}$ be a $(M+N_2+1)\times 1$ column vector as $g = (\underbrace{11\cdots1}_{M \text{ ones}} \underbrace{00\cdots0}_{(N_2+1) \text{ zeros}})^T$. Then an upper bound for the MP can be derived as: $\overline{\Omega}^T\vec{b} = (\Omega^T + \Delta_a\vec{g})\vec{b}$.*

Proof: Let $\widehat{\Omega}^T$ be a feasible solution to the dual problem of the MP (termed DMP), according to the relationship between the dual and primal problems [8], we have

$$\widehat{\Omega}^T H_i^q \geq \varpi_i^q, \ \ i = 1, \cdots, M, \ q = 1, \cdots, Q_i, \quad (45)$$

where H_i^q is given in (44). As discussed, at each iteration we can obtain a simplex multiplier vector Ω^T by solving the RMP, as well as the most positive reduced cost by solving the PP.

$$\Delta_a = \max_{i,q}\{\varpi_i^q - \Omega^T H_i^q\}$$
$$\Rightarrow \Omega^T H_i^q \geq \varpi_i^q - \Delta_a, i = 1, \cdots, M, q = 1, \cdots, Q_i, \quad (46)$$

where $\Delta_a > 0$. Denote $\overline{\Omega}^T = (\Omega^T + \Delta_a\vec{g})$ and multiply its both sides by H_i^q. We have

$$\overline{\Omega}^T H_i^q = \Omega^T H_i^q + \Delta_a\vec{g}H_i^q \Rightarrow \overline{\Omega}^T H_i^q = \Omega^T H_i^q + \Delta_a \cdot 1$$
$$\Rightarrow \overline{\Omega}^T H_i^q - \Delta_a = \Omega^T H_i^q \Rightarrow \overline{\Omega}^T H_i^q - \Delta_a \geq \varpi_i^q - \Delta_a$$
$$\Rightarrow \overline{\Omega}^T H_i^q \geq \varpi_i^q, \ i = 1, \cdots, M, \ q = 1, \cdots, Q_i.$$

The first inequality is from (46). This means that $\overline{\Omega}^T$ is a feasible solution to the DMP. By duality, the corresponding dual LP of a maximization LP is a minimization LP [8]. So the DMP is a minimization LP, and $\overline{\Omega}^T$ is a feasible solution to the DMP.

Let the optimal solution to DMP be $\underline{\Omega}^T$. It follows that

$$\overline{\Omega}^T\vec{b} \geq \underline{\Omega}^T\vec{b} = \Upsilon^*, \quad (47)$$

where Υ^* is the optimal objective value of the minimization LP DMP. Due to *Strong Duality*, Υ^* is also the optimal objective value of the MP. It follows that $\overline{\Omega}^T\vec{b} = (\Omega^T + \Delta_a\vec{g})\vec{b}$ is an upper bound for the MP according to (47). ■

B. Complexity Analysis

In the general problems solved by the CG method, the INP and PP problems are at least as hard as the one dimensional 0-1 Knapsack problem, which is NP-hard [17].

However, an interesting characteristic of the INP and PP in our case is that the coefficients of the constraint matrix in the INP and PP are either 0 or 1, such that the *unimodularity property* [16] is satisfied in both problems. As a result, both the INP and PP have the optimal solution with their LP relaxations, and thus they can be solved with the Simplex method [8], [17]. Again, the upper bound of 1 on y_{ijk} can be removed as in P0.

Lemma 2. *The INP and PP are indeed LPs and thus can be solved with the Simplex method with a polynomial-time average-case complexity.*

VII. Simulation Results and Analysis

In this section, Matlab simulation results are used to demonstrate the performance of the proposed algorithms. Unless specified, the value of simulation parameters are as shown in Table I. Each simulated point in the figures is obtained by repeating the simulation $r_p = 50$ times with different random seeds, and 95% confidence intervals are computed and plotted in the figures.

In Figs. 1 and 2 we compare the performance of the optimal sensing strategy and that of the Heuristic algorithm, in terms of sensing performance and the resulting overall Y-PSNR of the received videos, respectively, We consider four cases that the number of sensors a CU has in the Heuristic algorithm Θ_i = 3, 4, 5, and 6, where $\Theta_i \ll N_1 = 30$. Note that $M \cdot \Theta_i \geq N_1 \cdot \Lambda_j$ is a necessary condition to have any channel sensed by at least Λ_j sensors. In Fig. 1, the legend 'Idle channels' means the number of

Table I
SIMULATION PARAMETERS

Parameter	*Value*	*Parameter*	*Value*
M	30	$\bar{\gamma}$	-25 dB
N_1	30	Λ_j	3
K	10	r_p	50
κ	10^4	Θ_i	3
d_{ij}	$-15 \sim -9$ dB	γ_{ij}	$-100 \sim 0$ dB
C_i	3	G_{total}	50
G_k	$10^{-\frac{k}{10}}$	P_{req}	0.99
f_s	10^6 H_z	$\min_j \{P(r_j\!=\!0)\}$	0.2
n_0	10^{-6}	$\max_j \{P(r_j\!=\!0)\}$	0.9
B_j	10^6 H_z		

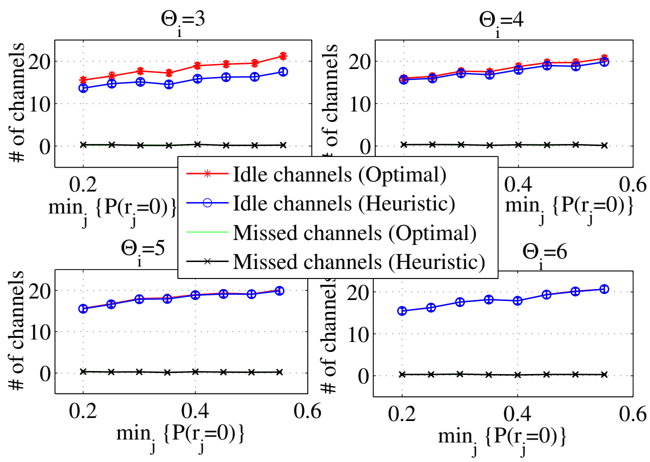

Figure 1. Optimal versus Heuristic Sensing in terms of sensing performance.

channels cooperatively sensed idle; ‘Missed channels’ means the number of channels cooperatively sensed idle while these channels are actually busy, so the number of channels that are sensed idle and are actually idle is the difference between the two.

From these two figures we can see that the heuristic algorithm achieves about 90% performance of the optimal sensing strategy when $\Theta_i = 3$. In this case, we have $M \cdot \Theta_i = N_1 \cdot \Lambda_j$. A channel j, may not have ‘good’ CUs (i.e., with a relative small $P_{f_{ij}}$) to sense it, since these CUs may have already been assigned to another channel j', which has a higher $P(r_{j'} = 0)$ and thus has a higher priority of being optimized. Thus channel j may be discarded due to insufficient CUs to sense it, resulting in a lower value of SP2 due the relationship between the value of SP2 and P_{f_j} as stated in Lemma (1). However, as the last three sub-figures shows, the Heuristic algorithm achieves almost the same performance as the optimal sensing strategy when Θ_i = 4, 5, and 6. Thus even the channels having a lower priority will have a higher chance to be sensed by ‘good’ CUs. Then P_{f_j}, $j = 1, 2, \ldots, N_2$ is more likely to be reduced, and the objective function value is improved.

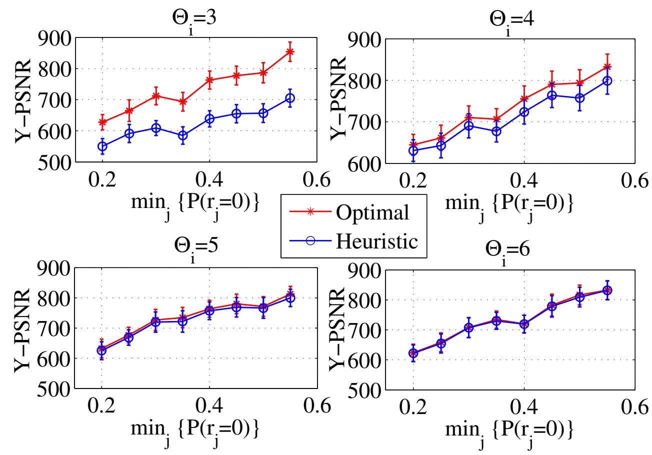

Figure 2. Optimal versus Heuristic Sensing in terms of overall Y-PSNR in dB.

Fig. 3 shows the convergence of the CG method in two cases: $M = N_1 = 30$ (the upper figure) and $M = N_1 = 60$ (the lower figure). We have the following observations. (i) The number of iterations is positively correlated to the problem size, since as the number of CUs and channels grows, there may be more feasible schemes to improves the current objective value at a specific iteration. (ii) The most positive reduced cost $\max_i \{\Delta_i\}$ tends to decrease over iterations. This trend is the result of the greedy approach of the CG algorithm, which means that the algorithm chooses the feasible scheme having the most positive reduced cost (thus possibly having the greatest potential to improve the current objective value of the MP) from the remaining candidate feasible schemes, to enter the RMP at each iteration. (iii) The increment of the objective function at a specific iteration is positive correlated to $\max_i \{\Delta_i\}$. This follows directly from the above discussions. (iv) The upper bound to the optimal objective function value converges quickly, and is also positively correlated to $\max_i \{\Delta_i\}$. From Theorem 2, the upper bound at a certain iteration $\overline{\Omega}^T b$ is actually a positive function of $\max_i \{\Delta_i\}$ (Δ_a in Theorem 2) at this iteration. Since $\max_i \{\Delta_i\}$ drops quickly and become very close to 0, $\overline{\Omega}^T b$ drops and then converges to the optimal objective function value quickly. Note that $\overline{\Omega}^T b$ is not necessary decreasing as the iteration goes. There are two main reasons: (a) $\max_i \{\Delta_i\}$ is not necessary decreasing as the iteration goes although it shows a trend of decreasing. (b) $\overline{\Omega}^T b$ also depends on the simplex multiplier Ω_T, whose convergence is hard to analysis.

VIII. CONCLUSION

In this paper, we investigated the problem of QoS-driven multi-user video streaming over cellular CRNs. We showed that there exists a *decomposition principle* in the optimal joint design of spectrum sensing, channel assignment, and

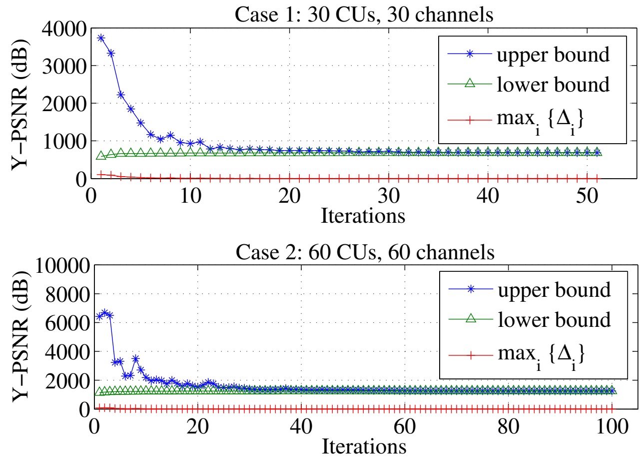

Figure 3. Convergence performance of the CG method for two cases: $M = N_1 = 30$ and $M = N_1 = 60$.

power allocation that circumvents the *curse of dimensionality* in general MINLPs. The decomposed spectrum sensing problem was solved with an optimal algorithm, along with a heuristic algorithm that is much less demanding on CUs' hardware. A CG-based decentralized channel assignment and power allocation algorithm was next developed to relieve the computation burden on the CBS. We analyzed the complexity and derived an upper bound for the CG-based algorithm, and validated its performance with simulations.

Acknowledgment

This work is supported in part by the US National Science Foundation (NSF) under Grant CNS-0953513, and through the NSF I/UCRC Broadband Wireless Access & Applications Center (BWAC) site at Auburn University. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

References

[1] Y. Zhao, S. Mao, J. Neel, and J. H. Reed, "Performance evaluation of cognitive radios: metrics, utility functions, and methodologies," *Proc. IEEE*, vol.97, no.4, pp.642–659, Apr. 2009.

[2] D. Hu, S. Mao, Y. T. Hou, and J. H. Reed, "Scalable video multicast in cognitive radio networks," *IEEE J. Sel. Areas Commun.*, Special Issue on Wireless Video Transmission, vol.29, no.3, pp.334–344, Apr. 2010.

[3] D. Hu, and S. Mao, "Streaming scalable videos over multi-hop cognitive radio networks," *IEEE Trans. Wireless. Commun.*, vol.11, no.9, pp.3501–3511, Nov. 2011.

[4] Y. Chen, Q. Zhao, and A. Swami, "Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors," *IEEE Trans. Inf. Theory*, vol.54, no.5, pp.2053–2071, May 2008.

[5] X. Zhang and H. Su "CREAM-MAC: Cognitive radio-enabled multi-channel MAC protocol over dynamic spectrum access networks," *IEEE J. Sel. Topics Signal Process.*, vol.17, no.9, pp.110–123, Feb. 2011.

[6] H. Su and X. Zhang, "Cross-layer based opportunistic MAC protocols for QoS provisionings over cognitive radio wireless networks," *IEEE J. Sel. Areas Commun.*, vol.26, no.1, pp.118–129, Jan. 2008.

[7] M. Wien, H. Schwarz, and T. Oelbaum "Performance Analysis of SVC," *IEEE Trans. Circuits Syst. Video Technol.*, vol.5, no.1, pp.1194–1203, Sept. 2007.

[8] M.S. Bazaraa, J.J. Jarvis, and H.D. Sherali *Linear Programming and Network Flows*, 4th edition, *Wiley*, New York, NY, 2009.

[9] Y.C. Liang, Y. Zheng, E. Peh, and A.T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Wireless. Commun.*, vol.7, no.4, pp.1326–1337, Apr. 2008.

[10] W. Wang, B. Kasiri, J. Cai, and A.S. Alfa. "Channel assignment of cooperative spectrum sensing in multi-channel cognitive radio networks," *Proc. IEEE ICC'11*, kyoto, Japan, June. 2011, pp.1–5.

[11] A. W. Min, Kang G. Shin, "Joint optimal sensor selection and scheduling in dynamic spectrum access networks," *IEEE Trans. Mobile Comput.*, vol.12, no.8, pp.1532–1545, Aug. 2013.

[12] Y. Chen, Y. Wu, B. Wang, and K. J. R. Liu, "Spectrum auction games for multimedia streaming over cognitive radio networks," *IEEE Trans. Commun.*, vol.58, no.8, pp.2381–2390, Aug. 2010.

[13] S. Ali and F. Yu, "Cross-layer QoS provisioning for multimedia transmissions in cognitive radio networks," *Proc. IEEE WCNC'09*, Budapest, Hungary, Apr. 2009, pp.1–5.

[14] G.B. Dantzig and P. Wolfe "The decomposition principle for linear programs," *Operations Research*, vol.8, no.1, pp.101–111, Feb. 1960.

[15] J.J. Lyu, H. Luh, and M.C. Lee "Solving linear programming problems on the parallel virtual machine environment," *American J. Appl. Sci.*, vol.1, no.2, pp.90–94, 2004.

[16] R. Garfinkel and G.L. Nemhauser *Integer Programming*, *Wiley*, New York, NY, 1972.

[17] I. Griva, S.G. Nash, and A. Sofer *Linear and Nonlinear Optimization*, 2nd edition, *Society for Industrial and Applied Mathematics*, Philadelphia, PA, 2009.

[18] A. Khan, L. Sun, and E. Ifeachor. "QoE prediction model and its application in video quality adaptation over UMTS networks," *IEEE Trans. Multimedia*, vol.14, no.2, pp.431–442, Apr. 2012.