Call Admission Control With Inter-Network Cooperation for Cognitive Heterogeneous Networks

Lei Xu, *Member, IEEE*, Ping Wang, *Senior Member, IEEE*, Qianmu Li, and Yinwei Jiang

***Abstract*— In this paper, a call admission control algorithm based on inter-network cooperation is proposed via a Stackelberg game framework for cognitive heterogeneous networks. The call admission control problem is subject to the variable bandwidth rate traffic, network service selection, feasible subchannel allocation, and call blocking probability. The call admission control algorithm is based on spectrum price at primary heterogeneous networks and subchannel allocation price and network selection at cognitive heterogeneous networks. In order to determine the call blocking probability, a probability upper bound of exceeding the maximum admission number for secondary mobile terminals (MTs) is analyzed based on M/M/∞ model. Then, the subchannel allocation price and network selection are designed via the dual decomposition method, and the vacant spectrum price is determined with Bertrand game theory. Finally, a call admission control algorithm is proposed. Simulation results demonstrate that the proposed algorithm not only improves quality of service at each secondary MT, but also reduces the call blocking probability for cognitive heterogeneous networks.**

***Index Terms*— Cognitive heterogeneous networks, call admission control, network selection, Stackelberg game, M/M/∞ model.**

I. INTRODUCTION

IN THE fifth generation (5G) mobile communication systems, heterogeneous wireless networks will use various wireless access technologies, e.g., macrocells providing low-to-medium rate services with a large coverage area, while microcells and Wi-Fi supporting high rate services in hotspots. In each wireless network, mobile terminals (MTs) do not always occupy all spectrum resources, and there exist vacant spectrum holes. Due to the spectrum resource scarcity, cognitive radio can be applied in each wireless network to utilize the vacant spectrum holes [1]. This leads to cognitive heterogeneous wireless networks, where the coverage areas of multiple cognitive wireless networks overlap and secondary MTs can opportunistically utilize the temporary spectrum holes [2]–[6]. There are international standards, e.g., IEEE 802.11af, IEEE 802.19 TG 1, IEEE 802.22, and LTE-U, in place to promote the development of cognitive wireless networks [7].

In cognitive heterogeneous wireless networks, the available spectrum resources for each network are dynamic and limited. It is difficult to provision quality of service (QoS) to secondary MTs. Hence, integrating cognitive heterogeneous wireless networks can help to provide various classes of service to secondary MTs and to support seamless secondary MT roaming. In order to guarantee QoS for MTs, multi-homing technology can be applied, where the data stream from an MT is split into multiple sub-streams, and transmitted over multiple networks by different radio interfaces simultaneously.

Efficient call admission control strategies are the key resource management components in such cognitive heterogeneous networks to guarantee QoS requirement for secondary MTs [8]. Unlike the fixed admitted users in wireless networks [9]–[11], there exist various works to investigate the call admission control problems in cognitive heterogeneous networks [8], [12]–[15]. These works can be classified into two categories, i.e., call admission control for single access network [8], [12]–[14] and that for multi-homing network [15]. For call admission control for single access network, each MT obtains its required resource from a single access network. A power-based call admission control scheme, to accommodate multi-class traffic, is proposed via directly extending the number-based call admission control scheme for code division multiple access (CDMA)-based wireless networks [12]. A two-tier call admission control scheme considering differentiated services is presented for OFDMA-based wireless networks [8]. A framework of a 2-dimension call admission control with the utility- and fairness-constrained optimal revenue policy is proposed for WiMAX wireless networks [13]. In [14], a novel call admission control scheme with multiple traffic models is investigated for mobile network. Different from [8], which focuses on the call admission control problem at the single access wireless network, our work focuses on that for cognitive heterogeneous networks with the multi-homing technology. Additionally, the references [9]–[11] address the subchannel and power allocation problem at the physical layer, while our work tackles the call admission control problem at the network layer.

In multi-homing network, multiple radio interfaces of each MT are used simultaneously to satisfy the MT's requirement, and each MT obtains its required resource from all available heterogeneous networks. In [15], a call admission control

Manuscript received July 19, 2016; revised December 2, 2016; accepted January 21, 2017. Date of publication January 25, 2017; date of current version March 8, 2017. This work was supported in part by the National Natural Science Foundation of China under Grant 61301108 and Grant 61671244, in part by the Fundamental Research Funds for the Central Universities under Grant 30915011320 and Grant 30916015104, in part by the Project of Jiangsu Provincial Six Talent Peaks under Grant XYDXXJS-033, in part by the National key Research and Development Program under Grant 2016YFE0108000, in part by CERNET Next Generation Internet Technology Innovation Project unde Grant NGII20160122, and in part by ZTE Cooperation Forum Project under Grant 2016ZTE04-11. The associate editor coordinating the review of this paper and approving it for publication was R. Zheng.

L. Xu, Q. Li, and Y. Jiang are with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: xuleimarcus@126.com).

P. Wang is with the School of Computer Engineering, Nanyang Technological University (e-mail: wangping@ntu.edu.sg).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TWC.2017.2657757

algorithm, based on multi-homing technology, is proposed for heterogeneous wireless access medium. However, how cognitive radio affects call admission control for cognitive heterogeneous networks has not been studied yet in the literature. In this paper, secondary MTs have multi-homing capabilities, and we study call admission control, based on $M/M/\infty$ queueing theory and Stackelberg game theory, for cognitive heterogeneous networks.

Specially, we summarize our contributions as follows: (i) Call admission control, based on inter-network cooperation, for cognitive heterogeneous networks is formulated as a Stackelberg game framework, and call blocking probability is analyzed with $M/M/\infty$ queueing theory; (ii) Stackelberg game analysis is proposed, joint subchannel allocation prices and network selections are analyzed via the dual decomposition method, and vacant spectrum prices are determined based on Bertrand game theory; (iii) A call admission control algorithm is presented for cognitive heterogeneous networks, and simulation results demonstrate that the proposed algorithm can reduce the call blocking probability at secondary MTs efficiently.

The rest of this paper is organized as follows. The system model is described in Section II. Section III presents the Stackelberg game framework, and the call admission control algorithm for cognitive heterogeneous networks is proposed in Section IV. Finally, performance evaluation and conclusions are given in Sections V and VI, respectively.

II. System Model

In this section, the system and channel models are described first. Then, the network service types and traffic model are presented. Finally, the mobility model is given, and the number of admitted secondary MTs is analyzed.

A. System Description

Consider a geographical region with $\mathcal{N} = \{1, 2, \cdots, N\}$ primary wireless access networks, based on different wireless access technologies and operated by different service operators. In network n, there is a set, $\mathcal{S}_n = \{1, 2, \cdots, S_n\}$, of base stations (BSs). Corresponding to network n BS s, there is a cognitive wireless network n BS s. The cognitive heterogeneous wireless networks are shown in Fig. 1. There is a set, $\mathcal{M} = \{1, 2, \cdots, M\}$, of secondary MTs in the geographical region, and $\mathcal{M}_{ns} = \{1, 2, \cdots, M_{ns}\} \in \mathcal{M}$ is a subset of the MTs, which reside in the coverage area of network n BS s. In primary network n BS s, the total spectrum is divided into K_{ns} subchannels, and each subchannel has the same bandwidth. Additionally, the transmission power at secondary BSs for downlink or at secondary MTs for uplink should be controlled to protect the transmission in primary networks according to the interference temperature model [16]. [1] In the same cognitive network, interference mitigation is achieved by interference management schemes [19], [20]. Using the multi-homing mechanism and multiple radio interfaces, each secondary MT can communicate with multiple secondary BSs simultaneously.

[1] In this work, we divide the spectrum into a number of subchannels. Primary networks use some subchannels, while cognitive networks utilize the vacant subchannels. The interference exists between primary networks and cognitive networks due to the cross-channel interference, e.g., [17], [18]. Hence, the interference introduced by cognitive networks to primary networks should be controlled via the interference temperature model.

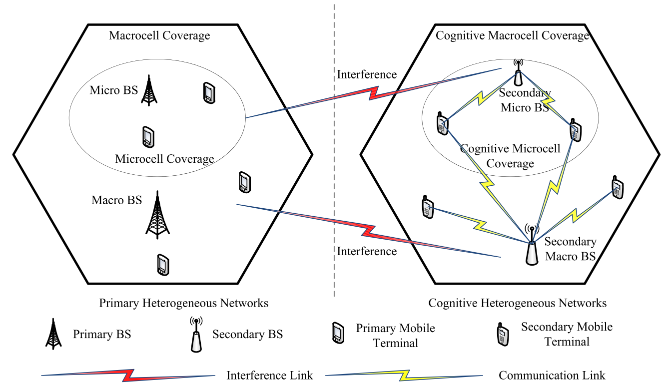

Fig. 1. Cognitive heterogeneous networks.

B. Channel Model

Consider K_{ns} time-varying subchannels licensed for primary network n BS s. The occupancy of each subchannel, $k \in \{1, 2, \cdots, K_{ns}\}$, is modeled as a time-homogeneous discrete Markov process, denoted by s_{ns}^k, where $s_{ns}^k = 0$ means that the subchannel is idle, and $s_{ns}^k = 1$ means that the subchannel is busy. Define α_{ns}^k (β_{ns}^k) as the probability that the licensed subchannel k at primary network n BS s changes from idle (busy) in the previous time slot to busy (idle) in the present slot. In cognitive heterogeneous networks, $1 - s_{ns}^k = 1$ indicates that the subchannel k is available for cognitive network n BS s, otherwise $1 - s_{ns}^k = 0$ [16].

The vacant subchannel allocation should satisfy the feasibility, i.e.,

$$\sum_{m \in \mathcal{M}_{ns}} \rho_{nsm}^k \left(1 - s_{ns}^k\right) \leq 1, \quad \forall n \in \mathcal{N},\ s \in \mathcal{S}_n,\ k \in \mathcal{K}_{ns} \tag{1}$$

where ρ_{nsm}^k is the indicator variable of subchannel allocation, e.g., if $\rho_{nsm}^k = 1$, the subchannel k is allocated to cognitive network n BS s MT m. $\mathcal{K}_{ns}$ is the vacant subchannel set for cognitive network n BS s.

C. Network Service Types

In a given service area, r, the subset of secondary MTs is $\mathcal{M}_r$. Besides its own home network, each secondary MT can get service from other secondary BSs available at its location. Secondary MTs can be classified into two categories: network subscribers and network users. Each network subscriber uses its own home network, while each network user utilizes the other networks. Consequently, p_{nsm} is defined as a priority parameter, representing service priority for secondary network n in allocating its resources to secondary MT m via secondary BS s, where $p_{nsm} \in [0, 1)$ denotes for low-priority network users, and $p_{nsm} = 1$ denotes for high-priority network subscribers. At each secondary MT, single network service and multi-homing service are considered. The subset of secondary

MTs using single network service in the service area r is $\mathcal{M}_{1r}$, while the subset of secondary MTs using multi-homing service is $\mathcal{M}_{2r}$. $\mathcal{S}_r$ is the subset of all secondary BSs in the service area r. The set of secondary MTs in cognitive network n BS s, including MTs using either multi-homing service or single-network service, is denoted by $\mathcal{M}_{ns}$ [21].

If secondary MT m uses single-network service in area r, the constraint is

$$\sum_{n\in\mathcal{N}}\sum_{s\in\mathcal{S}_n} x_{nsm}=1,\quad m\in\mathcal{M}_{1r},\ r\in\mathbf{R} \tag{2}$$

where x_{nsm} is a binary assignment variable, e.g., if secondary MT m communicates with cognitive network n BS s, $x_{nsm}=1$; otherwise $x_{nsm}=0$. $\mathbf{R}$ is the service area set.

If secondary MT m uses multi-homing network service in area r and cognitive network n BS s serves in the service area r, the constraint is

$$x_{nsm}=1,\quad m\in\mathcal{M}_{2r},\ r\in\mathbf{R}. \tag{3}$$

D. Traffic Model

Consider the video traffic at each secondary MT. In a geographic area, the video call arrivals are modeled as a Poisson process [7]. Specially, in service area r, the arrival process for both new and handoff video calls is modeled by a Poisson process with arrival rate υ_r. Additionally, the video call duration follows exponential distribution. Consequently, the probability density function (PDF) of the video call duration, T_c, with mean $\overline{\overline{T}}_c$, is

$$f_{T_c}(t)=\frac{1}{\overline{\overline{T}}_c}\exp\left(-\frac{t}{\overline{\overline{T}}_c}\right),\quad t\geq 0. \tag{4}$$

The video call is variable bandwidth rate (VBR) traffic. For secondary MT m, a VBR call requires a bandwidth allocation within a minimum value, $B_m^{\min}$, and a maximum value, $B_m^{\max}$, i.e.,

$$\sum_{n\in\mathcal{N}}\sum_{s\in\mathcal{S}_n}\sum_{k\in\mathcal{K}_{ns}} x_{nsm}\rho_{nsm}^k\left(1-s_{ns}^k\right)B_{ns}\geq B_m^{\min} \tag{5}$$

and

$$\sum_{n\in\mathcal{N}}\sum_{s\in\mathcal{S}_n}\sum_{k\in\mathcal{K}_{ns}} x_{nsm}\rho_{nsm}^k\left(1-s_{ns}^k\right)B_{ns}\leq B_m^{\max} \tag{6}$$

where B_{ns} is the subchannel bandwidth at primary network n BS s. In a given service area, if there are sufficient bandwidth resources, the VBR call will obtain the maximum required bandwidth, $B_m^{\max}$. When all secondary BSs reach their capacity limitation, the bandwidth allocation for each VBR call is degraded towards the minimum required bandwidth, $B_m^{\min}$ [21].

E. Mobility Model

In a given service area r, the user residence time is T_s^r, which follows an exponential distribution with mean $\overline{\overline{T}}_s^r$. Since the channel holding time is $T_h^r=\min\left(T_c,T_s^r\right)$, we can obtain

$$\begin{aligned}\Pr\left\{\min\left(T_s^r,T_c\right)>t\right\}&=\Pr\left\{T_s^r>t,T_c>t\right\}\\&=\Pr\left\{T_s^r>t\right\}\Pr\left\{T_c>t\right\}\end{aligned} \tag{7}$$

where T_s^r and T_c are independent with each other.

From (7), the PDF of T_h^r is

$$f_{T_h^r}(t)=\left(\frac{1}{\overline{\overline{T}}_c}+\frac{1}{\overline{\overline{T}}_s^r}\right)\exp\left[-\left(\frac{1}{\overline{\overline{T}}_c}+\frac{1}{\overline{\overline{T}}_s^r}\right)t\right],\quad t\geq 0. \tag{8}$$

F. The Number of Admitted MTs

In dynamic cognitive heterogeneous networks, we should avoid performing the resource allocation for every video call arrival or departure in service area r. Consequently, the objective of cognitive heterogeneous networks is to allocate the required resource for video calls with a certain call blocking probability. In service area r, the number of existing calls is M_r. The maximum number of calls which can be supported in service area r with the available capacities of secondary BSs is denoted by $\hat{C}_r$. Since the maximum number of calls can be determined using a capacity analysis, a call admission control procedure is in place, which guarantees that $M_r\leq\hat{C}_r$, so that feasible resource allocation solutions exist [15]. The call admission control is implemented via the subchannel allocation price and spectrum price. The spectrum price is announced by the primary network. With the spectrum price, cognitive networks decide the amount of vacant spectrum to buy from the primary network. Additionally, the subchannel price is determined by the secondary BSs, and secondary BS decides each new call whether to be admitted based on the subchannel price and the remaining bandwidth resource.

Define $\hat{M}_r$ as the target value of allowable admitted calls. The choice of the target value, $\hat{M}_r$, determines the overall performance of the geographical region in terms of the allocated resources per call and call blocking probability. Consequently, the value, $\hat{M}_r$, should be properly chosen to achieve satisfactory performance in the resource allocation. Given the PDF for M_r, we can represent $\hat{M}_r$ by a design parameter ε_r so that

$$\Pr\left(M_r>\hat{M}_r\right)\leq\varepsilon_r \tag{9}$$

where $\varepsilon_r\in[0,1]$. It is clear that the value of $\hat{M}_r$ depends on ε_r and the distribution of M_r.

Since video call arrival in service area r follows a Poisson process, the channel holding time follows an exponential distribution, and all video calls are served simultaneously, an M/M/∞ queueing model can be used to determine $\hat{M}_r$ with the steady-state call traffic and MT mobility statistics. Consequently, in service area r, M_r follows the Poisson distribution with mean $M_r^a=\upsilon_r E\left[T_h^r\right]$, where $E\left[T_h^r\right]$ is the average channel holding time, i.e.,

$$E\left[T_h^r\right]=\frac{T_cT_s^r}{T_c+T_s^r}. \tag{10}$$

With (10), $\hat{M}_r$ can be found by (11) as the minimum integer.

$$\sum_{i=0}^{\hat{M}_r}\frac{\left(M_r^a\right)^i\exp\left(-M_r^a\right)}{i!}\geq 1-\varepsilon_r. \tag{11}$$

When $\hat{M}_r$ is larger than $\hat{C}_r$, set $\hat{M}_r = \hat{C}_r$. Via $\hat{M}_r$, the number of admitted MTs, M_{ns}, at cognitive network n BS s is obtained. In this work, a call admission control (CAC) algorithm is designed to satisfy the QoS of the VBR traffic of video call for inter-network-based cognitive heterogeneous networks. We aim to maximize the profit for primary networks by selling the vacant spectrum resource and guarantee the QoS for the admitted secondary MTs via the CAC algorithm. For the primary heterogeneous networks, the decision variable is the spectrum price. For the cognitive heterogeneous networks, the decision variables are the subchannel allocation price, subchannel allocation and network selection.

III. Stackelberg Game Framework

In order to design the spectrum price for primary heterogeneous networks and the subchannel allocation price for the CAC at cognitive heterogeneous networks, the stackelberg game model is adopted. In this section, Stackelberg game formulation is given firstly. Then, joint subchannel allocation price and network selection are analyzed for secondary MTs at cognitive heterogeneous networks. Finally, spectrum price analysis for primary heterogeneous networks is given based on Bertrand game theory.

A. Stackelberg Game Formulation

In order to describe the relationship between primary heterogeneous networks and cognitive heterogeneous networks, we adopt the Stackelberg game to design the spectrum price strategy at the primary networks and subchannel allocation price and network selection policy at the cognitive networks. Game theory analyzes behavior in strategic situations, in which players influence each other's decision and performance, e.g., Stackelberg leadership model is a strategic game, where the leaders act first and the followers act sequentially [22], [23]. It is very suitable to analyze the influence of the spectrum price at primary heterogeneous networks to the subchannel allocation price and network selection at cognitive heterogeneous networks. The spectrum price is the price at which the primary networks sell the spectrum to the cognitive heterogeneous networks. On the other hand, the subchannel allocation price is the price that each secondary BS charges its serving secondary MTs with the aim to guarantee the call blocking probability below a certain threshold at each secondary network. In the Stackelberg game, primary heterogeneous networks are leaders who define the spectrum prices first, and cognitive heterogeneous networks are followers who observe the spectrum prices and perform the subchannel price and network selection.

In order to obtain the Stackelberg equilibrium, the back-ward induction method is adopted [24], [25]. In the back-ward induction method, the subchannel allocation price and network selection are studied firstly, which maximize the cognitive heterogeneous networks payoff. Then, given the optimal sub-channel allocation price and network selection, the spectrum price is investigated to maximize the primary heterogeneous networks' utility. The formulation of the proposed Stackelberg game is as follows.

1) Followers: Given the spectrum price at primary heterogeneous networks, the joint subchannel allocation and network selection can be mathematically formulated as

$$\mathbf{G}_c = \left\{\mathcal{M}_a, \Upsilon, \{u_m\}_{m\in\mathcal{M}_a}\right\} \tag{12}$$

where $\mathcal{M}_a$ is the set of active secondary MTs. The set of feasible subchannel allocation region is

$$\Upsilon = \left\{\rho_{nsm}^k \geq 0, x_{nsm} \in \{0,1\} \,|\, (1)-(3), (5), (6)\right\}. \tag{13}$$

The utility function, $u_m\left(x_{nsm}, \rho_{nsm}^k\right)$, at secondary MT m is

$$u_m\left(x_{nsm}, \rho_{nsm}^k\right) = \sum_{n\in\mathcal{N}}\sum_{s\in\mathcal{S}_n} u_{nsm}\left(x_{nsm}, \rho_{nsm}^k\right) \tag{14}$$

where

$$\begin{aligned} u_{nsm}&\left(x_{nsm}, \rho_{nsm}^k\right)\\ &= \log_2\left[1 + \eta_1 \sum_{k\in K_{ns}} x_{nsm}\rho_{nsm}^k\left(1 - s_{ns}^k\right)B_{ns}\right]\\ &\quad - \left[\lambda_n + \eta_2\left(1 - p_{nsm}\right)\right]\sum_{k\in K_{ns}} x_{nsm}\rho_{nsm}^k\left(1 - s_{ns}^k\right)B_{ns} \end{aligned} \tag{15}$$

where $u_{nsm}\left(x_{nsm}, \rho_{nsm}^k\right)$ is the utility function for secondary MT m to communicate with cognitive network n BS s, η_1 and η_2 are used for the scalability of ρ_{nsm}^k, λ_n is the vacant spectrum price for primary network n, and p_{nsm} is a priority parameter assigned by cognitive network n to secondary MT m on its resources in secondary BS s. The utility function, $u_m\left(x_{nsm}, \rho_{nsm}^k\right)$, at secondary MT m is the total reward, i.e., the transmission rate, for secondary MT m minus the cost of the bandwidth.

2) Leaders: In primary heterogeneous networks, the dynamic spectrum sharing problem under price competition is modeled as an oligopoly market. Additionally, Bertrand game can be applied to describe it [26], [27]. In the oligopoly market competition, there are N primary networks to provide the vacant spectrum resource, and these primary networks compete with each other by adjusting the price to maximize their profits. In Bertrand game, the players are the primary networks. The strategy for each primary network is the price offered per unit of spectrum. The payoff for each primary network is the profit of primary network n in selling vacant spectrum to secondary MTs. Given the subchannel allocation price and network selection, the primary networks' noncooperative subgame can be mathematically formulated as

$$\mathbf{G}_p = \left\{\mathcal{N}, \lambda, \{c_n\}_{n\in\mathcal{N}}\right\} \tag{16}$$

where $\mathcal{N}$ is the set of primary networks. The set of feasible spectrum price region is

$$\lambda = \{\lambda_n \geq 0, \forall n\}. \tag{17}$$

The utility function, c_n, for primary network n is

$$c_n = \lambda_n \sum_{s\in\mathcal{S}_n}\sum_{m\in\mathcal{M}_{ns}}\sum_{k\in\mathcal{K}_{ns}} x_{nsm}\rho_{nsm}^k\left(1 - s_{ns}^k\right)B_{ns} \tag{18}$$

where c_n is the total investment reward, i.e., the profit for selling bandwidth, for primary network n.

Since the Stackelberg game contains the leaders and followers, we can transform the analysis of the whole problem into the successive analysis of the leaders' problem and the followers' problem. Firstly, we solve the subchannel allocation price and network selection problem based on convex optimization theory. Once we obtain the solution of the followers' problem, the leaders' problem is a Bertrand game problem, and the successive analysis is enough for proposing the call admission control algorithm.

B. Subchannel Allocation and Network Selection

Subchannel allocation and network selection for secondary MTs with single-network and multi-homing services in the cognitive heterogeneous networks is formulated as

$$\begin{aligned} OP1: \quad & \max_{x_{nsm},\rho_{nsm}^k} \sum_{\mathcal{M}} u_m\left(x_{nsm},\rho_{nsm}^k\right) \\ & S.t.:\ x_{nsm},\rho_{nsm}^k \in \Upsilon. \end{aligned} \tag{19}$$

where problem (19) is mixed integer non-linear programming (MINLP). From (14)-(15), secondary MT m has the high priority to access its own serving secondary BS. If a secondary MT can not obtain enough resource, it will utilize the multi-homing service to obtain the resource from other secondary BSs. Problem (19) is divided into two suboptimal problems: subchannel allocation price problem and network selection problem.

In order to solve the subchannel allocation subproblem, assume all secondary MTs have the multi-homing ability, and set network selection variable $x_{nsm} = 1$, if secondary MT m can communicate with cognitive network n BS s. Consequently, problem (19) can be rewritten as

$$\begin{aligned} OP2: \quad & \max_{\rho_{nsm}^k} \sum_{\mathcal{M}} u_m\left(\rho_{nsm}^k\right) \\ & S.t.:\ \rho_{nsm}^k \in \Upsilon. \end{aligned} \tag{20}$$

Proposition 1: Problem (20) is a convex optimization problem.

Proof: See Appendix A.

Proposition 2: The optimal solution of (20) is

$$\rho_{nsm}^k = \frac{1}{\alpha_{nsm}^k} - \frac{1 + \sum\limits_{k'\neq k\in\mathcal{K}_{ns}} \eta_1 \rho_{nsm}^{k'}\left(1-s_{ns}^k\right)B_{ns}}{\eta_1\left(1-s_{ns}^k\right)B_{ns}} \tag{21}$$

and

$$\begin{aligned} \alpha_{nsm}^k = & \Big\{ v_{ns}^k + [\eta_2\left(1-p_{nsm}\right) \\ & + \lambda_n + \left(\beta_m^{(2)} - \beta_m^{(1)}\right)\Big] \left(1-s_{ns}^k\right)B_{ns}\Big\} \ln 2 \end{aligned} \tag{22}$$

Proof: See Appendix B.

The optimum values $\beta_m^{(1)}$, $\beta_m^{(2)}$ and v_{ns}^k can be calculated by solving the dual problem (37) in Appendix B. Therefore, a gradient descent method can be applied to calculate the optimal values for $\beta_m^{(1)}$, $\beta_m^{(2)}$ and v_{ns}^k, i.e.,

$$\beta_m^{(1)}(i+1) = \left[\beta_m^{(1)}(i) + \Delta\varepsilon_1\Delta\beta_1\right]^+ \tag{23}$$

$$\beta_m^{(2)}(i+1) = \left[\beta_m^{(2)}(i) + \Delta\varepsilon_2\Delta\beta_2\right]^+ \tag{24}$$

$$v_{ns}^k(i+1) = \left[v_{ns}^k(i) + \Delta\varepsilon_3\left(1 - \sum_{m\in\mathcal{M}_{ns}} \rho_{nsm}^k\right)\right]^+ \tag{25}$$

$$\Delta\beta_1 = \sum_{n\in N}\sum_{s\in S_n}\sum_{k\in K_{ns}} \rho_{nsm}^k\left(1-s_{ns}^k\right)B_{ns} - B_m^{\min} \tag{26}$$

and

$$\Delta\beta_2 = B_m^{\max} - \sum_{n\in N}\sum_{s\in S_n}\sum_{k\in K_{ns}} \rho_{nsm}^k\left(1-s_{ns}^k\right)B_{ns} \tag{27}$$

where i is the iteration index, and $\Delta\varepsilon_j$, $j = 1, 2, 3$, is a step size. Since the gradient of (37) satisfies the Lipchitz continuity condition, the convergence towards the optimum solution is guaranteed by (23)-(27). Consequently, the subchannel allocation price, $v_{ns}^k(i)$, converges to the optimum solution [28].[2]

When the subchannel allocation, ρ_{nsm}^k, is determined, the network selection, x_{nsm}, for each secondary MT is

$$x_{nsm} = \begin{cases} 1, & \sum\limits_{k\in\mathcal{K}_{ns}} \rho_{nsm}^k > 0 \\ 0, & \text{otherwise.} \end{cases} \tag{28}$$

C. Spectrum Price

Given a set of spectrum prices for other primary networks λ_{-n}, the best response function for primary network n is

$$D_n\left(\lambda_{-n}\right) = \arg\max_{\lambda_n} c_n\left(\lambda_{-n}\cup\lambda_n\right). \tag{29}$$

The set, $\lambda^* = \left\{\lambda_1^*, \cdots, \lambda_N^*\right\}$, denotes the Nash equilibrium of the Bertrand game for this spectrum price competition, if and only if

$$\lambda_n^* = D_n\left(\lambda_{-n}^*\right), \forall n \tag{30}$$

where λ_{-n}^* denotes the set of best responses for primary network j for $j \neq n$. Consequently, to obtain the sub-game Nash equilibrium, we have to solve

$$\frac{\partial c_n}{\partial \lambda_n} = 0, \quad \forall n \in \mathcal{N}. \tag{31}$$

From (31), we can obtain

$$\sum_{s\in\mathcal{S}_n}\sum_{m\in\mathcal{M}_{ns}}\sum_{k\in\mathcal{K}_{ns}} \left(\rho_{nsm}^k + \lambda_n\frac{\partial\rho_{nsm}^k}{\partial\lambda_n}\right)\left(1-s_{ns}^k\right)B_{ns} = 0. \tag{32}$$

From (32), we can obtain the optimal spectrum price, i.e.,

$$\lambda_n = \left[\phi\left(\rho_{nsm}^k, \beta_m^{(1)}, \beta_m^{(2)}, v_{ns}^k\right)\right]^+ \tag{33}$$

where $\phi(\bullet)$ is a mapping function which satisfies (32).

[2]Since $v_{ns}^k(i)$ has an impact on the subchannel allocation, we address the Lagrangian multiplier, $v_{ns}^k(i)$, as the subchannel allocation price.

Algorithm 1 Joint Spectrum Price and Subchannel Allocation Price Setting (JSPSAPS)

Require: $B_m^{\min}$, $B_m^{\max}$, and M_{ns}.
Ensure: $\upsilon_{ns}^k(i+1)$ and $\lambda_n(i+1)$.
1: Initialize $\beta_m^{(1)}(i) \geq 0$, $\beta_m^{(2)}(i) \geq 0$, $\upsilon_{ns}^k(i) \geq 0$, $\lambda_n(i)$, and $i = 1$.
2: **repeat**
3: Cognitive network n BS s calculates ρ_{nsm}^k according to (21), and updates $\beta_m^{(1)}(i+1)$, $\beta_m^{(2)}(i+1)$ and $\upsilon_{ns}^k(i+1)$ according to (23)-(27).
4: Primary network n updates $\lambda_n(i+1)$ according to (33).
5: **if** $B_m^{\min} \leq \sum\limits_{n\in\mathcal{N}} \sum\limits_{s\in\mathcal{S}_n} \sum\limits_{k\in\mathcal{K}_{as}} \rho_{nsm}^k \left(1 - s_{ns}^k\right) B_{ns} \leq B_m^{\max}$ and $|\lambda_n(i+1) - \lambda_n(i)| \leq \varepsilon_\lambda$ **then**
6: Go to step 11.
7: **else**
8: Set $i \leftarrow i+1$, and go to step 3.
9: **end if**
10: **until**
11: Output $\upsilon_{ns}^k(i+1)$ and $\lambda_n(i+1)$.

Algorithm 2 Hybrid Access Call Admission Control (HACAC)

Require: $B_m^{\min}$, $B_m^{\max}$, $\upsilon_{ns}^k(i+1)$ and $\lambda_n(i+1)$.
Ensure: x_{nsm} and ρ_{nsm}^k.
1: Initialize $\beta_m^{(1)}(i) \geq 0$, $\beta_m^{(2)}(i) \geq 0$, and $i = 1$.
2: **repeat**
3: Each secondary MT calculates ρ_{nsm}^k, $\beta_m^{(1)}(i+1)$, and $\beta_m^{(2)}(i+1)$ according to (23)-(27).
4: **if** $\left|\beta_m^{(1)}(i+1) - \beta_m^{(1)}(i)\right| \leq \varepsilon_\beta$ **then**
5: The secondary MT is blocked if the total allocated resources do not satisfy its required bandwidth. Calculate x_{nsm} according to (28), and go to step 10.
6: **else**
7: Set $i \leftarrow i+1$, and go to step 3.
8: **end if**
9: **until**
10: Output x_{nsm} and ρ_{nsm}^k.

IV. Call Admission Control for Cognitive Heterogeneous Networks

In this section, the call admission control for cognitive heterogeneous networks is designed with the optimal subchannel allocation price and the vacant spectrum price. Then, the computational complexity is analyzed for the proposed algorithms.

A. Call Admission Control

In order to design a call admission control algorithm, we first need to solve the optimal subchannel allocation price and the vacant spectrum price from the Stackelberg game analysis. The joint spectrum price and subchannel allocation price setting (JSPSAPS) is given in Algorithm 1. The subchannel allocation price, υ_{ns}^k, and the spectrum price, λ_n, obtained from Algorithm 1 will be broadcast by each secondary BS via its ID beacon. A new secondary MT uses its multiple radio interfaces to listen to the prices sent by the secondary BSs available at its location. Then a hybrid access call admission control (HACAC) algorithm is performed at the secondary MT. Specifically, the secondary MT (denoted as secondary MT m) uses these two prices to calculate the resource to be allocated from each secondary BS to itself, i.e., to calculate the network selection indicator, x_{nsm}, and the subchannel share indicator, ρ_{nsm}^k, from secondary BS s of cognitive network n. The secondary MT is blocked if the total allocated resources do not satisfy its required bandwidth. The details of the HACAC is given in Algorithm 2, and the flowchart of call admission control is depicted at Fig. 2. In Fig. 2, algorithm 1 first analyzes the Stackelberg game, and obtains the optimal spectrum price for primary heterogeneous networks and the optimal subchannel allocation price for cognitive heterogeneous networks. Then, algorithm 2 determines whether each secondary MT to be blocked or to be admitted via the optimal spectrum and subchannel allocation prices. ε_β and ε_λ are the arbitrarily small

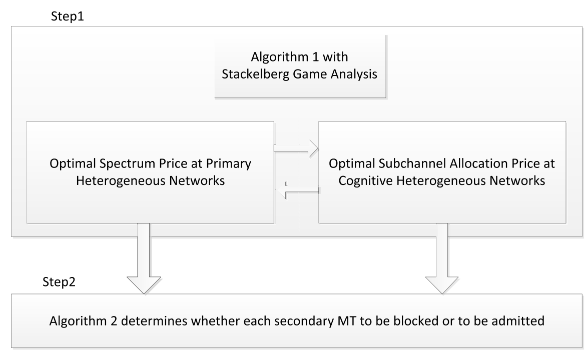

Fig. 2. Flowchart of call admission control for cognitive heterogeneous networks.

positive numbers. $\lambda_n(i+1)$ and $\lambda_n(i)$ are the variable values at the $(i+1)$th iteration and the ith iteration. $\beta_m^{(1)}(i+1)$, $\beta_m^{(2)}(i+1)$, and $\upsilon_{ns}^k(i+1)$ are the Lagrangian multipliers at the $(i+1)$th iteration. $\beta_m^{(1)}(i)$, $\beta_m^{(2)}(i)$ and $\upsilon_{ns}^k(i)$ are the Lagrangian multipliers at the ith iteration.

HACAC can obtain the optimal solution for cognitive heterogeneous networks. However, it has a relatively high computational complexity when supporting multi-homing network service. In order to reduce the computational complexity, a simple single access call admission control (SACAC) is proposed, in which each secondary MT can only access its serving secondary BS. In the serving secondary BS, if the remaining bandwidth resource can afford a new call, this secondary MT can be admitted; otherwise, it will be blocked. SACAC is a sub-optimal solution with low complexity as compared with HACAC.

Since the available spectrum at cognitive heterogeneous networks and the analysis for the number of admitted MTs have stochastic characteristics, Markov process and queueing theory are used to analyze them. Because the Markov process is a two-state time-homogeneous discrete Markov process and the number of admitted MTs can be computed with (11), the costs of using queueing theory and Markov process are limited.

Additionally, the Stackelberg game and Bertrand game are utilized to analyze the relationship between the primary heterogeneous networks and the cognitive heterogeneous networks. We have conducted the computational analysis for the proposed algorithms, i.e., JSPSAPS and HACAC to evaluate the cost of using Stackelberg game and Bertrand game models.

B. Computational Complexity

In JSPSAPS, the computational complexity is determined by the complexity of solving the joint subchannel allocation price and network selection problem and the spectrum price problem. The complexity of the gradient method for the joint subchannel allocation price and network selection problem is polynomial in number of dual variables. As a result, the computational complexity of JSPSAPS is given by $O\left(P_I O_I M^2 \sum_{n\in\mathcal{N}}\sum_{s\in\mathcal{S}_n}|\mathcal{K}_{as}|\right)$, where O_I is the number of iterations required for the convergence of subchannel allocation price and network selection, and P_I is the number of iterations required for the convergence of spectrum price. The computational complexity of HACAC is given by $O\left(A_I \sum_{n\in\mathcal{N}}\sum_{s\in\mathcal{S}_n}|\mathcal{K}_{as}|\right)$, where A_I is the number of iterations required for the convergence of subchannel allocation for each secondary MT. Consequently, the total computational complexity is $O\left(\left(P_I O_I M^2 + A_I\right)\sum_{n\in\mathcal{N}}\sum_{s\in\mathcal{S}_n}|\mathcal{K}_{as}|\right)$. In SACAC, the computational complexity is determined by $O\left(P_I M \sum_{n\in\mathcal{N}}\sum_{s\in\mathcal{S}_n}|\mathcal{K}_{as}|\right)$.

V. Performance Evaluation

This section presents the simulation results using the proposed algorithm for cognitive heterogeneous networks. A geographical region that is entirely covered by a secondary macrocell BS and partially covered by a secondary microcell BS. As a result, $\mathcal{N} = \{1, 2\}$ with cognitive macrocell and cognitive microcell indexed as 1 and 2, respectively. Two service areas can be distinguished, $\mathbf{r} = \{1, 2\}$. In area 1, only the service from the secondary macrocell BS is available. In area 2, both the secondary macrocell and microcell BSs services are available. The numbers of secondary MTs in the first area and in the second area are 5 and 3, respectively. The primary macrocell and microcell capacities are 50 Mbps and 30 Mbps, respectively. The minimum and maximum bandwidths for VBR are 0.256 Mbps and 0.512 Mbps, respectively. The number of subchannels for primary macrocell and microcell are 50 and 30, respectively. The average call duration is 10 Minutes. The other simulation parameters are $\eta_1 = 1$, and $\eta_2 = 1 \times 10^{-6}$.

Fig. 3-Fig. 5 show the impact of call arrival rate at cognitive microcell on total utility for primary heterogeneous networks, total utility for secondary MTs, and call blocking probability at cognitive microcell for different algorithms, respectively. There are two cases for the vacant subchannel state transition

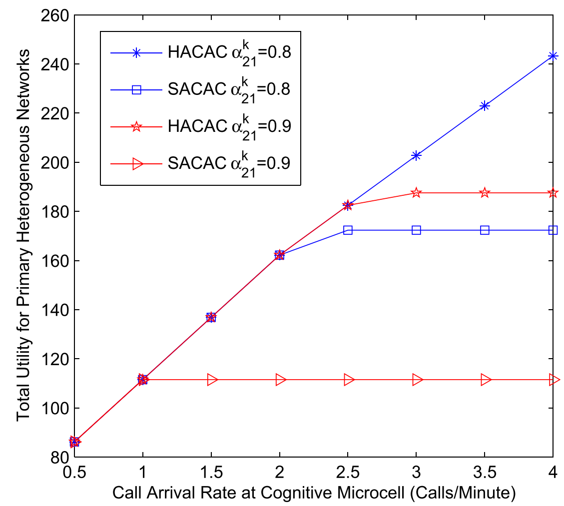

Fig. 3. Total utility for primary heterogeneous networks vs. call arrival rate at cognitive microcell.

probabilities at a macrocell. In the first case, the probability that in a macrocell the licensed subchannel becomes busy (idle) is 0.9 (0.1). In the second case, the probability that in a macrocell the licensed subchannel becomes busy (idle) is 0.8 (0.2). On the other hand, the probability that in a microcell the licensed subchannel becomes busy (idle) is 0.75 (0.25). The parameter ε_r is 0.01. The average residence time for each secondary MT is 20 Minutes. We can see that the total utilities at primary heterogenous networks for HACAC and SACAC increase with the call arrival rate at cognitive microcell. This is because increasing call arrival rate at cognitive microcell means increasing the number of secondary MTs, and secondary MTs will consume more spectrum resource. When call arrival rate at cognitive microcell reaches a threshold, the total utility for primary heterogeneous networks remains unchanged. This can be explained that the vacant spectrum resource at primary heterogeneous networks is sold out and secondary MTs can not buy spectrum resources from primary heterogenous networks any more, although new secondary MT needs spectrum resource to its VBR traffic requirement. We can also see that the total utility for primary heterogeneous networks with the case $\alpha^k_{21} = 0.8$ is larger than that with the case $\alpha^k_{21} = 0.9$. This is due to the fact that when the licensed subchannel becomes less busy, more vacant spectrum resource is available for secondary MTs. Since HACAC aggregates all vacant spectrum resource to support the VBR requirement of secondary MTs, and more secondary MTs can be admitted, HACAC can utilize the vacant spectrum resource efficiently, and the total utility for primary heterogeneous networks outperforms that with SACAC. Fig. 4 has the similar phenomenon as Fig. 3. From Fig. 5, we can see that the call blocking probability at cognitive microcell increases with call arrival rate at cognitive microcell. This is due to the fact that heterogeneous wireless networks can not support all the calls when the existing calls exceed the maximum number of admitted calls. Consequently, some new calls are blocked to maintain the QoS requirement of existing calls.

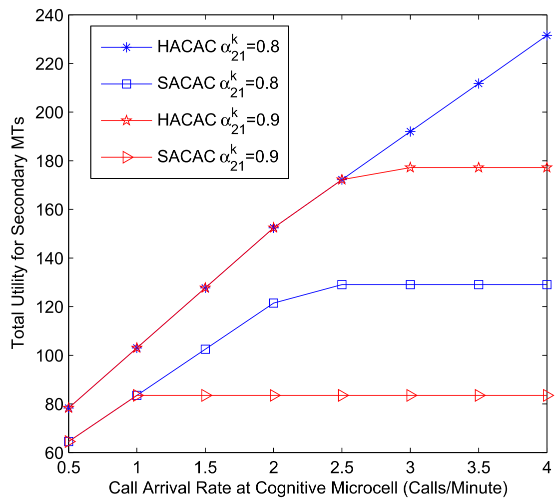

Fig. 4. Total utility for secondary MTs vs. call arrival rate at cognitive microcell.

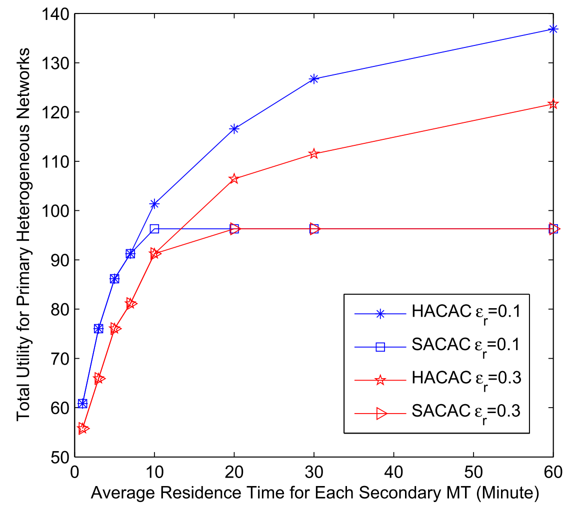

Fig. 6. Total utility for primary heterogeneous networks vs. average residence time for each secondary MT.

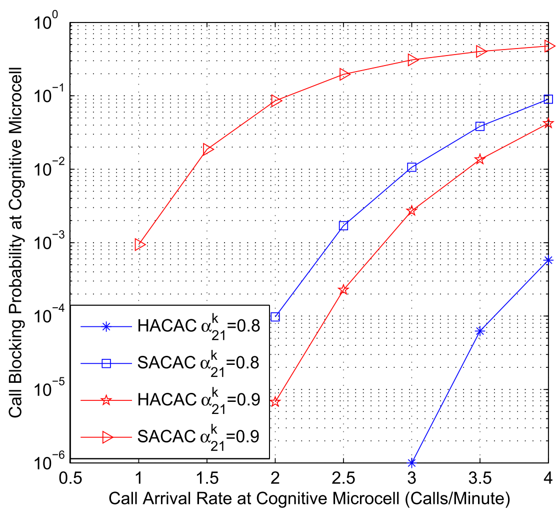

Fig. 5. Call blocking probability at cognitive microcell vs. call arrival rate at cognitive microcell.

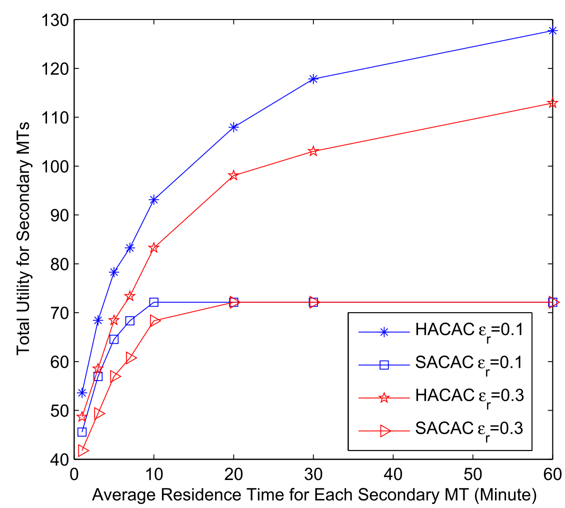

Fig. 7. Total utility for secondary MTs vs. average residence time for each secondary MT.

Fig. 6-Fig. 8 depict average residence time for each secondary MT with respect to the total utility for primary heterogeneous networks, the total utility for secondary MTs, and the call blocking probability at cognitive microcell for different algorithms, respectively. There are two cases for the parameter ε_r, i.e., $\varepsilon_r = 0.1$ and $\varepsilon_r = 0.3$. ε_r is the upper bound for the call blocking probability. The probability that in a macrocell the licensed subchannel becomes busy (idle) is 0.9 (0.1), whereas the probability that in a microcell the licensed subchannel becomes busy (idle) is 0.85 (0.15). The video call arrival rate at cognitive microcell is 1.5 Calls/Minute. It can be seen that the total utilities for primary heterogeneous networks for the two algorithms increase with the average residence time for each secondary MT. This is due to the fact that increasing the average residence time for each secondary MT is equivalent to increasing the number of secondary MTs at each area. This leads to more vacant spectrum resource being bought by secondary MTs. When the vacant spectrum resource is not enough to be allocated, the total utility for primary heterogeneous networks keeps constant. It can be also seen that the total utility for primary heterogeneous networks with the case $\varepsilon_r = 0.3$ is smaller than that with the case $\varepsilon_r = 0.1$. This is because reducing the parameter ε_r can admit more secondary MTs while guaranteeing the VBR requirement. Fig. 7 has the same phenomenon as Fig. 6. In Fig. 8, it can be seen that the call blocking probability at cognitive microcell increases as the average residence time for each secondary MT grows. This is because the longer the residence time for each secondary MT is, the longer the channel holding time for each secondary MT is. Consequently, the existing calls in the geographical area become more and more, and the call blocking probability at cognitive microcell becomes larger.

From Fig. 3 to Fig. 8, it can be concluded that the proposed call admission control not only enhances the total utility for primary heterogeneous networks, but also improves the total

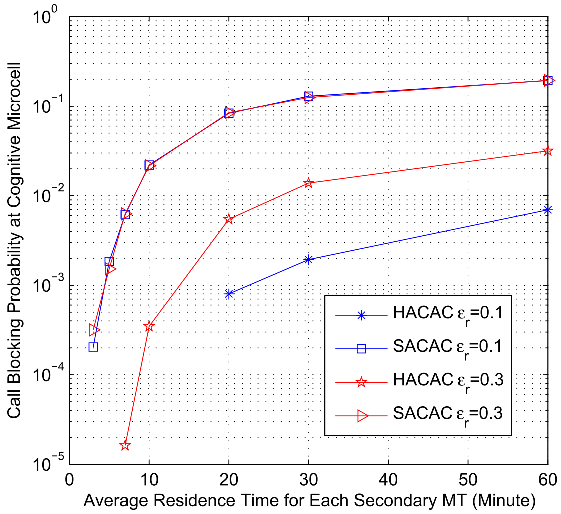

Fig. 8. Call blocking probability at cognitive microcell vs. average residence time for each secondary MT.

utility for secondary MTs. Additionally, the proposed call admission control can reduce the call blocking probability significantly.

VI. Conclusions

In this paper, we have studied the call admission control problem for cognitive heterogeneous networks. The radio subchannel allocation price and network service selection have been adjusted based on the vacant spectrum price at primary heterogeneous networks, so as to maximize the utility for cognitive heterogeneous networks under the QoS constraints. In order to solve the above call admission control problem, we have modeled it as a Stackelberg game framework, and analyzed the call blocking probability based on the queueing theory. In the Stackelberg game framework, primary heterogeneous networks are the multi-leaders and cognitive heterogeneous networks are multi-followers. Then, the joint subchannel allocation price and network selectionproblem at the cognitive heterogeneous networks via the dual-decomposition method, and vacant spectrum prices among primary heterogeneous networks have been designed with the Bertrand game theory. Finally, the equilibrium of the joint spectrum price and subchannel allocation price has been obtained based on the recursive manner, and a call admission control algorithm has been designed. Simulation results demonstrate the proposed algorithm can reduce the call blocking probability significantly.

In this work, we assume a perfect CSI and spectrum sensing. In a realistic scenario, the imperfect CSI is more reasonable [20], [21]. In our future work, we will take the imperfect CSI and its impact on the call admission control performance into consideration.

Appendix A
Proof of Proposition 1

Proof: We focus on proving the convexity of the objective function. Set $y = \sum_{m\in\mathcal{M}} u_m\left(\rho_{nsm}^k\right)$, and the second derivative of y with respective to ρ_{nsm}^k is

$$\frac{\partial^2 y}{\partial\left(\rho_{nsm}^k\right)^2} = -\frac{\left[\eta_1 \sum\limits_{k\in K_{ns}} x_{nsm}\left(1-s_{ns}^k\right)B_{ns}\right]^2}{\left[1+\eta_1 \sum\limits_{k\in K_{ns}} x_{nsm}\rho_{nsm}^k\left(1-s_{ns}^k\right)B_{ns}\right]^2 \ln 2} \leq 0. \tag{34}$$

From (31), y is concave on ρ_{nsm}^k. Since the objective function is convex, and the constraints in Υ constitute a convex set in ρ_{nsm}^k, (20) is a convex programming problem.

Appendix B
Proof of Proposition 2

Proof: Since (20) is a convex optimization problem and the constraints are linear, the slater's condition is satisfied, thus a strong duality exists [29]. Consequently, we adopt the dual decomposition method to obtain the optimal solution, and the Lagrangian function for problem (20) is

$$\begin{aligned} &f\left(\beta_m^{(1)},\beta_m^{(2)},\upsilon_{ns}^k,\rho_{nsm}^k\right) \\ &= \sum_{m\in\mathcal{M}} u_m\left(\rho_{nsm}^k\right) \\ &\quad + \sum_{m\in\mathcal{M}} \beta_m^{(1)}\left(\sum_{n\in\mathcal{N}}\sum_{s\in\mathcal{S}_n}\sum_{k\in\mathcal{K}_{ns}} \rho_{nsm}^k\left(1-s_{ns}^k\right)B_{ns} - B_m^{\min}\right) \\ &\quad + \sum_{m\in\mathcal{M}} \beta_m^{(2)}\left(B_m^{\max} - \sum_{n\in\mathcal{N}}\sum_{s\in\mathcal{S}_n}\sum_{k\in\mathcal{K}_{ns}} \rho_{nsm}^k\left(1-s_{ns}^k\right)B_{ns}\right) \\ &\quad + \sum_{n\in\mathcal{N}}\sum_{s\in\mathcal{S}_n}\sum_{k\in\mathcal{K}_{ns}} \upsilon_{ns}^k\left(1-\sum_{m\in\mathcal{M}_{ns}} \rho_{nsm}^k\left(1-s_{ns}^k\right)\right) \end{aligned} \tag{35}$$

where $\beta_m^{(1)}$, $\beta_m^{(2)}$ and υ_{ns}^k are Lagrangian multipliers.

The dual function, $h\left(\beta_m^{(1)},\beta_m^{(2)},\upsilon_{ns}^k\right)$, is

$$h\left(\beta_m^{(1)},\beta_m^{(2)},\upsilon_{ns}^k\right) = \begin{cases} \min\limits_{\rho_{nsm}^k} \ f\left(\beta_m^{(1)},\beta_m^{(2)},\upsilon_{ns}^k,\rho_{nsm}^k\right) \\ S.t.: \ \ \rho_{nsm}^k \geq 0. \end{cases} \tag{36}$$

With (36), the dual problem is

$$\begin{aligned} OP3: \quad &\max_{\beta_m^{(1)},\beta_m^{(2)},\upsilon_{ns}^k} h\left(\beta_m^{(1)},\beta_m^{(2)},\upsilon_{ns}^k\right) \\ &S.t.: \ \beta_m^{(1)} \geq 0, \beta_m^{(2)} \geq 0, \upsilon_{ns}^k \geq 0. \end{aligned} \tag{37}$$

The Lagrangian function in (35) can be divided into M sub-Lagrangian functions, and each secondary MT solves its Lagrangian function, i.e.,

$$\begin{aligned} &f_m\left(\beta_m^{(1)},\beta_m^{(2)},\upsilon_{ns}^k,\rho_{nsm}^k\right) \\ &\quad = u_m\left(\rho_{nsm}^k\right) \\ &\qquad + \left(\beta_m^{(1)}-\beta_m^{(2)}\right)\sum_{n\in N}\sum_{s\in S_n}\sum_{k\in K_{ns}} \rho_{nsm}^k\left(1-s_{ns}^k\right)B_{ns} \\ &\qquad - \sum_{n\in N}\sum_{s\in S_n}\sum_{k\in K_{ns}} \upsilon_{ns}^k\rho_{nsm}^k\left(1-s_{ns}^k\right). \end{aligned} \tag{38}$$

Consequently, each secondary MT can solve its own utility maximum problem, i.e.,

$$\begin{aligned} OP4: \quad & \max_{\rho_{nsm}^k}: \; f_m\left(\beta_m^{(1)}, \beta_m^{(2)}, \upsilon_{ns}^k, \rho_{nsm}^k\right) \\ & S.t.: \; \rho_{nsm}^k \geq 0. \end{aligned} \tag{39}$$

For the fixed values $\beta_m^{(1)}$, $\beta_m^{(2)}$ and υ_{ns}^k , the optimal subchannel allocation, ρ_{nsm}^k, can be calculated with (40) by applying KKT condition on (39). [3]

$$\frac{\partial f_m\left(\beta_m^{(1)}, \beta_m^{(2)}, \upsilon_{ns}^k, \rho_{nsm}^k\right)}{\partial \rho_{nsm}^k} = 0. \tag{40}$$

From (40), we can obtain

$$\rho_{nsm}^k = \frac{1}{\alpha_{nsm}^k} - \frac{1 + \sum\limits_{k' \neq k \in \mathcal{K}_{ns}} \eta_1 \rho_{nsm}^{k'} \left(1 - s_{ns}^k\right) B_{ns}}{\eta_1 \left(1 - s_{ns}^k\right) B_{ns}}. \tag{41}$$

References

[1] S. Gunawardena and W. Zhuang, "Capacity analysis and call admission control in distributed cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 9, pp. 3110–3120, Sep. 2011.

[2] N. D. Chatzidiamantis, E. Matskani, L. Georgiadis, I. Koutsopoulos, and L. Tassiulas, "Optimal primary-secondary user cooperation policies in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3443–3455, Jun. 2015.

[3] S. Mallick, R. Devarajan, R. A. Loodaricheh, and V. K. Bhargava, "Robust resource optimization for cooperative cognitive radio networks with imperfect CSI," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 907–920, Feb. 2015.

[4] Y. Xu, L. Wang, C. Fischione, and V. Fodor, "Distributed spectrum leasing via vertical cooperation in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 2, pp. 1588–1601, Feb. 2016.

[5] Z. He, S. Mao, and S. Kompella, "A decomposition approach to quality-driven multiuser video streaming in cellular cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 1, pp. 728–739, Jan. 2016.

[6] H. Zhang, X. Chu, W. Guo, and S. Wang, "Coexistence of Wi-Fi and heterogeneous small cell networks sharing unlicensed spectrum," *IEEE Commun. Mag.*, vol. 53, no. 3, pp. 158–164, Mar. 2015.

[7] C. Ghosh, S. Roy, and D. Cavalcanti, "Coexistence challenges for heterogeneous cognitive wireless networks in TV white spaces," *IEEE Wireless Commun.*, vol. 18, no. 4, pp. 22–31, Aug. 2011.

[8] D. Niyato and E. Hossain, "Call admission control for QoS provisioning in 4G wireless networks: Issues and approaches," *IEEE Netw.*, vol. 19, no. 5, pp. 5–11, Sep. 2005.

[9] D. Liu, H. Zhang, W. Zheng, and X. Wen, "The sub-channel allocation algorithm in femtocell networks based on ant colony optimization," in *Proc. IEEE Military Commun. Conf. (MILCOM)*, Oct. 2012, pp. 1–6.

[10] H. Zhang, C. Jiang, N. C. Beaulieu, X. Chu, X. Wen, and M. Tao, "Resource allocation in spectrum-sharing OFDMA femtocells with heterogeneous services," *IEEE Trans. Commun.*, vol. 62, no. 7, pp. 2366–2377, Jul. 2014.

[11] H. Zhang, C. Jiang, N. C. Beaulieu, X. Chu, X. Wang, and T. Q. S. Quek, "Resource allocation for cognitive small cell networks: A cooperative bargaining game theoretic approach," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3481–3493, Jun. 2015.

[12] J. G. Choi, Y. J. Choi, and S. Bahk, "Power-based admission control for multiclass calls in QoS-sensitive CDMA networks," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 468–472, Feb. 2007.

[13] B. Rong, Y. Qian, K. Lu, H. H. Chen, and M. Guizani, "Call admission control optimization in WiMAX networks," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2509–2522, Jul. 2008.

[14] W. Li and X. Chao, "Call admission control for an adaptive heterogeneous multimedia mobile network," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 515–525, Feb. 2007.

[3] In this work, ρ_{nsm}^k is relaxed to a continuous variable [30], i.e., $\rho_{nsm}^k \in [0, 1]$, which can be explained that a fraction of a timeslot at the vacant subchannel k for cognitive network n BS s is allocated to secondary MT m.

[15] M. Ismail, A. Abdrabou, and W. Zhuang, "Cooperative decentralized resource allocation in heterogeneous wireless access medium," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 714–724, Feb. 2013.

[16] O. Habachi, R. El-Azouzi, and Y. Hayel, "A Stackelberg model for opportunistic sensing in cognitive radio networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 5, pp. 2148–2159, May 2013.

[17] Y. Zhang and C. Leung, "Resource allocation in an OFDM-based cognitive radio system," *IEEE Trans. Commun.*, vol. 57, no. 7, pp. 1928–1931, Jul. 2009.

[18] M. Shaat and F. Bader, "Asymptotically optimal resource allocation in OFDM-based cognitive networks with multiple relays," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 892–897, Mar. 2012.

[19] K. Ahuja, Y. Xiao, and M. van der Schaar, "Efficient interference management policies for femtocell networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 9, pp. 4879–4893, Sep. 2015.

[20] B. Guler and A. Yener, "Uplink interference management for coexisting MIMO femtocell and macrocell networks: An interference alignment approach," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 2246–2257, Apr. 2014.

[21] M. Ismail and W. Zhuang, "A distributed multi-service resource allocation algorithm in heterogeneous wireless access medium," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 2, pp. 425–432, Feb. 2012.

[22] A. Zappone and E. A. Jorswieck, "Resource allocation in amplify-and-forward relay-assisted DS/CDMA systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1271–1276, Apr. 2012.

[23] R. Xie, F. R. Yu, H. Ji, and Y. Li, "Energy-efficient resource allocation for heterogeneous cognitive radio networks with femtocells," *IEEE Trans. Wireless Commun.*, vol. 11, no. 11, pp. 3910–3920, Nov. 2012.

[24] Q. Cao, H. V. Zhao, and Y. Jing, "Power allocation and pricing in multiuser relay networks using Stackelberg and bargaining games," *IEEE Trans. Veh. Technol.*, vol. 61, no. 7, pp. 3177–3190, Sep. 2012.

[25] T. Zhang, W. Chen, Z. Han, and Z. Cao, "Hierarchic power allocation for spectrum sharing in OFDM-based cognitive radio networks," *IEEE Trans. Veh. Technol.*, vol. 63, no. 8, pp. 4077–4091, Oct. 2014.

[26] S. Bu and F. R. Yu, "Green cognitive mobile networks with small cells for multimedia communications in the smart grid environment," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2115–2126, Jun. 2014.

[27] S. Y. Jung, S. M. Yu, and S. L. Kim, "Optimization of spectrum allocation and subsidization in mobile communication services," *IEEE Trans. Veh. Technol.*, vol. 65, no. 10, pp. 8432–8443, Oct. 2016.

[28] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

[29] M. Zhu and S. Martinez, "On distributed convex optimization under inequality and equality constraints," *IEEE Trans. Autom. Control*, vol. 57, no. 1, pp. 151–164, Jan. 2012.

[30] J. Huang, V. G. Subramanian, R. Agrawal, and R. Berry, "Joint scheduling and resource allocation in uplink OFDM systems for broadband wireless access networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 2, pp. 226–234, Feb. 2009.

Lei Xu (M'16) received the bachelor's, master's, and Ph.D. degrees from the Nanjing University of Aeronautics and Astronautics, China, in 2006, 2009, and 2012, respectively, all in communication and information system. He is currently an Associate Professor with the School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China. His research interests include 5G wireless network, network analysis, Internet of Things, satellite communication, and radar signal processing.

Ping Wang (M'08–SM'15) received the Ph.D. degree in electrical and computer engineering from the University of Waterloo, Canada, in 2008. She is currently an Associate Professor with the School of Computer Science and Engineering, Nanyang Technological University, Singapore. Her current research interests include resource allocation in wireless networks, cloud computing, and smart grid. She was a co-recipient of the Best Paper Award at the IEEE Wireless Communications and Networking Conference in 2012 and the IEEE International Conference on Communications in 2007.

Qianmu Li received the B.Sc. and Ph.D. degrees from the Nanjing University of Science and Technology, China, in 2001 and 2005, respectively. He is currently a Professor with the School of Computer Science and Engineering, Nanjing University of Science and Technology, China. His research interests include information security and data mining. He received the China Network and Information Security Outstanding Talent Award and multiple Education Ministry Science and Technology Awards.

Yinwei Jiang, photograph and biography not available at the time of publication.