

Research Article

Self-Organized Cognitive Sensor Networks: Distributed Channel Assignment for Pervasive Sensing

Li-Chuan Tseng,¹ Feng-Tsun Chien,¹ Abdelwaheb Marzouki,² Ronald Y. Chang,³
Wei-Ho Chung,³ and ChingYao Huang¹

¹ Department of Electronics Engineering, National Chiao Tung University, Hsinchu 30010, Taiwan

² Institut Mines-Télécom, Télécom SudParis, 91011 Évry, France

³ Research Center for Information Technology Innovation, Academia Sinica, Taipei 11529, Taiwan

Correspondence should be addressed to Feng-Tsun Chien; ftchien@mail.nctu.edu.tw

Received 21 October 2013; Accepted 9 February 2014; Published 17 March 2014

Academic Editor: Luis Javier Garcia Villalba

Copyright © 2014 Li-Chuan Tseng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We study the channel assignment strategy in multichannel wireless sensor networks (WSNs) where macrocells and sensor nodes are overlaid. The WSNs dynamically access the licensed spectrum owned by the macrocells to provide pervasive sensing services. We formulate the channel assignment problem as a potential game which has at least one pure strategy Nash equilibrium (NE). To achieve the NE, we propose a stochastic learning-based algorithm which does not require the information of other players' actions and the time-varying channel. Cluster heads as players in the game act as self-organized learning automata and adjust assignment strategies based on their own action-reward history. The convergence property of the proposed algorithm toward pure strategy NE points is shown theoretically and verified numerically. Simulation results demonstrate that the learning algorithm yields a 26% sensor node capacity improvement as compared to the random selection, and incurs less than 10% capacity loss compared to the exhaustive search.

1. Introduction

In wireless sensor networks [1], spatially distributed, low-power, and low-cost sensor nodes are deployed in a geographical area to monitor the environment. The sensor nodes usually form clusters, and in each cluster there is an energy-rich sensor node acting as the cluster head, while other sensor nodes are referred to as cluster members. A cluster head is a special sensor node with better cognitive radio (CR) [2, 3] functionality and is responsible for the spectrum sensing and the channel assignment among its cluster members.

To enable the various kinds of services [4–6] provided by a pervasive sensing system, proper radio resource management [7] is important. Due to the spectrum scarcity and the ad hoc nature of sensor network deployment, it could be hard to assign licensed bands to sensor networks. Therefore, the CR technology has been considered as a

promising solution to the channel assignment problem of sensor networks. CR technology enables dynamic spectrum access (DSA) of unlicensed users in distributed networks. The key idea for CR operation is to allow the active sensing of the dynamic radio environment so as to improve the spectrum utilization. Akan et al. [8] provided a survey on cognitive radio sensor networks. By utilizing the CR technology, the sensor networks are able to attain high data rate due to available spectrum holes. In addition, dynamic spectrum access helps mitigate the interference incurred by dense deployment of sensor nodes.

Despite the promising features of cognitive sensor networks, the deployment of such heterogeneous networks with sensor clusters underlying the same spectrum as macrocells and in the same geographical area brings new technical challenges. In particular, we are interested in the case of densely populated sensor networks where, due to extensive frequency

reuse, the cochannel interference (CCI) among sensor nodes and the cross-tier interference (between the macrocell and sensor networks) affect the system performance.

In the absence of a central controller, channel assignment in cognitive sensor networks is implemented in a distributed manner. In this paper, we consider the distributed channel assignment for self-organized cognitive sensor networks from a game-theoretic perspective. The main contributions of this paper are summarized as follows.

- (i) We model the femtocell channel assignment problem as an ordinal potential game (OPG). The game considers time-varying channel availability as its external state.
- (ii) We propose a fully decentralized channel assignment algorithm in which the channel is selected by each cluster head independently based on its action-reward history. The convergence property of the algorithm to a pure strategy NE point is verified theoretically as well as numerically.

This paper is organized as follows. We review the related work and compare it with our study in Section 2. In Section 3, the system model for a cognitive sensor network is presented. Section 4 describes the game-theoretic model of the distributed channel assignment problem. Section 5 presents the stochastic learning procedure carried out by the cluster heads. Finally, numerical results are given in Section 6, with the conclusion drawn in Section 7.

Notations. Normal letters represent scalar quantities; uppercase and lowercase boldface letters denote matrices and vectors, respectively. Given a finite set \mathcal{A} , $\Delta(\mathcal{A})$ represents the set of all probability distributions over the elements of \mathcal{A} . $\mathbb{1}_{\{cond\}}$ is the indicator function which equals one if the condition *cond* is satisfied and zero otherwise.

2. Related Works

Distributed channel assignment has been extensively investigated for different networking applications where concurrent transmissions among neighboring wireless links exist.

In an interference avoidance scenario, different channels must be assigned to neighboring links. In femtocell networks [9], different methods have been proposed to assign different spectrum to adjacent femtocells. Examples include distributed random access [9], dynamic frequency planning [10], and clustering [11]. For sensor networks with multiple channels, graph theory-based methods have also been considered [12]. These methods can be viewed as variations of frequency planning and usually rely on negotiations among neighboring links. Graph theory is also a popular approach as the interference condition can be represented as nodes and edges in a graph. In sensor networks, Chowdhury et al. [13] proposed the dynamic channel allocation (DCA) and studied the related protocol design. Yu et al. [14] considered a game theory-based approach which takes into account both the network topology and transmission routing information.

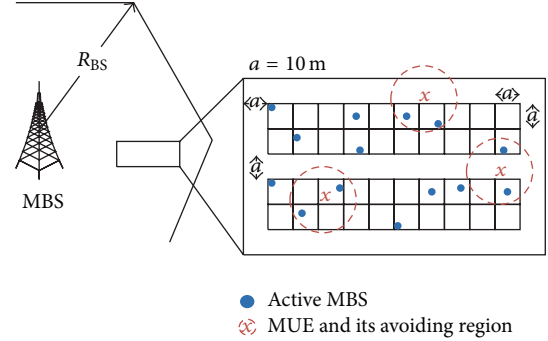


FIGURE 1: Dual-stripe deployment of sensor clusters.

Channel selection for multicell orthogonal frequency division multiple access- (OFDMA-) based networks using graph framework was considered in [15].

On the other hand, in an interference mitigation scenario, mutual interference is tolerated. Channel assignment for cognitive sensor networks has been studied in [16]. Recently, self-organization of distributed agents based upon reinforcement learning (RL) mechanisms [17, 18] has been shown to be effective in the literature. Multiagent Q-learning (MAQL) was applied to femtocell networks in [19, 20]. MAQL involves the actions of other agents as the external state and thus requires the sharing of the knowledge of all agents' actions. The stochastic learning (SL), in contrast, updates the actions of users based on their individual action-reward history. Nie and Comaniciu [21] considered the channel selection in cognitive radio using interference mitigation game formulation. SL was also applied to the opportunistic spectrum access in cognitive radio networks [22] to achieve the Nash equilibrium (NE) strategy. However, fully distributed resource allocation in cognitive sensor networks has not been extensively investigated.

3. System Model

We consider a cognitive radio sensor network consisting of one MBS and N sensor clusters under the coverage of the MBS. The method of sensor node clustering and cluster head selection [23] are also interesting topics but are out of the scope of this paper. Also we consider only the single-hop transmission and omit the multihop routing issue for ad hoc networks [24]. The sensors are deployed in an apartment block with a dual-stripe room layout, as shown in Figure 1.

In our considered system, the medium access control (MAC) function in a cluster resembles that of cellular systems. The time domain is divided into frames, and a frame is further divided into time slots. In each frame, a cluster head allocates its cluster members (i.e., sensor nodes) in different time slots following a time division multiple access (TDMA) rule. For simplicity, we assume that in each slot each cluster head allocates one sensor node over one of the available channels. We emphasize that the proposed method can be easily generalized to the cases with multiple sensor nodes per slot. A sensor node is in idle mode unless the current time

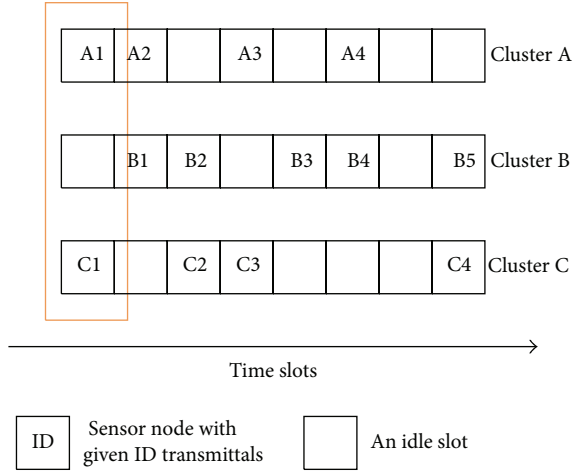


FIGURE 2: Exemplary time slot allocation in a frame. In the first slot, cluster heads A and C assign channels for sensor nodes A1 and C1, respectively.

slot is allocated for it. A cluster head is idle if none of its sensor nodes is transmitting data and active otherwise. We therefore introduce an *active ratio*, which is defined as the percentage of active clusters in a time slot. An exemplary time slot allocation is depicted in Figure 2.

The spectrum is divided into C channels, and the channels may be licensed to different macrocells (a.k.a spectrum owners). By utilizing CR, the sensor nodes access the same frequency band as the macrocell does. Since the sensor nodes are in an energy-tight situation and operate with ultralow power, we assume that the transmission power of a macrocell user equipment (MUE) is much higher than that of the sensors. Thus, the uplink transmission of an MUE will block the nearby sensor nodes using the same spectrum. For cross-tier interference mitigation, we define an *avoiding region* for each MUE. A channel is available to a cluster only if the channel is not assigned to an MUE whose avoiding region covers the cluster head. The channel availability for sensor clusters is expressed as a binary matrix $\mathbf{X} \in \{0, 1\}^{N \times C}$, in which the element $x_{i,c}$ equals one if channel c is available to link i and zero otherwise. The elements of \mathbf{X} follow the Bernoulli distribution and can be described by a probability matrix $\Theta \in [0, 1]^{N \times C}$, where the element $\theta_{i,c}$ is the probability that $x_{i,c} = 1$.

Assuming perfect synchronization in time and frequency, let P_i denote the power of sensor node i , and let $|h_{i,j}|^2$ indicate the link gain between cluster head i and sensor node j . The interference received by cluster head i from sensor node j is given by

$$I_{j \rightarrow i} = \mathbb{1}_{\{a_i(n)=a_j(n)\}} P_j |h_{i,j}|^2, \quad \forall i, j \in \mathcal{N}, \quad (1)$$

where $a_i(n)$ is the action (channel selection) of cluster head i in frame n . For notational brevity, we will hereafter discard the timing dependence of the action $a_i(n)$ in occasions

without ambiguity. Then, the signal-to-interference-and-noise ratio (SINR) at cluster head i can be expressed as

$$\gamma_i = \frac{P_i |h_{i,i}|^2}{\sum_{j=1, j \neq i}^N I_{j \rightarrow i} + \sigma^2}. \quad (2)$$

Consequently, the expected capacity for link i in bits/s/Hz is given by

$$R_i = \theta_{i,a_i} \log_2(1 + \gamma_i). \quad (3)$$

Let $\mathbf{a} = (a_1, \dots, a_N)$ be the channel assignment profile of all active clusters. The global objective of the system is to find the optimal channel selection profile \mathbf{a}_{opt} that maximizes the sum capacity. Formally, consider

$$\mathbf{a}_{\text{opt}} = \underset{\mathbf{a}}{\operatorname{argmax}} \sum_{i=1}^N \theta_{i,a_i} \log_2(1 + \gamma_i). \quad (4)$$

To reflect a practical cognitive sensor network, our system model incorporates the following considerations.

- (1) The uplink resource allocation for MUE is time-varying during the learning period, and the channel availability statistics (i.e., Θ) are fixed but unknown to any sensor nodes.
- (2) There is no centralized controller and the channel selection is performed independently by each cluster head.
- (3) The number of sensor clusters in the system, N , is unknown.

With these considerations, solving (4) is a challenging task, since the only available information for decision making at each individual player is its own action-reward history. Thus, a fully distributed channel selection scheme is proposed.

4. Game-Theoretic Model and Channel Selection

In this section, we present the game-theoretic formulation of the self-organized cognitive sensor network channel selection problem. Our objective is to devise for each cluster head a distributed channel assignment strategy that takes into account the effect of both the sensor-tier and cross-tier interference. We summarize our notations related to the game formulation in Table 1.

4.1. Problem Formulation and Game Model. The channel selection problem described in the previous section can be modeled by a normal-form game with external state, expressed as a 4-tuple:

$$\mathcal{G} = (\mathcal{X}, \mathcal{N}, \{\mathcal{A}_i\}_{i \in \mathcal{N}}, \{u_i\}_{i \in \mathcal{N}}), \quad (5)$$

where \mathcal{X} is the external state (channel availability) space, \mathcal{N} is the set of players (cluster heads), \mathcal{A}_i is the set of actions

TABLE 1: Summary of notations in game-theoretic formulation.

Symbol	Meaning
\mathcal{X}	External state (channel availability)
\mathbf{X}	A realization of external state (channel availability)
\mathcal{N}	Set of players
\mathcal{A}_i	Set of actions of player i
$s_i \in \mathcal{A}_i$	An element of \mathcal{A}_i
$a_i(n) \in \mathcal{A}_i$	Action (channel selection) of player i at slot n
$a_{-i}(n) \in \mathcal{A}_i$	Actions of players except for i at slot n
$\mathcal{P}_i := \Delta(\mathcal{A}_i)$	Set of probability distribution over \mathcal{A}_i
$\mathbf{p}_i(n) \in \mathcal{P}_i$	Mixed strategy of player i at slot n
$r_i(n) \in \mathbb{R}$	Observed utility of player i at slot n
$\tilde{\mathbf{u}}_i(n) \in \mathbb{R}^{ \mathcal{A}_i }$	Estimated utility vector of player i at slot n
(ϵ_i, λ_i)	Learning rates of player i

(selections of channels) that player i can take, and $\{u_i\}_{i \in \mathcal{N}}$ is the utility function of player i that depends on his own action as well as on the actions of other players.

Inspired by [21], the reward function is designed to consider the interference received (inward) and generated (outward) by each link. In this way, the cluster heads implicitly cooperate to reduce the interference generated toward other sensor nodes. We define the *generalized SINR* (gSINR) for player i as

$$\tilde{\gamma}_i = \frac{P_i |h_{i,i}|^2}{\sum_{j=1, j \neq i}^N (I_{j \rightarrow i} + I_{i \rightarrow j}) + \sigma^2}. \quad (6)$$

Then the instantaneous reward function of cluster head i is designed as

$$r_i = \begin{cases} \log_2(1 + \tilde{\gamma}_i), & \text{if } x_{i,a_i} = 1; \\ 0, & \text{otherwise.} \end{cases} \quad (7)$$

By the definition in (7), when the channel is available, the reward is given by Shannon's capacity formula where both inward and outward interference are accounted for. When the channel is not available, the reward is zero. Notice that the calculation of the reward function in (7) relies on the knowledge of other players' action. This leads to overhead due to the required information. The implementation is possible, and discussion on such protocol design can be found in [21]. The *self-organization* claimed in this paper is based on the fact that the action in each time instant is selected by each player independently and simultaneously.

For systems with the channel availability as the external state, the utility function is defined as the expected reward of player i over the external state (i.e., channel availability \mathbf{X}); that is,

$$u_i(a_i, a_{-i}) = \theta_{i,a_i} \log_2(1 + \tilde{\gamma}_i). \quad (8)$$

Furthermore, if the cluster heads are assumed to be selfish and rational players, they will compete to maximize their own individual utility. In fact, a selfish cluster head will not only

maximize the capacity of its own user but also reduce the interference. Formally, the game \mathcal{G} is expressed as

$$(\mathcal{G}) : \max_{a_i \in \mathcal{A}_i} u_i(a_i, a_{-i}), \quad \forall i \in \mathcal{N}. \quad (9)$$

4.2. Analysis of Nash Equilibrium. With the utility function defined in (8), we show the existence of an NE point for the proposed game in the following proposition.

Proposition 1. *The game \mathcal{G} is an ordinal potential game (OPG) which possesses at least one pure strategy NE.*

Proof. Consider the function $\Phi : \times_{i \in \mathcal{N}} \mathcal{A}_i \rightarrow \mathbb{R}_+$:

$$\Phi(\mathbf{a}) = \log_2 \left(1 + \frac{\sum_{k=1}^N P_k |h_{k,k}|^2}{\sum_{k=1}^N \sum_{j=1, j \neq k}^N I_{k \rightarrow j}} \right). \quad (10)$$

Now consider an improvement step made by cluster head i that changes its action unilaterally from a_i to \tilde{a}_i , so that $u_i(\tilde{a}_i, a_{-i}) > u_i(a_i, a_{-i})$. Defining $I_{\tilde{a}_i \rightarrow j} \triangleq \mathbb{1}_{\{\tilde{a}_i = a_j\}} P_i |h_{j,i}|^2$ and $I_{j \rightarrow \tilde{a}_i} \triangleq \mathbb{1}_{\{\tilde{a}_i = a_j\}} P_j |h_{i,j}|^2$, we have

$$\begin{aligned} u_i(\tilde{a}_i, a_{-i}) &> u_i(a_i, a_{-i}) \\ \iff \sum_{j=1, j \neq i}^N [I_{\tilde{a}_i \rightarrow j} + I_{j \rightarrow \tilde{a}_i}] &< \sum_{j=1, j \neq i}^N [I_{a_i \rightarrow j} + I_{j \rightarrow a_i}] \\ \iff \sum_{j=1, j \neq i}^N [I_{\tilde{a}_i \rightarrow j} + I_{j \rightarrow \tilde{a}_i}] + \sum_{j=1, j \neq i}^N \sum_{k=1, k \neq i, j}^N I_{j \rightarrow k} \\ &< \sum_{j=1, j \neq i}^N [I_{a_i \rightarrow j} + I_{j \rightarrow a_i}] + \sum_{j=1, j \neq i}^N \sum_{k=1, k \neq i, j}^N I_{j \rightarrow k}. \end{aligned} \quad (11)$$

Here we have used the fact that when cluster head i changes its action, the effects are only on the interference that it receives ($I_{j \rightarrow i}$) and generates ($I_{i \rightarrow j}$). From (10) and (11), we obtain

$$u_i(\tilde{a}_i, a_{-i}) - u_i(a_i, a_{-i}) > 0 \iff \Phi(\tilde{a}_i, a_{-i}) - \Phi(a_i, a_{-i}) > 0. \quad (12)$$

Therefore, \mathcal{G} is an OPG with potential function Φ , and the existence of a pure strategy NE is always guaranteed [26] since it coincides with the local maxima of the potential function. This completes the proof. \square

Notice that the term $\sum_{k=1}^N \sum_{j=1, j \neq k}^N I_{k \rightarrow j}$ in the potential function Φ denotes the summation of all mutual interference in the sensor network. Therefore, every NE point is the strategy profile, that is, a local maximum of the summed interference.

5. Stochastic Learning Procedure

Here, we discuss obtaining the NE via stochastic learning. As the channel state is time-varying and the action is selected

(1) Initially, set $n = 0$. Set the channel assignment probability vector and utility estimation as

$$p_{i,s_i}(0) = 1/|\mathcal{A}_i|, \quad \hat{u}_{i,s_i}(-1) = 0, \quad \forall i \in \mathcal{N}, s_i \in \mathcal{A}_i.$$

(2) At the beginning of the n th slot, each player selects an action $a_i(n)$ according to the current channel assignment probability $\mathbf{p}_i(n)$.

(3) In each slot, each BS transmits data. At the end of each slot, each BS receives the instantaneous reward $r_i(n)$ specified by (15) depending on the precoding scheme.

(4) All players update their channel assignment probability vector and utility estimation according to the rules:

$$\begin{aligned} \hat{u}_{i,s_i}(n) - \hat{u}_{i,s_i}(n-1) &= \eta_i \mathbb{1}_{\{a_i(n)=s_i\}} (r_i(n) - \hat{u}_{i,s_i}(n-1)), \\ p_{i,s_i}(n+1) &= \frac{p_{i,s_i}(n)(1 + \epsilon_i)^{\hat{u}_{i,s_i}(n)}}{\sum_{s'_i \in \mathcal{A}_i} p_{i,s'_i}(n)(1 + \epsilon_i)^{\hat{u}_{i,s'_i}(n)}}, \end{aligned} \quad (*)$$

where ϵ_i and η_i are the learning rates for action probability and utility estimation, respectively.

ALGORITHM 1: Distributed channel assignment (DCA).

by each player simultaneously and independently in each play, previous algorithms that require complete information (e.g., better response dynamics [26]) may not be applicable here. Thus, we propose a decentralized stochastic learning- (SL-) based algorithm by which the BSs learn toward the equilibrium strategy profile from their individual action-reward history.

To facilitate the development of the SL-based channel selection algorithm, we extend the channel selection game into a mixed strategy form. Let $\mathbf{p}_i(n) = [p_{i,1}(n), \dots, p_{i,C}(n)]^T$, for all $i \in \mathcal{N}$, be the channel selection probability vector for player i , where $p_{i,s_i}(n)$ is the probability that player i selects strategy $s_i \in \mathcal{A}_i$ at slot n . More precisely, using mixed strategies means that the channel assignment of cluster head i is the outcome of a *probabilistic experiment* based on the probability vector \mathbf{p}_i (imagine that each SU rolls a biased dice in each strategy update). The mixed-strategy extension of the utility function is defined upon $\times_{i \in \mathcal{N}} \mathcal{P}_i$, where \mathcal{P}_i is the set of probability distributions over the action space of player i . Let $\mathbf{P}(n) = [\mathbf{p}_1(n), \dots, \mathbf{p}_N(n)]$ be the mixed strategy profile of \mathcal{E} , and let $\psi_i(s_i, \mathbf{P})$ be the expected reward function of player i if he employs pure strategy s_i while other players j , for all $j \in \mathcal{N}, j \neq i$, employ a mixed strategy \mathbf{p}_j ; that is,

$$\psi_i(s_i, \mathbf{P}) = \sum_{a_l, l \neq i} u_i(a_1, \dots, s_i, \dots, a_N) \prod_{j \neq i} p_{j,a_j}. \quad (13)$$

The proposed distributed channel assignment (DCA) algorithm for cognitive sensor networks is described in Algorithm 1.

In each play, the channel selection is based on a probability distribution over the set of channels. After each play, cluster head i obtains the instantaneous reward and updates the mixed strategy (i.e., channel selection vector) $\mathbf{p}_i(n)$ and utility estimation $\hat{u}_i(n)$. Notably, the utility estimation serves as a reinforcement signal so that higher utility induces higher probability in the next play. Furthermore, the proposed learning algorithm is fully distributed, and the channel selection is solely based on individual action-reward experience without a centralized controller. In fact, the proposed algorithm belongs to the combined fully distributed payoff strategy

reinforcement learning (CODIPAS-RL) [27]. The evolution of the mixed strategies is described as follows.

Proposition 2. *If the learning rates are sufficiently small, the sequence $\{\mathbf{P}(n)\}$ converges to \mathbf{P}^* , which is the solution for the following ordinary differential equation (ODE):*

$$\frac{dp_{i,s_i}(t)}{dt} = p_{i,s_i}(t) \left[\psi_i(s_i, \mathbf{P}) - \sum_{s'_i \in \mathcal{A}_i} \psi_i(s'_i, \mathbf{P}) p_{i,s'_i}(t) \right]. \quad (14)$$

Proof. Please refer to [28, Section 4]. \square

The ODE in (14) is actually the ODE of the *replicator dynamics* [29]. An intuitive interpretation is that the probability of taking an action increases if the utility is higher than the average utility over all possible actions and decreases otherwise.

The convergence property of the proposed algorithm is discussed in the following proposition.

Proposition 3. *The SOCA algorithm converges to a pure strategy NE for OPGs if the learning rates are sufficiently small.*

Proof. First, we rewrite the ODE in (14) as follows:

$$\begin{aligned} \frac{dp_{i,s_i}(t)}{dt} &= p_{i,s_i}(t) \sum_{s'_i \in \mathcal{A}_i} p_{i,s'_i}(t) [\psi_i(\mathbf{e}_{s_i}, \mathbf{P}_{-i}) - \psi_i(\mathbf{e}_{s'_i}, \mathbf{P}_{-i})]. \end{aligned} \quad (15)$$

Let $\Psi(\mathbf{P})$ be the expected potential function; that is,

$$\Psi(\mathbf{P}) = \sum_{a_i \in \mathcal{A}_i} \Phi(a_1, \dots, a_N) \prod_{i=1}^N p_{i,a_i}. \quad (16)$$

For OPGs, $\Psi(\mathbf{e}_{s_i}, \mathbf{P}_{-i}) = \partial \Psi(\mathbf{P}) / \partial p_{i,s_i}$ is an increasing function of $\psi_i(\mathbf{e}_{s_i}, \mathbf{P}_{-i})$. Let

$$\begin{aligned} D_{i,s_i,s'_i} &= \psi_i(\mathbf{e}_{s_i}, \mathbf{P}_{-i}) - \psi_i(\mathbf{e}_{s'_i}, \mathbf{P}_{-i}), \\ E_{i,s_i,s'_i} &= \Psi(\mathbf{e}_{s_i}, \mathbf{P}_{-i}) - \Psi(\mathbf{e}_{s'_i}, \mathbf{P}_{-i}), \end{aligned} \quad (17)$$

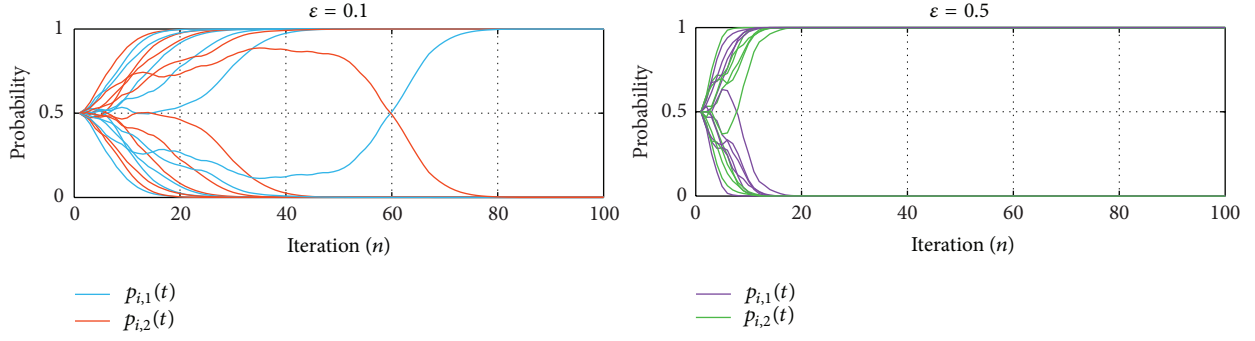


FIGURE 3: Evolution of the mixed strategies (probability of taking different actions) of all players. Each pair of $p_{i,1}(t)$ and $p_{i,2}(t)$ shows the behavior of player i .

and we may write

$$D_{i,s_i,s'_i} > 0 \iff E_{i,s_i,s'_i} > 0. \quad (18)$$

By applying (15) and (18), the derivation of $\Psi(\mathbf{P})$ with respect to t is given by

$$\begin{aligned} \frac{d\Psi(\mathbf{P})}{dt} &= \sum_{i \in \mathcal{N}} \sum_{s_i, s'_i \in \mathcal{A}_i} \frac{\partial \Psi(\mathbf{P})}{\partial p_{i,s_i}} \frac{dp_{i,s_i}}{dt} \\ &= \sum_{i \in \mathcal{N}} \sum_{s_i, s'_i \in \mathcal{A}_i} p_{i,s_i} p_{i,s'_i} \Psi(\mathbf{e}_{s_i}, \mathbf{P}_{-i}) \cdot D_{i,s_i,s'_i} \\ &= \frac{1}{2} \sum_{i \in \mathcal{N}} \sum_{\substack{s_i, s'_i \in \mathcal{A}_i \\ s_i < s'_i}} p_{i,s_i} p_{i,s'_i} E_{i,s_i,s'_i} \cdot D_{i,s_i,s'_i} \\ &\geq 0, \end{aligned} \quad (19)$$

where the last inequality holds since, given the condition in (18), D_{i,s_i,s'_i} and E_{i,s_i,s'_i} always have the same sign.

Thus, Ψ is nondecreasing along the trajectories of the ODE, and asymptotically all the trajectories will be in the set $\{\mathbf{P} \in \mathcal{P} : d\Psi(\mathbf{P})/dt = 0\}$. From (15) and (19), the following is known:

$$\begin{aligned} \frac{d\Psi(\mathbf{P})}{dt} &= 0 \\ \implies p_{i,s_i} p_{i,s'_i} [\psi_i(\mathbf{e}_{s_i}, \mathbf{P}_{-i}) - \psi_i(\mathbf{e}_{s'_i}, \mathbf{P}_{-i})]^2 &= 0 \\ \implies \frac{dp_{i,s_i}}{dt} &= 0, \quad \forall i, s_i, s'_i \\ \implies \mathbf{P} &\text{ is a stationary point of the ODE (14).} \end{aligned} \quad (20)$$

In other words, when starting from an interior point of the simplex of the mixed strategy space \mathcal{P} , the sequence $\mathbf{P}(n)$ converges to a stationary point of the ODE in (15). By Proposition 3, we complete the proof. \square

While the SL-based learning algorithm converges to an NE point when the learning rates approach zero, smaller learning rates lead to a slower convergence. Therefore, the

TABLE 2: Simulation parameters.

Parameter	Value
Minimum distance between nodes	3 m
Carrier frequency	2 GHz
Number of channels	2
Transmission bandwidth of each channel	180 kHz
Path loss and shadowing	Table A.2.1.1.2-8 [25]
Penetration loss	Table A.2.1.1.2-8 [25]
Sensor transmission power	1 mW
Thermal noise	-174 dBm/Hz
Learning rates (default)	$(\lambda_i, \epsilon_i) = (0.1, 0.1)$

choice of the learning rates strikes a trade-off between accuracy and speed and may be determined by training in practice.

6. Numerical Results

For system-level simulations, we consider a cognitive sensor network deployed within the coverage of a cellular network. As in Figure 1, the simulation environment includes one macrocell covering one dual-stripe apartment block. The apartment block contains 40 single-floor apartments. There is one sensor cluster in each apartment. When a sensor cluster is active, its cluster head assigns one channel to cluster members randomly located in the same apartment. Without loss of generality, we consider the channel assignment in the first slot of each frame, in which for each active cluster there is one cluster member. The simulation parameters are listed in Table 2.

6.1. Convergence of the Proposed SL-Based Learning Algorithm. We first study the time-evolving behaviors of the proposed stochastic learning method.

6.1.1. Evolution of Mixed Strategies. Figure 3 shows the evolutions of the channel assignment probabilities (i.e., mixed strategy) using the proposed SL-based algorithm. We consider different learning rates and study the convergence behaviors. It is observed that, with equal initial probability,

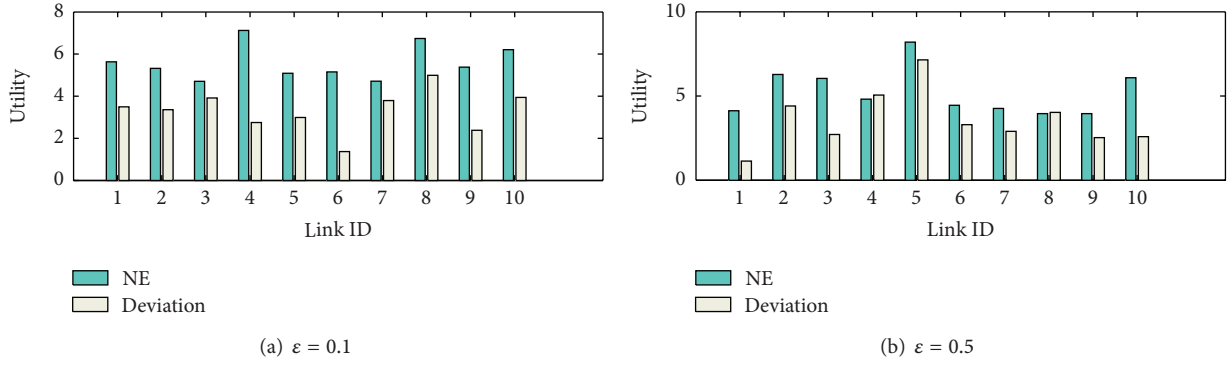


FIGURE 4: Test of unilateral deviation from the resulting strategy profile of each of the 10 players.

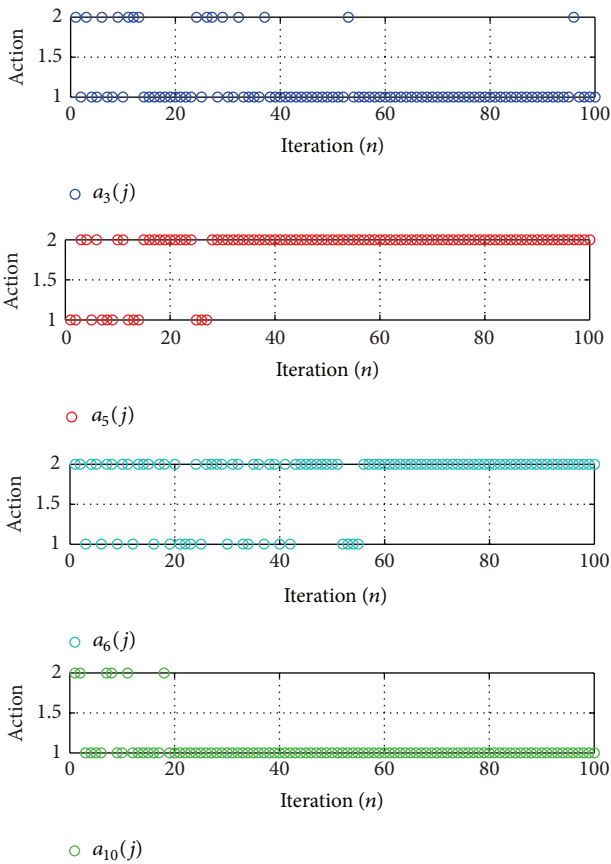


FIGURE 5: Evolution of the actions $a_i(j)$ for some players.

the channel assignment probability converges to a pure strategy (i.e., the probability of choosing one strategy approaches one) in around 80 and 20 iterations for $\epsilon = 0.1$ and $\epsilon = 0.5$, respectively. As expected, larger learning rate results in faster convergence.

6.1.2. Verification of NE. As shown in Figure 3, the convergence toward pure strategy is observed for both $\epsilon = 0.1$ and $\epsilon = 0.5$. An intuitive question to ask is as follows: does the resulting strategy profile achieve the Nash equilibrium? In

Figure 4, we verify the NE property by testing the unilateral deviation with a 25% active ratio and different learning rates. As can be seen from Figure 4(a), when $\epsilon = 0.1$, a unilateral deviation results in lower utility for all players. In other words, the outcome of the learning algorithm is an NE point. On the other hand, when $\epsilon = 0.5$, as shown in Figure 4(b), links number 4 and number 8 both achieve higher throughput by unilateral deviation, and thus the resulting strategy is no longer an NE point. These results reflect the trade-off between accuracy and convergence speed which we mentioned before.

6.1.3. Evolution of Actions. During the learning procedure, the channel assignment is based on probabilistic experiments. When the channel assignment changes in the next frame, the switching between different channels brings overhead since the sensor node needs to be reconfigured. The evolution of actions for selected players is shown in Figure 5. As can be seen, while Figure 3 (Left) reveals that it takes around 80 iterations for all players to converge to pure strategies, the actions seldom change after about 60 iterations in the learning procedure. This suggests that channel switching, if at all happens, usually happens only in the beginning of the entire learning procedure. Actually, our proposed learning algorithm aims at learning the equilibrium strategy in the long run. The channel switching and the incurred sensor node reconfiguration are manageable overheads compared to the long operation time.

6.1.4. Different Active Ratios. We further consider different active ratios and investigate the convergence behaviors under different levels of mutual interference. The results for active ratio of 50% and 75% are shown in Figure 6. We observe that the convergence toward pure strategy takes around 100 and 150 iterations for active ratio of 50% and 75%, respectively. Comparing the case of 25% active ratio in Figure 3 (Left), we see that it takes fewer iterations for densely active networks to converge than for sparsely active sensor networks.

6.2. Capacity Performance

6.2.1. Capacity under Unilateral Deviation. In Figure 4 we have shown that unilateral deviation leads to decreased utility.

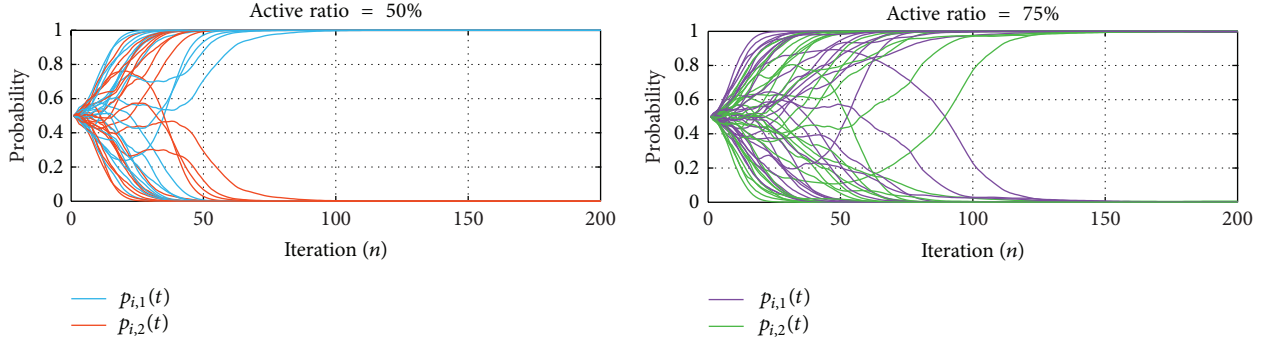


FIGURE 6: Evolution of the mixed strategies (probability of taking different actions) of all players with active ratios of 50% and 75%. Each pair of $p_{i,1}(t)$ and $p_{i,2}(t)$ shows the behavior of player i .

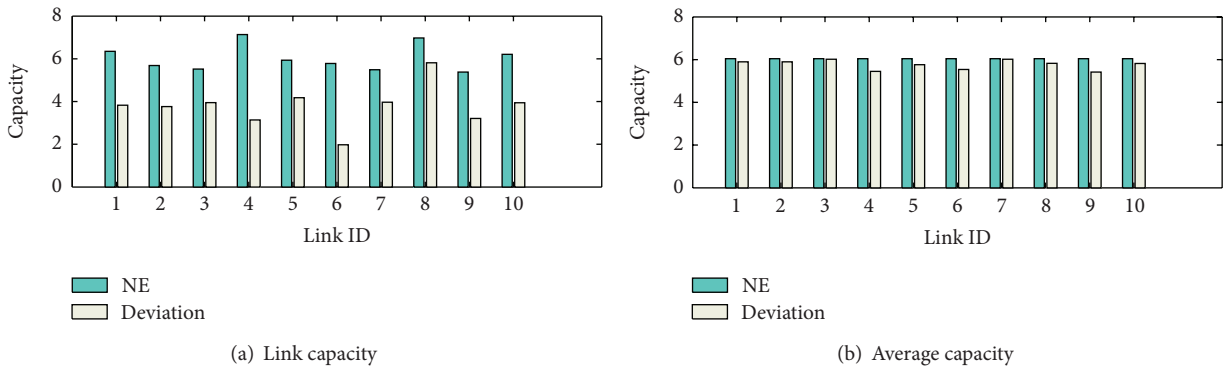


FIGURE 7: Test of unilateral deviation from the resulting strategy profile of each of the 10 players.

While the altruistic utility function design reduces the mutual interference, we are also interested in the performance of Nash equilibrium strategy in terms of the throughput of each cluster as well as the whole system. Therefore, in Figure 7, we test the change in capacity under unilateral deviation from the NE strategy for all players. As depicted in Figure 7(a), there is no significant change in the average capacity per sensor link when only one player unilaterally deviates from its NE strategy. From Figure 7(b) we observe that, for all players, deviation from NE strategy decreases their own capacity.

6.2.2. Comparison with Other Methods. We further compare the performance of the proposed channel selection scheme with two other approaches, namely, random allocation and exhaustive search, described as follows.

- (i) In the random allocation scheme, each cluster head randomly selects a channel for its sensor node in each frame. Neither learning algorithm nor centralized controller is implemented.
- (ii) In the exhaustive search scheme, it is assumed that there exists a centralized controller which knows all system information including the channel gains, the channel availability statistics, and the number of clusters. The channel assignment profile is determined by maximizing the expected sum capacity (i.e., solving (4)).

TABLE 3: Comparison of the capacity and fairness for different channel assignment schemes.

Number of SUs	Proposed	Exhaustive	Random
Active ratio = 25%, R_{avg}	6.0426	6.2433	4.7912
Active ratio = 25%, J	0.9370	0.8512	0.9516
Active ratio = 50%, R_{avg}	4.8375	4.9454	4.0955
Active ratio = 50%, J	0.8855	0.8235	0.9056

The performance of different channel selection schemes is evaluated by the average capacity per sensor node, $R_{\text{avg}} = (1/N) \sum_{i=1}^N R_i$, and the fairness among sensor nodes. In the literature, fairness of resource allocation is usually quantified by Jain's fairness index (JFI) [30], which is defined as

$$J = \frac{\left(\sum_{i=1}^N R_i\right)^2}{N \sum_{i=1}^N R_i^2}. \quad (21)$$

The value of JFI falls in the interval of $[1/N, 1]$, and a higher JFI value indicates better fairness.

The simulation results of average capacity and JFI for different active ratios are summarized in Table 3. We observe that the exhaustive search method results in the best average capacity with the worst fairness. The random selection scheme, in contrast, has the lowest average capacity but good fairness due to its randomness nature. The proposed

method shows well-balanced performance in terms of both average capacity and fairness. The results show the advantages of the proposed method; through the learning procedure toward equilibrium, the capacity of each player is considered and fewer players are sacrificed. If we examine the final channel selection profile, it is observed that, in the progress of convergence toward the NE point, the proposed learning algorithm allocates the mutually interfered users on different channels.

7. Conclusion

In this paper, we have studied the problem of distributed channel assignment in self-organized cognitive sensor networks with unknown channel and unknown number of clusters. We have presented a game-theoretic approach to distributively manage interference and enable the coexistence of sensor and macrocell operations in a scenario where sensor nodes operate in the same spectrum as a cellular system. We modeled channel assignment problem by means of an ordinal potential game. A decentralized stochastic learning algorithm has been proposed. Simulation results have demonstrated the convergence of the algorithm toward a pure strategy Nash equilibrium with sufficiently small learning rates. The proposed method outperforms the random selection scheme in terms of average capacity, while the performance loss compared to the exhaustive search is limited. In addition, its fairness level is comparable to that of the random selection and surpasses the exhaustive search scheme.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

This work was supported in part by the National Science Council, Taiwan, under Grant NSC 102-2218-E-001-001.

References

- [1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–105, 2002.
- [2] B. Wang, Y. Wu, and K. J. R. Liu, "Game theory for cognitive radio networks: an overview," *Computer Networks*, vol. 54, no. 14, pp. 2537–2561, 2010.
- [3] C. Liu, O. Granados, R. Duarte, J. Andrian, and C. Author, "Energy efficient architecture using hardware acceleration for software defined radio components," *Journal of Information Processing Systems*, vol. 8, no. 1, pp. 133–144, 2012.
- [4] J. Ng, "Ubiquitous healthcare: healthcare systems and application enabled by mobile and wireless technologies," *Journal of Convergence*, vol. 3, no. 2, pp. 31–36, 2012.
- [5] Y. Zhu and Q. Jin, "An adaptively emerging mechanism for context-aware service selections regulated by feedback distributions," *Human-Centric Computing and Information Sciences*, vol. 2, no. 1, pp. 1–15, 2012.
- [6] S. Silas, K. Ezra, and E. Blessing Rajasingh, "A novel fault tolerant service selection framework for pervasive computing," *Human-Centric Computing and Information Sciences*, vol. 2, no. 1, pp. 1–14, 2012.
- [7] G. H. Carvalho, A. Anpalagan, I. Woungang, and S. K. Dhurandher, "Energy-efficient radio resource management scheme for heterogeneous wireless networks: a queueing theory perspective," *Energy*, vol. 3, no. 4, 2012.
- [8] O. B. Akan, O. B. Karli, and O. Ergul, "Cognitive radio sensor networks," *IEEE Network*, vol. 23, no. 4, pp. 34–40, 2009.
- [9] K. Sundaresan and S. Rangarajan, "Efficient resource management in OFDMA femto cells," in *Proceedings of the 10th ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc '09)*, pp. 33–42, New York, NY, USA, May 2009.
- [10] D. López-Pérez, A. Valcarce, G. de la Roche, and J. Zhang, "OFDMA femtocells: a roadmap on interference avoidance," *IEEE Communications Magazine*, vol. 47, no. 9, pp. 41–48, 2009.
- [11] A. Hatoum, N. Aitsaadi, R. Langar, R. Boutaba, and G. Pujolle, "FCRA: femtocell cluster-based resource allocation scheme for OFDMA networks," in *Proceedings of the IEEE International Conference on Communications (ICC '11)*, pp. 1–6, Kyoto, Japan, June 2011.
- [12] C. Seneviratne and H. Leung, "A game theoretic approach for resource allocation in cognitive wireless sensor networks," in *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC '11)*, pp. 1992–1997, Anchorage, Alaska, USA, October 2011.
- [13] K. R. Chowdhury, N. Nandiraju, P. Chanda, D. P. Agrawal, and Q.-A. Zeng, "Channel allocation and medium access control for wireless sensor networks," *Ad Hoc Networks*, vol. 7, no. 2, pp. 307–321, 2009.
- [14] Q. Yu, J. Chen, Y. Fan, X. Shen, and Y. Sun, "Multi-channel assignment in wireless sensor networks: a game theoretic approach," in *Proceedings of the IEEE INFOCOM*, pp. 1–9, San Diego, Calif, USA, March 2010.
- [15] R. Y. Chang, Z. Tao, J. Zhang, and C.-C. J. Kuo, "Multicell OFDMA downlink resource allocation using a graphic framework," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 7, pp. 3494–3507, 2009.
- [16] J. Ansari and P. Mähönen, "Channel selection in spectrum agile and cognitive MAC protocols for wireless sensor networks," in *Proceedings of the 8th ACM International Workshop on Mobility Management and Wireless Access (MobiWac '10)*, pp. 83–90, Bodrum, Turkey, October 2010.
- [17] Y. Sung and K. Cho, "A method for learning macro-actions for virtual characters using programming by demonstration and reinforcement learning," *Journal of Information Processing Systems*, vol. 8, no. 3, pp. 409–420, 2012.
- [18] B. J. Oommen, A. Yazidi, and O.-C. Granmo, "An adaptive approach to learning the preferences of users in a social network using weak estimators," *Journal of Information Processing Systems*, vol. 8, no. 2, pp. 191–212, 2012.
- [19] M. Bennis and D. Niyato, "A Q-learning based approach to interference avoidance in self-organized femtocell networks," in *Proceedings of the IEEE GLOBECOM Workshops*, pp. 706–710, Miami, Fla, USA, December 2010.
- [20] A. Galindo-Serrano and L. Giupponi, "Femtocell systems with self organization capabilities," in *Proceedings of the 5th International Conference on Network Games, Control and Optimization (NetGCOP '11)*, pp. 1–7, Paris, France, October 2011.

- [21] N. Nie and C. Comaniciu, "Adaptive channel allocation spectrum etiquette for cognitive radio networks," in *Proceedings of the 1st IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN '05)*, pp. 269–278, Baltimore, Md, USA, November 2005.
- [22] Y. Xu, J. Wang, Q. Wu, A. Anpalagan, and Y.-D. Yao, "Opportunistic spectrum access in unknown dynamic environment: a game-theoretic stochastic learning solution," *IEEE Transactions on Wireless Communications*, vol. 11, no. 4, pp. 1380–1391, 2012.
- [23] B. Singh and D. Lobiyal, "A novel energy-aware cluster head selection based on particle swarm optimization for wireless sensor networks," *Human-Centric Computing and Information Sciences*, vol. 2, no. 1, pp. 1–18, 2012.
- [24] X. Li, N. Mitton, A. Nayak, and I. Stojmenovic, "Achieving load awareness in position-based wireless ad hoc routing," *Journal of Convergence*, vol. 3, no. 3, 2012.
- [25] 3GPP, "E-UTRA: further advancements for E-UTRA physical layer aspects," Tech. Rep. TR 36.814, v 9.0.0, 3GPP, Sophia Antipolis, France, 2010.
- [26] D. Monderer and L. S. Shapley, "Potential games," *Games and Economic Behavior*, vol. 14, no. 1, pp. 124–143, 1996.
- [27] M. A. Khan, H. Tembine, and A. V. Vasilakos, "Game dynamics and cost of learning in heterogeneous 4G networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 1, pp. 198–213, 2012.
- [28] H. Tembine, *Distributed Strategic Learning for Wireless Engineers*, CRC Press, New York, NY, USA, 2012.
- [29] D. Fudenberg and D. Levine, *The Theory of Learning in Games*, vol. 2, The MIT press, Cambridge, Mass, USA, 1998.
- [30] R. Jain, D. Chiu, and W. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," Tech. Rep. TR-301, DEC, Maynard, Mass, USA, 1984.

