Cognitive Base Stations in LTE/3GPP Femtocells: A Correlated Equilibrium Game-Theoretic Approach

Jane Wei Huang, *Student Member, IEEE*, and Vikram Krishnamurthy, *Fellow, IEEE*

***Abstract*—This paper considers downlink spectrum allocation in a long term evolution (LTE) system macrocell which contains multiple femtocells. By incorporating cognitive capabilities into femtocell base stations, the Home evolved Node Bs (HeNBs) can be formulated as secondary base stations seeking to maximize the spectrum utility while minimizing interference to primary base stations (evolved Node-Bs). The competition amongst cognitive HeNBs for spectrum resources is formulated as a non-cooperative game-theoretic learning problem where each agent (HeNB) seeks to adapt its strategy in real time. We formulate the resource block (RB) allocation among HeNBs in the downlink of a LTE system using a game-theoretic framework, where the correlated equilibrium solutions of the formulated game are being investigated. A distributed RB access algorithm is proposed to compute the correlated equilibrium RB allocation policy.**

***Abstract*—LTE/3GPP system, cognitive base stations, femtocells, self-organized network, correlated equilibrium, game-theoretic learning.**

GLOSSARY

3GPP	3rd Generation Partnership Project
4G	4th Generation
AWGN	Additive White Gaussian Noise
CCI	Co-Channel Interference
DFP	Dynamic Frequency Planning
DSL	Digital Subscriber Line
eNB	Evolved Node-B
HeNB	Home Evolved Node-B
LTE	Long Term Evolution
Mbps	Megabit Per Second
OFDMA	Orthogonal Frequency-Division Multiple Access
QoS	Quality of Service
RB	Resource Block
SON	Self-Organized Network
UMTS	Universal Mobile Telecommunication System
WLAN	Wireless Local Area Network

I. INTRODUCTION

An important feature of Long Term Evolution (LTE)/3rd Generation Partnership Project (3GPP) systems [1] is that it allows distributed implementation of femtocells to meet a variety of service requirements. The femtocell access points, denoted as Home evolved Node-B (HeNB) in 3GPP, are low-cost, low-power, plug-and-play cellular base stations. In order to provide broadband connectivity, these HeNBs will need to possess *adaptive/cognitive* facilities. In the October 2010 release of 3GPP, HeNBs are described as self-optimized nodes in a Self-Organized Network (SON) which need to maintain quality of service (QoS) with minimal intervention from the service operator [2]. HeNBs are equipped with cognitive functionalities for load balancing, interference management, random access channel optimization, capacity and coverage optimization, and handover parameter optimization.

With the above motivation, this paper considers spectrum resource allocation in an orthogonal frequency division multiple access (OFDMA) LTE downlink system which consists of a macrocell base station (evolved Node-B (eNB)) and multiple femtocell base stations (HeNBs). By incorporating cognitive capabilities into these self-optimized femtocell base stations, the cognitive HeNBs aim to maximize the spectrum utility by utilizing the unoccupied frequencies while minimizing interference to the eNB (primary base station) in a spectrum overlay LTE system. The unit of spectrum resource to be allocated in a LTE system is called a resource block (RB) and it is comprised of 12 subcarriers at a 15 kHz spacing.

Given the RB occupancy of the eNB, the competition for the spectrum resources among HeNBs can be formulated in a game-theoretic setting [3]. Instead of computing the Nash equilibrium policy of the formulated game, we seek to characterize and compute the *correlated equilibrium* policy set [4], [5]. The set of correlated equilibria is a convex polytope. It includes the set of Nash equilibria – indeed the Nash equilibria are isolated points at the extrema of this set [6], [7]. The set of correlated equilibria [5] is arguably the most natural attractive set for a decentralized adaptive algorithm such as the one considered here, and describes a condition of competitive optimality between agents (cognitive femtocell base stations). It is more preferable than Nash equilibria since it directly considers the ability of agents to coordinate their actions. This coordination can lead to higher performance than if each agent was required to act in isolation. Indeed, Hart and Mas-Colell observe in [8] that, for most simple adaptive procedures, "... there is a natural coordination device: the common history, observed by all players. It is thus not reasonable to expect that, at the end, independence among players will obtain." Since the set of correlated equilibria is convex, fairness between players can be addressed in this domain. Finally, decentralized, online adaptive procedures naturally converge to the correlated equilibria, whereas the same is not true for Nash equilibria (the so-called law of conservation of coordination [9]).

Paper approved by E. G. Larsson, the Editor for Game Theory and Communications Systems Optimization of the IEEE Communications Society. Manuscript received November 23, 2010; revised June 6, 2011.

The authors are with the Department of Electrical and Computer Engineering, The University of British Columbia, 5500-2332 Main Mall, Vancouver, BC V6T 1Z4, Canada (e-mail: {janeh, vikramk}@ece.ubc.ca).

Digital Object Identifier 10.1109/TCOMM.2011.093011.100693

Related work: There are several important issues being addressed in recent literature regarding the deployment of HeNB femtocells in a LTE system. One area of interest is how to mitigate the interferences among HeNBs so as to improve the system performance. In [10] and [11], Lopez-Perez *et al.* used Dynamic Frequency Planning (DFP), an interference avoidance technique, to decrease the inter-cell interference and increase the network capacity by dynamically adapting the radio frequency parameters to the specific scenario. They also verified the performance of the DFP technique in a OFDMA WiMAX macrocells and femtocells system. Choi *et al.* investigated in [12] how to minimize the interference caused by femtocells in an open access (spectrum underlay) network. They showed adaptive open access will maximize the value of a femtocell both to its owner and to the network using numerical results.

Attar *et al.* studied the benefits of developing cognitive base-stations in a UMTS LTE network [13]. Radio resource management protocols are not specified by standards, such as 3GPP's UMTS LTE. Thus, there is considerable flexibility in their design. The insufficiency of traditional coexistence solutions in LTE context is shown in [13]. It is argued that cognitive base-stations are crucial to an efficient and distributed radio resource management of LTE given its distributed architecture. One motivation for such argument, is the lessons learnt from wide-spread deployment of wireless local area network (WLAN) access points. The simple plug-and-play nature of WLAN routers, along with the unlicensed nature of WLAN spectrum access, alleviated the need for time and cost-intensive network plannings. This in turn helped a rapid proliferation of WLAN hotspots. However, as the number of coexisting WLAN networks increases, so does their mutual interfering effect, rendering such simple, selfish coexistence strategies problematic. By incorporating the main three cognitive radio capabilities into the LTE base-stations, which are

1) radio scene analysis
2) online learning based on the feedback from RF environment
3) and agile/dynamic resource access schemes

a successful coexistence strategy among eNBs and HeNBs can be achieved.

Main Results: We formulate the RB allocation among HeNBs in the downlink LTE system as a game in Section II. Given the RB usage of the eNB in a LTE macrocell, the cognitive HeNBs are modelled as selfish players competing to acquire the common frequency resources. This framework borrows the idea of cognitive radio systems [14], [15], where we formulate the cognitive HeNBs as secondary users and the eNB as the primary user in the shared spectrum system. A global utility function is defined to evaluate the overall LTE network performance. To achieve the optimal global utility value in a distributed set-up, we also define a local utility function for each cognitive HeNB. This local utility comprises of components that incorporate self-interest, fairness and power consumption of each HeNB.

Section III-A defines the correlated equilibrium of a game. A RB access algorithm (Algorithm 1) is proposed which converges to the correlated equilibrium solution. This RB access algorithm is based on the regret matching algorithm [8], [16], [17] and it has a distributed nature as each cognitive HeNB does not require the information of other HeNBs. We also prove that this proposed algorithm will converge to the correlated equilibrium set of the formulated game in Section III-C. Finally, numerical examples are give in Section IV.

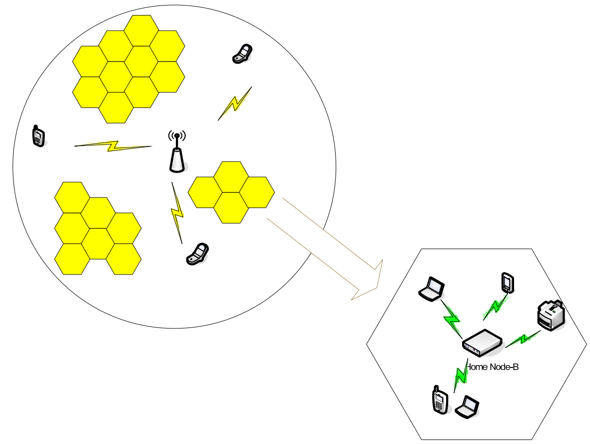

Fig. 1. System schema of a single eNB macrocell which contains multiple HeNB femtocells in a LTE/3GPP network.

II. Resource Allocation Among HeNBs: Problem Formulation

We consider a macrocell area with a number of femtocells randomly deployed by home and offices within a OFDMA LTE network (Fig. 1). By incorporating the cognitive capacities into femtocell base stations, the spectrum occupancy behaviour of the macrocell base station (eNB) can be formulated as a primary base station and that of the femtocell base stations (HeNBs) can be formulated as secondary base stations (cognitive base stations) in a spectrum overlay network[1]. Due to the selfish nature of each base station, the competition for the common spectrum resource among cognitive base stations can be formulated using a game-theoretic framework.

A. System Description

The resource allocation process in LTE networks follows a time slotted system model where each time slot length equals to that of a RB (0.5 ms), with RB being the smallest unit of resource that can be allocated. Let $t \in \{1, 2, \ldots, T\}$ denote the time slot index. T defines the time horizon of the formulated game. Let N denote the total number of available RBs in the system, N^t_{henb} denote the number of RBs occupied by HeNBs at time t and N^t_{eNB} denotes the number of RBs occupied by eNB at time t. As we consider a spectrum overlay system, $N^t_{henb} \leq (N - N^t_{eNB})$.

Let K denote the total number of HeNBs coexisting in the network, $k \in \mathcal{K} = \{1, 2, \ldots, K\}$ denote the user index and $f \in \{1, 2, \ldots N^t_{henb}\}$ denote the RB index occupied by

[1]As described in [13], it is also possible to study an *underlay* cognitive femtocell strategy in which RBs accessed by eNB and HeNBs are not orthogonal. However, as the objective of our analysis is modelling the competition among selfish HeNBs, the proposed solutions can be readily extended to the aforementioned spectrum underlay LTE systems.

cognitive base stations (HeNBs). We use $p_k^t(f) \in \{0,1\}$ to denote the action of the kth HeNB on the fth RB at time t, where 0 represents *not transmit* and 1 represents *transmit*.

Let $\mathbf{p}_k^t = \{p_k^t(1), \ldots, p_k^t(N_{henb}^t)\} \in \mathcal{P}_k$ denote the action of the kth HeNB over all the available RBs to HeNBs at time t, with $\mathcal{P}_k$ denoting the action space of the kth HeNB. $\mathbf{p}^t = \{\mathbf{p}_1^t, \ldots, \mathbf{p}_K^t\} \in \mathcal{P}$ is used to denote the composition of the actions from all the HeNBs at time t. $\mathcal{P}$ is the joint action space of all HeNBs.

Use $s_k^t(f) \in \{1, 2, \ldots, Q_s\}$ to denote channel quality state of the kth HeNB on the fth RB at time t after quantization. For example, the channel quality can be obtained by quantizing a continuous valued channel model comprising of circularly symmetric complex Gaussian random variables into Q_s different states. Let $\mathbf{s}_k^t = \{s_k^t(1), \ldots, s_k^t(N_{henb}^t)\}$ denote the channel state composition of the kth HeNB over all the available RBs and $\mathbf{s}^t = \{\mathbf{s}_1^t, \ldots, \mathbf{s}_K^t\}$ denote the channel state composition over all the HeNBs.

Let $I_k^t(f)$ denote the interference introduced to the kth HeNB at time t on the fth RB. The interference comprises two parts, namely, noise and co-channel interference (CCI). The noise $n_k^t(f)$ is assumed to be additive white Gaussian noise (AWGN) with a noise covariance of $\sigma_k^2(f)$ and the CCI is introduced by having different HeNBs sharing the same RB. An interference matrix $\mathbf{w}^t(f)$ is introduced to model the CCI among all HeNBs on the fth RB at time t. The elements of this interference matrix, i.e., $w_{i,j}^t(f)$ where $i, j \in \mathcal{K}$, denote the cross channel quality between the ith and jth HeNBs. Assuming channel reciprocity, $\mathbf{w}^t(f)$ is a symmetric matrix, i.e., $w_{i,j}^t(f) = w_{j,i}^t(f)$. It is clear that the interference matrix has zero-valued diagonal elements, i.e., $w_{i,i}^t(f) = 0$ for $\forall i \in \mathcal{K}$. The value of $\mathbf{w}^t(f)$ depends on the location of all the HeNBs.

We assume the kth ($k \in \mathcal{K}$) HeNB has the following information at the beginning of a time slot t.

1) N_{henb}^t: the number of available RBs for HeNBs,
2) $\mathbf{s}_k^t$, i.e., the channel state vector of the kth HeNB,
3) $I_k^t(f)$ which is the received interference at HeNB k on RB f at time t,
4) the current demand level of the kth HeNB which is denoted by d_k^t.

N_{henb}^t can be obtained through a fixed broadband access network (e.g., DSL, Cable) as described in [18]. $\mathbf{s}_k^t$ and $I_k^t(f)$ can be obtained by channel sensing mechanism during the guard interval at the beginning of each time slot. d_k^t is used to denote the system resource demand level of HeNB k at time t, it is of the same unit as that of system capacity. d_k^t is determined by the user requirement with in a HeNB cell at time t. In the case that there are more devices transmit data using frequency bandwidth within HeNB k at time t, d_k^t is of higher value. An important characteristic of this model is that radio-specific quantity d_k^t needs only to be known to the kth HeNB cell, thus allowing decentralized resource allocation algorithms. Based on the above information, HeNB k chooses its action vector $\mathbf{p}_k^t$, selfishly, to maximize its local utility function. The definition of utility will be presented in Section II-B.

Note that in the case that HeNBs are owned by different agents and they are so sophisticated to behave maliciously, they can opt not to reveal their true demand levels d_k^t ($k = 1, \ldots, K$) to optimize their own utilities at the cost of reducing the overall system performance. It requires mechanism design theory in order to prevent this from happening. Similar problems has been studies in [19] where the authors applied pricing mechanism to ensure each rational selfish user maximize its own utility function, at the same time optimizing the overall system utility. Reputation based mechanism design is another area of research, where system uses reputation as a tool to motivate cooperation between users and indicate a good behaviour within the network. If a user does not pay attention to its reputation and keep acting maliciously, it will be isolated and discarded. Such reputation based mechanism has been applied in ad hoc networks and sensor networks [20], [21]. However, this paper assumes all the malicious behaviours have been eliminated in the system and each HeNB uses its true demand level d_k^t to compute its utility function.

The distributed decision making process amongst HeNBs defines the action set $\mathbf{p}^t$, which in turn leads to a different realization of interference level $I_k^t(f)$, $\forall\ k \in \mathcal{K}$ and $f = \{1, \ldots, N_{henb}^t\}$. Therefore, the action of one femtocell base station (HeNB) affects the utilities of other femtocells, which motivates the use of game-theoretic approaches to analyze and compute the RB allocation policies among all the HeNBs. In the following subsection, we define the global system utility function and the local utility functions for femtocell base stations.

B. Utility Function

The goal of this paper is to optimize the global resource allocation problem using a decentralized approach. We should demonstrate a connection global utility function and the local utility function that will guide the allocation decision of each HeNB. This connection is presented through the derivation of global and local performance measures. This subsection defines a global utility function to evaluate the overall system performance, based on which a local utility function is defined for each cognitive HeNB. Each HeNB selfishly maximizes its local utility function which does not guarantee the global system performance. We aim to design local utility functions which ensure global system performance quality at the correlated equilibrium of the formulated game.

Let C_k^t denote the capacity of HeNB k at time t. If a capacity-achieving code such as turbo code or low-density parity-check code (LDPC) code is applied for error correction, C_k^t can be expressed as follows using Shannon-Hartley's theorem [22].

$$\begin{aligned} C_k^t &= \sum_{f=1}^{N_{henb}^t} \omega \cdot \log_2 \left[1 + \frac{p_k^t(f) \times s_k^t(f)}{I_k^t(f)} \right] \\ I_k^t(f) &= \sigma_k^2(f) + \sum_{i=1}^{K} w_{k,i}^t(f) \times p_i^t(f), \end{aligned} \tag{1}$$

where ω denotes the bandwidth of a RB. We assume that all the HeNBs treat the interferences as Gaussian noises. Note that $w_{k,k}^t = 0$ for $k = 1, \ldots, K$.

1) Global Utility Function: Since all cognitive HeNBs have equal priority in accessing system resources, a global objective is chosen to maximize the performance of the worst-off HeNB such that the available resources are fairly allocated among HeNBs. Therefore, given the action vector $\mathbf{p}^t$ of all the HeNBs, the global utility function at time t is defined as follows.

$$U_G(\mathbf{p}^t) = \min_{k\in\mathcal{K}} \big[\min(\frac{C_k^t}{d_k^t}, 1)\big]. \quad (2)$$

Here, the term $\min(\frac{C_k^t}{d_k^t}, 1)$ represents the *satisfaction level* of the kth HeNB and it is a function of the instantaneous capacity of the kth HeNB at time t divided by its current demand level. Note that mathematically (2) is equivalent to $U_G(\mathbf{p}^t) = \min_{k\in\mathcal{K}} \frac{C_k^t}{d_k^t}$, operation $\min(\frac{C_k^t}{d_k^t}, 1)$ is applied because the satisfaction level is within the range of $[0, 1]$

The action vector among all the HeNBs at time t $\mathbf{p}^t$ is chosen to maximize the minimum satisfaction level among all the K HeNBs. That is,

$$\mathbf{p}^t = \arg\max_{\mathbf{p}\in\mathcal{P}} U_G(\mathbf{p}) = \max_{\mathbf{p}\in\mathcal{P}} \min_{k\in\mathcal{K}} \big[\min(\frac{C_k^t}{d_k^t}, 1)\big]. \quad (3)$$

This global utility optimization problem (3) aims to maximize the minimum satisfaction level among all the K HeNBs. Note the global utility function can be of other forms, e.g., aiming to maximize the average satisfaction level among all the K HeNBs, in which case the global utility function can be specified as follows (4). However, this paper focuses on the scenario where the global utility is chosen to maximize the worst-off user (2).

$$U_G(\mathbf{p}^t) = \frac{1}{K}\sum_{k\in\mathcal{K}} \big[\min(\frac{C_k^t}{d_k^t}, 1)\big]. \quad (4)$$

The system objective is to find the action vector $\mathbf{p}^t$ so as to maximize the satisfaction level of the worst-off HeNB $U_G(\mathbf{p}^t)$. The proposed approach to achieve the correlated equilibrium policy in a decentralized way, is to allow each cognitive HeNB choosing its action $\mathbf{p}_k^t$ $(k \in \mathcal{K})$ based on the optimization of its local utility function. The question that remains to be answered is how to select a proper local utility function which also ensures a good overall system performance. In the next subsection, we will derive such a utility function and determine its relation to the global utility objective.

2) Local HeNB Utility Function: If each cognitive HeNB has a reasonable estimate of the global utility function, U_G, then the decentralized resource allocation policy can be directly realized by each cognitive HeNB choosing an action which maximizes its estimate of U_G. However, as the global utility function also depends on the private information of other players, i.e., the demand levels d_i^t and actions $\mathbf{p}_i^t$ $(i \in \mathcal{K}$ and $i \neq k)$ of other HeNBs, such an assumption is not practical.

Below, we construct a local utility function U_k for HeNB k $(k \in \mathcal{K})$ consisting of three parts, where each part addresses a certain aspect of the problem at hand.

The first part of the local utility function reflects the *self interest* component of a cognitive HeNB, given by,

$$U_k[1](\mathbf{p}_k^t) = \min\big(\frac{C_k^t}{d_k^t}, 1\big). \quad (5)$$

Maximizing $U_k[1](\mathbf{p}_k^t)$ is equivalent to maximizing HeNB k's portion of the global utility U_G (2). However, a game with (5) as the only component of the local utility function would resemble a congestion game, which may not be solvable in closed form. Moreover, (5) only shows the self interested part of the global objective (2) and it neglects the inter-relation of decisions of each player on the achieved utility of other players. Since each HeNB known only its own demand leven d_k^t and action $\mathbf{p}_k^t$, we induce such interaction through the following two principles.

1) At each time t, the capacity of the kth HeNB C_k^t should not exceeds its demand level d_k^t, as it leaves less resource to other HeNBs.
2) Each HeNB should minimize its transmission power, as higher transmission power decreases the performance of other HeNBs by introducing higher interferences.

The first principle is satisfied by introducing the second component of the local utility function where a penalty is introduced if the choice of resources of a HeNB exceeds its demand level. The details of the second component of the local utility function is shown as follows:

$$U_k[2](\mathbf{p}_k^t) = -\frac{1}{d_k^t}\big(C_k^t - d_k^t\big)^+, \quad (6)$$

where $(x)^+$ denotes the operation $\max(x, 0)$. The second local utility component suppresses greedy HeNB behaviour, and it brings C_k^k closer to its demand d_k^t. This local utility component also helps to maintain the fairness among all the HeNBs.

The third component of the local utility function is used to implement the second principle by considering power as part of the cost of a HeNB, thus, encouraging each HeNB to minimize its power consumption. The details of the third component is shown as follow:

$$U_k[3](\mathbf{p}_k^t) = -\sum_{f=1}^{N_{henb}^t} p_k^t(f). \quad (7)$$

Note that we assume unit power transmission of all the HeNBs on each of the RB in our system model. In this system model, each HeNB is assumed to be a selfish user aiming to maximize its own utility function with the minimum cost. By including the power consumption cost as part of the local utility function, it helps the local utility to represent the global utility.

Based on the above definitions, the local utility function of a HeNB k can be defined as follows.

$$U_k(\mathbf{p}_k^t) = U_k[1](\mathbf{p}_k^t) + \alpha_2 \cdot U_k[2](\mathbf{p}_k^t) + \alpha_3 \cdot U_k[3](\mathbf{p}_k^t), \quad (8)$$

where α_2 and α_3 are the weighting factors introduced to combine the three utility components assuming $U_k[1]$ has a unit weighting factor, i.e., $\alpha_1 = 1$. These weighting factors are necessary because the actual effect of each component is unknown. By carefully adjusting the values of (α_2, α_3), we can change the effect of each of the three utility components, which then enable the local utility to mimic the behaviour of the global utility (2) in the best way. Thus, the question remains to be how to choose (α_2, α_3) which lead to the best overall system performance (2). This paper does not provide a closed form solution to this question, instead, we choose the weight factors according to the numerical studies in (IV). The selected weighting factors (α_2, α_3) does not ensure the maximization of the instantaneous global utility function at time t, instead, it maximizes the expected system performance under different channel realizations, i.e., $\mathbb{E}_{s_1^t, s_2^t, \ldots s_K^t}\{U_G(\mathbf{P^t})\}$.

Recall that a cognitive HeNB k, $\forall k \in \mathcal{K}$, tries to maximize its utility function $U_k(\mathbf{p}_k^t)$ selfishly by choosing the action vector $\mathbf{p}_k^t$ at the beginning of each time slot. In the following sections, we will show the existence of the correlated equilibrium solution, given the distributed decision making process of HeNBs in a static environment.

III. Correlated Equilibrium Solutions with a Game-Theoretic Approach

This section uses game-theoretic approach to formulate the resource allocation among cognitive femtocell base stations (HeNBs) in a static environment, each of HeNB is formulated as a selfish game player aiming to maximize its local utility function (8). A static environment is where the system parameters (e.g., the channel statistics $\mathbf{s}_k^t$ ($k \in \mathcal{K}$), the primary base station behaviour and the number of HeNBs K) are constants or slowly evolve with time. We investigate the correlated equilibrium solution of this static game, which can be obtained via a distributed RB access algorithm. The algorithm is an adaptive variant of the regret matching procedure of [8]. It dynamically adapts the behaviour of HeNBs to time varying environment conditions. We also prove the RB access algorithm converges to the correlated equilibrium set of the defined game.

A. Definition of Correlated Equilibrium

In a K-player (HeNB) game set-up, each HeNB k ($k \in \mathcal{K}$) is a selfish game player aiming to devise a rule for selecting an action vector $\mathbf{p}_k^t$ at each time slot to maximize (the expected value of) its utility function $U_k(\mathbf{p}_k^t)$. Since each player only has control over its own action $\mathbf{p}_k^t$, the optimal action policy depends on the rational consideration of the action policies from other users. We focus on the *correlated equilibrium* solution of the considered game [4], [5], this solution is an important generalization of the Nash equilibrium and is defined as follows.

Definition 3.1: Define a joint policy $\boldsymbol{\pi}$ to be a probability distribution on the joint action space $\mathcal{P} = \mathcal{P}_1 \times \mathcal{P}_2 \ldots \mathcal{P}_K$. Given actions of other players $\mathbf{p}_{-k}^t$, the policy $\boldsymbol{\pi}$ is a correlated equilibrium, if for any alternative policy $\widehat{\mathbf{p}}_k^t \in \mathcal{P}_k$, it holds that,

$$\sum_{\mathbf{p}_{-k}^t \in \mathcal{P}_{-k}} \boldsymbol{\pi}(\mathbf{p}_{-k}^t, \mathbf{p}_k^t) U_k(\mathbf{p}_k^t) \geq \sum_{\mathbf{p}_{-k}^t \in \mathcal{P}_{-k}} \boldsymbol{\pi}(\mathbf{p}_{-k}^t, \widehat{\mathbf{p}}_k^t) U_k(\widehat{\mathbf{p}}_k^t).$$

Correlated equilibrium can be intuitively interpreted as if $\boldsymbol{\pi}$ provides the K players a strategy recommendation from the trusted third-party. The implicit assumption is that the $K-1$ other players follow this recommendation, and player k ask itself whether it is of its best interest to follow the recommendation as well. The equilibrium condition states that there is no deviation rule that could award player k a better expected utility than $\boldsymbol{\pi}$. Any Nash equilibrium can be represented as a correlated equilibrium when users can generate their recommendations independently. One of the advantages of using correlated equilibrium is that it permits coordination among users, generally through observation of a common signal, which leads to an improved performance over a Nash equilibrium [5].

B. Decentralized RB Access Algorithm

The RB access algorithm is an adaptive extension of the regret matching procedure [8], it enables HeNBs to adapt their policies to time varying system environment. Let $\mathbf{H}_k(\mathbf{p}^t)$ denote the regret matrix of the kth cognitive HeNB at time t, it is of size $|\mathcal{P}_k| \times |\mathcal{P}_k|$ with its $(|\mathbf{i}|, |\mathbf{j}|)$th entry $(\mathbf{i}, \mathbf{j} \in \mathcal{P}_k)$ specified as,

$$H_k^{(|\mathbf{i}|,|\mathbf{j}|)}(\mathbf{p}^t) = 1_{(\mathbf{p}_k^t = \mathbf{i})} \times [U_k(\mathbf{j}, \mathbf{p}_{-k}^t) - U_k(\mathbf{i}, \mathbf{p}_{-k}^t)], \tag{9}$$

where $1_{(\cdot)}$ is an indicator function. Furthermore, define $\boldsymbol{\theta}_k^t$ to be the overall regret matrix of HeNB k and it is also of size $|\mathcal{P}_k| \times |\mathcal{P}_k|$. The regret value $\boldsymbol{\theta}_k^t(|\mathbf{j}|, |\mathbf{i}|)$ measures the average gain of user k at time t if k had chosen action $\mathbf{i}$ in the past (from time 0 to t) instead of $\mathbf{j}$. If the gain is positive, k is more likely to switch to action $\mathbf{i}$ in the future, otherwise, k is more likely to stay with $\mathbf{j}$. Specifically, policy transition matrix is a function of the current regret matrix $\boldsymbol{\theta}_k^{t-1}$ as specified in (10). Note this regret based scheme requires users to know the reward for each action, even if that action is not taken. Users updates its policy based on the calculated regret matrix (10). A RB access algorithm is proposed to compute the correlated equilibrium resource allocation policy, the details of which are listed in Algorithm 1.

Algorithm 1 LTE Cognitive HeNB RB Access Algorithm

Step 1 Set $t = 0$; Initialize $\mathbf{p}^0$ and set $\boldsymbol{\theta}_k^0 = \mathbf{H}_k(\mathbf{p}^0)$ for $\forall k \in \mathcal{K}$.
Step 2
for $t = 1, 2, 3, \ldots$ **do**
Action Update: For $\forall k \in \mathcal{K}$, choose $\mathbf{p}_k^t = \mathbf{i}$ with probability

$$\mathbb{P}(\mathbf{p}_k^t = \mathbf{i} | \mathbf{p}_k^{t-1} = \mathbf{j}, \boldsymbol{\theta}_k^{t-1}) = \begin{cases} \frac{\max(\theta_k^{t-1}(|\mathbf{j}|,|\mathbf{i}|),0)}{\mu} & \text{if } \mathbf{i} \neq \mathbf{j} \\ 1 - \sum_{\mathbf{i}\neq\mathbf{j}} \frac{\max(\theta_k^{t-1}(|\mathbf{j}|,|\mathbf{i}|),0)}{\mu} & \text{if } \mathbf{i} = \mathbf{j} \end{cases} \tag{10}$$

Regret Matrices Update: Based on the new action, the overall regret matrices are updated according to the following stochastic approximation algorithm with step size ε^t.

$$\boldsymbol{\theta}_k^t = \boldsymbol{\theta}_k^{t-1} + \varepsilon^t \cdot (\mathbf{H}_k(\mathbf{p}^t) - \boldsymbol{\theta}_k^{t-1}), \quad \forall k \in \mathcal{K}. \tag{11}$$

end for

Algorithm 1 can be summarized as follows. Step 1 initializes the system by setting time index $t = 0$ and the initial values of $\mathbf{p}^0$ and $\boldsymbol{\theta}_k^0$. Step 2 is the main iteration of the algorithm which is composed of two parts: actions update and regret matrices update. A HeNB k chooses its action for time slot t according to the current action $\mathbf{p}_k^{t-1}$ and the regret matrix $\boldsymbol{\theta}_k^{t-1}$. In (10), μ is a constant parameter which is chosen to be $\mu \geq \sum_{\mathbf{i}\neq\mathbf{j}} \max(\theta_k^{t-1}(|\mathbf{j}|, |\mathbf{i}|), 0)$ to ensure the probabilities are non-negative. The choice of step size ε^t used in (11), can be either a decreasing step size $\varepsilon^t = 1/t$ or a small constant step size $\varepsilon^t = \varepsilon$ $(0 < \varepsilon \ll 1)$. If system parameters do not evolve with time, using decreasing step size convergences the algorithm to the correlated equilibrium set with probability one. Using a constant step size enables the algorithm to track of the correlated equilibrium set if system parameters slowly evolve with time.

Depending on the selection of the step size ε^t, the regret matrix $\boldsymbol{\theta}_k^t$ can be rewritten as follows.

$$\begin{aligned}\boldsymbol{\theta}_k^t(|\mathbf{j}|,|\mathbf{i}|) &= \frac{1}{t}\sum_{l\le t,\mathbf{p}_k^l=\mathbf{j}}\big(U_k(\mathbf{i},\mathbf{p}_{-k}^l)-U_k(\mathbf{j},\mathbf{p}_{-k}^l)\big),\\ \text{if} \quad & \varepsilon^t=\frac{1}{t};\\ \boldsymbol{\theta}_k^t(|\mathbf{j}|,|\mathbf{i}|) &= \sum_{l\le t,\mathbf{p}_k^l=\mathbf{j}}\varepsilon(1-\varepsilon)^{t-l}\big(U_k(\mathbf{i},\mathbf{p}_{-k}^l)-U_k(\mathbf{j},\mathbf{p}_{-k}^l)\big),\\ \text{if} \quad & \varepsilon^t=\varepsilon. \end{aligned} \tag{12}$$

The above RB access algorithm is a modification of the regret matching procedure in [8]. In the regret matching approach, the decisions of each HeNB is based on the average history of all past observed performance results. This choice, however, is not desirable in our scenario since the system parameters may vary over time. Instead, our algorithm adapts the regret matrices $\boldsymbol{\theta}_k^t$ ($k\in\mathcal{K}$) according to the updated system parameters which captures the time-varying nature of the system. The decentralized feature of this RB access algorithm permits its implementation among the distributed femtocells in a LTE/3GPP network.

Regret matrix $\boldsymbol{\theta}_k^t$ is one of the key parameters in Algorithm 1, based on which a cognitive HeNB adjusts its future action. The interpretation of the $(|\mathbf{i}|,|\mathbf{j}|)$th entry of $\boldsymbol{\theta}_k^t$ is that measures the average gain that a cognitive HeNB k would have, had it chosen action $\mathbf{j}$ in the past (i.e, in the $(t-1)$th time slot) instead of $\mathbf{i}$.

C. Convergence of RB Access Algorithm

By using the result from [23], we introduce Theorem 3.2 which proves that the RB access algorithm (Algorithm 1) converges to the correlated equilibrium under certain conditions. Let $\boldsymbol{\theta}_k$ denote the regret matrix of the kth HeNB when $t\to\infty$. Use $\Gamma_\Omega(\boldsymbol{\theta}_k)$ to denote the projection of parameter $\boldsymbol{\theta}_k$ on Ω, where Ω is the closed negative orthant of $\mathbb{R}^{|\mathbf{P}_k|\times|\mathbf{P}_k|}$. Let $<x,y>$ denote the inner product of x and y.

Theorem 3.2: The RB access algorithm (Algorithm 1) is ensured to converge to the correlated equilibrium set of the formulated game.

Proof: By using Proposition 3.8 and Corollary 5.7 in [23], we know if every HeNB follows the strategy in Algorithm 1, it is enough to prove that the following inequality, given by (13), holds in order to prove that Algorithm 1 converges to the set of correlated equilibria of the defined game.

$$<\boldsymbol{\theta}_k-\Gamma_\Omega(\boldsymbol{\theta}_k),\boldsymbol{\theta}_k^{'}-\Gamma_\Omega(\boldsymbol{\theta}_k)>\le 0. \tag{13}$$

□

Condition (13) is originated from the Blackwell's sufficient condition for approachability [24]. Therefore, we only need to demonstrate that (13) holds in order to prove the convergence of RB access algorithm.

Note that the negative orthant Ω is a convex set. The left hand side of (13) can be expressed as,

$$\begin{aligned}& <\boldsymbol{\theta}_k-\Gamma_\Omega(\boldsymbol{\theta}_k),\boldsymbol{\theta}_k^{'}-\Gamma_\Omega(\boldsymbol{\theta}_k)>\\ =\ & <\boldsymbol{\theta}_k-\Gamma_\Omega(\boldsymbol{\theta}_k),\boldsymbol{\theta}_k^{'}>-<\boldsymbol{\theta}_k-\Gamma_\Omega(\boldsymbol{\theta}_k),\Gamma_\Omega(\boldsymbol{\theta}_k)>,\end{aligned}$$

where $<\boldsymbol{\theta}_k-\Gamma_\Omega(\boldsymbol{\theta}_k),\Gamma_\Omega(\boldsymbol{\theta}_k)>=0$ due to the definition of projection. Thus, in order to establish (13) we need to prove $<\boldsymbol{\theta}_k-\Gamma_\Omega(\boldsymbol{\theta}_k),\boldsymbol{\theta}_k^{'}>\le 0$.

Let us construct a Markov chain with the transition probability specified in (10) and use $\tau_{|\mathbf{i}|}$ ($\mathbf{i}\in\mathcal{P}_k$) to denote the stationary distribution of such a Markov chain. $\tau_{|\mathbf{i}|}$ can be specified as follows.

$$\tau_{|\mathbf{i}|}=\sum_{|\mathbf{j}|\neq|\mathbf{i}|}\tau_{|\mathbf{j}|}\frac{\theta_k^+(|\mathbf{j}|,|\mathbf{i}|)}{\mu}+\tau_{|\mathbf{i}|}\cdot\big(1-\sum_{|\mathbf{i}|\neq|\mathbf{j}|}\frac{\theta_k^+(|\mathbf{i}|,|\mathbf{j}|)}{\mu}\big), \tag{14}$$

where μ is a constant chosen to be $\mu>\sum_{|\mathbf{j}|\neq|\mathbf{i}|}\theta_k^+(|\mathbf{j}|,|\mathbf{i}|)$. Then, (14) is equivalent to the following equation.

$$\sum_{|\mathbf{j}|\neq|\mathbf{i}|}\tau_{|\mathbf{j}|}\theta_k^+(|\mathbf{j}|,|\mathbf{i}|)=\tau_{|\mathbf{i}|}\sum_{|\mathbf{j}|\neq|\mathbf{i}|}\theta_k^+(|\mathbf{i}|,|\mathbf{j}|). \tag{15}$$

Since the projection is on the negative orthant Ω, $\boldsymbol{\theta}_k-\Gamma_\Omega(\boldsymbol{\theta}_k)=\boldsymbol{\theta}_k^+$. $<\boldsymbol{\theta}_k-\Gamma_\Omega(\boldsymbol{\theta}_k),\boldsymbol{\theta}_k^{'}>$ can then be written as

$$\begin{aligned}& <\boldsymbol{\theta}_k-\Gamma_\Omega(\boldsymbol{\theta}_k),\boldsymbol{\theta}_k^{'}>=<\boldsymbol{\theta}_k^+,\boldsymbol{\theta}_k^{'}>\\ =\ & \sum_{\mathbf{j}}\sum_{\mathbf{i}\neq\mathbf{j}}\theta_k^+(|\mathbf{j}|,|\mathbf{i}|)\big[U_k(\mathbf{i},\mathbf{p}_{-k})-U_k(\mathbf{j},\mathbf{p}_{-k})\big]\tau_{|\mathbf{j}|}\\ =\ & \sum_{\mathbf{i}\neq\mathbf{j}}\sum_{\mathbf{j}}\theta_k^+(|\mathbf{i}|,|\mathbf{j}|)U_k(\mathbf{j},\mathbf{p}_{-k})\tau_{|\mathbf{i}|}\\ & -\sum_{\mathbf{j}}\sum_{\mathbf{i}\neq\mathbf{j}}\theta_k^+(|\mathbf{j}|,|\mathbf{i}|)U_k(\mathbf{j},\mathbf{p}_{-k})\tau_{|\mathbf{j}|}\\ =\ & \sum_{\mathbf{j}}\Big[\sum_{\mathbf{i}\neq\mathbf{j}}\theta_k^+(|\mathbf{i}|,|\mathbf{j}|)\tau_{|\mathbf{i}|}-\sum_{\mathbf{i}\neq\mathbf{j}}\theta_k^+(|\mathbf{j}|,|\mathbf{i}|)\tau_{|\mathbf{j}|}\Big]U_k(\mathbf{j},\mathbf{p}_{-k})\\ =\ & 0.\end{aligned} \tag{16}$$

Therefore, the condition stated in (13) is proved to hold. This concludes the proof of Theorem 3.2.■

D. Correlated Equilibrium under Dynamic Environments and Curse of Dimensionality

In the case that system contains large number of active mobile users, it causes high dynamic macro base station behaviour. In which scenario, the problem can be described as: resource allocation among femtocells (HeNBs) in a OFDMA LTE downlink system under a dynamic environment where the resource occupancy behaviour of the primary base station (system state) is varying quickly while other system parameters (e.g., number of HeNBs) are constants or evolving slowing.

By formulating the dynamic of system state as a Markov chain, it requires Markov game-theoretic approach to formulate the resource allocation problem among femtocell base stations (HeNBs) as a Markov game. Different from static game, system state and state transition probabilities are important elements in dynamic games as they abstract the time-varying nature of a dynamic environment. A reasonable choice of the system state is $[N_{henb}^t,\mathbf{s}_1^t,\ldots,\mathbf{s}_K^t]$ which is composed of the number of available RBs for HeNBs and the channel states of HeNBs. By defining the correlated equilibrium or Nash equilibrium of such a dynamic stochastic game, different optimization algorithms can be used to compute the equilibrium transmission polices. E.g., [25] proposed iterative value optimization algorithm and stochastic approximation algorithm to compute the Nash equilibrium policies in the formulated Markovian game.

Potential applications notwithstanding, there remains substantial hurdles in the application of dynamic stochastic games as a modelling tools in practice. Discrete-time stochastic games with finite number of states are central to the analysis of strategic interactions among selfish HeNBs in dynamic environment. The usefulness of discrete-time games, however, is limited by their computational burden; in particular, there is "curse of dimensionality". In a discrete-time dynamic stochastic game, each game player (HeNB) is distinguished by an individual state at each time slot. The system state is a vector encoding the number of players with each possible value of the individual state variable $[N^t_{henb}, \mathbf{s}^t_1, \ldots, \mathbf{s}^t_K]$. The system state is exponential to the number of the HeNBs in the system. How to efficiently reduce the state space is yet an issue to solve before the implementation of stochastic games in LTE systems. One direction of research is to consider continuous-time stochastic game models. E.g., [26] aims to reduce the dimensionality by exploring the alternative continuous-time stochastic games with a finite number of states and show that the continuous time has substantial advantages.

IV. Numerical Examples

Algorithm 1 is designed to compute the correlated equilibrium policies for cognitive HeNBs in the downlink of a OFDMA LTE/3GPP system. This section illustrates the performance of the RB access algorithm (Algorithm 1) in a game set-up. For demonstration purpose, we consider $K = 6$ HeNBs. For the kth HeNB ($k \in \mathcal{K}$), its channel quality at any RB f ($f \in \mathcal{F}$) is $s_k(f) \in \{1, 2, 3\}$, its demand level belongs to the set $d_k \in \{10, 20, 30, 40\}$. The action of the kth HeNB at the fth RB is specified as $p_k(f) \in \{0, 1\}$, where 0 represents no transmission and 1 represents transmit. In the simulation set-up, we specify the noise covariance to be $\sigma^2_k(f) = 0.1$ ($\forall k \in \mathcal{K}$ and $\forall f \in \mathcal{F}$). In this simulation model, the number of available RB for HeNBs is fixed to be $N_{henb} = 10$.

Note that a LTE eNB macrocell usually consists large number of HeNB femtocells. The reason we specify $K = 6$ HeNBs and $N_{henb} = 10$ in our simulation is because the policy space grows exponentially with the number of available RBs for HeNBs under the current problem formulation. Thus, it is necessary to reduce the action space dimensionality before the widely application of the proposed algorithm. In the presence of large number of HeNBs, HeNBs can be lumped into a single other player (named a composite player). The action space of that composite player can be chosen appropriately and the RB access algorithm can then be applied.

In first example, we are going to investigate the impact of the pricing parameters (α_2, α_3) in the local utility function (8) on different global utilities (2) and (4). The pricing parameters is chosen to ensure the global system performance. This is an off-line calculation procedure. The simulation results in

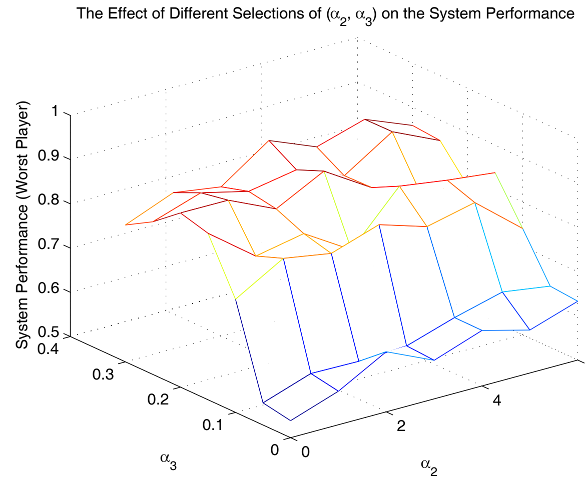

Fig. 2. The effect of different values of (α_2, α_3) defined in (8) on the global system performance specified in (2) .

Fig. 2 and Fig. 3 are averaged over 1000 iterations, where 20 different scenarios with different noise $n_k(f)$, channel states $\mathbf{s}_k$ are being considered in each iteration. In both figures, x-axis and y-axis denote α_2 and α_3, respectively. Z-axis denotes the global system performance specified in (2) in Fig. 2, while it denotes the system average performance specified in (4) in Fig. 3.

Based on Fig. 2, we notice the completely selfish behaviours from HeNBs ($\alpha_2 = 0, \alpha_3 = 0$) do not ensure the optimum of global system performance; while $\alpha_2 = 5$, $\alpha_3 = 0.25$ lead the least satisfaction level among HeNBs to 0.9, i.e., $U_G(\mathbf{p}^t) = 0.9$. We will specify $\alpha_2 = 5$ and $\alpha_3 = 0.25$ in the simulation for Fig. 4.

Fig. 3 shows the system average performance (4) is less sensitive to the change of pricing parameters (α_2, α_3). System average performance is of similar level when the pricing parameters are of the range $0 \le \alpha_2 \le 7$, $0.05 \le \alpha_3 \le 0.5$. Thus, in the case we choose (4) as the global utility function, the selection of (α_2, α_3) is not unique, they can be of any values within the above range. It can also be noticed from Fig. 3 that the system average performance decrease dramatically when $\alpha_2 > 7$. This can be explained as follows: if the power consumption cost is greater than a certain threshold, a HeNB will choose not to transmit as the payoff is less than the cost. Thus, a very high power consumption cost weighting factor (α_3) can have a negative effect on HeNB performances.

The next example (Fig. 4) compares the performance of the proposed RB access algorithm with the existing "Best Response" algorithm with a global utility function specified in (2). In the simulation, we use small constant step size and it is specified as $\varepsilon^t = \varepsilon = 0.05$. The results are averaged over 50 scenarios and there are 2000 iterations.

Best Response is a simple case where each HeNB chooses its action $\mathbf{p}^t_k$ at each time slot solely maximizing its local utility (8) and each HeNB assumes that the actions of other HeNBs are fixed. Thus, Best Response is a special case of the proposed RB access algorithm with the step size chosen to be $\varepsilon^t = 1$. The action update in the Best Response is not a function of the previous regret $\boldsymbol{\theta}^{t-1}_k$ but only a function of the current instantaneous regret matrix $\mathbf{H}_k(\mathbf{p}^t)$. From Fig. 4

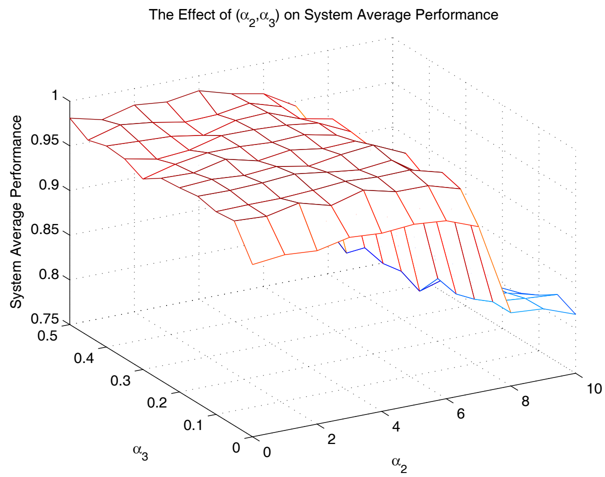

Fig. 3. The effect of different values of (α_2, α_3) defined in (8) on the global system average performance specified in (4).

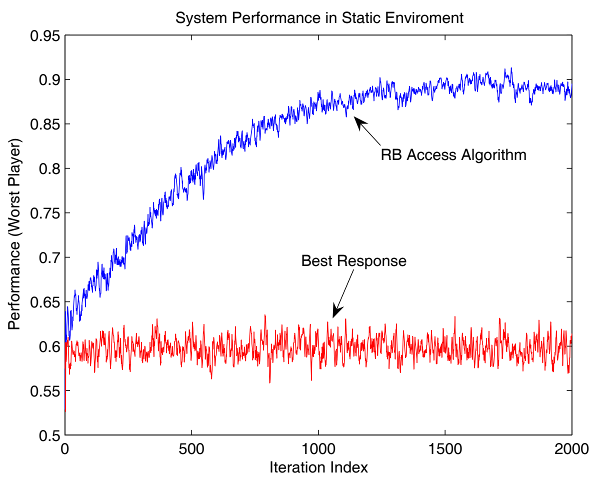

Fig. 4. Performance comparison between RB access algorithm (Algorithm 1) and the "Best Response" algorithm.

we can see that the system performance using the RB access algorithm reaches 0.9 after 1400 iterations and the result from the "Best Response" algorithm stays at around 0.6. It can be seen that the RB access algorithm (Algorithm 1) improves the system performance greatly compared to the "Best Response" algorithm. We can also observe from Fig. 4 that both RB access algorithm and "Best Response" algorithm do not converge to constant values, it is because transmission policy $\mathbf{p}^t$ converges to a correlated equilibrium set when $t \to \infty$ and the correlated equilibrium set has more than one correlated equilibrium policy.

V. Conclusions

We have proposed implementation of cognitive femtocell base stations for resource block allocation in the downlink of a eNB macrocell LTE/3GPP system. By considering the eNB as the primary base station, the HeNBs are formulated as multiple secondary base stations competing for spectrum resources. The RB allocation problem is formulated in a static environments, using static game framework. A RB access algorithm is proposed to compute the correlated equilibrium policy in such a environment. We also prove that the RB access algorithm converges to the correlated equilibrium set of the formulated game. Numerical examples are used to verify the performances of the proposed algorithm.

VI. Acknowledgements

The authors would like to thank Dr. Alireza Attar for his valuable suggestions and help.

References

[1] 3GPP, "Technical Specification and Technical Reports for a UTRAN-based 3GPP system," vol. TR21.101 v0.0.8, 2009. Available: http://www.3gpp.org/ftp/specs/html-info/FeatureOrStudyItemFile-400035.htm.
[2] ——, "3GPP work items on Self-Organizing Networks," Oct. 2010. Available: http://www.3gpp.org/ftp/Information.
[3] B. Wang, Y. Wu, and K. J. R. Liu, "Game theory for cognitive radio networks: an overview," *International J. Comput. Telecommun. Netw.*, vol. 54, no. 14, pp. 2537–2574, Oct. 2010.
[4] R. J. Aumann, "Subjectivity and correlation in randomized strategies," *J. Mathematical Economics*, vol. 1, pp. 67–96, Mar. 1974.
[5] ——, "Correlated equilibrium as an expression of Bayesian rationality," *Econometrica*, vol. 55, no. 1, pp. 1–18, 1987.
[6] O. Morgenstern and J. v. Neumann, *The Theory of Games and Economic Behavior*. Princeton University Press, 1947.
[7] J. Nash, "Equilibrium points in n-person games," in *Proc. National Academy Sciences*, 1950, pp. 48–49.
[8] S. Hart and A. Mas-Colell, "A simple adaptive procedure leading to correlated equilibrium," *Econometrica*, vol. 68, no. 5, pp. 1127–1150, Sep. 2000.
[9] ——, "Uncoupled dynamics do not lead to Nash equilibrium," *American Economic Review*, vol. 93, no. 5, pp. 1830–1836, Dec. 2003.
[10] D. Lopez-Perez, G. d. l. Roche, A. Valcarce, A. Juttner, and J. Zhang, "Interference avoidance and dynamic frequency planning for WiMAX femtocells network," in *Proc. IEEE Singapore International Conf.*, 2008, pp. 1579–1584.
[11] D. Lopez-Perez, A. Juttner, and J. Zhang, "Optimisation methods for dynamic frequency planning in OFDMA networks," in *Proc. Netw.*, Sep. 2008, pp. 1579–1584.
[12] D. Choi, P. Monajemi, S. Kang, and J. Villasenor, "Dealing with loud neighbors: the benefits and tradeoffs of adaptive femtocell access," in *Proc. IEEE GLOBECOM*, 2008, p. 15.
[13] A. Attar, V. Krishnamurthy, and O. Namvar, "Interference management using cognitive base-stations for UMTS LTE," to appear in *IEEE Commun. Mag.*, 2011.
[14] J. Mitola III, "Cognitive radio for flexible mobile multimedia communications," in *Proc. IEEE International Workshop Mobile Multimedia Commun.*, 1999, pp. 3–10.
[15] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
[16] V. Krishnamurthy, M. Maskery, and G. Yin, "Decentralized adaptive filtering algorithms for sensor activation in an unattended ground sensor network: a correlated equilibrium game theoretic analysis," *IEEE Trans. Signal Process.*, vol. 56, no. 12, pp. 6086–6101, Dec. 2008.
[17] M. Maskery, V. Krishnamurthy, and Q. Zhao, "Decentralized dynamic spectrum access for cognitive radios: cooperative design of a non-cooperative game," *IEEE Trans. Commun.*, vol. 57, no. 2, pp. 459–469, Feb. 2009.
[18] 3GPP, "Overview of 3GPP," vol. Release 9 V0.1.1, Sep. 2010. Available: http://www.3gpp.org/ftp/Information.
[19] J. W. Huang and V. Krishnamurthy, "Truth revealing opportunistic scheduling in cognitive radio systems," in *Proc. SPAWC*, June 2009, pp. 91–95.
[20] K. Ren, T. Li, Z. Wan, R. H. Deng, and K. Kim, "Highly reliable trust establishment scheme in ad hoc networks," *J. Comput. Telecommun. Netw.*, vol. 45, no. 6, pp. 687–699, Aug. 2004.
[21] H. Chen, "Task-based trust management for wireless sensor networks," *International J. Security Appl.*, vol. 3, no. 2, pp. 21–26, Apr. 2009.
[22] I. C. Wong and B. L. Evans, "Optimal downlink of OFDMA resource allocation with linear complexity to maximize ergodic rates," *IEEE Trans. Wireless Commun.*, vol. 7, no. 3, pp. 962–971, Mar. 2008.

[23] M. Benam, J. Hofbauer, and S. Sorin, "Stochastic approximations and differential inclusions—part II: applications," *Mathematics Operation Research*, vol. 31, no. 4, pp. 673–695, Nov. 2006.
[24] D. Blackwell, "An analog of the minmax theorem for vector payoffs," *Pacific J. Mathematics*, vol. 6, pp. 1–8, 1956.
[25] J. W. Huang and V. Krishnamurthy, "Transmission control in cognitive radio as a Markovian dynamic game—structural result on randomized threshold policies," *IEEE Trans. Commun.*, vol. 58, no. 2, pp. 301-310, Feb. 2010.
[26] U. Doraszelski and K. L. Judd, "Avoiding the curse of dimensionality in dynamic stochastic games," under revision at *Rand J. Economics*, 2010.

Jane Wei Huang (S'07) received her bachelor's degree from Zhejiang University, China in 2005, and M.Phil. from Hong Kong University of Science and Technology, Hong Kong in 2007. She is currently a Ph.D. student in the University of British Columbia. Her research interests include game theory, Markov decision process, cognitive radio, and sensor networks.

Vikram Krishnamurthy (S'90-M'91-SM'99-F'05) was born in 1966. He received his bachelor's degree from the University of Auckland, New Zealand in 1988, and Ph.D. from the Australian National University, Canberra, in 1992. He is currently a professor and holds the Canada Research Chair at the Department of Electrical Engineering, University of British Columbia, Vancouver, Canada. Prior to 2002, he was a chaired professor at the Department of Electrical and Electronic Engineering, University of Melbourne, Australia, where he also served as deputy head of department. His current research interests include computational game theory, stochastic dynamical systems for modeling of biological ion channels and stochastic optimization and scheduling.

Dr. Krishnamurthy has served as associate editor for several journals including IEEE TRANSACTIONS AUTOMATIC CONTROL, IEEE TRANSACTIONS ON SIGNAL PROCESSING, IEEE TRANSACTIONS AEROSPACE AND ELECTRONIC SYSTEMS, IEEE TRANSACTIONS NANOBIOSCIENCE, and *Systems and Control Letters*. From 2009-2010 he serves as distinguished lecturer for the IEEE signal processing society. From 2010, he serves as editor in chief of IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING.