Learning Based Mechanisms for Interference Mitigation in Self-Organized Femtocell Networks

Mohsin Nazir*, Mehdi Bennis*, Kaveh Ghaboosi*, Allen B. MacKenzie†, and Matti Latva-aho*
*Centre for Wireless Communications, University of Oulu, Finland
†Department of Electrical and Computer Engineering, Virginia Tech, USA
Email: {mohsin.nazir, mehdi.bennis, kaveh.ghaboosi, matti.latva-aho}@ee.oulu.fi, mackenab@vt.edu

Abstract—We introduce two mechanisms for interference mitigation, inspired by evolutionary game theory and machine learning to support the coexistence of a macrocell network underlaid with self-organized femtocell networks. In the first approach, stand-alone femtocells choose their strategies, observe the behavior of other players, and make the best decision based on their instantaneous payoff, as well as the *average* payoff of all other femtocells. We formulate the interactions among selfish femtocells using evolutionary games and demonstrate how the system converges to an equilibrium. In contrast, in the Reinforcement-Learning (RL) approach, information exchange among femtocells is no longer possible and hence each femtocell adapts its strategy and gradually learns by interacting with its environment (i.e., neighboring interferers) through trials-and-errors. Our investigations reveal that through learning, femtocells are able to self-organize by relying only on *local* information, while mitigating the interference towards the macrocell network. Besides, a trade-off exists where faster convergence is obtained in the evolutionary case as compared to the RL approach, at the expense of more side information. Finally, it is shown that femtocells face an interesting tradeoff of exploration versus exploitation in their learning processes.

I. INTRODUCTION

Recently, a new type of indoor Base Station (BS), called femtocell, has gained the attention of the industry [1], [6]. A femtocell is a low-cost and low-power BS deployed by the end-users, designed to extend indoor coverage. Femtocells are connected to the network of the operator over a backhaul connection such as Digital Subscriber Line (DSL) or optical fiber. Meanwhile, femtocells also provide coverage to the end customers using a cellular network standard, e.g., Universal Mobile Telecommunication System (UMTS), Wireless Interoperability for Microwave Access (WiMAX), and Long-Term Evolution (LTE).

Femtocells aim to enhance the poor indoor coverage, boost the overall spectral efficiency, and offload the overlay macrocell traffic [4]. Although femtocells provide significant benefits for mobile operators, their introduction comes with many challenges. Among these is the *cross-tier* interference between macro- and femtocells, as well as *co-tier* interference arising among femtocells. This calls for effective interference management strategies, such as distributed power allocation, resource partitioning, and other interference avoidance techniques. Undoubtedly, interference avoidance has never been a trivial task neither in macrocell deployments nor in femtocell networks. Due to the selfish nature of femtocells and uncertainty on
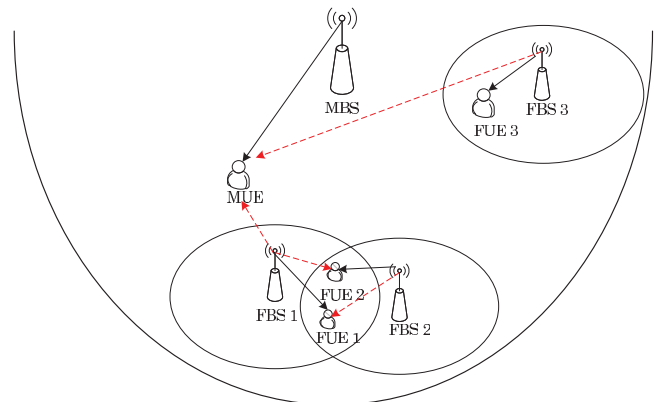

Fig. 1. Network topology with one macrocell underlaid with three femtocell networks. MUE and FUE stand for macro/femtocell user equipment, respectively. MBS and FBS stand for macro and femtocell base station, respectively.

their number and locations, operators must use optimal and dynamic approaches rather than the classical static network planning and optimization to avoid interference. In order to successfully react to the changes of the traffic, and minimize interference in femtocell deployments, the use of sophisticated self-organization techniques is paramount. Self-organization will allow femtocells to integrate themselves into the network of the operator, learn about their environment (such as neighboring femtocells and local interference map) and tune their parameters (transmit power and carrier frequency) accordingly.

A. Related work

different self-organization strategies for femtocells have been introduced within the framework of femtocell networks. In [2], a power control method was proposed for pilot and data channels in UMTS networks that ensures a constant coverage femtocell radius. Each femtocell sets the transmit power such that on average it is equal to the power received from the closest macrocell at a target femtocell radius. In [3], a method was presented for coverage adaptation for UMTS networks using information on mobility events of outdoor passing and indoor users. Each femtocell sets its power to a value that on average minimizes the total number of attempts of outdoor passing users to connect to such femtocell. In [5], a distributed utility-based signal to interference plus noise

ratio (SINR) adaptation at femtocells was proposed in order to alleviate cross-tier interference at the macrocell from cochannel femtocells. In [6], interference avoidance using a time-hopped Code Division Multiple Access (CDMA) physical layer and sectorial antennas is investigated. These approaches are mostly based on wide-band code division multiple Access (WCDMA) networks, and do not mitigate interference through sub-channel allocation, which is a very important feature of current Orthogonal Frequency Division Multiple Access (OFDMA) systems. Recent works investigating interference mitigation for OFDM femtocell networks can be found in [17], [19] for frequency resource partitioning, and [14] [18] using tools from game theory.

Existing research on Reinforcement-Learning (RL) [13] have been carried out in cognitive radio networks (e.g., see [22], [23], [16], [21]). In [22], the authors focused on the resource competition in a spectrum auction system, where the channel allocation is determined by the spectrum regulator, which is different from this paper in which no regulator exists. In [23], a distributed opportunistic spectrum access for cognitive radio using correlated equilibrium and no-regret learning was studied in which mutual communication among secondary users is assumed. A Q-learning based algorithm was investigated in [16] and [21] in the context of network selection for heterogeneous wireless networks, and channel selection in multi-user cognitive radios, respectively.

B. Contributions

The contributions of this paper can be summarized as follows:

- The strategic coexistence between the macrocell and femtocell networks is modeled and analyzed using tools from machine learning and evolutionary game theory, under different information assumption knowledge.
- *Distributed* algorithms relying on either *local* information or information exchange are proposed for different femtocell learning strategies, so as to mitigate interference towards the macrocell network.
- A comparison is provided between different strategic learning approaches, namely Q-learning, evolutionary-based approach.

The rest of the paper is organized as follows: Section II outlines the system model. Section III describes the investigated learning based mechanism, namely the Q-learning and evolutionary learning approach. Numerical results are given in Section IV and conclusions are drawn in Section V.

II. SYSTEM MODEL

We consider a wireless network consisting of one macrocell base station (MBS), and N_f femtocell base stations (FBSs) transmitting over N_{sub} subcarriers. In each time-slot, orthogonal downlink signaling is assumed, i.e., 1 user/slot/cell. Figure 1 depicts an illustration of the considered network topology where 1 MBS is underlaid with $N_f = 3$ FBSs. Let $p_0^{(n)}$ denote the macrocell's transmit power on subcarrier n towards its user. Likewise, the transmit power of FBS i in subcarrier n is

denoted by $p_i^{(n)}$. Let $|h_{i,j}^{(n)}|^2$ denote the channel gain between base station i and user j on subcarrier n, where $|h_{0,0}^{(n)}|^2$ is the channel gain between the MBS and its associated Macro User Equipment (MUE) in subcarrier n. Moreover, let σ_n^2 be the variance of Additive White Gaussian Noise (AWGN) at MUE, which is assumed to be constant over all subcarriers. At each time interval each FBS serves one FUE over one or a subset of the available subcarriers following a time division multiple access (TDMA) policy.

The signal to interference plus noise ratio of MBS at MUE (assuming Gaussian signalling) is given as:

$$\gamma_0^{(n)} = \frac{|h_{0,0}^{(n)}|^2 p_0^{(n)}}{\sigma^2 + \underbrace{\sum_i |h_{i,0}^{(n)}|^2 p_i^{(n)}}_{\text{femtocells}}}. \tag{1}$$

In this paper, the goal is to guarantee that the macrocell user meets its Quality of Service (QoS) requirement, that is $R_0 \geq \Gamma_0$, where Γ_0 is the MUE's target data rate. Similarly, the signal to interference plus noise ratio of FBS $i \in \{1, ..., N_f\}$ serving its FUE is given as:

$$\gamma_i^{(n)} = \frac{|h_{i,i}^{(n)}|^2 p_i^{(n)}}{\sigma^2 + \underbrace{|h_{0,i}^{(n)}|^2 p_0^{(n)}}_{\text{macrocell}} + \underbrace{\sum_{j \neq i} |h_{j,i}^{(n)}|^2 p_j^{(n)}}_{\text{femtocells}}}. \tag{2}$$

III. DISTRIBUTED LEARNING APPROACHES FOR MACRO-FEMTOCELL COEXISTENCE

In this section, two strategic learning mechanisms for interference mitigation towards the macrocell network are investigated, namely the evolutionary and Q-learning based approaches.

A. Evolutionary-based approach

The first learning mechanism for interference mitigation is based on the concept of evolutionary game theory, where each FBS chooses its strategy against other FBSs within the same network. FBSs observe the behavior of other competitors, learn from the observations, and make the best decision based on their instantaneous payoff, as well as the *average* payoff of all other femtocells. The game theoretic model is formulated as follows:

- **Players** FBS i, $\forall i \in \{1, ..., N_f\}$.
- **Action** $\mathcal{A}_i = \{a_i^{(n)}\}_{n \in \{1, ..., N_{sub}\}}$ where $a_i^{(n)}$ is the transmit power level used by FBS i in subcarrier n.
- **Payoff** $R_i = \mathbb{1}_{(\Gamma_0 - R_0)} \sum_n \log_2 \left(1 + \gamma_i^{(n)}\right)$ is the reward of FBS i transmitting in subcarrier n. Moreover, the average payoff of the entire population is defined as $\bar{R} = \frac{\sum_i R_i}{N_f}$.

In the aforementioned mechanism, an entity, which is referred to as home-eNodeB Gateway [1], collects the payoffs for all femtocells and calculates the average rate of the entire femtocell network. At time t, the payoff $R_i(t)$ of FBS i is

then compared with the average payoffs $\bar{R}(t)$ and in the case when it is less that the average rate of the femtocell network, a random strategy is chosen and the whole process is repeated again. Algorithm 1 describes the proposed evolutionary-based interference mitigation mechanism.

Let $\zeta_a^{(n)}(t) = \sum_{i=1}^{N_f} \mathbb{1}_{\left\{ p_i^{(n)}(t) = a_i^{(n)} \right\}}$ represent the total number of femtocells using strategy $a \in \mathcal{A}$, and $x_a(t) = \frac{\zeta_a^{(n)}(t)}{\sum_{a=1}^{|A|} \zeta_a^{(n)}(t)}$ the proportion of femtocells using strategy a, respectively. Thus, the replication dynamic equation can be defined as follows:

$$\dot{x}_a^{(n)}(t) = x_a^{(n)}(t) \Big(R_a^{(n)}(t) - \bar{R}^{(n)}(t) \Big), \qquad (3)$$

where $R_a^{(n)}(t)$ is the payoff of all femtocells transmitting at time t with strategy $a \in \mathcal{A}$ in subcarrier n, and $\bar{R}^{(n)}(t) = \sum_a x_a^{(n)} R_a^{(n)}(t)$ is the corresponding average payoff of the entire femtocell population over all strategies $a \in \mathcal{A}_i$. Based on the replicator dynamics in (3), the evolutionary equilibrium is defined as the set of fixed points of the replicator dynamics that are stable. This evolutionary equilibrium is a desirable solution to the evolutionary game since when the population of players evolves over time, it will converge to the evolutionary equilibrium. Furthermore, at this evolutionary equilibrium, none of the individuals wants to change its strategy since its payoff is equal to the average payoff of the population.

B. Q-learning based approach

Unlike the case where femtocells exchange their payoff information, Q-learning approach is based on the fact that information exchange among femtocells is no longer possible. The Q-learning formulation consists of a set of states and actions aiming at finding a policy that maximizes the observed rewards over the interaction time of the agents/players (i.e., femtocells). Every femtocell explores its environment, observes its current state, and takes a subsequent action, according to a decision policy. With their ability to learn, the knowledge about other players' strategies is not needed. Instead, a Q-table maintains the knowledge about other players in the network, based on which decisions are made. A number of works have shown that Q-learning converges to optimal values in Markov decision process environment [13]. Thus, the goal of an agent is to find an optimal policy $\pi^*(s)$ for each state s, which maximizes a cumulative measure of the rewards over time.

The state, agent, and reward associated with the Q-learning game are defined as follows:

- **Agent** i: FBS i, $\forall i \in \{1, ..., N_f\}$.
- **State** $S_i = \{s_i^{(n)}\}_{n \in \{1,...,N_{sub}\}}$ where $s_i^{(n)} \in \{0, 1\}$. The state of every FBS i in subcarrier n indicates whether FBS i generates interference towards the macrocell user above a given threshold, (i.e., $R_0 < \Gamma_0$).
- **Action** $A_i = \{a_i^{(n)}\}_{n \in \{1,...,N_{sub}\}}$ where $a_i^{(n)}$ is the transmit power level of FBS i over a set of subcarriers $\mathcal{C} \subseteq \{1, ..., N_{sub}\}$.

Algorithm 1 Evolutionary Based Algorithm
1: **Input:** Number of subcarriers N_{sub}, Number of femtocells N_f;
2: Choose a random strategy and obtain the reward $R_i(0)$;
3: **while** 1 **do**
4:     Provide $R_i(t)$ to the HNB-Gateway;
5:     Get $\bar{R}(t)$ from the HNB-Gateway;
6:     **if** $(R_i(t) < \bar{R}(t))$ **then**
7:         **if** $(rand(.) < (\bar{R}(t) - R_i(t))/(\bar{R}(t)))$ **then**
8:             Choose a random strategy and obtain the reward $R_i(t)$;
9:         **end if**
10:     **end if**
11: **end while**

Algorithm 2 Q-learning Based Algorithm
1: **Input:** Number of subcarriers N_{sub}, Number of femtocells N_f;
2: **Init:** $Q(.,.) = 0$ initialize Q-value for FBS$_i$
3: **while** 1 **do**
4:     **if** $(rand(.) < \beta)$ **then**
5:         Select an action randomly; [exploration step]
6:     **else**
7:         choose action $A_i(t+1) = \arg\max Q_{t+1}(S_i, A_i)$; [exploitation step]
8:     **end if**
9:     Receive immediate reward for femtocell i at time $t+1$: $R_i(t+1)$;
10:     Observe new state $S_i(t+1)$;
11:     Update Q-table as given in (4);
12: **end while**

- **Reward** $R_i = \mathbb{1}_{(\Gamma_0 - R_0)} \sum_{n \in \mathcal{C}} \log_2 \left(1 + \gamma_i^{(n)} \right)$ is the reward of FBS i transmitting in a subset $\mathcal{C}$ of all subcarriers, where:
$$\mathbb{1}_{(\Gamma_0 - R_0)} = \begin{cases} 0 & \text{for} \quad \Gamma_0 < R_0 \\ 1 & \text{for} \quad \Gamma_0 \geqslant R_0 \end{cases}.$$

Algorithm 2 describes the Q-learning process, in which FBS i performs the exploration step with probability β. A new Q-value, i.e., $Q_{t+1}(S_i, A_i)$, which is the expected payoff for the future iterations, is obtained based on previous value, i.e., $Q_t(S_i, A_i)$, along with the new observed payoff R_i at time $t+1$. Furthermore, the Q-learning equation is updated as follows:

$$Q_{t+1}(S_i, A_i) = (1 - \alpha)Q_t(S_i, A_i) + \qquad (4)$$
$$\alpha \Big[R_i + \gamma \max_{B_i \neq A_i} Q_t(S_i, B_i) \Big],$$

where α is the player's willingness to learn from its environment and γ is the discount factor.

IV. NUMERICAL RESULTS

We provide insight into the performance comparison of both learning strategies through numerical results. We consider *one* macrocell, $N_f = \{20, 40, 60\}$ femtocells and $N_{sub} = 15$ subcarriers. The Q-learning rate and discount factor are $\alpha = 0.5$ and $\gamma = 0.5$, respectively. Finally, the macrocell target rate is set to $\Gamma_0 = 50$ [bps/Hz]. We invesigate both uniform power allocation in which case the transmit power of femtocell i over subcarrier n is set to $p_i^{(n)} = P = 5$, whereas in the optimal power allocation using the Water-filling technique $\sum_{n=1}^{N_{sun}} p_i^{(n)} \leq \bar{P}$, with $\bar{P} = 25$.

Figure 2 shows the dynamics of the Q-learning approach with uniform power allocation where 2 femtocells share
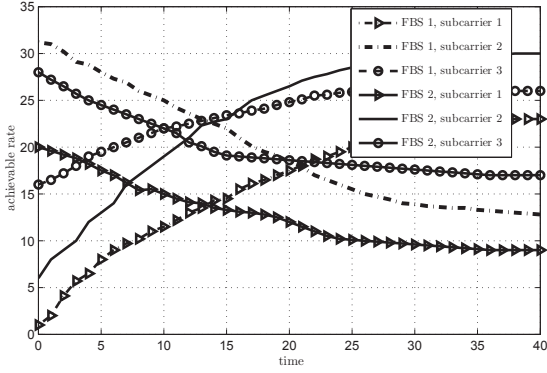
Fig. 2. Dynamics of the Q-learning algorithm in which 2 femtocells share 3 sub-carriers.

$N_{sub} = 3$ subcarriers, while at the same time coexisting with the macrocell network (i.e., $R_0 \geq \Gamma_0$). Although femtocells do not exchange information nor observe each other's actions, they implicitly coordinate their access when accessing the same spectrum. Figure 3 illustrates the dynamics of the considered system under different both Q-learning and evolutionary-based mechanisms for both uniform and optimal power allocations. It is shown that the proposed evolutionary approach outperforms the Q-learning approach in terms of convergence time and achievable rate. Furthermore, a comparison of both approaches with and without power control is given where the dynamic power control with water-filling technique yields higher capacity increase. More importantly, as time goes by, femtocells are able to coexist with the macrocell where the date rate of the macrocell user is satisfied.
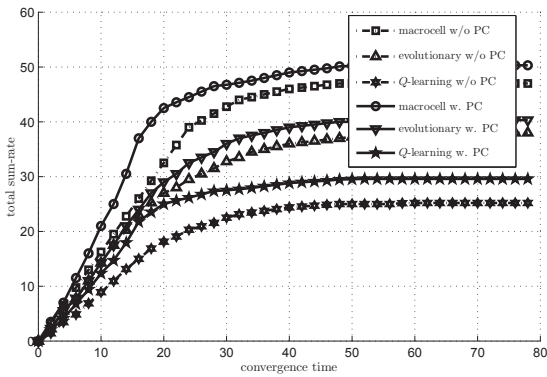

Fig. 3. Achievable rate R_0 of the macrocell user, the average capacity and the convergence of three different interference mitigation techniques: Evolutionary, Q-Learning, and random allocation.

Furthermore, Figure 4 depicts the total femtocell sum-rate with respect to the femtocell density N_f which satisfy the macrocell target rate constraint. It can be seen that there is an *optimal* point for the femtocell deployment density N_f^* whose effect is akin to the famous *Braess* paradox [10] in which increasing the number of deployed femtocells ends

up degenerating the global performance of the network. In addition, the evolutionary based approach yields a higher data rate.
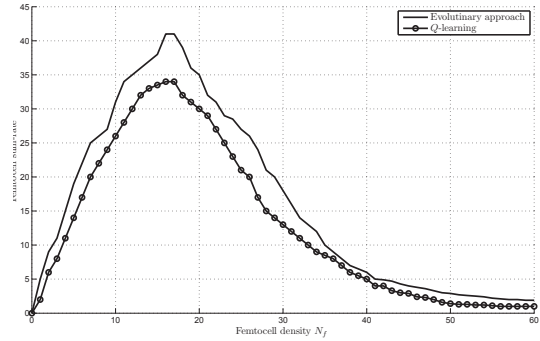

Fig. 4. The femtocell sum-rate with respect to the femtocell density N_f for both Q-learning and evolutionary based approach.

Figure 5 shows the increase in overall femtocell sum-rate with different values of the learning parameter α. Here, the learning rate determines to what extent the newly acquired information by each femtocell will update the outdated information. When α values are closer to zero, femtocells are unwilling to learn new actions, instead they focus on exploiting strategies which have been learnt so far, eventually yielding better payoffs. On the other hand, the closer the learning rate is to one, femtocells focus entirely on learning new strategies.
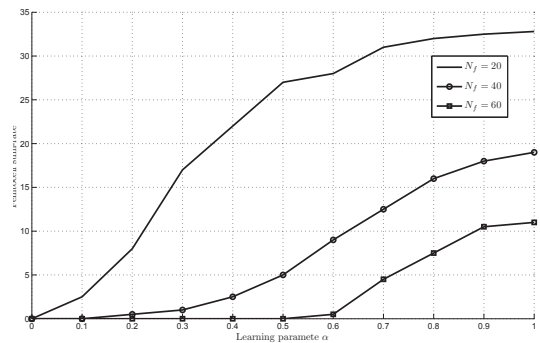

Fig. 5. The total femtocell sum-rate with respect to the learning rate α under the Q-learning approach.

Similarly, Figure 6 shows the behavior in terms of femtocell sum-rate with different values of exploration probability β. The total sum-rate increases with increasing β, nonetheless there exists an *optimal* point after which the payoff of the femtocell network starts decreasing until reaching zero when femtocells can no longer satisfy the macrocell target rate. Note that the optimal number of femtocells corresponding to Figure 4 correspond to the maximum available femtocell sum-rate compared to the femtocell density N_f. Clearly, an interesting exploitation versus exploitation tradeoff exists among femtocells in their coexistence with the macrocell network.
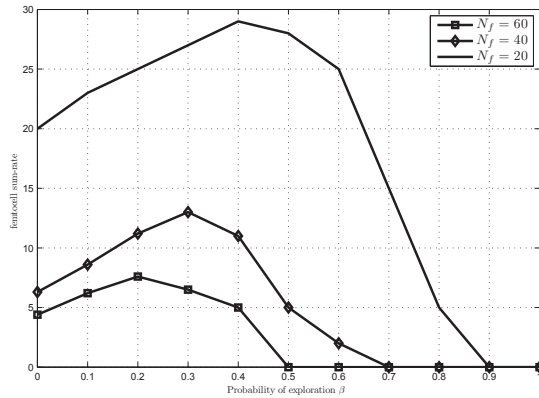
Fig. 6. The total femtocell sum-rate with respect to the Q-learning exploration probability β.

V. Conclusion

In this paper, we investigated and compared the convergence behavior of Q-learning with evolutionary game theoretic approach for the interference mitigation among macrocell and femtocells. The biologically-inspired evolutionary approach converges more rapidly to the desired equilibrium as compared to the reinforcement learning based (Q-learning) and random approach. This comes at the expense of more side information required at the femtocells. In both cases, it was shown that femtocells can coexist with the macrocell network. Future work will look into other aspects of strategic learning, as well as extending the current framework to other femtocell access policies (closed, open and hybrid access control).

VI. Acknowledgements

References

[1] 3GPP TR 25.820, "3G Home NodeB Study Item Technical Report (Release 8)," March 2008.

[2] H. Claussen, L. T. W. Ho, and L. G. Samuel, "An overview of the femtocell concept. Bell Labs Technical Journal," Wiley, 3(1): 221-245, May 2008.

[3] H. Claussen, L. T. W. Ho, and L. G. Samuel, "Self-optimization coverage for femtocell deployments," In Wireless Telecommunications Symposium, 24-26, pages 278 285, April 2008.

[4] V. Chandrasekhar and J. G. Andrews, "Femtocell networks: A survey," IEEE Communication Magazine, 46(9):5967, September 2008.

[5] V. Chandrasekhar, T. Muharemovic, Z. Shen, J. G. Andrews and A. Gatherer, "Power Control in Two-Tier Femtocell Networks," IEEE Transactions on Wireless Communications, October 2008.

[6] V. Chandrasekhar and J. Andrews, "Uplink capacity and interference avoidance for two-tier femtocell networks," IEEE Transactions on Wireless Communications, vol. 8, no. 7, July 2009.

[7] J. Nie and S. Haykin, "A Q-Learning-Based Dynamic Channel Assignment Technique for Mobile Communication Systems," IEEE Trans. On Vehicular Technology, Vol. 48, No. 5, Sept. 1999.

[8] J. Hofbauer and K. Sigmund, "Evolutionary Games and Population Dynamics,". Cambridge University Press, UK, 1998.

[9] B. Wang, K. J. R. Liu and T. C. Clancy, "Evolutionary game framework for behavior dynamics in cooperative spectrum sensing," in Proc. of IEEE Globecom, New Orleans, USA, 2008.

[10] D. Braess. "Uber ein paradoxon aus der verkehrsplanung," Unternehmensforschung, 24(5):258 268, May 1969.

[11] R. S. Sutton and A. G. Barto, "Reinforcement Learning: An Introduction," MIT Press, Cambridge, 1998.

[12] J. Hofbauer and K. Sigmund, "Evolutionary Games and Population Dynamics," Cambridge University Press, 1998.

[13] M. E. Harmon and S. S. Harmon, "Reinforcement learning: A tutorial," 2000.

[14] S. Guruacharya, D. Niyato, E. Hossain, and D. I. Kim, "Hierarchical competition in femtocell-based cellular networks," to be presented in IEEE GLOBECOM'10, Miami FL USA, 6-10 December 2010.

[15] D. Niyato and E. Hossain, "Competitive spectrum sharing in cognitive radio networks: A dynamic game approach," IEEE Transactions on Wireless Communications, vol. 7, no. 7, 2651-2660, July 2008.

[16] D. Niyato and E. Hossain, "Dynamic of network selection in heterogeneous wireless networks: an evolutionary game approach," IEEE Transactions on Vehicular Communications, vol. 58, no. 4, 2651-2660, May 2009.

[17] Z. Bharucha, H. Haas, G. Auer, and I. Cosovic, "Femtocell resource partitioning," IEEE Global Telecommunications Workshops (GLOBECOM'09), Honolulu, USA, November 2009.

[18] M. Bennis and M. Debbah, "On spectrum sharing with underlaid femtocell networks," the Second IEEE International Workshop on Indoor and Outdoor Femto Cells (IEEE PIMRC), Istanbul, 26 September, 2010.

[19] I. Guvenc, M. R. Jeong, F. Watanabe, and H. Inamura, "A Hybrid Frequency Assignment for Femtocells and Coverage Area Analysis for Co-Channel Operation," IEEE Commun. Lett., vol. 12, no. 12, pp. 880-882, Dec. 2008.

[20] D. Fudenberg and D. K. Levine, "The Theory of Learning in Games," The MIT Press, Cambridge, MA, 1998.

[21] H. Li, "Multi-agent Q-Learning of Channel Selection in Multi-user Cognitive Radio Systems: A Two by Two Case," IEEE Conference on System, Man and Cybernetics (SMC), 2009.

[22] F. Fu and M. van der Schaar, "Learning to compete for resources in wireless stochastic games," submitted to IEEE Trans. Veh. Technol., 2009.

[23] Z. Han, C. Pandana and K. Liu, "Distributive opportunistic spectrum access for cognitive radio using correlated equilibrium and no-regret learning," in Proc. of IEEE Wireless Communications and Networking Conference (WCNC), 2007.