Distributed Dynamic Spectrum Access with Adaptive Power Allocation: Energy Efficiency and Cross-Layer Awareness

Mahdi Ben Ghorbel*, Bechir Hamdaoui#, Rami Hamdi*, Mohsen Guizani*, and MohammadJavad NoroozOliaee#
*Qatar University, Doha, Qatar, #Oregon State University, Oregon, USA.
Email: mahdi.benghorbel@qu.edu.qa

***Abstract*—This paper proposes energy and cross-layer aware resource allocation techniques that allow dynamic spectrum access users, by means of learning algorithms, to locate and exploit unused spectrum opportunities effectively. Specifically, we design private objective functions for spectrum users with multiple channel access and adaptive power allocation capabilities. We also propose a simple, two-phase heuristic for allocating spectrum and power resources among users. The proposed heuristic splits the spectrum and power allocation problem into two sub-optimal problems, and solve each of them separately. The spectrum allocation problem is solved, during the first phase, using learning whereas, the power allocation is formulated as an optimization problem and solved, during the second phase, by traditional optimization solvers. Simulation results show that energy and cross-layer awareness and multiple channel access capability improve the performance of the system in terms of the per-user average rewards received from accessing the dynamic spectrum access system.**

***Index Terms*—Cross-layer resource allocation, dynamic spectrum access, distributed resource sharing, private objective functions, cognitive radio networks.**

I. Introduction

Dynamic Spectrum Allocation (DSA) [1] has been one of the hot topics in wireless communications in the last decade due to its potential for improving spectrum utilization efficiency, thus addressing the spectrum shortage problem. DSA has been an important catalysis for numerous research works, ranging from protocol design [2] to performance optimization [3] and spectrum sensing techniques [4]. One of the important factors in the design of efficient wireless systems is power consumption. Power and energy awareness has been generating continuous interest in the research community, as the importance of reducing energy consumption is becoming crucial not only in designing wireless systems, but also in any engineering systems due to other factors such as environmental concerns (global warming, pollution, etc.) and increased energy costs.

On the other hand, developing fully decentralized approaches is becoming more needed than ever due to the complexity of these emerging wireless systems. Though it can be very challenging to design them, decentralized approaches scale well, as they typically incur little to no communication and computational overhead while still performing relatively well. Recently, an efficient distributed technique for spectrum access and allocation based on learning was proposed [5]. The authors proposed a close-optimal, scalable, and highly learnable objective functions that can be used for enabling efficient DSA. Although the proposed technique is shown to perform well in terms of throughput, it does not account for power consumption.

In this work, we propose a joint dynamic multi-channel spectrum access and adaptive power allocation techniques that extend the technique proposed in [5] to account for power consumption and cross-layer couplings. Specifically, we develop learning-based, distributed energy and cross-layer aware resource allocation techniques that allow DSA users, by means of learning algorithms, to locate and exploit unused spectrum opportunities effectively. A key challenge of this work lies in how to propose an efficient algorithm that exploits channel diversity to enhance performance, but without suffering enormously from the added complexity of such an exploitation. To tackle this challenge, we propose a two-phase heuristic approach that combines learning and optimization in a way that reduces the computational complexity while still achieving good performances. The proposed heuristic splits the spectrum and power allocation problem into two sub-optimal problems, and solve each of them separately. The spectrum allocation problem is solved using learning algorithms during the first phase, whereas the power allocation is formulated as an optimization problem and solved by traditional optimization solvers during the second phase. Our simulation results show that the proposed energy and cross-layer aware techniques coupled with the multiple channel access capability improve DSA performances by increasing the per-user average rewards that users receive from accessing the DSA system.

The rest of this paper is organized as follows. We recall in Section II the main results of [5]. Section III introduces the proposed techniques. In Section IV, we present our formulation to the DSA resource allocation problem, discuss the complexity challenge of the proposed DSA technique, and present the proposed suboptimal approach to be used to overcome the complexity challenge. In Section V, we present simulation results and discuss the performance of the proposed approach under various different system parameters. Finally, we conclude the paper in Section VI.

II. Learning-Based Distributed DSA

In this section, we briefly overview and describe the objective function proposed in [5] for completeness. For this, we first begin by describing the problem setup, and then present and illustrate the technique.

A. DSA Problem: Assumptions and Objective

NoroozOliaee et al. [5] investigate distributed resource allocation techniques for large-scale DSA networks. They consider a large-scale spectrum allocation problem with n users all competing to access m spectrum bands ($n \gg m$), where each user selects and uses one spectrum band among the m available bands to carry out its data communication. The interfering users, those that end up selecting the same band, are assumed to share the spectrum band using a carrier sense multiple access (CSMA) scheme [7]. An elastic traffic model is considered where the intrinsic reward received by each user is proportional to the amount of received throughput provided that it exceeds a certain threshold, R_{th}. When the received throughput drops below this threshold, the reward drops exponentially. Explicitly, the reward received by user i can be written as

$$r_i(t) = \begin{cases} \frac{V_j}{|s_i(t)|+1} & \text{if } |s_i(t)|+1 \leq \frac{V_j}{R_{th}} \\ R_{th} e^{-\beta\left(\frac{(|s_i(t)|+1)R_{th}-V_j}{V_j}\right)} & \text{otherwise,} \end{cases} \tag{1}$$

where V_j is the capacity of band j and $s_i(t)$ is the set of users interfering with user i at time t.

The goal in [5] is then to design private objective functions that can be used, with any learning algorithm, to maximize the rewards that users receive from accessing the DSA system. The objective function is derived with four design goals in mind: it should allow users to achieve high rewards (optimality), it should be implementable in a distributed manner (distributivity), it should enable users to locate and find spectrum resources quickly (learnability), and it should be perform well for both small and large numbers of users (scalable).

B. The Difference Objective Function

To design an efficient objective function g_i for a user i that meets the above four design requirements, the authors start by analyzing the performance of two obvious and intuitive objective functions. The first consists of simply taking the intrinsic reward as the objective function ($g_i = r_i$ for each user i), whereas the second consists of using the global network reward as the objective function ($g_i = \sum_{k=1}^{n} r_k(t)$ for each user i). The intrinsic reward function choice results in an oscillating behavior of the performance with rapidly increasing and decreasing slopes. This behavior is due to the selfish nature of the intrinsic objective function, which does not take other users' actions into account when deciding on what actions a user i should take. On the other hand, the global objective function choice results in a much steadier behavior where the performance increases but at very low rate. This is due to the insensitivity of the function to the user's action. Thus, a good objective function should balance between the two conflicting requirements: $i)$ accounting for each other's actions when taking one's own actions to ensure objective alignment among users and $ii)$ being sensitive to one's own actions to increase users' learnability. The authors proposed to use an objective function that strikes a good balance between these two requirements. The basic idea lies in that removing the effects of all other users from the global objective gives us an objective function with higher learnability than the global objective function but while still ensuring objective alignment among users. Essentially, this function, referred to as *difference objective function*, measures user i's contribution to the total system received rewards, making it more learnable without compromising its alignedness quality. Formally, the difference function can be written as

$$D_i(t) = \sum_{k=1}^{n} r_k\left(s_k(t)\right) - \sum_{k=1, k\neq i}^{n} r_k\left(s_k(t) - \{i\}\right) \tag{2}$$

For the sake of illustration, we simulate, evaluate and present in Fig. 1 the performance of the difference function D_i and compare it against those of the two other functions: intrinsic r_i and global $G(t) = \sum_{k=1}^{n} r_k(t)$. Observe that the D_i function outperforms the other two significantly in terms of both optimality (it achieves high rewards) and learnability (it reaches up to high rewards very quickly). This objective function allows users to achieve around 90% of the maximal possible reward in less than 100 time episodes. In addition, it can be computed in a distributed manner when the network is fully connected. In the case of fully connected networks, D_i of user i (Eq. (2)) can be simplified to a function of only the number of interferers to that user i. That is, D_i can be expressed as

$$D_i(t) = \left(|s_i(t)|+1\right) r_i\left(s_i(t)\right) - |s_i(t)|\, r_i\left(s_i(t) - \{i\}\right) \tag{3}$$

III. The Proposed Reward Function: Power and Cross-Layer Awareness

Unlike [5] where each user is only allowed to select and communicate over one channel band, we assume in this work that each user is allowed to select (for example by means of a multicarrier scheme) and use more than one channel band to communicate. In addition, this work also employs power control to reduce energy consumption, a factor that has not been taken into consideration in previous works. In this work, we also use channel gains to compute the received throughput, allowing us to evaluate and analyze the performance under various channel conditions.

We consider a network topology that consists of n users (each user, also sometimes called agent, here refers to a transmitter-receiver pair) sharing m non-overlapping spectrum bands. We consider an Orthogonal Frequency-Division Multiple Access (OFDMA) system; we then divide each spectrum band into l equally distant sub-bands, where l is selected

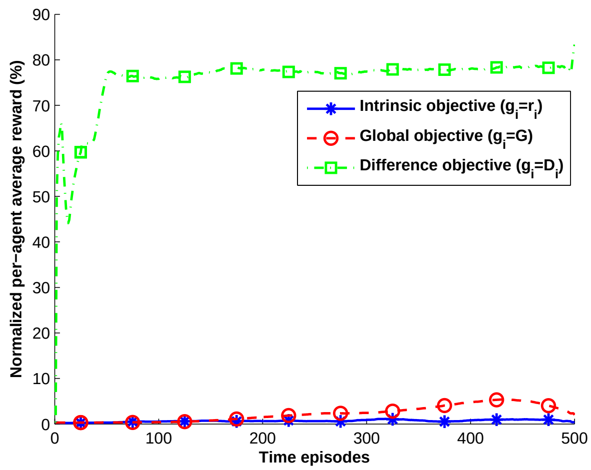

Fig. 1. Performance of the single band allocation with different private objective functions with the following system parameters: $n = 500$, $m = 10$,$V = 20$, $R = 2$, and $\beta = 2$.

to guarantee orthogonality between sub-bands. We denote by $g_i^{(j)}(t)$ the power of the instantaneous channel gain between user i's transmitter and receiver in the j^{th} channel sub-band.

One contribution of this work is to express the reward function explicitly in terms of the channel gains and allocated power. Let B_j be the bandwidth of each channel sub-band. B_js are selected such that the channel gains are constant over each channel sub-band. The throughput of user i at instant t is expressed in terms of the allocated power per band $P_i^{(j)}(t)$ as

$$R_i(t) = \sum_{j=1}^{ml} B_j a_i^{(j)} \, log_2 \left(1 + \frac{g_i^{(j)}(t)\, P_i^{(j)}(t)}{N_0 B_j} \right) \quad (4)$$

where $a_i^{(j)}$ is the user-band occupation mapping index (i.e., $a_i^{(j)} = 1$ if user i uses band j and $a_i^{(j)} = 0$ otherwise), and N_0 is the noise power level.

Thus, similarly to Eq. (1), the reward of user i at instant t, $r_i(t)$, can be expressed as

$$r_i(t) = \begin{cases} R_i(t) & \text{if } R_i(t) \geq R_{th} \\ R_{th} e^{-\beta\left(\frac{R_{th}}{R_i(t)} - 1\right)} & \text{otherwise} \end{cases}$$

As done in [5], three types of reward functions are studied in this work to show the effectiveness of our proposed technique:

- **Intrinsic reward (selfish behavior).** The reward for each user is equal to its own reward

$$g_i(t) = r_i(t)$$

- **Global reward (cooperative behavior).** The reward for each user is equal to the sum of all users' rewards.

$$g_i(t) = G(t) \triangleq \sum_{k=1}^{n} r_k(t)$$

- **Difference function reward.** In this reward type, each user aims to maximize the difference function described in Section II-B. That is,

$$g_i(t) = D_i(t)$$

where $D_i(t)$ is given in Eq. (3).

It is worth iterating that the difference between the three reward functions proposed in [5] and those proposed and studied in this work is three-fold: One, our proposed objective functions are cross-layer aware in that the reward a user receives depends on power level and channel characteristics (this is provided via Eq. (4)). Two, our proposed techniques are energy-aware in that the channel selection method (to be described later) accounts for power consumption of users via adaptive power control. Three, each user is allowed to access and use more than one spectrum band at the same time; that is, users have multi-channel access capabilities. Therefore, throughout the remaining of the paper, we will refer to our proposed techniques as *energy and cross-layer aware objective functions* to distinguish them from those proposed in [5].

IV. Joint Spectrum and Power Allocation

The joint spectrum and power resource allocation problem can be formulated as

$$\begin{array}{llr} \max\limits_{\{a_i^{(j)}, P_i^{(j)}\}} & \sum\limits_{i=1}^{n} r_i(t) & \\ \text{S.t} & 0 \leq P_i^{(j)} \leq P_i^{(j),max} & i \in \{1...n\},\ j \in \{1...ml\} \\ & \sum\limits_{j=1}^{ml} a_i^{(j)} P_i^{(j)} \leq P_i^{max} & i \in \{1...n\} \\ & 0 \leq \sum\limits_{i=1}^{n} a_i^{(j)} \leq 1 & j \in \{1...ml\} \\ & 1 \leq \sum\limits_{j=1}^{ml} a_i^{(j)} \leq m_i^{max} & i \in \{1...n\} \end{array} \quad (5)$$

where P_i^{max} is user i's maximum transmit power, $P_i^{(j),max}$ is the maximum allowed power per sub-band j for user i, and m_i^{max} is the maximum number of used sub-bands for user i.

In this problem, the objective is to maximize the total reward obtained by all users while respecting some constraints. The two first constraints concern energy consumption: the first is a budget power per user and the second is a maximum power per band which can serve to limit the generated interference. The third constraint guarantees non-interference between users as only one user is allowed to access each sub-band while the last constraint can be used to reduce the multi-carrier complexity by limiting the number of used sub-bands per user.

A. Complexity Challenge of the Learning Algorithm

Our proposed approach consists of extending the technique proposed in [5] to account for power consumption when allocating spectrum resources among users. We therefore evaluate our proposed technique while using the same learning algorithm that has been used in [5] (i.e., Q-learner [6]).

Accounting for power resources when allocating spectrum resources, though increases the spectrum efficiency as will be seen later, comes at a cost. The additional degree of freedom with the possibility of allocating different power levels on multiple channel sub-bands makes the resource allocation optimization problem a mixed integer programming (MIP) problem. Instead of searching in one allocated band per user (m possibilities per user), the new unknown variable is a vector of $l \times m$ scalars (power allocated per sub-band) which can take $\sum_{j=0}^{m_i^{max}} \binom{ml}{j} L_i^j$ possibilities, where L_i is the number of non-zero possible power levels for user i. This quantity is upper-bounded by $\sum_{j=0}^{ml} \binom{ml}{j} L_i^j = (L_i+1)^{ml}$ since $m_i^{max} \leq ml$. Given that the complexity of the learning algorithm is proportional to the search space of the unknown variables, this complexity is exponential on the number of channels, making it impractical for large systems. In addition, the discreteness of the power levels affects considerably the performance of the power allocation. For instance, increasing the number of levels L_i gives more freedom in allocating the power among used channel bands, which improves the performance, but on the other hand, increases the computational complexity also.

Therefore, in what follows, we propose a less complex, but sub-optimal approach for solving this problem.

B. Disjoint Channel and Power Allocation

To overcome the complexity issue of the learning algorithm, we propose, instead, a two-phase algorithm that consists of first using learning to determine channel-allocation mapping, and then solving the optimization problem formulated above to determine the best power allocation.

1) Learning-Based Channel Allocation: In this phase, we apply the Q-learner to find the best set of channel bands. Specifically, the best set of channel bands corresponds to the available sub-bands with maximal values in the Q-table. The available sub-bands are to be determined using a sensing phase that allows users to determine distributively whether a channel sub-band is used by other users or a primary user in case of cognitive systems. Taking into consideration the constraint on the maximal number of allowed bands per user, the total number of possible sets is $\sum_{j=0}^{m_i^{max}} \binom{ml}{j}$. This quantity is bounded from above by 2^{ml} (since $m_i^{max} \leq ml$ and $\sum_{j=0}^{ml} \binom{ml}{j} = 2^{ml}$). The considerable decrease in the search space during this first learning phase will result in a notable reduction of the computational complexity of the Learning algorithm .

2) Power Allocation Optimization: Having the channel sub-bands to be allocated to each user after the first phase, the problem of determining the allocated power per sub-band for each user can be formulated as a constrained convex optimization problem. In this sub-problem, the sub-bands allocation indexes $\{a_i^{(j)}\}$ are known and the objective for each user is to determine the power to be allocated for each used sub-band $P_i^{(j)}$. The optimization problem for each user i is formulated as

$$\begin{aligned} \max_{\{P_i^{(j)}\}_{a_i^{(j)}=1}} \quad & r_i(t) \\ \text{S.t} \quad & 0 \leq P_i^{(j)} \leq P_i^{(j),max} \quad j \in \{1...ml/a_i^{(j)}=1\} \\ & \sum_{j=1}^{ml} a_i^{(j)} P_i^{(j)} \leq P_i^{max} \end{aligned} \tag{6}$$

In ordinary resource allocation problems where the reward $r_i(t)$ is exactly the throughput $R_i(t)$, the solution to this problem can be found explicitly using a water-filling algorithm [9]. While in this problem, due to the elastic reward function, the optimization can not be solved analytically but it can be solved numerically using an ordinary optimization tool.

V. Simulation Results

We consider an uplink cellular network where $n = 100$ users are generated randomly inside a circular cell of radius $d = 1$ Km. The base station (receiver for all users) is located in the center of the cell. The channel gains are generated according to a Rayleigh distribution [10] of mean power equal to the distance-based pathloss ($\frac{1}{d^\eta}$) with a pathloss exponent $\eta = 3$. $m = 10$ spectrum bands are considered and $l = 50$ sub-bands for each band (i.e., a total of 500 sub-bands). The budget power per user is taken as $P_i^{max} = 20$ dBm while no maximum power per band is forced.

In Fig. 2, we show the performance (in terms of the per-user/agent average received rewards) of our proposed energy and cross-layer aware objective functions with multiple channel access and adaptive power allocation capabilities, and compare it with that achieved under the difference reward function D_i as proposed in [5]. For a fair comparison in terms of the available amount of service, we consider that the amount of service offered by a band depends on which user is using the band. Let $V_i^{(j)}$ denote the amount of service offered by a band j to user i; $V_i^{(j)}$ can be computed using Eq. (4) with $P_i(j) = P_i^{max}$ for a fair comparison in terms of the power consumption. We observe that the multiple channel allocation/access capability enhances the performance when compared with the single channel access by increasing the obtained average reward. This increased reward is a direct result of the benefit from the exploitation of channel diversity using this multiple channel access. In the proposed model, users are free to allocate more than one channel to maximize their reward. In addition, the proposed model allows the control of the energy consumption through the added power constraints.

Fig. 3 shows the performance obtained under each of the three studied functions when enabled with multiple bands access and adaptive power allocation capabilities. This figure confirms the conclusions drawn in [5] on the efficiency of the difference objective function in achieving better performance than the intrinsic and global objective functions when enabled with our proposed cross-layer and energy aware features.

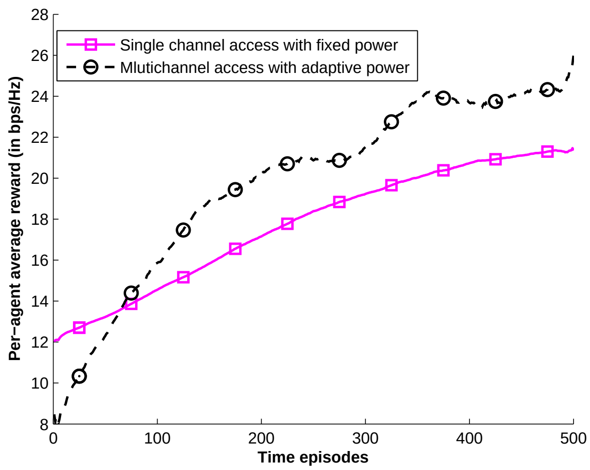

Fig. 2. Impact of multi-band spectrum access and adaptive power allocation capabilities on the achievable performance under the difference objective function.

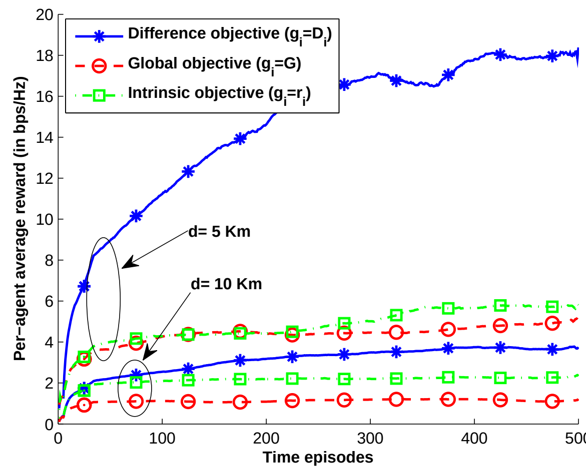

Fig. 4. Impact of network topology on the performance obtained under the three studied objective functions when enabled with multi-band spectrum access and adaptive power allocation capabilities.

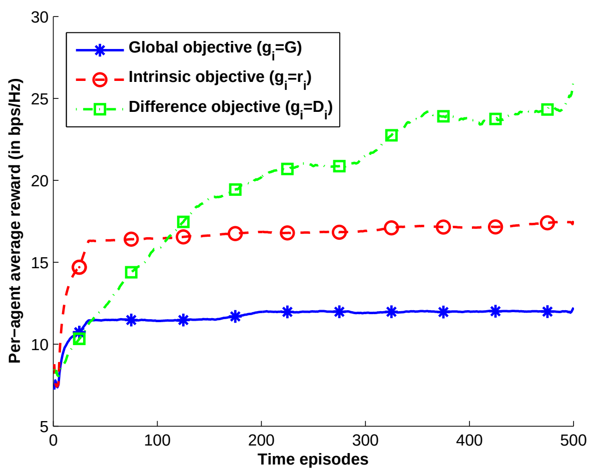

Fig. 3. Impact of multi-band spectrum access and adaptive power allocation capabilities on the achievable performance under the three studied objective functions: g_i, G, and D_i.

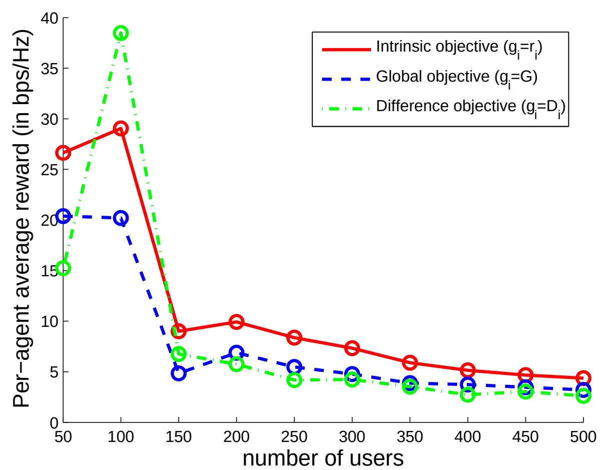

Fig. 5. Scalability study of the protocol: effect of increasing the number of users on the protocol performance.

In Fig. 4, we plot the performance of the proposed energy and cross-layer aware techniques with multiple channel access and adaptive power allocation capabilities for two different network topologies by varying the cell radius. First, we observe that the difference objective functions outperform the other two regardless of the network topology. Also, note that this performance amelioration increases as the cell radius decreases. This is simply because the reward is asymptotically inversely proportional to the distance.

In Fig. 5, we study the scalability of the proposed techniques by plotting the average achieved reward as a function of the number of users in the network. The figure shows that as the number of users increases, the per-user average received rewards go down rapidly. This is regardless of the used objective functions and regardless of the number of channels a user is allowed to use.

VI. Conclusion

This paper proposes learning-based, cross-layer and energy aware resource allocation techniques with multi-channel spectrum access and adaptive power allocation capabilities. It also proposes a heuristic for allocating spectrum and power resources among users. The proposed heuristic overcomes the complexity issues by splitting this resource allocation problem into two sub-optimal problems, spectrum allocation problem and power allocation problem, and solves each of them separately. The spectrum allocation problem is solved using learning methods whereas, the power allocation one is formulated as an optimization problem and solved by traditional numerical methods. Our simulation results show that proposed techniques perform well in terms of the per-user average achieved rewards because of their energy and cross-layer awareness and their multiple channel access capability.

VII. Acknowledgment

This work was made possible by NPRP grant # NPRP 5-319-2-121 from the Qatar National Research Fund (a member of Qatar Foundation). The statements made herein are solely the responsibility of the authors.

References

[1] Q. Zhao and Sadler B.M., "A Survey of Dynamic Spectrum Access", IEEE Signal Processing Magazine, vol. 24, no. 3, pp. 79 - 89, May 2007.

[2] B. Hamdaoui and K. G. Shin, "OS-MAC: An efficient MAC protocol for spectrum-agile wireless networks", IEEE Transactions on Mobile Computing, August 2008.

[3] Y. Ma, D. I. Kim, and Z. Wu, "Optimization of OFDMA-based cellular cognitive radio networks", IEEE Transactions on Communications, vol. 58, no. 8, 2010.

[4] Z. Quan, S. Cui, and A. H. Sayed, "Optimal linear cooperation for spectrum sensing in cognitive radio networks", IEEE Journal of Selected Topics in Signal Processing, Februray 2008.

[5] NoroozOliaee, MohammadJavad; Hamdaoui, Bechir; Tumer, Kagan, "Efficient Objective Functions for Coordinated Learning in Large-Scale Distributed OSA Systems", IEEE Transactions onMobile Computing, vol.12, no.5, pp. 931 - 944, May 2013.

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

[7] H. Kwon, H. Seo, S. Kim and B. Gi Lee, "Generalized CSMA/CA for OFDMA Systems: Protocol Design, Throughput Analysis, and Implementation Issues", IEEE Transactions on Wireless Communications, vol. 8, no. 8, August 2009.

[8] C. Watkins and P. Dayan, "Technical note: Q-learning", Machine Learning, vol. 8, no. 3-4, pp. 279-292, 1992.

[9] N. Papandreou and T. Antonakopoulos, "Bit and power allocation in constrained multicarrier systems: The single-user case", EURASIP J. Adv. Signal Process, vol. 2008, Jan. 2008.

[10] Bernard Sklar, "Rayleigh Fading Channels in Mobile Digital Communication Systems Part I: Characterization", IEEE Communications Magazine, vol. 35, no. 7, pp. 90-100, July 1997.