Dynamic Power Pricing using Distributed Resource Allocation for Large-Scale DSA Systems

Bassem Khalfi[1#], Mahdi Ben Ghorbel[2*], Bechir Hamdaoui[3#], Mohsen Guizani[4*]
[*]Qatar University, Doha, Qatar, [#]Oregon State University, Oregon, USA.
[1,3]{khalfib,hamdaoui}@eecs.orst.edu, [2]{mahdi.benghorbel}@qu.edu.qa,[4]mguizani@ieee.org

***Abstract*—In this paper, we propose dynamic power pricing for distributed resource allocation in large-scale Dynamic Spectrum Access (DSA) systems. The dynamic power pricing is considered to influence the users' spectrum assignment and power allocation in two resource allocation problems. In the first scenario, the objective is to maximize the reward of the obtained throughput over the time window while not exceeding a fixed budget for the power cost. The second problem consists of minimizing the total power cost while guaranteeing a minimum achieved throughput. Since the optimal solutions are of high computational complexity, we propose a distributed two-step algorithm to solve the optimization problems. In the first step, we rely on "learning" to determine the best channel selection for each user. In the second step, we optimize the allocated power to be used for the selected channels. Using simulations, we show that dynamic power pricing models allow achieving better DSA throughput when compared to the case of a static pricing for the same budget.**

***Index Terms*—Energy consumption awareness, spectrum access efficiency, dynamic spectrum access, dynamic power pricing.**

I. Introduction

Efficient power consumption has become one of the key design requirements of communication systems, motivated by the emergence of green communication [1] as well as smart grids. The latter has been seen as a promise for the future grid to create a distributed energy delivery network. It is anticipated to enhance the capacity and the efficiency of the grid by means of two-way communications between end-users and power plants, as well as by the inclusion and use of various types of renewable energy sources [2].

Smart grids are envisioned to provide, in the near future, a real-time power-usage pricing. It is used to prevent the energy concentration and flatten the peak loads. Hence, it could be a great opportunity for wireless communication systems, in general, to follow up the power-pricing and control their power consumption accordingly. In particular, this could provide a great opportunity for researches in Dynamic Spectrum Access (DSA) systems.

The concept of DSA has emerged as a key solution for the current spectrum scarcity caused by the static spectrum allocation policies. It allows an efficient use of the spectrum by allowing the coexistence of license-exempt users with legacy users. Due to its potential, DSA has created a significant research interest, ranging from spectrum awareness methods [3], [4] to spectrum sharing protocols [5]. With the emergence of smart grids, DSA became more pertinent due to the new opportunities given by dynamic power pricing. Therefore, it is important to revisit the resource allocation protocols proposed for DSA networks by taking into consideration the dynamic power pricing to further improve budget efficiency. Even though, recent trends in wireless communication systems are opting towards smaller cells with low power devices, power savings are still important to consider in the context of large-scale systems. Thus, of particular interest to us is to address scalability of the proposed resource allocation for DSA.

Although centralized approaches can be designed to achieve optimal performance, they are not suitable for large-scale DSA networks due to their lack of scalability. On the other hand, learning-based techniques have been viewed as potential candidates for decentralizing the allocation of spectrum resources, thereby enabling distributed DSA [6], [7], [8], [9]. In particular, the authors in [6] proposed an efficient private objective function that allows each user to maximize the reward that a user receives from accessing the DSA network. Although the proposed objective function is shown to be scalable, the analysis was conducted without taking into account the physical layer aspects. Authors in [7] revisited the objective function, proposed in [6], and enhanced it by generalizing the work to the case of multichannel access for every user. In addition, cross-layer awareness was taken into consideration by allowing each user to adapt the power allocated on each selected channel. To overcome the high complexity of the joint spectrum and power allocation approaches, the authors solved each problem separately, using learning for the former problem and water filling for the latter one.

With this in mind, we propose in this work a distributed multi-channel spectrum assignment and power allocation for large-scale DSA networks when using dynamic power pricing. We consider two resource allocation problems. In the first problem, the objective is to maximize the total throughput over a time window while meeting the budget limit for the power consumption cost. In the second problem, the objective is to minimize the total power cost of the consumed power while guaranteeing a minimum achieved throughput. To be in line with the actual standards, we assume that the power budget for each band is limited to a maximum transmit power. This limit can also serve to keep the interference under a limited threshold in case of underlay networks. To alleviate the joint spectrum and power allocation processing complexity at each user, we propose a two step approach. We first use reinforcement learning to perform the multi-band spectrum resource allocation. Then, we allocate the power budget to the selected bands optimally.

The rest of this paper is organized as follows. Section II describes the system model. In Section III, we first present the proposed dynamic power pricing model, and then formulate our resource allocation problems. In Section IV, we propose the disjoint spectrum and power allocation algorithm that is used to solve these problems. Simulation-based analysis is presented in Section V. Finally, the conclusion are presented in Section VI.

II. System Model

We consider a large-scale distributed DSA system where N users, called DSA agents, are competing to access l vacant bands. We assume that the users have accurately declared the bands as unused using spectrum sensing. Spectrum sensing process is out of the scope of this paper and it is assumed to be perfect. Each DSA agent represents a transmitter-receiver pair. Examples of transmitters could be femto-cell base stations, WiFi modems, wireless routers, etc.

We assume that each DSA agent i needs to communicate over a time window T. To do so, each agent i is allowed to select, at each instant t, up to $l_i^{\max}$ bands, and to use in each band j a maximum power $\hat{P}_i^{(j)}$. Let $a_i^{(j)}$ be the occupation mapping index for the band j. If user i has selected the band j, then $a_i^{(j)} = 1$, otherwise, $a_i^{(j)} = 0$. We denote by $g_i^{(j)}$ the j^{th} channel gain between the i^{th} user and its corresponding receiver and by $P_i^{(j)}$ the allocated power in each selected band j. Assuming that users who select the same band will share it orthogonally and equally to avoid interference (by using a carrier sense multiple access scheme), the achieved throughput by user i at instant t is expressed as

$$R_i(t) = \sum_{j=1}^{l} a_i^{(j)}(t) \frac{B_j}{n_j(t)} \log_2 \Big(1 + \frac{P_i^{(j)}(t) g_i^{(j)}(t)}{N_0 \frac{B_j}{n_j(t)}}\Big), \quad (1)$$

where N_0 is the noise's power spectral density, which is assumed to remain constant over time and to be equal for all bands, and $n_j(t) = \sum_{i=1}^{N} a_i^{(j)}(t)$ is the number of users sharing band j.

The power consumed by a user i at instant t, $P_i(t)$, is written as the sum of the allocated power over the selected bands $P_i^{(j)}(t)$ plus a constant component $P_i^{(0)}$ that models the non-radiated power, which is consumed independently regardless of the number of selected bands and the allocated power. It is essentially consumed by data processing, circuit RF chain, and cooling. Hence, $P_i(t)$ is written as

$$P_i(t) = \sum_{j=1}^{l} a_i^{(j)}(t) P_i^{(j)}(t) + P_i^{(0)}. \quad (2)$$

III. Problem Formulation

In this section, we start presenting the dynamic pricing model and describing the main difference compared to that of the static pricing scheme. Then, we formulate our resource allocation problem for dynamic pricing based DSA systems.

A. Power Pricing

In an ordinary scenario, the user's power unit cost is solely dependent on its own power consumption, and this is regardless of the global grid power demand. Hence, the total power cost scales linearly and could be written as

$$c_i(P_i) = \mu_i([P_1, P_2, ..., P_N]) \times P_i, \quad (3)$$

where P_i is the total power consumption of user i and $\mu_i([P_1, P_2, ..., P_N])$ is the power price per unit, which depends on the user's power consumption only. However, this model does not take into account the power grid load and the power provider's pricing policy which depend on many factors (peak hours, energy prices, etc.). With the use of dynamic pricing, the unit price not only scales with the user's consumption, P_i, but also depends on the other users' power consumptions and the whole system load. Hence, during the peak load time, the unit cost price is set sufficiently high to urge users to shift their consumption. However, when the whole power demand in the system is low, the unit price is set low to allow users to benefit from lower prices and avoid peak hours. However, the price in this situation does not scale well. Here, the total cost could be expressed as

$$c_i(P_i) = \mu_i([P_1, P_2, ..., P_N], D) \times P_i, \quad (4)$$

where D denotes the overall demand on the grid.

Note that in this context, each user is assumed to be equipped with a smart meter that captures instantaneously the power pricing parameters and will be used by the user to adapt its power consumption accordingly.

B. Resource Allocation Problems

As stated above, we consider two resource allocation problems, discussed separately in the following paragraphs.

1) Throughput Reward Maximization: In this problem, the main objective is to maximize the total throughput over all users while ensuring that each user reaches over the time window T a minimum target throughput R_i^{th} while not exceeding a maximum cost for the power at each time slot $\gamma_i(t)$ given the market price. This problem is suitable for elastic traffic, such as web browsing, file transfer, emails, etc.

Since we are targeting a distributed scheme for the resource allocation, and given that users' throughputs are mutually affected by the allocated spectrum by other users as shown in (1), we formulate our optimization problem as follows

$$\begin{aligned}
&\text{For each user } i \text{ in } \{1...N\} \\
&\max_{\{a_i^{(j)}, P_i^{(j)}\}} \quad r_i(R_i(t)) \\
&\text{S.t} \quad c_i(P_i(t)) \le \gamma_i(t), \\
&\qquad 0 \le P_i^{(j)} \le \hat{P}_i^{(j)}, \qquad \forall j \in \{1...l\} \\
&\qquad 1 \le \sum_{j=1}^{l} a_i^{(j)} \le l_i^{\max},
\end{aligned} \quad (5)$$

where $r_i(R_i(t))$ represents the reward of the achieved throughput $R_i(t)$ and can be written as

$$r_i(R_i(t)) = \begin{cases} R_i(t), & \text{if} R_i(t) \geq Q_i(t), \\ R_i(t)\exp(-\beta \frac{Q_i(t)-R_i(t)}{R_i(t)}), & \text{otherwise.} \end{cases} \tag{6}$$

Here r_i increases as the throughput increases, but it drops rapidly (exponentially) when the throughput is under a targeted threshold. $Q_i(t)$ is a targeted threshold at the current instant t, computed adaptively as a function of the obtained throughput in the previous time slots. Hence, it is written as

$$Q_i(t) = \frac{R_i^{\text{th}} - \sum_{t'=1}^{t} R_i(t')}{T-t}. \tag{7}$$

2) Power Cost Minimization: In this problem, the main objective is to minimize the total power cost while ensuring that the power cost for each user over the time window T does not exceed the user's budget Γ_i while achieving a minimum target throughput at each instant t, $R_i^{\text{th}}(t)$. This problem is suitable for inelastic applications requiring continuous minimum rate at each instant, such as voice and/or video streaming.

With this in mind, and using the same approach as in the previous problem, the distributed joint spectrum assignment and power allocation for this problem is given as follows

$$\begin{array}{ll} \multicolumn{2}{l}{\text{For each user } i \text{ in } \{1...N\}} \\ \min\limits_{\{a_i^{(j)}, P_i^{(j)}\}} & u_i(c_i(t)) \\ \text{S.t} & R_i(t) \geq R_i^{\text{th}}(t), \\ & 0 \leq P_i^{(j)} \leq \hat{P}_i^{(j)}, \qquad \forall j \in \{1...l\} \\ & 1 \leq \sum_{j=1}^{l} a_i^{(j)} \leq l_i^{\max}. \end{array} \tag{8}$$

Similarly to the first problem, $u_i(c_i(t))$ is a reward function associated with the power cost $c_i(t)$, and can be written as

$$u_i(c_i(t)) = \begin{cases} c_i(t), & \text{if } c_i(t) \leq q_i(t), \\ c_i(t)\exp(\beta \frac{c_i(t)-q_i(t)}{c_i(t)}), & \text{otherwise,} \end{cases} \tag{9}$$

where $q_i(t)$ is a targeted power cost at instant t and determined adaptively as a function of the cost of the total power in the previous time slots. $q_i(t)$ can be written as

$$q_i(t) = \frac{\Gamma_i - \sum_{t'=1}^{t} c_i(t')}{T-t}. \tag{10}$$

IV. Joint Spectrum-Assignment and Power-Allocation Algorithm

Using ordinary tools to solve the two previously presented problems could be computationally costly. In addition, they require a central entity to enable the coordination among users. Therefore, simple distributed approaches are needed to solve these problems. To do so, we rely on Q-learning which has already been shown to be a promising approach for solving spectrum assignment allocation problems in large-scale DSA systems [6]. However, solving the joint spectrum and power allocation problem is not possible using learning due to the very large learning set, which could deteriorate the learning performance and increase the computational complexity. For instance, in [7], we showed that the complexity is exponential on the number of channels as well as on the power levels. Therefore, approaches yielding suboptimal solutions, but with reasonable complexity, are more appealing. For this, we propose to di-associate the problem of the spectrum and power allocation. First, we solve the problem of the spectrum allocation using Q-learning. Then, once each user selected its bands, we solve the power optimization problem.

A. Spectrum Assignment

To alleviate the complexity issue of the joint spectrum and power resource allocation, we consider learning only to select the channels for each user. We use the ϵ-greedy Q-learner [11] to determine the best channels to select at each instant t based on their values in the Q-table at each user i. The Q-table values are updated recursively based on the observed rewards in the past time slots. Then, only the best l_i channels, such that $0 \leq l_i \leq l_i^{\max}$, among the available channels will be selected based on the associated value in the Q-table.

In this work, we also adopt the difference objective function, which is shown to ensure, in DSA systems, scalability, high learnability, and distributivity [6]. It is computed by removing the effects of other users in the global reward from the actual global reward; a detailed discussion of this function when applied to DSA can be found in [6].

In our case, we compute the difference function $D_i^{(j)}(t)$ for band j to be allocated to user i as follows

$$D_i^{(j)}(t) = \sum_{k=1}^{N} r_k^{(j)}(t) - \sum_{k=1}^{N} \hat{r}_{k,-i}^{(j)}(t), \tag{11}$$

where $\hat{r}_{k,-i}^{(j)}(t)$ stands for the received reward by user k by accessing band j when user i is supposed to be absent.

B. Power Allocation

Once the channels are selected for each user, i.e., $a_i^{(j)}$ are known, the complexity of the problem resulting from the mixed integer-real problem is removed; the power optimization problem turns out to be convex. The throughput maximization can be re-written as

$$\begin{array}{ll} \multicolumn{2}{l}{\forall i \in \{1...N\}} \\ \max\limits_{\{P_i^{(j)}\}} & r_i(R_i(t)) \\ \text{S.t} & c_i(P_i(t)) \leq b_i(t), \\ & 0 \leq P_i^{(j)} \leq \hat{P}_i^{(j)}, \quad \forall j \in \{1 \leq j \leq l \text{ and } a_i^{(j)} = 1\}. \end{array} \tag{12}$$

Whereas the power cost minimization can be written as

$$\begin{array}{ll} \multicolumn{2}{l}{\forall i \in \{1...N\}} \\ \min\limits_{\{P_i^{(j)}\}} & u_i(c_i(t)) \\ \text{S.t} & R_i(t) \geq R_i^{\text{th}}(t), \\ & 0 \leq P_i^{(j)} \leq \hat{P}_i^{(j)}, \quad \forall j \in \{1 \leq j \leq l \text{ and } a_i^{(j)} = 1\}. \end{array} \tag{13}$$

In simpler scenarios where the reward is exactly equal to the throughput or the consumed power, the problem can be solved analytically and the solution can be found via weighted water filling. Also, in the special case of single-band allocation (i.e., $l_i^{\max} = 1$), the problem can be directly solved by allocating all the possible powers deduced from the constraint in that band (the possible power is determined from the maximum budget for the power cost in the case of reward throughput maximization and the minimum required throughput in case of cost power minimization). In the general case, an ordinary optimization tool can be used to derive the optimal power allocations for each selected channel.

For the sake of illustration, we present in Algorithm 1 the different steps of solving the distributed resource allocation problem. We should emphasize that during the band allocation step for throughput reward maximization, the l_i bands with the highest values in the Q-table are selected while the l_i bands with the lowest Q-table values are selected in the problem of the power cost minimization.

Algorithm 1 Spectrum and power allocation for large scale DSA system.

INPUT: $b_i(t)$, $R_i^{\text{th}}(t)\ \forall i \in \{1...N\}$.
OUTPUT: $a_i^{(j)}(t)$ and $P_i^{(j)}(t)\ \forall i \in \{1...N\},\ j \in \{1...l\}$.

Initialize the Q-table:
$Q_i(1:l) = \mathbf{0}\ \ \forall i \in \{1...N\}$
for all episode t **do**

for all DSA agent i in the set of the agents **do**

1) **Bands' selection using ϵ greedy, Q-learner**
With a probability ϵ: select randomly l_i bands
With a probability $1-\epsilon$: select the the l_i bands available as follows:
Throughput maximization: select the highest values in the Q-table.
Power cost minimization: select the lowest values in the Q-table.

2) **Power allocation**
Throughput maximization: use equation (12).
Power cost minimization: use equation (13).

3) **Update the Q-table**
Compute the reward as follows:
Throughput maximization: use equation (6).
Power cost minimization: use equation (9).
Compute the difference function: $D_i^{(j)}(t)\ \forall j \in \{1...l\}$.
Update the Q-table:
$Q_i(j) = \alpha Q_i(j) + (1-\alpha) D_i^{(j)}(t)\ \ \forall j \in \{1...l\}$.

end for
end for

V. Simulation Results

We consider a set of $N = 500$ DSA users uniformly distributed in a cell with a radius $d_0 = 1$ Km. Each DSA user tries to communicate with its receiver over a slow Rayleigh fading channel. To capture the path loss effect on the different channels, we consider the average channel gain as $(d/d_j)^{\eta}$, where d_j represents the distance that separates the j^{th} transmitter and receiver whereas η is the pathloss exponent assumed to be equal to 3. We consider a total number of available bands that is equal to $m = 10$, where each band is assumed to have a bandwidth $B = 1$ MHz.

We assume that each DSA agent is equipped with a smart meter that could provide it with (instantaneous) unit pricing in real-time. Although we use a simple policy for the pricing of the consumed power where the unit price is a linear function of the consumed power, we consider two different system models for power pricing of the users. In the first one, each user is connected independently to the power grid and hence its power per unit price will depend only on its consumption as follows:

$$\mu_i([P_1, P_2, ..., P_N]) = \alpha(t) \times P_i. \tag{14}$$

In the second model, all DSA agents are connected together to the same power generator. Thus, the power per unit price will depend on their total consumption. In this case, the unit price is written as follows

$$\mu_i([P_1, P_2, ..., P_N]) = \alpha(t) \sum_{k=1}^{N} P_i. \tag{15}$$

In (14) and (15), $\alpha(t)$ captures the fluctuations of the price by the power provider that will depend on the total load and market energy prices. In Fig. 1, we show the used models for $\alpha(t)$. The case of $\alpha(t) = \alpha_1$ corresponds to the static power pricing where the unit cost varies only as a function of user i's power consumption. However, the cases of $\alpha(t) = \alpha_2$ and $\alpha(t) = \alpha_3$ correspond to the dynamic power pricing case. In the model $\alpha(t) = \alpha_2$, $\alpha(t)$ follows a uniform distribution where the mean is $E(\alpha(t)) = \alpha_1$ to ensure a fair comparison. In the model $\alpha(t) = \alpha_3$, there are mainly two regions: a region with high unit cost and a region with low unit cost. Likewise, in this model, the mean value $E(\alpha(t)) = \alpha_1$ guarantees a fair comparison.

In Fig. 2, we show the per-agent achieved throughput reward for the problem of throughput maximization for the first pricing policy and with $\alpha(t) = \alpha_1$ and $\alpha(t) = \alpha_2$. We conclude that the achieved throughput reward is higher in the case of dynamic power pricing than in the case of the static pricing. The per-agent fluctuation is explained by the fluctuation in the demand ($\alpha(t)$).

In Fig. 3, we consider the per-agent power cost in the problem of power cost minimization for the first pricing policy with $\alpha(t) = \alpha_1$, $\alpha(t) = \alpha_2$ and $\alpha(t) = \alpha_3$. First, we notice that for the three cases, the per-agent power price decreases over time (learnability effect). Second, with smart grid, i.e., $\alpha(t) = \alpha_2$ and $\alpha(t) = \alpha_3$, we achieve a lower power cost compared to the conventional power grid ($\alpha(t) = \alpha_1$).

In Fig. 4, we illustrate the per-agent power cost under the second proposed power consumption pricing policy given by (15). Like the previous result, here with dynamic pricing, we achieve lower power costs when compared with static power pricing.

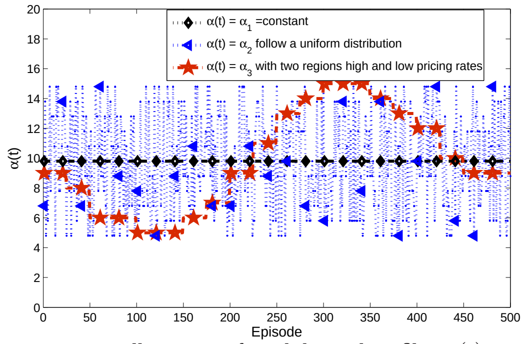

Fig. 1. Illustration of used demand profiles $\alpha(t)$.

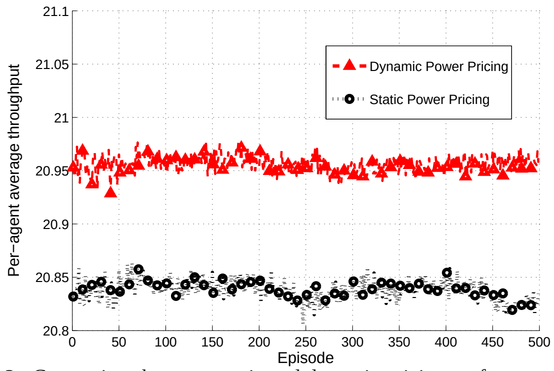

Fig. 2. Comparison between static and dynamic pricing performance in terms of reward per time slot for throughput reward maximization.

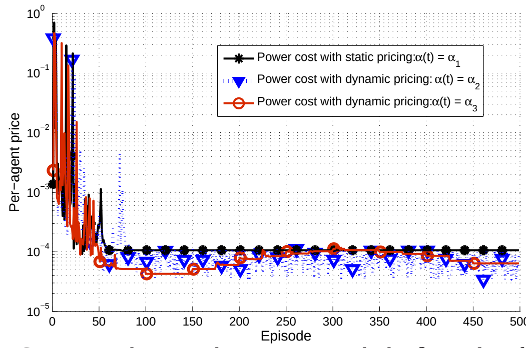

Fig. 3. Comparison between the price cost with the first policy for the two models of dynamic pricing with the model of static pricing.

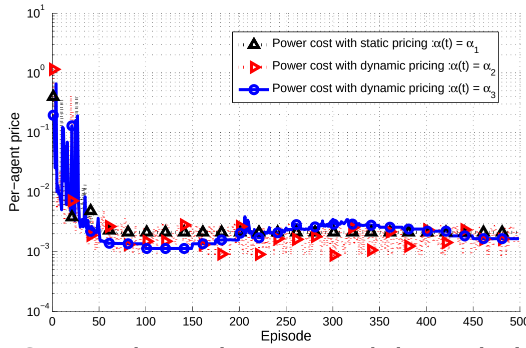

Fig. 4. Comparison between the price cost with the second policy for the two models of pricing using static and dynamic pricing.

VI. Conclusion

This paper proposed a joint power and spectrum allocation scheme based on learning for large-scale dynamic spectrum access with dynamic power pricing. Since the joint power-spectrum problem is of high computation complexity, we proposed a two-phase spectrum and power allocation approach to tackle the complexity issue. We rely on Q-learning to allocate the available spectrum among different users while relying on optimization to allocate power levels. We solve the problem for two different scenarios; one for throughput reward maximization and one for power cost minimization. We showed that dynamic pricing could be of great promise for the two scenarios since the users could adjust their resources according to the given price and profit to improve their performance and/or save the power budget. As of possible future research work in this area, we plan to improve the achieved results through studies with realistic data for power pricing and enhance the algorithm by predicting the power pricing.

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] Yan Chen; Shunqing Zhang; Shugong Xu; Li, G.Y., "Fundamental trade-offs on green wireless networks", IEEE Communications Magazine, vol.49, no.6, pp.30,37, June 2011.

[2] Fang, Xi; Misra, Satyajayant; Xue, Guoliang; Yang, Dejun, "Smart Grid - The New and Improved Power Grid: A Survey", IEEE Communications Surveys & Tutorials, vol.14, no.4, pp.944,980, Fourth Quarter 2012.

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

[4] Yucek, T.; Arslan, H., "A survey of spectrum sensing algorithms for cognitive radio applications", IEEE Communications Surveys & Tutorials, vol.11, no.1, pp.116,130, First Quarter 2009.

[5] Tragos, E.Z.; Zeadally, S.; Fragkiadakis, A.G.; Siris, V.A., "Spectrum Assignment in Cognitive Radio Networks: A Comprehensive Survey", IEEE Communications Surveys & Tutorials, vol.15, no.3, pp.1108,1135, Third Quarter 2013.

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

[7] Ben Ghorbel, M.; Hamdaoui, B.; Hamdi, R.; Guizani, M.; NoroozOliaee, M., "Distributed Dynamic Spectrum Access with adaptive Power allocation: Energy efficiecy and Cross layer awarness", IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pp.694,699, April 2014.

[8] Zhe Chen; Qiu, R.C., "Q-learning based bidding algorithm for spectrum auction in cognitive radio", in Proc. of IEEE SoutheastCon 2011, pp.409,412, March 2011.

[9] Yinglei Teng; Yong Zhang; Fang Niu; Chao Dai; Mei Song, "Reinforcement Learning Based Auction Algorithm for Dynamic Spectrum Access in Cognitive Radio Networks", in Proc. of IEEE Vehicular Technology Conference Fall (VTC-Fall'10), pp.1,5, September 2010.

[10] NoroozOliaee, MohammadJavad; Hamdaoui, Bechir; Guizani, Mohsen, "Adaptive service function for system reward maximization under elastic traffic model", in Proc. of The International Wireless Communications and Mobile Computing Conference (IWCMC'13), pp 1722,1726, July 2013.

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