

# Channel Sensing Order for Distributed Cognitive Radio Networks Using No-Regret Learning

Li-Wang Li, Jin-Cheng Ge

College of Communications Engineering, PLA University of Science and Technology, Nanjing, China  
E-mail: 18761681067@163.com, gejincheng@hotmail.com

**Abstract**-This work investigates the problem of channel sensing order in spatially distributed networks with multi-user. We model this problem into a non-cooperative game and define a generalized interference metric. In order to minimize the aggregate interference and maximize the normalized throughput, we employ the no-regret learning algorithm. However, this algorithm requires information exchange. Therefore, we propose a modified no-regret learning algorithm that does not require information exchange. Through simulations, we evaluate our proposal and compare its performance with other algorithms, and the results obtained are close to the no regret learning algorithm and superior to random selection algorithm.

**Keywords**-sensing order; spatially distributed networks; non-cooperative game ;no-regret learning

## I. INTRODUCTION

The demand for spectrum is speedy increasing and some licensed bands is used inefficiently [1] push the idea that those free underutilized bands is used for opportunistic spectrum access (OSA) [2]. OSA requires reconfigurable networks devices, called *cognitive radios* (CR), can adapt their behavior in response to environment stimuli [3]. For this, these cognitive devices or secondary users (SUs) need to determine by spectrum sensing when primary users are active in order to avoid causing them a harmful interference [4]. Therefore, OSA is important to utilize the licensed frequency spectrum more efficiently through opportunistic access to unused spectrum bands. At present, sensing-based OSA is widely studied because it does not require the licensed users to alter their hardware or behavior [5]. In sensing-based OSA, SU monitors the environment to detect the primary user (PU) signals and operate when the band is empty [5]. But in practice, it relies on a combination of sensing and the use of geo-location spectrum occupancy databases [6].

When multiple channels are available for opportunistic access, time-slotted multiple access is widely considered in these works [7]-[12]. The first portion of each time slot is used by SU to sense spectrum, and the second portion is used to access the free channel, if one is found [13]. However, when multiple secondary users have to search multiple available channels to access, these SUs will face competition from one another to access the channel. For example, if a channel is sensed free by two or more SUs and more than one of them decide to transmit on the channel, then collision occurs. Therefore, this paper investigates how SUs

autonomously select channel sensing orders to reduce the number of collisions.

In sensing-based OSA, the SUs are required to perform periodic spectrum sensing so that when a PU becomes active in a channel, the SUs can vacate that channel [14]. There are two broad categories of periodic sensing policies: a single-channel sensing policy and a sequential channel sensing policy. Under a single-channel sensing policy, a SU first select a channel to sense in any given time slot and transmits if the channel is free; otherwise, SU stay silent for the entire duration of that time slot. The works in [15]-[20] proposed distributed learning and allocation strategies that employed adaptive randomization based on feedback for the SUs to arrive at orthogonal channel selections. Under a sequential channel sensing policy, in any given time slot, a SU can sense the channels sequentially based on some sensing order. The works in [21]-[23] proposed optimal policies for the selection of a channel sensing order for a single SU. Unlike [21]-[23], the work in [13] took into account multiple SUs competition for channels. The selection of an optimal sensing order by a coordinator for a two-SU networks in [8]. Obviously, a coordinator was simple to implement for two-SU networks, but in practice a network comprised a large number of SUs so that it required significant signaling overhead to coordinate successful channel utilization. The works in [9] proposed a channel sensing order policy for distributed networks. However, this work assumed that SUs knew the gain of each channel, and each SU sensed channel in descending order of their achievable rates and transmitted in the first channel that was sensed free. Unlike [9], the work in [13] did not assume that SUs had knowledge of channel gains. Moreover, it proposes adaptive  $\gamma$ -persistent strategy to reduce the collision, but this work assumed that SUs were always active. The work in [24] proposed two non-cooperation game models for multiuser sequential channel sensing and access in dynamic cognitive radio networks, and proposed a stochastic learning algorithm.

In general, these works did not take into account the spatially distributed characteristic. Therefore, we consider it in this paper. Furthermore, in order to reduce the number of collisions in multi-user and multi-channel environment, we employ no-regret learning algorithm and propose a modified no-regret learning algorithm. Compared with no-regret learning algorithm, it does not require SUs to know the information of other SUs so that it can save lots of network resources. On the other hand, the performance obtained by modified no-regret learning algorithm is close to no-regret learning algorithm.

In reminder of this paper, In Section II, we present the system model and problem formulation. In Section III, we formulate channel sensing order selection game. In Section IV, we introduce the no-regret learning algorithm and propose a modified no-regret learning algorithm. In Section V, simulation results and discussion are presented. Finally, we present discussions and draw conclusions in Section VI.

## II. SYSTEM MODEL AND PROBLEM MODELING

### A. System Model

Considering a distributed CR network with one gateway,  $N$  SUs and  $M$  licensed channels. The set of SUs and channels is  $\mathcal{N} = \{1, 2, \dots, N\}$  and  $\mathcal{K} = \{1, 2, \dots, M\}$ . The licensed channels are owned by the PUs and can only be used by the SUs when they are not occupied by the PUs [13]. The PUs and SUs are both assumed to use a time slotted system, and each PU is present or absent for the entire time slot [7], [10], [17].

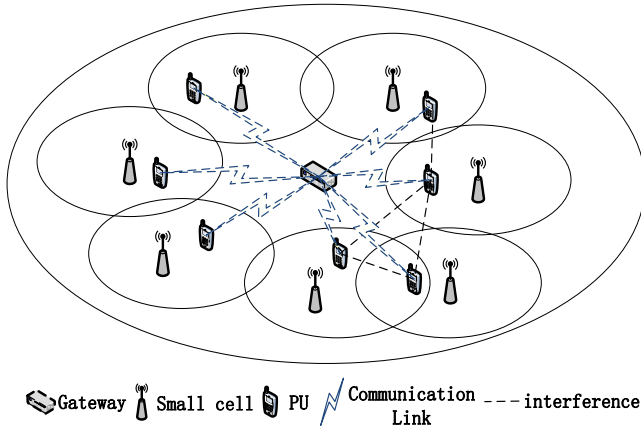


Figure 1. System model

The selection of channel for OSA is determined as follow: The SUs use the beginning of each time slot to sense the channels sequentially in some order to find a free channel. We refer to this as the sensing stage. Then the SUs access the first free channel, if one exists. We refer to this as the data transmission stage [13]. Let  $\mathcal{P}$  denote the set of available sensing orders. Note that the sensing order that a SU employs come from a Latin Square (LS), i.e., a  $M$  by  $M$  matrix of  $M$  channel indices in which every channel index occurs exactly once in each row and column of the matrix [25], [26].

Before sensing stage, all SUs first select sensing orders from LS. Let  $Q_n$  denote the channel sensing order of  $SU_n$ . If two or more than two SUs select a same channel sensing order, they will sense a same free channel and access the channel (false alarm probability are assumed to be zero), if one exists. Then collision occurs. However, in spatially distributed CR networks, this collision occurs only when neighbor SUs select a same channel sensing order and access a same free channel. In this paper, we define this collision as channel sensing order interference. Let  $d_0$  denote the

distance threshold of neighbor SU, and  $L_n$  denotes the set of  $SU_n$ 's neighbor SUs.

In sensing stage, each time slot is divided into a number of sensing steps. The SU senses a different channel in each sensing step. If SU finds a free channel in its  $i$ th sensing step and transmits data in that channel. However, if all channels are found busy, then SU stay silent in the reminding duration of that time slot. Let  $\tau$  denote the normalized time required to sense each channel, so the normalized time for data transmission is  $1 - i\tau$ . Here, we define the normalized time for data transmission as throughput of SU in each time slot, as follow:

$$\mathcal{T} = \begin{cases} 1 - i\tau, & \text{if } SU \text{ finds a free channel in } i\text{th step} \\ 0, & \text{if } SU \text{ finds all channels are busy} \end{cases} \quad (1)$$

When multiple SUs search multiple potentially available channels for OSA, the following three events maybe happen in each sensing step: 1) The SU senses a given channel and is the only one SU to find it free, then transmits data at the remaining time slot. 2) The SU senses a given channel and finds that channel is occupied by PU or SU, then it continues to sense other channels in next sensing steps. 3) The SU sense a given channel and finds it free, then accesses to that channel. At the same time, other SUs find that channel free and access to that channel too, a collision occurs and the SU is not able to transmit until the next time slot. Note that false alarm would have effect on channel state (free or busy). Let  $P_f$  denote the probability of false alarm. For example, Figure 2 (a)-(c) illustrates four different scenarios for sequential channel sensing using sensing orders.

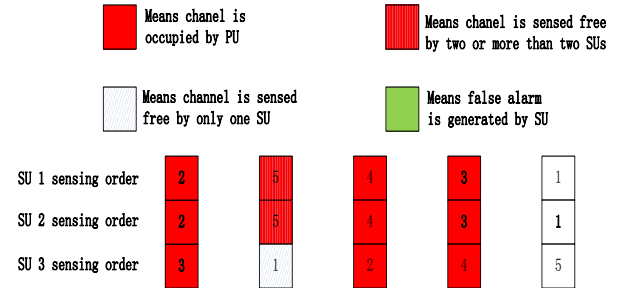


Figure 2. (a) Scenario 1

- Scenario 1: SU 1 and SU 2 select a same sensing order, so they both find channel 5 free in step 2 and collide. SU 3 finds channel 1 free in step 2.

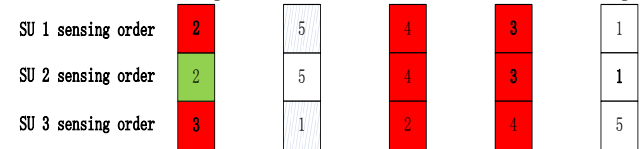


Figure 2. (b) Scenario 2

- Scenario 2: SU 1 and SU 2 select the same sensing order but they avoid collisions as SU 1 finds channel 2 busy in step 1 and SU 2 generates a false alarm in step 1 and finds channel 2 free (in fact, it is occupied by PU), then SU 2 collide with PU. SU1 finds channel 5 free in step 2, SU 3 finds channel 1 free in step 2.

SU 1 sensing order	2	5	4	3	1
SU 2 sensing order	1	3	5	2	4
SU 3 sensing order	3	1	2	4	5

Figure 2. (c) Scenario 3

- Scenario 3: SU1 finds channel 5 free in step 2. SU2 finds channel 1 free in step 1. SU3 finds channel 4 free in step 4.

### B. Problem Formulation

In a distributed CR network with multi-user. Let  $Q_n = (q_{n1}, q_{n2}, \dots, q_{nM})$  and  $Q_m = (q_{m1}, q_{m2}, \dots, q_{mM})$  denote the sensing order of SU  $n$  and SU  $m$ . Inspired by work in [24], we define the interference level of SU  $n$  is,

$$I_n(\mathcal{N}) = \sum_{m \in \mathcal{N}, m \neq n} \left[ \sum_{l=1}^M (q_{nl} \odot q_{ml}) P_{q_{nl}} (1 - P_f) \right] \bullet \delta_{nm} \quad (2)$$

where  $\odot$  is the bitwise XNOR operation,  $P_{q_{nl}}$  is the free probability of the channel  $q_{nl}$ ,  $P_f$  is the probability of false alarm, and  $\delta$  is the following indicator function between SU  $n$  and SU  $m$ ,

$$\delta_{nm} = \begin{cases} 1, & d_{nm} \leq d_o \\ 0, & d_{nm} > d_o \end{cases} \quad (3)$$

where  $d_{nm}$  denotes the distance from SU  $n$  to SU  $m$ ,  $d_o$  is also the distance threshold of interference.

From the SU side, a lower value of  $I_n(\mathcal{N})$  is desirable as it is expected to achieve higher throughput. From the system side, a lower value of the aggregate generalized interference level of all the SUs is also desirable. Therefore, we define the aggregate generalized interference level of SUs set  $\mathcal{N}$  as follow:

$$\mathcal{I}(\mathcal{N}) = \sum_{n \in \mathcal{N}} I_n(\mathcal{N}) \quad (4)$$

Mathematically, the optimal objective is to minimize the aggregate generalized interference level,

$$P1: \max \{-\mathcal{I}(\mathcal{N})\} \quad (5)$$

Generally, this problem  $P1$  is a challenge, because it is combinatorial optimization problem [24]. In the following, we will propose a game-theoretic learning framework to solve that problem.

### III. CHANNEL SENSING ORDER SELECTION GAME

The fact that the selections of channel sensing orders of SUs are interactive motivates us to formulate this problem as a non-cooperative game [24]. In this paper, we refer to it

as channel sensing order selection game and analyze the existence of Nash equilibrium.

#### A. Channel Sense Order Selection Game

The channel sensing order selection game is denoted as  $\mathcal{G} = [\mathcal{N}, \{\mathcal{A}_n, n \in \mathcal{N}\}, \{U_n, n \in \mathcal{N}\}]$ , where  $\mathcal{N}$  is the set of SUs,  $\mathcal{A}_n$  is the available action set of player  $n$  and  $U_n$  is the utility function of player  $n$ . Since the actions of players are interactive, the utility function is denoted as  $U_n(Q_n, Q_{-n})$ , as follow:

$$U_n(\mathcal{N}, Q_n, Q_{-n}) = -I_n(\mathcal{N}) = - \sum_{m \in \mathcal{N}, m \neq n} \left[ \sum_{l=1}^M (q_{nl} \odot q_{ml}) P_{q_{nl}} (1 - P_f) \right] \bullet \delta_{nm} \quad (6)$$

In non-cooperative game models, each player intends to maximize individual utility function [27]. So we can express it as:

$$\mathcal{G}_d: \max U_n(\mathcal{N}, Q_n, Q_{-n}), \forall n \in \mathcal{N} \quad (7)$$

Now we formulate problem  $P1$  as a non-cooperative game. Therefore, we need to analysis the existence of Nash equilibrium.

#### B. Analysis of the Nash Equilibrium

The properties of the channel sensing order selection game is  $\mathcal{G}_d$  characterized by the following theorem.

**Theorem 1** [24]: For any SU set  $\mathcal{N}$ , the channel sensing order selection game  $\mathcal{G}_d$  is an exact potential game that has at least one pure strategy NE point. Furthermore, any optimal solution of problem  $P1$  constitutes a pure strategy NE of the game.

### IV. CORRELATED EQUILIBRIUM AND DISTRIBUTED NO-REGRET LEARNING ALGORITHM

When multiple SUs have to search multiple potential available channels for OSA, they face competition from neighbor SUs. If two or more than two neighbor SUs select a same channel sensing order ignoring the influence of false alarm, collision will occur. In order to reduce that collision, we employ the no-regret learning algorithm. But the no-regret learning algorithm requires information exchange. So we propose a modified no-regret learning algorithm that does not require information exchange. Moreover, the two algorithms can both converge to Correlated Equilibrium

#### A. Correlated Equilibrium

The concept of Correlated Equilibrium (CE) was proposed by Nobel Prize winner, Robert J. Aumann [28]. CE is more general than NE, and the performance of CE is superior to NE. The definition of CE as follow.

**Definition 1** [28]: Let  $\mathcal{G} = [\mathcal{N}, \{\mathcal{A}_n\}_{n \in \mathcal{N}}, \{U_n\}_{n \in \mathcal{N}}]$  be a finite strategy game, where  $\mathcal{N}$  is the player set,  $\mathcal{A}_n$  is the strategy set of player  $n$  and  $U_n$  is the utility function of player  $n$ , a

probability distribution  $p$  is a Correlated Equilibrium of  $\mathcal{G}$  if and only if  $\forall n \in \mathcal{N}$ , it holds that

$$\sum_{r_{-n} \in \mathcal{A}_{-n}, r_n \in \mathcal{A}_n} p(r_n, r_{-n}) [U(r_n', r_{-n}) - U(r_n, r_{-n})] \leq 0 \quad (8)$$

We note that every NE is a CE and NE correspond to the special case where  $p(r_n, r_{-n})$  is a product of each individual players probability for different strategies, i.e., the play of the different players is independent [28]-[30].

### B. No-regret Learning Algorithm

In this subsection, we express the no-regret learning algorithm [30]. For any two actions  $r, r' \in LS, r \neq r'$  at time  $T$ , the regret of player  $n$  for not playing  $r'$  is,

$$R_n^T(r_n, r_n') = \max\{D_n^T(r_n, r_n'), 0\} \quad (9)$$

where  $D_n^T(r_n, r_n')$  is,

$$D_n^T(r_n, r_n') = \frac{1}{T} \sum_{t \leq T} [U_n^t(r_n', r_{-n}) - U_n^t(r_n, r_{-n})] \quad (10)$$

$D_n^T(r_n, r_n')$  is the average payoff that player  $n$  would have obtained, if it had played action  $r'$  every time in the past instead of choosing  $r$ . The detail of no-regret learning algorithm is shown in Table . However, the interference level is not too accurate because other SUs have an effect on channel state (free or busy) in the equation (2). And we can note that the aggregate normalized throughput is inversely proportional to the generalized interference level, So we redefine the utility function  $U_n(r_n, r_{-n})$ , as follow

$$U_n(r_n, r_{-n}) = \begin{cases} 1 - \tau, & \text{if player } n \text{ does not collide with other players} \\ 0, & \text{if player } n \text{ collides with other players} \end{cases} \quad (11)$$

Mathematically, the optimal objective is to maximize the aggregate normalized throughput, i.e.,

$$\max\left\{\sum_{n \in \mathcal{N}} U_n(r_n, r_{-n})\right\} \quad (12)$$

For every period  $T$ , we define the relative frequency of players strategy  $r$  played till  $T$  periods of time as follow:

$$z_T(r) = \frac{1}{T} S\{t \leq T : r_t = r\} \quad (13)$$

where  $S$  denotes the number of periods before  $T$  that the players strategy is  $r$ .

**Theorem 2**[30]:  $z_T$  is guaranteed to converge almost surely to a set of Correlated Equilibrium in no-regret learning algorithm.

TABLE I. NO-REGRET LEARNING ALGORITHM INTRODUCTION

### Algorithm 1 No-regret learning algorithm

**Initialization:** each player  $n$ , generate random probability for select an action  $r_n^0$ , for  $n \in \mathcal{N}, r \in LS$ .

**For**  $t = 1, 2, 3, \dots$  **do**

    Update the average regret  $R_n^t$  based on equations (9) (10).

    Let  $r_n^t$  denote the action which player  $n$  selects for iteration  $t$ , calculate  $p_n^{t+1}(r_n')$  as:

$$p_n^{t+1}(r_n') = \frac{1}{\mu} R_n^t(r_n', r_n), \text{ all } r_n' \neq r_n, \mu > 0$$

$$p_n^{t+1}(r_n) = 1 - \sum_{r_n' \neq r_n} p_n^{t+1}(r_n')$$

**End for**

Obviously, no-regret learning algorithm requires player  $n$  to know its own action and other players actions. It is difficult to realize in the distributed networks. Therefore, we proposed a modified no-regret learning algorithm.

### C. Modified No-Regret Learning Algorithm

In this subsection, we develop an algorithm based on no-regret learning algorithm. We study how the proposed modified no-regret learning algorithm can be implemented in the distributed networks. Here, we estimate player  $n$ 's utility function value for choosing other strategy  $r'$  as follow:

$$U_n(r_n', r_{-n}) = \frac{p_n(r_n')}{p_n(r_n)} U_n(r_n, r_{-n}), \text{ all } r_n' \neq r_n \quad (14)$$

But if  $U_n(r_n, r_{-n}) = 0$ , the estimation would fail. In the other hand, in order to reduce the probability of collision, we introduce a penalty mechanism, i.e., set the player  $n$ 's utility function value to -1 when player  $n$  collides with other players, as follow,

$$U_n(r_n, r_{-n}) = \begin{cases} 1 - \tau, & \text{if player } n \text{ does not collide with other players} \\ -1, & \text{if player } n \text{ collide with other players} \end{cases} \quad (15)$$

Consequently, the average regret payoff can be calculated based on only local information. The other steps of that algorithm are the same as algorithm 1. Furthermore, that algorithm can converge to CE in guaranteed by Theorem 2.

## V. SIMULATION RESULTS AND DISCUSSION

In the simulation, for convenience of discussion, we assume that the idle probabilities of all licensed channels are the same, i.e.,  $\theta_m = \theta, \forall m \in \{1, 2, \dots, M\}$ . The normalized



time used to sense each channel  $\tau=0.05$ , and the probability of false alarm is 0.05, i.e.,  $P_f=0.05$ . The learning parameter  $\mu$  is 4.

#### A. Coverage Property

In this subsection, we study the convergence property of the proposed no-regret learning algorithm in the distributed networks where any two SUs are neighbor users. It is assumed that there are nine channels and ten users, i.e.,  $M=9$  and  $N=10$ . The probability of any channel free is 1, i.e.,  $\theta=1$ . The convergence of throughput for two algorithms is shown in Fig.3. It is noted in the figure that two algorithms finally converge to a pure strategy. Moreover, modified no-regret learning algorithm requires 10 time slots more than no-regret learning algorithm.

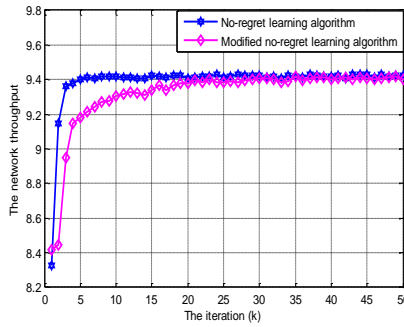


Figure 3. The convergence property of two algorithms, no-regret learning algorithm requires about 10 time slots, but modified algorithm requires about 20 time slots. ( $M=9$ ,  $N=10$ ,  $\theta=1$ )

#### B. Throughput Performance Comparison

In practice, the SUs randomly spread in the spatially distributed networks. Therefore, we take that consider into our work. Moreover, we illustrate the spatially distributed networks topology in Figure.4, ten SUs randomly distributed in the area of  $100\text{ m} \times 100\text{ m}$ , and the distance threshold of interference is 30m, i.e.,  $N=10$ ,  $d_0=30$ .

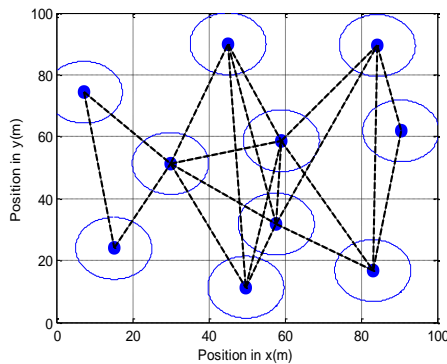


Figure 4. the topology of networks, the circle point denotes SUs and its radius of coverage is 10m, the dotted line represent interference between SUs.

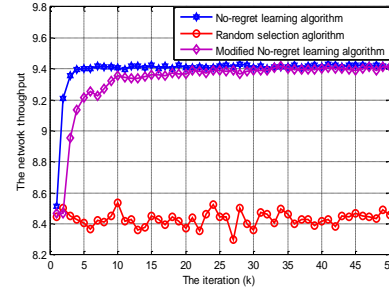


Figure 5. Throughput comparison of three algorithms in small-scale spatially distributed networks. ( $M=9$ ,  $N=10$ ,  $\theta=1$ )

Under this network topology, we compare the throughput performance of three algorithms including no-regret learning algorithm, modified no-regret learning algorithm and random selection algorithm. In the random selection algorithm, each player randomly select strategy. The comparison results of the network throughput as shown in Fig.5. There are ten SUs and nine channels, we call this small-scale networks because of the small number of SUs. At the same time, the probability of any channel free is 1, i.e.,  $\theta=1$ . It is noted in the Fig.5 that the normalized system throughput of modified no-regret learning algorithm is almost equal to no-regret learning algorithm. But compared with random selection algorithm, the throughput performance generated by two no-regret learning algorithms increases more than 10%.

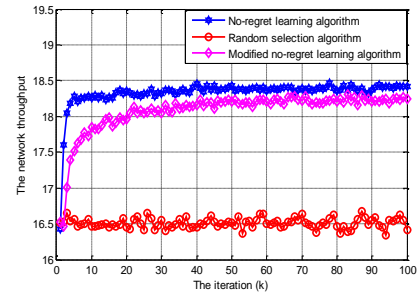


Figure 6. Throughput comparison of three algorithms in medium-scale spatially distributed networks. ( $M=9$ ,  $N=20$ ,  $\theta=1$ )

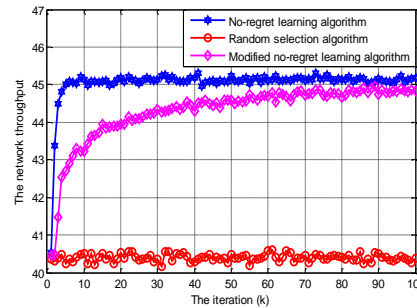


Figure 7. Throughput comparison of three algorithms in large-scale spatially distributed networks. ( $M=9$ ,  $N=50$ ,  $\theta=1$ )

Here, we also compare the throughput performance of three algorithms in medium-scale (the number of SUs is 20, the area is 200 m $\times$ 200 m) and large-scale (the number of SUs is 50, the area is 500 m $\times$ 500 m) networks. Meanwhile, other parameters keep constant. Fig.6 and Fig.7 show the throughput performance in medium-scale and large-scale networks. We can also see that the throughput performance by employing the two no-regret learning algorithms are obviously better than random selection algorithm.

## VI. CONCLUSION

In this paper, we studied the problem of multi-user sequential channel sensing and access in spatially distributed CR networks. In order to solve the problem, we modeled it into a non-cooperative game and defined a generalized interference metric. We established an optimization objective: maximize the aggregate normalized throughput. Considering that no-regret learning algorithm that required information exchange, we proposed a modified no-regret learning algorithm that did not require information exchange, and proved that two algorithms can both converge to CE. From the simulation results, the throughput performance of two algorithms was obviously better than random selection algorithm. In the future, we plan to consider users requirements in CR networks.

## REFERENCES

- [1] M. A. McHenry, NSF Spectrum Occupancy Measurements Project Summary, Shared Spectrum Company report, Tech. Rep., 2005.
- [2] FCC, FCC-03-322-NOTICE OF PROPOSED RULE MAKING AND ORDER, Federal Communications Commission, Tech. Rep., 30 Dec.2003.
- [3] J. Mitola III and G. Q. Maguire Jr., Cognitive Radio: Making Software Radio more Personal, *IEEE Pers. Communications*, vol. 6, no. 4, pp.1318, Aug. 1999.
- [4] A. Mendes, C. Augusto, M. Silva, R. M. Guedes, and J. F. de Rezende, Channel sensing order for cognitive radio networks using reinforcement learning, in *Proc. 36th Annu. IEEE Conf. Local Comput. Netw.*, 2011, pp. 546553.
- [5] S. Huang, X. Liu, and Z. Ding, Opportunistic spectrum access in cognitive radio networks, in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, Phoenix, AZ, USA, Apr. 2008, pp. 14271435.
- [6] A. Ghasemi and E. S. Sousa, Spectrum sensing in cognitive radio networks: requirements, challenges and design trade-offs, *IEEE Communications Magazine*, vol. 46, no. 4, pp. 3239, Apr. 2008.
- [7] R. Fan and H. Jiang, Optimal multi-channel cooperative sensing in cognitive radio networks, *IEEE Transactions on Wireless Communications*, vol. 9, no. 3, pp. 11281138, Mar. 2010.
- [8] R. Fan and H. Jiang, Channel sensing-order setting in cognitive radio networks: a two-user case, *IEEE Transactions on Vehicular Technology*, vol. 58, no. 9, pp. 49975008, Nov. 2009.
- [9] H. T. Cheng and W. Zhuang, Simple channel sensing order in cognitive radio networks, *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 4, pp. 676688, 2011.
- [10] A. Anandkumar, N. Michael, and A. Tang, Opportunistic spectrum access with multiple users: Learning under competition, in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, 2010, pp. 19.
- [11] N. B. Chang and M. Liu, Competitive analysis of opportunistic spectrum access strategies, in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, Apr. 2008, pp. 15351542.
- [12] K. Liu and Q. Zhao, Distributed learning in multi-armed bandit with multiple players, *IEEE Transactions on Signal Processing*, vol. 58, no. 11, pp. 56675681, Nov. 2010.
- [13] Z. Khan, J. Lehtomki, L. DaSilva, and M. Latva-aho, Autonomous sensing order selection strategies exploiting channel access information, *IEEE Trans. Mobile Comput.*, vol. 12, no. 2, pp. 274288, Feb. 2013.
- [14] Q. Zhao, S. Geirhofer, L. Tong, and B. M. Sadler, Opportunistic spectrum access via periodic channel sensing, *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 785796, Feb. 2008.
- [15] A. Anandkumar, N. Michael, and A. Tang, Opportunistic spectrum access with multiple users: Learning under competition, in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, 2010, pp. 19.
- [16] A. Anandkumar, N. Michael, A. K. Tang, and A. Swami, Distributed algorithms for learning and cognitive medium access with logarithmic regret, *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 4, pp. 731745, 2011.
- [17] H. Li, Multi-agent Q-learning for Aloha-like spectrum access in cognitive radio systems, *EURASIP Journal on Wireless Communications and Networking*, vol. 2010, pp. 115, Apr. 2010.
- [18] K. Liu and Q. Zhao, Distributed learning in cognitive radio networks: Multi-armed bandit with distributed multiple players, in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2010, pp. 30103013.
- [19] Y. Gai, B. Krishnamachari, and R. Jain, Learning multiuser channel allocations in cognitive radio networks: A combinatorial multiarmed bandit formulation, in *Proceedings of the IEEE International Dynamic Spectrum Access Networks (DySPAN)*, 2010, pp. 19.
- [20] H. Liu, B. Krishnamachari, and Q. Zhao, Cooperation and learning in multiuser opportunistic spectrum access, in *Proceedings of the IEEE International Conference on Communications Workshops (ICC workshops)*, 2008, pp. 487492.
- [21] N. B. Chang and M. Liu, Competitive analysis of opportunistic spectrum access strategies, in *Proceedings of the IEEE International Conference on Computer Communications (INFOCOM)*, Apr. 2008, pp. 15351542.
- [22] H. Jiang, L. Lai, R. Fan, and H. V. Poor, Optimal selection of channel sensing order in cognitive radio, *IEEE Transactions on Wireless Communications*, vol. 8, no. 1, pp. 297307, Jan. 2009.
- [23] N. B. Chang and M. Liu, Optimal channel probing and transmission scheduling for opportunistic spectrum access, in *Proceedings of the Second Workshop on Information Theory and Applications*, 2007, pp. 2738.
- [24] Y. Xu, Q. Wu, J. Wang, L. Shen, and A. Anpalagan, Robust multiuser sequential channel sensing and access in dynamic cognitive radio networks: Potential games and stochastic learning, *IEEE Transactions on Vehicular Technology*, vol. 64, no. 2, pp. 3594-3607, August. 2015.
- [25] C. F. Laywine and G. L. Mullen, Discrete Mathematics using Latin Squares, 1st ed., ser. Wiley-Interscience Series in Discrete Mathematics and Optimization. New York: John Wiley & Sons, 1998.
- [26] P. J. Davis, Circulant Matrices, 2nd ed. Chelsea Publishing Company, 1994.
- [27] R.B.Myerson, Game theory:analysis of conflict,1991.
- [28] R. J. Aumann, Subjectivity and correlation in randomized strategy, *Journal of Mathematical Economics*, vol. 1, no. 1, pp. 67-96, 1974
- [29] R. J. Aumann, Correlated equilibrium as an expression of Bayesian rationality, *Econometrica*, vol. 55, no. 1, pp. 1 - 18, January 1987.
- [30] S. Hart and A. Mas-Colell, A simple adaptive procedure leading to correlated equilibrium, *Econometrica*, vol. 68, no. 5, pp. 1127-1150, September 2000.