# Stable Multiuser Channel Allocations in Opportunistic Spectrum Access

Li Huang and Guangxi Zhu
Department of Electronic and Information Engineering
Huazhong University of Science and Technology, China
Email: szemmah@hotmail.com

Xiaojiang Du
Department of Computer and Information Sciences
Temple University
Philadelphia, USA
Email: dxj@ieee.org

Kaigui Bian
School of Electrical Engineering and Computer Science
Peking University, China
Email: bkg@pku.edu.cn

***Abstract*—We consider the distributed channel allocation problem in an asymmetrical opportunistic spectrum access (OSA) system where each secondary user possibly has different channel reward even in the same channel due to geographic dispersion. We formulate this problem as a Gale-Shapley stable theorem using game theory to optimize the sum reward of all secondary users. It is challenging to achieve the stable matching of user-channel pairs without centralized control and prior knowledge of channel availability statistics. In this paper, we present a novel decentralized order-optimal learning Gale-Shapley scheme (OLGS) in which secondary users learn from their local history data and individually adjust their behaviors in a time-varying OSA system. The proposed scheme eliminates collisions among secondary users by a one-to-one user-channel matching policy. It also achieves stable spectral allocations using learning method without assuming known channel parameters and independent of information exchange among secondary users. Simulation results show that the system regret of the OLGS solution grows with time at the logarithmic order with low complexity.**

## I. Introduction

Cognitive radio (CR) technology is a promising solution to resolve the dilemma between spectrum scarcity and inefficient spectrum usage in licensed spectrum bands [1]. One of underlying models is the interweave paradigm (or opportunistic spectrum access, OSA) where secondary users (i.e., unlicensed users or cognitive users) are allowed to access spectrum holes only when they do not interfere with high-priority primary users (i.e., licensed users) [2]. A lot of algorithms for opportunistic channel selection have been published [3]-[8]. Most of them are based on two assumptions. Firstly, existing work assumes that secondary users have symmetrical transmission rates in accessing the same channel [3]-[6]. Secondly, full knowledge of channel parameters or information about other users operations is assumed known [7]-[8]. However, these assumptions may not be valid in a realistic asymmetrical network. Due to geographic dispersion in proximity of different primary users, spectral opportunities available to secondary users may be dissimilar even in the same channel. Moreover, spectrum holes are time-varying due to burst traffic of primary users. Hence, it is difficult for secondary users to have complete knowledge of channel parameters and fully sense all channels in a limited sensing period. Furthermore, mass information exchange about actions among secondary users also leads to large communication overhead in a distributed CR network.

In this paper, we study the channel allocation problem without the above assumptions. We consider an OSA system with the following characteristics: (1) Secondary users are dispersed in overlay CR network and have asymmetrical channel rewards. (2) Primary traffic statistics are not available. (3) There is no centralized controller node and no information exchange among secondary users. In this case, each secondary user can only utilize their individual history decision and local observation to independently access a channel. In order to improve the system throughput, it is desirable to design an intelligent and distributed cognitive spectrum sensing and accessing policy.

In this paper, we present a novel decentralized order-optimal learning Gale-Shapley scheme (OLGS) to achieve this goal. The problem is formulated as a game theoretic Gale-Shapley stable theorem due to two reasons. The first reason is that a one-to-one matching of user-channel pairs minimizes collisions among secondary users if at most one secondary user gains reward from a channel. Since the Gale-Shapley (G-S) theorem always has a stable one-to-one matching for any preference function [9], it can avoid multi-user contention under this interference model. The second reason is that the G-S theorem has a unique stable matching when entries of the preference matrix are all different [8]. Hence, the G-S theorem is suitable to solve the spectral allocation problem in asymmetrical networks. Besides, we exploit the order-optimal learning rule to implement the stable matching in an unknown random environment because it is an efficient algorithm which achieves the logarithmic order of the regret.

Our main contributions are summarized as follows: (1) We apply the G-S theorem to allocate channels that effectively avoids multi-user collisions. (2) When channel rewards are not available and information exchange is restricted, we exploit the order-optimal learning algorithm to achieve the stable matching. This distributed OLGS algorithm neither needs prior knowledge of primary traffic, nor requires information exchange about secondary actions.

The rest of this paper is organized as follows. We describe the system model as well as the problem formulation in Section II. We propose the distributed OLGS scheme in section III. Simulation results are presented in Section IV. We conclude the paper in Section V.

## II. System Model and Problem Formulation

### A. System Model

We consider an OSA system with $N$ orthogonal channels in licensed bands and $M$ secondary users, where $N \geq M > 1$. In an asymmetrical CR network, different secondary users experience different transmission rates in the same channel due to widespread locations and diverse channel conditions. The above system model is more realistic than the symmetrical channel rate assumption in some previous works.

An asymmetrical OSA scenario is illustrated in Fig. 1. Three geographically dispersed secondary users, S1, S2, and S3, are assumed to be in three different hierarchical areas of three primary users labeled as P1, P2, and P3, who occupy channel 1, 2, 3 respectively. Each value in the table denotes the reward for a secondary user operating in a channel. If the reward table is available, channel allocation solution becomes straightforward. As the example shows, the optimal allocation policy is to assign channel 2 to user S1, channel 3 to user S2, and channel 1 to user S3. However, primary traffic information and secondary channel quality are time-varying such that rewards are difficult to be known in advance. Therefore, an online learning policy is required when secondary users only have partial channel sensing capabilities. In addition, since the optimal centralized allocation is to exhaustively search the maximum overall throughput in permutations $P(N, M)$, it is desirable to find a distributed policy that can reduce computational complexity.

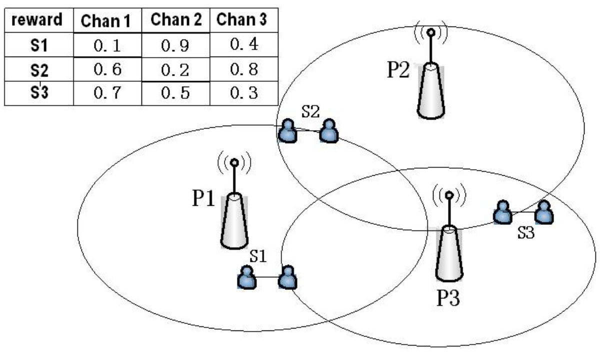


Fig. 1. An asymmetrical OSA scenario

### B. Problem Formulation

In our formulation, $M$ secondary user pairs (or equivalently $M$ pairs of secondary transmitters and receivers) share $N$ primary channels. A simple channel utility (or reward) is the ergodic capacity of a user on a channel. We denote $R_{i,j}(n)$ as the actual reward, i.e., the normalized available transmission rate or throughput obtained by secondary user $i$ on channel $j$ at time slot $n$.

$$R_{i,j}(n) = [\frac{T - T_s(n)}{T}] \cdot I_{i,j} \cdot E_i(K_j) \cdot C_{i,j} \tag{1}$$

where $T$ denotes the duration of a slot and $T_s(n)$ is channel sensing and waiting time in a slot. Sensing time is decided by the hardware detection period. Waiting time is related to the secondary users back-off duration that monotonically decreases with the channel reward. Hence, a secondary user with more channel reward has higher priority to access this channel. The channel reward depends on the primary traffic and interaction of secondary users. Let $I_{i,j}$ be the channel idle indication where "1" is idle and "0" is busy. $E_i(K_j)$ is the multi-user contention indication where "1" is successful and "0" is failed. Both $I_{i,j}$ and $E_i(K_j)$ are Bernoulli random variables, and their probability density functions (pdfs) are given below:

$$\rho_{\theta_{i,j}}(I_{i,j}) = \theta_{i,j} \cdot \delta(I_{i,j} - 1) + (1 - \theta_{i,j}) \cdot \delta(I_{i,j}) \tag{2}$$

where $\delta(\cdot)$ is the delta function and $\theta_{i,j}$ is the mean idle probability of the primary channel.

$$\rho(E_i(K_j)) = \frac{1}{K_j} \cdot \delta(E_i(K_j) - 1) + (1 - \frac{1}{K_j}) \cdot \delta(E_i(K_j)) \tag{3}$$

where $K_j$ indicates the number of competitive secondary users. The expectation $E_i(K_j)$ is inversely proportional to the number of competitive secondary users $K_j$. $C_{i,j}$ is the channel capability of secondary user $i$ under i.i.d Rayleigh fading channel $j$.

$$C_{i,j} = log(1 + SNR \cdot X) \tag{4}$$

where $X \sim \chi_2^2$ and $SNR = \frac{P_t}{\sigma^2}$.

Based on (1)-(4), the expected reward achieved by secondary user $i$ on channel $j$ is given by

$$E[R_{i,j}] = E\left\{[\frac{T - T_s}{T}] \cdot \frac{\theta_{i,j} \cdot C_{i,j}}{K_j}\right\} \tag{5}$$

The utility matrix $\mathbf{R}$ has $M$ row by $N$ column, in which each element is the preference of a secondary user to a channel. Since secondary users are located in widespread areas, channel utilities of secondary users are different. This characteristic is suitable to exploit the Gale-Shapley stable theorem.

$$\mathbf{R} = \begin{pmatrix} R_{1,1} & R_{1,2} & \dots & R_{1,N} \\ R_{2,1} & R_{2,2} & \dots & R_{2,N} \\ \vdots & \vdots & \vdots & \vdots \\ R_{M,1} & R_{M,2} & \dots & R_{M,N} \end{pmatrix} \tag{6}$$

The overarching utility function is to maximize

$$E[U] = E\Big[\sum_{n=1}^{K} \sum_{i=1}^{M} R_{i,\pi(i)}(n)\Big] \tag{7}$$

Where $\pi(i)$ is the channel selected by secondary user $i$ using strategy $\pi$. The objective function is to maximize overall throughput of all secondary users under primary interference restriction. If the utility matrix $\mathbf{R}$ is given, the optimal centralized scheme based on the Hungarian method is to search the maximum sum reward in collision-free permutation $P(N, M)$. The challenges are that secondary users are unaware of the channel-idle probability and no information exchange is available among secondary users.

The loss of a learning strategy $\pi$ can be expressed as:

$$\Gamma(R;\pi) = nR^* - E\Big[\sum_{n=1}^{K}\sum_{i=1}^{M} R_{i,\pi(i)}(n)\Big]$$

$$\leq R^* \cdot \left\{ \sum_{i=1}^{M} \sum_{j \in M-worst} E[T_{i,j}(n)] + E[M(n)] \right\} \quad (8)$$

where $R^* = max \sum_{(i,j \in U)} R_{i,j}$ is the expected maximum overall reward. $T_{i,j}(n)$ is the slot number that secondary user $i$ senses channel $j$. $M(n)$ is the collision number that is made by secondary users in $M$-best channels. The upper bound of regret under any distributed policy is represented in (8). The first term expresses performance loss for selection of the $M$-worst channels, while the second term involves lost transmission opportunities due to collisions among secondary users in the $M$-best channels [10].

## III. Order-optimal Learning Gale-Shapley Scheme

To minimize multi-user collision loss, we exploit the Gale-Shapley theorem to obtain a one-to-one stable matching between secondary users and primary channels. Furthermore, we adopt the order-optimal learning algorithm to implement the stable matching without prior channel knowledge and full sensing capabilities supported in a dynamic environment. Without central control, back-off timers are used to implement distributed spectral allocations. Next, we briefly describe the Gale-Shapley theorem and the learning algorithm.

### A. The Gale-Shapley Theorem

The Gale-Shapley theorem [9] is a well-known one-to-one stable matching algorithm concerned with the problem of college admission and marriage stability. There always exists an iterative procedure to find a stable set of marriage. The uniqueness is proved by induction [8] in the context of cognitive spectral allocation problem. The convergence rate of the Gale-Shapley algorithm can be accelerated by exploiting multi-channel sensing capabilities.

In Fig. 2, we give an example to illuminate the Gale-Shapley theorem. Table I shows a 3 by 5 utility matrix for three secondary users and five channels in an OSA network. Each element is the utility of a user-channel pair and the value in parentheses is the back-off timer value. Fig. 2 shows ping time from the utility matrix. Based on the Gale-Shapley theorem, secondary users 1, 2, and 3 are allocated to channel3 ($t$=0.6), channel2 ($t$=0.7) and channel5 ($t$=0.83), respectively.

TABLE I
Utility Matrix

| | CH1 | CH2 | CH3 | CH4 | CH5 |
|---|---|---|---|---|---|
| User1 | 5(0.95) | 11(0.89) | 40(0.60) | 22(0.78) | 13(0.87) |
| User2 | 25(0.75) | 30(0.70) | 35(0.65) | 14(0.86) | 29(0.71) |
| User3 | 12(0.88) | 28(0.72) | 21(0.79) | 4(0.96) | 17(0.83) |

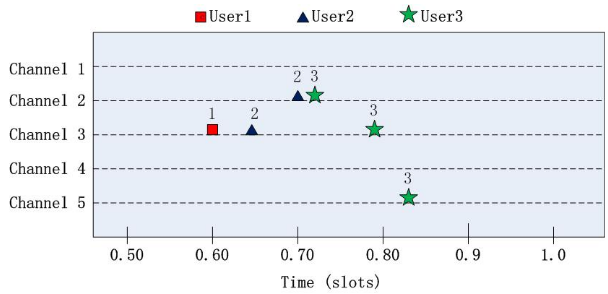


Fig. 2. The Gale-Shapley theorem ping time

### B. Order-optimal Learning Algorithm

In a dynamic CR network, it is difficult to obtain complete channel information and incapable to monitor all channels simultaneously due to hardware constraints. A learning method is necessary to achieve the stable matching of the Gale-Shapley theorem. The existing learning algorithms, such as regret learning, reinforcement learning and learning automata, can learn unknown channel parameters in a time-varying wireless environment. To achieve the consistency and less regret with fast convergence, we adopt a linear complexity learning algorithm to estimate the reward value. The order-optimal learning rule is simple and efficient which is extensively applied to a tradeoff between exploration and exploitation [5,7]. According to the finite-time analysis result, the estimated $\hat{R}_{m,n}(j)$ will converge to the real $R_{m,n}(j)$ as step $j$ increases as long as sample number of each channel grows as fast as $O(lnT)$ [11].

The notations are given below:

$X_{m,n}(j)$: number of times for which secondary user $m$ has successfully accessed channel $n$ in time slot $j$.

$Y_{m,n}(j)$: number of times for which secondary user $m$ has selected channel $n$ in time slot $j$.

$C_{m,n}(j)$: capability of channel $n$ to which secondary user $m$ has accessed in time slot $j$.

$\hat{R}_{m,n}(j)$: estimated channel reward of secondary user $m$ in channel $n$ in time slot $j$.

The order-optimal learning algorithm is given below:

- **Initialization**: At the beginning of each time slot, each secondary user senses every channel once.
- **Loop**: The secondary user chooses the channel that maximizes the reward $\hat{R}_{m,n}(j)$, then updates $X_{m,n}(j)$,

$Y_{m,n}(j)$ and $C_{m,n}(j)$ after each accessing.

$$\hat{R}_{m,n}(j) = \left[ (\frac{X_{m,n}(j)}{Y_{m,n}(j)}) + \sqrt{\frac{2lnj}{Y_{m,n}(j)}} \right] \cdot \bar{C}_{m,n}(j) \tag{9}$$

The learning algorithm can achieve the logarithmic regret uniformly over the iteration number $j$ without any knowledge about the reward distribution. In [11], the expected regret of the order-optimal learning algorithm is proved to be at most:

$$L(R;\Gamma_{UCB1}) \leq \left[ 8 \sum_{i:\mu_i<\mu^*} (\frac{lnj}{\Delta_{m,i}}) + (1+\frac{\pi^2}{3})(\sum_{k=1}^{N} \Delta_{m,k}) \right]$$

$$\Delta_{m,n} = R_m^* - R_{m,n} \qquad L(R;\Gamma_{UCB1}) \sim O(lnT) \tag{10}$$

The learning cost results from the time spent in sampling inferior channels during learning true value is $R_{m,n}(j)$.

### C. Order-optimal Learning Gale-Shapley Scheme

Now we present a joint order-optimal learning scheme with the Gale-Shapley theorem to allocate channels in an asymmetrical CR network. In order to realize the distributed algorithm, we set back-off timers to implement a rejection mechanism, which does not need information sharing in a central node. The OLGS solution is described below.

1) Initially: Set the monotonically decreasing function $\mathbf{T} = f(\mathbf{R})$ and calculate back-off timers in light of $f(\mathbf{R})$, $t$=0, $Y_{m,n}(1)$ =1, $X_{m,n}(1)$ =1, $C_{m,n}(1)$ =1, $j$=1.
   For the secondary user $m$, set $a(m)$=0, $done(m)$=0, $(m = 1, \ldots, M)$.
   For the primary channel $n$, set $idle(n)$=0, $c(n)$=0, $(n = 1, \ldots, N)$.
   the channel utility $R_{m,n}(1) = 1/N, \forall m \in M, n \in \{1, \ldots, N\}$.
2) At the beginning of each time slot, secondary user $m$ calculates and sets the back-off timer
   $$\{timer(R_{m,n}(j)) \mid m,n\}(m = 1, \ldots, M; n = 1, \ldots, N)$$
   in light of the function $\mathbf{T} = f(\mathbf{R})$.
3) At every time $t$ when a back-off timer expires, $t \in \{timer(R_{m,n}(j)) \mid m,n\}$ do:
   For secondary user $m$ s.t. $done(m)$=0 selects channel $n_m^*(t)$, where $n_m^*(t)$ is determined by:
   $$n_m^*(t) = \{n \mid timer(R_{m,n}(j)) = t\}.$$
   That is, $n_m^*(t)$ is the channel that leads to an unexpired minimal back-off timer or the residual maximal channel reward. $Y_{m,n_m^*(t)}(j) = Y_{m,n_m^*(t)}(j) + 1$.
4) If $idle(n_m^*(t))$=1 and $c(n_m^*(t))$=0, then $a(m) = n_m^*(t)$, $done(m) = 1, X_{m,n_m^*(t)}(j) = X_{m,n_m^*(t)}(j) + 1$.
   That is, if the selected channel $n_m^*(t)$ is idle and no collisions among secondary users on this channel, secondary user $m$ access channel $n_m^*(t)$ and get the channel transmission rate $C_{m,n_m^*(t)}(j)$.
5) If $idle(n_m^*(t))$=0 or $c(n_m^*(t))$=1, then $done(m)$=0, exclude $n_m^*(t)$ from the channel set $\mathbf{N}$, then go to step (3).
   That is, if the selected channel $n_m^*(t)$ is busy or has collision among secondary users over this channel, secondary user $m$ excludes $n_m^*(t)$ from the channel set $\mathbf{N}$ then go to step (3) to update $n_m^*(t)$ until finding suitable channel.
6) At the end of the $j$th slot, secondary user $m$ updates
   $$\hat{R}_{m,n}(j) = \left[ (\frac{X_{m,n}(j)}{Y_{m,n}(j)}) + \sqrt{\frac{2lnj}{Y_{m,n}(j)}} \right] \cdot \bar{C}_{m,n}(j)$$
   $$as: \bar{C}_{m,n}(j) = \sum_{t=1}^{j} C_{m,n}(t)/j.$$
   Secondary user updates each channel utility related with successful access probability and average channel capability.
7) Go to step (3) and repeat the process.

## IV. Simulation Results

We conduct simulation experiments with MATLAB software to evaluate the system throughput and the convergence rate of the OLGS scheme. The simulation results are shown in Fig. 3 and 4. Several algorithms are compared: optimal, random, order-optimal learning, and the OLGS algorithm. The optimal algorithm serves as an upper bound of the achievable throughput where a central controller implements exhaustive searching under ideal assumptions of known system parameters. The random allocation scheme is that secondary users access an arbitrary channel with equal probability. In the order-optimal algorithm, secondary users learn unknown channel parameters but randomly access a channel after sensing. The OLGS algorithm is different from other schemes in which it achieves a stable one-to-one user-channel matching with a collision-free mechanism among secondary users.

### A. Performance of the Asymmetrical OSA System

We consider an asymmetrical OSA system where different secondary users have diverse channel utilities. The system parameters are randomly set as follows: $M$=3, $N$=5,

$$\mathbf{R} = \begin{pmatrix} 0.8284 & 0.0253 & 0.7813 & 0.9182 & 0.9106 \\ 0.9161 & 0.3270 & 0.4243 & 0.8545 & 0.6396 \\ 0.0076 & 0.1803 & 0.8674 & 0.8286 & 0.4646 \end{pmatrix}$$

Simulation results are averaged over $10^5$ independent trials.

The optimal channel allocation pairs are: user1-to-channel4, user2-to-channel1 and user3-to-channel3. The overall throughput of the optimal scheme is about 2.70. Fig. 3 compares the system throughput of various algorithms and shows that the overall throughput of the order-optimal learning scheme (2.39) is more than that of the random selection approach (1.57), but less than that of the OLGS algorithm (2.62). The reason is that the single learning approach explores more idle opportunities than the random method but it cannot effectively avoid multi-user collisions.

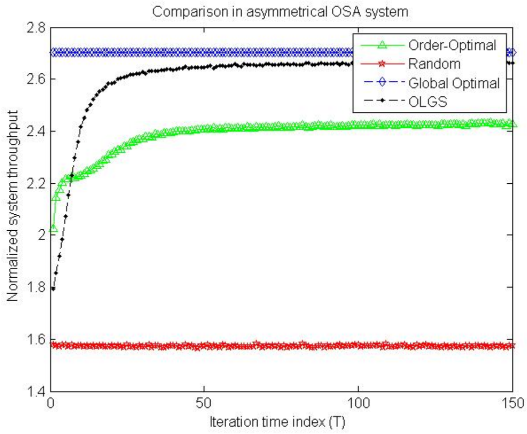


Fig. 3. Comparison of normalized system throughput of various schemes in an asymmetrical OSA system

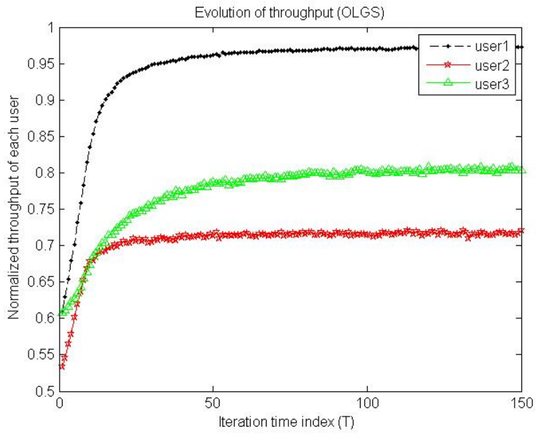


Fig. 4. Evolution of the throughput of each secondary user

In contrast, the OLGS algorithm exploits back-off timers to implement a one-to-one stable matching so that multi-user collision loss is negligible. Since the best channel of each secondary user is different in the case, all secondary users are spread over different channels with the OLGS algorithm. The proposed algorithm considerably outperforms the random allocation approach and the learning algorithm. The proposed OLGS algorithm gradually approaches the optimal allocation through the learning process. Fig. 3 shows that the throughput under the OLGS algorithm is improved owing to the one-to-one stable matching in an asymmetrical OSA system.

### *B. The Convergence Rate*

To measure the convergence rate of the OLGS scheme, channel reward matrix is randomly generated as follows:

$$\mathbf{R} = \begin{pmatrix} 0.6892 & 0.9845 & 0.2898 & 0.8752 & 0.2197 \\ 0.7371 & 0.7208 & 0.0156 & 0.4974 & 0.6887 \\ 0.5814 & 0.6440 & 0.5479 & 0.8395 & 0.4394 \end{pmatrix}$$

The stable matching policy based on the Gale-Shapley scheme is: user1-to-channel2, user2-to-channel1 and user3-to-channel4. Fig. 4 shows that the achievable throughput of the OLGS algorithm converges to the stabilization in about 45 iterations. Each secondary user selects its own best channel at the end and the average throughput converges to its largest channel reward.

In Table II, we compare the expected ratio to optimal rate and the convergence time of several algorithms. Note that the convergence time of the optimal scheme is $O(N^3)$. The OLGS algorithm not only converges to the stable matching status as quickly as the order-optimal learning algorithm but also achieves a higher throughput.

TABLE II
PROTOCOLS COMPARISON

| Algorithm | Type | Ratio to optimal rate | Convergence time(slots) |
|---|---|---|---|
| **Optimal** | centralized | 1 | $O(N^3)$ |
| **Random** | distributed | $\sim$ 0.58 (simulation) | / |
| **Order-optimal** | distributed | $\sim$ 0.80 (simulation) | $O(logN)$ |
| **OLGS** | distributed | $\sim$ 0.95 (simulation) | $O(logN)$ |

## V. CONCLUSION

In this paper, we studied the problem of distributed channel allocations in a time-varying asymmetrical OSA system, in which secondary users have limited sensing capabilities and lack prior channel information. We presented an order-optimal learning Gale-Shapley scheme in which secondary users learn from their individual history data and adapt to a dynamic network. The Gale-Shapley spectrum sharing provides a one-to-one user-channel matching policy to avoid collisions among secondary users. The order-optimal learning algorithm implements the stable spectral matching without prior channel knowledge and mass information exchange. Simulation results demonstrated that about 95% of the optimal rate is achieved by the stable matching for collision-free spatio-spectral reuse in an asymmetrical network. The OLGS scheme improves system throughput with fast convergence and achieves logarithmic regret over time. The proposed solution is completely distributed and it provides cognitive medium access without requiring full capabilities of sensing all channels. Furthermore, it does not require the powerful signal processing hardware and the incurred computation cost is low.


## ACKNOWLEDGMENT

This work was supported by the key project of National Natural Science Foundation in China (Grant No. 61231007,61172174), Major State Science and Technology Specific Projects of China (Grant No. 2010ZX03005-003,2010ZX03005-001,2011ZX03003-001-04), and by the US National Science Foundation under grants CNS-0963578, CNS-1022552 and CNS-1065444.


## REFERENCES


[1] I. F. Akyildiz, W. Y.Lee, M. C. Vuran and S. Mohanty, *"Next generation / dynamic spectrum access/cognitive radio wireless networks: a survey"*, Computer Network Journal, vol. 50, no. 13, pp. 2127-2159, 2006.

[2] A. Goldsmith, S. A. Jafar, I. Maric and S. Srinivasa, *“Breaking spectrum gridlock with cognitive radios: an information theoretic perspective“*, IEEE Proc.,vol. 97, no. 5, pp. 894-914, 2009.
[3] Y. Chen, Q. Zhao and A. Swami, *“Joint design and separation principle for opportunistic spectrum access in the presence of sensing erors“*, IEEE Trans. on Inform. Theory, vol. 54, no. 5, pp. 2053-2071, 2008.
[4] K. Liu and Q. Zhao, *“Distributed learning in multi-armed bandit with multiple players“*, IEEE Trans. on Signal. Proc., vol. 58, no. 11, pp. 5667-5681, 2010.
[5] L. Lai, H. E. Gamal, H. Jiang and H. V. Poor, *“Cognitive medium access: exploration, exploitation, and competition“*,IEEE Trans. on Mobile Comp., vol. 10, no. 2, pp. 239-253, 2011.
[6] Y. Xu, J. Wang, Q. Wu, A. Anpalagan and Y. D. Yao,*“Opportunistic spectrum access in unknown dynamic environment: a Game-Theoretic stochastic learning solution“*, IEEE Trans. on Wireless Comm., vol. 11, no.4, pp. 1380-1391, 2012.
[7] Y. Gai, B. Krishnamachari and R. Jain,*“Learning multiuser channel allocations in cognitive radio networks: a combinatorial multi-armed bandit formulation“*, IEEE DySPAN proc., pp. 1-9,2010
[8] A. Leshem, E. Zehavi and Y. Yaffe, *“Multichannel opportunistic carier sensing for stable channel access control in cognitive radio systems“*, IEEE JSAC., vol. 30, no. 1, pp. 82-95, 2012.
[9] D. Gale. and L. S. Shapley, *“College admissions and the stability of marriage“*,the American Mathematical Monthly, vol. 69, no. 1, pp. 9-15, 1962.
[10] A. Anandkumar, N. Michael, A. K. Tang and A. Swami, *“Distributed Algorithms for Learning and Cognitive Medium Access with Logarithmic Regret“*, IEEE JSAC.,vol. 29, no. 4, pp. 731-745, 2011
[11] P. Auer, N. C. Blanchi and P. Fischer,*“Finite-time analysis of the multiarmed bandit problem“*,Machine Learning, vol.47, pp. 235-256, 2002