Two-Dimensional POMDP-Based Opportunistic Spectrum Access in Time-Varying Environment with Fading Channels

Yumeng Wang, Yuhua Xu, Liang Shen, Chenglong Xu, and Yunpeng Cheng

Abstract: **In this research, we study the problem of opportunistic spectrum access (OSA) in a time-varying environment with fading channels, where the channel state is characterized by both channel quality and the occupancy of primary users (PUs). First, a finite-state Markov channel model is introduced to represent a fading channel. Second, by probing channel quality and exploring the activities of PUs jointly, a two-dimensional partially observable Markov decision process framework is proposed for OSA. In addition, a greedy strategy is designed, where a secondary user selects a channel that has the best-expected data transmission rate to maximize the instantaneous reward in the current slot. Compared with the optimal strategy that considers future reward, the greedy strategy brings low complexity and relatively ideal performance. Meanwhile, the spectrum sensing error that causes the collision between a PU and a secondary user (SU) is also discussed. Furthermore, we analyze the multiuser situation in which the proposed single-user strategy is adopted by every SU compared with the previous one. By observing the simulation results, the proposed strategy attains a larger throughput than the previous works under various parameter configurations.**

Index Terms: **Cognitive radio, finite-state Markov channel, opportunistic spectrum access, partially observable Markov decision process.**

I. INTRODUCTION

In cognitive radio networks (CRNs) [1]–[3], opportunistic spectrum access (OSA) [4]–[12] is one of the main functions, where the activities of primary users (PUs) are not interfered by secondary users (SUs) who are sensing and accessing the instantaneous available spectrum. In general, the vacant spectrum information is obtained by two main approaches: spectrum sensing and geolocation database. According to the current standardization rules in the US and Europe [13], [14], the spectrum sensing approach is used only in special cases such as ad hoc networks [22], where the channel availability is modeled as a threshold-based hypothesis test of licensed signals and where sensing errors exist [15], [16]. The geolocation database seems to provide a technically feasible and commercially viable solution in the future, which avoids the sensing errors and is adopted by Federal Communications Commission (FCC) while accessing television white space (TVWS) [17]. In [18]–[20], the "SenseLess" database system is presented, where TVWS availability is determined by unlicensed devices via a geolocation database service. Geolocation database approaches are mainly proposed to determine the TVWS availability in the space domain, i.e., the geographical area not covered by licensed PUs (such as digital television (DTV) users and wireless microphones). Spectrum sensing can be used to determine the TVWS in both the space domain and the time domain [16], i.e., the primary signal is absent, which is suitable for ad hoc networks and scenarios of SUs without location capability. In this study, threshold-based spectrum sensing is adopted in ad hoc networks as in [22].

Because of the hardware and energy constraints, SUs generally cannot sense all channels simultaneously. To address these challenges, the decision-theoretic solution of partially observable Markov decision process (POMDP) [21] is proposed to address this problem for OSA systems, where SUs can only observe a part of the system state. Recently, some strategies based on POMDP have been proposed [22]–[27]. In [22], the authors proposed a POMDP framework by investigating the decentralized medium access control (MAC) protocol to maximize the expected throughput. The authors of [23] proposed stationary optimal sensing and access policies based on bursty traffic in energy-constrained POMDP. The impacts of sensing errors were also investigated by the authors of [24], where the POMDP in the presence of sensing errors was formulated as a constrained POMDP. In [25], the authors tried to find an optimal policy to maximize the expected net reward considering the following problems: The delay cost connected with staying idle in a time slot, the energy cost connected with spectrum sensing and data transmissions, and the throughput gain connected with successful transmissions. To counter the view proposed by other scholars that POMDP is not suitable for unslotted PUs, in [26] the authors analyzed this situation in particular. In [27], based on cooperative spectrum sensing, the authors designed a multi-channel MAC protocol. However, most of them disregarded the practical problems in a fading environment, particularly the problem of channel quality. Therefore, we focus on the problem of channel quality using the framework of POMDP for OSA systems. When the channels are not occupied by the PUs, the SUs are willing to select the channel with the best quality to transmit data. Compared to the previous works, this study provides a solution to solve the problems related to channel quality in a fading environment.

It is significant to solve this problem because of the fading characteristic of wireless channels, which is caused by multiple

Manuscript received September 15, 2013.

This work was supported by the National Science Foundation of China under No.61172062, and in part by Jiangsu Province Natural Science Foundation of China under Grant No. BK2011116. The contact author is Yuhua Xu.

The authors are with the Institute of Communications Engineering,PLA University of Science and Technology, Nanjing, China, email: yongrenburao@163.com, yuhuaenator@gmail.com, ShenLiang671104@sina.com, chenglong_xu@yeah.net, chengyp2000@vip.sina.com.

Digital object identifier 10.1109/JCN.2014.000033

transmission paths between the source and the destination in a time-varying environment. Therefore, the channel quality cannot be neglected. However, this problem is difficult to solve mainly because of the following reasons. 1) As the channel is fading, the channel quality changes analogously over time. 2) The channel state is influenced by both the time-varying channel quality and the occupancy of PUs at any time. Moreover, they are independent of each other. The authors of [28] provided a solution to this problem, wherein the PU's activity is modeled as a two-state Markov chain, and the channel quality variation is modeled as a finite-state Markov chain. Inspired by the above references, we propose a two-dimension POMDP framework. Different from the works in [28], here, the term "two-dimensional" means that the channel quality is partially observed in the same manner as channel availability, and that the quality information obtained by the SU is updated when the belief vector is updated. The other difference and the detailed approach are as follows:

In this paper, a finite-state Markov channel (FSMC) [29], [30] model is introduced to characterize a fading channel in a time-varying environment for a data transmission system. Meanwhile, by probing the channel quality and exploring the PU's activity jointly, we propose a two-dimensional POMDP framework, i.e., the occupancy state of channels is partially observed by SUs, and the channel quality information is probed just when the sensing channel is idle. According to the sensing and probing outcomes, both the availability and the quality information of channels are updated. Compared with the one-dimensional framework, the two-dimensional framework obtains more information, including that on channel quality and availability, which is beneficial for predicting the information about channels more accurately. In addition, the greedy strategy is designed to maximize the instantaneous reward in the current slot, which also leads to low complexity and relatively ideal performance compared with the optimal strategy that considers the future reward. Further, the proposed strategy is compared with the previous strategy in [22] in the same cognitive network considering the channel quality. Different from the previous strategy in [22] that selects a channel with the largest probability of availability, the proposed strategy allows the SU to choose the channel with the best expected data transmission rate, which incorporates the channel availability and quality.

In addition, the two single user strategies, i.e., the proposed strategy and the previous strategy proposed in [22], are compared in different situations. First, the situation in the presence of the sensing error is studied roughly for discussing the collisions between PUs and SUs, where the spectrum sensing error interferes with the sensing and probing outcomes. Second, the situation of multiple SUs is also analyzed, where different SUs are located in different areas of the same network. The performance of the entire network will be different if the two single user strategies are applied to the network. On one hand, the availability states of the same channel sensed by them are the same in the same time slot. On the other hand, the quality states of the same channel probed by them may be different. Adopting the proposed single user strategy leads to more different selections and fewer collisions than adopting the previous strategy proposed in [22] because of different quality vectors and the same belief vectors. Therefore, if every SU adopts the proposed strategy, the performance of the network will be better than when every SU adopts the previous strategy proposed in [22].

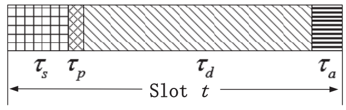


Fig. 1. The slot structure of SU.

According to the simulation results, the proposed strategy attains better performance than the one discussed in [22] in different situations. Moreover, some cases with different key parameters (e.g., number of channels and partitions of the FSMC) are compared, and the simulation results also verify our analysis.

The rest of this paper is organized as follows. In Section II the related work is mentioned. Section III describes the system model. Section IV presents the proposed two-dimensional POMDP framework. Then, in Section V, the greedy strategy is proposed considering the channel quality in a fading environment. The presented simulation results prove the advantage of the proposed strategy by observing the upgrade of the throughput in Section VI. Last the paper is concluded in Section VII.

II. SYSTEM MODEL

A. Transmission Model of SU

Suppose that in a time-varying environment there are N fading channels with the same bandwidth B Hz, which are licensed to PUs and form a cognitive wireless system. The cognitive users of the system communicate with a pair of synchronous transmitter and receiver. When the SUs explore the spectrum opportunities, they cannot sense all channels simultaneously because of the hardware and energy constraints. Thus, in this study, we assume that only one channel is selected at a time for the sake of simplicity.

To begin with, a SU selects a channel to sense with sensing time τ_s for its data transmission. If the channel is idle, the SU avails τ_p time to probe it and obtain the channel quality information θ. Otherwise, the SU waits for the next slot to proceed with the sensing again. If the SU decides to exploit the channel according to the sensing and probing outcomes, e.g., the channel is not occupied by PUs and its quality is desirable, it will cost time τ_d to transmit data packets based on the current data transmission rate, which is determined by the channel quality. At the end, after the data are successfully transmitted, the receiver acknowledges costing τ_a time and the summation of each data packet duration is $\tau_{da} = \tau_d + \tau_a$. The basic slot structure consists of these four time parameters, as shown in Fig. 1.

B. Traffic Model of PU

The traffic model of the PU's activities is described as the availability of all the N channels, which follows a discrete-time Markov process with $M = 2^N$ states [22]. In particular, in slot t, the availability state of channel $n \in \{1, 2, \cdots, N\}$ can be denoted by $S_n(t) \in \{b(\text{busy}), i(\text{idle})\}$ (Fig. 2), and the system

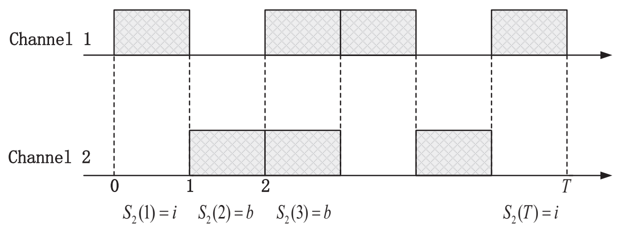


Fig. 2. An example of the spectrum occupancy.

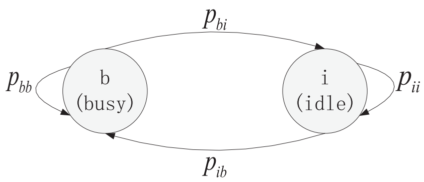


Fig. 3. The state transition of PU channel.

state at slot t is given by the vector:

$$S(t) = [S_1(t), S_2(t), \cdots, S_N(t)] \in \{b, i\}^N. \tag{1}$$

It is assumed that the traffic statistics of the PUs remain unchanged for T slots and the elements of $S(t)$ are independent of each other. The conditional transition probability of $S_n(t)$ is given as follows:

$$P_{kl}^n(\tau) = \Pr\{S_n(t+\tau) = l | S_n(t) = k\}, \forall k, l \in \{b, i\}. \tag{2}$$

The SU can achieve the transition probabilities from the estimation of the traffic statistics of PUs. In Fig. 3, p_{bi} and p_{ii} are probabilities that a busy channel becomes idle and an idle channel remains idle, respectively. This transition probability matrix can be written as $\mathbf{Q}_{\text{Occupancy}} = \begin{bmatrix} p_{bb} & p_{bi} \\ p_{ib} & p_{ii} \end{bmatrix}$, and this matrix is subject to the following stationary distribution: $[p_b, p_i] = [p_{ib}/(p_{ib}+p_{bi}), p_{bi}/(p_{ib}+p_{bi})]$.

C. FSMC Model

FSMC model (Fig. 4) is introduced to characterize a fading channel in a time-varying environment [29], [30]. Under normal circumstances, the thresholds are given as

$$\Gamma_k = e^{k\eta/B} - 1, k \in \{0, 1, \cdots, K-1\} \tag{3}$$

where $0 = \Gamma_0 < \Gamma_1 < \Gamma_2 < \cdots < \Gamma_{K-1} = \infty$ and η denotes the rate increment between adjacent channel quality states. The transmitter transmits a training sequence, and then, the receiver receives them and sends feedback [30]. According to the thresholds, the received instantaneous signal-to-noise ratio (SNR) γ is partitioned into K non-overlapping segments. If $\Gamma_k \le \gamma < \Gamma_{k+1}$, the quality state of the channel is regarded as k and the corresponding ideal data rate is $D_k = B\ln(1+\Gamma_k)$. The channel quality state is assumed to remain unchanged in each slot.

In a typical multipath propagation environment, the received SNR γ obeys the exponential distribution with the probability density function $p(\gamma) = \frac{1}{\gamma_0}e^{-\gamma/\gamma_0}$, where γ_0 denotes the average SNR [29]. The steady-state probability and the level crossing rate function are written as (4) and (5), respectively, by referring to [29].

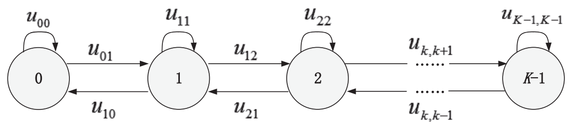


Fig. 4. The FSMC model.

$$\pi_k = \int_{\Gamma_k}^{\Gamma_{k+1}} p(\gamma)d\gamma = e^{-\Gamma_k/\gamma_0} - e^{-\Gamma_{k+1}/\gamma_0} \qquad k = 0, 1, \cdots, K-1 \tag{4}$$

$$\Lambda(\Gamma) = \sqrt{\frac{2\pi\Gamma}{\gamma_0}} f_d e^{-\Gamma/\gamma_0} \tag{5}$$

where f_d denotes the Doppler frequency and the level crossing rate function $\Lambda(\cdot)$ represents the average number of times per unit interval that a fading signal crosses a given signal level [29].

Here, the adjacent transfer (AT) method [29] is adopted, which assumes that transitions only happen between adjacent states, i.e., $P_{k,l} = 0, \text{ if } |k-l| > 1, \ (k,l) \in \{0, 1, \cdots K-1\}$.

Therefore, the state transition probability can be obtained as follows:

$$\begin{cases} u_{k,k+1} = \frac{\Lambda(\Gamma_{k+1})}{\pi_k}\tau_{da}, & k = 0, 1, \cdots, K-2; \\ u_{k,k-1} = \frac{\Lambda(\Gamma_k)}{\pi_k}\tau_{da}, & k = 1, 2, \cdots, K-1 \end{cases} \tag{6}$$

where τ_{da} denotes the packet duration time [30]. According to (6) the state transition probability matrix is obtained as follows:

$$\mathbf{Q}_{\text{Channel}} = \begin{bmatrix} u_{00} & u_{01} & 0 & \cdots & 0 \\ u_{10} & u_{11} & u_{12} & \cdots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \cdots & u_{K-3,K-2} & u_{K-2,K-2} & u_{K-2,K-1} \\ 0 & \cdots & 0 & u_{K-1,K-2} & u_{K-1,K-1} \end{bmatrix} \tag{7}$$

where $u_{k,k}$ is calculated as in (8) because the summation of every row of the matrix is 1.

$$u_{k,k} = \begin{cases} 1 - u_{01}, & k = 0; \\ 1 - u_{k,k+1} - u_{k,k-1}, & 1 < k < K-1; \\ 1 - u_{K-1,K-2}, & k = K-1. \end{cases} \tag{8}$$

III. TWO-DIMENSIONAL POMDP FRAMEWORK

Because of the hardware and energy constraints, the SUs generally cannot sense all channels simultaneously. The decision-theoretic solution of the POMDP is proposed to tackle the problem for OSA systems, where the SUs can only observe a part of the system state.

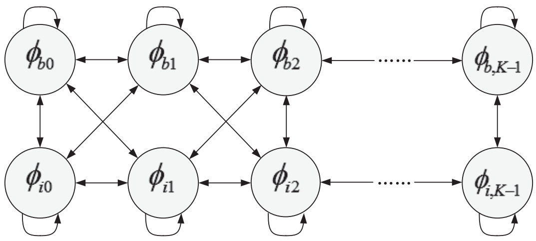


Fig. 5. The diagram of the integrated channel state transition.

Consequently, the realistic channel state is characterized by both channel quality and the occupancy of PUs; the original POMDP framework is inappropriate for this situation. Therefore, a new two-dimensional POMDP framework is proposed, where the channel quality state is probed and the PU's activity is explored jointly. By applying the FSMC model, we will discuss and define the core components of the new framework in this section.

A. State Space and Transition Probability

The system state consists of the states of all channels. After sensing and probing each channel, we can obtain the channel state information $\delta \in \{\varphi_b, \varphi_{i0}, \varphi_{i1}, \cdots, \varphi_{i,K-1}\}$. The channel state φ_b implies that the sensing channel is busy but the quality state cannot be probed, whereas $\varphi_{i,k}, k \in \{0, \cdots, K-1\}$ means that the channel is idle with quality state k.

Although the quality state is not observed when the channel is busy, it still exists and changes over time. Therefore, it is assumed that φ_b can be regarded as $\{\varphi_{b,0}, \varphi_{b,1}, \cdots, \varphi_{b,K-1}\}$, where $\varphi_{b,k}$ means that the channel is busy with quality state k; i.e., irrespective of the state of $\{\varphi_{b,0}, \varphi_{b,1}, \cdots, \varphi_{b,K-1}\}$ that the channel is in, the sensing outcomes are all φ_b. According to this assumption, the integrated channel state contains $\{\varphi_{b,0}, \varphi_{b,1}, \cdots, \varphi_{b,K-1}, \varphi_{i,0}, \varphi_{i,1}, \cdots, \varphi_{i,K-1}\}$ and the state transition proceeds among the $2K$ states as shown in Fig. 5. Therefore, the system state space is given as follows:

$$\begin{aligned} S(t) &= [S_1(t), S_2(t), \cdots, S_N(t)] \\ &\in \{\varphi_{b,0}, \varphi_{b,1}, \cdots, \varphi_{b,K-1}, \varphi_{i,0}, \varphi_{i,1}, \ldots \varphi_{i,K-1}\}^N. \end{aligned} \tag{9}$$

According to the steady-state probability of the channels, the distribution of $\varphi_{s,k}$ can be calculated as (10), where p_b and p_i are the probabilities that the channel is busy or idle, respectively.

$$\begin{aligned} &\Pr\{\delta = \varphi_{s,k}, s \in [b,i], k \in [0, 1, \cdots, K-1]\} = p_{s,k} \\ &= \begin{cases} p_b \pi_k, & 0 \le k \le K-1; \\ p_i \pi_k, & 0 \le k \le K-1. \end{cases} \end{aligned} \tag{10}$$

In Fig. 5, it is shown that if the channel state is φ_{b3}, it transfers to the states: $\varphi_{b,2}, \varphi_{b,3}, \varphi_{b,4}, \varphi_{i,2}, \varphi_{i,3}, \varphi_{i,4}$. In particular, when the quality state of the channel is 0 or K, there are only four states to transfer respectively: $\varphi_{b,0}, \varphi_{b,1}, \varphi_{i,0}, \varphi_{i,1}$ and $\varphi_{b,K-1}, \varphi_{b,K}, \varphi_{i,K-1}, \varphi_{i,K}$. The transition probabilities among the $2K$ states can be computed as

$$\begin{aligned} &\Pr\{\delta_n^{j+1} = \varphi_{s_2,l} | \delta_n^j = \varphi_{s_1,k}\} = q_{s_1 k, s_2 l}, \\ &(s_1, s_2) \in \{b, i\}, (k, l) \in \{0, 1, \cdots K-1\} \end{aligned} \tag{11}$$

and the specific explanations of these cases is given by the following four cases:

Case 1: If $|k-l| \ge 2$,

$$q_{s_1 k, s_2 l} = 0.$$

Case 2: If $k = l$,

$$q_{s_1 k, s_2 l} = \begin{cases} p_{bb} u_{k,k}, & s_1 = s_2 = b; \\ p_{bi} u_{k,k}, & s_1 = b, s_2 = i; \\ p_{ii} u_{k,k}, & s_1 = s_2 = i; \\ p_{ib} u_{k,k}, & s_1 = i, s_2 = b. \end{cases}$$

Case 3: If $l = k+1, k = 0, 1, \cdots K-2$,

$$q_{s_1 k, s_2 l} = \begin{cases} p_{bb} u_{k,k+1}, & s_1 = s_2 = b; \\ p_{bi} u_{k,k+1}, & s_1 = b, s_2 = i; \\ p_{ii} u_{k,k+1}, & s_1 = s_2 = i; \\ p_{ib} u_{k,k+1}, & s_1 = i, s_2 = b. \end{cases}$$

Case 4: If $l = k-1, k = 1, 2, \cdots K-1$,

$$q_{s_1 k, s_2 l} = \begin{cases} p_{bb} u_{k,k-1}, & s_1 = s_2 = b; \\ p_{bi} u_{k,k-1}, & s_1 = b, s_2 = i; \\ p_{ii} u_{k,k-1}, & s_1 = s_2 = i; \\ p_{ib} u_{k,k-1}, & s_1 = i, s_2 = b. \end{cases}$$

where $u_{k,k}$, $u_{k,k+1}$, and $u_{k,k-1}$ are given by (6) and (8), respectively.

According to (11), we can obtain the integrated transition probability matrix containing the occupancy of the PUs as follows:

$$\mathbf{Q} = \begin{bmatrix} p_{bb}\mathbf{Q}_{\text{Channel}} & p_{bi}\mathbf{Q}_{\text{Channel}} \\ p_{ib}\mathbf{Q}_{\text{Channel}} & p_{ii}\mathbf{Q}_{\text{Channel}} \end{bmatrix} \tag{12}$$

Then, by plugging (7) into (12), the matrix is calculated as (12). It is subject to the stationary distribution of $[p_b\pi_0, p_b\pi_1, \cdots, p_b\pi_{K-1}, p_i\pi_0, p_i\pi_1, \cdots, p_i\pi_{K-1}]$.

The matrix $\mathbf{Q}$ shown in (13) provides the transition probabilities that the channel state changes from one to another. Every row denotes the probabilities that the corresponding state of this row transfers to all states, while every column represents the probabilities that all states transfer to the corresponding state of this column. The existing situation of $\mathbf{Q}$ is explained as follows: The quality states of the FSMC and the available states are independent of one another, therefore, depending on whether the FSMC is idle the quality state of the FSMC can only transfer to the adjacent states or remain in the current state.

B. Action Space

At the beginning of every slot, the actions of the SU have two stages: Determining a channel ($a(t) \in \{1, 2, \cdots, N\}$) to sense and deciding whether to access it ($\Psi_a = \{0(\text{No access}), 1(\text{Access})\}$) according to the sensing outcome. The sensing space is given as $A = [a(1), a(2), \cdots, a(t), \cdots, a(T)]$ and the corresponding action space is $\{A, \Psi_a\}$.

In this paper, our emphasis is on the first stage, i.e., on determining a channel to sense according to the proposed strategy in order to obtain the maximum reward.

C. Reward

In each slot, after implementing the action $\{a(t) \in A, \Psi_a\}$ the SU receives the remaining reward denoted as $V_t(\Omega(t), \theta_a(t))$, where $\Omega(t)$ and $\theta_a(t)$ denote the belief vector and the quality vector, respectively. The reward consists

$$\mathbf{Q}=\begin{bmatrix} p_{bb}u_{00} & p_{bb}u_{01} & 0 & \cdots & 0 & p_{bi}u_{00} & p_{bi}u_{01} & 0 & \cdots & 0 \\ p_{bb}u_{10} & p_{bb}u_{11} & p_{bb}u_{12} & \cdots & 0 & p_{bi}u_{10} & p_{bi}u_{11} & p_{bi}u_{12} & \cdots & 0 \\ \vdots & \ddots & \ddots & & \vdots & \vdots & \ddots & \ddots & & \vdots \\ 0 & \cdots & p_{bb}u_{K-2,K-3} & p_{bb}u_{K-2,K-2} & p_{bb}u_{K-2,K-1} & 0 & & p_{bi}u_{K-2,K-3} & p_{bi}u_{K-2,K-2} & p_{bi}u_{K-2,K-1} \\ 0 & \cdots & 0 & p_{bb}u_{K-1,K-2} & p_{bb}u_{K-1,K-1} & 0 & & 0 & p_{bi}u_{K-1,K-2} & p_{bi}u_{K-1,K-1} \\ p_{ib}u_{00} & p_{ib}u_{01} & 0 & \cdots & 0 & p_{ii}u_{00} & p_{ii}u_{01} & 0 & \cdots & 0 \\ p_{ib}u_{10} & p_{ib}u_{11} & p_{ib}u_{12} & & 0 & p_{ii}u_{10} & p_{ii}u_{11} & p_{ii}u_{12} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \ddots & \ddots & & \vdots \\ 0 & \cdots & p_{ib}u_{K-2,K-3} & p_{ib}u_{K-2,K-2} & p_{ib}u_{K-2,K-1} & 0 & \cdots & p_{ii}u_{K-2,K-3} & p_{ii}u_{K-2,K-2} & p_{ii}u_{K-2,K-1} \\ 0 & \cdots & 0 & p_{ib}u_{K-1,K-2} & p_{ib}u_{K-1,K-1} & 0 & \cdots & 0 & p_{ii}u_{K-1,K-2} & p_{ii}u_{K-1,K-1} \end{bmatrix} \tag{13}$$

of two parts: (1) The immediate reward $R_a(t)$ achieved in slot t when channel a is sensed and the observing outcome is $\delta_a \in \{\varphi_b, \varphi_{i0}, \varphi_{i1}, \cdots, \varphi_{i,K-1}\}$; (2) the future expected reward $V_{t+1}(\Omega(t+1), \theta_a(t+1))$ that starts from slot $t+1$ with the updated belief vector $\Omega(t+1)$ and quality vector $\theta_a(t+1)$.

In this paper, a greedy strategy that just focuses on the immediate reward in order to reduce complexity is designed. Before the state transition in slot t, given the knowledge of the network state $\Omega(t)$ and the channel quality information $\theta(t)$, we can obtain the data rate $D_k = B\ln(1+\Gamma_k)$ when the channel quality state is k. If the SU selects channel a in slot t, it will obtain the expected reward as follows:

$$R_a(t) = \sum_{k=0}^{K-1} \omega_a(t)\beta_a^k D_k \tau_{da} = \sum_{k=0}^{K-1} \omega_a(t)\beta_a^k B\ln(1+\Gamma_k)\tau_{da} \tag{14}$$

where $\omega_a(t)$ and $\beta_a^k(t)$ are conditional probabilities that channel a is idle and its quality state is k, respectively, which are defined in the next part.

D. Belief Vector and Quality Vector

In a POMDP framework, the actual system state is generally unknown. According to all the past explorations and decisions, the estimation of the system state can be described as an important definition, which is called Belief Vector and written as follows:

$$\Omega(t) = [\omega_1(t), \omega_2(t), \cdots, \omega_n(t), \cdots, \omega_N(t)]^T, \tag{15}$$

where $\omega_n(t)$ denotes the conditional probability that the channel n is idle ($S_n(t) = i$). The initial belief vector $\Omega(1)$ is generally assumed to be the stationary distribution ω_0 shown in (16).

$$\omega_0 = p_i = \frac{p_{bi}}{p_{ib} + p_{bi}}. \tag{16}$$

However, the traditional belief vector cannot perfectly represent sufficient statistical information as the quality state is unknown when the channel is idle or busy. Therefore, another important concept called quality vector is defined to describe the channel quality information in this paper and is given by follows:

$$\begin{aligned} \theta_n(t) = [\beta_n^0(t), \beta_n^1(t), \cdots, \beta_n^k(t), \cdots, \beta_n^{K-1}(t)], \\ k \in \{0, 1, \cdots, K-1\}, n \in \{1, 2, \cdots, N\} \end{aligned} \tag{17}$$

where $\beta_n^k(t)$ denotes the probability that the channel n is in quality state k in slot t and $\sum_{k=0}^{K-1} \beta_n^k(t) = 1$. The vector is utilized to characterize the probability distribution of different quality states of channel n at the beginning of the slot and calculate the expected available capacity as (14) in the slot. Moreover, this vector is updated over time and is independent of the belief vector. Therefore, the system quality information can be forecast and is given by $\theta(t) = \{\theta_1(t), \theta_2(t), \cdots, \theta_n(t), \cdots, \theta_N(t)\}$. Based on the above mentioned two vectors, the SU selects the channel according to (14).

E. Policies

In this framework, the objective is essentially achieving an optimal policy. Define $\phi(t)$ as the policy which is the mapping from a belief vector $\Omega(t)$ and a quality vector $\theta(t)$ to an action $\{a(t) \in A, \Psi_a\}$. The policy $\phi(t)$ and the policy set of all slots are written as (18) and (19), respectively.

$$\phi(t)\colon \Omega(t), \theta(t) \to \{a(t) \in A, \Psi_a\}, \tag{18}$$

$$\Phi = [\phi(1), \phi(2), \cdots, \phi(t), \cdots, \phi(T)]. \tag{19}$$

IV. PROPOSED GREEDY STRATEGY

First, it is assumed that the greedy strategy is proposed without the spectrum sensing error. Second, the structure of the strategy is analyzed in the presence of the sensing error, which is beneficial to discuss the collision between a SU and a PU. Finally, the application of the proposed single user strategy is extended to the situation of multiple SUs, where the advantage of the proposed strategy compared with the previous strategy proposed in [22] is also clarified.

A. Proposed Strategy for Single User

In the proposed two-dimensional POMDP framework, the computational complexity of the optimal solution increases exponentially with the slots. A greedy strategy that maximizes the instantaneous expected reward is proposed; it has low complexity and relatively ideal performance compared with the optimal scheme. In this section, the proposed scheme is given, where one SU chooses a channel with the best expected data transmission rate supported by the channel quality.

Our objective is to choose the set of policies $\Phi = [\phi(1), \phi(2), \cdots, \phi(t), \cdots, \phi(T)]$ based on the belief vectors and the quality vectors to maximize the total expected reward during finite T slots as follows:

$$\Phi^* = \underset{\Phi=[\phi(1),\cdots,\phi(T)]}{\arg\max} \mathrm{E}_\Phi \left[\sum_{t=1}^{T} R_{\phi(t)}(t) \middle| \Omega(1), \theta(1) \right] \quad (20)$$

and the corresponding policy $\phi^*(t)$ is given by

$$\phi^*(t) = \arg\max_{\phi(t)} \mathrm{E}_{\phi(t)} \left[R_{\phi(t)}(t) \middle| \Omega(t), \theta(t) \right], \quad t = 1, 2, \cdots, T. \quad (21)$$

After performing the action, the expected best channel $a^*(t)$ is selected in slot t to maximize the expected immediate reward as follows:

$$\begin{aligned} a^*(t) &= \arg \max_{a=1,2,\cdots,N} R_a(t) \\ &= \arg \max_{a=1,2,\cdots,N} \sum_{k=0}^{K-1} \omega_a(t) \beta_a^k B \ln(1 + \Gamma_k) \tau_{da}. \end{aligned} \quad (22)$$

Given the cognitive scenario with fading channels, the previous strategy proposed in [22] still chooses the channel with the largest probability of availability according to the belief vector. Different from the previous strategy proposed in [22], the proposed strategy allows every SU to select a channel with the maximal expected reward according to the belief vector and the quality vector, which means choosing the channel with the maximal expected available data transmission capacity shown in (22). Because of the constant transmission time τ_{da}, the selection is equivalent to choose the channel with the best available expected data transmission rate, which is written as $\arg\max_{a=1,2,\cdots,N} \sum_{k=0}^{K-1} \omega_a(t) \beta_a^k B \ln(1 + \Gamma_k)$. Moreover, the proposed strategy obtains more information about the network containing the quality states and the availability states of channels over time, which makes the selection more efficient. If the data rates supported by the channel quality are used for calculating the actual throughput, the performance of the proposed strategy exceeds the previous strategy proposed in [22] in the same cognitive scenario that considers the channel quality.

At the end of slot t , the belief vector $\Omega(t)$ and the quality vector $\theta_n(t)$ are updated based on the action $a^*(t)$ and the observation δ_{a^*} (the state of channel a^*) as in (23) and (24), respectively.

$$\omega_n(t+1) = \begin{cases} p_{bi}, & if \;\; a^*(t) = n, \delta_{a^*}(t) = \varphi_b; \\ p_{ii}, & if \;\; a^*(t) = n, \; \delta_{a^*}(t) = \varphi_{i,k}; \\ \omega(t)p_{ii} + (1 - \omega(t))p_{bi}, & if \;\; a^*(t) \neq n; \end{cases} \quad (23)$$

$$\beta_n^l(t+1) = \begin{cases} u_{kl}, & if \;\; a^*(t) = n, \; \delta_{a^*}(t) = \varphi_{i,k}; \\ \sum\limits_{k=0}^{K-1} \beta_n^k(t) u_{kl}, & if \;\; a^*(t) = n, \; \delta_{a^*}(t) = \varphi_b \\ & or \;\; a^*(t) \neq n. \end{cases} \quad (24)$$

It is noted that in (23), there is no effect of the value of k when $\delta_{a^*}(t) = \varphi_{i,k}$. In particular, the proposed greedy strategy is described in algorithm 1.

Algorithm 1 Proposed OSA scheme

1: Initially, give the necessary parameters to calculate the transition probability matrix $\mathbf{Q}_{\text{Channel}}$ as (7) and the integrated transition probability matrix $\mathbf{Q}$ as (14).
2: Calculate the initial belief vector from the matrix $\mathbf{Q}$ and make the first choice according to (21) and (22) at the beginning of the first slot. (At the beginning, vectors are the same for every SU; therefore, choosing the largest is equal to choosing randomly.)
3: At the end of this slot, the belief vector and the quality vector are updated as (23) and (24) according to the sensing and probing outcome, respectively.
4: At the beginning of the next slot, make a new channel selection as (21) according to the new belief vector and quality vector obtained in step 3.
5: After the data are successfully transmitted, the receiver acknowledges the same. Repeat steps 3 and 4, and accumulate the reward according to (20).

B. Spectrum Access in the Presence of Sensing Error

In the actual scenario, spectrum-sensing errors can not be neglected, and can be mainly categorized into the following: false alarm and miss detection. The former means that the idle channel is sensed to be busy, which leads to a wastage of the spectrum resource. The latter is the opposite; i.e., the busy channel is explored to be idle, which causes the collision between a SU and a PU if the SU trusts the sensing outcome. In this study, the focus is on miss detection, which influences the process of probing quality states of the channels.

When miss detection occurs, the SU will probe the channel that is sensed to be idle, where the activities of PUs actually exist. As a result, the probing outcome, i.e., the quality state, is interfered. However, how the activities of PUs interfere with the probing outcome is beyond the scope of this paper. In addition, false alarm does not influence the probing outcome because of which the SU does not probe the channel.

At the end of the slot, the receiver will acknowledge it if the SU transmits data successfully. The acknowledgement is denoted by $A \in \{0(\text{Failing}), 1(\text{Successful})\}$. If $a^*(t) = n$, $\delta_{a^*}(t) = \varphi_{i,k}, A_{a^*} = 0$, it means that miss detection occurs and the probing outcome $\delta_{a^*}(t) = \varphi_{i,k}$ is wrong. The actual channel state should be $\delta_{a^*}(t) = \varphi_b$. In addition, the corresponding means of updating the belief vector and the quality vector should be changed owing to which the channel is busy. Miss detection and false alarm probabilities are denoted by p_m and p_f, respectively. According to the sensing outcome, the probing states and acknowledgement, belief vectors and quality vectors are updated as (25) and (26), as shown on the next page.

C. Application of the Greedy Strategy for Multiple SUs

In this subsection, the applications of two different single user strategies, i.e., the proposed strategy and the previous strategy proposed in [22], are compared in the same cognitive network with multiple SUs. The network performance will be different when two strategies are applied to the network. In general, the

$$\omega_n(t+1) = \begin{cases} \omega(t)p_{ii} + (1-\omega(t))p_{bi}, & if\ \ a^*(t) \neq n; \\ p_{ii}, & if\ \ a^*(t) = n,\ A_{a^*} = 1; \\ \frac{p_f(\omega_{a^*}p_{ii}+(1-\omega_{a^*})p_{bi})}{p_f(\omega_{a^*}p_{ii}+(1-\omega_{a^*})p_{bi})+(\omega_{a^*}(1-p_{ii})+(1-\omega_{a^*})(1-p_{bi}))}, & if\ \ a^*(t) = n,\ A_{a^*} = 0; \end{cases} \tag{25}$$

$$\beta_n^l(t+1) = \begin{cases} u_{kl}, & if\ \ a^*(t) = n,\ \delta_{a^*}(t) = \varphi_{i,k}, A_{a^*} = 1; \\ \sum\limits_{k=0}^{K-1} \beta_n^k(t) u_{kl}, & if\ \ a^*(t) = n,\ \delta_{a^*}(t) = \varphi_{i,k}, A_{a^*} = 0; \\ \sum\limits_{k=0}^{K-1} \beta_n^k(t) u_{kl}, & if\ \ a^*(t) = n,\ \delta_{a^*}(t) = \varphi_b\ or\ a^*(t) \neq n. \end{cases} \tag{26}$$

optimal channel is not unique; i.e., different SUs may select different "best" channels to sense. Therefore, the single-user strategies mentioned above could be suitable for the situation of multiple SUs. The proposed strategy allows every SU to choose an arbitrary channel among several channels with the best-expected data transmission rate, while the previous one in [22] lets every SU select a random channel among some channels with the largest probability of availability.

Considering the cognitive network with multiple SUs, the availability states of the same channel sensed by different SUs are the same in the same time slot, which is caused by the same observation on the PU's activities. However, the quality states of the same channel probed by different SUs may be different, because their different locations lead to different channel fading. The corresponding belief vectors are the same, but the quality vectors are different, which leads to the same probabilities of availability and different expected data transmission rates respectively. Therefore, the number of selections that SUs can make according to different data transmission rates is larger than they do according to the same probabilities of availability in [22]. Therefore, the proposed strategy brings less collisions and more network performance improvement than the previous one in [22].

In this study, we assume that the SU adopts carrier sense multiple access (CSMA) to avoid collisions with other SUs [22]. If two or more SUs want to access the same channel that is idle for transmission, one of them can succeed in transmitting data.

The reward of the system, which denotes the average throughput of the entire network during T slots, can be computed as

$$R = \sum_{m=1}^{M} \sum_{t=1}^{T} R_{a_m^*}^m(t)/T. \tag{27}$$

V. SIMULATION RESULTS

In this section, a typical scene is considered in which the channel is slowly fading and the environment is time-varying. The performance is evaluated by MATLAB simulations, whose parameters are as follows: Channel bandwidth $B = 6$ MHz, average received SNR $\gamma_0 = 15$ dB, carrier frequency $f_c = 50$ MHz, terminal mobile speed $v = 2$ m/s (Doppler frequency is vf_c/c), data transmission time and acknowledgment time $\tau_{da} = 100$ ms, and the rate interval $\eta = 3$ Mbps. The system model is established with five independent

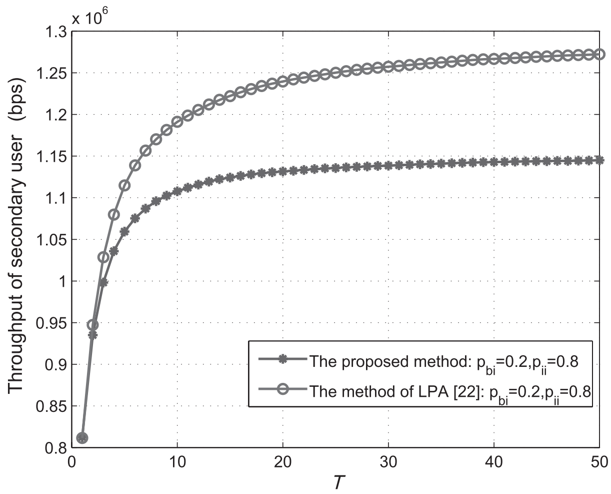


Fig. 6. Performance comparison of our approach and the previous approach of choosing channel with the largest probability of availability (LPA) in [22]. The transition probability is case 1: $\{p_{bi} = 0.2, p_{ii} = 0.8\}$

eight-states Markov channels, which describe the fading characteristic of the time-varying environment.

First, three cases of different transition probabilities are given: Case 1: $\{p_{bi} = 0.2, p_{ii} = 0.8\}$, case 2: $\{p_{bi} = 0.5, p_{ii} = 0.5\}$, and case 3: $\{p_{bi} = 0.8, p_{ii} = 0.2\}$, which have the same stationary distribution. The selection of these three cases makes it convenient to compare the performance of the proposed approach with that of the previous approach proposed in [22], which just chooses the channel with the largest probability of availability (LPA). In Figs. 6, 7, and 8, the simulation results demonstrate that the approach proposed in this paper prevails over the preceding approaches. In particular, in Fig. 7, the upgrade is approximately up to 20% because of the improved information gained from the observed PU's activities and the probed channel quality states. It is worth mentioning that the curve of the bottom in Fig. 7 also has an upward trend, which is only caused by the exploration results of the quality state. If the channel quality is not considered in case 2, the curve will almost become a straight line because of the same transition probabilities, which means that the accumulating observations provide little information and the channels are selected randomly.

Second, we discuss the influence caused by the change in N, which denotes the number of channels. Assume that there are

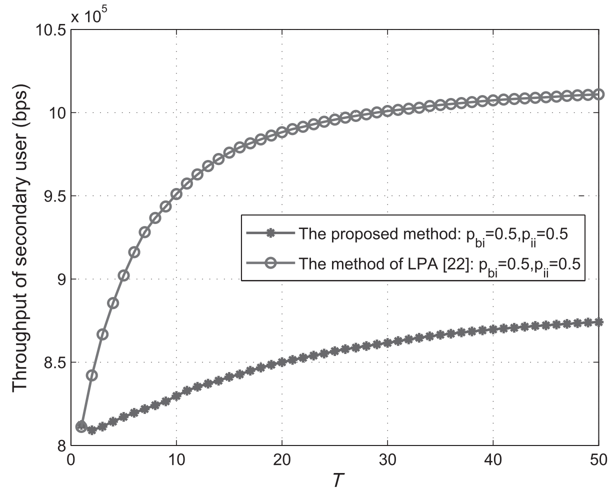


Fig. 7. Performance comparison of our approach and the previous approach of choosing channel with the largest probability of availability (LPA) in [22]. The transition probability is case2: $\{p_{bi} = 0.5, p_{ii} = 0.5\}$

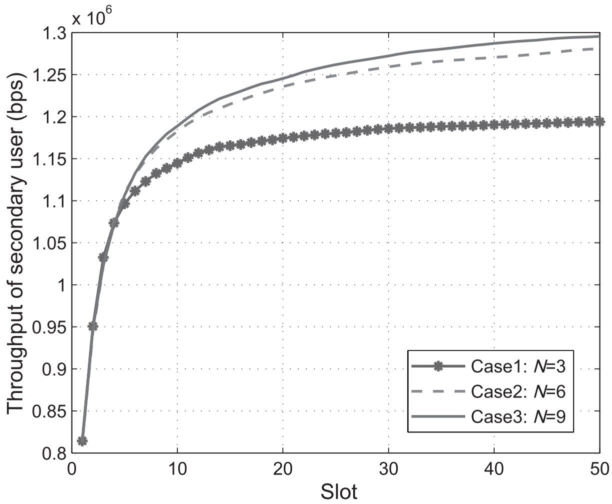


Fig. 9. The performance comparison based on three cases with different numbers of channels each of which has 8 states. There are $N = 3$, $N = 6$ and $N = 9$ independent channels in case 1, 2, and 3 respectively.

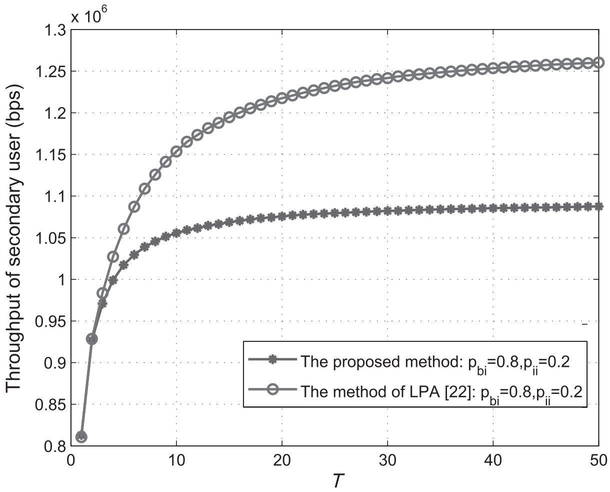


Fig. 8. Performance comparison of our approach and the previous approach of choosing channel with the largest probability of availability (LPA) in [22]. The transition probability is case 3: $\{p_{bi} = 0.8, p_{ii} = 0.2\}$

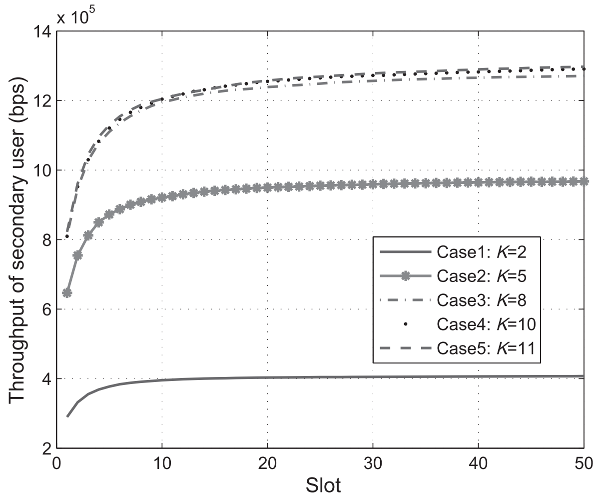


Fig. 10. The performance comparison based on the number of channels is constant with $N = 5$ and five cases with different number of states $K = 2$, $K = 5$, $K = 8$, $K = 10$, and $K = 11$, respectively.

$N = 3$, $N = 6$, and $N = 9$ independent FSMCs each of which has $K = 8$ states in cases 1, 2, and 3, respectively. The comparison among these three cases is shown in Fig. 9. The proposed approach can be more effective with a larger number of channels. Because the larger the value of N is, the greater is the number of channel choices made by the SU. However, when the number of channels N is sufficiently large, the upgrade caused by the increase in N can be almost neglected. There are two reasons for this: 1) The SU just makes one channel selection at a time. 2) As in the case of multichannel diversity, the performance increment decreases as the value of N increases.

Third, the number of channels is set to be constant, i.e., $N = 5$. We attempt to analyze five cases with different values for K, $K = 2$, $K = 5$, $K = 8$, $K = 10$, and $K = 11$. As is apparent from Fig. 10, the larger the value of K is, the more effective is the proposed approach. For a larger K, the probability of transmission through a better channel state and a higher data rate increases for the same received SNR. As shown in Fig. 9, if K is sufficiently large, the increase in performance is little. Because when K is sufficiently large, the same received SNR may be portioned into the same state or adjacent states, which leads to almost the same data transmission rate.

Fourth, the performance and the spectrum efficiency in the presence of a sensing error are studied. For comparing conveniently and convincingly with the previous strategy proposed in [22], the same spectrum sensor as that used in [22] is applied to the simulation in this study. Therefore, the function relationship between p_m and p_f is given as in [22]. The transition probabilities are also set to be $\{p_{bi} = 0.4, p_{ii} = 0.5\}$. The comparison results are illustrated in Fig. 11. As the collision probability be-

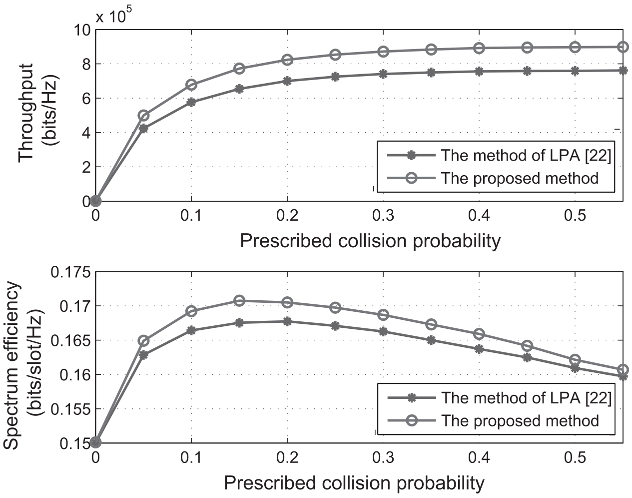


Fig. 11. The performance comparison in presence of sensing error. $\{p_{bi} = 0.4, p_{ii} = 0.5\}$

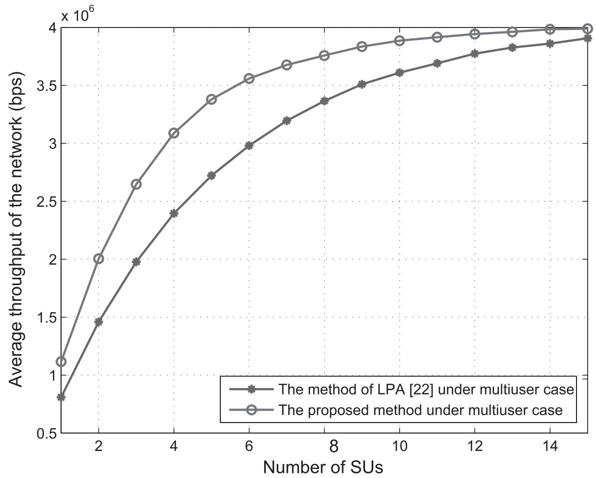


Fig. 12. The performance comparison under the case of multiuser adopting the proposed single user method and the previous LPA method in [22] respectively.

tween a PU and a SU increases, the performance of the proposed strategy is better than that of the previous strategy proposed in [22]. The proposed strategy can obtain more information, and thus, the SU can utilize the channels more efficiently.

Finally, set the parameters as follows: $N = 5$, $K = 8$, and $\{p_{bi} = 0.2, p_{ii} = 0.8\}$. As shown in Fig. 12, if the two strategies are applied to the network with multiple SUs, the performance of the proposed strategy exceeds that of the previous strategy proposed in [22]. On the one hand, the availability states of the same channel sensed by different SUs are the same in the same time slot. On the other hand, the quality states probed by them are different. While these two cases lead to the same belief vectors, they lead to different quality vectors. Consequently, the selections may be the same for the SUs adopting the previous strategy proposed in [22], while the selections may be different when adopting the proposed strategy. Therefore, the proposed strategy leads to more different selections and better performance than the previous strategy proposed in [22]. In addition, the two curves have the similar increasing trend with an increase in the SUs. Because when two or more SUs choose the same optimal channel, there is always one SU that can transmit data successfully adopting CSMA. In addition, the larger the number of SUs is, the more utilized are the channels and the better is the performance. If the number of SUs is sufficiently large, e.g., $N > 15$, the curves remain unchanged because of the limited spectrum resource.

VI. CONCLUSIONS

In this paper, a FSMC model was introduced to represent a fading channel in a time-varying environment. Because the FSMC was also affected by the PU's activity, we proposed a two-dimensional POMDP framework for OSA systems. Furthermore, a greedy strategy in which the SU selected one channel with the best-expected data transmission rate in order to maximize the instantaneous reward in every slot was designed. Further, the greedy strategy had low complexity and relatively ideal performance compared with the optimal strategy. The scenario in the presence of a sensing error and the application of the proposed single-user strategy in a multiuser situation were also analyzed. The performance of the proposed strategy was evaluated by simulations, which demonstrated that our strategy was better than the previous works. In addition, some cases with different parameters were compared and the simulation results verified our analysis.

For OSA systems, POMDP is only a type of decision-theoretic frameworks that contain some other theoretic frameworks, e.g., game theory, optimal stopping problem (OSP), and multi-armed bandit (MAB) problem [10]. This two-dimensional framework, which solves a situation by considering the channel quality, may be appropriate for other theoretic frameworks. This will be the focus of our next work.

REFERENCES

[1] J. Mitola and G. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[2] A. Sahai, N. Hoven, and R. Tandra, "Some fundamental limits on cognitive radio," in *Proc. Allerton Conf. on Commun., Control, and Comput.*, Oct. 2004.

[3] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, 2005.

[4] I. F. Akyildiz, W. Lee, M. Vuran, *et al.*, "Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Comput. Netw.*, vol. 50, no. 13, pp. 2127–2159, 2006.

[5] I. F. Akyildiz *et al.*, "A survey on spectrum management in cognitive radio networks," *IEEE Commun. Mag.*, vol. 46, no. 4, pp. 40–48, 2008.

[6] Q. Zhao and B. Sadler, "Dynamic spectrum access: signal processing, networking, and regulatory policy," *in IEEE Trans. Signal Process. Mag.*, May 2007.

[7] K. Shin, H. Kim, A. Min, and A. Kumar, "Cognitive radios for dynamic spectrum access: From concept to reality," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 64–74, 2010.

[8] Q. Zhao and A. Swami, "A Decision-theoretic framework for opportunistic spectrum access," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp.14–20, 2007.

[9] A. De Domenico, E. Strinati, and M. Di Benedetto, "A survey on MAC strategies for cognitive radio networks," *IEEE Commun. Surveys & Tutorials*, vol. 14, no. 1, pp. 21–44, 2012.

[10] Y. H. Xu *et al.*, "Decision-theoretic distributed channel selection for opportunistic spectrum access: Strategies, challenges and solutions," *IEEE commun. Survey & Tutorial*, vol. 15, no. 4, pp. 1689–1713, 2013.

[11] Y. H. Xu *et al.*, "Opportunistic spectrum access in cognitive radio networks: Global optimization using local interaction games," *IEEE J. Sel. Topics in Signal Process.*, vol. 6, no. 2, pp. 180–194, 2012.

[12] Y. H. Xu *et al.*, "Opportunistic spectrum access in unknown dynamic environment: A game-theoretic stochastic learning solution," *IEEE Trans. Wireless Commun.*, vol. 11, no. 4, pp. 1380–1391, 2012.

[13] C. Stevenson *et al.*, "IEEE 802.22: The first cognitive radio wireless regional area network standard," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 130–138, Jan. 2009.

[14] A. B. Flores *et al.*, "IEEE 802.11af: A standard for TV white space spectrum sharing," *IEEE Commun. Mag.*, vol. 51, no. 10, pp. 92–100, Oct. 2013.

[15] A. W. Min, K. G. Shin, and X. Hu, "Secure cooperative sensing in IEEE 802.22 WRANs using shadow fading correlation," *IEEE Trans. Mobile Comput.*, vol. 10, no. 10, pp. 1434–1447, Oct. 2011.

[16] Q. H. Wu, G. R. Ding, J. L. Wang, and Y. D. Yao, "Spatial-temporal opportunity detection in spectrum-heterogeneous cognitive radio networks: Two-dimensional sensing," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 516–526, Feb. 2013.

[17] W. Lee and D. H. Cho, "Comparison of channel information acquisition schemes in cognitive radio system," in *Proc. IEEE WCNC*, 2012.

[18] R. Murty, R. Chandra, T. Moscibroda, and P. Bahl, "Senseless: A databased-driven white spaces network," *IEEE Trans. Mobile Comput.*, vol. 11, no. 2, pp. 189–203, Feb. 2012.

[19] H. R. Karimi, "Geolocation databases for white space devices in the UHF TV bands: Specification of maximum permitted emission levels," in *Proc. IEEE DySPAN*, 2011.

[20] M. Barrie, S. Delaere, and G. Sukareviciene, "Geolocation Database beyond TV white spaces? Matching applications with Database requirements," in *Proc. IEEE DySPAN*, 2012, pp. 443–454.

[21] Q. Zhao and Sadler, "A survey of dynamic spectrum access," *IEEE Signal Process. Mag.*, vol. 2. no. 3, pp. 79–89, May 2007.

[22] Q. Zhao, L. Tong, and A. Swami, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE Sel. Areas Commun.*, vol. 25, no. 3, pp.589–600, Apr. 2007.

[23] Y. X. Chen, Q. Zhao, and A. Swami, "Bursty traffic in energy-constrained opportunistic spectrum access," in *Proc. IEEE GLOBECOM*, 2007, pp. 4641–4646.

[24] Y. X. Chen, Q. Zhao, and A. Swami, "Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors," *IEEE Trans. Inf. Theory*, vol. 54, pp. 2053–2071, May 2008.

[25] A. Honang, Y. H. Liang, and D. Wang, "Opportunistic spectrum access for energy-constrained cognitive radio," in *Proc. IEEE VTC*, 2008.

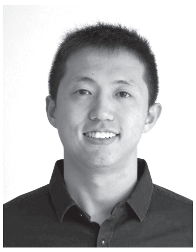
[26] Q. Zhao, B. Krishnamachari, and K. Q. Liu, "On myopic sensing for multichannel opportunistic access: Structure, optimality, and performance," *IEEE Trans. Wireless Commun.*, vol. 7, pp. 5431–5440, Dec. 2008.

[27] S. Ahmad, M. Y. Liu, and T. Javidi, "Optimality of myopic sensing in multichannel opportunistic access," *IEEE Trans. Inf. Theory*, vol. 55, pp. 4040–4050, Sept. 2009.

[28] M.M. Rashid, J. Hossain, E. Hossain, and V.K. Bhargava, "Opportunistic spectrum access in cognitive radio networks: A queueing analytic model and admission controller design," in *Proc. IEEE GLOBECOM*, 2007.

[29] Q. Q. Zhang and S. Kassam, "Finite-state Markov model for Rayleigh fading channels," *IEEE Trans. Commun.*, vol. 47, pp. 1688–1692, 1999.

[30] B. W. Li, P. L. Yang, and J. L. Wang, "Optimal frequency-temporal opportunity exploitation for multichannel ad hoc networks," *IEEE Trans. Parallel and Distrib. Syst.*, vol. 23, no. 12, pp. 2289–2302, Dec. 2012.



Yumeng Wang has received his B.E. degree in Communication Engineering from Hebei University of Science and Technology, Tianjin, China in 2011. He has been a master that studies Systems of Communications and Information in Wireless Communication Laboratory of PLA University of Science and Technology since 2011. His major research interests include wireless communication theory, cognitive radio, and signal processing.

Yuhua Xu has received his B.S. degree in Communications engineering, M.S. degree in Communications and Information System from Institute of Communications Engineering, Nanjing, China in 2006 and 2008, respectively. He is currently pursuing the Ph.D. degree in Communications and Information System in Institute of Communications Engineering, PLA University of Science and Technology. His research interests focus on opportunistic spectrum access, learning theory, game theory, and optimization techniques. He was an Exemplary Reviewer for the IEEE Communications Letters in 2011 and 2012.

Liang Shen has received his B.S. degree in Communications Engineering and M.S. degree in Communications and Information System from the Institute of Communications Engineering, Nanjing, China in 1988 and 1991 respectively. He is currently a Professor at the PLA University of Science and Technology, China. Recent years, he undertook and accomplished several Key Scientific Projects of the National Basic Research Programs of China and China 863 Communication High Technology Programs. The group headed by him has received the second prize of National Science and Technology Progress Prize twice. Due to his outstanding contribution, he was awarded as General Staff Excellent Young and Middle-aged Experts in 2001 and achieved special government allowance in 2004. His current research mainly focuses on HF communication, MODEM, signal processing, and information theory.

Chenglong Xu has received his B.E. degree in Communication Engineering from PLA University of Science and Technology in 2011 and has been a master that studies Military Equipment in Wireless Communication Laboratory in this university since 2011. His major research interests include wireless communication theory, cognitive radio, and military equipment.

Yunpeng Cheng has received B.E., M.S., and Ph.D. degrees in Communications and Electronic Systems in Institute of Communications Engineering, PLA University of Science and Technology, Nanjing, China in 1997, 2000, and 2003 respectively. He is currently an Associate Professor of Wireless Communication Laboratory in this school. His current research interests include signal processing, HF communication, spectrum aggregation, etc.