SPA: Almost Optimal Accessing of Nonstochastic Channels in Cognitive Radio Networks

Xiang-Yang Li*‡§, Panlong Yang†‡, Yubo Yan†
* Department of Computer Science, Illinois Institute of Technology, Chicago, USA
† PLA University of Science and Technology, Nanjing, China
‡ MOE Key Lab, School of Software, TNLIST, Tsinghua University, Beijing, China
§ College of Computers and Software, Nanjing University of Information Science and Technology, China

Abstract—In this work, we address the spectrum utilization problem in cognitive radio (CR) networks, in which a CR can only utilize spectrum opportunities when the channel is idle. One challenge for a CR is to balance exploring new channels and exploiting existing channel, due to the fact that the channel availability and channel quality, potentially heterogeneous and time-dependent, are often unknown in advance due to the large number of channels, and the limited hardware capability of single CR. In this work, we propose joint channel sensing, probing, and accessing schemes for secondary users in cognitive radio networks. Our method has time and space complexity $O(N \cdot u)$ for a network with N channels and u secondary users, while applying classic methods requires exponential time complexity. We prove that, even when channel states are selected by adversary (thus nonstochastic), it results in a total *regret* uniformly upper bounded by $\Theta(\sqrt{TN \log N})$, *w.h.p*, for communication lasts for T timeslots. Our protocol can be implemented in a distributed manner due to the *nonstochastic* channel assumption. Our experiments show that our schemes achieve almost optimal throughput compared with an optimal static strategy, and perform significantly better than previous methods in many settings.

Index Terms—Cognitive networks, online, throughput-efficient, competitive ratio, regret, optimization.

1 Introduction

In opportunistic spectrum usage, primary users (PU) will open under utilized portions of the licensed spectrum for secondary re-use. Secondary users observe the channel availability dynamically and explore/exploit it opportunistically. Secondary users (SU) are cognitive radio (CR) devices that can sense the environment and adapt to appropriate frequency, power, and transmission schemes. A conventional approach for a CR to select channels is to scan the channels to find an idle one and then transmit data over an idle channel. The scanning sequence can be decided by historical observations. However, the qualities (*e.g.*, the data rate supported) of different channels often are time dependent, and may differ significantly across all channels at the same time. Thus, it is possible that the first found idle channel may not have better quality, which in turn could reduce the overall throughput.

We investigate a joint sensing/probing/accessing mechanism that will result in (almost) optimal[1] expected throughput by assuming independent channels, which is valid [18]. In our approach, a CR not only senses the busy/idle status of the channel, but also probes the instaneous *quality* of the channel using short predefined probing packets. Based on this additional channel quality information, and possibly some historical observations, a CR will decide whether to transmit data over current channel (called *exploitation*) or to continue to sense/probe some other channels (called *exploration*). When there are multiple secondary users within vicinity, there is a potential competition among these users to access the channel with the best quality. Our goal in this work is to design efficient (using polynomial space and time) and effective (with small *regret*, *i.e.*, the gap between the overall throughput by a fixed best strategy, and the throughput achieved by our scheme) channel sensing/probing/accessing scheme for multiple users.

Designing better schemes for channel usage has been extensively studied in the literature. Some of these results, *e.g.*, [9], [10], [20] model the channel using two states, *busy*, or *idle*, thus, ignoring the different channel conditions. Then a number of schemes were proposed to exploit the dynamic channel qualities, *e.g.*, [18] assumed a homogeneous distribution of the channel quality. Recently, several results, *e.g.*, [1], [8], [15], were proposed by exploring the parallels between the cognitive medium access and the multi-armed bandit (MAB) problem. These methods can theoretically guarantee almost the optimum *regret* under various assumptions. However, these protocols require either exponential space, or exponential time, or both when multiple users present.

We propose several joint sensing/probing/accessing schemes which maximize the expected throughput. We first consider the system with heterogeneous and stochastic channel availability and qualities. For a network of single SU, we derive a throughput optimal strategy based on optimal stopping rule. Optimal centralized method is then proposed for a network with multiple users. We then consider networks with *nonstochastic*

1. The almost optimal means very close to the optimal results asymptotically, which will be formally addressed in Section 3.

channel qualities. We first propose a centralized scheme, called ϵ-SPA, which runs in polynomial space and polynomial time $O(N \cdot u)$ for a network of N channels and u SUs. Our method is inspired by results in [4] and will adaptively select channels for communication. We theoretically prove that the regret per round is $O(\sqrt{\frac{N \ln N}{T}})$, when the system lifetime is T. Since our protocol assumes nonstochastic channels, each SU can run our ϵ-SPA scheme individually in a distributed manner, which will also maximize the throughput in each user's view. We also discuss how to achieve better throughput by setting parameters such as sensing, probing and accessing time.

We conduct experiments using USRP and sensor networks to study the performance of our schemes. We also conduct simulations to study the performances under a large scale network. Our experimental and simulation results show that the total throughput achieved by our SPA scheme is indeed close to the optimum. Furthermore, our SPA scheme achieves the best throughput in majority cases, except when all channels' qualities are very poor. Also, we consider the settings with mixed channels, *i.e.*, the channel quality is divers, where channels are set with good, medium and bad quality. When there are less than 10 users, our schemes perform an order of magnitude better than a recently implemented method [18]. With bad channels only, we suggest to use a strategy mixed by SPA scheme and optimal stopping rule, which achieves the best performance in our study.

Paper organization: The remainder of the paper is organized as follows. In Section 2, we present the system model used in this work and formally define the questions studied. We then present our optimal sensing/probing/accessing (SPA) strategy in Section 3. We report our extensive evaluation results in Section 4, discuss the related work in Section 5, and conclude the paper in Section 6.

2 Network and System Models

2.1 Network and System Model

Consider a set $\mathcal{C} = \{\mathbf{c}_1, \mathbf{c}_2, \cdots, \mathbf{c}_N\}$ of N channels. These channels will be used by some primary user(s) sporadically. When a channel $\mathbf{c}_i$ is used by a primary user (or a secondary user if multiple secondary users exist), it is termed as *busy*; otherwise, it is termed as *idle*. Each channel $\mathbf{c}_i$ is modeled as a continuous time random process that alternates between idle and busy. Observe that the transmission of a PU is not slotted. We assume that when a channel $\mathbf{c}_i$ is observed randomly, the probability that it is busy is $p_B(i)$ and the probability that it is idle is $p_I(i)$. Note that the channel status sensing is not perfect: it may have false alarms and miss detection. We use P_{fB} (P_{fI} respectively) to denote that probability that the CR will decide the channel is busy (idle, respectively), while actually the channel is idle (busy, respectively). These two probabilities P_{fB}, P_{fI} are monotonically decreasing functions of the sensing time t_s. In practice, P_{fI} typically is about 1%. Since we can only measure the busy and idle probabilities of channels using sensing techniques, the actual busy and idle probabilities p_B, p_I should be computed from the measured busy and idle probabilities (denoted as $\widetilde{p_B}$, $\widetilde{p_I}$ respectively). Clearly, $\widetilde{p_I} = p_B \cdot P_{fI} + p_I \cdot (1 - P_{fB})$, and $\widetilde{p_B} = p_B \cdot (1 - P_{fI}) + p_I \cdot P_{fB}$. Thus, we can compute the probabilities p_I and $p_B = 1 - p_I$ as $p_I = \frac{\widetilde{p_I}(1-P_{fI}) - \widetilde{p_B} \cdot P_{fI}}{1-P_{fI}-P_{fB}}$.

The set of N channels can be used opportunistically by a set $\mathcal{S} = \{\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_u\}$ of u secondary users (SU) equipped with CR, along with the primary users (PU), where u is the number of secondary users. In this work, an SU is composed of a pair of transmitter node and receiver node. Before accessing a channel, each pair of transmitter/receiver needs to sense the channel to determine whether it is an idle channel, and if channel is idle then probe the channel to determine the estimated data rate, $R_{i,t}$, that can be achieved from next timeslot. In this work, the data rate $R_{i,t}$ could be arbitrary, while previous work [19] [14], [15] assumed it is drawn from a given (maybe unknown) distribution. Furthermore, the different users may have different observation on the data rate of the same channel at the same time instance. We generally assume that the data rate distribution R_i is *heterogeneous*, *time variant*, and *unknown* in advance. When there are multiple pairs of transmitter/receivers, we use $R_{j,i,t}$ to denote the data rate of channel $\mathbf{c}_i$ at time slot t observed by the j-th pair of transmitter/receiver $\mathbf{s}_j$.

We assume that each radio on CR needs time t_s for sensing the status of a channel and time t_p for probing the quality of a channel. The total time of sensing and probing a channel is $t_{sp} = t_s + t_p$. The actual time t_{sp} depends on the technology and device. Practical measurement [2] shows that t_s is about 10ms and t_p is from 10ms to 133ms. Here we assume that t_{sp} is homogeneous for all pairs of transmitter/receivers. For a pair of transmitter/receiver, we assume that the clock is synchronized among the transmitter and the receiver, and a common random number generator is shared by each pair of transmitter/receiver to determine the sequence of channels to be sensed/probed/accessed. Notice that even the channel is not idle, the sensing and probing time is still t_{sp} since the receiver needs to wait for a probing time to infer that the channel is busy. When a channel is idle, transmitter/receiver can only access it for at most t_a time continuously so it can detect the existence of a primary user. In practice, the value t_a is typically $2s$.

2.2 Problem Formulation

Since the channel quality is a time variable, and sensing/probing will take time, the overall average throughput over a time period by a pair of transmitter/receiver clearly depends on their sensing/probing/accessing strategies (when and which channel). It is more challenging if we also have to consider the contentions between different pairs of transmitter/receivers. Our approaches are built upon some classical results from multi-armed bandit problem (MAB) [6], in addition to the optimal stopping rule [7].

We first focus on the case when there is only one pair of transmitter/receiver equipped with only one radio. One major difference between MAB and the sequential channel sensing/probing/accessing problem is that, for MAB problem, at every round t, we will receive a reward and we repeat this for T rounds; while for SPA problem, at a round t, we will *not* have any gain if we only sense/probe a channel: the actual gain only happens when a CR transmits data using some channel. To address this challenge, we call the time spent for sensing/probing a chosen channel as a *round*. The time spent for transmitting data over a chosen channel is *not* counted as a round. Instead, we will treat the total data transmitted as the gain of the prior sensing/probing round. The strategy for sensing/probing/accessing is then composed of many sequential rounds of sensing/probing.

For every round t and every channel $\mathbf{c}_i$, let $X_{i,t} \in [0, M]$ be the random variable denoting the data rate of $\mathbf{c}_i$ at round t, where constant M is the known maximum data rate of any link. We use $X_{i,t} = 0$ to denote that the channel $\mathbf{c}_i$ is not idle at round t. We assume that the transmitter/receiver always has data to communicate. A sensing/probing/accessing strategy χ by the pair of CRs will decide its action. At every round t, let $\theta(t)$ be the channel index selected by the strategy χ for sensing/probing, and $R_{\theta(t),t}$ be the actual data rate of the channel probed. We use $\sigma(t)$ to denote whether it decides to transmit data using the probed channel $\mathbf{c}_{\theta(t)}$ after this round immediately: $\sigma(t) = 1$ if they will transmit using channel $\mathbf{c}_{\theta(t)}$ and $\sigma(t) = 0$ otherwise. Observe that, although the channel quality often has a coherence time t_c, to avoid the possible collisions with PUs and other CRs, we first assume that a CR will *not* recall (reuse) a previously probed channel. We will then extend our mechanisms to the case when recall is allowed and prove that recall does not improve the asymptotical throughput for all models.

For a sensing/probing strategy $\chi(t) = \langle\theta(t), \sigma(t)\rangle$ by a pair of transmitter/receiver, at round t, its gain is

$$g_\chi(t) = \sigma(t) \cdot R_{\theta(t),t} \cdot Z_t \cdot (1 - P_{loss}(t)), \tag{1}$$

where Z_t is a variable denoting the duration of the actual transmission and P_{loss} denotes the fraction of the transmissions that will be destroyed by the return of primary users or some other CR during the Z_t time-frame. In this work, for simplicity, we assume that $Z_t = t_a$. We assume that the channel busy and idle states are memoryless. We can compute the probability that the PU comes back within timeframe Z_t as $P_{loss}(t)$. For example, when the idle time $I_{\theta(t)}$ for a channel $\mathbf{c}_{\theta(t)}$ is an exponential distribution with mean $\alpha_{\theta(t)}$, we have $P_{loss}(t) = 1 - e^{-\frac{Z_t}{\alpha_{\theta(t)}}}$.

Let n be the number of sensing/probing rounds executed during the system's lifetime T, which should satisfy

$$\text{Condition: } \; n \cdot t_{sp} + \sum_{t=1}^{n} \sigma(t) \cdot Z_t \leq T. \tag{2}$$

The first part is time spent for sensing/probing and the second part is the time for data transmissions. Then our objective is to design a strategy that maximizes the expected throughput over the system lifetime T, *i.e.*, $\max_\chi \frac{\sum_{t=1}^{n} g_\chi(t)}{T}$. A channel with the highest expected throughput is called the *best channel* (or *optimal static strategy*). The regret ϱ after T rounds of an online strategy χ is defined as the difference between the reward sum associated with an optimal static strategy and the sum of the data rates in T rounds by χ. A strategy whose average regret per round $\varrho/T \to 0$ with probability 1 when $T \to \infty$ is a *zero-regret strategy*. Intuitively, zero-regret strategies are guaranteed to converge to an optimal static strategy, not necessarily unique, if enough rounds are played.

3 Almost Optimal SPA for Single User

We first propose two methods when centralized decision is possible. The first method assumes channels' statistics follow some unknown distributions and the number of possible rates is limited. The strategy is *learning then transmitting*. A CR will first collect the channel statistics by sensing/probing all channels for some rounds, then transmit data based on the collected information. The second method works for a more general case when the channel states are nonstochastic. The strategy is sort of *learning when transmitting*. The main idea is as follows. We guess an optimal strategy at the beginning. With certain probability, we execute the strategy we guessed; otherwise, we try some new strategies. Based on the feedback, *i.e.*, the throughput, we adjust our guess dynamically. For both methods, we first consider the simple case where the network has only one SU. Then we extend our methods to cases with u users. We also discuss how to optimize the parameters used in our methods.

3.1 OSP SPA for Stochastic Channels

Assume that each channel $\mathbf{c}_i$ has rates drawn from a set $\mathcal{R} = \{R_1, R_2, \cdots, R_r\}$ with $R_j < R_{j+1}$ for $j \in [1, r-1]$, following some distributions with mean μ_i. After learning rate distributions, we can then use the optimal stopping rule (OSP) [7] for the single channel. Intuitively, the *static* optimal strategy should always sense/probe/access a fixed channel that will maximize the expected throughput $p_I(i) \cdot \mu_i$. Unfortunately, this may not be always optimal as we will show later.

When we choose a *fixed* channel $\mathbf{c}_i$ to sense/probe, the *rate of return*, *i.e.*, the achieved data rate for next transmission, is $\frac{R_{i,h} t_a (1-P_{loss})}{h \cdot t_{sp} + t_a}$ if we stop at the h-th probing after previous transmission. To maximize mean rate of return, it is easy to show that the optimal strategy is as follows (1) keep sensing/probing the channel $\mathbf{c}_i$ until the probed data rate $R_{i,t}$ at t-th probing after previous transmission is at least $\frac{\lambda_i^*}{1-P_{loss}}$; (2) transmit using data rate $R_{i,t}$. Here λ_i^* is the solution of

$$\sum_{X \geq \frac{\lambda_i^*}{1-P_{loss}}} \mathbf{Pr}_i(X)\left(X \cdot (1 - P_{loss}) - \lambda_i^*\right) = \lambda_i^* t_{sp}/t_a \tag{3}$$

where $\mathbf{Pr}_i(X)$ is the probability that the channel $\mathbf{c}_i$ will have rate X (by integrating the availability $p_I(i)$ of the channel $\mathbf{c}_i$). Notice that λ_i^* is different from μ_i.

We then design the best strategy if we have heterogeneous channels. We focus on one secondary user $\mathbf{s}_u$. Let λ_i^* denote the expected data rate achieved if we stick with sensing/probing channel $\mathbf{c}_i$ using the optimal stopping rule (see Eq. (3)). Without loss of generality, we assume $\lambda_i^* \geq \lambda_{i+1}^*$, for $i \in [1, N-1]$. Let Y denote any possible mixed strategy, *i.e.*, at any time slot t, we can randomly sense/probe a channel $\theta(t)$. Assume that with probability q_i it will sense/probe channel $\mathbf{c}_i$, where $\sum_{i=1}^{N} q_i = 1$. Let variable Y_t denote the data rate that can be observed at time t and λ_Y^* be the expected rate of the return by this strategy Y. Then $\Pr(Y_t = R_j) = \sum_{i=1}^{N} q_i \cdot \mathbf{Pr}_i(R_j)$. Notice that it is not necessary that $\lambda_Y^* = \sum_i q_i \lambda_i^*$. Let $\lambda_i^*(u)$ be the expected average data rate an SU $\mathbf{s}_u$ will achieve when it keeps sensing/probing and accessing the channel i. We show that this mixed strategy is no better than sticking with the best single channel.

Theorem 1: For any mixed strategy Y by an SU $\mathbf{s}_u$,

$$\lambda_1^*(u) \geq \lambda_Y^*(u) \geq \lambda_N^*(u) \tag{4}$$

Proof: Let $P_{loss,i}$ be the probability that the transmission by strategy using channel $\mathbf{c}_i$ will get lost due to the return of PU. Notice that $P_{loss,i}$ will be a fixed constant when the channel access time t_a is a fixed value. Let $p_i(x)$ be the probability density distribution (PDF) of the data rate of channel $\mathbf{c}_i$. Let $p_Y(x)$ be the PDF of the data rate observed by strategy Y. Then $p_Y(x) = \sum_i q_i p_i(x)$. Then $\frac{t_{sp}}{t_a}\lambda_i^* = E(\max(X_{i,t}(1 - P_{loss,i}) - \lambda_i^*, 0)) = \int_{x \geq \lambda_i^*/(1-P_{loss,i})} p_i(x) \cdot (x \cdot (1 - P_{loss,i}) - \lambda_i^*) \mathbf{d}x = \frac{1}{1-P_{loss,i}} \int_{u \geq \lambda_i^*} \bar{p}_i(u)(u - \lambda_i^*)\mathbf{d}u$. Then for strategy Y, its expected rate of return λ_Y^* is the solution of the following equation. $\lambda_Y^* \cdot \frac{t_{sp}}{t_a} = E[\max(Y_t(1 - P_{lossY}) - \lambda_Y^*, 0)] = \sum_i q_i E(\max(X_{i,t}(1 - P_{loss,i}) - \lambda_Y^*, 0)) = \sum_i q_i \int_{x \cdot (1-P_{loss,i}) \geq \lambda_Y^*} (x \cdot (1 - P_{loss,i}) - \lambda_Y^*) p_i(x) \mathbf{d}x$. Here variable P_{lossY} denotes the loss probability under strategy Y.

First assume by contradiction that $\lambda_Y^* < \lambda_N^*$. Then we have $\lambda_Y^* \cdot \frac{t_{sp}}{t_a} = \sum_i q_i \cdot E[\max((X_{i,t}(1 - P_{loss,i}) - \lambda_Y^*), 0)] \geq \sum_i q_i \lambda_i^* \cdot \frac{t_{sp}}{t_a} \geq \lambda_N^* \cdot \frac{t_{sp}}{t_a}$. The second last inequality comes from the fact that function $f(\lambda) = E[\max((X_{i,t}(1-P_{loss,i}) - \lambda), 0)]$ is monotonically decreasing function of λ and $\lambda_i^* > \lambda_Y^*$ for all i. Thus, we have $\lambda_Y^* \geq \lambda_N^*$, which contradicts to the initial assumption that $\lambda_Y^* < \lambda_N$.

We then show that $\lambda_1^* \geq \lambda_Y^*$ also by contradiction. Assume that $\lambda_1^* < \lambda_Y^*$. Then $\lambda_Y^* \cdot \frac{t_{sp}}{t_a} = \sum_i q_i \cdot E[\max((X_{i,t}(1 - P_{loss,i}) - \lambda_Y^*), 0)] \leq \sum_i q_i \lambda_i^* \cdot \frac{t_{sp}}{t_a} \leq \lambda_1^* \cdot \frac{t_{sp}}{t_a}$. Thus, we have $\lambda_Y^* \leq \lambda_1^*$, which contradicts to the initial assumption that $\lambda_Y^* > \lambda_1$. Consequently, we have $\lambda_1^* \geq \lambda_Y^* \geq \lambda_N^*$. □

For multiple secondary users, we define a weighted bipartite graph H over two sets of vertices SUs $\mathcal{S}$ and channels $\mathcal{C}$, and the weight of an edge $(\mathbf{s}_u, \mathbf{c}_i)$ is $\lambda_i^*(u)$. Let Π be a maximum weighted matching in graph H. Then, to maximize overall throughput, the best strategy for a user $\mathbf{s}_u$ is to use the optimal stopping method on channel i if $(\mathbf{s}_u, \mathbf{c}_i)$ is in the maximum weighted matching Π.

3.2 SPA for Nonstochastic Channels

Previous channel accessing methods often assume i.i.d. distributions of channel data rates (*e.g.*, [1], [8], [15], [18]), or Markovian [19]. Here we design a method whose expected throughput is almost optimal when the channel qualities may not have stochastic distributions.

3.2.1 Our Protocol Overview

The challenge is that at any time instance, we do not know whether our current strategy is *good* enough or not. Inspired by [4], we use parameter γ to adjust the fraction of exploration and exploitation. While in Exp3 [4], only the weight of the selected arm is updated, here the weight of the arms that are not selected are also updated (deterministically). Also, another difference is the addition of channel access which is done with probability epsilon. With probability $1 - \gamma$, we will *exploit* and just adapt the strategy used in previous round. Here γ is a relatively small parameter depending on the number of total rounds n (which is mainly decided by the lifetime T and our strategy). And exploitation will promise an almost optimal performance when previously used strategy is almost optimal, where the virtual rate for a candidate channel is defined as:

$$R'_{i,t} = \begin{cases} \frac{R_{i,t}+\beta}{p_{i,t}} & \text{if } \theta(t) = c_i \\ \frac{\beta}{p_{i,t}} & \text{otherwise.} \end{cases} \tag{5}$$

With a probability γ we will *explore* new channels, to estimate the rate of each channel with same probability $\frac{1}{N}$. The process of exploring is also important, which eventually improves our strategy to the optimal solution. In all our methods, for technical convenience, all actual data rates are scaled with maximum value 1. The virtual rates used in our method are used to compute the channel sensing/probing probability for next round.

3.2.2 Protocol for Single Secondary User

Algorithm 1 summarizes our protocol inspired by methods in [4], when there is only one secondary user. We present this method for the completeness of presentation and discussion of an efficient method for multiple users.

Algorithm 1 ϵ-SPA Scheme for Single User

Parameters: real number $\beta > 0$, $0 < \eta$, $\gamma < 1/2$.
Initialization: Set $w_{i,0} = 1$ for all $1 \leq i \leq N$, and $W = N$. Divide all data rates by the maximum possible data rate M.

1: At tth round, randomly select a channel $\theta(t) = c_i$ according to the following distribution $p_{i,t}$ $\forall i \in [1, N]$:

$$p_{i,t} = (1 - \gamma)\frac{w_{i,t-1}}{W_{t-1}} + \frac{\gamma}{N} \tag{6}$$

2: Sense and probe the channel $\theta(t)$, get the scaled data rate (*i.e.*, in the range $[0, 1]$), denoted as $R_{\theta(t),t}$, of channel $\theta(t)$ at time t. Calculate virtual rates $R'_{i,t}$, $\forall i$, according to Equ. 5.

3: Update

$$\begin{cases} w_{i,t} = w_{i,t-1}e^{\eta R'_{i,t}}, \\ W_t = \sum_{i=1}^{N} w_{i,t} \end{cases}$$

4: Access the channel $\theta(t)$ with probability ϵ, *i.e.*, set $\sigma(t) = 1$ with probability ϵ.

To study the performance of our algorithm, we analyze its *regret*. For this purpose, we define the accumulated data rate, denoted as $\mathbb{R}_{i,n}$, of channel $\mathbf{c}_i$, and the accumulated virtual data rate, denoted as $\mathbb{R}'_{i,n}$, as follows

$$\begin{cases} \mathbb{R}_{i,n} = \sum_{t=1}^{n} R_{i,t} \\ \mathbb{R}'_{i,n} = \sum_{t=1}^{n} R'_{i,t} \end{cases} \tag{7}$$

Lemma 2: For any $\delta \in (0, 1)$, $\beta \in [0, 1)$ and $i \in [1, N]$, we have $\mathbf{Pr}\left(\mathbb{R}_{i,n} > \mathbb{R}'_{i,n} + \frac{1}{\beta}\ln\frac{N}{\delta}\right) \geq \frac{\delta}{N}$

Proof: For any fixed i, $u > 0$ and $c > 0$, by the Chernoff bound, we have

$$\mathbb{P}[\mathbb{R}_{i,n} > \mathbb{R}'_{i,n} + u] \leq e^{-cu}\mathbb{E}e^{c(\mathbb{R}_{i,n}-\mathbb{R}'_{i,n}}$$

Let $u = \ln\frac{N}{\delta}/\beta$ and $c = \beta$, we get

$$e^{-cu}\mathbb{E}e^{c(\mathbb{R}_{i,n}-\mathbb{R}'_{i,n}} = \frac{\delta}{N}\mathbb{E}e^{\beta(\mathbb{R}_{i,n}-\mathbb{R}'_{i,n})}$$

Let $Z_n = e^{\beta(\mathbb{R}_{i,n}-\mathbb{R}'_{i,n})}$. To finish the proof, here we show that $\mathbb{E}[Z_n] \leq 1$ by showing $\mathbb{E}[Z_t] \leq Z_{t-1}$ for all $t \geq 2, \cdots, n$, and $\mathbb{E}[Z_1] \leq 1$. Let $X_{i,t} = R_{i,t} - \frac{R_{i,t}}{p_{i,t}}$ if $\theta(t) = c_i$, and $X_{i,t} = R_{i,t}$ otherwise. By definition, we have

$$Z_t = Z_{t-1}e^{\beta(X_{i,t}-\frac{\beta}{p_{i,t}})}$$

Let $\mathbb{E}_t$ denote the conditional expectation $\mathbb{E}[|\theta(1),\cdots,\theta(t-1)]$, we have

$$\mathbb{E}_t[Z_t] = Z_{t-1}\mathbb{E}_t[e^{\beta(X_{i,t}-\frac{\beta}{p_{i,t}})}] = Z_{t-1}e^{-\frac{\beta^2}{p_{i,t}}}\mathbb{E}_t[e^{\beta X_{i,t}}]$$

Since $\beta < 1$, $X_{i,t} \leq 1$ and $e^x \leq 1 + x + x^2$ for $x \leq 1$, we have

$$\begin{aligned}\mathbb{E}_t[Z_t] &\leq Z_{t-1}e^{-\frac{\beta^2}{p_{i,t}}}\mathbb{E}_t[1+\beta X_{i,t}+\beta^2X_{i,t}^2]\\ &= Z_{t-1}e^{-\frac{\beta^2}{p_{i,t}}}\mathbb{E}_t[1+\beta^2X_{i,t}^2]\end{aligned}$$

from the fact that $\mathbb{E}_t[X_{i,t}] = 0$. Finally, we can get

$$\mathbb{E}_t[Z_t] \leq Z_{t-1}e^{-\frac{\beta^2}{p_{i,t}}}(1+\frac{\beta^2}{p_{i,t}}) \leq Z_{t-1}$$

because the inequality $1 + x \leq e^x$ holds. A similar argument could show $\mathbb{E}[Z_1] \leq 1$ which finishes the proof. □

Define $\hat{\mathbb{R}}_n(\epsilon)$ as the expected total rates that can be achieved by an ϵ-SPA scheme over n rounds. Theorem 3 bounds the regret, $\max_{1\leq i\leq N}\mathbb{R}_{i,n} - \hat{\mathbb{R}}_n(1)$ of Algorithm 1 when $\epsilon = 1$.

Theorem 3: For any real value $\delta \in (0,1)$, when the real value $\beta = \sqrt{\frac{\ln\frac{N}{\delta}}{Nn}}$, $\gamma = 2\eta N$, $\eta = \sqrt{\frac{\ln N}{4nN}}$ and $n \geq \max\{\frac{\ln\frac{N}{\delta}}{N}, 4N\ln N\}$, we have

$$\mathbf{Pr}\left(\max_{1\leq i\leq N}\mathbb{R}_{i,n} - \hat{\mathbb{R}}_n(1) \leq 6\sqrt{nN\ln N}\right) \geq 1-\delta$$

Proof: The main idea of the proof is to bound the quantity $\ln\frac{W_n}{W}$. The lower bound is given as

$$\ln\frac{W_n}{W} \geq \ln\sum_{i=1}^{N}e^{\eta\mathbb{R}'_{i,n}} - \ln N \geq \eta\max_{1\leq i\leq N}\mathbb{R}'_{i,n} - \ln N$$

For the upper bound, for all $t = 1, 2, \cdots$ we have

$$\ln\frac{W_t}{W_{t-1}} = \ln\sum_{i=1}^{N}\frac{w_{i,t-1}}{W_{t-1}}e^{\eta R'_{i,t}} = \ln\sum_{i=1}^{N}\frac{p_{i,t}-\frac{\gamma}{N}}{1-\gamma}e^{\eta R'_{i,t}}$$

From the fact that $e^x \leq 1 + x + x^2$ for all $x \leq 1$, we get

$$\ln\frac{W_t}{W_{t-1}} \leq \ln(1+\sum_{i=1}^{N}\frac{p_{i,t}}{1-\gamma}(\eta R'_{i,t}+\eta^2R'^2_{i,t}))$$

Since $\ln(1 + x) \leq x$ for all $x > -1$, we have

$$\ln\frac{W_t}{W_{t-1}} \leq \frac{\eta}{1-\gamma}\sum_{i=1}^{N}p_{i,t}R'_{i,t} + \frac{\eta^2}{1-\gamma}\sum_{i=1}^{N}p_{i,t}R'^2_{i,t}$$

On one hand,

$$\sum_{i=1}^{N}p_{i,t}R'_{i,t} = R_{\theta(t),t} + N\beta$$

On the other hand,

$$\sum_{i=1}^{N}p_{i,t}R'^2_{i,t} \leq \sum_{i=1}^{N}p_{i,t}\frac{1+\beta}{p_{i,t}}R'_{i,t} = (1+\beta)\sum_{i=1}^{N}R'_{i,t}$$

Combining these results, we obtain

$$\ln\frac{W_t}{W_{t-1}} \leq \frac{\eta}{1-\gamma}(R_{\theta(t),t}+N\beta) + \frac{\eta^2(1+\beta)}{1-\gamma}\sum_{i=1}^{N}R'_{i,t}$$

Summing for $t = 1, \cdots, n$, we get

$$\ln\frac{W_n}{W} \leq \frac{\eta}{1-\gamma}(\hat{\mathbb{R}}_n(1)+nN\beta) + \frac{\eta^2(1+\beta)}{1-\gamma}\sum_{i=1}^{N}\mathbb{R}'_{i,n}$$

Since $\sum_{i=1}^{N}\mathbb{R}'_{i,n} \leq N\max_{1\leq i\leq N}\mathbb{R}'_{i,n}$ and

$$\ln\frac{W_n}{W} \leq \frac{\eta}{1-\gamma}(\hat{\mathbb{R}}_n(1)+nN\beta) + \frac{\eta^2(1+\beta)}{1-\gamma}N\max_{1\leq i\leq N}\mathbb{R}'_{i,n}$$

Combining the upper bound and lower bound, we get

$$\hat{\mathbb{R}}_n(1) \geq (1-\gamma-\eta(1+\beta)N)\max_{1\leq i\leq N}\mathbb{R}'_{i,n} - \frac{1-\gamma}{\eta}\ln N - nN\beta$$

Apply Lemma 2, with probability at least $1 - \delta$,

$$\begin{aligned}\hat{\mathbb{R}}_n(1) \geq (1-\gamma-\eta(1+\beta)N)(\max_{1\leq i\leq N}\mathbb{R}_{i,n} - \frac{1}{\beta}\ln\frac{N}{\delta})\\ -\frac{1-\gamma}{\eta}\ln N - nN\beta\end{aligned}$$

Let $\hat{\mathbb{L}}_n(1) = n - \hat{\mathbb{R}}_n(1)$ and $\min_{1\leq i\leq N}\mathbb{L}_{i,n} = n - \max_{1\leq i\leq N}\mathbb{R}_{i,n}$. Use the fact $1 - \gamma - \eta(1 + \beta)N > 0$ which follows the assumptions of the theorem, with probability at least $1 - \delta$, we have

$$\begin{aligned}\hat{\mathbb{L}}_n(1) \leq n(\gamma+\eta(1+\beta)N) + (1-\gamma-\eta(1+\beta)N)\min_{1\leq i\leq N}\mathbb{L}_{i,n}\\ +(1-\gamma-\eta(1+\beta)N)\frac{1}{\beta}\ln\frac{N}{\delta} + \frac{1-\gamma}{\eta}\ln N + nN\beta\end{aligned}$$

which implies

$$\hat{\mathbb{L}}_n(1) - \min_{1\leq i\leq N}\mathbb{L}_{i,n} \leq n\gamma + 2\eta nN + \frac{1}{\beta}\ln\frac{N}{\delta} + \frac{\ln N}{\eta} + nN\beta$$

By setting $\beta = \sqrt{\frac{1}{nN}\ln\frac{N}{\delta}}$, $\gamma = 2\eta N$, and $\eta = \sqrt{\frac{\ln N}{4nN}}$, the above regret is minimized: $\max_{1\leq i\leq N}\mathbb{R}_{i,n} - \hat{\mathbb{R}}_n(1) = \hat{\mathbb{L}}_n(1) - \min_{1\leq i\leq N}\mathbb{L}_{i,n} \leq 6\sqrt{nN\ln N}$. This finishes the proof. □

According to our ϵ-SPA scheme, CR will transmit ϵn times in expectation during n rounds. It is easy to show $E[\hat{\mathbb{R}}_n(\epsilon)] = \epsilon E[\hat{\mathbb{R}}_n(1)]$, which implies

$$E[\hat{\mathbb{R}}_n(\epsilon)] \geq \epsilon\max_{1\leq i\leq N}E[\mathbb{R}_{i,n}] - 6\epsilon\sqrt{nN\ln N} \tag{8}$$

Let R_{max} be the largest expected data rate among all channels. We have $\max_{1\leq i\leq N}E(\mathbb{R}_{i,n}) = n \cdot R_{max}$. Assume $t_a = \alpha t_{sp}$ where constant $\alpha \gg 1$. Then we have the expected throughput of ϵ-SPA is at least

$$\frac{\hat{\mathbb{R}}_n(\epsilon)t_a}{T} \geq \frac{R_{max} - 6\sqrt{\frac{N\ln N}{n}}}{\frac{1}{\epsilon\alpha}+1} = \frac{R_{max} - 6\sqrt{\frac{(1+\alpha\epsilon)t_{sp}N\ln N}{T}}}{1+\frac{1}{\alpha\epsilon}}$$

with probability $1-\delta$, where $T = nt_{sp} + \epsilon nt_a$.

When T is sufficiently large, the expected throughput achieved is at least $\frac{R_{max}}{1+\frac{1}{\alpha\epsilon}}$, which is maximized when $\epsilon = 1$. Clearly, the expected throughput that can be achieved is no more than $\frac{R_{max}t_a}{t_{sp}+t_a} = \frac{R_{max}}{1+\frac{1}{\alpha}}$, because each transmission takes at least $t_{sp}+t_a$ time while the expected data rate is no more than R_{max}. Thus, when T is sufficiently large, our ϵ-SPA scheme is almost optimal.

3.2.3 Multiple Secondary Users

Here we consider the case when there are u secondary users with a centralized decision. For our SPA scheme, its main idea is still to *explore/exploit* channels. The difference is that our new method needs to select k channels in each round, k is the number of channels one could sense in each round for future selection. Therefore, there are totally $\binom{N}{k}$ different strategies when $k \le N$ (here we view each channel as an arm in the MAB problem, and each SU as a player). When $k \ge N$, each channel should be sensed/probed/accessed. Then we will view each SU as an arm, and each channel as a player, thus, there are $\binom{k}{N}$ different strategies when $k \ge N$. The rest of discussions focuses on the case $k \le N$. All our results can be converted easily when $k \ge N$.

Let χ_j denote the jth strategy which includes a set of k channels, where $1 \le j \le \binom{N}{k}$. Here we still use $\theta(t)$ to denote the set of channels we choose at time t. A simple method to address multiple users is to modify Algorithm 1 with new settings: each strategy is a combination of k channels, and each strategy χ is associated with a weight $w_{\chi,t}$ which is recursively updated as

$$\begin{cases} w_{\chi,t} &= w_{\chi,t-1}\prod_{i\in\chi} e^{\eta R'_{i,t}}; \\ W_t &= \sum_{\chi} w_{\chi,t} \end{cases}$$

Then at time t, a strategy χ is chosen with probability

$$p_{\chi,t} = (1-\gamma)\frac{w_{\chi,t-1}}{W_{t-1}} + \gamma\frac{\sum_{S\in\mathbb{C}} I((\chi\cap S)\neq\emptyset)}{|\mathcal{C}|}. \quad (9)$$

Here the set $\mathbb{C} = \{\{\mathbf{c}_1,\cdots,\mathbf{c}_k\},\{\mathbf{c}_{k+1},\cdots,\mathbf{c}_{2k}\},\cdots,\}$ is called a covering set of $\mathcal{C}$. Notice $\mathbb{C}$ includes strategies such that each channel appears in at least one of the strategies. So the size of a covering set is $\lceil N/k\rceil$. This simple implementation will have the same regret bound as our new method (to be discussed), but it will have time and space complexity $O(N^k)$, exponential in number of users, which could be very expensive if we have a large number of users.

To reduce the complexity, we propose a novel approach that utilizes the internal structure of our method. We will not choose a strategy from the strategy set (*i.e.*, any combination of k channels) directly. In our new method (Algorithm 2), we make decision on each channel one by one. Assume each channel $\mathbf{c}_i$ is still associated with a weight $w_{i,t}$ at time t. And the weight of a strategy χ_j is defined as the product of weights of all channels in that strategy. Let $S(p,q,k)$ denote the set of all strategies that will choose exactly k channels from subset $\{\mathbf{c}_p,\mathbf{c}_{p+1},\cdots,\mathbf{c}_q\}$. Define the weight of $S(p,q,k)$ at time t as $W_t(p,q,k) = \sum_{\forall\chi_j\in S(p,q,k)}\prod_{\mathbf{c}_i\in\chi_j} w_{i,t}$. For example, the total weight of all strategies choosing k channels from N channels is $W_t(1,N,k)$. Here the probability that we choose channel $\mathbf{c}_1$ in our strategy at time t is $\frac{w_{1,t}W_t(2,N,k-1)}{W_t(1,N,k)}$, which is the total weight of strategies choosing channel $\mathbf{c}_1$ over the total weight of all strategies. Similarly, the probability that channel $\mathbf{c}_1$ is not chosen at time t is $\frac{W_t(2,N,k)}{W_t(1,N,k)}$.

For channel $\mathbf{c}_2$, if channel $\mathbf{c}_1$ is chosen, the probability that channel $\mathbf{c}_2$ is in the strategy is $\frac{w_{1,t}w_{2,t}W_t(3,N,k-2)}{w_{1,t}W_t(2,N,k-1)} = \frac{w_{2,t}W_t(3,N,k-2)}{W_t(2,N,k-1)}$; if channel $\mathbf{c}_1$ is not chosen, the probability that channel $\mathbf{c}_2$ is in the strategy is $\frac{w_{2,t}W_t(3,N,k-1)}{W_t(2,N,k)}$. Thus, for channel $\mathbf{c}_i$, if k' channels has been chosen among $\mathbf{c}_1, \cdots, \mathbf{c}_{i-1}$, the probability that $\mathbf{c}_i$ is in the strategy is $\frac{w_{i,t}W_t(i+1,N,k-k'-1)}{W_t(i,N,k-k')}$.

For current time slot, repeat the previous steps for all channels. We can show that, given that exactly k channels will be chosen, the time and space complexity at each timeslot is $O(u\cdot N)$ and the probability that a strategy χ will be chosen is $p_{\chi,t}$.

For regret analysis, similarly, we define

$$\begin{cases} \mathbb{R}_{\chi_j,n} = \sum_{t=1}^{n} R_{\chi_j,t} = \sum_{t=1}^{n}\sum_{c_i\in\chi_j} R_{i,t} \\ \mathbb{R}'_{\chi_j,n} = \sum_{t=1}^{n} R'_{\chi_j,t} = \sum_{t=1}^{n}\sum_{c_i\in\chi_j} R'_{i,t} \end{cases} \quad (10)$$

Lemma 2 still holds and similarly, we have following theorem.

Theorem 4: For $\delta \in (0,1)$, $\beta = \sqrt{\frac{\ln\frac{kN}{\delta}}{Nn}}$, $\gamma = 2\eta N$, $\eta = \sqrt{\frac{\ln N}{4nkN}}$ and $n \ge \max\{\frac{k\ln\frac{N}{\delta}}{N}, 4\frac{N}{k}\ln N\}$, we have

$$\mathbf{Pr}\left(\max_{\chi_j}\mathbb{R}_{\chi_j,n} - \hat{\mathbb{R}}_n(1) \le 6k\sqrt{nN\ln N}\right) \ge 1-\delta \quad (11)$$

where $\hat{\mathbb{R}}_n(1)$ is the expected total rates that can be achieved by an 1-SPA scheme over n rounds.

The proof is similar to that of theorem 3. Let R^k_{max} be the largest total expected data rates of k channels among all channels. When T is sufficiently large, our ϵ-SPA scheme for u users is almost optimal.

3.3 Further Discussions on SPA

3.3.1 Exploit Channel Coherence

Here we consider the affect of channel coherence time. Within the coherence time t_c we assume the data rate does not change with high probability. Recalling a previously sensed/probed channel is possible when t_c is long enough. Assume that $t_c \ge m(t_{sp}+\epsilon t_a)+t_a$ for some integer $m \ge 1$. Then we can recall m previously sensed/probed channels whose data rates are assumed to remain same.

Our method with recall is similar to our ϵ-SPA scheme. The difference is as follows. At each round t, we randomly select a channel $\theta(t)$, sense and probe it. At same time, we "virtually" sense and probe the last $m-1$ channels we just sensed and probed, *i.e.*, channels $\theta(t-1),\cdots,\theta(t-m+1)$. Here "virtually" means we don't really sense and probe them. We update the weights just as if we sensed and probed them. Thus, in each round, we sense and probe m channels, and choose the best one with probability ϵ. Then,

Algorithm 2 Coordinated ϵ-SPA Scheme for k Users

Parameters: real number $\beta > 0$, $0 < \eta, \gamma < 1$.

Initialization: Set $w_{i,0} = 1\ \forall i \in [1, N]$; $W_0(N, N, 1) = 1$; $W_0(N, N, k') = 0\ \forall k' \geq 2$; $W_0(i, N, 1) = N - i + 1\ \forall i \geq 1$.

1: At tth round, initialize $\theta(t) = \emptyset$, $k' = k$, and update $W_t(i, N, k')$, $\forall i \in [1, N]$ and $\forall k' \in [1, k]$ as follows.

$$W_t(i, N, k') = w_{i,t}W_t(i+1, N, k'-1) + W_t(i+1, N, k')$$

2: **for** $i = 1$ to N **do**

3: With prob. $\frac{w_{i,t}W_t(i+1,N,k'-1)}{W_t(i+1,N,k')}$, choose channel $\mathbf{c}_i$,

$$\theta(t) = \theta(t) \bigcup \mathbf{c}_i \quad \text{and} \quad k' = k' - 1$$

4: Compute $p_{i,t} = \gamma \frac{\sum_{S\in\mathcal{C}} I((\chi\cap S)\neq\emptyset)}{|\mathcal{C}|} + (1 - \gamma) \sum_{k'=0}^{k} \frac{W_{t-1}(1,i-1,k')w_{i,t-1}W_{t-1}(i+1,N,k-k'-1)}{W_{t-1}(1,N,k)}$.

5: Sense and probe all channels in $\theta(t)$, get the scaled data rates. Calculate virtual rates $R'_{i,t}$, $\forall i \in [1, N]$:

$$R'_{i,t} = \begin{cases} \frac{R_{i,t}+\beta}{p_{i,t}} & \text{if } c_i \in \theta(t) \\ \frac{\beta}{p_{i,t}} & \text{oththerwise.} \end{cases}$$

6: Update the weights, $\forall i$, $w_{i,t} = w_{i,t-1}e^{\eta R'_{i,t}}$.

7: Access all channels in $\theta(t)$ with probability ϵ, *i.e.*, set $\sigma(t) = 1$ with probability ϵ.

Theorem 5: When $\beta = \sqrt{\frac{\ln\frac{N}{\delta}}{Nnm}}$, $\gamma = 2\eta N$, $\eta = \sqrt{\frac{\ln N}{4nmN}}$ and $nm \geq \max\{\frac{\ln\frac{N}{\delta}}{N}, 4N\ln N\}$, the regret of our algorithm 1-SPA with m channel recall satisfies $\mathbf{Pr}\left(\max_{1\leq i\leq N}\mathbb{R}_{i,n} - \hat{\mathbb{R}}_n(1) \leq 6\sqrt{\frac{nN\ln N}{m}}\right) \geq 1 - \delta$.

Proof: The proof is similar to that of theorem 3. We can consider each round as a m virtual sub-rounds, where the $(jm + k)$th sub-round, for $1 \leq j \leq n$, $0 \leq k \leq m - 1$, is mapped to the $(j + m - k)$th round in the problem without recall. Then the access probability of each sub-round is $1/m$ for our 1-SPA scheme and we have $n \cdot m$ sub-rounds totally. By replacing ϵ with $\frac{1}{m}$, and n with nm in Eq. (8), we have $\mathbf{Pr}\left(E[\hat{\mathbb{R}}_{nm}(1)] \geq \frac{1}{m}\max_{1\leq i\leq N} E[\mathbb{R}_{i,nm}] - 6\sqrt{\frac{nN\ln N}{m}}\right) \geq 1 - \delta$. The theorem then follows from $E[\hat{\mathbb{R}}_n(1)] = E[\hat{\mathbb{R}}_{nm}(1)]$, and $\max_{1\leq i\leq N} E[\mathbb{R}_{i,n}] = \frac{1}{m}\max_{1\leq i\leq N} E[\mathbb{R}_{i,nm}]$. □

Similarly with a proper parameter setting we have $\mathbf{Pr}\left(E[\hat{\mathbb{R}}_n(\epsilon)] \geq \epsilon\max_{1\leq i\leq N} E[\mathbb{R}_{i,n}] - 6\epsilon\sqrt{\frac{nN\ln N}{m}}\right) \geq 1-\delta$ for our ϵ-SPA scheme. Comparing with Eq. (8), allowing recall improves the convergence speed of our method.

3.3.2 Impact of sensing time

Notice that we don't consider the false alarm probability of sensing in previous analysis. Since we consider energy detector for channel sensing, the false alarm probability [13] is approximated by $P_{fa}(t_s) = Q((\frac{\epsilon_0}{\sigma_u^2} - 1)\sqrt{t_s f_s})$, where $\frac{\epsilon_0}{\sigma_u^2}$ is the decision threshold for sensing, f_s is the channel bandwidth, and $Q()$ is the Q-function for the tail probability of the standard normal distribution. Consider the false alarm probability, we have

Lemma 6: With probability $1 - \delta$, the expected throughput of ϵ-SPA scheme (Algorithm 1) is at least

$$\left(R_{max} - 6\sqrt{\frac{(1 + \alpha\epsilon_0)t_{sp}N\ln N}{(1 - P_{fa})T}}\right) / \left(1 + \frac{1}{\alpha\epsilon_0}\right)$$

The proof is similar to that of Theorem 5. For each round, we sense and probe channel successfully $1 - P_{fa}$ in expectation. Replacing m with $1-P_{fa}$, we get the above result. Here P_{fa} is a function of t_s, and $\alpha = \frac{t_a}{t_s+t_p}$. Treating t_s as a variable, we can compute the optimal t_s which maximizes the expected throughput by numerical analysis.

3.3.3 Impact of probing time and others

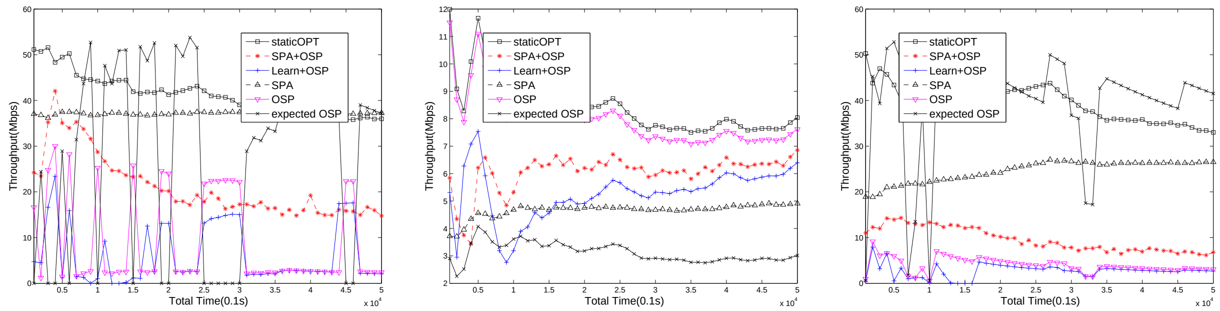
The step of probing is not necessary in our problem. The reason why we need to probe the channel is that we want to make sure the data rate is good enough. This is important when the qualities of the channels are not good. On the other hand, when the qualities of the channels are good enough, a sense/access scheme may achieve better throughput since there is no probing overhead. Our ϵ-SPA scheme also can be extended to a simplified ϵ-SA scheme without probing steps. In ϵ-SA scheme, we can only get the observation on the data rate after each successful transmission and ACK. In other word, in n rounds, ϵn data rates will be observed in expectation by the ϵ-SA scheme. Replacing m with ϵ in Theorem 5, we can show that the expected throughput of ϵ-SA scheme is $\left(R_{max} - 6\sqrt{\frac{(1+\alpha'\epsilon)t_sN\ln N}{\epsilon T}}\right) / \left(1 + \frac{1}{\alpha'\epsilon}\right)$, where $\alpha' = \frac{t_a}{t_s}$. Let t_p^* be the probing time which satisfies $(R_{max} - 6\sqrt{\frac{(1+\alpha'\epsilon)t_sN\ln N}{\epsilon T}})/(1 + \frac{1}{\alpha'\epsilon}) = (R_{max} - 6\sqrt{\frac{(1+\alpha\epsilon)(t_s+t_p^*)N\ln N}{T}})/(1 + \frac{1}{\alpha\epsilon})$. When $t_p \leq t_p^*$, we will use ϵ-SPA; when $t_p \geq t_p^*$, we will use ϵ-SA.

The transmission may be destroyed by the return of PUs. Thus with the knowledge of p_I we can optimize t_a to maximize the expected throughput. By numerical analysis, we can also find ϵ that maximizes the expected throughput.

3.4 Decentralized Protocol With Multiple Users

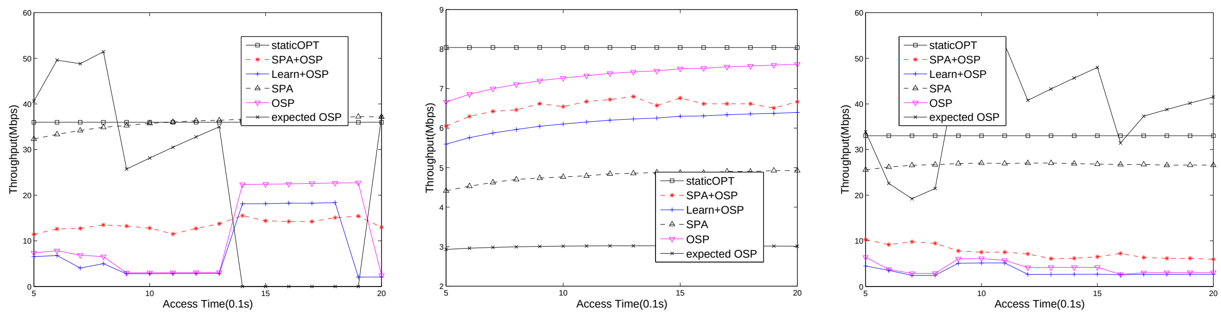
It is challenging to design optimal decentralized protocol without using a common control channel (CCC). Since the energy-detection cannot differentiate spectrum usage of PUs and SUs, the view of each SU is also affected by other SUs. The channels quality and availability are thus dynamic and nonstochastic. Notice that the protocol ϵ-SPA developed previously applies to a more general observation model as long as SUs have the common set of u best channels and each of these channels has the same mean across players.

Our method is to let each user in the network run the ϵ-SPA method based on their own view. Eventually, each user will almost maximize its own throughput when others do not change their strategies. We expect to prove that the regret of each user is upper-bounded in this nonstochastic setting. This could be proved by applying an approach similar to the TDFS scheme [15]. We leave the detailed analysis as a future work.



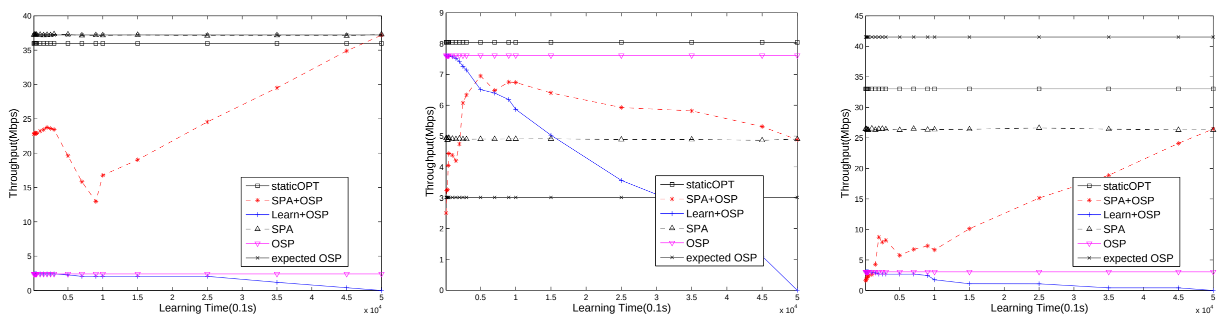
(a) Good Channels, (b) Bad Channels, (c) Mixed Channels.

Fig. 1. Performance when total lifetime T varies.



(a) Good Channels, (b) Bad Channels, (c) Mixed Channels.

Fig. 2. Performance when accessing time t_a varies.



(a) Good Channels, (b) Bad Channels, (c) Mixed Channels.

Fig. 3. Performance when learning time varies

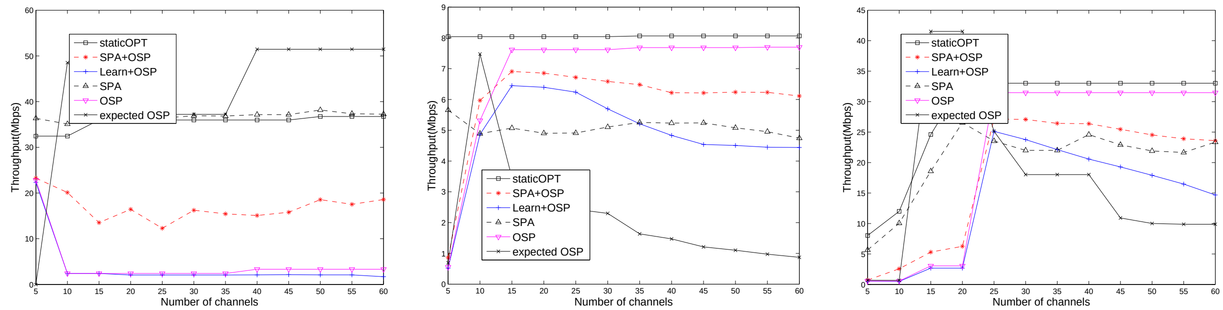
4 Performance Verifications

4.1 Large Scale Simulation Studies

4.1.1 Notations

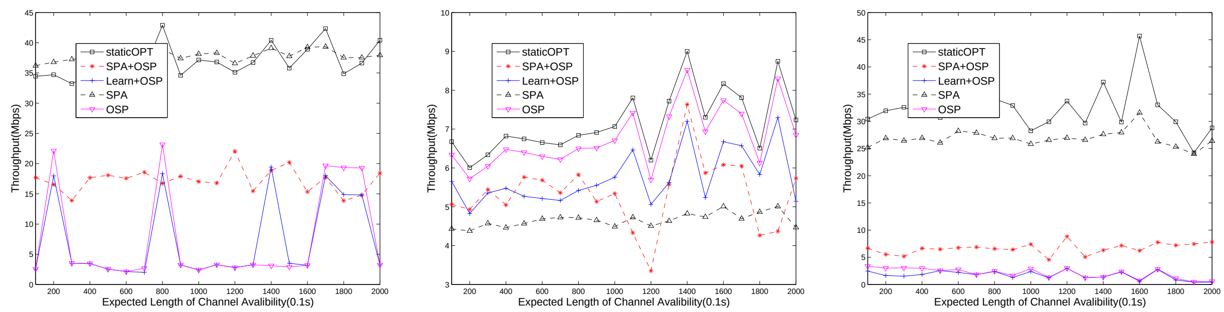
We conduct extensive simulations to study the throughput of our methods. "SPA" in figures denotes our SPA scheme. "Learn+OSP" in figures denotes the strategy that first learn the distributions for certain time duration, then use optimal stopping method in [18]. We also study the performance of a method that first apply our SPA scheme for certain time duration in which distributions are learned and then use optimal stopping method, which is denoted as "SPA+OSP" in figures. To evaluate the performance, we compare our methods with some other methods. The static optimum ("staticOPT" in figures) denotes the maximum possible data rate among all channels, which, although cannot be implemented in practice, will be used as a comparison reference for all methods. Notice that the static optimum cannot be achieved since it does not consider the sensing and probing time. We also implement optimal stopping method ("OSP" in figures) in [18], which knows distributions of channels' data rates and availability. In [18], channel qualities and channel availabilities are homogeneous; in our work, they could be heterogeneous, which may affect the performance of optimal stopping method. Therefore, we also plot "expected OSP" in our figures, which denotes the expected performance of optimal stopping method when each channel has enough duplications (channels with same availability and distribution of data rates). The "expected OSP" is simply the value $\max_i \lambda_i^*$.

In our simulation, we consider the affect of coherence time and channel availability. When a transmission starts at time t



(a) Good Channels, (b) Bad Channels, (c) Mixed Channels.

Fig. 4. Performance when number of channels N varies.



(a) Good Channels, (b) Bad Channels, (c) Mixed Channels.

Fig. 5. Performance when expected length of channel availability varies.

with data rate R, it will last t_a timeslots. From time t to time $t+t_a-1$, we only count the total data transmitted in timeslots whose data rate is at least R since the transmission will fail when the actual data rate is less than R.

4.1.2 Simulation Setup

The channel availability is simulated as that the primary users are using the channels in Poisson distribution; thus the duration of ideal state and busy state follow Exponential distribution with mean $\alpha = 50s$ and $\beta = 50s$. The system lifetime is 5000 seconds and each timeslot is 100ms. The sensing time and probing time $t_s + t_p$ is 100ms.

We use the data rates of IEEE 802.11g, *i.e.* the possible data rates are $\{1, 2, 6, 9, 12, 18, 24, 36, 48, 54\}$ Mbps. We map the ten data rates to 10 intervals $[0.5+i, 1.5+i], i = 0, 1, 2, \ldots, 9$, where the ith data rate is mapped to the ith interval. The probability for each data rate is generated by Normal distribution with mean μ uniformly distributed in $[0.5, 10.5]$ and the standard deviation σ is 2. We generate a good (or median, bad, respectively) channel by choosing μ uniformly from $[6.5, 10.5]$ (or $[3.5, 7.5]$, $[0.5, 4.5]$, respectively). The frequency coherence time follows Exponential distribution with mean value 5 seconds. And the learning time for "learn+OSP" and "SPA+OSP" is $700s$.

We study 4 static scenarios: 20 good channels, 20 median channels, 20 bad channels, and 20 mixed channels (composed of 6 bad channels, 7 median channels and 7 good channels), where the quality of each channel follows some corresponding static distribution. We also study 1 dynamic scenario: 20 mixed channels, at half of the total lifetime, bad (median, good respectively) channels becomes median (good, bad respectively) channels. Due to space limitation, we did not plot the scenario of median channels since it did not tell significant differences.

4.1.3 Results and Findings for Single User

We compare the performance of different methods for single user case. Figure 1 (Fig. 2, Fig. 3, Fig. 4, Fig. 5 respectively) shows the results when total lifetime T (accessing time t_a, learning time, number of channels N, expected length of channel availability respectively) varies. Our key observations are as follows.

First, our SPA scheme performs well in all cases. When the channel qualities are not bad, SPA scheme is the best one among all methods (staticOPT and expected OSP are not actual protocols, but used just as reference). SPA scheme is even better than the optimal stopping method, which knows the distributions of data rates in advance. This is because SPA scheme takes the advantage of channel hopping, while OSP sticks on the best channel and wastes a lot of time in waiting for channel value to be above the threshold.

Second, the optimal stopping method does not work well when channels are not bad. Actually, in our heterogeneous setting, its performance is much worse than its expected performance in homogeneous setting (see "expected OSP"). The throughput of OSP even decreases when the channel qualities are better. The reasons are as follows. When channel qualities are good, OSP will have a higher threshold (which is equal to the expected throughput). In the homogeneous setting, if one channel is busy, OSP may find another one immediately, then a higher threshold results in better performance. However, in heterogeneous setting, OSP will stick on the best channel. If that channel is busy, OSP has to wait since there are no other replacements. Thus, higher threshold reduces the performance actually because too much time is wasted in waiting. When the distributions of channels' data rates are not known in advance, the performance of OSP is worse (see "learn+OSP").

Third, the performance of "SPA+OSP", a mixed strategy of our SPA scheme and optimal stopping method, is always between that of SPA and OSP, and better than "learn+OSP" although "SPA+OSP" learns less data rates than "learn+OSP" (the learning times are same, "SPA+OSP" takes $t_s + t_p + \epsilon t_a$ times to learn one data rate in expectation, while "learn+OSP" only takes $t_s + t_p$ times). This is because our SPA scheme focuses more on those channels with better qualities eventually. Therefore, the learning is more accurate in those good channels, which will significantly improve the performance.

Therefore, our conclusion for single user system is: use SPA scheme when the channel qualities are not bad; use "SPA+OSP" when the channel qualities are bad.

4.1.4 Results and Findings for Multiple Users

We then study the performance of our schemes in a network of CRs, and study the affect of the number of users. Assume each user is equipped with one radio. In results reported here, "SPA+C" denotes the performance when control channel is available (this case can be addressed by Algorithm 2); "SPA+NC" denotes the performance when control channel is not available, *i.e.*, each user does not know the existence of each other, thus uses Algorithm 1 individually. The definitions of "staticOPT", "OSP", "learn+OSP" are as same as those in previous subsection. The main difference is that there are multiple users. For a network of k users, the strategy is to stick on the best k channels (each user for one channel).

4.2 Testbed Implementation and Results

We build a cognitive radio system which is composed of 20 sensor nodes and 6 USRP devices and conduct extensive experiment studies on the performance of our system. The maximum transmission speed for each sensor node is $250kbps$. Using this software programmable platform by USRP, we can adaptively select the channel and transmission power. We use these USRPs as cognitive nodes and use the other sensor nodes as primary users. More importantly, we can also monitor and diagnose the wireless network using USRPs. There are 16 channels (from channel 11 to 26) available for the primary user network, and in each channel, the transmission bandwidth of the primary users is 2MHz, while the bandwidth of USRPs is 500kHz. We use the software suit built upon GNU radio, and add our SPA model into the model. There are two additional modules for running the SPA algorithm efficiently.

4.2.1 Overall Design, Workflow, and Parameters

We assume that all the secondary users will agree upon a common control channel (CCC). In our testing, we use channel 26 as the CCC. If the sender wants to initiate a communication, it will send a packet INIT_PKT to the receiver using CCC and then start the SPA process. The INIT_PKT also contains a seed that will be used to generate the sequence of random channel numbers used by SPA. The receiver will wait for the communication request from the sender. It will start the SPA process after it received the packet INIT_PKT.

TABLE 1
Parameter values for Implementations

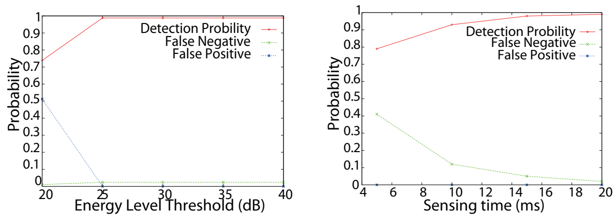
Parameters	Values	Parameters	Values
Sensing_time	200ms	Probing_time	800ms
Accessing_time	2000ms	Running_time	100s to 600s
switch_delay	50ms	proc_delay	3ms
update_delay	50ms	Packet_interval	10ms
probe_num	50	probe_RT	200ms

Channel Sensing: We use the energy threshold method to detect if a given channel is busy or idle. We perform signal smooth and filtering such that it works well under the unique characteristics of sensor networks. The channel is considered to be busy if the energy level of the signal is above a certain threshold. Fig. 6(a) shows the different success probability using different energy threshold for detecting the busy/idle status of a channel in our testing.

For this experiment, using $25dB$ may be the best choice. The success probability also depends on the duration we examine a channel. From Fig. 6 (b), we find that when we exam a channel for about $15ms$, the success probability is already good enough. We choose $200ms$ in experiments for better accuracy.

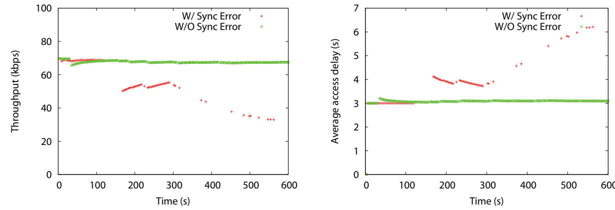
Channel Probing: If the sender determines that the channel is idle using channel sensing method described previously, it then starts probing the channel quality and uses a timer, **PT**, for probing time. Timer **PT** starts with the maximum value $Probing\_time$. The sender will send Probe_num of probing packets when **PT** is reduced to $Probing\_time - switch\_delay$. Here we use a delay, $switch\_delay$, to avoid the negative impact caused by the processing delay of USRP. The receiver could send back the ACKs for the probing packets. Such feedback messages can be collected for the SNR and PRR (packet reception ratio) information. Before the expiration of the timer **PT**, when **PT** is reduced to a value $update\_delay + switch\_delay$, the sender updates the channel observations based on the probing results. Here $update\_delay$ is the delay for the SPA algorithm to update its internal parameters. This could be negligible for USRP node, but not for sensor nodes when our algorithm runs in sensor nodes. They decide whether to access some channel or continue sensing/probing other channels.

For the receiver, it also starts a timer **PT** with starting value $Probing\_time$, when it knows that the sender/receiver starts the probing phase. It will send an ACK packet to the sender when **PT** is reduced to time $prob\_RT + update\_delay + switch\_delay$. Here $prob\_RT$ is the round-trip delay for the probing packet. In our experiments, we set $prob\_RT$ to be a reasonable value that can also cover the errors of clock synchronization between the sender and the receiver, the clock shifting of the sender and receiver. A proper choice of this parameter can let the sender receive the ACK packet



(a) energy threshold (b) sensing time

Fig. 6. Energy and time threshold for channel sensing.



(a) throughput (b) delay

Fig. 7. Impact of clock synchronization error on throughput and delay.

in time. It also updates its own SPA internal parameters. Then it will decide to either access some channel or continue sensing/probing other channels based on SPA method using the common random number generator as the sender.

If the sender or the receiver determines the channel is busy, they will start another round of sensing/probing after delay of $probing\_time$.

Parameters: Table 1 summarizes some parameters used for our experiments. These parameters are selected according to the performance of the working USRP system, and to improve the stability of the system.

Channel Accessing: During the channel accessing period, the sender will send data packets to the receiver using the maximum data rate under the current channel condition. Transmission of two consecutive packets are separated by a delay $proc\_delay$. This delay is introduced to avoid the packet-drops by the receiver when the receiver has a processing speed lower than that of the sender.

Coding and modulation effects: We made extensive tests on multiple coding and modulation combinations over different SNRs. The interferences are from the external and internal. The external interferences are generated by the sensor nodes, where packet transmissions occurred periodically around the USRP nodes. The internal "interferences" are generated by the power control ability on the USRP mother board. With different levels of the output gain at the transmitter side, the SNR value changes accordingly. We can use different combination tests for the optimal transmission rate. As shown in Fig. 8 (a), we adjust the transmission power to different levels and achieve the mapping table of SNR and transmission power. We then conduct extensive experimental study to investigate the SNR-PER relationship. Fig. 8(c) illustrates the relationship between SNR and the PER.

In our experiments, we have 8 different combinations of modulation and coding rate: (BPSK, 1/2), (BPSK, 3/4), (QPSK, 1/2), (QPSK, 3/4), (16-QAM, 1/2), (16-QAM, 3/4), (64-QAM, 1/2), (64-QAM, 3/4). The corresponding data rates (units kbps) are 150, 225, 300, 450, 600, 900, 1200, and 1350 respectively. We implement this based on rawOFDM (see http://people.csail.mit.edu/szym/rawofdm/README.html).

We applied two different approaches to build heterogeneous channel qualities: (1) adjust the transmission power of the transmitter (We use RFX2400 daughter board in the experiment, which does not allow us to change output gain, therefore we adjust the amplitude of transmission signal to change power) (2) using a separate wireless sensor network that will transmit the collected data where the data arrival follows a Poisson distribution with arrival rate λ_0.

Fig. 8 shows the impact of different amplifications on SNR and PER.

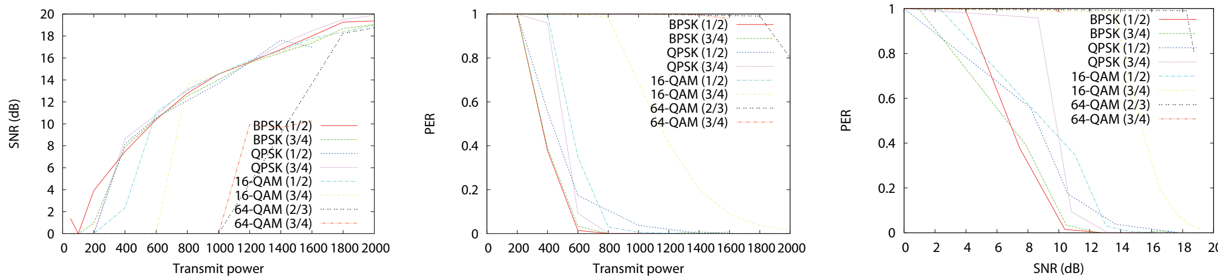
4.2.2 Experimental Results

In our experiment, we use 6 USRPs (numbered 1 to 6) to form 3 pairs, denoted as (1,2), (3,4), (5,6), of sender/receiver as secondary users. We first measure the energy level on different channels when the SPA is applied (this is also used for idle/busy detection).

Impact of Synchronization: Due to different processing abilities and asynchronous clocks by the sender and the receiver, they may have different views on the current system status, which could greatly impact the throughput of the system and the delay of accessing the channels. In our testbed implementation, we introduced several delay mechanisms by the sender and receiver, such as $switch\_delay$, $proc\_delay$, and $update\_delay$ to address this notoriously challenging issue. Unfortunately, these delay mechanisms will reduce the ideal achievable maximum throughput and increase the ideal minimum accessing-delay (time duration between two consecutive successful transmissions) to some extent. Accessing-delay is called delay hereafter. To study the impact of time synchronization on the throughput and delay, we conduct two separate experiments (1) the first experiment will connect each pair of the communicating USRPs to a computer, which we assume that time synchronization error is negligible; (2) the second experiment will let each pair of the communicating USRPs to connect to two separate computers, which we assume that the synchronization error is not negligible. Fig. 7 shows our experimental results of the impact of synchronization error on the throughput and delay. The throughput drops significantly and the delay increases significantly after 300s when synchronization errors exist. On the other hand, the throughput and delay remain stable when synchronization error is negligible.

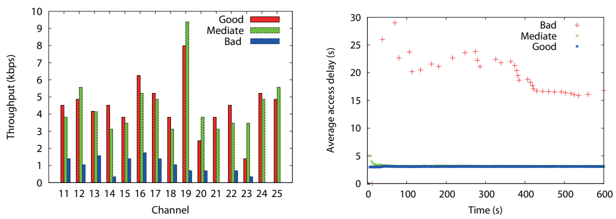
Impact of different channel qualities: We then study the impact of different channel qualities on the performance. Recall that two different approaches are used in controlling the channel qualities: adjusting the signal amplitude, or the number of primary users already using some channels.

For the first approach, we apply three different groups of amplifications, denoted as *bad*, *mediate*, and *good*. The amplifications of 15 channels are as follows: (1) Bad: all channels use amplification 500, (2) Mediate: [500, 500, 500, 500, 500, 1000, 1000, 1000, 1000, 1000, 2000, 2000, 2000,



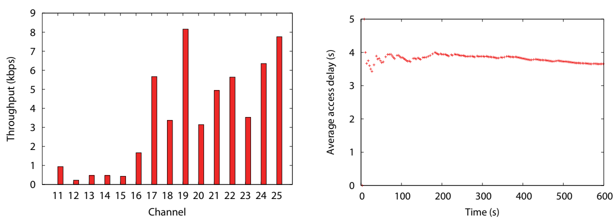
(a) SNR (b) PER (c) SNR-PER

Fig. 8. Impact on SNR and PER by different amplifications, and the relationship between SNR and PER.



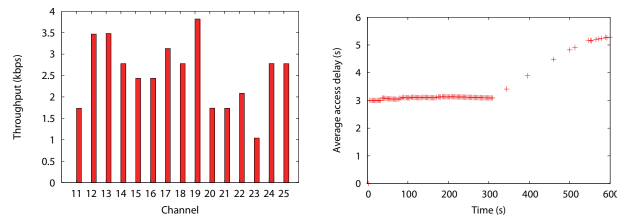
(a) throughput (b) delay

Fig. 9. Impact of amplification-controlled channel qualities on throughput and delay.



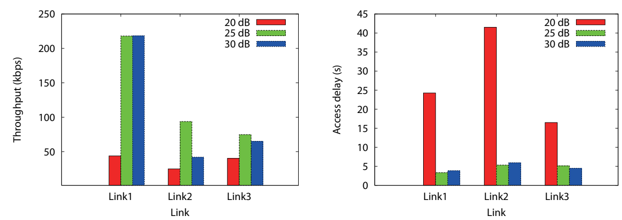
(a) throughput (b) delay

Fig. 10. Impact of the number of primary users on throughput and delay.



(a) throughput (b) delay

Fig. 11. Adaptivity of SPA scheme.



(a) throughput (b) delay

Fig. 12. Impact of different sensing thresholds.

2000, 2000]; (3) Good: all channels use amplification 2000. Fig. 9 shows our experimental results of the impact of different channel qualities on the throughput and delay. There is a significant throughput and delay dropoff for the case “Bad”, compared with cases “Good” and “Mediate”.

For the second approach, we adjust the number of primary users (sensor nodes in our experiment) in each of the 15 channels $11, 12, \cdots, 25$. Channel 26 is reserved as CCC. Fig. 10 shows the experimental results of the number of existing primary users on the throughput and delay. Here the number of primary users for each channel $\in [11, 25]$ is set as follows: [3,3,3,3,3,1,1,1,1,1,0,0,0,0,0]. We found that SUs still achieve good throughput and delay when there is one PU, while the throughput drops significantly in channels with 3 PUs.

Adaptivity of SPA system: We then study whether the SPA system can adjust its strategy when the channel qualities change dramatically after a certain time duration, such that the historical channel observation will have negative impact on the performance. We conduct the following experiment. In the first 300s, the amplification of each channel is as follows: [500,500,500, 500, 500, 1000, 1000, 1000, 1000, 1000, 2000, 2000, 2000, 2000, 2000]. After the system is run for 300s, we change the amplification of each channel as follows: [2000, 2000, 2000, 2000, 2000, 1000, 1000, 1000, 1000, 1000, 500, 500, 500, 500, 500]. In other words, good channels will become bad and bad channels will become good. Fig. 11 demonstrates that our SPA scheme can quickly adapt to the new environment.

Impact of channel status sensing threshold: We then study the impact of different channel status sensing thresholds on the final achievable throughput and delay. Recall that all previous experimental results assume that the energy threshold for channel status sensing is set as $25dB$. Fig. 12 shows that our choice of energy-detection thresholds is indeed the best for all links.

5 Related Work

There is a rich body of results for allocating spectrum channels. Li *et al.* [12], [22], [26] designed efficient and truthful mechanisms for various spectrum assignment problems. Zhou *et al.* [30] proposed a truthful and efficient spectrum auction system to serve many small players. Several results [17], [31] designed truthful double spectrum auctions with provable performance. All these results are based on offline models. Xu *et al.* [21], [23]–[25], [27], [28] studied online spectrum allocation and truthful mechanisms when secondary users could bid arbitrarily.

Many results have been developed for dynamic spectrum access in cognitive radio networks [29]. Huang *et al.* [10] presented a threshold-based sensing-transmission structure that is optimal under a technical constraint to maximize the SU's utility. Xu and Liu [20] proposed an optimal transmitting, sensing, and sleeping structure. However, all these results ignored different data rates across all channels and over a time period. Recently several results [18], [29] were presented by using the connections between channel access and the multi-armed bandits problem. Shu and Krunz [18] proposed a throughput-efficient sensing/probing/access scheme with sensing errors. The difference with our work is that they assume stochastic homogeneous channels while we consider non-stochastic channels. Gai *et al.* [8] used a combinatorial MAB formulation to address the multi-user channel allocation. Their method is centralized and assumes i.i.d. stochastic channels. Anandkumar *et al.* [1] designed distributed policy for learning and allocation and it achieves logarithmic growth of regret. It also assumes i.i.d. stochastic channels and cooperation among users in distributed implementation. Liu *et al.* [14], [15] presented a distributed learning in MAB problem with multiple players when the reward is i.i.d. from some distribution. The time and space requirement is exponential in the number of users. Proutiere *et al.* [16] presented a decentralized channel access protocol using learning approach. Recently, Tekin and Liu [19] modeled each channel as a restless Markov chain. They presented an algorithm using a sample-mean based index policy, and showed that under mild conditions this algorithm achieves logarithmic regret uniformly over time.

Our problem is related to classical MAB problem [6]. When the awards are i.i.d. according to an unknown law with unknown expectation, several protocols [11] can achieve the optimum logarithmic regret asymptotically. When the awards of actions are chosen by adversary, logarithmic regret cannot be achieved. For example, when the results of *all* N possible actions at each round are known, the regret per round can be bounded by $O(\sqrt{\ln N/n})$ [3]. When only the result of the action that the decision maker performs is known, the best possible regret per round is bounded by $\Theta(\sqrt{N\ln(Nn/\delta)/n})$ with probability at least $1-\delta$ [4]. Cesa-Bianchi *et al.* [5] shows that the regret per round is $\Theta(\sqrt{\ln N/m})$ where m queries are allowed during n rounds.

6 Conclusions

We proposed efficient channel sensing/accessing/probing methods for optimizing the throughput achieved by secondary users. Our methods have negligible regret even the channel data rates and channel availabilities are arbitrary. Our methods can be extended to deal with several settings not specifically discussed here. For example, when the idle probabilities of all channels are low, it is better for the secondary user not to always sense/probe/access channels. To reflect such choices, we can introduce a third action, *idle*, by a secondary user. Some interesting future work is to study the Nash equilibriums of the system when multiple users exist and design distributed protocol with optimal regret.

Acknowledgment

This research is partially supported by NSF CNS-1035894, NSF EECS-1247944, NSF EECS-1343306, NSF China under Grants No. 61232018, 61272487, 61170216, 61172062, 61228202, and 61273210.

References

[1] Anandkumar, A., Michael, N., and Tang, A. Opportunistic Spectrum Access with Multiple Users: Learning under Competition. In *IEEE INFOCOM* (2010).

[2] Arbaugh, W. Improving the latency of the probe phase during 802.11 handoff. online at www.umiacs.umd.edu/partnerships/ltsdocs/Arbaug_talk2.pdf.

[3] Auer, P., Cesa-Bianchi, N., Freund, Y., and Schapire, R. E. Gambling in a rigged casino: The adversarial multi-armed bandit problem. In *FOCS '95: Proceedings of the 36th Annual Symposium on Foundations of Computer Science* (Washington, DC, USA, 1995), IEEE Computer Society, p. 322.

[4] Auer, P., Cesa-Bianchi, N., Freund, Y., and Schapire, R. E. The nonstochastic multiarmed bandit problem. *SIAM J. Comput. 32*, 1 (2003), 48–77.

[5] Cesa-Bianchi, N., Lugosi, G., and Stoltz, G. Minimizing regret with label efficient prediction. *Information Theory, IEEE Transactions on 51*, 6 (june 2005), 2152 – 2162.

[6] Even-Dar, E., Mannor, S., and Mansour, Y. PAC Bounds for Multi-armed Bandit and Markov Decision Processes. In *15th Annual Conference on Computational Learning Theory: COLT 2002.* (2002), Springer, p. 255.

[7] Ferguson, T. S. Optimal stopping and applications. available on line at http://www.math.ucla.edu/∼tom/Stopping/Contents.html.

[8] Gai, Y., Krishnamachari, B., and Jain, R. Learning multiuser channel allocations in cognitive radio networks: A combinatorial multi-armed bandit formulation. In *IEEE DySPAN* (2010).

[9] Huang, S., Liu, X., and Ding, Z. Opportunistic spectrum access in cognitive radio networks. In *Proc. IEEE INFOCOM* (2008), vol. 13, pp. 1427–1435.

[10] Huang, S., Liu, X., and Ding, Z. Optimal sensing-transmission structure for dynamic spectrum access. In *IEEE INFOCOM* (2009).

[11] Lai, T., and Robbins, H. Asymptotically efficient adaptive allocation rules. *Advances in applied mathematics 6*, 1 (1985), 4–22.

[12] Li, X.-Y., Xu, P., Tang, S., and Chu, X. Spectrum bidding in wireless networks and related. In *COCOON* (2008), pp. 558–567.

[13] Liang, Y.-C., Zeng, Y., Peh, E., and Hoang, A. T. Sensing-throughput tradeoff for cognitive radio networks. pp. 5330 –5335.

[14] Liu, K., and Zhao, Q. Decentralized multi-armed bandit with multiple distributed players. In *Information Theory and Applications Workshop (ITA), 2010* (2010), IEEE, pp. 1–10.

[15] Liu, K., and Zhao, Q. Distributed Learning in Multi-Armed Bandit with Multiple Players. *Signal Processing, IEEE Transactions on*, 99 (2010), 1.

[16] Proutiere, A., and Kasbekar, G. Decentralized opportunistic medium access in multi-channel wireless systems: A learning approach. In *Allerton* (2010).

[17] ShiGuang Wang, Ping Xu, X. X. S.-J. T. X.-Y. L., and Liu, X. Oda: Truthful online double auction for spectrum allocation in wireless networks. *IEEE DySpan* (2010).

[18] Shu, T., and Krunz, M. Throughput-efficient sequential channel sensing and probing in cognitive radio networks under sensing errors. In *ACM MobiCom* (2009), pp. 37–48.

[19] Tekin, C., and Liu, M. Online learning in opportunistic spectrum access: a restless bandit approach. In *IEEE INFOCOM* (April 2011).

[20] Xu, D., and Liu, X. Opportunistic spectrum access in cognitive radio networks: when to turn off the spectrum sensors. In *Proceedings of the 4th Annual International Conference on Wireless Internet* (2008), ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering) ICST, Brussels, Belgium, Belgium.

[21] Xu, P., and Li, X. SOFA: Strategyproof Online Frequency Allocation for Multihop Wireless Networks. In *Algorithms and Computation: 20th International Symposium, Isaac 2009, Honolulu, Hawaii, USA, December 16-18, 2009. Proceedings* (2009), Springer-Verlag New York Inc, p. 311.

[22] Xu, P., and Li, X.-Y. Online market driven spectrum scheduling and auction. In *MOBICOM-CoRoNet* (2009), pp. 49–54.

[23] Xu, P., and Li, X.-Y. Online market driven spectrum scheduling and auction. In *ACM CoRoNet* (2009), pp. 49–54.

[24] Xu, P., and Li, X.-Y. Sofa: Strategyproof online frequency allocation for multihop wireless networks. In *ISAAC '09: Proceedings of the 20th International Symposium on Algorithms and Computation* (Berlin, Heidelberg, 2009), Springer-Verlag, pp. 311–320.

[25] Xu, P., and Li, X.-Y. Tofu: semi-truthful online frequency allocation mechanism for wireless network. *IEEE/ACM Trans. Netw. 19*, 2 (2011), 433–446.

[26] Xu, P., Li, X.-Y., Tang, S., and Zhao, J. Efficient and strategyproof spectrum allocations in multi-channel wireless networks. *IEEE Transactions on Computers* (Nov 2009).

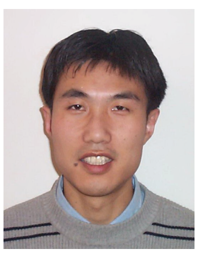
[27] Xu, P., Wang, S., and Li, X.-Y. Salsa: Strategyproof online spectrum admissions for wireless networks. *IEEE Trans. Computers 59*, 12 (2010), 1691–1702.

[28] Xu, P., Xu, X., Tang, S., and Li, X.-Y. Truthful online spectrum allocation and auction in multi-channel wireless networks. In *INFOCOM* (2011), pp. 26–30.

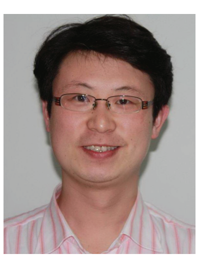
[29] Zhou, Y, Huang, Q., Li, F., Li, X.-Y., Liu, M., Li, Z., and Yin, Z. Almost Optimal Channel Access in Multi-Hop Networks With Unknown Channel Variables, In *IEEE ICDCS* (2014).

[30] Zhou, X., Gandhi, S., Suri, S., and Zheng, H. eBay in the Sky: strategy-proof wireless spectrum auctions. In *Proceedings of the 14th ACM international conference on Mobile computing and networking* (2008), ACM New York, NY, USA, pp. 2–13.

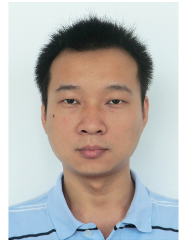
[31] Zhou, X., and Zheng, H. Trust: A general framework for truthful double spectrum auctions. In *IEEE INFOCOM* (2009).



Xiang-Yang Li (M'99-SM'08-F'15) is a professor at the Illinois Institute of Technology. He is an IEEE Fellow and an ACM Distinguished Scientist. He holds EMC-Endowed Visiting Chair Professorship at Tsinghua University. He is a recipient of China NSF Outstanding Overseas Young Researcher (B). Dr. Li received MS (2000) and PhD (2001) degree at Department of Computer Science from University of Illinois at Urbana-Champaign, a Bachelor degree at Department of Computer Science and a Bachelor degree at Department of Business Management from Tsinghua University, P.R. China, both in 1995. His research interests include wireless networking, mobile computing, security and privacy, cyber physical systems, social networking, and algorithms. He and his students won four best paper awards, one best demo award and was nominated for best paper awards twice (ACM MobiCom 2005, 2008). He published a monograph "Wireless Ad Hoc and Sensor Networks: Theory and Applications".



Panlong Yang (M'02) received his B.S. degree, M.S. degree, and Ph.D. degree in communication and information system from Nanjing Institute of Communication Engineering, China, in 1999, 2002, and 2005 respectively. During September 2010 to September 2011, he was a visiting scholar in HKUST. Dr. Yang is now an associate professor in the Nanjing Institute of Communication Engineering, PLA University of Science and Technology. His research interests include wireless mesh networks, wireless sensor networks and cognitive radio networks. Dr. Yang has published more than 50 papers in peer-reviewed journals and refereed conference proceedings in the areas of mobile ad hoc networks, wireless mesh networks and wireless sensor networks. He has also served as a member of program committees for several international conferences. He is a member of the IEEE Computer Society and ACM SIGMOBILE Society.



Yubo Yan (S'10) received the B.S. degree and M.S. degree in communication and information system from the College of Communications Engineering, PLA University of Science and Technology, China, in 2006 and 2011 respectively. He is currently working towards the Ph.D. degree at the PLA University of Science and Technology. His current research interests include cognitive radio networks, software radio systems and wireless sensor networks. He is a student member of the IEEE and the IEEE Computer Society.