This full text paper was peer reviewed at the direction of IEEE Communications Society subject matter experts for publication in the WCNC 2010 proceedings.

Sensing and Communication Tradeoff for Cognitive Access of Continues-Time Markov Channels

Xin Li, Qianchuan Zhao, Xiaohong Guan
Center for Intelligent and Networked System
Department of Automation and TNLIST
Tsinghua University, Beijing 100084, China
{lxin05@mails,zhaoqc@,xhguan@}tsinghua.edu.cn

Lang Tong
School of Electrical and Computer Engineering
Cornell University, Ithaca
NY 14853, USA
Email: ltong@ece.cornell.edu

***Abstract*—Dynamic spectrum access (DSA) aims to improve spectrum efficiency via spectrum sensing and optimal spectrum access. An essential component in DSA is the joint design of sensing and access strategies. This paper focuses on dynamic spectrum access in the time domain. To maximize channel utilization while limiting interference to primary users, a framework of linear programming is presented based on the stationary distribution of the primary user channels. It is shown that the optimal tradeoff between sensing and transmitting is achieved with required limit on the interference to the primary users.**

***Index Terms*—Cognitive radio, Dynamic spectrum access, Multichannel sensing, Markov decision processes.**

I. Introduction

Dynamic spectrum access (DSA) is a technique for increasing radio spectrum utilization. In the context of cognitive radio communications, DSA enables the so-called secondary users (SUs) to transmit in channels assigned to the primary users (PUs) while guarantees that the PUs are not interferred beyond the prescribed level. A survey of related work can be found in [1].

We consider an overlay cognitive network in which the secondary user exploits temporal opportunity through the design of sensing and transmission policies. One of the first such DSA protocols was presented in [2] based on the framework of Partially Observable Markov Decision Processes (POMDP) in which a SU decides which channel to sense based on the "belief" of the network state. In [3], a periodic sensing technique is introduced for continuous-time traffic models. While restricting to periodic sensing is suboptimal, such sensing schemes lead to a reformulation of DSA as a problem of constrained Markov Decision Processes (MDP). More recent work on optimal DSA using dynamic programming techniques can be found in [3], [4], [5], [6], [7], [8] and references therein.

Challenges of spectrum sensing in cognitive radio networks are widely recognized. See [9] for a survey. A typical model used in spectrum sensing for cognitive radio is one of Gaussian signal in Gaussian noise, and the optimal detector and its performance are well known [10]. A practical difficulty, however, is that hierarchical cognitive radio networks often impose tight constraints on detection error performance. To this end, a number of sophisticated collaborative sensing techniques have been proposed [11], [12], [13]. Furthermore, of particular importance in practice is model uncertainties. In [14], authors consider the problem of spectrum sensing under model uncertainties. They introduce the notion of "SNR wall" to capture the fundamental limits on detection error probabilities when the signal and noise models are not known precisely.

The accuracy of channel sensing can affect significantly the performance of DSA. Channel sensing carries a cost; a longer sensing period provides better sensing accuracy at the expense of transmission time. Thus there is a fundamental tradeoff between the cost of sensing and the rate of communications. The tradeoff between sensing and rate has been studied in [15], [16], [17] for the single channel scenario. In [15], optimal tradeoff between throughput and sensing cost is solved using an MDP formulation whereas in [16], [17], the authors considered a one stage formulation where the throughput of SU is maximized under the constraint that detection probability should be no less than a given level. For the multiple channel case, analytical characterization of the tradeoff between channel sensing time and DSA throughput is more challenging; only limited numerical results have been reported. See, e.g., [7] where a numerical example is used to demonstrate that there is an optimal sensing time for DSA. In [18], the authors consider the tradeoffs between channel probing and communications in the context of multichannel opportunistic spectrum access. Using dynamic programming techniques, the authors presented strategies for selecting channels to sense and channels for transmission. The main difference between [18] and the current work is that the channels considered in [18] are temporally independent whereas the channel model in this paper is Markovian. In addition, we model sensing errors explicitly in this paper. Under the same model, the problem of sensing and communication tradeoff can also be formulated as a multi-armed bandit problem [19] where channels for primary users may have unknown parameters, and the problem becomes one of learning PU channels while capturing transmission opportunities [20], [21], [22]. These formulations do not take into account sensor errors.

This paper extends the work [3] where the framework of periodic sensing is adopted for multichannel DSA under the continuous-time Markovian traffic model. Representing the sensing cost, the duration of sensing is determined for the optimal channel utilization. The optimal stationary sensing and transmitting strategy is derived by solving a linear programming problem. Performance is evaluated numerically.

The rest of the paper is organized as follows. Section II describes the primary and the secondary system. Section III introduces the sensing model for the SUs. Section IV gives the Markov property of the channel and the estimation criterion. The problem formulation is presented and the optimal stationary sensing policy is derived in Section V. The numerical testing results for performance evaluation are shown in Section VI. A short discussion on the results and concluding remarks then follow.

II. Primary and Secondary Systems

In this section, we will describe the primary and the secondary systems. The PUs are the licensed users who are authorized to access N parallel PU channels that evolve independently in time. The SUs, on the other hand, seek transmission opportunities to access channels that are not in use by the PUs. We model each channel by a two-state continuous time Markov process with the i-th channel state $l_i(t) = 1$ indicating that there is the primary user transmitting in channel i at time t, while the idle state is denoted by $l_i(t) = 0$. By the Markovian assumption, the holding time lengths in the two states are exponentially distributed with the parameters λ_i(idle) and μ_i(busy) respectively.

The SUs aim at maximizing the overall channel utilization while limiting their interference to the PUs. We indicate the interference as the case that an SU collides with a PU, an unavoidable scenario for the continuous-time Markovian channels. Assume that an SU uses a slotted sensing and transmission structure (see in Fig.1) with the time slot size T. The SU employs a round-robin periodic sensing strategy, sensing each channel in turn at the beginning of the time slot for a duration time αT, where $\alpha \in (0, 1)$. At the end of a sensing in a time slot, according to the sensing result, the SU can either transmit data during the rest of the time slot or not transmit and wait until the next time slot sensing.

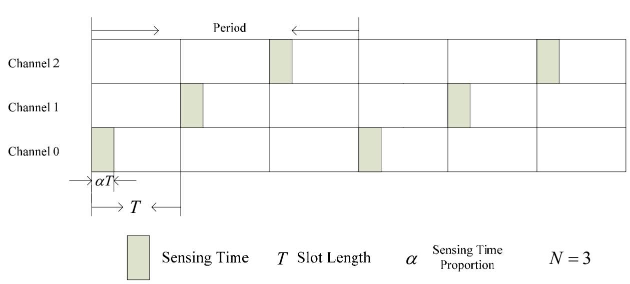

Fig. 1. Periodic Sensing Protocol

III. Sensing Model

We assume that the secondary system employs a single channel spectrum sensor. At the beginning of each time slot, it senses the presence of a PU in a particular channel. As mentioned in Section II and shown in Fig. 1, the spectrum sensor senses the channel for the time duration of αT. It is also assumed that, for the ease of analysis, T is sufficiently small relative to the mean idle and busy holding time of the PU so that no state transition of the channel occurs during the sensing period.

We formulate the sensing process as a hypothesis testing between the null-hypothesis H_0 of observing only noise and the alternative hypothesis H_1 of observing a signal in noise. Specifically, the discrete-time model for the sensor measurements is given by

$$H_0 : y(n) = u(n) \quad \text{vs} \quad H_1 : y(n) = s(n) + u(n)$$

where the noise $u(n)$ is complex-valued Gaussian, independent and identically distributed (i.i.d) sequence with mean zero and variance σ_u^2; the PU's signal $s(n)$, independent of $u(n)$, is also assumed a complex-valued Gaussian, i.i.d sequence with mean zero and variance σ_s^2. For convenient expression, we denote this relationship as $\rho = \sigma_s^2/\sigma_u^2$ which is used in many scenarios and means the received signal-to-noise ratio (SNR). Usually it can also be represented in the logarithmic form as $\rho_{\text{dB}} = 10\log(\sigma_s^2/\sigma_u^2)$ if necessary.

The optimal test for the above binary hypotheses is an energy detector using sufficient statistic:

$$T(y) = \frac{1}{N_s}\sum_{n=1}^{N_s} |y(n)|^2 \underset{H_0}{\overset{H_1}{\gtrless}} \varepsilon \tag{1}$$

where ε is the decision threshold designed for specific target detection performance and N_s is the total number of samples. It is known that the test statistic $T(y)$ is χ^2 distributed with $2N_s$ degrees of freedom and is also influenced by the relationship between σ_u and σ_s.

This strategy brings the problem of sensing accuracy, and therefore two kinds of error probabilities are of interest. Let D_0 and D_1 be the idle and busy sensing decision respectively. We then have the false alarm and miss detection probability, respectively, given by

$$P_f \triangleq \Pr(D_1|H_0) \quad \text{vs} \quad P_m \triangleq \Pr(D_0|H_1). \tag{2}$$

Given α and ρ, the false alarm probability P_f and miss detection probability P_m can be expressed as follows [10]:

$$P_f = \Pr(T(y) > \varepsilon|H_0) = 1 - \Gamma(N_s, \frac{\varepsilon}{\sigma_u^2}), \tag{3}$$

$$P_m = \Pr(T(y) \leq \varepsilon|H_1) = \Gamma(N_s, \frac{\varepsilon}{\sigma_u^2 + \sigma_s^2}), \tag{4}$$

where $\Gamma(x;t) \triangleq \int_0^t e^{-y}y^{x-1}dy/\Gamma(x)$ is the *incomplete gamma function* and $\Gamma(x) = \int_0^\infty e^{-y}y^{x-1}dy$ is the *gamma function*. Further more, ε can be expressed as:

$$\varepsilon = \sigma_u^2\Gamma^{-1}(N_s; 1 - P_f), \tag{5}$$

where $\Gamma^{-1}(x;\cdot)$ is the inverse function of $\Gamma(x;\cdot)$ in its second variable.

Therefore P_m can be represented by P_f

$$P_m = \Gamma(N_s; \frac{1}{1+\rho}\Gamma^{-1}(N_s; 1-P_f)). \quad (6)$$

Denote $\eta = P_f$ for convenience. It can be seen that the "miss detection" probability P_m is a variable of the parameters α, ρ and η. For the reason that the three parameters often appear together as a combination, we denote $\Theta \stackrel{\Delta}{=} (\alpha, \rho, \eta)$ as the representation of such combination for simplified expression.

IV. Markov Channel Model

We limit the sensing capability of the SU to sensing one channel in each time slot, which makes it challenging to find an optimal access policy since the state of the channel is only partially observable. The suboptimal periodic sensing scheme converts the problem to a fully observable Markov Decision process, which simplifies the problem considerably [3].

Under the periodic sensing framework, channels are sensed periodically in an increasing order. At the beginning of each time slot, the detector makes a decision on the state of the channel, either it is busy or idle. Let $Z(k) = [z_0(k), z_1(k), \ldots, z_{N-1}(k)]$ denote the vector that contains the sensing outcomes in past N slots with $z_i(k) = 0$ indicating that the most recently sensing outcome of channel i is idle. Specifically, in slot k, let $q = k \bmod N$ be the channel sensed in the current slot. We then update $Z(k)$ from $Z(k-1)$ by replacing the q-th entry of $Z(k-1)$ with the sensing outcome in the current slot. Similarly, as shown in [3], when $\Theta = (\alpha, \rho, \eta)$ keeps constant, the existence of the stationary distribution $f(Z; \theta)$ is guaranteed.

To find the mathematical expression of $f_q(Z, \Theta)$, we should first obtain the probability of the observations in each channel. For channel i, let η_i and ξ_i denote the false alarm probability and the miss detection probability respectively. We then have

$$\Pr(z_i = 0) = \frac{\mu_i}{\lambda_i + \mu_i}(1-\eta_i) + \frac{\lambda_i}{\lambda_i + \mu_i}\xi_i \quad (7)$$

and

$$\Pr(z_i = 1) = \frac{\mu_i}{\lambda_i + \mu_i}\eta_i + \frac{\lambda_i}{\lambda_i + \mu_i}(1-\xi_i), \quad (8)$$

where we have used the stationary distribution of the i-th channel state.

Since the PU's activities in different channels are independent, the stationary probability of the observations has a product form:

$$f(Z; \Theta) = \prod_{i=0}^{N-1} (1_{[z_i=0]} \cdot \Pr(z_i = 0) + (1_{[z_i=1]} \cdot \Pr(z_i = 1)), \quad (9)$$

where $1_{[\cdot]}$ denotes the indicator function.

V. Optimal Stationary Access Policy

In this section, we will focus only on stationary randomized channel access policies of the SU. That means the SU needs to decide whether to transmit and the transmission probability based on the value of the latest observation $Z(k)$.

A. Actions and Rewards

Specifically, after sensing operation, the SU will either choose one of the N channels to transmit, or alternatively not transmit at all. Denote the action in the k-th time slot as $a_k \in \Lambda = \{-1, 0, \ldots, N-1\}$, where $a_k \geq 0$ means that the SU chooses to transmit in the a_k-th channel, or otherwise $a_k = -1$ means no transmission at all.

When choosing to transmit, the SU can get a reward or pay a price when it collides with a PU. Note that even if the actual channel state is idle when the SU transmits, it cannot be promised that there will be no collisions during the transmission since the PU's access is not time slotted.

Define the reward $r(Z, k, i; \Theta)$ obtained from a successful transmission in k-th time slot, with sensing result Z and actions $a_k = i$ as

$$r(Z, k, i; \Theta) = \begin{cases} (1-\alpha)g(Z, k, i; \Theta), & 0 \leq i \leq N-1; \\ 0, & i = -1. \end{cases} \quad (10)$$

where α is the fraction of time used for sensing

$$g(Z, k, i; \Theta) = \Pr(l_i(t) = 0, t \in I_k^{1-\alpha} | z_i(k) = z_i),$$

and $z_i \in \{0, 1\}$. Here $I_k^{1-\alpha} = [kT + \alpha T, (k+1)T]$ is defined as the transmission time segment in the k-th time slot.

Then we define $c(Z, k, i; \Theta)$ as the probability of the collision with the PU when the SU chooses to transmit as follows

$$c(Z, k, i; \Theta) = \begin{cases} 1 - g(Z, k, i; \Theta), & 0 \leq i \leq N-1; \\ 0, & i = -1. \end{cases} \quad (11)$$

Before evaluating the performance of the secondary system, we need mathematical expression of the rewards which is equivalent to get the expression of $g(Z, k, i; \Theta)$.

From the property of the Markov process, without difficulty, we can get

$$g(Z, k, i; \Theta) = \mathrm{Prob}_1 \cdot \mathrm{Prob}_2, \quad (12)$$

where $\mathrm{Prob}_1 = \Pr(l_i(t) = 0, t \in I_k^{1-\alpha} | l_i(kT + \alpha T) = 0)$ and $\mathrm{Prob}_2 = \Pr(l_i(kT + \alpha T) = 0 | z_i(\varphi(i,k)) = z_i)$. Here $\varphi(i,k)$ is the time slot index when channel i is last sensed before the k-th slot. This is a product of two items. The first one is immediately obtained because of the property of the continuous time Markov chain,

$$\begin{aligned} \mathrm{Prob}_1 &= \Pr(l_i(t) = 0, t \in I_k^{1-\alpha} | l_i(kT + \alpha T) = 0) \\ &= e^{-\lambda_i(1-\alpha)T}. \end{aligned} \quad (13)$$

The second one can also be obtained according to the property of continuous time Markov chain, and its mathematical expression is a little complex:

$$\begin{aligned} \mathrm{Prob}_2 &= \Pr(l_i(kT + \alpha T) = 0 | z_i(\varphi(i,k) = z_i)) \\ &= 1_{z_i(\varphi(i,k)=0)}(\Pr(l_i = 0 | z_i = 0)Q_{00}(k,i) \\ &\quad + \Pr(l_i = 1 | z_i = 0)Q_{10}(k,i)) \\ &\quad + 1_{z_i(\varphi(i,k)=1)}(\Pr(l_i = 0 | z_i = 1)Q_{00}(k,i) \\ &\quad + \Pr(l_i = 1 | z_i = 1)Q_{10}(k,i)), \end{aligned} \quad (14)$$

where

$$\begin{aligned} Q_{l_i 0}(k,i) &= (\Pr(l_i=0) + (1_{[l_i=0]}\Pr(l_i=1) \\ &\quad -1_{[l_i=1]}\Pr(l_i=0))) \\ &\quad \cdot \exp(-(\lambda_i+\mu_i)((N+k-i) \bmod N)T) \end{aligned}$$

is the transition probability of the channel i from state $l_i(\varphi(i,k)T+\alpha T)=l_i$ to state $l_i(kT+\alpha T)=0$.

Therefore the final result is

$$\begin{aligned} g(Z,k,i;\Theta) &= e^{-\lambda_i(1-\alpha)T} \\ &\quad \cdot(1_{[z_i(\varphi(i,k))=0]}(\Pr(l_i=0|z_i=0)Q_{00}(k,i) \\ &\quad +\Pr(l_i=1|z_i=0)Q_{10}(k,i)) \\ &\quad +1_{[z_i(\varphi(i,k))=1]}(\Pr(l_i=0|z_i=1)Q_{00}(k,i) \\ &\quad +\Pr(l_i=1|z_i=1)Q_{10}(k,i))). \end{aligned} \tag{15}$$

B. Policy Optimization Problem

Our aim is to maximize the throughput of the secondary system while abiding by hard constraints on the interference. Specifically we can formulate the problem as maximizing the average reward of successful transmissions of the SU

$$J(\pi;\Theta) = \lim_{K\to\infty}\frac{1}{K}\sum_{k=1}^{K}E_\pi(r(Z(k),k,a_k;\Theta)), \tag{16}$$

where the expectation is taken over the access probability distribution of policy π. At the meantime, the SU should abide by the constraints on the interference to the PUs. For the PU in channel i, define the asymptotic ratio of the collision and transmission number of the slots as a measurement for the degree of the interference due to the presence of the PU. In particular,

$$\begin{aligned} &C_i(\pi;\Theta) = \\ &\lim_{K\to\infty}\frac{\sum_{k=1}^{K}E[c(Z(k),i,k;\Theta)\Pr(a_k=i|\pi,Z(k),k)]}{E(B_i(K))}, \end{aligned} \tag{17}$$

where $B_i(K)$ is the total number of slots occupied by PU in channel i up to time kT, and $\Pr(a_k=i|\pi,Z(k),k)$ is the probability that channel i is chosen by policy π, given the sensing result $Z(k)$.

Mathematically the optimization problem is thus,

$$\max_{\pi\in\mathcal{P}} J(\pi;\Theta) \tag{18}$$

subject to

$$C_i(\pi;\Theta)\le\gamma_i, \quad \forall i\in\{0,\dots,N-1\},$$

where $0\le\gamma_i\le 1$ and Θ are the given constraints and the given parameters respectively. Here $\mathcal{P}$ is the set of stationary randomized polices.

C. Linear Program Solution

To find the optimal stationary policy to the optimization problem (18) formulated in V-B, we first consider a fixed optimal stationary policy π^*. When we classify the transmission according to the position q in a period, denote $\beta_{q,a}^{\pi^*,P}(Z)\in[0,1]$ as the frequency of action $a\in\Lambda$ chosen by policy π^* in the time slot I_{pN+q} when the observation $Z(pN+q)$ is in a sample path with $p=0,1,2,\dots,P$. Specifically,

$$\beta_{q,a}^{\pi^*,P}(Z) = P^{-1}\sum_{p=1}^{P}\Pr(a_{pN+q}=a|Z(pN+q)=Z,\pi^*). \tag{19}$$

Since the optimal policy π^* is stationary, the frequency $\lim_{P\to\infty}\beta_{q,a}^{\pi^*,P}(Z)$ of the state-action pair (Z,q,a) exists. Let us denote collectively these frequencies as

$$\beta^{\pi^*} = (\beta_{q,a}^{\pi^*}(Z), a\in\Lambda, q\in\{0,1,\dots,N-1\}, Z\in V), \tag{20}$$

where β^{π^*} represents the policy π^* which the SU applies specifically. When the SU needs to make a decision to choose which channel to transmit, it is just like flipping a biased coin with probability $\beta_{q,i}^{\pi^*}(Z)$ to transmit in channel $i\ge 0$. View them as the decision variables, and we will try to figure out the optimal policy that maximizes the throughput.

We know that a reward $r(Z,q,i;\Theta)$ can be obtained when the SU transmits in channel i, in time slot I_{pN+q}, given the observation Z where $p=0,1,\dots,P$. Because Z is a stochastic process, considering the possibilities the observation Z occurs and using the stationary policy π, we can get the expected reward at one time (assuming this time slot index is q)

$$E_\pi(r(Z(q)),q,i;\Theta) = \sum_{Z\in V}\sum_{i=0}^{N-1}R(Z,q,i;\Theta)\beta_{q,i}(Z;\Theta), \tag{21}$$

where $R(Z,q,i;\Theta)=r(Z,q,i;\Theta)f_q(Z;\Theta)$.

Therefore the average throughput over a long enough time can be obtained by summing over q in only one period time length NT, and it is:

$$J(\pi;\Theta) = N^{-1}\sum_{q=0}^{N-1}\sum_{Z\in V}\sum_{i=0}^{N-1}R(Z,q,i;\Theta)\beta_{q,i}(Z;\Theta), \tag{22}$$

where the decision variable $\beta_{q,i}(Z;\Theta)$ is also a function of the parameters Θ just as expressed in (22) since different Θ can lead to different accessing policies.

As to the constraints, similarly, the expectation probability of a failed transmission in channel i in the k-th time slot under the stationary policy π is

$$\begin{aligned} &E[c(Z(k),i,k;\Theta)\Pr(a_k=i|\pi,Z(k),k)] \\ &= \textstyle\sum_{Z\in V}c(Z,q,i;\Theta)f_q(Z;\Theta)\beta_{q,i}(Z;\Theta). \end{aligned} \tag{23}$$

The busy probability in time slot I_{pN+q} in channel i is given by $1-\Pr(l_i=0)e^{-\lambda_i T}$. Because of the same reason that we use the stationary policy, the average cost over a long time length can also be represented in a period time NT. Then, the mean collision rate in channel i under the policy π is given by

$$C_i(\pi;\Theta) = \frac{\sum_{q=0}^{N-1}\sum_{Z\in V}c(Z,q,i;\Theta)f_q(Z;\Theta)\beta_{q,i}(Z;\Theta)}{N(1-\Pr(l_i=0)e^{-\lambda_i T})}. \tag{24}$$

When we denote $Co(Z,q,i;\Theta) = \frac{c(Z,q,i;\Theta)f_q(Z;\Theta)}{1-\Pr(l_i=0)e^{-\lambda_i T}}$ for simplicity, the collision constraint in each channel can be represented as:

$$C_i(\pi;\Theta) = N^{-1}\sum_{q=0}^{N-1}\sum_{Z\in V} Co(Z,q,i;\Theta)\beta_{q,i}(Z;\Theta) \le \gamma_i. \tag{25}$$

We also note that the decision variables have their own constraints:

$$\sum_{i\in\Lambda, i\neq -1} \beta_{q,i}(Z;\Theta) \le 1, \forall q, Z, \Theta,\ \beta_{q,i}(Z;\Theta)\in[0,1], \forall q, Z, \Theta. \tag{26}$$

Finally we come to the linear programming problem:

$$\max_{\beta} N^{-1}\sum_{q=0}^{N-1}\sum_{Z\in V}\sum_{i=0}^{N-1} R(Z,q,i;\Theta)\beta_{q,i}(Z;\Theta) \tag{27}$$

s.t. $N^{-1}\sum_{q=0}^{N-1}\sum_{Z\in V} Co(Z,q,i;\Theta)\beta_{q,i}(Z;\Theta) \le \gamma_i$,

$$\sum_{i=0}^{N-1}\beta_{q,i}(Z;\Theta) \le 1, \forall q, Z, \Theta,\ \beta_{q,i}(Z;\Theta)\in[0,1], \forall q, Z, \Theta.$$

The linear programming problem (27) provides an equivalent form of policy optimization problem (18) as long as we explain each solution β as a parameterization representation of a stationary randomized policy.

Note that (27) is a parameterized linear programming problem, and the optimal policy can also be viewed as a function of the parameters $\Theta = (\alpha, \eta, \rho)$. Denote the optimal objective as $J_{\rm opt}(\Theta)$. Our aim is to find out the optimal combination of the parameters to get the maximum throughput. When Θ are all determined, it will degenerate to a linear program without the variation of the parameters. This means that when the wireless environment, the sensing time and the sensing decision rules are given, SU can figure out the optimal access policy to get the maximum throughput $J_{\rm opt}(\Theta)$. If just part of Θ is determined, such as ρ specifically, we should also find out the optimal combination of the other parameters α and η to get the maximum throughput among all the possible values of $J_{\rm opt}(\Theta)$ when given ρ, which is actually the joint design of optimal sensing and access strategies.

VI. Numerical Results

In this section, we evaluate the performance of the optimal policy numerically. We focus on the case of $N = 6$ independent channels, and each channel has the identical parameter $\lambda_i = \lambda$ and $\mu_i = \mu$, $i = 0, 1, \cdots, 5$. The values $\lambda^{-1} = 4.2$ ms and $\mu^{-1} = 1$ms are motivated from the actual applications [5]. The time slot T is set 0.25ms and the sample time interval is $0.25\mu s$. As to the constraint of the model, we assume $\gamma_i = 0.03, \forall i = 0, \cdots, 5$.

After the setting of these parameters, we want to show the relationship between the optimal throughput and the parameters. Specifically, we range η from 0 to 1 at each combination of α and ρ to choose the optimal η which leads to the maximum throughput at that α and ρ combination. When we figure out the maximum throughputs of all the parameter combinations, the relationship between the optimal throughput and the parameters can be found. The corresponding results are clearly shown in following figures Fig. 2, Fig. 3, and Fig. 4.

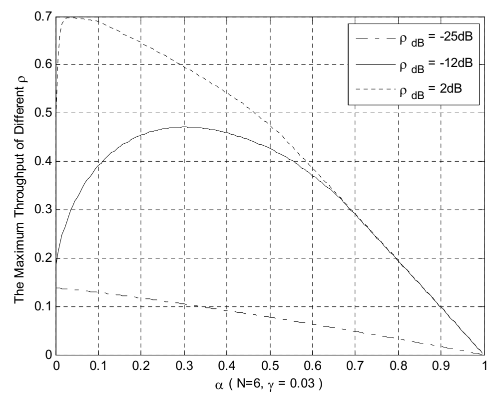

Fig. 2. The maximum throughput vs. sensing proportion α

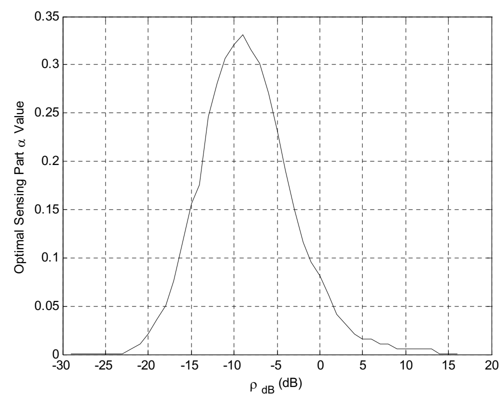

Fig. 3. The optimal sensing part with ρ

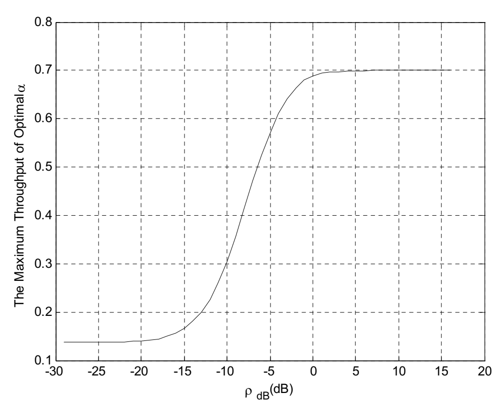

Fig. 4. The maximum throughput with ρ

Note that, in Fig. 2, when ρ is very small as $\rho_{\rm dB} = -25{\rm dB}$, which means bad wireless environment and certainly imprecise sensing, the maximum throughput decreases all the time along with the increasing of α. The reason is that we can't sense the channels accurately, the time used for sensing would have been better utilized for transmission. But when sensing is improved as $\rho_{\mathrm{dB}} = -12\mathrm{dB}$, spending an appropriate amount of time for sensing improves the overall throughput. There exists an optimal sensing time for SU to achieve the maximum throughput, which stimulates us to find out it. When ρ is much better as $\rho_{\mathrm{dB}} = 2\mathrm{dB}$, implying that even if a very short sensing time, we can still sense the channel accurate enough to get the maximum throughput. In Fig. 3, it shows how ρ affects the optimal sensing time αT during striving to obtain the maximum throughput. In Fig. 4, as ρ increases, equivalent to the improving of the wireless environment, the maximum throughput increases monotonically which is not surprising.

Different parameters bring different optimal accessing policies, and furthermore different maximum throughputs. It makes sense, when the wireless environment is given, how we will try to figure out the optimal combination of the sensing time and sensing decision rules to get the optimal accessing policy to achieve the goal of making maximum use of the channels. Longer sensing time means more accuracy but relatively less transmission time, and therefore there exists such a tradeoff between the sensing accuracy and the transmission opportunity.

VII. Conclusion

We have considered the problem of the sharing spectrum in the time domain by exploiting idle periods between burst transmissions of the PUs. Introducing sensing time and sensing noise model to the periodic sensing scheme makes the system more practical. Aiming to make maximum use of the channel, we formulate the policy optimization problem and find the optimal stationary randomized control policy by solving a parameterized linear programming. After evaluating the method's performance numerically, our results show that there exists such a trade off between the sensing accuracy and the transmission opportunity.

Acknowledgment

This work was supported by NSFC Grant (60574067, 60736027, 60721003) and by the Programme of Introducing Talents of Discipline to Universities (the National 111 International Collaboration Project, B06002) and the US Army Research Office MURI Program under award W911NF-08-1-0238.

References

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

[2] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework," *IEEE Journal on Selected Areas in Communications (JSAC): Special Issue on Adaptive, Spectrum Agile and Cognitive Wireless Networks*, April, 2007, vol. 25, no. 3, pp. 589-600.

[3] Qianchuan Zhao, S. Geirhofer, L. Tong, and B.M. Sadler, "Opportunistic Spectrum Access via periodic Channel Sensing", *IEEE Trans. on Signal Processing*. Jan. 2008

[4] S. Geirhofer, L. Tong, and B. M. Sadler, "Dynamic Spectrum Access in the Time Domain: Modeling and Exploiting Whitespace," *IEEE Communications Magazine*, Vol. 45, No. 5, May 2007.

[5] S. Geirhofer, L. Tong, and B. M. Sadler, "Cognitive Medium Access: Constraining Interference Based on Experimental Models," *IEEE Journal on Selected Areas in Communication*, Vol. 26, No. 1, Jan. 2008.

[6] Y. Chen, Q. Zhao, and A. Swami, "Joint Design and Separation Principle for Opportunistic Spectrum Access," in *Proc. of IEEE Asilomar Conference on Signals, Systems, and Computers*, November, 2006

[7] Y. Chen, Q. Zhao, and A. Swami, "Joint Design and Separation Principle for Opportunistic Spectrum Access in the Presence of Sensing Errors" *IEEE Transactions on Information Theory*, vol. 54, no. 5, pp. 2053-2071, May, 2008.

[8] Q. Zhao, B. Krishnamachari, and K. Liu, "On Myopic Sensing for Multi-Channel Opportunistic Access: Structure, Optimality, and Performance," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 5431-5440, December, 2008.

[9] A. Ghasemi and E. S Sousa, "Spectrum sensing in cognitive radio networks: requirements, challenges and design trade-offs," *IEEE Communications Magazine*, Volume 46, Issue 4, pp. 32-39, April, 2008

[10] H. Poor, *An Introduction to Signal Detection and Estimation, 2Ed*. New York: Springer-Verlag, 1994.

[11] G. Ganesan and Y. Li, Cooperative spectrum sensing in cognitive radio networks, in *Proc. IEEE Dyspan*, pp.137-143, Nov, 2005

[12] Won-Yeol Lee and Akyildiz I.F., "Optimal spectrum sensing framework for cognitive radio networks," *IEEE Transactions on Wireless Communications*, Volume 7, Issue 10, pp. 3845-3857, Oct, 2008

[13] Zhi Quan, Shuguang Cui, and Sayed, A.H., "Optimal Linear Cooperation for Spectrum Sensing in Cognitive Radio Networks," *IEEE Journal of Selected Topics in Signal Processing*, Volume 2, Issue 1, pp. 28-40 , Feb. 2008

[14] Rahul Tandra and Anant Sahai, "SNR Walls for signal detection," *IEEE Journal on Special Topics in Signal Processing*, Volume 2, Issue 1, pp 4-17, Feb, 2008

[15] A. Laourine and L. Tong, "Betting on Gilbert-Elliot Channels," to appear in *IEEE Transactions on Wireless Communications*.

[16] Y. C. Liang, Y. Zeng, E. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Trans. Commun.*, vol. 7, no. 4, pp. 1326-1337, Apr. 2008.

[17] Senhua Huang, Xin Liu, and Zhi Ding, "Optimal Sensing-Transmission Structure for Dynamic Spectrum Access," in *INFOCOM 2009*, pp.2295-2303,April 2009

[18] Chang, N. B. and Liu, M., "Optimal channel probing and transmission scheduling for opportunistic spectrum access," in *ACM MobiCom 2007*, pp.27-38., 2007

[19] T. Lai and H. Robbins, "Asymptotically efficient adaptive allocation rules," *Advances in Applied Mathematics*, vol. 6, no. 1, pp. 4C22, 1985.

[20] A. Anandkumar, N. Michael, and A.K. Tang, "Opportunistic Spectrum Access with Multiple Players: Learning under Competition," to appear in *Proc. of IEEE INFOCOM*, Mar. 2010.

[21] Hai Jiang, Lifeng Lai, Rongfei Fan, and Poor H.V., "Optimal selection of channel sensing order in cognitive radio," *IEEE Transactions on Wireless Communications*, Volume 8, Issue 1, pp.297-307, Jan, 2009

[22] K. Liu and Q. Zhao, "Distributed Learning in Cognitive Radio Networks: Multi-Armed Bandit with Distributed Multiple Players," submitted to *2010 IEEE International Conference on Acoustics , Speech, and Signal Processing (ICASSP)*. See http://arxiv.org/PS_cache/arxiv/pdf/0910/0910.2065v2.pdf