Opportunistic Spectrum Access in Cognitive Radio Networks: When to Turn off the Spectrum Sensors[*]

[Invited Paper]

Dan Xu and Xin Liu
Department of Computer Science
UNiversity of California, Davis
{danxu, xinliu}@ucdavis.edu

ABSTRACT

In cognitive radio networks, spectrum sensing is critical to both protecting the primary users and creating spectrum access opportunities of secondary users. Channel sensing itself, including active probing and passive listening, often incurs cost, in terms of time overhead, energy consumption, or intrusion to primary users. It is thus not desirable to sense the channel arbitrarily. In this paper, we are motivated to consider the following problem. A secondary user, equipped with spectrum sensors, dynamically accesses a channel. If it transmits without/with colliding with primary users, a certain reward/penalty is obtained. If it senses the channel, accurate channel information is obtained, but a given channel sensing cost incurs. The third option for the user is to turn off the sensor/transmitter and go to sleep mode, where no cost/gain incurs. So when should the secondary user transmit, sense, or sleep, to maximize the total gain? We derive the optimal transmitting, sensing, and sleeping structure, which is a threshold-based policy. Our work sheds light on designing sensing and transmitting scheduling protocols for cognitive radio networks, especially the in-band sensing mechanism in 802.22 networks.

Keywords

cognitive radio, dynamic spectrum access, spectrum sensing stochastic dynamic programming

1. INTRODUCTION

Current static spectrum allocation policy of FCC has resulted in obvious imbalance of spectrum usage between licensed users and numerous unlicensed users crowded in the ISM band. To alleviate such a imbalance, cognitive radio[2]. has been considered as a key technology to enable the unli-

censed users or secondary users (SU) to dynamically access the licensed spectrum. Cognitive radio networks has attracted a lot of attention from both industry and academy. The IEEE standard 802.22, Wireless Regional Area Networks (WRAN) [1] is a highlight on this area.

Cognitive radio can dynamically and promptly adjust itself to the environments. For example, it can dynamically bind channels, switch channels, and control transmit power and modulation schemes[2, 3]. However, the power of cognitive radio cannot make it succeed solely. Cognitive radio networks can only dynamically access the spectrum holes where primary users (PU) are absent. How to protect the PUs communication is a key issue on the research of cognitive radio networks.

1.1 Motivation

Spectrum sensing and so called listen-before-talk strategy are vital for protecting PUs. The IEEE 802.22 standard has proposed a set of spectrum sensing schemes. We brief them as follows.

In 802.22 WRAN, spectrum sensors are deployed to detect PU signals. There are two objectives of spectrum monitor[1, 4]. First, if the channel is currently not used by the SUs, which is referred as out-of-band channel, channel sensing is to explore the spectrum hole that satisfies the demand of the SUs. Second, if the channel is being used by the SUs, referred as in-band channel, channel sensing is to detect the return of PUs and therefore protect PUs communication. This is also referred as in-band channel sensing, which is the focus of this paper. For in-band channel sensing, the IEEE 802.22 considers periodic sensing schemes, using both energy detection and feature detection. To detect the primary users' signal with accuracy. The sensing time is not negligible compared to the transmission time of a packet. For example, in feature detection of a standard DTV system, the fine sensing time is 25ms, which is much longer than the sub-millisecond transmission time of a packet. In addition, in 802.22 standard, cooperative sensing is proposed to decrease the probability of false alarm and miss detection. The sensing information of each sensor should be reported to a CPE or Base Station, which also incurs overhead.

There are already research works[6, 7, 8, 9] identifying the overhead of spectrum sensing, and the tradeoff between sensing overhead and PU protection. Most of the previous works assume periodic sensing and optimize the sensing time or periods, to achieve the balance between protecting PU communication and creating transmitting opportunity for SUs. With the same objective, there also exists work [9]

optimizing the sensing-transmitting structure. However, all these works simply treat the sensing cost as the time overhead and study the SUs' sensing and transmitting behaviors only. Generally, channel sensing cost can also be the energy cost of the spectrum sensors, or the collision penalty if the SU uses active probing. In some situation, the SU should neither transmit nor sense the channel. For example, if PUs are transmitting, although keeping sensing can find the beginning of PU idle time timely, the energy cost of spectrum sensors may be considerable. In this case, it may be better to consider the sleeping behavior of SUs. While channel switching can also be an option besides sensing, transmitting, and sleeping. It is out of the scope of the paper.

Therefore, an interesting and challenging question aries: in opportunistic spectrum access, when the SU should sleep, sense the channel, or transmit, considering the spectrum sensing cost, the requirements of of protecting PUs communications, and the demands of transmitting opportunities of SUs? We are motivated to study this problem in this paper. Intuitively, if the channel is more likely to be idle, the SU should transmit, if the channel is more probable to be busy, the SU should sleep rather than transmit or sense the channel. Sensing happens when it not clear whether the channel is idle or not. We start from this intuition and study the optimal channel sensing, transmitting, and sleeping structure of SU considering channel sensing cost. We make the following contributions.

1.2 Contribution

We consider channel sensing cost in opportunistic spectrum access. We study sleeping, sensing, and transmitting behaviors of SUs. We formulate the sleeping, sensing, and transmitting scheduling problem as a Markov Decision Problem (MDP). We show that the optimal sensing, sleeping, and transmitting has a simple threshold policy. In most cases, the SUs should sleep when the channel idle probability is below a certain threshold, transmit when the channel idle probability is larger than a certain threshold, and sense the channel otherwise. Our work sheds light on designing inband channel sensing protocol for cognitive radio networks.

1.3 Roadmap

The rest of paper is organized as follows. We describe the PU and SU models, as well as reward models in Section 2. In Section 3, we formulate the sleeping, sensing, and transmitting scheduling problem using discounted dynamic programming. We proceed to propose the optimal sleeping, sensing, and transmitting policy in Section 4, followed by the discussions in Section 5. After reviewing related work in Section 6, we draw conclusions in Section 7.

2. SYSTEM MODEL

We consider a channel where the PUs are licensed to access. At any moment, the channel can be in either a busy or idle state, which refers respectively to the times the PU transmits on the channel or not. We assume the time duration of the busy and idle state are exponentially distributed. Note that the PU transmission is not slotted.

We consider a SU as an opportunistic user equipped with a cognitive radio and a number of spectrum sensors. The SU operates in a slotted system, which is not synchronized to PU behavior. We use $S(n)$ to denote the channel state at slot n. Let $S(n) = 1$ if the channel is in idle state, and

$S(n) = 0$ when it is busy. Therefore, we can use a discrete markov chain to model the channel state transition. We have

$$P(S(n+1) = 0|S(n) = 1) = q,$$
$$P(S(n+1) = 1|S(n) = 0) = p. \tag{1}$$

where $p(q)$ is the transition probability from busy(idle) state to idle(busy) state. $p/(p+q)$ is the steady state probability of channel being idle. A reasonable assumption is $1-p-q > 0$, which implies that two adjacent slots are more likely to have the same channel state.

Spectrum sensors detect PUs' signal by sensing the channel. Cooperative sensing is desirable. How the spectrum sensors are organized to sense the channel is beyond the scope of the paper. Spectrum sensors report channel information to the cognitive radio which makes channel access decisions. The decisions are from the action space A, where A = $\{D : 0 \text{ (transmitting)}, 1 \text{ (sensing)}, 2 \text{ (sleeping)}\}$. For example, at the beginning of slot n, $D_n=2$ means the SU transmits. The decisions are made based on channel state information, at slot n, where X_n is the estimated probability of channel being idle at slot n. The value of X_n is from the space $\{\tau : 0 \leq \tau \leq 1\}$, which is a countable set (We can only consider the rational numbers in $[0, 1]$). X_n is updated according to the following rules. When $D_n = 0$ or 2,

$$X_{n+1} = X_n(1 - q) + (1 - X_n)p. \tag{2}$$

when $D_n = 1$,

$$X_{n+1} = \begin{cases} 1, & \text{w.p. } X_n(1-q) + (1-X_n)p; \\ 0, & \text{w.p. } 1 - X_n(1-q) - (1-X_n)p. \end{cases} \tag{3}$$

By (2), when transmitting or sleeping, the SU updates information state by markov transition. In (3), we assume perfect channel sensing results. That is, the sensing result is either 1 (idle) or 0 (busy). $X_n(1 - q) + (1 - X_n)p$ is the probability that slot $n + 1$ is sensed idle.

We assume one packet can be sent in a slot. In slot n, if $S(n) = 1$, the packet is sent successfully. In this case, the SU can get a certain throughput, which can be considered as the reward of the transmission, denoted by R_t. If $S(n) = 0$, the SU's transmission collides with the PU's packet. In this case, the SU is penalized by a cost, C_c. The larger the value of C_c, the better protection of the PU, but the less transmission opportunity for the SU. At the beginning of a time slot, if the information state X_n is τ, the expected total reward for transmission is $\tau R_t - (1 - \tau)C_c$. Note the average probability of channel being idle is $\frac{p}{p+q}$, which can be measured or queried from a database. Before accessing the channel, the SU often has no information on channel state. If $\frac{p}{p+q}R_t - \frac{q}{p+q}C_c > 0$, the SU can always transmit and receive a positive expected reward without channel sensing. This is not desirable for PU protection. Therefore, it is reasonable to assume $\frac{C_c}{R_t+C_c} \geq \frac{p}{p+q}$, although our paper does not restrict to it.

If $D_n = 1$, the SU senses the channel at slot n. Here channel sensing can be passive channel listening or active channel probing. In this paper, we assume perfect sensing model. Channel sensing provides the SU accurate channel information and therefore is helpful to make the decisions on transmitting or not. However, channel sensing has cost. If the SU is energy constraint, channel sensing depletes energy resource for both passive sensing or active probing. If the SU uses active probing, it may collide with PU and therefore

receives penalty. For simplicity, we assume the sensing cost is a constant, denoted by C_s.

We also consider the sleeping action of the SU. There is no reward and cost incurred by the sleeping action. Intuitively, the SU chooses sleeping action in the case that transmitting brings a large penalty or sensing the channel results in a large cost.

To sum up, the gain of a SU at slot n depends on information state X_n and action D_n and therefore is a function of them. We can write

$$G(X_n, D_n) = \begin{cases} \tau(R_t + C_c) - C_c, & X_n = \tau, D_n = 0; \\ -C_s, & D_n = 1; \\ 0, & D_n = 2. \end{cases} \tag{4}$$

Intuitively, if τ is large, the SU should transmit to get certain reward; if τ is small, the SU should choose to sleep to save sensing cost.

3. PROBLEM FORMULATION

The rewards from both the current slot and future slots should be considered to make a decision. The future reward is discounted by a factor α, $0 < \alpha < 1$. Suppose a policy π determines the action in each slot. Let $\pi = \{D_0, D_1, \ldots, D_n, \ldots\}$. Under policy π, starting from information state $X_0 = \tau$, we can write the total gain as

$$V_\pi(\tau) = E_\pi \left[\sum_{n=0}^{\infty} G(X_n, D_n)\alpha^n | X_0 = \tau \right], \tag{5}$$

Our objective is to find a policy π such that $V_\pi(\tau)$ is maximized. That is

$$V(\tau) = \max_\pi V_\pi(\tau), \forall \tau. \tag{6}$$

For maximizing discounted reward problem, we can write the Optimality Equation as

$$V(\tau) = \max_{\{0,1,2\}} (TR(\tau), SE(\tau), SL(\tau)). \tag{7}$$

where $TR(\pi)$, $SE(\tau)$, and $SL(\pi)$ are the expected rewards when the SU chooses to TRansmit, SEense the channel, and SLeep, respectively. Specifically, we have

$$TR(\tau) = \alpha V(\tau') + \tau(R + C_c) - C_c,$$
$$SE(\tau) = \alpha(\tau' V(0) + (1 - \tau')V(1)) - C_s, \tag{8}$$
$$SL(\tau) = \alpha V(\tau').$$

where $\tau' = \tau(1-q) + (1-\tau)p$. At a slot, if the information state is τ, a stationary policy that chooses the action to maximize the right side of (7) is optimal. To derive such a stationary optimal policy, let us first examine the properties of $V(\tau)$.

We have two lemmas on $V(\tau)$ as follows.

Lemma 1: $V(\tau)$ *is a convex function of* $\tau \in [0,1]$.

Proof: To prove it is a convex function. We need to show for $\forall \tau_1$ and $\tau_2 \in [0,1]$ and any $\beta \in [0,1]$, there is

$$V(\beta\tau_1 + (1-\beta)\tau_2) \leq \beta V(\tau_1) + (1-\beta)V(\tau_2). \tag{9}$$

Similar to the arguments in Lemma 3.1 in [5], a coin with probability β of heads is flipped. If the coin comes up head, the prior probability τ_1 is used. Otherwise the prior probability τ_2 is chosen. If the SU has no information on the outcome of the flip, the maximum expected reward should

be $V(\beta\tau_1 + (1-\beta)\tau_2)$. On the other hand, if the SU has the information of the outcome, the maximum expected reward should be $\beta V(\tau_1) + (1-\beta)V(\tau_2)$. Because this must be as good as the case that the SU ignores the information of the outcome of the flip. (9) holds and lemma 1 follows.

Lemma 2: $V(\tau)$ *is a monotonic increasing function of* τ.

Proof: To prove $V(\tau)$ is a mono-increasing function, we need to show for $\tau_1 \leq \tau_2$, there is $V(\tau_1) \leq V(\tau_2)$.

Since we consider discounted reward model, $\forall \varepsilon > 0$, there exists N large enough such that for any policy π,

$$\left| \sum_{n=N+1}^{\infty} G(X_n, D_n)\alpha^n | X_0 = \tau_i \right| < \varepsilon, \quad i = 1, 2, \ldots$$

Define

$$\hat{V}_k(\tau) = \max_\pi E \left[\sum_{n=k}^{N} G(X_n, D_n)\alpha^n | X_k = \tau \right], k = 0, 1, 2, \ldots, N.$$

We have

$$\left| V(\tau_i) - \hat{V}_0(\tau_i) \right| \leq \varepsilon, \quad i = 1, 2, \ldots$$

Next, we show that

$$\hat{V}_k(\tau_1) \leq \hat{V}_k(\tau_2), \quad k = 0, 1, 2, \ldots, N \tag{10}$$

First, let us consider $\hat{V}_N(\tau)$. $\hat{V}_N(\tau)$ is the maximal reward from a single slot N with an information state τ. If with τ_1, the optimal action is to sleep, the reward $\hat{V}_N(\tau_1)$ is 0. By also sleeping, $\hat{V}_N(\tau_2)$ is at least 0 and no less than $\hat{V}_N(\tau_1)$.

With τ_1, if the optimal action is transmitting, we have

$$\hat{V}_N(\tau_1) = \tau_1(R_t + C_c) - C_c$$
$$\leq \tau_2(R_t + C_c) - C_c \leq \hat{V}_N(\tau_2).$$

The optimal action at slot N cannot be sensing, because it only incurs cost and no reward will be obtained. Therefore, we have $\hat{V}_N(\tau_1) \leq \hat{V}_N(\tau_2)$ for $\tau_1 \leq \tau_2$. We next use induction to prove (10). Let us assume for $k = M$, there is $\hat{V}_M(\tau_1) \leq \hat{V}_M(\tau_2)$ for $\tau_1 \leq \tau_2$. Now consider slot $M-1$. By taking the action of sensing the channel, for information state τ_1 and τ_2, the total rewards from $M-1$ to N are respectively $\alpha(\tau_1'\hat{V}_M(1) + (1-\tau_1')\hat{V}_M(0)) - C_s$ and $\alpha(\tau_2'\hat{V}_M(1) + (1-\tau_2')\hat{V}_M(0)) - C_s$. Since $\hat{V}_M(\tau)$ is an increasing function, we have $\hat{V}_M(0) \leq \hat{V}_M(1)$. Meanwhile, because $\tau_1 \leq \tau_2$, we have $\tau_1' \leq \tau_2'$. Thus

$$\alpha(\tau_1'\varphi\hat{V}_M(1) + (1-\tau_1')\hat{V}_M(0)) - C_s$$
$$\leq \alpha(\tau_2'\hat{V}_M(1) + (1-\tau_2')\hat{V}_M(0)) - C_s.$$

If the SU transmits at slot $M-1$, the total reward from slot $M-1$ to N is $\alpha\hat{V}_M(\tau_1') + \tau_1(R_t + C_c) - C_c$ and $\alpha\hat{V}_M(\tau_2') + \tau_2(R_t + C_c) - C_c$ for τ_1 and τ_2 respectively. Since we have $\hat{V}_M(\tau)$ is an increasing function and $\tau_1' \leq \tau_2'$, there is

$$\alpha\hat{V}_M(\tau_1') + \tau_1(R_t + C_c) - C_c$$
$$\leq \alpha\hat{V}_M(\tau_2') + \tau_2(R_t + C_c) - C_c.$$

Taking sleeping action, the reward from slot $M-1$ to N has the relation $\alpha\hat{V}_M(\tau_1') \leq \alpha\hat{V}_M(\tau_2')$. Therefore, at slot $M-1$, for information state τ_1, whatever action it takes, we can guarantee a equal or larger reward for τ_2 by taking the same action. Then we have $\hat{V}_{M-1}(\tau_1) \leq \hat{V}_{M-1}(\tau_2)$ and therefore

$V_{M-1}(\tau)$ is an increasing function. Then by induction, we proved $V_k(\tau)$ is an increasing function for $k = 0, 1, 2, \ldots, N$.

Since from slot $N + 1$, the total reward is smaller than ε for any π. We have

$$V(\tau_1) - \varepsilon \le \hat{V}_0(\tau_1),$$
$$V(\tau_2) \le \hat{V}_0(\tau_2) + \varepsilon.$$

We have

$$V(\tau_1) - 2\varepsilon \le \hat{V}_0(\tau_1) - \varepsilon \le V(\tau_2).$$

Because ε is arbitrarily small, we have $V(\tau_1) \le V(\tau_2)$. Lemma 2 follows.

4. OPTIMAL SENSING, SLEEPING AND TRANSMITTING: A THRESHOLD BASED POLICY

With the two lemmas, we have Theorem 1 on the optimal transmitting, sensing, and sleeping policy.

Theorem 1: When $\frac{p}{p+q} \le \frac{C_c}{C_c+R_t}$, the following policy is optimal

$$D_n^* = \begin{cases} 0(\text{Transmitting}), & \text{when } \tau > \tau_2; \\ 1(\text{Sensing}), & \text{when } \tau_1 < \tau \le \tau_2; \\ 0(\text{Transmitting}), & \text{when } \frac{C_c}{C_c+R_t} < \tau \le \tau_1; \\ 2(\text{Sleeping}), & \text{when } \tau \le \frac{C_c}{C_c+R_t}; \end{cases}$$
(11)

or

$$D_n^* = \begin{cases} 0(\text{Transmitting}), & \text{when } \tau > \tau_2; \\ 1(\text{Sensing}), & \text{when } \tau_1 < \tau \le \tau_2; \\ 2(\text{Sleeping}), & \text{when } \tau \le \tau_1; \end{cases}$$
(12)

when $\tau_1 < \frac{C_c}{R_t+C_c} \le \tau_2$;

When $\frac{p}{p+q} > \frac{C_c}{C_c+R_t}$, the following policy can also be optimal

$$D_n^* = \begin{cases} 0(\text{Transmitting}), & \text{when } \tau > \frac{C_c}{R_t+C_c}; \\ 2(\text{Sleeping}), & \text{when } \tau_2 < \tau \le \frac{C_c}{R_t+C_c}; \\ 1(\text{Sensing}), & \text{when } \tau_1 < \tau \le \tau_2; \\ 2(\text{Sleeping}), & \text{when } \tau \le \tau_1; \end{cases}$$
(13)

where τ_1 and τ_2 are thresholds and will be discussed later. Theorem 1 states that the optimal policy is a threshold based policy. The structure of the policy can either be sleeping, sensing, sleeping, and transmitting or sleeping, transmitting, sensing, and transmitting. By examining the relations among τ_1, τ_2, $\frac{C_c}{R_t+C_c}$, and $\frac{p}{p+q}$, we simplify the structure of the optimal policy in some cases. The proof of Theorem 1 is as follows.

First, let us define a function $F(\tau)$,

$$F(\tau) = \max\{TR(\tau), SL(\tau)\}.$$
(14)

According to (8), $F(\tau) = SL(\tau)$ if $\tau \le \frac{p}{p+q}$, and $F(\tau) = TR(\tau)$ otherwise. $F(\tau)$ is an increasing convex function, because $SL(\tau)$ and $TR(\tau)$ are both increasing convex. Second, let us consider $SE(\tau)$, since $SE(\tau) = \tau' V(1) + (1-\tau')V(0) - C_s$, and $\tau' = p + (1-p-q)\tau$, we have

$$\frac{dSE(\tau)}{d\tau} = (1-p-q)(V(1) - V(0)),$$

which is constant. Therefore, $SE(\tau)$ is a linear function. It is also increasing because $V(0) \le V(1)$.

Consider the relation between $F(\tau)$ and $SE(\tau)$. There are three cases. First, $F(\tau)$ and $SE(\tau)$ have no intersection. In this case, there either be $SE(\tau) < F(\tau)$ or $SE(\tau) > F(\tau)$, $\forall \tau$. The latter case cannot be true. Suppose $SE(\tau) > F(\tau)$, $\forall \tau$, by optimality equation, the SU will always choose to sense the channel, which receives a negative reward. However, $F(\tau) \ge 0$ by taking the sleeping action. Therefore, in this case, the optimal policy is

$$D_n^* = \begin{cases} 0(\text{Transmitting}), & \text{when } \tau > \frac{C_c}{R_t+C_c}; \\ 2(\text{Sleeping}), & \text{when } \tau \le \frac{C_c}{R_t+C_c}; \end{cases}$$
(15)

The intuition that the SU never senses the channel is because C_s is relatively high compared to R_t and C_c.

The second case is that $F(\tau)$ and $SE(\tau)$ has one intersection. The intersection can be a point or a segment. If it is a segment, define τ_1 as

$$\tau_1 = \min\{\tau, F(\tau) = SE(\tau)\}.$$

Then, when $\tau \le \tau_1$, there must be $SE(\tau) \le F(\tau)$. The reason is, suppose $SE(\tau) > F(\tau)$ when $\tau \le \tau_1$. The slope of $F(\tau)$ at τ_1 must be larger than that of $SE(\tau)$. Since the slope of $F(\tau)$ is not decreasing, there must be $F(\tau) > SE(\tau)$ when $\tau > \tau_1$. Then the intersection between $SE(\tau)$ and $F(\tau)$ cannot be a segment. Therefore, when the intersection between $SE(\tau)$ and $F(\tau)$ is a segment, $SE(\tau) \le F(\tau), \forall \tau$. Thus the optimal policy is the same as there is no intersection. We will discuss the case the intersection is a point below.

The third case is that $F(\tau)$ and $SE(\tau)$ have two intersections, shown by Fig. 1. In this case, either intersection cannot be a segment. Let τ_1 and τ_2 denote the intersection points of $F(\tau)$ and $SE(\tau)$. The second case can be considered as a special case of the third case when $\tau_1 = 0$ or $\tau_2 = 1$. Therefore, we only discuss the third case. Since $F(\tau)$ is a convex increasing function, and $SE(\tau)$ is a linear increasing function, there cannot be three intersections between $F(\tau)$ and $SE(\tau)$. We have

$$SE(\tau) < F(\tau), \quad \tau < \tau_1;$$
$$SE(\tau) > F(\tau), \quad \tau_1 < \tau < \tau_2;$$
$$SE(\tau) < F(\tau), \quad \tau_2 < \tau.$$
$$SE(\tau) = F(\tau), \quad \tau = \tau_1 \text{ or } \tau_2;$$

Therefore, between τ_1 and τ_2, the optimal action is to sense the channel. Outside the region, the optimal action is either sleeping or transmitting, depending on the location of $\tau_3 = \frac{C_c}{R_t+C_c}$. There are three cases, $\tau_2 \le \tau_3$, $\tau_1 < \tau_3 \le \tau_2$ and $\tau_3 \le \tau_1$, as Fig. 2 shows, which are discussed as follows.

Case 1: $\tau_2 \le \tau_3$.

In this case, as showed by Fig. 2(a), first, if $\tau > \tau_3$, the SU must transmit and this is the only transmission region. There are two sleeping regions. One is the case when $\tau \le \tau_1$, the other is when $\tau_2 < \tau \le \tau_3$. Formally, the optimal action is

$$D_n^* = \begin{cases} 0(\text{Transmitting}), & \text{when } \tau > \tau_3; \\ 2(\text{Sleeping}), & \text{when } \tau_2 < \tau \le \tau_3; \\ 1(\text{Sensing}), & \text{when } \tau_1 < \tau \le \tau_2; \\ 2(\text{Sleeping}), & \text{when } \tau \le \tau_1; \end{cases}$$
(16)

Case 2: $\tau_1 < \tau_3 \le \tau_2$.

In this case, obviously, the optimal policy is sleeping when τ is smaller than τ_1, sensing the channel when $\tau_1 < \tau \le \tau_2$
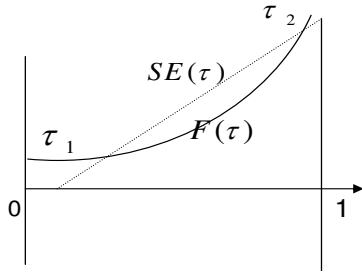
Figure 1: When $\tau_1 \leq \tau \leq \tau_2$, the SU should sense the channel. Otherwise it either transmits or sleeps.
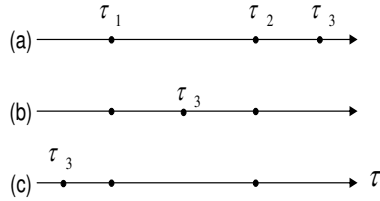

Figure 2: When $\tau_1 \leq \tau \leq \tau_2$, the SU should sense the channel. Otherwise it either transmits or sleeps.

and transmitting if $\tau > \tau_2$. Formally, we have

$$D_n^* = \begin{cases} 0(\text{Transmitting}), & \text{when } \tau > \tau_2; \\ 1(\text{Sensing}), & \text{when } \tau_1 < \tau \leq \tau_2; \\ 2(\text{Sleeping}), & \text{when } \tau \leq \tau_1; \end{cases} \quad (17)$$

Case 3: $\tau_3 \leq \tau_1$.

In this case, the optimal policy is sleeping when τ is smaller than τ_3, transmitting when $\tau_3 < \tau \leq \tau_1$, sensing the channel when $\tau_1 < \tau \leq \tau_2$, and transmitting when τ is larger than τ_2. That is

$$D_n^* = \begin{cases} 0(\text{Transmitting}), & \text{when } \tau > \tau_2; \\ 1(\text{Sensing}), & \text{when } \tau_1 < \tau \leq \tau_2; \\ 0(\text{Transmitting}), & \text{when } \tau_3 < \tau \leq \tau_1; \\ 2(\text{Sleeping}), & \text{when } \tau \leq \tau_3; \end{cases} \quad (18)$$

The above are the optimal policies corresponding to different locations of τ_3. Next, we show under certain condition, the first case does not exist.

Lemma 3: When $\frac{p}{p+q} \leq \frac{C_c}{C_c+R_t}$, τ_3 cannot be larger than τ_2.

Proof: Suppose when $\frac{p}{p+q} \leq \frac{C_c}{C_c+R_t}$, τ_3 is larger than τ_2. Let SL_2 to denote the region $\tau_2 < \tau \leq \tau_3$, where the optimal action is sleeping. There are three cases as below.

- $\tau_2 < \frac{p}{p+q} \leq \tau_3$. In this case, at $\tau = \frac{p}{p+q}$, the optimal action is sleeping, therefore, $\frac{p}{p+q}$ becomes the absorbing state. $V(\frac{p}{p+q}) = 0$. Since $V(\tau)$ is a monotonically increasing function, for all information state $\tau_1 < \tau \leq \tau_2$, we have $V(\tau) = 0$. For $\tau_1 < \tau \leq \tau_2$, if the action is sleeping, we have $V(\tau) = 0$. Therefore, between τ_1 and τ_2, the optimal action is also sleeping. In this case, the optimal policy is to transmit when $\tau > \tau_3$, sleep otherwise. There is no sensing range. It contradicts the hypothesis that between τ_1 and τ_2, it is optimal to sense rather than sleep or transmit.

- $\tau_1 < \frac{p}{p+q} \leq \tau_2$. In this case, the optimal action for $\tau = \frac{p}{p+q}$ is to sense the channel. Consider an information state $\tau_2 < \tau \leq \tau_3$. The optimal action for τ is to sleep. At the next slot, the information state is τ'. Since $\tau_1 < \frac{p}{p+q} \leq \tau_2$ and $\tau > \frac{p}{p+q}$, we have $\tau_1 < \tau' \leq \tau_2$ or $\tau_2 < \tau' \leq \tau_3$. Consider the former case first. In this case, with τ', the optimal action is to sleep. Therefore we have

$$V(\tau) = \alpha V(\tau') = \alpha^2(\tau'' V(1) + (1-\tau'')V(0)) - \alpha C_s,$$

where τ'' is the probability of the channel being sensed idle when the sensing action is taken at the slot with an information state τ'. If the SU chooses to sense the channel with τ, then the reward is $\alpha(\tau' V(1) + (1-\tau')V(0) - C_s))$. Since $\tau' > \tau'' > \frac{p}{p+q}$, $V(1) > V(0)$ (according to Lemma 2), we have $V(\tau') \leq \alpha(\tau' V(1) + (1-\tau')V(0) - C_s)$. Therefore, it is optimal to sense the channel with τ rather than sleeping. In the case $\tau_2 < \tau' \leq \tau_3$, the optimal action for τ_2' is to sleep. If at the next slot, τ'' is no larger than τ_2. Then we have proved that with information state τ', the optimal action is to sense the channel, therefore, with τ, by induction, the optimal action is also to sense the channel. If τ'' is still larger than τ_2, by keeping sleeping, finally, the probability of channel being idle will fall below τ_2 while is lager than $\frac{p}{p+q}$. Therefore, by induction, with information state τ, the optimal action is to sense the channel. Then in this case we have $\tau_2 = \tau_3$. SL_2 does not exist. The optimal policy is sleeping if $\tau \leq \tau_1$, sense the channel if $\tau_1 < \tau \leq \tau_3$, and transmit if $\tau > \tau_3$. It contradicts the hypothesis that between τ_1 and τ_2, it is optimal to sense rather than sleep or transmit.

- $\frac{p}{p+q} \leq \tau_1$. In this case, consider a τ which satisfies $\tau_2 < \tau \leq \tau_3$. At the next slot, the information state τ' satisfies $\tau_1 < \tau' \leq \tau_2$ or $\tau' \leq \tau_1$. In the former case, we can prove the optimal action for τ is to sense the channel, following the same steps as the above case. And thus $\tau_2 = \tau_3$. In the later case, by taking the action of sleeping, from τ, the SU will reach the absorbing state $\frac{p}{p+q}$, then $V(\tau) = 0$. Therefore, we have when $\tau_1 < \tau \leq \tau_3$, the optimal action is also to sleep. There is no sensing region. It contradicts the hypothesis that between τ_1 and τ_2, it is optimal to sense rather than sleep or transmit.

Therefore, in summary, when $\frac{p}{p+q} \leq \frac{R_t}{C_c+R_t}$, the optimal policy for Case 1 is sleeping when $\tau \leq \tau_1$, sensing when $\tau_1 < \tau \leq \tau_3$, and transmitting when $\tau > \tau_3$. Thus the optimal policy of Case 2 covers the optimal policy of Case 1 when $\tau_2 = \tau_3$. We get (11) and (12) hold. When $\frac{p}{p+q} > \frac{R_t}{C_c+R_t}$, (13) holds according to (16).

5. DISCUSSION

We have presented the structure of the optimal sleeping, sensing, and transmitting policies. The two values $\frac{C_c}{R_t+C_c}$ and $\frac{p}{p+q}$ are important in determining the structure of optimal policy. Before the SU accesses the channel, the best estimation of the channel state is the steady state probability of channel being idle, i.e., $\frac{p}{p+q}$. That is, the SU starts with

the initial information state $\frac{p}{p+q}$. To protect PU communication, $\frac{p}{p+q} \leq \frac{C_c}{R_t+C_c}$ is desirable to prevent the SU from always transmitting and still receiving positive reward. In this case, at $\frac{p}{p+q}$, the SU either always sleeps or senses the channel. If it senses the channel, according to our results, the optimal policy is to sleep when τ is less than a threshold, transmit when τ is larger than a threshold, and sense otherwise.

The discount factor α does not change the structure of the optimal policy. Suppose with a certain α, the optimal policy is to sleep, sense, and transmit, when α changes, the policy will not change to other structure such as sleeping, transmitting, sensing, and transmitting. However, the change of α will change the values of the thresholds. Suppose the optimal policy has a sensing region for a given alpha. The sensing region will diminish as α decreases. If α is very small, the SU will not sense the channel, since the future rewards resulting from accurate channel information cannot compensate the channel sensing cost at the current slot. .

6. RELATED WORK

We briefly review related work in this section. Our work is different from [6, 7, 8] since those works consider periodic sensing schemes. In [6], the authors propose spectrum sensors clustering, and study energy detection and feature detection for in-band spectrum sensing. The frequency of channel sensing is minimized while the miss detection probability requirement is satisfied. The authors do not consider the energy consumption of spectrum sensors. They do not consider PU protection mechanism in case SUs collide with PUs either. In [7, 8], the authors optimize sensing time to achieve optimal tradeoff between the false-alarm probability and throughput. However, only optimizing sensing time duration is not enough for reducing sensing cost since it can be better to turn off spectrum sensors for some time when PUs are using the channels. In [9], the authors study optimal sensing-transmitting structure for secondary users. Our work in sensing and transmitting scheduling is similar to this work. However, the authors in [9] do not consider the energy cost of channel sensing. They propose to always sense the channel during the PU transmission. In [10], the authors consider channel sensing (probing) cost and explore the multi-channel diversity. The authors show that the optimal strategy on channel sensing (probing), guessing, and transmission has a threshold-based structure. Our work studies the structure of sleeping, sensing and transmitting, with random PU behaviors considered. In [11, 12, 13], the authors study channel selection and access problems in multi-channel cognitive radio networks. Different from our work, they assume slotted PU activities. In [11], the authors also consider the energy cost of spectrum sensors. They study the problem that whether to sense, which channel to sense, and whether to access, and formulate it as a POMDP problem. However, by assuming slotted PU traffic, they do not consider the collision cost with PUs. Their work involves exploring spectrum holes while our work focus on in-band channel sensing. In [12, 13], the authors use collision probability constraint to protect PUs. The two works consider limited sensing ability of SUs but do not consider sensing cost.

7. CONCLUSIONS

In this paper, we study channel sensing and transmitting scheduling in cognitive radio networks. A SU reaps a reward when transmits successfully, pays a penalty when collides with PU, and incurs a sensing cost when sensing the channel, e.g., sensing energy consumption or probing penalty. The SU needs to determine when to sense the channel, transmit, and sleep. We formulate the problem using discounted dynamic programming. We derive the optimal sensing, sleeping, and transmitting structure. In most cases, the SU should transmit if the probability of channel being idle is larger than a certain threshold, sleep when it is smaller than a threshold, and sense the channel otherwise. In our future work, we will extend the study to other reward models.

8. REFERENCES

[1] IEEE 802.22 WRAN WG, www.ieee802.org/22/.

[2] S. Haykin. Cognitive radio: Brain-empowered wireless communications. *IEEE Journal on Selected Areas in Communications.*, 23(2):201–220, Febraury 2005.

[3] R. W. Thomas, L. A. DaSilva, A. B. MacKenzie, Cognitive networks, In *IEEE DySPAN* 2005.

[4] Cordeiro, K. Challapali, and M. Ghosh. Cognitive PHY and MAC layers for dynamic spectrum access and sharing of TV bands. In *ACM TAPAS*, Aug. 2006.

[5] S. M. Ross. *Introduction to Stochastic Dynamic Programming.* Orlando, FL, USA: Academic Press, Inc., 1983,

[6] H. Kim and K. G. Shin. In-band Spectrum Sensing in Cognitive Radio Networks: Energy Detection or Feature Detection? In *ACM Mobicom* 2008.

[7] P. Wang, L. Xiao, S. Zhou, and J. Wang. Optimization of detection time for channel efficiency in cognitive radio systems. In *IEEE WCNC* 2007.

[8] Y.-C. Liang, Y. Zeng, E. Peh, and A. T. Hoang. Sensing-throughput tradeoff for cognitive radio networks. *IEEE Transactions on Wireless Communications.*, vol. 7, no. 4, pp. 1326Í C1337, April 2008.

[9] S. Huang, X. Liu and Z. Ding. Optimal Sensing-Transmission Structure for Dynamic Spectrum Access. available at $http : //www.cs.ucdavis.edu/\ liu/preprint/huang_sensing.pdf$.

[10] N. B. Chang and M. Liu. Optimal channel probing and transmission scheduling for opportunistic spectrum access In *ACM MobiCom* , 2007.

[11] Y. Chen, Q. Zhao, and A. Swami. Distributed Spectrum Sensing and Access in Cognitive Radio Networks with Energy Constraint. To appear in IEEE Transactions on Signal Processing.

[12] Q. Zhao, L. Tong, A. Swami, and Y. Chen. Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework. In *IEEE Journal on Selected Areas in Communications (JSAC): Special Issue on Adaptive, Spectrum Agile and Cognitive Wireles Networks,* vol. 25, no. 3, pp. 589Í C600, 2007.

[13] 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. 2053Í C 2071, 2008.