What Should Secondary Users Do Upon Incumbents' Return?

Caoxie Zhang and Kang G. Shin, *Fellow, IEEE*

***Abstract*—In a cognitive radio network (CRN), secondary users (SUs) opportunistically utilize idle licensed spectrum bands. We address the natural questions that arise when the incumbents or primary users (PUs) return to the channel the SUs are using opportunistically. Instead of immediately switching to another idle channel as proposed in almost all existing approaches, the SUs may opt to wait silently in their current channel until the PUs depart. This option would be beneficial to the SUs if the returned PUs stay at the channel only for a short period of time and the SUs' channel-switching incurs a non-negligible overhead. We determine how long the SUs should wait in their current channel before switching to a new idle channel.**

The SUs should also occasionally sense those (called *out-of-band*) channels currently not in use for sensing the availability of spectrum opportunities. We propose an efficient, adaptive spectrum-sensing technique to detect when a busy out-of-band channel becomes idle. We also present a spectrum-management architecture that integrates the SUs' strategies and facilitates fast discovery of spectrum opportunities.

***Index Terms*—Cognitive radio, channel switching, adaptive sensing, Markov decision process.**

I. Introduction

COGNITIVE radio (CR) has opened a new way to improve the poor spectrum-utilization problem under the current *static* allocation policy. It allows cognitive or secondary users (SUs) to opportunistically utilize the legacy or primary channels allocated to licensed or primary users (PUs). Today's primary channels are the UHF bands used for TV broadcast and wireless microphone transmissions. For the secondary devices, although FCC advocates the incorporation of geo-location databases and eliminates the requirements of spectrum sensing in TV bands, it still encourages the development of spectrum sensing for opportunistic utilization of other bands [1]. Here we are mainly interested in, but not limited to, primary channels that are more dynamic than TV transmissions, on which even cellular or WiMAX networks may operate. Spectrum sensing is still under development for secondary devices (with a single radio) and necessary for highly dynamic bands, for example, with data-intensive incumbent networks.

The central issue in CRNs is how SUs can make efficient use of the licensed band while keeping their interference to the PUs below a specified level. In this paper, we focus on the design of spectrum access and sensing algorithms for

Manuscript received 1 January 2012; revised 12 May 2012 and 22 July 2012. The work reported in this paper was supported in part by the NSF under Grant CNS-1160775.

The authors are with Department of Electrical Engineering and Computer Science, The University of Michigan, Ann Arbor, MI, 48109-2121, USA (e-mail: {caoxiezh,kgshin}@umich.edu).

Digital Object Identifier 10.1109/JSAC.2013.130308.

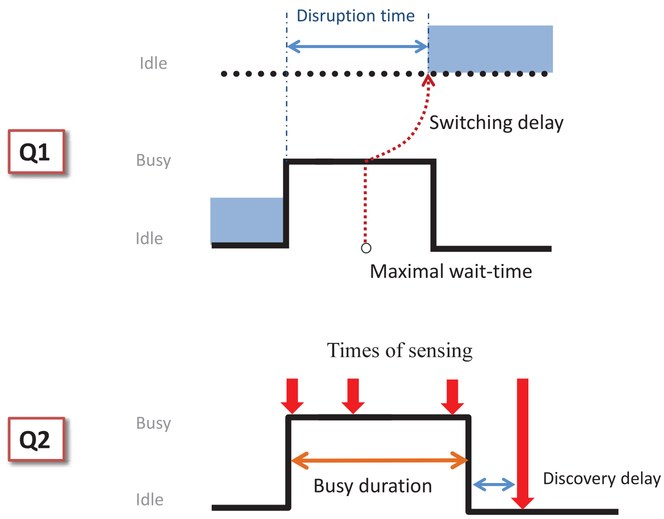


Fig. 1. Upon return of PUs, there are two key questions for the SUs to answer: (Q1) Should the SUs switch or not? (Q2) If the SUs leave the current busy channel, how to detect the PUs' departure at low cost? The grey color in (Q1) represents SUs' utilization of the primary channels.

fast discovery of dynamically available spectrum resources. In particular, we are interested in answering an important, practical question "what should SUs do when the PU returns to the channel that they are utilizing opportunistically?" Upon return of the PUs, the SUs must vacate the channel within a certain time limit to prevent harmful interference to the PUs' communications. To enable the SUs to quickly discover available spectrum, we consider two important questions regarding their choice of actions: (Q1) if the returned PUs finish their session quickly, should the SUs stay silently on their current channel or switch to another channel?, (Q2) if the SUs leave the currently busy (due to the PUs' return) channel, how often should they sense this channel to efficiently discover the PUs' departure? Fig. 1 illustrates insights behind these two questions and the corresponding SUs' strategies. We now introduce novel and efficient answers to both questions.

A. Wait or Switch

If the PUs' active session ends quickly enough (e.g., after transmitting a few data packets in a cellular/WiMAX network), the SUs may opt to stay silently in the channel until the PUs leave and regain opportunistic access to the channel. It could be beneficial for the SUs to wait in the current (busy) channel as the PUs may depart soon and the SUs' switching to another channel is time-consuming. In fact, a channel-switch can involve searching for an idle channel and setting up connections among the SUs. Existing approaches have not considered this wait-or-switch option, despite its practical importance. In an unslotted structure of the SUs' spectrum sensing and access, once the SUs sense the PUs' return, they will immediately switch to another channel [2–4]. In a slotted structure where the spectrum sensing and access are done at pre-determined discrete times, the SUs are allowed to stay in a busy (due to the PUs' return) band but are forced to wait until the beginning of the next slot, thus depriving their freedom to sense and access other channels during the slot [5–8]. We address this problem by empowering the SUs' an option to wait for some time before switching to another idle channel in an unslotted structure, as illustrated in Q1 of Fig. 1. We determine how long the SUs should wait in order to minimize the time that the secondary network services are disrupted until a new spectrum opportunity is found. We obtain this result for an arbitrary distribution of the PUs' occupancy when the distribution is known to the SUs, e.g., by accessing a geo-location database.

The problem becomes more challenging if the statistics of the PUs' activity are unknown to the SUs. The SUs must learn the distribution of the PUs' occupancy on-the-fly, i.e., estimating the distribution from the SUs' observations given the amount of time they have already waited. The SUs' goal then becomes minimization of their overall disruption time while learning the PUs' activity pattern. We study how the SUs make a sequential decision on the wait-time to capture the tradeoff between learning of the PUs' activity pattern and reduction of the SUs' disruption time.

B. Cost-Effective Detection of PUs' Busy $\rightarrow$ Idle Transition

When the PUs' active period is relatively long, the SUs need to switch to another idle channel. A remaining question is then how the SUs detect the availability of the channel they had opportunistically occupied until the PUs returned. Since the PUs' departure from that channel is unknown to the SUs, the SUs should occasionally sense the channel to determine when it becomes idle again, as illustrated in Q2 of Fig. 1. This spectrum-sensing should be scheduled economically since SUs must spend time that would otherwise be used in accessing other idle channels. Kim and Shin [9] optimized periodic sensing, which cannot achieve the best performance and may sometimes perform poorly. For example, suppose a primary voice service may either last for a short period of time, say 10s, or last very long, say 30min. The periodic sensing will inevitably schedule the sensing during the period from the 10-th second to the 30-th minute, thus unnecessarily sensing the channel many times. Moreover, the SUs may not quickly detect the PUs' departure from the channel. The sensing intervals need not be identical if the SUs knew possible durations of this voice service. We can sense the channel more often around the 10-th second and the 30-th minute, and much less often in between. Such aperiodic or adaptive sensing can save more time and energy for the SUs. We study how to schedule sensing to minimize the latency in detecting the PUs' departure while limiting the total fraction of time spent on sensing.

C. Spectrum-Management Architecture

We have considered whether SUs should switch or wait (Q1) and how they efficiently detect the transition of a busy channel to become idle (Q2). We also propose a spectrum-managemetn architecture to allow these SUs' startegies to work synergistically under all circumstances. Our spectrum-management architecture specifies the details of all the operations SUs need to take, with the goal of fast discovery of spectrum opportunities upon the PUs' return.

D. Contributions

We propose novel algorithms for spectrum access and sensing, addressing several important problems in fast discovery of spectrum opportunities.

In the wait-or-switch problem, unlike existing approaches, we allow SUs to wait silently in a busy (due to the PUs' return) channel as a proactive means for them to acquire a spectrum opportunity in the future. We derive interesting properties of how long the SUs should wait when the distribution of the PUs' occupancy duration has a monotonic hazard rate, including nearly all of well-known positive distributions. When the distribution is unknown, we formulate a problem as a Markov decision process to determine the optimal wait-time while learning the distribution using a Bayesian approach. We advocate the use of a simple myopic policy for an unknown exponential distribution of the PUs' occupancy duration, as it is easy to implement and performs very close to the optimal policy obtained by using dynamic programming.

In the problem of detecting PUs' busy-to-idle transition, we propose to use adaptive sensing intervals. Obtaining the optimal adaptive sensing intervals turns out to be intractable as it introduces an infinite number of variables. So, we introduce the concept of *smooth sensing rate* to approximate the times of sensing and use the variations of calculus to obtain the optimal sensing rate. Our numerical results show that this approach yields much better performance than existing periodic sensing, and is also computationally efficient. We also discuss the impact of sensing errors on the scheduling of this adaptive sensing.

Finally, we propose a spectrum-management architecture that allows SUs to discover spectrum opportunities as quickly as possible when the PUs return. In this architecture, system designers can tune the fraction of time for out-of-band sensing to strike a balance between the exploration and exploitation of dynamic spectrum opportunities.

II. Wait or Switch?

When a PU's occupancy period is short, how long should the evicted SUs wait in the current channel before switching to another idle channel? Specifically, we will determine the maximal wait-time to minimize the average disruption time experienced by the SUs until they discover a new opportunity. This maximal-wait strategy is inspired by the early work in [10] on the use of retry to recover from a failure in computer systems.

We use a probabilistic model for a primary channel's busy period. Suppose the primary channel becomes busy due to some PU's return at $t = 0$, the busy duration is a random duration X, with p.d.f. $f(x)$ and c.d.f. $F(x)$. While the SUs are waiting silently on this primary channel, they continuously sense the channel. During the wait, they can switch to another channel, taking an additional (random) time T_{switch} to find it and set up connections there. Here we assume the expected switching delay $\mathbb{E}[T_{\text{switch}}]$ is known as *a priori*[1]. As for the knowledge of the primary channel's active distribution $F(x)$, we consider two different scenarios. In the first scenario, $F(x)$ is constantly provided to the SUs by a trusted third party, e.g., the SUs can access parameters of the distribution via a geo-location database. In this scenario, we characterize the properties of the optimal wait-time in Sec. II-A. In the other scenario, the SUs do not have the knowledge of the PU's activity pattern initially. The SUs need to estimate the distribution while minimizing the disruption time. We present optimal and suboptimal solutions and their evaluations in Sec. II-B.

A. When Primary Busy Distribution $F(x)$ Is Known

Here we assume that the SUs know *a priori* the primary channel's active distribution, $F(x)$. In the case of cellular networks, since the stationary period is reported to be 1 hour [11], the SUs can access the geo-location database on an hourly basis. Upon return of a PU, the SUs will wait for a maximal wait-time, t_w, before switching. If the primary channel becomes idle due to the PU's departure before t_w, i.e., $X < t_w$, since the SUs are constantly sensing the channel, they can regain opportunistic use and are disrupted only for the period of time, X. Otherwise (i.e., $X \geq t_w$), the SUs will leave the channel at t_w, find an idle channel and set up connections on the channel at $t_w + T_{\text{switch}}$. So, the disruption time is a random variable and can be written as

$$D(t_w) = \begin{cases} X & \text{if} \quad X \leq t_w \\ t_w + T_{\text{switch}} & \text{if} \quad X > t_w. \end{cases} \tag{1}$$

We would like to minimize its expected value over the maximal wait-time t_w:

$$\underset{t_w \geq 0}{\text{minimize}} \quad \mathbb{E}[D(t_w)] = \int_0^{t_w} x f(x) \mathrm{d}x + P(X \geq t_w)(t_w + \mathbb{E}[T_{\text{switch}}]). \tag{2}$$

We now study properties of the solution to the above problem. Taking the first derivative of the objective function, we get

$$\frac{\mathrm{d}\mathbb{E}[D(t_w)]}{\mathrm{d}t_w} = (1 - F(t_w))(1 - u(t_w)\mathbb{E}[T_{\text{switch}}]), \tag{3}$$

where $u(x) = \frac{f(x)}{1-F(x)}$ is the *hazard rate function*, also known as the *failure rate function* [12], of the primary channel's active duration. We will show that the monotonicity of the hazard rate function decides the structure of the optimal wait strategy. We will also study the properties of certain distributions with non-monotonic hazard rates, such as Pareto and lognormal distributions. Our analysis will then cover almost all well-known distributions.

[1] In Sec. IV we will describe how to obtain $\mathbb{E}[T_{\text{switch}}]$.

1) Occupancy Distribution with Monotonic Hazard Rate: There are many distributions with monotonic hazard rate functions. Some are nondecreasing, such as the gamma and Weibull distributions, when the shape parameters are greater than or equal to 1, uniform distribution and Gaussian distribution[2]. Some are decreasing, such as the gamma and Weibull distributions when the shape parameters are less than 1. We now present two theorems regarding monotonic hazard rate functions.

Theorem 2.1: When the hazard rate function, $u(x)$, is nondecreasing in its domain, the optimal maximal wait-time is:

$$t_w^* = \begin{cases} 0 & \text{if} \quad \mathbb{E}[X] > \mathbb{E}[T_{\text{switch}}] \\ +\infty & \text{if} \quad \mathbb{E}[X] \leq \mathbb{E}[T_{\text{switch}}]. \end{cases} \tag{4}$$

The corresponding minimum average disruption time is then $\min\{\mathbb{E}[X], \mathbb{E}[T_{\text{switch}}]\}$.
An increasing hazard rate means that the PU will be more likely to depart its channel as its stay extends. It is therefore reasonable for the SUs to make a wait-or-switch decision at the very beginning. The above theorem indicates by only comparing the two means the SUs can either switch instantly upon the PU's return or wait in-band forever. The theorem is valid for exponential (the hazard rate is a constant) and Erlang distributions since both correspond to the case of the gamma distribution's shape parameters being positive integers.

Proof: Since $u(x)$ is nondecreasing, if there exists any x^* such that $u(x^*) = \frac{1}{\mathbb{E}[T_{\text{switch}}]}$, we have $\frac{\mathrm{d}\mathbb{E}[D(t_w)]}{\mathrm{d}t_w} \geq 0$ for $t_w < x^*$ and $\frac{\mathrm{d}\mathbb{E}[D(t_w)]}{\mathrm{d}t_w} \leq 0$ for $t_w \geq x^*$. This suggests that $\mathbb{E}[D(t_w)]$ be a unimodal function with the maximum being $\mathbb{E}[D(x^*)]$. Hence, the minimum of $\mathbb{E}[D(t_w)]$ occurs only at $t_w = 0$ or $t_w = +\infty$. ■

Theorem 2.2: When the hazard rate function, $u(x)$, is decreasing in its domain, the optimal maximal wait-time t_w^* is obtained as follows. If there exists any x^* such that $u(x^*) = \frac{1}{\mathbb{E}[T_{\text{switch}}]}$, then $t_w^* = x^*$, and the minimum average disruption time is $\mathbb{E}[D(x^*)]$. Otherwise (i.e., there does not exist x^* such that $u(x^*) = \frac{1}{\mathbb{E}[T_{\text{switch}}]}$), we have the same result as in (4).

Proof: Since the hazard rate function $u(x)$ is decreasing, there can exist at most one x^* such that $u(x^*) = \frac{1}{\mathbb{E}[T_{\text{switch}}]}$. If there exists such an x^*, by inspecting the first derivative, we can see that $\mathbb{E}[D(t_w)]$ is a unimodal function that first decreases at $[0, t^*]$ and then increases at $(t^*, +\infty)$. If there does not exist such x^*, the first derivative of $\mathbb{E}[D(t_w)]$ should always be greater or less than 0, meaning that $\mathbb{E}[D(t_w)]$ is an increasing or decreasing function. ■

2) Results on Occupancy Distributions with Non-Monotonic Hazard Rates: The actual primary channel occupancy distribution might not have a monotonic hazard rate [13]. We study two popular distributions of this class: Pareto and lognormal. The Pareto distribution belongs to the family of heavy-tail/long-tail distribution. We can obtain its hazard rate function as

$$u(t) = \begin{cases} 0 & \text{if} \quad 0 \leq t < x_m \\ \frac{\alpha}{t} & \text{if} \quad t \geq x_m, \end{cases} \tag{5}$$

[2] To model positive durations, the Gaussian distribution is left truncated at 0.

where x_m and α are the scale and shape parameters, respectively. One can easily show that only $t_w = 0$ and $t_w = \alpha E[T_{\text{switch}}]$ can be the local minimum of $D_0(t_w)$. Thus, the minimum disruption time is $\min\{E[T_{\text{switch}}], D_0(\alpha E[T_{\text{switch}}])\}$.

Dealing with the hazard rate of lognormal distribution is more complicated analytically. We have the hazard rate expressed as

$$u(t) = \frac{\exp\left(-\frac{(\ln t-\mu)^2}{2\sigma^2}\right)}{\sqrt{2\pi}t\sigma Q\left(\frac{\ln t-\mu}{\sigma}\right)}. \tag{6}$$

where μ and σ are the mean and standard deviation of the logarithm transform of X, and $Q(t) = \frac{1}{2\pi}\int_t^{+\infty} \exp\left(-\frac{x^2}{2}\right) dx$. It is shown in [14] that the hazard rate first increases from 0 to the maximum and then decreases asymptotically to 0. Thus, the equation $u(x) = \frac{1}{E[T_{\text{switch}}]}$ can have at most two roots, which can be obtained numerically. From this property, one can see that 0 and the second root are the local minimum of $\mathbb{E}[D(t_w)]$, and are the only candidates for the optimal maximal wait-time.

B. The Primary Occupancy Distribution $F(x)$ Is Unknown

We now discuss the scenario in which the SUs do not know the PU's active distribution. This may be more suitable if the geo-location database cannot provide the distribution, and hence, the SUs must estimate it while minimizing the disruption time on-the-fly. To make an accurate estimation, the SUs can choose to wait for a long time until the PU leaves, but suffering from a long disruption time. Or, the SUs can choose to switch shortly after the PU returned, thus achieving a short disruption time but lacking accurate estimation due to incomplete samples of the PU's busy duration. We formulate this problem in the form of sequential decision process to capture the tradeoff between exploration and exploitation, and seek to minimize the average disruption time of the SUs in the long run.

Let a decision epoch be the time when the primary channel the SUs are using opportunistically becomes busy. At each decision epoch, the SUs make a decision on the maximal wait-time. Whenever the SUs use this channel again, the next decision epoch would be the time when the primary channel becomes busy again. Since the PU's occupancy duration is a random variable X, given a maximal wait-time t_w, we have the following two possible outcomes. When $X \le t_w$, since the SUs are constantly sensing the channel, they can learn the exact realization of X. We call this outcome a *type*-0 observation. When $X > t_w$, the SUs will switch to a new channel at time t_w, so they only know the fact of $X > t_w$, which is an incomplete sample of X. We term this outcome a *type*-1 observation. We can therefore represent all the outcomes as a 2-tuple (I, t), where I is a single-bit flag and t a time duration. $I = 0$ represents a type-0 observation, under which t is the realization of X and $t \le t_w$. $I = 1$ represents a type-1 outcome, under which $t = t_w$. Let (I_i, t_i) $(i = 1, 2, \ldots, n)$ be the past outcomes up to n decision epochs.

For tractability, we assume that the PU's occupancy distribution $f(x)$ is exponential with an unknown parameter θ, i.e., $f_\theta(x) = \theta e^{-\theta x}$. We assume a Bayesian approach: the exponential parameter θ is a random variable, and its distribution, $f_I(\theta)$, also known as a prior distribution, follows the gamma distribution:

$$\begin{aligned} f_I(\theta) &\triangleq \text{Ga}(\theta; \alpha, \beta) \\ &= \theta^{\alpha-1}\beta^\alpha e^{-\theta\beta}\Gamma(\alpha)^{-1};\ \alpha > 0, \beta > 0. \end{aligned} \tag{7}$$

where (α, β) are the hyperparameters, and $\Gamma(x)$ is the gamma function. The gamma prior is a conjugate prior [15], i.e., the posterior distribution still follows the gamma distribution in our problem. To show this, let us first look at the likelihood function for the i-th outcome (I_i, t_i)

$$L(\theta|(I_i, t_i)) \propto \begin{cases} f_\theta(t_i) = \theta e^{-\theta t_i} & \text{if} \quad I_i = 0 \\ P_\theta(X > t_i) = e^{-\theta t_i} & \text{if} \quad I_i = 1. \end{cases}$$

The posterior distribution after the i-th outcome (I_i, t_i) can then be obtained by the Bayes formula:

$$\begin{aligned} f(\theta|(I_i, t_i)) &\propto f_I(\theta) L(\theta|I_i, t_i) \\ &\propto \begin{cases} \theta^\alpha e^{-\theta(t_i+\beta)} & \text{if} \quad I_i = 0 \\ \theta^{\alpha-1} e^{-\theta(t_i+\beta)} & \text{if} \quad I_i = 1. \end{cases} \end{aligned} \tag{8}$$

The form of the posterior density implies that it can only be a gamma density. The gamma hyperparameters are adjusted by

$$\begin{aligned} (\alpha, \beta) &\leftarrow (\alpha + 1, \beta + t_i) & \text{if} \quad I_i = 0 \\ (\alpha, \beta) &\leftarrow (\alpha, \beta + t_i) & \text{if} \quad I_i = 1 \end{aligned} \tag{9}$$

Thus, the gamma prior is a conjugate prior under the likelihood function. This conjugate property of the gamma prior in our problem can greatly facilitate subsequent analysis.

We now formulate the problem in the framework of Markov decision process (e.g., see [16],[17]). Let the decision epochs be $T = \{1, 2, \ldots, N\}$, $N \le \infty$. To represent the system state, the Bayesian approach allows us to represent it as the gamma hyperparameters (α, β) at each decision epoch, as it represents the prior distribution of θ and therefore, all the available information the SUs can know about the PU's occupancy distribution. The action the SUs can take is the maximally waiting at each decision epoch. The state-transition dynamic is the likelihood of transiting to the next state, i.e., another gamma hyperparameters, given current state and current action. We represent the dynamic as follows: given current state (α, β) and current maximal wait-time t_w, we may have two types of observation.

For type-0 observations, suppose the outcome is $(I, t) = (0, x)$ where $x \le t_w$, then the next state would be $(\alpha+1, \beta + x)$, with the transition probability density function averaged over the prior distribution as:

$$\begin{aligned} f((\alpha + 1, \beta + x)|(\alpha, \beta)) &= \int_0^\infty f_\theta(x) f_I(\theta) \mathrm{d}\theta \\ &= \alpha(x + \beta)^{-\alpha-1}\beta^\alpha;\ x \le t_w. \end{aligned} \tag{10}$$

For type-1 observations, the outcome can only be $(I, t) = (1, t_w)$, and the next state is $(\alpha, \beta + t + w)$ with the transition probability averaged over the prior distribution as:

$$\begin{aligned} P((\alpha, \beta + t_w)|(\alpha, \beta)) &= \int_0^\infty P_\theta(X > t_w) f_I(\theta) \mathrm{d}\theta \\ &= (t_w + \beta)^{-\alpha}\beta^\alpha. \end{aligned} \tag{11}$$

The immediate cost is the expected disruption time in current state (α, β) given an action t_w, averaged over the prior distribution:

$$d(t_w, \alpha, \beta) = \int_0^\infty \mathbb{E}_\theta[D(t_w)] f_I(\theta) \mathrm{d}\theta, \tag{12}$$

where $\mathbb{E}_\theta[D(t_w)]$ is the immediate disruption time under an exponential distribution with parameter θ from (2):

$$\mathbb{E}_\theta[D(t_w)] = e^{-\theta t_w} \left(\mathbb{E}[T_{\text{switch}}] - \frac{1}{\theta} \right) + \frac{1}{\theta}. \tag{13}$$

After some algebraic manipulations, we can represent $d(t_w, \alpha, \beta)$ as

$$\begin{cases} \frac{\beta}{\alpha-1} + \left(\frac{\beta}{t_w+\beta}\right)^\alpha \left(\mathbb{E}[T_{\text{switch}}] - \frac{t_w+\beta}{\alpha-1}\right) & \text{if } \alpha \neq 1 \\ \frac{\beta \mathbb{E}[T_{\text{switch}}]}{t_w+\beta} + \beta \log \frac{t_w+\beta}{\beta} & \text{if } \alpha = 1. \end{cases} \tag{14}$$

Now, we have the decision epochs $(i = 1, 2, \ldots, N)$, states, denoted as $S_i \triangleq (\alpha_i, \beta_i)$ ((α_i, β_i) are the gamma hyper-parameters at the i-th decision epoch), actions t_w^i (maximal wait-time at the i-th decision epoch), the system dynamics ((10) and (11)) and immediate expected cost $d(t_w^i, S_i)$, the only remaining part to completely describe a Markov decision process is the objective: find a sequence of actions, i.e., a policy, to minimize the overall cost, i.e., disruption time:

$$\underset{\pi}{\text{minimize}} \quad \mathbb{E}\left[\sum_{i=1}^{N} \gamma^{i-1} d(t_w^i, S_i) \middle| S_1 = (\alpha_1, \beta_1) \right], \tag{15}$$

where $\pi \triangleq (t_w^1, t_w^2, \ldots, t_w^N)$, γ $(0 \leq \gamma \leq 1)$ is the discount factor, and S_1 is the initial prior. The Markov property enables us to focus on policies that only depend on current state, instead of the entire history.

The standard technique for solving Markov decision process is to use dynamic programming. Let the optimal value function of state (α, β) starting at the n-th decision epoch be:

$$V_n(\alpha, \beta) = \min_\pi \mathbb{E}\left[\sum_{i=n}^{N} \gamma^{i-n} d(t_w^i, S_i) \middle| S_n = (\alpha, \beta) \right].$$

We can then characterize $V_n(\alpha, \beta)$ by relating it to the next states' value functions, considering all possible type 0 and 1 observations. (16) is the Bellman optimality equation, and the optimal policy, i.e., maximal wait-time is the minimizer of (16). It is trivial to modify (16) for the case of an infinite horizon by eliminating the subscripts of the value function, so it is omitted.

1) Myopic Policy: Solving (16) as well as storing the optimal policy suffers the curse of dimensionality in dynamic programming. The state and the action spaces are continuous and unbounded, so obtaining the exact solution is generally intractable. Therefore, we use a suboptimal but easily computable method. Our idea is to use a myopic strategy without considering the future actions. We will also show that this myopic policy is surprisingly close to the optimal policy in our evaluations.

Suppose at each decision epoch, the gamma hyperparameters are $S = (\alpha, \beta)$, we would like to find a maximal wait-time to minimize the immediate expected disruption time:

$$t_w^*(\alpha, \beta) = \arg\min_{t_w \geq 0} d(t_w, \alpha, \beta) = (\alpha \mathbb{E}[T_{\text{switch}}] - \beta)^+, \tag{17}$$

where $(.)^+$ is the projection on nonnegative real numbers. The minimizer in (17) is true for any $\alpha > 0$ in (14).

The policy can proceed as follows: at each decision epoch with a prior (α, β), choose the maximal wait-time determined by (17), then update the hyperparameters according to the observations by the rule (9), and then repeat the policy at the next decision epoch. We choose the initial prior to be $(\alpha, \beta) = (1, 0)$. The SUs only need to store two variables (α, β) in the memory and the update only takes 5 operations (9) and (17)). This myopic policy is therefore much simpler than the optimal policy and still preserves the learning of the PU's activity pattern.

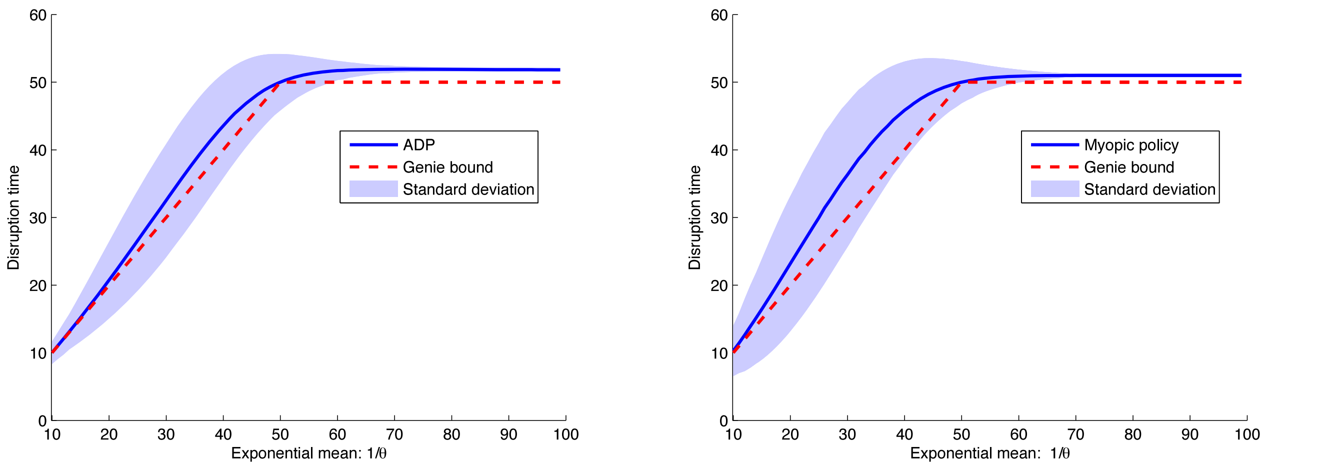
2) Evaluation of Optimal and Myopic Policies: Since obtaining the exact solutions of the optimal policy is intractable, we try to obtain approximate solutions as close to the optimal solution as possible. First, we discretize and bound all continuous variables. We let the action (i.e., maximal wait-time t_w) be in the range $[0, T_{\max}]$ and uniformly divide the interval into K points. Likewise, we discretize the state space (α, β) and also approximate the integration in the Bellman optimality equation (16). By choosing $T_{\max}$ and K large enough, we can get a close approximation to the exact solutions. We use backward induction to obtain the value function, from which the optimal policy is obtained. Under our approximate dynamic programming (ADP), the run-time complexity is $O(N^3 K^2)$ and space complexity is $O(N^3 K)$.

We first vary the exponential parameter θ and let the optimal and myopic policies run over the horizon. Therefore, they have to estimate the distribution while minimizing the average disruption time. Their performance is plotted in Fig. 2. The solid lines represent the long-term average disruption time, i.e., the mean of all N disruption time. We also compare them with a genie bound, in which there is a genie who knows the exact θ at no cost, and therefore, can obtain the smallest disruption time as $\min(1/\theta, 1/\mathbb{E}[T_{\text{switch}}])$ by Theorem 2.1. As one can see from the two figures, both policies' performance are close to that of the genie bound. The optimal policy has in general a smaller variance than the myopic policy. It is surprising that the myopic policy has better performance when $1/\theta$ is greater than $\mathbb{E}[T_{\text{switch}}]$.

For a direct comparison of the two policies, we generate θ from a gamma distribution with $(\alpha, \beta) = (1, \mathbb{E}[T_{\text{switch}}])$, and then run the two policies to obtain the average disruption time. We vary $\mathbb{E}[T_{\text{switch}}]$ and compare them in Fig. 3. We can see that the myopic policy is always within 1% of the approximated optimal policy.

We also evaluate its effectiveness on three non-exponential distributions, Weibull, gamma and truncated Gaussian, although the myopic policy is derived from an exponential distribution. Surprisingly, as can be seen from Fig. (4), the myopic algorithm yields a disruption time very close to the genie bound obtained via Theorem 2.1 (as the three distributions under study have increasing hazard rates). Considering the fact

$$V_n(\alpha,\beta)=\min_{t_w\geq 0}\{d(t_w,S_n)+\gamma\int_0^{t_w}V_{n+1}(\alpha+1,\beta+x)f((\alpha+1,\beta+x)|S_n)\mathrm{d}x+\gamma V_{n+1}(\alpha,\beta+t_w)P((\alpha,\beta+t_w)|S_n)\}. \quad (16)$$



(a) ADP (b) Myopic policy

Fig. 2. SUs' long-term disruption time achieved by using approximate dynamic programming (ADP) and myopic policies. The two policies run without knowing the PU's occupancy distribution. The parameters of the system are: $N = 50$, $\mathbb{E}[T_{\text{switch}}] = 50$ (unit of time), and $\gamma = 1$. The parameters of ADP are $T_{\max} = 1000$ (unit of time) and $K = 1000$. We choose the initial prior to be $(\alpha,\beta) = (1, \mathbb{E}[T_{\text{switch}}])$ when solving ADP.

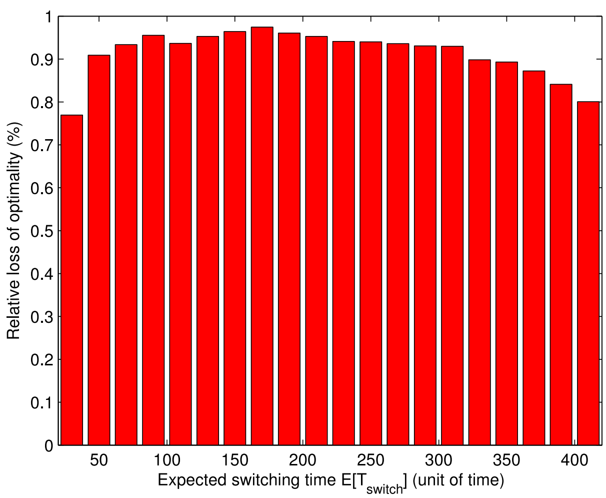


Fig. 3. Direct comparison of the myopic policy with the policy obtained by approximate dynamic programming in terms of long-term average disruption time $\overline{V}$. The relative loss stands for $(\overline{V}_{\text{Myopic}} - \overline{V}_{\text{ADP}})/\overline{V}_{\text{ADP}}$.

that (1) the computation overhead of ADP is large enough to inhibit practical usage and (2) the myopic policy is very simple to implement and works well even for non-exponential distributions, we advocate its use for this problem.

III. Efficient Detection of PUs' Busy → Idle Transition

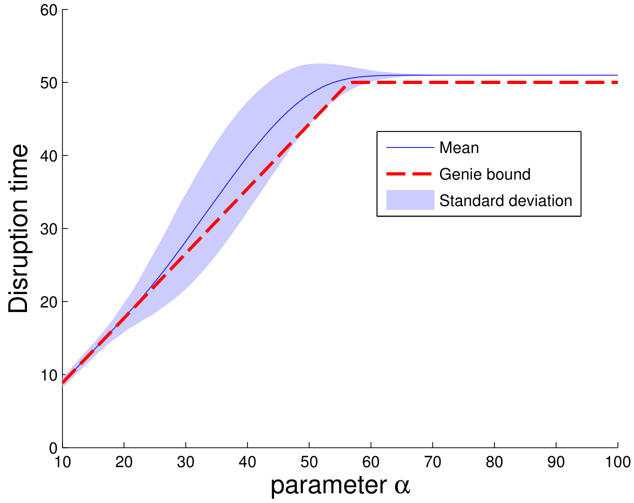
So far, we have discussed how long the SUs should wait before switching to an idle channel in case the PUs return to the channel the SUs are utilizing opportunistically. In case the SUs decide not to wait but to leave the current channel, they need to efficiently sense this out-of-band channel to detect when it becomes idle again.

The PU's departure from its channel can only be discovered by its occasional sensing by the SUs. The time gap between the departure and its detection is called the *discovery latency* as shown in Fig. 1. On the other hand, since SUs use time that would otherwise be used for in-band spectrum access, we should limit the average fraction of time used for the out-of-band sensing. Periodic sensing is inadequate for this since it only has one degree-of-freedom, i.e., using the same sensing interval. A better option is to use different sensing intervals to detect the time instants that are likely to be the end of the PU's active period, which we call *adaptive sensing*.

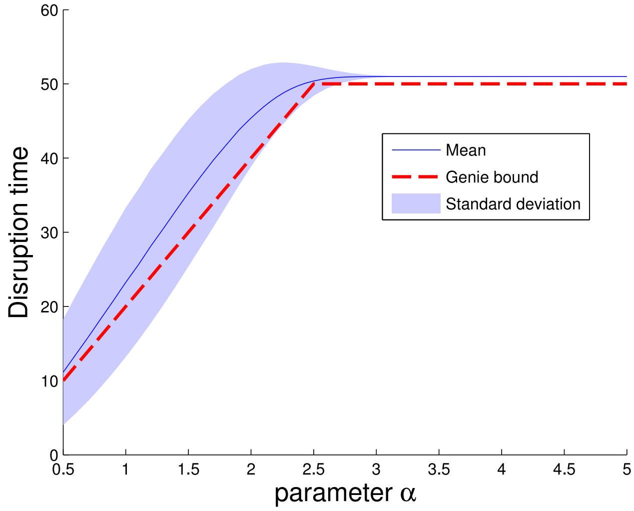
Beginning at time 0 when the SUs leave the current busy channel, we can characterize the remaining duration of the PU's stay, $\tilde{X}$, with c.d.f. $\tilde{F}(x)$:

$$\tilde{F}(x) = P(X \leq x + t_w | X > t_w) = \frac{F(x+t_w) - F(t_w)}{1 - F(t_w)} \quad (18)$$

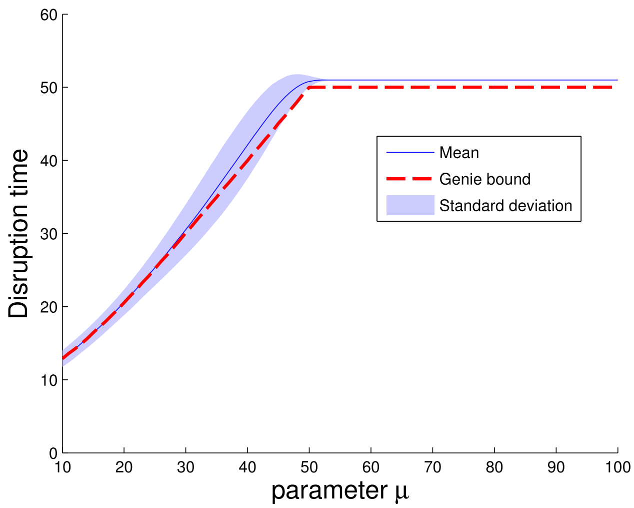
where t_w is the time when the SUs leave this channel. The SUs are not using this channel but sense the channel at time instants $t_0 = 0 < t_1 < t_2 < \ldots$. The p.d.f. $\tilde{f}(x)$ is supported at $[0, \lim_{n\to\infty} t_n)$. Every time the SUs sense the channel, they spend τ, and the average fraction of time spent on sensing should not exceed r_1, which is a parameter determined by the system designer. We assume the SUs sense and report the PU's presence/absence accurately, given sufficient sensing time τ. We will later discuss the impact of sensing errors. Here we focus on the impact of the PU's activity pattern on the scheduling of sensing.



(a) Weibull ($\beta = 2$)



(b) Gamma ($\beta^{-1} = 20$)



(c) Truncated Gaussian ($\sigma = 10$)

Fig. 4. Running the myopic algorithm on non-exponential distributions: Weibull, gamma and truncated Gaussian.

When the PU's remaining active period, $\tilde{X}$, ends within the interval $[t_{k-1}, t_k)$, the SUs detect the PU's departure at t_k with the discovery latency $t_k - \tilde{X}$. Since the SUs have sensed this channel k times, the fraction of time spent on sensing until then is $\frac{k\tau}{t_k}$. Considering all possibilities of $\tilde{X}$, we would like to minimize the average discovery delay upon detection of the channel to be idle:

$$\underset{0<t_1<t_2<\ldots}{\text{minimize}} \quad \sum_{k=1}^{\infty} \int_{t_{k-1}}^{t_k} (t_k - x)\tilde{f}(x)\mathrm{d}x \tag{19}$$

$$\text{subject to} \quad \sum_{k=1}^{\infty} \int_{t_{k-1}}^{t_k} \frac{k\tau}{t_k}\tilde{f}(x)\mathrm{d}x \le r_1. \tag{20}$$

Inequality (20) suggests the average fraction of time spent on sensing should be no greater than r_1. The above optimization is, in general, difficult to achieve since the channel can be sensed infinitely frequently, and is also non-convex for most of the distributions. In what follows, we take a suboptimal approach by introducing the concept of smooth sensing rate. We use the calculus of variations to obtain the optimal sensing rate and thus to approximate the optimal discrete times of sensing. The idea of using a smooth sensing rate is similar to that in [18, 19] using the continuous inspection rate to obtain the optimum inspection policies in reliability theory [12]. Let us introduce $n(x)$, a smooth function, to approximate the number of times to sense per unit of time. Then, $\int_0^t n(x)dx$ is the number of times to sense until time t. The key approximation is that two successive sensing times, t_k and t_{k+1}, are close enough such that $t_{k+1} - t_k \approx 1/n(t_k)$. So, we have the following approximation of the average discovery delay:

$$\sum_{k=1}^{\infty} \int_{t_{k-1}}^{t_k} (t_k - x)\tilde{f}(x)\mathrm{d}x \approx \sum_{k=1}^{\infty} \int_{t_{k-1}}^{t_k} \frac{1}{2n(x)}\tilde{f}(x)\mathrm{d}x = \int_0^{\infty} \frac{\tilde{f}(x)}{2n(x)}\mathrm{d}x. \tag{21}$$

The average fraction of time for the out-of-band sensing of this busy channel is approximately:

$$\int_0^{\infty} \frac{\tau \int_0^x n(t)\mathrm{d}t}{x}\tilde{f}(x)\mathrm{d}x = \tau \int_0^{\infty} n(t) \int_t^{\infty} \frac{\tilde{f}(x)}{x}\mathrm{d}x\mathrm{d}t. \tag{22}$$

By letting $g(x) = \int_x^{\infty} \frac{\tilde{f}(t)}{t}\mathrm{d}t$, we can formulate a functional optimization problem as

$$\underset{n(x)>0}{\text{minimize}} \quad \int_0^{\infty} \frac{\tilde{f}(x)}{2n(x)}\mathrm{d}x$$

$$\text{subject to} \quad \tau \int_0^{\infty} n(x)g(x)\mathrm{d}x \le r_1. \tag{23}$$

The above constrained optimization can be solved by minimizing the Lagrangian $L(n(x), \gamma)$ over a nonnegative continuous function $n(x)$:

$$L(n(x), \gamma) = \int_0^{\infty} \left(\frac{\tilde{f}(x)}{2n(x)} + \gamma\tau n(x)g(x) \right) \mathrm{d}x - \gamma r_1, \tag{24}$$

where γ is the Lagrange multiplier. The integrand has a lower bound achieved at

$$\frac{\tilde{f}(x)}{2n(x)} = \gamma\tau n(x)g(x) \quad \text{or} \quad n(x) = \sqrt{\frac{\tilde{f}(x)}{2\gamma\tau g(x)}}. \tag{25}$$

$n(x)$ should also exhaust the constraint (23), and substitute $n(x)$ into (23). We can then obtain the Lagrange multiplier γ as

$$\gamma = \frac{\tau \left(\int_0^\infty \sqrt{\tilde{f}(x) g(x)} \mathrm{d}x \right)^2}{2r_1^2}. \tag{26}$$

Therefore, $n(x)$ can be explicitly expressed as a function of $\tilde{f}(x)$. The sequence of discrete sensing times $\{t_k\}$ are obtained from $n(x)$:

$$\int_0^{t_k} n(x) \mathrm{d}x = k; \qquad k = 1, 2, \ldots \tag{27}$$

A. Evaluation of Adaptive Sensing

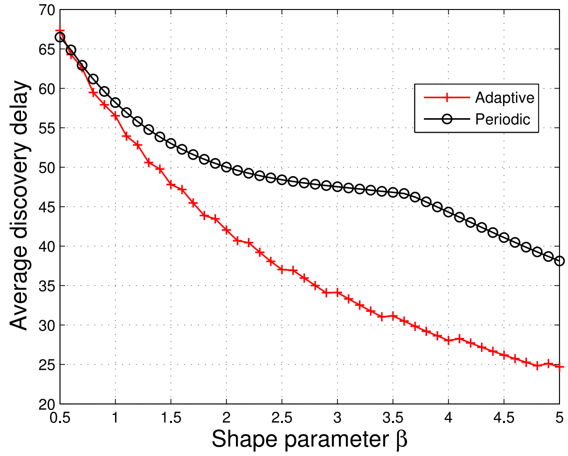
We now evaluate the performance of the above adaptive sensing by comparing it with periodic sensing. First, we seek to obtain an optimal periodic sensing strategy. Suppose the sensing period is x, then the sensing times are $t_k = kx$ ($k = 1, 2, \ldots$). Substituting t_k into the original optimization problem (20), we can solve a single variable nonlinear optimization.

Let us consider the cases when the PU's remaining busy duration follows three commonly used distributions for modeling positive duration: the Weibull distribution, the gamma distribution and the truncated Gaussian distribution:

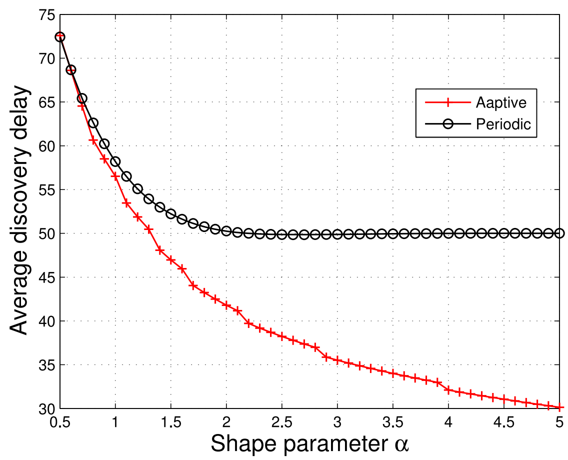
$$\begin{aligned} &\text{Weibull:} && \tilde{f}(x) = \beta \alpha^{-\beta} x^{\beta-1} e^{-\left(\frac{x}{\alpha}\right)^\beta}, \\ &\text{Gamma:} && \tilde{f}(x) = x^{\alpha-1} \beta^\alpha e^{-x\beta} \Gamma(\alpha)^{-1}, \\ &\text{Truncated Gaussian:} && \tilde{f}(x) = \frac{f_{\text{Gaussian}}(x; \mu, \sigma)}{1 - F_{\text{Gaussian}}(0; \mu, \sigma)}, \end{aligned}$$

where $f_{\text{Gaussian}}(x; \mu, \sigma)$ is a Gaussian density with mean μ and standard deviation σ.

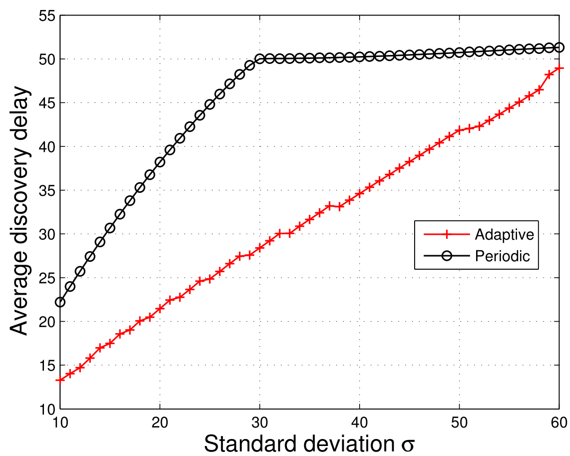
In our numerical examples, we fix the scale parameters and vary the shape parameters of all the distributions. The sensing time is set to $\tau = 1$ (unit of time), and we require the average sensing overhead not greater than 1%. We compare the proposed adaptive sensing with the optimal periodic sensing in Fig. 5. The adaptive sensing is shown to outperform the periodic sensing significantly in most situations. It is interesting to observe that when the shape parameters of the Weibull and gamma distributions are less than 1 (when the shape parameter is 1, both distributions reduce to exponential distributions), both distributions have a decreasing hazard rate and the performance gap turns out to be insignificant.



(a) Weibull ($\alpha = 100$)



(b) Gamma ($\beta^{-1} = 100$)



(c) Truncated Gaussian ($\mu = 100$)

Fig. 5. Comparison of adaptive sensing with periodic sensing in terms of discovery latency of the PU's busy→idle transition. The PU's busy duration follows three different distributions.

B. On the Impact of Sensing Errors

Either energy or feature detection, even with the sophisticated collaborative sensing, inevitably introduces sensing errors due to the receiver noise and channel fading, resulting in miss detection (reporting idle when the primary channel is in fact busy) and false alarm (reporting busy while the primary channel is idle). Here we discuss the impact of sensing errors on the scheduling of adaptive sensing. First, we will set the miss-detection error to be negligible by moving along the receiver operating characteristic curve. Thus, we only need to focus on false alarms, not on both, which would otherwise make the problem much more difficult to analyze. In order to combat false alarms, the SUs have to sense more frequently to capture the possibly missed opportunities. Increasing the sensing frequency incurs more overhead, so we need to re-design the sensing intervals to re-balance between minimizing discovery delay and limiting the sensing overhead.

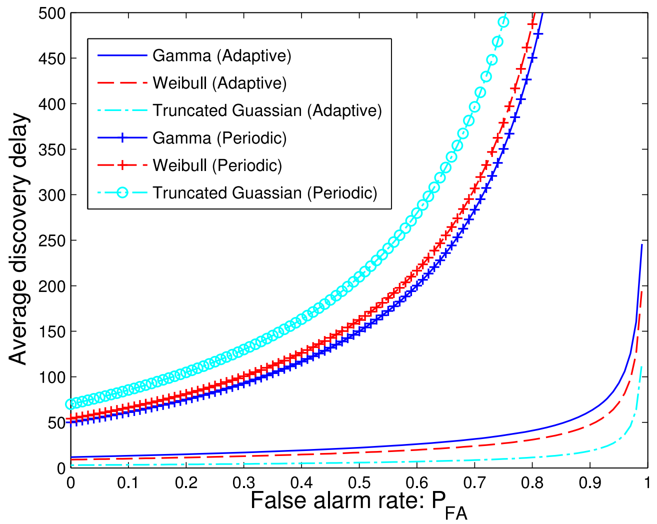


Fig. 6. Varying false alarm rate for three distributions: gamma ($\alpha = 2, \beta^{-1} = 200$), Weibull ($\alpha = 200, \beta = 1$), Truncated Gaussian ($\mu = 30, \sigma = 10$)).

Let P_{FA} ($0 \leq P_{FA} < 1$) be the false alarm probability. Here we continue to use the smooth sensing rate to approximate the discrete sensing times. Due to space limit, we omit the intermediate derivations and present the following final optimization problem:

$$\begin{aligned} \underset{n(x)>0}{\text{minimize}} \quad & \int_0^{\infty} \frac{1+P_{FA}}{2n(x)(1-P_{FA})}\tilde{f}(x)\mathrm{d}x \\ \text{subject to} \quad & \int_0^{\infty} n(x)(1-\tilde{F}(x))\mathrm{d}x + \frac{P_{FA}}{1-P_{FA}} \leq \\ & r_1\left(\mathbb{E}[X] + \int_0^{\infty} \frac{1+P_{FA}}{2n(x)(1-P_{FA})}\tilde{f}(x)\mathrm{d}x\right). \end{aligned} \tag{28}$$

The solution for the sensing rate is the minimizer of the Lagrangian and is obtained as

$$n(x) = \sqrt{\frac{(1-\gamma r_1)(1+P_{FA})\tilde{f}(x)}{2\gamma(1-P_{FA})(1-\tilde{F}(x))}}. \tag{29}$$

The Lagrange multiplier γ can be computed by substituting $n(x)$ from (29) into (28). Therefore, $n(x)$ can be explicitly expressed and the discrete sensing times can be obtained as in (27).

We evaluated the impact of sensing error for three different PUs' active distributions by plotting the solution of (29) and also periodic sensing in Fig. 6. If the false alarm rate is low (less than 0.5), the discovery delay under adaptive sensing is found not affected much. However, if the false alarm rate is very high, the discover delay is found to increase exponentially. As for periodic sensing the discovery delay is shown to be more sensitive to false alarm rate.

IV. Out-of-Band Sensing of Idle Primary Channels

Thus far, we discussed the sensing of an out-of-band busy channel. When this channel is detected to be idle but not chosen for opportunistic access, it should be used as a backup for future access. The idle channel still needs to be sensed occasionally, but more frequently than a busy channel in order to provide reliable estimation of the channel's availability for future usage. Periodic sensing thus becomes attractive here, as it can only incur a marginal loss in discovery latency compared to adaptive sensing. Therefore, all out-of-band idle channels are used as backups, and sensed periodically to estimate the channels' availability. We focus on how to efficiently allocate the sensing periods of different idle channels such that the evicted SUs can use the sensing-based estimation to switch to an idle channel very quickly.

When the evicted SUs decide to switch to an idle channel, they can use various algorithms to search for the target channel [2–4] based on the statistics of the channels' availability. Here we do not address the search algorithms. We address how to optimize the sensing in order to benefit from these algorithms.

The key assumption here is that the duration of an idle channel is i.i.d. exponentially-distributed, which reflects that the PUs' arrivals at the channel follow a Poisson process. In actual cellular networks, inter-arrival times of calls have also been reported to be exponentially-distributed [11].

Suppose there are M idle primary channels and let $\frac{1}{\lambda_i}$ be the mean duration of each idle channel. Channel i is sensed at the interval of T_i, and each sensing takes τ_i with a perfect report. For each idle channel i, let $\tilde{t_i}$, ($0 < \tilde{t_i} < T_i$) denote the time elapsed since the last sample. Then, the probability of detecting the channel to be idle at current time is $P_i(\tilde{t_i}) = e^{-\lambda_i \tilde{t_i}}$. We are interested in its average over the elapsed time $\tilde{t_i}$, and the average probability of detecting channel i to be idle is

$$\overline{P_i} = \frac{1}{T_i}\int_0^{T_i} P_i(\tilde{t_i})d\tilde{t_i} = \frac{1-e^{-\lambda_i T_i}}{\lambda_i T_i}. \tag{30}$$

Without loss of generality, suppose the SUs that decide to switch will sense and search for the out-of-band idle channels in the order of $\{1, 2, \ldots, M\}$. Once the SUs discover an idle channel, it takes a constant time t_{setup} to set up new connections among themselves and the search process ends. If all searched idle channels are found busy, the SUs will wait for a mandatory retry time, t_{retry}, and then search again for an idle channel. The retry time should be much greater than the subsequent search time. Let $\tau_{M+1} = t_{\text{retry}}$.

When the SUs search and detect channel i first to be idle, it takes a total time $\sum_{j=1}^{i}\tau_j + t_{\text{setup}}$ with probability $\prod_{j=1}^{i-1}(1-\overline{P_j})\overline{P_i}$. The switching delay, which is the time for the SUs to set up new connections when they decide to switch, has the expected value as

$$\mathbb{E}[T_{\text{switch}}] = \sum_{i=1}^{M+1}\left(\prod_{j=1}^{i-1}(1-\overline{P_j})\right)\overline{P_i}\sum_{j=1}^{i}\tau_j + t_{\text{setup}}. \tag{31}$$

We would like to minimize this average search time $\mathbb{E}[T_{\text{switch}}]$, while limiting the overall sensing overhead of all out-of-band idle channels to the ratio r_2. We can then formulate the optimization problem as

$$\begin{aligned} \underset{T_1, T_2, \ldots, T_M}{\text{minimize}} \quad & \mathbb{E}[T_{\text{switch}}] = \sum_{i=1}^{M+1}\tau_i\prod_{j=1}^{i-1}(1-\overline{P_j}) + t_{\text{setup}} \\ \text{subject to} \quad & \sum_{i=1}^{M}\frac{\tau_i}{T_i} \leq r_2. \end{aligned} \tag{32}$$

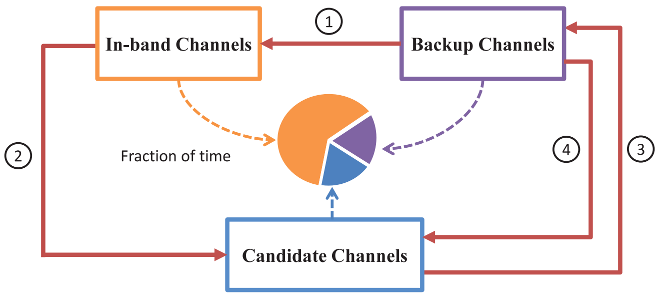


Fig. 7. Spectrum-management architecture for CRNs. Backup (candidate) channels are out-of-band idle (busy) channels.

$1-\overline{P_i}$ is found to be a concave and increasing function of T_i. Therefore, the objective function is not convex. Minimizing such a function can be very difficult. However, since the sensing period should be much smaller than the mean idle duration, we have $\lambda_i T_i \ll 1$ and a very good approximation for $e^{-\lambda_i T_i}$ as $1-\lambda_i T_i + \frac{(\lambda_i T_i)^2}{2}$). The objective function is then approximated as

$$\mathbb{E}[T_{\text{switch}}] \approx \sum_{i=1}^{M+1} \frac{\tau_i}{2^{i-1}} \prod_{j=1}^{i-1} \lambda_j T_j + t_{\text{setup}}. \quad (33)$$

Both the objective (33) and the constraint (32) functions are posynomials, and the problem becomes a geometric programming (GP) problem, which is in fact a convex optimization. GP can be solved efficiently and reliably using, for example, the interior-point method [20]. On the other hand, the original problem (32) without the approximation is a non-convex optimization. Solving it requires tricky tuning of parameters and always leads to a local optimal solution.

V. Putting Them All Together

We have shown that the state of a primary channel—active or idle, in-band or out-of-band—can lead to very different strategies. We now present a spectrum-management architecture for CRNs that deal with different channel states by integrating individual modules. The overall architecture is depicted in Fig. 7.

All primary channels are managed by one of the three modules: the *in-band channels module* for in-band spectrum access and sensing, the *backup channels module* for out-of-band sensing of idle channels, and the *candidate channels module* for out-of-band sensing of busy channels. The names of backup and candidate channels are borrowed from IEEE 802.22, a standard near completion for opportunistic utilization of TV channels [21].

The in-band channels module manages all primary channels that are currently being used opportunistically by the SUs. This module requires SUs to access the primary channels and limit interference to the returning PUs as specified in IEEE 802.22. We do not address the in-band sensing and access here (e.g., see [22]). Once the SUs detect a primary signal, they wait for a maximal amount of time before switching to a new channel. The maximal wait-time is computed as in Section II. During the wait, if the current evicted channel becomes idle, the discovery process ends. Otherwise, the SUs will switch to a new idle channel according to the estimated channels' availability provided by the backup channels module that performs the out-of-band sensing of idle channels. Arrow (1) in Fig. 7 indicates such an information flow and the new idle channel that is imported to the in-band channels module. At the same time, if SUs decide to switch to another channel, the original busy channel will be sent to the candidate channels module for out-of-band sensing of busy channels, as indicated in arrow (2).

The candidate channels module allows SUs to sense every busy channel economically, in the order of time, computed in Section III to minimize the delay in discovering the PU's actual departure. Once a channel is sensed to be idle, it will be sent to the backup channels module as indicated in arrow (3).

The backup channel module periodically senses all idle channels, and the sensing periods are optimized, as in Section IV, to provide reliable estimation of channel availability to the in-band channel module and minimize the switching delay when the SUs decide to switch channel. Once a channel is sensed busy, it will be sent to the candidate channels module in arrow (4).

The objective of the architecture is to minimize the disruption time of the SUs when the PUs return. Each module seeks to optimize this objective directly or indirectly. Specifically, the in-band channels module tries to minimize the disruption time when the SUs are evicted, which, in turn, depends on the switching delay. The backup channels module, therefore, attempts to minimize the switching delay, which can be reduced if there are more idle channels in the backup channels module. The candidate channels module, therefore, tries to discover a busy channel that becomes idle, as quickly as possible. This decomposition of tasks is inherently not optimal, but provides manageable modularity that allows us to optimize each module independently.

The three modules are operating in a time-division manner. The fraction of time allocated to each module is controllable. We can adjust the sensing overhead for each out-of-band busy channel as well as the overall fraction of time in sensing out-of-band idle channels. Therefore, system designers can make a tradeoff between exploitation (the in-band channels module) and exploration (the backup and candidate channels modules) of dynamic spectrum opportunities.

VI. Related Work

Finding temporal opportunities is a classical problem in CRNs. Existing approaches to this problem may be divided into two categories. The first category considered a slotted structure of the SUs' spectrum sensing and access [5–8, 23–25]. The objective is to maximize the SUs' expected rewards in terms of effective utilization of idle channels. They mainly focus on the SUs' response to the primary's activity, whether being one-slot memory [7, 8, 23] or memory-less [5, 6, 24, 25], whether the statistics are known [7, 23, 25] or unknown [5, 6, 8, 24]. Under this slotted structure, when the SUs sense the channel to be busy, they will be forced, by the mathematical model, to wait until the next slot to make a decision. This can be very inefficient in gathering the channels' availability information if the SUs' sensing time is much smaller than the data transmission time. Our work belongs to the second category which considered use of an unslotted sensing and transmission strategy, [2–4]. However, all the previous approaches assume that whenever the SUs sense a channel that is busy (or probe the data rate of the channel to be undesirable), the SUs will immediately leave the channel. In fact, the SUs may choose to stay silently in the evicted band for future reuse if the primary busy duration lasts relatively short. We derive the optimal time for the SUs to switch and propose a learning strategy to estimate the primary's activity pattern while minimizing the disruption time of the SUs on-the-fly.

Out-of-band sensing is necessary to explore dynamic spectrum opportunity. Rather than using periodic sensing in [9], we use an adaptive sensing to efficiently detect a busy primary channel to be idle as quickly as possible.

The in-band spectrum sensing and access has been addressed in [22, 26–28] for efficient utilization of the currently-available white space while protecting the legacy users. Many sensing techniques have been proposed to accurately detect the presence/absence of a primary signal, such as collaborative sensing [29, 30]. We do not address the in-band access and sensing here.

Note that the measurement results in cellular networks [11] guided our approach to sense the primary's busy channel adaptively and economically. A CRN is implemented in [31] on top of the DTV white space, with an additional scanner sensing all channels. In this paper, we assume the SUs are equipped with a single radio that can either sense or access a channel at a time, not both.

VII. Conclusion

In this paper, we have developed strategies for SUs when the PUs return to the channel the SUs are using opportunistically. It would be interesting to consider an extension of the proposed approach for multiple secondary networks. One challenging issue associated with this extension is how to share the sensing and control information among different secondary networks. If there is a dedicated (e.g., control) channel for SUs to exchange information, one may be able to achieve a better utilization. For example, secondary networks can exploit receiver diversity [29], and different secondary networks can switch to different primary channels when the PUs return. These are matters of our future inquiry.

References

[1] FCC, "The Second Memorandum Opinion and Order," Sep 2009.

[2] Hyoil Kim and Kang G. Shin, "Fast Discovery of Spectrum Opportunities in Cognitive Radio Networks," in *Proc. IEEE DySPAN*, Oct 2008, pp. 1–12.

[3] Nicholas B. Chang and Mingyan Liu, "Optimal Channel Probing and Transmission Scheduling for Opportunistic Spectrum Access," in *Proc. ACM MobiCom*, Sep 2007.

[4] Tao Shu and Marwan Krunz, "Throughput-efficient Sequential Channel Sensing and Probing in Cognitive Radio Networks under Sensing Errors," in *Proc. ACM MobiCom*, Sep 2009.

[5] Animashree Anandkumar, Nithin Michael, and Ao Tang, "Opportunistic Spectrum Access with Multiple Users: Learning under Competition," in *Proc. of IEEE INFOCOM*, San Deigo, USA, Mar 2010.

[6] Lifeng Lai, Hesham El Gamal, Hai Jiang, and H. Vincent Poor, "Cognitive Medium Access: Exploration, Exploitation and Competition," *IEEE Transactions on le Computing*, vol. 10, no. 2, pp. 239–253, Feb 2011.

[7] Qing Zhao, Lang Tong, Ananthram Swami, and Yunxia Chen, "Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework," *IEEE Journal on Selected Areas in Communications (JSAC)*, vol. 25, no. 3, pp. 589–600, Apr 2007.

[8] Cem Tekin and Mingyan Liu, "Online Learning in Opportunistic Spectrum Access: A Restless Bandit Approach," in *Proc. of IEEE INFOCOM*, Shanghai, China, Apr 2011.

[9] Hyoil Kim and Kang G. Shin, "Efficient Discovery of Spectrum Opportunities with mac-layer Sensing in Cognitive Radio Networks," *IEEE Transactions on Mobile Computing*, vol. 7, no. 5, pp. 533–545, May 2008.

[10] Yann-Hang Lee and Kang G. Shin, "Optimal Design and Use of Retry in Fault-tolerant Computer Systems," *Journal of the ACM*, vol. 1, no. 35, pp. 45–69, Jan 1988.

[11] Daniel Willkomm, Sridhar Machiraju, Jean Bolot, and Adam Wolisz, "Primary Users in Cellular Networks: A Large-scale Measurement Study," in *Proc. IEEE DySPAN*, Oct 2008, pp. 1–11.

[12] Richard E. Barlow and Frank Proschan, *Mathematical Theory of Reliability*. Wiley, New York, 1965.

[13] Dawei Chen, Sixing Yin, Qian Zhang, Mingyan Liu, and Shufang Li, "Mining Spectrum Usage Data: A Large-scale Spectrum Measurement Study," in *Proc. ACM MobiCom*, Sep 2009.

[14] Arnold L. Sweet, "On the Hazard Rate of the Lognormal Distribution," *IEEE Transactions on Reliability*, vol. 39, no. 3, pp. 325–328, Aug 1990.

[15] Jose M. Bernardo and Adrian F. M. Smith, *Bayesian Theory*. John Wiley, 1994.

[16] Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning: An Introduction*. MIT Press, 1998.

[17] Martin L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley, 1994.

[18] Joseph B. Keller, "Optimum Checking Schedules for Systems Subject to Random Failure," *Management Science*, vol. 21, no. 3, pp. 256–260, Nov 1974.

[19] Naoto Kaio and Shunji Osaki, "Some Remarks on Optimum Inspection Policies," *IEEE Transactions on Reliability*, vol. R-33, no. 4, pp. 277–279, Oct 1984.

[20] Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[21] IEEE 802.22 Working Group on Wireless Regional Area Networks, http://www.ieee802.org/22/.

[22] Hyoil Kim and Kang G. Shin, "In-band Spectrum Sensing in Cognitive Radio Networks: Energy Detection or Feature Detection?" in *Proc. ACM MobiCom*, Sep 2008.

[23] Sahand Haji Ali Ahmad, Mingyan Liu, Tara Javidi, Qing Zhao, and Bhaskar Krishnamachari, "Optimality of Myopic Sensing in Multichannel Opportunistic Access," *IEEE Transactions on Information Theory*, vol. 55, no. 9, pp. 4040–4050, Sep 2009.

[24] Beycan Kahraman and Feza Buzluca, "A Novel Channel Handover Strategy To Improve the Throughput in Cognitive Radio Networks," in *Wireless Communications and Mobile Computing Conference (IWCMC)*, Jul 2011.

[25] Xue Feng, Qu Daiming, Zhu Guangxi, and Li Yanchun, "Smart Channel Switching in Cognitive Radio Networks," in *2nd Intl. Congress on Image and Signal Processing (CISP)*, Oct 2009.

[26] Ying Chang Liang, Yonghong Zeng, Edward C.Y. Peh, and Anh Tuan Hoang, "Sensing-throughput Tradeoff for Cognitive Radio Networks," *IEEE Transactions on Wireless Communications*, vol. 7, no. 5, pp. 1326–1337, Apr 2008.

[27] Senhua Huang, Xin Liu, and Zhi Ding, "Opportunistic Spectrum access in Cognitive Radio Networks," in *Proc. of IEEE INFOCOM*, 2008.

[28] Ashwini Kumar, "Enhancing Coexistence, Quality of Service, and Energy Performance in Dynamic Spectrum Access Networks," Ph.D. dissertation, The University of Michigan, 2010.

[29] Danijela Cabric, Artem Tkachenko, and Robert W. Brodersen, "Experimental Study of Spectrum Sensing based on Energy Detection and Network Cooperation," in *ACM TAPAS*, 2006.

[30] Amir Ghasemi and Elvino S. Sousa, "Collaborative Spectrum Sensing for Opportunistic Access in Fading Environments," in *Proc. IEEE DySPAN*, Nov 2005, pp. 131–136.

[31] Paramvir Bahl, Ranveer Chandra, Thomas Moscibroda, Rohan Murty, and Matt Welsh, "White Space Networking with Wi-Fi like Connectivity," in *ACM SIGCOMM*, Aug 2009.

REFERENCES

[1] FCC, "The Second Memorandum Opinion and Order," Sep 2009.

[2] Hyoil Kim and Kang G. Shin, "Fast Discovery of Spectrum Opportunities in Cognitive Radio Networks," in *Proc. IEEE DySPAN*, Oct 2008, pp. 1–12.

[3] Nicholas B. Chang and Mingyan Liu, "Optimal Channel Probing and Transmission Scheduling for Opportunistic Spectrum Access," in *Proc. ACM MobiCom*, Sep 2007.

[4] Tao Shu and Marwan Krunz, "Throughput-efficient Sequential Channel Sensing and Probing in Cognitive Radio Networks under Sensing Errors," in *Proc. ACM MobiCom*, Sep 2009.

[5] Animashree Anandkumar, Nithin Michael, and Ao Tang, "Opportunistic Spectrum Access with Multiple Users: Learning under Competition," in *Proc. IEEE INFOCOM*, San Deigo, USA, Mar 2010.

[6] Lifeng Lai, Hesham El Gamal, Hai Jiang, and H. Vincent Poor, "Cognitive Medium Access: Exploration, Exploitation and Competition," *IEEE Trans. le Computing*, vol. 10, no. 2, pp. 239–253, Feb 2011.

[7] Qing Zhao, Lang Tong, Ananthram Swami, and Yunxia Chen, "Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework," *IEEE J. Sel. Areas Commun. (JSAC)*, vol. 25, no. 3, pp. 589–600, Apr 2007.

[8] Cem Tekin and Mingyan Liu, "Online Learning in Opportunistic Spectrum Access: A Restless Bandit Approach," in *Proc. IEEE INFOCOM*, Shanghai, China, Apr 2011.

[9] Hyoil Kim and Kang G. Shin, "Efficient Discovery of Spectrum Opportunities with mac-layer Sensing in Cognitive Radio Networks," *IEEE Trans. Mobile Computing*, vol. 7, no. 5, pp. 533–545, May 2008.

[10] Yann-Hang Lee and Kang G. Shin, "Optimal Design and Use of Retry in Fault-tolerant Computer Systems," *J. ACM*, vol. 1, no. 35, pp. 45–69, Jan 1988.

[11] Daniel Willkomm, Sridhar Machiraju, Jean Bolot, and Adam Wolisz, "Primary Users in Cellular Networks: A Large-scale Measurement Study," in *Proc. IEEE DySPAN*, Oct 2008, pp. 1–11.

[12] Richard E. Barlow and Frank Proschan, *Mathematical Theory of Reliability*. Wiley, New York, 1965.

[13] Dawei Chen, Sixing Yin, Qian Zhang, Mingyan Liu, and Shufang Li, "Mining Spectrum Usage Data: A Large-scale Spectrum Measurement Study," in *Proc. ACM MobiCom*, Sep 2009.

[14] Arnold L. Sweet, "On the Hazard Rate of the Lognormal Distribution," *IEEE Trans. Reliab.*, vol. 39, no. 3, pp. 325–328, Aug 1990.

[15] Jose M. Bernardo and Adrian F. M. Smith, *Bayesian Theory*. John Wiley, 1994.

[16] Richard S. Sutton and Andrew G. Barto, *Reinforcement Learning: An Introduction*. MIT Press, 1998.

[17] Martin L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley, 1994.

[18] Joseph B. Keller, "Optimum Checking Schedules for Systems Subject to Random Failure," *Management Science*, vol. 21, no. 3, pp. 256–260, Nov 1974.

[19] Naoto Kaio and Shunji Osaki, "Some Remarks on Optimum Inspection Policies," *IEEE Trans. Reliab.*, vol. R-33, no. 4, pp. 277–279, Oct 1984.

[20] Stephen Boyd and Lieven Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.

[21] IEEE 802.22 Working Group on Wireless Regional Area Networks, http://www.ieee802.org/22/.

[22] Hyoil Kim and Kang G. Shin, "In-band Spectrum Sensing in Cognitive Radio Networks: Energy Detection or Feature Detection?" in *Proc. ACM MobiCom*, Sep 2008.

[23] Sahand Haji Ali Ahmad, Mingyan Liu, Tara Javidi, Qing Zhao, and Bhaskar Krishnamachari, "Optimality of Myopic Sensing in Multichannel Opportunistic Access," *IEEE Trans. Inf. Theory*, vol. 55, no. 9, pp. 4040–4050, Sep 2009.

[24] Beycan Kahraman and Feza Buzluca, "A Novel Channel Handover Strategy To Improve the Throughput in Cognitive Radio Networks," in *Wireless Communications and Mobile Computing Conference (IWCMC)*, Jul 2011.

[25] Xue Feng, Qu Daiming, Zhu Guangxi, and Li Yanchun, "Smart Channel Switching in Cognitive Radio Networks," in *2nd Intl. Congress on Image and Signal Processing (CISP)*, Oct 2009.

[26] Ying Chang Liang, Yonghong Zeng, Edward C.Y. Peh, and Anh Tuan Hoang, "Sensing-throughput Tradeoff for Cognitive Radio Networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 5, pp. 1326–1337, Apr 2008.

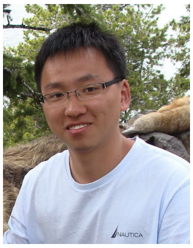
[27] Senhua Huang, Xin Liu, and Zhi Ding, "Opportunistic Spectrum access in Cognitive Radio Networks," in *Proc. IEEE INFOCOM*, 2008.

[28] Ashwini Kumar, "Enhancing Coexistence, Quality of Service, and Energy Performance in Dynamic Spectrum Access Networks," Ph.D. dissertation, The University of Michigan, 2010.

[29] Danijela Cabric, Artem Tkachenko, and Robert W. Brodersen, "Experimental Study of Spectrum Sensing based on Energy Detection and Network Cooperation," in *ACM TAPAS*, 2006.

[30] Amir Ghasemi and Elvino S. Sousa, "Collaborative Spectrum Sensing for Opportunistic Access in Fading Environments," in *Proc. IEEE DySPAN*, Nov 2005, pp. 131–136.

[31] Paramvir Bahl, Ranveer Chandra, Thomas Moscibroda, Rohan Murty, and Matt Welsh, "White Space Networking with Wi-Fi like Connectivity," in *ACM SIGCOMM*, Aug 2009.



Caoxie Zhang received his B.Eng. degree in 2009 from Shanghai Jiaotong University, China, and his M.S. degree in 2011 from the University of Michigan, Ann Arbor. He is currently a Ph.D. student in the Department of Electrical Engineering and Computer Science, University of Michigan.

Kang G. Shin is the Kevin and Nancy O'Connor Professor of Computer Science in the Department of Electrical Engineering and Computer Science, The University of Michigan. His current research focuses on QoS-sensitive computing and networking as well as on embedded real-time and cyber-physical systems. He has supervised the completion of 73 PhDs, and authored/coauthored about 800 technical articles, a textbook, and more than 20 patents or invention disclosures, and received numerous best paper awards, including the Best Paper Awards from the 2011 ACM International Conference on Mobile Computing and Networking (MobiCom'11), the 2011 IEEE International Conference on Autonomic Computing, the 2010 and 2000 USENIX Annual Technical Conferences, as well as the 2003 IEEE Communications Society William R. Bennett Prize Paper Award and the 1987 Outstanding IEEE Transactions of Automatic Control Paper Award. He has also received several institutional awards, including the Research Excellence Award in 1989, Outstanding Achievement Award in 1999, Distinguished Faculty Achievement Award in 2001, and Stephen Attwood Award in 2004 from The University of Michigan (the highest honor bestowed to Michigan Engineering faculty); a Distinguished Alumni Award of the College of Engineering, Seoul National University in 2002; 2003 IEEE RTC Technical Achievement Award; and 2006 Ho-Am Prize in Engineering (the highest honor bestowed to Korean-origin engineers).