Load-Balancing Spectrum Decision for Cognitive Radio Networks

Li-Chun Wang, *Fellow, IEEE,* Chung-Wei Wang, *Student Member, IEEE,* and Fumiyuki Adachi, *Fellow, IEEE*

***Abstract*—In this paper, we present an analytical framework to design system parameters for load-balancing multiuser spectrum decision schemes in cognitive radio (CR) networks. Unlike the non-load-balancing methods that multiple secondary users may contend for the same channel, the considered load-balancing schemes can distribute the traffic loads of secondary users to multiple channels. Based on the preemptive resume priority (PRP) M/G/1 queueing theory, a spectrum decision analytical model is proposed to evaluate the effects of multiple interruptions from the primary user during each link connection, the sensing errors (i.e., missed detection and false alarm) of the secondary users, and the heterogeneous channel capacity. With the objective of minimizing the overall system time of the secondary users, we derive the optimal number of candidate channels and the optimal channel selection probability for the sensing-based and the probability-based spectrum decision schemes, respectively. We find that the probability-based scheme can yield a shorter overall system time compared to the sensing-based scheme when the traffic loads of the secondary users is light, whereas the sensing-based scheme performs better in the condition of heavy traffic loads. If the secondary users can intelligently adopt the best spectrum decision scheme according to sensing time and traffic conditions, the overall system time can be improved by 50% compared to the existing methods.**

***Index Terms*—Cognitive radio, spectrum decision, channel selection, overall system time, preemption, queueing theory.**

I. Introduction

COGNITIVE radio (CR) techniques improve spectrum efficiency by allowing the low-priority secondary users to temporarily utilize the unused licensed spectrum of the high-priority primary users [1]–[6]. However, the secondary users need to vacate the occupied channel when the primary users have data to transmit on this channel because the primary users have the preemptive priority to access channel. In order to provide reliable transmission for the secondary users, *spectrum handoff* procedures are initiated to help the secondary user return the channel to the primary user and resume the secondary user's unfinished transmission at other channel or at the same channel after the completion of the primary transmissions [7]–[9].

Manuscript received 3 December 2009; revised 1 June 2010. This work was supported in part by the MoE ATU Plan (in Taiwan), by the National Science Council (NSC) grants 96-2628-E-009-004-MY3, 97-2221-E-009-099-MY3, and 97-2917-I-009-109 (in Taiwan), and by the Globe COE Program (in Japan). Part of this work was presented at the IEEE Personal, Indoor and Mobile Radio Communications Symposium (PIMRC), 2009.

L.-C. Wang and C.-W. Wang are with the Department of Electrical Engineering, National Chiao Tung University, Hsinchu, Taiwan (e-mail: lichun@cc.nctu.edu.tw and hyper.cm91g@nctu.edu.tw).

F. Adachi is with the Department of Electrical and Communication Engineering at the Graduate School of Engineering, Tohoku University, Sendai, Japan (e-mail: adachi@ecei.tohoku.ac.jp).

Digital Object Identifier 10.1109/JSAC.2011.110408.

Spectrum decision is a crucial process in CR networks [10], which helps the secondary user select the best channel to transmit data from candidate channels. In order to distribute the traffic loads of the secondary users evenly to these candidate channels, an effective spectrum decision scheme should take the traffic statistics of the primary users as well as the secondary users into account. In this paper, we introduce a performance measure for evaluating various spectrum decision schemes – the overall system time of the secondary connection, which is defined as the duration from the instant that data arrives at system until the instant of finishing the whole transmission.

The overall system time of the secondary users' connections is affected by the multiple interruptions from the primary users, the sensing errors like missed detection and false alarm for the primary users, and the heterogeneous channel capacity. Within the transmission period of a secondary connection, it is likely to have multiple spectrum handoffs due to the interruptions from the primary users. Clearly, multiple spectrum handoffs will increase the overall system time [11]. In the meanwhile, false alarm occurs when the detector mistakenly reports the presence of a primary user. In this situation, the overall system time of the secondary user's connections becomes longer because the secondary users cannot transmit data even with an idle channel. When the detection of a primary user is missed, data collision of both the primary user and the secondary user occurs, resulting in retransmitting and prolonging the overall system time of the secondary users' connections. Furthermore, different channels may have various capacity and data transmission rate, thereby resulting in different service time and overall system time for the secondary users. Hence, it is crucial to incorporate the effects of multiple handoffs, sensing errors, and heterogeneous channel capacity in spectrum decision methods for CR networks.

In this paper, two kinds of spectrum decision schemes are considered: (1) the sensing-based spectrum decision scheme; and (2) the probability-based spectrum decision scheme. For the sensing-based spectrum decision method, a secondary user selects its operating channel according to the *instantaneous* sensing results from scanning the wideband spectrum. For the probability-based spectrum decision method, the operating channel is selected based on the predetermined probabilities which are determined according to traffic statistics from the *long-term* observation. Note that the goal of load-balancing spectrum decision can be achieved by considering the sensing outcomes when designing system parameters because these sensing outcomes in both the methods are related to the traffic statistics of both the primary users and the secondary users.

The two considered spectrum decision schemes have different design issues. For the sensing-based spectrum decision scheme, the total number of candidate channels for channel selection significantly affects the overall system time because this scheme requires scanning all the candidate channels. Intuitively, a narrowband sensing (or a smaller number of candidate channels) can reduce the total sensing time. However, it is difficult to find one idle channel from a small number of candidate channels. Hence, one challenge is to determine the optimal number of candidate channels to minimize the overall system time. On the other hand, the probability-based spectrum decision scheme needs to prevent the secondary users from selecting a busy channel. Hence, the most important issue is to determine the optimal channel selection probability to minimize the overall system time.

In this paper, we investigate how to evaluate the overall system time for the sensing-based and the probability-based spectrum decision schemes in the CR network when multiple interruptions from the primary user, sensing errors, and channel capacity are taken into account. To this end, we design our multiuser spectrum decision schemes on top of the preemptive resume priority (PRP) M/G/1 queueing model. Based on the proposed analysis-based framework, we can design the suitable parameters to shorten the overall system time. Unlike the non-load-balancing methods that multiple secondary users may contend for the same channel, the channel selection schemes based on the designed parameters of the proposed analytical model can evenly distribute the traffic loads of secondary users to multiple channels, thereby reducing the average overall system time. The major contributions of this paper are summarized in the following:

- Derive the optimal selection probability for the probability-based channel selection method.
- Develop a method to determine the optimal number of candidate channels for the sensing-based channel selection method.
- Compare the sensing-based and the probability-based channel selection methods and suggest which spectrum decision scheme can result in shorter overall system time with various sensing error probabilities and traffic parameters.
- Characterize the effects of sensing errors on the spectrum decision schemes of CR networks in terms of the overall system time of the primary and the secondary connections.

The rest of this paper is organized as follows. Section II reviews the spectrum decision schemes in the literature. The system model and the problem formulations are presented in Sections III and IV. Next, we investigate the effects of multiple interruptions, sensing errors, channel capacity, and spectrum sharing on the overall system time in Sections V and VI. Numerical results are shown in Section VII. Finally, we give our concluding remarks in Section VIII.

II. Related Work

Most of the spectrum decision schemes proposed in the literature can be classified into two categories: the non-load-balancing and the load-balancing schemes. The load-balancing spectrum decision schemes can be further categorized into two methods: the sensing-based spectrum decision and the probability-based spectrum decision. Table I compares the existing load-balancing spectrum decision schemes, where "○" and "×" indicate that the proposed method "does" and "does not" consider the corresponding feature, respectively. In the following, we discuss the features of these spectrum decision methods in more details.

A. Non-load-balancing Spectrum Decision

For the non-load-balancing spectrum decision, the secondary user selects its operating channel based on certain conditions, such as traffic load [26], [27], channel idle probability [28], the expected waiting time [29], [30], the expected remaining idle period [31], [32], or the expected throughput [33], [34]. Most of these methods ignored the necessity of sharing spectrum with other secondary users. When all the secondary users select the same channel, the channel contention issue arises. To solve this problem, the load-balancing spectrum decision methods, including sensing-based channel selection and probability-based channel selection, are proposed to evenly distribute the traffic loads of the secondary users to multiple channels.

B. Probability-based Spectrum Decision

In the literature, many probability-based spectrum decision schemes were proposed to balance the traffic loads of secondary users in multi-channel CR networks, which can be categorized into three types: (1) packet-wise probabilistic (PP) approach [12]–[16]; (2) game-theoretic (GT) approach [17]–[19]; and (3) learning automata (LA) approach [20].

- Packet-wise probabilistic spectrum decision approaches [12]–[16] aim at maximizing the expected throughput of the secondary users at each slot by determining the probability of selecting each channel from the pool of candidate channel. Based on busy probability and capacity of each channel, [12] suggested a method to determine the probability for selecting channels on top of p-persistent carrier sense multiple access (CSMA) medium access control (MAC) protocol in a decentralized manner. They claimed that their proposed sub-optimal channel probability assignment can achieve the Nash equilibrium as the number of secondary users tends to infinity. Furthermore, [13], [14] considered the effects of sensing errors in terms of false alarm and missed detection on the throughput of the secondary users in a two-channel system, and proposed a probabilistic channel selection approach to maximize the throughput of the secondary users in each slot while maintaining the latency constraint of the primary users. Moreover, [15] formulated an optimization problem for channel selection probability to maximize the throughput of the secondary users in each slot while maintaining the interference constraint of the primary users when the primary and secondary networks are asynchronous. Unlike [13]–[15] considered only the case that one single secondary user can select the channel at each time instant, [16] further extended the probabilistic channel selection approach of [13], [14] to the case that multiple users can

TABLE I
COMPARISON OF VARIOUS LOAD-BALANCING SPECTRUM DECISION SCHEMES FOR COGNITIVE RADIO NETWORKS, WHERE PP, GT, AND LA STAND FOR THE PACKET-WISE PROBABILISTIC, GAME-THEORETIC, AND LEARNING AUTOMATA APPROACHES, RESPECTIVELY.

		Channel Occupancy Model of a Primary Network	Multiple Interruptions	Sensing Errors	Heterogeneous Channel Capacity
Probability-based Methods	PP	Bernoulli Process [12]	$\times$	$\times$	$\circ$
		Bernoulli Process [13]–[16]	$\times$	$\circ$	$\times$
	GT	Deterministic Process [17]	$\times$	$\times$	$\times$
		M/M/1 [18] or M/G/1 [19]	$\circ$	$\times$	$\circ$
	LA	General Distribution [20]	$\circ$	$\times$	$\times$
Sensing-based Methods		Deterministic Process [21]	$\times$	$\times$	$\times$
		Two-state Markov Chain [22]	$\times$	$\times$	$\times$
		Bernoulli Process [23]–[25]	$\times$	$\times$	$\times$
Our Proposed Model		M/G/1	$\circ$	$\circ$	$\circ$

simultaneously select their operating channels from the pool of candidate channels, and analyzed the throughput of the secondary users based on the probabilistic channel selection approach taking into account of the effects of channel contention as well as sensing errors. Note that the packet-wise probabilistic spectrum decision approaches in [12]–[16] were executed in a slot-by-slot manner, which may lead to many channel-switching behaviors during each secondary user's link connection. Moreover, it is assumed that the traffic loads of the secondary users are saturated. Further, the channel occupancy model of a primary network is modeled as a Bernoulli process and thus the length of busy and idle periods are exponentially distributed.

- Game-theoretic approaches were proposed to solve the spectrum decision problem in CR networks [17]–[19]. Based on the game theory model, each player (secondary user) can decide the best strategy (channel selection probability) to maximize its utility function. [17] proposed a game-theoretic load-balancing approach to find a set of channel selection probabilities so that no secondary user has incentive to unilaterally change his/her action. To converge to such the Nash equilibrium, a best-reply algorithm was designed for each user to calculate each channel's selection probability as well as its transmission duration based on a utility function related to the load-balancing channel selection. Beside the load-balancing issue, [18] suggested that the utility function in the game-theoretic spectrum decision should also incorporate the channel bandwidth and its idle period as well as the cost of spectrum handoff because the spectrum decision procedure must be executed many times due to multiple interruptions. They emphasized that the channel selection game shall be repeated many times to capture the scenario when primary users stochastically activate or deactivate at each epoch. Unlike the pervious work that considered the homogeneous secondary users, [19] assumed that the secondary users can have different priorities. They proposed a dynamic strategy learning algorithm to determine the channel selection strategies that can converge to the Nash equilibrium. Noteworthily, the Nash equilibrium solution of the game-theoretic approach is not necessary the globally optimal solution from the viewpoint of the overall network [35].
- In [20], a learning automata (LA) approach was suggested to determine the channel selection probabilities by exploring the uncertainty of traffic patterns in CR networks. After a huge number of trials, the secondary users can estimate the optimal channel selection probability. However, the problem for this method is its converging speed, especially for a large number of users.

C. Sensing-based Spectrum Decision

As mentioned in Section I, the sensing-based spectrum decision scheme requires scanning all the candidate channels to determine the most suitable operating channel. Thus, the total number of candidate channels significantly affects the overall system time in the sensing-based spectrum decision scheme. In [21]–[23], the optimal number of candidate channels to maximize the spectrum accessibility and the procedures to determine the optimal set of candidate channels were investigated. Furthermore, the authors in [24], [25] formulated the sequential channel sensing problem as an optimal stopping problem with the objective of maximizing the throughput of the secondary users. They studied when the secondary users shall stop sensing and start transmitting data. Nevertheless, the effects of multiple interruptions from the primary user and the sensing errors for the primary user's occurrence on the overall system time of the secondary users in the CR networks have not been addressed in these existing sensing-based spectrum decision methods.

D. Objectives of This Paper

The objective of this paper is to develop an analytical framework for the connection-based load-balancing spectrum decision schemes that can take into account of all the effects of (1) the primary users' multiple interruptions, (2) the secondary users' sensing errors, and (3) heterogeneous channel capacity. Furthermore, it is preferred that the developed analytical framework can be used in both the probability-based as well as the sensing-based spectrum decision schemes with general service time distributions of the primary and the secondary users. To the best of our knowledge, such an analytical framework has rarely been seen in the literature. We will detail the proposed modeling techniques for such general spectrum decision behaviors in the next section.

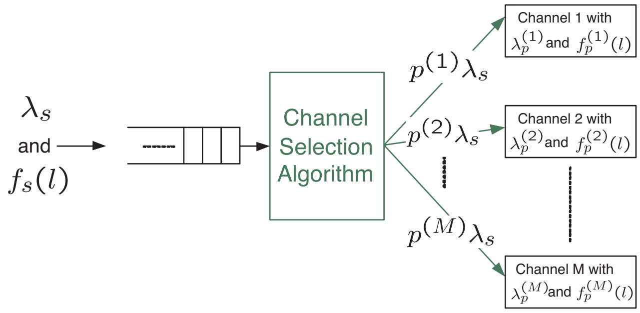

Fig. 1. Spectrum decision behavior model.

III. System Model

A. Assumptions

We consider a time-slotted CR network where the slot structure was also adopted in [33], [36]–[39]. In order to detect and protect the primary users, the spectrum sensing procedure must be executed by the secondary users at the beginning of each time slot. If the current operating channel is idle, the secondary user can transmit data in this time slot. By contrary if the current operating channel is busy, the secondary user must perform spectrum handoff procedures to resume its unfinished transmission when the current channel becomes idle. This kind of listen-before-talk channel access scheme has been adopted in many wireless techniques, such as the clear channel assessment (CCA) of the IEEE 802.11 standard [40] and the quiet period of the IEEE 802.22 standard [41].

B. Spectrum Decision Behavior Model

Fig. 1 illustrates the spectrum decision behavior model, which will be used to evaluate the overall system time of a secondary connection for different channel selection schemes. We assume that the arrival processes of the primary and the secondary connections[1] are Poisson. Let $\lambda_p^{(k)}$ (arrivals/slot) and λ_s (arrivals/slot) be the average arrival rates of the primary connections at channel k and the secondary connections of CR network, respectively. Also, denote $L_p^{(k)}$ (bits/arrival) and L_s (bits/arrival) the sizes of the primary connections of channel k and the secondary connections, respectively; and let $f_p^{(k)}(l)$ and $f_s(l)$ be probability mass functions (pmf) of $L_p^{(k)}$ and L_s, respectively. It is assumed that $\lambda_p^{(k)}$, λ_s, $f_p^{(k)}(l)$, and $f_s(l)$, which can be estimated by the existing methods [43], are known to all the secondary users. Furthermore, denote $R_p^{(k)}$ (bits/slot) and $R_s^{(k)}$ (bits/slot) as the data rate of the primary and the secondary connections at channel k, respectively. Hence, the service time of the primary and the secondary connections at channel k is $X_p^{(k)} \triangleq L_p^{(k)}/R_p^{(k)}$ (slots/arrival) and $X_s^{(k)} \triangleq L_s/R_s^{(k)}$ (slots/arrival), respectively.

As shown in Fig. 1, each secondary connection can select one of M candidate channels for its operating channel. Based on our proposed analytical framework, which will be discussed in more detail later, all the secondary users can dynamically select their operating channels with suitable probability that can balance the traffic loads of secondary users in multiple channels. The distribution probability vector (denoted by $\boldsymbol{p} = (p^{(1)}, p^{(2)}, \cdots, p^{(M)})$) represents the set of probabilities for selecting all the candidate channels, in which $p^{(k)}$ denotes the probability of a secondary connection selecting channel k for its operating channel. Thus, the effective arrival rate of the secondary connection at channel k is $\lambda_s^{(k)} = p^{(k)}\lambda_s$. Note that various channel selection algorithms yield different distribution probability vectors.

[1]When a secondary transmitter has data to send, how to establish a secondary connection to its intended receiver has been investigated in [42].

C. Sensing Errors

The service time of the primary and secondary connections will be extended due to missed detection and false alarm probabilities, respectively[2]. When missed detections occur, the primary user must retransmit these stained data frames in the next slots. Thus, the service time of a primary connection will be extended from $X_p^{(k)}$ (slots/arrival) to $\widetilde{X}_p^{(k)}$ (slots/arrival). Furthermore, a secondary user cannot transmit data even with an idle channel when a false alarm occurs. Hence, a secondary user needs to spend more time to complete its connection transmission. Then, the service time of a secondary connection will be extended to $\widetilde{X}_s^{(k)}$ (slots/arrival) from $X_s^{(k)}$ (slots/arrival). In the remaining part of this paper, $\widetilde{X}_p^{(k)}$ and $\widetilde{X}_s^{(k)}$ are called the actual service time of the primary and secondary connections. $\widetilde{X}_p^{(k)}$ and $\widetilde{X}_s^{(k)}$ will be derived in Section VI.

IV. Problem Formulation

A. Performance Metric: Overall System Time

The overall system time (denoted by S) is an important quality of service (QoS) metric for the connection-based service of the secondary users. It consists of the waiting time (denoted by W) and the extended data delivery time (denoted by T) as shown in Fig. 2. Hence, we have

$$\mathbf{E}[S] = \mathbf{E}[W] + \mathbf{E}[T] \ , \tag{1}$$

where $\mathbf{E}[\cdot]$ is the expectation function. Here, the waiting time is defined as the duration from the instant that a data transmission request arrives at the system until the instant of starting transmitting data. The duration of waiting time depends on the channel selection scheme that the secondary users adopt. Furthermore, the extended data delivery time is defined as the duration from the beginning of transmitting the data in the first time slot until the completion of the data in the last time slot. Clearly, multiple handoff behaviors significantly affect the extended data delivery time.

B. Overall System Time Minimization Problem for Probability-based Channel Selection Scheme

For the probability-based channel selection method, each secondary user selects its operating channel from all the M candidate channels based on a predetermined distribution probability vector $\boldsymbol{p}_{pb}$. In this case, an **Overall System**

[2]The relationship between the missed detection probability and the false alarm probability can be characterized by the receiver operating characteristic (ROC) curve [44].

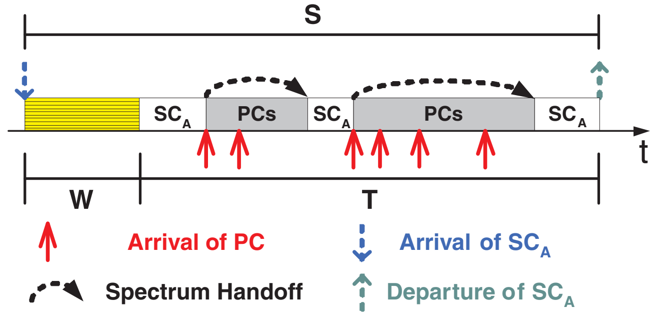

Fig. 2. Example of the overall system time of the secondary connection SC_A. The white areas indicate that channel is occupied by SC_A. Furthermore, the gray areas indicate that channel is occupied by the primary connections (PCs) and its duration is the busy period resulting from transmissions of the primary connections. Here, SC_A experiences two interruptions from the primary connections during its transmission period.

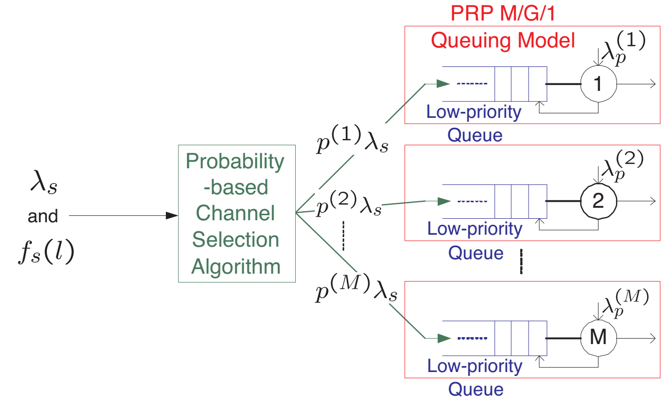

Fig. 3. Performance model for the probability-based channel selection scheme where the channel usage behaviors are characterized by the PRP M/G/1 queueing systems.

Time Minimization Problem for Probability-based Channel Selection Scheme can be formulated as follows. Given the set of candidate channels $\Omega = \{1, 2, \ldots, M\}$, we aim to find the optimal distribution probability vector (denoted by $\boldsymbol{p}^*$) to minimize the average overall system time of the secondary connections (denoted by $\mathbf{E}[S_{pb}]$). Formally,

$$\boldsymbol{p}^* = \underset{\forall\, \boldsymbol{p}_{pb}}{\arg\min}\, \mathbf{E}[S_{pb}(\boldsymbol{p}_{pb})] \ , \tag{2}$$

subject to:

$$0 \le p_{pb}^{(k)} \le 1, \quad \forall\ k \in \Omega \ , \tag{3}$$

$$\sum_{k\in\Omega} p_{pb}^{(k)} = \sum_{k=1}^{M} p_{pb}^{(k)} = 1 \ , \tag{4}$$

and

$$\rho^{(k)} = \rho_p^{(k)} + \rho_s^{(k)} < 1 \ , \tag{5}$$

where $\rho^{(k)}$ is the busy probability of channel k. Furthermore, $\rho_p^{(k)}$ and $\rho_s^{(k)}$ are the busy probabilities resulting from the primary and the secondary connections at channel k when sensing errors are considered, respectively. We can have $\rho_p^{(k)} = \lambda_p^{(k)}\mathbf{E}[\widetilde{X}_p^{(k)}]$ and $\rho_s^{(k)} = \lambda_s^{(k)}\mathbf{E}[\widetilde{X}_s^{(k)}]$.

C. Overall System Time Minimization Problem for Sensing-based Channel Selection Scheme

For the sensing-based channel selection scheme, the secondary users perform wideband sensing to find an idle channel from all the candidate channels. If more than one idle channel is found, the secondary user randomly selects one channel from the idle channels for its operating channel. Furthermore, if all the candidate channels are busy, the secondary user still randomly selects one channel from all the candidate channels and wait for the available time slot of this selected channel.

In order to decrease the total sensing time, the secondary users shall reduce the number of candidate channels by sensing only the best n channels among M channels. Without loss of generality, we assume that the channel preference of the secondary users follows the lexicographic order. That is, channel i is not better than channel j if $i > j$. Note that the ordering issue for channel preference has been discussed in [45]. Let Ω be the set of candidate channels. Then, we can have $\Omega = \{1, 2, \ldots, n\}$, where $n = |\Omega| \le M$. Next, we formulate an **Overall System Time Minimization Problem for Sensing-based Channel Selection Scheme** as follows. Given the total number of channels M, we aim to find the optimal number of candidate channels (denoted by n^*) to minimize the average overall system time of the secondary connections (denoted by $\mathbf{E}[S_{sb}]$). Formally,

$$n^* = \underset{1\le n\le M}{\arg\min}\, \mathbf{E}[S_{sb}(n)] \ . \tag{6}$$

D. Performance Model

In order to calculate the overall system time of various spectrum decision schemes, we extend the general model in Fig. 1 to characterize the probability- and the sensing-based channel selection schemes. Fig. 3 shows the performance model for the probability-based scheme. When the traffic of the secondary user (i.e., the secondary connection) arrives at the system, it can be directly connected to the selected channel based on the predetermined distribution probability vector. On the other hand, Fig. 4 shows the performance model of the sensing-based scheme. When the traffic of a secondary user arrives at the system, the secondary user performs spectrum sensing to find idle channels. The total sensing time can be modeled by a tapped delay line $\boxed{\mathrm{S}}$. In this case, $\boxed{\mathrm{S}}$ can be regarded as a server with constant service time, which equals to sensing time. If an idle channel can be found, the secondary connection can be served immediately.

As shown in Figs. 3 and 4, the proposed channel selection model is integrated with the preemptive resume priority (PRP) M/G/1 queueing systems in order to characterize the effects of multiple interruptions, sensing errors, and channel capacity on the overall system time. Some important properties for the PRP M/G/1 queueing model are listed below [46]:

- Each server (channel) has two types of customers (connections). The connections of the primary and secondary users are connected to the high-priority queue and the low-priority queue, respectively. Note that the high-priority queue has not been plotted in Figs. 3 and 4 due to space limitations.

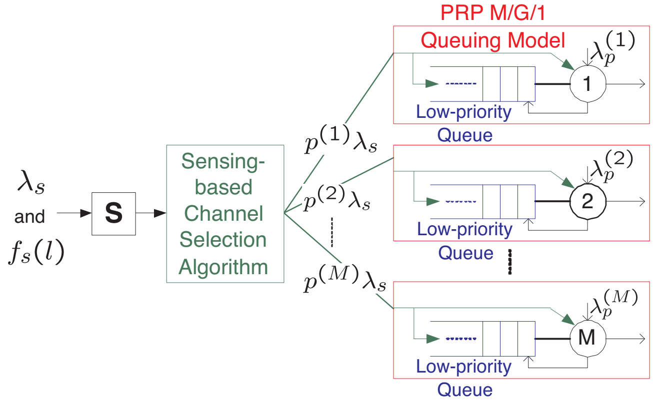

Fig. 4. Performance model for the sensing-based channel selection scheme where the channel usage behaviors are characterized by the PRP M/G/1 queueing systems.

- The primary users have the preemptive priority to interrupt the transmission of the secondary user. The remaining transmission of the interrupted secondary user will be put into the head of the low-priority queue of the current operating channel. Furthermore, the interrupted secondary user can resume the unfinished transmission when the current channel becomes idle, instead of re-transmitting the whole data.
- A secondary connection may experience multiple interruptions from the primary connections during its transmission period. This model can characterize the effects of multiple spectrum handoffs.

Here, we assume that connections which have the same priority access channels with the first-come-first-served (FCFS) scheduling discipline.

Based on the proposed performance models, we can analytically compare the overall system time resulting from both the spectrum decision schemes for various sensing time and traffic parameters. Then, each secondary user can intelligently adopt the best channel selection scheme to minimize its overall system time. Thus, the optimal overall system time (denoted by S^*) can be expressed as follows:

$$S^* = \min\left(\mathbf{E}[S_{pb}], \mathbf{E}[S_{sb}]\right) \ . \tag{7}$$

In the next section, we will show how to derive $\mathbf{E}[S_{pb}]$ and $\mathbf{E}[S_{sb}]$.

V. Analysis of Overall System Time

As discussed in Section IV-A, the overall system time consists of the waiting time and the extended data delivery time. Let $\mathbf{E}[T_{pb}]$ and $\mathbf{E}[T_{sb}]$ be the average data delivery time for the probability- and sensing-based spectrum decision methods, respectively. Furthermore, denote $\mathbf{E}[W_{pb}]$ and $\mathbf{E}[W_{sb}]$ as the average waiting time for the probability- and sensing-based spectrum decision methods, respectively. Then, we can have

$$\mathbf{E}[S_{pb}] = \mathbf{E}[W_{pb}] + \mathbf{E}[T_{pb}] \ , \tag{8}$$

and

$$\mathbf{E}[S_{sb}] = \mathbf{E}[W_{sb}] + \mathbf{E}[T_{sb}] \ . \tag{9}$$

In the following, we will investigate how to obtain the average extended data delivery time and the average waiting time.

A. Extended Data Delivery Time

First, we investigate the effects of multiple interruptions on the extended data delivery time. Within the transmission period of a secondary connection, it is likely to have multiple spectrum handoffs due to the interruptions from the primary users. The spectrum handoff procedure helps the secondary users vacate the occupied channel and then resume the unfinished transmission when this channel becomes idle. Clearly, multiple spectrum handoffs will increase the extended data delivery time and degrade the QoS for the latency-sensitive traffic of the secondary users.

Based on the PRP M/G/1 queueing model, we can derive the extended data delivery time of the secondary connections as follows. Let $N^{(k)}$ be the total number of interruptions for a secondary connection at channel k. Furthermore, denote $Y_p^{(k)}$ as the duration from the time instant that channel k is occupied by the primary connections until the time instant that the high-priority queue becomes empty. This duration is called the *busy period* resulting from transmissions of primary connections at channel k. When a secondary connection is interrupted by the primary user, it must stop transmitting on the current operating channel until all the primary connections in the high-priority queue have been served. In this case, this secondary connection of channel k must wait for the duration of $\mathbf{E}[Y_p^{(k)}]$ on average after the interruption event occurs. Denote $T^{(k)}$ as the extended data delivery time of the secondary connections at channel k. We can have

$$\mathbf{E}[T^{(k)}] = \mathbf{E}[\widetilde{X}_s^{(k)}] + \mathbf{E}[N^{(k)}]\mathbf{E}[Y_p^{(k)}] \ , \tag{10}$$

where $\mathbf{E}[N^{(k)}] = \lambda_p^{(k)}\mathbf{E}[\widetilde{X}_s^{(k)}]$ and $\mathbf{E}[Y_p^{(k)}] = \frac{\mathbf{E}[\widetilde{X}_p^{(k)}]}{1-\lambda_p^{(k)}\mathbf{E}[\widetilde{X}_p^{(k)}]}$ according to to [8].

Finally, the average extended data delivery time for the probability- and sensing-based channel selection methods can be expressed as follows:

$$\mathbf{E}[T_{pb}] = \sum_{k=1}^{M} p_{pb}^{(k)}\mathbf{E}[T^{(k)}] \ , \tag{11}$$

and

$$\mathbf{E}[T_{sb}] = \sum_{k=1}^{n} p_{sb}^{(k)}\mathbf{E}[T^{(k)}] \ . \tag{12}$$

For various channel selection algorithms, we use different methods to evaluate the corresponding distribution probability vectors $\boldsymbol{p}$. For the probability-based scheme, the distribution probability vector $\boldsymbol{p}_{pb}$ can be designed by solving the **Overall System Time Minimization Problem for Probability-based Channel Selection Scheme** in (2). For the sensing-based scheme, the distribution probability vector $\boldsymbol{p}_{sb}$ is determined inherently based on the given traffic patterns. Intuitively, a channel with larger idle probability will be selected more frequently through spectrum sensing. How to derive $\boldsymbol{p}_{sb}$ from the given traffic parameters will be discussed in Appendix A.

B. Waiting Time

1) Probability-based Channel Selection Scheme: In this case, a secondary connection selects its operating channel based on the predetermined probability. Then, it is directly connected to the low-priority queue of the selected channel. It cannot be served until all the primary and the secondary connections in the high-priority queue and the present low-priority queue of the selected channel have been served. Hence, the waiting time is the required duration from the time instant that a secondary connection arrives at the low-priority queue of the selected channel until the time instant that the selected channel becomes idle. That is, the waiting time is the duration spent in the waiting queue by a secondary connection. Hence, we have

$$\mathbf{E}[W_{pb}] = \sum_{k=1}^{M} p_{pb}^{(k)} \mathbf{E}[W_{pb}^{(k)}] \ , \tag{13}$$

where $W_{pb}^{(k)}$ is the waiting time of the secondary connections at channel k for the probability-based channel selection scheme. Applying the PRP M/G/1 queueing theory [47], one can obtain

$$\mathbf{E}[W_{pb}^{(k)}] = \frac{\mathbf{E}[R^{(k)}]}{(1-\rho_p^{(k)})(1-\rho_p^{(k)}-\rho_s^{(k)})} \ , \tag{14}$$

where $\mathbf{E}[R^{(k)}]$ is the average remaining time to complete the service of the connection being served at channel k. Referring to [47], we have

$$\mathbf{E}[R^{(k)}] = \frac{1}{2}\lambda_p^{(k)}\mathbf{E}[(\widetilde{X}_p^{(k)})^2] + \frac{1}{2}p_{pb}^{(k)}\lambda_s\mathbf{E}[(\widetilde{X}_s^{(k)})^2] \ . \tag{15}$$

Then, substituting (14) and (15) into (13), we can obtain the closed-from expression for $\mathbf{E}[W_{pb}]$.

Finally, substituting (11) and (13) into (8), we can obtain the relationship between the average overall system time and the distribution probability vector $\boldsymbol{p}_{pb}$ for the probability-based channel selection scheme. Then, the optimal distribution probability vector $\boldsymbol{p}^*$ can be determined by solving the **Overall System Time Minimization Problem for Probability-based Channel Selection** in (2).

2) Sensing-based Channel Selection Scheme: The waiting time W_{sb} for the sensing-based channel selection method consists of the total sensing time and the queueing time (denoted by W'_{sb}). Let τ be the sensing time for scanning one candidate channel. Hence, $n\tau$ is the total sensing time for scanning all the n candidate channels. After wideband sensing, the secondary user can decide channel availability and then transmits data at one of the idle channels. Moreover, if the idle channel cannot be found, the secondary user cannot transmit immediately. In this case, the secondary connection will be put into the low-priority queue of the randomly selected channel. Hence, we can have

$$\mathbf{E}[W_{sb}] = n\tau + \mathbf{Pr}\{\mathcal{E}\} \times 0 + \mathbf{Pr}\{\mathcal{E}^c\} \times \mathbf{E}[W'_{sb}] \ , \tag{16}$$

where $\mathcal{E}$ is the event that at least one idle channel can be found after sensing, and $\mathcal{E}^c$ is the compliment of $\mathcal{E}$.

Next, the closed-form expressions for $\mathbf{Pr}\{\mathcal{E}\}$ and $\mathbf{Pr}\{\mathcal{E}^c\}$ can be derived by the following two observations. First, a channel is called actually idle if and only if (1) this channel is not occupied by the primary connections and (2) the low-priority queue of this channel is empty. Note that the second condition should be contained because the FCFS scheduling discipline is adopted. Secondly, an idle channel is assessed as idle through spectrum sensing if and only if a false alarm does not occur. Hence, we can have

$$\begin{aligned}\mathbf{Pr}\{\mathcal{E}\} &= \sum_{k=1}^{n}[\mathbf{Pr}\{\mathcal{E}|\text{k channels are actually idle}\} \times \\ &\qquad \mathbf{Pr}\{\text{k channels are actually idle}\}] \\ &= \sum_{k=1}^{n}\Big[[1-(P_F)^k] \times \\ &\qquad \sum_{\Im\subseteq\Omega,|\Im|=k}\Big[\prod_{i\in\Im}(1-\rho^{(i)})\prod_{j\in\Omega-\Im}\rho^{(j)}\Big]\Big]\end{aligned} \tag{17}$$

where $\rho^{(k)} = \rho_p^{(k)} + \rho_s^{(k)}$ and P_F is the false alarm probability. On the other hand, $\mathcal{E}^c$ is the compliment of $\mathcal{E}$. That is,

$$\mathbf{Pr}\{\mathcal{E}^c\} = 1 - \mathbf{Pr}\{\mathcal{E}\} \ . \tag{18}$$

Moreover, when all channels are assessed as busy, each channel is selected by the secondary users with probability $1/n$. Hence, in this case, one can derive the average queueing time based on the PRP M/G/1 queueing theory as follows [47] :

$$\mathbf{E}[W'_{sb}] = \sum_{k=1}^{n}\left[\frac{1}{n}\cdot\frac{\mathbf{E}[R^{(k)}]}{(1-\rho_p^{(k)})(1-\rho_p^{(k)}-\rho_s^{(k)})}\right] \ . \tag{19}$$

Finally, substituting (12) and (16) into (9), we can obtain the relationship between the average overall system time and the number of candidate channels n for the sensing-based channel selection scheme.

Determining the optimal number of candidate channels (denoted by n^*) is the key issue for the sensing-based spectrum decision scheme. Intuitively, a small number of candidate channels can reduce the total sensing time $n\tau$ in (16). However, it is harder to find one idle channel from fewer candidate channels, resulting in a larger value of $\mathbf{Pr}\{\mathcal{E}^c\}$ in (16) and thus increasing the overall system time. The optimal number of candidate channels n^* can be determined by solving the **Overall System Time Minimization Problem for Sensing-based Channel Selection** in (6).

VI. Effects of Sensing Errors

Sensing errors such as false alarm and missed detection will degrade the performance of the secondary users and the primary users. This section investigates the effects of false alarm and missed detection on the actual service time of the secondary and the primary connections. Specifically, we will show how to derive the first and the second moments of $\widetilde{X}_s^{(k)}$ and $\widetilde{X}_p^{(k)}$.

A. False Alarm

First, we study the effect of false alarm on the actual service time of the secondary connections. When a false alarm occurs, the secondary user cannot transmit data even with an idle channel. Hence, the actual service time of a secondary connection will be extended to $\widetilde{X}_s^{(k)}$ (slots/arrival) from $X_s^{(k)}$

(slots/arrival). The first and the second moments of $\widetilde{X}_s^{(k)}$ can be expressed as follows:

$$\mathbf{E}[\widetilde{X}_s^{(k)}] = \sum_{x=1}^{\infty} \mathbf{E}[\widetilde{X}_s^{(k)}|X_s^{(k)} = x]\mathbf{Pr}\{X_s^{(k)} = x\} \ , \tag{20}$$

and

$$\mathbf{E}[(\widetilde{X}_s^{(k)})^2] = \sum_{x=1}^{\infty} \mathbf{E}[(\widetilde{X}_s^{(k)})^2|X_s^{(k)} = x]\mathbf{Pr}\{X_s^{(k)} = x\} \ . \tag{21}$$

Note that because the false-alarm slot cannot be exploited by any secondary or primary connections, it can be regarded as a busy slot. Hence, we can have $\rho_s^{(k)} = \lambda_s^{(k)}\mathbf{E}[\widetilde{X}_s^{(k)}]$.

When a false alarm occurs, the data transmission is postponed to the next slot. Hence, for a connection with x slots, its actual service time will be extended to $x+i$ slots if and only if false alarms occur in i slots out of the first $x+i-1$ slots and false alarm does not occur at the $(x+i)^{th}$ slot. Thus, the conditional expectation of the actual service time follows the negative binomial distribution with parameter P_F. That is,

$$\begin{aligned} &\mathbf{E}[\widetilde{X}_s^{(k)}|X_s^{(k)} = x] \\ = \ &\sum_{i=0}^{\infty}(x+i)\binom{x+i-1}{i}(1-P_F)^x(P_F)^i \ , \end{aligned} \tag{22}$$

and

$$\begin{aligned} &\mathbf{E}[(\widetilde{X}_s^{(k)})^2|X_s^{(k)} = x] \\ = \ &\sum_{i=0}^{\infty}(x+i)^2\binom{x+i-1}{i}(1-P_F)^x(P_F)^i \ , \end{aligned} \tag{23}$$

where P_F is the false alarm probability. Because $\mathbf{Pr}\{X_s^{(k)} = x\}$ is given by $f_s(l)$, we can obtain $\mathbf{E}[\widetilde{X}_s^{(k)}]$ and $\mathbf{E}[(\widetilde{X}_s^{(k)})^2]$ by substituting (22) and (23) into (20) and (21), respectively. For example, if $f_s(l)$ is the geometric distribution, i.e.,

$$f_s(l) = \left(1 - \frac{1}{\mathbf{E}[L_s]}\right)^{x-1}\left(\frac{1}{\mathbf{E}[L_s]}\right) \ , \tag{24}$$

we can have

$$\mathbf{E}[\widetilde{X}_s^{(k)}] = \frac{\mathbf{E}[X_s^{(k)}]}{1-P_F} \ , \tag{25}$$

and

$$\mathbf{E}[(\widetilde{X}_s^{(k)})^2] = \frac{\mathbf{E}[X_s^{(k)}](2\mathbf{E}[X_s^{(k)}] - 1 + P_F)}{(1-P_F)^2} \ , \tag{26}$$

where $\mathbf{E}[X_s^{(k)}] = \mathbf{E}[L_s]/R_s^{(k)}$.

B. Missed Detection

The data frame of the primary connection will be stained by the secondary connection when a missed detection occurs. Thus, the primary user will request to retransmit this stained data frame in the next slot. Hence, the actual service time of a primary connection will be extended to $\widetilde{X}_p^{(k)}$ (slots/arrival) from $X_p^{(k)}$ (slots/arrival). The first and the second moments of $\widetilde{X}_p^{(k)}$ can be expressed as follows:

$$\mathbf{E}[\widetilde{X}_p^{(k)}] = \sum_{x=1}^{\infty} \mathbf{E}[\widetilde{X}_p^{(k)}|X_p^{(k)} = x]\mathbf{Pr}\{X_p^{(k)} = x\} \ , \tag{27}$$

and

$$\mathbf{E}[(\widetilde{X}_p^{(k)})^2] = \sum_{x=1}^{\infty} \mathbf{E}[(\widetilde{X}_p^{(k)})^2|X_p^{(k)} = x]\mathbf{Pr}\{X_p^{(k)} = x\} \ . \tag{28}$$

Basically, there are two types of missed detections in CR networks [48], [49]. Firstly, when a primary user transmits data, a newly arriving secondary connection may incorrectly determine that this specific channel is available in its first sensing phase. We call this situation the class-A missed detection. After a secondary user arrives at a CR network for a while, it may also fail to detect the presence of primary users. In this case, the class-B missed detection occurs. It was found that the class-B missed detection is small because the sensing results at the first sensing phase can be employed to improve the accuracy of the sensing results at the following sensing phases.

Next, we explain the effect of class-A missed detection on the actual service time of the primary connection at channel k. We consider a transmission slot of this primary connection. During this slot, more than one arrival of the secondary connection appears with probability $1 - e^{-\lambda_s^{(k)}\Delta}$, where Δ is the slot duration. For these arrivals of secondary connections, each of them will assess this busy slot as idle if and only if (1) a missed detection occurs and (2) the low-priority queue of channel k is empty. Let $Q_s^{(k)}$ be the length of the low-priority queue at channel k. Hence, the first arrival at the considered slot will make an error channel assessment with probability $P_M\mathbf{Pr}\{Q_s^{(k)} = 0\}$, where P_M is the missed detection probability for spectrum sensing and $\mathbf{Pr}\{Q_s^{(k)} = 0\}$ has been derived in [50]. However, for the remaining arrivals in the considered slot, we have $\mathbf{Pr}\{Q_s^{(k)} = 0\} = 0$ because the first arrival has been put into the low-priority queue of channel k. Thus, the remaining arrivals do not make the error channel assessment. From above observations, we can conclude that a primary connection's transmission slot is stained by the arrivals of the secondary connections with probability

$$P_I^{(k)} = (1 - e^{-\lambda_s^{(k)}\Delta})P_M\mathbf{Pr}\{Q_s^{(k)} = 0\} \ . \tag{29}$$

Similar to the case of false alarm, we find that the random variables $\widetilde{X}_p^{(k)}$ and $(\widetilde{X}_p^{(k)})^2$ follows the negative binomial distribution with parameter $P_I^{(k)}$ when $X_p^{(k)} = x$. Then, because $\mathbf{Pr}\{X_p^{(k)} = x\}$ can be determined by $f_p^{(k)}(l)$, we can calculate the values of $\mathbf{E}[\widetilde{X}_p^{(k)}]$ and $\mathbf{E}[(\widetilde{X}_p^{(k)})^2]$ in (27) and (28), respectively. For example, if $f_p^{(k)}(l)$ is the geometric distribution, i.e.,

$$f_p^{(k)}(l) = \left(1 - \frac{1}{\mathbf{E}[L_p^{(k)}]}\right)^{x-1}\left(\frac{1}{\mathbf{E}[L_p^{(k)}]}\right) \ , \tag{30}$$

we can have

$$\mathbf{E}[\widetilde{X}_p^{(k)}] = \frac{\mathbf{E}[X_p^{(k)}]}{1-P_I^{(k)}} \ , \tag{31}$$

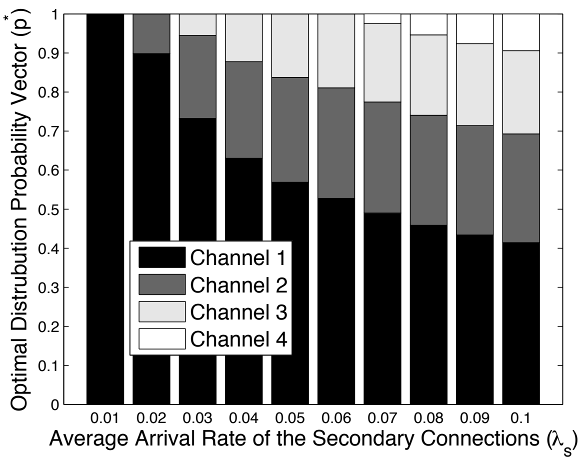

Fig. 5. Optimal distribution probability vector for the probability-based spectrum decision with various arrival rates of the secondary connections, where $P_F = 0.1$, $P_M = 0.1$, and $\mathbf{E}[X_s] = 10$.

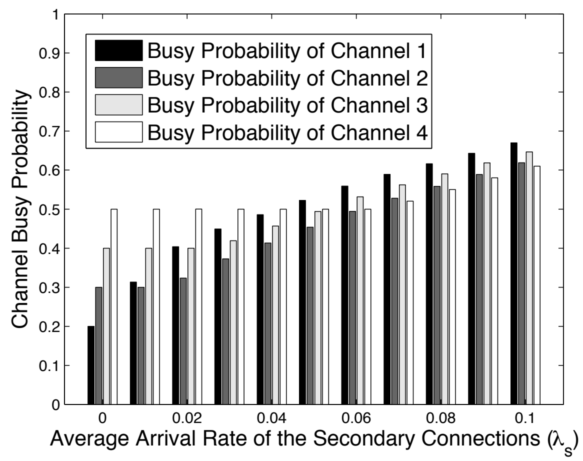

Fig. 6. Channel busy probability for the probability-based spectrum decision with various arrival rates of the secondary connections, where $P_F = 0.1$, $P_M = 0.1$, and $\mathbf{E}[X_s] = 10$.

and

$$\mathbf{E}[(\widetilde{X}_p^{(k)})^2] = \frac{\mathbf{E}[X_p^{(k)}](2\mathbf{E}[X_p^{(k)}] - 1 + P_I^{(k)})}{(1 - P_I^{(k)})^2} , \qquad (32)$$

where $\mathbf{E}[X_p^{(k)}] = \mathbf{E}[L_p^{(k)}]/R_p^{(k)}$.

VII. Numerical Results

In this section, numerical results are presented to show how to design the system parameters for the load-balancing spectrum decision methods, including the probability-based and the sensing-based spectrum decision schemes. We adopt the system parameters in the IEEE 802.22 standard in our simulation [51], where the time slot duration Δ is 10 msec, $P_M = 0.1$, and $P_F = 0.1$. Furthermore, we assume that $R_p^{(k)} = R_s^{(k)} = 1$ for any k. Hence, we have $\mathbf{E}[X_p^{(k)}] = \mathbf{E}[L_p^{(k)}]$ and $\mathbf{E}[X_s^{(k)}] = \mathbf{E}[L_s]$ for any k. To ease the notation, let $\mathbf{E}[X_s^{(k)}] \triangleq \mathbf{E}[X_s]$ for any k. Moreover, because this paper focuses on the latency-sensitive traffic, we can assume that the connection length (or equivalently service time) distributions of primary and secondary connections are geometrically distributed (see page 135 in [47]). Note that we only use the geometric distribution as an example here. Indeed, the proposed analytical framework can be applied to any distributions. It only requires the knowledge of the first and the second moments of the service time distributions for the primary and the secondary connections.

A. Probability-based Spectrum Decision Scheme

Figure 5 shows the effect of various arrival rates of the secondary connections on the optimal distribution probability vector, where the distribution probability vector is plotted in each bar and the summation of all probabilities in each bar is 1. In the figure, we consider a four-channel system with the following traffic parameters: $\lambda_p^{(1)} = 0.01$, $\lambda_p^{(2)} = 0.01$, $\lambda_p^{(3)} = 0.02$, and $\lambda_p^{(4)} = 0.02$ as well as $\mathbf{E}[X_p^{(1)}] = 20$, $\mathbf{E}[X_p^{(2)}] = 30$, $\mathbf{E}[X_p^{(3)}] = 20$, and $\mathbf{E}[X_p^{(4)}] = 25$. When $\lambda_s = 0.01$, all the secondary users prefer selecting channel 1 to be their operating channels because channel 1 has the lightest traffic loads. Furthermore, as λ_s increases, some secondary users tend to select other channels to transmit data in order to balance the traffic loads in each channel. For example, when $\lambda_s = 0.1$, the optimal distribution probability vector is $(0.4142, 0.2784, 0.2131, 0.0943)$. Inevitably, channel 1 is still selected to be the operating channel with the largest probability.

Furthermore, Fig. 6 shows the channel busy probability under various arrival rates of the secondary connections. When $\lambda_s = 0$, channel 1 has the lowest busy probability. However, when $\lambda_s \geq 0.05$, channel 1 has the highest busy probability because most secondary users prefer to select channel 1 to transmit data. Although channel 1 has the highest busy probability when $\lambda_s \geq 0.05$, one can find that the secondary users still favor channel 1 from Fig. 5.

Figure 7 shows that most secondary connections prefer selecting a channel with the largest arrival rate and the shortest service time of the primary connections even though all the channels have the same busy probability of the primary connections. Here, we consider the following traffic parameters: $\lambda_p^{(1)} = 0.01$, $\lambda_p^{(2)} = 0.02$, $\lambda_p^{(3)} = 0.04$, and $\lambda_p^{(4)} = 0.08$ as well as $\mathbf{E}[X_p^{(1)}] = 40$, $\mathbf{E}[X_p^{(2)}] = 20$, $\mathbf{E}[X_p^{(3)}] = 10$, and $\mathbf{E}[X_p^{(4)}] = 5$. Hence, all channels have the same busy probability, which is equal to 0.4, when $\lambda_s = 0$. According to (13) and (14), we know that selecting channel 4 can result in shorter average waiting time ($\mathbf{E}[W_{pb}]$) because channel 4 has the smallest value of $\mathbf{E}[R^{(k)}]$. Consequently, most of the secondary connections prefer selecting channel 4 and thus it has the highest busy probability when $\lambda_s > 0$.

Figure 8 shows the effects of false alarms on the optimal distribution probability vector. When $P_F = 0.05$, only three channels can be the candidate channels. However, all the four channels can be the candidate channels when $P_F \geq 0.1$. This phenomenon can be interpreted as follows. When P_F becomes higher, $\mathbf{E}[\widetilde{X}_s]$ increases due to more false alarms. Hence, the actual traffic loads ($\rho_s = \lambda_s \mathbf{E}[\widetilde{X}_s]$) of the secondary connections become heavy. Then, the secondary connections

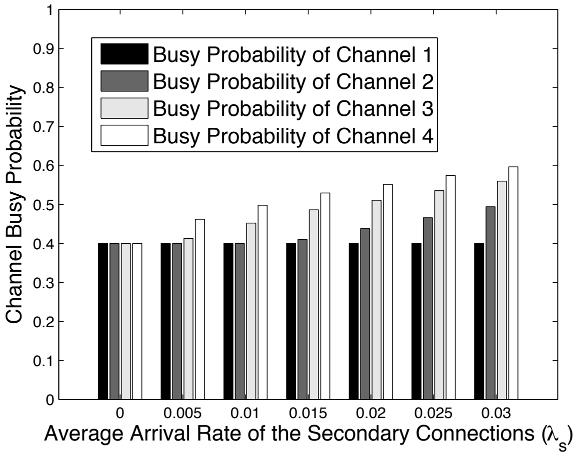

Fig. 7. Channel busy probability for the probability-based spectrum decision with various arrival rates of the secondary connections, where $P_F = 0$, $P_M = 0$, and $\mathbf{E}[X_s] = 15$.

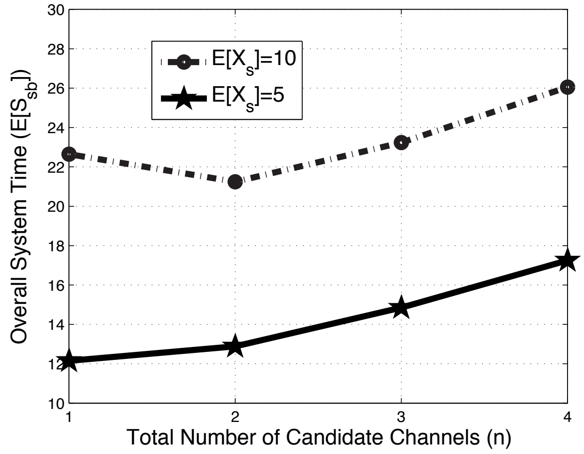

Fig. 9. Overall system time for the sensing-based spectrum decision with various numbers of candidate channels n, where $P_F = 0.1$, $P_M = 0.1$, $\tau = 2$, and $\mathbf{E}[X_p] = 20$.

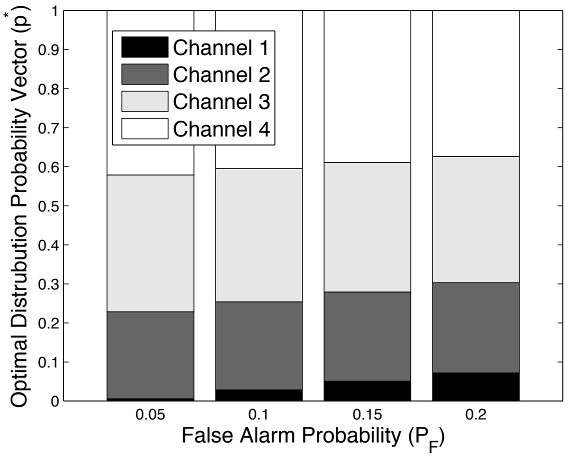

Fig. 8. Optimal distribution probability vector for the probability-based spectrum decision with various arrival rates of the secondary connections, where $P_M = 0.1$, $\lambda_s = 0.03$, and $\mathbf{E}[X_s] = 15$.

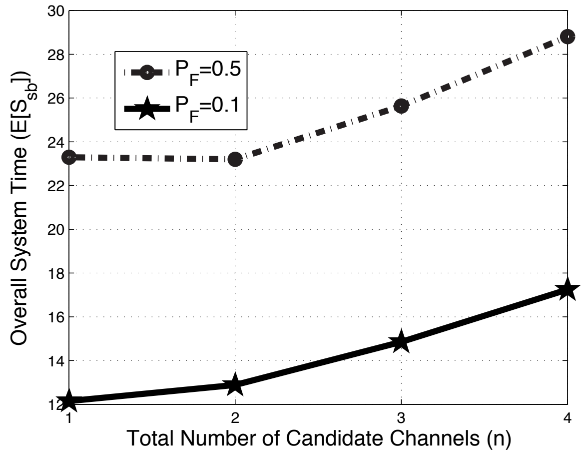

Fig. 10. Overall system time for the sensing-based spectrum decision with various numbers of candidate channels n, where $P_M = 0.1$, $\tau = 2$, $\mathbf{E}[X_p] = 20$, and $\mathbf{E}[X_s] = 5$.

must distribute overall traffic loads to more channels in order to prevent channel contention.

B. Sensing-based Spectrum Decision Scheme

Figures 9 and 10 show the effects of $\mathbf{E}[X_s]$ and P_F on the optimal number of candidate channels n^*, respectively. Here, we consider a four-channel system with the following traffic parameters: $(\lambda_p^{(1)}, \lambda_p^{(2)}, \lambda_p^{(3)}, \lambda_p^{(4)}) = (0.01, 0.015, 0.02, 0.025)$, $\lambda_s = 0.02$ and $\tau = 2$. Moreover, $\mathbf{E}[X_p^{(k)}] = 20$ for any k. From Fig. 9, one can see that when $P_F = 0.1$, $n^* = 1$ and 2 for $\mathbf{E}[X_s] = 5$ and 10, respectively. In Fig. 10, we see that $n^* = 1$ and 2 for $P_F = 0.1$ and 0.5 when $\mathbf{E}[X_s] = 5$. It is observed that the optimal value of n monotonically increases as $\mathbf{E}[X_s]$ or P_F increases. This is because a larger value of $\mathbf{E}[X_s]$ or P_F can lead to a larger value of $\mathbf{E}[\widetilde{X}_s]$ according to (25). From (19) one can expect that the queueing time will become longer for a larger value of $\mathbf{E}[\widetilde{X}_s]$. In this case, the secondary users shall sense more channels to increase the probability of finding idle channels $\mathbf{Pr}\{\mathcal{E}\}$, which will reduce the waiting time.

C. Comparison between Different Spectrum Decision Schemes

Figure 11 shows the effects of λ_s on the average overall system time for three different channel selection schemes: (1) sensing-based method; (2) probability-based method; and (3) non-load-balancing method. Consider a three-channel system with the following traffic parameters: $(\lambda_p^{(1)}, \lambda_p^{(2)}, \lambda_p^{(3)}) = (0.02, 0.02, 0.03)$, $(\mathbf{E}[X_p^{(1)}], \mathbf{E}[X_p^{(2)}], \mathbf{E}[X_p^{(3)}]) = (20, 25, 20)$, and $\mathbf{E}[X_s] = 10$. The overall system time of the probability-based and sensing-based channel selection schemes are calculated from (8) and (9), respectively. For the non-load-balancing method, all the secondary connections will select channel 1 to be their operating channels because channel 1 has the lowest busy probability. One can find that both the load-balancing channel selection schemes can significantly reduce the average overall system time compared to the non-load-balancing scheme, especially for larger λ_s. When τ is small

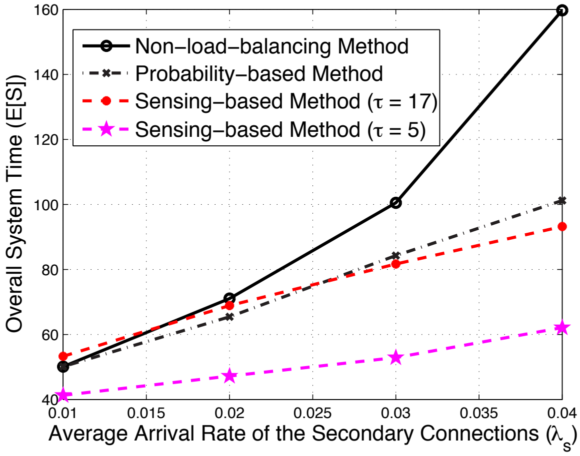

Fig. 11. Comparison of the overall system time for three considered spectrum decision schemes, where $P_F = 0.1$, $P_M = 0.1$, and $\mathbf{E}[X_s] = 10$.

(e.g. 5 slots), the sensing-based spectrum decision scheme can result in the shortest overall system time. As τ increases, the improvement of the sensing-based spectrum decision over other schemes decreases. In addition, we also observe that when $\tau = 17$ and $\lambda_s < 0.026$, the probability-based scheme has better overall system time performance than the sensing-based scheme. This is because the probability-based spectrum decision scheme can select the channels with lower interrupted probability. By contrast, if $\lambda_s > 0.026$, the sensing-based scheme can result in shorter overall system time because the sensing-based scheme can significantly reduce waiting time through wideband sensing. Based on (7), each secondary user can intelligently adopt the best channel selection scheme to minimize its overall system time. The two considered load-balancing spectrum decision methods can reduce the overall system time by over 50% compared to the existing non-load-balancing method when $\lambda_s = 0.04$.

VIII. Conclusions

In this paper, an analytical framework has been proposed to design the system parameters for the sensing- and the probability-based spectrum decision schemes. The proposed model integrated with the PRP M/G/1 queueing systems can evaluate the effects of multiple interruptions, sensing errors, and channel capacity on the overall system time of the secondary connections. Based on this analytical model, the optimal number of candidate channels for the sensing-based spectrum selection method and the optimal channel selection probability for the probability-based spectrum selection method can be obtained analytically for various sensing time and traffic parameters. We found that the probability-based scheme can reduce the overall system time compared to the sensing-based scheme when the traffic loads of the secondary users is light, whereas the sensing-based scheme performs better in the condition of heavy traffic loads. This observation provide an important insight into design a traffic-adaptive spectrum decision scheme in the presence of sensing errors.

Some interesting research issues that can be extended from this paper include the following. First, it is worthwhile to determine the optimal distribution probability vector for the probability-based spectrum decision method when the secondary connections may have different opinions on the observed traffic statistics $\lambda_p^{(k)}$, λ_s, $f_p^{(k)}(l)$, and $f_s(l)$. Secondly, it would be interesting to see how to analyze the overall system time for both considered load-balancing spectrum decision methods in the hopping mode, i.e., the operating channel after the primary user's interruption is different from the current channel.

Appendix A
Distribution Probability Vector for the Sensing-based Channel Selection Scheme

The probability that a secondary user can select channel k for its operating channel is determined inherently based on the traffic patterns for the sensing-based spectrum decision scheme. According to the sensing outcomes, this probability consists of three components. First, we consider the case that false alarm dose not occur at the idle channel k. When the channels in $\Im \subseteq \Omega - \{k\}$ are also actually idle and false alarms do not occur at the channels in $\Re \subseteq \Im$, channel k will be selected with probability $\frac{1}{1+|\Re|}$. Secondly, we consider the case when a false alarm occurs at the idle channel k. If false alarms also occur at all the remaining idle channels, the secondary user will randomly select one channel from all candidate channels to be its operating channel. In this case, channel k is selected with probability $1/|\Omega|$. Thirdly, we consider the case when channel k is actually busy. With the similar argument in the previous case, the secondary user will randomly select one channel if false alarms occur at all the idle channels. In this case, channel k will be selected with probability $1/|\Omega|$. On the other hand, channel k cannot be selected when $k \notin \Omega$. From these observations, we can have (33).

References

[1] J. Mitola and G. Q. Maguire, "Cognitive Radio: Making Software Radios More Personal," *IEEE Personal Commun.*, vol. 6, pp. 13–18, Aug. 1999.

[2] S. Haykin, "Cognitive Radio: Brain-empowered Wireless Communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.

[3] R. W. Thomas, L. A. DaSilva, and A. B. MacKenzie, "Cognitive Networks," *IEEE International Symposium on Dynamic Spectrum Access Networks (DYSPAN)*, Nov. 2005.

[4] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "NeXt Generation/Dynamic Spectrum Access/Cognitive Radio Wireless Networks: A Survey," *Computer Networks Journal (Elsevier)*, vol. 50, pp. 2127–2159, Sep. 2006.

[5] C.-W. Wang, L.-C. Wang, and F. Adachi, "Performance Gains for Spectrum Utilization in Cognitive Radio Networks with Spectrum Handoff," *International Symposium on Wireless Personal Multimedia Communications (WPMC)*, Sep. 2009.

[6] L.-C. Wang, C.-W. Wang, Y.-C. Lu, and C.-M. Liu, "A Concurrent Transmission MAC Protocol for Enhancing Throughput and Avoiding Spectrum Sensing in Cognitive Radio," *IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2007.

[7] L.-C. Wang and C.-W. Wang, "Spectrum Handoff for Cognitive Radio Networks: Reactive-sensing or Proactive-sensing?" *IEEE International Performance Computing and Communications Conference (IPCCC)*, Dec. 2008.

[8] C.-W. Wang and L.-C. Wang, "Modeling and Analysis for Proactive-decision Spectrum Handoff in Cognitive Radio Networks," *IEEE International Conference on Communications (ICC)*, Jun. 2009.

[9] C.-W. Wang, L.-C. Wang, and F. Adachi, "Modeling and Analysis for Rective-decision Spectrum Handoff in Cognitive Radio Networks," *IEEE GLOBECOM*, Dec. 2010.

$$p_{sb}^{(k)} = \begin{cases} (1-\rho^{(k)})(1-P_F)\times \\ \displaystyle\sum_{\Im\subseteq\Omega-\{k\}}\left[\prod_{i\in\Im}(1-\rho^{(i)})\prod_{j\in\Omega-\{k\}-\Im}\rho^{(j)}\sum_{\Re\subseteq\Im}\frac{1}{1+|\Re|}(1-P_F)^{|\Re|}(P_F)^{|\Im|-|\Re|}\right]+ \\ \left[(1-\rho^{(k)})P_F+\rho^{(k)}\right]\times \\ \displaystyle\sum_{\Im\subseteq\Omega-\{k\}}\left[\prod_{i\in\Im}(1-\rho^{(i)})\prod_{j\in\Omega-\{k\}-\Im}\rho^{(j)}(P_F)^{|\Im|}\right]\times\frac{1}{|\Omega|}\ , & k\in\Omega \\ 0\ , & k\notin\Omega \end{cases} \tag{33}$$

[10] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "A Survey on Spectrum Management in Cognitive Radio Networks," *IEEE Commun. Mag.*, vol. 46, no. 4, pp. 40–48, Apr. 2008.

[11] H.-J. Liu, Z.-X. Wang, S.-F. Li, and M. Yi, "Study on the Performance of Spectrum Mobility in Cognitive Wireless Network," *IEEE Singapore International Conference on Communication Systems (ICCS)*, Jun. 2008.

[12] A. Banaei and C. N. Georghiades, "Throughput Analysis of a Randomized Sensing Scheme in Cell-based Ad-hoc Cognitive Networks," *IEEE International Conference on Communications (ICC)*, Jun. 2009.

[13] J. Gambini, O. Simeone, Y. Bar-Ness, U. Spagnolini, and T. Yu, "Packet-wise Vertical Handover for Unlicensed Multi-standard Spectrum Access with Cognitive Radios," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 5172–5176, Dec. 2008.

[14] J. Gambini, O. Simeone, U. Spagnolini, Y. Bar-Ness, and Y. Kim, "Cognitive Radio with Secondary Packet-By-Packet Vertical Handover," *IEEE International Conference on Communications (ICC)*, May 2008.

[15] G. Noh, J. Lee, and D. Hong, "Stochastic Multichannel Sensing for Cognitive Radio Systems: Optimal Channel Selection for Sensing with Interference Constraints," *IEEE Vehicular Technology Conference Fall*, Sep. 2009.

[16] C.-M. Lee, J.-S. Lin, Y.-P. Hsu, and K.-T. Feng, "Design and Analysis of Optimal Channel-hopping Sequence for Cognitive Radio Networks," *IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2010.

[17] A. T. Chronopoulos, M. R. Musku, S. Penmatsa, and D. C. Popescu, "Spectrum Load Balancing for Medium Access in Cognitive Radio Systems," *IEEE Commun. Lett.*, vol. 12, no. 5, pp. 353–355, May 2008.

[18] I. Malanchini, M. Cesana, and N. Gatti, "On Spectrum Selection Games in Cognitive Radio Networks," *IEEE Global Communications Conference (GLOBECOM)*, Nov. 2009.

[19] H.-P. Shiang and M. van der Schaar, "Queuing-based Dynamic Channel Selection for Heterogeneous Multimedia Applications Over Cognitive Radio Networks," *IEEE Trans. Multimedia*, vol. 10, no. 5, pp. 896–909, Aug. 2008.

[20] Y. Song, Y. Fang, and Y. Zhang, "Stochastic Channel Selection in Cognitive Radio Networks," *IEEE Global Communications Conference (GLOBECOM)*, Nov. 2007.

[21] H.-J. Liu, S.-F. Li, Z.-X. Wang, W.-J. Hong, and M. Yi, "Strategy of Dynamic Spectrum Access Based-on Spectrum Pool," *IEEE International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM)*, Sep. 2008.

[22] A. W. Min and K. G. Shin, "Exploiting Multi-channel Diversity in Spectrum-agile Networks," *IEEE International Conference on Computer Communications (INFOCOM)*, Apr. 2008.

[23] B. Hamdaoui, "Adaptive Spectrum Assessment for Opportunistic Access in Cognitive Radio Networks," *IEEE Trans. Wireless Commun.*, vol. 8, no. 2, pp. 922–930, Feb. 2009.

[24] A. Sabharwal, A. Khoshnevis, and E. Knightly, "Opportunistic Spectral Usage: Bounds and a Multi-band CSMA/CA Protocol," *IEEE/ACM Trans. Netw.*, vol. 15, no. 3, pp. 533–545, Jun. 2007.

[25] J. Jia, Q. Zhang, and X. S. Shen, "HC-MAC: A Hardware-constrained Cognitive MAC for Efficient Spectrum Management," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 1, pp. 106–117, Jan. 2008.

[26] S.-C. Lo and C.-W. Tseng, "A Novel Multi-channel MAC Protocol for Wireless Ad Hoc Networks," *IEEE Vehicular Technology Conference Fall*, Sep. 2007.

[27] P. Zhu, J. Li, and X. Wang, "A New Channel Parameter for Cognitive Radio," *International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom)*, Aug. 2007.

[28] T.-S. Kim, H. Lim, and J. C. Hou, "Understanding and Improving the Spatial Reuse in Multihop Wireless Networks," *IEEE Trans. Mobile Comput.*, vol. 7, no. 10, pp. 1200–1212, Oct. 2008.

[29] A. C.-C. Hsu, D. S.-L. Wei, and C.-C. J. Kuo, "A Cognitive MAC Protocol Using Statistical Channel Allocation for Wireless Ad-hoc Networks," *IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2007.

[30] R.-T. Ma, Y.-P. Hsu, and K.-T. Feng, "A POMDP-based Spectrum Handoff Protocol for Partially Observable Cognitive Radio Networks," *IEEE Wireless Communications and Networking Conference (WCNC)*, Apr. 2009.

[31] S.-U. Yoon and E. Ekici, "Voluntary Spectrum Handoff: A Novel Approach to Spectrum Management in CRNs," *IEEE International Conference on Communications (ICC)*, May 2010.

[32] Y. Song and J. Xie, "Common Hopping Based Proactive Spectrum Handoff in Cognitive Radio Ad Hoc Networks," *IEEE Global Communications Conference (GLOBECOM)*, Dec. 2010.

[33] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized Cognitive MAC for Opportunistic Spectrum Access in Ad Hoc Networks: A POMDP Framework," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 3, pp. 589–600, Apr. 2007.

[34] Q. Zhao, S. Geirhofer, L. Tong, and B. M. Sadler, "Opportunistic Spectrum Access via Periodic Channel Sensing," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 785–796, Feb. 2008.

[35] M. Felegyhazi and J.-P. Hubaux, "Game Theory in Wireless Networks: A Tutorial," *EPFL Technical Report: LCA-REPORT-2006-002*, 2006.

[36] H. Su and X. Zhang, "Channel-hopping Based Single Transceiver MAC for Cognitive Radio Networks," *IEEE Annual Conference on Information Sciences and Systems (CISS)*, Mar. 2008.

[37] Y.-C. Liang, Y. Zeng, E. C. Peh, and A. T. Hoang, "Sensing-throughput Tradeoff for Cognitive Radio Networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 4, Apr. 2008.

[38] P. Wang, L. Xiao, S. Zhou, and J. Wang, "Optimization of Detection Time for Channel Efficiency in Cognitive Radio Systems," *IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2007.

[39] W.-Y. Lee and I. F. Akyildiz, "Optimal Spectrum Sensing Framework for Cognitive Radio Networks," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3845–3857, Oct. 2008.

[40] IEEE 802.11, *Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, supplement to IEEE 802.11 Standard, Sep. 1999.

[41] C. R. Stevenson, G. Chouinard, Z. Lei, W. Hu, S. J. Shellhammer, and W. Caldwell, "IEEE 802.22: The First Cognitive Radio Wireless Regional Area Network Standard," *IEEE Commun. Mag.*, vol. 47, no. 1, pp. 130–138, Jan. 2009.

[42] L.-C. Wang, Y.-C. Lu, C.-W. Wang, and D. S.-L. Wei, "Latency Analysis for Dynamic Spectrum Access in Cognitive Radio: Dedicated or Embedded Control Channel?" *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sep. 2007.

[43] X. Li and S. A. Zekavat, "Traffic Pattern Prediction and Performance Investigation for Cognitive Radio Systems," *IEEE Wireless Communications and Networking Conference (WCNC)*, Mar. 2008.

[44] Q. Zhao and B. M. Sadler, "A Survey of Dynamic Spectrum Access: Signal Processing, Networking, and Regulatory Policy," *IEEE Signal Processing Mag.*, pp. 79–89, May 2007.

[45] H. Jiang, L. Lai, R. Fan, and H. V. Poor, "Optimal Selection of Channel Sensing Order in Cognitive Radio," *IEEE Trans. Wireless Commun.*, vol. 8, no. 1, pp. 297–307, Jan. 2009.

[46] L.-C. Wang and C.-W. Wang, "Spectrum Management Techniques with QoS Provisioning in Cognitive Radio Networks," *International Symposium on Wireless and Pervasive Computing (ISWPC)*, May 2010.

[47] Chee-Hock Ng and Boon-Hee Soong, *Queueing Modelling Fundamentals with Applications in Communication Networks, 2nd*. John Wiley & Sons Inc., 2008.
[48] S. Tang and B. L. Mark, "Modeling and Analysis of Opportunistic Spectrum Sharing with Unreliable Spectrum Sensing," *IEEE Trans. Wireless Commun.*, vol. 8, pp. 1934–1943, Apr. 2009.
[49] I. Suliman, J. Lehtomaki, T. Braysy, and K. Umebayashi, "Analysis of Cognitive Radio Networks with Imperfect Sensing," *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Sep. 2009.
[50] N. K. Jaiswal, "Preemptive Resume Priority Queue," *Operations Research*, vol. 9, no. 5, pp. 732–742, Sep./Oct. 1961.
[51] *Draft Standard for Wireless Regional Area Networks Part 22: Cognitive Wireless RAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications*, IEEE 802.22 Working Group.

Li-Chun Wang (S'92-M'96-SM'06-F'11) received the B.S. degree in electrical engineering from the National Chiao-Tung University, Hsinchu, Taiwan, in 1986, the M.S. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, in 1988, and the M.Sc. and Ph.D. degrees in electrical engineering from Georgia Institute of Technology, Atlanta, in 1995 and 1996, respectively.

In 1995, he was affiliated with Northern Telecom in Richardson, Texas. From 1996 to 2000, he was with AT&T Laboratories, where he was a Senior Technical Staff Member in the Wireless Communications Research Department. Since August 2000, he has joined the Department of Communication Engineering of National Chiao-Tung University in Taiwan as an Associate Professor and has been promoted to a full professor since August 2005. Dr. Wang was a corecipient of the Jack Neubauer Best Paper Award from the IEEE Vehicular Technology Society in 1997. His current research interests are in the areas of cellular architectures, radio network resource management, cross-layer optimization for cooperative and cognitive wireless networks. He is the holder of three U.S. patents with three more pending.

Chung-Wei Wang (S'07) received the B.S. degree in electrical engineering from Tamkang University, Taipei, Taiwan, in 2003, and the Minor M.S. and Ph.D. degrees in applied mathematics and communication engineering from the National Chiao-Tung University, Hsinchu, Taiwan, in 2007 and 2010, respectively. From 2009 to 2010, he was also a visiting scholar in Tohoku University, Sendai, Japan. He was awarded the student travel grant from IEEE ICC 2009. His current research interests include cross-layer optimization, MAC protocols design, and radio resource management in wireless sensor networks, ad hoc networks, and cognitive radio networks.

Fumiyuki Adachi (M'79-SM'90-F'00) received the B.S. and Dr. Eng degrees in electrical engineering from Tohoku University, Sendai, Japan, in 1973 and 1984, respectively. In April 1973, he joined the Electrical Communications Laboratories of Nippon Telegraph & Telephone Corporation (now NTT) and conducted various types of research related to digital cellular mobile communications. From July 1992 to December 1999, he was with NTT Mobile Communications Network, Inc. (now NTT DoCoMo, Inc.), where he lead a research group on wideband/broadband CDMA wireless access for IMT-2000 and beyond. Since January 2000, he has been with Tohoku University, Sendai, Japan, where he is a Professor of Electrical and Communication Engineering at the Graduate School of Engineering. His research interests are in CDMA wireless access techniques, equalization, transmit/receive antenna diversity, MIMO, adaptive transmission, and channel coding, with particular application to broadband wireless communications systems.