Wireless Netw (2014) 20:1003–1022
DOI 10.1007/s11276-013-0658-x

Stochastic spectrum handoff protocols for partially observable cognitive radio networks

Jui-Hung Chu · Rui-Ting Ma · Kai-Ten Feng

Abstract Recent studies have been conducted to indicate the ineffective usage of licensed bands due to static spectrum allocation. In order to improve spectrum utilization, cognitive radio (CR) is therefore suggested to dynamically exploit the opportunistic primary frequency spectrums. How to provide efficient spectrum handoff has been considered a crucial issue in the CR networks. Existing spectrum handoff algorithms assume that all the channels can be correctly sensed by the CR users in order to perform appropriate spectrum handoff process. However, this assumption is impractical since excessive time will be required for the CR user to sense the entire spectrum space. In this paper, the partially observable Markov decision process (POMDP) is applied to estimate the network information by partially sensing the frequency spectrums. A POMDP-based spectrum handoff (POSH) scheme is proposed to determine the optimal target channel for spectrum handoff according to the partially observable channel state information. Moreover, a POMDP-based multi-user spectrum handoff (M-POSH) protocol is proposed to exploit the POMDP policy into multi-user CR networks by distributing CR users to frequency spectrum bins opportunistically. By adopting the policies resulted from the POSH and M-POSH algorithms for target channel selection, minimal waiting time at each occurrence of spectrum handoff can be achieved which will be feasible for multimedia applications. Numerical results illustrate that the proposed spectrum handoff protocols can effectively minimize the required waiting time for spectrum handoff in the CR networks.

Keywords Cognitive radio · Partially observable Markov decision process (POMDP) · Spectrum handoff · Channel selection

1 Introduction

According to the research conducted by FCC [1, 2], a large portion of the auctioned frequency spectrum remains idle at any given time and location. It has been indicated that the spectrum shortage problem is primarily resulted from the spectrum management policy rather than the physical scarcity of frequency spectrum. Consequently, a great amount of research is devoted to cognitive radio (CR) in recent years [3–6]. The CR user, i.e. the secondary user, is capable of sensing the channel condition and can adapt its internal parameters to access the licensed channels while these channels are not being utilized by the primary users. The IEEE 802.22 [7, 8] is a standard that allocates the TV broadcast spectrum on a license-exempt basis, which is considered a realization of the CR concept. The white space of TV spectrum can be utilized by the cognitive radio users to efficiently improve the utility rate of the TV bands. The CR users are allowed to share the unused spectrum while assuring that no interference is caused to the primary users, which is feasible for multimedia applications such as digital or analog TV broadcasting. Since there is no promise for a CR user to finish its transmission on a certain spectrum, a mechanism called spectrum handoff has been introduced to allow the CR user to select another channel to

J.-H. Chu · R.-T. Ma · K.-T. Feng (✉)
Department of Electrical and Computer Engineering, National Chiao Tung University, Hsinchu, Taiwan
e-mail: ktfeng@mail.nctu.edu.tw

J.-H. Chu
e-mail: jhchu.cm98g@g2.nctu.edu.tw

R.-T. Ma
e-mail: scotie.cm96g@nctu.edu.tw

maintain its data transmission. Consequently, the main objective for spectrum handoff is to select a feasible target channel which the CR user can be switched into in order to retain its on-going transmissions. The performance of spectrum handoff will primarily be dominated by the feasibility of conducting channel selection.

Channel selection schemes can be classified into pre-sensing and post-sensing strategies. The pre-sensing scheme benefits from pre-determination of target channel selection without inducing excessive waiting time for spectrum handoff. However, inefficient channel selection can be occurred due to fast-changing environments for channel variations. On the other hand, the post-sensing methods allow the CR users to determine the target channels based on both instantaneous observations and previous updates of channel states. More precise channel conditions can be acquired such that better spectrum handoff can be performed by CR users. However, post-sensing strategies require a great amount of time for spectrum sensing which is considered inefficient for CR users to conduct spectrum handoff. In this paper, without the necessity of obtaining all the correct channel information, the partially observable Markov decision process (POMDP) [9] will be utilized to reveal network information by partially sensing available frequency channels.

A POMDP-based spectrum handoff (POSH) mechanism is proposed as a post-sensing strategy in order to acquire the policy such that optimal channel can be obtained with minimal waiting time at each occurrence of spectrum handoff. The transition probabilities for the channel states are derived by considering the channel as an M/G/1 system with given packet arrival rate and service rate. In order to observe the behavior of spectrum handoff, analytical models for the proposed POSH protocol and other existing channel selection schemes are derived, validated, and compared via simulations in this paper. Furthermore, practical considerations on the required phases for channel selection are also discussed and compared between the spectrum handoff schemes. It is demonstrated via simulations that reduced waiting time can be obtained from the proposed POSH protocol comparing with other existing schemes.

Furthermore, multiple CR users can exist concurrently and access the licensed spectrums in a more realistic environment. Unlike the single user scenario, a specific frequency spectrum can be simultaneously chosen by more than one CR users as the target channel in multi-user scenarios. Therefore, it is crucial to provide a channel selection policy by considering information from other CR users, and also to alleviate the overhead while multiple CR users are contending for the target channel. In this paper, a POMDP-based multi-user spectrum handoff (M-POSH) protocol is proposed to extend the original POSH scheme into the multi-user CR networks. In addition to merely observing the primary user's traffic, the M-POSH scheme also monitors the required information from CR users in the network. A negotiation procedure is employed to facilitate the determination process for a feasible CR user to operate within a specific target channel. In comparison with other existing protocols, the proposed M-POSH scheme can effectively minimize the waiting time for spectrum handoff. Note that a short version of proposed POSH protocol has be presented in our previous work in [10]. Only a single CR user is considered in the network scenario to perform spectrum handoff. Moreover, theoretical models for proposed POSH scheme have not been developed in the short paper for performance validation and comparison.

The rest of this paper is organized as follows. The related works are discussed in Sect. 1. The proposed POSH scheme is modeled and derived in Sect. 3. Performance analysis of the POSH protocol associated with practical considerations is conducted in Sect. 4. Considering the network scenario with multiple CR users, the M-POSH protocol is proposed in Sect. 5. Section 6 illustrates the performance evaluation for both the proposed POSH and M-POSH approaches; while the conclusions are drawn in Sect. 7.

2 Related work

Target channel selection can be categorized into two different types of schemes according to their sensing strategies [11], i.e. the pre-sensing and post-sensing methods. The pre-sensing scheme indicates that the secondary user will sense the frequency spectrums and consequently choose a sequence of selected target channels before the beginning of its data transmission. Once the secondary user is interrupted by the primary user, the secondary user will be switched into a channel which was determined in sequence in the pre-sensing phase. In general, pre-sensing techniques can reduce the waiting time for spectrum handoff since the target channel is selected based on a pre-determined channel list. However, since the stochastic characteristics of channel can vary drastically in realistic environments, the pre-determined channel list can become inefficient to be adopted in target channel selection for spectrum handoff. The channel reservation scheme as proposed in [12] conducts pre-sensing by exploiting the balance between blocking probability and forced termination in order to reserve idle channels for spectrum handoff. Analytical models have been studied in [13, 14] to illustrate the beneficial aspects of pre-sensing strategy. However, those reserved idle channels cannot be ensured available at the time for spectrum handoff. As a result, the performance of pre-sensing strategies cannot be guaranteed especially under fast-fading channel environments.

On the other hand, the post-sensing technique is implemented while the secondary user is forced to terminate its transmission by the primary user. The CR user will start to sense the spectrum in order to verify if there are available channels that can be accessed and consequently become its target channels. Available network information will be updated at each time slot and the CR user can determine its target channel in the next slot based on current observation and updated information state. Compared to pre-sensing schemes, the post-sensing methods require additional time overhead by conducting spectrum sensing and implementing sensing policy. However, a more feasible and accurate channel can be selected by exploiting the post-sensing schemes since the target channel is determined based on previous updated believable information and instantaneous observation when the secondary user is interrupted. Compared to post-sensing approaches, pre-sensing schemes only consider historical and stationary channel states for spectrum handoff. Since instantaneous channel information is not taken into consideration, the selected target channel will not always be optimal even though decent channel is selected based on historical information. Without considering the impact of time for sensing strategies, the pre-sensing methods will outperform random channel selection scheme on either throughput or waiting time because user will select channel based on the historical and stationary channel information, rather than just randomly selecting channel. On the other hand, the post-sensing methods will outperform pre-sensing scheme because the post-sensing schemes update channel state information for every time slot. The channel selection strategy is based on the latest channel states and the optimal channel can therefore be selected.

Nevertheless, the post-sensing methods in general require excessive time in spectrum sensing, especially under crowded network traffic. This situation is not permissible in spectrum handoff since the allowable time duration is considered limited for the CR user to vacate its current channel for the primary user. Furthermore, it is assumed in most of the existing pre-sensing and post-sensing strategies that all the channels within the network can be correctly sensed, which is not realistic in practice. In other words, the transition probabilities of all the channel states are not always available to the CR users within the network. The algorithm in [15] was proposed to estimate the transition probability by adopting the maximum likelihood function. However, the converging speed for the estimation algorithm becomes intolerably slow under small value of transition probability.

Existing works have been proposed to design feasible sensing time for CR users in order to achieve better system performance [3, 16–18]. By partially sensing available frequency channels, the POMDP framework proposed in [9, 19–21] can be adopted to acquire network information. Optimal channel sensing is acquired in [16] to achieve maximum throughput for secondary user with consideration of sensing time overhead in the formulated POMDP problem. However, the proposed policy is designed for a single CR user which cannot be applied to multiple CR users. The scheme proposed in [17] obtains optimal allocation of sensing time in CR networks. However, the proposed approach can only be exploited in the case that there are two primary channels, which is not suitable for generalized networks with multi-channel operations. The POMDP-based dynamic spectrum access (DSA) MAC protocols are proposed in [3, 18] to provide CR users with the optimal policies for spectrum handoff in order to achieve maximum throughput. However, in [3, 18], considering the cases with multiple CR users, only one CR user can win and access a single channel while the remaining CR users are not allowed to adopt the other available channels. In other words, unselected idle channels will become wasted spectrum holes since these channels are not considered optimal channel to be selected by the remaining CR users.

3 Proposed POMDP-based spectrum handoff (POSH) scheme

In this section, the POMDP framework will be utilized to model the spectrum handoff problem in a slotted overlay CR network. The proposed POSH scheme is exploited under the single CR user scenario in Sect 3.1. The protocol implementation of proposed POSH scheme will be explained in Sect. 3.2.

3.1 System model for POSH scheme

A Markov decision process (MDP) refers to a discrete time stochastic control process that conducts decision-making based on the present state information $s_k \in \boldsymbol{S}_k$, where $\boldsymbol{S}_k$ represents the entire state space at the kth time slot. Considering the realistic case that not all the current states are obtainable, the partially observable Markov decision process (POMDP) [9] is utilized to determine the decision policy based on the partially available information and the observations from external environment. The schematic diagram of POMDP framework is illustrated in Fig. 1. In general, optimization techniques are exploited in order to obtain the solution for the POMDP-based problem.

In the considered CR network, there are N channels that are available to be accessed by both the primary and secondary users. Based on the secondary user's point of view,

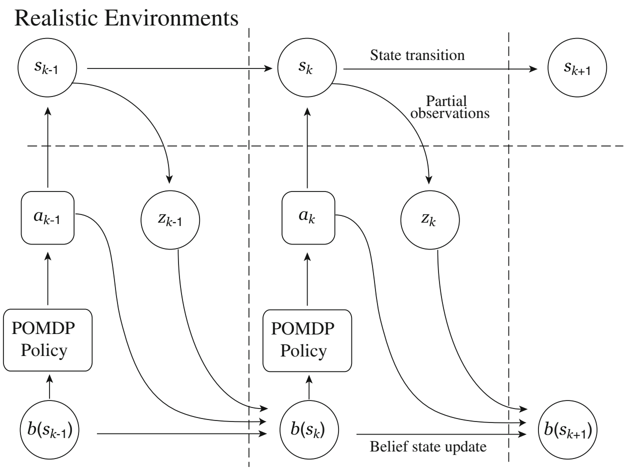

Fig. 1 The schematic diagram of the POMDP framework

each channel is assumed to be either in the busy state, i.e. occupied by the primary user, or in the idle state, i.e. free to be accessed. In other words, the channel state information is the channel occupancy condition of primary user, i.e. zero represents the idle state and one stands for the busy state. Due to the excessive time for sensing entire spectrum, the secondary user only partially senses one channel at once, i.e. only the channel condition of a single channel can be obtained. The channel conditions of the other channels that are unobservable will be estimated and updated based on prediction via the belief state. Considering that $c_{i,k}$ denotes the state of the ith channel in time slot k, the state of the entire network in the kth time slot can be written as

$$s_k = [c_{1,k}, \ldots, c_{i,k}, \ldots, c_{N,k}], c_{i,k} \in \{0,1\}, \quad \forall s_k \in \boldsymbol{S}_k \tag{1}$$

where $c_{i,k} = 0$ indicates the idle state and $c_{i,k} = 1$ represents the busy state. Since not all the states are directly observable within the POMDP setting, a set of observations $z_k \in \boldsymbol{Z}_k$ is essential to provide an indication about which state the environment should be located. The observations can be considered with probabilistic nature, where an observation function o is defined as a probability distribution over all possible observations z_k for each action a_k and resulting states s_{k+1}, i.e.

$$o(s_{k+1}, a_k, z_k) = P_r(z_k | a_k, s_{k+1}) \tag{2}$$

$\forall z_k \in \boldsymbol{Z}_k, a_k \in \boldsymbol{A}_k, s_{k+1} \in \boldsymbol{S}_{k+1}$ where $\boldsymbol{A}_k$ stands for the action set at the kth time slot. The parameter a_k denotes the action chosen by the POMDP formulation and s_{k+1} is the resulting state after executing action a_k. The most essential part in spectrum handoff is the target channel selection, which is defined as the action set within the POMDP framework. In other words, an action a_k at the time instant k is to appropriately choose the target handoff channel from the entire N channels within the CR network, i.e. $a_k = \{1, \ldots, N\}$. After the execution of an action, the channel state can consequently be observed. The set of observations $z_k \in \{0,1\}$ can be defined as the sensing outcome, where 0 represents the idle state and 1 stands for the busy state.

The belief state, i.e. the information state, is developed to reveal the state of environment and help to behave truly and effectively in a partially observable world. The belief state $b(\boldsymbol{S}_k)$ is a statistic distribution over the state space $\boldsymbol{S}_k$; while $b(s_k)$ corresponds to the probability of state s_k with $\sum_{s_k \in \boldsymbol{S}_k} b(s_k) \leq 1$. It is noticed that the belief state comprises a sufficient statistical information for the past history, including all the actions and observations that can provide a basis for decision-making under environmental uncertainties. Furthermore, the essential part of belief state is that it can be updated after each corresponding action in order to incorporate one additional step of information into the history. It is considered beneficial to capture the variations from dynamic environment and consequently obtain more accurate information of the environment. As shown in Fig. 1, the updated belief state is acquired as the outcome of state estimator, which is composed by the inputs of observation, action, and previous belief state. Therefore, $b(s_{k+1})$ that corresponds to probability of state s_{k+1} at time slot $k+1$ will be equal to

$$\frac{P_r\{\text{state transition from all possible states} \, \forall \text{s}_\text{k} \in \boldsymbol{S}_\text{k} \text{ to } \text{s}_{\text{k}+1}\}}{P_r\{\text{state transition from all possible states } \forall s_k \in \boldsymbol{S}_k \text{ to all possible states } \forall s_{k+1} \in \boldsymbol{S}_{k+1}\}}.$$

It can be obtained as

$$\begin{aligned} b(s_{k+1}) &= P_r(s_{k+1} | b(\boldsymbol{S}_k), a_k, z_k) \\ &= \frac{o(s_{k+1}, a_k, z_k) \sum_{s_k \in \boldsymbol{S}_k} \Gamma(s_k, a_k, s_{k+1}) b(s_k)}{P_r(z_k | b(\boldsymbol{S}_k), a_k)} \end{aligned} \tag{3}$$

where $b(s_k)$ indicates the former belief state of s_k. The parameter $\Gamma(s_k, a_k, s_{k+1})$ represents the state transition probability from s_k to s_{k+1} according to the action a_k, i.e. $\Gamma(s_k, a_k, s_{k+1}) = P_r(s_{k+1} | a_k, s_k)$. In the numerator of (3), the summation of $\Gamma(s_k, a_k, s_{k+1})$ updates the belief state by prediction based on state transition probabilities; while the observation $o(s_{k+1}, a_k, z_k)$ updates the belief state by observing the sensing outcome z_k with executing action a_k and resulting state s_{k+1}. The denominator of (3) can be considered as a normalizing factor, which is obtained as

$$P_r(z_k|b(\boldsymbol{S}_k), a_k) = \sum_{s_k \in \boldsymbol{S}_k,} \sum_{s_{k+1} \in \mathbf{S}_{k+1}} o(s_{k+1}, a_k, z_k) \cdot \Gamma(s_k, a_k, s_{k+1}) b(s_k) \tag{4}$$

In the proposed POSH scheme, the transition probability can be determined by modeling a channel as a single-server queue with interarrival time and service time are exponentially distributed with arrival rate λ and service rate μ. Note that exponential distribution is often assumed in the modeling of CR networks as in [12, 22, 23]. With the assumption of exponential interarrival time and service time, the inherent Markov property can be utilized in the mathematical analysis. By assuming Poisson traffic for the arrival packets, the probability distribution of arriving packets can be represented as

$$P_r(n_{\lambda,k} = x) = \frac{e^{-\lambda}\lambda^x}{x!} \tag{5}$$

where $n_{\lambda,k}$ denotes the number of arriving packets in the kth time slot. With the execution of action a_k, the channel transition probability $\tau(c_{i,k}, a_k, c_{i,k+1})$ represents the transition from the present channel state $c_{i,k}$ to the channel state $c_{i,k+1}$ at the next time slot. By adopting the result from (5), the transition probability from idle to idle state for a channel $c_{i,k}$ after executing action a_k can be acquired as

$$\begin{aligned} \tau(c_{i,k} = 0, a_k, c_{i,k+1} = 0) &= \frac{P_r(n_{\lambda,k} = 0, c_{i,k} = 0)}{P_r(c_{i,k} = 0)} \\ &= \frac{P_r(n_{\lambda,k} = 0) \cdot P_r(c_{i,k} = 0)}{P_r(c_{i,k} = 0)} = e^{-\lambda} \end{aligned} \tag{6}$$

for $i = \{1, \ldots, N\}$. It is noted that the second equality in (6) is attributed to the fact that the availability of a channel in any kth time slot is independent to the total number of arrival packets within the same slot. On the other hand, the transition probability for the channel $c_{i,k}$ coming from busy to idle state can be expressed as

$$\begin{aligned} \tau(c_{i,k} = 1, a_k, c_{i,k+1} = 0) &= \frac{P_r(n_{\lambda,k} = 0, n_{\lambda,k-1} > 0, T_{s,k} \le 1)}{P_r(c_{i,k} = 1)} \\ &= \frac{P_r(n_{\lambda,k} = 0) \cdot P_r(n_{\lambda,k-1} > 0) \cdot P_r(T_{s,k} \le 1)}{P_r(c_{i,k} = 1)} \end{aligned} \tag{7}$$

where the second equality is also contributed to the independency of the three probabilities within the numerator of (7). The parameter $T_{s,k}$ is the total service time in time slot k which includes both the time durations for serving packets coming into this kth time slot and the remaining packets acquired from the previous $(k-1)$th slot. It is assumed that $\gamma_1, \gamma_2, \ldots, \gamma_\alpha$ are the exponential random variables of a packet service time with mean value of 1/μ, where α represents the number of packets arrived from the previous $(k-1)$th slot, i.e. $n_{\lambda,k-1} = \alpha$. The time server takes for serving these α packets within the kth slot is denoted as $T_{\alpha,k} = \sum_{j=0}^{\alpha} \gamma_j P_r(n_{\lambda,k-1} = \alpha)$. Therefore, the third term in the numerator of (7) can be rewritten as

$$\begin{aligned} P_r(T_{s,k} \le 1) = P_r(T_{s,k-1} \le 1) \cdot P_r(T_{\alpha,k} \le 1) \\ + P_r(T_{s,k-1} > 1) \cdot P_r(T_{s,k-1} - 1 + T_{\alpha,k} \le 1) \end{aligned} \tag{8}$$

which is the combination of two cases as follows: (a) the packets can be served in both previous and this time slots; and (b) the packets have not been entirely served in the previous slot but are able to be served in this time slot. Furthermore, the denominator of (7) that represents the probability for a busy channel can be expressed as

$$\begin{aligned} P_r(c_{i,k} = 1) = P_r(n_{\lambda,k-1} > 0) + P_r(n_{\lambda,k-1} = 0) \\ \cdot P_r(T_{s,k-1} > 1) \end{aligned} \tag{9}$$

Based on (6) and (7), the transition probabilities from idle to busy state and from busy to busy state can be respectively obtained as $\tau(c_{i,k} = 0, a_k, c_{i,k+1} = 1) = 1 - \tau(c_{i,k} = 0, a_k, c_{i,k+1} = 0)$ and $\tau(c_{i,k} = 1, a_k, c_{i,k+1} = 1) = 1 - \tau(c_{i,k} = 1, a_k, c_{i,k+1} = 0)$. As a result, by assuming that each channel $c_{i,k}$ is independent with each other for $i = 1$ to N, the transition probability for the entire network $\Gamma(s_k, a_k, s_{k+1})$ can be obtained as

$$\Gamma(s_k, a_k, s_{k+1}) = \prod_{i=1}^{N} \tau(c_{i,k} = \varsigma_1, a_k, c_{i,k+1} = \varsigma_2) \tag{10}$$

where $c_{i,k} \in s_k$, $c_{i,k+1} \in s_{k+1}$, and $\varsigma_1, \varsigma_2 \in \{0, 1\}$. In order to ensure optimal decision is made by adopting the POMDP-based approach, it is necessary to provide a measurement such as to evaluate the cost or to reward the update from each state. An immediate reward function $r(s_k, a_k, s_{k+1}, z_k)$ is defined to represent the reward by executing action a_k which turns from state s_k to s_{k+1} associated with the observation z_k. Since both the state transition and observation function are probabilistic, the expected immediate reward $R(s_k, a_k)$ can be obtained as

$$\begin{aligned} R(s_k, a_k) = \sum_{s_{k+1} \in \boldsymbol{S}_{k+1},} \sum_{z_k \in \boldsymbol{Z}_k} \Gamma(s_k, a_k, s_{k+1}) \cdot o(s_{k+1}, a_k, z_k) \\ \cdot r(s_k, a_k, s_{k+1}, z_k) \end{aligned} \tag{11}$$

$R(s_k, a_k)$ is the expected reward deriving by executing the action a_k with current state s_k. It is equivalent to the state transition probabilities from current state s_k to every possible state $\forall s_{k+1} \in \boldsymbol{S}_{k+1}$ times the reward function by executing action a_k which turns from state s_k to s_{k+1}. All possible states transition and corresponding rewards are considered. It is noticed that the expected immediate reward function $R(s_k, a_k)$ is denoted as the one-step value function since only the present reward is the major concern. The optimal policy can be directly determined by adopting the reward function as in (11). However, certain period of time is always considered to evaluate the value of reward. Therefore, the decision policy by exploiting the POMDP is determined by optimizing the L-step value function with $L \geq 1$.

The main objective of proposed POSH scheme is to select a target channel that has the minimum waiting time, i.e. the smallest number of waiting slots required for the CR user if the target channel is still occupied by the primary user. Note that the time required for spectrum handoff of a CR user is defined as the time duration from the termination of packet transmission in one channel to the starting time of retransmission in another channel. The waiting time, which is served as the cost function within the POMDP framework, will be minimized by the POSH scheme with the selection of an optimal channel in the spectrum handoff process. The immediate cost is considered as the total number of waiting slots n_w required by the secondary user while executing a specific action, i.e. $a^* = a_k$, for spectrum handoff. Consequently, we define the expected cost $C(s_k, a_k)$ as

$$\begin{aligned}
C(s_k,a_k) &= \frac{1}{R(s_k,a_k)} = n_w^{a_k} \\
&= E[n_w = \ell | a^* = a_k] = \sum_{\ell=0}^{\infty} \ell \cdot P_r(n_w = \ell | a^* = a_k) \\
&= \sum_{\ell=1}^{\infty} \ell \cdot P_r\left(\bigcap_{p=1}^{\ell} c_{a_k,k+p-1} = 1, c_{a_k,k+\ell} = 0, \right) \\
&= \sum_{\ell=1}^{\infty} \ell \cdot \tau(c_{a_k,k+\ell-1} = 1, a_k, c_{a_k,k+\ell} = 0) \\
&\quad \cdot P_r\left(\bigcap_{p=1}^{\ell} c_{a_k,k+p-1} = 1 \right) \\
&= \sum_{\ell=1}^{\infty} \ell \cdot \ \tau(c_{a_k,k+\ell-1} = 1, a_k, c_{a_k,k+\ell} = 0) \\
&\quad \cdot P_r(c_{a_k,k} = 1) \cdot \tau(c_{a_k,k} = 1, a_k, c_{a_k,k+1} = 1)^{\ell-1}
\end{aligned} \tag{12}$$

where $c_{a_k,k}$ denotes the channel state after selecting channel a_k at the time instant k. Noted that the cost function $C(s_k, a_k)$ is the expected waiting time after selecting channel a_k based on the present state s_k and is defined as the inverse of the expected immediate reward $R(s_k, a_k)$ in (11) that will be utilized in the POMDP formulation.

3.2 Protocol implementation of POSH scheme

An overlay slotted CR network with partially observable information is considered for the POSH scheme, which indicates that the secondary user is not allowed to coexist with the primary user while the time duration for packet transmission is divided into time slots. As shown in Fig. 1, partial channel information $o(s_{k+1}, a_k, z_k)$ is assumed available to be observed by the secondary users, which will be exploited for the update of belief state $b(s_{k+1})$ as in (3). The secondary user utilizes the updated belief state in order to estimate the channel state of CR network.

According to POMDP framework, an L-step value function will be adopted to obtain corresponding action that results in minimal waiting time after handoff process. Note that L-step indicates that the current action is determined based on the future information from the predicted L time slots. In other words, based on the L-step value function $V_L^*[b(s_k)]$, which is mapped from the belief state space, CR user will determine feasible action to take for achieving the highest reward. The L-step value function for CR user can be obtained as

$$\begin{aligned}
V_L^*[b(s_k)] = \max_{a_k \in \boldsymbol{A}_k} & \sum_{s_k \in \boldsymbol{S}_k} b(s_k) \cdot R(s_k, a_k) \\
& + \rho \sum_{z_k \in \boldsymbol{Z}_k} P_r(z_k | b(\mathbf{S}_k), a_k) \cdot V_{L+1}^*[b(s_{k+1})]
\end{aligned} \tag{13}$$

where ρ is denoted as a discount factor for convergence control of the value function. The probability $P_r(z_k | b(\mathbf{S}_k), a_k)$ is defined as in (4). Notice that the computation of L-step value function in (13) is considered complex and in general difficult to solve. The one-step value function as was adopted in [3, 18] is usually employed, i.e., $V_1^*[b(s_k)] = \max_{a_k \in \boldsymbol{A}_k} \sum_{s_k \in \boldsymbol{S}_k} b(s_k) \cdot R(s_k, a_k)$, since it is feasible and sufficient for generic network circumstance.

Furthermore, the dimension of belief state $\boldsymbol{S}_k$, where $\boldsymbol{S}_k = [s_{k,1}, \ldots, s_{k,M}]$ with $M = 2^N$, grows exponentially as the number of channel is augmented, which makes it difficult to be adopted for practical implementation. Based on the property of independent channels, a linear state vector as proposed in [24] can be utilized to denote the set of busy channel probability, i.e., $\Omega_k = [\omega_{1,k}, \ldots \omega_{i,k}, \ldots \omega_{N,k}]$ where $\omega_{i,k}$ represents the probability that channel i is busy in time slot k. Therefore, the belief state $b(s_k)$ within the value function can be replaced by $\omega_{i,k}$ such that the statistics dimension of the considered CR network can be effectively reduced from 2^N to N. Therefore, the one-step value function can be rewritten as

$$\begin{aligned}
V_1^*[b(s_k)] &= \max_{a_k \in \boldsymbol{A}_k} \sum_{s_k \in \boldsymbol{S}_k} b(s_k) \cdot R(s_k, a_k) = V_1^*[\omega_{i,k}] \\
&= \max_{a_k \in \boldsymbol{A}_k} \sum_{\omega_{i,k} \in \Omega_k} \omega_{i,k} \cdot R(\omega_{i,k}, a_k) \\
&= \min_{a_k \in \boldsymbol{A}_k} \sum_{\omega_{i,k} \in \Omega_k} \omega_{i,k} \cdot C(\omega_{i,k}, a_k) \\
&= \min_{a_k \in \boldsymbol{A}_k} \sum_{\omega_{i,k} \in \Omega_k} \omega_{i,k} \cdot n_w^{a_k}
\end{aligned} \tag{14}$$

where $n_w^{a_k}$ can be obtained from (12) as the expected number of waiting time slots if action a_k is executed. Consequently, the action taken in time slot k for minimizing the expected waiting time slots is chosen as

$$a_k^* = \arg\min_{a_k \in \boldsymbol{A}_k} \omega_{a_k,k} \cdot \sum_{\ell=1}^{\infty} \ell \cdot \tau(c_{a_k,k+\ell-1} = 1, a_k, c_{a_k,k+\ell} = 0) \cdot \tau(c_{a_k,k} = 1, a_k, c_{a_k,k+1} = 1)^{\ell-1} \tag{15}$$

Note that the computational complexity for obtaining (15) will be $O(N \cdot \ell)$ where ℓ corresponds to the total number of time slots required for the computation of expected waiting time. In practice, finite value of ℓ will be adopted and the convergence behavior of ℓ will be evaluated in Sect. 6. At the beginning of time slot where spectrum handoff occurred, the CR user will choose a target channel that possesses minimum waiting slots based on the results obtained from (15). After switching to the target channel, the CR user will conduct the sensing task for observing the newly updated channel state even though only partial state information is obtainable. After waiting for the required time slots that are determined by the POSH scheme, the secondary user can start to conduct its packet transmission within the target channel.

4 Performance analysis of proposed POSH scheme

In this section, the analytical model for proposed POSH scheme is derived in Sect. 4.1 in order to analyze its performance. The models for both the non-spectrum handoff (NSH) scheme and the randomly choose strategy (RCS) will also be demonstrated. As the NSH scheme is performed, CR user will remain to stay at the current channel even though this channel is sensed busy at the beginning of time slot. The CR user has to wait for the next spectrum hole within the same channel for potential data transmission. On the other hand, with the RCS scheme, CR user will randomly select and switch to another channel for channel access if the current channel is occupied by a primary user. As the proposed POSH scheme is adopted, the CR user will perform the POMDP policy to determine its target channel when the current channel is unavailable due to the existence of primary user. Based on the POMDP policy, optimal channel can be obtained with minimal waiting time during spectrum handoff for CR users. The effectiveness of analytical models is to serve as validation purpose for these schemes, which will be compared with simulation results as in Sect. 6. Furthermore, the analytical models for these three handoff algorithms under practical considerations will also be derived and explained in Sect. 4.2.

4.1 Analytical modeling of spectrum handoff schemes

The analytical models for the three spectrum handoff schemes, including NSH, RCS, and POSH mechanisms, will be derived and studied in this subsection. It is noted that the parameter a_k defined in the previous section emphasizes the action executed at the time slot k, and consequently to acquire the reward function $R(s_k, a_k)$ at time k as in (12). However, it is not essential to point out a particular moment in analytical expressions, i.e. the subscript k for each action a_k will be neglected in the remaining discussion. Instead, the action a_s is defined as the selected target channel which can be either the current channel a_c or the destination channel a_d, where $a_d \in \{1, \ldots, N\}$ and $a_d \neq a_c$. The channel selected by the action a_s will be retained for a time period until another action is executed.

4.1.1 NSH scheme

Let $n_{w,nsh}$ be the expected waiting time if the NSH scheme is performed, the CR user will not switch to other channels but stay at its current channel a_c to wait for the next spectrum hole. Similar to (12), the expected waiting time $n_{w,nsh}^{a_c}$ can be obtained as

$$\begin{aligned} n_{w,nsh} &= E[n_w = \ell | a^* = a_c] \\ &= \sum_{\ell=1}^{\infty} \ell \cdot P_r\left(\bigcap_{p=1}^{\ell} c_{a_c,k+p-1} = 1, c_{a_c,k+\ell} = 0 \right) \\ &= \sum_{\ell=1}^{\infty} \ell \cdot \tau(c_{a_c,k+\ell-1} = 1, a_c, c_{a_c,k+\ell} = 0) \\ &\quad \cdot \tau(c_{a_c,k} = 1, a_c, c_{a_c,k+1} = 1)^{\ell-1} \end{aligned} \tag{16}$$

It is noted that the third equality in (16) indicates that the NSH method is adopted under the condition that no spectrum handoff is executed at the current time k even though the current channel state is busy, i.e. $c_{a_c,k} = 1$. Therefore, the expected waiting time by exploiting NSH scheme $n_{w,nsh}$ can be obtained by assigning $P_r(c_{a_c,k} = 1) = 1$ in (12).

4.1.2 RCS scheme

Let $n_{w,rcs}$ be the expected waiting time as the RCS scheme is performed, the CR user will randomly switch to a target channel a_s from the network spectrums 1 to N to acquire channel access. Therefore, for the calculation of $n_{w,rcs}$, there exists probability $P_r(a^* = a_s)$ for the action a_s to select one channel within all the N channels as below:

$$\begin{aligned} n_{w,rcs} &= \sum_{a_s=1}^{N} E[n_w = \ell | a^* = a_s] \cdot P_r(a^* = a_s) \\ &= \frac{1}{N} \cdot E[n_w = \ell | a^* = a_c] + \frac{1}{N} \cdot \\ &\quad \sum_{a_d=1, a_d \neq a_c}^{N} E[n_w = \ell | a^* = a_d] \\ &= \frac{1}{N} \cdot n_{w,nsh} + \frac{1}{N} \cdot \sum_{a_d=1, a_d \neq a_c}^{N} n_w^{a_d} \end{aligned} \tag{17}$$

where $n_{w,nsh}$ is defined in (16). The parameter $n_w^{a_d}$ denotes the expected waiting time by selecting the channel a_d without staying at the current channel a_c, which can be computed from (12) by assigning $a_k = a_d$. In the case that the traffic pattern of all the N channels are identical, (17) can be reformulated by incorporating (16) as

$$\begin{aligned} n_{w,rcs} &= \frac{1}{N} \left[1 + \sum_{a_d=1, a_d \neq a_c}^{N} P_r(c_{a_d,k} = 1) \right] \cdot \\ &\quad \sum_{\ell=1}^{\infty} \ell \cdot \tau(c_{a_s,k+\ell-1} = 1, a_s, c_{a_s,k+\ell} = 0) \\ &\quad \cdot \tau(c_{a_s,k} = 1, a_s, c_{a_s,k+1} = 1)^{\ell-1} \end{aligned} \tag{18}$$

As illustrated in (18), the expected waiting time obtained from the RCS scheme should be less or at least equal to that acquired from NSH method, i.e. $n_{w,rcs} \leq n_{w,nsh}$, since the probability $P_r(c_{a_d,k} = 1) \leq 1 \; \forall a_d \neq a_c$. The benefit for using the random channel selection scheme over the case to remain at the current channel can therefore be analytically revealed.

4.1.3 POSH scheme

The channel selection behavior of both the NSH and RCS schemes are straightforward such that their analytical models can be directly expressed by stationary probabilities and transition probabilities. The proposed POSH scheme, on the other hand, determines its target channel by the POMDP policy which is derived from the belief state at each step as shown in Fig. 1. Since the value of belief state is obtained from the observation, action, and former belief state at each step, it is not possible to predict and obtain the target channel in advance. Nevertheless, the analytical model of POSH scheme can still be approximately estimated since the updated belief state will gradually approach to stationary probability considering that the network is not rapidly varying. Let $n_{w,posh}$ be the expected waiting time as the POSH scheme is adopted, the resulting formulation is similar to the RCS scheme except that the probabilistic distribution $\tilde{P}_r(a^* = a_s)$ of action a^* is considered non-uniform in this case, i.e.

$$n_{w,posh} = \sum_{a_s=1}^{N} E[n_w = \ell | a^* = a_s] \cdot \tilde{P}_r(a^* = a_s) \tag{19}$$

Let $n_w^{a_{d1}}$ and $n_w^{a_{d2}}$ be the first and second minimum expected waiting time resulted from action a_{d1} and a_{d2} by adopting the proposed POSH scheme, respectively. These two actions can be expressed as

$$\begin{aligned} a_{d1} &= \arg \min_{\forall a_s} n_w^{a_s} \\ a_{d2} &= \arg \min_{\forall a_s, a_s \neq a_{d1}} n_w^{a_s} \end{aligned} \tag{20}$$

where the expected waiting time by conducting spectrum handoff $n_w^{a_s}$ can be obtained from (12) by assigning $a_k = a_s$. The statistical distribution of the chosen action a^* in proposed POSH scheme can be acquired as

$$\tilde{P}_r(a^* = a_s) = \begin{cases} 1, & \text{for } a_s = a_{d1}, \text{and } n_{w,nsh}^{a_{d1}} \leq n_w^{a_{d2}} \\ 0.5, & \text{for } a_s = a_{d1} \text{ or } a_{d2}, \text{ and } n_{w,nsh}^{a_{d1}} > n_w^{a_{d2}} \\ 0, & \text{otherwise} \end{cases} \tag{21}$$

where $n_{w,nsh}^{a_{d1}}$ denotes the expected waiting time by staying at the same channel a_{d1} that was determined in the previous time slot. It can be observed that the first case $\tilde{P}_r(a^* = a_s) = 1$ in (21) happens under the situation that a_s is selected as a_{d1} and $n_{w,nsh}^{a_{d1}} \leq n_w^{a_{d2}}$. This case indicates that the expected waiting time obtained by staying at the channel a_{d1} will be comparably smaller or equal to that from the POSH-based spectrum handoff scheme $n_w^{a_{d2}}$. It is suggested not to conduct spectrum handoff to another channel a_{d2}, which results in the probability of staying at the channel $a^* = a_{d1}$ to be equal to 1, i.e. $\tilde{P}_r(a^* = a_{d1}) = 1$. On the other hand, the other case with $n_{w,nsh}^{a_{d1}} > n_w^{a_{d2}}$ should also be taken into account. It is considered that the previous selected channel is a_{d1}. If the expected waiting time resulted from a_{d1} without spectrum handoff is larger than that from a_{d2} by conducting spectrum handoff, the corresponding CR user will decide to switch into channel a_{d2}. Nevertheless, instead of remaining at channel a_{d2}, CR user will choose a_{d1} again at the next spectrum handoff since the waiting time $n_w^{a_{d1}}$ is smaller than $n_{w,nsh}^{a_{d2}}$ as depicted in (20). As a result, the channels a_{d1} and a_{d2} will alternatively be chosen depending on the previous action of the CR user, which results in $\tilde{P}_r(a^* = a_s) = 0.5$ for either $a^* = a_{d1}$ or a_{d2} if $n_{w,nsh}^{a_{d1}} > n_w^{a_{d2}}$.

Note that (21) presents the generic methodology to analyze the optimal policy. If the channel observed at the current time still results in the minimum waiting time as the selected optimal channel in the previous time slot, that channel will be chosen by the optimal policy at the current time slot. On the other hand, if the optimal channel selected at the previous time slot does not result in the minimum waiting time in the current time, the channel with second minimum waiting time observed at the previous time slot will be considered as the smallest one and is chosen at the current time. Consequently, the channels with the minimum and the second minimum waiting time slots will be alternatively selected in order to achieve the optimal policy. For example, it is considered that the minimum expected waiting time resulted from action a_{d1} at the $(k-1)$th time slot is assumed to occur at channel 1, i.e. $a_{d1}(k-1) = CH_1$. The expected waiting time slots observed at time k to stay at the same channel $a_{d1}(k-1) = CH_1$ is denoted as $n_{w,nsh}^{a_{d1}(k-1)=CH_1}(k)$. In the case that $n_{w,nsh}^{a_{d1}(k-1)=CH_1}(k) < n_{w}^{a_{d2}(k-1)=CH_2}(k)$, the optimal channel at the kth time slot will be selected as $a_{d1}(k) = a_{d1}(k-1) = CH_1$ as denoted in the first case of (21). On the other hand, if $n_{w,nsh}^{a_{d1}(k-1)=CH_1}(k) > n_{w}^{a_{d2}(k-1)=CH_2}(k)$, the target channel at time k will be chosen as $a_{d1}(k) = a_{d2}(k-1) = CH_2$ since CH_2 is observed to possess smaller waiting time at time k. Progressively, considering the $(k+1)$th time slot with $n_{w,nsh}^{a_{d1}(k)=CH_2}(k+1) > n_{w}^{a_{d2}(k)=CH_1}(k+1)$, the optimal channel at time $(k+1)$ will be selected as $a_{d1}(k+1) = a_{d2}(k) = a_{d1}(k-1) = CH_1$ owing to the reason that the expected waiting time resulting from CH_1 is smaller than that to stay at CH_2. It can be observed that the target channel will be alternatively selected between CH_1 and CH_2, which reflects the second case of (21).

Although there exist overheads such as channel switching time while implementing the POMDP policy, it will be beneficial to decrease the expected waiting time after conducting channel switching. The secondary user has to select the channel with currently expected minimum waiting time as the target channel, otherwise, the expected waiting time will arise because of the increased probability of channel occupancy by primary user. The penalty, i.e. waiting time, of staying at the current channel which is not the optimal target is much larger than the delay of channel switching to an optimal channel. Consequently, according to the probabilistic distribution of the action a^* as obtained in (21), the expected waiting time of proposed POSH scheme can be reformulated from (19) as

$$n_{w,posh} = \begin{cases} n_{w,nsh}^{a_{d1}}, & \text{if } n_{w,nsh}^{a_{d1}} \leq n_{w}^{a_{d2}} \\ \frac{1}{2} n_{w}^{a_{d1}} + \frac{1}{2} n_{w}^{a_{d2}}, & \text{if } n_{w,nsh}^{a_{d1}} > n_{w}^{a_{d2}} \end{cases} \tag{22}$$

The benefits of these analytical models can provide the flexibility to evaluate the performance of different spectrum handoff schemes in advance. A feasible mechanism can be selected by the CR user that will be most beneficial in specific circumstance. Performance validation and comparison between these schemes will be illustrated later in Sect. 6

4.2 Practical considerations for spectrum handoff schemes

In the previous subsection, according to the POMDP-based channel selection policies, the analytical results show that the proposed POSH scheme should provide reduced expected waiting time than that from both the RCS and NSH algorithms at the occurrence of spectrum handoff. However, the ideal circumstance is assumed where only data transmission is considered in a specific time slot. In practice, it is inevitably to spend a period of time to detect the network condition and to make sure that spectrum handoff can be successfully executed. It is intuitive to recognize that additional periods of time can be required by the proposed POSH scheme for conducting the POMDP-based channel selection with partially observable channel information. Therefore, practical consideration for spectrum handoff that involves channel sensing and handshaking will be discussed in this subsection. The performance difference between these schemes under the practical circumstance can consequently be observed.

As mentioned in previous sections, the time slotted system is considered in this paper as shown in Fig. 2. It is required for the CR user to perform spectrum sensing on its current channel at the beginning of each time slot in order to assure the availability of the present channel. Note that perfect channel sensing is assumed for each CR user. If the outcome of spectrum sensing is observed to be idle, the CR user can remain in the same slot to conduct packet transmissions. The receiver will consequently return an acknowledgement (ACK) frame at the end of this slot to acknowledge the reception of data packet. On the other hand, in the case that the sensing outcome is busy, additional messages will be delivered in order to perform different spectrum handoff schemes which are stated as follows.

4.2.1 NSH scheme

As shown in Fig. 2(a), the behavior for the time-slotted channel with the adoption of NSH scheme is depicted. The time slots t_k and $t_{k+\ell}$ $(for\, \ell \geq 0)$ are utilized to represent the initial and final slots that are considered in the various handoff schemes. At the slot t_k, the intention from primary user to utilize this channel is observed during the spectrum sensing period. Based on the NSH scheme, the CR user will remain silent on the current channel and wait until the primary user to finish its transmissions. Considering that the traffic from primary user has not been observed during the sensing period of slot $t_{k+\ell}$, the CR user can consequently conduct its data transmission. Therefore, the total waiting time by adopting the NSH scheme (i.e. $T_{w,nsh}$) can be expressed as

$$T_{w,nsh} = n_{w,nsh} \cdot T_{slot} + T_{sens} \tag{23}$$

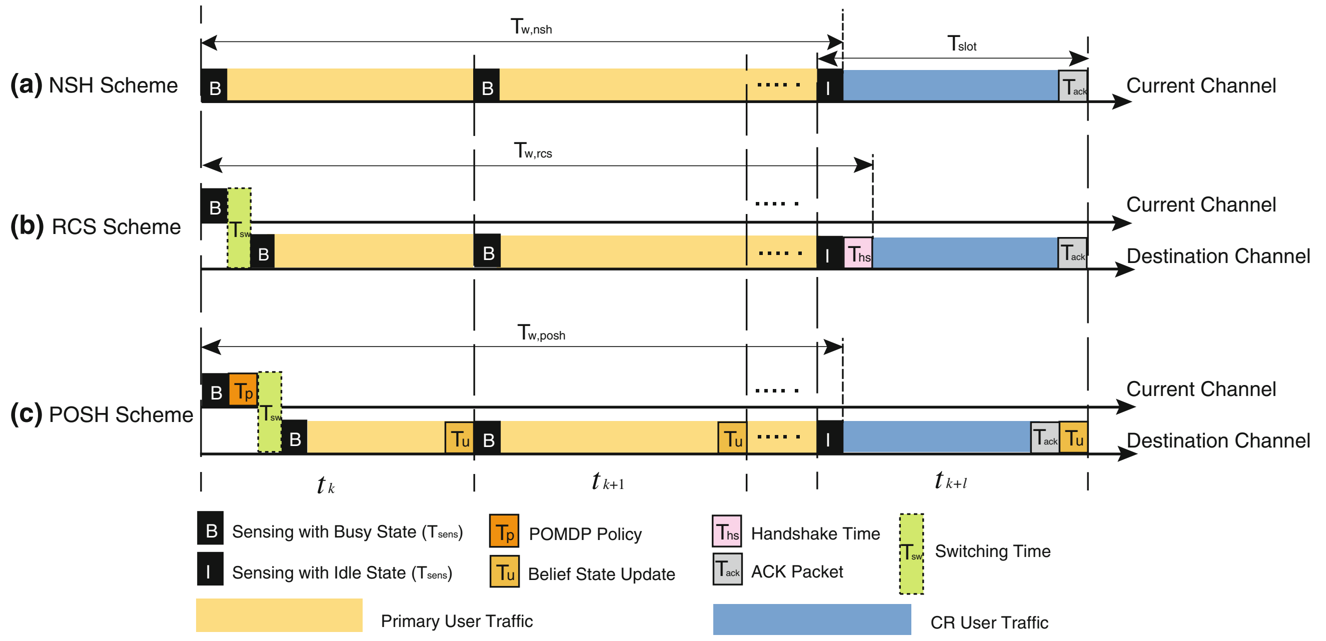

Fig. 2 Practical consideration for the time-slotted channels with different spectrum handoff protocols: **a** NSH scheme, **b** RCS scheme, and **c** POSH scheme

where the expected number of waiting time slots $n_{w,nsh}$ can be obtained from (16). The parameter T_{slot} represents the time duration of each slot and T_{sens} denotes the time interval of sensing period. Furthermore, the net transmission time within a time slot by adopting the NSH method (i.e. $T_{s,nsh}$) can be acquired as

$$T_{s,nsh} = \frac{T_{slot} - T_{sens} - T_{ack}}{n_{w,nsh} + 1} \tag{24}$$

where T_{ack} indicates the required time for returning the ACK packet to conduct handshake for data transmission. Noted that the net transmission time within a time slot $T_{s,nsh}$ in (24) is defined as the data transmission time in a time slot divided by the total time slots required for a transmission. Based on the definition, the total time slots required for a transmission includes the total waiting time $n_{w,nsh}$ and the one time slot for data transmission. It can be observed that the net transmission time $T_{s,nsh}$ will be reduced with larger number of waiting slots $n_{w,nsh}$.

4.2.2 RCS scheme

The practical consideration for the RCS scheme is illustrated in Fig. 2(b). After the initial sensing, the CR transmitter will randomly select a target channel for spectrum handoff, where the determination process is assumed short and negligible. In the RCS scheme, the CR user does not know the channel conditions and it has no idea how long the waiting time will be no matter it stays at the newly sensed channel or switches every time slot to find a free channel. It will not be beneficial for CR user to switch channel at every time slot, which can cause more overheads such as channel switching time for the CR user. Therefore, the RCS scheme adopts the case that CR user randomly selects a new channel and stays in that channel until it becomes available as shown in Fig. 2(b). Consider the RCS scheme with N channels in the network, there are $1/N$ probability that the CR user will remain at its current channel; while there are $(N-1)/N$ possibility switching to another spectrum. The average waiting time of RCS scheme $T_{w,rcs}$ can be obtained as

$$T_{w,rcs} = \frac{1}{N} \cdot T_{w,nsh} + \frac{1}{N} \sum_{a_d=1, a_d \neq a_c}^{N} \Big[P_r(c_{a_d,k} = 1) \cdot n_w^{a_d} \cdot T_{slot} + P_r(c_{a_d,k} = 0) \cdot (T_{sw} + T_{sens}) + T_{sens} + T_{hs} \Big] \tag{25}$$

where $T_{w,nsh}$ is obtained from (23) if the CR user is determined to stay at the existing channel. In the case that the CR user is suggested to switch to another channel, both the switching time T_{sw} and an additional sensing time T_{sens} will be required in order to observe if the destination channel is busy or not. If the target channel is occupied by the primary user with probability $P_r(c_{a_d,k} = 1)$, additional waiting time slots $n_w^{a_d}$ are required until the channel becomes idle to be utilized. On the other hand, if the randomly selected channel is found to be idle with probability $P_r(c_{a_d,k} = 0)$, both the CR transmitter and receiver will spend the time interval T_{hs} for exchanging handshake messages in order to confirm the utilization of target channel. Moreover, the resulting net transmission time in a slot $T_{s,rcs}$ can therefore be acquired as

$$
\begin{aligned}
T_{s,rcs} = {} & \frac{T_{slot} - T_{sens} - T_{hs} - T_{ack}}{N(n_{w,nsh}+1)} + \frac{1}{N}\sum_{a_d=1,a_d\neq a_c}^{N}\frac{1}{n_w^{a_d}+1} \\
& \cdot [P_r(c_{a_d,k}=1)\cdot(T_{slot}-T_{sens}-T_{hs}-T_{ack}) \\
& + P_r(c_{a_d,k}=0)\cdot(T_{slot}-T_{sens}-T_{sw}-T_{sens}-T_{hs}-T_{ack})]
\end{aligned}
\tag{26}
$$

where the computation of (26) can be depicted from Fig. 2(b).

4.2.3 *POSH scheme*

The proposed POSH scheme with practical consideration is shown in Fig. 2(c). It is assumed that all the CR users share identical network condition and observe the same sensing consequence. Since the POMDP policy is simultaneously performed by both CR transmitter and receiver, it is unnecessary to exchange handshake messages with required time duration T_{hs} in order to inform the receiver which spectrum is selected as the target channel. Nevertheless, it is considered that an additional time interval T_p is required by the CR user to implement the POMDP-based policy, and the time T_u to update the belief state at the end of each slot. Consequently, the overall waiting time $T_{w,posh}$ can be represented as

$$
\begin{aligned}
T_{w,posh} = {} & \frac{1}{N}\cdot T_{w,nsh} + \frac{N-1}{N}[P_r(c_{a_s,k}=1)\cdot n_{w,posh}\cdot T_{slot} \\
& + P_r(c_{a_s,k}=0)\cdot(T_{sens}+T_p+T_{sw})+T_{sens}]
\end{aligned}
\tag{27}
$$

where a_s will either be a_{d1} or a_{d2} according to the conditions as stated in (21). The net transmission time in a slot $T_{s,posh}$ becomes

$$
\begin{aligned}
T_{s,posh} = {} & \frac{1}{N(n_{w,posh}+1)}\{(T_{slot}-T_{sens}-T_{ack}-T_u) \\
& + (N-1)[P_r(c_{a_d,k}=1)\cdot(T_{slot}-T_{sens}-T_{ack}-T_u) \\
& + P_r(c_{a_d,k}=0)\cdot(T_{slot}-T_{sens}-T_p-T_{sw}-T_{sens} \\
& - T_{ack}-T_u)]\}
\end{aligned}
\tag{28}
$$

By comparing (27) and (28) from the POSH scheme with (25) and (26) from the RCS scheme, it can be observed that the performance difference between these two protocols will be dominated by (a) the parameters $n_{w,posh}$ and $n_w^{a_d}; \forall a_d = 1, a_d \neq a_c$, and (b) the parameter set (T_p, T_u) and T_{hs}. It is apparently that $n_{w,posh}$ should be smaller than $n_w^{a_d}$ from the RCS algorithm since the POSH scheme is designed to select the minimal waiting time for spectrum handoff. On the other hand, the parameters (T_p, T_u) that are implemented in the POSH scheme should be feasibly chosen in order not to result in excessive computation overheads.

5 Proposed POMDP-based multi-user spectrum handoff (M-POSH) protocol

In Sect. 3, the POSH scheme is introduced to provide the POMDP-based policy for a single CR user in order to determine the target channel for spectrum handoff. In the case that the network exists more than one CR user whose transmissions are interrupted by the primary users at a specific time slot, it is possible that those CR users will select an identical frequency channel for spectrum handoff. Therefore, the consequence of performing the original POSH scheme in multi-user network will result in packet collisions among the CR users that intend to perform spectrum handoff. In this section, a multi-user POSH (M-POSH) protocol is proposed to distribute the CR users to different opportunistic spectrums rather than contending the access right in the same channel. The main purpose of proposed scheme is to reduce the waiting time of entire network by ensuring that every possible spectrum hole can be fully exploited, whereas the fairness for channel access among all CR users can still be maintained. In the following two subsections, the system model and implementation of the proposed M-POSH scheme will be described.

5.1 System model of M-POSH protocol

The system model of proposed M-POSH protocol is described as follows. Unlike the POSH scheme, the M-POSH protocol not only needs to consider the traffic of primary user but also requires to coordinate the channel access among the secondary users. Moreover, instead of exchanging messages among the CR users in distributed manner, a common control channel is utilized to exchange required information between the CR users. There has been arguments regarding whether to utilize a dedicate channel for delivering control messages. The investigation in [25] provides analytical results in order to illustrate the benefits for adopting an dedicated control channel in the network.

The belief state for multi-user can be extended from that for single user as defined in Sect. 3.1 and (1). The belief state of the rth CR user is represented as $b_r(c_{i,k})$, where $c_{i,k} \in \{0, 1\}$ denotes the state of the ith channel in time slot k. Let $\tilde{\mathcal{A}}_k$ denotes the action set representing the channels that are utilized by the CR users at time slot k, i.e. $\tilde{\mathcal{A}}_k \in \boldsymbol{A}_k$ where $\boldsymbol{A}_k$ represents the entire action set which corresponds to all the channels in the network as defined in Sect. 3.1. The update for belief state in multi-user scenario can be obtained as

$$b_r(c_{i,k+1}) = \begin{cases} 0, & \text{if}(z_{r,k} = 0, a_{r,k} = i) \text{ and } a_{r,k} \notin \tilde{\mathcal{A}}_k \\ 1, & \text{if } (z_{r,k} = 1 \text{ and } a_{r,k} = i) \text{ or } a_{r,k} \in \tilde{\mathcal{A}}_k \\ b_r(c_{i,k}) \cdot \tau(c_{i,k} = 0, a_{r,k}, c_{i,k+1} = 0) & \\ +[1 - b_r(c_{i,k})] \cdot \tau(c_{i,k} = 1, a_{r,k}, c_{i,k+1} = 0), & \text{if } a_{r,k} \neq i \end{cases} \quad (29)$$

where $a_{r,k}$ represents the action of the rth user in time slot k, and $z_{r,k}$ denotes the observation of the rth user in slot k. The update process of belief state in (29) is derived via the reduced strategy as in [3] for POMDP-based formulation. For the rth user that decides to take action $a_{r,k}$ to access the ith channel, i.e. $a_{r,k} = i$, the update of the rth user's belief state $b_r(c_{i,k+1})$ at time $(k+1)$ will be determined by its corresponding observation $z_{r,k}$ on the same ith channel at time k. In other words, the update of belief state $b_r(c_{i,k+1}) = 0$ if the ith channel is observed to be in the idle state at time k, i.e. $z_{r,k} = 0$; while $b_r(c_{i,k+1}) = 1$ in the case that $z_{r,k} = 1$. Another condition for the belief state at the $(k+1)$th time slot to be in the busy state is that the action taken at time slot k belongs to the busy channel set, i.e. $a_{r,k} \in \tilde{\mathcal{A}}_k$. Furthermore, if the rth user is using a channel other than the ith channel at current time k, i.e. $a_{r,k} \neq i$, the update process of belief state for the ith channel to remain in the idle state at time $(k+1)$ will be determined as shown in the third term of (29). The idle probability for the belief state at the $(k+1)$th time slot is equal to the idle transition probability $\tau(c_{i,k} = 0, a_{r,k}, c_{i,k+1} = 0)$ times the idle probability at time k plus the busy to idle probability $\tau(c_{i,k} = 1, a_{r,k}, c_{i,k+1} = 0)$ times the busy probability at the same kth time slot.

5.2 Implementation of M-POSH protocol

Without loss of generality, it is assumed that all the existing CR users are initially located and operated on different frequency spectrums. After the channel sensing period, a handshaking time interval is designed in the proposed M-POSH protocol that all the CR users will listen to a common control channel for potential handoff messages. In the case that a CR user intends to conduct spectrum handoff, it will broadcast a handoff message on the control channel in order to announce all the other CR users regarding the change of network condition. Since there may exist more than one CR users that need to broadcast their handoff messages simultaneously in the same time slot, it is possible that these messages from different CR users will collide with each other. A random backoff contention window [26] is therefore utilized in the proposed M-POSH scheme in order to resolve the potential packet collision problem. Each CR user that intends to deliver the handoff message on the control channel will wait for a random number in a pre-specified interval such as to ensure the success of message transmission. It is noted that prioritized channel access between the CR users can also be implemented by assigning different ranges of contention windows among the CR users, e.g. CR users with higher priority channel access should be designed with a smaller range of contention window in order to possess high opportunity to obtain channel access.

The handoff message delivered from each CR user mainly includes the prioritized destination channel list, i.e. $\boldsymbol{D}_{r,k} = [d_{r,1,k}, \ldots, d_{r,i,k}, \ldots, d_{r,N,k}]$ where $d_{r,i,k}$ represents the channel with ith priority for user r to select at time slot k. The determination of channel priorities within the $\boldsymbol{D}_{r,k}$ set is based on the required waiting time for user r to handoff to that specific spectrum, which can be individually computed from (12) as

$$d_{r,i,k} = \arg \min_{\forall a_{r,k} \in \boldsymbol{A}_k, a_{r,k} \neq d_{r,j,k} \forall j < i} n_w^{a_{r,k}} \quad (30)$$

for $i = 1$ to N. Note that the channel with smaller waiting time will be assigned with higher priority in the prioritized destination channel list $\boldsymbol{D}_{r,k}$, i.e. with smaller value of index i. After the handoff messages are broadcast in the control channel during the handshaking time interval, all the CR users that intend to conduct spectrum handoff will determine their destination channel based on the prioritized destination channel lists $\boldsymbol{D}_{r,k}$ from the received handoff messages. The CR user that possesses smaller backoff counter will have higher priority to select the channel with comparably smaller waiting time slots.

The following example is utilized to explain the mechanism of proposed M-POSH protocol. It is considered that there are total of five channels in the network, i.e. $N = 5$. There exists a primary user resides in channel 4 and three CR users r_1, r_2, and r_3 initially locate in channel 1, 2, and 3, respectively. Each user conducts data transmission with its distinct packet arrival rate. In the case that two primary users intend to utilize their original channels 1 and 2 at time slot k, the secondary users r_1 and r_2 will yield from their corresponding channels 1 and 2 after conducting channel sensing. Handoff messages will be broadcast to the control channel by r_1 and r_2 during the handshaking time interval based on the random backoff mechanism. It is assumed that the backoff counter for r_1 and r_2 is randomly selected to be 4 and 7 respectively, which indicates that handoff message from r_1 will be delivered at the 4th minislot and that from r_2 will be broadcast at the 7th minislot. Note that packet collision may happen when the users select the same number of backoff counter, and they will go to the next backoff stage with the window size doubled and continue until the maximum backoff stage is reached. The prioritized destination channel lists within the handoff messages for r_1 and r_2 can be obtained based on the required waiting time slots in that channel, which are assumed as $\boldsymbol{D}_{r_1,k} = \{5, 1, 2, 3, 4\}$ and $\boldsymbol{D}_{r_2,k} = \{5, 3, 1, 2, 4\}$ in this example. Channel 5 is selected as the first entry in both of the lists since it is not occupied by any user at the time slot k. Note that the different channel priorities between $\boldsymbol{D}_{r_1,k}$ and $\boldsymbol{D}_{r_2,k}$ can happen due to their difference in channel observation and update based on the POMDP-based policy. According to the value of backoff counter, user r_1 will have higher priority than r_2 in choosing the destination channel for spectrum handoff, i.e. channel 5 will be selected since it is in the first entry of its channel list $\boldsymbol{D}_{r_1,k}$. On the other hand, user r_2 will notice that r_1 has higher priority in selecting the destination channel and both users possess the same channel 5 as the first entry in the prioritized destination channel list. The second entry, i.e. channel 3, in $\boldsymbol{D}_{r_2,k}$ will be selected by r_2 for spectrum handoff. Therefore, user r_1 will utilize the empty channel number 5 and r_2 will handoff to channel 3 waiting the required time slots for r_3 before conducting packet transmission. The performance of proposed M-POSH protocol for multiple CR user scenarios will be evaluated via simulations in next section.

5.3 Analytical model of M-POSH protocol

In order to broadcast handoff messages during the time interval T_{hs}, all CR users have to contend for the access to the control channel. The contention behavior is analyzed based on [26] which studied the process of backoff operations with a Markov chain model. Let τ be the probability that the CR user transmits packet, and p be the probability that each packet collides. According to [26], the relationship between τ and p can be obtained as

$$\tau(p) = \frac{2(1-2p)}{(1-2p)(W+1) + p \cdot W(1-(2p)^m)} \tag{31}$$

$$p(\tau) = 1 - (1-\tau)^{n-1} \tag{32}$$

where $W = CW_{min}$ is the minimum contention window, m is the maximum backoff stage, and n is the number of contending CR users. By iteratively solving these two equations, the value of τ and p can be obtained. Let P_{tr} be the probability that there is at least one transmission in a slot time, and P_s be the probability for a successful transmission conditioned on the fact that at least one user transmits. These two parameters can be derived based on (31) and (32) as follows: $P_{tr} = 1 - (1-\tau)^{n\text{-}1}$ and $P_s = \frac{n\tau(1-\tau)^{n-1}}{P_{tr}}$. Therefore, the average time required for a successful handshake on the control channel, denoted as $\mathrm{E}[T_{ave}]$, can be acquired as

$$E[T_{ave}] = T_s + \sigma\frac{1-P_{tr}}{P_{tr}P_s} + T_c\left(\frac{1}{P_s} - 1\right) \tag{33}$$

where σ is the slot time size. The parameters T_s and T_c are the time required for a successful transmission and the time duration while a collision happened, respectively, which can be obtained as

$$\begin{aligned} T_s = {} & T_{RTS} + T_{SIFS} + \delta + T_{CTS} + T_{SIFS} + \delta + E[P] + T_{SIFS} \\ & + \delta + T_{ACK} + T_{DIFS} + \delta \end{aligned} \tag{34}$$

$$T_c = T_{RTS} + T_{DIFS} + \delta \tag{35}$$

where δ is the propagation delay and $E[P]$ is the average packet payload size. The other parameters in (34) and (35) are the time durations denoted by their corresponding subscripts, e.g. T_{SIFS} represents the time interval for short interframe space (SIFS). Let n_{max} be the maximum number of CR users that complete the handshaking in the duration of T_{hs}, which can be obtained as $n_{max} = T_{hs}/E[T_{ave}]$ in average. For ease of analysis, it is assumed that all the CR users who need to utilize the control channel can successfully exchange the handshaking packets in the duration of T_{hs}, which corresponds to the result that $n < n_{max}$. Furthermore, the number of CR users n is assumed smaller than or equal to the number of total licensed spectrums in a slotted CR network in this analysis, i.e., $n \leq N$, which will be explained in the next paragraph. Note that the operation of proposed M-POSH protocol is not limited to these two assumptions.

The channel selection behavior of each CR user in the M-POSH scheme is based on the proposed POSH protocol for a single CR user. The case for the number of CR users to be equal to the number of total licensed spectrums, i.e. $n = N$, will first be considered in the analysis. If all the network channels are occupied by the CR users, all the users will decide to stay at the original channel in order to reduce potential collision due to spectrum handoff, which is denoted by the first case in (21) as $n_{w,nsh}^{a_{d1}} \leq n_w^{a_{d2}}$. The expected waiting time by staying at the current channel will be comparably smaller than or equal to that from the spectrum handoff scheme. Therefore, as $n = N$, we can acquire the expected waiting time for the M-POSH scheme to be equal to that for the NSH scheme in (16), i.e. $n_{w,mposh} = n_{w,nsh}$.

Furthermore, consider the case of $n = N - 1$, the advantage of the M-POSH protocol is to distribute the $N - 1$ CR users to exploit the one more available channel.

It indicates that there is one more chance for the CR user to switch to the other channel if $n_{w,nsh}^{a_{d1}} > n_{w}^{a_{d2}}$. According to (22), the expected waiting time for the M-POSH scheme can be obtained as $n_{w,mposh} = \frac{1}{2} n_{w,nsh} + \frac{1}{2} n_{w,posh}$ if $n = N - 1$. Consider the cases with $n = N - 2, N - 3, \ldots, 1$, one more channel will become available for the CR users to execute spectrum handoff with one less CR user in the network. The expected waiting time will become $\frac{1}{2} n_{w,mposh} + \frac{1}{2} n_{w,posh}$ every time the number of CR users is decremented by one. Consequently, the expected waiting time by adopting the M-POSH protocol $n_{w,mposh}$ can be obtained as

$$n_{w,mposh} = n_{w,nsh} \frac{1}{2^{N-n}} + n_{w,posh} \sum_{\ell=1}^{N-n} \frac{1}{2^{\ell}} \qquad (36)$$

for all $n \leq N$. The analytical result obtained from (36) will be validated via simulations in the next section.

6 Performance evaluation

In this section, simulations are presented to demonstrate the performance of proposed POSH and M-POSH protocols based on a C/C++ simulator. The network scenario consists of one primary user in each of the N channels. The traffic of primary user follows the Poisson distribution with packet arrival rate λ, and the service time is assumed to be a uniform distribution with mean 1/μ = 1. The major focus in the simulations is to obtain the required waiting time and the net transmission time for the secondary user under the occurrence of spectrum handoff. For the POSH protocol, one single CR user is considered to conduct spectrum handoff within the N channels in the network. On the other hand, the M-POSH scheme considers n multiple CR users contending the channel and perform spectrum handoff in the N channels. The parameters for channel contention are listed in Table 1. Since full channel state information is required by all of the existing spectrum handoff algorithms, it is considered unfair to compare the existing schemes within the environment adopted by the proposed POSH and M-POSH algorithms, where only partial channel information is observable. Therefore, the proposed scheme will be compared with two different cases as mentioned in Sect. 4, including both the NSH scheme and the RCS mechanism.

Table 1 Parameters

Parameter	Value
PHY header	192 bits
MAC header	224 bits
T_{RTS}	(160 bits + PHY header)/R
T_{CTS}	(112 bits + PHY header)/R
T_{ACK}	(112 bits + PHY header)/R
σ	20 μs
δ	1 μs
SIFS	16 μs
DIFS	56 μs
CW_{min}	32
Maximum backoff stage	5
channel bit rate (R)	1 Mbps

6.1 Scenarios with single CR user

6.1.1 *Parameter convergence and model validation*

The total number of time slots ℓ required for the computation of expected waiting time as obtained from (15) will be evaluated. Three channels with channel states $s_k = [c_{1,k}, c_{2,k}, c_{3,k}]$ are considered in the simulations. Figure 3 shows the variance of value function acquired by (14) versus the total number of required time slots ℓ under different transition probabilities from busy to idle in (15) as $\tau(1, a_k, 0) = 0.3, 0.5,$ and 0.7.

The purpose of this figure is to observe the convergence of ℓ within finite value, which is original infinite as shown in (15) to achieve the expected waiting time. It can be seen from Fig. 3 that the required time slots can be limited within 20 slots under different values transition probability $\tau(1, a_k, 0)$. The computational complexity associated with (15) for the proposed POSH scheme can therefore be feasibly confined.

The analytical models for required waiting time slots of the three schemes, including NSH, RCS, and POSH algorithms, as presented in (16), (18), and (22) are validated via simulations as shown in Fig. 4. Noted that the results under

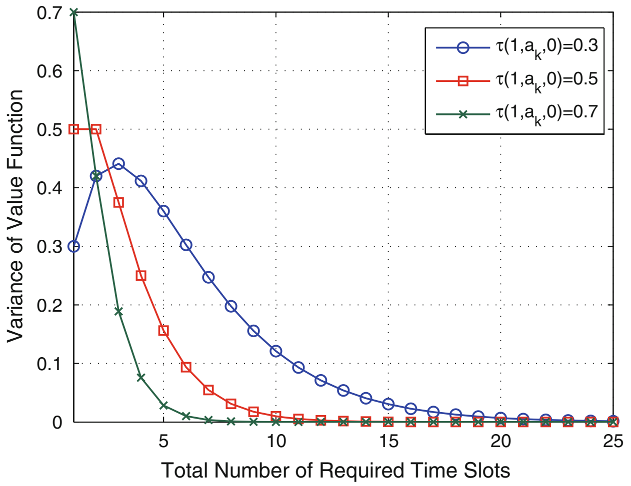

Fig. 3 Parameter convergence: variance of value function versus total number of required time slots ℓ under different transition probabilities $\tau(1, a_k, 0) = 0.3,\ 0.5,$ and 0.7

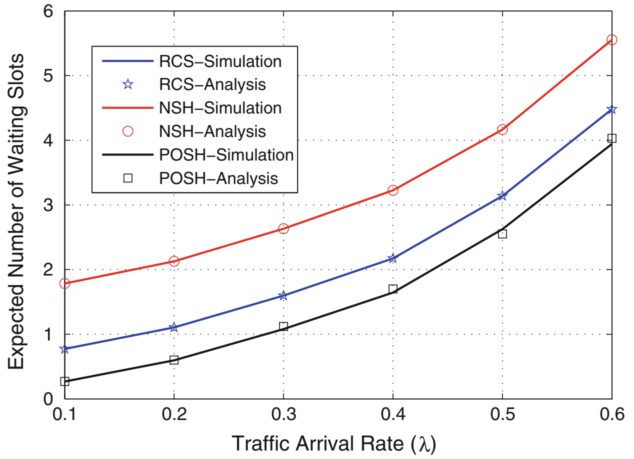

Fig. 4 Performance validation: the expected number of waiting time slots for a single spectrum handoff versus traffic arrival rate of primary user

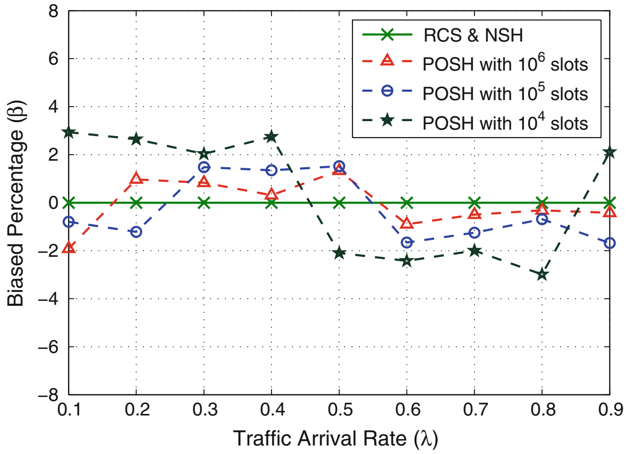

Fig. 5 Performance validation: biased percentage β versus traffic arrival rate of primary user

traffic arrival rate λ from 0.1 to 0.6 is illustrated since the contribution of proposed POSH scheme can be revealed under low and moderate traffic arrival rates. In the case that with larger λ, there is not much chance for the CR users to transmit date no matter which scheme is utilized, i.e. all the curves from the three schemes will be converged. It can be observed that the expected number of waiting time slots for a single spectrum handoff from all these three schemes increase as the traffic arrival rate (λ) of the primary user is augmented. Under different arrival rates, the proposed POSH algorithm can provide the smallest waiting time slots comparing with the other two schemes. Furthermore, it can also be seen that the simulation results of both NSH and RCS schemes match with their corresponding analytical results. On the other hand, there exists slight difference between the analytical and simulation results of the proposed POSH scheme. The major reason for this deviation can be contributed to the definition of non-uniform probability $\tilde{P}_r(a^* = a_s)$ in (21) as described in its analytical model, which is designed to be a stationary probability. However, the POSH scheme involves non-stationary updating process of the belief state for the estimation of channel state. The deviation resulting from the stationary model of POSH scheme will further be studied as illustrated in Fig. 5.

In order to clearly illustrate the difference between the analytical and simulation results, the biased percentage β is introduced and is defined as $\beta = \frac{n^a_{w,\zeta} - n^s_{w,\zeta}}{n^s_{w,\zeta}} \times 100$ where $n^a_{w,\zeta}$ and $n^s_{w,\zeta}$ correspond to the expected waiting time slots obtained from analysis and simulation respectively. The parameter ζ indicates either one of the three spectrum handoff schemes is adopted. Figure 5 illustrates the biased percentage β for the three schemes under different traffic arrival rates of the primary user. It is noted that the proposed POSH method is implemented under the simulation runs with different numbers of transmission time slots, i.e. $T = 10^4$, 10^5, and 10^6. It can be observed that even the non-stationary behavior of belief state cannot be exactly modeled and analyzed, the analytical results derived by stationary probability can still approach to the simulation values within 3 % of estimation difference. It can also be seen from Fig. 5 that the bias will be diminished as the number of transmission time slots is increased for the simulations. This results reveal the case that with longer transmission time slots, the incomplete and non-stationary network information can be updated more accurately. The simulation results will tend to possess stationary behaviors as is presented by the analytical model in (22).

6.1.2 Performance comparison

Figure 6 shows the performance comparison of the number of waiting time slots versus the total number of spectrum handoff for both the proposed POSH scheme and the RCS method. Two different channel conditions are considered for comparison purpose as follows. A better channel condition C_b is chosen with the transition probability from idle to idle state for each channel as $\tau(c_{i,k} = 0, a_k, c_{i,k+1} = 0) = 0.8, 0.7, 0.65$ for $i = 1, 2, 3$; while that from busy to idle state is selected as $\tau(c_{i,k} = 1, a_k, c_{i,k+1} = 0) = 0.4, 0.5, 0.55$ for $i = 1, 2, 3$. On the other hand, a worse channel condition C_w is determined with the transition probability from idle to idle state as $\tau(c_{i,k} = 0, a_k, c_{i,k+1} = 0) = 0.4, 0.3, 0.35$, and that from busy to idle state is set as $\tau(c_{i,k} = 1, a_k, c_{i,k+1} = 0) = 0.1, 0.2, 0.15$. For fair comparison, the NSH scheme is not implemented in this case since the CR user can always stay at the channel with better

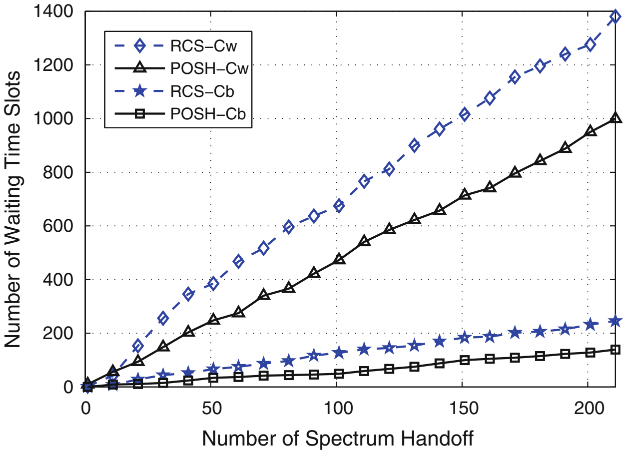

Fig. 6 Performance comparison: number of waiting time slots versus number of spectrum handoff

condition. It is intuitive to observe from Fig. 6 that the total number of waiting time slots is increased as the number of spectrum handoff is augmented. Furthermore, the secondary user has to wait for comparably more time slots in the worse channel case by adopting both schemes. Nevertheless, the total waiting time slots acquired from the proposed POSH protocol is smaller than that from the RCS scheme under both channel conditions. It is also observed that the POSH protocol can provide better performance as the number of spectrum handoff is increased. The reason can be contributed to the situation that more updated belief states are acquired by the POSH scheme as the number of handoff is augmented.

Figure 7 illustrates the performance comparison between the number of waiting time slots and the total number of transmission time slots. It is noticed that different numbers of waiting time slots and handoff numbers will be resulted by each scheme at every specific number of transmission time slots. In other words, the combining effects from both the waiting time slots and the handoff numbers will be revealed in Fig. 7 at each horizontal data point. It can be observed that the proposed POSH algorithm still outperforms the RCS scheme under both the C_b and C_w channel conditions. Even though the effect from the total number of spectrum handoff has not been considered in the value function as in (13), the POSH scheme can still provide smaller waiting time comparing with the RCS method.

Figures 8 and 9 illustrate the performance comparison among the POSH, the RCS, and the NSH schemes under different values of packet arrival rate λ of the primary user. Comparing the case in Fig. 4 with a single spectrum

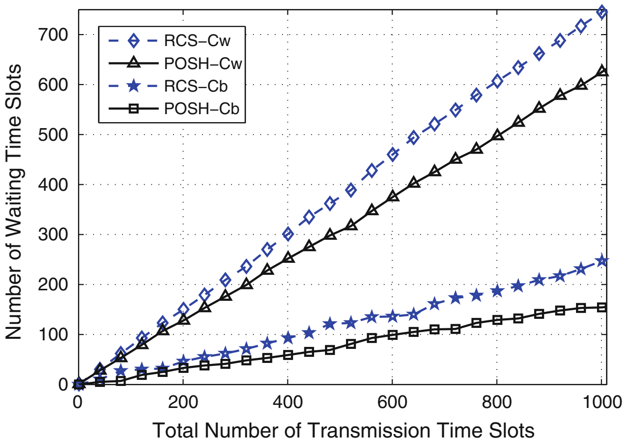

Fig. 7 Performance comparison: number of waiting time slots versus total number of transmission time slots

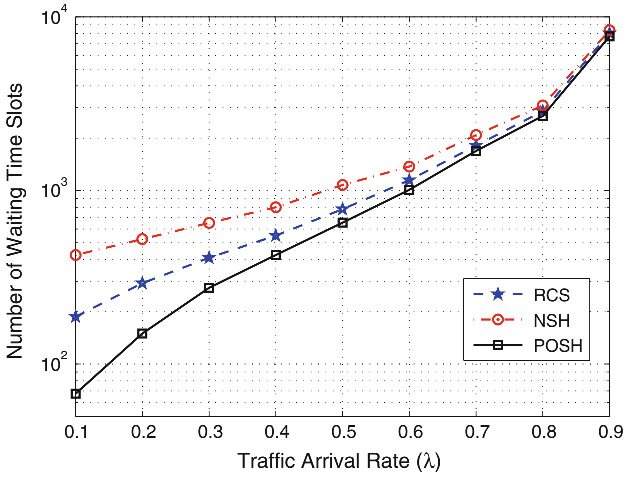

Fig. 8 Performance comparison: number of waiting time slots versus traffic arrival rate of primary user under numbers of spectrum handoff = 250

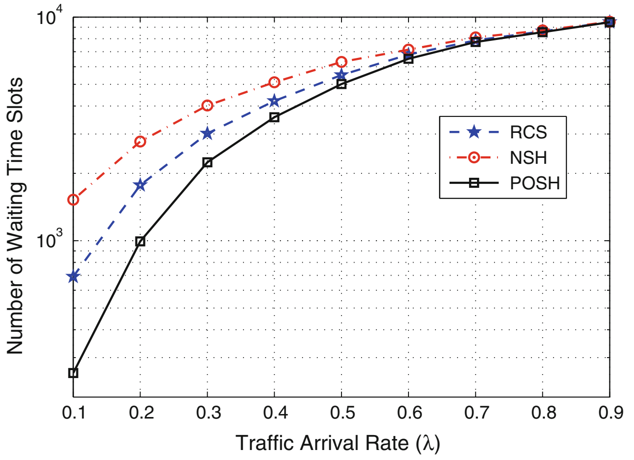

Fig. 9 Performance comparison: number of waiting time slots versus traffic arrival rate of primary user under number of transmission time slots = 1,200

handoff, Fig. 8 shows the averaged performance comparison under a larger fixed number of spectrum handoff equal to 250. On the other hand, Fig. 9 illustrates the comparison between these three protocols under the number of transmission time slots equal to 1,200. It can be observed from both figures that the proposed POSH scheme can outperform the other two methods under different packet arrival rates. The benefits from the adoption of proposed POSH algorithm is especially revealed at smaller packet arrival rates since there can be more opportunity for the POSH scheme to select a feasible target channel to conduct spectrum handoff.

Practical consideration for the three handoff schemes are presented as shown in Figs. 10 and 11. The corresponding parameters are listed as follows: $T_{slot} = 100$ ms, $T_{sens} = 5$ ms, $T_{sw} = 10$ ms, $T_{hs} = 1$ ms, and $T_{ack} = 1$ ms. Figure 10 illustrates the waiting time obtained from (23), (25), and (27) for the NSH, RCS, and POSH scheme respectively. Two cases with required computation time $T_p = 0$ and 5 ms for implementing the POMDP-based policy are considered; while the update time for belief state $T_u = 3$ ms. With practical consideration, the POSH algorithm is demonstrated to result in the smallest waiting time under different packet arrival rate of primary user. Even though the calculation of POMDP-based policy requires additional computation time, the overall performance is still better in comparison with the NSH and RCS methods. Furthermore, it is observed in Fig. 10 that the waiting time $T_{w,posh}$ will only be affected by the computation time T_p if the destination channel of spectrum handoff has high probability to be in the idle state. In other words, T_p will only degrade the performance of proposed POSH scheme at lower packet arrival rate of primary user, and will become uninfluential under higher primary traffic rate. Nevertheless, since the POSH scheme can provide better performance among the three handoff schemes at lower primary traffic (as shown in Figs. 8 and 9), it can be concluded that the degraded effect from T_p will not be significant by adopting the POSH scheme.

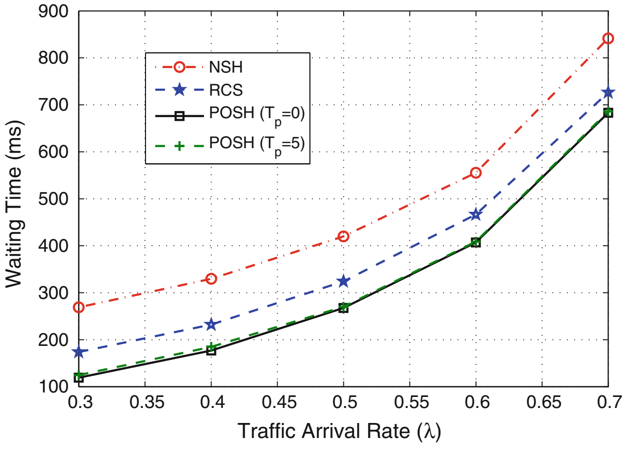

Fig. 10 Performance comparison with practical consideration: waiting time versus traffic arrival rate of primary user

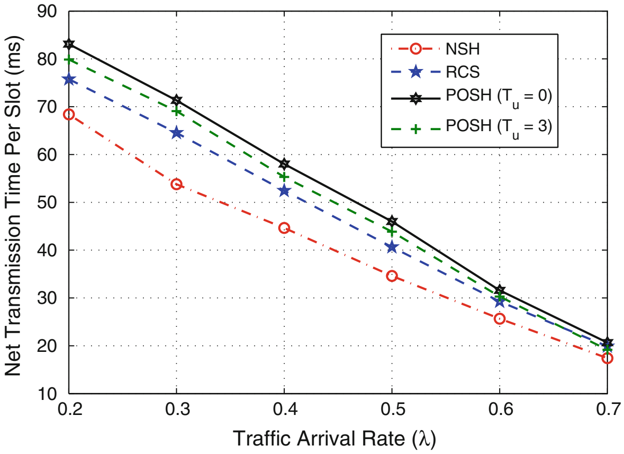

Fig. 11 Performance comparison with practical consideration: net transmission time per slot versus traffic arrival rate of primary user

In order to better present the utilization of licensed spectrum, Fig. 11 is exploited to illustrate the net transmission time per time slot as was computed from (24), (26), and (28) for the NSH, RCS, and POSH scheme respectively. Two cases with the update time of belief state $T_p = 0$ and 3 ms for the proposed POSH scheme are considered; while the required time for POMDP-based policy $T_p = 5$ ms. It is noticed that the update time T_u is required for the proposed POSH scheme at the end of each time slot; while the computation time for POMDP-based policy T_p will only be utilized as spectrum handoff occurs. It is intuitive to observe that the net transmission time resulted from all the three schemes decreases as the packet arrival rate λ of primary traffic is increased. Furthermore, as λ is incremented, the net transmission time of proposed POSH protocol can become similar to or even worse than that obtained from the RCS and NSH schemes, e.g. the net transmission time acquired from the POSH scheme with $T_u = 3$ ms is worse than that from the RCS protocol as $\lambda > 0.65$. Therefore, practical consideration for these handoff schemes can provide a channel selection criterion for the CR users to determine the feasible handoff scheme to be applied in order to obtain their destination channel for spectrum handoff.

6.2 Scenarios with multiple CR users

6.2.1 Model validation

Figure 12 illustrates the results for validating the analytical model of the proposed M-POSH scheme in (36) via

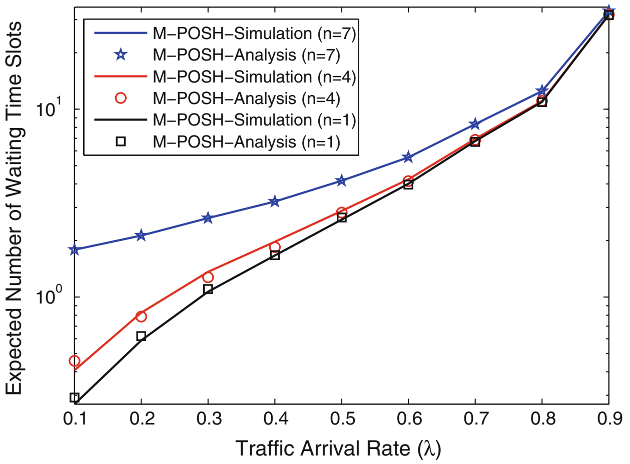

Fig. 12 Performance validation: the expected number of waiting time slots versus traffic arrival rate of primary user

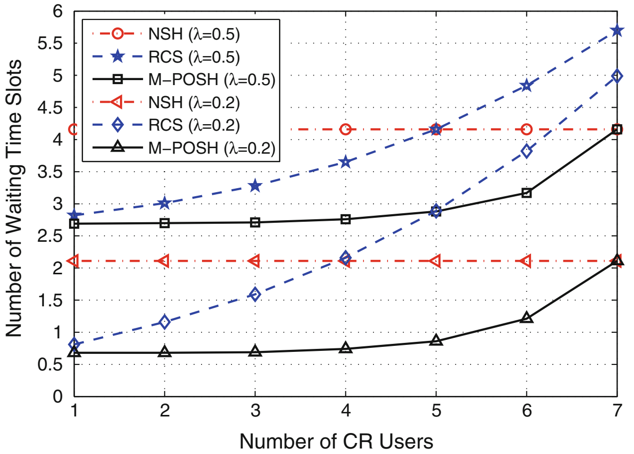

Fig. 14 Performance comparison: number of waiting time slots versus number of CR users (with 7 channels)

simulations. The performance validation is conducted by observing the expected waiting time slots versus the traffic arrival rate of primary user under different numbers of CR users, i.e. $n = 1, 4, 7$ with $N = 7$ channels. It can be observed that the proposed analytical model can effectively predict the behavior of the M-POSH scheme. It is apparently that the expected number of waiting time slots increases with the augmentation of the number of CR users and the traffic arrival rate of primary user. More CR users indicates that there are less available channels to be utilized while executing spectrum handoff, which requires more waiting time slots.

6.2.2 *Performance comparison*

Figures 13 and 14 compare the performance of NSH, RCS, and M-POSH protocols under the circumstance of multiple CR users within multi-channels network. Figure 13 shows the performance comparison with two CR users under different numbers of available channels; while Fig. 14 illustrates the comparison with 7 channels under different numbers of CR users. Two different arrival rates of primary traffic are considered for all the three schemes, i.e. $\lambda = 0.2$ and 0.5. It is apparently to observe that the NSH scheme results in the same performance under different numbers of channels and CR users since it does not perform any spectrum handoff activity. In other words, the NSH scheme in multi-user scenario can be regarded as the combination of single user without mutual interactions. On the other hand, regarding the RCS and M-POSH schemes, the expected number of waiting time slots is decreased as the number of available channels is augmented; while it is increased with the augmentation of total number of CR users due to potential packet collisions happened in the network. The performance of RCS scheme can becomes worse than that from the NSH method under certain circumstances since the NSH scheme at least guarantees that packet collisions will not happen in the network. From example as in Fig. 14, the RCS scheme with $\lambda = 5$ results in higher number of waiting time slots comparing with the NSH method if the number of CR users is greater than 5.

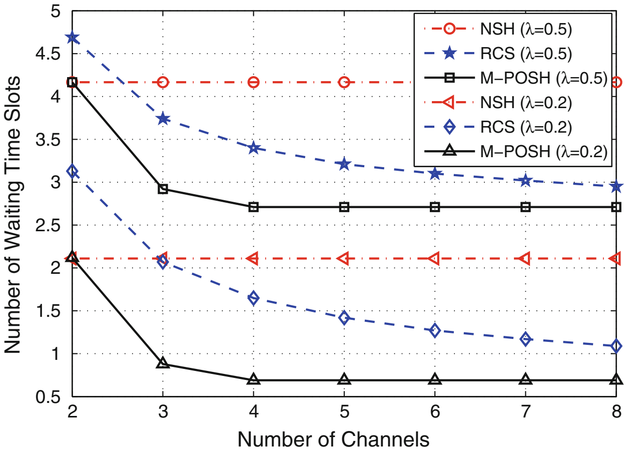

Fig. 13 Performance comparison: number of waiting time slots versus number of channels (with 2 CR users)

Nevertheless, with the consideration of potential packet collisions between the CR users, the proposed M-POSH protocol can adaptively select the destination channel based on the availability of network channels. As shown in both Figs. 13 and 14, the M-POSH scheme outperforms both the RCS and NSH algorithms under different numbers of channels and CR users. It is also observed from certain end data points of both figures that the M-POSH and the NSH schemes will result in the same number of waiting time slots, i.e. the most left data points in Fig. 13 under number of channels = 2 with two CR users, and the most right data points in Fig. 14 under number of CR users $= 7$ with 7 channels. This indicates that if all the network channels are occupied, the CR users will decide to stay at the current channel by adopting the proposed M-POSH protocol in order to reduce the probability of packet collision between the CR users. On the other hand, as long as the network contains more available channels that can be utilized, the M-POSH protocol will distribute the CR users that intend to conduct handoff to exploit those available channels. Consequently, the performance of M-POSH scheme can achieve the optimal performance as that obtained from the POSH scheme for single user case. The merits of the proposed POSH and M-POSH protocols can therefore be observed.

7 Conclusion

This paper proposes a spectrum handoff strategy based on partially observable Markov decision process (POMDP) in the overlay cognitive radio (CR) networks. With only partially observable state information, the proposed POMDP-based spectrum handoff (POSH) scheme selects the optimal destination channel in order to achieve the minimal waiting time for packet transmission. Furthermore, in order to consider the network with multiple CR users, the multi-user POSH (M-POSH) protocol is proposed to resolve the packet collision problem among multiple CR users that intend to conduct spectrum handoff. It is observed from both the simulation and analytical results that both the proposed POSH and M-POSH protocols can effectively reduce the waiting time of spectrum handoff for a partially observable CR network.

References

1. F.C.C. (2006). Unlicensed operation in the TV broadcast bands. *ET Docket No. 04-186*, pp. 4–186.
2. F.C.C. (2002). FCC spectrum policy task force report. *ET Docket No. 02-135*.
3. Zhao, Q., Tong, L., Swam, A. & Chen Y. (2007). Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework. *IEEE Journal on Selected Areas Communications 25*, 589–600.
4. Bhargava, V.K. (2007). Research challenges in cognitive radio networks. *Proceeding of IEEE ICCCN* pp. 1–6.
5. Shanker, S. (2007). Squeezing the most out of cognitive radio: A joint MAC/PHY perspective. *Proceeding of IEEE ICASSP*
6. Sahai, A., Hoven, N., & Tandra, R. (2004) Some fundamental limits on cognitive radio. *Proceeding of Allerton Conference on Signals, Systems, and Computers*.
7. Sengupta, S., Brahma, S., Chatterjee, M.& Shankar, N.S. (2007). Enhancements to cognitive radio based IEEE 802.22 air-interface. Proceeding of IEEE ICC.
8. Cordeiro, C. Challapali, K. Birru, D. & Shankar, N.S. (2005) IEEE 802.22: The first worldwide wireless standard based on cognitive radios. *Proceding on IEEE DySPAN*, pp. 328–337.
9. Cassandra, A.R. (1994). Exact and approximate algorithms for partially observable Markov decision processes. (Ph.D. dissertation, Brown University, 1994).
10. Ma, R.-T., Hsu, Y.-P. & Feng, K.-T. (2009). A POMDP-based spectrum handoff protocol for partially observable cognitive radio networks. *Proceeding of IEEE WCNC*, pp. 1–6.
11. Wang, L.C. & Chen A. (2008). On the performance of spectrum handoff for link maintenance in cognitive radio. *Proceeding of IEEE ISWPC*, pp. 670–674.
12. Zhu, X., Shen, L. & Yum, T.S.P. (2007). Analysis of cognitive radio spectrum access with optimal channel reservation. *IEEE Communications Letters 11*, 304–306.
13. Wang, L.C. & Wang C.W. (2009). Modeling and analysis for proactive-decision spectrum handoff in cognitive radio networks. *Proceeding of IEEE ICC*
14. Wang L.C. & Wang C.W. (2008). Spectrum handoff for cognitive radio networks: Reactive-sensing or proactive-sensing?. *Proceeding of IEEE IPCCC*
15. Long, X., Gan, X., Xu, Y., Liu, J. & Tao, M. (2008). An estimation algorithm of channel state transition probabilities for cognitive radio systems. *Proceeding of IEEE CROWN COM*.
16. Gong, S. Wang, P. Liu, W. & Yuan, W. (2010). Maximize secondary user throughput via optimal sensing in multi-channel cognitive radio networks. *Proceeding of IEEE GLOBECOM*, pp. 1–5.
17. Kim, I. & Kim, D. (2010). Optimal allocation of sensing time between two primary channels in cognitive radio networks. *IEEE Communications Letters*, *14*(4), 297–299.
18. Zhao Q., Krishnamachari B. & Liu K. (2008). On myopic sensing for multi-channel opportunistic access: Structure, optimality, and performance. *IEEE Transactions on Wireless Communications*, *7*(12), 5431–5440.
19. Kaelbling L.P., Littman M.L., Cassandra A.R. (1998) Planning and acting in partially observable stochastic domains. Artificial Intelligence 101:99–134
20. Astrom K.J. (1965). Optimal control of markov decision processes with incomplete state estimation. *Journal of Mathematical Analysis and Applications*, *10*, 174–205.
21. Cheng H.-T. (1988). Algorithms for partially observable Markov decision processes. (Ph.D. dissertation, University of British Columbia, British Columbia, 1988).
22. Tang, P.K., Chew, Y.H. & Ong, L.C. (2009). On the distribution of opportunity time for the secondary usage of spectrum. *IEEE Transactions on Vehicular Technology 58*, 1517–1527.
23. Zhang Y. (2008) Dynamic spectrum access in cognitive radio wireless networks. *Proceeding of IEEE ICC*.
24. Zhao, Q., Tong, L. & Swami, A. (2005). Decentralized cognitive MAC for dynamic spectrum access. *Proceeding of IEEE DY-SPAN*, pp. 224–232.
25. Wang, L.C., Lu, Y.C., & Wang, C.W. (2007). Latency analysis for dynamic spectrum access in cognitive radio: Dedicated or embedded control channel?. *Proceeding of IEEE PIMRC*
26. Bianchi, G. (2000). Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications, 18*(3), 535–547.

Author Biographies

Jui-Hung Chu received his B.S. degree in 2009 from the Department of Communication Engineering, National Chiao Tung University, Hsinchu, Taiwan. Since 2009, he has been pursuing his Ph.D. degree in the Institute of Communication Engineering, National Chiao Tung University, where he is currently a Ph.D. candidate. His research interests include MAC protocol design, vehicular communications, and cognitive networks.

Rui-Ting Ma received his B.S. and M.S. degrees from the department of Communication Engineering, National Chiao Tung University, Hsinchu, Taiwan, in 2007 and 2009, respectively. Since 2009, he has been with the HTC Corporation, Taiwan, as a system engineer. His research interests include MAC protocol design, wireless local area networks, and Cognitive Radio networks.

Kai-Ten Feng received the B.S. degree from National Taiwan University, Taipei, in 1992, the M.S. degree from the University of Michigan, Ann Arbor, in 1996, and the Ph.D. degree from the University of California, Berkeley, in 2000. Since August 2007, he has been with the Department of Electrical Engineering, National Chiao Tung University, Hsinchu, Taiwan, as an associate professor. He was an assistant professor with the same department between February 2003 and July 2007. He was with the OnStar Corp., a subsidiary of General Motors Corporation, as an in-vehicle development manager/ senior technologist between 2000 and 2003, working on the design of future Telematics platforms and the in-vehicle networks. His current research interests include cooperative and cognitive networks, mobile ad hoc and sensor networks, embedded system design, wireless location technologies, and Intelligent Transportation Systems (ITSs). He received the Best Paper Award from the IEEE Vehicular Technology Conference Spring 2006, which ranked his paper first among the 615 accepted papers. He is also the recipient of the Outstanding Young Electrical Engineer Award in 2007 from the Chinese Institute of Electrical Engineering (CIEE). He has served on the technical program committees of VTC, ICC, and APWCS.