國 立 交 通 大 學

電信工程學系

碩 士 論 文

在感知性網路中以部分探測馬可夫決策過程為基礎的
頻帶換手機制

A POMDP-based Spectrum Handoff Protocol for
Partially Observable Cognitive Radio Networks

研究生：馬瑞廷

指導教授：方凱田 教授

中 華 民 國 九十八 年 六 月

在感知性網路中以部分探測馬可夫決策過程為基礎的頻帶換手機制

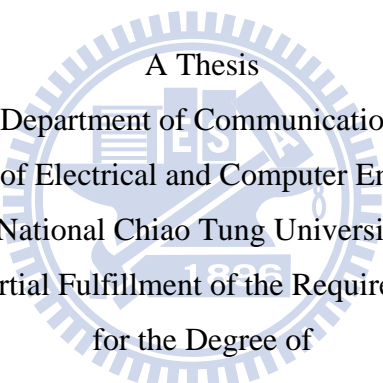A POMDP-based Spectrum Handoff Protocol for Partially Observable Cognitive Radio Networks

研 究 生：馬瑞廷　　　　Student：Rui-Ting Ma

指導教授：方凱田　　　　Advisor：Kai-Ten Feng

國 立 交 通 大 學
電 信 工 程 學 系
碩 士 論 文

A Thesis
Submitted to Department of Communications Engineering
College of Electrical and Computer Engineering
National Chiao Tung University
in partial Fulfillment of the Requirements
for the Degree of
Master of Science
in

Communications Engineering

June 2009

Hsinchu, Taiwan, Republic of China

中華民國九十八年六月

在感知性網路中以部分探測馬可夫決策過程為基礎的頻帶換手機制

學生：馬瑞廷　　　　　　　　　　　　指導教授：方凱田 教授

國立交通大學電信工程學系碩士班

摘　　　要

　　近年的研究說明了靜態的頻帶分配是造成頻帶使用缺少效率的主因，為了增進頻帶使用率，可動態偵測且使用認證頻帶的感知無線電(CR)因應而生。 如何提供有效率的頻帶換手在 CR 中是個很重要的議題。現存的頻帶換手方法假定感知無線電的使用者(CR user)可以正確的偵測每一個頻帶以便找到適合的頻帶進行換手。然而，這個假設在實際的情況下是不實際的，因為 CR 使用者偵測頻帶所花費的時間將會太高而影響主要使用者的品質。 在這篇論文中，藉由部分可知的環境下的 Marcov 決策過程(POMDP)的幫助，可以透過探測部分的頻帶來估測整個網路環境 。此篇論文提出以 POMDP 為基準的頻帶換手機制(POSH)，其目的為藉由部分的通道狀態來找出最適合進行換手的頻帶。除此之外，為了適應多位使用者的環境，此篇論文本著將頻寬資源充分分配給各使用者的概念提出了一個以 POMDP 為基準的多使用者的換手機制(M-POSH)。 藉由 POMDP 為基準的頻帶換手機制所選出的頻帶，可達到在每次換手時 CR 使用者所需等待的時間最短，數據結果顯示出此方法可有效率地讓 CR 使用者在每次頻帶換手時達到最少的等待時間。

A POMDP-based Spectrum Handoff Protocol for Partially Observable Cognitive Radio Networks

student：Rui-Ting Ma   Advisors：Dr. Kai-Ten Feng

Department of Communication Engineering
National Chiao Tung University

ABSTRACT

Recent studies have been conducted to indicate the ineffective usage of licensed bands due to the static spectrum allocation. In order to improve the spectrum utilization, the 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 within the network can be correctly sensed by the CR users in order to perform appropriate spectrum handoff process. However, this assumption is considered impracticable in realistic circumstances primarily due to the excessive time required for the CR user to sense the entire spectrum space. In this paper, the partially observable Markov decision process (POMDP) is exploited 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 adapt the POMDP policy into multi-user CR network by distributing CR users to opportunistic frequency spectrums. 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. Numerical results illustrate that the proposed spectrum handoff protocols can effectively minimize the required waiting time for spectrum handoff in the CR networks.

最後要感謝我的家人，他們給我毫無保留的支持，讓我能在外地生活而無後顧之憂。有了他們的支持，我才能經歷這許多美好的一切。套句人家說的話，我們家雖不富有，但他們願意給我最好的。

研究成果可以用這本論文記載，這兩年經歷獲得的助益怎樣也訴說不盡。感謝這一路走來幫助過我的朋友們，有了你們才造就了今天的我。雖然我無法在這一一表達所有朋友曾經帶給我的幫助，但我將懷抱著這份感激，邁向我下一段的旅程。
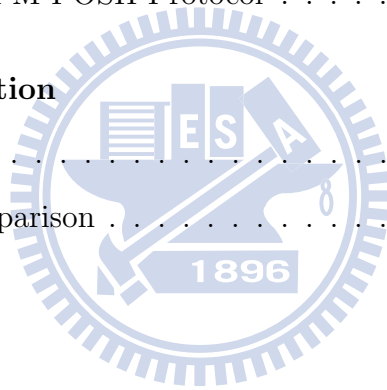

馬瑞廷謹誌　于新竹國立交通大學

Contents

List of Figures

Chapter 1

Introduction

According to the research conducted by FCC [1; 2], a large portion of the priced frequency spectrums 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 the study of 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], considered as a realistic implementation of the CR concept, is an emerging standard that allocates spectrums for TV broadcast services via a license-exempt basis. 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 maintain its data transmission. Consequently, the main objective for spectrum handoff is to select a feasible target channel such that the CR user can be switched into such as to retain its on-going transmission. The performance of spectrum handoff will primarily be dominated by the feasibility of conducting channel selection.

Target channel selection can be categorized into two different types of schemes according to their sensing strategies [9], 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 to a channel which was determined in sequence from 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 situations, the pre-determined channel list can become infeasible to be adopted in target channel selection for spectrum handoff. The channel reservation scheme as proposed in [10] 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 [11; 12] to illustrate the beneficial aspects of the pre-sensing strategy. However, those reserved idle channels can not be ensured available at the time for spectrum handoff. As a result, the performance of pre-sensing strategies can not be guaranteed especially under the fast fading channel environments.

On the other hand, the post-sensing techniques 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 becomes its target channel. Compared to the pre-sensing approaches, a more feasible and accurate channel can be selected by exploiting the post-sensing schemes since the target channel is determined at the time while the secondary user is interrupted. Nevertheless, the post-sensing methods in general spend 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 considered impracticable in realistic environments. In other words, the transition probabilities of all the channel states are not always served as the available information to the CR users within the network. An estimation algorithm has been proposed in [13] to estimate the transition probability by adopting the maximum likelihood function. However, the converging speed for the estimation can become intolerable while small value of transition probability has been occurred.

In this paper, without the necessity of obtaining all the correct channel information, the partially observable Markov decision process (POMDP) [14–17] is utilized to reveal the network information by partially sensing the 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 in this paper 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 method and other existing channel selection schemes are derived, validated, and compared via simulations. Furthermore, one may doubt even though the proposed POSH scheme can effectively reduce the number of waiting time slots, the complicated calculation process may bring out excessive overhead which will actually degrade its performance. Therefore, the required phases for precise channel selection behavior is discussed and compared under realistic environment. The computation complexity of the precise POSH scheme associated with the update of belief state update are also presented. It

is demonstrated in the simulation results that reduced waiting time can be obtained from the proposed POSH mechanism under the consideration of precise behaviors.

Furthermore, multiple 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. Therefore, it is crucial to provide channel selection policy by considering information from other CR users, and also to alleviate the overhead while multiple CR users are contending to access 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 the observe the traffic of primary user, that from the CR users is required to be monitored by the M-POSH scheme. A negotiation procedure is employed in order to facilitate the determination of a feasible CR user to operate within a specific target channel. In comparison with other existing algorithm, the proposed M-POSH scheme can effectively minimize the waiting time for spectrum handoff.

The rest of this paper is organized as follows. Chapter 2 briefly summarizes the concept of the POMDP approach. The proposed POSH scheme is modeled and derived in Chapter 3. Performance analysis of the POSH protocol is conducted in Chapter 4, while precise schemes of channel selection is presented in Chapter 4.2. Considering the network scenario with multi-users, the M-POSH protocol is proposed in Chapter 5. Chapter 6 illustrates the performance evaluation for both the proposed POSH and M-POSH mechanisms; while the conclusions are drawn in Chapter 7.

Chapter 2

POMDP framework

A Markov decision process (MDP) refers to a discrete time stochastic control process that conducts decision-making based on the present state information, i.e. $s_k \in \boldsymbol{S}_k$. It is noted that the subscript k is utilized to denote the kth time slot in consideration; while $\boldsymbol{S}_k$ 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) [14] is utilized to determine the decision policy based on the partially available information and the observations from the environment. The schematic diagram of the POMDP framework is illustrated in Fig. 2.1. In general, optimization techniques are required to be exploited in order to obtain the solution for the POMDP-based problem.

2.1 Observation

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
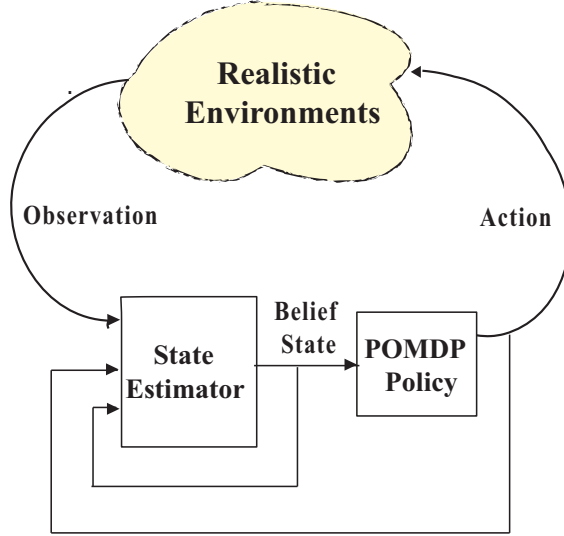
Figure 2.1: The schematic diagram of the POMDP framework.

possible observations z_k for each action a_k and resulting states s_{k+1}, i.e.

$$o(s_{k+1}, a_k, z_k) = P_r(z_k | a_k, s_{k+1}) \tag{2.1}$$

$\forall z_k \in \boldsymbol{Z}_k, a_k \in \boldsymbol{A}_k, s_{k+1} \in \boldsymbol{S}_{k+1}$ 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. Considering the MDP case, a policy is determined to map from the current state to the corresponding action since the present state of MDP is fully observable. On the other hand, the POMDP can only map from the latest observation to the corresponding action (as in Fig. 2.1), which is in general considered insufficient to represent the history of the process. Therefore, the belief state is utilized to reveal the statistic distribution of the current state information, which will be explained in the next section.

6

2.2 Belief State

The concept of 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 $\sum_{s_k \in \mathbf{S}_k} b(s_k) \leq 1$. 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 a more accurate information of the environment. As shown in the Fig. 2.1, the updated belief state is acquired as the outcome of the state estimator, which is consisted by the inputs of observation, action, and the former belief state. Consequently, the resulting belief state $b(s_{k+1})$ w.r.t. the state s_{k+1} can be obtained as

$$b(s_{k+1}) = P_r(s_{k+1}|b(\boldsymbol{S}_k), a_k, z_k) = \frac{o(s_{k+1}, a_k, z_k)\sum_{s_k \in \boldsymbol{S}_k}\Gamma(s_k, a_k, s_{k+1})b(s_k)}{P_r(z_k|b(\boldsymbol{S}_k), a_k)} \quad (2.2)$$

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. $\Gamma(s_k, a_k, s_{k+1}) = P_r(s_{k+1}|a_k, s_k)$. The denominator of (2.2) can be considered as a normalizing factor, which is obtained as

$$P_r(z_k|b(\boldsymbol{S}_k), a_k) = \sum_{s_k \in \boldsymbol{S}_k, s_{k+1} \in \mathbf{S}_{k+1}} o(s_{k+1}, a_k, z_k) \cdot \Gamma(s_k, a_k, s_{k+1})b(s_k) \quad (2.3)$$

7

2.3 Reward and Value Functions

In order to ensure optimal decision is made by adopting the POMDP, 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 to turn 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 reward $R(s_k, a_k)$ can be obtained as

$$R(s_k, a_k) = \sum_{s_{k+1} \in \boldsymbol{S}_{k+1},\, z_k \in \boldsymbol{Z}_k} \sum \Gamma(s_k, a_k, s_{k+1}) \cdot o(s_{k+1}, a_k, z_k) \cdot r(s_k, a_k, s_{k+1}, z_k) \qquad (2.4)$$

It is noticed that the immediate reward function 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 (2.4). However, certain period of time is in general considered for evaluating the value of the reward. Therefore, the decision policy by exploiting the POMDP is determined by optimizing the L-step value function with $L \geq 1$.

Chapter 3

Proposed POMDP-based Spectrum Handoff (POSH) Scheme

In this chapter, the POMDP framework will be utilized to model the spectrum handoff problem in a slotted overlay CR network. The proposed POSH scheme will be exploited under the single CR user scenario. The value function will consequently be formulated in order to obtain the optimal policy for spectrum handoff.

3.1 System Model for POSH Scheme

In the considered CR network, there are N channels that are available to be accessed by both the primary and the secondary users. Based on the secondary user's point of view, 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. Considering that $c_{i,k}$ denotes the state of the ith channel in time slot k (i.e. $c_{i,k} = 0$ indicates the idle state and $c_{i,k} = 1$ represents the busy state), the state of the entire network in the kth time slot

can be written as

$$s_k = [c_{1,k}, \ldots c_{i,k}, \ldots c_{N,k}], c_{i,k} \in \{0,1\}, \forall s_k \in \boldsymbol{S}_k \tag{3.1}$$

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 stand for the busy state.

Furthermore, the transition probability can be determined by modeling a channel as an M/G/1 system with arrival rate λ and service rate μ. By assuming poisson traffic for the arrival packets, the probability distribution of arriving packets can be represented as

$$P_r(n_{\lambda,k} = x) = \frac{e^{-\lambda}\lambda^x}{x!} \tag{3.2}$$

where $n_{\lambda,k}$ is denoted as 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 (3.2), the transition probability from idle to idle state for a channel $c_{i,k}$ (for $i = \{1, \ldots, N\}$) after executing action a_k can be acquired as

$$\tau(c_{i,k} = 0, a_k, c_{i,k+1} = 0) = \frac{P_r(n_{\lambda,k} = 0, c_{i,k} = 0)}{P_r(c_{i,k} = 0)} = \frac{P_r(n_{\lambda,k} = 0) \cdot P_r(c_{i,k} = 0)}{P_r(c_{i,k} = 0)} = e^{-\lambda} \tag{3.3}$$

It is noted that the second equality in (3.3) 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 represented as

$$\tau(c_{i,k} = 1, a_k, c_{i,k+1} = 0) = \frac{P_r(n_{\lambda,k} = 0, n_{\lambda,k-1} > 0, T_{s,k} \leq 1)}{P_r(c_{i,k} = 1)}$$
$$= \frac{P_r(n_{\lambda,k} = 0) \cdot P_r(n_{\lambda,k-1} > 0) \cdot P_r(T_{s,k} \leq 1)}{P_r(c_{i,k} = 1)} \quad (3.4)$$

where the second equality is also contributed to the independency of the three probabilities within the numerator of (3.4). 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 random variables of a packet service time with mean value of $1/\mu$, 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 take 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 (3.4) can be rewritten as

$$P_r(T_{s,k} \leq 1) = P_r(T_{s,k-1} \leq 1) \cdot P_r(T_{\alpha,k} \leq 1) + P_r(T_{s,k-1} > 1) \cdot P_r(T_{s,k-1} - 1 + T_{\alpha,k} \leq 1)$$
$$(3.5)$$

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 (3.4) that representing the probability for a busy channel can be expressed as

$$P_r(c_{i,k}=1)=P_r(n_{\lambda,k-1}>0)+P_r(n_{\lambda,k-1}=0)\cdot P_r(T_{s,k-1}>1) \tag{3.6}$$

Based on (3.3) and (3.4), 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

$$\Gamma(s_k, a_k, s_{k+1}) = \prod_{i=1}^{N} \tau(c_{i,k} = \varsigma_1, a_k, c_{i,k+1} = \varsigma_2) \tag{3.7}$$

where $c_{i,k} \in s_k$, $c_{i,k+1} \in s_{k+1}$, and $\varsigma_1, \varsigma_2 \in \{0, 1\}$, 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. Three factors are considered for the spectrum handoff time including the switching time, the handshaking time, and the waiting time. In general, both the switching and the handshaking time intervals are assumed fixed with comparably smaller values. Therefore, the main objective of the 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 in the case that the target channel is still occupied by the primary user.

The waiting time, which is served as the reward 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 reward is considered as the total number of waiting slots n_w required by the secondary user while executing a certain action
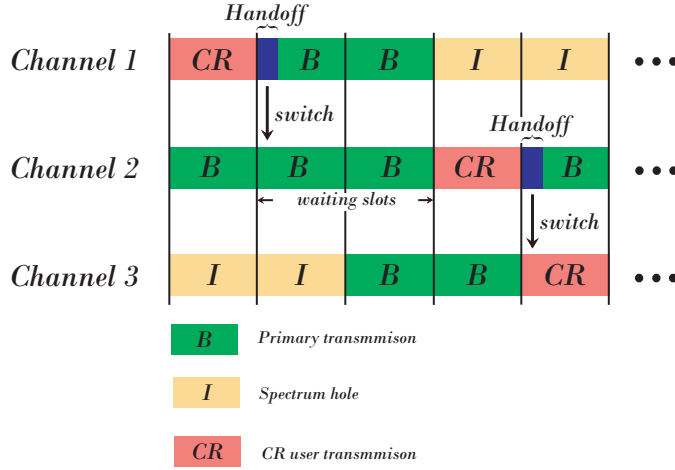
Figure 3.1: Performance comparison: the number of waiting time slots versus the number of spectrum handoff.

(i.e. $a^* = a_k$) for spectrum handoff. Consequently, the expected reward $R(s_k, a_k)$ can be written as

$$
\begin{aligned}
R(s_k, a_k) &= E[n_w = \ell | a^* = a_k] = \sum_{\ell=0}^{\infty} \ell \cdot P_r(n_w = \ell | a^* = a_k) = \sum_{\ell=1}^{\infty} \ell \cdot P_r\left(\bigcap_{p=1}^{\ell} c_{a_k, k+p-1} = 1, c_{a_k, k+\ell} = 0,\right) \\
&= \sum_{\ell=1}^{\infty} \ell \cdot \tau(c_{a_k, \ell+k-1} = 1, a_k, c_{a_k, \ell+k} = 0) \cdot P_r\left(\bigcap_{p=1}^{\ell} c_{a_k, k+p-1} = 1\right) \\
&= \sum_{\ell=1}^{\infty} \ell \cdot \tau(c_{a_k, \ell+k-1} = 1, a_k, c_{a_k, \ell+k} = 0) \cdot P_r(c_{a_k, k} = 1) \cdot \tau(c_{a_k, k} = 1, a_k, c_{a_k, k+1} = 1)^{\ell-1}
\end{aligned}
$$
(3.8)

$$
= \sum_{\ell=1}^{\infty} \ell \cdot c_{a_k, k} \cdot \tau(c_{a_k, \ell+k-1} = 1, a_k, c_{a_k, \ell+k} = 0) \cdot \tau(c_{a_k, k} = 1, a_k, c_{a_k, k+1} = 1)^{\ell-1} \tag{3.9}
$$

where $c_{a_k, k}$ denotes the channel state after selecting channel a_k at the time instant k. It is noted that the relationship from (3.8) to (3.9) is due to the fact that $P_r(c_{a_k, k} = 1) = c_{a_k, k}$ since a specific channel a_k is considered for the calculation of the waiting time duration.

13

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 allow to coexist with the primary user while the time duration for packet transmission is divided into time slots. As shown in Fig. 2.1, 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 the belief state $b(s_{k+1})$ as in (2.2). The secondary user utilizes the updated belief state in order to estimate the channel state of the CR network.

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

$$V_L^*[b(s_k)] = \max_{a_k \in \boldsymbol{A}_k} \sum_{s_k \in \boldsymbol{S}_k} b(s_k) \cdot R(s_k, a_k) + \rho \sum_{z_k \in \boldsymbol{Z}_k} P_r(z_k | b(\mathbf{S}_k), a_k) \cdot V_{L+1}^*[b(s_{k+1})] \quad (3.10)$$

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 (2.3). At the beginning of the time slot where the spectrum handoff occurred, the CR user will choose a target channel that possesses the minimum waiting slots based on the results obtained from the L-step value function as in (3.10). 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.

Moreover, it is noticed that the computation of L-step value function (in (3.10)) is considered complex and in general difficult to solve. The dimension of the belief state can grow exponentially as the number of channel is augmented, which makes it difficult to be adopted for practical implementation. A reduced state strategy has been proposed in [18] to establish an approximated linear state vector, which can effectively decrease the computation complexity of the value function. The complex optimization problem associated with (3.10) can therefore be resolved in an efficient manner.
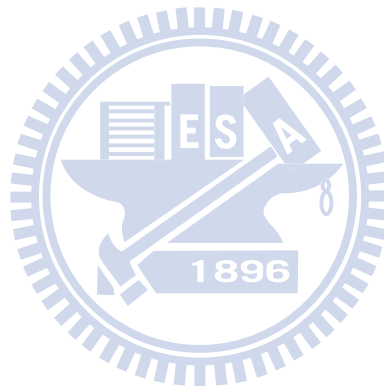
Chapter 4

Performance Analysis of Proposed POSH Scheme

In this chapter, the analytical model for proposed POSH scheme is derived in section A in order to analyze its performance. The models for both the no spectrum handoff (NSH) scheme and the randomly choose strategy (RCS) will also be demonstrated. The effectiveness of the analytical models is to serve as the validation purpose for these schemes, which will be compared with simulation results as in Section6. Furthermore, the analytical models for these three handoff algorithms under practical considerations will also be derived and explained in section B.

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 section. It's noticed that the parameter a_k defined in Section 2 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 (3.9).

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. Instead, two different types of actions are defined, including selecting the current channel a_c and the destination channel a_d where $a_c, a_d \in A_k$ and $a_c \neq a_d$. It is noted that the channel (either a_c or a_d) selected by the action will be retained for a time period until the 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 (3.8), the expected waiting time $n_{w,NSH}^{a_c}$ can be obtained as

$$
\begin{aligned}
n_{w,nsh} &= E[n_w = \ell | a^* = a_c] \\
&= \sum_{\ell=1}^{\infty} \ell \cdot P_r \left(\bigcap_{p=1}^{\ell-1} c_{a_c,k+p} = 1, c_{a_c,k+\ell} = 0 | c_{a_c,k} = 1 \right) \\
&= \sum_{\ell=1}^{\infty} \ell \cdot \tau(c_{a_c,k+\ell-1} = 1, a_c, c_{a_c,k+\ell} = 0) \cdot \tau(c_{a_c,k} = 1, a_c, c_{a_c,k+1} = 1)^{\ell-1} \quad (4.1)
\end{aligned}
$$

It is noted that the second equality in (4.1) 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 $n_{w,nsh}$ is to assign $P_r(c_{a_c,k} = 1) = 1$ in (3.8).

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 is probability

$P_r(a^* = a_s)$ for the action a_s to select one channel within all the N channels as below:

$$n_{w,rcs} = \sum_{a_s=1}^{N} E[n_w = \ell | a^* = a_s] \cdot P_r(a^* = a_s)$$

$$= \frac{1}{N} \cdot E[n_w = \ell | a^* = a_c] + \frac{1}{N} \cdot \sum_{a_s=1, a_s \neq a_c}^{N} E[n_w = \ell | a^* = a_s]$$

$$= \frac{1}{N} \cdot n_{w,nsh} + \frac{1}{N} \cdot \sum_{a_s=1, a_s \neq a_c}^{N} n_w^{a_s} \qquad (4.2)$$

where $n_{w,nsh}$ is defined in (4.1) and $n_w^{a_s}$ denotes the expected waiting time by selecting the channel a_s excepting to stay at the current channel a_c. In the case that the traffic pattern of all the N channels are identical, (4.2) can be reformulated by incorporating (4.1) and (4.2) as

$$n_{w,rcs} = \frac{1}{N}[1 + P_r(c_{a_d,k} = 1)] \cdot$$
$$\sum_{\ell=1}^{\infty} \ell \cdot \tau(c_{a_d,k+\ell-1} = 1, a_d, c_{a_d,k+\ell} = 0) \cdot \tau(c_{a_d,k} = 1, a_d, c_{a_d,k+1} = 1)^{\ell-1} nwNSH$$

$$(4.3)$$

where the selected destination channel a_d is composed by both the randomly selected channel a_s and the current channel a_c. Furthermore, as illustrated in (4.3), 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 $P_r(c_{a_s,k} = 1) \leq 1$ for $\forall a_s \neq a_c$. 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 mainly mapped by the belief state at each step as shown in Fig. 2.1. Since the value of belief state is obtained from the observation, action, and former belief state at each step, it becomes impossible to predict and obtained 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 varied rapidly. The effectiveness of the analytical model for the POSH scheme will be validated and evaluated via simulations as will be presented in Section 6.

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.

$$n_{w,posh} = \sum_{a_s=1}^{N} E[n_w = \ell | a^* = a_s] \cdot \tilde{P}_r(a^* = a_s) \tag{4.4}$$

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} respectively, where these two values are obtained as

$$a_{d1} = \arg \min_{\forall a_s} n_w^{a_s}$$

$$a_{d2} = \arg \min_{\forall a_s, a_s \neq a_{d1}} n_w^{a_s} \tag{4.5}$$

where the expected waiting time by conducting spectrum handoff $n_w^{a_s}$ is defined in (4.2) and can be obtained from (4.3). Considering that the channel conditions are

19

identical, switching to another channel will result in either smaller or at least equal expected waiting time than that with staying at the current channel, i.e. the probability $P_r(c_{a_s,k} = 1) \leq 1$ for $\forall a_s \neq a_c$ which results in $n_w^{a_s} \leq n_{w,nsh}^{a_c}$ as can be observed from (4.3). Therefore, the statistical distribution of the chosen action a^* in proposed POSH scheme can be acquired as

$$
\tilde{P}_r(a^* = a_s) = \begin{cases} 1, & \text{for } a_s = a_{d1} \text{ and } n_{w,nsh}^{a_{d1}} \leq n_w^{a_{d2}} \\ 0.5, & \text{for } a_s = a_{d1} \text{ and } n_{w,nsh}^{a_{d1}} > n_w^{a_{d2}} \\ 0.5, & \text{for } a_s = a_{d2} \text{ and } n_{w,nsh}^{a_{d1}} > n_w^{a_{d2}} \\ 0, & \text{otherwise.} \end{cases} \tag{4.6}
$$

It can be observed that the first case $\tilde{P}_r(a^* = a_s) = 1$ happens under the situations when a_s is selected as a_{d1} and $n_{w,nsh}^{a_{d1}} \leq n_w^{a_{d2}}$. It indicates that in the case of staying at the the channel a_{d1} with minimal expected waiting time will still result in comparably smaller or equal value than that from POSH-based spectrum handoff $n_w^{a_{d2}}$, it is suggested not to conduct spectrum handoff to another channel a_{d2}. Therefore, the probability to stay at the channel $a^* = a_{d1}$ will be equal to 1, i.e. $\tilde{P}_r(a^* = a_{d1}) = 1$. On the other hand, it is also required to consider the other case with $n_{w,nsh}^{a_{d1}} > n_w^{a_{d2}}$. Under the situation that the previous selected channel is a_{d1}, if the expected waiting time resulted from a_{d1} without spectrum handoff is worse than that from a_{d2} by conducting spectrum handoff, the corresponding CR user will decide to switch into channel a_{d2} at its spectrum handoff. Nevertheless, instead of remaining at channel a_{d2}, the CR user will choose a_{d1} again at the next spectrum handoff since the waiting time $n_w^{a_{d1}}$ is definitely smaller than $n_{w,nsh}^{a_{d2}}$. 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 $a^* = a_{d1}$ or a_{d2}. According to the probabilistic distribution of the action a^* as obtained in (4.6), the

expected waiting time of proposed POSH scheme can be reformulated from (4.4) as

$$n_{w,posh} = \begin{cases} n_{w,nsh}^{a_{d1}}, & \text{if } n_{w,nsh}^{a_{d1}} \leq n_w^{a_{d2}} \\ \frac{1}{2}n_w^{a_{d1}} + \frac{1}{2}n_w^{a_{d2}}, & \text{if } n_{w,nsh}^{a_{d1}} > n_w^{a_{d2}} \end{cases} \quad (4.7)$$

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. The performance validation and comparison between these schemes will be illustrated in Section 6.

4.2 Practical Considerations for Spectrum Handoff Schemes

In the previous section, according to the POMDP-based channel selection policies, the analytical results show that the proposed POSH scheme results in reduced expected waiting time slots 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 practical, it's inevitably to spend a period of time in order 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 necessitate channel sensing and handshaking will be discussed in this chapter. The performance difference between these schemes under the practical circumstance can consequently be observed.
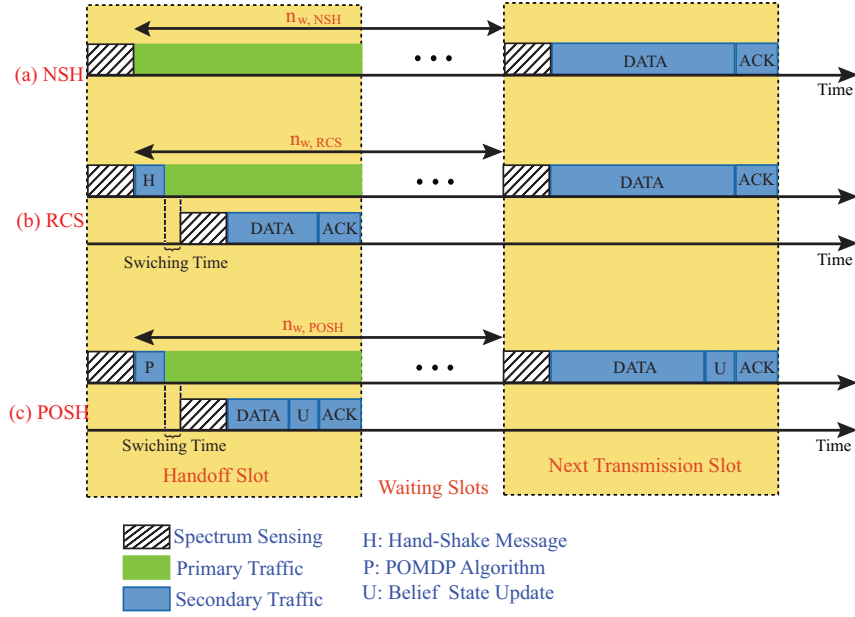
Figure 4.1: Practical design for the time-slotted channels under different spectrum handoff schemes.

As mentioned in previous chapters, the time slotted system is considered in this paper as shown in Fig. 4.1. It is required for the CR user to perform initial sensing on its current spectrum at the beginning of each time slot in order to assure the availability of the present channel. If the outcome of initial 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 ACK frame at the end of this slot to acknowledge to reception of the 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 will be stated as follows.

4.2.1 NSH Scheme

As shown in Fig. 4.1.(a), the behavior for the time-slotted channel with the adoption of the NSH scheme is depicted. The slots S_k and $S_{k+\ell}$ are utilized to represent the initial and final slots that are considered in the various handoff schemes for $\ell > 0$.

At the slot S_k, the intention from primary user to utilized this channel is observed during the spectrum sensing period. Based on the NSH scheme, the CR user will remain silent on the existing channel and wait until the primary user has finished its transmissions. Considering that the traffic from primary user has not been observed during the sensing period of slot S_{k+p}, 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

$$T_{w,nsh} = n_{w,nsh} \cdot T_{slot} \tag{4.8}$$

where $n_{w,nsh}$ is obtained from (4.1), and T_{slot} represents the time duration of each slot. Furthermore, the net transmission time within a slot by adopting the NSH method (i.e. $T_{s,nsh}$) can be acquired as

$$T_{s,nsh} = T_{slot} - T_{ack} - T_{sens} \tag{4.9}$$

where T_{sens} dentes the time interval of sensing period, and T_{ack} indicates the required time for returning the ACK packet in order to conduct handshake for data transmission.

4.2.2 RCS Scheme

The practical consideration for the RCS scheme is illustrated in Fig. 4.1.(b). After the initial sensing, the CR transmitter will randomly select a target channel for spectrum handoff, where the determination process is considered short and ignored in the designed. Considering that there are N channels to be selected in the network. According to the RCS method, there are $1/N$ probability that the CR user will remain at the current channel; while there are $(N-1)/N$ chances to switch to another spectrum.

If the CR user is determined to stay at the existing channel, the required waiting time acquired in (4.1) will be adopted. 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 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 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 $P_r(c_{a_d,k} = 0)$, the CR transmitter and receiver will spend T_{hs} time for exchanging handshake messages in order to confirm the utilization of target channel. As a result, the average waiting time of the RCS scheme $T_{w,rcs}$ becomes

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

$$(4.10)$$

where $n_{w,nsh}$ is obtained from (4.1). The resulting net transmission time in a slot $T_{s,rcs}$ can therefore be obtained as

$$T_{s,RCS} = T_{slot} - T_{ACK} - T_{sens} - (T_{hs} + T_{sw}) * P_r(c_{a_d,k} = 0) \qquad (4.11)$$

4.2.3 Precise POSH Scheme

The precise calculation for proposed POSH scheme is illustrated as shown in Fig. 4.1.(c). The POSH scheme does not need to spend handshake time between the transmitter and the receiver as required by the RCS method; while additional calculation time for conducting the POMDP policy is necessitated by the POSH scheme. It is assumed that all the CR users shared identical network condition and observed the same sensing consequence. Since the POMDP policy is simultaneously performed by both CR transmitter

and receiver, it is unnecessary to exchange messages in order to inform the receiver which spectrum is selected as the target channel. Nevertheless, CR user needs to spend additional time T_p to implement the POMDP algorithm, and the time T_u to update the belief state at the end of each slot. Combining all the compartments mentioned above, the overall waiting time $T_{w,posh}$ can be represented as

$$T_{w,posh} = P_r(c_{a_d,k} = 1) * (n_{w,posh} * T_{slot} + T_{sens}) + P_r(c_{a_d,k} = 0) * (T_p + T_{sw} + T_{sens})$$

(4.12)

The net transmission time in a slot $T_{s,posh}$ becomes

$$T_{s,posh} = T_{slot} - T_{ack} - T_{sens} - T_u - (T_p + T_{sw}) * P_r(c_{a^d,k} = 0) \qquad (4.13)$$

Due to the large amounts of system states and the complexity of witness algorithm [19–21], it is in general considered a complicated process to calculate a POMDP-based algorithm. It can be found in section 3.1 that as the number of channel increases, the number of system state is augmented exponentially, which makes the update process of belief state becomes complicated. A reduced form of system state was introduced in [3; 18] to be adapted in the problems with POMDP-based formulation. It has been proved to reduce the total number of states which effectively decrease the computation complexity. The reduced form successfully reduce the number of system state which becomes equivalent to the number of channel in the system. Furthermore, a efficient approach for belief state update is also exploited in order to effectively simplify the updated procedure. Since the number of reduced belief state is equal to the number of channels as N and each of the state is updated by a specific equation, it can be concluded that the time complexity of belief state update becomes $O(N)$.

Chapter 5

Proposed POMDP-based Multi-user Spectrum Handoff (M-POSH) Protocol

In Section 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 user at a specific time slot, those CR users will likely to select an identical frequency channel for spectrum handoff since all the secondary users observe the same behavior of primary user. Therefore, the consequence of performing the POSH scheme in multi-user network will result in packet collisions among those CR users that intend to perform spectrum handoff. A simple solution to deal with the collision problem is to adopt the CSMA-based contention resolution [**?**]. The CR users that choose the identical channel will pick a random number to backoff before actually conducting spectrum handoff. However, the CSMA-based scheme is considered an inefficient approach since there will only be one user that has the right to access the

chosen spectrum within a specific time slot. All the other CR users will turn out to lose the opportunity to transmit data in this slot.

In this chapter, 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 spectrum. The main purpose of the 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 sections, 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. It is assumed that there are n CR users which conduct saturated data transmission in a slotted CR network with N $(n < N)$ licensed spectrum. All the CR users are considered to observe the same channel behavior in the network. 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 exchange 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 [22] 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 section II.B and (3.1). 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

27

the set of 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 total available channels in the network as defined in section III.A. The update for belief state in multi-user scenario can be obtained as

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

$$(5.1)$$

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 (5.1) is derived via the reduced strategy as in [3] for POMDP-based formulation. For the rth user that determines to take action $a_{r,k}$ to access the ith channel, i.e. $a_{r,k} = 1$, the update of the belief state $b_r(c_{i,k+1})$ is determined by the observation $z_{r,k}$. In other words, the update of belief state $b_r(c_{i,k+1}) = 0$ if the observation $z_{r,k}$ is equal to 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 illustrate 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, the update process of belief state for the other $n - 1$ channel (i.e. $a_{r,k} \neq i$) is determined by the the probability computation as shown in the third item of (5.1). In other words, the idle probability for the next time slot is equal to the idle to the idle probability $\tau(c_{i,k} = 0, a_{r,k}, c_{i,k+1} = 0)$ times idle probability at this time slot plus busy to idle probability $\tau(c_{i,k} = 1, a_{r,k}, c_{i,k+1} = 0)$ times busy probability at this 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. The information that the occupancy of each channel by its corresponding CR user is available to all the CR users. In the case that a CR user intends to conduct spectrum handoff, it will broadcast a handoff message on the common control channel in order to announce all the users regarding the change of network condition. Since there may exist more than one CR user that need to broadcast its handoff message simultaneously in the same time slot, there is potential that these messages from different CR users can collide with each other. A random backoff contention window is therefore utilized in the proposed M-POSH scheme in order to resolve the potential packet collision problem. Each CR user that intends to delivery the handoff message on the control channel will wait for a random number in a pre-specified interval such as to ensure the message can be successfully received.

The handoff message includes three fields as follows: (a) access number, (b) original channel, and (c) candidate set of target channels. The access number is assigned to each CR user that intends to perform spectrum handoff, whose value is randomly selected by the CR user within a predefined interval. The original channel field records the ID of the channel that the CR user is currently in use before handoff. The candidate set of target channels for each specific CR user denotes the available channels to potentially conduct spectrum handoff, i.e. $\boldsymbol{D}_r = [d_{r,1}, ..., d_{r,n}]$ where $d_{r,i}$ represents the ith priority in the target channel set of user r. The procedure for user r to acquire the candidate set of target channels $\boldsymbol{D}_r$ can be expressed in Algorithm 1 as follows.

The concept of the candidate set of target channels for each CR user is to acquire the formation of unoccupied channels based on the criterion of minimal waiting time. As shown in Algorithm 1, the target candidates are selected sequentially based on the
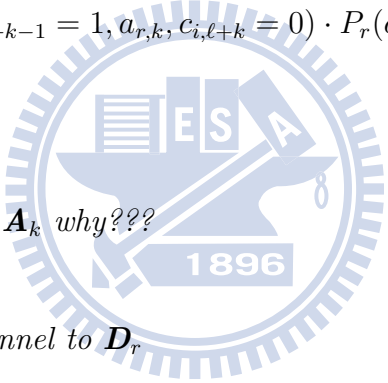
corresponding cost function Ψ. It is noted that the parameter m denotes the number of CR users that intend to conduct spectrum handoff at time slot k, which consist of the user set $\boldsymbol{R}_m$. If the selected action is observed to exist in $\tilde{\mathcal{A}}_k$, the chosen a^* will be replaced by the original channel and recorded in the candidate set. As a result, this replacement criterion can ensure that none of the CR user's candidate set $\boldsymbol{D}_r$ will be identical. This can facilitate the negotiation process will for the CR users to select a designate target channel from their candidate set, which will be described as in Algorithm 2.

Algorithm 1: *Acquiring Candidate Set of Target Channels $\boldsymbol{D}_r$ for User r*

Input : $\tilde{\mathcal{A}}_k$ with size (?? n), $\boldsymbol{A}_k$ with size N
Output: $\boldsymbol{D}_r$ with size n

1 **for** $i \leftarrow 1$ **to** n **do**
2     $\Psi \leftarrow \sum_{\ell=1}^{m} \ell \cdot \tau(c_{i,\ell+k-1}=1, a_{r,k}, c_{i,\ell+k}=0) \cdot P_r(c_{i,k}=1) \cdot \tau(c_{i,k}=1, a_{r,k}, c_{i,k+1}=1)^{\ell-1}$
3     $a_r^* = \arg\min_{a_{r,k}} \Psi$
4   **if** $a_r^* \notin \tilde{\mathcal{A}}_k$ **then**
5   *Add a_r^* to $\boldsymbol{D}_r$*
6   *Remove a^* from $\boldsymbol{A}_k$ why???*
7   **end**
8   **else**
9   *Add original channel to $\boldsymbol{D}_r$*
10   **end**
11 **end**

As illustrated in Algorithm 2, after collecting the information of handoff message from the m CR users, the target channel for each user can therefore be determined by a negotiation procedure between the users. Let $\boldsymbol{T}_k$ denotes the set of target channel that are selected by the CR users at the occurrence of spectrum handoff in time slot k. The target channels for the entire m CR users will be sequentially selected based on the access number of each CR user. The selection criterion is based on the situation if the potential target channel $d_{r,i}$ is not in both sets of $\boldsymbol{T}_r$ and $\tilde{\mathcal{A}}_k$. Therefore, every

CR users within the network can choose its target channel without collision with the other CR users. It is noticed that with the consideration of other CR traffic within the network, the occupied channel sets of CR users $\tilde{\mathcal{A}}_k$ will also be updated based on the information from both the original channel and the target channel of the user set $\boldsymbol{R}_m$ that conduct spectrum handoff. In summary, the proposed M-POSH protocol can provide a negotiation mechanism for each CR user to select its target channel without being collide with the selection made by other CR users in the network.

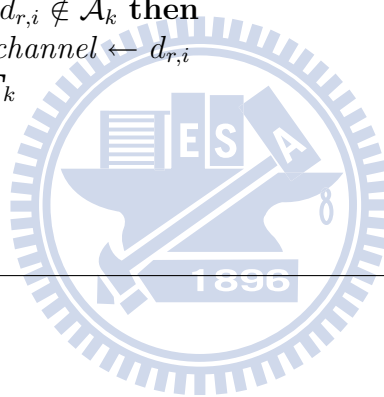**Algorithm 2**: *Selection of Target Channel*

    Input   : $\boldsymbol{D}_r$ with size n, $\boldsymbol{R}_m$ with size m
    Output: *Target channel of each CR user $a_{k,r}$*

1 **foreach** *CR users in $\boldsymbol{R}_m$* **do**
2   **for** $i \leftarrow 1$ **to** m **do**
3   **if** $d_{r,i} \notin \boldsymbol{T}_r$ *and* $d_{r,i} \notin \tilde{\mathcal{A}}_k$ **then**
4   *Destination channel* $\leftarrow d_{r,i}$
5   *Add $d_{r,i}$ to $\boldsymbol{T}_k$*
6   *break*
7   **end**
8   **end**
9 **end**

Chapter 6

Performance Evaluation

In this chapter, simulations are presented to demonstrate the performance of proposed POSH and M-POSH protocols. The major focus in the simulations is to obtain the required waiting time slots for the secondary user while it has been directed to the target channel. 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 Section 4, including both the NSH scheme and the RCS mechanism. The traffic of the primary user follows the poison distribution, and the service time is assumed to be a uniform distribution with mean $1/\mu = 1$. Three channels are considered in the simulations, i.e. $s_k = [c_{1,k}, c_{2,k}, c_{3,k}]$; while the discount factor in (3.10) is selected as $\rho = 1$. The reduced state strategy is utilized in the simulations to obtain the numerical results of the POMDP-based optimization problem.
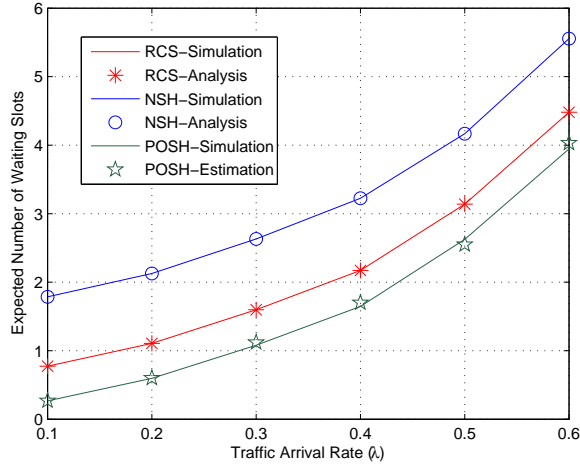
Figure 6.1: Performance validation: the number of waiting time slots versus traffic arrival rate.

6.1 Model Validation

The analytical models for required waiting time slots of the three schemes, including POSH, NSH, and RCS algorithms, as presented in (4.1), (4.3), and (4.7) are validated via simulations results as shown in Fig. 6.1. It can be observed that the expected number of waiting time slots 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 imprecise modeling of POSH scheme based on stationary probabilities.

In order to clearly illustrate the difference between these two results, the biased percentage β is introduced and is defined as $\beta = \frac{n_{w,\zeta}^a - n_{w,\zeta}^s}{n_{w,\zeta}^s} \times 100$ where $n_{w,\zeta}^a$ and $n_{w,\zeta}^s$ correspond to the expected waiting time slots obtained from analysis and simulation
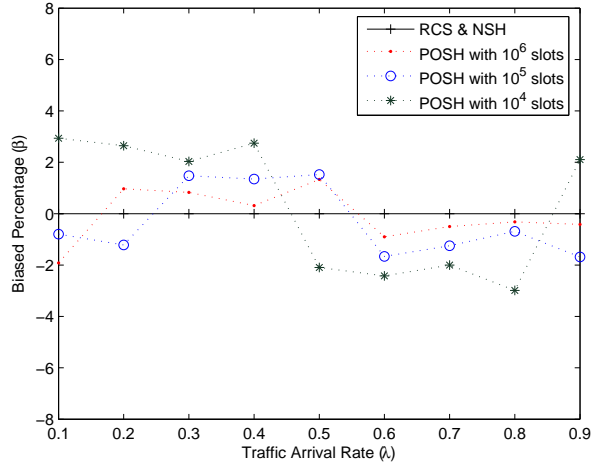
Figure 6.2: Performance validation: the biased percentage β versus traffic arrival rate.

respectively. The parameter ζ indicates the either one of the three spectrum hand-off schemes. Fig. 6.2 illustrates the biased percentage β for the three schemes under different traffic arrival rates. It is noted that the proposed POSH method is implemented under the simulation runs with different number of transmission time slots, i.e. $T = 10^4$, 10^5, and 10^6. It can be observed that even the behavior of belief state can not 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. 6.2 that the bias can be diminished as the number of transmission time slots for the simulations are increased. This results reveal the case that with longer time of simulation, the incomplete network information can be updated more accurately. The simulation results will tend to possess stationary behaviors as is presented by the analytical model in (4.7).
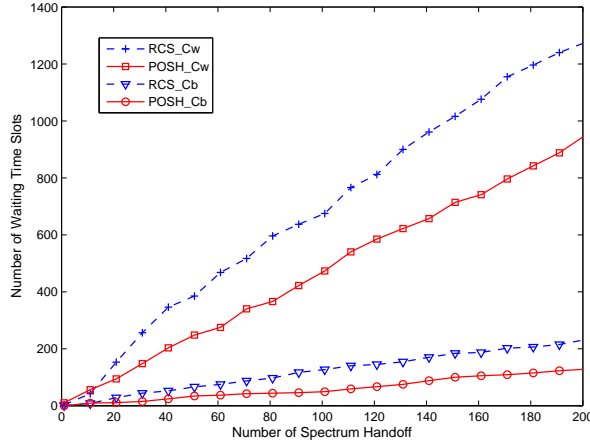
Figure 6.3: Performance comparison: the number of waiting time slots versus the number of spectrum handoff.

6.2 Performance Comparison

Fig. 6.3 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 method is not implemented in this case since the CR user can always stay at the channel with better condition. It is intuitive to see 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 under both schemes. Nevertheless, the total waiting time
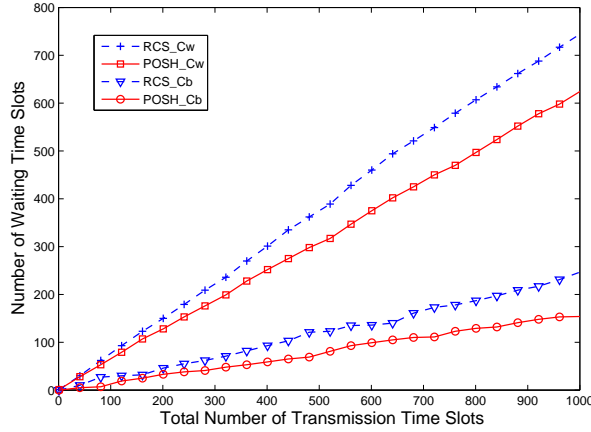
35

Figure 6.4: Performance comparison: the number of waiting time slots versus the total number of transmission time slots.

slots acquired from the proposed POSH scheme is comparably smaller than that from the RCS method under both channel conditions. It is also observed that the POSH scheme performs better as the number of spectrum handoff is increased. The reason can be attributed to the situation that more updated belief states are acquired by the POSH scheme as the number of handoff is augmented.

Fig. 6.4 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. 6.4 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 (in (3.10)), the POSH scheme can still provide smaller waiting time comparing with the RCS method.

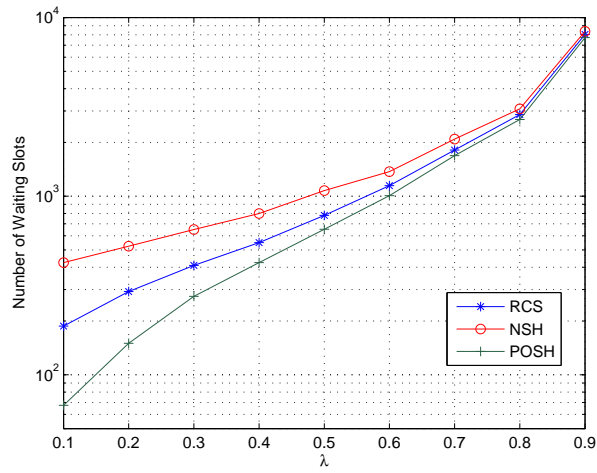Figs. 6.5 and 6.6 illustrate the performance comparison among the POSH, the RCS,

Figure 6.5: Performance comparison: the number of waiting time slots versus traffic arrival rate (numbers of spectrum handoff = 250).
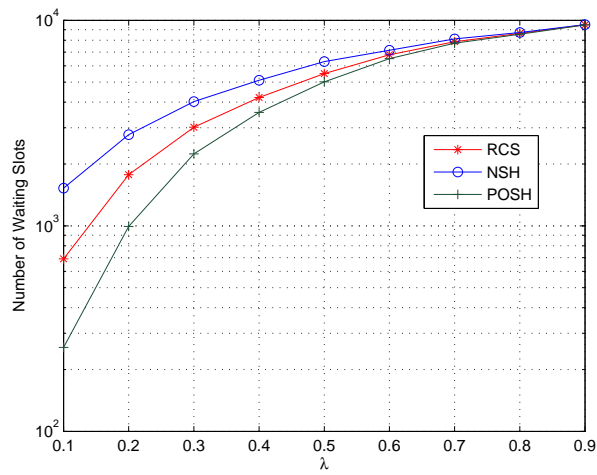

Figure 6.6: Performance comparison: the number of waiting time slots versus traffic arrival rate (number of transmission time slots = 10000).
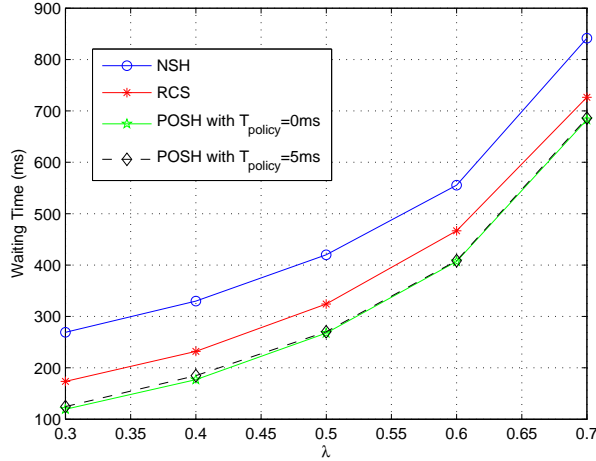
Figure 6.7: Performance comparison: the waiting time with practical consideration versus different arrival rates.

and the NSH schemes under different values of packet arrival rate λ of the primary user. Fig. 6.5 shows the comparison under fix numbers of spectrum handoff equal to 250 and Fig. 6.6 is performed under number of transmission time slots equal to 1200. It can be observed that the proposed POSH scheme can outperform the other two methods under different packet arrival rates. The benefits from the adoption of the POSH algorithm is especially revealed at smaller values of packet arrival rate since there can be more opportunity for the POSH scheme to select a feasible target channel.

The practical consideration for the three handoff schemes are presented as in Figs. 6.7 and 6.8. 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. Fig. 6.7 illustrates the expected waiting time as was obtained from (4.8), (4.10), and (4.12) from the NSH, RCS, and POSH scheme respectively. Two cases with policy time $T_p = 0$ and 5 ms for the proposed POSH scheme are considered; while the update time $T_u =$??. With practical consideration, the proposed POSH scheme is demonstrated to provide smaller expected waiting time slots under different arrival rate of primary user. Even though the calculation of POMDP-based policy requires additional computation time,
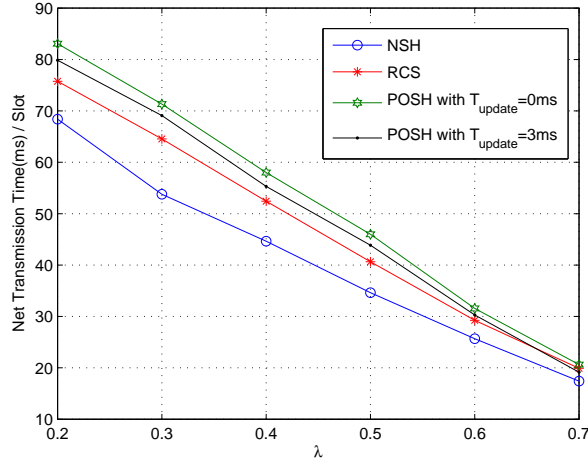
Figure 6.8: Performance comparison: the net transmission time per slot versus different arrival rates.

the overall performance is still better in comparison with the NSH and RCS methods. Furthermore, it is observed in Fig. 6.7 that the waiting time $T_{w,posh}$ will only be affected by T_p while the target channel of spectrum handoff has high probability in idle state. In other words, T_p will degrade the performance of proposed POSH scheme at lower primary traffic circumstance, and will become insignificant with higher primary traffic. Nevertheless, since POSH scheme can provide better performance among the three schemes at low primary traffic (as shown in Figs. 6.5 and 6.6), it is concluded that the degraded effect from T_p will not be significant by adopting the POSH scheme.

In order to better present the utilization of licensed spectrum, Fig. 6.8 is exploited to illustrate the net transmission time within a time slot. Two cases with update time $T_p = 0$ and 3 ms for the proposed POSH scheme are considered; while the policy time $T_p =$??. It is noticed that the update time T_u is considered additional time to expense for the proposed POSH scheme; while the policy time T_p will only occur as spectrum handoff happens. It is intuitive to observe that the net transmission time obtained from all three schemes decreases as the arrival rate of primary traffic is increased. Furthermore, as λ is increased, the net transmission time of proposed POSH scheme
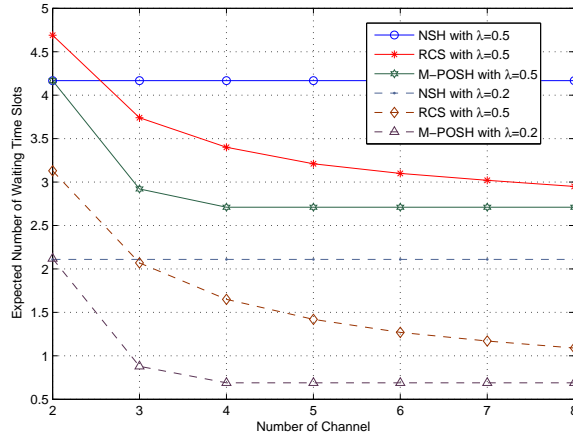
39

Figure 6.9: Performance comparison: the expected number of waiting slots versus different number of channels (with 2 CR users).

becomes closer or even worse than the RCS and NSH schemes, e.g. the POSH scheme with $T_u = 3$ ms is worse than the RCS algorithm as $\lambda > 0.75$. Therefore, practical consideration for these handoff schemes can provide a channel selection criteria for the CR users to determine which handoff scheme should be applied to obtain their target channel.

Figs. 6.9 and 6.10 compare the performance of NSH scheme, RCS scheme, and the proposed M-POSH protocol under the circumstance of multiple CR users within multi-channels network. Fig. 6.9 shows the performance comparison under two CR users; while Fig. 6.10 illustrates the circumstance with 7 channels in the networks. Two different arrival rates of primary traffic are considered for all 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 case without the consideration of multiple users in the network. On the other hand, regarding the RCS and M-POSH schemes, the expected number of waiting slots is decreased as the number of available
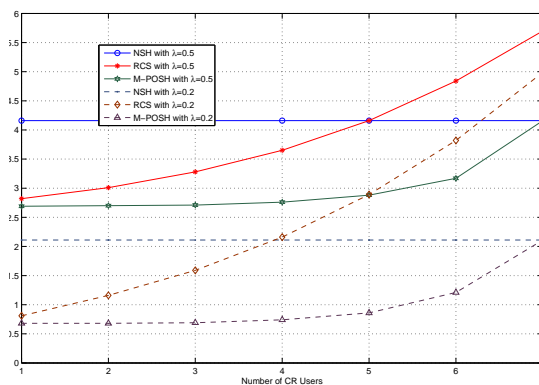
Figure 6.10: Performance comparison: the expected number of waiting slots versus different number of CR users (with 7 available channels).

channels is augmented; while it is increased with the augmentation of total CR users due to the potential collisions happened in the network. The performance of RCS scheme can becomes worse than that from NSH method in certain circumstances accounting to NSH scheme at least guarantee no collisions will happen in the network. From example as in Fig. 6.10, the RCS scheme with $\lambda = 5$ results in higher number of waiting slots as the number of CR users greater than 5 in comparison with the NSH method.

Nevertheless, the proposed M-POSH protocol can adaptively selecting the target channel based on the availability of network channels with the consideration of collisions between the CR users. As shown in both Figs. 6.9 and 6.10, the proposed M-POSH scheme outperforms both the RCS and NSH algorithms under different scenarios. As the network channels are in occupied conditions, e.g. the number of CR users is equal to the number of licensed spectrum as the most left data point in Fig. 6.9, the proposed M-POSH protocol will decide to stay at the current channel in order to reduce the probability of 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

41

performance as that obtained from the POSH scheme for single user case (as presented in Fig. 6.1). The merits of the proposed POSH and M-POSH schemes can therefore be observed.

Chapter 7

Conclusion

This paper proposes a strategy for post-sensing spectrum handoff 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 target channel in order to achieve the minimal waiting time for packet transmission. Furthermore, in order to consider the multi-user CR network, POMDP-based multi-user spectrum handoff (M-POSH) protocol is proposed to resolve the collision problem among multiple users that intend to conduct spectrum handoff. It is observed from both the simulation and analytical results that the proposed POSH M-POSH protocols can effectively reduce the waiting time of spectrum handoff for a partially observable CR network.

Bibliography

[1] F.C.C, "Unlicensed operation in the TV broadcast bands," *ET Docket*, pp. 4–186, 2006.

[2] FCC, "Fcc spectrum policy task force report," *ET Docket No. 02-135.*, 2002.

[3] Q. Zhao, L. Tong, A. Swam, and Y. Chen, "Decentralized cognitive MAC for opportunistic spectrum access in ad hoc networks: A POMDP framework," *IEEE J. Sel. Areas Commun.*, vol. 25, pp. 589–600, 2007.

[4] V. K. Bhargava, "Research challenges in cognitive radio networks," *Proc. IEEE ICCCN*, pp. 1–6, 2007.

[5] S. Shanker, "Squeezing the most out of cognitive radio: A joint mac/phy perspective," *Proc. IEEE ICASSP*, 2007.

[6] A. Sahai, N. Hoven, and R. Tandra, "Some fundamental limits on cognitive radio," *Proc of Alleston Conference*, 2004.

[7] S. Sengupta, S. Brahma, M. Chatterjee, and N. S. Shankar, "Enhancements to cognitive radio based IEEE 802.22 air-interface," *Proc. IEEE ICC*, 2007.

[8] C. Cordeiro, K. Challapali, D. Birru, and N. S. Shankar, "IEEE 802.22: The first worldwide wireless standard based on cognitive radios," *Proc. IEEE DySPAN*, pp. 328–337, 2005.

[9] L. C. Wang and A. Chen, "On the performance of spectrum handoff for link maintenance in cognitive radio," *Proc. IEEE ISWPC*, pp. 670–674, 2008.

[10] X. Zhu, L. Shen, and T. S. P. Yum, "Analysis of cognitive radio spectrum access with optimal channel reservation," *IEEE Communications Letters*, vol. 11, pp. 304–306, 2007.

[11] L.-C. Wang and C. W. Wang, "Modeling and analysis for proactive-decision spectrum handoff in cognitive radio networks," *Proc. IEEE ICC*, 2009.

[12] L. C. Wang and C. W. Wang, "Spectrum handoff for cognitive radio networks: Reactive-sensing or proactive-sensing?" *Proc. IEEE IPCCC*, 2008.

[13] X. Long, X. Gan, Y. Xu, J. Liu, and M. Tao, "An estimation algorithm of channel state transition probabilities for cognitive radio systems," *Proc. IEEE CROWNCOM*, 2008.

[14] A. R. Cassandra, "Exact and approximate algorithms for partially observable markov decision processes," Ph.D. dissertation, Brown University, 1994.

[15] L. P. Kaelbling, M. L. Littman, and A. R. Cassandra, "Planning and acting in partially observable stochastic domains," *Artificial Intelligence*, vol. 101, pp. 99–134, 1998.

[16] K. J. Astrom., "Optimal control of markov decision processes with incomplete state estimation," *Journal of Mathematical Analysis and Applications*, pp. 174–205, 1965.

[17] H.-T. Cheng, "Algorithms for partially observable markov decision processes," Ph.D. dissertation, University of British Columbia, British Columbia, Canada, 1988.

[18] Q. Zhao, L. Tong, and A. Swami, "Decentralized cognitive mac for dynamic spectrum access," *Proc. IEEE DYSPAN*, pp. 224–232, 2005.

[19] M. L. Littman, "The witness algorithm: Solving partially observable macov decision process," *Technical Report: CS-94-40*, 1994.

[20] R. D. Smallwood and E. J. Sondik, "The optimal control of partially observable markov processes over a finite horizon." *Operations Research. 21*, pp. 1071–1088, 1973.

[21] R. S. Sutton and A. G. Barto, "Reinforcement learning: An introduction." *MIT Press, Cambridge, Massachusetts*, 1998.

[22] L. C. Wang, Y. C. Lu, and C. W. Wang, "Latency analysis for dynamic spectrum access in cognitive radio: Dedicated or embedded control channel?" *Proc. IEEE PIMRC*, 2007.

[23] D. Gross and C. M. Harris, *Fundamentals of Queueing Theory*, 3rd ed. Wiley, 1998.

[24] W. Hu, D. Willkomm, M. Abusubaih, J. Gross, G. Vlantis, M. Gerlaand, and A. Wolisz, "Cognitive radios for dynamic spectrum access - dynamic frequency hopping communities for efficient IEEE 802.22 operation," *IEEE Communications Magazine*, vol. 45, pp. 80–87, 2007.

[25] Z. Kostic and I. Maric, "Dynamic frequency hopping in wireless cellular systems-simulations of full-replacement and reduced-overhead methods," *Proc. IEEE VTC*, pp. 914–918, 1999.