Dynamic Spectrum Sharing for Future Wireless Communications

Xueyuan Jiang
Wadham College

Communications Research Group
Department of Engineering Science
University of Oxford

A thesis submitted for the degree of *Doctor of Philosophy*

Hilary Term 2013

Declaration

I hereby declare that this thesis entitled 'Dynamic Spectrum Sharing for Future Wireless Communications' is not substantially the same as any that I have submitted for a degree or diploma or other qualification at any other University. I further state that no part of my thesis has already been or is being concurrently submitted for any such degree, diploma or other qualification. Except where noted in the context, this thesis is the result of my own work during my Ph.D. study.

I hereby declare that my thesis does not exceed the limit of length prescribed in the Special Regulations of the Ph.D. examination for which I am a candidate. The length of my thesis does not exceed 65,000 words.

Signed

Xueyuan Jiang

I would like to dedicate this thesis to my loving parents and wife.

Acknowledgements

First, I would like to express my gratitude to my supervisor, Prof. David J. Edwards, for his guidance, encouragement, patience and supervision during my three-year DPhil study. His deep insights, knowledge and valuable suggestions make my research work always go in the right track.

I am also grateful to Dr. Yangyang Zhang and Dr. Kai-Kit Wong for their valuable advice, helpful comments and encouragement.

Thanks to my collaborators and good friends, Dr. Fan Zhang, Dr. Riqing Chen, Dr. Lubin Zeng, and Dr. Jingjing Liu, for the interesting discussions and sharing ideas during the last couple of years. Thanks to Mr. Grahame Faulkner for his technical support.

I would like to thank all my friends who have enriched my colorful life in Oxford. Thanks to Wenjia Bai, Po Wang, Qifeng Cheng, Hao (Edgar) Wang, Yimin Wu, Fan Xu, Mengchen Hu, Huayong Zhao, Min Wu, Luyun Jiang, and many others.

I would like to acknowledge the financial support from the China Scholarship Council.

Last but not least, I thank my parents for their endless support and encouragement. I would like to express my sincere gratitude to my wife for her love, patience, understanding and being with me in the whole life journey.

Abstract

The spectrum has become one of the most important and scarce resources for future wireless communications. However, the current static spectrum policy cannot meet the increasing demands for spectrum access. To improve spectrum efficiency, dynamic spectrum access (DSA) attempts to allocate the spectrum to users in an intelligent manner. Cognitive radio (CR) is an enabling technology for DSA, and can maximize spectrum utilization by introducing unlicensed or secondary users (SUs) to the primary system. The key component of DSA is dynamic spectrum sharing (DSS), which is responsible for providing efficient and fair spectrum allocation or scheduling solutions among licensed or primary users (PUs) and SUs. This thesis focuses on the design of efficient DSS schemes for the future wireless communication networks.

Firstly, based on the coordinated DSS model, this thesis proposes a heterogeneous-prioritized spectrum sharing policy for coordinated dynamic spectrum access networks. Secondly, based on the uncoordinated DSS model, a novel partial spectrum sharing strategy and the cross-layer optimization method have been proposed to achieve efficient spectrum sharing between two licensed networks. Then, a hybrid strategy which combines the overlay and underlay schemes is proposed under uncoordinated DSS model. The proposed analytical methods can provide efficient and accurate modeling to predict the behaviors of the PUs and SUs in DSS systems. This thesis presents the performance prediction of the proposed novel DSS schemes that achieve efficient spectrum sharing for coordinated and uncoordinated future wireless networks.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Research Motivation

The wireless communication industry has experienced an exponential increase of mobile services, especially in the last decade. In 2011, there were 6 billion mobile cellular subscribers and 1.2 billion mobile broadband subscribers around the world. With the launching of the new wireless networks and services, such as wireless regional area networks (WRAN), wireless local area networks (WLAN), and wireless regional personal networks (WPAN), wireless has become a required technology, which changes our life style-the way we work, live and play. It is expected that wireless technology will take over personal computers to become the premier technology driver for the next decades, and will have profound impact on our economy and society.

In wireless communications, the data transmission rate and the spectrum utilization rate are the two important parameters to evaluate the performance of wireless systems. Future wireless communication networks are expected to provide much higher data rates to meet the requirements of increasing media services [1]. Based on the Shannon–Hartley theorem [2], the data transmission rate is bounded by the bandwidth of the spectrum resources. Some physical layer technologies, such as Multiple-in-multiple-out (MIMO) [3] and Orthogonal frequency-division multiplexing (OFDM) [4], were proposed to improve the spectral efficiency. However, due to scarce frequency spectrum and increasing demands for high quality communication services, the required capacity has exceeded that provided by traditional wireless communication systems [5].

On the other hand, with the rapid development of wireless technologies, various systems coexist and therefore the problem of crowded frequency spectrum and severe interference will arise. To solve these problems, the current wireless systems are regulated by static and “exclusive-use” spectrum assignment policy, i.e., the spectrum resources are governed and regulated by the governmental agencies, and are allocated to different wireless applications, such as mobile phone, broadcast, radar and television networks, on a long term basis for a large geographical regions [6]. The static spectrum assignment policy can prevent the interference among different wireless services. However, it also causes the inefficient utilization of spectrum resources [7]. Based on the spectrum measurement report by the Shared Spectrum Company [8] for downtown area in Washington D. C., there are about 62% spectrum resources that have not been utilized. Furthermore, the federal communications commission (FCC) reported that there are a lot of temporal and geographical spectrum holes (unused spectrum) in the assigned 3GHz-below frequency bands, in which the spectrum utilization is range from 15% to 85% [7].

Dynamic spectrum access (DSA) on the other hand attempts to allocate the spectrum to users in an intelligent manner, has received most attention as the solution for better spectrum utilization [9]. Cognitive radio (CR) is an enabling technology for DSA, and can maximize spectrum utilization by introducing unlicensed or secondary users (SUs) to the primary system [10; 11]. With CR, SUs are aware of the radio frequency environment, and dynamically selects the operating parameters (such as frequency, modulation scheme, and transmission power), to opportunistically access the spectrum holes (i.e., the unused frequency bands) without causing interference to the licensed or primary users (PUs).

1.2 Research Objectives

Dynamic spectrum sharing (DSS) is the key component of DSA, and is responsible for providing efficient and fair spectrum allocation or scheduling solutions among primary and secondary users. To do so, there are generally two main approaches [9]:

- Coordinated sharing (the spectrum property-rights model or spectrum leasing)—Under this framework, certain information exchange between the PUs and the SUs is allowed. PUs are the owners of the spectrum, and can lease their spectrum to the SUs via a centralized spectrum manager (e.g., spectrum broker in [12]). Hence, PUs are incentivized in spectrum sharing to gain revenues.

- Uncoordinated sharing (the hierarchical spectrum access model)—In this model, SUs constantly sense the spectrum utilization and access any available spectrum holes in which PUs are not active. Moreover, SUs are also required to detect the PU activities to ensure no harmful interference causing to the PUs. As such, PUs will be operating as if SUs are not there, provided that spectrum sharing by the SUs is perfect.

Hence, this research project focuses on the design of efficient DSS schemes under coordinated and uncoordinated sharing scenario for the future wireless communication networks. For coordinated sharing, since there is spectrum manager to coordinate the spectrum allocation and coexistence of PUs and SUs, SUs do not need to perform sensing to get spectrum occupation information. In this scenario, hierarchical multiple-class SUs can be considered and accommodated in the system. PUs may adopt different prioritized policies to solve the channel collision with SUs. Markov chain models will be used to analyze the overall system performance and based on this model, the maximum admitted traffic of SUs, also referred as the system serving capability for SUs, can be obtained. Furthermore, Call Admission Control (CAC) strategies can be implemented in the system to enhance the maximum admitted traffic for SUs. For uncoordinated sharing, to exploit spectrum holes efficiently and avoid harmful interference to PUs, spectrum sensing becomes one of the most important and fundamental physical layer tasks. Due to the environmental noise and interference, sensing errors are inevitable and should be considered in the analysis. The parameters selection during spectrum sensing at the physical layer will also affect the MAC layer performance such as the blocking probability and interference probability, which will further affect the grade-of-service (GoS) performance the spectrum sharing users. Therefore, the cross layer design framework will be investigated to enhance the overall DSS system performance. In uncoordinated sharing, the overlay and underlay access are the two common sharing models. As the interference tolerance of PUs is ignored for overlay access, and the opportunity for SUs to transmit at higher power when PUs are inactive is missed for underlay access, mixed or hybrid overlay-underlay strategies are desirable and have been proposed to improve spectrum efficiency, which can also achieve a better trade-off between the network throughput and the interfering to PUs.

1.3 Contributions and Thesis Outline

The work described in this thesis has investigated how to efficiently utilize the limited spectrum resources in the manner of time-frequency-space domains using CR technologies. In particular, three important issues have been addressed: 1) how to achieve the collaboration between PUs-SUs and SUs-SUs for spectrum sharing, with the guarantee of PUs' access priority. 2) how to efficiently and accurately model the users access process for the new DSS strategy, and predict the system performance using the analytical model. 3) how to adjust the parameters in different communication network layers to optimize the performance metrics.

This thesis contains 7 main chapters which will provide further descriptions for the research on coordinated and uncoordinated DSS networks mentioned above. A brief outline of the contributions in each chapter is as follows.

Chapter 2 gives an overview of DSA, CR, Markov process, and related work on coordinated and uncoordinated spectrum sharing.

Chapter 3 proposes a heterogeneous-prioritized spectrum sharing policy for coordinated dynamic spectrum access networks, where a centralized spectrum manager coordinates the access of PUs and SUs to the spectrum. Through modeling the access of PUs and multiple classes of SUs as CTMC models, the overall system performance with consideration of GoS guarantee for the users is analyzed. In addition, two new CAC strategies are devised in the models to enhance the maximum admitted traffic of SUs for the system. Numerical results show that the proposed heterogeneous-prioritized policy achieves higher maximum admitted traffic for SUs. The tradeoff between the system's serving capability and the serving fairness for multiple classes of SUs is also studied. Moreover, the proposed CAC strategies can achieve better performance under max-sum, proportional, and max-min fairness criteria than the conventional CAC strategies.

Chapter 4 first proposes a new partial spectrum sharing policy, which achieves efficient spectrum sharing between two licensed networks. Then, a CTMC model is developed to analyze the proposed spectrum sharing policy considering the effects of on-demand sensing errors. A cross layer design framework is also constructed, in which the parameters of spectrum sharing policy at the MAC layer and the spectrum sensing parameters at the physical layer are simultaneously coordinated to maximize the overall throughput of the networks, while satisfying the GoS constraints of the users. Particle Swarm Optimization (PSO) algorithm is introduced to solve the cross layer optimization problem. Numerical results show that the proposed spectrum sharing policy and cross layer design strategy achieves a higher overall throughput of the two networks.

Chapter 5 further explores the influence of sensing errors during periodic sensing. A DTMC model is developed for partial spectrum sharing policy. The main contribution is a DTMC analysis characterizing the impact of imperfect sensing and sensing periodicity on the network performance. A cross-layer design framework is also developed, where the parameters in spectrum sensing and spectrum sharing are jointly optimized. Results show that the proposed policy achieves a higher throughput than conventional strategies.

Chapter 6 studies the traditional overlay and underlay access schemes in DSA. A hybrid access strategy which combines the advantages of these two traditional schemes is proposed. Based on a double-threshold energy detection method, SUs can switch between full-access and partial-access operation mode dynamically. A Markov chain model is developed to derive performance metrics for evaluating the proposed hybrid access strategy. Numerical results show that the proposed strategy can improve the interfering performance for PUs. Moreover, SUs can adjust the access probability of partial-access mode to achieve tradeoff between the throughput and interfering performance.

Chapter 7 concludes the thesis and discusses potential directions for future research.

Chapter 2

Overview of Research Background

In this chapter, an overview of DSA, CR and Markov chain models are given. The related work for coordinated and uncoordinated spectrum sharing are also analyzed and discussed.

2.1 Basic Concepts of DSA and CR

The concept of DSA contains various approaches to achieve intelligent spectrum assignment. As illustrated in Figure 2.1, DSA strategies can be broadly categorized under three models [9].

- Dynamic Exclusive Use Model—This model keeps the frameworks of the current static spectrum assignment policy, i.e., the spectrum band will be allocated to specific wireless services for exclusive use. The aim of this model is to improve spectrum efficiency by introducing flexibility. Two approaches can be used to achieve this aim under this model, which are spectrum property rights and dynamic spectrum allocation. In the former approach, the owner of the spectrum can lease his spectrum to the unlicensed users who would like to access it. However, although the owner has the rights to lease or share his spectrum freely, it is not an obligatory activity, and thus cannot be treated as a spectrum assignment policy. For the second approach, based on the statistics of communication traffic under spatial and temporal basis, the spectrum will be allocated to different services for exclusive use. Due to the burstiness and variation of the communication traffic, dynamic spectrum allocation cannot improve the spectrum efficiency.

- Open Sharing Model—In this framework, a certain frequency band is defined and referred to as the unlicensed spectrum band, where users can use such band without the spectrum licenses. Due to the popularity of WiFi, this model has received much attention. However, based on coexistence among peer-users, the open sharing comes with the risk of mutual interference. In today's unlicensed bands, there is no limit to the number of devices in a given location, and no limit to the potential congestion [14].

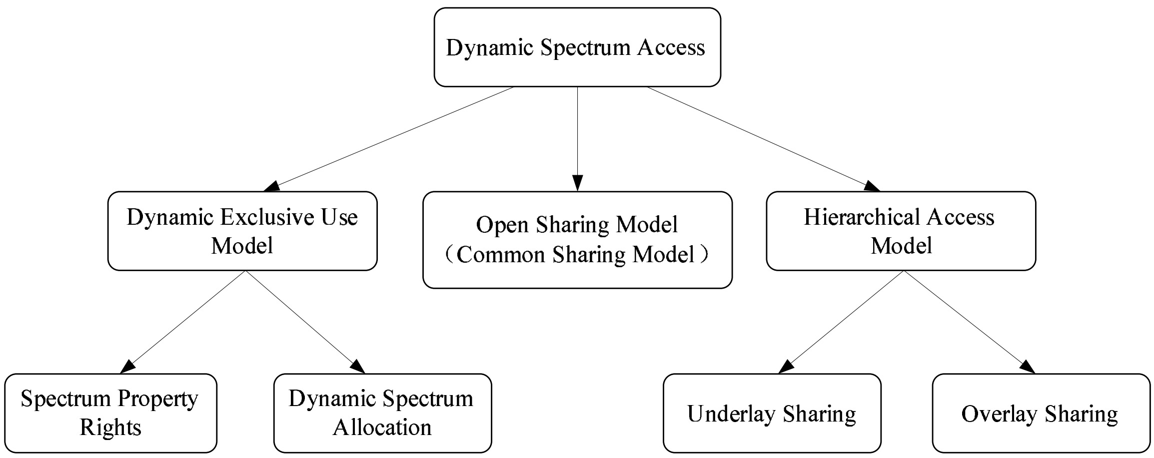


Figure 2.1: Dynamic Spectrum Access Models.

- Hierarchical Access Model—This model is also referred to as primary-secondary sharing model. The PUs are the owner of the spectrum, while SUs are allowed to transmit in the spectrum without causing harmful interference to PUs. Furthermore, SUs can either be allowed to transmit at lower power that they never cause harmful interference to the PUs, as with ultra-wideband [13], or they can be allowed to transmit opportunistically when they determine that transmissions will not cause harmful interference [14].

In the above three models, since Hierarchical Access Model does not require hardware updates for the current systems and is capable of controlling potential interference, and thus be treated as the most promising model. CR technology is the key technology that enables DSA. The term, cognitive radio, can formally be defined as [7]: A "Cognitive Radio" is a radio that can change its transmitter parameters based on interaction with the environment in which it operates. Cognitive capability and reconfigurability are the two main characteristics of CR, which are referred to as the ability to capture radio environment information and to coordinate its parameters according to the radio environment, respectively. The main functions for CR in DSA networks can be summarized as follows [6]:

- Spectrum sensing: Detecting spectrum holes (unused spectrum) and the activities of PUs, access the spectrum without causing harmful interference with other users.

- Spectrum management: Capturing the best available spectrum to meet user communication requirements.

- Spectrum mobility: Maintaining seamless communication requirements during the transition to better spectrum.

- Spectrum sharing: Providing the fair spectrum scheduling method among SUs.

2.2 Markov Process and Markov Chain

This section is mainly based on [15]–[17], which deals with the fundamental concepts related to Markov processes and Markov chains. To start with, the notion of Markov process, and then the state transitions of the Markov chain model is introduced.

2.2.1 Definition

The Markov processes appear in numerous applications, including queuing theory, computer communications networks, statistics, biological systems, and a wide array of other applications. As a result of their frequent occurrence, these processes have been studied extensively and a wealth of theory exists to solve problems related to these processes.

Definition 2.1 A stochastic process $\{X_t : t \in T\}$ constitutes a Markov process if for all $t_0 < t_1 < \cdots < t_n < t_{n+1}$ and all $s_i \in S$ the conditional CDF of $X_{t_{n+1}}$, depends only on the last previous value X_{t_n}, and not on the earlier values $X_{t_0}, X_{t_1} \ldots, X_{t_{n-1}}$:

$$Pr\left(X_{t_{n+1}} \leq s_{n+1} \left| X_{t_n} = s_n, X_{t_{n-1}} = s_{n-1}, \ldots, X_{t_0} = s_0\right.\right) = Pr\left(X_{t_{n+1}} \leq s_{n+1} \left| X_{t_n} = s_n\right.\right) \quad (2.1)$$

Notice that in Markov process, given the present, the future is conditionally independent of the past. Hence, the critical variables of a Markov process is the current state X_{t_n}.

Definition 2.2 A Markov chain is a sequence of random variables $\{X_t : t \in T\}$ for all $t_0 < t_1 < \cdots < t_n < t_{n+1}$ and all $s_i \in S$ with the Markov property, namely that, given the present state, the future and past states are independent. Mathematically,

$$Pr\left(X_{t_{n+1}} = s_{n+1} \left| X_{t_n} = s_n, X_{t_{n-1}} = s_{n-1}, \ldots, X_{t_0} = s_0\right.\right) = Pr\left(X_{t_{n+1}} = s_{n+1} \left| X_{t_n} = s_n\right.\right) \quad (2.2)$$

2.2.2 State Transition Probabilities

It is considered that Markov chains with discrete state spaces only, however, the state transitions can be discrete or continuous time based, respectively. The basic requirement for defining a Markov chain is to specify the probability of determining the next transition state for each state in the process and for each transition time.

Definition 2.3 Let X_t be a Markov chain with states $(s_0, s_1 \ldots, s_n)$, then the probability of transition from state s_i to state s_j in one time instant is

$$P_{s_i,s_j} = Pr\left(X_{t+1} = s_j \left| X_t = s_i\right.\right) \quad (2.3)$$

Then, the transition probability matrix can be defined as,

$$\mathbf{P} = \begin{bmatrix} P_{s_1,s_1} & P_{s_1,s_2} & \cdots & P_{s_1,s_n} \\ P_{s_2,s_1} & P_{s_2,s_2} & \cdots & P_{s_2,s_n} \\ \vdots & \vdots & \vdots & \vdots \\ P_{s_n,s_1} & P_{s_n,s_2} & \cdots & P_{s_n,s_n} \end{bmatrix} \quad (2.4)$$

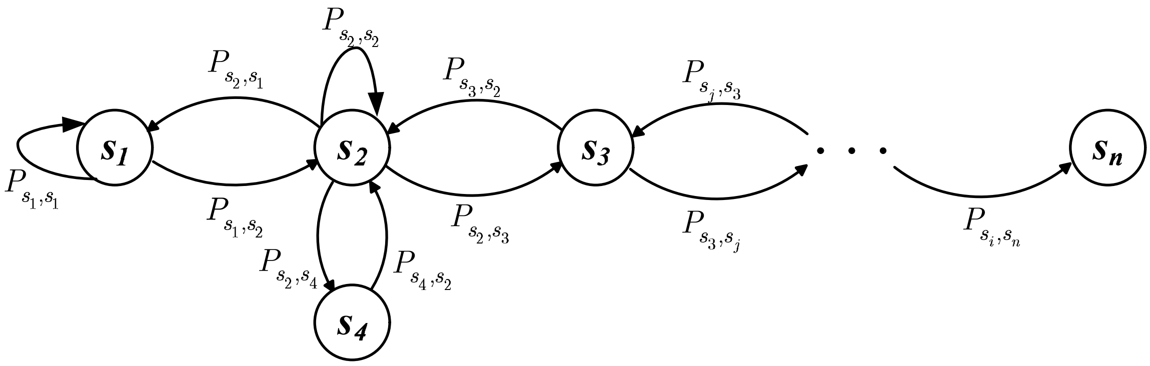


Figure 2.2: An example of state diagram for the Markov chain.

where $0 \leq P_{s_i,s_j} \leq 1$ for $1 \leq i, j \leq n$ and $\sum_{j=1}^{n} P_{s_i,s_j} = 1$ for $i = 1, 2 \ldots, n$. Since the rows of the transition probability matrix sum to one, $n(n-1)$ parameters are necessary to specify the transition probabilities behavior of a n state Markov chain.

The transition process of a Markov chain can also be graphically represented by a state diagram. An example of such a diagram is illustrated in Figure 2.2. In the figure, each directed arrow represents a possible transition and the label on each arrow represents the probability of making that transition.

Define $\mathbf{\Pi}_t$ as the state probability distribution at time t, respectively. Specifically:

$$\mathbf{\Pi}_t = [\Pr(s_1, t), \Pr(s_2, t), \ldots, \Pr(s_n, t)] \tag{2.5}$$

in which $\Pr(s_i, t)$ $(i = 1, 2, \ldots, n)$ represents the probability of finding the model in the state s_i at time t. By the definition of the state transition probability matrix (2.4), the state distribution at time $t + 1$ is then given by,

$$\mathbf{\Pi}_{t+1} = \mathbf{\Pi}_t \mathbf{P} \tag{2.6}$$

Therefore, in general,

$$\mathbf{\Pi}_{t+k} = \mathbf{\Pi}_t \mathbf{P}^k \tag{2.7}$$

where $\mathbf{P}^k$ represents the k step transition matrix.

Markov processes settle to a steady-state probability distribution as time evolves. Assuming that the Markov process converges to a steady state value distribution $\mathbf{\Pi}_{t+k} = \mathbf{\Pi}_t$ for sufficiently large t and arbitrary k. Hence, by solving the linear equations $\mathbf{\Pi} = \mathbf{\Pi P}$ and $\mathbf{\Pi e} = 1$ (where $\mathbf{e}$ is the column vector with all ones), the steady state probabilities can be obtained for the Markov chain models.

2.3 Literature Review

DSS is the key component of DSA, and is responsible for providing efficient and fair spectrum allocation or scheduling solutions among primary and secondary users. To do so, there are generally two main approaches: coordinated sharing and uncoordinated sharing.

2.3.1 Coordinated Sharing

In the coordinated sharing model (spectrum property rights or spectrum leasing) [9], PUs are the owners of the spectrum, which may not utilize their spectrum resources at all times or geographical areas. As such, PUs may wish to grant the so-called SUs with the spectrum access rights, in exchange for some revenue. Such spectrum leasing is coordinated via a centralized spectrum manager [12; 18], and thus users' quality-of-service (QoS) can be well coordinated. However, coordinated sharing is likely to require considerable hardware upgrades for the primary systems to accommodate SUs operating in their licensed spectrum.

In recent studies for coordinated DSS, channel borrowing schemes have been proposed for cellular networks [19]. However, this sharing scheme can only be implemented within the same network. When the spectrum is shared among multiple radio networks [20], a centralized spectrum manager may be needed. In this case, various studies have been conducted, including cooperation between PUs and SUs in a game theoretic model [18], the study of access fairness among different users [21], and admission control mechanisms [22].

For coordinated spectrum sharing, CAC can be performed to regulate each user's admitted traffic [22]. In [23]–[25], guard channel strategies are employed to decrease the forced termination probability at the cost of slight increasing of the blocking probability, to achieve efficient utilization of spectrum and guarantee the SUs' GoS requirements. In [22], three spectrum admission control strategies and their performance are evaluated, to maximize the supported SUs traffic in the multi-radio sharing system.

2.3.2 Uncoordinated Sharing

In uncoordinated sharing, CR permits spectrum sharing without any hardware updates for the primary systems. With CR, SUs have awareness of the radio frequency environment, and can dynamically select the operating parameters (such as frequency, modulation scheme, and transmission power) to opportunistically access the spectrum holes (i.e., the unused frequency bands) without causing interference to the PUs.

In order to exploit spectrum holes efficiently while avoiding harmful interference to PUs, spectrum sensing becomes one of the most important and fundamental PHY layer tasks for CR. A number of spectrum sensing methods have been proposed, such as matched filters, cyclostationary feature detection [26; 27], entropy-based sensing [28], and energy detection [29; 30]. The energy detection method is simple to implement and has been extensively used when there is no prior knowledge of the PU signals. However, when experiencing fading, using only local observations at an SU for sensing may be inaccurate.

To overcome this, cooperative sensing methods [31; 32] have been proposed, in which spectrum observations from multiple SUs are incorporated to detect the presence of PUs. Optimal data fusion for cooperative sensing was studied in [33; 34]. However, with noise and fading, errors are inevitable in sensing and an SU may mistakenly conclude one channel being occupied by a PU when in fact the channel is available, leading to a false alarm event. On the other hand, an SU may fail to detect the presence of PUs, leading to a so-called miss-detection event. In the case of false alarm events, spectrum access opportunities are wasted, while in the events of miss-detection, more collisions are caused. The relationship between false alarm and miss-detection is captured by the receiver operating characteristics (ROC) of the SUs' sensing capability, and their tradeoff can be determined by the operating point [35]. The operating point selection in spectrum sensing at the PHY layer affects the MAC layer performance metrics such as blocking probability (BP) and interference probability (IP), which will worsen the GoS performance of the spectrum sharing users. As a result, a cross-layer design of the PHY spectrum sensing technique and the MAC spectrum sharing policy can be used for performance optimization. To this end, Gelabert *et al.* [36] considered the use of energy detection in their spectrum sensing model, and studied the influence of operating point selection to the aggregate GoS performance for the system. The optimal operating points selection is addressed to improve service quality for PUs and SUs. In [37], sensing parameters (i.e., sensing period) and fractional CAC parameters were jointly optimized to minimize the dropping rate of SUs.

2.3.3 Overlay and Underlay Sharing

Hierarchical overlay and underlay sharing schemes have been shown to achieve a good tradeoff between spectrum efficiency and interference control for DSA [9]. For overlay access, SUs perform spectrum sensing to obtain the current spectrum occupancy information and access the spectrum not used by PUs opportunistically. Reliable spectrum sensing algorithms, and optimal sensing-access mechanism have been widely addressed in [32; 38; 39]. For underlay access schemes, by constraining the transmission power of SUs and minimizing interference to PUs, SUs can access the spectrum even if PUs are active. Efficient power allocation and adaptive interference threshold constraint methods have been presented in [40; 41].

Interference tolerance of PUs is ignored for overlay access, and the opportunity for SUs to transmit at higher power when PUs are inactive is missed for underlay access. Therefore, mixed or hybrid strategies are desirable and have been proposed to improve spectrum efficiency. In [42], a hybrid CR system was proposed where the system is normally working in overlay mode but may switch to underlay mode based on a probabilistic control method for the departure rate of SU transmitters. Also, Chakravarthy *et al.* [43] designed a hybrid overlay-underlay waveform to enable soft decision in CR and to exploit both unused and under-used spectrum resources. Later, [44] further considered the interference- or power-constrained spectrum sharing system, and has proposed a mixed access strategy in which SUs attempt to transmit with a pre-defined probability to ensure the overall interference under certain threshold.

2.3.4 Markov Modeling

The Markov chain model has been widely used for modeling lower and upper layer problems in wireless networks [45]–[47]. For the MAC layer problem, users' access modeling in uncoordinated DSA have also received so much attention recently [50]–[55]. In particular, Xing *et al.* [49] analyzed the manipulation of wideband and narrowband users in open spectrum bands, and used a Markov chain model to predict the user behaviors. A similar work appears in [50] in which the blocking probability for the wideband and narrowband users is also evaluated. In [25], a continuous-time Markov chain model was presented to study the spectrum access of SUs with and without spectrum handoff capability. Zhang [51] later proposed and analyzed DSA schemes with SU buffering. Wong *et al.* [53] studied the delay performance of SUs in CR-based DSA systems with a finite user population. Gelabert *et al.* [36] more recently studied the aggregate GoS performance for the DSS system by a DTMC model. Chung *et al.* [20] proposed the Markov model for DSS based on spectrum level and call level analysis. In [21], a partial spectrum sharing policy was proposed and a CTMC model was established to study the performance metrics of the DSS system in the ideal case of perfect spectrum sensing. In [54], a 3-dimensional Markov chain model was used to model the access of real-time and non-real-time users and find optimal reservation channels. Ngatched *et al.* [55] more recently studied the primary channel assembling schemes and explored the impacts of imperfect sensing using a Markov model.

2.4 Chapter Conclusions

In this chapter, the basic concepts of DSA and CR were introduced. The models for DSA, i.e., Dynamic Exclusive Use Model, Open Sharing Model, and Hierarchical Access Model were analyzed. Due to the various application scenarios for future DSA networks, Open Sharing Model will be used in license-free wireless services (e.g., Wi-Fi), while a Dynamic Exclusive Use Model and Hierarchical Access Model will be used in license-controlled networks (e.g., cellular, Digital TV networks) for secondary access. Since a Hierarchical Access Model does not require hardware updates for current systems and is capable of controlling potential interference, and thus be treated as the most promising model. The basic concepts of Markov processes and Markov chains were introduced. Mathematical descriptions of the state transition matrix and the steady state probabilities were discussed.

Related work for different DSS strategies have been discussed. However, there are many problems regarding the design of future DSS networks that have not been well explored. For coordinated sharing networks, since SUs may have different QoS requirements, the coordination of their access process needs to be studied to enhance spectrum sharing. For an uncoordinated sharing model, the potential spectrum sharing between two licensed networks were ignored by the literature, and the advantages of PHY-MAC cross layer optimization in DSS were not well addressed and analyzed. Furthermore, interference tolerance for overlay and underlay sharing was ignored. Hence, in the following chapters, novel DSS strategies and modeling methods will be investigated to solve the above problems, which can be used for spectrum sharing enhancement in various scenarios of future wireless networks.

Chapter 3

Quality-of-Service-Aware Coordinated DSS: Prioritized Markov Model and Call Admission Control

In this chapter, the focus is on coordinated spectrum sharing for a DSA wireless network where PUs and hierarchical SUs are accommodated. In particular, this work consider multiple classes of SUs each with a specific GoS [1] constraint indicating their tolerance (e.g., based on some pricing models). Blocking probability (BP), interfering probability (IP) and forced termination probability (FTP) are adopted as the metrics for SUs' GoS whereas blocking probability and interfering probability are chosen as the GoS for PUs. A prioritized policy is proposed to achieve efficient spectrum sharing between PUs and SUs. A Markov chain model is used to analyze the overall system performance. Based on this model, the maximum admitted traffic of SUs (referred to as the system serving capa-

[1] The term "GoS" here refers to as the probability that a call is unsuccessful due to the blocking, forced termination, or interference.

bility for SUs), can be derived. Furthermore, two CAC strategies are devised to enhance the maximum admitted traffic for SUs. Moreover, the access fairness of the multiple classes of SUs are also investigated. To summarize, the main contributions can be given as follows:

- An analytic Markov chain model is developed to analyze the proposed prioritized DSS policy with GoS guarantees for PUs and multiple classes of hierarchical SUs.

- By considering various optimality criteria such as max-sum, max-min and proportional-fair, the fairness of spectrum sharing among multi-class SUs is studied.

- Two new CAC strategies are proposed to improve the admitted traffic for SUs.

3.1 System Model

Consider a wireless network which provides a set of channels over a given area, as shown in Figure 3.1. Here, the term “channel” indicates the basic unit of capacity, which could be a time-slot in time-division multiple-acess (TDMA), or a frequency channel in frequency-division multiple-acess (FDMA) systems. This network is the primary owner (referred to as the cellular network) of the spectrum, and its users are regarded as the PUs. At the same time, there are secondary wireless networks (referred to as the WLAN networks or individual ad-hoc users) operating in the same area, whose users are regarded as the SUs. It is assumed that the SUs share the same spectrum with the PUs managed by a centralized

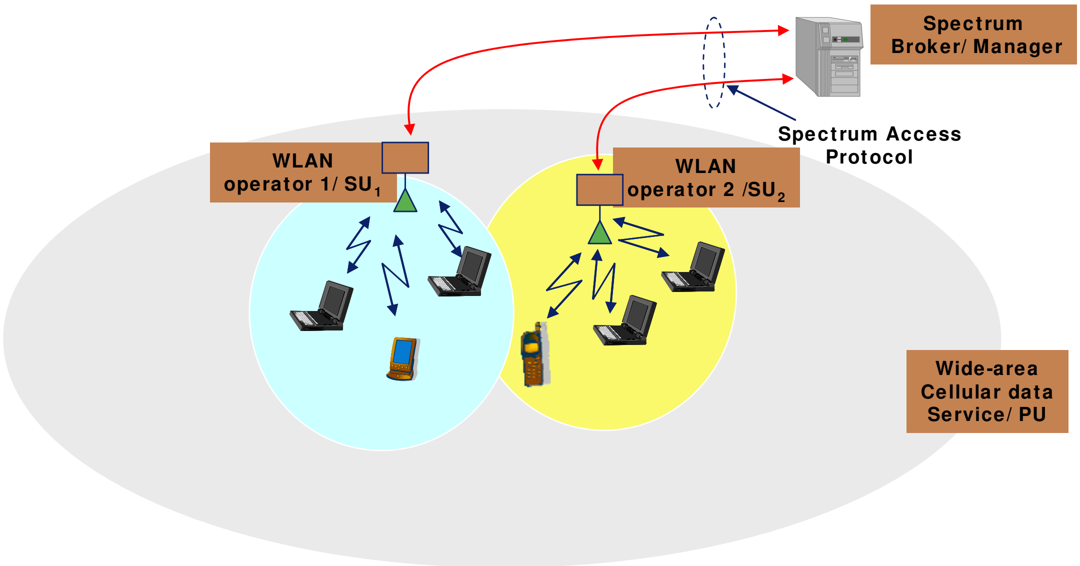


Figure 3.1: System Model.

spectrum manager/broker. This will be the scenario if the primary network leases whole or part of the spectrum to the secondary network to generate revenues. Depending on access pricing, there can be M classes of SUs (namely $\{\mathsf{SU}_m\}_{m=1}^M$) which are differentiated by their GoS constraints. In particular, it is assumed, without loss of generality, that SU_m has a more stringent GoS constraint than SU_n if $m < n$.

Under this model, the total number of orthogonal channels shared by the PUs and SUs is assumed to be N. The arrivals of PU and multiple classes of SUs are all assumed to follow a Poisson process with their respective mean arrival rates λ_{PU}, and $\{\lambda_{\mathsf{SU}_m}\}_{m=1}^M$. In what follows, their channel occupation times have a negative exponential distribution with the mean holding times $\frac{1}{\mu_{\mathsf{PU}}}$ and $\{\frac{1}{\mu_{\mathsf{SU}_m}}\}_{m=1}^M$, respectively. Hence, their traffic loads can be defined as $\tau_{\mathsf{PU}} \triangleq \frac{\lambda_{\mathsf{PU}}}{\mu_{\mathsf{PU}}}$

and $\{\tau_{\mathsf{SU}_m} \triangleq \frac{\lambda_{\mathsf{SU}_m}}{\mu_{\mathsf{SU}_m}}\}_{m=1}^{M}$, respectively. To make the analysis tractable, it is further assumed that every user regardless of PU and SU requires only a single channel for transmission. More general scenarios can be considered but this simplification preserves the essence of the problem.

3.2 Probabilistic Prioritized DSS

In this section, a heterogeneous-prioritized spectrum sharing policy is proposed and an analytical model for system performance evaluation is derived. Then, two new CAC strategies, which can be used to improve the GoS for both the PUs and SUs, are presented.

3.2.1 Heterogeneous-prioritized Markov Model

In coordinated DSS, according to the traffic loads and the GoS of PUs and SUs, the spectrum manager can have different priorities to different classes of SUs. Figure 3.2 shows the flow chart of the spectrum manager's actions in the proposed model. In this prioritized spectrum sharing policy, the spectrum manager processes the SUs' access as follows:

- When there is a vacant channel in the spectrum, if the channel being occupied by an SU, say SU_m, is needed by the PU, then the spectrum manager will stop the SU's transmission and allocate the channel to the PU in need, with probability α_m.
- In the case where the channel is taken over by the PU, the SU will hand off to a vacant channel if there is any or otherwise be terminated. Since this is a probabilistic action, the spectrum manager may delay the PU's

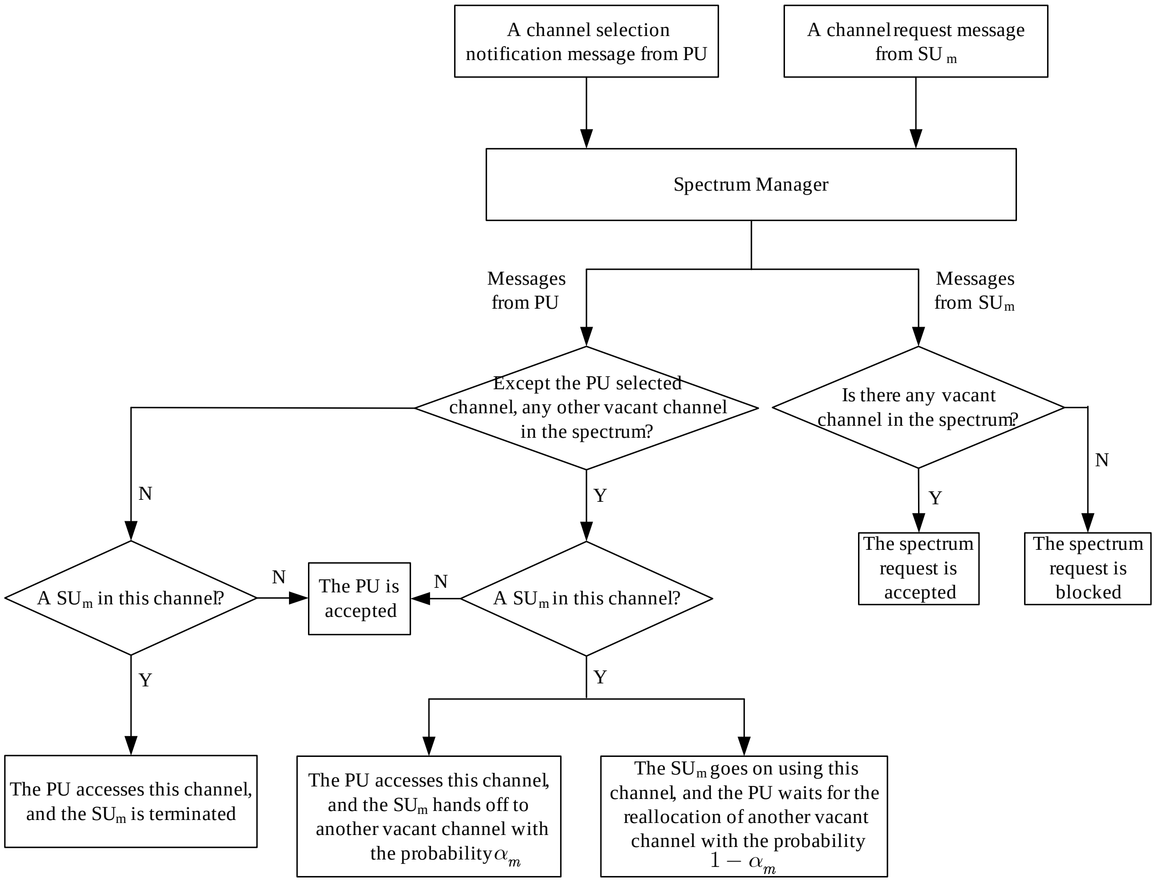


Figure 3.2: The flow chart of the spectrum manager's actions.

access request and reallocate another vacant channel for this PU, without terminating SU_m, which occurs with probability $1 - \alpha_m$. In this case, the PU is said to be interfered in terms of a service request delay.

- When the spectrum is full, if the channel being occupied by an SU_m is needed by the PU, SU_m will give up the channel (with probability 1), which will be used by the PU in need.

The delay time for interfering PUs or SUs are ignored in this model. It is assumed that there are M different types of SUs in the system, and thus $(M+1)$-

tuple $(i, j_1, \ldots, j_M)$ are used to denote the channel states in the system, where i, j_m $(m = 1, 2, \ldots, M)$ denote, respectively, the number of PUs and the number of SU_m operating in the system. The state space of the Markov model is given by

$$\mathcal{S} = \left\{ (i, j_1, j_2, \ldots, j_M) \;\middle|\; \begin{array}{l} 0 \leq i \leq N, j_m \geq 0, \\ \text{for } m = 1, 2, \ldots, M, \\ \text{and } \displaystyle\sum_{m=1}^{M} j_m \leq N - i \end{array} \right\}. \tag{3.1}$$

Let $\pi(i, j_1, j_2, \ldots, j_M)$ represent the steady state probability of the state $(i, j_1, j_2, \ldots, j_M)$, and $\Upsilon_{(i', j'_1, j'_2, \ldots, j'_M)}^{(i, j_1, j_2, \ldots, j_M)}$ denote the state transition rate from $(i, j_1, j_2, \ldots, j_M)$ to $(i', j'_1, j'_2, \ldots, j'_M)$. The state transitions from the state $(i, j_1, j_2, \ldots, j_M)$ are triggered by the following events.

(i) A PU request arrives but no SU is terminated: Upon a PU call request, if the number of active PUs is less than N, the PU's request is accepted, and no SUs will be terminated.

Although no SUs are terminated, it may still cause the *interfering events* of the PU and SUs. An interfering event is defined as the access of PUs or SUs being delayed due to access collision. In particular, when the channel being occupied by an SU_m is needed by the PU, the spectrum manager will stop the SU's transmission with probability α_m, interfering SU_m. Similarly, the spectrum manager may delay the PU's request with probability $1-\alpha_m$, interfering the PU. Hence, the state transition may arise from one of the following: (a) A PU arrives and selects a vacant channel, with probability $\frac{N-i-\sum_{m=1}^{M} j_m}{N-i}$; (b) a PU arrives and selects the channel being occupied by SU_m, then SU_m (or PU) is interfered with probability α_m (or $1 - \alpha_m$), and hands off to another vacant channel for transmission. Hence, this case happens with probability $\frac{\alpha_m j_m}{N-i}$ (resp. $\frac{(1-\alpha_m)j_m}{N-i}$). Therefore, this event gives rise to the state transition to $(i+1, j_1, j_2, \ldots, j_M)$ with the transition rate

$$\Upsilon_{(i+1,j_1,j_2,\ldots,j_M)}^{(i,j_1,j_2,\ldots,j_M)} = \lambda_{\mathsf{PU}}. \tag{3.2}$$

(ii) A PU request arrives, causing an SU to terminate: This occurs when the spectrum is full (i.e., $i+\sum_{m=1}^{M} j_m = N$) and a PU request arrival will cause an SU to terminate. Therefore, the termination probability of SU_m is $\frac{j_m}{N-i}$ and the state is transited to $(i+1, j_1, \ldots, j_m-1, \ldots, j_M)$ with the transition rate

$$\Upsilon_{(i+1,j_1,\ldots,j_m-1,\ldots,j_M)}^{(i,j_1,j_2,\ldots,j_M)} = \frac{j_m}{N-i}\lambda_{\mathsf{PU}}. \tag{3.3}$$

(iii) The service of a PU is completed: This occurs when a PU finishes its service, and thus the state is transited to $(i-1, j_1, j_2, \ldots, j_M)$ with the transition rate

$$\Upsilon_{(i-1,j_1,j_2,\ldots,j_M)}^{(i,j_1,j_2,\ldots,j_M)} = i\mu_{\mathsf{PU}}. \tag{3.4}$$

(iv) An SU request arrives: The access request of SU_m is accepted when the spectrum is not full. The state is transited to $(i, j_1, \ldots, j_m+1, \ldots, j_M)$, with the state transition rate

$$\Upsilon_{(i,j_1,\ldots,j_m+1,\ldots,j_M)}^{(i,j_1,j_2,\ldots,j_M)} = \lambda_{\mathsf{SU}_m}. \tag{3.5}$$

(v) The service of an SU is completed: This occurs when SU_m finishes its service, and thus the state is transited to $(i, j_1, \ldots, j_m-1, \ldots, j_M)$ with the transition rate

$$\Upsilon_{(i,j_1,\ldots,j_m-1,\ldots,j_M)}^{(i,j_1,\ldots,j_m,\ldots,j_M)} = j_m\mu_{\mathsf{SU}_m}. \tag{3.6}$$

Let $\mathbf{\Pi}$ denote the row vector of steady state probabilities with elements $\pi(i, j_1, j_2, \ldots, j_M)$, and $\mathbf{Y}$ denote the transition rate matrix, whose elements $\Upsilon_{(i', j_1', j_2', \ldots, j_M')}^{(i, j_1, j_2, \ldots, j_M)}$ can be obtained from the analysis above. The steady state matrix $\mathbf{\Pi}$ can be obtained by solving the linear equations $\mathbf{\Pi} = \mathbf{\Pi}\mathbf{Y}$ and $\mathbf{\Pi}\mathbf{e} = 1$, where $\mathbf{e}$ is the column vector with all ones.

3.2.2 Proposed CAC Strategies

Since the transmission of an SU could be interfered or terminated by a PU arrival, the GoS of SU cannot be guaranteed. Similarly, a PU could also be interfered in such heterogeneous-prioritized sharing policy, and the protection for the GoS of PU should also be enhanced as well. To achieve this, two *new* CAC strategies are proposed which shall be referred to as guard channels and restrict functions (GCRES) and guard channels and restricted channels (GCRC), respectively. The objective is to enhance the maximum admitted traffic of SUs that the spectrum sharing system with limited resources can be supported. For the sake of analysis, simplifying the above $(M+1)$-dimensional Markov chain model by setting $M = 2$, i.e., considering the scenario with only two classes of SUs (SU_1 and SU_2), in which SU_1 has a more stringent GoS constraint than SU_2.

In GCRES, guard channels (GCs) are those reserved only for use of PUs, whereas shared channels (SCs) are available for all users. Furthermore, a restrict function is used to balance the resource allocation between the two types of SUs, by accepting a user's request in SU_2 with a probability that decreases linearly with the increasing number of users in SU_2. N_G denotes the number of GCs, and the linear restrict function $f(j_2)$ is defined as

$$f(j_2) = 1 - c_{\mathsf{RES}} j_2, \quad \text{for } 0 \leq c_{\mathsf{RES}} < \frac{1}{N - N_G - 1}. \tag{3.7}$$

Note that the restrict function is valid when the system still has the capacity to admit SU_2. The access scheme of GCRES is as follows:

- When the number of PUs occupying the spectrum, i, is less than N_G, the newly activating PU will choose a GC. If $i \geq N_G$, the activating PU will attempt to access an SC.
- A user in SU_1 can access the spectrum if there is a vacant SC, and the maximum number of supported users of SU_1 is $N - N_G$.
- A user in SU_2 can access the SCs with probability $f(j_2)$, and the maximum number of supported users of SU_2 is $N - N_G$.

In this scheme, GCs are used to decrease the chance for collisions between PUs and SUs. To improve the GoS of SU_1, specifically, the restrict function is chosen to decrease the blocking probability for users in SU_1 by intentionally blocking spectrum requests from SU_2.

In GCRC, the channels are divided into three groups, as depicted in Figure 3.3. Restricted channels (RCs) are reserved for SU_1 and cannot be used by SU_2, while the definition of GCs and SCs are the same as within GCRES. GCRC can be implemented by a central spectrum manager, who is capable of denying requests from SU_1 or SU_2. The numbers of GCs and RCs are denoted by N_G and N_R, respectively. The access scheme of GCRC is described as follows:

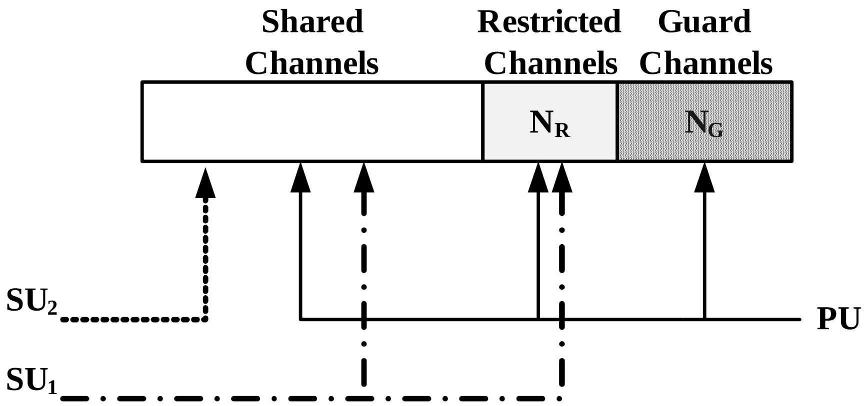


Figure 3.3: The GCRC strategy.

- When the number of PUs occupying the spectrum is less than N_G, the newly activating PU will choose a GC. If $i \geq N_G$ and $i < N - N_R$, the activating PU will access an SC. Finally, if $i \geq N - N_R$, then it will choose to access a RC.

- When the number of users in SU_1 occupying the spectrum, j_1, is less than N_R, the newly activating ones will choose a RC. If $j_1 \geq N_R$, the users choose to access SCs.

- SU_2 can only access an SC and the maximum number of supported SU_2 is $N - N_G - N_R$.

In this scheme, RCs are used to decrease the blocking probability, interfering probability and forced termination probability for the users in SU_1, and thus can provide better GoS guarantee to SU_1.

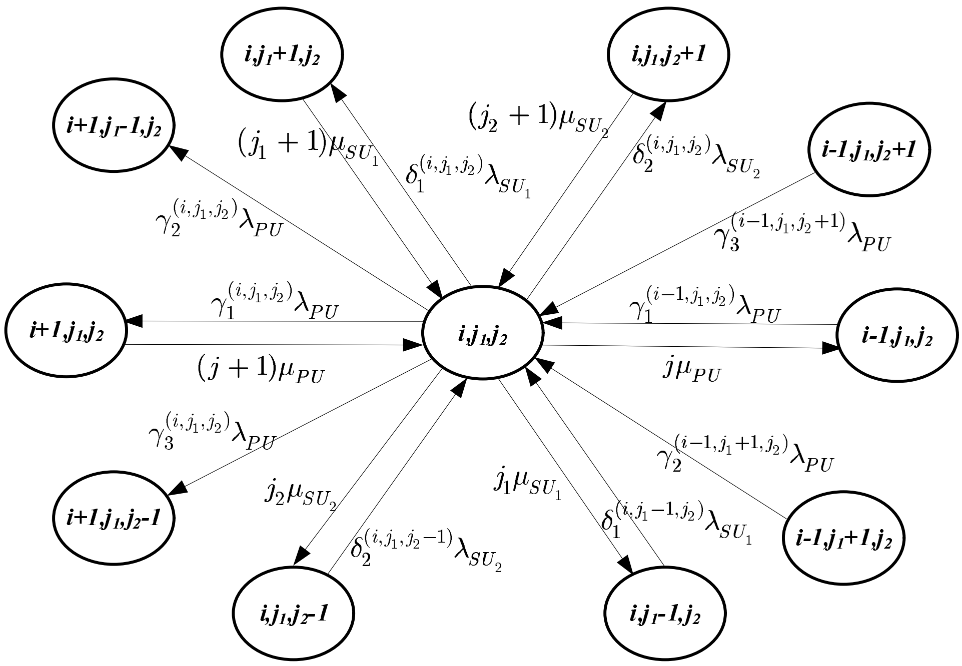


Figure 3.4: The state transition diagram of the heterogeneous-prioritized Markov model with CAC strategies. i, j_1, and j_2 denote, respectively, the number of PUs, SU_1, and SU_2 operating in the system.

3.2.3 Analysis

Adopting the 3-dimensional Markov chain, let $\pi(i, j_1, j_2)$ represent the steady state probability of the state (i, j_1, j_2). The state transition diagram for the heterogeneous-prioritized Markov model with CAC strategies can be derived, as shown in Figure 3.4. Let C_G, C_R and C_S denote, respectively, the numbers of used GCs, RCs and SCs. Then, (3.8) can be obtained.

$$(C_G, C_R, C_S) = \begin{cases} (i, j_1, j_2) & \text{for } 0 \le i \le N_G \text{ and } j_1 \le N_R, \\ (i, N_R, j_1 - N_R + j_2) & \text{for } 0 \le i \le N_G \text{ and } j_1 > N_R, \\ (N_G, j_1, i - N_G + j_2) & \text{for } N_G < i \le N - N_R \text{ and } j_1 \le N_R, \\ (N_G, N_R, i - N_G + j_1 - N_R + j_2) & \text{for } N_G < i \le N - N_R \text{ and } j_1 > N_R, \\ (N_G, i - (N - N_R) + j_1, N - N_G - N_R) & \text{for } N - N_R < i \le N. \end{cases} \tag{3.8}$$

The state space of the Markov model is then given by

$$\Omega = \left\{ (i, j_1, j_2) \left| \begin{array}{l} 0 \le i \le N, \\ 0 \le j_1 \le \min\left(N - N_G, N - i\right), \\ C_S \le N - N_G - N_R \end{array} \right. \right\}. \tag{3.9}$$

Before proceeding, it is useful to define the indicator functions $\{\gamma_k\}_{k=1}^{3}$ and $\{\delta_k\}_{k=1}^{2}$ to ensure that only feasible states are considered in the state transitions. The transitions from state (i, j_1, j_2) are now described as follows.

(i) To state $(i+1, j_1, j_2)$: This happens when a PU arrives and occupies a channel from the spectrum. The indicator function is $\gamma_1^{(i,j_1,j_2)} = 1$, if $i + j_1 + j_2 < N$, and 0 otherwise.

(ii) To state $(i+1, j_1 - 1, j_2)$: This happens when the spectrum is full, a PU arrives and a user in SU_1 is terminated. The indicator function $\gamma_2^{(i,j,k)}$ in this case is given by

$$\gamma_2^{(i,j_1,j_2)} = \begin{cases} \frac{j_1-N_R}{N-N_R-i} & \text{for} \begin{cases} N_G \leq i < N-N_R, \\ j_1 \geq N_R, \text{and } \mathcal{C} = N, \end{cases} \\ \frac{j_1}{N-i} & \text{for } N-N_R \leq i < N,\ \mathcal{C}=N, \\ 0 & \text{for } \mathcal{C} < N \text{ or } \mathcal{C}=N,\ j_1{=}0. \end{cases} \tag{3.10}$$

where $\mathcal{C}$ is defined as the total number of channels in use, i.e., $\mathcal{C} = C_G + C_R + C_S$.

(iii) To state $(i+1, j_1, j_2-1)$: This happens when both the GCs and SCs are full and a PU arrives. This will cause a user's transmission in SU_2's to be terminated. As such, the indicator function becomes $\gamma_3^{(i,j_1,j_2)} = \frac{j_2}{N-N_R-i}$, if $N_G \leq i < N-N_R$, $C_S = N-N_G-N_R$, and 0 otherwise.

(iv) To state (i, j_1+1, j_2): This happens when there are vacant channels in SCs or RCs, the incoming user in SU_1 will be accepted. In this case, the indicator function becomes $\delta_1^{(i,j_1,j_2)} = 1$, if $C_R + C_S < N - N_G$, and 0 otherwise.

(v) To state (i, j_1, j_2+1): This occurs when there are vacant channels in SCs, the incoming user in SU_2 is accepted with probability defined by the restrict function $f(j_2)$. The indicator function in this case will be $\delta_2^{(i,j_1,j_2)} = f(j_2)$, if $C_S < N - N_G - N_R$, and 0 otherwise.

(vi) To states $(i-1, j_1, j_2)$, (i, j_1-1, j_2) and (i, j_1, j_2-1): These cases happen, respectively, when PU, SU_1, and SU_2 complete their services successfully.

Let $\mathbf{\Pi}'$ denote the row vector of steady state probabilities with elements $\pi(i, j_1, j_2)$. From the above analysis and the transition diagram in Figure 3.4, the transition rate matrix $\mathbf{Y}'$ can be obtained. By solving the linear equations $\mathbf{\Pi}' = \mathbf{\Pi}'\mathbf{Y}'$ and $\mathbf{\Pi}'\mathbf{e} = 1$, the steady state probabilities $\pi(i, j_1, j_2)$ for the 3-dimensional Markov model using the GCRES and GCRC strategies can be obtained.

3.3 Derivation of Performance Metrics

In this section, performance metrics of the heterogeneous-prioritized spectrum sharing model with the proposed CAC strategies are presented, which can all be derived from $\pi(i, j_1, j_2)$.

3.3.1 Blocking Probability (BP)

For PUs and users in SU_1, blocking occurs when there are no channels available for newly arriving users. The blocking probability for PUs and users in SU_1 can be derived, respectively, as

$$P_{\mathsf{BP}}^{(0)} = \pi(N, 0, 0), \tag{3.11}$$

and

$$P_{\mathsf{BP}}^{(1)} = \sum_{\substack{(i,j_1,j_2)\in\Omega \\ C_R+C_S=N-N_G}} \pi\left(i, j_1, j_2\right). \tag{3.12}$$

Due to the inclusion of $f(j_2)$, the blocking probability for users in SU_2 should not only consider the case when there are no available channels for newly arriving users from SU_2, but also the blocking caused by this restrict function. Hence, the blocking probability for users in SU_2 can be derived as

$$P_{\mathsf{BP}}^{(2)} = \sum_{\substack{(i,j_1,j_2)\in\Omega \\ C_S = N - N_G - N_R}} \pi\left(i, j_1, j_2\right) + \sum_{\substack{(i,j_1,j_2)\in\Omega \\ C_S < N - N_G - N_R}} \left[1 - f(j_2)\right] \pi\left(i, j_1, j_2\right). \tag{3.13}$$

3.3.2 Interfering Probability (IP)

If there is a vacant channel and the PU's target channel is occupied by an SU in SU_m $(m = 1, 2)$, the PU (resp. SU_m) would wait for the reallocation of a new vacant channel with probability $1 - \alpha_m$ (resp. α_m). The interfering probability is defined as the probability that PU access or SU transmission is delayed by the spectrum manager when there is a channel collision. Firstly, three subspaces of Ω in which interfering events will occur. The subspaces Ω_1, Ω_2, Ω_3 are:

$$\Omega_1 = \left\{ (i, j_1, j_2) \,\middle|\, \begin{array}{l} (i, j_1, j_2) \in \Omega, \mathcal{C} < N, \\ N_G \leq i < N - N_R, j_1 > N_R \end{array} \right\}, \tag{3.14}$$

$$\Omega_2 = \left\{ (i, j_1, j_2) \,\middle|\, \begin{array}{l} (i, j_1, j_2) \in \Omega, \mathcal{C} < N, \\ i \geq N - N_R, j_1 \geq 1, \end{array} \right\}, \tag{3.15}$$

$$\Omega_3 = \left\{ (i, j_1, j_2) \,\middle|\, \begin{array}{l} (i, j_1, j_2) \in \Omega, N_G \leq i < N - N_R, \\ j_2 \geq 1, C_S < N - N_G - N_R \end{array} \right\}, \tag{3.16}$$

where the PU and users in SU_1 could be interfered if the state falls into the spaces Ω_1 and Ω_2, while the PU and users in SU_2 could be interfered if the state belongs to Ω_3.

The interfering probabilities for the PU, and the users in SU_1 and SU_2 can be derived, respectively, as

$$P_{\mathsf{IP}}^{(0)} = \frac{\left[\begin{array}{l} \sum_{(i,j_1,j_2)\in\Omega_1} \frac{(j_1-N_R)(1-\alpha_1)}{N-N_R-i}\pi\left(i,j_1,j_2\right)+ \\ \sum_{(i,j_1,j_2)\in\Omega_2} \frac{j_1(1-\alpha_1)}{N-i}\pi\left(i,j_1,j_2\right)+ \\ \sum_{(i,j_1,j_2)\in\Omega_3} \frac{j_2(1-\alpha_2)}{N-N_R-i}\pi\left(i,j_1,j_2\right) \end{array}\right]}{1-P_{\mathsf{BP}}^{(0)}}, \tag{3.17}$$

$$P_{\mathsf{IP}}^{(1)} = \frac{\left[\begin{array}{l} \sum_{(i,j_1,j_2)\in\Omega_1} \frac{(j_1-N_R)\alpha_1}{N-N_R-i}\pi\left(i,j_1,j_2\right)+ \\ \sum_{(i,j_1,j_2)\in\Omega_2} \frac{j_1\alpha_1}{N-i}\pi\left(i,j_1,j_2\right) \end{array}\right]\lambda_{\mathsf{PU}}}{\left(1-P_{\mathsf{BP}}^{(1)}\right)\lambda_{\mathsf{SU}_1}}, \tag{3.18}$$

and

$$P_{\mathsf{IP}}^{(2)} = \frac{\sum_{(i,j_1,j_2)\in\Omega_3} \frac{j_2\alpha_2}{N-N_R-i}\pi\left(i,j_1,j_2\right)\lambda_{\mathsf{PU}}}{\left(1-P_{\mathsf{BP}}^{(2)}\right)\lambda_{\mathsf{SU}_2}}. \tag{3.19}$$

3.3.3 Forced Termination Probability (FTP)

The transmission of an SU may be terminated by a PU, and the forced termination probability for SU is defined as the ratio between the total SU forced termination rate and the total SU connection rate. In particular, the forced termination probability for users in SU_1 and SU_2 can be derived, respectively, as

$$P_{\mathsf{FTP}}^{(1)} = \frac{\sum_{(i,j_1,j_2)\in\Omega} \gamma_2^{(i,j_1,j_2)}\lambda_{\mathsf{PU}}\pi\left(i,j_1,j_2\right)}{\left(1-P_{\mathsf{BP}}^{(1)}\right)\lambda_{\mathsf{SU}_1}}, \tag{3.20}$$

and

$$P_{\mathsf{FTP}}^{(2)} = \frac{\sum_{(i,j_1,j_2)\in\Omega} \gamma_3^{(i,j_1,j_2)}\lambda_{\mathsf{PU}}\pi\left(i,j_1,j_2\right)}{\left(1-P_{\mathsf{BP}}^{(2)}\right)\lambda_{\mathsf{SU}_2}}. \tag{3.21}$$

3.3.4 GoS

The GoS of an SU can be defined based on the blocking probability, interfering probability and forced termination probability. On the other hand, the GoS for a PU can be defined as the combination of the blocking probability and interfering probability. To exemplify their application, define the GoS as a normalized value [36], in which a larger GoS value can be explained as a tighter GoS constraint. The GoS of a PU as can be defined as

$$\mathsf{GoS}^{(0)} = \frac{P_{\mathsf{BP}}^{(0)} + \omega_0 P_{\mathsf{IP}}^{(0)}}{1+\omega_0}, \tag{3.22}$$

and the GoS of a user in SU_m as

$$\mathsf{GoS}^{(m)} = \frac{P_{\mathsf{IP}}^{(m)} + \omega_1 P_{\mathsf{BP}}^{(m)} + \omega_2 P_{\mathsf{FTP}}^{(m)}}{1+\omega_1+\omega_2}, \text{ for } m \in \{1,2\}\,, \tag{3.23}$$

where ω_k, $k \in \{0,1,2\}$ is the parameter to balance the impact of interfering, blocking, and forced termination. Note that ω_k should be empirical and depend on the subjective perception of the GoS metric [36]. For example, $\omega_2 > 1$ means that forced termination is more harmful with respect to interfering for SU_2's GoS.

3.3.5 Throughput of SUs

The throughput of SUs can be defined as the average number of completed services per unit time. It can be evaluated by

$$\eta_{\mathsf{SU}} = \sum_{(i,j_1,j_2)\in\Omega} (j_1+j_2)\pi\left(i,j_1,j_2\right). \tag{3.24}$$

3.3.6 Maximum Admitted Traffic of SUs

In the proposed heterogeneous-prioritized sharing policy, the goal is to determine the optimal choice of $\{\alpha_m\}$ and $\{\tau_{\mathsf{SU_m}}\}$ for $m \in \{1,2\}$, such that the system serving capability for SUs can be maximized, while satisfying the GoS constraints of the PU and SUs. The optimization problem can therefore be formulated as

$$\max_{\substack{\alpha_1,\alpha_2\\ \tau_{\mathsf{SU}_1},\tau_{\mathsf{SU}_2}}} \tau_{\mathsf{SU}} \text{ s.t. } \mathsf{GoS}^{(k)} \leq \mathsf{GoS}^{(k)*}, \text{ for } k \in \{0,1,2\}, \tag{3.25}$$

where $\tau_{\mathsf{SU}} = \tau_{\mathsf{SU}_1} + \tau_{\mathsf{SU}_2}$, and $\mathsf{GoS}^{(k)*}$ $(k \in \{0,1,2\})$ are the GoS constraints of PU, SU_1, and SU_2, respectively. The following three optimization criteria are considered for the system performance evaluation [56]:

- Max-Sum: This aims to maximize the sum of the admitted traffic of SUs (system's serving capability for SUs), i.e.,

$$\tau_{\mathsf{SU}}^{\mathsf{maxsum}} = \max_{\substack{\alpha_1,\alpha_2\\ \tau_{\mathsf{SU}_1},\tau_{\mathsf{SU}_2}}} \tau_{\mathsf{SU}_1} + \tau_{\mathsf{SU}_2}. \tag{3.26}$$

  In this criterion, the SU with a more stringent GoS constraint will be more difficult to admit, leading to unfairness.

- Max-Min: This aims to maximize the admitted traffic of the SUs with more stringent GoS (i.e., SU_1), so that

$$\tau_{\mathsf{SU}}^{\mathsf{maxmin}} = \max_{\substack{\alpha_1,\alpha_2\\ \tau_{\mathsf{SU}_1},\tau_{\mathsf{SU}_2}}} \min\left(\tau_{\mathsf{SU}_1}, \tau_{\mathsf{SU}_2}\right). \tag{3.27}$$

  For Max-Min criteria, the SU with a more stringent GoS constraint is optimized, i.e., the equal admitted traffic of SU_1 and SU_2 can be achieved, at the cost of inferior SU serving capability.

- Max-Proportional-Fair (Max-PF): In this scheme, the product of the admitted traffic of SUs is maximized, and the tradeoff between serving capability and fairness for SUs can be achieved. The optimization obejective under this criterion can be given as

$$\tau_{\mathsf{SU}}^{\mathsf{maxPF}} = \max_{\substack{\alpha_1,\alpha_2 \\ \tau_{\mathsf{SU}_1},\tau_{\mathsf{SU}_2}}} \tau_{\mathsf{SU}_1}\tau_{\mathsf{SU}_2}. \tag{3.28}$$

As a result, for the proposed GCRES and GCRC CAC strategies, the optimization problems can be formulated, respectively, as

$$\begin{aligned} &\max_{\substack{\alpha_1,\alpha_2,\tau_{\mathsf{SU}_1},\tau_{\mathsf{SU}_2} \\ N_G,c_{\mathsf{RES}}}} \tau_{\mathsf{SU}} \\ &\text{s.t. } \mathsf{GoS}^{(k)} \leq \mathsf{GoS}^{(k)*}, \text{ for } k \in \{0,1,2\}, \end{aligned} \tag{3.29}$$

and

$$\begin{aligned} &\max_{\substack{\alpha_1,\alpha_2,\tau_{\mathsf{SU}_1},\tau_{\mathsf{SU}_2} \\ N_G,N_R}} \tau_{\mathsf{SU}} \\ &\text{s.t. } \mathsf{GoS}^{(k)} \leq \mathsf{GoS}^{(k)*}, \text{ for } k \in \{0,1,2\}. \end{aligned} \tag{3.30}$$

The optimization problem (3.25) can be easily solved by the continuous PSO algorithm [57]. For the optimization problems (3.29) and (3.30), which contain the continuous and discrete variables. The optimal value of N_G and N_R can be obtained using an exhaustive search method, while the PSO algorithm is used to obtain the optimal values for the continuous variables, i.e., τ_{SU_1}, τ_{SU_2}, α_1, α_2 and c_{RES}. The details of solving these optimization problems can be found in Appendix B.

3.4 Numerical Results

Numerical results are obtained from the analytical model to evaluate the performance of the proposed schemes. For comparison, the results for first-come-first-served (FCFS) non-CAC scheme, GCRES, and GCRC CAC strategies can also be easily obtained by setting $N_G = N_R = 0, c_{\mathsf{RES}} = 0$; $N_R = 0$; and $c_{\mathsf{RES}} = 0$, respectively, in the analytical model. A MATLAB based discrete event simulation is set up to verify the analytical model. A large number (i.e., 10^6) PU and SU access requests were generated. The simulator accepts or denies the requests based on the access schemes described in the above section. The number of blocked, interfered, and forced terminated requests were recorded, and then were used to calculate the corresponding probabilities.

The flow chart of the discrete event simulator is defined as follows.

The flow chart of the discrete event simulator

Step 1: Initialize the number of channels N, the arrival rates of PU, SU_1, and SU_2 as λ_{PU}, and λ_{SU_1}, λ_{SU_2}, and their channel occupation times $\frac{1}{\mu_{\mathsf{PU}}}$, and $\frac{1}{\mu_{\mathsf{SU}_1}}$, and $\frac{1}{\mu_{\mathsf{SU}_2}}$, respectively. Set the start time of the simulation as $\mathsf{Current\_Time} = 0$;

Step 2: Generate the event times for the six types of events, i.e., the arrivals and departures of PU, SU_1, and SU_2, based on an exponential distribution;

Step 3: Sort these events times for the increasing order, and save to the $\mathsf{Event\_List[i]}$, where i is the index of event times, and set $i = 1$;

Step 4: Set $\mathsf{Current\_Time} = \mathsf{Event\_List[i]}$, update model status based on the type of event encountered and the access process defined in Section 3.2.2.

Step 5: If $\mathsf{Event\_List[i]} \leq \mathsf{Max\_Simulation}$, where $\mathsf{Max\_Simulation}$ is the maximum simulation time, set $i = i + 1$. Repeat **Step 4**. Otherwise, stop the simulator.

Step 6: Output the blocking, forced termination, and interfering probabilities.

The performance of the proposed heterogeneous-prioritized spectrum sharing policy will be studied first, and compared with the common sharing [23; 49; 50] and hierarchical sharing [24; 25] policies. Then, the performance of the proposed GCRES and GCRC CAC strategies is compared with the FCFS non-CAC and conventional GC-only CAC schemes [24; 25]. If not otherwise specified, it is assumed that $N = 10$ and $\mu_{\mathsf{PU}} = \mu_{\mathsf{SU}_1} = \mu_{\mathsf{SU}_2} = 0.5$ users per unit time. For the balancing weights in GoS calculation, since interfering events only cause the PU or SU waiting for the reallocation of the channels, thus it is assumed that blocking and forced termination events have larger impacts on users' GoS, as $\omega_0 = 1$, $\omega_1 = 2$, $\omega_2 = 5$. The GoS constraints for PU, SU_1, and SU_2 are defined as $\mathsf{GoS}^{(0)*} = 0.02$, $\mathsf{GoS}^{(1)*} = 0.05$ and $\mathsf{GoS}^{(2)*} = 0.08$, respectively. These values are empirical and depend on the subjective perception of GoS for each user type [36]. The analytical framework is practicable for various traffic parameters. The settings of simulation parameters are summarized in Table 3.1.

3.4.1 Performance for Heterogeneous-prioritized Sharing Policy

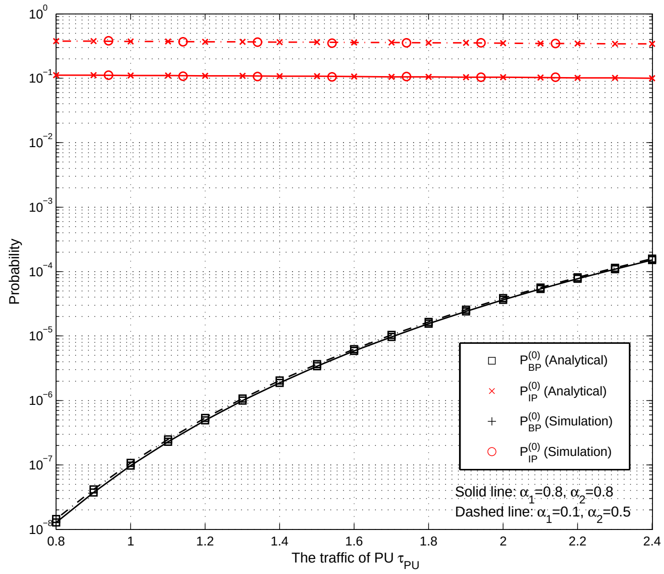
Figure 3.5 shows the results for the blocking, interfering, and forced termination probability for the PUs and SUs against the traffic of PU, τ_{PU}, with $\lambda_{\mathsf{SU}_1} = 2.0$ and $\lambda_{\mathsf{SU}_2} = 2.5$. The solid curve and dashed curve denote the cases with various SUs' priorities, $\alpha_1 = \alpha_2 = 0.8$ and $\alpha_1 = 0.1, \alpha_2 = 0.5$, respectively. As we can see in Figure 3.5(a), simulation results coincide with the analytical results and the PU blocking probability does not change with α_1 and α_2. On the other hand, with a

Table 3.1: Simulation Parameters.

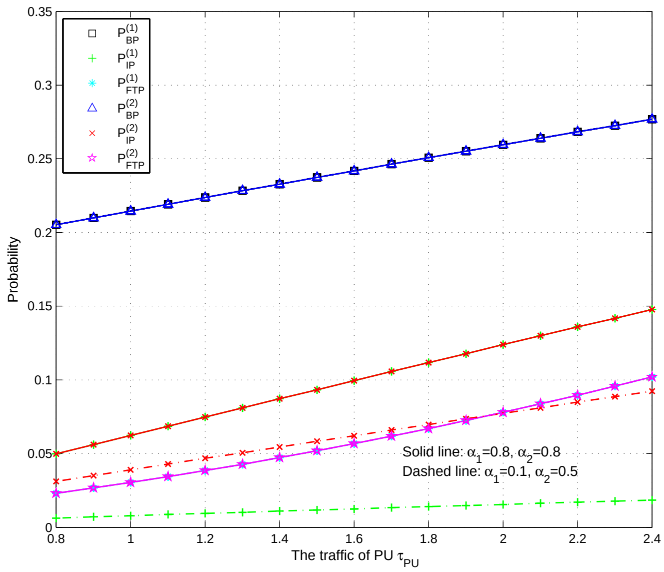
Arrival rate of PU (users per unit time)	$0.4 \leq \lambda_{\mathsf{PU}} \leq 1.2$
Arrival rates of SU_1 and SU_2 (users per unit time)	$\lambda_{\mathsf{SU}_1} = 2.0$, $\lambda_{\mathsf{SU}_1} = 2.5$
Departure rates (users per unit time)	$\mu_{\mathsf{PU}} = \mu_{\mathsf{SU}_1} = \mu_{\mathsf{SU}_2} = 0.5$
Weight factor of interfering and blocking for GoS of PU	$\omega_0 = 1$
Weight factor of interfering and blocking, forced termination for GoS of SU	$\omega_1 = 2$ $\omega_2 = 5$
GoS constraints	$\mathsf{GoS}^{(0)*} = 0.02$, $\mathsf{GoS}^{(1)*} = 0.05$, $\mathsf{GoS}^{(2)*} = 0.08$

larger value of α_1 and α_2, a lower PU interfering probability, $P_{\mathsf{IP}}^{(0)}$, can be achieved. In Figure 3.5(b), for both cases with $\alpha_1 = \alpha_2$ and $\alpha_1 \neq \alpha_2$, SU_1 and SU_2 have the same blocking probability and forced termination probability performance. However, with a lower α_1 or α_2, a better $P_{\mathsf{IP}}^{(1)}$ or $P_{\mathsf{IP}}^{(2)}$ can be achieved, and such improvement will be more significant with the decreases of α_1 or α_2.

Figure 3.6(a) compares the results of the maximum admitted traffic of SUs achieved by the proposed heterogeneous-prioritized (labeled as "HP" in the figure), common and hierarchical sharing policies, in terms of the offered traffic of PU, τ_{PU}. The results of common and hierarchical sharing policies can be obtained by using the analytical model with $\alpha_1 = \alpha_2 = 0$ and $\alpha_1 = \alpha_2 = 1$, respectively. For the proposed HP sharing policy, the maximum admitted traffic of SUs following different optimization objectives (i.e., Max-Sum, Max-Min, Max-PF) can be obtained by solving the optimization problems (3.25)–(3.28). It is observed that the proposed HP sharing can admit more SU traffic under all three optimization criteria, as compared with common and hierarchical sharing policies.



(a)



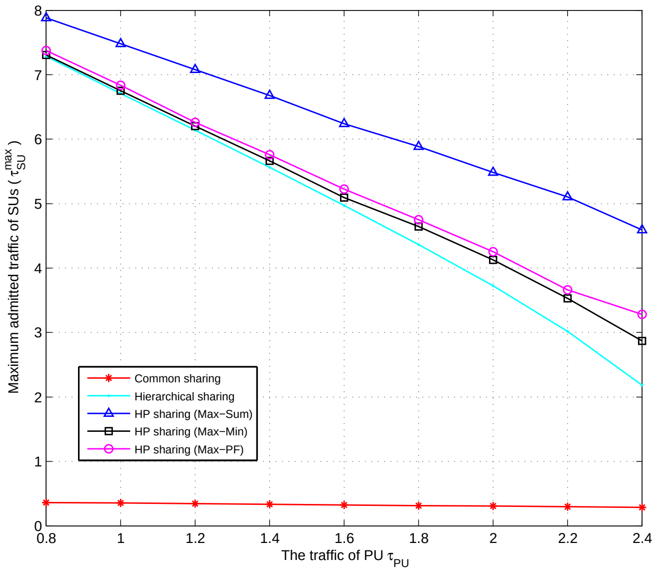
(b)

Figure 3.5: The results for the (a) PU blocking probability $P_{\mathsf{BP}}^{(0)}$ and interfering probability $P_{\mathsf{IP}}^{(0)}$, and (b) SU blocking probability $P_{\mathsf{BP}}^{(1)}$, $P_{\mathsf{BP}}^{(2)}$, interfering probability $P_{\mathsf{IP}}^{(1)}$, $P_{\mathsf{IP}}^{(2)}$, and forced termination probability $P_{\mathsf{FTP}}^{(1)}$, $P_{\mathsf{FTP}}^{(2)}$ against the traffic of PU τ_{PU} ($\tau_{\mathsf{PU}} = \lambda_{\mathsf{PU}}/\mu_{\mathsf{PU}}$). α_1 and α_2 denote the priorities of SU_1 and SU_2, respectively, which can be defined as the probabilities that SU_1 and SU_2 hand off to another vacant channel when they has collision with a PU.

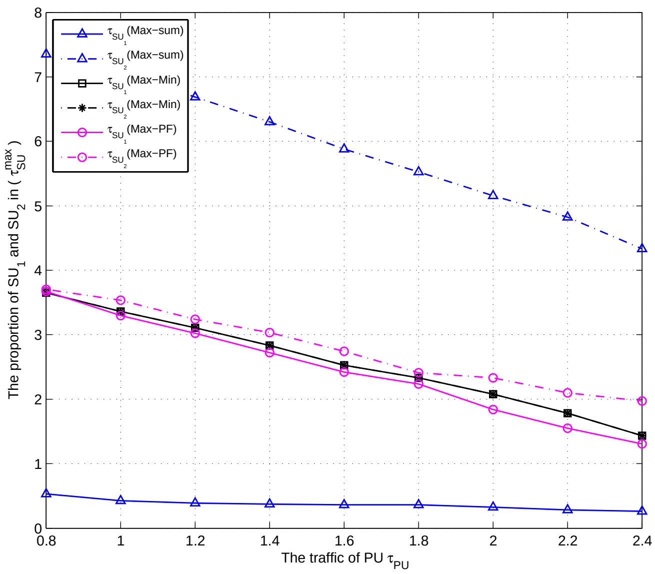
Figure 3.6(b) further compares the proportion of SU_1 and SU_2 traffic in the total admitted traffic of SUs for HP sharing, following the three optimization criteria. Since SU_1 has a more stringent GoS constraint, for the Max-Sum criterion, τ_{SU_2} is much greater than τ_{SU_1}, which is unfair for different classes of SUs. For Max-Min criterion, it achieves the best fairness between τ_{SU_1} and τ_{SU_2}, i.e., τ_{SU_1} equals to τ_{SU_2}, at the cost of the greatest loss of total admitted traffic for SUs. The Max-PF scheme is a tradeoff of the two which can achieve better fairness than Max-Sum, and has better spectrum utilization (SU serving capability) than Max-Min, and is therefore a better approach.

3.4.2 Performance for GCRES and GCRC CAC

Figure 3.7 compares the proposed CAC strategies with the FCFS non-CAC and the GC-only CAC schemes under Max-Sum criterion, for the maximum admitted traffic of SUs. For FCFS non-CAC and the GC-only CAC results can be obtained by setting $N_G = N_R = 0$ and $N_R = 0$ in the proposed analytical model, respectively. The results are obtained by solving problems (3.26), (3.29) and (3.30) by the PSO algorithm. In this figure, it is observed that the GCRES and GCRC achieve better SU serving capability, as compared with FCFS and GC-only CAC. This can be explained as follows. In the considered scenario, PU and SU_1 have more stringent GoS constraints, and thus easier to violate their GoS constraints. In GCRES scheme, the restrict function $f(j_2)$ is used to block some traffic of SU_2 intentionally, and thus more resources can be allocated to SU_1, leading to a better GoS performance of SU_1. While in GCRC scheme, the use of RCs will reduce the interfering probability for PU and the interfering probability, forced termination



(a)



(b)

Figure 3.6: (a) The maximum admitted traffic of SUs $\tau_{\mathsf{SU}}^{\mathsf{max}}$ for heterogeneous-prioritized (HP), common and hierarchical sharing policies, and (b) the proportion of SU_1 and SU_2 traffic in $\tau_{\mathsf{SU}}^{\mathsf{max}}$, against the traffic of PU τ_{PU}. The common and hierarchical sharing policies can be obtained by using the analytical model with $\alpha_1 = \alpha_2 = 0$ and $\alpha_1 = \alpha_2 = 1$, respectively. Max-Sum, Max-Min, Max-PF denote the optimization criteria, which can be obtained by solving the problems (3.25)–(3.28).

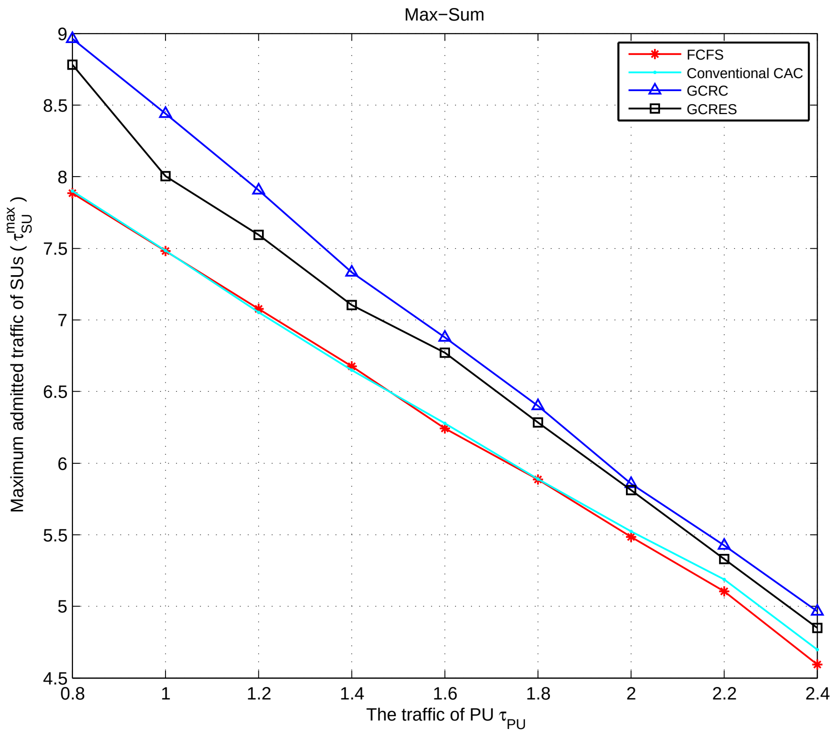
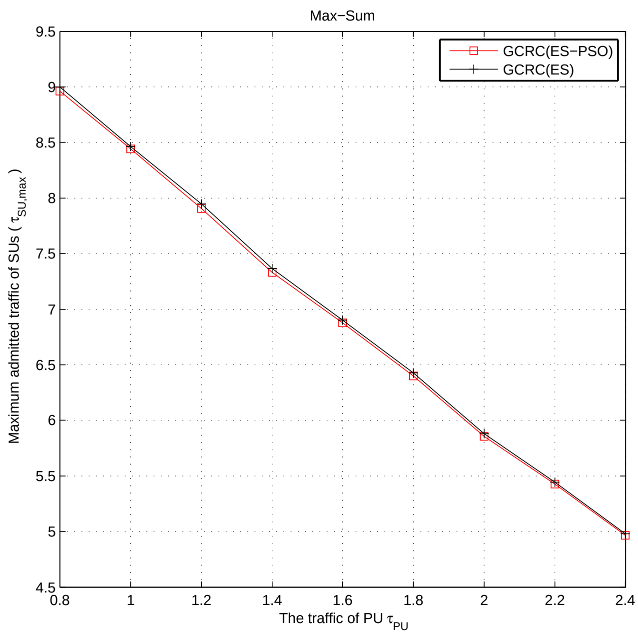


Figure 3.7: Maximum admitted traffic of SUs, $\tau_{\mathsf{SU}}^{\mathsf{max}}$, under Max-Sum criterion against the traffic of PU τ_{PU}. The proposed GCRC and GCRES can achieve better performance than FCFS non-CAC and the GC-only CAC schemes. The results are obtained by solving problems (3.26), (3.29) and (3.30) by the PSO algorithm.

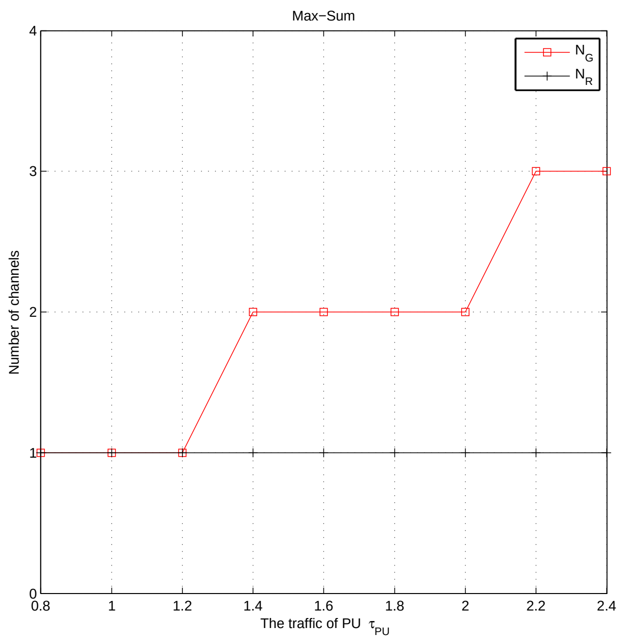
probability and blocking probability for SU_1, at the cost of the increasing the blocking probability and forced termination probability for SU_2.

Figure 3.8(a) compares achieved the maximum admitted traffic of SUs using exhaustive search (ES) only (ES for continuous and discrete variables, with the step size 0.01 for continuous variables) and the ES-PSO (ES for discrete variables, and the PSO for continuous variables) schemes. For the maximum admitted traffic of SUs, the performance obtained by the ES-PSO algorithm is nearly the same as the results obtained by ES. In Figure 3.8(b), the optimal number of guard channels (GCs) and restricted channels (RCs) for GCRC schemes are further presented. Notice that the optimal number of GCs increases with τ_{PU}, while the optimal number of RCs does not change. The reason is that with a larger τ_{PU}, more GCs should be used to guarantee the GoS constraint of PU. On the other hand, to admit more traffic of SUs in Max-Sum criterion, more resources should be allocated to SU_2 rather than SU_1. Since RCs are used to provide better GoS guarantee to SU_1, thus N_R cannot increase with τ_{PU}.

Figure 3.9 compares the proposed CAC strategies with the FCFS and the GC-only CAC schemes under Max-min and Max-PF criteria, for the maximum admitted traffic of SUs. From Figure 3.7 and Figure 3.9, it can be concluded that the GCRES and GCRC achieve better SU serving capability under all the three optimization criteria. Note that, this conclusion is also valid for different range of values for τ_{PU}. This is due to the fact that the GCRES and GCRC schemes can provide more flexible operating parameters, and thus the GC-only scheme performs as the lower bound of the proposed schemes. In particular, for the considered range of traffic for PU, $0.8 \leq \tau_{PU} \leq 2.4$, the proposed GCRES can achieve on average 16.4% and 6.8% admitted SUs' traffic gain over FCFS and

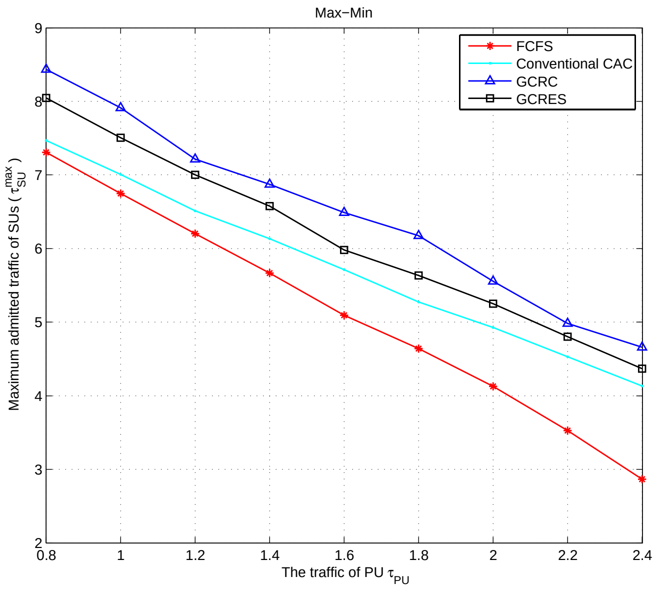


(a)

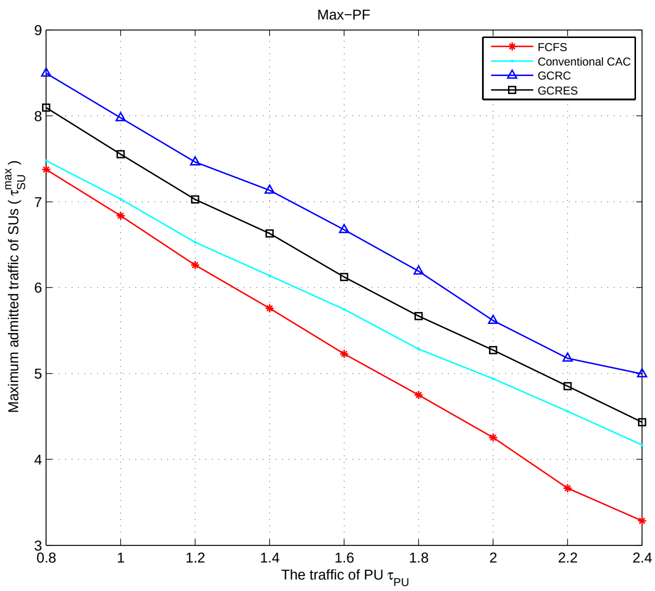


(b)

Figure 3.8: Maximum admitted traffic of SUs, $\tau_{\mathsf{SU}}^{\mathsf{max}}$ achieved by exhaustive search (ES) only and ES-PSO schemes under Max-Sum criterion. (b) The optimal number of guard channels (GCs) and restricted channels (RCs) for GCRC scheme.



(a)



(b)

Figure 3.9: Maximum admitted traffic of SUs, $\tau_{\mathsf{SU}}^{\mathsf{max}}$, under (a) Max-Min, and (b) Max-PF criteria, against the traffic of PU τ_{PU}. The proposed GCRC and GCRES can achieve better performance than FCFS non-CAC and the GC-only CAC schemes. The results are obtained by solving problems (3.26), (3.29) and (3.30) by the PSO algorithm.

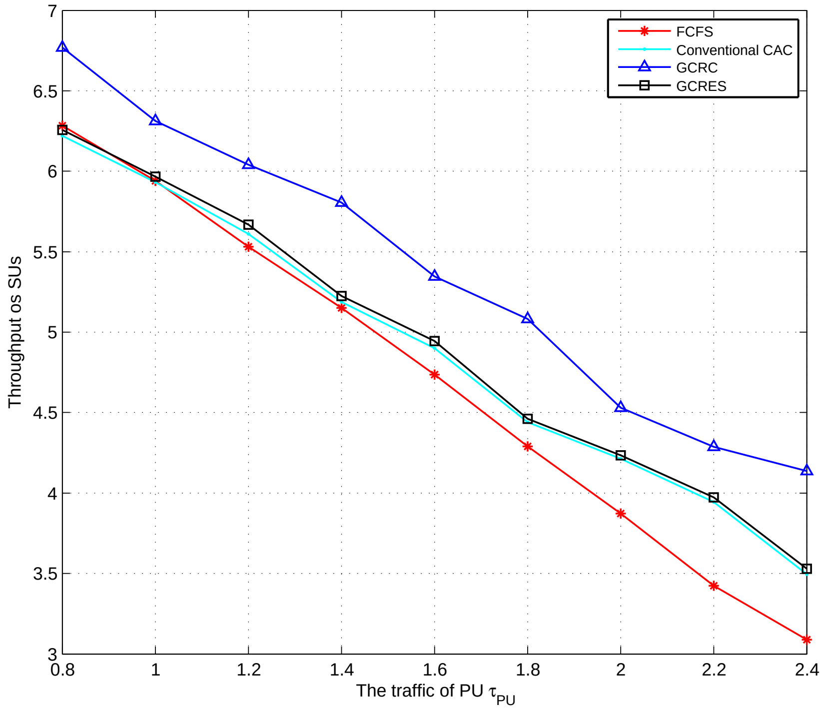


Figure 3.10: Throughput of SUs against the traffic of PU τ_{PU} under Max-PF criterion.

GC-only CAC, respectively. Furthermore, the GCRC achieves 22.8% and 12.5% improvement, respectively.

Since the Max-PF criterion achieves a better tradeoff between the serving capability and the fairness of SU_1 and SU_2, the throughput of SUs for different CAC strategies under PF criterion is also studied. In Figure 3.10, it is observed that the throughput of SUs decreases with τ_{PU}. The reason is that a higher PU traffic results in lower admitted traffic of SUs and lower SU completion probability. Also, it is observed that the proposed GCRC achieves the highest throughput of SUs and the performance gain over FCFS is 16.0%, and 10.3% over the conventional

GC-only CAC.

3.5 Chapter Conclusions

This chapter proposed a heterogeneous-prioritized spectrum sharing policy and model the interactions between PUs and multiple classes of SUs as continuous-time Markov chain models. The SU blocking probability (BP), interfering probability (IP), and forced termination probability (FTP), and PU interfering probability and blocking probability can be derived from the analytical models. In addition, two new guard channels and restrict functions (GCRES) and guard channels and restricted channels (GCRC) CAC strategies were proposed to enhance the system serving capability for SUs. Results illustrated that the proposed heterogeneous-prioritized sharing policy achieves higher maximum admitted traffic of SUs, as compared with common and hierarchical sharing. Moreover, the proposed GCRES and GCRC CAC strategies are superior to conventional CAC schemes under Max-Sum, Max-Min and Max-PF criteria. In particular, the proposed GCRC achieves the highest throughput of SUs, and the performance gain over FCFS non-CAC is 16.0%, and 10.3% over the conventional GC-only CAC. It is also found that the Max-PF fairness criterion achieves a better tradeoff between serving capability and fairness for multiple classes of SUs.

For coordinated DSS networks in this chapter, spectrum sharing between PUs and SUs is coordinated via a centralized spectrum manager, thus the users' QoS can be well coordinated. However, coordinated sharing is likely to require considerable hardware upgrades for the primary systems to accommodate SUs operating in their licensed spectrum. On the other hand, in uncoordinated sharing, CR permits spectrum sharing without any hardware updates for the primary systems, which can be flexible for DSS networks in some scenarios. Hence, in Chapter 4, new schemes for uncoordinated sharing models will be studied.

Chapter 4

Cross-layer Design of Uncoordinated Partial Spectrum-Shared Licensed Networks - A Continuous Time Markov Model

The recent literature [12; 18; 19; 21; 22; 25; 36; 37; 49; 51; 53] all focused on either coordinated or uncoordinated PU-SU sharing model, i.e., primary system opens its entire spectrum for SUs' access. The possible spectrum sharing between two types of licensed networks in the context of uncoordinated sharing, e.g., WiMAX and cellular networks, which is practical for future software-based and CR-equipped wireless devices, has not been well studied yet. Considering the benefits of no hardware updates for launching such uncoordinated sharing, current wireless service operators with licensed spectrum would have interests to apply opportunistic spectrum sharing scheme for enhancing the spectrum efficiency.

The main contributions of this work are summarized as follows:

1. A *new* partial sharing policy for licensed networks is proposed. Based on the traffic loads and the GoS requirements, the under-utilized network can decide on how much spectrum is shared by the over-utilized network.

2. A Markov chain model is developed to analyze the proposed partial sharing policy with imperfect sensing observations in terms of metrics such as blocking probability (BP), interfering probability (IP), forced termination probability (FTP), handoff probability (HP), and overall throughput.

3. A cross-layer optimization framework is designed, in which the parameters of spectrum sharing policy and spectrum sensing are jointly coordinated to maximize the overall throughput of the networks, while satisfying the GoS requirements of the users. In particular, the operating point (jointly defined by the normalized energy detection threshold and the voting threshold rules) which decides the tradeoff between miss-detection and false alarm probabilities for cooperative spectrum sensing will be optimized.

4.1 System Model

Consider two licensed wireless networks (Networks A and B) operating over a given service area and that both have respective licenses of spectrum usage locally, as well as their own base stations for network management and maintenance. The system model is shown in Figure 4.1. In the proposed model, it is assumed that Network A (N_A) is under-utilized, and its users are referred to as PU_A. On the contrary, Network B (N_B) is over-utilized, and its users are referred to as PU_B. In this chapter, consider that Network A grants temporary access rights for $\{\mathsf{PU}_B\}$

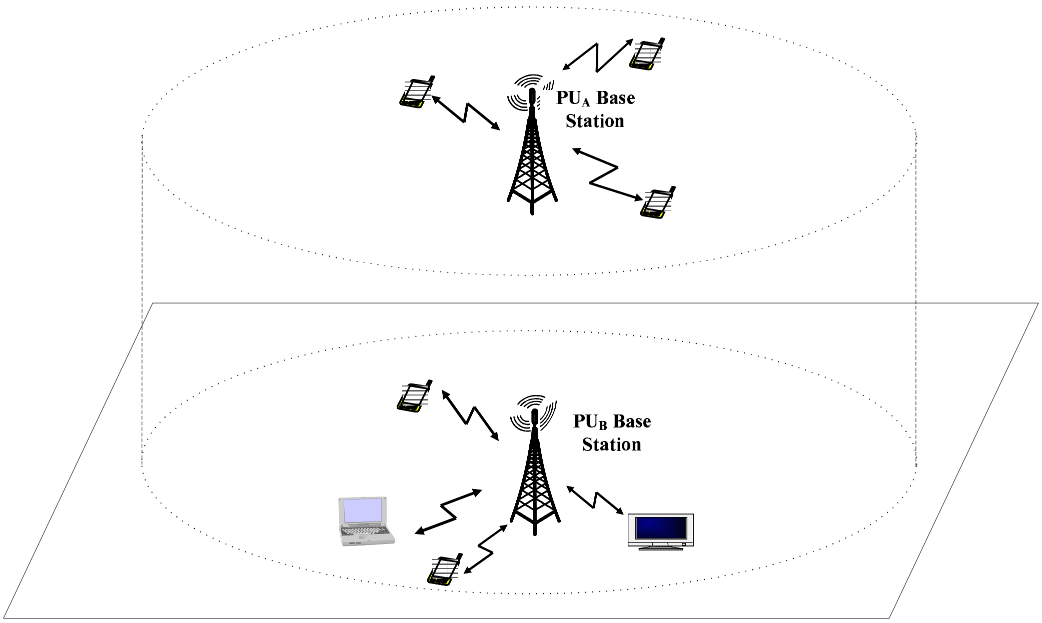


Figure 4.1: System Model.

in its licensed spectrum. Using CR technologies, PU_B are able to access some of N_A's channels[1] opportunistically.

4.1.1 Partial Spectrum Sharing Policy

Figure 4.2 depicts the proposed partial spectrum sharing policy between the two licensed networks. The pre-allocated numbers of channels for N_A and N_B in the static spectrum allocation policy are denoted by C_A and C_B, respectively, whereas the sets containing those respective channels are denoted as $\mathcal{C}_\mathcal{A}$ and $\mathcal{C}_\mathcal{B}$. In the proposed partial sharing policy, it is assumed that PU_A and PU_B will use their own licensed resources first, and channel renting happens only when needed. Note

[1]The term "channel" here refers to any basic unit of capacity, which could be a time-slot, a frequency channel, a tone, a code or a combination of those.

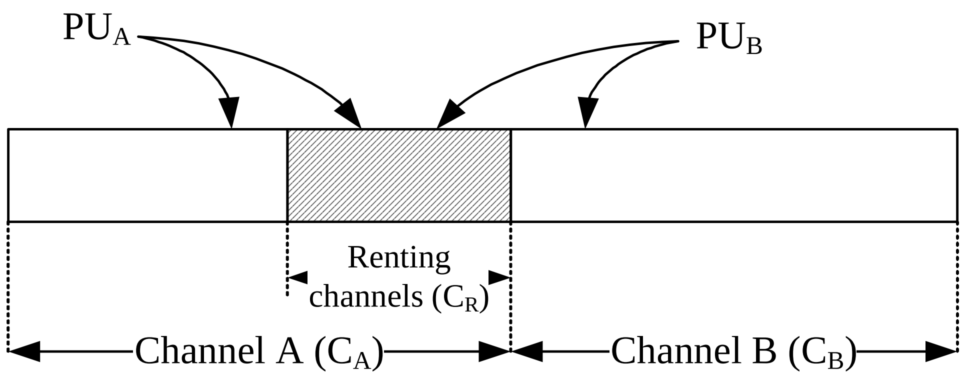


Figure 4.2: A partial spectrum sharing policy.

that this assumption requires less changes in the current spectrum policy, and is easier to be accepted by the service providers in a real world implementation.

According to the traffic loads and GoS requirements, N_A can decide on the number of channels (denoted by C_R and the set containing these channels by $\mathcal{C}_{\mathcal{R}}$) that can be rented by N_B. Here, the term "rent" means that N_A allows PU_B to opportunistically access parts of its channels in order to gain profit from the N_B rental of the channel. The spectrum sharing policy is described as follows.

- PU_A: For any PU_A's access request, if the number of PU_A is less than $C_A - C_R$, then the base station of N_A will allocate the non-renting channels to the upcoming PU_A. Otherwise, the renting channels in $\mathcal{C}_{\mathcal{R}}$ will be used.

- PU_B: For any PU_B's access request, the base station of N_B first allocates its licensed channels, i.e., the channels in $\mathcal{C}_{\mathcal{B}}$. If there is no vacant channels in $\mathcal{C}_{\mathcal{B}}$ available, N_B base station will coordinate cooperative spectrum sensing to explore the vacant channels in $\mathcal{C}_{\mathcal{R}}$. PU_A have access priority over PU_B for the channels in $\mathcal{C}_{\mathcal{R}}$. In the case that the total number of serviced PU_A users is greater than $C_A - C_R$, those PU_B operating in $\mathcal{C}_{\mathfrak{R}}$ may be interrupted by PU_A, forcing service termination of PU_B.

Note that this scheme takes advantages of both coordinated and uncoordinated sharing models. In particular, less modifications will be needed in N_A and N_B infrastructures, where a control channel is used by N_A and N_B base stations. Through the control channel, N_A base station can inform the renting channels to N_B, i.e., which channels can be sensed and rented by PU_B. Hence, the renting channels can be separate in the spectrum, and can change from time to time. Also, N_B base station will use the control channel to coordinate cooperative sensing among multiple PU_Bs in renting channels. Comparing with the coordinated sharing model, the N_A base station in the proposed policy does not allocate the renting channels to PU_B directly, thus needing less hardware updates. In comparison with the uncoordinated sharing model, i.e., N_A opens all its licensed channels for PU_B. The proposed policy can maintain the number of renting channels in real-time, and thus can achieve a better QoS and interference management for both PU_A and PU_B.

4.1.2 Cooperative Sensing with Voting Threshold

The overall performance of a Dynamic Spectrum Sharing (DSS) system depends on the sensing accuracy of CR-enabled PU_B. Using energy-detection-based spectrum sensing, the received energy over a fixed bandwidth W during an observation time window T is measured and the following hypothesis testing is established with the null hypothesis H_0 and the alternative hypothesis H_1:

$$H_0 : \mathsf{PU}_A \text{ is absent in the channel} \tag{4.1a}$$

$$H_1 : \mathsf{PU}_A \text{ is present in the channel} \tag{4.1b}$$

For single spectrum observation of PU_B (i.e., local sensing), when the spectrum sensing fails to detect the existence of PU, this can be referred to as a *miss detection* event. On the other hand, if the channel is vacant but the sensing mistakenly indicates that it is occupied by PU, this case is referred to as a false alarm event. The false alarm probability, P_{fa}, and detection probability P_{d} in Rayleigh fading environments are, respectively, [32]

$$P_{\mathsf{fa}} = \Pr\left\{Y > \lambda_{\mathsf{th}} \left| H_0 \right.\right\} = \frac{\Gamma\left(u, \frac{\lambda_{\mathsf{th}}}{2}\right)}{\Gamma\left(u\right)} \tag{4.2}$$

and

$$\begin{aligned} P_{\mathsf{d}} &= \Pr\left\{Y > \lambda_{\mathsf{th}} \left| H_1 \right.\right\} \\ &= e^{-\frac{\lambda_{\mathsf{th}}}{2}} \sum_{k=0}^{u-2} \frac{1}{k!} \left(\frac{\lambda_{\mathsf{th}}}{2}\right)^k + \left(\frac{1+\overline{\gamma}}{\overline{\gamma}}\right) \left(e^{-\frac{\lambda_{\mathsf{th}}}{2(1+\overline{\gamma})}} - e^{-\frac{\lambda_{\mathsf{th}}}{2}} \sum_{k=0}^{u-2} \frac{1}{k!} \left(\frac{\lambda_{\mathsf{th}}\overline{\gamma}}{2\left(1+\overline{\gamma}\right)}\right)^k\right), \end{aligned} \tag{4.3}$$

where Y and λ_{th} are the decision statistic and decision threshold of the energy detection, $\overline{\gamma}$ is the average signal-to-noise ratio (SNR), and $u = TW$ is referred to as the time-bandwidth product. In addition, $\Gamma\left(\cdot\right)$ and $\Gamma\left(\cdot, \cdot\right)$ are the complete and uncomplete gamma functions, respectively. The miss-detection probability can be computed by $P_{\mathsf{md}} = 1 - P_{\mathsf{d}}$.

To enhance sensing performance, multiple PU_Bs can be coordinated by the N_B base station to obtain spectrum observations simultaneously by sharing their

1-bit decisions (H_0 or H_1 in (4.1)) [31; 32]. Let K be the number of PU_B for cooperative sensing and assume that the K PU_Bs undergo independent and identically distributed (i.i.d.) Rayleigh fading with the same average SNR, $\overline{\gamma}$. An overall decision from all K PU_Bs can be obtained using voting threshold. In particular, the percentage rule in [61] is used, with the voting threshold $0 \leq \zeta \leq 1$, in which if more than $100\zeta\%$ of the cooperating PU_B report the presence of PU_A in the channel, then the final decision will be H_1. With the energy decision threshold, λ_{th}, for all PU_B, the false alarm and miss-detection probabilities are, respectively, given by [61]

$$Q_{\mathsf{fa}} = \sum_{l=\lceil \zeta K \rceil}^{K} \binom{K}{l} (1 - P_{\mathsf{fa}})^{K-l} P_{\mathsf{fa}}^{l}, \tag{4.4}$$

$$Q_{\mathsf{md}} = \sum_{l=0}^{\lceil \zeta K \rceil - 1} \binom{K}{l} (1 - P_{\mathsf{md}})^{l} P_{\mathsf{md}}^{K-l}, \tag{4.5}$$

where $\lceil \zeta K \rceil$ denotes the smallest integer that is no less than ζK.

The tradeoff between false alarm and miss-detection probabilities can be understood by the ROC curves, which plot Q_{md} against Q_{fa}. From (4.2) and (4.4), Q_{fa} is independent of $\overline{\gamma}$, but determined by λ_{th}, ζ and u. Once λ_{th} and ζ are decided, P_{md} and Q_{md} can be calculated for various u and $\overline{\gamma}$. By selecting λ_{th} and ζ, a pair of $(Q_{\mathsf{fa}}, Q_{\mathsf{md}})$ (referred to as an operating point) can be obtained. The normalized energy detection threshold can be defined as,

$$\Theta \triangleq \frac{\lambda_{\mathsf{th}}}{\lambda^{*}_{\mathsf{th}}}, \tag{4.6}$$

where $\lambda^{*}_{\mathsf{th}}$ can be obtained by inverting (4.2) when the minimum value for P_{fa} is achieved. The minimum P_{fa}, which can be referred to as the sensing resolution [36], is determined by the sensing equipment hardware characteristics. The results $(Q_{\mathsf{fa}}, Q_{\mathsf{md}})$ under different voting threshold rules against Θ are plotted in

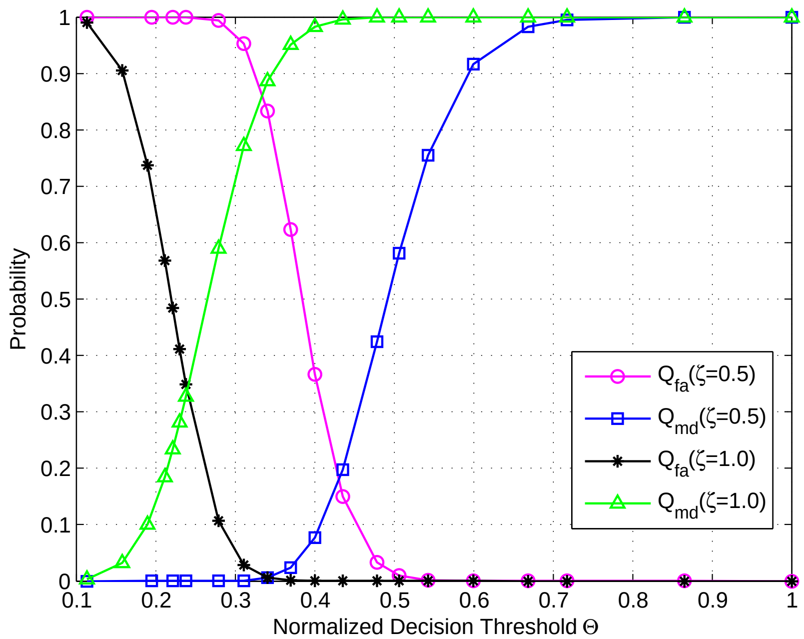


Figure 4.3: The false alarm and miss-detection probabilities against the normalized energy decision threshold Θ.

Figure 4.3. In this work, the fixed number of PU_Bs for cooperative sensing are defined, as $K = 10$, $u = 10$, and $\overline{\gamma} = 5$ dB. It is observed that with an increase in Θ decreases Q_{fa} but increases Q_{md}, showing a fundamental tradeoff between them. Moreover, under various voting thresholds ζ, different $(Q_{\mathsf{fa}}, Q_{\mathsf{md}})$ against Θ can be obtained, which indicates that there is a one-to-one correspondence between $(\lambda_{\mathsf{th}}, \zeta)$ and $(Q_{\mathsf{fa}}, Q_{\mathsf{md}})$. Hence, Θ and ζ can be used as the PHY-layer parameters in CR, which determine the operating point for cooperative spectrum sensing.

4.1.3 Assumptions

- PU_B can perform spectrum sensing cooperatively, which is coordinated and synchronized by N_B's base station, and can be of two types [37]: (1) on-demand and (2) periodic. When a new or handed-off PU_B arrives, on-demand sensing is performed to examine the availability of the channels. On the other hand, when a PU_B is already operating in the renting channels, periodic sensing is performed at a predefined time interval to discover the appearance of PU_A. In this work, the aim is to study the imperfect on-demand sensing, but periodic sensing is assumed to be perfect.

- Since channels in $\mathcal{C}_\mathcal{B}$ cannot be used by PU_A, false alarm and miss-detection events can only occur in $\mathcal{C}_\mathcal{R}$. For the miss-detection events, consider the scenarios that PU_A gives up the channel due to harmful interference when colliding with PU_B.

- The arrivals of PU_A and PU_B are assumed to follow a Poisson process with the mean arrival rate λ_A and λ_B, respectively. A Poisson process can well-model the events such as the arrival requests for documents on a web server, and the joining of customers in a queue at a checkout counter. However, on the other hand, Tang *et al.* [60] found that Poisson process fails to model internet traffic, which is typically bursty and has self-similar characteristics. Despite this, this topic is beyond the scope of this chapter. The channel occupying times of PU_A and PU_B are also assumed to follow a negative exponential distribution with mean μ_A^{-1} and μ_B^{-1}, respectively. For convenience, it is assumed that each user requires only a single channel for transmission.

4.2 Markov Chain Model

In this section, an analytical model is developed, which takes into account the effects of sensing errors for the partial spectrum sharing policy. Specifically, a three-dimensional Markov chain model is proposed to analyze the network. To start with, let $S(i,j,k)$ represent the system state with i PU_A users being served by channels in $\mathcal{C}_{\mathcal{A}}$, j PU_Bs being served by channels in $\mathcal{C}_{\mathcal{R}}$, and k PU_Bs being served by channels in $\mathcal{C}_{\mathcal{B}}$. The state space is given by

$$\mathcal{S} = \left\{ S(i,j,k) \left| \begin{array}{l} 0 \le i \le C_A, \\ \\ 0 \le j \le \min\left(C_A - i, C_R\right) \\ \\ 0 \le k \le C_B \end{array} \right. \right\}. \tag{4.7}$$

Let $\pi(i,j,k)$ be the steady state probability of the state $S(i,j,k)$. Figure 4.4 shows the state transition diagram for the Markov chain model for the partial spectrum sharing policy. The transition of system states is caused by the following events: 1) PU_A request arrival; 2) PU_B request arrival; 3) PU_A service completion; and 4) PU_B service completion.

Before deriving the state transition probabilities for this model, it is useful to consider the case when there are s vacant channels and l PU_As in $\mathcal{C}_{\mathcal{R}}$, and define the probability of having m false alarm events and n miss-detection events as

$$p_{m,n}^{(s,l)} = \binom{s}{m} Q_{\mathsf{fa}}^m \left(1 - Q_{\mathsf{fa}}\right)^{s-m} \binom{l}{n} Q_{\mathsf{md}}^n \left(1 - Q_{\mathsf{md}}\right)^{l-n}, \tag{4.8}$$

where $0 \le m \le s \le C_R$ and $0 \le n \le l \le C_R - s$.

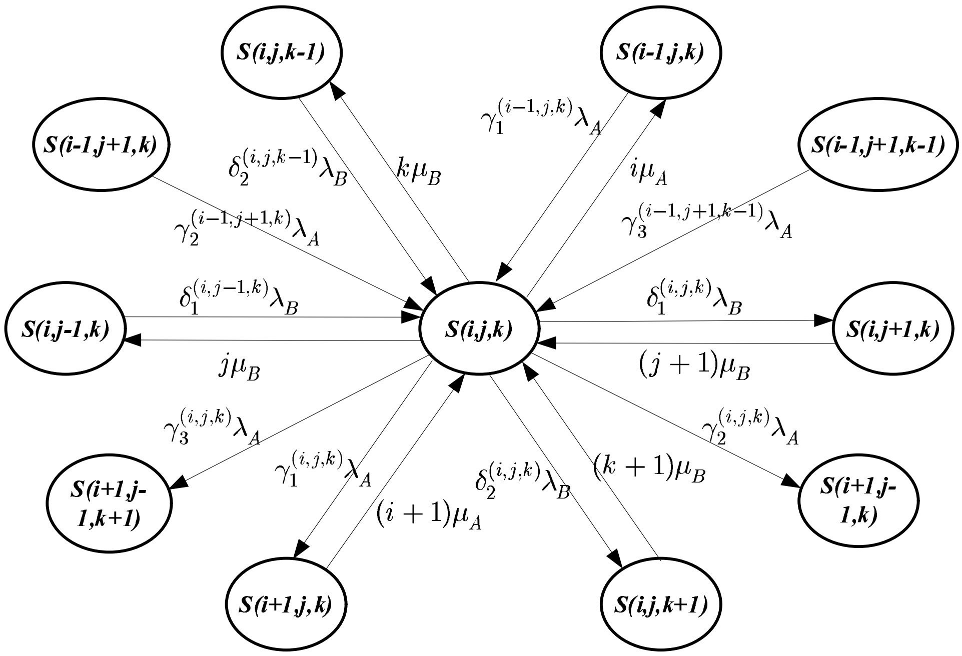


Figure 4.4: A Markov chain model for the partial spectrum sharing. $S(i,j,k)$ represent the system state with i PU_A users in $\mathcal{C}_{\mathcal{A}}$, j PU_B users in $\mathcal{C}_{\mathcal{R}}$, and k PU_B users in $\mathcal{C}_{\mathcal{B}}$.

4.2.1 State Transitions upon PU_A Request Arrival

For the event of a PU_A request arrival, the possible transitions from state $S(i,j,k)$ are states $S(i+1,j,k)$, $S(i+1,j-1,k)$, and $S(i+1,j-1,k+1)$, respectively, where $\{S(i+1,j,k), S(i+1,j-1,k), S(i+1,j-1,k+1)\} \in \mathcal{S}$.

- To state $S(i+1,j,k)$: This occurs when a PU_A arrives, and does not have channel collision with any PU_B. This may arise either when the number of PU_As is smaller than $C_A - C_R$, which does not need any channel being occupied by PU_Bs; or when the number of PU_As is greater than $C_A - C_R$ but the interrupted PU_B can perform on-demand sensing and successfully hand off to another vacant channel in $\mathcal{C}_\mathcal{R}$. As such, the state transition rate can be given by $\gamma_1^{(i,j,k)}\lambda_A$, where $\gamma_1^{(i,j,k)}$ is defined to ensure that only feasible states are considered in the state transitions,

$$\gamma_1^{(i,j,k)} = \begin{cases} 1 & \text{for } i < C_A - C_R, \\ \frac{(C_A-i-j)+j\beta_1^{(i,j,k)}}{C_A-i} & \text{for } C_A - C_R \le i < C_A, \\ 0 & \text{for } i = C_A, \end{cases} \tag{4.9}$$

  where $\beta_1^{(i,j,k)}$ is the probability that the interrupted PU_B successfully hands off to another vacant channel in $\mathcal{C}_\mathcal{R}$ based on its on-demand sensing result, given by

$$\beta_1^{(i,j,k)} = \begin{cases} \sum_{m=0}^{s-1}\sum_{n=0}^{l} p_{m,n}^{(s,l)}\frac{s-m}{n+s-m} & \text{if } s>0, l\ge 0 \text{ and } k = C_B, \\ 0 & \text{otherwise}, \end{cases} \tag{4.10}$$

  where $s \triangleq C_A - i - j$ and $l \triangleq i - (C_A - C_R)$ are defined for convenience.

- To state $S(i+1, j-1, k)$: This occurs when the arrival PU_A interrupts a PU_B and the PU_B fails to find any vacant channels in $\mathcal{C}_{\mathcal{R}}$ or $\mathcal{C}_{\mathcal{B}}$. Hence, the PU_B is forcefully terminated. In this case, the state transition rate is $\gamma_2^{(i,j,k)}\lambda_A$, where $\gamma_2^{(i,j,k)}$ is defined and given by

$$\gamma_2^{(i,j,k)} = \begin{cases} \frac{jp_{s,0}^{(s,l)}}{C_A - i} & \text{for } C_A - C_R \leq i < C_A,\ j \geq 1 \text{ and } k = C_B, \\ 0 & \text{for } i < C_A - C_R \text{ or } i = C_A \text{ or } j = 0 \text{ or } k < C_B, \end{cases} \tag{4.11}$$

in which $p_{s,0}^{(s,l)}$ denotes the probability that when $\mathcal{C}_{\mathcal{B}}$ is fully occupied, the interrupted PU_B misunderstands that all the vacant channels in $\mathcal{C}_{\mathcal{R}}$ are occupied, and thus fails to handoff successfully.

- To state $S(i+1, j-1, k+1)$: This happens when there are vacant channels in $\mathcal{C}_{\mathcal{B}}$ and that the arriving PU_A interrupts the PU_B's transmission but this PU_B succeeds to hand off to the vacant channel in $\mathcal{C}_{\mathcal{B}}$. The state transition rate in this case becomes $\gamma_3^{(i,j,k)}\lambda_A$, where

$$\gamma_3^{(i,j,k)} = \begin{cases} \frac{j}{C_A - i} & \text{for } C_A - C_R \leq i < C_A,\ j \geq 1 \text{ and } k < C_B, \\ 0 & \text{for } i < C_A - C_R \text{ or } i = C_A \text{ or } j = 0 \text{ or } k = C_B. \end{cases} \tag{4.12}$$

4.2.2 State Transitions upon PU_B Request Arrival

For the event of a PU_B request arrival, the possible transitions from state $S(i, j, k)$ are states $S(i, j+1, k)$, $S(i, j, k+1)$, and $S(i-1, j+1, k)$, respectively, where $\{S(i, j+1, k), S(i, j, k+1), S(i-1, j+1, k)\} \in \mathcal{S}$.

- To state $S(i, j+1, k)$: This happens when all the channels in $\mathcal{C}_{\mathcal{B}}$ are occupied, a PU_B arrives and occupies a channel from $\mathcal{C}_{\mathcal{R}}$. The state transition rate is $\delta_1^{(i,j,k)}\lambda_B$, where $\delta_1^{(i,j,k)}$ can be found as

$$\delta_1^{(i,j,k)} = \begin{cases} \sum_{m=0}^{C_R-j-1} p_{m,0}^{(C_R-j,0)} & \text{for } i < C_A - C_R, j < C_R, \text{and } k = C_B, \\ \sum_{m=0}^{s} \sum_{n=0}^{l} p_{m,n}^{(s,l)} \frac{s-m}{n+s-m} & \text{for } i \geq C_A - C_R, i+j < C_A, \text{and } k = C_B, \\ 0 & \text{for } i+j = C_A, \text{or } j = C_R, \text{or } k < C_B. \end{cases} \tag{4.13}$$

- To state $S(i,j,k+1)$: When the number of PU_B in the system is less than C_B, the arriving PU_B does not need a renting channel, as it can be allocated a licensed channel in $\mathcal{C}_\mathcal{B}$ by the N_B base station. Hence, in this case, the state transition rate $\delta_2^{(i,j,k)}\lambda_B$, where $\delta_2^{(i,j,k)}$ is given by

$$\delta_2^{(i,j,k)} = \begin{cases} 1 & \text{for } k < C_B, \\ 0 & \text{for } k = C_B. \end{cases} \tag{4.14}$$

- To state $S(i-1,j+1,k)$: This occurs when a PU_B arrives and jams a channel occupied by PU_A in $\mathcal{C}_\mathcal{R}$, due to miss-detection. In this case, the state transition rate becomes $\delta_3^{(i,j,k)}\lambda_B$, where $\delta_3^{(i,j,k)}$ is

$$\delta_3^{(i,j,k)} = \begin{cases} \sum_{m=0}^{s} \sum_{n=1}^{l} p_{m,n}^{(s,l)} \frac{n}{n+s-m} & \text{for } i > C_A - C_R \text{ and } k = C_B, \\ 0 & \text{for } i \leq C_A - C_R \text{ or } k < C_B. \end{cases} \tag{4.15}$$

4.2.3 State Transitions upon PU_A or PU_B Service Completion

The state transitions to states $S(i-1,j,k)$, $S(i,j-1,k)$ and $S(i,j,k-1)$ occur, respectively, when PU_As, PU_Bs are served by the channels in $\mathcal{C}_\mathcal{R}$, and PU_Bs in

$\mathcal{C}_{\mathcal{B}}$ complete their services successfully. The state transitions rates in these three cases can be given as $i\mu_A$, $j\mu_B$, and $k\mu_B$, respectively.

Based on the above analysis, the balance equation (4.16) can be derived as

$$\begin{bmatrix} \gamma_1^{(i,j,k)}\lambda_A + i\mu_A + \gamma_2^{(i,j,k)}\lambda_A + \gamma_3^{(i,j,k)}\lambda_A + \\ \delta_1^{(i,j,k)}\lambda_B + j\mu_B + \delta_2^{(i,j,k)}\lambda_B + k\mu_B + \delta_3^{(i,j,k)}\lambda_B \end{bmatrix} \pi(i,j,k) =$$
$$\gamma_1^{(i-1,j,k)}\lambda_A\pi(i-1,j,k) + (i+1)\mu_A\pi(i+1,j,k) + \gamma_2^{(i-1,j+1,k)}\lambda_A\pi(i-1,j+1,k) +$$
$$\gamma_3^{(i-1,j+1,k-1)}\lambda_A\pi(i-1,j+1,k-1) + \delta_1^{(i,j-1,k)}\lambda_B\pi(i,j-1,k) + (j+1)\mu_B\pi(i,j+1,k) +$$
$$\delta_2^{(i,j,k-1)}\lambda_B\pi(i,j,k-1) + (k+1)\mu_B\pi(i,j,k+1) + \delta_3^{(i+1,j-1,k)}\lambda_B\pi(i+1,j-1,k). \tag{4.16}$$

Combining the above equations with the normalization equation $\sum_{(i,j,k)\in\mathcal{S}} \pi(i,j,k) = 1$, the steady state probabilities can then be solved.

4.3 Performance Analysis

In this section, given the steady state probabilities $\pi(i,j,k)$, performance metrics are derived, such as the blocking probability and the interfering probability for PU_A, the blocking probability, the forced termination probability and the handoff probability for PU_B, and the total network throughput. Finally, a cross layer optimization problem is formulated.

4.3.1 BP and IP for PU_A

Blocking of PU_As occurs when all the channels in $\mathcal{C}_A$ are occupied by PU_As. Therefore, the blocking probability for PU_A, denoted by $P_{\mathsf{b}}^{(A)}$, is given by

$$P_{\mathsf{b}}^{(A)} = \sum_{k=0}^{C_B} \pi(C_A, 0, k). \tag{4.17}$$

Due to sensing errors, the interference for PU_A appears when the new arrival or handoff PU_B miss-detects the PU_A's transmission in the channel, and PU_A has to cease using the channel. Hence, the interfering probability for PU_A is defined as the ratio between the total PU_A interference rate and the total PU_A connection rate, which can be expressed as

$$P_{\mathsf{i}}^{(A)} = \frac{\sum_{S(i,j,k)\in\mathcal{S}} \left(\delta_3^{(i,j,k)}\lambda_B + \frac{j\beta_3^{(i,j,k)}}{C_A-i}\lambda_A\right)\pi(i,j,k)}{\lambda_A(1-P_{\mathsf{b}}^{(A)})}, \tag{4.18}$$

where

$$\beta_3^{(i,j,k)} = \begin{cases} \sum_{m=0}^{s}\sum_{n=1}^{l} p_{m,n}^{(s,l)} \frac{n}{n+s-m} & \text{for } C_A - C_R < i < C_A \text{ and } k = C_B, \\ 0 & \text{otherwise,} \end{cases} \tag{4.19}$$

is the probability that the handoff PU_B miss-detects, causing PU_A to terminate its transmission.

4.3.2 BP, FTP and HP for PU_B

Blocking of PU_B occurs when 1) all the channels in $\mathcal{C}_\mathcal{B}$ are occupied, and 2) the sensing results indicate that there is no vacant channel in $\mathcal{C}_\mathcal{R}$. From the analytical model, the average number of serviced PU_Bs is defined as $\sum_{S(i,j,k)\in\mathcal{S}} (j+k)\,\pi(i,j,k)$, which can also be interpreted as the average served traffic of PU_B [67]. Therefore,

$$\sum_{S(i,j,k)\in\mathcal{S}} (j+k)\,\pi(i,j,k) = \frac{\lambda_B}{\mu_B}\left(1-P_{\mathsf{b}}^{(B)}\right)\left(1-P_{\mathsf{ft}}^{(B)}\right), \tag{4.20}$$

where $P_{\mathsf{ft}}^{(B)}$ is the forced termination probability for PU_B, and is given by

$$P_{\mathsf{ft}}^{(B)} = \frac{\sum_{S(i,j,k)\in\mathcal{S}} \gamma_2^{(i,j,k)}\lambda_A\pi(i,j,k)}{\lambda_B\left(1-P_{\mathsf{b}}^{(B)}\right)}. \tag{4.21}$$

Combining (4.20) and (4.21), have the blocking probability for PU_B as

$$P_{\mathsf{b}}^{(B)} = 1-\left(\frac{\sum\limits_{S(i,j,k)\in S}\left((j+k)\,\mu_B+\gamma_2^{(i,j,k)}\lambda_A\right)\pi(i,j,k)}{\lambda_B}\right). \tag{4.22}$$

With a PU_A arrival, a PU_B may need to hand off to another channel when its current operating channel is required. As a result, the hand-off probability can be derived as

$$P_{\mathsf{h}}^{(B)} = \sum_{S(i,j,k)\in\mathcal{S}} \frac{\frac{j}{C_A-i}\lambda_A\pi(i,j,k)}{\lambda_B\left(1-P_{\mathsf{b}}^{(B)}\right)}. \tag{4.23}$$

C. GoS Definition

A possible definition for the GoS of N_A, which considers both blocking probability and interfering probability, is given by [36]

$$\mathsf{GoS}_A = \frac{P_{\mathsf{b}}^{(A)}+\omega_A P_{\mathsf{i}}^{(A)}}{1+\omega_A}, \tag{4.24}$$

where ω_A is the parameter that can be carefully chosen to provide a desired tradeoff between interference and blocking. Similarly, the GoS of N_B can be defined as

$$\mathsf{GoS}_B = \frac{P_{\mathsf{b}}^{(B)}+\omega_B P_{\mathsf{ft}}^{(B)}}{1+\omega_B}. \tag{4.25}$$

Note that the lower the values for GoS_A and GoS_B, the better the GoS performance.

4.3.3 Throughput

The overall network throughput can be defined as the average number of bits achieved by successfully completed PU_A and PU_B services per unit time [68]. Assuming that all the PU_A and PU_B have the same normalized data rate, i.e., 1 bit per unit time, the network throughput is written as

$$\tau = \left(1 - P_{\mathsf{nc}}^{(A)}\right) \lambda_A \left(\frac{\mu_A}{1 - P_{\mathsf{i}}^{(A)}}\right)^{-1} + \left(1 - P_{\mathsf{nc}}^{(B)}\right) \lambda_B \left(\frac{\mu_B}{1 - P_{\mathsf{ft}}^{(B)}}\right)^{-1}, \tag{4.26}$$

where

$$P_{\mathsf{nc}}^{(A)} = P_{\mathsf{b}}^{(A)} + (1 - P_{\mathsf{b}}^{(A)}) P_{\mathsf{i}}^{(A)}, \tag{4.27}$$

$$P_{\mathsf{nc}}^{(B)} = P_{\mathsf{b}}^{(B)} + (1 - P_{\mathsf{b}}^{(B)}) P_{\mathsf{ft}}^{(B)}, \tag{4.28}$$

are the non-completion probabilities for PU_A and PU_B, respectively. In the above, for $m \in \{A, B\}$, $(1 - P_{\mathsf{nc}}^{(m)})\lambda_m$ gives the service completion rate, and $\left(\frac{\mu_A}{1-P_{\mathsf{i}}^{(A)}}\right)^{-1}$ and $\left(\frac{\mu_B}{1-P_{\mathsf{ft}}^{(B)}}\right)^{-1}$ are the average durations of the successfully completed PU_A and PU_B services, respectively. Hence, the throughput achieved by incomplete connections will be not counted.

4.3.4 Cross-layer Optimization

The objective is to find the optimal energy decision threshold Θ and the voting threshold ζ as the PHY layer parameters, and the optimal number of C_R as the MAC layer parameter, in maximizing τ, given the GoS constraints for N_A (G_A) and N_B (G_B), i.e.,

$$\max_{\Theta,\zeta,C_R} \tau \quad \text{s.t.} \quad \begin{cases} \mathsf{GoS}_A \leq G_A, \\ \mathsf{GoS}_B \leq G_B. \end{cases} \tag{4.29}$$

This is a mixed-integer nonlinear programming problem, which contains integer and continuous variables. Particle swarm optimization (PSO) [71] is a useful method to find an optimal or quasi-optimal for such kind of problems. For each possible value of C_R, the PSO algorithm is used to obtain the optimal values for the continuous variables, i.e., Θ and ζ. Then for the discrete variable C_R, its optimal value can be obtained using exhaustive search methods. The process of solving this optimization problem is similar to the method described in Chapter 3.3.6, where the details can be found in Appendix B.

4.4 Numerical Results

In this section, the numerical results for studying the performance of the partial spectrum sharing policy in the presence of sensing errors is presented. Then the cross-layer design strategy is also evaluated.

It is considered the scenario that N_A is under-utilized and N_B is over-utilized in the pre-allocated scheme. Hence, if not specified otherwise, it is assumed that the arrival rate of PU_As within the range of $0.6 \leq \lambda_A \leq 1.2$, which corresponds to the N_A's average utilization rate at $20.0\% - 39.0\%$. The arrival rate of PU_Bs is set to $\lambda_B = 3.5$ calls per unit time, equivalent to an average utilization rate at 78.2%. Note that these assumptions of the users' traffic rates are practical in real world wireless systems. For the sensing parameters, the settings in Figure4.3 are adopted, as $u = 10$, $K = 10$ and $\overline{\gamma} = 5$ dB. The settings of parameters are summarized in Table 4.1.

Table 4.1: Parameter Settings.

Arrival rate of PU_As	$0.6 \leq \lambda_A \leq 1.2$
Arrival rate of PU_Bs	$\lambda_B = 3.5$
Channel occupying times	$\mu_A^{-1} = \mu_B^{-1} = 2$
Weight factor of interference and blocking for GoS of N_A	$\omega_A = 10$
Weight factor of forced termination and blocking for GoS of N_B	$\omega_A = 10$
GoS constraints for N_A and N_B	$G_A = 0.01$, $G_B = 0.02$
Time-bandwidth product	$u = 10$
The number of PU_B for cooperative sensing	$K = 10$
Average SNR	$\overline{\gamma} = 5$

4.4.1 Evaluating Partial Spectrum Sharing Policy

To verify the analytical model, a MATLAB-based discrete event simulator was used. A large number (i.e., 10^5) of PU_A and PU_B access requests were generated, and these requests were accepted or denied based on the access schemes described in the chapter. The number of blocked requests were recorded and then were used to calculate the corresponding probabilities. The simulations were repeated for 100 times, and calculated the mean and the 95% confidence interval of the probabilities. The results in Figure 4.5 are provided for the blocking probabilities of PU_A and PU_B obtained from the proposed analytical model and the simulation results, with $C_R = 3$, and the false alarm and miss-detection probabilities $Q_{\mathsf{fa}} = 0.1$ and $Q_{\mathsf{md}} = 0.1$. It is observed that the simulation results coincide with the analytical results.

Figure 4.6 provides the blocking probability results for PU_A against the arrival

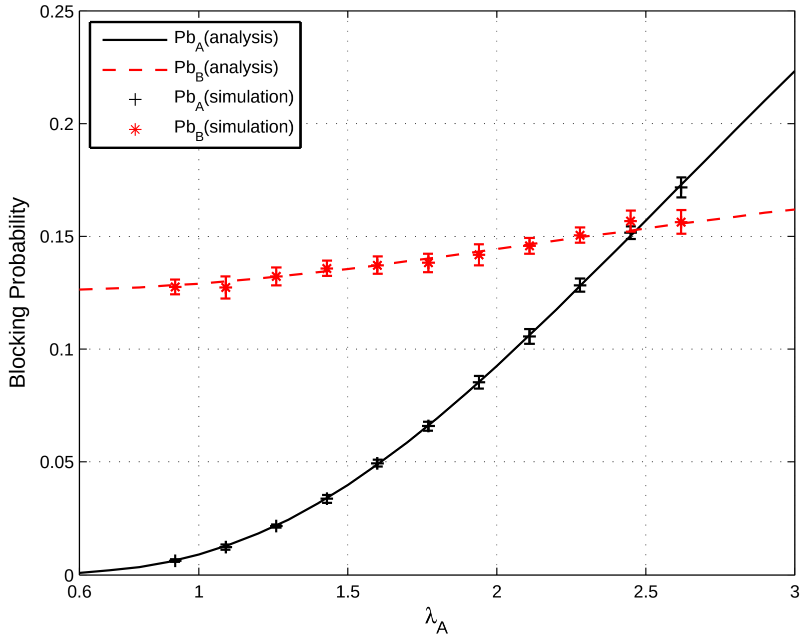


Figure 4.5: Analytical and simulation results for PU_A and PU_B blocking probabilities.

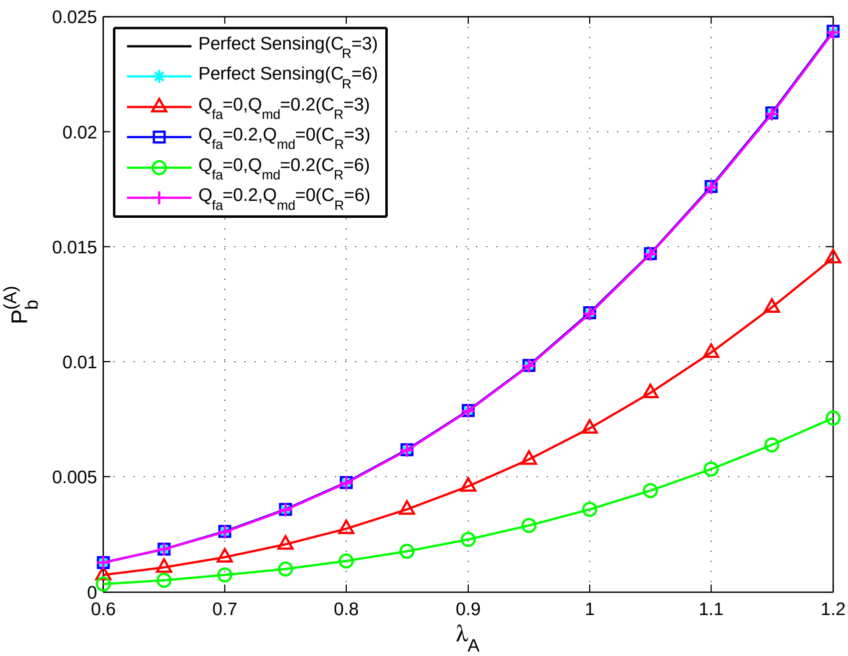


Figure 4.6: The blocking probabilities for PU_A for different values of false alarm probability (Q_{fa}), miss-detection probability (Q_{md}) and the number of sharing channels (C_R), against the arrival rate of PU_A λ_A.

rate λ_A. It is observed that PU_A blocking probability increases with λ_A with or without sensing errors, due to the increasing access requests from PU_A. Results also show that the blocking probability for PU_A does not change with an increase of Q_{fa}. On the other hand, with an increases of Q_{md}, the blocking probability for PU_A decreases. This can be explained by the fact that Q_{fa} only wastes channel access opportunities for PU_B, while Q_{md} interferes the current PU_A to give up the channel, which also leads to an additional access opportunity for upcoming PU_As' access requests. Since PU_A's licensed channels in $\mathcal{C}_{\mathcal{A}}$ are unchanged, with perfect sensing, the blocking probability for PU_A is independent of $\mathcal{C}_{\mathcal{R}}$. However, in the case of imperfect sensing, with more shared channels in $\mathcal{C}_{\mathcal{R}}$, a larger impact of Q_{md} on the blocking probability for PU_A can be observed. Figure 4.7 illustrates the interfering probability results for PU_A. It is observed that only Q_{md} causes the interference of PU_As, and the interference is more noticeable for larger C_R or larger arrival rate λ_A.

Figure 4.8 shows the blocking probability results for PU_B against the arrival rate λ_A. Results illustrate that with more shared channels in $\mathcal{C}_{\mathcal{R}}$, PU_B can sense and access more renting channels, thus a much lower blocking probability for PU_B can be obtained. We can also observe that a larger Q_{fa} causes a higher blocking probability for PU_B. The reason is that false alarm events would interpret the vacant channel as being occupied, and therefore waste the channel resources for PU_Bs. On the other hand, a larger Q_{md} leads to a lower blocking probability for PU_B because miss-detection events result in PU_As to give up the channel, giving more access opportunities for upcoming PU_B request.

In Figure 4.9, the forced termination probability for PU_B against various λ_A is plotted. As we can see, with more shared channels in $\mathcal{C}_{\mathcal{R}}$, PU_B is more likely to

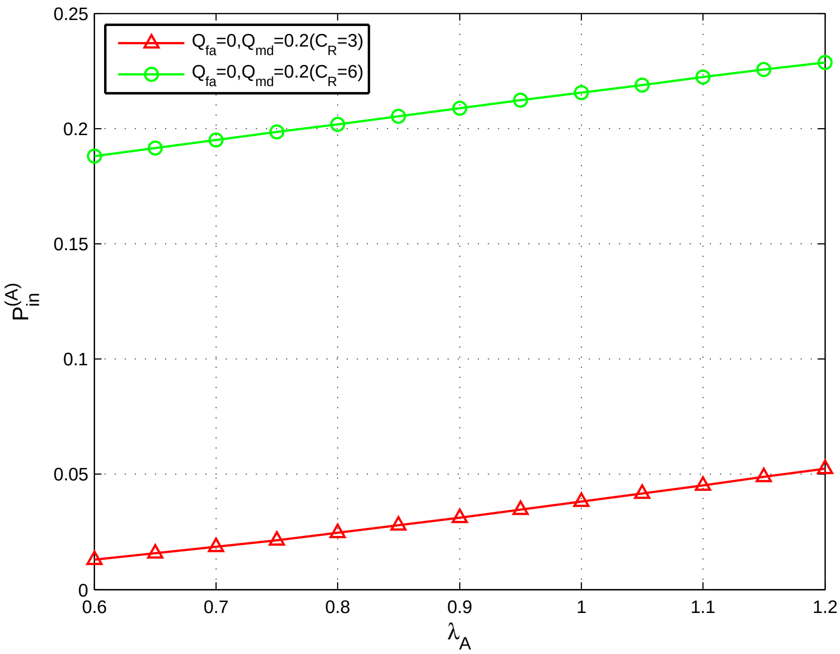


Figure 4.7: The interference probability for PU_A for different values of false alarm probability (Q_{fa}), miss-detection probability (Q_{md}) and the number of sharing channels (C_R), against the arrival rate of PU_A λ_A. Note that the interference probability for PU_A for the Perfect Sensing and $Q_{fa} = 0.2, Q_{md} = 0$ cases are $P_{\mathsf{in}}^{(A)} = 0$.

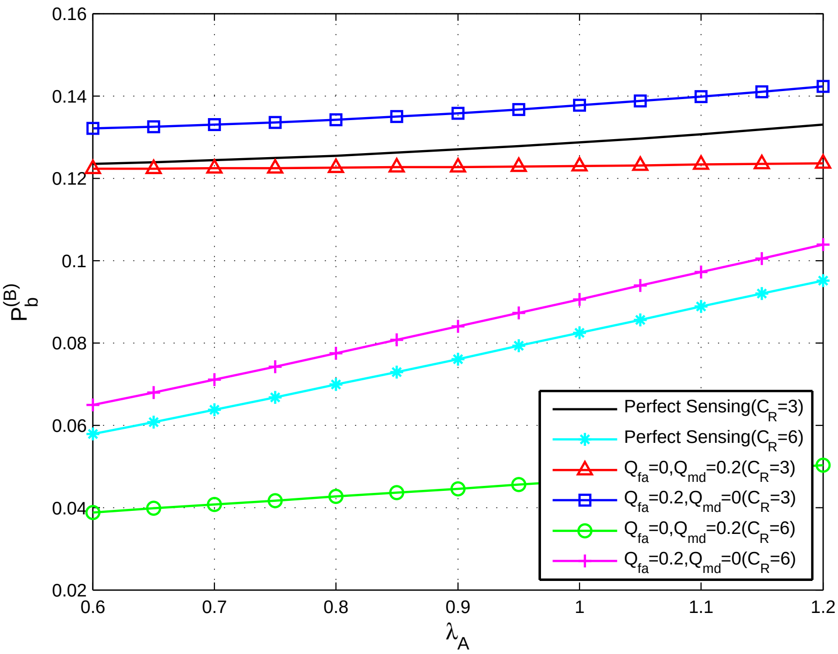


Figure 4.8: The blocking probability for PU_B for different values of false alarm probability (Q_{fa}), miss-detection probability (Q_{md}) and the number of sharing channels (C_R), against the arrival rate of PU_A λ_A.

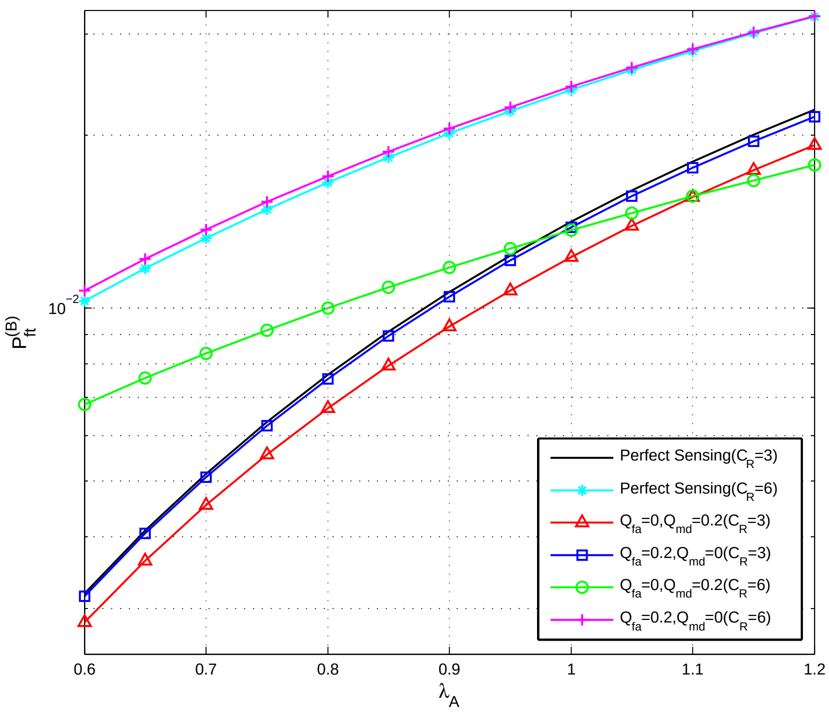


Figure 4.9: The forced termination probability for PU_B for different values of false alarm probability (Q_{fa}), miss-detection probability (Q_{md}) and the number of sharing channels (C_R), against the arrival rate of PU_A λ_A.

be terminated forcefully. Note that in the case of a larger Q_{fa}, the forced termination probability for PU_B can be greater or smaller than that in the case of perfect sensing. The reason is that false alarm events cause less PU_Bs operating in the spectrum, leading to a lower overall forced PU_B termination rate, as compared with the perfect sensing case. However, false alarm events also cause a higher blocking probability for PU_B, and therefore provide a lower PU_B total connection rate, while for a larger Q_{md}, miss-detection events provide PU_Bs with more opportunities to achieve successful handoffs, when they have collisions with PU_As. Hence, the forced termination probability for PU_B decreases with an increase of Q_{md}.

In Figure 4.10, it is observed that more handoff operations are needed for PU_B with a larger C_R. By comparing with the cases $Q_{fa} = 0.1, Q_{md} = 0$ and $Q_{fa} = 0, Q_{md} = 0.1$, results in this figure illustrate that changing Q_{fa} has greater impact on the handoff probability. This is reasonable since Q_{fa} decreases the number of PU_B in the spectrum directly, while Q_{md} may decrease the number of operating PU_A. If the number of PU_A is less than $C_A - C_R$, PU_A will not cause PU_B spectrum handoff.

Figure 4.11(a) and (b) plot the GoS performance for N_A and N_B against the energy decision threshold Θ, respectively. As we can see in Figure 4.11(a), an increase of Θ results in GoS performance degradation for N_A. The reason is that a large Θ incurs a greater miss-detection probability, and thus causes more interference to PU_As. With a lower PU_B arrival rate λ_B, the GoS performance for N_A is slightly improved due to the lower interfering probability, when Θ increases. However, with a larger λ_A or C_R, the GoS degradation is mainly due to the increasing of interfering probability which is particularly significant in the case

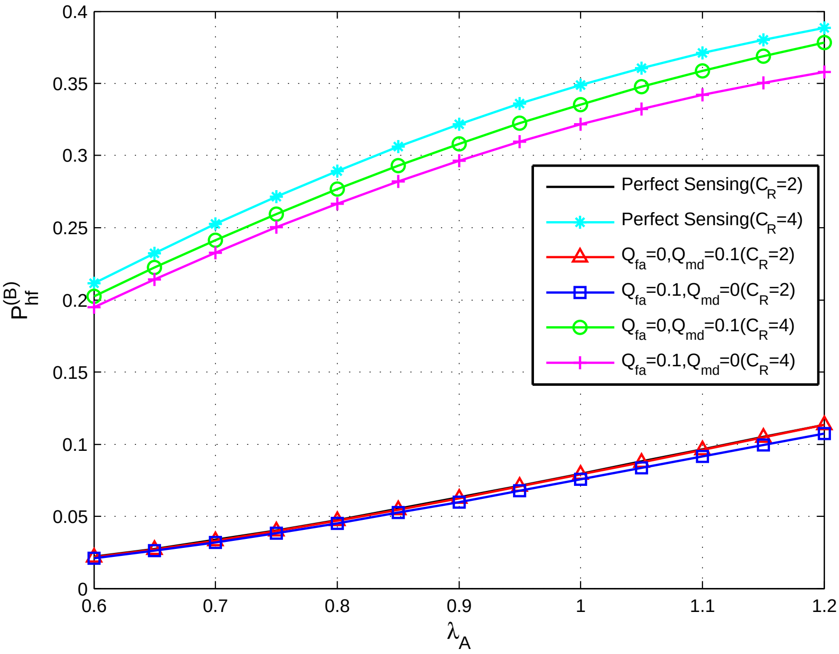
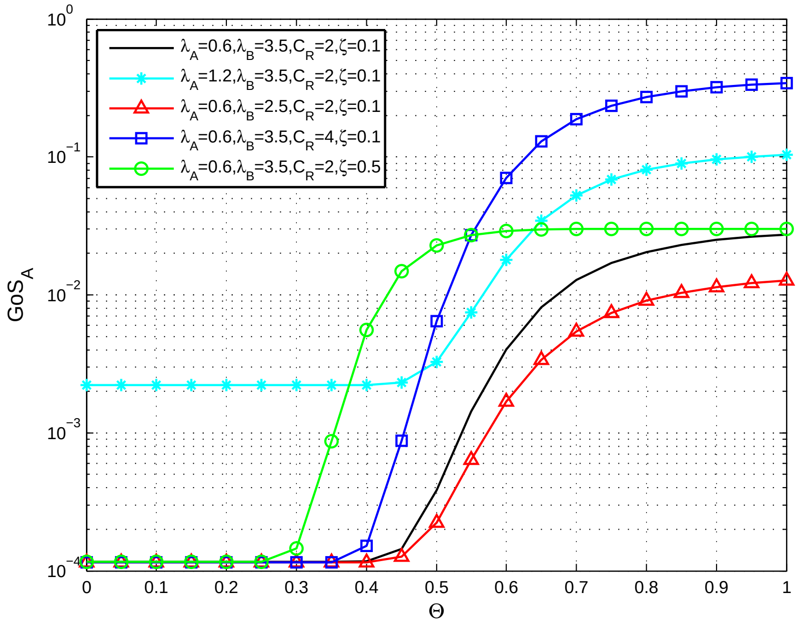
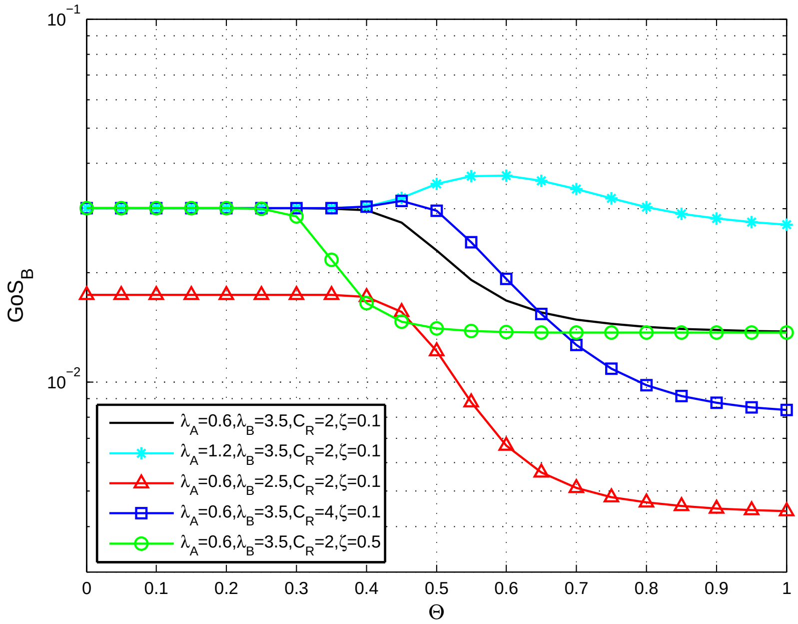


Figure 4.10: The hand-off probability for $P_{hf}^{(B)}$ for different values of false alarm probability (Q_{fa}), miss-detection probability (Q_{md}) and the number of sharing channels (C_R), against the arrival rate of PU_A λ_A.



(a)



(b)

Figure 4.11: The GoS performance of (a) PU_A, (b) PU_B, for different values of the arrival rate of PU_A λ_A, the arrival rate of PU_B λ_B, the number of sharing channels (C_R), and the voting threshold ζ, against the normalized energy decision threshold Θ.

with a larger C_R. It is also observed that for a lower voting threshold ζ, a better GoS performance for N_A can be achieved. The reason is that with less number of votes needed to make a final decision of PU_A appearance, a lower ζ achieves lower miss-detection probability, when Θ is large.

In Figure 4.11(b), an increase of Θ results in a lower false alarm probability, which helps PU_B to access channels and thus leads to less blocking and forced termination for PU_Bs. On the other hand, this also incurs a larger miss-detection probability, which leads to more forced terminations for PU_Bs. Since forced termination events have a greater impact than blocking on the GoS performance for N_B, GoS_B will increase first and then decrease, in the case $\lambda_A = 1.2$. With a lower λ_B or a larger C_R, when Θ increases, PU_Bs benefit from higher miss-detection probability because they are less likely to be forced termination. For different voting threshold rules, a greater threshold ζ results in lower false alarm probability for PU_B, and thus achieves better GoS performance for N_B.

4.4.2 Performance of Cross-Layer Optimization Scheme

In this section, the cross-layer design strategy is evaluated by comparing with the uncoordinated overlay hierarchical sharing policy [25] without cross-layer optimization, i.e., the selection of MAC layer parameter is independent of the PHY layer spectrum sensing parameters. The basic idea of hierarchical sharing policy in CR networks is to open all licensed spectrum to SUs, while limiting the interference perceived by PUs. To compare the sharing policies under the same conditions, we assume the pre-allocated number of channels for N_A and B are $C_A = 6$ and $C_B = 6$, respectively. For the uncoordinated hierarchical sharing policy, PU_A are considered as the PUs, while PU_B are referred to as the SUs, which corresponds to the setting $C_A = 6$, $C_R = 6$, and $C_B = 6$ in the model.

Figure 4.12 shows the overall throughput achieved by N_A and N_B. With perfect sensing, the proposed partial sharing model achieves a higher overall throughput compared to the hierarchical sharing model. This can be explained by recognizing the fact that the number of sharing channels can adapt with various λ_A in the proposed policy, which can better coincide with the GoS constraints of PU_A and PU_B. Since the PHY layer (i.e., Θ and ζ in spectrum sensing) optimization is independent of MAC layer in the hierarchical model, the number of shared channels cannot match with the various PU_A or PU_B arrival rates. However, in the proposed scheme, spectrum sensing parameters Θ and ζ are jointly optimized with MAC layer parameter C_R. Therefore, for the case of imperfect sensing, the throughput gain in the proposed model with the cross layer optimized strategy becomes much more significant. In particular, the gain over the hierarchical sharing model is 11.2% on average.

Figure 4.13 shows the optimal values of PHY and MAC layer parameters for the imperfect sensing case when the maximum overall throughput is achieved. For the partial sharing policy, with an increase of λ_A, the optimal value of Q_{fa} increases to achieve a lower Q_{md}, and thus can guarantee the GoS constraint of PU_A and adapt to a larger traffic of PU_A. Note that the optimal value of C_R changes from 3 to 2 when $\lambda_A > 1.05$, to guarantee the interference constraints of PU_A. For the hierarchical sharing, a much lower Q_{md} is required due to a fixed and larger number of shared channels. Furthermore, with an increase of λ_A, the optimal value of Q_{fa} decreases. This is because that the GoS constraint of PU_B becomes the dominate factor in (4.29), i.e., the network that first violates

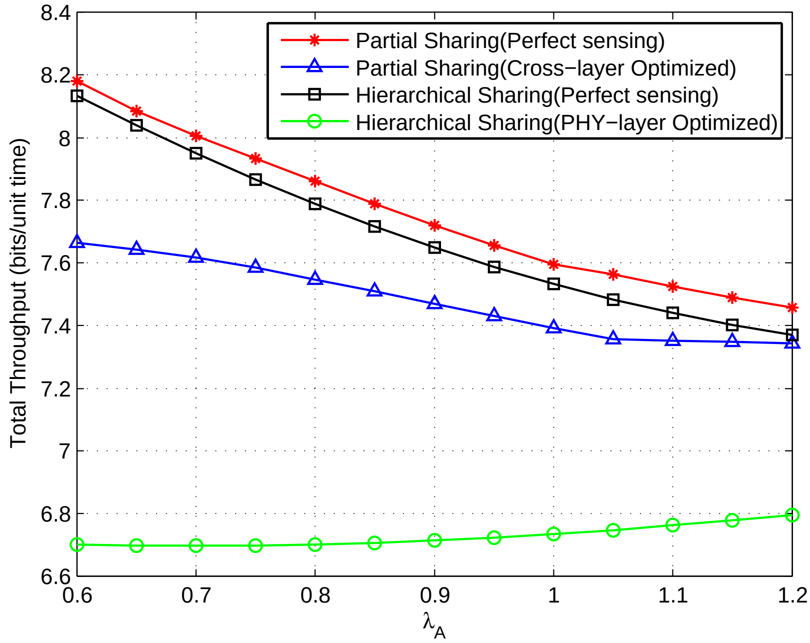


Figure 4.12: The total throughput τ against the arrival rate of PU_A λ_A.

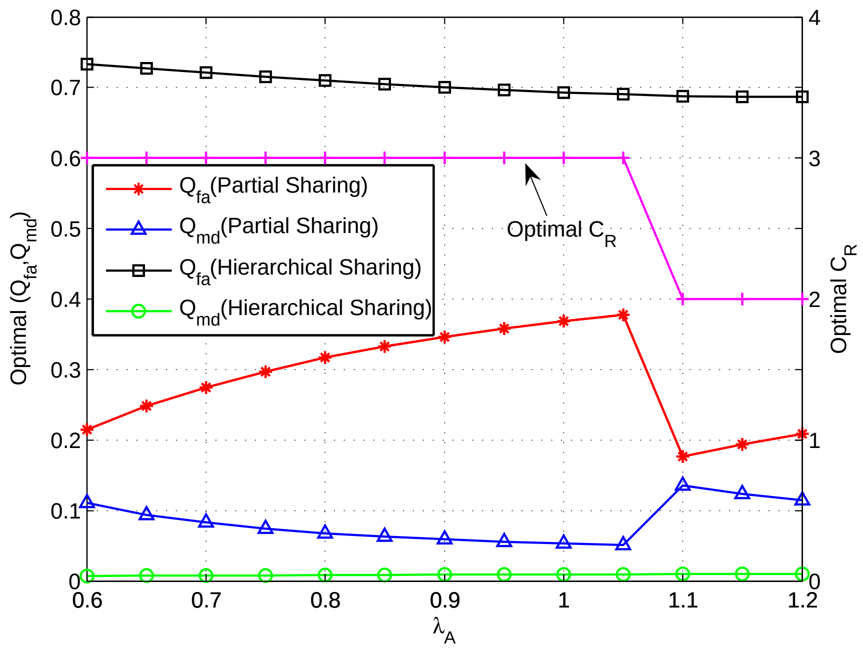


Figure 4.13: The optimal values of PHY and MAC layer parameters against the arrival rate of PU_A λ_A.

its prescribed GoS constraint. Hence, in this case, an increasing Q_{fa} is needed t guarantee the GoS constraint of PU_B.

4.5 Chapter Conclusions

This work studied spectrum sharing between two licensed wireless networks and proposed a new partial sharing policy for DSS. A Markov chain model was used to model the access process, with consideration of imperfect spectrum sensing. The blocking probability, forced termination probability, interfering probability, handoff probability, and the throughput for the proposed system have been derived. A cross-layer design was also presented, where the MAC layer spectrum sharing parameter and the PHY layer spectrum sensing parameters are jointly optimized to maximize the overall network throughput. Numerical results have illustrated that the partial sharing policy with the cross-layer design achieves 0.9% and 11.2% gains in the overall throughput in the perfect and imperfect sensing scenarios, respectively, while satisfying the GoS constraints of the users. The proposed spectrum sharing scheme and cross-layer design framework take advantages of coordinated and uncoordinated sharing models, which can achieve a better trade-off between system complexity and flexibility. Hence, this scheme is practical and can improve the system throughput in a real world uncoordinated DSS network efficiently.

The imperfect on-demand sensing was studied in this work, while periodic sensing is a also very important sensing scheme for CR. Chapter 5 will further study the imperfect periodic sensing and its impacts on performance metrics, which is practical for future DSS networks design.

Chapter 5

On Imperfect Spectrum Sensing for Uncoordinated Partial Spectrum-Shared Licensed Networks - A Discrete Time Markov Model

As mentioned in Chapter 4, opportunistic spectrum sharing between two licensed networks, e.g., WiMAX and a cellular network, is much less understood. In the scenario such as during a public event, there are large crowds and the traffic demands of cellular network will be locally extremely large. In this case, a cellular network might rent spectrum from WiMAX to increase its capacity and guarantee the users' GoS performance.

In this chapter, the DTMC modeling approach is adopted to extend the analysis of the partial spectrum sharing system in Chapter 4 to cope with imperfect and periodic spectrum sensing. Using this analytical model, the DSS system with partial sharing and imperfect spectrum sensing for real environments can be modeled. To study the capacity performance of the proposed scheme, a cross-layer design framework is also constructed, where the spectrum sensing parameters at the PHY layer and the spectrum sharing parameters at MAC layer are simultaneously coordinated to optimize the throughput, while satisfying the GoS constraints of the users.

5.1 System Model

5.1.1 Network Architecture

Two infrastructure-based licensed wireless networks (labeled as Networks A and B, whose users are referred to as PU_A and PU_B, respectively) are considered operating over a given area. Each of the networks has a base station for network management. It is assumed that the channels for Network B are not sufficient for the traffic demands of PU_B but PU_B can access parts of the channels of Network A opportunistically. The term "channel" here refers to a basic unit of capacity, which can be a time-slot, a frequency channel, a code or a combination of those.

The pre-allocated numbers of channels for Networks A and B are denoted by C_A and C_B, respectively. The sets containing the respective channels are denoted as $\mathcal{C}_\mathcal{A}$ and $\mathcal{C}_\mathcal{B}$. The renting or sharing channels are some of the channels in $\mathcal{C}_\mathcal{A}$ that can be rented by PU_B if needed. The number of renting channels is denoted by $C_R(\leq C_A)$, included in the set $\mathcal{C}_\mathcal{R}$. The following partial spectrum sharing policy is considered.

- PU_A: For any PU_A's access request, if the number of PU_As is less than $C_A - C_R$, PU_A will not occupy the renting channels in $\mathcal{C}_\mathcal{R}$; otherwise, the renting channels will be used.

- PU_B: For any PU_B's access request, if the number of PU_Bs is less than C_B, PU_B will occupy the channels in $\mathcal{C}_B$; otherwise, PU_B senses the spectrum to identify any vacant channels in $\mathcal{C}_\mathcal{R}$. If there are, then the PU_B will be admitted.

All users have spectrum handoff capability. Hence, a user of PU_A or PU_B can hand off from the renting channel to its pre-allocated non-renting channels, as soon as available.

5.1.2 Spectrum Sensing

The periodic spectrum sensing technique for DSS networks is shown in Figure 5.1. Specifically, PU_B will monitor the occupancy of renting channels periodically at time instances $t_1, t_2, \dots, t_n$. The sensing interval $\Delta T = t_n - t_{n-1}$ defines the time at which the occupancy information of the renting channels is available to PU_B. Based on the observations, PU_B can decide whether or not to access or release any of the renting channels. The access of PU_A and PU_B to $\mathcal{C}_\mathcal{A}$ and $\mathcal{C}_\mathcal{B}$ are coordinated by their base stations, respectively. Therefore, sensing errors appear only when PU_B are accessing the channels in $\mathcal{C}_\mathcal{R}$ and they are characterized by false alarm and miss detection probabilities. When the spectrum sensing fails to detect the appearance of PU, this case can be referred to as a *miss detection* event. However, if the channel is vacant but the sensing mistakenly indicates that it is occupied by PU, this case is referred to as a *false alarm* event.

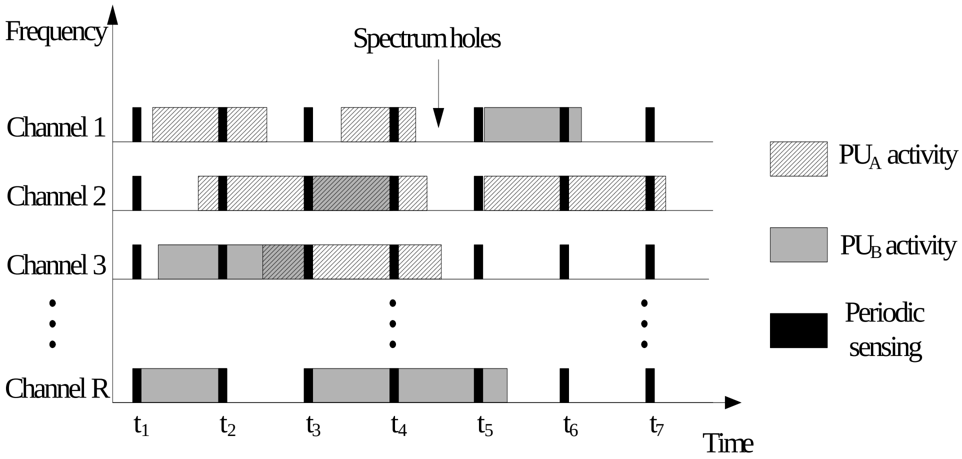


Figure 5.1: The periodic sensing scheme for the DSS system.

5.2 Discrete Time Markov Chain Model

5.2.1 Assumptions

The arrivals of PU_A and PU_B are assumed to experience a Poisson process with mean arrival rate λ_A and λ_B, respectively. Their channel occupying times are assumed to follow a negative exponential distribution with mean μ_A^{-1} and μ_B^{-1}, respectively. To make the following analysis tractable, it is assumed that the observation time interval of the DTMC model is also ΔT. PU_A and PU_B are assumed to require a single channel each time, and a user arriving at the interval between t_{n-1} and t_n will not depart at the same interval, and the orders of arrival and departure events during an observation time interval will be disregarded as in [36].

5.2.2 DTMC Formulation

In this work, a three-dimensional DTMC model is adopted to analyze the DSS system with imperfect spectrum sensing. To proceed, let (i, k, j) represent the system state with i PU_A users operating in $\mathcal{C}_{\mathcal{A}}$, k PU_A users operating in $\mathcal{C}_{\mathcal{R}}$, and also j PU_B users operating in $\mathcal{C}_{\mathcal{B}}$ and $\mathcal{C}_{\mathcal{R}}$. As mentioned in Section II, PU_A will not occupy the renting channels unless needed. Hence, the state parameter k depends on i, and $k = 0$ when $i \leq C_A - C_R$, where $C_A - C_R$ is the number of non-sharing channels in $\mathcal{C}_{\mathcal{A}}$. The state space is given by

$$\mathcal{S} = \left\{ (i,k,j) \;\middle|\; \begin{array}{l} 0 \leq i \leq C_A \\ \\ k = \max(0, i - (C_A - C_R)) \\ \\ 0 \leq j \leq C_B + C_R \end{array} \right\}. \tag{5.1}$$

Since the arrival of PU_A follows a Poisson distribution, the probability that s PU_A users arrive during ΔT is given by

$$p^{[\mathsf{A}]}_{\mathsf{arrive}}(s) = \frac{(\lambda_A \Delta T)^s}{s!} e^{-\lambda_A \Delta T}. \tag{5.2}$$

Define $p^{[\mathsf{B}]}_{\mathsf{arrive}}(s)$ similarly with λ_B for PU_B users. In addition, the probability of l departures out of s PU_A users (similarly for PU_B users) during ΔT is given by

$$p^{[\mathsf{A}]}_{\mathsf{depart}}(l, s) = \binom{s}{l} \left(1 - e^{-\mu_A \Delta T}\right)^l \left(e^{-\mu_A \Delta T}\right)^{s-l}. \tag{5.3}$$

The probability of believing m PU_A users in $\mathcal{C}_{\mathcal{R}}$ when there are actually k PU_A in $\mathcal{C}_{\mathcal{R}}$ is defined as [36]

$$\mathbb{P}_{m,k} = \sum_{n=\max(0,k-m)}^{\min(k, C_R - m)} \binom{C_R - k}{n + m - k} Q_{\mathsf{fa}}^{n+m-k} (1 - Q_{\mathsf{fa}})^{C_R - n - m} \binom{k}{n} Q_{\mathsf{md}}^{n} (1 - Q_{\mathsf{md}})^{k-n}, \tag{5.4}$$

where Q_{fa} and Q_{md} denote the false alarm probability and miss detection probability, respectively. In this equation, n denotes the number of miss detected channels, and miss detection can only happen on PU occupied channels, thus $n \leq k$.

The transition of system states is triggered by the following events:

1) Channel assignment for PU_A: As PU_A is the owner of $\mathcal{C}_A$, its channel assignment events will not be affected by PU_B's operation. Hence, the conditional probability of assigning s PU_A channels during a sensing interval ΔT, provided $l(\leq i)$ PU_A channels are released, is given by

$$\alpha^{[\mathsf{A}]}(s|i,j,l) = \begin{cases} p_{\mathsf{arrive}}^{[\mathsf{A}]}(s) & \text{for } s < C_A - i + l, \\ 1 - \sum_{n=0}^{s-1} p_{\mathsf{arrive}}^{[\mathsf{A}]}(n) & \text{for } s = C_A - i + l. \end{cases} \tag{5.5}$$

In (5.5), when the number of arriving PU_A is less than or equal to the number of channels available, i.e., the case $s < C_A - i + l$, all PU_A will be accepted. Otherwise, if the number of arriving PU_A is more than the number of channels available, only $C_A - i + l$ PU_A users will be assigned channels.

2) Channel release for PU_A: This happens when PU_A users complete their services. Hence, the probability of releasing $l(\leq i)$ PU_A channels during ΔT is given as

$$\beta^{[\mathsf{A}]}(l|i,j) = p_{\mathsf{depart}}^{[\mathsf{A}]}(l,i). \tag{5.6}$$

3) Channel assignment for PU_B: This event depends upon the observed spectrum occupancy from PU_B, which may be erroneous, i.e., false alarm and

miss detection events. Similar to (5.5), define $\tilde{\alpha}^{[\mathsf{B}]}(s|i,m,j,l)$ as the conditional probability of assigning s PU_B channels during ΔT provided there are m observed PU_A users operating in $\mathcal{C}_{\mathcal{R}}$, and $l(\leq j)$ PU_B channels released. Therefore, by definition,

$$\tilde{\alpha}^{[\mathsf{B}]}(s|i,m,j,l) = \begin{cases} p^{[\mathsf{B}]}_{\mathsf{arrive}}(s) & \text{for } \tilde{s}+m < C_B + C_R, \\ 1 - \sum_{n=0}^{s-1} p^{[\mathsf{B}]}_{\mathsf{arrive}}(n) & \text{for } \tilde{s}+m = C_B + C_R. \end{cases} \tag{5.7}$$

where $\tilde{s} \triangleq j-l+s$, is the number of PU_B users in the system after a channel assignment process. Hence, $\tilde{s}$ is less than or equal to C_B, the number of observed PU_A users operating in $\mathcal{C}_{\mathcal{R}}$ m, can be $m \in [0, C_R]$. While if $\tilde{s}$ is larger than C_B and $s > 0$, $m \in [0, C_B+C_R-\tilde{s}]$. For the special case $\tilde{s} > C_B$ and $s = 0$, the probability of no PU_B channels assigned is composed by the probability that PU_B has the equal number of arrival and leaving users when there are observed vacant channels, and the probability that there are no observed vacant channels. Based on (5.4), relationship between the actual and observed states can be mapped. As a result, the conditional probability of assigning s PU_B channels during ΔT, given $l(\leq j)$ PU_B channels released, is given by

$$\alpha^{[\mathsf{B}]}(s|i,k,j,l) = \begin{cases} \sum_{m=0}^{C_R} \tilde{\alpha}^{[\mathsf{B}]}(s|i,m,j,l)\mathbb{P}_{m,k} & \text{if } \tilde{s} \leq C_B \text{ and } s \geq 0, \\ \sum_{m=0}^{C_B+C_R-\tilde{s}} \tilde{\alpha}^{[\mathsf{B}]}(s|i,m,j,l)\mathbb{P}_{m,k} & \text{if } \tilde{s} > C_B \text{ and } s > 0, \\ \sum_{m=0}^{C_B+C_R-\tilde{s}} \tilde{\alpha}^{[\mathsf{B}]}(s|i,m,j,l)\mathbb{P}_{m,k} + \sum_{m=C_B+C_R-\tilde{s}+1}^{C_R} \mathbb{P}_{m,k} & \text{if } \tilde{s} > C_B \text{ and } s = 0. \end{cases} \tag{5.8}$$

4) Channel release for PU_B: Due to sensing errors, those PU_B users in $\mathcal{C}_{\mathcal{R}}$ may mistakenly depart their channels, i.e., false alarm events. Hence, the channel release happens when there are false alarms or service completions of PU_B users. In particular, when the number of PU_B releasing channels l, satisfying $j-l<C_B$, the PU_B channel release is only caused by the service completion. On the other hand, when $j-l\geq C_B$, the channel release might also be caused by the false alarm events. Therefore, the probability of releasing $l(\leq j)$ PU_B channels during ΔT is

$$\beta^{[\mathsf{B}]}(l|i,k,j)=\begin{cases} p_{\mathsf{depart}}^{[\mathsf{B}]}(l,j) & \text{for } j-l<C_B,\\ \sum_{r=0}^{l}\tilde{\beta}^{[\mathsf{B}]}(r|i,k,j,l-r)p_{\mathsf{depart}}^{[\mathsf{B}]}(r,j) & \text{for } j-l\geq C_B, \end{cases} \tag{5.9}$$

in which $\tilde{\beta}^{[\mathsf{B}]}(s|i,k,j,r)$ denotes the conditional probability that there are s PU_B channels released by false alarm events, given that $r(\leq l)$ PU_B users complete their service, and is given by (5.10).

$$\tilde{\beta}^{[\mathsf{B}]}(s|i,k,j,r)=$$
$$\begin{cases} \mathbb{P}_{C_B+C_R+s+r-j,k} & \text{for } j-r\geq C_B \text{ and } 0<s\leq j-r-C_B,\\ 1-\sum_{n=1}^{j-r-C_B}\mathbb{P}_{C_B+C_R+n+r-j,k} & \text{for } j-r\geq C_B \text{ and } s=0. \end{cases} \tag{5.10}$$

Let $\pi_{(i,k,j)}$ denote the steady state probability of the state (i,k,j). Based on the analysis, the state transition probability from (i,k,j) to $(i+a,k',j+b)$ ($k'=\max(0,i+a-(C_A-C_R))$, $-i\leq a\leq C_A-i$, $-j\leq b\leq C_B+C_R-j$) is given by

$$\varepsilon(i+a,k',j+b|i,k,j) = \sum_{l=\max(-a,0)}^{i} \alpha^{[\mathsf{A}]}(a+l|i,j,l)\beta^{[\mathsf{A}]}(l|i,j) + \sum_{l=\max(-b,0)}^{j} \alpha^{[\mathsf{B}]}(b+l|i,k,j,l)\beta^{[\mathsf{B}]}(l|i,k,j). \tag{5.11}$$

Let $\boldsymbol{\pi}$ be the row vector of steady state probabilities with elements $\pi_{(i,k,j)}$, and $\boldsymbol{\varepsilon}$ be the transition probability matrix of elements $\varepsilon(i+a,k',j+b|i,k,j)$ in (5.11). Then, the steady state $\boldsymbol{\pi}$ can be obtained by solving $\boldsymbol{\pi} = \boldsymbol{\pi\varepsilon}$ and $\boldsymbol{\pi}\mathbf{1} = 1$, where $\mathbf{1}$ is the column vector with all ones.

Note that $\pi_{(i,k,j)}$ denotes the actual steady state probability, while the "observed" steady state probability of PU_B may have sensing errors in determining the channel occupancy in $\mathcal{C}_{\mathcal{R}}$. From PU_B's viewpoint, only the number of PU_A in $\mathcal{C}_{\mathcal{R}}$ is of interest. Hence, from definition of i and k in (5.1), the three dimensional steady state probabilities can be converted into the two dimensional equivalent,

$$\pi'_{(k,j)} = \sum_{i=0}^{C_A} \pi_{(i,k,j)}. \tag{5.12}$$

Hence, the observed steady state probability for PU_B becomes

$$\pi^{\mathsf{observe}}_{(m,j)} = \sum_{k=0}^{C_R} \mathbb{P}_{m,k}\pi'_{(k,j)}, \text{for } (m,j) \in \mathcal{S}', \tag{5.13}$$

where $\mathcal{S}' \triangleq \{(m,j)|0 \le m \le C_R; 0 \le j \le C_B + C_R\}$.

5.3 Performance Analysis

The performance metrics are derived, such as blocking probabilities for PU_A and PU_B, forced termination probability for PU_B, interfering probability and throughput based on the actual and observed steady state probabilities $\pi_{(i,k,j)}, \pi^{\mathsf{observe}}_{(m,j)}$.

5.3.1 Blocking Probability

For PU_A, blocking events occur when no channels are available for upcoming PU_A users. Therefore, the blocking probability for PU_A is given by

$$P_{\mathsf{block}}^{[\mathsf{A}]} = \sum_{j=0}^{C_B+C_R} \pi_{(C_A,C_R,j)}. \tag{5.14}$$

Also, blocking events for PU_B occur if there are no vacant channels in $\mathcal{C}_{\mathcal{B}}$ and no observed channels in $\mathcal{C}_{\mathcal{R}}$. Thus, the blocking probability for PU_B is given by

$$P_{\mathsf{block}}^{[\mathsf{B}]} = \sum_{\substack{j+m \geq C_B+C_R \\ (m,j)\in \mathcal{S}'}} \pi_{(m,j)}^{\mathsf{observe}}. \tag{5.15}$$

5.3.2 Forced termination probability

In $\mathcal{C}_{\mathcal{R}}$, when an operating PU_B user detects a PU_A's activity, the PU_B user may be forced to terminate its transmission, if there are no vacant channels to hand off to. To derive the forced termination probability, the average number of serviced connections can be interpreted as the the average served traffic [67]. As a result,

$$\sum_{(i,k,j)\in \mathcal{S}} j\pi_{(i,k,j)} = \frac{\lambda_B}{\mu_B}\left(1 - P_{\mathsf{force\_terminate}}^{[\mathsf{B}]}\right)\left(1 - P_{\mathsf{block}}^{[\mathsf{B}]}\right), \tag{5.16}$$

which then gives

$$P_{\mathsf{force\_terminate}}^{[\mathsf{B}]} = 1 - \frac{\sum_{(i,k,j)\in \mathcal{S}} j\pi_{(i,k,j)}}{\frac{\lambda_B}{\mu_B}\left(1 - P_{\mathsf{block}}^{[\mathsf{B}]}\right)}. \tag{5.17}$$

5.3.3 Interfering probability

Due to the miss detection events, PU_A and PU_B users may simultaneously operate on the same $\mathcal{C}_{\mathcal{R}}$ channel, which is referred to as an interfering event. In the state

$\pi_{(i,k,j)}$, the average number of interfering channels can be found as

$$
\begin{aligned}
N_{\mathsf{interfering}}(i,k,j) &= N'_{(k,j)} + \\
&\sum_{r=\max(1,k+j-(C_R+C_B)-N'_{(k,j)})}^{\min(k,j-C_B)-N'_{(k,j)}} \frac{r Q_{\mathsf{md}}^{r} \binom{C_R-N'_{(k,j)}}{\max(k,j-C_B)-N'_{(k,j)}} \binom{\max(k,j-C_B)-N'_{(k,j)}}{r} \binom{C_R-\max(k,j-C_B)-N'_{(k,j)}}{\min\left(\substack{\min(k,j-C_B)-N'_{(k,j)}-r \\ C_R-(\max(k,j-C_B)-N'_{(k,j)})}\right)}}{\binom{C_R-N'_{(k,j)}}{k-N'_{(k,j)}} \binom{C_R-N'_{(k,j)}}{j-C_B-N'_{(k,j)}}}
\end{aligned}
\tag{5.18}
$$

where the first part of the right-hand-side of the equation denotes the lower bound of the interfering channels. Likewise, $N'_{(k,j)} = k+j-C_B-C_R$ if $k+j > C_B+C_R$, and 0 otherwise. In particular, when the number of PU_A and PU_B users in $\mathcal{C}_{\mathfrak{R}}$ is larger than C_R, there are at least $k+j-C_B-C_R$ channel collisions. Further, the remaining $k-N'_{(k,j)}$ PU_A and $j-C_B-N'_{(k,j)}$ PU_B may still have collisions and interfere with each other. In this case, the average number of interfering channels is considered by the second part of the right-hand-side of (5.18). The interfering probability is thus

$$
P_{\mathsf{interfering}} = \sum_{\substack{(i,k,j)\in\mathcal{S} \\ k>0, j>C_B}} \frac{N_{\mathsf{interfering}}(i,k,j)}{C_R} \pi_{(i,k,j)}. \tag{5.19}
$$

5.3.4 Throughput

The throughput can be defined as the number of completed connections. It is considered that during the interfering event, both PU_A and PU_B cannot transmit, and thus the throughput completed during the interfering event should be subtracted from the total throughput. Hence, the throughput of PU_B is given by

$$
\tau^{[\mathsf{B}]} = \sum_{(i,k,j)\in\mathcal{S}} \left(j - N_{\mathsf{interfering}}(i,k,j)\right) \pi_{(i,k,j)}. \tag{5.20}
$$

5.3.5 Cross-Layer Optimization

In the cross-layer design framework, the throughput performance can be maximized, while satisfying the GoS constraints of the users. It is assumed that the initial arrival rate of PU_B λ_B is greater to the available channels in the system, i.e., if accepting all PU_B's requests, its GoS constraints cannot be satisfied. Therefore, the objective is to find the optimal sensing interval ΔT at the PHY layer, and the admissible arrival rate of PU_B λ_B and the number of sharing channels C_R at the MAC layer, to maximize the throughput of PU_B under the GoS constraints of PU_A and PU_B. The optimization problem can be formulated as

$$\max_{\Delta T, \lambda_B, C_R} \quad \tau^{[\mathsf{B}]}$$
$$\text{s.t.} \quad \begin{cases} P_{\mathsf{interfering}} \leq P^{*}_{\mathsf{interfering}}, \\ P^{[\mathsf{B}]}_{\mathsf{block}} \leq P^{[\mathsf{B}]*}_{\mathsf{block}}, \\ P^{[\mathsf{B}]}_{\mathsf{force\_terminate}} \leq P^{[\mathsf{B}]*}_{\mathsf{force\_terminate}}. \end{cases} \tag{5.21}$$

where $P^{*}_{\mathsf{interfering}}, P^{[\mathsf{B}]*}_{\mathsf{block}}, P^{[\mathsf{B}]*}_{\mathsf{force\_terminate}}$ are the corresponding GoS constraints. The challenge is that (5.21) is a non-linear programming problem. PSO [58] is a useful algorithm that can find an optimal or quasi-optimal for such kind of problems. Therefore, first fix the value of the discrete variable C_R. For each value of C_R, the PSO algorithm in Chapter 3.3.6 is used to obtain optimal λ_B and ΔT. Then the optimal value of C_R can be obtained using an exhaustive search method. The process of solving this optimization problem is similar to the method described in Chapter 3.3.6, where the implement steps can be found in Appendix B.

Table 5.1: Parameter Settings.

Arrival rate of PU_As	$0.01 \leq \lambda_A \leq 0.1$ calls/s
Arrival rate of PU_Bs	$\lambda_B = 0.1$ calls/s
Channel occupying times	$\mu_A^{-1} = \mu_B^{-1} = 50$ s
The number of channels for PU_A and PU_B	$C_A = C_B = 4$
The number of sharing channels	$C_R = 2$
GoS constraints	$P^*_{\mathsf{interfering}} = 0.003$, $P^{[\mathsf{B}]*}_{\mathsf{block}} = 0.2$, $P^{[\mathsf{B}]*}_{\mathsf{force\_terminate}} = 0.05$
The sensing interval	$\Delta T = 0.5$s

5.4 Numerical Results

This section provides analytical and simulation results to show the performance of the CR-based DSS system with the partial spectrum sharing policy for various sensing errors and sensing intervals conditions. A MATLAB-based discrete event simulator was used to verify the analytical DTMC model. A large number (i.e., 10^5) of PU_A and PU_B access requests were generated, and these requests were accepted or denied based on the proposed access schemes. The number of blocked requests were recorded and then were used to calculate the corresponding probabilities. The parameters $C_A = C_B = 4$, and $\mu_A^{-1} = \mu_B^{-1} = 50$s. If not specified otherwise, define the arrival rate of PU_A within the range of $0.01 \leq \lambda_A \leq 0.1$, and the arrival rate of PU_B is set to $\lambda_B = 0.1$ calls/s. The parameters settings are summarized in Table 5.1.

Figure 5.2–Figure 5.5 demonstrate the performance metrics derived in the analytical model in the cases of perfect and imperfect sensing against the arrival rate of PU_A λ_A. The number of sharing channels is $C_R = 2$, and the sensing

interval $\Delta T = 0.5$ s as in Table 5.1. In Figure 5.2, it is observed that the numerical results for blocking probability of the DTMC model coincide with the derived CTMC model in Chapter 4 for the perfect sensing case. Furthermore, the results coincide with the analytical results for the imperfect sensing cases. It is observed that with a larger Q_{md}, PU_B will benefit from miss detection events by having more access opportunities, leading to a lower blocking probability than in the perfect sensing case. It is also observed that with a greater Q_{fa}, a lower $P^{[\mathsf{B}]}_{\mathsf{block}}$ can be obtained. The reason is that false alarm events lead to less operating PU_B users to hand off successfully, and thus provide more access opportunities for upcoming PU_B users.

For the forced termination results as shown in Figure 5.3, since miss detection events cause more PU_B users operating in $\mathcal{C}_{\mathcal{R}}$, more forced termination events for PU_B will occur, when compared with the perfect sensing case. In addition, with a higher false alarm probability, more handoff opportunities for PU_B users will be wasted, giving a higher $P^{[\mathsf{B}]}_{\mathsf{force\_terminate}}$. In this case, results also indicate that with a larger λ_A, less PU_B users can access channels, and therefore $P^{[\mathsf{B}]}_{\mathsf{force\_terminate}}$ decreases.

On the other hand, Figure 5.4 shows that miss detection events will cause a larger $P_{\mathsf{interfering}}$. This can be explained by recognizing the fact that with a large likelihood that SU fails to detect the appearance of PU, more PU will be interfered.

Figure 5.5 shows the throughput for PU_B against the arrival rate λ_A. It is observed that a larger false alarm probability leads to a lower throughput for PU_B. However, with a larger miss detection probability, higher throughput for PU_B can

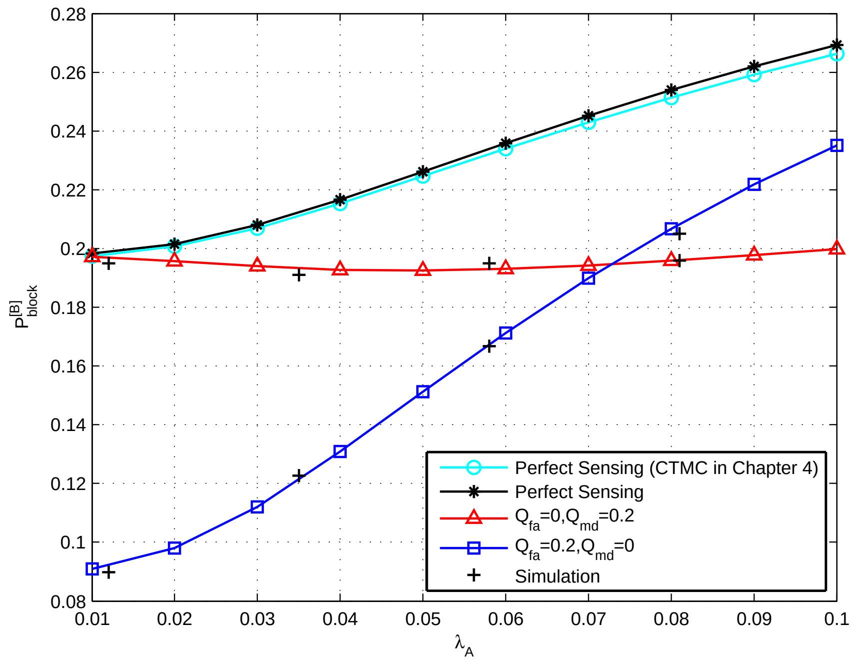


Figure 5.2: The blocking probability for PU_B against the arrival rate of PU_A λ_A. "Perfect sensing (CTMC in Chapter 4)" denotes the analytical results derived in Chapter 4. "Perfect sensing", "$Q_{fa} = 0, Q_{md} = 0.2$", "$Q_{fa} = 0.2, Q_{md} = 0$" cases are obtained from the analytical model in this chapter, where Q_{fa} and Q_{md} are the false alarm probability and miss-detection probability, respectively. "Simulation" denotes the simulation results, which are obtained from discrete event simulator and can verify the analytical results for imperfect sensing cases.

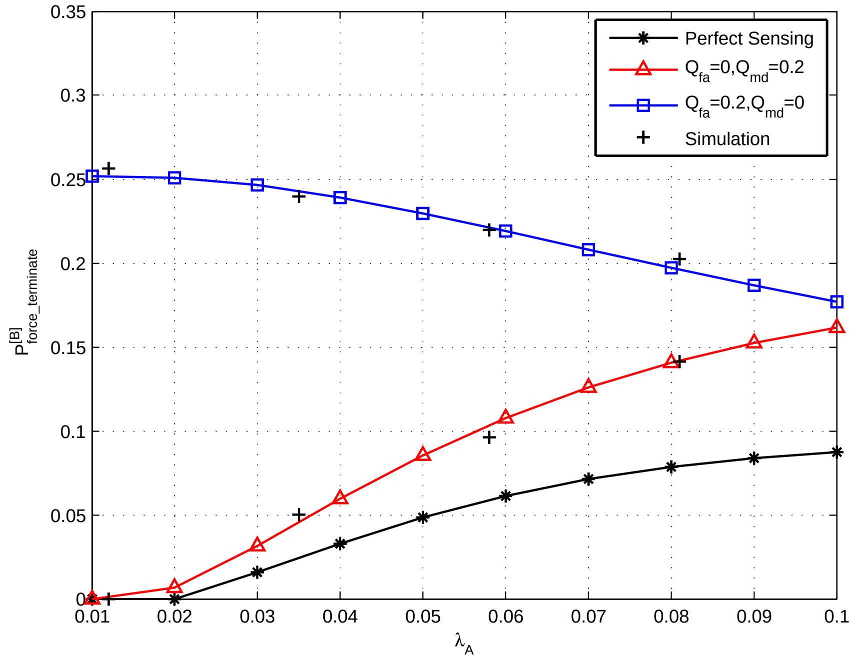


Figure 5.3: The forced termination probability for PU_B against the arrival rate of PU_A λ_A. The solid lines are obtained from the analytical model. Q_{fa} and Q_{md} are the false alarm probability and miss-detection probability, respectively. "Simulation" denotes the simulation results are obtained to verify the analytical model for imperfect sensing cases.

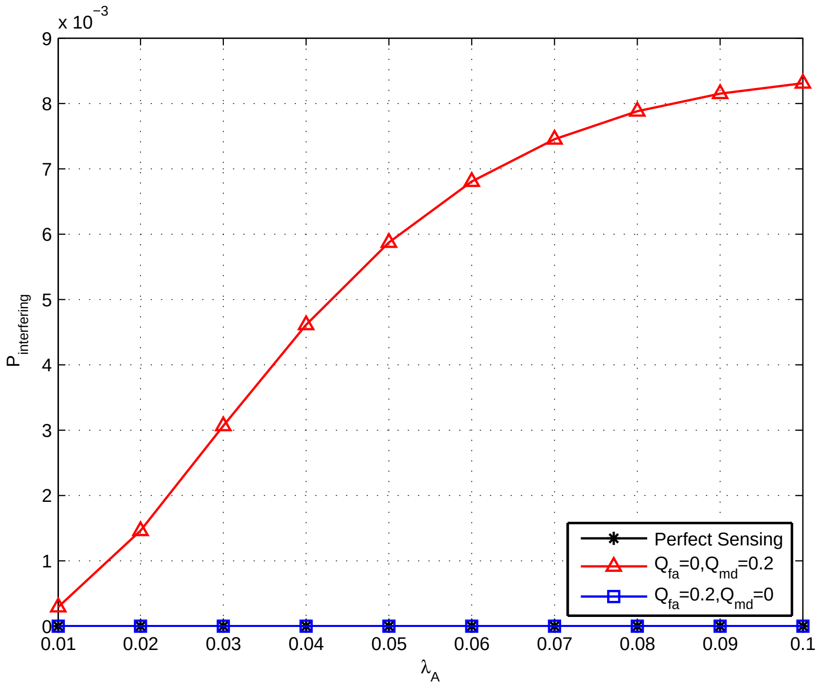


Figure 5.4: The interfering probability against the arrival rate of PU_A λ_A. The results are obtained from the analytical model. Q_{fa} and Q_{md} are the false alarm probability and miss-detection probability, respectively.

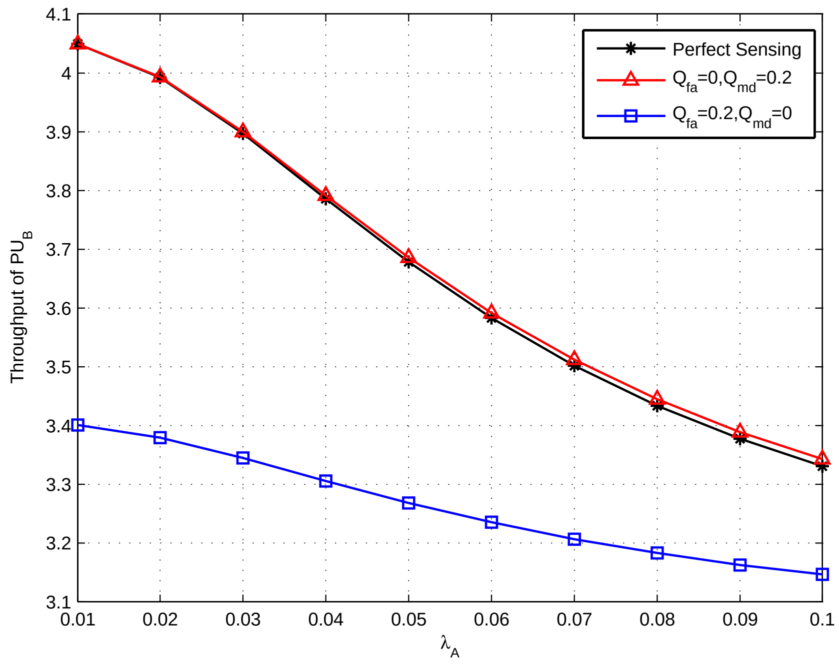
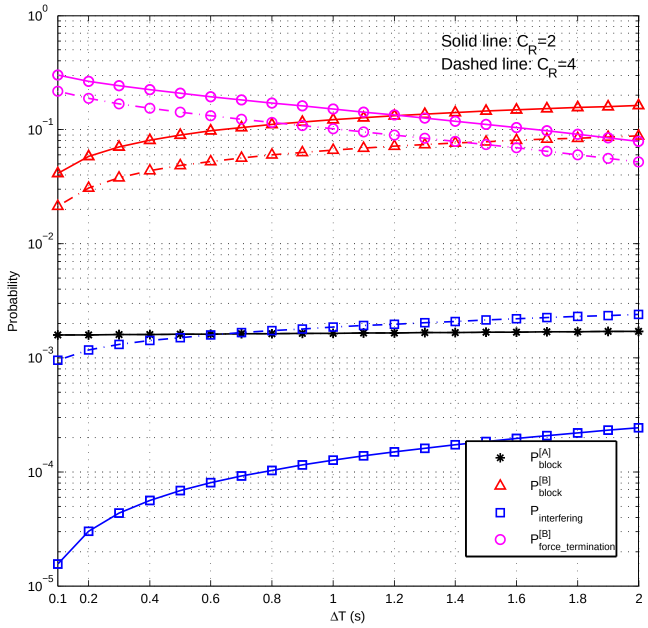


Figure 5.5: The throughput for PU_B against the arrival rate of PU_A λ_A. The results are obtained from the analytical model. C_R denotes the number of sharing channels. Q_{fa} and Q_{md} are the false alarm probability and miss-detection probability, respectively.

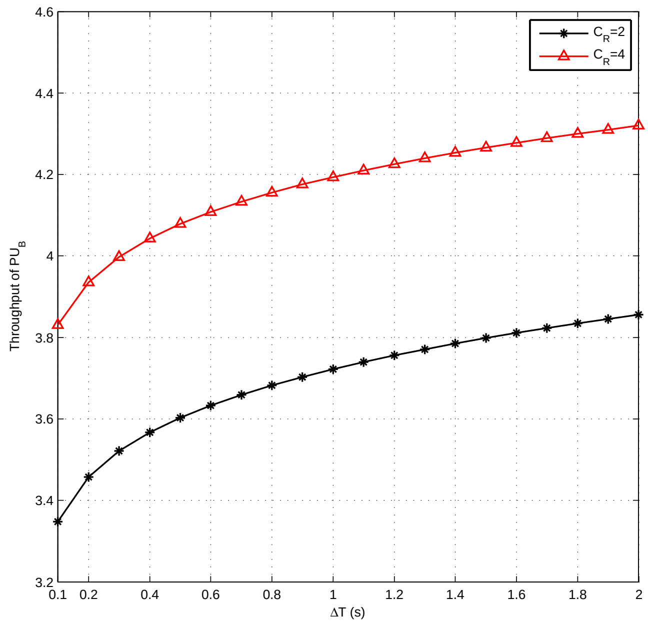
be obtained, due to the excessive channel access opportunities introduced by miss detection events.

Figure 5.6 shows the impact of sensing interval ΔT on the performance metrics. In Figure 5.6(a), $P^{[\mathsf{A}]}_{\mathsf{block}}$ does not change with ΔT. With a smaller sensing interval, a lower $P^{[\mathsf{B}]}_{\mathsf{block}}$ and $P^{[\mathsf{B}]}_{\mathsf{interfering}}$ can be obtained, due to up-to-date spectrum observation results. Also, $P^{[\mathsf{B}]}_{\mathsf{force\_terminate}}$ becomes larger due to more frequent spectrum observations, thus less PU_B users can complete their services. For the cases with different number of sharing channels C_R, i.e., the cases $C_R = 2$ and $C_R = 4$, $C_R = 4$ leads to a lower $P^{[\mathsf{B}]}_{\mathsf{block}}$ and $P^{[\mathsf{B}]}_{\mathsf{interfering}}$, and a larger $\tau^{[\mathsf{B}]}$, because of more accessing opportunities for newly arrived or handed off PU_B users. However, this also causes a higher interfering probability, due to more PU_B users operating in $\mathcal{C}_{\mathcal{R}}$. In Figure 5.6(b), with an increase of sensing interval, miss detection probability will increase, and thus a higher throughput of PU_B $\tau^{[\mathsf{B}]}$ can be obtained. This can be explained by recognizing the fact that SUs can benefit from more miss detection events, which can provide more access opportunities.

Figure 5.7 shows the near-optimal throughput of PU_B with its corresponding optimal parameters ΔT^*, λ_B^* and $\mathcal{C}_{\mathcal{R}}^*$, against λ_A. The hierarchical sharing [25; 36; 53] and non-sharing policies are used to compare with the proposed partial sharing schemes. In hierarchical sharing, PU_A are considered as the PUs, while PU_B are referred to as the SUs, which corresponds to the setting $C_A = 4$, $C_R = 4$, and $C_B = 4$ in the analytical model. The non-sharing case can be obtained by setting $C_A = 4$, $C_R = 0$, and $C_B = 4$ in the model. For the GoS constraints in (5.21), set $P^*_{\mathsf{interfering}} = 0.003, P^{[\mathsf{B}]*}_{\mathsf{block}} = 0.2, P^{[\mathsf{B}]*}_{\mathsf{force\_terminate}} = 0.05$. In Figure 5.7(a), the PSO algorithm achieves nearly the same throughput capacity as the exhaustive search (ES) method. It is also observed that the proposed partial sharing policy can achieve a higher throughput of PU_B than the conventional strategies. For the optimal parameters shown in Figure 5.7(b), with an increase of λ_A, λ_B^* decreases from 0.154 to 0.077, and thus can adapt a larger traffic of PU_A. Further, $\mathcal{C}_{\mathcal{R}}^*$ decreases from 3 to 0 to satisfy the interfering constraint. ΔT^* also decreases to provide up-to-date spectrum observation results. Note that for $\lambda_A \geq 0.07$, $\mathcal{C}_{\mathcal{R}}^* = 0$, which means that the large traffic of PU_A and the tight interfering constraints prevent PU_A sharing channels with PU_B. In this case, no spectrum sensing is required for PU_B, so ΔT^* only indicates the DTMC model observation interval.

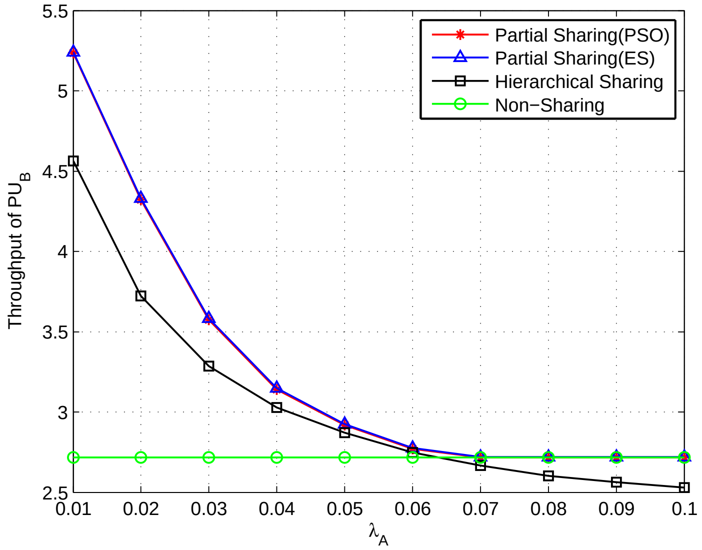


(a)

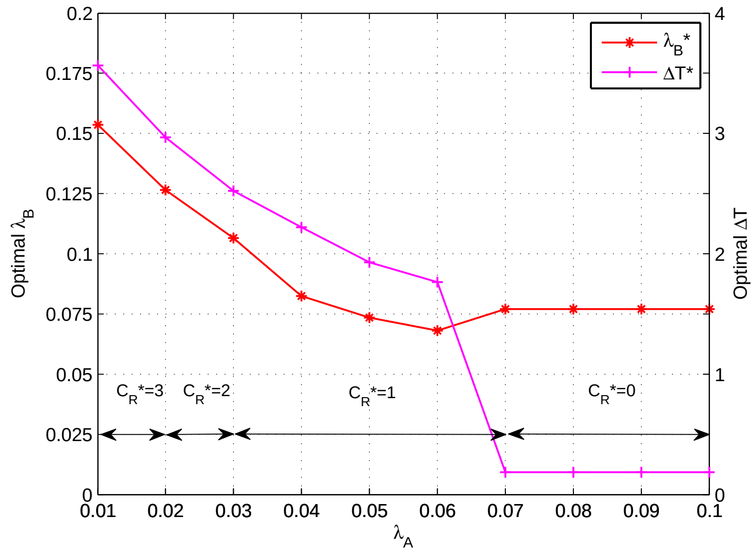


(b)

Figure 5.6: The impact of sensing interval ΔT on the (a) blocking, forced termination, interfering probabilities, and (b) throughput for PU_B with $Q_{\mathsf{fa}} = 0.1, Q_{\mathsf{md}} = 0.1$. Q_{fa} and Q_{md} are the false alarm probability and miss-detection probability, respectively.



(a)



(b)

Figure 5.7: Results for the (a) optimized throughput of PU_B $\tau^{[\mathsf{B}]}$. "Partial sharing" results are obtained by PSO algorithm and exhaustive search (ES), respectively. The hierarchical sharing and non-sharing schemes are obtained from the analytical model with PSO algorithm and used for comparison. (b) The optimal values for the cross layer optimization problem. The optimal arrival rate of PU_B λ_B^*, and the optimal sensing interval ΔT^* are obtained by PSO algorithm, while the optimal sharing channel $\mathcal{C}_{\mathcal{R}}^*$ is obtained by ES.

5.5 Chapter Conclusions

This work developed a DTMC model to study the proposed partial spectrum sharing scheme in the presence of imperfect spectrum sensing. Using the proposed analytical model, the impacts of sensing errors and periodic sensing intervals can be investigated. Our model and analytical method can also be used to study CR-based DSS systems with various sharing policies. A PHY-MAC cross-layer design framework was also devised, where the parameters in spectrum sensing and spectrum sharing are jointly optimized. Results show that the proposed partial sharing policy can achieve 7.4% and 20.8% gains in throughput, comparing with traditional hierarchical sharing and non-sharing strategies, respectively. By comparing with the results in Chapter 4, it can be concluded that the imperfect periodic sensing studied in this chapter can cause more severe deterioration in throughput than imperfect on-demand sensing.

The uncoordinated sharing does not require hardware updates for the primary systems, which can provide a good solution for future DSS networks. However, in uncoordinated sharing, the interference tolerance of PUs is ignored for overlay access, and the opportunity for SUs to transmit at higher power when PUs are inactive is missed for underlay access. In Chapter 6, a mixed or hybrid overlay-underlay strategies are proposed to further improve spectrum efficiency, which can also achieve a better trade-off between the network throughput and the interfering to PUs.

Chapter 6

On Hybrid Overlay-Underlay Dynamic Spectrum Access: Double-Threshold Energy Detection and Markov Model

In traditional uncoordinated DSS strategies, the interference tolerance of PUs is ignored for overlay sharing, while the opportunity for SUs to transmit at higher power when PUs are inactive is forbidden for underlay sharing. Therefore, mixed or hybrid strategies are preferred and have been used to improve spectrum efficiency. In this chapter, a hybrid full-partial access strategy for DSS is proposed. A double-threshold energy detection method is proposed to identify PU's operating location. Based on the spectrum observations, SUs can dynamically switch between full and partial access modes. A Markov chain model is developed to analyze the proposed strategy. Performance metrics such as the throughput of SUs and interfering probability to PU will be derived. The contribution is that this work jointly model and analyze spectrum sensing in the PHY layer with the hybrid access in the MAC layer.

6.1 System Model

6.1.1 Network Model

Consider a DSS network, consisting of N primary channels and a number of PUs and SUs. To make the analysis tractable, the work focus on a single primary channel with bandwidth W. The objective is to study the achievable throughput and the average interfering probability on this channel by introducing the SUs. In particular, there is a pair of SU transmitter (Tx) and receiver (Rx) in a given area of interest. From the SU's viewpoint, PUs may appear at random locations in the area with a radius D_1, as shown in Figure 6.1. It is also specified the shaded region within the radius D_2 as Region 2 in the figure and its operational meaning will be explained later. It is assumed that a PU can be referred to as a pair of PU Tx and Rx, which are located in the same region (Region 1 or 2), but their locations are independent and randomly distributed within that region. It is assumed that PU and SU transmit at fixed power levels P^t_{PU} and P^t_{SU}, respectively.

6.1.2 Proposed Hybrid Access Strategy

The received energy level over the channel with bandwidth W during a sensing time T_s is measured. For Rayleigh fading, the false alarm and detection probabilities are, respectively, given by [32]

$$\mathcal{Q}(v) = \Pr\left\{Y > v\,|H_0\right\} = \frac{\Gamma\left(u, \frac{v}{2}\right)}{\Gamma\left(u\right)} \tag{6.1}$$

$$\begin{aligned}\mathcal{O}(v,\overline{\gamma}) &= \Pr\left\{Y > v \,|H_1\right\}\\ &= e^{-\frac{v}{2}}\sum_{k=0}^{u-2}\frac{1}{k!}\left(\frac{v}{2}\right)^k + \left(\frac{1+\overline{\gamma}}{\overline{\gamma}}\right)\left(e^{-\frac{v}{2(1+\overline{\gamma})}} - e^{-\frac{v}{2}}\sum_{k=0}^{u-2}\frac{1}{k!}\left(\frac{v\overline{\gamma}}{2\left(1+\overline{\gamma}\right)}\right)^k\right),\end{aligned} \tag{6.2}$$

where H_0 and H_1 denote the hypothesis that PU is inactive and active in the channel, respectively, Y is the decision statistics, v is the threshold for energy detection, $\overline{\gamma}$ is the average received signal-to-noise ratio (SNR) at the SU Rx,

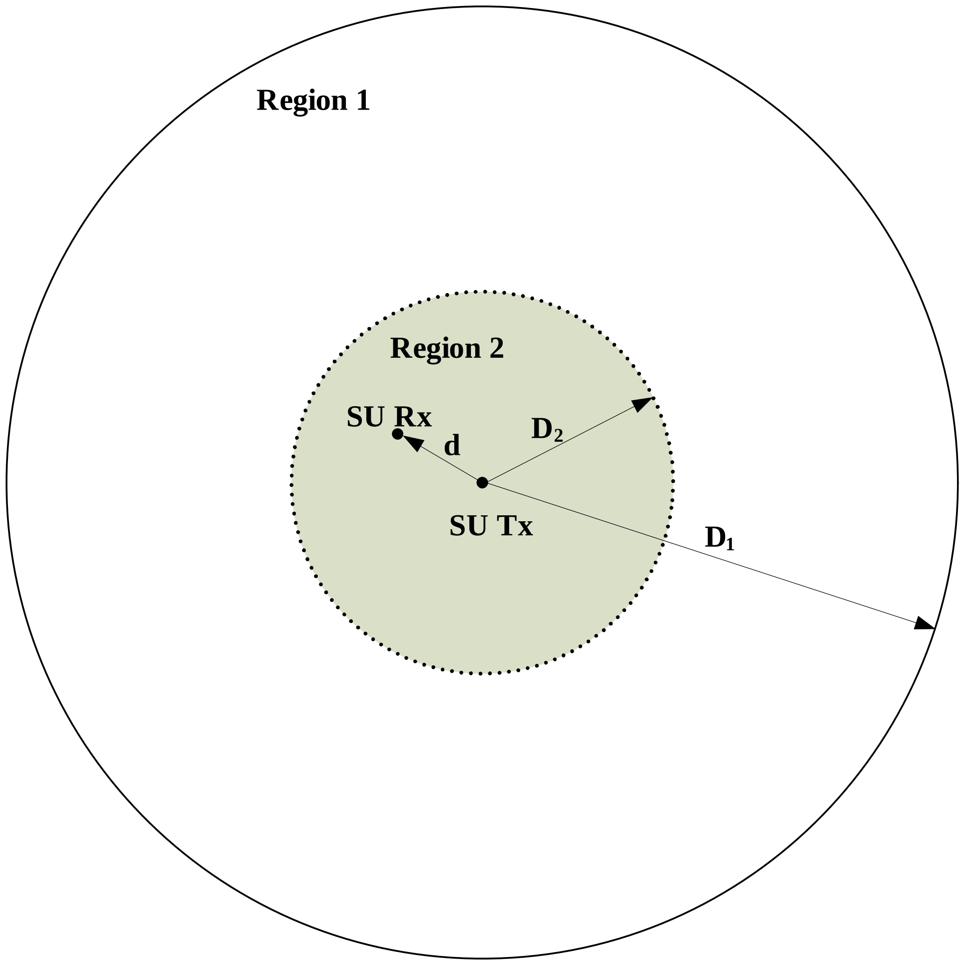


Figure 6.1: The DSS system model with the hybrid access strategy.

and $u = T_s W$ is the time-bandwidth product. In addition, $\Gamma(\cdot)$ and $\Gamma(\cdot,\cdot)$ are the complete and incomplete gamma functions, respectively.

When the PU is near the SU (i.e., PU in Region 2), the average received SNR at the SU Tx is high and hence the PU detection probability will be much greater than the case when the PU appears in Region 1. Moreover, if a miss-detection event happens when the PU is in Region 2, it is more likely that the SU's power will violate the interference constraints of the PU. In the proposed hybrid access strategy, a double-threshold energy detection method is therefore used to identify the operating location of the PU (Region 1 or 2). Based on the identification, the SU may switch between full and partial access modes.

In the considered scenario, double-thresholds, υ_{L} and υ_{H}, are used to identify the operating location of the PU. As a result, the following three-hypothesis testing is established:

$$H_0 : \mathsf{PU} \text{ is inactive in the channel;} \tag{6.3a}$$

$$H_1 : \mathsf{PU} \text{ is active and transmitting in Region 1;} \tag{6.3b}$$

$$H_2 : \mathsf{PU} \text{ is active and transmitting in Region 2.} \tag{6.3c}$$

The following probabilities can be defined:

$$P_0(\mathsf{H}) = \Pr\{Y > v_{\mathsf{H}} \,|H_0\} = \mathcal{Q}(v_{\mathsf{H}}), \tag{6.4a}$$

$$P_0(\mathsf{L}) = \Pr\{v_{\mathsf{L}} < Y \leq v_{\mathsf{H}} \,|H_0\} = \mathcal{Q}(v_{\mathsf{L}}) - P_0(\mathsf{H}), \tag{6.4b}$$

$$P_1(\mathsf{H}) = \Pr\{Y > v_{\mathsf{H}} \,|H_1\} = \mathcal{O}(v_{\mathsf{H}}, \overline{\gamma}_{\mathsf{L}}), \tag{6.4c}$$

$$P_1(\mathsf{L}) = \Pr\{v_{\mathsf{L}} < Y \leq v_{\mathsf{H}} \,|H_1\} = \mathcal{O}(v_{\mathsf{L}}, \overline{\gamma}_{\mathsf{L}}) - P_1(\mathsf{H}), \tag{6.4d}$$

$$P_2(\mathsf{H}) = \Pr\{Y > v_{\mathsf{H}} \,|H_2\} = \mathcal{O}(v_{\mathsf{H}}, \overline{\gamma}_{\mathsf{H}}), \tag{6.4e}$$

$$P_2(\mathsf{L}) = \Pr\{v_{\mathsf{L}} < Y \leq v_{\mathsf{H}} \,|H_2\} = \mathcal{O}(v_{\mathsf{L}}, \overline{\gamma}_{\mathsf{H}}) - P_2(\mathsf{H}), \tag{6.4f}$$

where $\overline{\gamma}_{\mathsf{L}}$ and $\overline{\gamma}_{\mathsf{H}}$ denote the average received SNR when PU is, respectively, operating in Region 1 and Region 2. As a consequence, when the measured energy level in the channel is lower than v_{L}, SU is operating in the full access mode, i.e., can access the channel with probability 1. If the energy level exceeds v_{L} but is lower than v_{H}, then the SU operates in the partial access mode, i.e., can access the channel with the probability P_{access}. Otherwise, SU should stop transmitting and stay inactive in this channel.

The spectrum sensing interval is defined as the time slot between two consecutive spectrum observations. To achieve a tradeoff between SU throughput and interference of PU, two types of sensing intervals are considered, i.e., the long and short intervals. In the proposed scheme, when the measured energy level exceeds v_{L} but is lower than v_{H}, the SU operates in the partial access mode and a short sensing interval is scheduled. Otherwise, a long sensing interval is carried out, when the energy level is lower than v_{L} or higher than v_{H}.

6.2 Model Formulation and Analysis

Here, the proposed hybrid access strategy in the presence of imperfect spectrum sensing is analyzed. A DTMC model is used to model the PHY layer sensing and the MAC layer access.

6.2.1 Assumptions

- For convenience, the long and short spectrum sensing interval are defined as $\Delta T_L = M\tau$ and $\Delta T_S = M'\tau$, respectively, where M, M' are positive integers. τ denotes a basic unit for the system observation interval.
- Consider a single licensed frequency channel and model the PU's access process as a two-state ON/OFF process, which specifies the time slots when PU is active (ON) or inactive (OFF).
- At the beginning of each model observation, the probability that PU transfers from inactive to active is λ, while the probability that PU transfers from active to inactive is μ.
- When PU becomes active, the probability that PU appears in Region 1 is δ. It is assumed that the model observation interval is much smaller than the average PU active period. Therefore, the probability that PU appears in two regions successively in one observation interval is negligible.

6.2.2 Model Formulation

A DTMC model is developed to analyze the proposed hybrid access strategy. Let $E_{(i,j)}$ and $T_{(i,j,k)}$ denote the energy detection state and normal transmission (non-sensing) state of the system model, respectively. Specifically, i indicates the status of PU in this channel. If PU is inactive, $i = 0$, and when PU is in Region 1 and 2, $i = 1$ and $i = 2$, respectively. Additionally, j represents the access mode of SU. If SU is inactive, $j = 0$, and when SU is operating in the partial and full access mode, $j = 1$ and $j = 2$, respectively. To make the analysis tractable, the short sensing interval is defined as $\Delta T_S = \tau$. Likewise, k denotes the number of observation intervals passed during a long sensing interval ΔT_L. The state space is then given by

$$\Psi = \left\{ E_{(i,j)}, T_{(i,j,k)} \middle| \begin{array}{c} i, j \in \{0, 1, 2\}, \\ \\ k = 1, 2, \ldots, M-1 \end{array} \right\}. \tag{6.5}$$

Figure 6.2 depicts the state transition diagrams for the proposed system. To keep the figure simple, only the outflows of the state transitions are shown in this figure. Figure 6.3(upper) shows an example of the state transitions from the energy detection state to the next energy detection or normal transmission state, i.e., from the state $E_{(i,j)}$ to $E_{(i+s,j+l)}$ or $T_{(i+s,j+l,1)}$, where $i, i+s, j, j+l \in \{0, 1, 2\}$. In this case, the PU operating mode transfers from i to $i+s$, which is caused by the status changes of communicating with PU Rx and is independent of spectrum sensing. As mentioned in the previous section, it is assumed that the probability that PU appears in two regions successively in one observation interval is negligible, i.e., do not consider the transitions from $i = 1$ to $i = 2$ or from $i = 2$ to $i = 1$ in a single observation interval τ. Hence, s can be defined as, for $i = 0$, $s \in \{0, 1, 2\}$; for $i = 1$, $s \in \{-1, 0\}$; and for $i = 2$, $s \in \{-2, 0\}$. Similarly, the SU operating mode can transfer from j to $j + l$, depending on the spectrum sensing results. For $j+l = 1$, the SU is operating in partial access mode

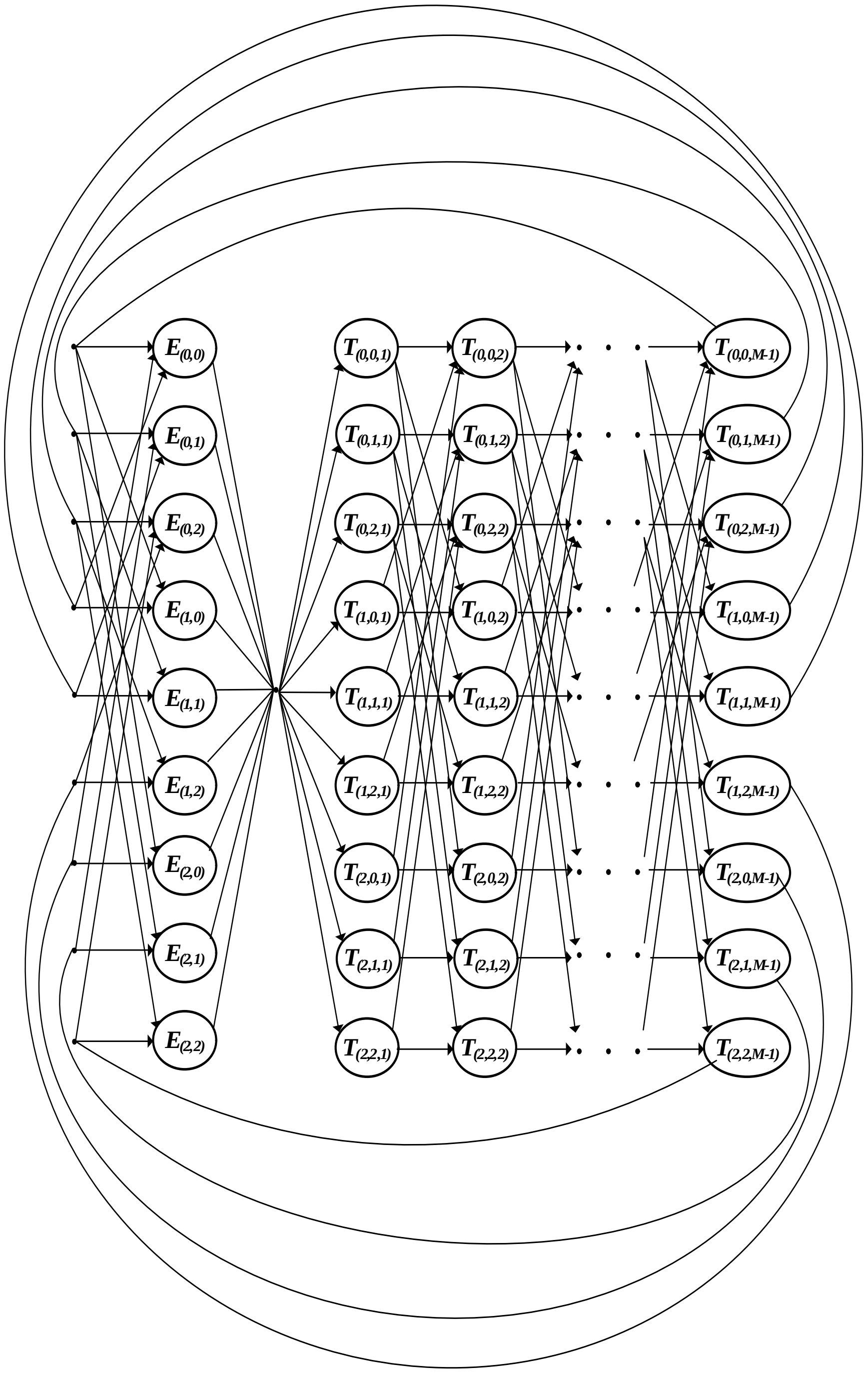


Figure 6.2: State transition diagrams for the hybrid access Markov model. $E_{(i,j)}$ and $T_{(i,j,k)}$ denote the energy detection state and normal transmission state, respectively. i is the status of PU, i.e., if PU is inactive, $i = 0$, and when PU is in Region 1 and 2, $i = 1$ and $i = 2$, respectively. j represents the access mode of SU, i.e., if SU is inactive, $j = 0$, and when SU is in the partial and full access mode, $j = 1$ and $j = 2$, respectively.

and a short sensing interval is scheduled. Hence, $E_{(i,j)}$ will transfer to $E_{(i+s,j+l)}$. For $j+l \in \{0,2\}$, the SU is inactive or operating in full access mode, and thus a long sensing interval is scheduled. Hence, $E_{(i,j)}$ transfers to $T_{(i+s,j+l,1)}$. It is further assumed that energy detection is performed at the end of the transitions from $E_{(i,j)}$ to $E_{(i+s,j+l)}$ or $T_{(i+s,j+l,1)}$.

For example, in the case when both PU and SU are inactive, i.e., the state $E_{(0,0)}$, if a PU in Region 2 starts transmitting and the measured energy level is higher than v_{H}, the system will transfer to the state $T_{(2,0,1)}$, with the probability $\delta\lambda P_2(\mathsf{H})$, where $P_2(\mathsf{H})$ is given by (6.4e). On the other hand, if the measured energy level is lower than v_{L} and PU stays inactive, the system will transfer to the state $T_{(0,2,1)}$, with the probability $(1-\lambda)\,(1-P_0(\mathsf{H})-P_0(\mathsf{L}))$. For the short sensing interval case, if a PU in Region 1 starts transmitting and the measured energy level is larger than v_{L} but lower than v_{H}, the system will transfer to the state $E_{(1,1)}$, with the probability $(1-\delta)\lambda P_1(\mathsf{L})$.

As such, the state transition probabilities from $E_{(i,j)}$ to $E_{(i+s,j+l)}$ and $T_{(i+s,j+l,1)}$, respectively, $\gamma_1^{(i,s,j,l)}$ and $\gamma_2^{(i,s,j,l)}$, can be expressed as

$$\gamma_1^{(i,s,j,l)} = \theta_{i,s}\phi_{i,j+l} \text{ for } j+l=1, \tag{6.6}$$

$$\gamma_2^{(i,s,j,l)} = \theta_{i,s}\phi_{i,j+l} \text{ for } j+l \in \{0,2\}, \tag{6.7}$$

where $\theta_{i,s}$ and $\phi_{i,j}$ are the functions related to the changes of PU's operating

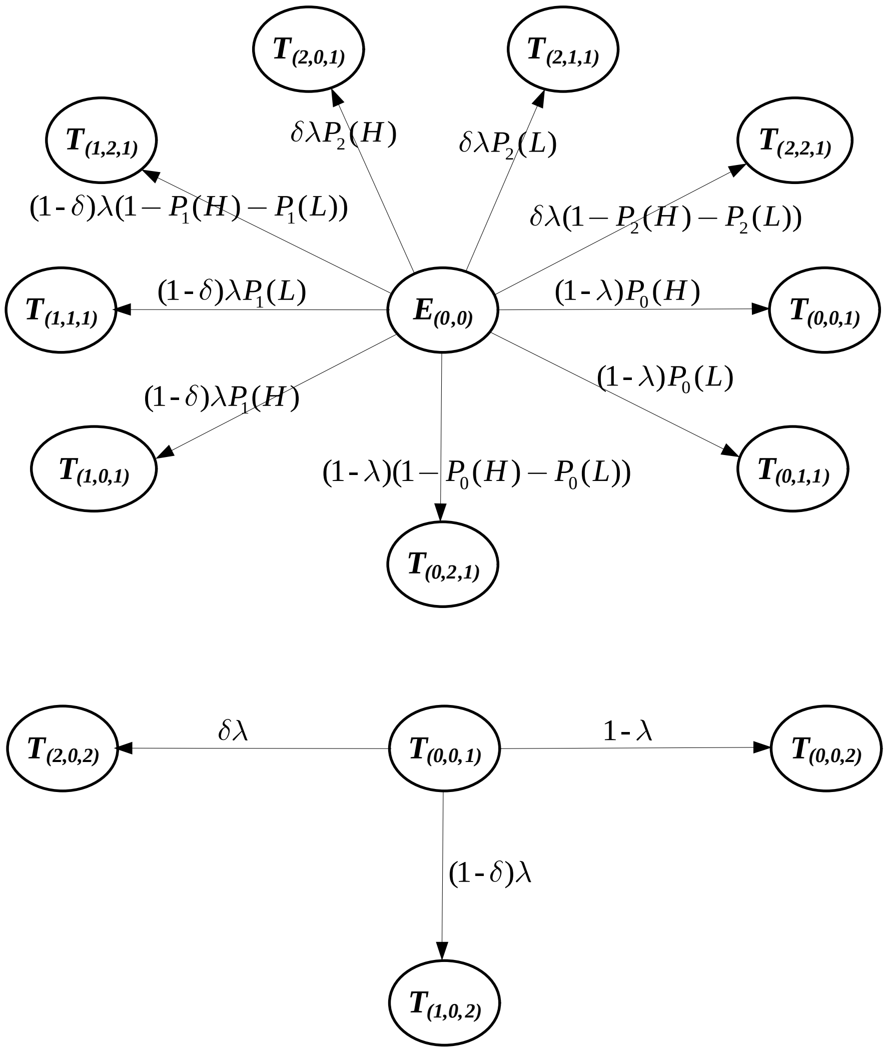


Figure 6.3: Examples of the detailed state transition diagrams. Upper: Transitions from the state $E_{(i,j)}$ to $E_{(i+s,j+l)}$ and $T_{(i+s,j+l,1)}$. Bottom: Transitions from the state $T_{(i,j,k)}$ to $T_{(i+s,j,k+1)}$. i is the status of PU, i.e., if PU is inactive, $i = 0$, and when PU is in Region 1 and 2, $i = 1$ and $i = 2$, respectively. j represents the access mode of SU, i.e., if SU is inactive, $j = 0$, and when SU is in the partial and full access mode, $j = 1$ and $j = 2$, respectively.

location mode and the energy detection results, which are, respectively, given by

$$\theta_{i,s} = \begin{cases} 1-\lambda & \text{for } i=0,\ s=0, \\ (1-\delta)\lambda & \text{for } i=0,\ s=1, \\ \delta\lambda & \text{for } i=0,\ s=2, \\ (1-\mu) & \text{for } i\in\{1,2\},\ s=0, \\ \mu & \text{for } i\in\{1,2\},\ s=-i, \end{cases} \tag{6.8}$$

$$\phi_{i,j+l} = \begin{cases} P_i(\mathsf{H}) & \text{for } j+l=0, \\ P_i(\mathsf{L}) & \text{for } j+l=1, \\ 1-P_i(\mathsf{H})-P_i(\mathsf{L}) & \text{for } j+l=2, \end{cases} \tag{6.9}$$

where $P_i(\mathsf{H})$ and $P_i(\mathsf{L})$ can be obtained using (6.4a)–(6.4f).

Figure 6.3(bottom) shows an example of the state transitions between normal transmission states, i.e., from the state $T_{(i,j,k)}$ to $T_{(i+s,j,k+1)}$ during long sensing interval, where $i, i+s \in \{0,1,2\}$, $j \in \{0,2\}$ and $1 \le k < M-1$. Since the changes of SU's operating mode only happen at the energy detection states, j will not change in this case. Therefore, to summarize the transition probabilities in this case, $\gamma_3^{(i,s)}$ is given by

$$\gamma_3^{(i,s)} = \theta_{i,s}. \tag{6.10}$$

For the state transitions from normal transmission states to energy detection state, i.e., from the state $T_{(i,j,M-1)}$ to $E_{(i+s,j)}$, where $i, i+s \in \{0,1,2\}$, $j \in \{0,2\}$. Similar to the above case, j will not change in such state transitions. Therefore, the transition probabilities can be defined by $\gamma_3^{(i,s)}$, which is given by

$$\gamma_4^{(i,s)} = \theta_{i,s}. \tag{6.11}$$

Let $\pi(E_{(i,j)})$ and $\pi(T_{(i,j,k)})$ denote the steady state probabilities of $E_{(i,j)}$ and $T_{(i,j,k)}$, respectively. Let $\mathbf{\Pi}$ denote the row vector of steady state probabilities with elements $\left\{\pi(E_{(i,j)}), \pi(T_{(i,j,k)})\right\}$, and $\mathbf{Y}$ denote the transition rate matrix, whose elements $\left\{\gamma_1^{(i,s,j,l)}, \gamma_2^{(i,s,j,l)}\gamma_3^{(i,s)}, \gamma_4^{(i,s)}\right\}$ can be obtained from the analysis above. The steady state matrix $\mathbf{\Pi}$ can be obtained by solving the linear equations $\mathbf{\Pi} = \mathbf{\Pi}\mathbf{Y}$ and $\mathbf{\Pi}\mathbf{e} = 1$, where $\mathbf{e}$ is the column vector with all ones.

6.2.3 Performance Analysis

Here, some important performance metrics are derived.

1) Throughput of SU: The throughput is measured by the average achievable transmission rate in bits per second (bps). However, if the SU causes harmful interference to PU, i.e., when the PU's interference constraints are violated, the throughput achieved during these periods will not be counted. Considering the sensing efficiency, the Shannon bound is used to obtain the achievable rate in a given bandwidth as

$$R = \eta W \log_2(1 + \gamma_{\mathsf{SU}}), \tag{6.12}$$

where γ_{SU} is the SNR of the SU, and η is the sensing efficiency, as $\eta = 1 - \frac{T_s}{\Delta T}$, where $\Delta T \in \{\Delta T_L, \Delta T_S\}$, $\eta \in \{\eta_L, \eta_S\}$ correspond to the long and short sensing intervals, respectively. Let P_{SU}^r denote the average SU received power. When PU is inactive, the SNR of SU in this case can be calculated as

$$\overline{\gamma}_{\mathsf{SU}} = \frac{P_{\mathsf{SU}}^r}{N_0 W}, \tag{6.13}$$

where N_0 is the one-sided noise power. On the other hand, if PU is active in Region 1 or Region 2, the SNR of SU in the two cases can be found as

$$\overline{\gamma}^1_{\mathsf{SU}} = \frac{P^r_{\mathsf{SU}}}{N_0 W + P^{r_1}_{\mathsf{PU}}}, \quad \text{(PU in Region 1)}, \tag{6.14}$$

$$\overline{\gamma}^2_{\mathsf{SU}} = \frac{P^r_{\mathsf{SU}}}{N_0 W + P^{r_2}_{\mathsf{PU}}}, \quad \text{(PU in Region 2)}, \tag{6.15}$$

where $P^{r_1}_{\mathsf{PU}}$ and $P^{r_2}_{\mathsf{PU}}$ are the average received PU signal powers at the SU Rx in the two cases.

The average throughput of SU can be derived as

$$Th = \sum_{i\in\{0,1,2\}} \varepsilon_i \left[\sum_{j\in\{0,2\},1\le k\le M} \sigma_j \eta_L \left(\pi(E_{(i,j)}) + \pi(T_{(i,j,k)})\right) + \sigma_1 \eta_S \pi(E_{(i,1)}) \right], \tag{6.16}$$

where

$$\varepsilon_i = \begin{cases} R_{\mathsf{SU}} & \text{for } i = 0, \\ R^1_{\mathsf{SU}}(1 - P^1_{\mathsf{ICV}}) & \text{for } i = 1, \\ R^2_{\mathsf{SU}}(1 - P^2_{\mathsf{ICV}}) & \text{for } i = 2, \end{cases} \tag{6.17}$$

in which R_{SU}, R^1_{SU}, and R^2_{SU} are the achievable rates when the SU SNR are, respectively, $\overline{\gamma}_{\mathsf{SU}}$, $\overline{\gamma}^1_{\mathsf{SU}}$, and $\overline{\gamma}^2_{\mathsf{SU}}$. In addition, interference constraints violation probability is defined as the probability that PU's received SU's power exceeds its interference constraint. P^1_{ICV} and P^2_{ICV} are the average interference constraints violation probability to PU in Region 1 and 2, respectively, as [69, (2.52)]. Also, σ_j is the function that defines the access probability when SU is operating in full and partial access mode, which is given by

$$\sigma_j = \begin{cases} 0 & \text{for } j = 0, \\ P_{\mathsf{access}} & \text{for } j = 1, \\ 1 & \text{for } j = 2. \end{cases} \tag{6.18}$$

2) Interfering Probability: Interfering is said to occur when SU causes harmful interference to PU, i.e., when SU fails to detect PU's signal, and SU's transmission power violates the interference constraints of PU. Hence, the interfering probability is given by

$$P_{\mathsf{In}} = P^1_{\mathsf{ICV}} \sum_{j\in\{1,2\},1\le k\le M} \sigma_j \left[\pi\left(E_{(1,j)}\right) + \pi\left(T_{(1,j,k)}\right)\right] + \\ P^2_{\mathsf{ICV}} \sum_{j\in\{1,2\},1\le k\le M} \sigma_j \left[\pi\left(E_{(2,j)}\right) + \pi\left(T_{(2,j,k)}\right)\right]. \tag{6.19}$$

6.3 Numerical Results

In this section, the proposed hybrid access strategy is evaluated, by comparing with the conventional overlay access schemes. In particular, the performance metrics of the overlay scheme can be obtained using the analytical model by using a single energy detection threshold.

6.3.1 Simulation Parameters

The radii of Region 1 and Region 2 are set as $D_1 = 300$ m and $D_2 = 100$ m, which can be referred to as the outer and inner areas of a building, respectively. Note that D_2 is an SU-defined parameter, which can be viewed as a threshold for the received PU power or the size of a building. For practical network implementation, when D_2 is defined, the probability of the PU appears in Region 2, δ, can be measured. The active PU Tx and Rx are assumed to be independent and uniformly distributed in the same region (Region 1 or 2), respectively. In the simulations, when a PU is active, $\delta = 0.7$ and SU Tx and Rx have fixed locations, with $d_{SU} = 50$ m distance between them. The bandwidth of a single channel is $W = 1$ MHz, and the noise power $N_0 = -163$ dBm/Hz. Also, the combined path loss and shadowing model [69] is used to consider the effects of path loss and shadowing. To model shadowing, a log-normal random variable with mean zero and variance $\sigma_{\psi_{dB}} = 3.65$ dB is used. For PU appearing at different locations, the average received power by SU Tx can be calculated using [69, (2.40),(2.45)], as

$$P^r_{SU}(dB) = P^t_{PU}(dB) + K(dB) - 10\gamma \log_{10} \frac{d}{d_0} + \frac{\sigma^2_{\psi_{dB}}}{20} \ln 10 \tag{6.20}$$

where the d is the distance between PU Rx and SU Tx, and the parameters are defined as the unitless constant $K = -31.54$ dB, PU transmission power $P^t_{\mathsf{PU}} = 13$ dBm, the path loss exponent $\gamma = 3.71$, and the reference distance $d_0 = 1$ m. The parameters of path loss and shadowing model are based on [69, Example 2.4], which are typical settings for indoor or outdoor wireless propagation. For the double-threshold energy detection, v_{H} and v_{L} are obtained for various false alarm probabilities, 0.001 and 0.1, respectively, i.e., $v_{\mathsf{H}} = \mathcal{Q}^{-1}(0.001)$ and $v_{\mathsf{L}} = \mathcal{Q}^{-1}(0.1)$ using (6.1). Using [69, (2.52)], the probability that SU's power exceeds the interference constraints at a given distance between SU Tx and PU Rx is found as,

$$\begin{aligned} P_{\mathsf{ICV}}(d) &= \Pr(P^r_{PU}(d) \geq P_{constraint}) \\ &= Q\left(\frac{P_{constraint} - (P^t_{SU} + K - 10\gamma \log 10(d/d_0))}{\sigma_{\psi_{dB}}}\right) \end{aligned} \tag{6.21}$$

where Q(.) denotes the Q-function, the SU transmission power $P^t_{\mathsf{SU}} = 13$ dBm, and the interference constraint to PU is set as -100 dBm. In the above equation, we can see that P_{ICV} is the function of the distance between PU Rx and SU Tx, and P_{ICV} will decrease when d increases. Hence, SU would have a very large probability to cause harmful interference in the case of miss detecting PU in Region 2.

10^5 locations of PU are generated uniformly and independently in Region 1 and Region 2, respectively. Using (6.20), the average received PU power by SU Tx in Region 1 and 2 can be calculated, respectively. Then, calculate the detection probabilities in (6.4a)–(6.4f), and the steady state probabilities of the DTMC model can be obtained. Based on the state probabilities, the average received PU power at SU Rx in (6.14) and (6.15), and the average interference constraints violation probabilities in (6.16) can be derived. It is defined the model observation interval as $\tau = 100$ ms, and the time for conducting energy detection is $T_s = 50$ μs. For the sensing intervals, the settings are $\Delta T_L = 5\tau = 500$ ms, $\Delta T_S = \tau = 100$ ms, respectively. The probability that PU transfers from active to inactive during a model observation is $\mu = \frac{1}{200}$. If not specified otherwise, $\lambda = \frac{1}{600}$, which corresponds to the mean PU channel utilization rate 25%. P_{access} is set to 0.9 for the hybrid access scheme. The simulation parameters are summarized in Table 6.1.

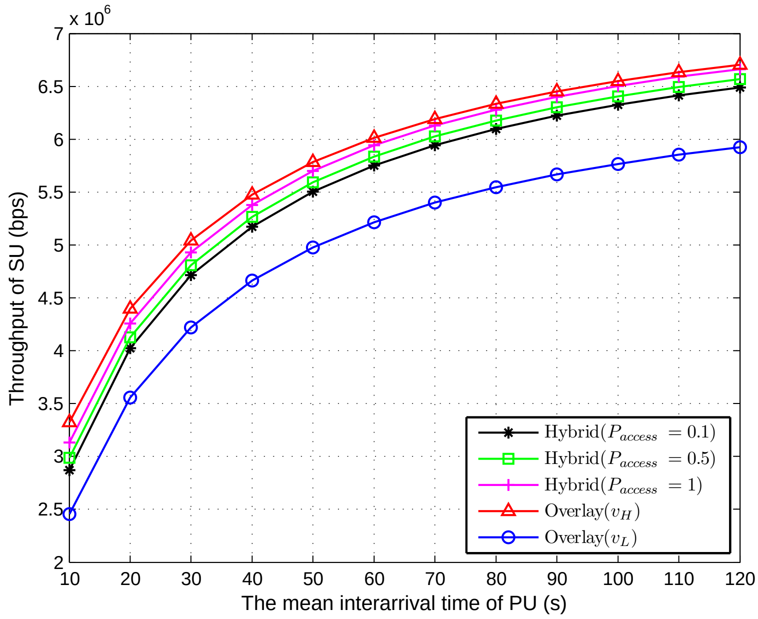
Table 6.1: Simulation Parameters.

The radii of Region 1 and Region 2	$D_1 = 300$ m, $D_2 = 100$ m
The probability appears in Region 2	$\delta = 0.7$
The distance between SU Tx and Rx	$d_{SU} = 50$ m
Channel bandwidth	$W = 1$ MHz
Noise power	$N_0 = -163$ dBm/Hz
Log-normal random variable for shadowing	mean zero and variance $\sigma_{\psi_{dB}} = 3.65$
Path loss exponent in (6.20),(6.21)	$\gamma = 3.71$
Unitless constant in (6.20),(6.21)	$K = -31.54$ dB
The reference distance in (6.20),(6.21)	$d_0 = 1$ m
PU transmission power	$P^t_{\mathsf{PU}} = 13$ dBm
SU transmission power	$P^t_{\mathsf{SU}} = 13$ dBm
Interference constraint to PU	$P_{constraint} = -100$ dbm
The model observation interval	$\tau = 100$ ms
The time for conducting energy detection	$T_s = 50$ μs
The long and short sensing intervals	$\Delta T_L = 500$ ms, $\Delta T_S = 100$ ms
P_{access} for partial access mode $P_{\mathsf{access}} = 0.9$	
The probability that PU transfers from active to inactive during τ	$\mu = \frac{1}{200}$
The probability that PU transfers from inactive to active during τ	$\lambda = \frac{1}{600}$

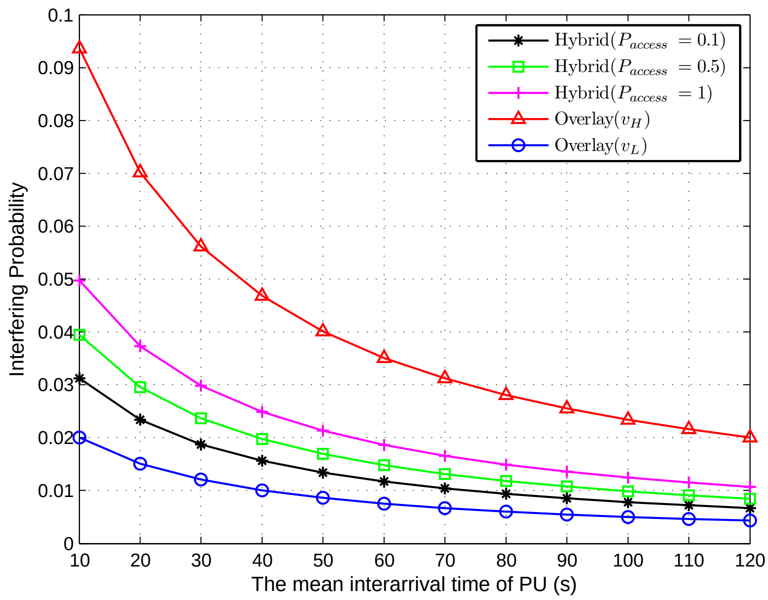
6.3.2 Simulation Results

It is defined that the probability that PU transfers from inactive to active at each observation interval as $\frac{1}{1200} \leq \lambda \leq \frac{1}{100}$, which corresponds to a mean PU channel utilization rate of 14.3%∼66.7%. Hence, the mean interarrival time of PU is $10 \leq \tau/\lambda \leq 120$ s. Figure 6.4 provides the results for throughput of SU and interfering probability against the mean interarrival time of PU. In Figure 6.4, it is observed that the proposed hybrid strategy achieves greater throughput than the overlay schemes with the single detection threshold v_{L}. In contrast to this overlay scheme, the SU in the hybrid strategy can share the channel with an active PU via partial access mode, in which SU will be less likely to cause harmful interference to the PU. As shown in Figure 6.4, it is also observed that the hybrid strategy achieves a much lower interfering probability than the overlay scheme with v_{H}. The reason is that the double-threshold energy detection scheme can provide a false alarm-detection performance tradeoff to near-side (Region 2) and far-end (Region 1) PUs by arranging the long and short sensing intervals. Furthermore, the performance gain over overlay schemes is more noticeable when there is a lower mean interarrival time of PU, i.e., higher PU channel utilization rate. In particular, the hybrid strategy with $P_{\mathsf{access}} = 1$ can achieve 46.86% improvement on interfering performance, at the cost of 1.28% decreasing of throughput, as compared with the overlay scheme with v_{H}. By adjusting the partial access probability P_{access}, a tradeoff between throughput and interfering probability can be adjusted, following the system QoS requirements.

Figure 6.5 shows the throughput and interfering results against the sensing interval ΔT. As we can see, with a shorter sensing interval, a better through-

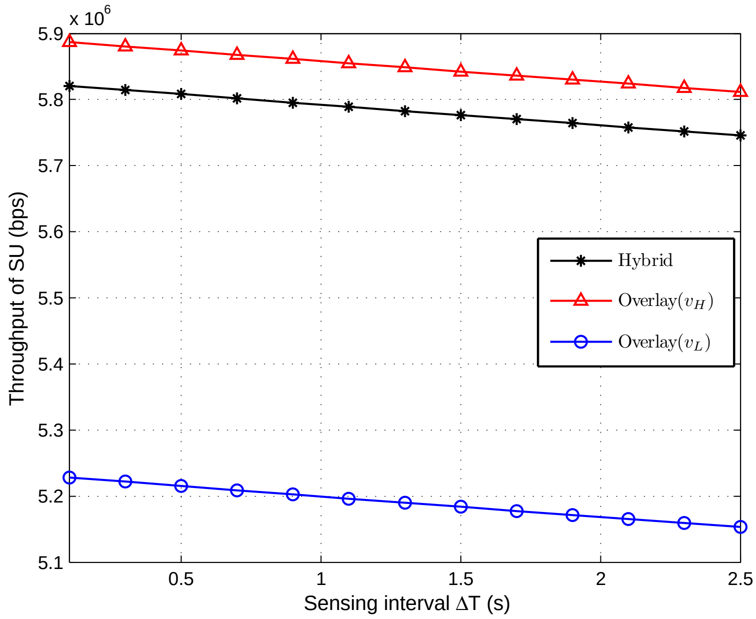


(a)

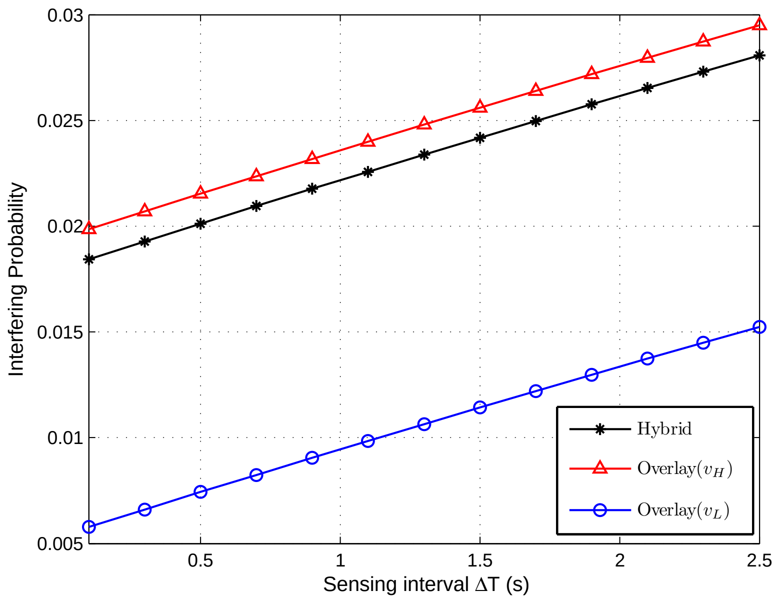


(b)

Figure 6.4: Results for (a) the throughput of SU and, (b) interfering probability against the mean interarrival time of PU. The results for the Hybrid scheme with various probabilities of SU transmitting in partial access mode are plotted. The overlay schemes with the single detection threshold v_{H} and v_{L} are plotted for comparison.



(a)



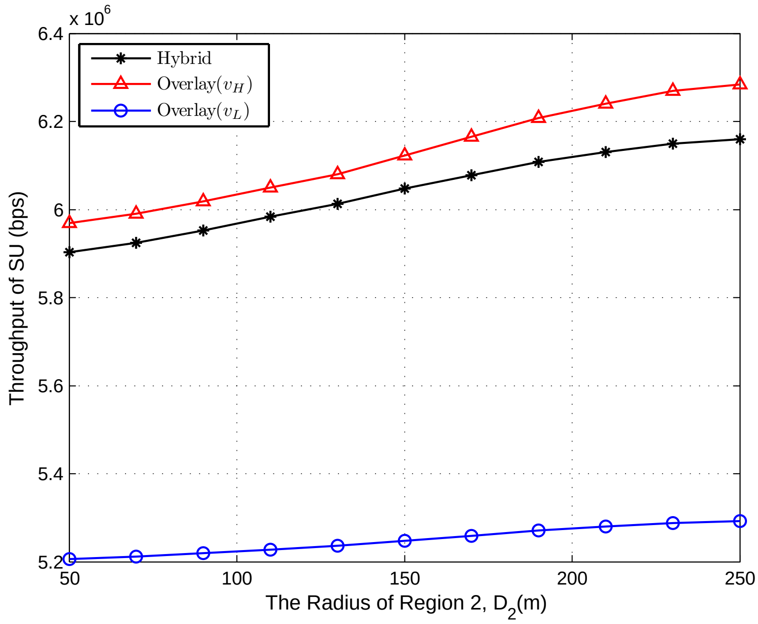
(b)

Figure 6.5: Results for (a) the throughput of SU and, (b) interfering probability against the sensing interval. The results for the Hybrid scheme with $P_{\mathsf{access}} = 0.9$ are compared with the traditional overlay schemes with the single detection threshold v_{H} and v_{L}.

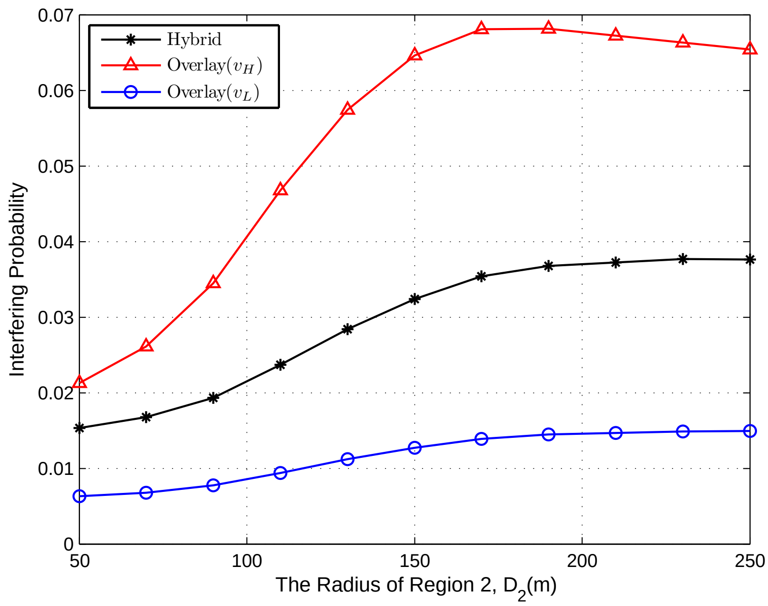
put and interfering performance can be obtained, due to the updated spectrum observations, and results in less miss detection events.

In Figure 6.6, the impact of the radius of Region 2 D_2 is studied. With a larger value of D_2, the average interference constraints violation probability in Region 1 and Region 2 decreases. As a result, a larger throughput can be achieved, as shown in Figure 6.6(a). On the other hand, in Figure 6.6(b), the interfering probability increases with D_2 first, and then decreases, in all three hybrid and overlay access schemes. This can be explained by recognizing the fact that interfering probability is jointly determined by the miss-detection probability and average interference constraints violation probability. With D_2 increasing, SU has a greater chance to miss detect the PU's signal, which results in the increasing of miss-detection probability. However, although SU may miss detect PU's signal and start transmission, it is less likely that the transmission will cause harmful interference to PU, due to a lower PU's interference constraint violation probability.

Figure 6.7 illustrates the effect of the probability that PU appears in Region 2 δ. It is observed that with a larger probability that PU appears in Region 2, both interfering and throughput decrease for the proposed hybrid strategy. Due to the lower average miss detection probability in Region 2, if more PUs appear, there will be less chance that PU will be interfered. On the other hand, it is more likely that SU stays inactive rather than operates in partial access mode, and therefore a lower throughput can be observed as expected.

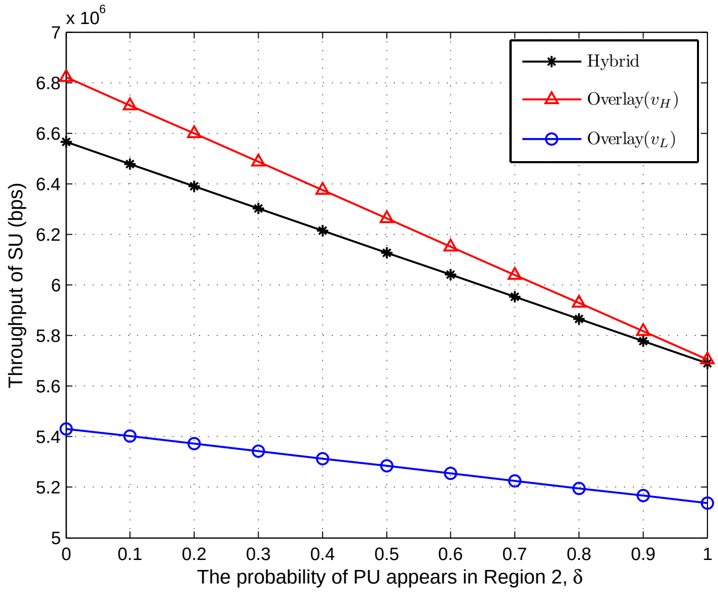


(a)

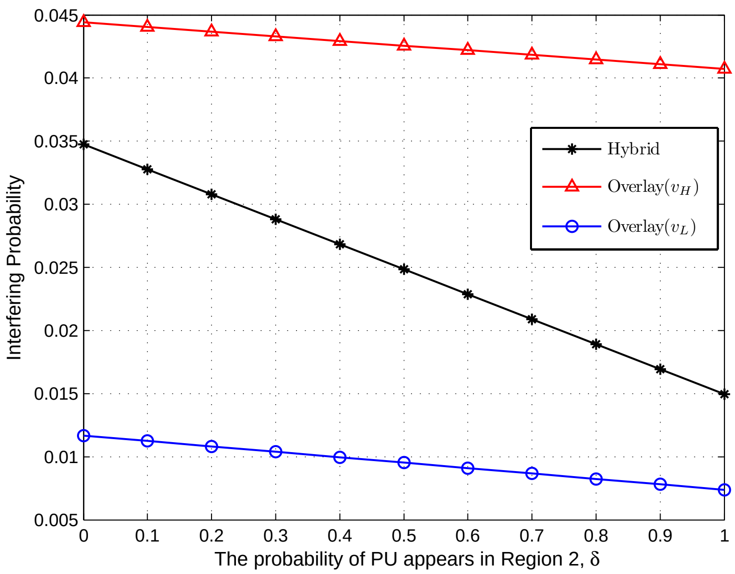


(b)

Figure 6.6: Results for (a) the throughput of SU and, (b) interfering probability against the the radius of Region 2. The results for the Hybrid scheme with $P_{\text{access}} = 0.9$ are compared with the traditional overlay schemes with the single detection threshold v_{H} and v_{L}.



(a)



(b)

Figure 6.7: Results for (a) the throughput of SU and, (b) interfering probability against the the radius of Region 2. The results for the Hybrid scheme with $P_{\mathsf{access}} = 0.9$ is compared with the traditional overlay schemes with the single detection threshold v_{H} and v_{L}.

6.4 Chapter Conclusions

In this chapter, a new hybrid access strategy was proposed and analyzed for uncoordinated DSS networks. A CR-based double-threshold energy detection method was employed to identify the PU's location. The spectrum sensing at the PHY layer and users' access process at the MAC layer were jointly modeled by a DTMC. Numerical results revealed that the hybrid strategy can achieve a 46.86% lowering of the interfering probability, as compared with the conventional overlay scheme. Moreover, SUs can adjust the access probability of the partial-access mode and coordinate the long and short sensing intervals. Hence, this hybrid access strategy can achieve a better tradeoff between throughput and interfering performance for uncoordinated DSS networks, comparing with the traditional overlay schemes with single energy detection fusion threshold.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

This thesis has focused on the performance improvement of future wireless communication systems from the point of view of DSS. A number of novel DSS methods have been developed in both coordinated and uncoordinated sharing scenarios. Our proposed methods are the general models and strategies for enhancement of the spectrum efficiency, which then can be applied to different wireless standards and systems.

In Chapter 2, the basic concepts of DSA and CR were introduced. We also gave an overview of Markov processes and Markov chains. Related work for coordinated and uncoordinated sharing, and overlay-underlay sharing models were analyzed and discussed, and problems with existing approaches were identified.

Chapter 3 dealt with the DSS problems in coordinated sharing scenarios. A heterogeneous-prioritized spectrum sharing policy was proposed and the interactions between PUs and multiple classes of SUs were modeled as Markov chain models. The SU blocking probability, interfering probability, and forced termination probability, and PU interfering probability and blocking probability can be derived from the models. In addition, we proposed two new GCRES and GCRC CAC strategies to enhance the system's serving capability for SUs. Results have demonstrated that the proposed heterogeneous-prioritized sharing policy achieves higher maximum admitted traffic of SUs, as compared with common and hierarchical sharing. Moreover, the proposed GCRES and GCRC CAC strategies are superior to conventional CAC schemes under max-sum, max-min and max-PF criterions. We also found that the max-PF fairness criterion achieves a better tradeoff between serving capability and fairness for multiple classes of SUs.

Chapter 4 described the DSS problems in a CR based uncoordinated sharing scenario. This chapter studied spectrum sharing between two licensed wireless networks and proposed a new partial sharing policy for DSS. A CTMC model was used to modeling the access process, considering the imperfect on-demand spectrum sensing results. We derived the blocking probability, the forced termination probability, the interference probability, the handoff probability, and the overall throughput for the proposed partial spectrum sharing policy. A cross layer design framework was also presented in this chapter, in which the parameters of spectrum sharing policy at the MAC layer and the spectrum sensing parameters at the physical layer are simultaneously coordinated to maximize the overall throughput of the networks, while satisfying the GoS constraints of the users. Numerical results have illustrated that the partial sharing policy with the cross layer design strategy achieves higher overall throughput performance, while satisfying the GoS constraints of the users. In particular, the gain over the hierarchical sharing model is 11.2% on average.

To further explore the influence of errors during periodic sensing, in Chapter 5, a DTMC model is further developed for partial spectrum sharing policy. Using the proposed analytical model, the impacts of sensing errors and periodic sensing intervals can be investigated. A PHY-MAC cross-layer design framework was also devised, where the parameters in spectrum sensing and spectrum sharing are jointly optimized. Results show that the proposed partial sharing policy can achieve 7.4% and 20.8% gains in throughput than traditional hierarchical sharing and non-sharing strategies, respectively.

In Chapter 6, based on the traditional overlay and underlay sharing schemes, a new hybrid access strategy was proposed and analyzed DSS networks. A CR-based double-threshold energy detection method is employed to identify the PU's location. The spectrum sensing at the PHY layer and users' access process at the MAC layer were jointly modeled by a DTMC. Numerical results revealed that the hybrid strategy can achieve a 46.86% lowering of the interfering probability, as compared with the conventional overlay scheme. Moreover, SUs can adjust the access probability of the partial-access mode and coordinate the long and short sensing intervals to achieve a better tradeoff between the throughput and interfering performance.

As spectrum has became the vital resource in telecommunication market, the ability to achieve greater spectrum utilization can create huge economic impacts in our society. For example, in the US, wireless service operators have to pay $200 million per MHz for the exclusive rights of spectrum license. According to spectrum management agency OFCOM in the UK [73], "The regulator envisions market forces will manage 72% of spectrum, while 7% will fall under license-exempt use, and the other 21% will be managed under current command and control approaches." Hence, it is expected that a spectrum reform and spectrum secondary market are inevitable in the near future. As described in this thesis, a number of novel DSS schemes with stochastic optimization methods have been developed. The uncoordinated schemes with less hardware updates, might be firstly adopted by the current wireless operators, i.e., using the proposed scheme in Chapter 4 to 6, PU can lease its licensed spectrum to SUs to generate profits. The coordinated scheme with a centralized spectrum manager may also appear in real world wireless systems later, in which wireless operators and individual wireless handsets can bid for spectrum access in real time. The proposed scheme in Chapter 3 can then be used to coordinate the access priorities for different users.

Comparing with the uncoordinated schemes in Chapter 4 and 5, the proposed CAC strategy in the coordinated scenario in Chapter 3 can achieve a 22.8% improvement on the throughput, at the cost of extra hardware updates. For the uncoordinated scenario in Chapter 4 and 5, by comparing with the tradition Hierarchical access scheme, it is concluded that the proposed cross-layer design methods can achieve a 11.2% and 7.4% throughput enhancement, in the cases of imperfect on-demand and periodic sensing, respectively. By further applying the Hybrid access scheme in Chapter 6 to the scheme, the interfering probability in the uncoordinated scenario can be lowered by 46.9%. In general, these proposed schemes can achieve a much higher spectrum efficiency and a better trade-off between the throughput and the users' GoS constraints.

7.2 Future Work

Current spectrum occupancy studies showed that most of the licensed bands are under-utilized, where the average utilization rate is lower than 20% [74], [75]. By using DSS and CR technologies, it is expected that the overall spectrum utilization can increase to 80%–90% level. The future generation wireless technologies are focusing on cooperative and competitive CR users with self-optimized wireless networks, which can automatically adapt to the changing environment. The research work in this thesis can be applied to different application scenarios in future wireless systems. In the following, a number of challenging research directions will be introduced.

Coordinated DSS: In this area, consideration of the different characteristics of the wireless services, i.e., real-time and non-real-time services, can be studied. The performance metrics of the real-time service would focus on blocking, forced termination probabilities, while for the non-real-time services, data transmission delay is an important performance metric. For the CAC schemes, the theory of Markov decision process (MDP) [70] needs to be studied, which is an extension of Markov chain, by introducing the actions (allowing choice) and rewards (giving motivation) to the modeling process. Furthermore, geo-location based spectrum database method can be investigated for coordinated DSS networks, in which database management will become the main issue.

Uncoordinated DSS: Besides the studies of PU-PU sharing in Chapter 4 and 5, the PU-SU and SU-SU sharing in the proposed model can be further investigated. During the performance assessment, the performance metrics at the PHY layer are detection probability, false alarm probability and sensing period, while at the MAC layer, blocking probability, forced termination probability, detection delay and throughput could be studied. The underlay case of PU-SU coexistence can also be further investigated, in which game theory could be used to achieve the tradeoff of power allocation for users. Hence, multi-objective optimization algorithms and cross-layer design frameworks can be further studied and used to obtain the optimized overall system performance. Furthermore, future work regarding DSS should be devoted to investigate practical cases, with specific technologies, such as over TV bands based on the IEEE 802.22 standard.

Hybrid Overlay-Underlay DSS: In future work, a general analytical model can be developed, by considering the various number of PUs and SUs, and multiple number of sharing channels, with the cross-layer optimization framework on the parameters such as sensing interval, sensing time, and partial-access probability, etc. Moreover, for some wireless services, e.g., real-time applications, the variance or deviation of throughput is also important to users' QoS. Hence, Portfolio theory based mean-variance optimization can be applied by the wireless service operators to provide user specific QoS guarantee, in terms of the average and variance of the achieved throughput. Statistics-based method can also be used to explore the historical frequency channel quality and availability, and achieve a better channel access decision.

In uncoordinated sharing, CR permits spectrum sharing without any hardware updates for the primary systems. Considering the benefits of no hardware updates for launching such uncoordinated sharing, current wireless service operators with licensed spectrum would have the most interests to apply the spectrum sharing and hybrid access schemes for enhancing the spectrum efficiency and making profits for spectrum leasing. Through further investigation of the above topics in future work, a general analytical framework, efficient DSS strategies, and intelligent access schemes, can be applied for future wireless communication systems.

Bibliography

[1] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[2] C. E. Shannon. Communication in the presence of noise. *Proceedings Institute of Radio Engineers*, 37(1):10–21, 1949.

[3] G. G. Raleigh and J. M. Cioffi. Spatio-temporal coding for wireless communication. *IEEE Transaction on Communications*, 46(3):357–366, 1998.

[4] S. Weinstein and P. Ebert. Data transmission by frequency-division multiplexing using the discrete Fourier transform. *IEE Transaction on Communication Technology*, 19(5):628–634, 1971.

[5] A. Paulraj, R. Nabar, and D. Gore. *Introduction to Space - Time Wireless Communications*. Cambridge University Press, 2003.

[6] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty. NeXt generation/dynamic spectrum access/cognitive radio wireless networks: A survey. *Computer Networks*, 50(13):2127–2159, 2006.

[7] FCC. Notice of proposed rule making and order. *ET Docket No 03-222*, 2003.

[8] M. McHenry. Spectrum white space measurements, *New America Foundation BroadBand Forum*, Jun. 2003.

[9] Q. Zhao and B. M. Sadler. A survey of dynamic spectrum access. *IEEE Signal Processing Magazine*, 24(3):79–89, 2007.

[10] J. Mitola Cognitive radio for flexible mobile multimedia communications. In *Procedings of IEEE Workshop Mobile Multimedia Communication*, 1999.

[11] J. Mitola. Cognitive radio for flexible mobile multimedia communications. *Mobile Networks and Application*, 6(5):435–441, 2001.

[12] M. M. Buddhikot, and K. Ryan. Spectrum management in coordinated dynamic spectrum access based cellular networks. In *Proceedings of First IEEE International Symposium New Frontiers in Dynamic Spectrum Access Networks*, 2005.

[13] FCC. BFirst Report and order, revision of Part 15 of the commissions rules regarding ultra-wideband transmission systems. *ET Docket 98–153*, Feb. 2002.

[14] J. M. Peha. Sharing Spectrum Through Spectrum Policy Reform and Cognitive Radio. *Proceedings of the IEEE*, 97(4):708–719, 2009.

[15] T. S. Rappaport, K. S. Shanmugan, and K. L. Kosbar, *Principles of Communication Systems Simulation: with Wireless Applications*. Prentice Hall, 2004.

[16] G. Bolch, S. Greiner, H. de Meer, and K. S. Trivedi. *Queueing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications*. John Wiley and Sons, INC., 2006.

[17] Y. Viniotis. *Probability and Random Processes for Electrical Engineers*. WCB/McGraw-Hill, 1998.

[18] O. Simeone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U. Spagnolini, and R. Pickholtz. Spectrum leasing to cooperating secondary ad hoc networks. *IEEE Journal on Selected Areas in Communication*, 26(1):203–213, 2008.

[19] X. Wu, and K. L. Yeung. Efficient channel borrowing strategy for real-time services in multimedia wireless networks. *IEEE Transactions on Vehicular Technology*, 49(4):1273–1284, 2000.

[20] Y. L. Chung, and Z. Tsai. Performance evaluation of dynamic spectrum sharing for two wireless communication networks. *IET Communications*, 4(4):452–462, 2010.

[21] X. Jiang, Y. Zhang, K. K. Wong, J. M. Kim, and D. Edwards. Quality of service-aware coordinated dynamic spectrum access: prioritized Markov model and call admission control. *Wireless Communications and Mobile Computing* 2011. doi: 10.1002/wcm.1118.

[22] P. K. Tang, Y. H. Chew, W. L. Yeow, and L. C. Ong. Performance comparison of three spectrum admission control policies in coordinated dynamic spectrum sharing systems. *IEEE Transactions on Vehicular Technology*, 58(7):3674–3683, 2009.

[23] M. Raspopovic and C. Thompson. Finite population model for performance evaluation between narrowband and wideband users in the shared radio spectrum. In *Proceedings of IEEE International Symposium New Frontiers in Dynamic Spectrum Access Networks*, 2007.

[24] P. K. Tang, Y. H. Chew, L. C. Ong, and M. K. Halder. Performance of secondary radios in spectrum sharing with prioritized primary access. in *Proceedings IEEE Military Communication Conference*, 2006.

[25] X. Zhu, L. Shen, T. P. Yum. Analysis of cognitive radio spectrum access with optimal channel reservation. *IEEE Communication Letters*, 11(4):304–306, 2007.

[26] K. Kim, I. A. Akbar, K. K. Bae, J. Urn, C. M. Spooner, and J. H. Reed. Cyclostationary approaches to signal detection and classification in cognitive radio. In *Proceedings of Second IEEE International Symposium New Frontiers in Dynamic Spectrum Access Networks*, 2007.

[27] G. Gelli, L. Izzo, and L. Paura. Cyclostationarity-based signal detection and source location in non-Gaussian noise. *IEEE Transactions on Communication*, 44(3):368–376, 1996.

[28] Y. Zhang, Q. Zhang, and S. Wu. Entropy-based robust spectrum sensing in cognitive radio. *IET Communications*, 4(4):428–436, 2010.

[29] H. Urkowitz. Energy detection of unknown deterministic signals. *Proceeding of IEEE*, 55:523–231, 1967.

[30] F. F. Digham, M. S. Alouini, and M. K. Simon. On the energy detection of unknown signals over fading channels. In *Procedings of IEEE International Conference on Communications*, 2003.

[31] S. M. Mishra, A. Aahai, and R. W. Brodersen. Cooperative sensing among cognitive radios. In *Procedings of IEEE International Conference on Communications*, 2006.

[32] A. Ghasemi, E. S. Sousa. Spectrum sensing in cognitive radio networks: The cooperation-processing tradeoff. *Wireless Communications and Mobile Computing*, 7(9):1049–1060, 2007.

[33] W. Zhang, R. Mallik, and K. Letaief. Optimization of cooperative spectrum sensing with energy detection in cognitive radio networks. *IEEE Transactions on Wireless Communications*, 8(12):5761–5766, 2009.

[34] J. Shen, S. Liu, L. Zeng, G. Xie, J. Gao, and Y. Liu. Optimisation of cooperative spectrum sensing in cognitive radio network. *IET Communications*, 3(7):1170–1178, 2009.

[35] Y. Chen, Q. Zhao, and A. Swami. Joint design and separation principle for opportunistic spectrum access in the presence of sensing errors. *IEEE Transaction on Information Theory*, 54(5):2053–2071, 2008.

[36] X. Gelabert, I. F. Akyildiz, O. Sallent, R. Agusti. Operating point selection for primary and secondary users in cognitive radio networks. *Computer Networks*, 53(8):1158–1170, 2009.

[37] Y. Yu, Y. Zhang, M. Huang, and S. Xie. Cross-layer optimized call admission control in cognitive radio networks. *Mobile Networks and Applications*, 15(5):610–626, 2010.

[38] T. Cui, F. Gao, and A. Nallanathan. Optimization of cooperative spectrum sensing in cognitive radio. *IEEE Transaction Vehicular Technology*, 60(4):1578–1589, 2011.

[39] Y.-C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang. Sensing-throughput tradeoff for cognitive radio networks. *IEEE Transaction Wireless Communication*, 7(4):1326–1337, 2008.

[40] X. Kang, R. Zhang, Y.-C. Liang, and H. K. Garg. Optimal power allocation strategies for fading cognitive radio channels with primary user outage constraint. *IEEE Journal on Selected Areas in Communication*, 29(2):374–383, 2011.

[41] S. Stotas, and A. Nallanathan. Optimal sensing time and power allocation in multiband cognitive radio networks. *IEEE Transaction Communication*, 59(1):226–235, 2011.

[42] J. Oh, and W. Choi. A hybrid cognitive radio system: A combination of underlay and overlay approaches. In *Proceeding of IEEE Vehicular Technology Conference*, Sep. 2010.

[43] V. Chakravarthy, X. Li, Z. Wu, M. Temple, F. Garber, R. Kannan, and A. Vasilakos. Novel overlay/underlay cognitive radio waveforms using SD-SMSE framework to enhance spectrum efficiency-part I: Theoretical framework and

analysis in AWGN channel. *IEEE Transaction Communication*, 57(12):3794–3804, 2009.

[44] M. G. Khoshkholgh, K. Navaie, and H. Yanikomeroglu. Access strategies for spectrum sharing in fading environment: Overlay, underlay, and mixed. *IEEE Transaction Mobile Computing*, 9(12):1780–1793, 2010.

[45] D. Malone, K. Duffy, and D. Leith. Modeling the 802.11 Distributed Coordination Function in Nonsaturated Heterogeneous Conditions. *IEEE/ACM Transactions on Networking*, 15(1):159–172, 2007.

[46] L. Badia, N. Baldo, M. Levorato, and M. Zorzi. A Markov framework for error control techniques based on selective retransmission in video transmission over wireless channels. *IEEE Journal on Selected Areas in Communications*, 28(3):488–500, 2010.

[47] Z. Jin, S. Anand, and S. P. Subbalakshmi Mitigating primary user emulation attacks in dynamic spectrum access networks using hypothesis testing. *ACM SIGMOBILE Mobile Computing and Communications Review*, 13(2):74–85, 2009.

[48] P. K. Tang, Y. H. Chew, L. C. Ong, and F. Chin. On the grade-of-services in the sharing of radio spectrum. In *Proceeding of International Conference on Cognitive Radio Oriented Wireless Networks and Communication*, Aug. 2007.

[49] Y. Xing, R. Chandramouli, S. Mangold, N SS. Dynamic spectrum access in open spectrum wireless networks. *IEEE Journal on Selected Areas in Communications*, 24(3):626–637, 2006.

[50] M. Raspopovic, C. Thompson, and K. Chandra. Performance models for wireless spectrum shared by wideband and narrowband sources. In *Proceeding of IEEE Military Communication Conference*, Oct. 2005.

[51] Zhang Y. Dynamic spectrum access in cognitive radio wireless networks. In *Proceeding of IEEE International Conference on Communications*, May 2008.

[52] D. T. C. Wong, A. T. Hoang, Y.-C. Liang, and F. P. S. Chin. Dynamic spectrum access with imperfect sensing in open spectrum wireless networks. In *Proceeding of IEEE Wireless Communication and Networks Conference*, Apr. 2008.

[53] E. W. M. Wong, and F. Chuan. Analysis of cognitive radio spectrum access with finite user population. *IEEE Communication Letters*, 13(5):294–296, 2009.

[54] S. S. Tzeng, Y. J. Lin, Y. C. Hsu. Radio resource reservation for heterogeneous traffic in cellular networks with spectrum leasing. *EURASIP Journal on Wireless Communications and Networking*, 2012(1):206–, 2012.

[55] T. M.N. Ngatched, S. Dong, S. A. Alfa, and J. Cai. Performance Analysis of Cognitive Radio Networks with Channel Assembling and Imperfect Sensing. In *Procedings of IEEE International Conference on Communications*, June 2012.

[56] H. Zheng, and C. Peng. Collaboration and fairness in opportunistic spectrum access. In *Procedings of IEEE International Conference on Communications*, May 2005.

[57] J. Kennedy, and R. Eberhart. Particle Swarm Optimization. In *Proceedings of IEEE International Conference on Neural Networks*, 1995.

[58] S. Kitayama, and K. Yasuda. A method for mixed integer programming problems by particle swarm optimization. *Electrical Engineering in Japan*, 157(2):40–49, 2006.

[59] Tang S, Mark BL. Modeling and analysis of opportunistic spectrum sharing with unreliable spectrum sensing. *IEEE Transactions on Wireless Communications* 2009; 8(4): pp. 1934–1943.

[60] S. Tang, and B. L. Mark. Analysis of opportunistic spectrum sharing with markovian arrivals and phase-type service. *IEEE Transactions on Wireless Communications*, 8(6):3142-3150, 2009.

[61] Y. Zhang, J. Xiang, Q. Xin, and G. E. Oien. Optimal sensing cooperation for spectrum sharing in cognitive radio networks. In *Proceeding of European Wireless Conference*, May 2009.

[71] J. Kennedy, R. C. Eberhart, and Y. Shi. *Swarm Intelligence*. Morgan Kaufmann, 2001.

[63] X. Jiang, N. Han, G. Zheng, S. H. Sohn, Z. Shi, and J. M. Kim. Analysis of opportunistic access in the spectrum sharing system with heterogeneous secondary users. In *Proceedings of International Conference on Communication, Circuits and Systems*, May 2008.

[64] X. Jiang, Z. Shi, W. Wang, and L. Huang. Performance of secondary access in the spectrum sharing system with call admission control strategy.

In *Proceedings of International Conference on Network Security, Wireless Communication and Trusted Computing*, Apr. 2009.

[65] Y. H. Zeng and Y.-C. Liang. Spectrum-sensing algorithms for cognitive radio based on statistical covariances. *IEEE Transaction on Vehicular Thchnology*, 58(4):1804–1815, 2009.

[66] E. Peh and Y. C. Liang. Optimization for Cooperative Sensing in Cognitive Radio Networks. In *Proceedings of Wireless Communication and Network Conference*, Mar. 2007.

[67] D. Bertsekas and R. Gallager. *Data Networks*. Prentice-Hall, 1992.

[68] W. Ahmed, J. Gao, H. Suraweera, and M. Faulkner. Comments on "Analysis of cognitive radio spectrum access with optimal channel reservation. *IEEE Transaction Wireless Communication*, 8(9):4488–4491, Sep. 2009.

[69] A. Goldsmith. *Wireless Communications* Cambridge University Press, 2005.

[70] M. L. Puterman. *Markov Decision Processes*. Wiley, 1994.

[71] J. Kennedy and R. C. Eberhart. *Swarm Intelligence*. Morgan Kaufmann, 2001.

[72] F. van Den Bergh. *An Analysis of Particle Swarm Optimizers*. PhD thesis, University of Pretoria, Pretoria, South Africa, 2002.

[73] OFCOM. Spectrum Framework Review Statement. 2005 [Online]: http://stakeholders.ofcom.org.uk/binaries/consultations/sfr/statement/sfr_statement.

[74] R. Bacchus, T. Taher, K. Zdunek, and D Roberson. Spectrum Utilization Study in Support of Dynamic Spectrum Access for Public Safety. In *Proceedings of IEEE International Symposium New Frontiers in Dynamic Spectrum Access Networks*, 2010.

[75] OFCOM. Capture of Spectrum Utilisation Information Using Moving Vehicles C Final Report. 2009 [Online]: http://stakeholders.ofcom.org.uk/binaries/research/technology-research/vehicles.pdf.

Publications

Journal Papers

[1] **X. Jiang**, K.-K. Wong, Y. Zhang, and D. J. Edwards, "Mean-Variance Optimization Based Dynamic Power Allocation for Cognitive Radio Networks", (Manuscript in Preparation).

[2] **X. Jiang**, K.-K. Wong, Y. Zhang, and D. J. Edwards, "On Hybrid Overlay-Underlay Dynamic Spectrum Access: Double-Threshold Energy Detection and Markov Model", *IEEE Transaction on Vehicular Technology*, 62(8), 2013 (In Press).

[3] **X. Jiang**, K.-K. Wong, Y. Zhang, and D. J. Edwards, "On Imperfect Spectrum Sensing for Partial Spectrum-Shared Licensed Networks ", *IET Communications*, 6(17):2894–2899, 2012.

[4] **X. Jiang**, K.-K. Wong, Y. Zhang, and D. J. Edwards, "Cross-layer Design of Partial Spectrum Sharing for Two Licensed Networks using Cognitive Radios", *Wiley Wireless Communication and Mobile Computing*, DOI: 10.1002/wcm.2345, 2013.

[5] **X. Jiang**, Y. Zhang, K.-K. Wong, J. M. Kim and D. J. Edwards, "Quality of service-aware coordinated dynamic spectrum access: prioritized Markov model and call admission control", *Wiley Wireless Communication and Mobile Computing*, DOI: 10.1002/wcm.1118, 2011.

Conference Proceedings

[6] **X. Jiang**, Y. Zhang, K.-K. Wong, and D. J. Edwards, "On Partial Spectrum Sharing of Two Licensed Networks Using Cognitive Radios", *2011 IEEE Conference on Vehicle Technology (VTC)-Spring*, May 2011.

[7] C. Zhao, L. Xie, **X. Jiang**, L. Huang, Y. Yao, "A PHY-layer Authentication Approach for Transmitter Identification in Cognitive Radio Networks", *IEEE Conference on Communications and Mobile Computing (CMC)*, April 2010.

Appendix A

List of Abbreviations

CAC	Call Admission Control
CR	Cognitive Radio
CTMC	Continuous-time Markov Chain
DSA	Dynamic Spectrum Access
DSS	Dynamic Spectrum Sharing
DTMC	Discrete-time Markov chain
ES	Exhaustive Search
FCC	Federal Communications Commission
FCFS	First-Come-First-Served
FDMA	Frequency Division Multiple Access
GCRC	Guard Channels and Restricted Channels
GCRES	Guard Channels and Restrict Functions
GCs	Guard Channels

GoS	Grade-of-Service
HP	Heterogeneous-Prioritized
MAC	Multiple-access Control
MIMO	Multiple-in-Multiple-out
OFDM	Orthogonal Frequency Division Multiplexing
PHY	Physical
PSO	Particle Swarm Optimization
PUs	Primary Users
QoS	Quality-of-service
RCs	Restricted channels
ROC	Receiver Operating Characteristics
Rx	Receiver
SCs	Shared Channels
SNR	Signal-to-noise Ratio
SUs	Secondary Users
Tx	Transmitter
WRAN	Wireless Regional Area Network
WLAN	Wireless Local Area Network
WPAN	Wireless Personal Area Network

Appendix B

Particle Swarm Optimization

Particle swarm optimization (PSO) is a population and bio-inspired based stochastic optimization algorithm, which was first proposed by Eberhart and Kennedy in 1995 [57], [71]. This algorithm is inspired by social behavior of bird flocking or fish schooling, which can be initialized with a population of random solutions and probes the search space by update, cooperation and competition among individuals (referred to as particles).

B.1 Basic PSO

PSO algorithm does not need gradient information of the objective function, thus this algorithm can be used in a broader range of optimization problems, such as the case in which the gradient information is unavailable [72]. Since PSO has no overlapping and mutation calculation and each particle is updated their velocity based on the current "personal best" and "global best", thus the speed of space searching is fast. However, the disadvantage of PSO is that it can be trapped into a local optimal.

The basic PSO algorithm is defined based on the following equations [72].

$$\begin{aligned} \mathbf{v}_i(k+1) = \mathbf{v}_i(k) + c_1\varphi_1(\mathbf{p}_{best,i}(k) - \mathbf{x}_i(k)) \\ + c_2\varphi_2(\mathbf{g}_{best}(k) - \mathbf{x}_i(k)) \end{aligned} \tag{B.1}$$

$$\mathbf{x}_i(k+1) = \mathbf{x}_i(k) + \mathbf{v}_i(k+1) \tag{B.2}$$

where

i	the index of particles;
k	the index of iterations;
$\mathbf{v}_i(k)$	the velocity vector of the i-th particle at k-th iteration ;
$\mathbf{p}_{best,i}$	the best position found by particle i till the current iteration;
$\mathbf{g}_{best}$	the best position found by all particles;
φ_1, φ_2	independent random numbers and $\varphi_1 \sim U(0,1)$ and $\varphi_2 \sim U(0,1)$;
c_1, c_2	the learning factors, and are often defined as $c_1 + c_2 = 4$;.

The flow chart of the basic PSO algorithm for a generalized minimization problem is described as follows:

Basic PSO algorithm for a generalized minimization problem

Step 1: The initial location and velocity are set random for each particle in the solution space. Set the iteration counter k as $k = 1$;

Step 2: For each particle i, evaluate the optimization objective function $f(\cdot)$. If $f(\mathbf{x}_i(k)) < f(\mathbf{p}_{best,i})$, then set $\mathbf{p}_{best,i} = \mathbf{x}_i(k)$;

Step 3: Find the $\mathbf{g}_{best}$ for all particles as $\mathbf{g}_{best} = \min\{\mathbf{p}_{best,i}\}$;

Step 4: Based on (B.1) and (B.2), the location and velocity for each particle are updated;

Step 5: If $k \leq k_{max}$, where k_{max} is the maximum number of iterations, set $k = k+1$. Repeat **Step 2** to **Step 4**. Otherwise, the search is terminated. The stopping criterion can be either defined as the predefined maximum number of iterations is achieved or a specified error bound is attained [72].

B.2 The PSO Algorithm For Problems (3.25),(3.29) and (3.30)

The optimization problem (3.25) can be solved by the basic PSO algorithm. Since the problems (3.29) and (3.30) contain the discrete variables, the optimal value of N_G and N_R can be obtained using exhaustive search method, while the PSO algorithm is used to obtain the optimal values for the continuous variables, i.e., τ_{SU_1}, τ_{SU_2}, α_1, α_2 and c_{RES}.

For the parameter settings, the population of particles is set to 50, the maximum number of iterations is set to 100, and the learning factors are $c_1 = c_2 = 2$ [72]. The flow chart of the PSO algorithm to the admitted traffic of SUs in (3.25), (3.29) and (3.30) is given as follows.

PSO algorithm for maximum admitted traffic of SUs problem

Step 1: The initial values τ_{SU_1}, τ_{SU_2}, α_1, α_2 and c_{RES} are generated, such as $\tau_{\mathsf{SU}_1} \sim U(0, 0.5)$, $\tau_{\mathsf{SU}_1} \sim U(0, 0.5)$, $\alpha_1 \sim U(0, 1)$, $\alpha_2 \sim U(0, 1)$, and $c_{\mathsf{RES}} \sim U(0, 1)$. Set the iteration counter k as $k = 1$.

Step 2: For each particle i, evaluate the optimization objective functions (3.25), (3.29) and (3.30), i.e., $f(\mathbf{x}_i(k)) = \tau_{\mathsf{SU}}$, which can be obtained by using α_1, α_2, τ_{SU_1}, τ_{SU_2} and c_{RES} values in the model.

Step 3: If $f(\mathbf{x}_i(k)) > f(\mathbf{p}_{best,i})$, and the GoS of PU, SU_1, and SU_2 are lower than their GoS constraints, then set $\mathbf{p}_{best,i} = \mathbf{x}_i(k)$;

Step 4: Find the $\mathbf{g}_{best}$ for all particles as $\mathbf{g}_{best} = \max\{\mathbf{p}_{best,i}\}$;

Step 5: Based on (B.1) and (B.2), the $\alpha_1, \alpha_2, \tau_{\mathsf{SU}_1}, \tau_{\mathsf{SU}_2}, c_{\mathsf{RES}}$ values for each particle are updated;

Step 6: If $k \leq k_{max}$, where k_{max} is the maximum number of iterations, set $k = k+1$. Repeat **Step 2** to **Step 5**. Otherwise, the search is terminated.

For a pair of N_G and N_R, the PSO is performed to obtain the optimized continuous variables. The process is then repeated for different values of N_G and N_R, and the maximum admitted traffic of SUs is obtained.