SPECTRUM MANAGEMENT IN COGNITIVE RADIO WIRELESS NETWORKS

A Thesis
Presented to
The Academic Faculty

by

Won Yeol Lee

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Electrical and Computer Engineering

Georgia Institute of Technology
August 2009

SPECTRUM MANAGEMENT IN COGNITIVE RADIO WIRELESS NETWORKS

Approved by:

Professor Ian F. Akyildiz, Advisor
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Ye (Geoffrey) Li
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Raghupathy Sivakumar
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Joy Laskar
School of Electrical and Computer Engineering
Georgia Institute of Technology

Professor Mostafa Ammar
College of Computing
Georgia Institute of Technology

Date Approved: 4 August 2009

To my parents

for their endless love and support.

ACKNOWLEDGEMENTS

First of all, I sincerely want to thank my advisor Dr. Ian F. Akyildiz for his guidance, support and encouragement. He gave me constructive criticisms as well as rewarding praises, which motivated and trained me to improve the quality of my research. He has also been a great role model and I undoubtedly intend to use what I have learned from him as I move forward in my career. He is someone I definitely hope to live up to someday.

I would like to acknowledge Dr. Ye (Geoffrey) Li, Dr. Raghupathy Sivakumar, Dr. Joy Laskar, and Dr. Mostafa Ammar for being on my dissertation defense committee. Their invaluable comments and enlightening suggestions have helped improve the quality for this dissertation.

I also want to acknowledge all former and current members of the Broadband Networking Laboratory for their support and friendship. The unique family-like environment made me to survive the tough Ph.D. student life with comfort.

Last but not least, I can never find enough words to express how grateful I am for the endless support of my family. My parents have always provided me with unconditional love, sacrifices, and support. Without their faith in me, I could not have made it this far. I also like to thank my younger brother and sister for their love, encouragement, and concern for me. I also want to thank my mother-in-law for her support and pride in me. Finally, and most importantly, I would like to express much gratitude and love to my wife, Yang Hee Jeon. Without her patience, sacrifices, and love, I would not have been able to complete this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS OR ABBREVIATIONS

A/D Analog-to-Digital.

AMPS Advanced Mobile Phone System.

AWGN Additive White Gaussian Noise.

BER Bit Error Rate.

BS Base-Station.

BSC Base-Station Controller.

CDMA Code Division Multiple Access.

CR Cognitive Radio.

FCC Federal Communications Commission.

FCFS First-Come First-Serve.

FDD Frequency Division Duplex.

FDMA Frequency Division Multiple Access.

GSM Global System for Mobile communications.

LAN Local Area Networks.

LCFS Least-Cost First-Serve.

LO Local Oscillator.

MAC Medium Access Control.

MCSD Miaximum Capacity-based Spectrum Decision.

ML Maximum Likelihood.

MS Multiple Selections.

MVSD Minimum Variance-based Spectrum Decision.

OFDM Orthogonal Frequency Division Multiplexing.

PU Primary User.

QoS Quality-of-Service.

RF Radio Frequency.

SDR Software-Defined Radio.

SINR Signal-to-Interference Plus Noise Ratio.

SNR Signal-to-Noise Ratio.

SS Single Selection.

TDD Time-Division Duplex.

TDMA Time Division Multiple Access.

UHF Ultra High Frequency.

W-CDMA Wideband Code Division Multiple Access.

WRAN Wireless Regional Area Network.

SUMMARY

The wireless spectrum is currently regulated by government agencies and is assigned to license holders or services on a long-term basis over vast geographical regions. Recent research has shown that a large portion of the assigned spectrum is used sporadically, leading to underutilization and waste of valuable frequency resources. Consequently, dynamic spectrum access techniques are proposed to solve these current spectrum inefficiency problems. This new area of research foresees the development of cognitive radio (CR) networks to further improve spectrum efficiency.

The basic idea of CR networks is that the unlicensed devices (also called CR users) share wireless channels with the licensed devices (also known as primary users) that are already using an assigned spectrum. CR networks, however, impose unique challenges resulting from high fluctuation in the available spectrum, as well as diverse quality-of-service (QoS) requirements. These challenges necessitate novel cross-layer techniques that simultaneously address a wide range of communication problems from radio frequency (RF) design to communication protocols, which can be realized through spectrum management functions as follows: (1) determine the portions of the spectrum currently available (spectrum sensing), (2) select the best available channel (spectrum decision), (3) coordinate access to this channel with other users (spectrum sharing), and (4) effectively vacate the channel when a primary user is detected (spectrum mobility).

In this thesis, a spectrum management framework for CR networks is investigated that enables seamless integration of CR technology with existing networks. First, an optimal spectrum sensing framework is developed to achieve maximum spectrum opportunities while satisfying interference constraints, which can be extended

to multi-spectrum/multi-user CR networks through the proposed sensing scheduling and adaptive cooperation methods. Second, a QoS-aware spectrum decision framework is proposed where spectrum bands are determined by considering the application requirements as well as the dynamic nature of the spectrum bands. Moreover, a dynamic resource management scheme is developed to decide on the spectrum bands adaptively dependent on the time-varying CR network capacity. Next, for spectrum sharing in infrastructure-based CR networks, a joint spectrum and power allocation scheme is proposed to achieve fair resource allocation as well as maximum capacity by opportunistically negotiating additional spectrum based on the licensed user activity (exclusive allocation) and having a share of reserved spectrum for each cell (common use sharing). Finally, we propose a novel CR cellular network architecture based on the spectrum-pooling concept, which mitigates the heterogeneous spectrum availability. Based on this architecture, a unified mobility management framework is devised to support both user and spectrum mobilities in CR networks.

CHAPTER I

INTRODUCTION

1.1 Background

Current wireless networks are characterized by a static spectrum assignment policy where government agencies assign wireless spectrum to license holders on a long-term basis for large geographical regions. Recently, because of the increase in spectrum demand, this policy has been faced with spectrum scarcity at particular spectrum bands. On the contrary, a large portion of the assigned spectrum is still used sporadically, leading to underutilization of a significant amount of the spectrum [21]. The limited available spectrum and inefficient spectrum utilization make it necessary to develop a new communication paradigm to exploit the existing wireless spectrum opportunistically. To address these critical problems, the Federal Communications Commission (FCC) recently approved the use of unlicensed devices in licensed bands [21]. Consequently, dynamic spectrum access techniques are proposed to solve these current spectrum inefficiency problems [3].

The key enabling technology for dynamic spectrum access techniques is cognitive radio (CR) networking, which allows intelligent spectrum-aware devices to opportunistically use the licensed spectrum bands for transmission [53]. The term *cognitive radio* can formally be defined as follows [22]:

> A *cognitive radio* is a radio that can change its transmitter parameters based on interaction with the environment in which it operates.

From this definition, two main characteristics of the cognitive radio can be defined as follows [29]:

- **Cognitive Capability:** Cognitive capability refers to the ability of the radio

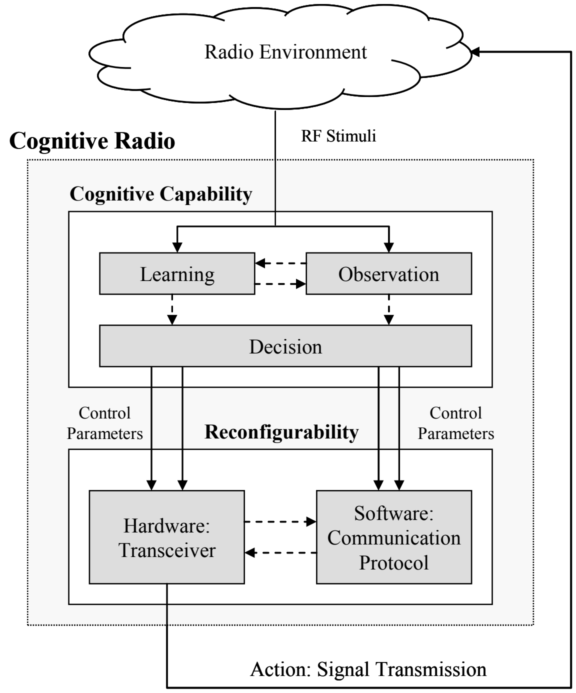


Figure 1: Cognitive radio concept.

technology to capture or sense the information from its radio environment. This capability cannot simply be realized by monitoring the power in some frequency band of interest, but more sophisticated techniques such as autonomous learning and action decision are required to capture the temporal and spatial variations in the radio environment and avoid interference to other users.

- **Reconfigurability:** The cognitive radio can be programmed to transmit and receive on a variety of frequencies and to use different transmission access technologies supported by its hardware design [36].

Figure 1 depicts how the cognitive radio concept can be realized through cognitive capability and reconfigurability. First, the cognitive radio identifies radio information through observation and learning processes and makes proper decisions accordingly. Based on these decisions, the cognitive radio reconfigures its software (e.g., communication protocols) and hardware (e.g., an radio frequency (RF) front-end and an antenna).

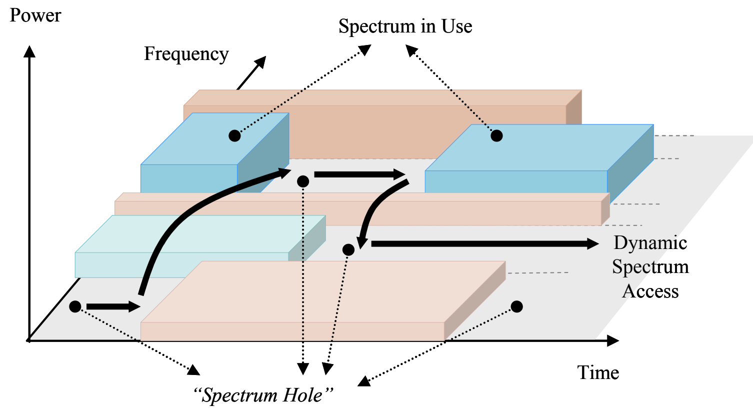


Figure 2: Spectrum hole and dynamic spectrum access.

Through cognitive capability and reconfigurability, the cognitive radio enables the usage of temporally unused spectrum, which is referred to as a *spectrum hole* or *white space* [29]. If this band is further used by a licensed user, the cognitive radio moves to another spectrum hole to avoid interference to the licensed users, as shown in Figure 2. This new area of research foresees the development of cognitive radio (CR) networks to further improve spectrum efficiency.

The components of the CR network architecture, as shown in Figure 3, can be classified in two groups as the *primary network* and the *cognitive radio network* [3]. The *primary network* is referred to as an existing network, where the *primary users* have a license to operate in a certain spectrum band. If the primary network has an infrastructure, primary user (PU) activities are controlled through the *primary base-stations*. Because of their priority in spectrum access, the operations of primary users should not be affected by any other unlicensed users.

The *CR network* does not have a license to operate in a desired band. Hence, additional functionalities are required for *CR users* to share the licensed spectrum band with primary networks. CR networks can be deployed as either an infrastructure-based network or an ad hoc network. CR infrastructure-based networks can be equipped with a central network entity such as a *CR base-stations*, which provide a single-hop connection to CR users. On the other hand, the CR ad hoc network

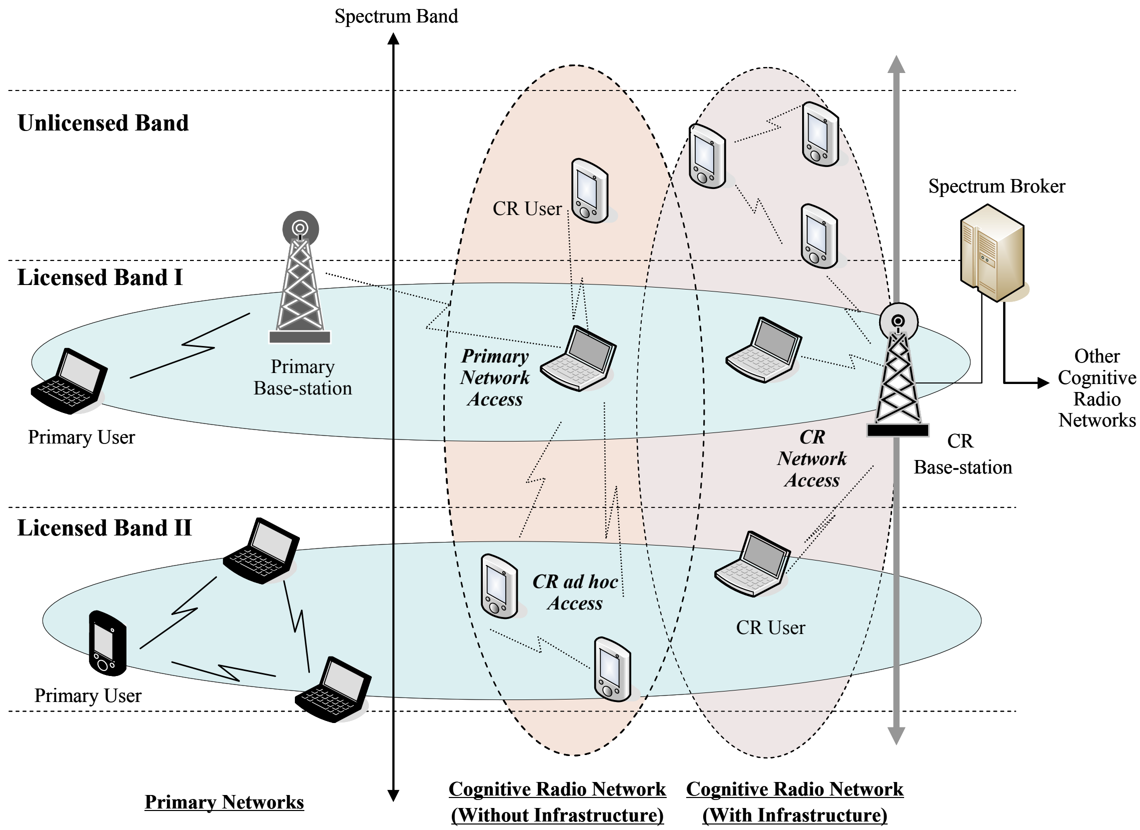


Figure 3: Cognitive radio network architecture.

does not have any infrastructure backbone. Thus, a CR user can communicate with other CR users through ad hoc connection on both licensed and unlicensed spectrum bands. Furthermore, CR networks may include *spectrum brokers* that play a role in sharing spectrum resources among different CR networks.

1.2 Research Objectives and Solutions

Cognitive radio provides the capability to share wireless channels with primary in an opportunistic manner. To this end, CR users need to continuously monitor the spectrum for the presence of the primary users and reconfigure the RF front-end according to the demands and requirements of the higher layers. CR networks, however, impose unique challenges because of the high fluctuation in the available spectrum, as well as the diverse quality of service (QoS) requirements of various applications. To address

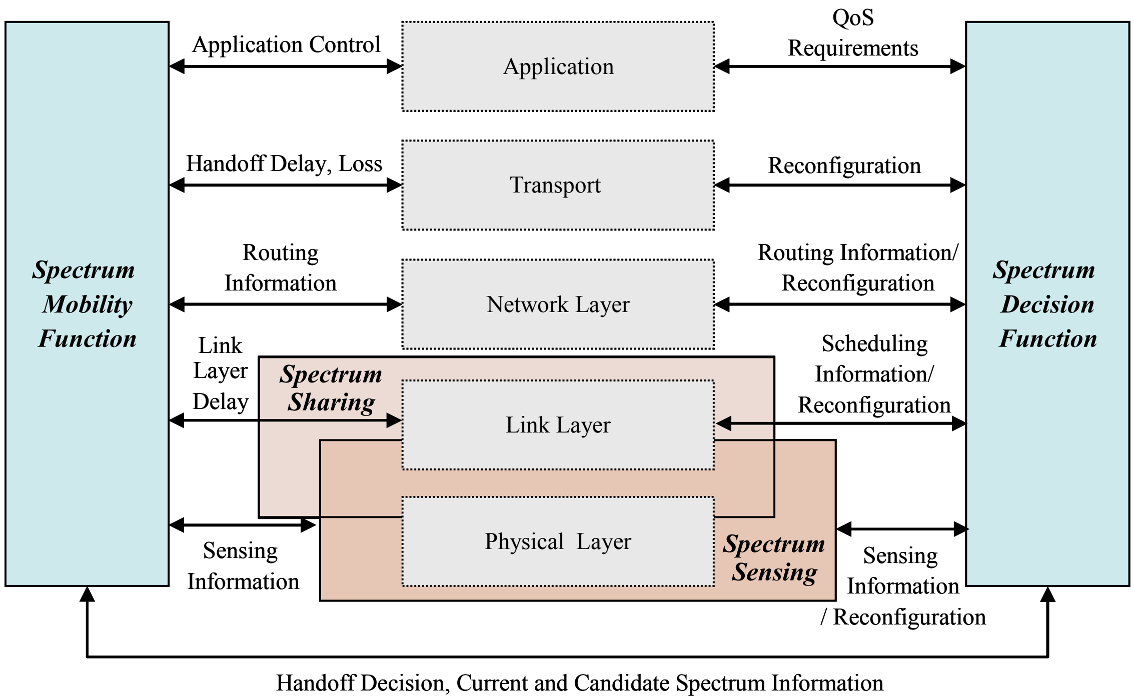


Figure 4: Spectrum management framework for cognitive radio networks.

these challenges, each CR user in the CR network must: 1) determine which portions of the spectrum are available, 2) select the best available channel, 3) coordinate access to this channel with other users, and 4) vacate the channel when a licensed user is detected. These capabilities can be realized through novel cross-layer design techniques that simultaneously address a wide range of communication problems from RF design to communication protocols, referred to as a *spectrum management framework* [3].

The ultimate objective of this research is to develop the spectrum management framework that exploits the dynamic spectrum environment and the cross-layer design advantages in CR networks to address the unique challenges posed by the dynamic spectrum access paradigm. The proposed spectrum management framework can be mainly classified into four topics: *spectrum sensing*, *spectrum decision*, *spectrum sharing*, and *spectrum mobility*, as shown in Figure 4. More specifically, the unique characteristics of the spectrum management framework and the proposed solutions for each topic addressed in this thesis can be summarized as follows:

1.2.1 Spectrum Management Framework in Cognitive Radio Networks

In this thesis, intrinsic properties and current research challenges of CR networks are presented. First, novel spectrum management functionalities such as spectrum sensing, spectrum sharing, and spectrum decision, and spectrum mobility are introduced. A particular emphasis is given to cross-layer design approaches from the viewpoint of both infrastructure-based network requiring central network entities and ad hoc networks based on distributed coordination. The main challenge in CR networks is to integrate these functions in the layers of the protocol stack, so that the CR users can communicate reliably over a dynamic spectrum environment. Thus, the influence of these functions on the performance of the upper layer protocols, such as the network layer, and transport layer protocols are investigated, and open research issues in these areas are also outlined.

1.2.2 Optimal Spectrum Sensing Framework for Cognitive Radio Networks

Spectrum sensing is the key enabling technology for cognitive radio networks. The main objective of spectrum sensing is to provide more spectrum access opportunities to cognitive radio users without interfering with the operations of the licensed network. Hence, recent research has focused on the *interference avoidance* problem. Moreover, current radio frequency (RF) front-ends cannot perform sensing and transmission at the same time, which inevitably decreases their transmission opportunities, leading to the so-called *sensing efficiency* problem. In this thesis, to solve both the interference avoidance and the spectrum efficiency problem, an optimal spectrum sensing framework is developed. More specifically, first a theoretical framework is developed to optimize the sensing parameters in such a way as to maximize sensing efficiency subject to interference avoidance constraints. Second, to exploit multiple spectrum bands, spectrum selection and scheduling methods are proposed where the best spectrum band for sensing are selected to maximize the sensing capacity. Finally, an adaptive and cooperative spectrum sensing method is proposed where the sensing parameters are optimized adaptively to the number of cooperating users. Simulation results show that the proposed sensing framework can achieve maximum sensing efficiency and opportunities in multi-user/multi-spectrum environments, satisfying interference constraints.

1.2.3 QoS-Aware Spectrum Decision Framework for Cognitive Radio Networks

Since CR networks can have multiple available spectrum bands with different channel characteristics, they should be capable of selecting the proper spectrum bands according to the application requirements, called *spectrum decision*. In this thesis, a spectrum decision framework is proposed to determine a set of spectrum bands by considering the application requirements as well as the dynamic nature of the spectrum bands. To this end, first, each spectrum is characterized by jointly considering primary user activity and spectrum sensing operations. Based on this, a minimum variance-based spectrum decision is proposed for real-time applications, which minimizes the capacity variance of the decided spectrums subject to the capacity constraints. For best effort applications, a maximum capacity-based spectrum decision is proposed where spectrum bands are decided to maximize the total network capacity. Moreover, a dynamic resource management scheme is developed to coordinate the spectrum decision adaptively dependent on the time-varying cognitive radio network capacity. Simulation results show that the proposed methods provide efficient bandwidth utilization while satisfying service requirements.

1.2.4 Spectrum Sharing Framework for Infrastructure-Based Cognitive Radio Networks

Since the spectrum availability varies over time and space, CR networks are required to have a dynamic spectrum sharing capability. This allows fair resource allocation as well as capacity maximization and avoids the starvation problems seen in the classical spectrum sharing approaches. In this thesis, a spectrum sharing framework for infrastructure-based CR networks is proposed that addresses these concerns by (i) opportunistically negotiating additional spectrum based on the licensed user activity (*exclusive allocation*), and (ii) having a share of reserved spectrum for each cell (*common use sharing*). Our algorithm consists of inter-cell and intra-cell spectrum sharing schemes, which account for the maximum cell capacity, minimize the interference caused to neighboring cells, and protect the licensed users through a sophisticated power allocation method. Simulation results reveal that the proposed spectrum sharing framework achieves better fairness and higher network capacity than the conventional spectrum sharing methods.

1.2.5 Spectrum-Aware Mobility Management in Cognitive Radio Cellular Networks

In CR cellular networks, CR users are traversing across multiple cells having different spectrum availability. Furthermore, they should switch to a new spectrum band when primary users appear in the spectrum, which is called *spectrum mobility*. Because of these heterogenous and dynamic spectrum environments, it is challenging to provide reliable communication channels to mobile CR users. In this thesis, a spectrum-aware mobility management scheme is proposed for CR cellular networks to enable seamless mobile communications by considering both user mobility and PU activity. This can be achieved by an intelligent switching of mobile users to the best combination of a target cell and spectrum, which leads to reconfiguration of the network to maximize capacity with the minimum switching latency. More specifically, a novel network architecture is introduced to mitigate the heterogeneous spectrum availability. Based on this architecture, a unified mobility management framework is developed to support diverse mobility events in CR networks that consists of *spectrum mobility management*, *user mobility management*, and *inter-cell resource allocation*. The spectrum mobility management scheme increases cell capacity by allowing CR users to select target cells and spectrums adaptively dependent on current spectrum utilization. In the user mobility management scheme, a switching cost-based handoff decision mechanism is developed to minimize quality degradation resulting from user mobility. Inter-cell resource allocation helps to improve the performance of both mobility management schemes by efficiently sharing spectrums with multiple cells. Simulation results show that the proposed method can achieve better performance than conventional handoff schemes in terms of both cell capacity as well as mobility support in communications.

1.3 Thesis Outline

This thesis is organized as follows: Chapter 2 presents a novel spectrum management framework along with its research challenges, which is necessary to realize efficient and reliable communications in CR networks. In Chapter 3, an optimal spectrum sensing framework is developed to achieve maximum spectrum opportunity while satisfying interference constraints. This new scheme can be extended to multi-spectrum/multi-user CR networks through the proposed sensing scheduling and adaptive cooperation methods. In Chapter 4, a QoS-aware spectrum decision framework is proposed where spectrum bands are determined by considering application requirements as well as the dynamic nature of the spectrum bands. In addition, a novel dynamic resource management scheme is developed to support the proposed decision framework by maintaining the QoS in the presence of time-varying spectrum resources. For spectrum sharing in infrastructure-based CR networks, a joint spectrum and power allocation scheme is proposed in Chapter 5, which achieves fair resource allocation as well as maximum capacity by opportunistically negotiating additional spectrum based on the licensed user activity and having a share of reserved spectrum for each cell. Chapter 6 introduces a novel mobility management scheme for CR cellular networks, where a spectrum pool-based network architecture is presented to mitigate the heterogeneity in spectrum availability. Based on this architecture, a unified handoff framework is devised to support both user and spectrum mobilities in CR networks. Finally, Chapter 7 summarizes the research contributions and identifies several future research directions.

CHAPTER II

SPECTRUM MANAGEMENT IN COGNITIVE RADIO NETWORKS

2.1 Introduction

CR networks impose unique challenges because of the coexistence with primary networks as well as diverse QoS requirements. Thus, new spectrum management functions are required for CR networks with the following critical design challenges:

- *Interference Avoidance:* CR network should avoid interference with primary networks.
- *QoS Awareness:* To decide an appropriate spectrum band, CR networks should support QoS-aware communication, considering dynamic and heterogeneous spectrum environment.
- *Seamless Communication:* CR networks should provide seamless communication, regardless of the appearance of the primary users.

To address these challenges, CR networks necessitate the spectrum-aware operations, which form a cognitive cycle. As shown in Figure 5, the steps of the cognitive cycle consist of four spectrum management functions: *spectrum sensing*, *spectrum decision*, *spectrum sharing*, and *spectrum mobility.* To implement CR networks, each function needs to be incorporated into the classical layering protocols, as shown in 4. The following are the main features of spectrum management functions [4]:

1. *Spectrum Sensing:* A CR user can allocate only an unused portion of the spectrum. Therefore, the CR user should monitor the available spectrum bands, capture their information, and then detect the spectrum holes.

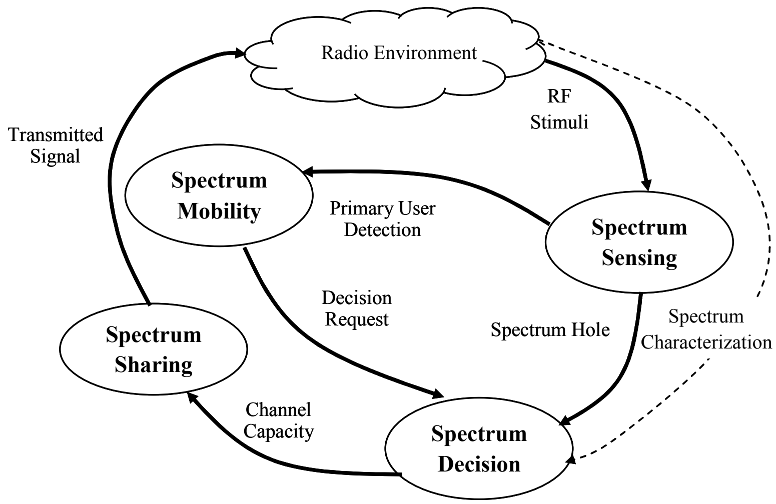


Figure 5: Cognitive cycle.

2. *Spectrum Decision:* Based on spectrum availability, CR users decide on the best spectrum band. This decision not only depends on spectrum availability, but it is also determined based on internal (and possibly external) policies.

3. *Spectrum Sharing:* Since there may be multiple CR users trying to access the spectrum, CR network access should be coordinated to prevent multiple users colliding in overlapping portions of the spectrum.

4. *Spectrum Mobility:* CR users are regarded as visitors to the spectrum. Hence, if the specific portion of the spectrum in use is required by a primary user, the communication needs to be continued in another vacant portion of the spectrum.

This spectrum management framework needs to be implemented differently according to the network architecture. In the infrastructure-based CR networks, the observations and analysis performed by each CR user feed the central CR base-station, so that it can make decisions on how to avoid interfering with primary networks. According to this decision, each CR user reconfigures its communication parameters, as shown in Figure 6 (a). On the contrary, in CR ad hoc networks, each user needs to have all CR capabilities and is responsible for determining its actions based on the

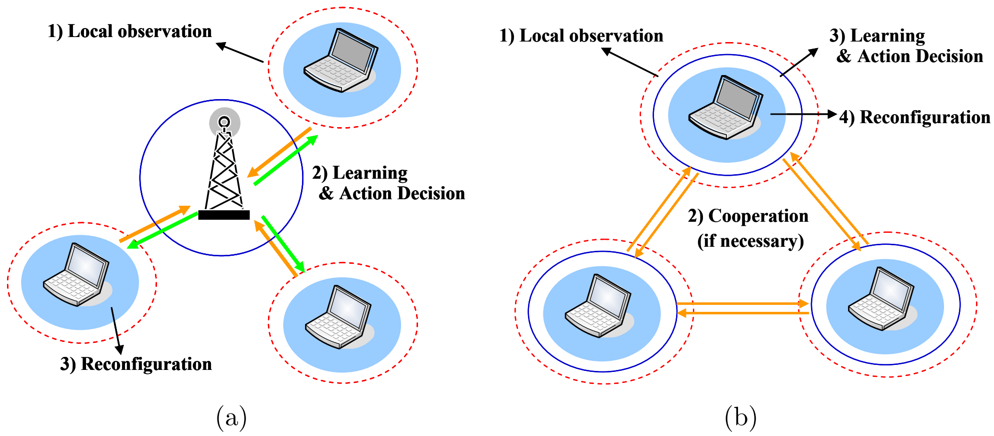


Figure 6: Comparison between CR capabilities for (a) infrastructure-based CR networks, and (b) CR ad hoc networks.

local observation, as shown in Figure 6 (b). Since the CR user cannot predict the influence of its actions on the entire network with its local observation, all of spectrum management functions are based on cooperative operation to broaden the knowledge on the network. In this scheme, all decisions are made based on the observed information that is gathered from their neighbors [1] [2].

In the following sections, we investigate how these spectrum management functions are integrated into the existing layering functionalities in CR networks and address the challenges of them. In this thesis, all proposed solutions are focused on the development of CR networks that require no modification of primary networks.

2.2 Spectrum Sensing

2.2.1 Basic Functionalities

A cognitive radio is designed to be aware of and sensitive to the changes in its surrounding, which makes spectrum sensing an important requirement for the realization of CR networks. Spectrum sensing enables CR users to exploit the unused spectrum portion adaptively to the radio environment. This capability is required in the following cases: (1) CR users find available spectrum holes over a wide frequency range for their transmission (*out-of-band sensing*), and (2) CR users monitor the spectrum

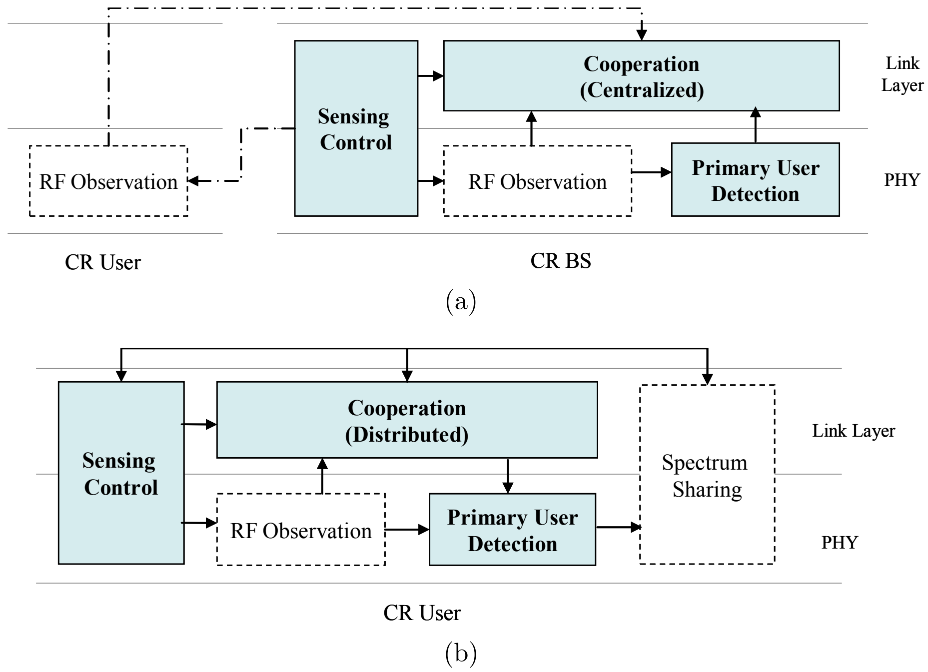


Figure 7: Functional block diagram for spectrum sensing: (a) infrastructure-based CR networks, and (b) CR ad hoc networks.

band during the transmission and detect the presence of primary networks to avoid interference (*in-band sensing*). As shown in Figure 7, the CR network necessitates the following functionalities for spectrum sensing:

- *PU Detection:* The CR user observes and analyzes its local radio environment. Based on these location observations of itself and its neighbors, CR users determine the presence of PU transmissions, and accordingly identify the current spectrum availability.

- *Cooperation:* The observed information in each CR user is sent to base-station or exchanged with its neighbors, and spectrum availability is determined accordingly. Through this cooperation, sensing accuracy is significantly improved.

- *Sensing Control:* The PU detection functionality is controlled and coordinated by a sensing controller, which considers two main issues on i) how quickly a CR user can find the available spectrum band over a wide frequency range for their transmissions [41] [42] [50], and ii) how long and how frequently a CR user should sense the spectrum to achieve sufficient sensing accuracy during the transmission and detect the presence of transmission in primary networks to avoid interference [27] [37] [55] [70].

Since CR networks are responsible for avoiding interference to primary networks, recent research has focused on improving sensing accuracy in PU detection. In [8], three different detection methods are investigated: *matched filter detection*, *energy detection*, and *feature detection*. A matched filter can perform coherent detection. On the contrary, energy detection is a non-coherent method that uses the energy of the received signal to determine the presence of primary signals. Feature detection exploits the inherent periodicity in the received signal [54]. To mitigate the multi-path fading and shadowing effects, cooperative detection methods among multiple CR users are proposed in [23] [52]. All these detection methods are based on transmitter detection to determine if a signal from a primary transmitter is locally present in a certain spectrum through the local observations of CR users. Unlike transmitter detection, a direct receiver detection method considers the location of primary receivers by exploiting the local oscillator (LO) leakage power of the primary receiver [74].

In infrastructure-based networks, the base-station plays a role in coordinating the operations of sensing operation through the synchronized sensing schedule. Sensing parameters determined through sensing control are applied to the sensing operations of all CR users. By considering all sensing information gathering from CR users, the base-station determines spectrum availability in its coverage, as shown in Figure 7 (a). On the other hand, due to the lack of strict coordination, CR ad hoc users perform sensing operations independently of each other, leading to an adverse influence on sensing performance. In the worst case, the sensing operations of one CR user may be interfered by the transmission of neighboring CR users, i.e. CR users cannot distinguish the signals from primary and CR users. Thus, spectrum sensing is closely coupled with spectrum sharing, especially medium access control (MAC) protocols, as depicted in Figure 7 (b).

2.2.2 Research Challenges

Although most of recent research in CR networks have explored spectrum sensing, the following issues need to be investigated further:

- *Optimization of Cooperative Sensing:* Cooperative sensing introduces another crucial issue. By requesting the sensing information from several CR users, the user that initiates the cooperative sensing, improves the accuracy. However, this also results in higher latency in collecting this information because of channel contention and packet re-transmissions. Thus, CR networks are required to consider these factors which must be optimized for correct and efficient sensing.

- *Support of Asynchronous Sensing:* If each user has independent and asynchronous sensing and transmission schedules, it can detect the transmissions of other CR users as well as primary users during its sensing period. However, with the energy detection, which is most commonly used for spectrum sensing, CR user cannot distinguish the transmission of CR and Primary users, and can detect only the presence of a transmission. As a result, the transmission of CR users detected during sensing operations causes false alarm in spectrum sensing, which leads to an increase in spectrum opportunities. Thus, how to coordinate the sensing cooperation of each CR user to reduce these false alarms is an important issue in spectrum sensing.

2.3 Spectrum Decision

2.3.1 Basic Functionalities

CR networks require capabilities to decide on the best spectrum band among the available bands according to the QoS requirements of the applications. This notion is called *spectrum decision* and constitutes a rather important but yet unexplored topic. Spectrum decision is closely related to the channel characteristics and the operations of primary users. Spectrum decision usually consists of two steps: First, each spectrum band is characterized based on not only local observations of CR users but also statistical information of primary networks. Then, based on this characterization, the most appropriate spectrum band can be chosen.

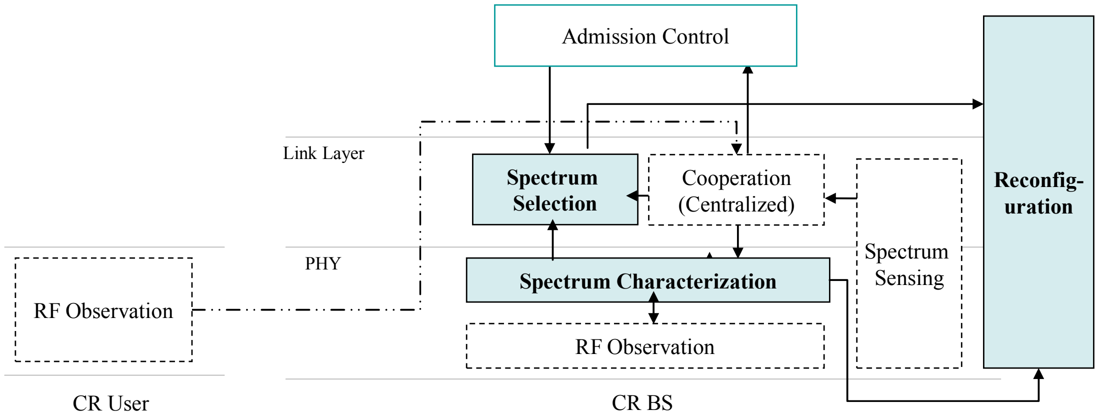
The following are main functionalities required for spectrum decision:

- *Spectrum Characterization:* Based on the observation, the CR users determine not only the characteristics of each available spectrum but also its PU activity model.

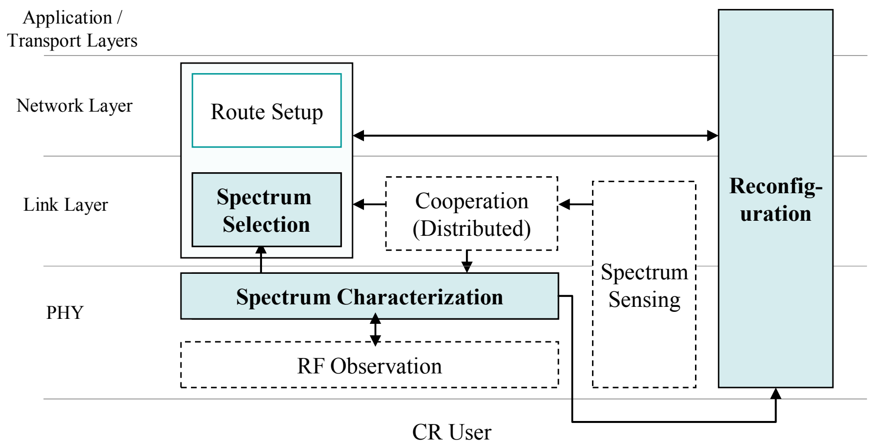
- *Spectrum Selection:* The CR user finds the best spectrum band to satisfy user QoS requirements.

- *Reconfiguration:* The CR users reconfigure communication protocol as well as communication hardware and RF front-end according to the radio environment and user QoS requirements.

CR users require spectrum decision in the beginning of the transmission. Through *RF observation*, CR users characterize available spectrum bands by considering the received signal strength, interference, and the number of users currently residing in the spectrum, which are also used for resource allocation in classical wireless networks. However, in CR networks, each user observes heterogeneous spectrum availability that is varying over time and space resulting from PU activities. This changing nature of the spectrum usage needs to be considered in the spectrum characterization. Based on this characterization, CR users determine the best available spectrum band to satisfy its QoS requirements. Furthermore, quality degradation of the current transmission can also initiate spectrum decision to maintain the quality of a current session.



(a)



(b)

Figure 8: Functional block diagram for spectrum decision: (a)infrastructure-based CR networks, and (b) CR ad hoc networks.

In infrastructure-based network, spectrum decision mainly focuses on allocating spectrum for a single hop to the base-station by considering current network utilization and the QoS requirements of a new incoming user. If the base-station cannot find the spectrum to satisfy the QoS requirements of the incoming user or adding the incoming user will expect significant quality degradation of current users, the base-station does not accept this incoming users through the *admission control*. Once the base-station admits the user, it allocates the best spectrum to the user as explained in Figure 8 (a). Unlike infrastructure-based CR networks, CR ad hoc networks have unique characteristics in spectrum decision due to the nature of multi-hop communication. Spectrum decision needs to consider the end-to-end route consisting of multiple hops. Furthermore, available spectrum bands in CR networks differ from one hop to the other. As a result, the connectivity is spectrum-dependent, which makes it challenging to determine the best combination of the routing path and spectrum. Thus, spectrum decision in ad hoc networks should interact with routing protocols [51] [71], which will be explained in Figure 8 (b).

2.3.2 Research Challenges

The following are open research issue in spectrum decision:

- *PU Activity Modeling:* Most of the current research on spectrum sensing are based on a simple ON-OFF model for PU activities, which cannot capture the diverse characteristics of all existing primary networks. This inaccurate model for primary networks leads to an adverse influence on spectrum sensing resulting in either lower spectrum access opportunities or higher interference to the primary networks. Some of the empirical models on PU activities [25] [75] are not computationally feasible in practical situations. Thus, we need to develop more practical PU activity models by considering the characteristics of access technologies as well as traffic types.

- *Joint Spectrum Decision and Reconfiguration Framework:* Once the available spectrum bands are characterized, the most appropriate spectrum band should be selected by considering the QoS requirements (sustainable rate, delay, jitter, average session time, acceptable loss rate, etc) and the spectrum characteristics. However, according to the reconfigurable transmission parameters such as modulation type, error control scheme, and communication protocol, these spectrum characteristics change significantly. Sometimes, with only reconfiguration, CR users can maintain the quality of the current session. For example, even if a signal-to-noise ratio (SNR) is changed, both bitrate and bit error rate (BER) can be maintained by exploiting an adaptive modulation, instead of changing spectrum and route. Hence, there is a need for a joint spectrum decision and reconfiguration framework to find the optimal combination of the spectrum band and parameter configuration according to applications with diverse QoS requirements.

2.4 Spectrum Sharing

2.4.1 Basic Functionalities

The shared nature of the wireless channel necessitates coordination of transmission attempts between CR users. In this respect, spectrum sharing provides the capability to maintain the QoS of CR users without causing interference to the primary users by coordinating multiple access of CR users as well as allocating communication resources adaptively to the changes of radio environment. Thus, spectrum sharing is performed in the middle of a communication session and within the spectrum band, and includes many functionalities of a medium access control (MAC) protocol and resource allocation in classical ad hoc networks. However, the unique characteristics of cognitive radios such as the coexistence of CR users with primary users and the wide range of available spectrum incur substantially different challenges for spectrum sharing in CR ad hoc networks.

Spectrum sharing techniques are generally focused on two types of solutions, i.e., spectrum sharing inside a CR network (intra-network spectrum sharing), and among multiple coexisting CR networks (inter-network spectrum sharing) [3]. Inter-network spectrum sharing can be implemented either based on a *spectrum broker* that is connected to the base-station [7] [24] [33] [79] or in a distributed approach without support of the central network entity [43] [47].

Figure 9 depicts the functional blocks for spectrum sharing in CR networks. Unlike spectrum decision, spectrum sharing mainly focuses on resource management within the same spectrum with the following functionalities:

- *Resource Allocation:* Based on the QoS monitoring results, CR users select the proper channels (channel allocation) [10] [11] [56] and adjust their transmission power (power control) [19] [30] [77] to achieve QoS requirements as well as resource fairness. Especially in power control, sensing results need to be considered so as not to violate the interference constraints.

- *Spectrum Access:* It enables multiple CR users to share spectrum resources by determining who will access the channel or when a user may access the channel [16] [32] [35].

Once a proper spectrum band is selected in spectrum decision, communication channels in that spectrum need to be assigned to a CR user while determining its transmission power to avoid the interference to the primary network (*resource allocation*). Then, the CR user decides when the spectrum should be accessed to avoid collisions with other CR users (*spectrum access*).

The infrastructure-based network can provide sophisticated spectrum sharing method with support of the base-station. Thus, it can exploit time slot-based scheduling and dynamic channel allocation to maximize the total network capacity as well as achieve fair resource allocation over CR users. Furthermore, through the synchronization in sensing operation, the transmission of CR users and primary users can be detected separately, which decouples sensing operation with spectrum sharing. Generally, CR networks uses a periodic sensing scheme where CR users are allowed to transmit only during the transmission period followed by sensing (observation) period. In this architecture, the transmission period is synchronized over all CR users. Thus, spectrum sharing needs to focus on channel allocation or time-slot-based scheduling within this transmission period. Also spectrum sharing just exploits the just exploits spectrum

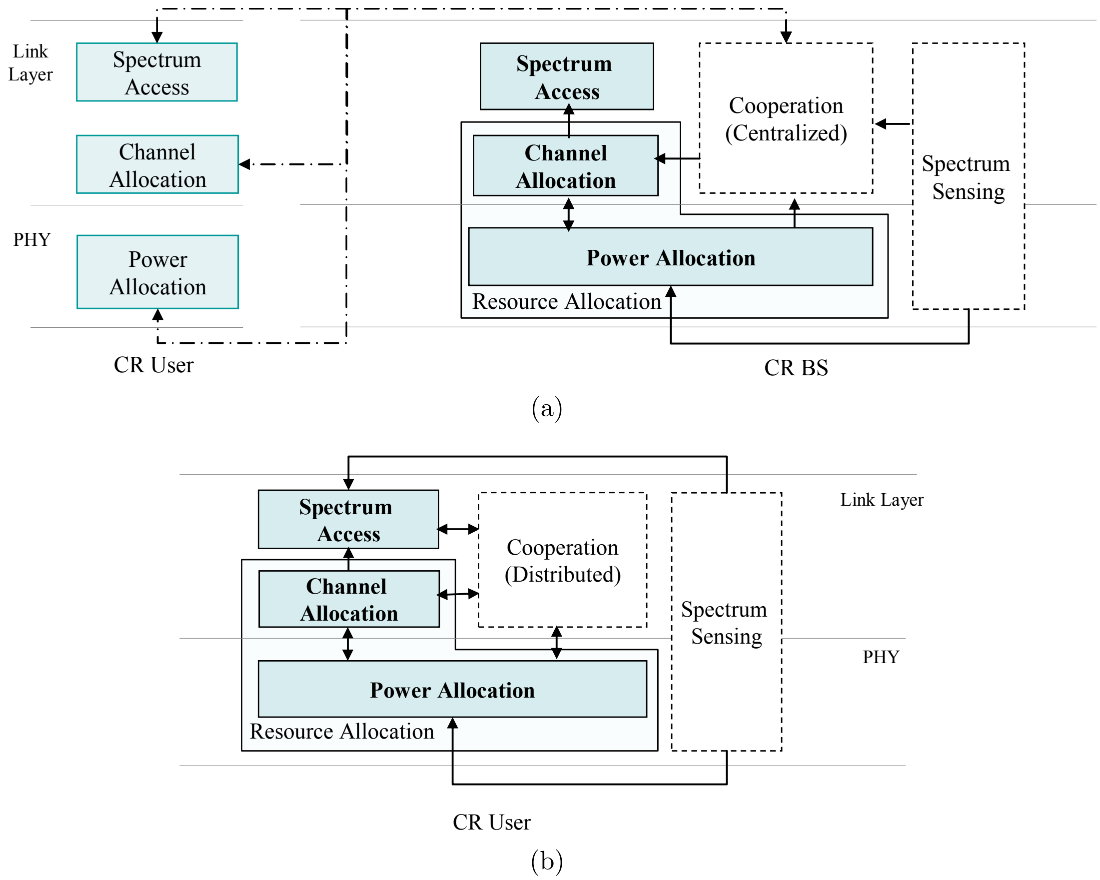


Figure 9: Functional block diagram for spectrum sharing: (a) infrastructure-based CR networks, and (b) CR ad hoc networks.

availability determined in the sensing and is not directly related to spectrum sensing. Similar to spectrum sensing and decision, all sharing operations in CR users are coordinated by the base-station, as illustrated in Figure 9 (a).

On the contrary, in CR ad hoc networks, the sensing schedules are determined and controlled by each user and not being controlled and synchronized by the central network entity. Thus, instead of this periodic sensing, CR ad hoc users may adopt the aperiodic or on-demand sensing triggered by only spectrum sharing operations can trigger spectrum sensing, i.e., when CR users want to transmit or are requested their spectrum availability by neighbor users. Furthermore, sensing and transmission intervals, determined by the sensing control in spectrum sensing, influence the performance of spectrum access. As a result, spectrum sensing should be integrated into spectrum sharing, especially in spectrum access functionality, which is shown in Figure 9 (b).

2.4.2 Research Challenges

Current research challenges in spectrum sharing are presented as follows:

- *Distributed Power Allocation:* The CR ad hoc user determines the transmission power in a distributed manner without support of the central entity. Infrastructure-based networks also need this distributed power allocation scheme for inter-network spectrum sharing among neighbor base-stations or other CR networks. However, these operations may cause interference because of the limitation of sensing area even if it does not detect any transmission in its observation range. Thus, spectrum sharing necessitates sophisticated power control methods for adapting to the time-varying radio environment to maximize capacity with the protection of the transmissions of primary users.

- *Reliable Control Channel:* To share spectrum resources efficiently, CR transmitter should have feedback information regarding channel condition and QoS status from its receiver. Thus, each CR user necessitates a reliable control channel for exchanging control information. The control channel can be established through either out-of-band or in-band signalling. However, with the in-band signaling, it is not easy to find the neighbor users tuning different spectrum band and exchange information. We may use the dedicated control channel based on out-of-band signalling, which is not reliable due to PU activities. Especially in CR ad hoc networks, asynchronous sensing and transmission schedules make it more difficult to exchange information with its neighbors. As a result, how to reliably obtain the channel and QoS information from the receiver or neighbor users is still unsolved in networks.

2.5 Spectrum Mobility

2.5.1 Basic Functionalities

CR users are generally regarded as 'visitors' to the spectrum. Hence, if the specific portion of the spectrum in use is required by a PU, the communication needs to be continued in another vacant portion of the spectrum. This notion is called *spectrum mobility*. Spectrum mobility gives rise to a new type of handoff in CR networks, the so-called *spectrum handoff*, in which, the users transfer their connections to an unused spectrum band. In CR ad hoc networks, spectrum handoff occurs 1) when PU is detected, 2) the CR user loses its connection resulting from the mobility of users involved in an on-going communication, or 3) with a current spectrum band cannot provide the QoS requirements.

In spectrum handoff, temporary communication break is inevitable because of the process for discovering a new available spectrum band. Since available spectrums are dis-contiguous and distributed over a wide frequency range, CR users may require the reconfiguration of operation frequency in its RF front-end, which leads to a significantly longer switching time. The purpose of the spectrum mobility management in CR ad hoc networks is to ensure smooth and fast transition leading to minimum performance degradation during a spectrum handoff. Furthermore, in spectrum mobility, the protocols for different layers of the network stack should be transparent to the spectrum handoff and the associated latency, and adapt to the channel parameters of the operating frequency. We describe this adaptation in the routing and transport protocols

Another intrinsic characteristic of spectrum mobility in CR networks is the interdependency with the routing protocols. Similar to spectrum decision, spectrum mobility needs to involve the recovery of link failure on the end-to-end route. Thus, it needs to interact with routing protocols to detect the link failure resulting from either user mobility or PU appearance, which is explained in Figure 10.

In the following, the main functionalities required for spectrum mobility in the CR ad hoc network are described:

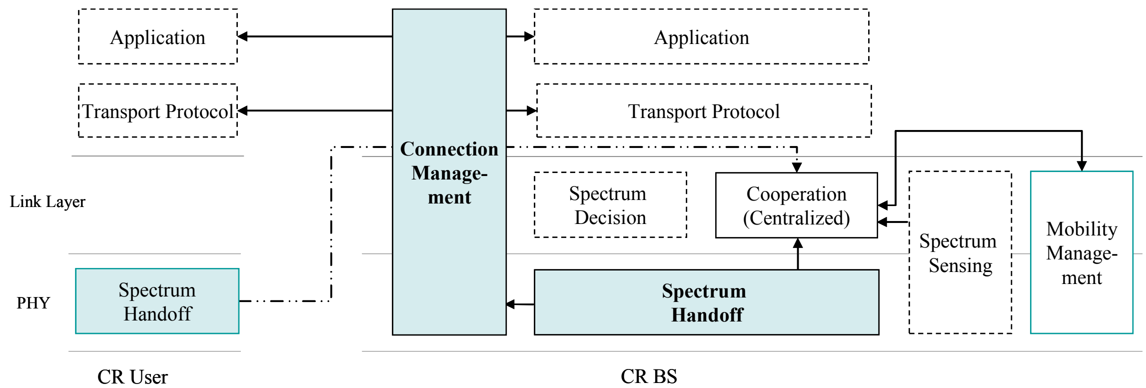
- *Spectrum Handoff:* The CR user switches the spectrum band physically and reconfigures the communication parameters for an RF front-end (e.g. operating frequency, modulation type).

- *Connection Management:* The CR user sustains the QoS or minimizes quality degradation during the spectrum switching by interacting with each layering protocols.

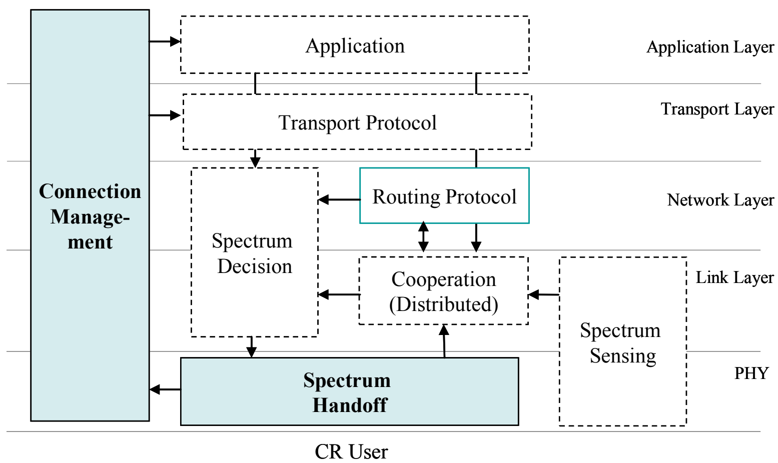
As stated previously, spectrum mobility events can be detected as a link failure caused by user mobility as well as PU detection. Furthermore, the quality degradation of the current transmission also initiates spectrum mobility. When these spectrum mobility events are detected through spectrum sensing, neighbor discovery, routing protocol, and mobility management function, they trigger spectrum mobility procedures. By collaborating with spectrum decision, a CR user determines a new spectrum band on the determined route, and switch its current session to the new spectrum (*spectrum handoff*). During the spectrum handoff, the CR user need to maintain current transmission not to be interfered by the switching latency.

Figure 10 (a) shows spectrum mobility functionalities for infrastructure-based networks. In this architecture, once the base-station detects the primary user appearance in spectrum sensing, user mobility in the mobility management, or quality degradation, it vacates current spectrum and moves to the new re-allocated spectrum or to new base-station if necessary. During the spectrum switching time, the base-station minimizes the influence on performance of upper-layer protocols and sustain the level of qualities required by user application through connection management function.

On the contrary, spectrum mobility in ad hoc networks needs to consider mainly focuses on link failure on the end-to-end route. Furthermore, compared to the



(a)



(b)

Figure 10: Functional block diagram for spectrum mobility: (a) infrastructure-based CR networks, and (b) CR ad hoc networks.

infrastructure-based network, the CR ad hoc network has more dynamic and complicated topology dependent on both spectrum and user mobilities. Furthermore, as shown in Figure 10 (b), the CR ad hoc network uses routing protocol to recover the link failure on its end-to-end route, but cannot manage the mobility events as efficiently as the infrastructure-based networks due to the lack of the central entity as well as more complicated topology. For these reasons, it is much more difficult to design spectrum mobility in CR ad hoc networks compared to the infrastructure-based networks.

2.5.2 Research Challenges

To the best of our knowledge, there exists no research effort to address the problems of spectrum mobility in CR networks to date. Although the routing mechanisms that have been investigated in the classical ad hoc networks may lay the groundwork in this area, there still exist many open research topics:

- *Switching Delay Management:* The spectrum switching delay is closely related to not only hardware, such as an RF front-end, but also to algorithm development for spectrums sensing, spectrum decision, link layer, and routing. Thus, it is desirable to design spectrum mobility in a cross-layer approach to reduce the operational overhead among each functionalities and to achieve a faster switching time. Furthermore, the estimation of accurate latency in spectrum handoff is essential for reliable connection management.

- *Flexible Spectrum Handoff Framework:* CR networks have two different spectrum handoff strategies: *reactive* and *proactive spectrum handoffs*, which show different influence on the communication performance. Furthermore, according to the mobility event, a spectrum switching time will change. For example, since a PU activity region is typically larger than the transmission range of CR users, multiple hops may be influenced by spectrum mobility events at the same time, which makes the recovery time much longer. Furthermore, in the case of delay-sensitive applications, CR users can use a proactive switching, instead of a reactive switching. In this method, through the prediction of PU activities, CR users switch the spectrum before PUs appear, which helps to reduce the spectrum switching time significantly [76]. On the other hand, energy constrained devices such as sensors need reactive spectrum switching. Thus, we need to develop a flexible spectrum handoff framework to exploit different switching strategies adapting to the type of applications and network environment.

CHAPTER III

OPTIMAL SPECTRUM SENSING FRAMEWORK FOR COGNITIVE RADIO NETWORKS

3.1 Introduction

A cognitive radio is designed to be aware of and sensitive to the changes in its surrounding, which makes spectrum sensing an important requirement for the realization of cognitive radio networks. *Spectrum sensing* enables unlicensed users, referred to as *CR users*, to adapt to the environment by detecting unused spectrum portions without causing interference to the licensed network, referred to as the *primary network*.

The main objective of spectrum sensing is to provide more spectrum access opportunities to CR users without interference to the primary networks. Since CR networks are responsible for detecting the transmission of primary networks and avoiding interference to them, CR networks should intelligently sense the primary band to avoid missing the transmission of primary users. Thus, *sensing accuracy* has been considered as the most important factor to determine the performance of CR networks. Hence, recent research has focused on improving sensing accuracy for interference avoidance.

Although all these efforts enable CR users to enhance the sensing accuracy, the hardware limitations of CR users introduce a new critical issue on spectrum sensing. Ideally, to avoid interference to the primary users, CR users should monitor the spectrum continuously through the RF front-end. However, in reality, the RF front-end cannot differentiate between the primary user signals and CR user signals [63]. While feature detection is known to be capable of identifying the modulation types of the primary signal, it requires a longer processing time as well as higher computational complexity [31]. With energy detection, mostly used in spectrum sensing, CR users are not able to perform the transmission and sensing tasks at the same time. Thus, owing to this hardware limitation, CR users necessitate a *periodic sensing* structure where sensing and transmission operations are performed in a periodic manner with separate observation period and transmission period. In this structure, CR users should stop their transmissions during the sensing time to prevent false alarms triggered by unintended CR signals.

This periodic sensing structure introduces the following design issues:

- *Interference Avoidance:* Interference in CR networks depends on sensing accuracy, which is determined by the observation time. However, in periodic sensing, CR users cannot sense the spectrum bands during the transmission time, which leads to the increase in interference. Thus, for the interference avoidance, both the observation time and the transmission time need to be considered in the periodic spectrum sensing method.

- *Sensing Efficiency:* The main objective of CR networks is efficient spectrum utilization. Thus, the spectrum sensing functionality should provide more transmission opportunities to CR users. However, during the observation period, the transmission of CR users is not allowed, which inevitably decreases the transmission opportunities of CR users, leading to the so-called *sensing efficiency* issue.

As explained above, there is a trade-off between interference and sensing efficiency. For interference avoidance, the observation time needs to be long enough to achieve sufficient detection accuracy, i.e., a longer observation time leads to higher sensing accuracy, and hence to less interference. But as the observation time becomes longer, the transmission time of CR users will be decreased. Conversely, while a longer transmission time enhances sensing efficiency, it causes higher interference due to the lack of sensing information. Hence, *observation time* and *transmission time* are the sensing parameters that mainly influence both the spectrum efficiency and interference avoidance. Thus, the proper selection of these sensing parameters is the most critical factor influencing the performance of CR networks.

Besides spectrum sensing parameters, there are two more crucial factors to be considered if the spectrum sensing method is applied to multi-spectrum/multi-user networks. Usually, CR users are allowed to exploit multiple spectrum bands. However, practically, CR users do not have enough sensing transceivers to sense all the available spectrum bands. To maximize the spectrum access opportunities of CR users subject to the transceiver constraint, a well-defined spectrum selection method is essential.

Furthermore, there exists a high spatial correlation among sensing data detected from different locations in CR networks since neighboring CR users are highly likely to be located in the same transmission range of the primary network. Cooperative sensing is the traditional approach to exploit this spatial correlation in multi-user networks by allowing CR users to exchange their sensing information. In cooperative sensing, the number of cooperating users affects sensing accuracy, and hence the sensing parameters. Since the number of users varies over time, it is essential for CR networks to adaptively decide the optimal sensing parameters with varying number of users.

As mentioned above, spectrum sensing primarily requires the decision of the proper sensing parameters by considering both spectrum efficiency and interference avoidance. However, in multi-spectrum/multi-user network environments, the spectrum sensing method is required to provide additional functionalities such as spectrum selection and multi-user cooperation. Thus, a unified spectrum sensing framework needs to be developed to consider all possible network environments and define inter-operations of all functionalities.

Hence, in this chapter, to solve both the interference avoidance and sensing efficiency problems, we develop an optimal sensing framework to maximize spectrum access opportunities considering interference and sensing resource limitations. More specifically, a theoretical framework is developed for the optimization of sensing parameters to maximize spectrum efficiency subject to interference constraints in a single spectrum band. For multi-spectrum environments, based on the optimal sensing parameters, a novel sensing resource allocation method is developed to maximize the spectrum access opportunities of CR users. Finally, to exploit sensing accuracy gain obtained by the multi-user cooperation, we propose an adaptive and cooperative decision method for the sensing parameters, where the transmission time can be optimized adaptively to the number of users.

The remainder of the chapter is organized as follows. The system model used in this chapter is presented in Section 3.2. In Section 3.3, we introduce a theoretical framework for sensing parameter optimization along with detection and the interference models. Then, we describe spectrum selection and resource scheduling for multi-spectrum sensing in Section 3.4. In Section 3.5, we investigate how cooperation gain influences sensing parameter optimization and propose an adaptive and cooperative sensing method to exploit the cooperation gain. Performance evaluation and simulation results are presented in Section 3.6.

3.2 System Model

3.2.1 System Description

The design objective of CR networks is to exploit the best available spectrum bands. To achieve this goal, spectrum sensing needs to consider the requirements on the network architecture, terminal hardware capabilities, and the radio environment as explained below.

3.2.1.1 Network Architecture

In this chapter, we assume CR networks have a centralized network entity such as a base-station in infrastructure-based networks. Ad hoc networks are assumed to have a cluster head node. This centralized network entity can communicate with all CR users within its coverage and decide the spectrum availability of its coverage.

There are two main reasons to adopt a centralized network architecture. The first reason is the receiver uncertainty problem. With the transmitter detection, CR networks cannot avoid interference at the nearby primary receivers since the transmitter detection relies only on local observations of CR users and does not consider the location information of the primary receivers [3]. Hence, to reduce the receiver uncertainty, CR networks require the base-station[1] to collect sensing information from CR users inside its coverage. The second reason is the limitation in sensing capabilities. All CR users have the same sensing cycles not to interfere with sensing operations, which means that CR networks should be synchronized to schedule spectrum sensing. Thus, CR networks need to have the base-station to synchronize the scheduling.

3.2.1.2 CR User Requirements

Here, CR users are assumed to use energy detection for spectrum sensing. Furthermore, CR users may have multiple software-defined radio (SDR) transceivers to exploit multiple spectrum bands over wide frequency ranges by adjusting the operating frequency through software operations. Each transceiver can be used for the purpose of both transmission and sensing.

3.2.1.3 Radio Environment

In CR networks, all available spectrum bands are spread over a wide frequency range, and hence exhibit different characteristics. In this chapter, CR networks are assumed

[1] In the remainder of the chapter we will use the term "base-station" to refer to the centralized network entity both in infrastructure-based networks and in ad hoc networks.

to be aware of the following *a priori* spectrum information of primary networks:

- *Operating Frequency Range:* CR users are aware of the bandwidth and of the frequency range of the primary networks.

- *Minimum Signal-to-Noise Ratio (SNR):* To determine spectrum availability, CR users need statistical information on the received primary signals. The minimum SNR is the least signal level needed to decode the received signals, depending on the modulation type, channel coding and multiple access methods of primary user networks.

- *Primary User Activity:* This is defined as the traffic statistics of the primary networks, which will be explained more in detail in Section 3.2.2.

- *Interference Constraint:* Since CR users cannot monitor the spectrum continuously, CR networks do not guarantee interference-free transmissions. Instead, CR networks exploit the interference constraint, which can be defined as either maximum interference level or maximum interference probability that primary networks can tolerate. Although the former is the most suitable for the objective of the opportunistic transmission, the latter is more practical since there is no practical way to measure the amount of the interference at the nearby primary receivers.

3.2.2 Primary User Activity Model

Since PU activity is closely related to the performance of CR networks, the estimation of this activity is a very crucial issue in spectrum sensing. We assume that PU activity can be modeled as exponentially distributed inter-arrivals. In this model, the primary user traffic can be modeled as a two state birth-death process with death rate α and birth rate β . An ON (Busy) state represents the period used by primary users and an OFF (Idle) state represents the unused period [14] [15]. Since each user arrival is

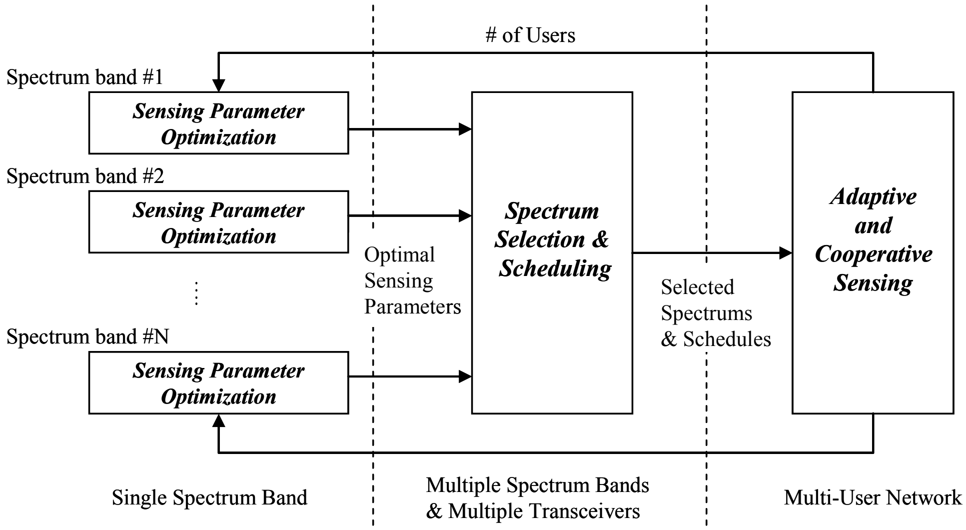


Figure 11: The proposed optimal spectrum sensing framework.

independent, each transition follows the Poisson arrival process. Thus, the length of ON and OFF periods are exponentially distributed [66].

Since primary user activity is closely related to the performance of CR networks, the estimation of this activity is a very crucial issue in spectrum sensing. We assume that primary user activity can be modeled as exponentially distributed inter-arrivals. In this model, the primary user traffic can be modeled as a two state birth-death process with death rate α and birth rate β. An ON (Busy) state represents the period used by primary users and an OFF (Idle) state represents the unused period [14], [15]. Since each user arrival is independent, each transition follows the Poisson arrival process. Thus, the length of ON and OFF periods are exponentially distributed [66].

3.2.3 Optimal Spectrum Sensing Framework

In this chapter we develop an optimal spectrum sensing framework, which is illustrated in Figure 11. The proposed framework consists of the *optimization of sensing parameters* in a single spectrum band, *spectrum selection and scheduling*, and an *adaptive and cooperative sensing* method.

The detailed scenario for the optimal sensing framework is as follows. According to the radio characteristics, base-stations initially determine the optimal sensing parameters of each spectrum band through the *sensing parameter optimization*. When CR users join the CR networks, they select the best spectrum bands for sensing and configure sensing schedules according to the number of transceivers and the optimized sensing parameters by using *spectrum selection and scheduling* methods. Then, CR users begin to monitor spectrum bands continuously with the optimized sensing schedule and report sensing results to the base-station. Using these sensing results, the base-station determines spectrum availability. If the base-station detects any changes which affect the sensing performance, sensing parameters need to be re-optimized and announced to its CR users through the *adaptive and cooperative sensing*.

3.3 Sensing Parameter Optimization in a Single Spectrum Band

In the preceding discussions, we defined the *a priori* information for spectrum sensing and introduced the optimal sensing framework consisting of three functionalities, namely *sensing parameter optimization*, *spectrum selection and scheduling* for multiple spectrum bands, and *adaptive and cooperative sensing* in multi-user networks. In this section, we first propose a sensing parameter optimization method to maximize the spectrum efficiency subject to the interference constraint.

3.3.1 Problem Definition

Consider a typical sensing scenario in which a single CR user monitors a single spectrum band. The CR user alternately senses the spectrum and transmits data with observation time t_{s} and transmission time T. To determine these sensing parameters accurately, we need to consider the interference constraint and sensing efficiency at the same time. Therefore, we introduce the following definitions:

Definition 1: The *interference ratio* T_{I} is the expected fraction of the ON state (i.e., the transmission time of primary networks) interrupted by the transmission of CR users, which will be derived in Eq. (13).

Definition 2: The *lost spectrum opportunity ratio* T_{L} is the expected fraction of the OFF state (i.e, idle time) undetected by CR users, which will be derived in Eq. (14).

Definition 3: The *maximum outage ratio* T_{P} is the maximum fraction of interference that primary networks can tolerate.

Definition 4: The *sensing efficiency* η is the ratio of the transmission time over the entire sensing cycle, defined as follows:

$$\eta = \frac{T}{T + t_{\mathrm{s}}} \tag{1}$$

The objective of spectrum sensing is to achieve accurate detection probability as well as high sensing efficiency. Since both metrics are related to the sensing parameters T and t_{s}, the sensing parameter decision can be expressed as the optimization problem to maximize the spectrum efficiency satisfying interference constraint T_{P} as follows:

$$\begin{aligned} &\text{Find: } T^*, t_{\mathrm{s}}^* \\ &\text{Maximize: } \eta = \frac{T}{T + t_{\mathrm{s}}} \\ &\text{Subject to: } T_{\mathrm{I}} \leq T_{\mathrm{P}} \end{aligned} \tag{2}$$

where t_{s}^*, T^* are optimal observation and transmission times, respectively.

In the following subsections, we first explain a maximum a posteriori (MAP) based energy detection model, and then we propose an analytical interference model. Finally, we show how to optimize sensing parameters based on the MAP based energy detector and the interference model.

3.3.2 Maximum A Posteriori (MAP) Energy Detection for Spectrum Sensing

Because of the interference constraints in CR networks, spectrum sensing method needs to develop a more accurate detection scheme. Although a MAP detector is known to be optimal [57], a maximum likelihood (ML) detection has been widely used for the energy detection without considering the probabilities of ON and OFF states [52], [68], [14]. In this chapter, we propose MAP-based energy detection and its decision criterion based on the primary user activities.

When CR users observe the spectrum to detect the primary user activity, the received signal $r(t)$ takes the following form [18]:

$$r(t) = \begin{cases} n(t) & H_0 \\ s(t) + n(t) & H_1 \end{cases} \tag{3}$$

where H_0 represents the hypothesis corresponding to "no signal transmitted", and H_1 to "signal transmitted". $s(t)$ is the signal waveform, and $n(t)$ is a zero-mean additive white Gaussian noise (AWGN).

Assume the spectrum has bandwidth W and the primary user activities with death rate α and birth rate β. From the primary user activity model, we can estimate the *a posteriori* probabilities as follows [17]:

$$\begin{aligned} P_{\text{on}} &= \frac{\beta}{\alpha + \beta} \\ P_{\text{off}} &= \frac{\alpha}{\alpha + \beta} \end{aligned} \tag{4}$$

where P_{on} is the probability of the period used by primary users and P_{off} is the probability of the idle period. From the definition of MAP detection, the detection probability P_{d} and false alarm probability P_{f} can be expressed as follows:

$$P_{\text{d}}(\lambda) = \Pr[Y > \lambda | H_1] P_{\text{on}} = \bar{P}_{\text{d}} \cdot P_{\text{on}} \tag{5}$$

$$P_{\text{f}}(\lambda) = \Pr[Y > \lambda | H_0] P_{\text{off}} = \bar{P}_{\text{f}} \cdot P_{\text{off}} \tag{6}$$

where λ is a decision threshold of MAP detection.

Generally, the decision threshold, λ can be determined by the minimum probability of error decision rule as $f(\lambda|H_1)P_{\text{on}} = f(\lambda|H_0)P_{\text{off}}$ where $f(y|H_1)$ and $f(y|H_0)$ are probability density functions of the received signal through the occupied spectrum and the idle spectrum, respectively. This method minimizes the total error probabilities, including false alarm and miss detection. However, in this method, sometimes one of the error probabilities may be greater than the other. In [14], to achieve the best trade-off between false alarm and detection error, this decision rule is dynamically exploited by considering the interference constraint which is assumed to be equal to the detection error probability. However, in reality, the false alarm probability also affects the interference, which will be explained in Section 3.3.3. Furthermore, in spectrum sensing, the detection of opportunities is as much important as that of the primary signals. Hence, instead of minimizing the total error probability, in this chapter, we emphasize the balance of both error probabilities as follows:

$$P_{\mathrm{on}} - P_{\mathrm{d}}(\lambda) = P_{\mathrm{f}}(\lambda) \tag{7}$$

This method enables balancing between the interference T_{I} and the lost spectrum opportunity T_{L} caused by the detection errors and the false alarms.

Based on the MAP detection model explained above, we derive the detection and false alarm probabilities of energy detection. In order to measure the energy of the received signal, the output signal of bandpass filter with bandwidth W is squared and integrated over the observation interval t_{s}. Finally, the output of the integrator, Y, is compared with a threshold, λ, to decide whether a licensed user is present or not. The output of the integrator in the energy detector is known as the Chi-square distribution [18]. However, if the number of samples is large, we can use the central limit theorem to approximate the Chi-square distribution as Gaussian distribution [68].

$$Y \sim \begin{cases} \mathcal{N}(n{\sigma_{\mathrm{n}}}^2, 2n{\sigma_{\mathrm{n}}}^4), & H_0 \\ \mathcal{N}(n({\sigma_{\mathrm{n}}}^2 + {\sigma_{\mathrm{s}}}^2), 2n({\sigma_{\mathrm{n}}}^2 + {\sigma_{\mathrm{s}}}^2)^2), & H_1 \end{cases} \tag{8}$$

where n is the number of samples, ${\sigma_{\mathrm{n}}}^2$ is the variance of the noise, and ${\sigma_{\mathrm{s}}}^2$ is the variance of the received signal $s(t)$. According to the Nyquist sampling theorem, the minimum sampling rate should be $2W$. Hence n can be represented as $2t_s W$ where t_s is the observation time.

From Eq. (5), (6), and (8), P_f and P_d in MAP-based energy detection can be derived in terms of the Q function as follows:

$$P_f(W, t_s, \alpha, \beta) = \frac{\alpha}{\alpha + \beta} \cdot Q(\frac{\lambda - 2t_s W {\sigma_n}^2}{\sqrt{4t_s W {\sigma_n}^4}}) \quad (9)$$

$$P_d(W, t_s, \alpha, \beta) = \frac{\beta}{\alpha + \beta} \cdot Q(\frac{\lambda - 2t_s W ({\sigma_s}^2 + {\sigma_n}^2)}{\sqrt{4t_s W ({\sigma_s}^2 + {\sigma_n}^2)^2}}) \quad (10)$$

From Eq. (9) and (10), we can see that each spectrum band has different detection and false alarm probabilities according to the spectrum information, α, β, and W, as well as the observation time t_s.

The decision threshold λ can be obtained by means of numerical methods. However, since λ is independent of the observation time t_s, it is not required to find optimal sensing parameters, T^* and t_s^*, which is explained in Appendix B.

3.3.3 Analytical Model for Interference

To optimize sensing parameters satisfying the interference constraint, we need to specify the relation between the interference ratio T_I and sensing parameters, as explained in Section 3.3.1. Hence, we propose an analytical interference model as a function of primary user activities and detection statistics derived in Section 3.3.2.

In periodic sensing, interference can be expected to occur in the following cases:

- *Interference on busy state sensing, I_{on}:* In this case, the spectrum band is busy, but the CR user does not detect the primary user signals and begins to transmit. As a result, interference can occur during the transmission period T, as shown in Figure 12 (a).

- *Interference on idle state sensing, I_{off}:* Even though the spectrum band is idle and CR users detect it correctly, there still exists the possibility that a

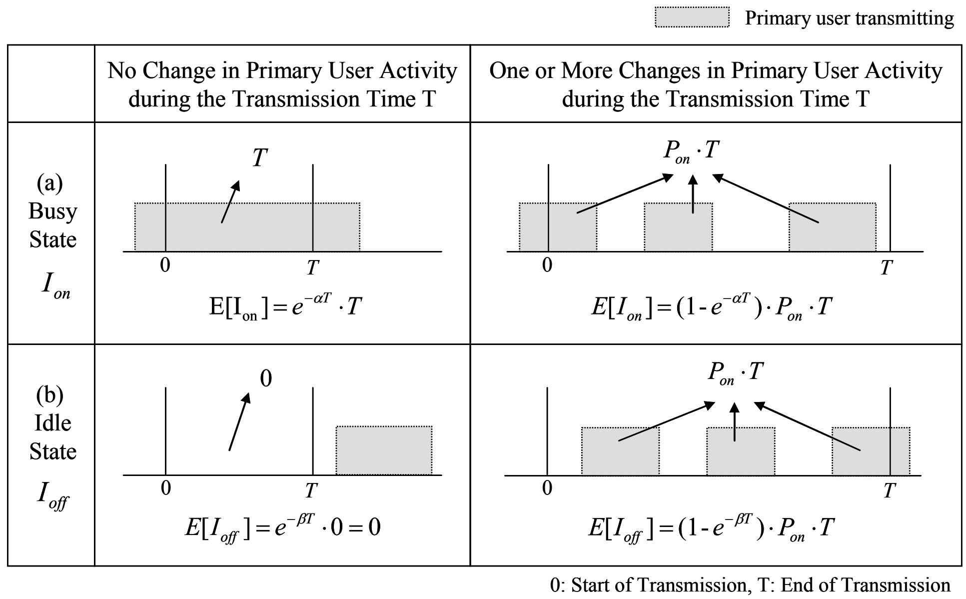


Figure 12: Interference model in busy state and idle state sensings.

primary user activity may appear during the transmission period T, as shown in Figure 12 (b).

As shown in Figure 12 (a), the interference I_{on} has two different patterns according to the transmission time T. The left figure depicts the interference over the entire transmission period T. The right figure describes the interference in case there are one or more changes in primary user activities during T. From the primary user activity model explained in Section 3.2.2, the probability that the spectrum band is busy during the entire transmission time T, can be obtained as $e^{-\alpha T}$, and the probability with one or more transition of primary user activities during T is $1 - e^{-\alpha T}$.

If T is relatively short, the spectrum state does not change during the transmission time T. Thus, the interference is highly likely to persist over the entire transmission period with probability $e^{-\alpha T}$, as shown in the left column of Figure 12 (a). However, if T is long enough, busy and idle states occur alternately during T and hence, interference converges to $P_{\text{on}} \cdot T$ with probability $1 - e^{-\alpha T}$, as shown in the right column of Figure 12 (a). Thus, the expected interference on the busy state sensing

$\mathrm{E}[I_{\text{on}}]$ during the transmission time T, can be expressed as follows:

$$\begin{aligned}\mathrm{E}[I_{\text{on}}] &= (P_{\text{on}} - P_{\text{d}})(e^{-\alpha T}T + (1 - e^{-\alpha T})P_{\text{on}}T) \\ &= P_{\text{off}}\bar{P}_{\text{f}}(\frac{\alpha}{\alpha+\beta}Te^{-\alpha T} + \frac{\beta}{\alpha+\beta}T)\end{aligned} \tag{11}$$

Similarly, in the case of interference in the idle state, I_{off}, the interference only occurs when one or more primary user activities occur during the transmission time, which converges approximately to $P_{\text{on}} \cdot T$ with the probability $1 - e^{-\beta T}$, as shown in Figure 12 (b).

$$\begin{aligned}\mathrm{E}[I_{\text{off}}] &= (P_{\text{off}} - P_{\text{f}})(e^{-\beta T} \cdot 0 + (1 - e^{-\beta T})P_{\text{on}}T) \\ &= P_{\text{off}}(1 - \bar{P}_{\text{f}})(1 - e^{-\beta T})\frac{\beta}{\alpha+\beta}T\end{aligned} \tag{12}$$

While the proposed models provide a close approximation in the expected interference over an entire transmission time range, they may show a finite approximation error when the transmission time T is shorter compared to the average busy time $1/\alpha$ or the average idle time $1/\beta$, which is more realistic assumption in CR networks. For example, if $\alpha > \beta$ and $T < 1/\beta$, the interference in the idle state will be much greater than $\mathrm{E}[I_{\text{off}}]$ given in Eq. (12) since a higher primary user activity α is a more dominant factor in determining interference in the above short transmission time. This approximation error can be mitigated as the average interference free period, i.e., idle time in Figure 12 (b), approaches the average busy time $1/\alpha$. As a result, the exponents α and β in Eq. (11) and (12) can be replaced with $\mu = \max(\alpha, \beta)$. By combining $\mathrm{E}[I_{\text{on}}]$ and $\mathrm{E}[I_{\text{off}}]$, we obtain the expected interference ratio T_{I} as follows:

$$\begin{aligned}T_{\text{I}} &= \frac{\mathrm{E}[I_{\text{on}}] + \mathrm{E}[I_{\text{off}}]}{T \cdot P_{\text{on}}} \\ &= \frac{\alpha}{\beta}[e^{-\mu T}\bar{P}_{\text{f}} + (1 - e^{-\mu T})\frac{\beta}{\alpha+\beta}]\end{aligned} \tag{13}$$

In Eq. (13), the range of T_{I} is determined as $\frac{P_{\text{off}}}{P_{\text{on}}}\bar{P}_{\text{f}} \leq T_{\text{I}} \leq P_{\text{off}}$. When the interference limit T_{P} is greater than P_{off}, this spectrum bands always satisfy the interference limit and can be used for CR transmission without any coordination of the sensing parameters. On the contrary, when the T_{P} is less than $\frac{P_{\mathrm{off}}}{P_{\mathrm{on}}}\bar{P}_{\mathrm{f}}$, this spectrum band cannot be used since the interference constraint is always violated.

This model shows another advantage in balancing the interference and the lost spectrum opportunity. Using the proposed interference model, the expected lost spectrum opportunity T_{L} can be obtained as follows:

$$T_{\mathrm{L}} = \frac{\beta}{\alpha}[e^{-\mu T}\bar{P}_{\mathrm{f}} + (1 - e^{-\mu T})\frac{\alpha}{\alpha + \beta}] \tag{14}$$

More details are given in Appendix A.

Since T_{I} and T_{L} have the duality characteristics of α and β, the interference and the lost spectrum opportunity can be balanced. From Eq. (14), we can see that the range of T_{L} is $\frac{P_{\mathrm{on}}}{P_{\mathrm{off}}}\bar{P}_{\mathrm{f}} \leq T_{\mathrm{L}} \leq P_{\mathrm{on}}$, which shows a similar trend to that of T_{I}. Only the primary user activity can determine the difference.

3.3.4 Sensing Parameter Optimization

In this section, based on the proposed MAP-based energy detection and interference model, we show how to solve the sensing parameter optimization problem defined in the beginning of this section.

3.3.4.1 Observation Time

To solve the optimization problem, we first specify the relation between the false alarm probability $\bar{P}_{\mathrm{f}}$ and the observation time t_{s}. Through the calculations given in Appendix B, t_{s} can be represented as follows:

$$t_{\mathrm{s}} = \frac{1}{W \cdot \gamma^2}[Q^{-1}(\bar{P}_{\mathrm{f}}) + (\gamma + 1)Q^{-1}(\frac{P_{\mathrm{off}}\bar{P}_{\mathrm{f}}}{P_{\mathrm{on}}})]^2 \tag{15}$$

where W is the bandwidth of the spectrum band and $\gamma = {\sigma_r}^2/{\sigma_{\mathrm{n}}}^2$ represents the signal-to-noise ratio (SNR).

Since this function is the sum of two different inverse-Q functions, it is obvious that this is a monotonically decreasing function.

3.3.4.2 Operating Region for Transmission Time

From Eqs. (2) and (13), the transmission time T has the following operating region:

$$\begin{aligned}\bar{P}_{\mathrm{f}} &< \frac{\frac{T_{\mathrm{P}} \cdot P_{\mathrm{on}}}{P_{\mathrm{off}}} - P_{\mathrm{on}}(1 - e^{-\mu T})}{e^{-\mu T}} \\ &= P_{\mathrm{on}} - P_{\mathrm{on}}(1 - \frac{T_{\mathrm{P}}}{P_{\mathrm{off}}})e^{\mu T} = \bar{P}_{\mathrm{f}}(T)\end{aligned} \tag{16}$$

where $\bar{P}_{\mathrm{f}}(T)$ is the boundary function of the operating region. Since T_{P} is less than P_{off}, as shown in Section 3.3.3, $\bar{P}_{\mathrm{f}}(T)$ is monotonically decreasing. In addition, $\bar{P}_{\mathrm{f}}$ is bounded by $min(0.5, 0.5 \cdot \frac{P_{\mathrm{on}}}{P_{\mathrm{off}}})$ since the false alarm and detection error probabilities are assumed to be the same. Furthermore, from Eq. (16), we can see that the maximum T is bounded by $-\frac{1}{\mu} \cdot log(1 - \frac{T_{\mathrm{P}}}{P_{\mathrm{off}}})$, which means that if T is greater than this value, this spectrum band cannot satisfy the interference constraint T_{P}, regardless of $\bar{P}_{\mathrm{f}}$.

Figure 13 shows the operating region given in Eq. (16) and the inverse function of Eq. (15), $\bar{P}_{\mathrm{f}}(t_{\mathrm{s}})$. The operating region, which is illustrated in gray in Figure 13, is the area of $\bar{P}_{\mathrm{f}}$ and T where the interference constraint T_{P} is always satisfied. The operating region and $\bar{P}_{\mathrm{f}}(t_{\mathrm{s}})$ are used in determining the optimal sensing parameters T^* and t_{s}^*, which will be explained in the next subsection.

3.3.4.3 Optimization Procedure

The optimization problem defined in the beginning of this section is not easy to be solved numerically since the objective function and the constraints are combined with the false alarm probability $\bar{P}_{\mathrm{f}}$. Instead, we introduce an iterative method to exploit $\bar{P}_{\mathrm{f}}(t_{\mathrm{s}})$, the inverse function of Eq. (15) and $\bar{P}_{\mathrm{f}}(T)$ given in Eq. (16).

In Figure 13, we show how to find the optimized parameters. As shown in Figure 13, T and t_{s} have the same false alarm probability $\bar{P}_{\mathrm{f}}$. Furthermore T, t_{s} and $\bar{P}_{\mathrm{f}}$ should be placed inside the operating region to satisfy the interference constraints. Thus, this optimization can be simplified to the problem to find an optimal false alarm

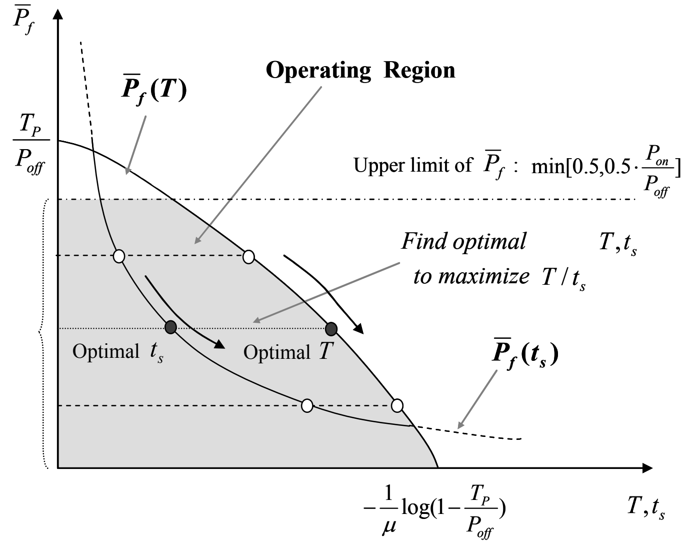


Figure 13: The operating region of optimal transmission and observation times.

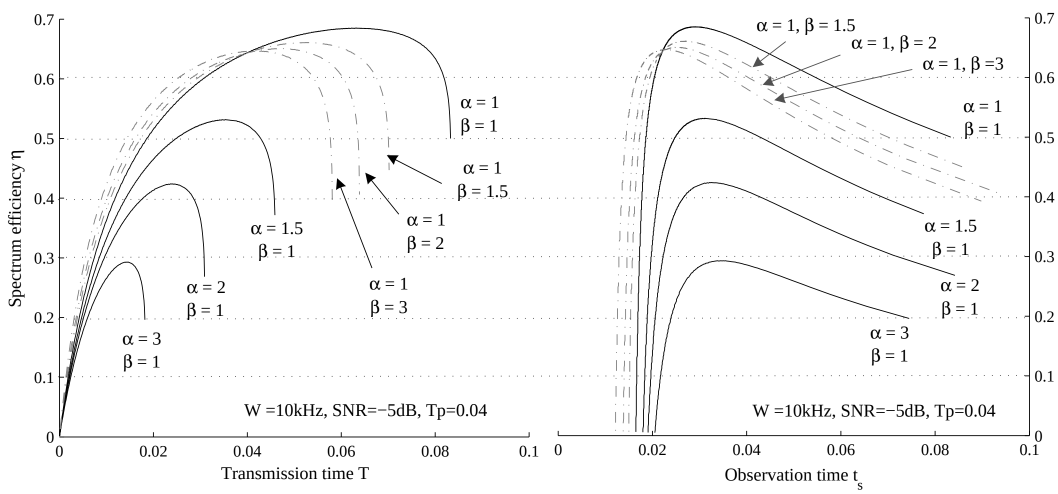


Figure 14: The relation between spectrum efficiency and sensing parameters (transmission and observation times).

probability $\bar{P}_{\rm f}$ to maximize sensing efficiency, which can be easily obtained through an iterative numerical method. In this method, first, $\bar{P}_{\rm f}$ is calculated according to the T using the boundary function $\bar{P}_{\rm f}(T)$. According to the $\bar{P}_{\rm f}$, $t_{\rm s}$ is obtained from Eq. (15), and then the spectrum efficiency is calculated using T and $t_{\rm s}$. As depicted in Figure 13, by searching all possible transmission times T within the operating region, we can obtain an optimal $\bar{P}_{\rm f}$ which provides a maximum sensing efficiency.

Figure 14 depicts the results of the numerical analysis on spectrum efficiency and sensing parameters where we can see that there exist optimal sensing parameters to maximize sensing efficiency. Furthermore, as shown in Fig 14, optimal sensing parameters and sensing efficiency are more sensitive to the changes of α than of β.

In this section, we proposed an MAP-energy detection and an analytical interference model for the periodic spectrum sensing. Then, we derived optimal observation and transmission times, which maximize sensing efficiency under the interference constraint. To extend this optimization method to multi-spectrum/multi-user network environment, additional functionalities need to be developed, which will be explained in the following sections.

3.4 Spectrum Selection and Scheduling for Spectrum Sensing on Multiple Spectrum Bands

In the previous section, we explained how to find the optimized parameters for single-band/single-user sensing. However, in reality, to mitigate the fluctuating nature of the opportunistic spectrum access, CR users are supposed to exploit multiple available spectrum bands showing different characteristics. To handle multiple spectrum bands, two different types of sensing strategies can be exploited: wideband sensing and sequential sensing. In *wideband sensing*, the sensing transceiver can sense multiple spectrum bands over a wide frequency range at a time. Although wideband sensing method requires only a single sensing transceiver, it uses identical observation and transmission times over multiple spectrum bands without considering their different characteristics, which cause the violation of interference limit. Furthermore, it requires a high-speed analog-to-digital (A/D) converter [8]. On the contrary, in *sequential sensing*, the sensing transceiver monitors only a single spectrum band at a time, which enables CR users to use sensing parameters adaptively to the characteristics of each spectrum band. However, CR users may not have enough transceivers to exploit all available spectrum bands, which introduces spectrum selection and scheduling problems in multi-spectrum CR networks

Here, we assume all CR users use sequential sensing. In the following subsections, we explain how to extend our proposed optimal sensing method to multiple spectrum bands.

3.4.1 Problem Definition

As explained in Section 3.3, multiple spectrum bands have different optimal observation and transmission times according to their characteristics. If CR users are required to exploit all available spectrum bands, the number of sensing transceivers can be expressed as $\sum_{i \in A} \frac{t^*_{\mathrm{s},i}}{T^*_i + t^*_{\mathrm{s},i}}$ where A is a set of all available spectrum bands and $t^*_{\mathrm{s},i}$ and T^*_i represent optimal observation and transmission times of spectrum band i. However, since CR users generally have a finite number of transceivers, it is not practical to monitor all available spectrum bands. Hence, instead of exhaustive sensing, selective sensing is more feasible in CR networks. To select the spectrum bands properly under the sensing resource constraint, we introduce a new notion, *opportunistic sensing capacity* as follows:

Definition 5: The *opportunistic sensing capacity* C^{op}_i represents the expected transmission capacity of spectrum band i that CR users can achieve, which can be defined as follows:

$$C^{\mathrm{op}}_i = \eta_i \cdot \rho_i \cdot W_i \cdot P_{\mathrm{off},i} \qquad (17)$$

where η_i, W_i, and $P_{\mathrm{off},i}$ represent sensing efficiency, the bandwidth, and the idle state probability of the spectrum band i. ρ_i is the spectral efficiency of the spectrum band i (bit/sec/Hz) depending on the modulation and channel coding schemes. $\rho_i \cdot W_i$ represents how much transmission rate this spectrum band can support. To reflect the dynamic nature of spectrum bands in CR networks, C^{op}_i also consider the spectrum efficiency and the probability of the idle state.

Another practical sensing problem in multi-spectrum networks is that each spectrum band has different optimized sensing cycles $T_i^* + t_{s,i}^*$. Once spectrum bands are selected, the sensing transceiver is required to be scheduled for spectrum sensing. However, heterogeneous sensing cycles of each spectrum cause the collision of the sensing operations, which degrades the transmission capacity in CR networks. Thus, a novel sensing scheduling method needs to be developed to reduce the collisions of the sensing schedules.

3.4.2 Spectrum Selection for Selective Sensing

Since the number of sensing transceivers is finite, CR users require a selective sensing method to exploit multiple available spectrum bands, which show different capacities according to the spectrum characteristics. To consider the dynamic and heterogenous nature of underlying spectrum bands in CR networks, we propose a spectrum selection method to maximize *opportunistic sensing capacity* of CR networks, which can be expressed as the following optimization problem:

$$\begin{aligned} &\text{Maximize:} \quad \sum_{i \in A} \eta_i \cdot \rho_i \cdot W_i \cdot P_{off,i} \cdot x_i \\ &\text{Subject to:} \quad \sum_{i \in A} \frac{t_{s,i}^*}{T_i^* + t_{s,i}^*} \cdot x_i \leq N_{\text{sen}} \end{aligned} \tag{18}$$

where A is a set of all available spectrum bands, N_{sen} represents the maximum number of transceivers for spectrum sensing, and $x_i \in \{0, 1\}$ represents the spectrum selection parameter. This optimization can be easily solved by the binary integer programming [62]. Once spectrum bands are selected, the transceiver is required to be scheduled for spectrum sensing, which is explained in the following subsection.

3.4.3 Sensing Scheduling for Multiple Spectrum Bands

The proposed spectrum selection method shows an ideal and theoretical sensing capacity bound of the sensing transceiver. However, in reality, it is impossible to assign multiple sensing tasks with different periods into one resource schedule without collision. If the sensing cycle is fixed over all multiple heterogeneous spectrum bands, sensing efficiency will be surely degraded. Thus, in this section, we propose a practical approach for sensing scheduling on multiple spectrum bands. While traditional scheduling methods in wireless networks have explored how multiple users can access the wireless channel considering fairness and channel throughput, the proposed scheduling is focusing on how the sensing transceiver is scheduled to sense multiple spectrum bands satisfying optimal sensing cycles of each spectrum. In this chapter, we assume the CR networks adopt a time-slotted sensing scheduling where a time slot is used as the minimum time unit of the observation time and the transmission time.

If multiple spectrum bands compete for the sensing slot at the same time, CR users determine one of the spectrum bands through the proposed sensing scheduling based on the *opportunity cost*. The *opportunity cost* is defined as the sum of the expected opportunistic sensing capacities of the spectrum bands to be blocked if one of the competing spectrum bands is selected. In the proposed method, the current time slot is assigned to one of the competing spectrum bands to minimize the opportunity cost, referred to as the *least-cost first-serve (LCFS)* scheduling algorithm. The following equation explains how to assign the sensing slot to the best spectrum band j^* in the LCFS scheduling.

$$j^* = \arg\min_{j \in B} \, (t^*_{\mathrm{s},j} \sum_{i \in B, i \neq j} \rho_i W_i P_{\mathrm{off},i} + \sum_{i \in B, i \neq j} t^{\mathrm{b}}_i \rho_i W_i P_{\mathrm{off},i}) \tag{19}$$

where B is a set of competing spectrum bands and t^{b}_i is the blocked time of the spectrum band i. ρ_i, W_i, and $P_{\mathrm{off},i}$ represent the spectral efficiency, the bandwidth, and the idle state probability of the spectrum band i, respectively. The first term represents the opportunity cost of spectrum band j. The second term represents the sum of the opportunistic capacities of the blocked spectrum bands during the past blocked time t_i^{b}. For the fair scheduling among competing spectrum bands, the proposed method considers not only the opportunity cost for the future sensing time but also the opportunistic capacity blocked in the past. Through these procedures, the LCFS algorithm assigns the current time slot to the spectrum band in such a way as to minimize the sum of the opportunity cost and the blocked opportunistic capacity of other spectrum bands.

The detailed procedure for sensing scheduling is as follows. When a sensing cycle starts, CR users check the state of the current time slot. If the current time slot is already occupied by the other spectrum band, all competing bands go to the *blocked period*. When the time slot is available, CR users assign the current time slot to one of the competing spectrum bands. The rest of the spectrum bands should block their sensing operations to the next available time slot. When the *observation period* ends after the observation time t_{s}, the spectrum band goes to the *transmission period* and the current time slot is available to the other spectrum bands.

3.5 Adaptive and Cooperative Spectrum Sensing in Multiuser Networks

The most important and unsolved issue in spectrum sensing is a receiver uncertainty problem [3]. With the local observation, CR users cannot avoid the interference to the primary receivers because of lack of location information. Generally, a cooperative sensing scheme method is known to be more effective in mitigating the receiver uncertainty problem. In this section, we extend our proposed optimal sensing method to the multi-user environment and propose an adaptive and cooperative sensing, especially focusing on the functionalities of the base-station.

3.5.1 Problem Definition

Assume CR networks have a base-station. CR users sense spectrum bands at each location and report the sensing results to the base-station periodically. Then, the base-station decides the availability of the spectrum bands inside its coverage and allocates the available spectrum bands to the users. These sensing data have a spatial correlation which can be used to enhance spectrum sensing accuracy through cooperation.

However, to exploit this cooperative gain, the base-station should consider the following issues. First, since the cooperative scheme can enhance the detection probability, the expected interference ratio is less than the originally estimated in the sensing parameter optimization, which means the optimal parameters are no longer valid. Second, the cooperation gain has the time-varying characteristic according to the number of users involved in the cooperation. Furthermore, the number of primary user activity regions will affect the cooperative gain. Considering all of the above issues, we propose an adaptive and cooperative sensing framework in the following subsections.

3.5.2 Availability Decision using Cooperative Gain

In traditional cooperative sensing, the spectrum band is decided to be available only if no primary user activity is detected out of all sensing data. Even if only one primary user activity is detected, CR users cannot use this spectrum band [52]. From this detection criterion, the cooperation gain of N sensing data is obtained by $\bar{P}_{\mathrm{d}}^{c} = 1 - (1 - \bar{P}_{\mathrm{d}})^N$ where $\bar{P}_{\mathrm{d}}^{\mathrm{c}}$ and $\bar{P}_{\mathrm{f}}^{\mathrm{c}}$ are the cooperative detection and false alarm probabilities, respectively. While this decision strategy surely increases the detection probability, it increases the lost spectrum opportunities as a result of the increase in cooperative false alarm probability, $\bar{P}_{\mathrm{f}}^{\mathrm{c}} = 1 - (1 - \bar{P}_{\mathrm{f}})^N$.

Thus, we define a new cooperative gain for the decision of spectrum availability. The number of detections follows the binomial distribution $\mathcal{B}(N, \bar{P}_{\mathrm{d}})$. Similarly, the number of false alarms also shows the binomial distribution $\mathcal{B}(N, \bar{P}_{\mathrm{f}})$. Thus, to determine the detection threshold N_{th} to balance between the detection error probability and the false alarm probability, we exploit the same strategy as explained in Section 3.3.2.

$$P_{\text{on}}(1 - P_{\text{bd}}(N_{\text{th}})) = P_{\text{off}} \cdot P_{\text{bf}}(N_{\text{th}}) \tag{20}$$

where P_{bd} is the binomial cumulative distribution function (CDF) of the number of detections, and P_{bf} is the binomial CDF of the number of false alarms.

To use this cooperative scheme, all CR users should be located in the same primary user activity region. In other words, the spatial correlation of primary user activities at each location affects the performance of the cooperative sensing significantly. If there are multiple primary user activities, the base-station should calculate cooperative detection probability of each region separately. Then, the cooperation gain is obtained as follows:

$$\bar{P}_{\text{d}}^{\text{Pc}} = 1 - \prod_{i=1}^{N_{\text{corr}}} (1 - \bar{P}_{\text{d},i}^{\text{c}}) \tag{21}$$

$$\bar{P}_{\text{f}}^{\text{c}} = 1 - \prod_{i=1}^{N_{\text{corr}}} (1 - \bar{P}_{\text{f},i}^{\text{c}}) \tag{22}$$

where N_{corr} is the number of the primary user activity regions in the CR network coverage. $\bar{P}_{\text{d},i}^{\text{c}}$ and $\bar{P}_{\text{f},i}^{\text{c}}$ represent the cooperative detection and false alarm probabilities of the primary user activity region i, respectively. In this case, only if none of the regions detects the primary signals, the spectrum is determined to be available, and hence the detection error probability and the false alarm probability are not the same any longer. For this reason, while the detection probability increases, the lost spectrum opportunity T_{L} increases owing to the increase in the false alarm probability, which shows the same pattern to the traditional cooperation approach explained in Section 3.5.2.

3.5.3 Sensing Parameter Adaptation

Through the proposed cooperative detection method explained above, both detection and false alarm probabilities can be improved as follows:

$$P_d^{\mathrm{c}} = P_{\mathrm{on}} \bar{P}_d^{\mathrm{c}} = P_{\mathrm{on}} \sum_{i=N_{\mathrm{th}}}^{N} \binom{N}{i} \bar{P}_{\mathrm{d}}{}^{i} (1 - \bar{P}_{\mathrm{d}})^{N-i} \tag{23}$$

$$P_f^{\mathrm{c}} = P_{\mathrm{off}} \bar{P}_f^{\mathrm{c}} = P_{\mathrm{off}} \sum_{i=N_{\mathrm{th}}}^{N} \binom{N}{i} \bar{P}_{\mathrm{f}}{}^{i} (1 - \bar{P}_{\mathrm{f}})^{N-i} \tag{24}$$

Since both detection and false alarm probabilities change, the optimal sensing parameters need to be re-optimized. However, the optimal observation time t_{s}^* is already considered for the false alarm probability of each user, which is used for calculating the cooperation gain. Hence, the cooperation gain only affects the transmission time T^*, which needs to be re-optimized using the Eq. (16). Usually the number of sensing data varies over time because of user mobility and transmission. Whenever it changes, the base-station re-optimizes the transmission time, which improves the transceiver utilization maintaining the same interference level as the non-cooperative sensing. Since the proposed method exploits the cooperation gain to reduce the sensing resources of the spectrum band, it enables CR users to have more spectrum access opportunities.

3.6 Performance Evaluation

In the previous sections, we developed the sensing parameter optimization scheme, spectrum selection, sensing scheduling, and the adaptive and cooperative sensing method. In this section, we present both analytical and simulation results on the performance of our proposed sensing framework.

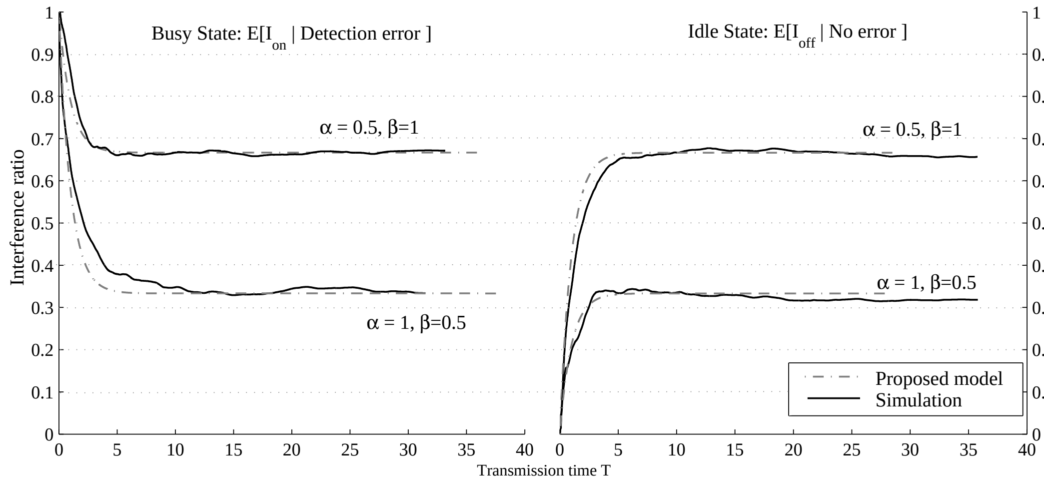


Figure 15: Comparison between the proposed interference model and simulation results.

3.6.1 Sensing Parameter Optimization in a Single Band

To evaluate the performance of the proposed optimal sensing algorithm explained in Section 3.3, we implement the primary traffic generator based on the ON-OFF Poisson arrival model and measure the expected interference ratio T_{I} on various sensing parameters.

First, in Figure 15, our proposed interference model, given in Section 3.3, is compared to the interference measurement through simulations. In Figure 15, we can see the proposed interference model is valid for both busy and idle states.

Based on the optimal sensing parameters obtained from Section 3.3, we simulate the periodic sensing procedure on the randomly generated primary user traffic. To demonstrate the optimality of the selected sensing parameters. we compare the optimal sensing parameters with two other non-optimal sensing parameter pairs selected from the operating region and the non-operating region, respectively. Figure 16 shows the moving average of the interference T_{I} measured in the simulations. While both optimal and non-optimal sensing parameters from the operating region satisfy the interference limit, optimal sensing parameters show a better sensing efficiency. In the

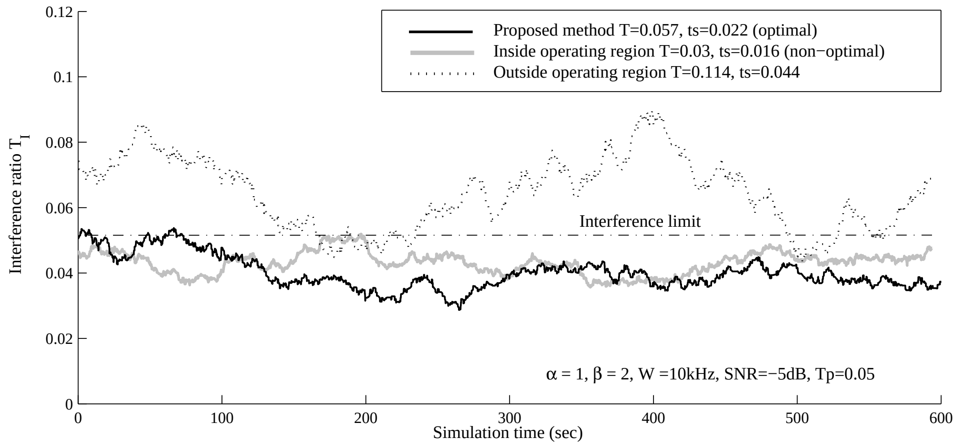


Figure 16: The simulation results of the proposed optimal sensing in a single band: interference T_{I}.

Table 1: Spectrum information for simulation.

Parameter	Low Opportunity									
	Low Activity					High Activity				
Spectrum #	1	2	3	4	5	6	7	8	9	10
α	0.2	0.3	0.4	0.3	0.8	1.5	2	1	1	2
β	0.4	0.6	0.5	0.9	1	4	5	2	5	3
SNR(dB)	-20	-15	-10	-5	0	-20	-10	-5	0	-15
BW(kHz)	250	100	70	40	10	250	100	70	40	10
T_{P}(%)	0.03	0.05	0.04	0.02	0.01	0.05	0.01	0.04	0.02	0.03

Parameter	High Opportunity									
	Low Activity					High Activity				
Spectrum #	11	12	13	14	15	16	17	18	19	20
α	0.2	0.8	0.7	1	0.3	4	3	2	3	5
β	0.1	0.1	0.2	0.7	0.2	1.5	2	1	3	2
SNR(dB)	-20	-10	-5	0	-15	-20	-10	-5	0	-15
BW(kHz)	250	100	70	40	10	250	70	100	40	10
T_{P}(%)	0.05	0.01	0.03	0.02	0.04	0.05	0.04	0.01	0.03	0.02

case of the sensing parameters obtained from the non-operating region, while sensing efficiency is the same as that of optimal parameters, they violate the interference constraint.

3.6.2 Resource Allocation on Multiple Spectrum Band

For simulations of spectrum sensing on multiple spectrum bands, we first define on scenario of the spectrum environments. According to the primary user activity and

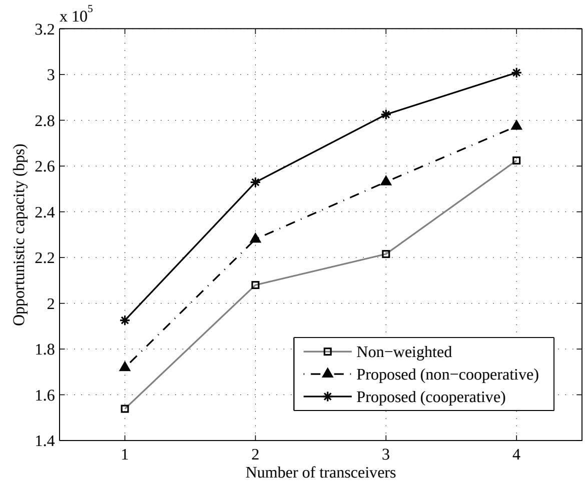


Figure 17: The opportunistic capacity of the proposed spectrum selection.

the portions of opportunities on the spectrum band, we classify the available spectrum bands in 4 classes: *high-opportunity/high-activity*, *high-opportunity/low-activity*, *low-opportunity/high-activity*, and *low-opportunity/low-activity*. High-opportunity represents the spectrum bands with $P_{\text{on}} < P_{\text{off}}$ and low-opportunity represents the spectrum bands with $P_{\text{on}} > P_{\text{off}}$. High-activity represents the spectrum with $\alpha > 1$ or $\beta > 1$, and low-activity represents the spectrum with $\alpha < 1$ and $\beta < 1$. According to this classification, we generate the spectrum information as explained in Table 1. In this simulation, we assume that bandwidth efficiency $\rho = 1$ over all spectrum bands.

First, in Figure 17, the proposed spectrum selection method is compared to the non-weighted method, where spectrum bands are determined to maximize the number of selected spectrum bands. In this simulation, our selection algorithm shows more capacity than the non-weighted methods, since our method considers the potential opportunistic capacities as well as traffic activities.

For the spectrum bands chosen by our proposed selection method, we evaluate the performance of the proposed sensing scheduling algorithm and compare it with the ideal scheduling and with First Come First Serve (FCFS) scheduling. Here, we assume that the CR user has a single transceiver. The ideal scheduling is assumed

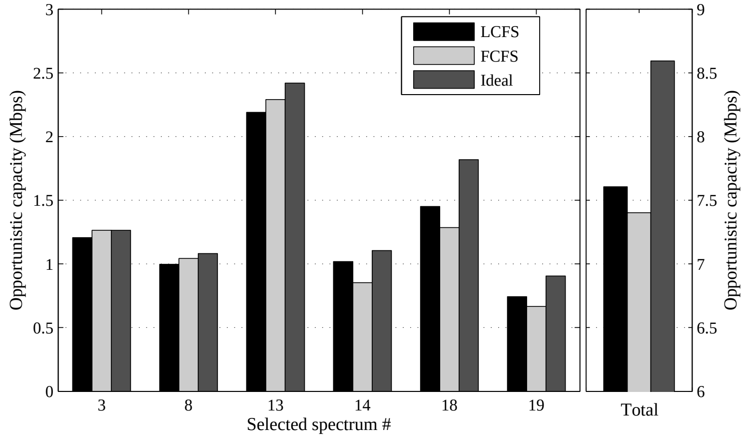


Figure 18: The performance of the proposed sensing scheduling.

to achieve the optimal sensing efficiency given in Section 3.3. In FCFS scheduling, the time slot is assigned to the spectrum band with the longest blocked time. In Figure 18, we show the allocated capacity of each spectrum band. As shown in Figure 18, our LCFS scheduling provides higher capacity in total than that of the FCFS, since our LCFS method assigns the sensing slot to minimize the opportunity cost, as explained in Section 3.4.3. Although high capacity is emphasized in the LCFS method, the fairness in allocating sensing resources is maintained by exploiting the blocked capacities in the past, as shown in Figure 18.

3.6.3 Cooperative Sensing in Multi-User Networks

To investigate how the proposed optimal sensing algorithm works in the cooperative sensing, we simulate the adaptive and cooperative sensing method in the multi-user environment. First, we evaluate the proposed cooperative sensing gain in terms of optimal transmission time. In Figure 19, according to the number of cooperating users, we recalculate optimal transmission times of each spectrum band (#2, #3, #4, #8, #11, #15, #17) given in Table 1. As depicted in Figure 19, the cooperation gain increases the optimal transmission time of the spectrum bands, which improves sensing efficiency. As shown in Figure 19, as the number of users increases, T^*

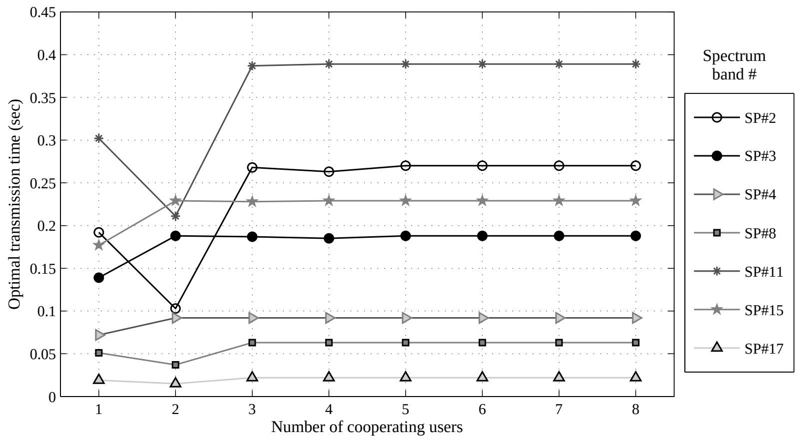


Figure 19: The optimal transmission time in the proposed cooperative sensing.

increases and is finally converged to $-\frac{1}{\mu}log(1-\frac{T_{\mathrm{P}}}{P_{\mathrm{off}}})$. Some of the spectrum bands show the degradation of the cooperation gain at the small number of users depending on the primary user activities. With the small number of cooperating users, our availability decision method given in Eq. (20) may increase both detection error and false alarm probabilities. In the case of small number of users, therefore, the traditional approach given in Section 3.5.2 is recommended.

To evaluate the performance of the proposed cooperative sensing scheme, given in Section 3.5.3, we use the same simulation explained in Section 3.6.1. Here, we assume there are 4 cooperating users in the same primary user activity region. In Figures 20 and 21, we show the T_{I} and T_{L} measured through the simulation based on the re-optimized sensing parameters. Although the transmission time increases due to the cooperation gain, our adaptive and cooperative method maintains the interference limit. However, same sensing parameters without the cooperation lead to the violation of the interference limit. We also compare our proposed algorithm with the traditional cooperation approach, given in Section 3.5.2. As shown in Figures 20 and 21, while the traditional approach satisfies the interference constraint with better spectrum efficiency, it shows much more lost spectrum opportunities due

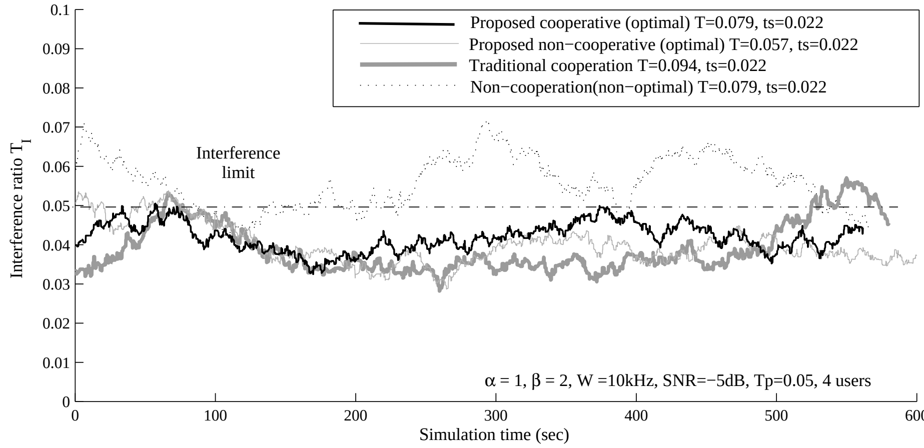


Figure 20: The simulation results of the cooperative sensing method: interference T_I.

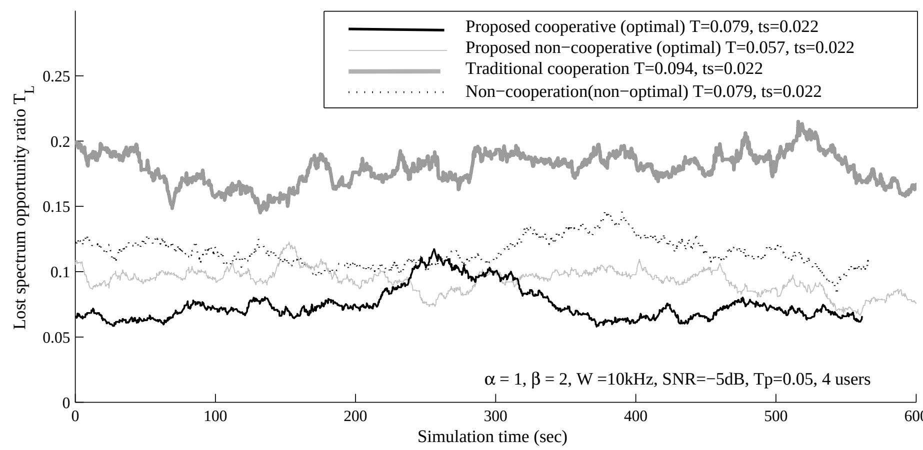


Figure 21: The simulation results of the cooperative sensing method: lost opportunity T_L.

to the increase in the false alarm probability. In Figure 17, we show how the proposed adaptive and cooperative sensing method can improve the sensing capacity by simulating the proposed spectrum selection method, given in Section 3.4.2. From the Figure 17, we can see that the proposed cooperative sensing can improve the total sensing capacity since it increases sensing efficiency of each spectrum band, i.e., the proposed cooperative sensing method enables the sensing transceiver to sense more spectrum bands without violation of the interference constraints.

CHAPTER IV

QOS-AWARE SPECTRUM DECISION FRAMEWORK FOR COGNITIVE RADIO NETWORKS

4.1 Introduction

Although spectrum sensing enables CR users to exploit spectrum opportunities effectively, the heterogenous spectrum environment introduces a new critical issue for CR networks. Generally, CR networks have multiple available spectrum bands over a wide frequency range that show different channel characteristics and need to support applications with diverse service requirements. Therefore, once available spectrum bands are identified through spectrum sensing, CR networks need to select the proper spectrum bands according to the application requirements. This process is referred to as *spectrum decision* which constitutes an important but yet unexplored topic in CR networks. To decide on spectrum bands properly, CR networks need to consider the following issues:

- All available spectrum bands show different characteristics in the CR network. To select the proper spectrum, the CR network needs to characterize the available spectrum bands by considering current radio conditions as well as the primary user (PU) activity.

- The CR network cannot guarantee a reliable and permanent communication channel because of PU activities. Thus, the CR network needs to provide a dynamic decision framework to consider all possible events to prevent reliable transmissions by closely interacting with other CR capabilities such as spectrum sensing and spectrum sharing.

- According to the PU activities, the total capacity in CR networks varies over time, which makes it more difficult to decide on spectrum bands while maintaining the service quality of other CR users. Thus, the CR network should perform the spectrum decision adaptively dependent on time-varying spectrum resources.

Thus, a unified framework for spectrum decision needs to be developed to consider all possible CR network environments and to define the inter-operations of other network functionalities.

In this chapter, an adaptive spectrum decision framework is proposed with the consideration of all decision events and application types [46]. First, a novel capacity model is developed to describe unique characteristics in CR networks by considering PU activity as well as sensing capability. Based on this, two different decision schemes are introduced. To satisfy the delay constraints in real-time applications, we propose a minimum variance-based spectrum decision (MVSD) scheme to select spectrum band to minimize the capacity variation. For the best-effort application, we propose a maximum capacity-based spectrum decision (MCSD) scheme to maximize the total network capacity. Both decision schemes are controlled by a proposed resource manager based on the current network condition.

The remainder of the chapter is organized as follows. In Section 4.2, we propose a novel framework for spectrum decision. In Section 4.3, we present a spectrum capacity model used in this chapter. Spectrum decision methods for real-time application and best effort application are proposed in Sections 4.4 and 4.5, respectively. Then, we develop dynamic resource management for the CR network in Section 4.6. Performance evaluation and simulation results are presented in Section 4.7.

4.2 A Framework for Spectrum Decision in Cognitive Radio Networks

4.2.1 System Model

The components of the CR network architecture, we consider here, can be classified in two groups as the *primary network* and the *cognitive radio (CR) network*. The *primary network* (or licensed network) is referred to as an existing network, where the *primary users* have a license to operate in a certain spectrum band. The *CR network* (or unlicensed network) does not have license to operate in a desired band. Here, we consider an infrastructure-based CR network which has a centralized network entity such as a base-station. The base-station exerts control over all *CR users* within its transmission range. CR users perform the observations and analysis and feed them to the central base-station which decides on spectrum availability and spectrum allocation. Each CR user has multiple software-defined radio (SDR) transceivers to exploit multiple spectrum bands over wide frequency ranges by adjusting the operating frequency through software operations. Here we assume a frequency division duplex (FDD) wireless system where uplink and downlink channels are separated. Thus, the proposed decision scheme can be applied to each link independently.

When primary users appear in the spectrum band, CR users need to move to a new available spectrum band, which may cause a temporary communication break. Moreover, CR users may not be able to detect any single spectrum band to meet the QoS requirements of users. To solve this problem, we assume that multiple non-contiguous spectrum bands can be simultaneously used for the transmission in the CR network. This method can create a signal that is not only capable of high data throughput, but is also immune to the the PU activity. Even if a primary user appears in one of the current spectrum bands, the rest of the spectrum bands will maintain current transmissions [3].

Another important architectural issue is how to establish a control channel in CR networks . The control channel plays an important role in exchanging information regarding sensing and resource allocation. Several methods are presented in [9], one of which is assumed to be used as the common control channel in our proposed method.

4.2.2 Framework Overview

Based on the system model above, we develop a novel framework for spectrum decision. Since CR users can have multiple available spectrum bands, the CR network requires capabilities to decide the best (set of) available band(s) among them according to the service requirements of the applications, referred to as *spectrum decision*. Spectrum decision is an event-based functionality, i.e., the CR network needs to decide on the spectrum bands in the case of following events:

- *CR user appearance:* When a new CR user appears in the CR network, it needs to be assigned to new spectrum bands for its transmission.
- *Primary user appearance:* When a primary user appears in the spectrum band, CR users should move to the new spectrum bands.
- *Channel quality degradation:* When channel condition becomes worse, CR users want to switch to better spectrum band.

To consider all decision events effectively, the CR network necessitates a unified framework for spectrum decision. Figure 22 shows the proposed framework for spectrum decision. A detailed description of the framework is as follows:

By considering current spectrum conditions, a *resource manager* determines if the CR network accepts a new incoming CR user or not. If a new CR user is allowed to transmit, it is assigned to the proper spectrum bands through *spectrum decision*. Since there may be the multiple CR users trying to share the same spectrum, *spectrum sharing* coordinates those multiple accesses to prevent the collisions and accordingly achieve the maximum capacity. In the *event detection*, the current spectrum bands

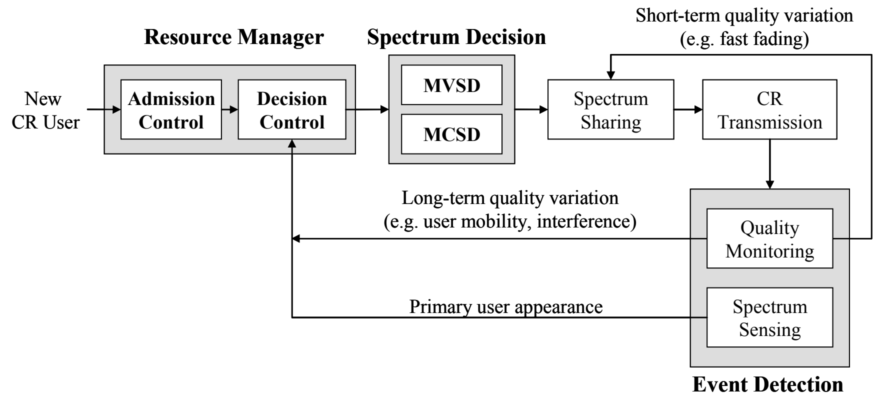


Figure 22: Spectrum decision framework for cognitive radio networks.

and users connections are monitored to detect decision events. The event detection consists of two main tasks: *spectrum sensing* and *quality monitoring*. When events are detected, the CR network reconfigures the resource allocation to maintain the service quality. In case of short-term channel variations such as fast fading, the CR network re-allocates the resources within the spectrum band through spectrum sharing. If a primary user is detected or the current spectrum band cannot provide the predetermined service quality any longer over a long-term period, the CR network switches the spectrum through the resource manager and the spectrum decision. All functionalities of the decision framework are performed in the base-station. CR users perform only the event detection. Based on the information gathering from CR users, the base-station decides on the spectrum availability and performs the spectrum decision as explained above.

The objective of the spectrum decision is similar to that of spectrum sharing in the sense that spectrum decision performs resource allocation based on service requirements. Most of the recent research in spectrum sharing have explored QoS issues [65] [67]. Despite this similarity, the spectrum decision has unique features to be distinguished from spectrum sharing. Generally, spectrum sharing is considered as a short-term operation, such as a packet-based or a time-slot based scheduling. On the contrary, spectrum decision is a connection-based and event-based operation.

Hence, compared to the spectrum sharing, the spectrum decision considers longer-term channel characteristics. In addition, although spectrum sharing is usually an intra-spectrum operation where all operations are performed within the spectrum band, spectrum decision is an inter-spectrum operation. Since available spectrums are distributed over a wide frequency range, the inter-spectrum operation inevitably introduces an additional switching delay leading to service quality degradation. Thus, it is not desirable to extend spectrum sharing designed to adapt to the fast time-varying channel to the long-term inter-spectrum operation. Here our design objective of the spectrum decision framework is to decouple all inter-spectrum functionalities totally from spectrum sharing.

Consequently, the proposed spectrum decision framework provides a hierarchical QoS guaranteeing scheme: spectrum sharing to allocate the channel and transmission power for short-term service qualities and spectrum decision to determine the best spectrum for maintaining service quality over a long term period. In the proposed framework, any conventional medium access control (MAC) protocol can be used for spectrum sharing functionality. Thus, in this chapter, we mainly focus on the decision functionalities: *spectrum decision* and *resource management*. Spectrum sharing and event detection functionalities are out of the scope in this chapter.

4.2.3 Spectrum Decision Functionalities

In the proposed framework given in Figure 22, we consider two types of applications: *real-time* and *best-effort* (In this chapter, the terms "application" and "user" are interchangeably used.). According to the application types, the proposed spectrum decision can be classified into a *minimum variance-based spectrum decision (MVSD)* method for real-time applications, and a *maximum capacity-based spectrum decision (MCSD)* for best-effort applications.

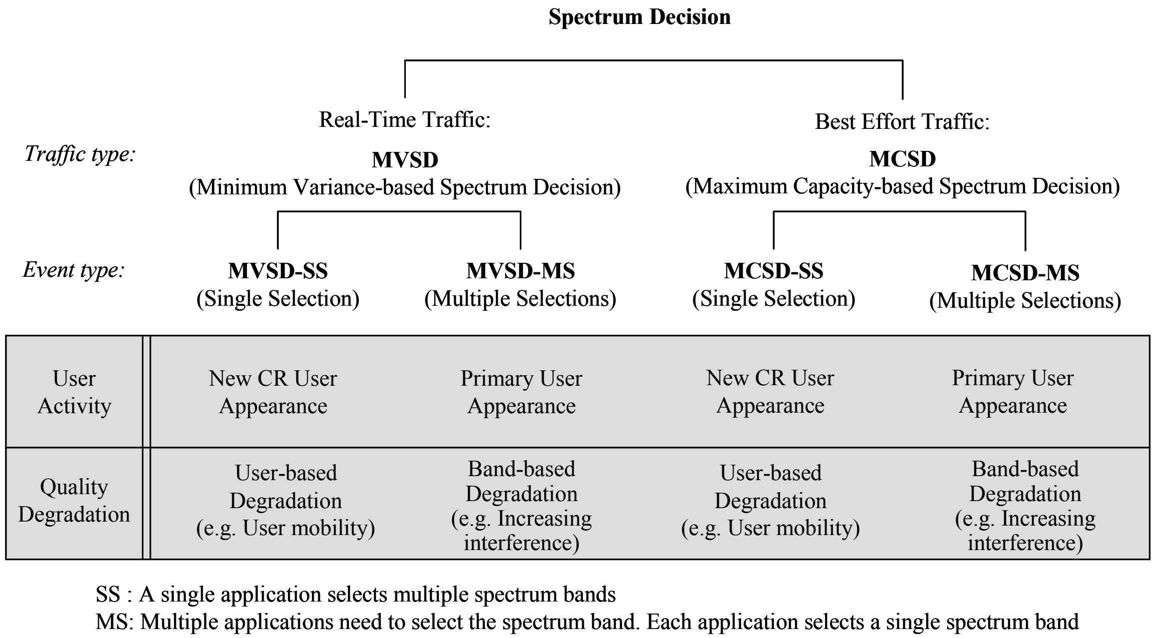


Figure 23: Classification of the proposed spectrum decision methods.

Decision events mainly occur owing to either user activities or quality degradations. When primary user appears in the spectrum band or a new CR user begins to transmit, spectrum decision needs to be triggered. Moreover, the quality degradation of either an entire spectrum band, (e,g, increase in interference) or a specific user connection (e.g. moving far from the base-station) can also trigger spectrum decision. If a CR user exploits multiple spectrum bands, the spectrum decision method becomes more complicated according to the event types. When a new CR user appears or the service quality of a CR user becomes worse, multiple spectrum bands need to be determined for a single user at a time, called *single selection (SS)*. On the other hand, when a primary user appears or the quality of the entire spectrum band becomes worse, multiple CR users residing in that spectrum band lose one of their current spectrum bands, which requires multiple spectrum decisions for each CR user, called *multiple selections (MS)*.

As shown in Figure 23, according to the traffic and event types, spectrum decision can be classified into four categories: MVSD-SS, MVSD-MS, MCSD-SS, and MCSD-MS, which are proposed in Sections 4.4.1, 4.4.2, 4.5.1, and 4.5.2, respectively.

4.3 Spectrum Characterization

To determine spectrum properly, it is important to identify the characteristics of each spectrum, which is mainly influenced by both channel condition and PU activity. To this end, in this section, we define the PU activity, and accordingly propose a novel CR capacity model.

4.3.1 Primary User Activity

For an efficient spectrum utilization, the CR network needs to be aware of the PU activities of each spectrum band, which are defined as the traffic statistics of the primary network, Since PU activity is closely related to the performance of the CR network, the estimation of this activity is a very crucial issue in spectrum decision. The PU activity can be modeled as exponentially distributed inter-arrivals [66]. In this model, the PU activity in spectrum i is defined as a two state birth-death process with death rate α_i and birth rate β_i. An ON (busy) state represents the period used by primary users and an OFF (idle) state represents the unused period [15] [44].

4.3.2 Cognitive Radio Capacity Model

In the CR network, the available spectrum bands are not contiguous and may be distributed over a wide frequency range with different bandwidth. Here, we assume the CR network has multiple orthogonal non-interfering spectrum bands. For more flexible manipulation of heterogenous spectrum bands, we employ an orthogonal frequency division multiplexing (OFDM) as the physical layer technology. In this chapter, we assume each spectrum band i has a different bandwidth B_i Hz, consisting of multiple sub-carriers. Each sub-carrier can be assigned to different CR users. Moreover, each user can be allocated to the different number of sub-carriers in every time slot to control the data rate and error probability individually for each user. If a user k can be assigned to all sub-carriers in spectrum i with bandwidth B_i, the channel capacity in OFDM can be obtained as follows:

$$r_i(k) = \int_0^{B_i} \log_2(1 + \frac{|H_i^k(f)|^2}{N_i^k(f) + I_i^k(f)} P_i^k(f))df \tag{25}$$

where $H_i^k(f)$, $P_i^k(f)$, $N_i^k(f)$, and $I_i^k(f)$ denote the channel frequency response, the transmission power spectral density, the noise power spectral density, and interference corresponding to a user k at a spectrum band i, respectively.

Usually, each sub-carrier has a different channel gain and a noise level which are time-varying. However, when we consider the long-term spectrum characteristics, both fast and frequency selective fading effects are mitigated, and hence we can say $H_k(f)/(N_k(f) + I_k(f))$ in the same spectrum band is identical over a long-term period. If $P_i^k(f)$ is also identical in frequency, a normalized channel capacity $c_i(k)$ (bits/sec/Hz) of spectrum band i can be expressed as $c_i(k) = r_i(k)/B_i$.

However, in CR networks, each spectrum i cannot provide its original capacity $c_i(k)$. First, CR users cannot have a reliable spectrum permanently and need to move from one spectrum to another according to the PU activity, which introduces the so-called *spectrum switching delay*. During the switching time, the transmission of CR user is temporarily disconnected, which causes an adverse effect on the channel capacity. Here, the spectrum switching delay includes times for the spectrum decision process in the base-station, signaling for the new channel establishment, and RF front-end reconfiguration. IEEE 802.22 Wireless Regional Area Network (WRAN) standard requires the switching delay is less than 2 sec [32]. Also the conventional mobile broadcasting system, for example, Qualcomm's MediaFLO, shows an average physical layer channel switching delay up to 1.5 sec [12]. Depending on the development of the hardware technology, we believe that it will be much shorter but still be a significant factor to influence the network performance. Furthermore, CR users are not allowed to transmit during sensing operations, leading to the periodic transmissions with the sensing efficiency η_i [44].

These unique features in CR networks shows a significant influence on the spectrum capacity $\mathbf{C_i(k)}$. To describe all these stochastic activities, we define a new capacity notion, the so-called *CR capacity* $C_i^{\mathrm{CR}}(k)$, which is defined as the expected normalized capacity of spectrum i at user k as follows:

$$C_i^{\mathrm{CR}}(k) = E[\mathbf{C_i(k)}] = \frac{T_i^{\mathrm{off}}}{T_i^{\mathrm{off}} + \tau} \cdot \eta_i \cdot c_i(k) \tag{26}$$

where τ represents the spectrum switching delay, and T_i^{off} is the expected transmission time without switching in spectrum i. Since CR users face to the spectrum switching after the idle period, the first term in Eq. (26) represents the transmission efficiency when CR users occupy the spectrum i.

If we consider perfect sensing, i.e., both false alarm and detection error probabilities are are zero, T_i^{off} is obtained by $1/\beta_i$, which is the average idle period based on the ON-OFF model in Section 4.2. On the contrary, in case of imperfect sensing, we should account for the influence of sensing capability. Let Δt be a sensing period. Then, the average number of sensing slots in the idle period n_{s} is $\lceil 1/\beta_i/\Delta t \rceil$. From this, the expected transmission time can be obtained as follows:

$$\begin{aligned} T_i^{\mathrm{off}} &= \Delta t \cdot \sum_{k=1}^{n_{\mathrm{s}}-1} k \cdot (1 - P_i^{\mathrm{f}})^k \cdot P_i^{\mathrm{f}} + \frac{1}{\beta_i} \cdot (1 - P_i^{\mathrm{f}})^{n_{\mathrm{s}}} \\ &= \Delta t \cdot [\frac{(1 - P_i^{\mathrm{f}})(1 - (1 - P_i^{\mathrm{f}})^{n_{\mathrm{s}}-1})}{P_i^{\mathrm{f}}} \\ &\quad - (n_{\mathrm{s}} - 1) \cdot (1 - P_i^{\mathrm{f}})^{n_{\mathrm{s}}}] + \frac{1}{\beta_i}(1 - P_i^{\mathrm{f}})^{n_{\mathrm{s}}} \end{aligned} \tag{27}$$

where P_i^{f} represents a false alarm probability of spectrum i at each sensing slot. Here T_i^{off} can be expressed as the sum of the expected duration that the false alarm is first detected after k^{th} sensing slot. In this chapter, we consider a cooperative sensing scheme based on decision fusion, where its detection error probability converges to 0 as the number of users increases [49]. Thus, the detection error probability can be ignored in estimating CR capacity.

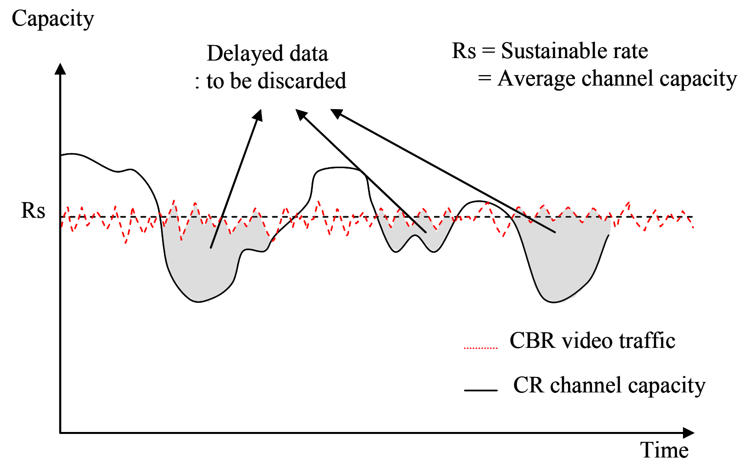


Figure 24: Data loss resulting from channel capacity fluctuation in real-time applications.

4.4 Spectrum Decision for Real-time Applications

Real-time applications are sensitive to delay and jitter. Moreover, they require a reliable channel to support a sustainable rate during the session time. Thus, real-time applications have strict constraints on the delay bound and the sustainable rate. Generally real-time applications drop the packets not arrived within the delay bound, leading to packet losses. Figure 24 illustrates the relationship between the channel capacity and data losses in real-time video transmission, where the dotted and solid lines represent the constant bit rate (CBR) video traffic and the CR channel capacity respectively. Even though the network can support the sustainable rate R_s on the average, because of the variation of channel capacity, packets can be delayed and finally discarded in the receiver, as shown in Figure 24.

Unlike conventional wireless networks, the CR network has a unique delay factor. Because of the low priority in spectrum access, the CR users cannot have a reliable communication channel for a long time. This variable channel characteristic prevents the real-time application from maintaining its sustainable rate, leading to delay and jitter. Moreover, when CR users either sense or switch the spectrum, they lose the connection temporarily, which also causes delay in real-time applications. To observe

Table 2: Symbols used for the analytical modeling in spectrum decision.

Notations	Descriptions
N	number of transceivers in CR users
$c_i(k)$	normalized capacity of spectrum i at user k
$C_i^{\mathrm{CR}}(k)$	normalized CR capacity of spectrum i at user k
α_i, β_i	primary user departure and arrival rates in spectrum i, respectively
T_i^{off}	expected transmission duration without switching at spectrum i
τ	spectrum switching delay
η_i	sensing efficiency
$R_{\mathrm{s}}(k)$	sustainable rate at user k
$P_{\mathrm{loss}}^{\mathrm{th}}$	target data loss rate
B_i	total bandwidth of spectrum i
W_i	currently available (idle) bandwidth of spectrum i
W_{av}	total bandwidth currently available in the network
W_{R}	total bandwidth currently used by real-time users,
W_{req}	expected bandwidth required for spectrum decision
$W_{\min}$	minimum bandwidth for QoS guarantee in the network
ϵ, π, ρ	operational parameters for overload, outage, and balance threshold

the effect of delay uniquely shown in CR networks, we assume that the buffering scheme is optimized to absorb the delay factors in the conventional wireless networks such as application layer, link layer, and transmission delays. Then, the additional delay factors introduced by the CR network can directly lead to data losses. For this reason, in this chapter, we use the data loss rate to evaluate the service quality of real-time applications. Also real-time applications are assumed to have a set of discrete sustainable rates and to adjust their rates through the negotiation flexibly.

In the chapter, we introduce a spectrum decision method called a minimum variance-based spectrum decision (MVSD). According to the decision events, as explained in Section 4.2.3, the proposed spectrum decision for real-time application can be classified into an MVSD - single selection (SS) and an MVSD - multiple selections (MS). For ease of representation, the important notations used in the subsequent discussion are summarized in Table 2.

4.4.1 Minimum Variance-based Spectrum Decision - Single Selection (MVSD-SS)

Real-time applications need to have more reliable and time-invariant communication channels to satisfy strict service requirements, such as delay constraints and sustainable rates. However, how to maximize the total network capacity is still a crucial problem. To address these issues together, it is necessary to guarantee the service quality of real-time applications with the minimum spectrum resources. Thus, the spectrum decision problem can be formulated as an optimization problem to minimize the bandwidth utilization subject to the constraint of the the sustainable rate, the data loss rate, and the number of transceivers. However, this problem is mixed with the discrete optimization for spectrum selection and the continuous optimization for bandwidth allocation, which is difficult to solve. Instead, we introduce a three stage spectrum decision method as follows:

4.4.1.1 Step 1: Spectrum Selection

From the view of the date loss rate caused by delay, the CR network prefers the spectrum bands with less PU activities. On the other hand, for the network capacity, the channel quality needs to be considered in spectrum decision. Thus, to maintain service quality and achieve the maximum capacity, a CR user k selects the spectrum bands according to the following linear integer optimization:

$$\textit{Maximize:} \sum_{i \in \mathcal{A}} \frac{C_i^{\mathrm{CR}}(k)}{\beta_i} x_i \tag{28}$$

$$\textit{subject to:} \sum_{i \in \mathcal{A}} x_i = N \tag{29}$$

$$C_i^{CR}(k) \cdot W_i \cdot x_i \geq \frac{R_{\mathrm{S}}(k)}{N} \quad (\forall i \in \mathcal{A}) \tag{30}$$

where W_i is the currently available bandwidth of spectrum i which is equal or less than the total bandwidth B_i, $\mathcal{A}$ is the set of the currently available spectrum bands, and $x_i \in \{0, 1\}$ represents the spectrum selection parameter that equals 1 if the spectrum i is selected in the binary integer optimization.

This optimization considers both the PU activity β_i and the CR capacity $C_i^{\mathrm{CR}}(k)$ simultaneously, as shown in Eq. (28). The number of the selected bands is restricted to the number of transceivers N as given in Eq. (29). The last constraint on the sustainable rate $R_{\mathrm{S}}(k)$ as given in Eq. (30) ensures that the selected spectrum bands have enough bandwidth for resource allocation which will be explained in the following subsection.

Since real-time applications usually require much stricter service requirements than the best-effort applications, they have a higher priority for resource allocation. Thus, the available bandwidth W_i includes the portions currently occupied by best-effort applications as well as the unused portion of the spectrum.

4.4.1.2 Step 2: Resource Allocation

Here, the CR network determines the bandwidth, i.e., a set of sub-carriers, of the selected spectrum bands to meet the constraints on both the sustainable rate $R_{\mathrm{S}}(k)$ and the target data loss rate $P_{\mathrm{loss}}^{\mathrm{th}}$. To allocate the bandwidth properly, first we derive the total capacity $\mathbf{R_T(k)}$ and the data loss rate $P_{\mathrm{loss}}(k)$ of a user k. When the bandwidth $w_i(k)$ is allocated to the selected spectrum i for a user k, the expected total capacity can be obtained as follows:

$$E[\mathbf{R_T(k)}] = \sum_{i \in \mathcal{S}} C_i^{\mathrm{CR}}(k) \cdot w_i(k) \tag{31}$$

where $\mathcal{S}$ is the set of the selected spectrum bands. To satisfy the service requirement on the sustainable rate, $E[\mathbf{R_T(k)}]$ should be equal to $R_{\mathrm{S}}(k)$.

Since the data loss rate $P_{\mathrm{loss}}(k)$, unlike the total capacity, is expressed in a complicated form, as derived in Appendix C, it cannot be easily used for the optimization. However, since the variance of the total capacity is proportional to the data loss rate, as shown in Appendix D, we can use the following the variance of the total capacity for the resource allocation, instead of the data loss rate.

$$Var[\mathbf{R_T(k)}] = \sum_{i \in \mathcal{S}} \frac{T_i^{\mathrm{off}} \eta_i \cdot (T_i^{\mathrm{off}} + \tau - T_i^{\mathrm{off}} \eta_i)}{(T_i^{\mathrm{off}} + \tau)^2} \cdot c_i(k)^2 \cdot w_i(k)^2 \tag{32}$$

where $\mathcal{S}$ is the set of the selected spectrum bands. $c_i(k)$ and $w_i(k)$ are the normalized capacity and the bandwidth of the spectrum band i for a user k, respectively. τ and

η_i are the spectrum switching delay and the transmission efficiency of the spectrum band i, respectively.

Based on the capacity variance derived in Eq. (32), the CR network determines the optimal bandwidth $w_i(k)$ of the selected spectrum bands to minimize the variance of the total capacity as follows:

$$\textit{Minimize: } Var[\mathbf{R_T(k)}] \tag{33}$$

$$\textit{subject to: } \sum_{i=1}^{M} C_i^{\text{CR}}(k) \cdot w_i(k) = R_{\text{S}}(k) \tag{34}$$

$$w_i(k) < W_i \quad (\forall i \in \mathcal{S}) \tag{35}$$

Eq. (34) and (35) represent the constraints on the sustainable rate and the available bandwidth, respectively.

By the Lagrange multiplier method, the optimal bandwidth $w_i(k)$ can be obtained as follows:

$$w_i(k) = \frac{R_{\text{s}}(k) \cdot (T_i^{\text{off}} + \tau)}{c_i(k) \cdot \eta_i (T_i^{\text{off}} + \tau - T_i^{\text{off}} \eta_i) \cdot \sum_{i \in \mathcal{S}} \frac{T_i^{\text{off}}}{T_i^{\text{off}} + \tau - T_i^{\text{off}} \eta_i}} \tag{36}$$

4.4.1.3 Step 3: QoS Checkup

This optimization is based on the minimum variance, which guarantees the minimum data loss rate but may not satisfy the target loss rate $P_{\text{loss}}^{\text{th}}$. If the data loss rate $P_{\text{loss}}(k)$ given in Eq. (109) is still higher than $P_{\text{loss}}^{\text{th}}$ after this optimization, we need to perform one of the following approaches to satisfy the target loss rate:

- *Aggressive approach:* By sacrificing bandwidth efficiency, the CR network tries to find the proper spectrum bands to meet the service requirements. To achieve this, the selected band having the highest PU activity needs to be replaced by the one with the highest $C_i^{\text{CR}}(k)/\beta_i$ among the unselected bands which have a lower PU activity than the original one. If CR network cannot find the proper spectrum band in the aggressive approach, it switches the decision method to the conservative approach as explained below.

- *Conservative approach:* In this chapter, real-time applications are assumed to support multiple sustainable rates and to adjust their rates adaptively. Thus, in this approach, instead of increasing the bandwidth to satisfy the current service requirements, the CR network reduces the current sustainable rate to the one step lower rate through the renegotiation of service quality and repeats the MVSD-SS while maintaining the bandwidth efficiency.

These aggressive and conservative approaches are applied in spectrum decision combined with resource management, which will be explained in Section 4.6.

4.4.2 Minimum Variance based Spectrum Decision - Multiple Selections (MVSD-MS)

MVSD-MS is performed when a primary user appears in the spectrum band or the channel quality of the entire spectrum band becomes worse because of the increase in interference. In these cases, all CR users in that spectrum band lose one of their connections.

Since multiple users lose their spectrum at the same time, first, they need to determine the order of spectrum decision. In MVSD-MS, CR user k with highest loss ratio $R_{\text{lost}}(k)/R_{\text{S}}(k)$ is selected as the first user to be re-assigned, where $R_{\text{lost}}(k)$ is the lost capacity of user k owing to the spectrum switching. According to the order of the loss ratio from the highest to the lowest, CR users perform spectrum selection and resource allocation, explained in Section 4.4.1. Since CR users lose one spectrum band in this case, they select a single spectrum band with highest $C_i^{CR}(k)/\beta_i$ to meet the sustainable rate and then do the resource allocation by considering all spectrum bands assigned to user k.

4.5 Spectrum Decision for Best Effort Applications

The objective of a typical scheduling method for the best-effort application is to maximize the network capacity. The spectrum decision for best-effort applications has the same objective as that of scheduling but it additionally needs to exploit the PU activity and long-term channel characteristics. Similar to the real-time application, the spectrum decision for the best-effort application can be classified into a maximum capacity based spectrum decision - single selection (MCSD-SS) and a multiple selections (MCSD-MS).

4.5.1 Maximum Capacity-based Spectrum Decision - Single Selection (MCSD-SS)

Optimally, for the maximum capacity, the CR network has to perform an entire spectrum decision over all current transmissions at every decision event, which requires high computational complexity. Also, the entire resource re-allocation leads to the spectrum switching of multiple users at the same time resulting in the abrupt quality degradation. Instead, we introduce a sub-optimal method for best-effort applications. If current resource allocation is optimal, the spectrum decision for the maximization of network capacity can be simplified as the following selection problem to choose spectrum bands so that the decision gain can be maximized.

$$\textit{Maximize: } \sum_{i \in \mathcal{A}} (\mathcal{G}(i, C_i^{\mathrm{CR}}(k), W_i) - \mathcal{L}(i, C_i^{\mathrm{CR}}(k), W_i)) x_i \tag{37}$$

$$\textit{subject to: } \sum_{i \in \mathcal{A}} x_i = N \tag{38}$$

where $\mathcal{G}(i, C_i^{\mathrm{CR}}(k), W_i)$ is the expected capacity gain when a new user k with the normalized CR capacity $C_i^{\mathrm{CR}}(k)$ joins the spectrum i with the available bandwidth W_i and $\mathcal{L}(i, C_i^{\mathrm{CR}}(k), W_i)$ is the expected capacity loss of other users in that spectrum band. $\mathcal{A}$ is the set of the currently available spectrum bands and N is the number of the transceivers of a CR user. $x_i \in \{0, 1\}$ represents the spectrum selection parameter used in the binary integer optimization. The decision gain can be defined as the sum of the difference between capacity gain and capacity loss caused by the addition of a new user.

Assume the spectrum sharing algorithm assigns the bandwidth to the users fairly. Then the capacity of each user competing for the same spectrum can be approximated as $C_i^{\text{CR}}(k) \cdot W_i / n_i^{\text{b}}$ where $n_{\text{b},i}$ represents the number of best effort users currently residing in spectrum i. Based on this capacity, the decision gain can be derived as follows:

$$\mathcal{G}_i - \mathcal{L}_i = \frac{C_i^{\text{CR}}(k) \cdot W_i}{n_{\text{b},i} + 1} - \sum_{j \in \mathcal{E}_i} \left(\frac{1}{n_{\text{b},i}} - \frac{1}{n_{\text{b},i} + 1}\right) \cdot C_i^{\text{CR}}(j) \cdot W_i \tag{39}$$

where $\mathcal{E}_i$ is the set of the best-effort CR users currently residing in spectrum band i. The first term represents the capacity gain of a new CR user k and the second term describes the total capacity loss of $n_{\text{b},i}$ CR users in spectrum i caused by the addition of a new CR user.

4.5.2 Maximum Capacity based Spectrum Decision - Multiple Selections (MCSD-MS)

Similar to the MVSD-MS, MCSD-MS enables multiple CR users to select a single spectrum band. Thus, the CR network first determines the order of spectrum decision, and then chooses the spectrum band for each CR user as follows:

- Each CR user who loses its spectrum band, finds a candidate spectrum band with the highest decision gain.
- A CR user with the highest decision gain is assigned to the spectrum first through the optimization given in Eq. (37).
- According to the optimization result, the CR network updates the current bandwidth allocation and repeats the MCSD-MS for the remaining CR users who need to be assigned to the new spectrum band.

4.6 Dynamic Resource Management for Spectrum Decision

Because of the PU activities, the available spectrum bands show the time-varying characteristics in the CR network. Thus, with the only proposed decision schemes, the CR network is not able to exploit spectrum resources efficiently, and hence results in the violation of the guaranteed service quality. As a result, the CR network necessitates an additional resource management scheme to coordinate the proposed spectrum decision methods adaptively with the bandwidth fluctuations. The main objectives of the proposed resource management are as follows:

- The CR network is capable of determining the acceptance of a new incoming CR user without any effect on the service quality of currently transmitting users.
- During the transmission, the CR network needs to maintain the service quality of currently transmitting users by considering the fluctuation of the available bandwidth.
- Since real-time users usually have a higher priority in spectrum access, best effort users may not have enough resources. Thus, the CR network may be required to balance the bandwidth between real-time and best effort users.

According to the current bandwidth utilization, the proposed resource management method specifies three different network states where different spectrum decision schemes are applied to satisfy the service requirements adaptively as illustrated in Figure 25. In the following subsections, we define these network states to describe the current spectrum utilization. Based on them, first we present an admission control scheme, and then propose decision control methods for two different decision events: CR user and primary user appearances.

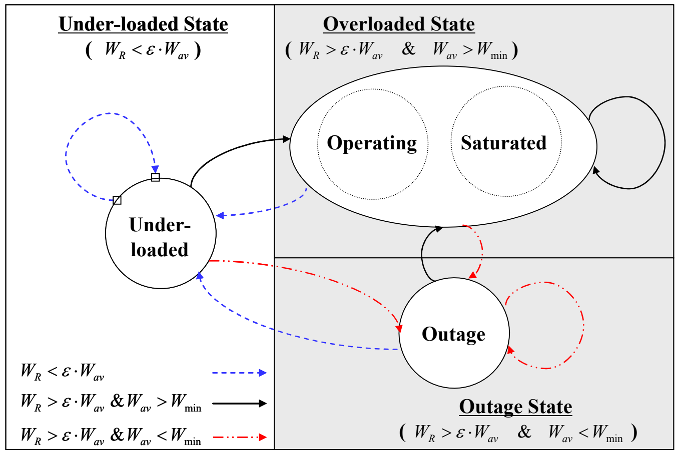


Figure 25: The state diagram for resource management.

4.6.1 Spectrum States for Admission Control

To exploit the spectrum resources efficiently, the proposed spectrum decision needs to adapt to the time-varying network conditions. Thus, we classify the network condition into three states according to the bandwidth utilization. Let W_{R} be the bandwidth currently assigned to real-time applications, and W_{av} be the total available bandwidth not occupied by primary users. W_{R} and W_{av} are time-varying according to the spectrum decision results and PU activities, respectively. $W_{\min}$ represents a minimum bandwidth to guarantee the service requirements of current users. Since the best-effort applications do not have strict service requirements, we consider only the bandwidth assigned to the real-time applications in determining the network state. As shown in Figure 25, the network states are classified as follows:

- *Under-loaded state:* If the current occupancy of real-time users, $W_{\mathrm{R}}/W_{\mathrm{av}}$ is less than ϵ, the CR network is under-loaded. ϵ is the pre-defined overload threshold to determine if the network is overloaded or not.

- *Overloaded state:* When $W_{\mathrm{R}}/W_{\mathrm{av}} > \epsilon$, the CR network is now overloaded. According to the amount of the remaining bandwidth, this state can be classified into two sub-states. If the expected bandwidth required for the spectrum decision, W_{req}, is less than the currently unused bandwidth $W_{\text{av}} - W_{\text{R}}$, the CR network is in the beginning of the overloaded state and still has enough resources (*operating state*). Otherwise, the CR network is almost saturated and does not have enough bandwidth for the current spectrum decision operation (*saturated state*). W_{req} is given in Section 4.6.3.2.

- *Outage state:* If the available bandwidth W_{av} is below W_{min}, the CR network cannot provide the guaranteed service quality to the currently active CR users.

Based on these network states, we propose the resource management schemes for the appearances of CR users and primary users in the following subsections.

4.6.2 Admission Control

The CR network is responsible for guaranteeing the service requirements of current CR users regardless of both bandwidth fluctuations and the appearance of new CR users. Thus, if the CR network cannot maintain the service requirements, it should reject a new incoming CR user, referred to as an *admission control*. The proposed admission control method requires the following procedures:

- *User characterization:* Each CR user requires different bandwidth to achieve the same service requirements, which mainly depends on its spectrum condition. The spectrum condition of each user k can be represented as its normalized capacity over all spectrum bands $C(k)$ as follows:

$$C(k) = \frac{\sum_{i=1}^{M} C_i^{\text{CR}}(k) \cdot B_i}{\sum_{i=1}^{M} B_i} \tag{40}$$

  where M is the number of all spectrum bands and B_i is the total bandwidth of spectrum i.

- *Bandwidth for guaranteeing the service quality:* Since the available bandwidth

W_{av} varies over time, the CR network cannot always satisfy the service requirements. Thus, we introduce the lower limit of the bandwidth W_{min} to guarantee the service requirements of the current CR users. Assume that regardless of the bandwidth fluctuation, the CR network should guarantee an average sustainable rate, $R_{\text{min}}(k)$, over an entire session of real-time user k. Then the minimum bandwidth of user k to support $R_{\text{min}}(k)$ is expressed as $R_{\text{min}}(k)/C(k)$. When a new CR user appears, W_{min} can be expressed as the sum of the minimum bandwidth for all CR users including both current users and an incoming user.

- *Admission criterion:* The proposed spectrum decision is designed for the network state when the available bandwidth W_{av} is above W_{min}. Thus, to maintain service requirements of the current CR users, W_{av} should not exceed W_{min}. Otherwise, the network condition is in the outage state. However, W_{min} is time-varying according to the current users and spectrum availability. To mitigate this temporal resource fluctuation, we first determine the stable interval T_{min} ,which is defined as the average period where no CR user disappearance is detected, and accordingly W_{min} does not change. Assume that the departure rate of CR users is μ. Then, T_{min} can be obtained as $1/(\mu \cdot n_{\text{r}})$ on average, where n_{r} is the number of current real-time users. To avoid resource outage resulting from the addition of a user, the proposed scheme accepts a new incoming user only if a *resource outage probability* during this interval is greater than the predetermined acceptable outage probability π. Otherwise, it is rejected. The resource outage probability, P_{out}, is the probability that $W_{\text{av}} > W_{\text{min}}$, which is derived in Appendix E.

The performance of the admission control method depends on the acceptable outage probability π. If the CR network has a higher π, it can accept more users while resulting in a higher quality degradation since it becomes highly probable that W_{av} is below W_{min}. As a result, we need to consider the quality degradation when

$W_{\mathrm{av}} < W_{\mathrm{min}}$ in the spectrum decision. Since the network capacity is proportional to the available bandwidth, the data loss rate newly introduced by the admission control can be approximately estimated by considering the expected available bandwidth as follows:

$$\widehat{P}_{\mathrm{loss}} = \frac{W_{\mathrm{min}} - E[\mathbf{W}_{\mathrm{av}} | \mathbf{W}_{\mathrm{av}} < W_{\mathrm{min}}]}{W_{\mathrm{min}}} \cdot P_{\mathrm{out}} \tag{41}$$

Here, the first term represents the ratio of the amount of bandwidth shortage for the minimum service quality to the bandwidth limit W_{min}.

The proposed MVSD method explained in Section 4.4 tries to satisfy the target data loss rate $P^{\mathrm{th}}_{\mathrm{loss}}$ on the assumption of the networks condition with the sufficient bandwidth. Since the spectrum decision also needs to consider the additional data loss rate resulting from the bandwidth shortage, the actual data loss rate should be expressed as the sum of the $P^{\mathrm{th}}_{\mathrm{loss}}$ and $\widehat{P}_{\mathrm{loss}}$. Assume that real-time users have the maximum allowable data loss rate P_{loss}. To satisfy this service requirement, the target rate $P^{\mathrm{th}}_{\mathrm{loss}}$ should be decided as follows:

$$P^{\mathrm{th}}_{\mathrm{loss}} = P_{\mathrm{loss}} - \widehat{P}_{\mathrm{loss}} \tag{42}$$

The proposed admission control method is originally designed only for real-time users. Since the best-effort users do not have strict service requirements, they do not need to have the admission control method.

4.6.3 Decision Control

Here, we propose the decision control schemes for both CR and primary user appearances, which enables spectrum decision to adapt to the different network states.

4.6.3.1 Decision Control in CR User Appearance

One of the important roles of decision control is to allocate spectrum resources with the minimum influence on current CR users when a new CR user appears. Figure 26 shows the procedures of the proposed decision control when a new CR user appears.

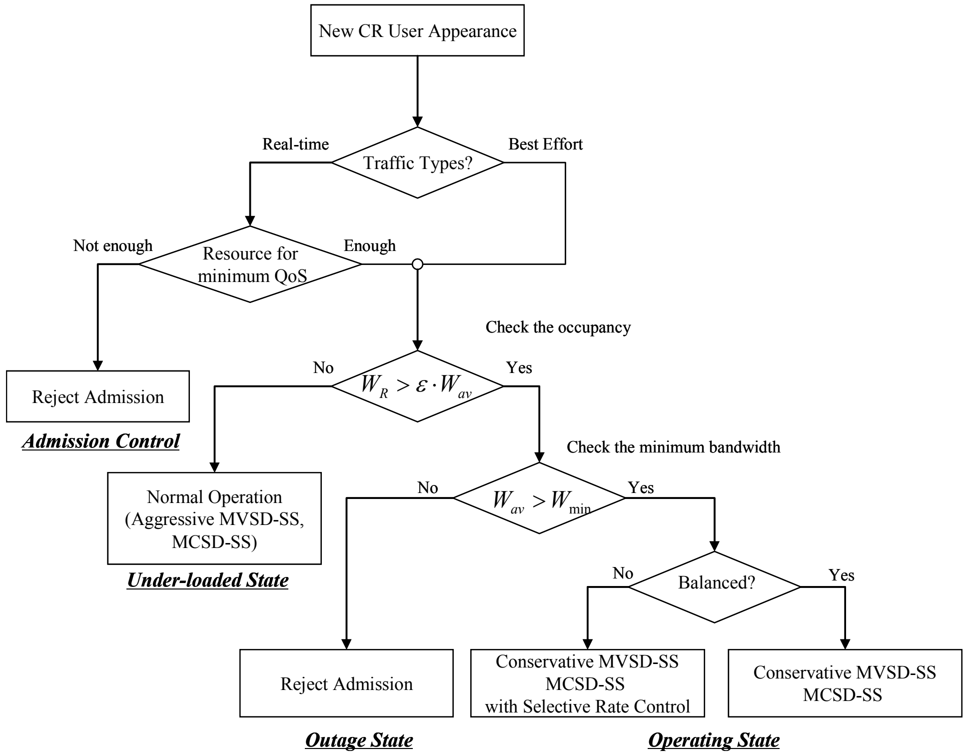


Figure 26: The flow chart for the proposed decision control - CR user appearance.

According to the states, the proposed control scheme coordinates spectrum decision as follows:

- ***Under-loaded State:*** Since the available bandwidth is sufficient in the under-loaded state, the CR network performs the spectrum decision aggressively. For a real-time user, the aggressive MVSD-SS is used whereas for a best-effort user, the MCSD-SS is applied.

- ***Overloaded State:*** Since the available bandwidth becomes scarce in this state, the spectrum decision needs to be more spectrum-efficient. Thus, the CR network performs the conservative MVSD-SS for the real-time user. However, since real-time users occupy much higher bandwidth through this operation, best-effort users may experience the bandwidth starvation in the overloaded state.

If the CR network is required to balance the resource allocation between real-time and best effort users, the resource management needs to check the current bandwidth utilization of both applications before it performs the MCSD-SS. Let δ be a balance coefficient pre-determined by the CR network. If the average bandwidth of current real-time users, $W_{\mathrm{R}}/n_{\mathrm{r}}$ is greater than the weighted average bandwidth for best-effort users, $\delta \cdot (W_{\mathrm{av}} - W_{\mathrm{R}})/n_{\mathrm{b}}$, current resource allocation is considered to be unbalanced where n_{r} and n_{b} is the number of the current real-time users and the current best-effort users, respectively. To solve the resource starvation problem in best effort users, we propose a *selective rate control* which maintains the resource balance in the overloaded state by reducing the sustainable rate of the selected real-time users.

When each real-time user k reduces its sustainable rate to a one-step lower rate, the expected bandwidth gain is expressed as $\Delta R(k)/C(k)$ where $\Delta R(k)$ and $C(k)$ is the rate decrement and the normalized capacity of a real-time user k, respectively. Based on the bandwidth gain, the CR network selects real-time users for the selective rate control to minimize total rate reduction subject to the balance constraint, which can be expressed as the following linear integer optimization problem:

$$\textit{Minimize:} \quad \sum_{k \in \mathcal{R}} \Delta R(k) \cdot x_k \tag{43}$$

$$\textit{subject to:} \ \frac{W_{\mathrm{R}} - \Delta W}{n_{\mathrm{r}}} - \delta \cdot \frac{W_{\mathrm{av}} - W_{\mathrm{R}} + \Delta W}{n_{\mathrm{b}}} \leq 0 \tag{44}$$

$$\Delta W \geq \sum_{k \in \mathcal{R}} \frac{\Delta R(k)}{C(k)} \cdot x_k, \qquad x_k \in \{0, 1\} \tag{45}$$

where $\mathcal{R}$ is the set of real-time users currently active and ΔW is the bandwidth required for the balance. $x_k \in \{0, 1\}$ represents the user selection parameter used in the binary integer optimization.

The real-time users selected through the above optimization reduce their sustainable rates to the one-step lower rates and then perform the resource allocation explained in Section 4.4.1. If $W_{\mathrm{R}}/n_{\mathrm{r}} < \delta \cdot (W_{\mathrm{av}} - W_{\mathrm{R}})/n_{\mathrm{b}}$, best-effort users have enough bandwidth to satisfy the balance condition. In this case, spectrum decision does not need to perform selective rate control.

Whenever the best-effort user appears in this state, the resource management scheme tries to satisfy the balanced condition. However, to avoid the abrupt quality degradation of real-time users, a selective rate control can change the sustainable rate of real-time users to only a one-step lower rate.

- ***Outage State:*** The service requirements of CR users cannot be guaranteed because of the shortage of spectrum resources. In principle, all new incoming users should be rejected in this state to avoid the overall quality degradation. However, since new real-time users in this state are already rejected through the admission control, this state is meaningless in real-time user appearances. Thus, only the best effort users are rejected in this state.

4.6.3.2 Decision Control in Primary User Appearance

Once the CR network accepts the users, it should guarantee their service requirements during the transmission regardless of the bandwidth variation. Thus, the decision control scheme should assign the bandwidth for the spectrum decision adaptively to the current network state. Figure 27 shows the decision control procedure in the primary user appearance. According to the network states, the proposed scheme can be performed as follows:

- ***Under-loaded State:*** Similar to the CR user appearance, the CR network performs the spectrum decision aggressively. For a real-time user, the aggressive MVSD-MS is used whereas the MCSD-MS is executed for a best-effort user.

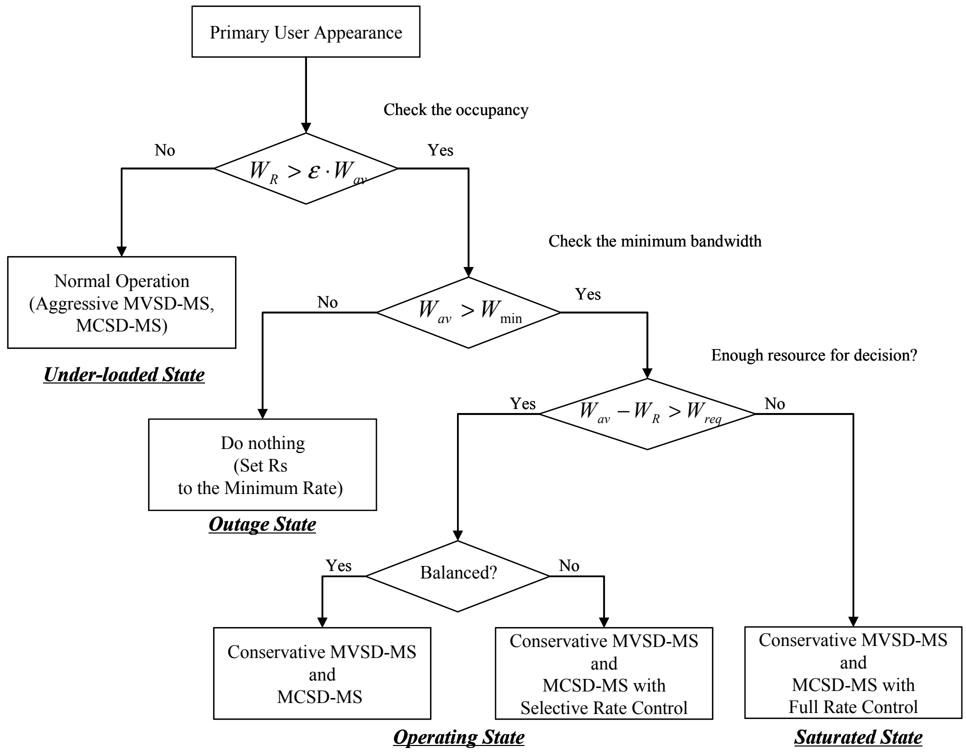


Figure 27: The flow chart for the proposed decision control - PU appearance.

Operating State: In the overloaded state, the decision control starts to coordinate the bandwidth allocation to maintain the service quality. In the primary user appearance, the overloaded state can be divided into two different sub-states according to the remaining spectrum resources. In the operating state, the CR network is considered to be overloaded but still has enough resources for spectrum decision,, i.e., the available bandwidth W_{av} is greater than the bandwidth required for spectrum decision, W_{req}.

When a primary user appears, the expected bandwidth requested in MVSD-MS, W_{req} can be derived as follows:

$$W_{\text{req}} = \sum_{k \in \mathcal{R}_l} \frac{R_{\text{lost}}(k)}{\sum_{i \in \mathcal{A}} C_i^{\text{CR}}(k) \cdot W_i} \cdot (W_{\text{av}} - W_{\text{R}}) \tag{46}$$

where W_i is the available bandwidth of spectrum i currently unused by both primary and real-time CR users and $R_{\text{lost}}(k)$ is the lost capacity of user k due to the PU activities. $\mathcal{A}$ is the set of the currently available spectrum bands and

$\mathcal{R}_l$ is the set of the real-time users who lose their spectrum band, respectively. Here, W_{req} is expressed as the sum of the expected bandwidth of user $k \in \mathcal{R}_l$ required to support $R_{\text{lost}}(k)$. The denominator in the summation in Eq. (46) represents the total expected capacity of user k over all currently available spectrum bands.

If the bandwidth of both applications is balanced, the CR network performs a conservative MVSD-MS and an MCSD-MS. If it is not balanced, the CR network needs a selective rate control before the spectrum decision similar to the CR user appearance. The only difference is that a selective rate control is just applied to the real-time users losing one of their spectrum bands to minimize the influence on the on-going transmissions of real-time users.

- ***Saturated State:*** The other overloaded state in primary user appearance is the saturated state where the remaining available bandwidth is less than the bandwidth required for the spectrum decision. In this case, real-time CR users cannot find the new spectrum bands to maintain their service requirements, which necessitates the re-negotiation of their service requirements.

  Let all possible sustainable rates for user k be $\{R_{\text{s},1}(k), R_{\text{s},2}(k), \ldots, R_{\text{s},n_k}(k)\}$ where n_k is the number of all possible sustainable rates. Then the expected bandwidth of each possible sustainable rate can be obtained as $R_{\text{s},i}(k)/C(k)$ where $C(k)$ is the normalized capacity of user k as given in Eq. (40).

  Based on the expected bandwidth gains in re-negotiation, we propose a *full rate control* where the sustainable rate of the real-time users currently requesting spectrum decision is optimized to satisfy both bandwidth and balance constraints. This optimization problem is expressed as the following linear integer optimization, the so-called lockbox problem [20].

$$\begin{aligned} \textit{Maximize:} \ & \sum_{k \in \mathcal{R}_l} R_{\mathrm{s}}(k) \\ &= \sum_{k \in \mathcal{R}_l} \sum_{i=1}^{n_k} R_{\mathrm{s},n_k}(k) \cdot x_i(k) \end{aligned} \tag{47}$$

$$\textit{subject to:} \ \frac{W_{\mathrm{R}}^s + \widehat{W}_{\mathrm{R}}}{n_{\mathrm{r}}} - \delta \cdot \frac{W_{\mathrm{av}} - W_{\mathrm{R}}^s - \widehat{W}_{\mathrm{R}}}{n_{\mathrm{b}}} < 0 \tag{48}$$

$$\widehat{W}_R < W_{\mathrm{av}} - W_{\mathrm{R}}^s \tag{49}$$

$$\widehat{W}_R = \sum_{k \in \mathcal{R}_l} \sum_{i=1}^{n_k} \frac{R_{\mathrm{s},i}(k)}{C(k)} \cdot x_i(k) \tag{50}$$

$$\sum_{i=1}^{n_k} x_i(k) = 1 \qquad x_i(k) \in \{0, 1\} \tag{51}$$

where $\mathcal{R}_l$ is the set of the real-time users who lose their spectrum bands, $\widehat{W}_{\mathrm{R}}$ is the expected bandwidth for the real-time users who lose their spectrum band, and W_{R}^s is the bandwidth of the real-time users not affected by the PU activities. Eq. (48) is the constraint on the resource balance explained in Section 4.6.3.1. Eq. (49) is the constraint on the available bandwidth required for the spectrum decision.

- ***Outage State:*** This state cannot provide a guaranteed service quality any longer. Thus, even though the CR network needs the spectrum decision, all CR users who lose their connections reduce their sustainable rate to the minimum and just wait until the network condition becomes better.

4.7 Performance Evaluation

In this section, we provide simulation experiments to evaluate the performance of the proposed spectrum decision framework.

4.7.1 Simulation Setup

Here we simulate an infrastructure-based CR network consisting of one base-station and multiple CR users. Each user is uniformly distributed over the network coverage with the radius of 2 km. The CR network is assumed to operate in 20 licensed spectrum bands consisting of 4 very high frequency (UHF)/ultra high frequency (UHF) TV, 4 advanced mobile phone system (AMPS), 4 global system for mobile communications (GSM), 4 code division multiple access (CDMA), and 4 wideband CDMA (W-CDMA) bands. The bandwidth of these bands are 6MHz (TV), 30kHz (AMPS), 200kHz (GSM), 1.25MHz (CDMA), and 5MHz (W-CDMA), respectively. The PU activities of each spectrum band, α_i and β_i, are randomly selected over $[0, 1]$. The service rate of CR traffic μ is 0.02, and its arrival rate can be determined according to the average number of users. In the simulations, we assume a log-normal fading channel model, where the noise power is -115dBm, the shadowing deviation is 4, and the path loss coefficient is set to 4 [60]. Transmission power $P_i^k(f)$ is unity over all frequencies.

Through the spectrum sensing, the base-station is already aware of the spectrum availability in its coverage. Sensing efficiency η_i, and a false alarm probability P_i^{f} are set to 0.9 and 0.99, respectively. These sensing capabilities are assumed to be identical over all spectrum bands. Since the user-based and the band-based quality degradations explained in Section 4.2.3 use the same strategies as the primary user appearance and the CR user appearance respectively, we do not consider the quality monitoring in the simulations.

For the simulations, we use the real-time application supporting 5 different bi-trates: 64kbps, 128kbps, 256kbps, 512kbps, and 1.2Mbps. For the resource management, $W_{\min}$ and $R_{\min}$ are set to 10MHz and 512Kbps, respectively. The overloaded threshold ϵ is set to 0.5, the balance coefficient δ is 1. The acceptable data loss rate, P_{loss}, and the acceptable outage threshold π are set to 0.05 and 0.03, respectively.

To evaluate the performance of our spectrum decision framework, we introduce three different cases as follows:

- *Case 1:* CR users exploit all functionalities of the entire spectrum decision framework including MVSD, MCSD, and all resource management functions explained in Sections 4.4, 4.5, and 4.6, respectively.

- *Case 2:* CR users perform the proposed spectrum decision framework without the admission control scheme in the resource management.

- *Case 3:* CR users use only MVSD and MCSD methods (Case 1 without both admission and decision control schemes).

- *Case 4:* Instead of the optimization schemes in Section 4.4.1, the proposed MVSD scheme adopts exhaustive search to determine proper spectrum bands and their bandwidth, which achieves optimal performance in real-time users.

Since there are no previous work related to spectrum decision, we compare our decision framework with two straightforward decision criteria as follows:

- *Case 5 Capacity-based decision:* CR users select the spectrum bands with the highest channel capacity, as follows:

$$\begin{aligned} \textit{Maximize:}\ & \sum_{i \in \mathcal{A}} C_i^{\mathrm{CR}}(k) \cdot x_i \\ \textit{subject to:}\ & \sum_{i \in \mathcal{A}} x_i \leq N \quad x_i \in \{0,1\} \\ & \sum_{i \in \mathcal{A}} C_i^{\mathrm{CR}}(k) \cdot W_i \cdot x_i \geq \ R_{\mathrm{s}}(k) \end{aligned} \tag{52}$$

  $\mathcal{A}$ is the set of the currently available spectrum bands. The last constraint is applied only to the real-time applications.

- *Case 6 Primary user (PU) activity-based decision:* CR users select the spectrum bands with the lowest PU activity. Instead of the objective function to be maximized in the case 5, the case 6 uses $\sum_{i \in \mathcal{A}} 1/\beta_i \cdot x_i$.

In the following subsections, we show the simulation results in three different scenarios (only real-time users, only best effort users, and both of them).

4.7.2 Real-time Applications

First, we consider the scenario with only real-time users to validate the proposed MVSD-SS and MVSD-MS described in Section 4.4.

Figure 28 (a) shows how the average number of users influences the data loss rate. Here we assume 3 spectrum bands and 0.1 sec for the switching delay. For this simulation, we generate CR user traffic from 10 to 80 on average. When a small number of users are transmitting, each case shows relatively low data loss rate. However, as the number of users increases, while other methods (cases 2, 5, 6) increase the data loss rate, the case 1 still maintains a certain level of the data loss rate where the admission control controls the addition of new users adaptively dependent on current network utilization. However, the case 1 shows little higher data loss rate than the acceptable data loss rate. The reason is that during the transmission the MVSD-MS scheme maintains all on-going transmissions even though they cannot find the spectrum bands to satisfy the acceptable data loss rate, which causes slight increase in the data loss rate. Even though the proposed method does not use admission control(cases 2), it still shows better data loss rate than the case 5 and 6.

In Figure 28 (b), we investigate the performance of the spectrum decision under four PU activity scenarios - low/low, low/high, high/low, and high/high. Low PU activity (either α_i or β_i) is uniformly distributed between 0 and 0.5, and high PU activity is between 0.5 and 1. The average number of users, the number of spectrum bands, and switching delay are set to 50, 3, and 0.1sec respectively. In all cases, the case 1 shows better performance in data loss than other method (cases 2, 3, 5, 6), and the similar loss rate to the case 4. Also, is is shown that β_i is a more dominant factor to determine the loss rate than α_i since a higher β_i introduces more frequent

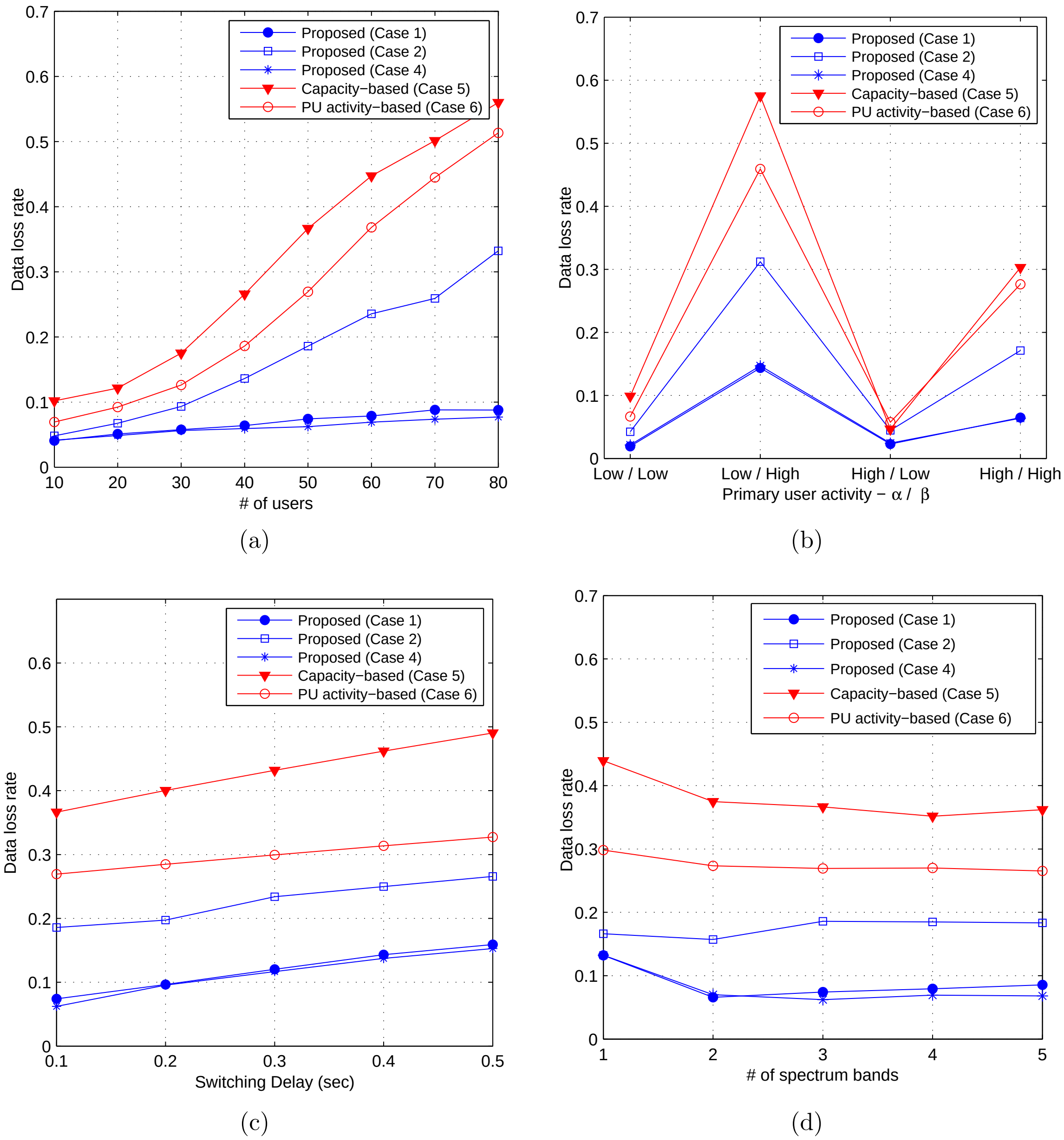


Figure 28: Data loss rate in real-time applications: (a) the number of users, (b) PU activities, (c) switching delay, and (d) the number of spectrums.

switching, leading to significant performance degradation.

We also show the relationship between the data loss rate and the switching delay in Figure 28 (c). Here we assume 50 users and 3 spectrum bands. Although longer switching delay increases the data loss rate in all cases, the proposed method (case 1) shows the lower loss rate than other method by both rejecting users before the transmission and reducing sustainable rate during the transmission. The case 2 still shows better performance than cases 5 and 6 because of both MVSD-SS and MVSD-MS.

As explained in Section 4.2.1, the transmission with multiple transceivers can mitigate the capacity fluctuation effects as well as prevent a temporary disconnection of communication channels. This phenomena are observed in Figure 28 (d). Here we assume 0.1 sec for the switching delay and 50 real-time users. An interesting point is that more spectrum bands do not always lead to good performance regarding the data loss rate. As the number of spectrum bands increases, the total amount of PU activities over multiple spectrum bands increases, which may cause an adverse effect on the data loss rate. In this simulation, each does not improve its data loss rate significantly when it has more than 2 spectrums.

In all simulation results, the proposed method (case 1) shows almost same performance as the optimal method (case 4), but requires less complexity as explained in Section 4.4.1

4.7.3 Best Effort Applications

To evaluate the performance of MCSD-MS and MCSD-SS described in Section 4.5, we compare the proposed method (case 1) with the cases 5 and 6 we introduced above. Since the admission and decision control functionalities are not needed in the only best effort scenario, we do not consider the cases 2, 3, and 4 here. In this simulation, we also show how the number of users, PU activity, the switching delay, and the number of spectrum bands influence the total network capacity. As shown in Figure 29, the case 1 shows higher capacity compared to the capacity-based and PU activity-based methods. Figure 29 (a) indicates the relationship between the number of users and total network capacity where we can observe the case 1 provides a better performance over cases 5 and 6 by exploiting the PU activity and the channel

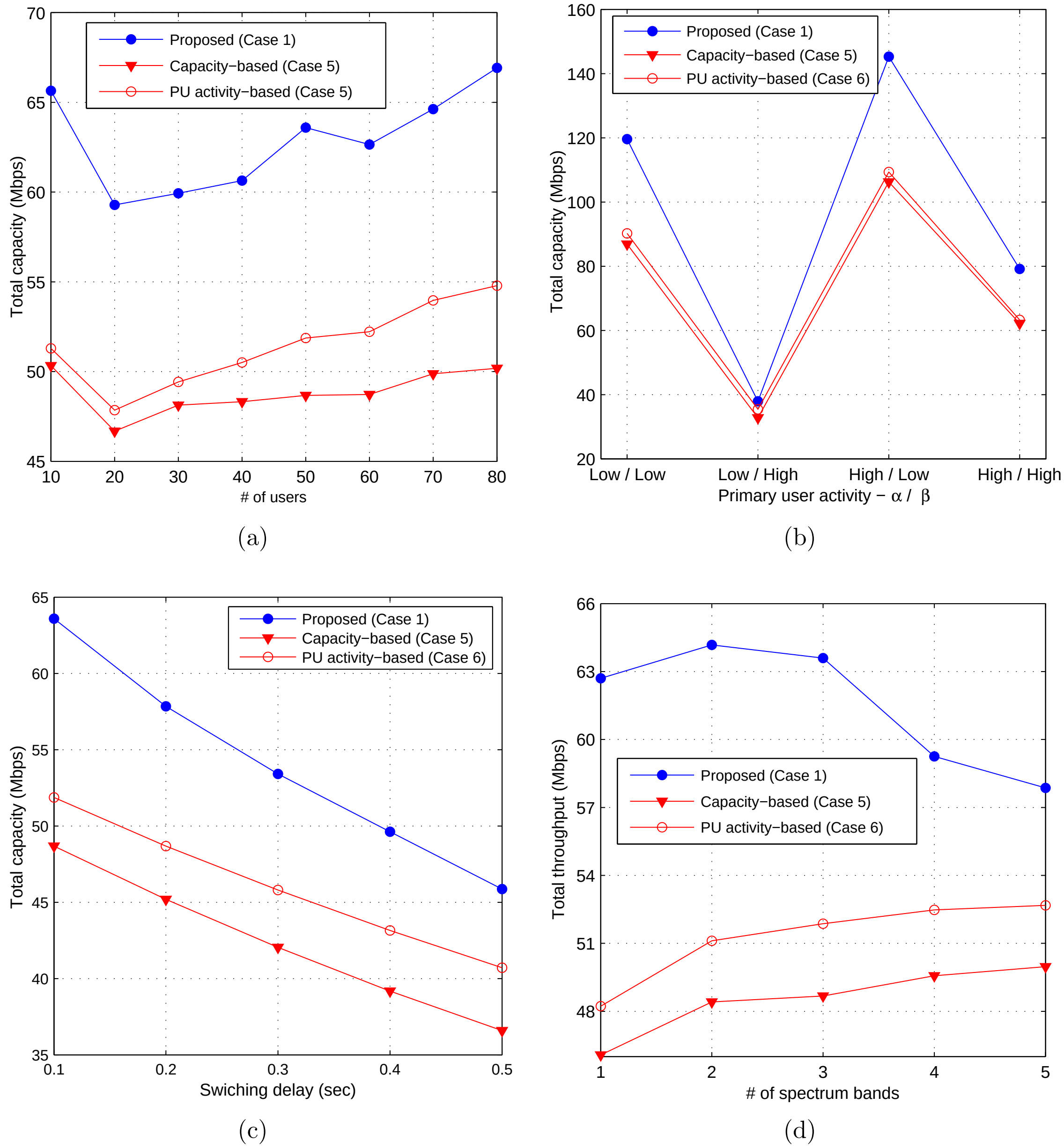


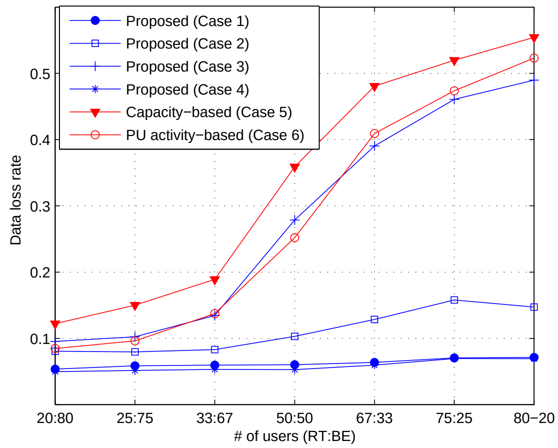
Figure 29: Total network capacity in best effort applications: (a) the number of users, (b) PU activities, (c) switching delay , and (d) the number of spectrums.

condition at the same time. In Figure 29 (b), we investigate how PU activities influence the performance of total capacity. The case 1 shows better performance than other cases. Especially when β_i is low, the case 1 shows more improvement because of less capacity degradation caused by switching delay. In Figure 29 (c), the simulation results on the total network capacity is examined when 50 best effort users with 3 spectrum bands are applied simultaneously. The results show that the increase in the switching delay has an adverse influence on network capacity. Figure 29 (d) shows the simulation results on the total network capacity when 50 best effort users with 0.1 sec switching delay are applied simultaneously. Similar to the simulation on real-time users, the case 1 shows the best performance in 2 spectrum bands, but less total capacity in more than 2 spectrum bands, since it causes a higher spectrum switching as well as prevents exploiting channel diversity.

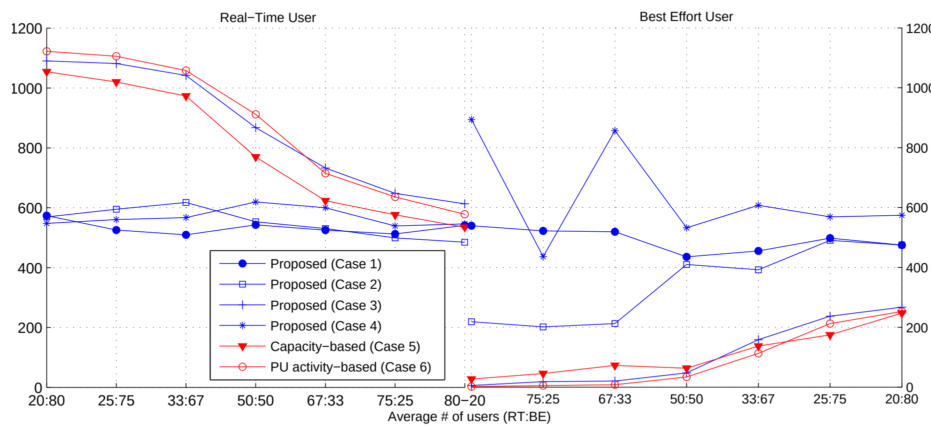
4.7.4 Hybrid Scenario

Here, we consider a hybrid scenario where both real-time and best effort users co-exist. Similarly, we assume 3 spectrum bands and 0.1 sec switching delay for this simulation. Here we set the total number of active users to 100 and vary the number of best effort and real-time users to investigate the performance according to the network state. In Figure 30 (a), we show the data loss rate of the real-time users on a hybrid scenario where we apply best effort and real-time users simultaneously. In the under-loaded state, i.e., when there are a small number of real-time users in the CR network, we can see each method shows lower data loss rate. On the other hand, overloaded conditions lead to considerably different performance according to the decision methods. Through admission and decision controls, the case 1 admits real-time users when it can guarantee the service requirement of current users, and hence maintains the lowest data loss rate. When the proposed method does not use the admission and decision controls (cases 2 and 3), the CR network accepts much more real-time users than it can provide with the guaranteed service quality, leading to the increase in the data loss rate and the decrease in the average capacity of the real-time user as described in Figure 30 (b). On the contrary, cases 5 and 6 show the worst data loss rates.

In Figure 30 (b), we show the average user capacity of the real-time and best



(a)



(b)

Figure 30: Data loss rate in the hybrid scenario: (a) data loss rate , and (b) user capacity.

effort users in the hybrid scenario. When real-time users are less than the best effort users, cases 5 and 6 show the highest user capacity in real-time users while maintaining slightly higher data loss rate as that of the proposed methods (cases 1, 2, 3). On the contrary, as the number of real-time users increases, while the proposed method (case 1) still provide enough capacity to best effort users, cases 5 and 6 show low capacity of best-effort users because of the lack of resource management. Even though real-time users occupy the most of bandwidth resources the cases 5 and 6, they cannot satisfy the service requirements and show the highest data loss

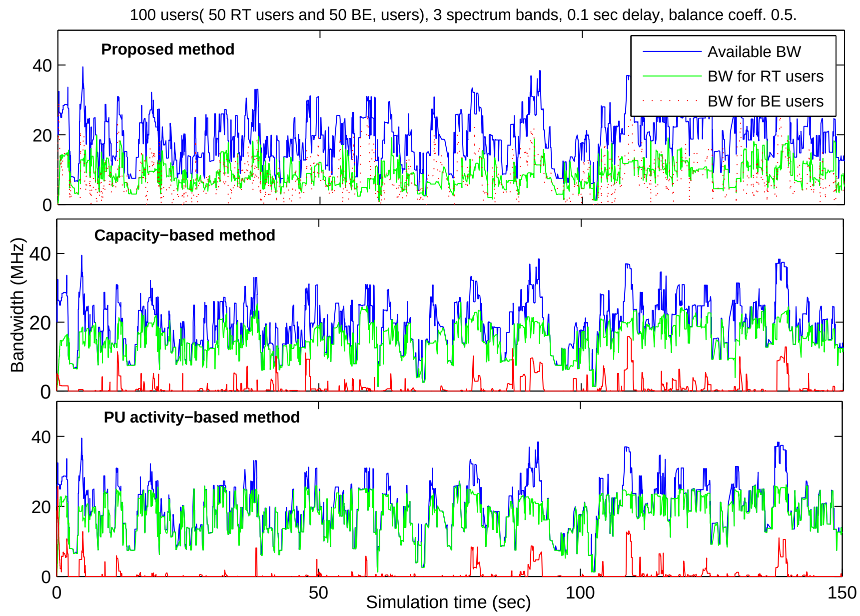


Figure 31: The comparison of the bandwidth utilization in the hybrid scenario.

rate as observed in Figure 30 (a). By exploiting the admission control scheme, the cases 1 and 2 shows better fairness in capacity between both application types while maintaining the low data loss rate in real-time users. Though the case 3 does not use both admission and decision control schemes, it shows slightly higher capacity of best effort users than cases 5 and 6 since the MVSD scheme provides bandwidth-efficient resource allocations, leading to increase in available bandwidth for best effort users as explained in Section 4.4. Similarly, the optimal method (case 4) selects the bandwidth efficient spectrum for real-time users, leading to slightly higher capacity of best effort users than the proposed method (case 1), while it achieves almost same data loss rate and user capacity in real-time users as the proposed method.

Figure 31 shows how the proposed admission control exploits bandwidth resources when 50 real-time users and 50 best-effort users are transmitting simultaneously. From the simulation results, we can see the proposed admission control (case 1) balances the bandwidth between both applications over the entire simulation time. On the contrary, in cases 5 and 6, real-time applications occupy the most of the available bandwidth to satisfy their service requirements, which leads to the bandwidth starvation of best effort users. From these simulations, we can observe that by exploiting spectrum resources efficiently, our spectrum decision framework provides the guaranteed service quality while balancing bandwidth allocation between real-time and best effort users.

CHAPTER V

JOINT SPECTRUM AND POWER ALLOCATION FOR SPECTRUM SHARING IN INFRASTRUCTURE-BASED CR NETWORKS

5.1 Introduction

Among the above spectrum management functions, spectrum sharing plays an important role in determining the performance of the CR network. Especially in infrastructure-based CR networks, their total network capacity mainly depends on the spectrum sharing schemes, which is comprised of resource allocations among base-stations, called *inter-cell spectrum sharing*, and among CR users residing in the same cell, called *intra-cell spectrum sharing*. Recent research on spectrum sharing has explored two different sharing models: *exclusive allocation* and *common use*. The *exclusive allocation* approach allows the CR user to use the spectrum exclusively to its neighbor users. Although the exclusive approach is known to be optimal [19], it has unfair resource allocation, especially in CR networks where the spectrum availability varies significantly over time and location. On the contrary, the *common use* approach enables each CR user to share the same spectrum with its neighbors by exploiting a sophisticated power allocation method [19] [58]. Although this method can mitigate the unfairness in resource allocation, it achieves lower total capacity than exclusive allocation because of the existence of higher co-channel interference. Since most of the research on spectrum sharing has focused on only one sharing model (either exclusive or common use models), spectrum and power allocations have not been considered together to date. With either of these approaches, the infrastructure-based CR network cannot achieve its objectives, high spectrum utilization and fair resource allocation with interference avoidance.

To address these challenges, we propose a hierarchical spectrum sharing framework for infrastructure-based CR networks, consisting of inter-cell and intra-cell spectrum sharing schemes. More specifically, in inter-cell spectrum sharing, each cell exploits the exclusive and common use approaches dynamically according to the spectrum utilization in its vicinity. In exclusive allocation, the base-station determines spectrum bands to achieve the highest expected cell capacity. This is characterized by the permissible transmission power based on the primary user (PU) activities. If there is no spectrum available for exclusive allocation, our framework switches to the common use approach, where spectrum selection is based on the interference and PU activities in the neighbor cells. This helps to realize 1) maximum cell capacity, 2) less influence to neighbor cells, and 3) interference-free uplink transmission. Furthermore, to protect the transmission of primary networks, inter-cell spectrum sharing necessitates a sophisticated power allocation scheme in both exclusive and common use methods. In addition, we propose an intra-cell spectrum sharing method where the base-station assigns the spectrums, obtained through inter-cell spectrum sharing, to its CR users so as to maximize the cell capacity as well as to avoid interference to primary networks. The rest of the chapter is organized as follows: In Section 5.2, we describe the limitations of conventional spectrum sharing methods and motivate a hybrid spectrum sharing approach. Section 5.3 presents the network architecture. In Section 5.4, we propose a framework for spectrum sharing in infrastructure-based CR networks. In Sections 5.5 and 5.6, we develop the spectrum allocation methods for exclusive and common use models, respectively. In Section 5.7, we explain a power allocation method for inter-cell spectrum sharing. In Section 5.8, an intra-cell spectrum sharing scheme is introduced. Performance evaluation and simulation results are presented in Section 5.9.

5.2 Motivation

In this section, we present conventional spectrum sharing methods, and describe the practical considerations for inter-cell spectrum sharing which are the motivations of our proposed work.

5.2.1 Related Work

Spectrum sharing has been considered as a main functionality to determine the total capacity of CR networks [3]. There are two different classical approaches in spectrum sharing: spectrum allocation for an exclusive model and power allocation for a common use model, which will be explained in the following subsections.

5.2.1.1 Exclusive Allocation Approach

Spectrum resource can be assigned to only one user to avoid interference to other neighbor users. In [56], a graph coloring based collaborative spectrum access scheme is proposed, where a topology-optimized allocation algorithm is used. In mobile networks, however, the network topology changes according to the node mobility. Using this global optimization approach, the network needs to completely recompute spectrum assignments for all users after each change, resulting in a high computational and communication overhead. Thus, a distributed spectrum allocation based on local bargaining is proposed in [10], where CR users negotiate spectrum assignment within local self-organized groups. For the resource constrained networks such as sensor and ad hoc networks, a rule-based device centric spectrum management is proposed, where CR users access the spectrum independently according to both local observation and predetermined rules, leading to minimizing the communication overhead [11].

5.2.1.2 Common Use Approach

This solution allows multiple users to access the same spectrum at the same time. Thus, in this approach, power allocation is the most important part to increase the capacity with less interference to other users. Game theory has been exploited to determine the transmission power of each user [19] [34]. Although this approach can achieve the Nash equilibrium, it cannot always guarantee the Pareto optimum, leading to lower network capacity compared to the exclusive allocation model. In [19], orthogonal power allocation, i.e., exclusive allocation, is shown to be optimal to maximize the entire network capacity. However, the common use model achieves more fair resource allocation, especially in networks with few available spectrums. In [58], a centralized power allocation method is proposed that uses a spectrum server to coordinate the transmissions of a group of links sharing a common spectrum. In [39], an optimal power allocation scheme is proposed to achieve ergodic and outage capacity of the fading channel under different types of power constraints and fading models. In [77], joint beamforming and power allocation techniques are presented to maximize the capacity of CR users while ensuring the QoS of primary users. However, all of these methods necessitate a perfect knowledge of channels, i.e., the channel gains of all possible links including channels between primary and CR users.

In [30], both single channel and multi-channel asynchronous distributed pricing (SC/MC-ADP) schemes are proposed, where each node announces its interference price to other nodes. Using this information from its neighbors, the node can first allocate a channel and in case there exist users in that channel, then, determine its transmit power. While both methods consider the spectrum and power allocation at the same time, they do not address the heterogeneous spectrum availability which is a unique characteristic in CR networks.

5.2.2 Considerations in Infrastructure-based CR Networks

Recent work on spectrum sharing has mainly focused on the distributed ad-hoc networks. However, the infrastructure-based CR networks have unique challenges in spectrum sharing, which have been unexplored so far. Since the infrastructure-based

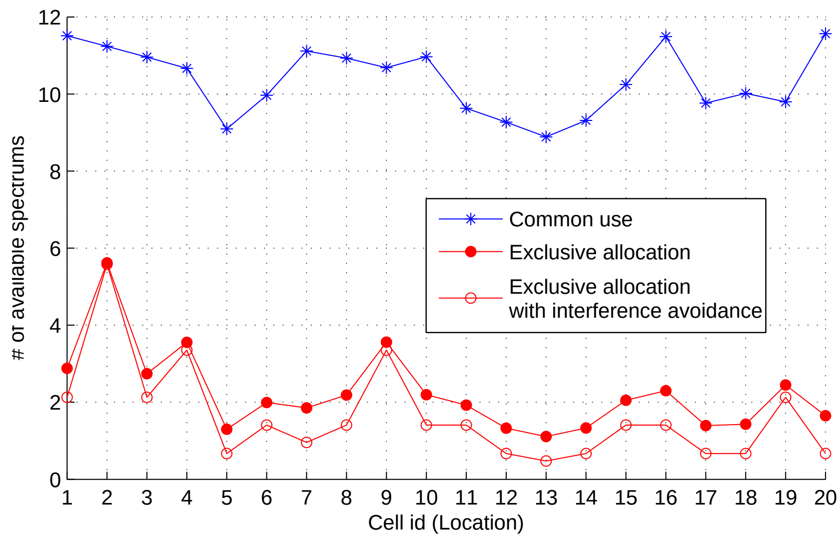


Figure 32: Available spectrum bands at different locations.

network consists of multiple cells, we need to consider not only spectrum sharing inside one cell, i.e., among users, called *intra-cell spectrum sharing* but also spectrum sharing among multiple cells, called *inter-cell spectrum sharing*. Furthermore, this network requires more strict fairness in resource allocation to provide communication channels to their users with the guaranteed service quality. Here are several practical issues we need to consider for spectrum sharing schemes in infrastructure-based CR networks.

5.2.2.1 Heterogeneous Resource Availability

The main difference between conventional wireless networks and CR networks lies in the PU activities. Since CR networks do not have spectrum license, they exploit the spectrum opportunistically and vacate the spectrum immediately when a primary user appears. According to the time and location, each cell experiences different PU activities, leading to the heterogeneous resource availability. Also the number of neighbor cells influences the performance of spectrum sharing. Figure 32 shows the number of available spectrum bands at each cell in the network topology used for our simulations in Section 5.9. The spectrum band is available only when all PU activity regions in the cell are idle. Thus, the expected number of available spectrum bands at a cell j can be expressed as $\sum_{i=1}^{N}[\prod_{m\in\widetilde{\mathcal{A}}_i(j)} p_{i,m}^{\text{off}}]$ where N is the total number of spectrum bands, $p_{i,m}^{\text{off}}$ is the probability that the spectrum i is idle at the primary activity region m, and $\widetilde{\mathcal{A}}_i(j)$ is a set of primary activity regions in the cell j at spectrum i. This value represents the spectrum availability of the common use approach. If the exclusive approach is used for spectrum sharing, the number of neighbor cells competing for the same spectrums should be considered as well. In [10], the lower bound of available spectrum resource for exclusive allocation, the so-called *poverty line*, is derived as $[\sum_{i=1}^{N}(\prod_{m\in\widetilde{\mathcal{A}}_i(j)} p_{i,m}^{\text{off}})]/(L+1)$ where L is the number of neighbor cells. As shown in Figure 32, spectrum availability varies significantly according to the cell locations. Furthermore, it shows different patterns in both common use and exclusive approaches. As a result, for the efficient spectrum utilization, we need to mitigate this heterogeneous spectrum availability by exploiting common use and exclusive approaches dynamically, i.e., in the limited spectrum environment, the common use approach helps to increase fairness in user capacity while the exclusive approach is much advantageous in the environment with sufficient spectrum resources.

5.2.2.2 Inter-Cell Interference

Since the interference range is generally larger than the transmission range [38], the transmission in the current cell influences its neighbor cells. For this reason, although the current cell does not detect any PU activity in its transmission range, its transmission may cause interference in the neighbor cell detecting PU activities. The simplest way to avoid this problem is not to use the spectrum where neighbor cells detect the transmission of primary networks. If we consider this constraint in the exclusive model, the available spectrum resources become lower as shown in Figure 32. Consequently, exclusive allocation shows an inefficient spectrum utilization in CR networks although it is theoretically optimal in classical wireless environments. To solve this problem while satisfying interference condition, exclusive allocation is also required to have a power allocation method adapting to spatial and temporal characteristics of PU activities.

5.2.2.3 Imperfect Knowledge of Neighbor Cells

Most of the power allocation schemes based on the common use approach assume that every CR user or a central network entity is aware of all radio information, such as the channel gains of all possible links and all interference information in the networks [58] [39] [77] [30]. However, in the infrastructure-based networks, it is impossible to obtain all necessary information for power allocation since there is no direct communication channel among CR users located in different cells. To get the information of neighbor cells, inter-cell spectrum sharing requires a cooperation mechanism among the cells. In addition, for a more practical spectrum sharing method, we need to estimate cell capacity with the minimum amount of information exchanged with neighbor cells.

5.3 System Model

5.3.1 CR Network Architecture

In this chapter, we consider the infrastructure-based CR network which has centralized network entities such as a base-station in cellular networks or an access point in wireless local area networks (LAN) [1]. The CR base-station forms a cell and have their own users, which are uniformly distributed in their coverage. To detect the transmission of primary networks, all CR users observe their local radio information and report them to their base-station. Based on these local measurements, CR base-stations determine the spectrum availability and allocate the spectrum resource to CR users [44].

[1] In the remainder of the chapter we will use the term "base-station" to refer to the central network entity

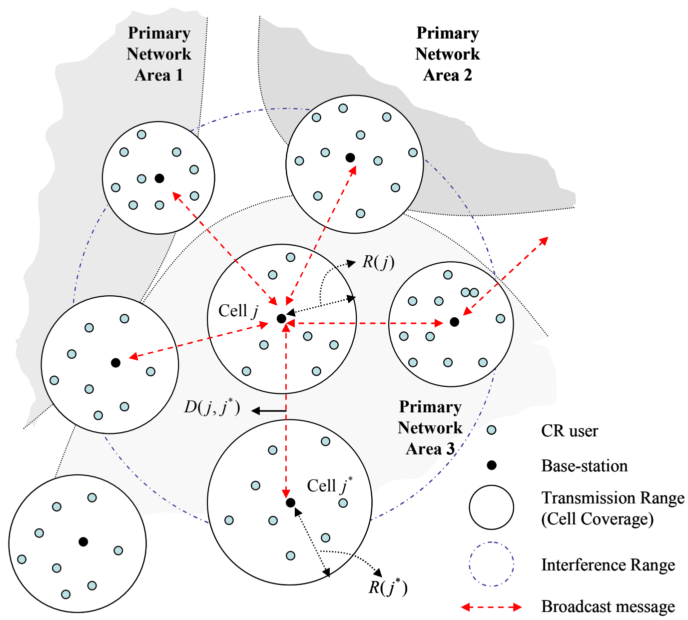


Figure 33: Network Architecture.

Figure 33 shows the network model that we consider in this chapter, where each base-station j has its transmission range with radius $R(j)$. $D(j, j^*)$ is the distance between base-stations j and j^*. Here, the transmission range (or cell coverage) is defined as the range within which a transmitted signal should be successfully received. The interference range is the area within which other unrelated users or base-stations will be interfered by the transmission of the current cell [38]. Each cell considers other cells in the interference range as its neighbors. In this architecture, each base-station is assumed to be aware of the information of its neighbors, such as distance, radius, and location of the base-station. Furthermore. it is capable of communicating with its neighbor base-stations.

For the bi-directional communication, we assume CR networks use a time-division duplex (TDD), which has been adopted in an IEEE 802.22 [32]. Thus, CR networks have separate time frames for uplink and downlink transmissions in the same spectrum band. Furthermore, each base-station j has the transmission power budget $P^{\text{tot}}(j)$ that can be allocated over its spectrum bands. Another important architectural issue in CR networks is how to establish a control channel. The control channel plays an important role in exchanging information regarding sensing and resource allocation. Several methods are presented in [9], one of which is assumed to be used as the common control channel in our proposed method.

5.3.2 Primary Network Model

All spectrum bands that CR networks can access are assumed to be licensed to different primary networks. We assume that the PU activity of spectrum i at PU activity region m can be modeled as a two state birth-death process with death rate $\alpha_{i,m}$ and birth rate $\beta_{i,m}$. An ON (Busy) state represents the period used by PUs and an OFF (Idle) state represents the unused period [44] [15] [46] [78]. Since each user arrival is independent, each transition follows the Poisson arrival process. Thus, the length of ON and OFF periods are exponentially distributed [66]. Based on this model, busy and idle probabilities of the spectrum i can be obtained as follows:

$$p_{i,m}^{\text{on}} = \frac{\beta_{i,m}}{\alpha_{i,m} + \beta_{i,m}}, \quad p_{i,m}^{\text{off}} = \frac{\alpha_{i,m}}{\alpha_{i,m} + \beta_{i,m}} \tag{53}$$

Here we assume that the CR network has M available licensed bands and is already aware of PU activities. Furthermore, each spectrum band can have multiple PU activities according to the location as illustrated in Figure 33.

5.4 Inter-Cell Spectrum Sharing Framework

5.4.1 Overview

As explained in Section 5.2.2, infrastructure-based CR networks are required to provide two different types of spectrum sharing schemes: *intra-spectrum sharing* and *inter-spectrum sharing*. To share spectrum efficiently, CR networks necessitate a unified framework to coordinate inter-cell and intra-cell spectrum sharing schemes and other spectrum management functions. Figure 34 shows the proposed framework for spectrum sharing in infrastructure-based CR networks, which consists of *event monitoring*, *inter-cell spectrum sharing*, and *intra-cell spectrum sharing*.

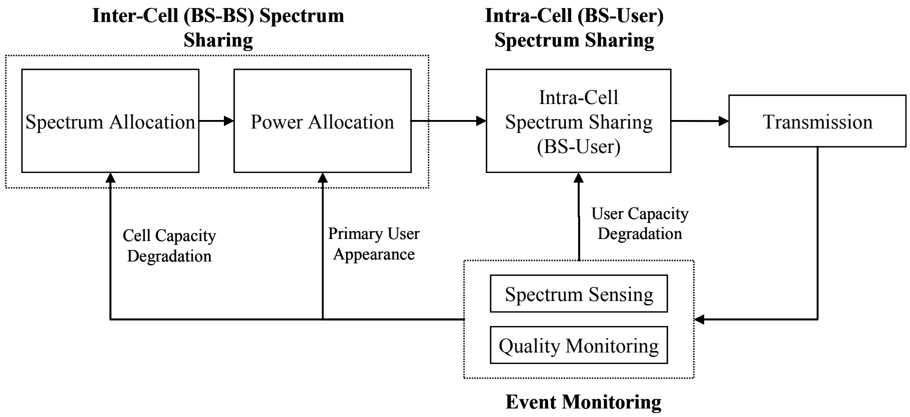


Figure 34: Spectrum sharing framework.

5.4.1.1 Event Monitoring

Event monitoring has two subfunctions. One is to detect the PU activities, called *spectrum sensing*. Here we consider a periodic sensing scheme that requires separate time slots for sensing and transmission [44]. In addition, CR users monitor the quality-of-service (QoS) of their transmission. According to the detected event type, the base-station determines spectrum sharing strategies and allocates the spectrums to each user adaptively based on the radio environments.

5.4.1.2 Inter-Cell Spectrum Sharing

In the proposed framework, inter-cell spectrum sharing is comprised of two subfunctions: *spectrum allocation* and *power allocation*. When the service quality of the cell becomes worse or is below the guaranteed level, the base-station initiates inter-cell spectrum sharing and adjusts its spectrum allocation. In the *spectrum allocation*, the base-station determines its spectrum band by considering the geographical information of primary networks and current radio activities. Here, inter-cell spectrum sharing exploits both *exclusive allocation* and *common use* approaches adaptively based on time-varying radio environment. After that, the base-station performs *power allocation* by determining the transmission power of its assigned spectrum bands to maximize cell capacity without interference to the primary network.

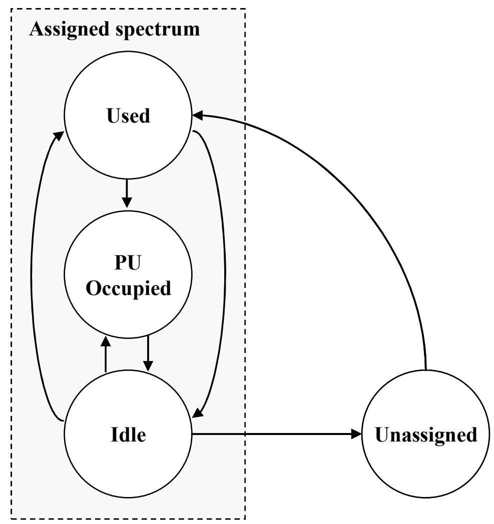
5.4.1.3 Intra-Cell Spectrum Sharing

Intra-cell spectrum sharing enables the base-station to avoid interference to the primary networks as well as to maintain the QoS of its CR users by allocating spectrum adaptively dependent on the event detected inside its coverage area. In this chapter, the proposed sharing scheme mainly focuses on the influence of neighbor cells and spectrum switching overhead. Also, intra-cell spectrum sharing necessitates a CR MAC protocol that allows multiple CR users to access to the same spectrum band. However this functionality has been widely investigated before [32] [78] [13] [69], and is out of scope in this chapter.

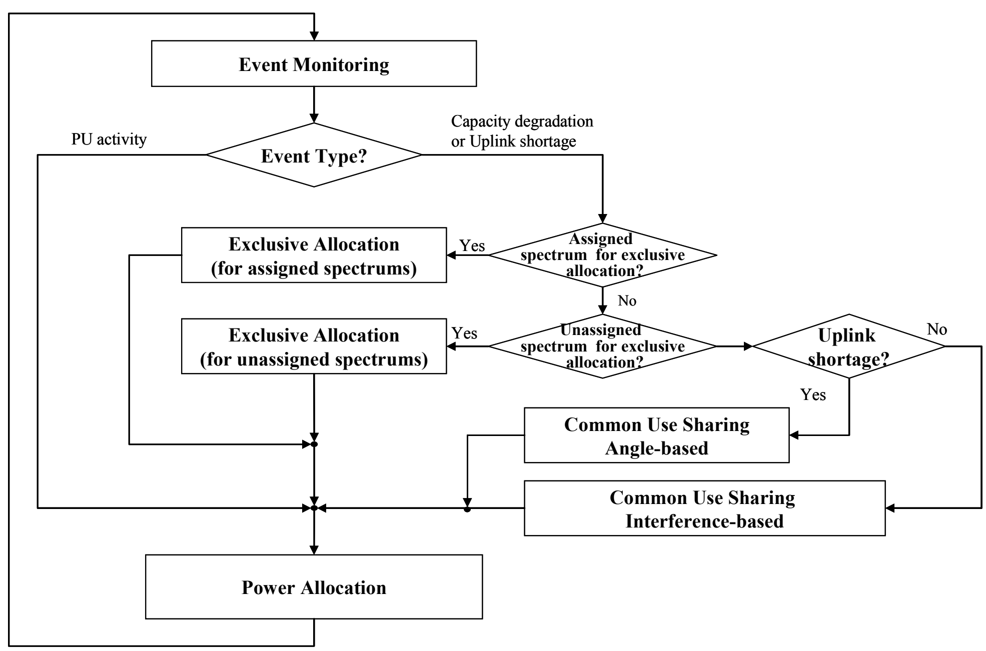
5.4.2 Spectrum Sharing Procedures

In spectrum sharing, it is desirable that spectrum allocation in the current cell has less influence on the transmission of other cells. In the worst case, spectrum allocation may lead to capacity degradation because of frequent interruption with the transmission of neighbors. Therefore, inter-cell spectrum sharing necessitates a coordination mechanism to reduce unnecessary influence on the entire networks.

To this end, we classify spectrum bands as the *assigned spectrum* and the *unassigned spectrum*. Figure 35 (a) shows the state-diagram for inter-cell spectrum sharing. The assigned spectrum bands are allowed to be accessed by the current cell while the unassigned bands are assigned to other neighbor cells. The assigned spectrum can have three sub-states, *used*, *PU occupied*, and *idle* according to its utilization. In *used* and *PU occupied* states, the spectrum is used by the current cell and by the primary network, respectively. The spectrum in *idle* state has been assigned to the cell but is not currently used. The assigned spectrum can be released to the unassigned only when it is currently idle and is requested by neighbor cells for their exclusive allocation. This helps to allocate spectrum among multiple neighbor cells without negotiation procedures that are generally required in conventional local



(a)



(b)

Figure 35: Spectrum sharing procedures: (a) spectrum status diagram, and (b) flow chart.

bargaining scheme [10]. Therefore, each base-station can determine its spectrum allocation immediately by considering information broadcasted from its neighbors, which significantly reduces communication overhead significantly.

Based on these states, we develop the following procedures for spectrum sharing.

Figure 35 (b) describes the flowchart of the proposed spectrum sharing scheme. Each cell continuously monitors the network status and radio environment through the local observations of its users. When one of the following events is detected, the base-station initiates the spectrum sharing procedure: 1) primary user appearance and 2) capacity degradation. If the detected event is related only to the PU activities in the coverage, the base-station turns off its transmission power on that spectrum. In case of PU activities in its neighbor cells, the base-station adjusts its transmission power not to violate the interference constraints. When the base-station detects the quality degradation in the cell, it performs exclusive allocation where the base-station first considers its assigned spectrum bands. Here it considers spectrum bands that that are not used by any primary users or neighbor cells in its interference range. If the base-station cannot find a proper one, it extends its search to the unassigned spectrum bands. If there is no idle spectrum band available for exclusive allocation, it switches to the common use sharing and performs the interference-based allocation method. In this case, the base-station can choose any available spectrum regardless of its state. Instead, the proposed common use allocation provides the capability to select the less influencing spectrum band, which is explained in Section 5.6. If the quality degradation is due to a resource shortage for uplink transmission, it selects the spectrum having the proper idle angle through the angle-based allocation. Once spectrum is allocated, each base-station allocates the transmission power over the assigned spectrum bands by considering the total power budget and transmission power constraints derived from spectrum allocation.

5.4.3 Distributed Spectrum Sharing Capability

Since each cell has a base-station, intra-cell spectrum sharing can be implemented in a centralized manner. However, it is not practical to develop inter-cell spectrum sharing as a centralized method because of the scalability problem. Furthermore, an additional network entity is required to coordinate resource allocation over the entire network. Instead, we introduce a distributed method for inter-cell spectrum sharing.

For distributed sharing operations, each cell should be aware of the exact information of its neighbor cells. As mentioned in Section 5.2.2, however, it is not feasible for the CR user to obtain all the necessary information from other users located in the other cells. Instead, in the proposed method, the base-station exchanges local cell information with its neighbor base-stations through the broadcast messages, called the distributed spectrum sharing messages (DSSMs). This message is assumed to be exchanged through the dedicated control channel or the backbone network connecting each base-station.

Each base-station broadcasts the DSSM to its neighbor cells periodically. Any conventional distributed MAC protocols can be used for the transmission of DSSMs. The DSSM can be used not only to exchange sensing and interference information but also to announce the initiation of inter-cell spectrum sharing. Once receiving the DSSM with spectrum sharing initiation, the cell prohibits itself and its neighbors from initiating another inter-cell spectrum sharing procedure until it receives the sharing completion message. These operations enable the conflict-free sharing in the multi-cell environment. The following information are conveyed through the DSSMs:

- *Spectrum availability:* The base-station determines the availability of all spectrum bands in its transmission range by considering the sensing information of its users, and broadcasts this availability to its neighbor base-stations to protect PU activities from the transmission of neighbor cells.

- *Spectrum utilization:* To take into account the influence of inter-cell interference, the base-station needs to have information regarding which spectrum is currently used by its neighbor cells. This, current spectrum utilization should be exchanged with each base-station for its spectrum allocation.

Table 3: Symbols used for the analytical modeling in spectrum sharing.

Symbols	Descriptions
N	Total number of spectrum bands
W_i	Bandwidth of spectrum i
$P^{\text{tot}}(j)$	Total transmission power budget of CR cell j
$P_i^{\max}(j)$	Maximum permissible transmission power of CR cell j at spectrum i
$P_{i,m}^{\text{pu}}(j)$	PU restricted power of CR cell cell j at spectrum i and region m
$P_i(j)$	Transmission power assigned to spectrum i in cell j
$R(j)$	Radius of CR cell j
$P_i^{\text{off}}(m)$	Idle probability of spectrum i at region m
P_{temp}	Interference temperature of spectrum i
$I_i^{\max}(j)$	Maximum cell interference measured at spectrum i of cell j
$I_i^{\min}(j)$	Minimum cell interference measured at spectrum i of cell j
$D(j, j^*)$	Distance between base-stations of cell j and j^*
$d_j(j^*, k)$	Distance between base-stations of cell j^* and user k in cell j

- *Minimum and Maximum cell interferences, $I_i^{\min}(j)$ and $I_i^{\max}(j)$:* Local information measured in the neighbor cells is essential to estimate the influence on the neighbor cells when a current cell uses a certain spectrum. However, it is not practical to exchange all local information with its neighbors. To reduce communication overhead, we use two representative information among all sensing results. The base-station j sends both minimum and maximum signal strengthes among all sensing data observed in its users, $I_i^{\min}(j)$ and $I_i^{\max}(j)$. These information includes the interference from other CR neighbors and noise. If the current cell detect the PU activity, they contains the primary signal strength as well.

Above information are used for spectrum and power allocations which will be explained in Sections 5.5, 5.6, and 5.7. To simplify the representation, the important symbols used in the subsequent discussion are summarized in Table 3.

5.5 Spectrum Allocation for an Exclusive Model

In wireless communications, the interference range is known to be larger than the transmission range [38]. Thus, for the interference-free communications, spectrum band needs to be allocated exclusively to each cell not to be overlapped with the

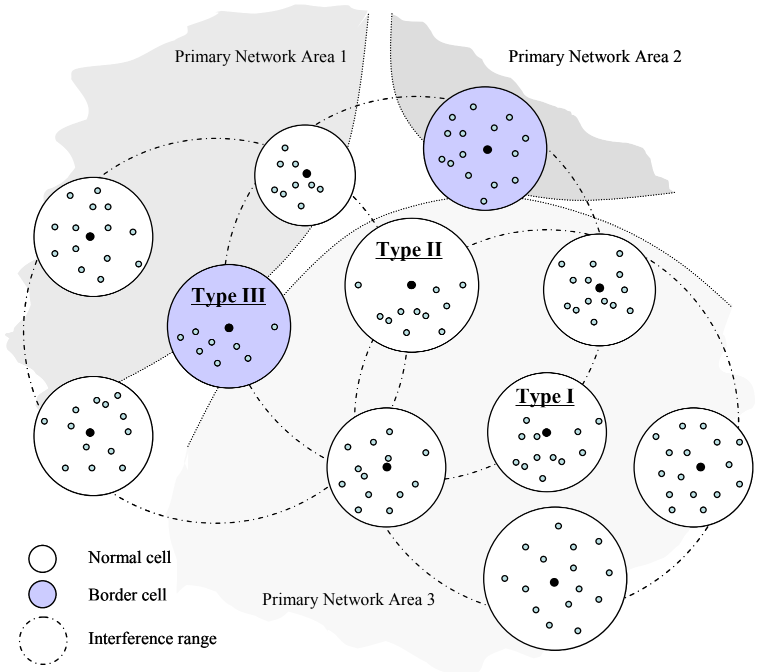


Figure 36: Cell characterization.

spectrums of its neighbors, which is the traditional approach to avoid interference in cellular networks. Furthermore, exclusive allocation needs to consider the spectrum with no PU activities in neighbor cells. However, as described in Section 5.2, the traditional exclusive approach is not suitable to CR networks, leading to inefficient and unfair spectrum utilization. To solve these problems, we propose a novel spectrum allocation to improve the spectrum availability in the exclusive model by considering the permissible transmission power derived from spatio-temporal characteristics of the PU activity.

5.5.1 Cell Characterization

As we explained in Section 5.2, even in the same spectrum band, PU activities show different characteristics according to the location. Based on this spatial characteristic of PU activities, the cells in CR networks can be classified as a *normal cell* and a *border cell*. While the normal cell has the same PU activity region inside its transmission range, the border cell is placed at the border of the PU activity region, and hence may have multiple PU activities, as shown in Figure 36. In the proposed method, according to the types of cells in the interference range, we classify three different scenarios for exclusive spectrum allocation as follows:

5.5.1.1 Type I. Homogeneous PU activity in the interference range

All cells in the interference range are placed in the same PU activity region, i.e., the spectrum availability is identical over the current cell and all its neighbors, as shown in Figure 36. If no primary user is detected in Type I cell j, the base-station can transmit on spectrum i with the maximum power $P_i^{\max}(j)$. Otherwise the transmission power is zero. Thus, the probabilities $\Pr(\cdot)$ of both cases can be derived as follows:

$$\begin{aligned} &\Pr(P_i^{\max}(j)) = p_{i,\widetilde{m}}^{\text{off}} \\ &\Pr(0) = 1 - p_{i,\widetilde{m}}^{\text{off}} \end{aligned} \tag{54}$$

where $\widetilde{m}$ is the PU activity region in the interference range.

5.5.1.2 Type II. Heterogeneous PU activity in the interference range

Some of the neighbors are border cells or located in the different PU activity regions, which can restrict the transmission power of the current cell even though no PU activities are detected in its transmission range. If the cell j has n PU activity regions in its interference range, it can have n different permissible transmission powers including one maximum transmission power $P_i^{\max}(j)$ and $n-1$ powers $P_{i,m}^{\text{pu}}(j)$ restricted by a region m.

Let $\mathcal{A}_i(j)$ be a set of PU activity regions in the interference range of cell j at spectrum i, and $\widetilde{m}$ be the region in the transmission range. In Type II, its transmission power can be $P_i^{\max}(j)$ when all PU activity regions in $\mathcal{A}_i(j)$ are idle. If one of the regions is busy, i.e., not available because of PU appearance, this region restricts the transmission power of current base-station. If multiple PU activities are detected in the interference range at the same time, transmission power is determined by the region which allows the current cell to use the smallest transmission power not to violate the interference constraint of other regions (more details are explained in Section 5.5.2). Here, we define dominant regions $\mathcal{A}^*_{i,m}(j)$ as a set of PU activity regions that allow smaller transmission power of cell j at spectrum i than region m when primary users are detected in all regions. Please note that $\widetilde{m}$ is not included in $\mathcal{A}^*_{i,m}(j)$. Then, the probabilities of each permissible transmission power can be determined as follows:

$$\begin{aligned}
&\Pr(P_i^{\max}(j)) = \prod_{m \in \mathcal{A}_i(j)} p^{\text{off}}_{i,m} \\
&\Pr(P^{\text{pu}}_{i,m}(j)) = p^{\text{off}}_{i,\widetilde{m}} \cdot (1 - p^{\text{off}}_{i,m}) \cdot \prod_{m^* \in \mathcal{A}^*_{i,m}(j)} p^{\text{off}}_{i,m^*} \\
&\qquad\qquad m \in \mathcal{A}_i(j),\ m \neq \widetilde{m} \\
&\Pr(0) = 1 - p^{\text{off}}_{i,\widetilde{m}}
\end{aligned} \tag{55}$$

As explained in Eq. (55), region m can determine the transmission power of current cell j, $P^{\text{pu}}_{i,m}(j)$ only when all dominant regions in $\mathcal{A}^*_{i,m}(j)$ are idle. In this case, the state of non-dominant regions does not affect the transmission power.

5.5.1.3 *Type III. Heterogeneous PU activity in the transmission range*

The cell is placed at the border of region with multiple PU activities. The probability of $P_i^{\max}(j)$ is the same as that of Type II. Let $\widetilde{\mathcal{A}}_i(j)$ be a set of PU activity regions in the transmission range. Then the probabilities of $P^{\text{pu}}_{i,m}(j)$ and zero power can be derived as follows:

$$\begin{aligned}
&\Pr(P^{\text{pu}}_{i,m}(j)) = (1 - p^{\text{off}}_{i,m}) \cdot \prod_{\widetilde{m} \in \widetilde{\mathcal{A}}_i(j)} p^{\text{off}}_{i,\widetilde{m}} \cdot \prod_{m^* \in \mathcal{A}^*_{i,m}(j)} p^{\text{off}}_{i,m^*} \\
&\qquad\qquad m \in \mathcal{A}_i(j) - \widetilde{\mathcal{A}}_i(j) \\
&\Pr(0) = 1 - \prod_{\widetilde{m} \in \widetilde{\mathcal{A}}_i(j)} p^{\text{off}}_{i,\widetilde{m}}
\end{aligned} \tag{56}$$

In this case, the CR network can use the spectrum only when all regions in the transmission range are idle. In the following subsection, we present how to determine the permissible transmission powers.

5.5.2 Permissible Transmission Power

For efficient spectrum allocation, the CR base-station should be aware of the permissible transmission power at each spectrum preventing interference to primary networks. Optimally, the permissible transmission power should be determined based on channel gains and the interference information of all possible link to users in current cell and its neighbors, which requires a significant amount information exchanges among base-stations and CR users. Instead, in the proposed method, the CR base-station estimates its permissible transmission power by considering representative information of each neighbor, i.e., maximum interference $I_i^{\max}(j)$ and the minimum interference $I_i^{\min}(j)$ conveyed in the DSSMs.

5.5.2.1 Maximum Transmission Power

When no PU activity is detected in any neighbors of the cell j at spectrum i, the maximum transmission power $P_i^{\max}(j)$ can be used in the cell. Let $\mathcal{F}(.)$ be a power propagation function that is determined by transmission power and the distance between a transmitter and a receiver. Then, the maximum transmission power $P_i^{\max}(j)$ can be obtained as follows:

$$I_\Delta(j^*) = P_{\text{temp}} W_i - I_i^{\max}(j^*), \quad j^* \in \mathcal{N}(j) \tag{57}$$

$$P_i^{\max}(j, j^*) = \mathcal{F}^{-1}(I_\Delta(j^*), D(j, j^*) + R(j^*)) \tag{58}$$

$$P_i^{\max}(j) = \min_{j^* \in \mathcal{N}(j)} P_i^{\max}(j, j^*) \tag{59}$$

where $I_\Delta(j^*)$ is the available power for CR users at a neighbor cell j^*, $P_i^{\max}(j, j^*)$ is the possible transmission power of cell j derived from $I_\Delta(j^*)$, and $\mathcal{N}(j)$ is a set of neighbors of cell j. In the PU activity region, the total interference should be less than $P_{\text{temp}} \cdot W_i$. Here P_{temp} represents the interference temperature (dBm/Hz), which is the amount of interference plus noise that primary networks can tolerate. From this constraint, we can obtain the interference margin, $P_{\text{temp}}W_i - I_i^{\max}(j^*)$ available to CR networks. Since there is no interference source, such as the activities of either primary or CR users, within its interference range, this interference margin is highly probable to be measured at the farthest border of neighbor cell j^*. $D(j, j^*) + R(j^*)$ from the current base-station. From this, the maximum possible power $P_i^{\max}(j, j^*)$ can be derived as Eq (58). To satisfy the interference condition in all neighbor cells, the base-station chooses the minimum transmission power among all $P_i^{\max}(j, j^*)$ as shown in Eq. (59).

5.5.2.2 PU Restricted Transmission Power

In exclusive allocation, the base-station can use maximum transmission power $P_i^{\max}(j, j^*)$. However, current transmission power may change according to the future PU activity in the interference range, which needs to be also considered in exclusive allocation. In Types II and III, neighbor cells located in the border or different regions will be boundary of PU activity, which is is likely to be the nearest border of neighbor cell j^* from the current base-station, i.e., $D(j, j^*) - R(j^*)$ from the base-station. To avoid interference to primary networks in neighbor cells, the permissible transmission power can be determined so that the received power at the border of neighbor cell does not exceed available power $I_\Delta(j^*)$. This available power can be estimated by using $I_i^{\min}(j)$ in Eq. (57) instead of $I_i^{\max}(j)$ since the minimum cell interference is highly probable to be measured at this nearest border. Then, the restricted transmission power $P_{i,m}^{\text{pu}}(j)$ can be obtained as follows:

$$P_{i,m}^{\text{pu}}(j) = \min_{j^* \in \mathcal{N}_{i,m}(j)} \mathcal{F}^{-1}(I_\Delta(j^*), D(j, j^*) - R(j^*)) \tag{60}$$

where $\mathcal{N}_{i,m}(j)$ is a set of neighbors of cell j located at region m of spectrum i. Similar to $P_i^{\max}(j)$ in Eq. (59), the minimum transmission power needs to be chosen for $P_{i,m}^{\text{pu}}(j)$ not to violate the interference constraint in any neighbor cells. This procedure can be also used to estimate $P_i^{\max}(j)$ when $I_\Delta(j, j^*)$ in Eq. (57) is less than zero at any neighbors.

5.5.3 Spectrum Selection

Based on the cell characterization and the permissible power, the capacities of all available spectrum bands can be estimated for spectrum selection, referred to as *opportunistic cell capacity.* The opportunistic cell capacity $C_i(j)$ is defined as the capacity of spectrum i at the boundary of the transmission range of cell j, which represents the minimum capacity to be provided by the base-station. According to the cell type, the opportunistic capacity can be derived as follows:

Type I:

$$C_i(j) = W_i \log_2(1 + \frac{\mathcal{F}(P_i^{\max}(j), R(j))}{I_i^{\max}(j)}) \cdot \Pr(P_i^{\max}(j)) \quad (61)$$

Type II & III:

$$\begin{aligned} C_i(j) = W_i[&\log_2(1 + \frac{\mathcal{F}(P_i^{\max}(j), R(j))}{I_i^{\max}(j)}) \cdot \Pr(P_i^{\max}(j)) \\ &+ \sum_{m \in \mathcal{A}_i(j) - \widetilde{\mathcal{A}}} \log_2(1 + \frac{\mathcal{F}(P_{i,m}^{\mathrm{pu}}(j), R(j))}{I_i^{\max}(j)}) \cdot \Pr(P_{i,m}^{\mathrm{pu}}(j)] \end{aligned} \quad (62)$$

If the base-station has multiple available spectrum bands for exclusive allocation, it selects the one with the highest opportunistic capacity.

5.6 Spectrum Allocation for Common Use Model

Although exclusive allocation is known to be optimal in terms of total network capacity, it is not suitable to CR networks because of unfair resource allocation, as explained in Section 5.2. On the contrary, a common use approach allows each cell to share the same spectrum with its neighbor cells, which improves fairness but causes capacity degradation owing to the inter-cell interference. To mitigate this effect, the following issues should be considered in the common use approach:

- *Cell capacity maximization:* The common use approach aims at finding a spectrum to maximize cell capacity by exploiting spectrum bands adaptively dependent on PU activities.

- *Less inter-cell interference:* Spectrum allocation in the current cell may cause inter-cell interference on its neighbors, leading to the degradation of total network capacity. Thus, the spectrum needs to be selected to minimize an adverse influence on other cells.

- *Uplink transmission:* Unlike the downlink (from base-station to CR users), the uplink shows the different interference range according to the location of the users. Since the interference range of the uplink is extended much farther than that of downlink, the uplink transmission causes higher interference to the neighbor cells. Furthermore, the uplink transmission is highly probable to interfere with the PU activity detected in its neighbor cells.

To address these issues, we propose two different spectrum sharing schemes for the common use model, which is explained in the following subsections.

5.6.1 Angle-Based Allocation for Uplink Transmission

As explained above, the uplink transmission can cause more significant interference to the PU activities at its neighbor cells. Figure 37 shows PU idle and busy regions based on the location of its neighbors who detect PU activities. When CR users in the busy region begin to transmit, they interfere with the transmission of primary networks in its neighbor cells. The best way to reduce interference in uplink transmission is to use the spectrum that does not have any PU activities in neighbor cells. If the base-station cannot find this spectrum, alternatively it can exploit multiple spectrum bands to allow all directions to be covered with their idle regions, referred to as *angle-based allocation*.

Let $\Theta_i(j)$ be the range of angles for PU idle regions at cell j in spectrum i. Then, to avoid the resource shortage of uplink transmission, the cell should satisfy the following angle condition:

$$\Theta_i(j) = \{\theta | \text{no PU activity in the direction } \theta\} \tag{63}$$

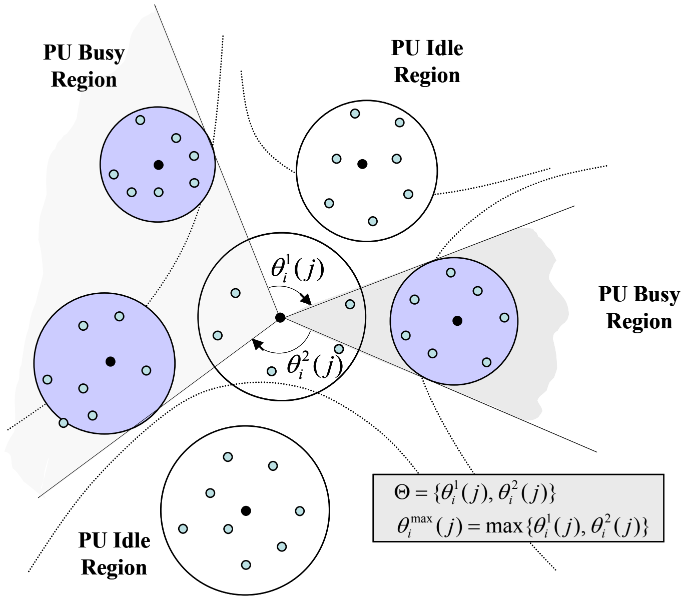


Figure 37: Busy and idle regions based on primary user activities.

$$\bigcup_{i \in \mathcal{S}(j)} \Theta_i(j) = \{\theta | 0 \leq \theta \leq 2\pi\} \tag{64}$$

where $\mathcal{S}(j)$ is the set of spectrums assigned to a cell j. The angle of the idle region can be estimated by the base-station, which is aware of the location information of its neighbors.

If the cell does not satisfy the angle condition for uplink transmission, the base-station initiates inter-cell spectrum sharing immediately and finds the proper spectrum band to satisfy the condition in Eq. (64).

5.6.2 Interference-Based Spectrum Allocation

If current spectrum allocation already meets the condition in Eq. (64), the CR network should consider the capacity maximization in terms of both cell and total network, as explained in the beginning of this section. This approach is closely related to the interference information in its neighbors, which determines transmission power in common use sharing.

5.6.2.1 Transmission Power in Common Use Sharing

In common use sharing, each neighbor cell may show different status according to the activities of both primary and CR users. Thus, the maximum permissible power restricted by each neighbor cell j^*, $P_i^{\lim}(j, j^*)$, can be obtained differently in the following three conditions:

- *Idle neighbor cells:* If there is no activity of either primary or CR users in neighbor cell j^*, the upper power limit $P_i^{\lim}(j, j^*)$ is the same as Eq. (58).

- *Neighbor cells with PU activities:* If a neighbor cell j^* detect the PU activity, $I_i^{\min}(j^*)$ can be considered as a reference interference to estimate $P_i^{\lim}(j)$, as explained in Section 5.5.2. In this case, $I_i^{\min}(j^*)$ includes the primary signal strength and interference components. Assume primary networks maintain the minimum signal-to-interference plus noise ratio (SINR) γ at their borders in the presence of interference temperature P_{temp}. The current interference at the border of the neighbor cell can be estimated as the difference between the measured signal strength and the expected primary signal strength, $I_i^{\min}(j^*) - \gamma \cdot P_{\text{temp}} \cdot W_i$. Then, the available power $I_\Delta(j^*)$ can be obtained by this current interference from maximum tolerable interference, $P_{\text{temp}} \cdot W_i$. With the similar procedure used in Eq. (60), $P_i^{\lim}(j, j^*)$ can be obtained as $\mathcal{F}^{-1}(I_\Delta(j^*), D(j, j^*) - R(j^*))$.

- *Neighbor cells in use:* If the neighbor cell j^* currently uses the spectrum i, we need to consider the transmission of cell j^* since $I_i^{\max}(j^*)$ does not contain its own signal strength. In this case, we can assume that the most portion of maximum interference $I_i^{\max}(j)$ measured in the current cell comes from cell j^*. From this, the transmission power of the neighbor cell can be estimated as $\mathcal{F}^{-1}(I_i^{\max}(j), D(j, j^*) - R(j))$. Then, the total interference at the farthest border of cell j^* from base-station j can be expressed as the sum of the interference from outside cells, $I_i^{\max}(j^*)$, and interference from its own transmission. Then, the available power at neighbor cell j^* can be obtained as follows:

$$\begin{aligned} I_\Delta(j^*) = P_{\text{temp}} \cdot W_i - [I_i^{\max}(j^*) \\ + \mathcal{F}(\mathcal{F}^{-1}(I_i^{\max}(j), D(j,j^*) - R(j)), R(j^*))] \end{aligned} \tag{65}$$

Then, $P_i^{\lim}(j,j^*)$ can be derived using Eq. (58).

Among all $P_i^{\lim}(j,j^*)$ obtained above, the base-station j chooses the minimum as an upper power limit of spectrum i, $P_i^{\lim}(j)$.

5.6.2.2 Selection Criterion

Based on the maximum permissible power, the principle of spectrum allocation is determined as follows: first, for maximum cell capacity, the cell should find the spectrum with the lowest interference in its transmission range, i.e., with the highest SINR. However, to maximize total network capacity, the cell needs to consider the influence on its neighbor cells when it determines a certain spectrum band. Optimally, we can allocate the spectrum to maximize total network capacity if each cell is aware of the channel gain of all possible links to the users both in neighbor and current cells as well as the interference at those users, which is not practical in infrastructure-based networks.

Instead, we propose a more practical and intuitive approach. Usually the cell with higher interference shows less influence on its capacity compared to one with lower interference, as shown in Figure 38. When new interference is added to the spectrum having low interference, it causes comparatively high capacity degradation. On the other hand, in case of the cell having higher interference, the degradation of capacity is relatively small even though additional interference is applied. If the capacity becomes below the threshold because of the additional interference, the base-station will initiate the inter-cell spectrum sharing procedure, and finally release the spectrum with low capacity, which helps to increase fairness in resource allocation as

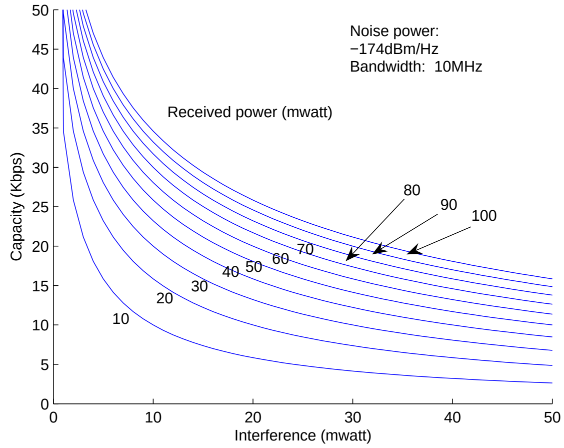


Figure 38: Capacity sensitivity to interference.

well as total network capacity. Thus, the cell needs to choose the spectrum bands with the highest interference in its neighbors.

From these observations, we devise the selection criterion to consider both cell capacity and total network capacity together, where the base-station j chooses the spectrum to maximize the product of SINR ratio of the product of the expected SINR of cell j, i.e., the ratio of maximum permissible power $P_i^{\text{lim}}(j)$ to its own interference $I_i^{\text{max}}(j)$, and interference in its neighbors $I_i^{\text{max}}(j^*)$.

Furthermore, even though the cell satisfies the condition for uplink transmission in Eq. (64), it is highly probable that primary users will re-appear in the assigned spectrum, which may violate the condition for uplink transmission. Thus, it is much advantageous for the cell to choose the spectrum with the widest idle angle range. By combining this idea with the above criterion, we can derive the following selection principle for the common use approach, called an *interference-based spectrum allocation*:

$$i^* = \arg\max_{i \in \mathcal{S}(j)} [\frac{\theta_i^{\text{mim}}(j)}{2\pi} \cdot \min_{j^* \in \mathcal{N}(j)} I_i^{\text{max}}(j^*) \cdot \frac{P_i^{\text{lim}}(j)}{I_i^{\text{max}}(j)}] \tag{66}$$

where $\theta_i^{\text{max}}(j)$ is the maximum idle angle in spectrum i at cell j, as shown in Figure 37. $\mathcal{S}(j)$ is a set of available spectrum bands at base-station j. Here, to minimize the influence on all neighbors, the proposed selection criterion exploit the lowest $I_i^{\max}(j^*)$ among the interference measured in neighbors. If the neighbor cell is in the busy PU activity region, we can use $I_i^{\min}(j^*) - \gamma \cdot P_{\text{temp}} \cdot W_i$ instead of $I_i^{\text{mim}}(j^*)$ as explained above.

5.7 Power Allocation for Inter-Cell Spectrum Sharing

In the previous sections, we introduced two different spectrum allocation methods for inter-cell spectrum sharing. Furthermore, they should have a capability to determine the transmission power of each base-station over the allocated spectrum bands to maximize cell capacity as well as to satisfy interference constraints. In this section, we propose a power allocation scheme, combining with spectrum allocation methods presented in the previous sections.

5.7.1 Upper Limits for Transmission Power

Once the spectrum is selected, the base-station determines its downlink transmission power over all currently used spectrum bands. Generally, a water filling scheme is used to optimally allocate power resource in the presence of noise. In this method, capacity is maximized when the sum of transmission power and interference are same over all frequencies in the spectrum bands [64]. However, unlike the general water filling method, power allocation in inter-cell spectrum sharing needs to account for the upper limits of transmission power in allocated spectrums, which depend on the PU activities and spectrum utilization in the vicinity of the current cell. This upper limit is exactly same as $P_i^{\text{lim}}(j)$ that is derived in Section 5.6.2. In case of exclusive allocation, $P_i^{\max}(j)$ in Eq. (59) can be used for the upper limit

5.7.2 Constrained Water Filling Method

Based on the constraints of all assigned spectrums in $\mathcal{S}(j)$, we introduce a constrained water filling method for power allocation, as shown in Algorithm 1.

Algorithm 1 Constrained Water Filling Method.

$P^{\mathrm{r}} = P^{\mathrm{tot}}(j)$, $\mathcal{S}^{\mathrm{r}}(j) = \mathcal{S}(j)$, $P_i^{\mathrm{dn}}(j) = 0, \forall i \in \mathcal{S}(j)$
$L_i^{\max} = I_i^{\max}(j)/W_i, \forall i \in \mathcal{S}(j)$
while $P^{\mathrm{r}} > 0$ and $\mathcal{S}^{\mathrm{r}}(j) \neq \emptyset$ **do**
  $i^* = \arg\min_{i \in \mathcal{S}^{\mathrm{r}}(j)} [L_i^{\max}]$
  $L^{\mathrm{tar}} = \arg\min_{i \in \mathcal{S}^{\mathrm{r}}(j) - i^*} [L_i^{\max}]$
  $\mathcal{S}^{\mathrm{c}}(j) = \{i | (P_i^{\mathrm{dn}}(j) + I_i^{\max}(j))/W_i < L^{\mathrm{tar}}, i \in \mathcal{S}^{\mathrm{r}}(j)\}$
  $P^{\mathrm{c}}(j) = (L^{\mathrm{tar}} - (I_{i^*}^{\max}(j) + P_{i^*}^{\mathrm{dn}}(j))/W_{i^*} \cdot \sum_{i \in \mathcal{S}^{\mathrm{c}}(j)} W_i$
  if $P^{\mathrm{c}}(j) <= P^{\mathrm{r}}(j)$ **then**
    $P_i^{\mathrm{dn}}(j) = L^{\mathrm{tar}} \cdot W_i - I_i^{\max}(j), \forall i \in \mathcal{S}^{\mathrm{c}}(j)$
    if $L_{i^*}^{\max} == I_{i^*}^{\max}(j)/W_i$ **then**
      $L_{i^*}^{\max} = (I_{i^*}^{\max}(j) + P_{i^*}^{\mathrm{lim}})/W_i$
    else
      $\mathcal{S}^{\mathrm{r}}(j) = \mathcal{S}^{\mathrm{r}}(j) - \{i^*\}$
    end if
    $P^{\mathrm{r}}(j) = P^{\mathrm{r}}(j) - P^{\mathrm{c}}(j)$
  else
    $P_i(j) = (P^{\mathrm{c}}(j) - P^{\mathrm{r}}(j)) \cdot W_i / \sum_{i \in \mathcal{S}^{\mathrm{c}}(j)} W_i, \forall i \in \mathcal{S}^{\mathrm{c}}(j)$
    $P^{\mathrm{r}}(j) = 0$
  end if
end while

Since each spectrum has different upper limits, $P_i^{\mathrm{lim}}(j)$, and interference levels, $I_i^{\max}(j)$, this method uses the iterative algorithm to achieve optimal power allocation. Here $\mathcal{S}^{\mathrm{c}}(j)$ represents a set of candidate spectrum bands selected for the current water filling stage, $\mathcal{S}^{\mathrm{r}}(j)$ is a set of remaining spectrums. $P^{\mathrm{r}}(j)$ and $P^{\mathrm{c}}(j)$ are a remaining power budget and the total power required for candidate spectrum bands in the current stage, respectively. L^{tar} is the expected target power level (mwatt/Hz) in the current water filling operation stage, and $L_i^{\max}$ is the maximum level of each spectrum i (mwatt/Hz) that is the normalized sum of the upper power limit $P_i^{\mathrm{lim}}(j)$ and the interference $I_i^{\max}(j)$ for a given bandwidth of spectrum i.

As explained in Algorithm 1, first, the lowest target level L^{tar} among $L_i^{\max}$ is determined in each stage, and then the transmission power is allocated to each candidate spectrum in $\mathcal{S}^{\mathrm{c}}(j)$ to satisfy the target level L^{tar}. These water filling operations are performed until either there is no power budget left or all available spectrums reach their upper power limits. In the constrained water filling method, the allocated downlink power $P_i^{\mathrm{dn}}(j)$ cannot exceed the upper limit of the transmission power, $P_i^{\mathrm{lim}}(j)$, which enables interference avoidance with primary networks while maximizing the cell capacity.

5.8 Intra-Cell Spectrum Sharing

Once the base-station determines spectrums and their corresponding transmission powers, it needs to allocate communication resources to its CR users, referred to as an *intra-cell spectrum sharing* (or inter-user spectrum sharing). Since each cell can exploit multiple spectrum bands in the proposed method, this functionality is classified into *inter-spectrum sharing* and *intra-spectrum sharing* according to the scope of sharing operations.

In inter-spectrum sharing, the base-station assigns its CR users to the proper spectrum. This operation is required when 1) the PU activity is detected, 2) a new spectrum is obtained through inter-cell spectrum sharing, and 3) transmission power is adjusted because of the PU activities detected in neighbor cells. Inter-spectrum sharing generally aims at maximizing total network capacity while maintaining fairness in resource allocation. However, optimal spectrum allocation is known as an NP hard problem, causing a high computational complexity [56]. To overcome this shortcoming, a heuristic graph coloring method has been proposed in [56], which is based on only exclusive allocation. In [46], a QoS-aware spectrum decision scheme is proposed to determine the proper spectrum band for both real-time and best effort applications. However, this method focuses on the single-cell network, and does not consider the influence of neighbor cells. In this chapter, we extend these solutions to the multi-cell network by considering additional characteristics as follows:

- In infrastructure-based CR networks, uplink and downlink transmissions show different interference characteristics. Especially, uplink communications are highly probable to interfere with the PU transmission detected in the neighbor cells, as explained in Section 5.6. To avoid interference violation, each user has different constraints on uplink transmission power according to its location and PU activities in its neighbors.

- Since the cell may have discontiguous spectrum bands distributed over a wide frequency range, CR users need relatively long spectrum switching latency to reconfigure their RF frontend as well as to re-establish communication channels. Hence, spectrum switching causes an adverse influence on network performance resulting from the temporary disconnection of their transmissions.

After spectrum assignment through inter-spectrum sharing, the base-station performs intra-spectrum sharing to coordinate multiple access among CR users assigned to the same spectrum band. Here we assume that CR networks can adopt any conventional multiple access schemes for intra-spectrum sharing scheme, such as CDMA, time division multiple access (TDMA), or frequency division multiple access (FDMA). Thus, in this section, we mainly focus on the inter-spectrum sharing method.

5.8.1 User Capacity Model

To assign CR users to the proper spectrum bands, first, we need to evaluate their expected capacities over all available spectrum bands. In downlink channels, the transmission power of cell j at spectrum i, $P_i^{\mathrm{dn}}(j)$, is obtained through the constrained water filling method, as explained in Section 5.7. On the contrary, the uplink transmission power of user k, $P_i^{\mathrm{up}}(j,k)$ is determined by its base-station j, with the similar procedures explained in Section 5.7.1 as follows:

$$\begin{aligned} P_i^{\mathrm{up}}(j,k) = \min[&\min_{j^* \in \mathcal{N}_i^{\mathrm{on}}(j)} \mathcal{F}^{-1}(I_\Delta(j,j^*), d_j(j^*,k) - R(j^*)), \\ &\min_{j^* \in \mathcal{N}_i^{\mathrm{off}}(j)} \mathcal{F}^{-1}(I_\Delta(j,j^*), d_j(j^*,k) + R(j^*))] \end{aligned} \tag{67}$$

where $d_j(j^*, k)$ is the distance between the base-station of cell j^* and user k in cell j. $\mathcal{N}_i^{\text{on}}(j)$ is a set of cell j's neighbors where spectrum i is currently busy because of the PU activity. $\mathcal{N}_i^{\text{off}}(j)$ is a set of idle neighbors of cell j in spectrum i, where no PU activity is detected. $I_\Delta(j, j^*)$ can be obtained according to the status of neighbor cell j^*, which is explained in Sections 5.5.2 and 5.7.1. Here the base-station estimates the permissible uplink transmission powers, determined by the interference condition in each neighbor cell as well as user location, and chooses the lowest one as the transmission power for uplink channels at user k in spectrum i.

The spectrums, which have $I_\Delta(j^*)$ less than zero in any of neighbor cells, are not suitable for CR transmission, and hence not considered in this spectrum allocation.

From the transmission powers of the base-station and users obtained above, the expected spectrum capacities of both uplink and downlink channels at CR user k can be derived as follows:

$$C_i^{\text{up}}(j,k) = W_i \cdot (1-\epsilon) \cdot \rho \cdot \log_2(1 + \frac{\mathcal{F}(P_i^{\text{up}}(j,k), d_j(j,k))}{I_i^{\text{bs}}(j)}) \tag{68}$$

$$C_i^{\text{dn}}(j,k) = W_i \cdot \epsilon \cdot \rho \cdot \log_2(1 + \frac{\mathcal{F}(P_i^{\text{dn}}(j), d_j(j,k))}{I_i^{\text{usr}}(j,k)}) \tag{69}$$

where ϵ is the fraction of the downlink frame, $I_i^{\text{bs}}(j)$ is the interference plus noise at base-station j, $I_i^{\text{usr}}(j,k)$ is the interference plus noise measured at user k of cell j, and ρ is a scaling factor to describe the dynamic spectrum switching influence, which is expressed as follows [46]:

$$\rho(i^*, i) = \begin{cases} \dfrac{\frac{1}{\sum_{m \in \mathcal{A}_i(j)} \beta_{i,m}}}{\frac{1}{\sum_{m \in \mathcal{A}_i(j)} \beta_{i,m}} + \tau} & i \neq i^* \\ 1 & i = i^* \end{cases} \tag{70}$$

where i^* and i is current and new spectrum bands, respectively. τ is the switching latency from one spectrum to another, and $1/\sum_{m \in \mathcal{A}_i(j)} \beta_{i,m}$ is the average idle period of spectrum i. $\mathcal{A}_i(j)$ is a set of PU activity regions in cell j at spectrum i. If newly assigned spectrum i^* is different from the previous one i, switching latency is inevitable, leading to quality degradation.

5.8.2 Intra-Cell Spectrum Sharing Procedures

Based on these user capacities, the base-station j performs inter-spectrum sharing with the following procedures:

- *Step 1:* Let $\mathcal{U}(j)$ be a set of remaining CR users that are not assigned to the spectrum through this inter-spectrum sharing procedure. Initially, $\mathcal{U}(j)$ includes all CR user in cell j.

- *Step 2:* Each user in $\mathcal{U}(j)$ calculates the expected total capacity at all available spectrum bands. Here, we consider a proportional fairness as a principle of intra-cell spectrum sharing. Thus, when user k' is added to spectrum i', the metric of total capacity is expressed as a sum of logarithmic capacities of user k', and users already assigned to either spectrum i' or other spectrum bands as follows:

$$\begin{aligned}\mathcal{G}_{i'}(j,k') &= \log(\frac{C_{i'}^{\mathrm{up}}(j,k') + C_{i'}^{\mathrm{dn}}(j,k')}{n_{i'}+1}) \\ &+ \sum_{k\in\mathcal{K}_{i'}(j)} \log(\frac{C_{i'}^{\mathrm{up}}(j,k) + C_{i'}^{\mathrm{dn}}(j,k)}{n_{i'}+1}) \\ &+ \sum_{i\in\mathcal{S}(j),i\neq i'} \sum_{k\in\mathcal{K}_i(j)} \log(\frac{C_{i}^{\mathrm{up}}(j,k) + C_{i}^{\mathrm{dn}}(j,k)}{n_{i}})\end{aligned} \tag{71}$$

  where n_i is the number of CR users in spectrum i and $\mathcal{K}_i(j)$ represents a set of CR users in cell j that are assigned to spectrum i. Here we assume that communication resources, such as time-slot or bandwidth, are fairly assigned to CR users in the same spectrum through multiple access schemes.

- *Step 3:* Let $\mathcal{S}_{\mathrm{q}}(j,k)$ be a set of available spectrum bands to support the minimum QoS requirement of CR user k. If there are any spectrums not to satisfy the minimum QoS requirement, the base-station removes them from $\mathcal{S}_{\mathrm{q}}(j,k)$. If any of the current users shows lower capacity than the minimum requirement, this spectrum is also removed from the set.

- *Step 4:* Each user determines its most preferable spectrum band, called a *color*, which shows the highest capacity among all candidate spectrums in $\mathcal{S}_{\mathrm{q}}(j,k)$, called a *label*. Then the base-station selects the one with the highest label among the remaining CR users in $\mathcal{U}(j)$, and assigns it to its color. If there is any user m having only one spectrum in $\mathcal{S}_{\mathrm{q}}(j,k)$, the base-station assigns it to its available spectrum preemptively.

- *Step 5:* The selected user is removed from $\mathcal{U}(j)$.

- *Step 6:* Based on the updated spectrum allocation, the base-station repeats these procedures (Steps 2-5) until $\mathcal{U}(j)$ is empty.

If CR users have multiple transceivers to use different spectrums for uplink and downlink communications, the base-station performs the above operations for each link separately.

5.9 *Performance Evaluation*

In this section, we present simulation results on the performance of the proposed spectrum sharing method.

5.9.1 Simulation Setup

To evaluate the performance of the proposed sharing method, we implement the network simulator to support the network topology consisting of multiple cells in 10km x 10km area. Figure 39 shows the network topology used in the simulation. Here we assume 20 cells that have different number of users from 20 to 40. The transmission range of each cell is uniformly distributed from 1 to 1.5km. The interference range is set to twice larger than the transmission range. Furthermore, we consider 20

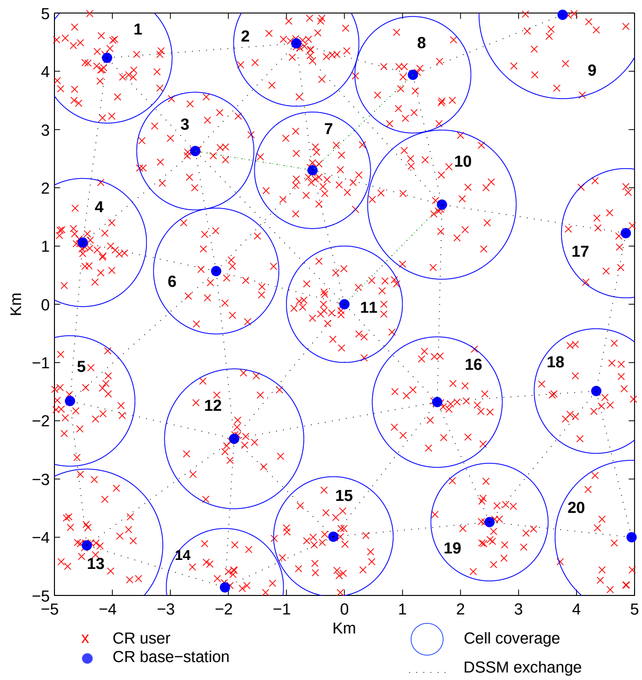


Figure 39: Network topology for simulation in spectrum sharing.

10MHz licensed spectrum bands with different PU activities, $\alpha_{i,m}$ and $\beta_{i,m}$, which are uniformly distributed in $[0.1, 0.5]$. Each spectrum band has 2-5 PU activity regions. Spectrum switching delay is set to 0.1sec.

In this simulation, we use a free space power attenuation model where the channel gain is set to -31.54dB, the reference distance is 1m, and the path loss coefficient is 3.5. The base-station has 1250mW transmission power in total and can allocate up to 250mW to each spectrum. The maximum uplink transmission power of the CR user is also set to 250mW. Noise power in the receiver is -174dBm/Hz. For the protection of primary networks, we set the interference temperature to 6dB greater than the noise power. The CR network uses the TDD with the same length of uplink and down link time slots. While base-stations can use the multiple spectrum bands at the same time, CR users can use only one spectrum for both uplink and down link transmissions.

To evaluate the proposed method, we use three different existing spectrum sharing methods as follows:

- *Fixed spectrum allocation:* Spectrum allocation can be obtained by the coloring method with a maximum proportional fairness criterion [56]. Here, each cell is assigned to the pre-determined spectrum bands and does not change them regardless of time-varying spectrum availability. Instead, this method considers the number of neighbor cells and PU activities.

- *Dynamic spectrum allocation:* This method is also based on the same coloring method used in the fixed allocation. However, in this method, spectrum allocation is dynamically updated over the entire network whenever spectrum availability changes.

- *Local bargaining:* In this method, each cell can negotiate with its neighbor to obtain spectrum bands when its capacity is below a threshold [10].

These existing methods use the maximum transmission power in the assigned spectrums. Also they adopt the same intra-cell spectrum sharing strategy used in the proposed method (Section 5.8), but do not consider the permissible transmission power and switching delay effect described in Eqs (67) and (70). In this simulation, we do not consider existing common use sharing methods since they are not suitable for the infrastructure-based networks as explained in Section 5.2.2.

5.9.2 Total Capacity

In Figure 40, we evaluate total network capacity for both downlink and uplink transmissions. Fixed allocation uses the graph-based optimization with a global topology knowledge, which leads to the highest downlink and uplink capacities among existing methods. Although dynamic allocation supports channel adaptation, it shows lower capacity than the fixed allocation since global optimization in every spectrum change results in frequent interruption of communications. In the bargaining method, each

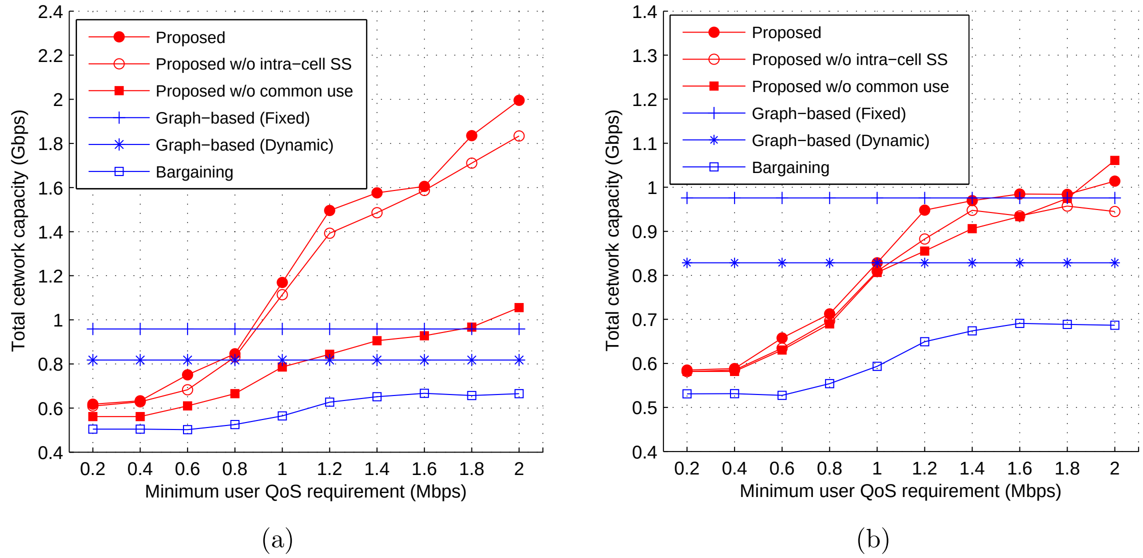


Figure 40: Performance comparison in total capacity: (a) total downlink capacity, and (b) total uplink capacity.

cell takes the spectrum band from other neighbor cells if it cannot satisfy the QoS. However, since this method cannot perform spectrum allocation if neighbors are currently involving in other bargaining process, it shows the lowest spectrum utilization. Although all these methods are based on exclusive allocation, the proposed method exploits both exclusive and common use models adaptively dependent on network environments, and hence achieves the highest downlink capacity in the limited spectrum requirements, which is shown in Figure 40 (a). On the contrary, since the uplink channel has more strict transmission power constraints, the dynamic mode adaptation scheme does not help to improve its total capacity as much as that of downlink. Thus, as shown in Figure 40 (b), the proposed method achieves slightly better performance in total capacity than fixed allocation in high QoS requirements.

Furthermore, we also evaluate the proposed method with two conditions: 1) without common use sharing and 2) without the switching delay factor in intra-spectrum sharing (Eq. (70). In these cases, their downlink capacities become lower than the original capacity, which shows that our exclusive allocation scheme can improve its

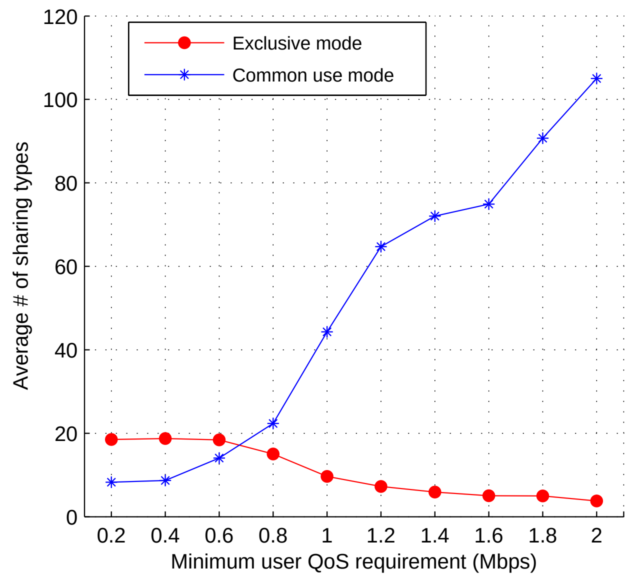


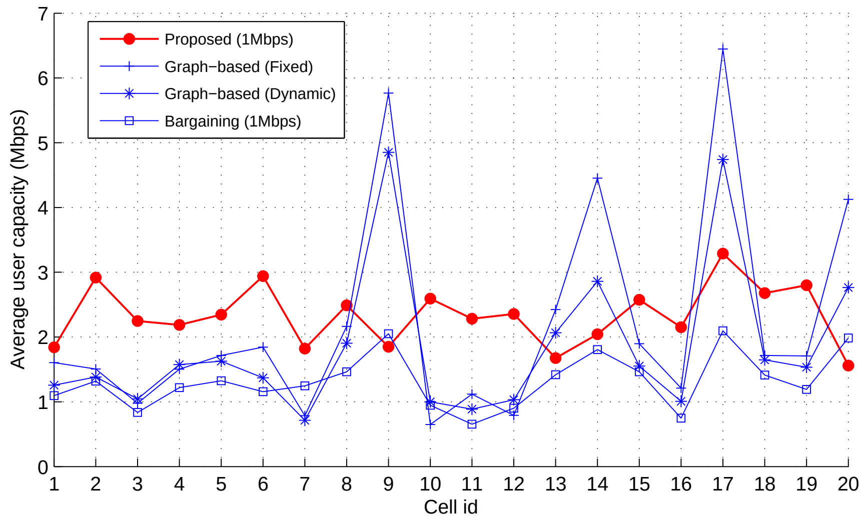
Figure 41: Spectrum sharing types.

total capacity by collaborating with the proposed common use sharing and intra-spectrum sharing schemes. Especially, common use sharing shows much higher influence on total downlink capacity, as shown in Figure 40 (a).

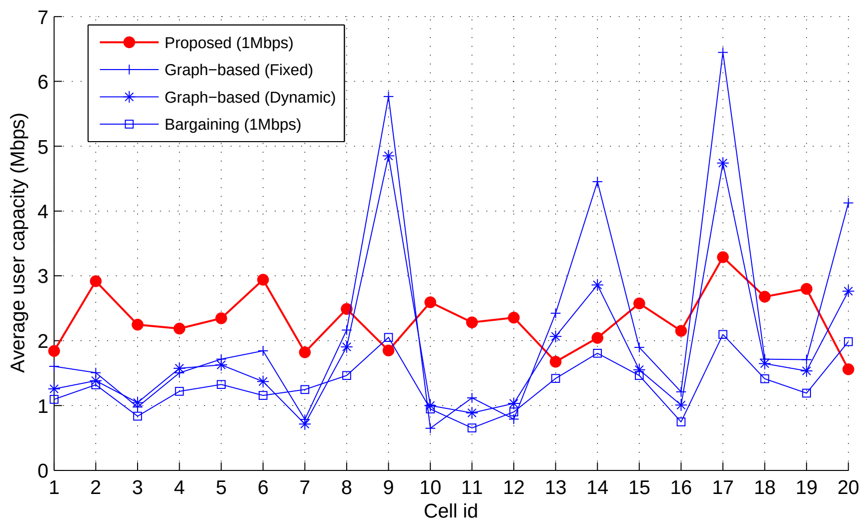
Figure 41 presents the dynamic mode selection in the proposed method. If the required capacity is relatively low, most of spectrum bands are used for exclusive allocation. As the QoS requirement increases, i.e., spectrum availability becomes lower, more spectrum bands are selected for common use sharing.

5.9.3 Fairness

Here, we investigate capacity fairness in both spatial and temporal domains, which are also important objectives in inter-cell spectrum sharing. As shown in Figures 42, both dynamic and fixed allocation methods show high capacity fluctuation over different locations since it does not have a QoS mechanism. Especially, cells #9 and #17 achieve much higher capacities in both uplink and downlink than other cells. However, through the dynamic mode selection, the proposed method maintains better fairness in cell capacity than other methods. While the bargaining method can also support



(a)



(b)

Figure 42: Performance comparison in fairness: (a) spatial fairness in downlink, and (b) spatial fairness in uplink.

capacity fairness over different locations through negotiation processes, it shows lower capacity than the proposed method over an entire network because of inefficient spectrum utilization.

Furthermore, the proposed method shows better performance in avoiding temporary resource starvation in a certain cell, as depicted in Figure 43. In existing methods, temporary resource starvation is inevitable since all of them are based on

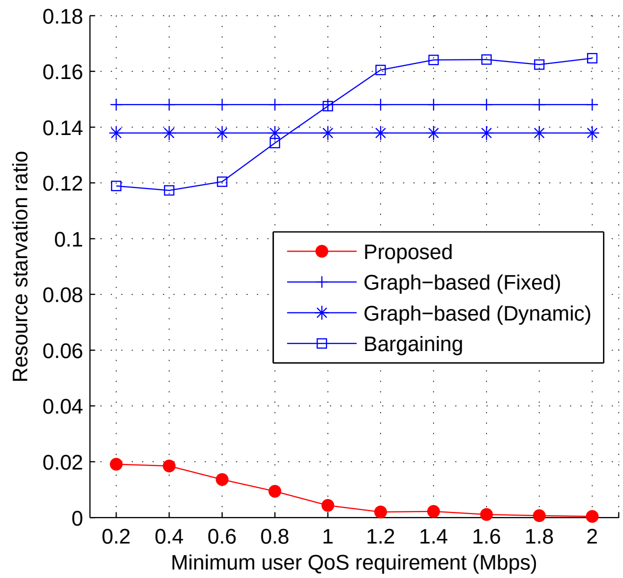


Figure 43: Average resource starvation ratio.

exclusive allocation. These methods may not have enough available spectrum bands according to spectrum utilization. Regardless of adaptation capability, dynamic allocation shows little higher starvation ratio than the fixed allocation because of frequent spectrum switching. As QoS requirements become higher in the bargaining method, spectrum utilization becomes higher, leading to increase in resource starvation ratio. However, the proposed method mitigates this temporary resource starvation by exploiting common use sharing adaptively dependent on spectrum utilization.

5.9.4 QoS and Complexity

In Figure 44, we observe the QoS violation ratio in both uplink and downlink transmissions according to user QoS requirements, which is defined as the fraction of time when the QoS of the cell is below the minimum requirement. Unlike the proposed method, the dynamic and fixed spectrum allocations do not have any QoS guarantee mechanism and just aim at maximizing total capacity and fairness under the exclusive sharing mode. Although they show higher uplink capacity in lower QoS requirements, as shown in Figure 40 (b), their violation ratios are much worse than

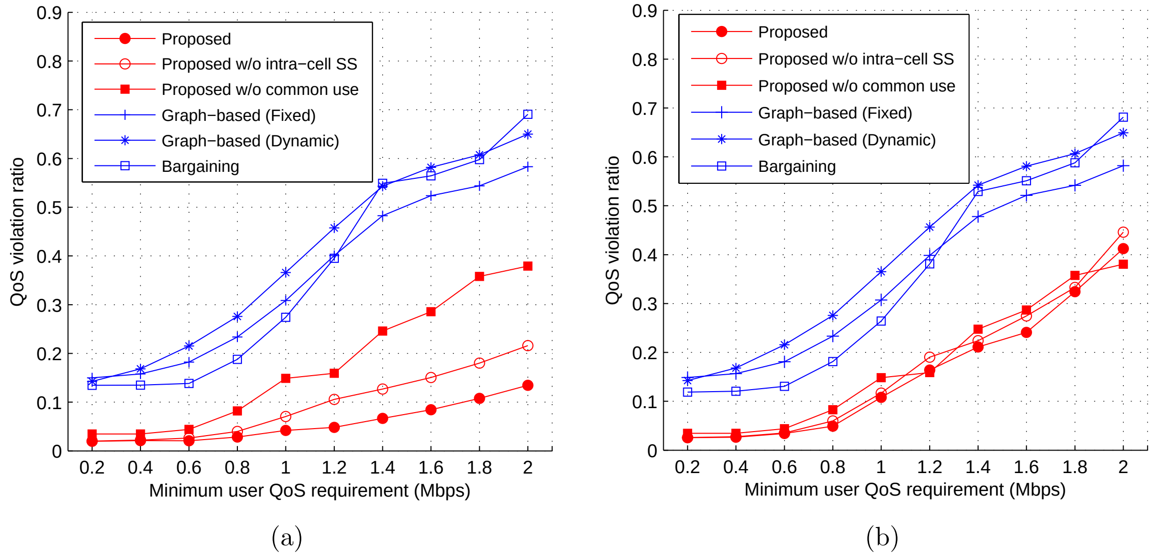


Figure 44: Performance comparison in QoS violation: (a) QoS violation ratio in downlink, and (b) QoS violation ratio in uplink.

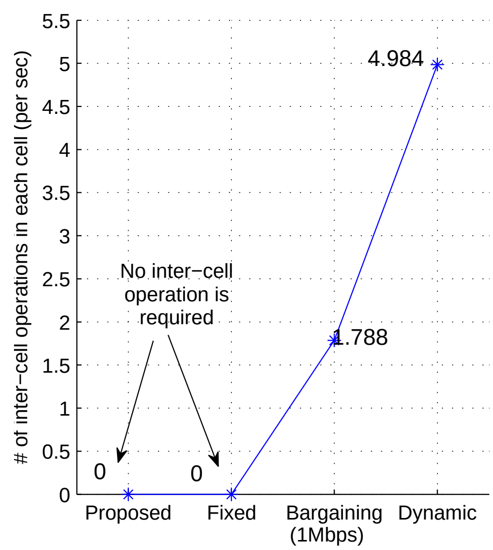


Figure 45: The average number of inter-cell opoerations of each cell.

the proposed method in all conditions. The bargaining method is based on classical exclusive allocation, and hence does not consider adaptive power allocation. Thus, while the bargaining method supports provides the QoS mechanism, it shows lower spectrum utilization, and hence a higher QoS violation ratio. If either common use sharing or switching delay influence is not considered in the proposed method, the

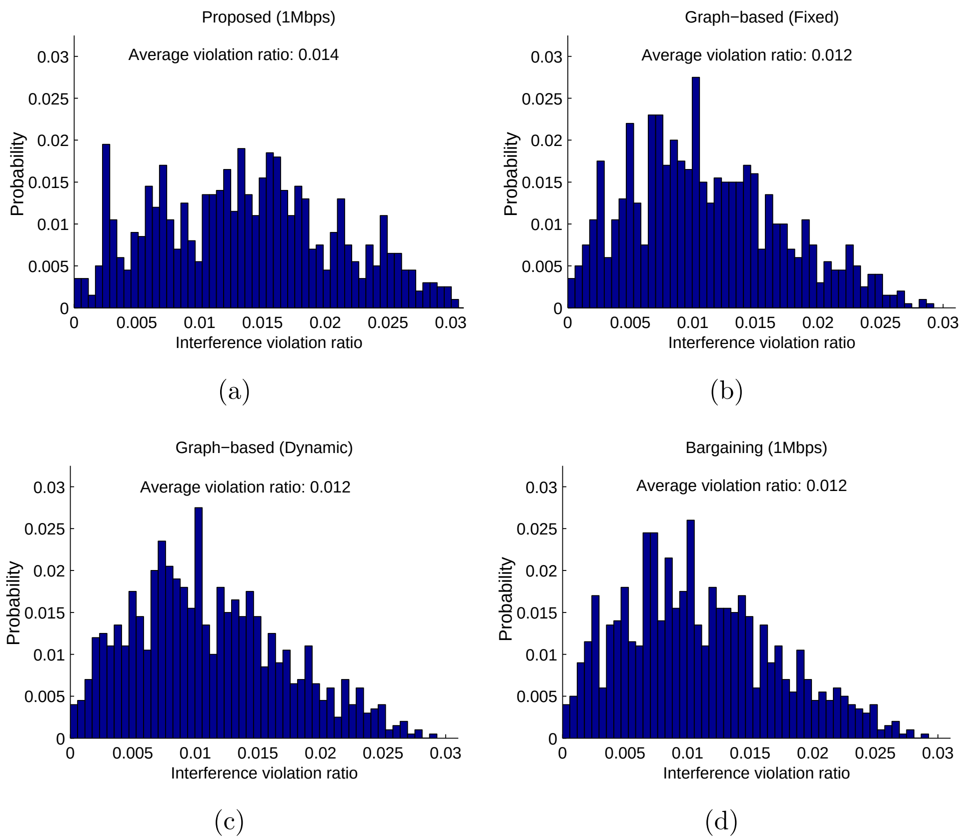


Figure 46: Histogram for interference violation ratio: (a) proposed method, (b) dynamic allocation, (c) fixed allocation, and (d) local bargaining.

violation ratio becomes higher than the original result, but is still significantly lower than that in classical methods. Similar to total capacity, the QoS violation ratio in downlink is much more sensitive to these two functions than that in uplink.

In Figure 45, we investigate the operational overhead for each method. Unlike the proposed method and fixed allocation, both dynamic allocation and bargaining methods require inter-cell operations for spectrum negotiations, which increase communication overhead significantly.

5.9.5 Interference Avoidance

Another crucial issue in CR networks is interference avoidance with primary networks, which has not been addressed in previous methods. Figure 46 shows an interference ratio under different sharing schemes, which is defined as the ratio of the area violating an interference temperature limit to the total area occupied by primary networks. As shown in Figure 46, our proposed method shows the similar interference ratio to other methods. To protect primary networks, existing methods do not allow CR cells to the spectrum where PU activity is detected by either current or any neighbor cells, which leads to inefficient spectrum utilization. However, since the proposed method flexibly determines the transmission power not to exceed the interference temperature, it achieves both higher capacity and better fairness while maintaining similar interference avoidance performance to other previous methods.

CHAPTER VI

SPECTRUM-AWARE MOBILITY MANAGEMENT IN CR CELLULAR NETWORKS

6.1 Introduction

Among these spectrum management functions, spectrum mobility imposes unique characteristics in mobility management for CR cellular networks. Mobility management, especially a handoff scheme, is one of the most important functions in classical cellular networks. Thus, much research on cellular networks have explored the handoff issues, mainly focusing on cell selection and resource management in the last couple of decades [5]. Although diverse cell selection methods have been proposed to support seamless handoff schemes while maximizing the network capacity [6] [28] [59] [61], they are based on the classical multi-cell based networks and do not consider the fluctuating nature of spectrum resources in CR networks. Especially, no special attention is given to either time and location-varying spectrum availability or switching delay in traversing the spectrum distributed over a wide frequency range.

The main difference between classical wireless networks and CR networks lies in the PU activities. Because of the PU activity, CR networks necessitate a new type of handoff, the so-called *spectrum handoff*, which also must be considered in designing mobility management schemes. Thus, mobility management constitutes an important but unexplored topic in CR networks to date. There are the following challenges:

- *Heterogeneous mobility events:* CR networks are required to provide two different types of handoff schemes: classical inter-cell handoff resultig from physical user mobility and spectrum handoff owing to spectrum mobility. Thus, it necessitates a unified mobility management framework to exploit different handoff types adaptively to mobility events.

- *Dynamic spectrum availability:* According to the PU activities, spectrum availability varies over time and space in the CR network, which makes it more difficult to provide seamless and reliable communication to mobile users traversing across multiple cells. For an efficient mobility management, CR networks need to mitigate this heterogeneous spectrum availability by performing mobility management adaptively dependent on the heterogeneous network conditions.

- *Broad range of available spectrum:* In CR networks, available spectrums are not contiguous and found over a wide frequency range. Thus, when CR users switch their spectrums, they need to reconfigure the operating frequency of the RF front-end to tune to the new spectrum band, leading to increase in switching delay. This delay is much longer than that in classical wireless networks.

To address the challenges mentioned above, we propose a spectrum-aware mobility management scheme for CR cellular networks [45]. First, we propose a novel CR cellular network architecture based on the spectrum-pooling concept, which mitigates the heterogeneity in spectrum availability. Based on this architecture, a unified mobility management framework is defined to support diverse mobility events in CR networks, consisting of inter-cell resource allocation, and spectrum and user mobility management functions. Through inter-cell resource allocation, each cell determines its spectrum configuration to improve the mobility as well as total capacity. To support spectrum mobility while maintaining maximum cell capacity, the spectrum mobility management is developed where both spectrum utilization and the stochastic connectivity model are exploited to determine the proper handoff types and target cells for CR users experiencing PU activities. In user mobility management, the switching cost-based handoff decision mechanism is proposed to minimize service quality degradation because of switching operations as well as to maximize cell capacity.

The rest of the chapter is organized as follows: Sections 6.2 and 6.3 present the proposed network architecture and mobility management framework for CR networks, respectively. Handoff models in the proposed framework are presented in Section 6.4 In Sections 6.5 and 6.6, a novel spectrum and user mobility management methods are proposed, respectively. Performance evaluation and simulation results are presented in Section 6.7.

6.2 The Proposed System Model

6.2.1 Network Architecture

In this chapter, we consider infrastructure-based CR networks consisting of multiple cells. Each cell has a single base-station (BS) and its CR users. In this architecture, CR users observe the radio environment and report the results to their BSs. Then, the BS determines the proper actions accordingly. Each BS can be controlled by the central network entities such as a base-station controller (BSC). This central entity is responsible for inter-cell resource allocation and mobility management. CR users have a single wideband RF transceiver that can cover an entire spectrum pool without reconfiguration. Thus, CR users can sense all spectrum bands in the pool at the same time. The spectrum pool will be explained in the following subsections. Each CR user m needs K_m channels to satisfy its QoS requirement.

All spectrum bands are assumed to be licensed to different primary networks. We assume that the PU activity of the spectrum can be modeled as a two state birth-death process [15] [66] [78]. An ON (Busy) state represents the period used by primary users and an OFF (Idle) state represents the unused period. Since each user arrival is independent, each transition follows the Poisson arrival process. Thus, the length of ON and OFF periods are exponentially distributed. Let PU departure and arrival rates at PU area k in spectrum j be $\alpha(j,k)$ and $\beta(j,k)$, respectively. Then, its idle probability, $P^{\text{off}}(j,k)$ can be expressed as $\alpha(j,k)/(\alpha(j,k)+\beta(j,k))$.

6.2.2 Spectrum Pool Structure

In the classical cellular networks, each cell uses different spectrum bands with its neighbor to prevent inter-cell interference. This concept can be also applied to CR cellular networks. Since the spectrum bands in the classical wireless networks such as wireless LANs are contiguous and located in the relatively narrow frequency range, mobile CR users can switch the spectrum without changing their RF front-ends. On the contrary, as explained in Section 6.1, CR users need to reconfigure their operation frequency at the RF front-end whenever available spectrum bands changes, which causes significant switching latency.

To solve this problem, we introduce the spectrum-pooling concept which is considered to be suitable to adapt to the dynamic radio environment in CR networks [9] [73]. We extend this concept to multi-cell environment to consider both spectrum and user mobilities. In the proposed architecture, the spectrum pool is defined as a set of contiguous licensed spectrum bands, each of which consists of multiple channels. Each channel is assumed to maintain the same QoS by exploiting power control and adaptive modulation schemes. Here, we assume that the spectrum pools are assigned to each cell exclusively of its neighbor cells while maintaining the frequency reuse factor as f, as shown in Figure 47. Although this architecture provides the seamless transition between spectrum bands within the pool, it still has difficulty in supporting seamless communication in CR users moving across different cells. To address this problem, in the proposed architecture, each cell has two different cell coverage types: basic area (BA) and extended area (EA). While the basic area is not overlapped with the coverage of its neighbor cells, the extended area has much larger coverage extended to the basic area of its neighbors. As a result, the spectrum pool consists of multiple *basic spectrum* bands that support only the basic area, and a single *extended spectrum* providing both the basic and the extended areas. The basic spectrum supports $N_i^{\max}(j)$ for users in the basic area.

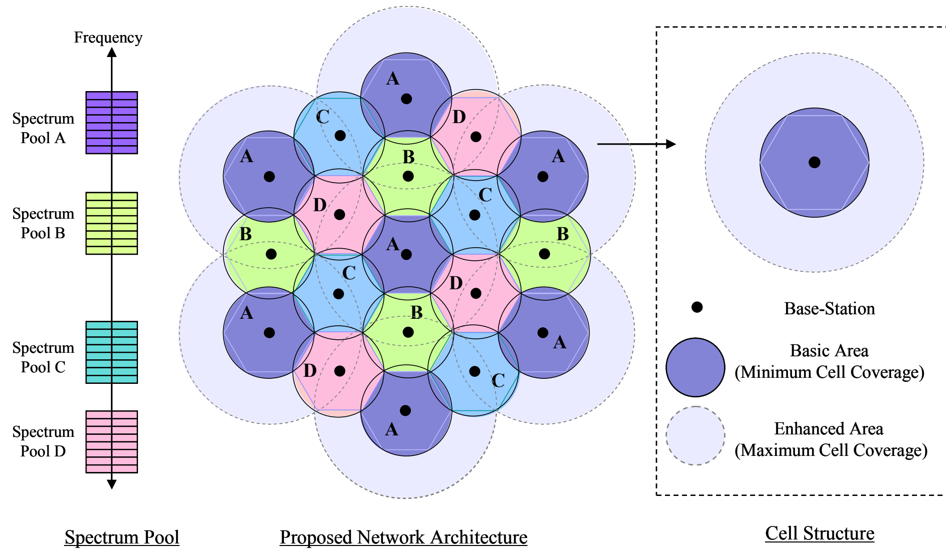


Figure 47: Spectrum pool based CR network architecture (a frequency reuse factor f is assumed to be 4).

Although a large coverage improves the mobility support in CR networks, the users in the extended area require more spectrum resources than those in the basic area, leading to degradation of cell capacity. Assume that the extended spectrum band j at spectrum pool i can support $\rho N_i^{\max}(j)$ channels for the users in the basic area. Then, it supports the $N_i^{\max}(j)$ channels for the users in the extended area because of the longer distant from the BS where ρ is greater than unity and can be determined dependent on the transmission power and the minimum signal strength for decoding. Furthermore, due to the extended spectrum, the current cell has another type of neighbor, referred to as *extended neighbors*. The extended neighbors are the cells which have the same spectrum pool within the interference range of the extended spectrum. In this architecture, unlike the basic spectrum, the extended spectrum in the current cell cannot be used in its extended neighbors to avoid inter-cell interference.

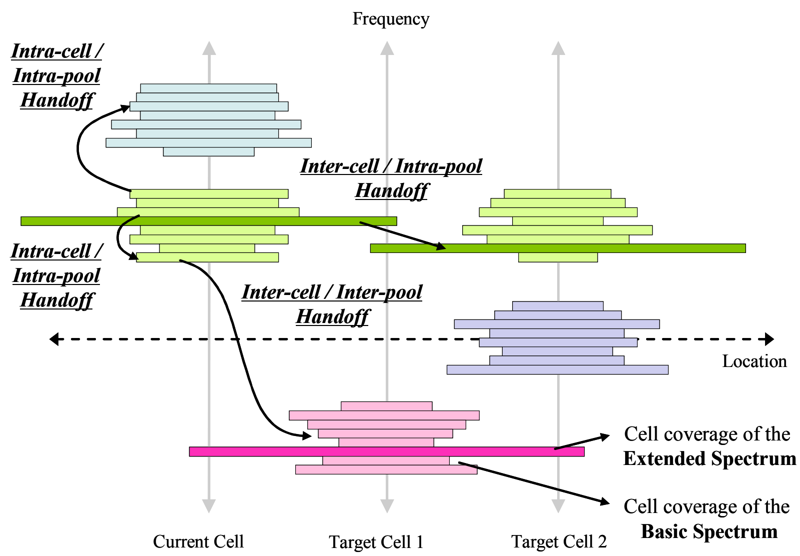


Figure 48: Different handoff types in CR networks.

6.2.3 Handoff Types

Mobility management in classical cellular networks is closely related to user mobility. However, CR networks have another unique mobility event, the so-called *spectrum mobility*. By taking into account both mobility events based on the proposed network architecture, we define following four different types of handoff schemes, as shown in Figure 48:

- *Intra-cell/intra-pool handoff:* The CR user moves to the spectrum band in the same spectrum pool without switching a serving BS.
- *Inter-cell/intra-pool handoff:* The CR user switches its serving BS to the neighbor BS without changing the spectrum pool.
- *Inter-cell/inter-pool handoff:* The CR user switches its serving BS to the neighbor BS, which has a different spectrum pool.
- *Intra-cell/inter-pool handoff:* CR users change their spectrum bands from one spectrum pool to another within the current cell.

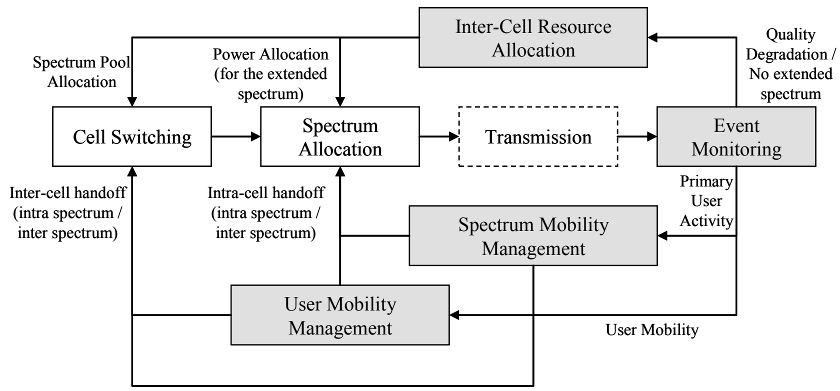


Figure 49: The proposed mobility management framework.

Each handoff type is related to different mobility event, and its performance is mainly dependent on both network and user conditions, such as resource availability, network capacity, user location, etc. Thus, CR networks need a unified mobility management scheme to exploit different handoff types adaptively to the dynamic nature of underlying spectrums, which will be explained in Section 6.3.

6.3 Mobility Management Framework

6.3.1 Overview

Because of the dynamic spectrum environment and heterogeneous handoff types, CR networks require more complicated mobility management functionalities. These functionalities are initiated by three different events: user mobility, spectrum mobility, and quality degradation. Here, *user mobility* is defined as the event that CR users are approaching the cell boundary. On the contrary, *spectrum mobility* is referred to as the event that CR users switch their spectrum resulting from the PU activity. Each BS detects one of these events by monitoring current spectrum availability and the quality variation of current transmissions, and perform a proper mobility management function accordingly.

In the case of user and spectrum mobility events, CR networks decide on a proper handoff type for their mobile users by performing user and spectrum mobility management functions, respectively. According to the decision, CR users need either to select a target cell (*cell selection*) or to determine the best available spectrum (*spectrum allocation*), as described in Figure 49.

If a current cell does not have enough spectrum resources owing to either PU activity or increase in users, the BS performs an *inter-cell resource allocation* through the negotiation with its neighbor cells, Through this operation, the cell can obtain additional spectrum pools, which increases its capacity. This concept has been widely studied in [10], [43], and [56]. Thus, in this chapter, we assume that each cell has a single spectrum pool to mainly focus on mobility management.

Furthermore, because of the unique architectural characteristics, the proposed mobility management framework requires a unique feature for inter-cell resource allocation, which will be explained in the following subsection.

6.3.2 Inter-Cell Resource Allocation

The use of the extended spectrum leads to increase in current cell capacity. However, as explained in Section 6.2, an extended spectrum in a current cell cannot be used in its extended neighbor cells, leading to decrease in their capacity. Since each spectrum shows different characteristics in capacity according to cell locations, how to select the extended spectrum for the current cell is a critical issue to determine the performance of proposed framework. Furthermore, each cell has time-varying spectrum resources because of the dynamic nature of underlying spectrum in CR networks. Thus, each cell cannot have a permanent extended spectrum band. As a result, CR networks necessitate a dynamic inter-cell resource allocation scheme to maintain the extended spectrum over time. Although a global optimization method in every spectrum change achieves optimal allocation, it requires high computational complexity and may cause high communication overhead as a result of frequency spectrum switching. Instead of global optimization, we consider the stochastic characteristics of spatial and temporal spectrum availabilities, and develop a distributed inter-cell resource allocation method, which improves total network capacity as well as mobility support, i.e., the availability of the extended spectrum. The following are the procedures of the proposed method.

1. Initially, all available spectrums in a current cell i, $\mathcal{S}_i$, are considered as basic spectrum bands.

2. The capacity of the current cell is defined as the sum of the expected idle duration in each spectrum as follows:

$$C_i(j) = N_i^{\max}(j) \cdot \prod_{k \in \mathcal{A}_i^{\mathrm{B}}(j)} P^{\mathrm{off}}(j,k) \tag{72}$$

   where $\mathcal{A}_i^{\mathrm{B}}(j)$ is a set of the PU activity regions of spectrum j in the basic area of cell i.

3. Once the extended spectrum is lost due to the PU activity, inter-cell spectrum sharing is performed to find a new spectrum, which takes time because of information exchange with its neighbor cells. Thus, reliability in the extended spectrum can be expresses as the ratio of an average idle time in the extended spectrum band to total time, including inter-cell spectrum sharing delay as follows:

$$R_i(j) = \frac{\frac{1}{\sum_{k \in \mathcal{A}_i^{\mathrm{E}}(j)} \beta(j,k)}}{\frac{1}{\sum_{k \in \mathcal{A}_i^{\mathrm{E}}(j)} \beta(j,k)} + T^{\mathrm{inter}}} \tag{73}$$

   where $1/\sum_{k \in \mathcal{A}_i^{\mathrm{B}}(j)} \beta(j,k)$ represents the average idle period of spectrum j at cell i, and T^{inter} is an inter-cell resource allocation delay.

4. Each cell prefers an extended spectrum with higher reliability. However, once the current cell determines the extended spectrum, its extended neighbor cannot use that spectrum, and hence lose their capacity. To describe these features,

Table 4: Symbols used for the analytical modeling in spectrum mobility.

Symbols	Descriptions
N_i^{b}	Total number of channels used in the basic area (BA) of cell i
N_i^{e}	Total number of channels used in the extended area (EA) of cell i
$N_i^{\max}(j)$	Maximum number of channels in spectrum band j at the basic area of cell i
$\alpha(j,k)$	PU activity (busy $\rightarrow$ idle) at area k of the spectrum band j
$\beta(j,k)$	PU activity (idle $\rightarrow$ busy) at area k of spectrum band j
ρ	Channel gain of users in the basic area at the extended spectrum
Δt	Sensing interval (sensing operation in every Δt)

we develop a novel metric for the expected gain, which can be expressed as the product of the spectrum reliability of the extended spectrum in the current cell and a ratio of the capacity gain in current cell to the sum of capacity loss in extended neighbors as follows:

$$G_i(j) = R_i(j) \cdot \frac{\rho N_i^{\max}(j) \prod_{k \in \mathcal{A}_i^{\mathrm{E}}(j)} P^{\mathrm{off}}(j,k) - C_i(j)}{\sum_{i' \in \mathcal{N}_i^E} [C_{i'}(j) \cdot N_{i'}^{\max}(j)]} \tag{74}$$

where $\mathcal{A}_i^{\mathrm{E}}(j)$ is a set of the PU activity regions of spectrum j in the extended area of cell i. $\mathcal{N}_i^E$ is a set of the extended neighbors of cell i.

5. The current cell considers the expected capacity gain over all available spectrum bands and chooses the extended spectrum j^* to satisfy the following condition.

$$j^* = \arg_{j \in \mathcal{S}_i} \max G_i(j) \tag{75}$$

In the following sections, we introduce handoff models in term of switching latency, and then propose spectrum and user mobility management schemes. For ease of presentation, the important symbols used in the subsequent discussion are summarized in Table 4.

6.4 Spectrum Handoff Modeling

According to the mobility events, each handoff scheme necessitates different strategies as follows:

- *Proactive handoff:* When CR users detect handoff events, they perform the handoff procedures while maintaining communications. After CR users make decision on handoff, they cut off communication channels and switch to a new spectrum band or a new BS. User mobility and cell overload are the examples of proactive handoff events. Most of classical handoff schemes are based on the proactive approach.

- *Reactive handoff:* CR users should stop the transmission immediately in the reactive handoff event. Then, they make decisions and perform the handoff. As a result, this handoff has an additional handoff delay, unlike the proactive approach. The PU activity is a reactive handoff event.

Based on these strategies, handoff schemes defined in Section 6.2 can be modeled as follows:

6.4.1 Intra-Cell/Intra-Pool Handoff

Intra-cell/Intra-pool handoff occurs when primary users are detected in the spectrum. Thus, it is implemented in a reactive approach. First, this handoff approach requires a preparation time to determine the handoff type (d_{prep}). After that, for sensing operations, CR users need to wait to the next sensing cycle, called a sensing synchronization time ($d_{\text{syn}}^{\text{sen}}$). Then, they sense the spectrum bands in the pool (d_{sen}), and determine the proper spectrum (d_{dec}). Finally, CR users move to the new spectrum band and resume transmission after the synchronization to the transmission schedule on that spectrum ($d_{\text{syn}}^{\text{tx}}$). Since spectrum bands in the pool are contiguous, CR users can switch the spectrum without reconfiguring their RF frontends, and hence the physical spectrum switching delay is negligible. In summary, the latency for intra-cell/intra-pool handoff (Type 1) can be expressed as follows:

$$D_1 = d_{\text{prep}} + d_{\text{syn}}^{\text{sen}} + d_{\text{sen}} + d_{\text{dec}} + d_{\text{syn}}^{\text{tx}} \tag{76}$$

6.4.2 Intra-Cell/Inter-Pool Handoff

If CR BSs can exploit multiple spectrum pools, they may use intra-cell/inter-pool handoff in the following case: If the current spectrum pool does not have enough spectrum resources because of PU activity, CR users detecting PU activities switch to another spectrum pool in the current cell. This is also a reactive handoff. Thus, its handoff latency is similar to that of the intra-cell/intra-pool handoff as follows (Type 2):

$$D_2 = d_{\text{prep}} + d_{\text{recfg}} + d_{\text{syn}}^{\text{sen}} + d_{\text{sen}} + d_{\text{dec}} + d_{\text{syn}}^{\text{tx}} \tag{77}$$

However, unlike the intra-cell/intra-pool handoff, this scheme requires the reconfiguration of RF frontend since each spectrum pool is placed in the different frequency range. Usually reconfiguration takes longer than other delay factors.

6.4.3 Inter-Cell/Inter-Pool Handoff

This handoff scheme is similar to that in classical cellular networks, which is required for CR users moving across multiple cells. To determine a target cell, a mobile CR user is required to observe the signals from neighbor cells during its transmission. However, since neighbor cells use different spectrum pools, the mobile CR user needs to stop its transmission and reconfigure its RF front-end in every observation of neighbor cells, which is a tremendous overhead in handoff. Thus, instead of this mobile station-controlled method, a network-controlled approach is more feasible for inter-cell/inter-pool handoff, where the BS determines the target cell based on the stochastic user information, which will be explained in Section 6.5. As a result, mobile CR users need a single reconfiguration time. In this case, the BS can prepare the handoff in advance according to user mobility. Thus, this is a proactive handoff and does not requires the handoff preparation time d_{prep} used in previous reactive handoff types as follows (Type 3):

$$D_3 = d_{\text{recfg}} + d_{\text{syn}}^{\text{sen}} + d_{\text{sen}} + d_{\text{dec}} + d_{\text{sych}}^{\text{tx}} \quad (78)$$

Furthermore, PU activities can initiate this handoff scheme in special reactive events. First, when all spectrum pools in the current cell are overloaded because of PU activity, the BS forces CR users to move to neighbor cells. This is exactly the same procedures as the intra-cell/inter-pool handoff, and requires D_2 handoff latency. Second, if a PU activity is detected in the extended spectrum, CR users in the extended spectrum should switch to the neighbor cells. Since there is no other available spectrum in the extended area after PU activity, they lose a control channel as well. To solve this problem, the BS determines handoff information and sends it to a selected target cell. Then, the target cell broadcasts the advertisement message for the CR user through its control channel. In this scenario, CR users need one or more reconfigurations of the RF frontend until it hears the advertisement message. Also in every reconfiguration, CR users monitor the control channel for a certain time (d_{lis}). The latency in this case (Type 4) can be expressed as follows:

$$D_4 = d_{\text{prep}} + \gamma(d_{\text{recfg}} + d_{\text{lis}}) + d_{\text{syn}}^{\text{sen}} + d_{\text{sen}} + d_{\text{dec}} + d_{\text{syn}}^{\text{tx}} \quad (79)$$

Because of multiple reconfigurations, inter-cell/inter-pool handoff in this case shows the worst performance. γ is dependent on the searching order of neighbor cells. In this chapter, the order is randomly chosen, and hence γ is considered as $(f + 1)/2$ on average where f is a frequency reuse factor.

6.4.4 Inter-Cell/Intra-Pool Handoff

This handoff is performed when mobile CR users in extended areas successfully switch to the extended neighbors. This is also a proactive handoff. Furthermore, a new target cell is an extended neighbor which uses the same spectrum pool as the current cell. Thus, reconfiguration is not required. Therefore, the latency for inter-cell/intra-pool handoff scheme (Type 5) can be expressed as follows:

$$D_5 = d_{\text{syn}}^{\text{sen}} + d_{\text{sen}} + d_{\text{dec}} + d_{\text{syn}}^{\text{tx}} \tag{80}$$

In this handoff, the latency is significantly reduced compared to that in other cases. Thus, this type of handoff is more advantageous to mobile CR users, and hence improves mobility in CR networks.

6.5 Spectrum Mobility Management in Cognitive Radio Networks

6.5.1 Overview

Spectrum mobility is the unique characteristic in CR networks. When primary users appears in the spectrum, CR users generally change its spectrum band without switching the BS. However, since CR networks have time-varying spectrum availability, each cell may not have enough spectrum band to support current users. To solve this problem, an admission control scheme is proposed in [46]. However, in the proposed architecture, CR users can have another option, cell switching, because of the hierarchical spectrum structure described in Section 6.2.2. Here, we propose the spectrum mobility framework by considering both spectrum and cell switching methods.

When the PU activity is detected in the cell, the BS needs to check if it has enough spectrum resources for intra-cell/intra-pool handoff. If the cell has sufficient resources, the BS performs the intra-cell/intra-pool handoff for all users requiring new spectrum bands (Type 1). Otherwise, some of current users are forced to move to the neighbor cells. Since the users in the extended area occupy much more spectrum resources, the BS selects some of them and performs the inter-cell/inter-pool handoff (Type 2). If this operation is not enough to ensure spectrum mobility, the BS extends its selection to the users in the basic area. If the PU activity is detected in the extended spectrum, all users in the extended areas need to perform inter-cell/inter-pool handoff (Type 2), regardless of current spectrum resources since they cannot find other available spectrum bands for switching in that area. After the user selection, the selected users need to find the proper target cell. Unlike the classical handoff, CR users cannot observe the signal strength from other neighbor cells while maintaining the connection to current cells. Instead, CR users select the new BS based on the stochastic connectivity model.

The intra-cell/intra-pool handoff is exactly same as the spectrum decision proposed in [46], and hence out of scope in this chapter. In the following subsection we will describe the user and cell selections for the inter-cell/inter-pool handoff scheme in more detail.

6.5.2 User Selection

Let $\mathcal{S}_i$ be a set of the currently available spectrum in cell i. Then, the number of unused channels in the available spectrum bands at cell i, N_i^{av}, can be expressed as $\sum_{j \in \mathcal{S}_i} N_i^{\max}(j) - (N_i^{\text{b}} + \rho N_i^{\text{e}})$. Here, N_i^{b} and N_i^{e} are the numbers of channels used in the basic area and the extended area of cell i, respectively, and can be obtained as follows:

$$N_i^{\text{b}} = \sum_{m \in \mathcal{U}_i^{\text{b}}} K_m, \quad N_i^{\text{e}} = \sum_{m \in \mathcal{U}_i^{\text{e}}} K_m \tag{81}$$

If the number of the channels required for spectrum switching, N^{req} is less than N_i^{av}, CR users just perform the intra-cell intra-pool handoff. As explained in Section 6.5.1, the users in extended area should move to the neighbor cells when they detect the PU activity, and hence are not counted in N^{req}.

If $N_i^{\text{av}} < N^{\text{req}} \leq N_i^{\text{av}} + \rho N_j^{\text{e}}$, current cell is overloaded and forces some of the users to be out to their neighbor cells. In this case, CR users using $\lceil \frac{N^{\text{req}} - N_i^{\text{av}}}{\rho} \rceil$ channels in the extended area need to be selected and moved to the neighbor cells. As the users stay in the extended area for a longer time, cell capacity becomes lower. Also these users have a higher probability to be interrupted by the PU activity. Furthermore, a user with more channels reduces the number of users that the current cell can admit.

Thus, the BS selects the users in the extended area with the longest expected staying time as well as the highest channel occupancy. As a result, a decision metric can be obtained as $K_m \cdot d_m/v_m$ where d_m is the expected moving distance of user m to the cell boundary, which is dependent on the user mobility model. v_m is the velocity of user m. The BS chooses users in the extended area with the largest decision metric, repeatedly until it can avoid cell overload state.

If $N^{\mathrm{req}} > N_i^{\mathrm{av}} + \rho N_i^{\mathrm{e}}$, it is not enough to select all users in the extended area. To avoid dropping or blocking connections resulting from cell overload, the BS hands over some of the users in the basic area to their neighbor cells. Unlike in the previous case, the BS selects CR users using $N^{\mathrm{req}} - (N_i^{\mathrm{av}} + \rho N_i^{\mathrm{e}})$ channels with the shortest expected staying time in the basic area since they are highly likely to move to the extended area, which will require more spectrum resources. Similar to the previous case, it is more advantageous to kick off the users with more channels. Thus, the BS chooses CR users in the basic area with the smallest decision metric, $d_m/(v_m \cdot K_m)$.

6.5.3 Cell Selection

One of main challenges in CR mobility management is how to determine a proper target cell. Since each spectrum pool is distributed over a wide frequency range, CR users need to reconfigure their RF front-ends for monitoring the signals from neighbor cells, leading to relatively long temporary disconnection of the transmission. In this chapter, instead of the received signal strength, we propose stochastic connectivity estimation for selecting a proper target cell. The user connectivity to the BS is mainly related to the distance from the transmitter. Furthermore, stochastic factors such as shadowing and multi-path fading influence the connectivity. If the received signal needs to be greater than $p_{0,\mathrm{dB}}$ for decoding data reliably, the connection probability can be obtained as follows [72]:

$$\begin{aligned} P_i^{\mathrm{c}} &= Pr[p_{\mathrm{t,dB}} - \bar{L}_{0,\mathrm{dB}} - 10\log_{10} E[\chi^2] \\ &\quad - 10\xi \log_{10} D - X_{\sigma_{\mathrm{s}}} \geq p_{0,\mathrm{dB}}] \\ &= \frac{1}{2}(1 - \mathrm{erf}[(10\xi\log_{10} D + p_{0,\mathrm{dB}} - p_{\mathrm{t,dB}} \\ &\quad - \bar{L}_{0,\mathrm{dB}} - 10\log_{10} E[\chi^2])/\sqrt{2}\sigma_{\mathrm{s}}]) \end{aligned} \tag{82}$$

where $p_{\mathrm{t,dB}}$ is the transmission power, $\bar{L}_{0,\mathrm{dB}}$ is the average path loss at the reference distance, 1 meter, D is the distance from the BS, $10\log_{10} E[\chi^2]$ is the average multi-path fading in dB, ξ is the path loss exponent, $X_{\sigma_{\mathrm{s}}}$ is shadowing, χ^2 is multi-path fading, $\mathrm{erf}[z]$ is the error function defined by $\int_0^z \frac{2}{\sqrt{\pi}} e^{-}x^2 dx$.

Since the spectrum pool consists of multiple spectrum bands, the connectivity of spectrum pool i, P_i^{c}, can be defined as the probability that at least one spectrum band provides the valid connection, which can be expressed as $1 - \prod_{j\in\mathcal{S}_i}(1 - P_i^{\mathrm{c}}(j))$, where $P_i^{\mathrm{c}}(j)$ is the connection probability of spectrum j in pool j. Besides connectivity, spectrum utilization is also an important factor in determining the target cell. Thus, CR users select target cell i^* with the highest weighted connectivity, P_i^{w}, which can be obtained by considering both connectivity and spectrum utilization as follows:

$$P_i^{\mathrm{w}} = (1 - \prod_{j\in\mathcal{S}_i}(1 - P_i^{\mathrm{c}}(j))) \cdot (1 - \frac{N_i^{\mathrm{b}} + \rho N_i^{\mathrm{e}}}{\sum_{j\in\mathcal{S}_i} N_i^{\mathrm{max}}(j)}) \tag{83}$$

6.6 User Mobility Management in Cognitive Radio Networks

6.6.1 Overview

User mobility is another main reason to initiate handoff in CR networks, which occurs at the boundary of either basic or extended areas.

When CR users approach the boundary of the extended area, they try to perform inter-cell/intra-pool handoff (Type 5) first. Unlike the inter-cell/inter-pool handoff, CR users can measure the signal strength from other BS directly, which is exactly the same as classical handoff schemes. If CR users cannot find the proper target cell for inter-cell/intra-pool handoff, they need to perform the inter-cell/inter-pool handoff to find a cell having a different spectrum pool. This procedure is same as the cell selection scheme but does not require a preparation time (Type 3), which is explained in Section 6.5.3.

Compared to the handoff strategy at the boundary of the extended area, CR networks need to have more complicated mobility management scheme at the boundary of the basic area. When CR users approach the basic cell boundary, they need to determine whether they will stay in the extended area of the current cell. Generally, for mobile users, larger cell coverage is known to be much more advantageous since it reduces the number of handoffs [59]. However, in CR networks, large cell coverage is not always desirable for mobile users. As the cell coverage becomes larger, the PU activity becomes higher since it is highly probable to include multiple PU activity regions. This PU activities in the extended area result in significantly long switching latency, as described in Section 6.4. In addition, since CR users in the basic area are allowed to have higher priority in channel access, as presented in Section 6.5, cell overload also influences the use of extended spectrum band. As a result, CR networks need a sophisticated algorithm to select the best handoff type for the mobile users.

Thus, in this section, we focus on mobility management in the boundary of the basic area. When CR users become closer to the boundary, the BS initiates the handoff procedures and gather the neighbor cell information from a central network entity. Based on the information, the BS estimates the connectivity of the candidate cells and determines the handoff timing t^* and target cell i^* as follows:

$$[t^*, i^*] = \arg_{i \in \mathcal{C}, t>0} [P^{\mathrm{c}}_{i_{\mathrm{c}}}[d^0_{i_{\mathrm{c}}} + v^{\mathrm{r}}_{i_{\mathrm{c}}} t] \leq \max_{i \in \mathcal{C}} [P^{\mathrm{c}}_i [d^0_i + v^{\mathrm{r}}_i t]]] \tag{84}$$

where P^{c}_i is a connectivity of cell i, which is a function of the distance d^0_i and the relevant velocity v^{r}_i to its BS. $\mathcal{C}$ is a set of candidate cells, i_{c} represents the current cell, and t is the moving time.

Once a target cell is determined, the BS determines the handoff type by considering the expected switching costs of both intra-cell/intra-pool (Type 1) and inter-cell/inter-pool handoff schemes (Type 3) at the boundary of the basic area. The expected switching costs can be determined by estimating the probability of mobility events after the decision. After the decision, CR users may experience the unexpected inter-cell/inter-pool handoff resulting from the following reasons: 1) PU activity in the extended area, 2) capacity overload in the extended area, and 3) capacity overload in the basic area. In the following subsections, first, we analyze these future events after the decision and accordingly propose an intelligent handoff decision scheme.

6.6.2 Primary User Activity in the Extended Area

If CR users are determined to perform the intra-cell/intra-pool handoff (Type 1) at the boundary of the basic area, they can stay in the current cell, which does not require long switching latency for inter-cell/inter-pool handoff. However, in the extended area, CR users may experience mobility events which cause inter-cell/inter-pool handoff (Type 4). One of those events is the PU activity. Since CR users in the extended area cannot find other available spectrums when they detect the PU activity, the inter-cell/inter-pool handoff is inevitable.

As shown in Figure 50, more PU activity regions can be involved in determining spectrum availability in the extended area, which leads to higher PU activity. Furthermore, the interference range of the extended spectrum is larger than its coverage and hence is overlapped with the coverages of the extended neighbors. Thus, for the accurate detection, all extended neighbors need to be involved in detecting the PU activity with its own detection and false probabilities. Assume that cooperative detection is performed according to an 'OR' logic, referred to as decision fusion [49]. Then, a cooperative detection probability converges to 1 as the number of cells increases [52]. Thus, the detection probability can be ignored when estimating spectrum

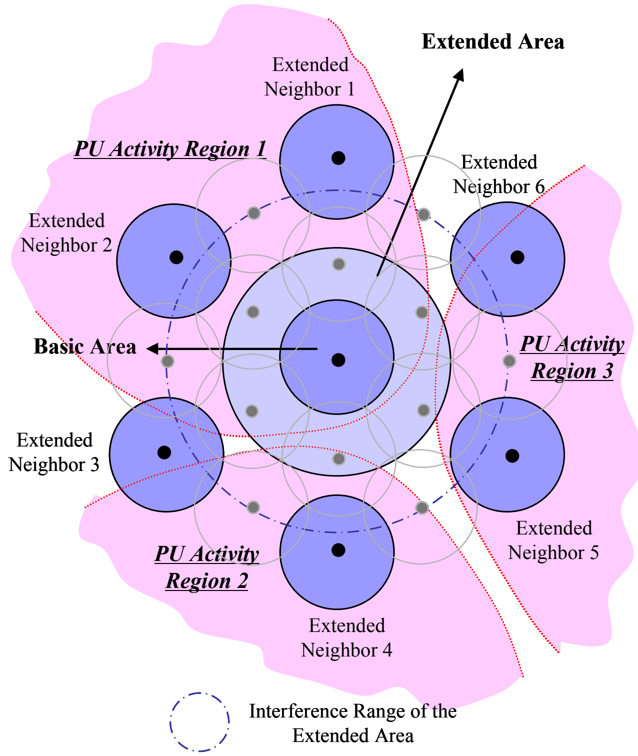


Figure 50: The influence of primary user activities in the extended area.

availability. On the contrary, the false alarm probability increases as the number of cells increases [52], which influences spectrum availability significantly in the extended area. Even though the spectrum band is idle, the spectrum is determined to be unavailable if the false alarm is detected.

Thus, to avoid the inter-cell/inter-pool handoff, neither the PU activity nor false alarm should not be detected in the extended area. Based on these observations, we can derive the probability that no primary user can be detected during the expected staying time $T_m = d_m/v_m$ as follows:

$$P_i^{\mathrm{av}}(T_m) = \prod_{i' \in \mathcal{N}_i^{\mathrm{E}}} (1 - P_{i'}^{\mathrm{f}})^{\lceil \frac{T_m}{\Delta t} \rceil} \cdot \prod_{k \in \mathrm{A}_i^{\mathrm{E}}(j)} e^{-\beta(j,k) T_m} \tag{85}$$

The first term represents the probability that all extended neighbors do not generate any false alarm during T_m where $\lceil T_m/\Delta t \rceil$ spectrum sensing operations are performed. This is based on decision fusion, and will change if other cooperative decision criteria are used. Here, sensing operation is assumed to be performed in every Δt sensing period. Second term represents the probability that no PU activity appears in the extended area. Then, the probability of detecting the PU activity can be obtained as $1 - P_i^{\mathrm{av}}(T_m)$.

6.6.3 Capacity Overload in the Extended Area

As explained in Section 6.5, when the current cell is overloaded, CR users in the extended area may need to perform an inter-cell/inter-pool handoff (Type 2). In this section, we derive the probability of cell overload. The PU activity in the extended spectrum leads to the inter-cell/inter-pool handoff, regardless of cell overload, which is already considered in Section 6.6.2. Thus, we assume that cell overload results from PU activities only in the basic spectrums, and the extended spectrum is considered to be idle in this case.

First, since each PU activity area in the spectrum has can have 2 states, busy and idle, we can model a transition matrix $\mathbf{X}(j,k)$ with following transition probabilities:

$$\begin{aligned} x_{1,1}(j,k) &= e^{-\beta(j,k)\Delta t} \\ x_{1,2}(j,k) &= 1 - e^{-\beta(j,k)\Delta t} \\ x_{2,1}(j,k) &= 1 - e^{-\alpha(j,k)\Delta t} \\ x_{2,2}(j,k) &= e^{-\alpha(j,k)\Delta t} \end{aligned} \tag{86}$$

where $x_{1,1}(j,k)$ and $x_{1,2}(j,k)$ are the transition probabilities from idle to idle and from idle to busy, respectively. $x_{2,1}(j,k)$ and $x_{2,2}(j,k)$ represent the transition probabilities from busy to idle and from busy to busy, respectively.

From this, the transition matrix after $r\Delta t$ can be obtained as $[\mathbf{X}(j,k)]^r$. Let $\mathbf{x_0}(j,k) \in \{(1,0),(0,1)\}$ be an initial vector to describe a current spectrum status where $(1,0)$ and $(0,1)$ denote that an area k at spectrumm j is currently idle and busy, respectively. Then, the idle probability of area k after $r\Delta t$, $P_i^{\mathrm{idle}}(j,k,r\Delta)$ is the first element of the vector, $x_0[\mathbf{X}(j,k)]^r$, which can be obtained by Eq. (87) [40].

$$P_i^{\mathrm{idle}}(j,k,r\Delta t) = \begin{cases} \dfrac{x_{2,1}(j,k)}{x_{1,2}(j,k)+x_{2,1}(j,k)} + \\ (1-x_{1,2}(j,k)-x_{2,1}(j,k))^r \cdot \dfrac{x_{1,2}(j,k)}{x_{1,2}(j,k)+x_{2,1}(j,k)}, & x_0=(1,0) \\ \dfrac{x_{2,1}(j,k)}{x_{1,2}(j,k)+x_{2,1}(j,k)} - \\ (1-x_{1,2}(j,k)-x_{2,1}(j,k))^r \cdot \dfrac{x_{2,1}(j,k)}{x_{1,2}(j,k)+x_{2,1}(j,k)}, & x_0=(0,1) \end{cases} \tag{87}$$

Based on idle probabilities at each PU activity area, we can derive the idle and busy probabilities of spectrum j. Assume that a current cell i has multiple PU activity regions in spectrum j. Then, it can use that spectrum only when all of these regions should be idle, and hence the idle and busy probabilities, is expressed as follows:

$$\begin{aligned} P_i^{\mathrm{idle}}(j,r\Delta t) &= \prod_{k\in\mathcal{A}_i^{\mathrm{B}}(j)} P_i^{\mathrm{idle}}(j,k,r\Delta t) \\ P_i^{\mathrm{busy}}(j,r\Delta t) &= 1-P_i^{\mathrm{idle}}(j,r\Delta t) \end{aligned} \tag{88}$$

where $\mathcal{A}_i^{\mathrm{B}}(j)$ is the set of PU activity regions of the basic area in spectrum j at cell i.

Based on both probabilities of each spectrum in the pool, we derive the expected spectrum availability of cell i as follows: The current cell i has $|S_i|$ assigned spectrum bands in the pool, which are either busy or idle. Since the extended spectrum is not considered as explained earlier, it has $2^{|S_i|-1}$ states according to spectrum availability. Among spectrum states, some of the states cannot satisfy the channel requirements to support current users, and finally result in inter-cell/inter-pool handoff of some of users in the extended area, which can be defined as follows:

$$\mathcal{L}_{\mathrm{E}} = \{\arg_n[\sum_{j\in\mathcal{I}_n} N_i^{\max}(j) < N_i^{\mathrm{b}} + \rho N_i^{\mathrm{e}}],\ \text{for } \forall n\} \tag{89}$$

where $\mathcal{I}_n$ is a set of idle spectrum bands at state n $(n=1,\ldots,2^{|S_i|-1})$.

To resolve cell overload at each state in $\mathcal{L}_{\mathrm{E}}$, current cell needs to obtain additional channels by switching CR users to neighbor cells. The numbers of required channels are different from one state to another. The following is the probability that users in the extended area are switched to neighbor cell as a result of cell overload at state $n \in \mathcal{L}_{\mathrm{E}}$:

$$u_{\mathrm{E}}^{\mathrm{req}}(n) = \frac{\min\left[\max\left[0, N_i^{\mathrm{b}} + \rho N_i^{\mathrm{e}} - B_i^{\max}(n)\right], \rho N_i^{\mathrm{e}}\right]}{\rho N_i^{\mathrm{e}}} \tag{90}$$

where $B_i^{\max}(n)$ is the number of available channels in spectrum pool i at state n, which can be obtained as $\sum_{j \in \mathcal{I}_n} N_i^{\max}(j)$. $N_i^{\mathrm{b}} + \rho N_i^{\mathrm{e}} - B_i^{\max}(n)$ represents the number of channels requiring the inter-cell/inter-pool handoff to prevent cell overload.

To order to maintain the underload state after $r\Delta t$, all spectrum bands in $\mathcal{I}_n, n \in \mathcal{L}_E$ should be idle and the rest of spectrums $j \notin \mathcal{I}_n$ should be busy. Furthermore, the cell should not have any cell overload and any PU activities in the extended area before $r\Delta t$. By considering these conditions, we derive the underload probability of cell i, $P_{\mathrm{E},i}^{\mathrm{under}}$, as follows:

$$P_{\mathrm{E},i}^{\mathrm{under}}(\Delta t) = \sum_{n \in \mathcal{L}_{\mathrm{E}}} \left[\prod_{j \in \mathcal{I}_n} P_i^{\mathrm{idle}}(j, \Delta t) \prod_{j \notin \mathcal{I}_n} P_i^{\mathrm{busy}}(j, \Delta t)\right] \tag{91}$$

$$\begin{aligned} P_{\mathrm{E},i}^{\mathrm{under}}(r\Delta t) &= P_{\mathrm{E},i}^{\mathrm{under}}((r-1)\Delta t) \\ &\cdot \sum_{n \in \mathcal{L}_{\mathrm{E}}} \left[\prod_{j \in \mathcal{I}_n} P_i^{\mathrm{idle}}(j, r\Delta t) \prod_{j \notin \mathcal{I}_n} P_i^{\mathrm{busy}}(j, r\Delta t)\right] \\ &(r = 2, 3, ...) \end{aligned} \tag{92}$$

For cell overload after $r\Delta t$, we should consider the states not in $\mathcal{L}_E$. Furthermore, the cell should not experience any cell overload as well as any PU activity before. Then, cell overload probability can be expressed as follows:

$$\begin{aligned} &P_{\mathrm{E},i}^{\mathrm{over}}(\Delta t) = \\ &\sum_{n \notin \mathcal{L}_{\mathrm{E}}} \left[\prod_{j \in \mathcal{I}_n} P_i^{\mathrm{idle}}(j, \Delta t) \prod_{j \notin \mathcal{I}_n} P_i^{\mathrm{busy}}(j, \Delta t) \cdot u_{\mathrm{E}}^{\mathrm{req}}(n)\right] \end{aligned} \tag{93}$$

$$\begin{aligned} P_{\mathrm{E},i}^{\mathrm{over}}(r\Delta t) &= P_{\mathrm{E},i}^{\mathrm{under}}((r-1)\Delta t) \\ &\cdot \sum_{n \notin \mathcal{L}_{\mathrm{E}}} \left[\prod_{j \in \mathcal{I}_n} P_i^{\mathrm{idle}}(j, r\Delta t) \prod_{j \notin \mathcal{I}_n} P_i^{\mathrm{busy}}(j, r\Delta t) \cdot u_{\mathrm{E}}^{\mathrm{req}}(n)\right] \\ &(r = 2, 3, ...) \end{aligned} \tag{94}$$

Based on both overload and underload probabilities, we obtain a probability that a CR user in the extended area initiates inter-cell/inter-pool handoff caused by cell overload, P_E^{over}, as follows:

$$P_{\text{E}}^{\text{over}} = \sum_{r=1}^{R} [P_{\text{E},i}^{\text{over}}(r\Delta t) \cdot P_i^{\text{av}}(r\Delta t)] \quad (95)$$

where $R = \lceil T_m/\Delta t \rceil$ where T_m is the expected time of user m to stay in the extended area. Note that the extended spectrum is assumed to be available in estimating $P_{\text{E}}^{\text{over}}$ as mentioned in the beginning of this section.

6.6.4 Capacity Overload in the Basic Area

If the BS is determined to perform the inter-cell/inter-pool handoff (Type 3) at the boundary of the basic area, mobile CR users may experience the capacity overload in the basic area of the target cell, which causes inter-cell/inter-pool handoff. This cell overload probability can be determined with a procedure similar to the one used in deriving $P_{\text{E}}^{\text{over}}$ in Section 6.6.3.

First, spectrum availability states for cell overload in the basic area can be derived as follows:

$$\mathcal{L}_{\text{B}} = \{\arg_n [\sum_{j \in \mathcal{I}_n} B_i^{\max}(j) < N_i^{\text{b}}], \text{ for } \forall n\} \quad (96)$$

Based on overload states $n \in \mathcal{L}_{\text{B}}$, we derive the probability of inter-cell/inter-pool handoff in the basic area to resolve cell overload as follows:

$$u_{\text{B}}^{\text{req}}(n) = \frac{\min\left[\max\left[0, N_i^{\text{b}} - N_i^{\max}(n)\right], N_i^{\text{b}}\right]}{N_i^{\text{b}}} \quad (97)$$

The probabilities of cell underload and overload in the basic area, $P_{\text{B},i}^{\text{under}}$ and $P_{\text{B},i}^{\text{over}}$, can be obtained by replacing $\mathcal{L}_{\text{E}}$ and $u_{\text{E}}^{\text{req}}(n)$ with $\mathcal{L}_{\text{B}}$ and $u_{\text{B}}^{\text{req}}(n)$, respectively in Eqs. (91), (92), (93) and (94). Accordingly, the probability that the CR users in the basic area performs inter-cell/inter-poll handoff, $P_{\text{B}}^{\text{over}}$, is estimated as follow.

$$P_{\text{B}}^{\text{over}} = \sum_{r=1}^{R} P_{\text{B},i}^{\text{over}}(r\Delta t) \quad (98)$$

Unlike Eq. (95), we consider all spectrum bands, including the extended spectrum in this case. Thus, we do not need to consider the probability of spectrum availability in the extended area, $P^{\text{av}}(\cdot)$, separately.

6.6.5 Switching Cost

According to the probability on future mobility events, we estimate the switching cost of two possible options in the boundary of the basic area. First, when CR users stay in the current cell by performing intra-cell/intra-pool handoff to the extended area, the expected switching cost T_{EA} can be obtained as follows:

$$\begin{aligned} T_{\text{EA}} = D_1 &+ P_{\text{E}}^{\text{over}} \cdot D_2 \\ &+ (1 - P_{\text{E}}^{\text{over}})(1 - P_i^{\text{av}}(T_m)) \cdot D_4 \\ &+ P_i^{\text{av}}(T_m)(1 - P_i^{\text{av}}(T_m)) \cdot D_5 \end{aligned} \tag{99}$$

The total delay includes the intra-cell/intra-pool handoff when the CR user switches to the extended spectrum, inter-cell/inter-pool handoffs caused by cell overload and and PU activity, and inter-cell/intra-pool handoff when it is successfully handed over to the extended neighbors.

Second, when CR users move to the neighbor cell by performing inter-cell/inter-pool handoff, the expected switching cost can be expressed as the sum of the instant switching delay and the expected switching delay resulting from overload in that neighbor cell as follows:

$$T_{\text{BA}} = D_3 + D_1 \frac{T_m}{\overline{T}_{\text{off},i} + D_1} + P_{\text{B}}^{\text{over}} \cdot D_2 \tag{100}$$

The latency in this case includes the inter-cell/inter-pool handoff to a new target cell, intra-cell/intra-pool handoff in the target cell, and inter-cell/inter-pool handoff caused by cell overload. Here, the average number of intra-cell/intra-pool handoff is obtained as $T_m/(\overline{T}_{\text{off},i} + D_1)$. $\overline{T}_{\text{off},i}$ is the average idle period of the spectrum in cell i, which is expressed as the average of $1/\sum_{k \in \mathcal{A}_i^{\text{B}}(j)} \beta(j,k)$ over all spectrum j.

Table 5: Configuration of handoff delay components in simulations.

Components	d_{prep}	d_{recfg}	d_{lis}	$d_{\text{syn}}^{\text{sen}}$	d_{sen}	d_{dec}	$d_{\text{syn}}^{\text{tx}}$
Delay (sec)	0.1	0.3	0.1	0.025	0.025	0.025	0.025

Based on the analysis above, the BS determines the handoff type with the lower expected spectrum cost.

6.7 Performance Evaluation

6.7.1 Simulation Setup

To evaluate the performance of the proposed mobility management framework, we implement a network simulator to support the network topology consisting of multiple cells in 10km x 10km area. Here, we assume 59 cells which have different channel utilization. The transmission range of each cell is set to 750m. The interference range is set to twice larger than the transmission range. The transmission range of the extended spectrums is also twice larger than that of basic spectrums. Furthermore, we consider 4 spectrum pools that consists of 10 spectrum bands. The basic and extended spectrums can support 10 and 40 channels for users in the basic area, respectively (i.e., ρ is set to 4). Furthermore, each spectrum band has 3-5 PU activity regions, which have different PU activities, $\alpha(j,k)$ and $\beta(j,k)$ uniformly distributed in [0.01, 0.05]. The BSs are assumed to generate a false alarm every two hour on average when they sense the availability of each spectrum.

Furthermore, based on the delay components in Table 5, the handoff delays defined in Section 6.4, D_1, D_2, D_3, D_4, and D_5 are set to 0.2 0.5, 0.4, 1.2, and 0.1sec, respectively. An operational delay for inter-cell resource allocation, T^{inter}, is assumed to be 5sec.

In this simulation, we use a free space power attenuation model where the channel gain is set to -31.54dB, the reference distance is 1m, and the path loss coefficient is 3.5. The BS uses -56.21dBm/Hz transmission power on average for the basic spectrum and -47.18dBm/Hz for the extended spectrum. Noise power in the receiver is -174dBm/Hz. The minimum decodable SNR is set to 0dB. To describe user mobility, we consider a Gauss-Markov mobility model proposed in [48].

6.7.2 Performance of Inter-Cell Resource Allocation for Extended Spectrums

In this simulation, we evaluate the performance of the proposed inter-cell resource allocation by comparing with the following methods.

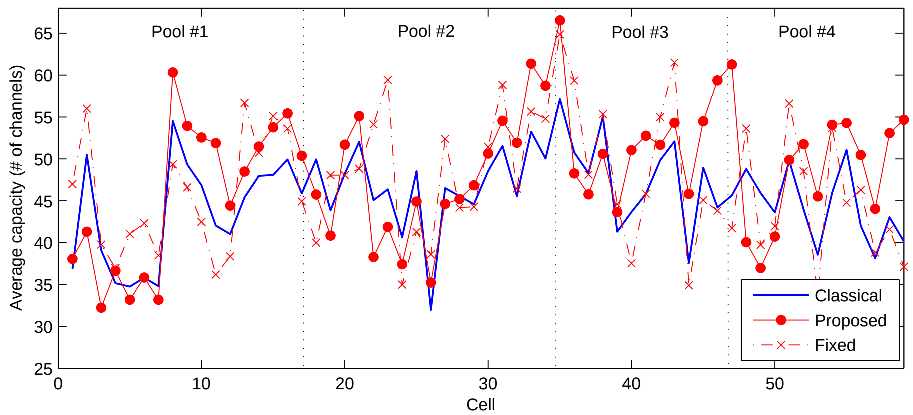
- *Classical handoff scheme:* This scheme does not support the extended spectrum. Thus, each cell is able to access all available spectrums in the pool without influence on its extended neighbors.

- *Highest capacity preferred scheme:* The BS selects the extended spectrum to maximize the total number of available channels in the network. A decision principle of this scheme is similar to Eq. (74), but does not consider a reliability metric $R_i(j)$.

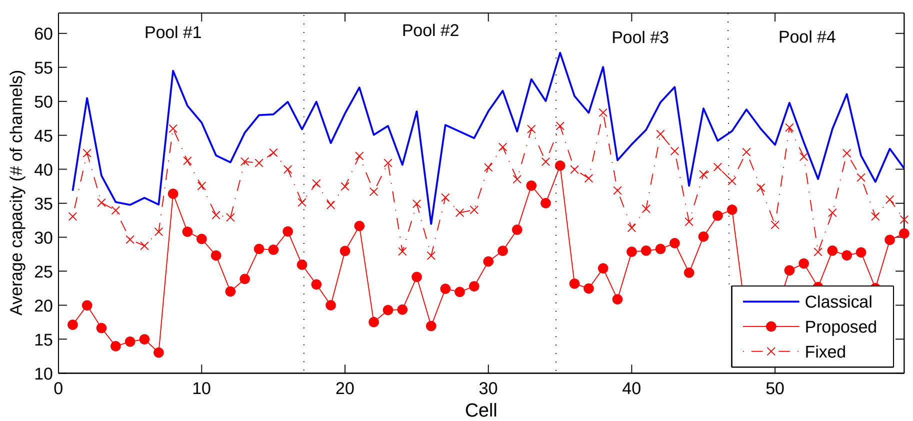
- *Highest availability preferred scheme:* The spectrum with the highest idle probability is selected for the extended spectrum, i.e., $G_i(j)$ is set to $\prod_{k\in\mathcal{A}_i^{\mathrm{E}}(j)} P_i^{\mathrm{off}}(j,k)$ in Eq. (75).

- *Fixed allocation:* This scheme is based the same decision criterion as the proposed method in Eq. (74). However, each cell is assigned to the predetermined extended spectrum bands based on the proposed method but does not change them, regardless of time-varying spectrum availability.

In Figure 51, we investigate the total spectrum availability of each scheme, i.e., total network capacity. If the extended spectrums are used, total network capacity is dependent on the location of users. Figure 51 (a) and (b) show the best case (i.e., all users using the extended spectrum are located on the basic area), and the



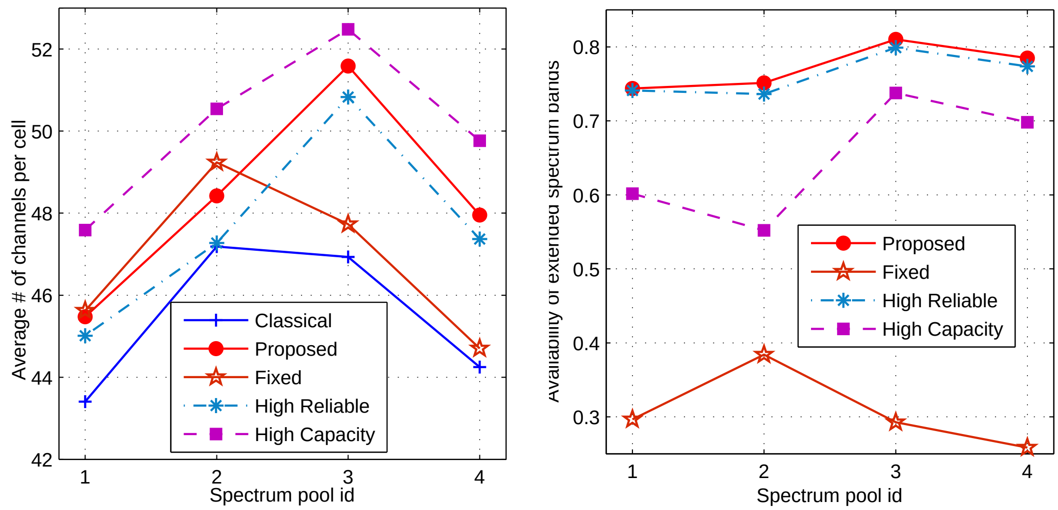
(a)



(b)

Figure 51: Average channel availability: (a) best case, and (b) worst case.

worst case (i.e., all users on the extended spectrum are located on the extended area), respectively. In the best case, both proposed and fixed methods show slightly higher capacity than the classical approach since the extended spectrum supports more channels to users in the basic area although it restrict the use of that spectrum in its extended neighbors. On the contrary, in the worst case, the classical method has much more available channels because the use of the extended spectrum in both proposed and fixed methods reduce the channel utilization in extended neighbor cells while users in the extended area require more channel resource for the same quality of service as users in basic area. In this case, since the proposed method has higher



(a) (b)

Figure 52: Performance comparison with other allocation schemes: (a) total available channels, and (b) availability in extended spectrums.

utilization of the extended spectrum, it shows the lowest number of available channels.

In Figure 52, we compare the proposed method with other decision principles in terms of total capacity (best case) and the availability of the extended spectrums. The availability of the extended spectrum is defined as the ratio of the time that the extended spectrum is valid for the cell to total simulation time. The highest capacity preferred method shows the highest total channel availability by reducing the effect on the rest of networks, but has trouble with finding more reliable extended spectrum. The highest availability preferred scheme shows lower capacity since it cause adverse influence on neighbor cells. In addition, since it only focuses on an overall idle probability of the spectrum without consideration of inter-cell operation delay, it may choose the spectrum requiring more frequent switching, leading to lower reliability in the extended spectrum than the proposed method. On the contrary, the proposed method shows the highest availability of the extended spectrum while maintaining higher capacity compared to the highest availability preferred and fixed methods by jointly considering capacity gain and reliability in the extended spectrum.

In summary, the use of the extended spectrum leads to lower network capacity compared with the classical methods but higher availability in the extended spectrum. However, it improves mobility support in CR cellular network, and hence allows the proposed method to achieve higher actual total capacity by reducing adverse effects of dynamic network environments, which will be shown in the subsequent simulations.

6.7.3 Performance of Spectrum and User Mobility Management Schemes

In this simulation, we investigate transmission statistics in mobile users under different network environments to evaluate the performance of both spectrum and mobility management schemes. To do this, we perform 30 one hour-simulations for each case and obtain average values. Here, we analyze the performance of mobility management in terms of three factors: user QoS requirement (i.e., how many channels are required for a current communication), current network load (i.e., how many channels are currently occupied by other users), and the velocity of mobile users.

Figure 53 shows the number of different mobility events in the proposed method. As user QoS requirement increases, the number of handoff types 2, 4, and 5, all of which involve in activities in the extended area, decrease since it reduces the probability to find enough resource in the extended area (Figure 53 (a)). Figure 53 (b) and (c) show the changes in handoff types according to network load. If the network is under-loaded, the number of a type 2 handoff is relatively lower than other type 3 because of lower handoff probability resulting from cell overload, as shown in Figure 53 (b). However, in a highly-loaded network, while the number of the type 1 handoff decreases, a drop rate becomes higher since some of PU activities may cause cell overload instead of successful spectrum switching. Furthermore, as network load increases, the number of all types of handoff decrease and conversely a drop rate increases because increase in cell overload probability reduces transmission opportunity, which is explained in Figure 53 (c). If user velocity increases, type 5

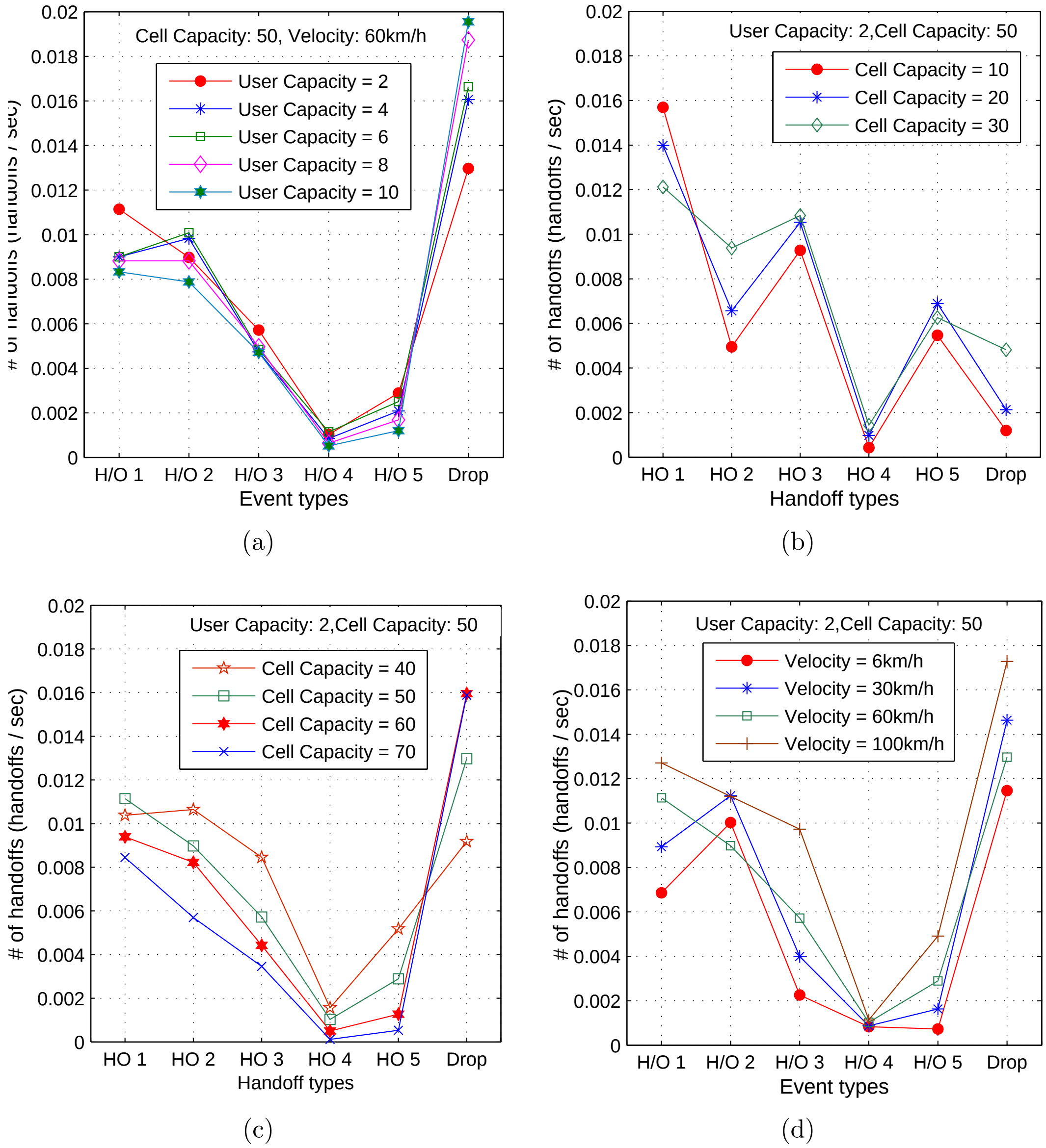


Figure 53: Handoff types in the proposed method (a) with different user capacity, (b) different cell occupancy (lower occupancy), (c) different cell occupancy (higher occupancy), and (d) different user velocity.

handoff to the extended area increases to reduce the abrupt quality degradation owing to frequent inter-cell/inter-pool handoff. In all cases, the proposed method keep the number of the worst handoff (type 4) a certain low level by intelligently choosing proper handoff types based on the expected switching delay.

One of the most important statistics in mobility management is the probability of a

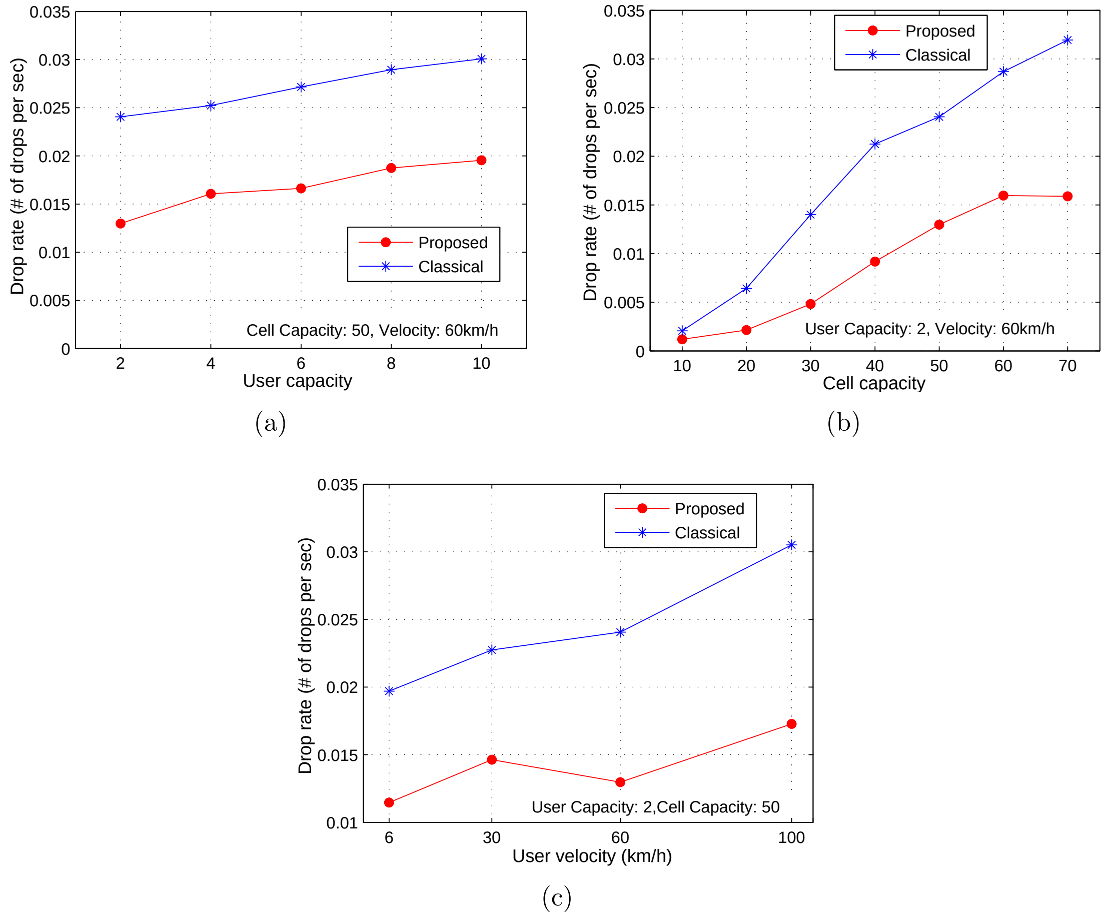


Figure 54: Drop rate in the proposed method (a) with different user capacity, (b) different cell occupancy, and (c) different user velocity.

call drop. The call drop occurs when a mobile user cannot find any available spectrum resources in both serving and target cells, which degrades the service quality of mobile users significantly. Here, we do not consider a call blocking probability. Figure 54 shows that the proposed method shows better performance in the drop rate than classical and fixed allocation methods. As shown in Figure 54 (a), although the user QoS requirement increases, the proposed method maintain a certain level of drop rate by exploiting spectrum mobility management. If the network load increases, a drop rate becomes higher due to the lack of available spectrum resources, but is still lower than classical method by exploiting different handoff types adaptively to cell conditions. Furthermore, the proposed method allows mobile users to adaptively use

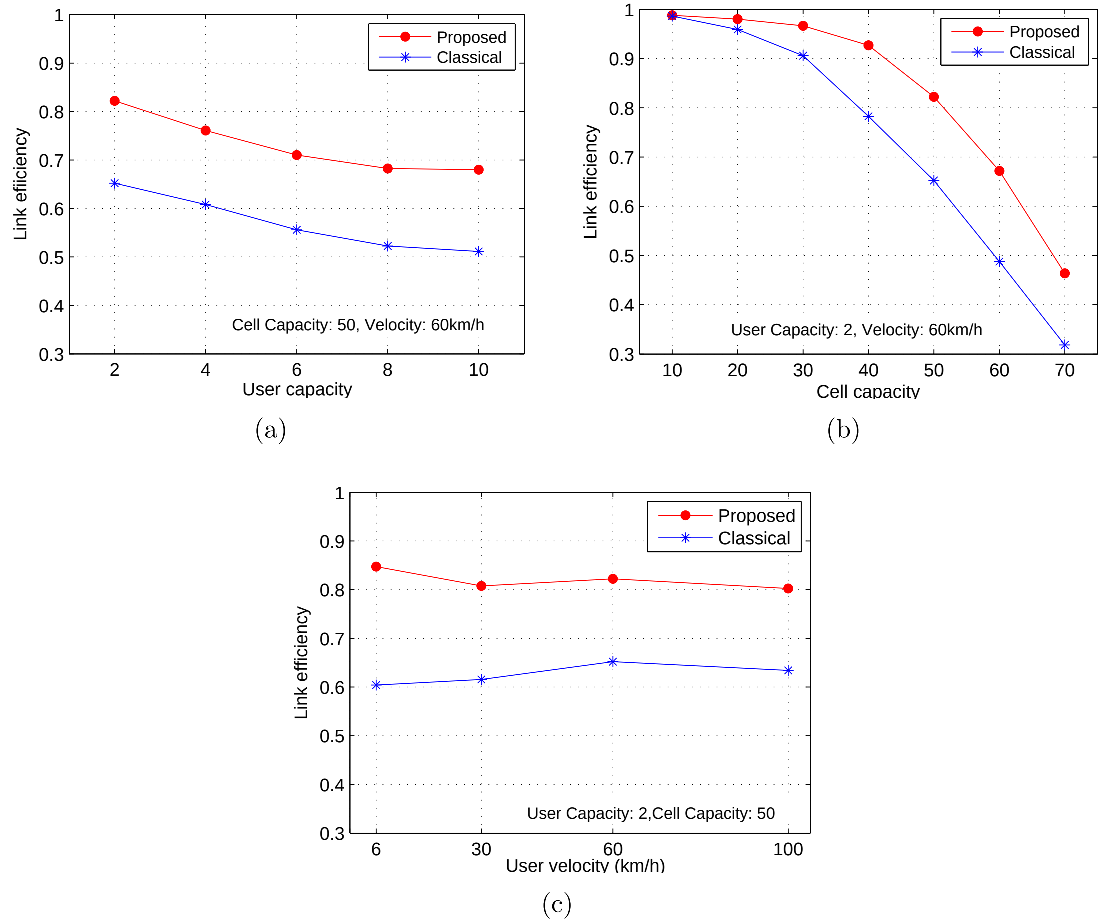


Figure 55: Link efficiency in the proposed method (a) with different user capacity, (b) different cell occupancy, and (c) different user velocity.

the extended area while reducing the number of inter-cell/inter-pool handoff through a user mobility management scheme. As a result, the proposed method sustains a lower drop rate although a mobile user traverse across wider areas and more cells boundaries with higher velocity, as shown in Figure 54 (c).

Figure 55 shows the link efficiency, which is defined as a real transmission time over an entire simulation time. In this simulation, the classical method shows lower link efficiency over all cases because of quality degradation caused by frequent inter-cell/inter-pool handoffs. Furthermore, when current cell is overloaded. some of mobile users cannot use spectrum resources until spectrum availability changes or they move into a new target cell area, which also reduces the link efficiency. On the contrary, the proposed method shows higher link efficiency by intelligently determine the handoff types to reduce the latency as well as a drop rate.

From these simulations, we can see that the proposed method achieves more actual transmission opportunity as well as less quality degradation during the transmission to mobile users, regardless of user and network conditions, although it shows lower network capacity theoretically because of the use of extended spectrums.

CHAPTER VII

CONCLUSIONS

7.1 Research Contributions

Cognitive radio (CR) technology is envisaged to solve the problems in wireless networks caused by the limited available spectrum and the inefficiency in the spectrum usage by exploiting the existing wireless spectrum opportunistically. CR networks, equipped with the intrinsic capabilities of the cognitive radio, will provide an ultimate spectrum-aware communication paradigm in wireless communications. CR networks, however, impose unique challenges because of the high fluctuation in the available spectrum as well as diverse quality-of-service (QoS) requirements. These key distinguishing factors necessitates a rethinking of the existing solutions developed for classical wireless networks.

In this thesis, a novel spectrum management framework is proposed to realize the goals of truly ubiquitous spectrum-aware communication. This framework enables CR devices to incorporate spectrum sensing, spectrum decision, spectrum sharing, and spectrum mobility functionalities. Research contributions have been made in the following areas:

1. Cross-layer spectrum management framework to enable seamless integration of unlicensed CR networks and pre-existing licensed primary networks without harmful interference,

2. Optimal spectrum sensing framework to optimize sensing and transmission times by considering both sensing efficiency and interference constraints, capacity of CR users while satisfying interference constraints to protect primary networks

3. QoS-aware spectrum decision framework to determine the best available spectrum according to QoS requirements,

4. Spectrum sharing framework in infrastructure-based CR networks to allocate the limited available spectrum resources efficiently to each cell,

5. Spectrum-aware mobility management in CR cellular networks to enable an intelligent switching of mobile users to the best combination of a target cell and spectrum.

In Chapter 2, intrinsic properties and current research challenges of spectrum management in CR networks are presented. In particular, we investigate novel spectrum management functionalities such as spectrum sensing, spectrum decision, spectrum sharing, and spectrum mobility from the viewpoint of both infrastructure-based network and ad hoc networks. In the infrastructure-based CR networks, all spectrum management functionalities are coordinated by a central network entities, i.e., CR base-stations make all decision on their actions based on the observations of each CR user. On the contrary, due to the lack of central network entities, CR ad hoc networks necessitate each CR user to have all the spectrum related CR capabilities, and determine its actions based on the local observation, leading to distributed operation. To overcome the drawback caused by the limited knowledge of the network, each of the these spectrum management functions relies on cooperative operations where CR users determine their actions based on exchanging information between the CR users.

In Chapter 3, we introduced the optimal sensing framework for cognitive radio networks that consists of three different functionalities. First, we proposed the sensing parameter optimization, which leads to the optimal transmission and observation time to maximize sensing efficiency satisfying the strict interference constraint of primary networks. Second, for the extension of multi-spectrum environment, we introduced a spectrum selection and scheduling algorithm based on the opportunistic capacity concept. Finally, we investigated how the cooperation sensing affects the performance of the proposed optimal sensing framework. To exploit the cooperative gain, we proposed an adaptive and cooperative sensing functionality mainly running on the centralized network entities such as a base-station. Furthermore, the simulation experiments show that the proposed sensing framework can achieve maximum sensing efficiency and opportunities in multi-user/multi-spectrum environments satisfying the interference constraints.

Chapter 4 addresses the problem of the spectrum decision in CR networks. We introduced a framework for spectrum decision to determine a set of spectrum bands by considering the channel dynamics in the CR network as well as application requirements. To this end, first, a novel spectrum capacity model is proposed that considers unique features in CR networks. Based on this capacity model, a minimum variance-based spectrum decision (MVSD) is developed for real-time applications, which determines the spectrums to minimize the capacity variance. For the best effort applications, a maximum capacity-based spectrum decision (MCSD) is proposed where spectrum bands are decided to maximize the total capacity. Moreover, a dynamic resource management scheme is introduced to enable the CR network to coordinate spectrum decision adaptively dependent on the time-varying spectrum resources. Simulation results show that the proposed spectrum decision framework provides efficient bandwidth utilization while guaranteeing the service quality.

In Chapter 5, we present a spectrum sharing framework in infrastructure-based CR networks. Although the exclusive method theoretically achieves optimal capacity, this approach cannot guarantee fair resource allocation that is also an important issue in inter-cell spectrum sharing. Furthermore, for optimal allocation, it requires spectrum utilization and topology information of the entire network, which leads to tremendous overhead and computational complexity. To solve these problems, first, we proposed novel spectrum allocation methods for both exclusive and common use models, which are dynamically exploited according to the QoS requirements, PU activities and current spectrum utilization. Both spectrum allocation schemes are closely collaborating with the proposed power allocation to protect the transmission of primary networks. In addition, we propose an intra-cell spectrum sharing method, where the base-station assigns the spectrum bands obtained through inter-cell spectrum sharing, to its CR users to maximize cell capacity as well as to avoid interference to primary networks. Simulation results show that the proposed framework achieves better performance in terms of network capacity, fairness. and QoS guarantees than classical methods.

In Chapter 6, we present a spectrum aware mobility management scheme for CR cellular networks. Because of the heterogeneous spectrum availability over time and space and discontiguous spectrum distribution over a wide frequency range, multi-cell based CR networks necessitate a novel mobility management framework to provide seamless and reliable communications to their mobile CR users. To this end, first, we propose the spectrum pool-based network architecture, which mitigates the heterogeneity in spectrum availability. Based on this architecture, a unified mobility management framework is defined to support diverse mobility events in CR networks, consisting of inter-cell resource allocation, and spectrum and user mobility management functions. Through inter-cell resource allocation, each cell determines its spectrum configuration to improve the mobility as well as total capacity. For the PU activity, spectrum mobility management is developed where the network determines the proper spectrums and target cells of each user according to both current spectrum utilization and the stochastic connectivity model. In user mobility management, the switching cost-based handoff decision mechanism is proposed to minimize quality degradation caused by user mobility as well as to maximize cell capacity. Simulation results show that the proposed methods provide maximum cell capacity while providing minimum quality degradation in mobile users.

7.2 *Future Research Directions*

Future wireless networking will be characterized by the increased presence of ubiquitous devices seamlessly embedded in the environment. These devices will constitute a cognitive and self-optimizing entity that senses, responds and adapts to the presence of people, objects, and to varying environmental characteristics. This new feature is enabled by extending current CR concept beyond spectrum management. Our future research covers the evolution into intelligent and self-optimizing CR networks from the perspective of each communication entity: network, service and user. Some of our planned research directions are outlined below.

- **Enabling Wireless Network Technologies (Network Perspective):** In this research, we plan to investigate the problems in designing an intelligent and self-optimizing wireless system, mainly focusing on the theoretical modeling of radio behaviors, architectural challenge, and deployment issue: First, we will develop a new PU activity model to capture the diverse characteristics of all the different types of existing primary networks. Based on this new model, we will develop an adaptive spectrum management scheme, which reconfigures its CR capabilities according to the physical layer technologies of primary networks. Second, the CR network architecture should support a well-established common control channel (CCC), which not only assists in disseminating broadcast messages, but also facilitates the cooperation among CR users. However, such a channel is difficult to be reliably established in CR networks because of the transmission of primary users. Therefore, in this research, I will systematically devise a mechanism to establish reliable CCCs, which is robust to PU activity and rapid topology changes. Finally, we plan to integrating the CR paradigm with current wireless technologies. The proliferation of wireless access points, and more recently the idea of femto-cells, has created an urgent need for interference coordination techniques. Interestingly, in the case of deployment of emto-cells, the interference should be self-managed in a distributed manner in licensed bands. In this research, we will develop CR femto-cell solutions, which will enable each femto-cell to control itself, to fit in with its network environment, and to provide seamless mobility along with interference management.

- **Service.Aware Protocol Design (Service Perspective:** Another concern is the efficient support of rich multimedia applications and services over dynamic wireless environment. Especially, the delivery and transport of multimedia to heterogeneous mobile users is very challenging, mainly due to the wireless channel unreliability, interference constraints, resource sharing among many users, limited bandwidth, different protocols and standards, etc. As a result, future multimedia services require innovation and advances in MAC and routing protocols, cross-layer interaction and optimization, QoS provisioning, adaptive transmission techniques, scalability support, and error correcting schemes. Drawing on my previous industrial experience in wireless multimedia services, we will develop a service-aware protocol design, which supports QoS guaranteed services with an intelligent adaptation to heterogeneous mobile device and network environments.

- **Auction Framework for Ubiquitous Connectivity (User Perspective):** The evolution in wireless communications has enabled the realization of various network architectures based on different technologies such as cellular networks, mesh networks, and wireless LANs. To extend the applicability of these architectures and provide useful information anytime and anywhere, the integration of these networks with Internet is an important challenge. Recently, an open mobile internet concept has been emerged to address this challenge, which enables devices to attach to any compatible network rather than being tied to a single provider. Furthermore, advances in the hardware technologies allow mobile devices automatically to detect the best network in range (cheapest, fastest, best optimized for a specific application, and so on). In this respect, we aim to develop an efficient, intelligent, and real-time auction and trading system to provide spectrum as well as wireless services driven by dynamic user demand and willingness-to-pay. In this framework, mobile devices broadcast their requests for service, and all available networks automatically return a list of price offers for that service.

APPENDIX A

CALCULATION OF THE LOST SPECTRUM OPPORTUNITY

The lost spectrum opportunity T_{L} can be obtained by the same procedure explained in Section 3.3.3. In the case of idle state sensing, the false alarm can introduce the loss of opportunities during transmission period T. If T is short, the opportunity is highly likely to be lost over the entire transmission period. Conversely, if T is long enough, the lost spectrum opportunity converges to $P_{\mathrm{off}} \cdot T$. Thus, the expected lost spectrum opportunity $\mathrm{E}[L_{\mathrm{off}}]$ can be obtained as follows:

$$\begin{aligned} \mathrm{E}[L_{\mathrm{off}}] &= P_{\mathrm{f}}(e^{-\mu T}T + (1-e^{-\mu T})P_{\mathrm{off}}T) \\ &= P_{\mathrm{off}}\bar{P}_{\mathrm{f}}(\frac{\beta}{\alpha+\beta}e^{-\mu T}T + \frac{\alpha}{\alpha+\beta}) \end{aligned} \tag{101}$$

where α and β represent the death and birth rates, respectively, and μ is $\max(\alpha, \beta)$. Similarly, the opportunity can be lost on busy state sensing only if there are one or more primary user activities during T, which converges approximately to the $P_{\mathrm{off}} \cdot T$ as follows:

$$\begin{aligned} \mathrm{E}[L_{on}] &= P_{\mathrm{d}}(e^{-\mu T}\cdot 0 + (1-e^{-\mu T})P_{\mathrm{off}}T) \\ &= (P_{\mathrm{on}} - \bar{P}_{\mathrm{f}})(1-e^{-\mu T})\frac{\alpha}{\alpha+\beta}T \end{aligned} \tag{102}$$

Thus, the expected lost spectrum opportunity, T_{L}, can be obtained as follows:

$$\begin{aligned} T_{\mathrm{L}} &= \frac{\mathrm{E}[L_{\mathrm{on}}] + \mathrm{E}[L_{\mathrm{off}}]}{T \cdot P_{\mathrm{off}}} \\ &= \frac{\beta}{\alpha}[e^{-\mu T}\bar{P}_{\mathrm{f}} + (1-e^{-\mu T})\frac{\alpha}{\alpha+\beta}] \end{aligned} \tag{103}$$

APPENDIX B

CALCULATION OF THE OBSERVATION TIME

Since we determine the threshold λ as the value to equalize both error probabilities, the detection error probability P_m can be represented as follows:

$$\begin{aligned} P_m &= P_{\text{on}}(1 - Q(\frac{\lambda - 2t_{\text{s}}W({\sigma_{\text{s}}}^2 + {\sigma_{\text{n}}}^2)}{\sqrt{4t_{\text{s}}W({\sigma_{\text{s}}}^2 + {\sigma_{\text{n}}}^2)^2}})) \\ &= P_{\text{on}}Q(\frac{2t_{\text{s}}W({\sigma_{\text{s}}}^2 + {\sigma_{\text{n}}}^2) - \lambda}{\sqrt{4t_{\text{s}}W({\sigma_{\text{s}}}^2 + {\sigma_{\text{n}}}^2)^2}}) \end{aligned} \tag{104}$$

From the false alarm probability P_{f} in Eq. (9), the threshold λ can be obtained as follows: From

$$\begin{aligned} \lambda &= \sqrt{4t_{\text{s}}W{\sigma_{\text{n}}}^4} \cdot Q^{-1}(\frac{P_{\text{f}}}{P_{\text{off}}}) + 2t_{\text{s}}W \\ &= \sqrt{4t_{\text{s}}W{\sigma_{\text{n}}}^4} \cdot Q^{-1}(\bar{P}_{\text{f}}) + 2t_{\text{s}}W \end{aligned} \tag{105}$$

Assume signal-to-noise ratio (SNR) $\gamma = {\sigma_{\text{s}}}^2/{\sigma_{\text{n}}}^2$. We can get another equation for threshold λ from the detection error probability P_{m} in Eq. (104) as follows:

$$\lambda = 2t_{\text{s}}W(\gamma + 1){\sigma_{\text{n}}}^2 - \sqrt{4t_{\text{s}}W}(\gamma + 1){\sigma_{\text{n}}}^2 Q^{-1}(\frac{P_{\text{off}}\bar{P}_{\text{f}}}{P_{\text{on}}}) \tag{106}$$

Since both equation should be the same, t_{s} can be represented as follows:

$$t_{\text{s}} = \frac{1}{W \cdot \gamma^2}[Q^{-1}(\bar{P}_{\text{f}}) + (\gamma + 1)Q^{-1}(\frac{P_{\text{off}}\bar{P}_{\text{f}}}{P_{\text{on}}})]^2 \tag{107}$$

APPENDIX C

DERIVATION OF THE DATA LOSS RATE IN COGNITIVE RADIO NETWORKS

In the CR network, each spectrum band has two discrete capacity states, 0 and $c_i(k) \cdot w_i(k)$ according to its PU activity, as explained in Section 4.3. Here $c_i(k)$ and $w_i(k)$ are the normalized capacity and the bandwidth of spectrum i for user k, respectively. Thus, when N spectrum bands are assigned to a CR user k, the total capacity $\mathbf{R_T(k)}$ has 2^N states according to the PU activities of the selected spectrum bands. Thus, each state m has the following state probability:

$$P_m(k) = \prod_{i \in \mathcal{I}_m} \frac{T_i^{\text{off}}}{T_i^{\text{off}} + \tau} \cdot \eta_i \prod_{i \in \mathcal{B}_m} (1 - \frac{T_i^{\text{off}}}{T_i^{\text{off}} + \tau} \cdot \eta_i) \tag{108}$$

where $\mathcal{I}_m$ and $\mathcal{B}_m$ are the sets of idle spectrum bands and busy spectrum bands at state m, respectively. T_i^{off} and η_i are the expected transmission time without switching and the transmission efficiency of spectrum band i, respectively. τ represents the spectrum switching delay.

Let the sustainable rate of user k be $R_{\text{s}}(k)$ and the capacity of each state m be $\widehat{R}_m(k)$. From the assumption that the data loss occurs when channel capacity is below $R_{\text{s}}(k)$, the data loss rate can be defined as the ratio of the expected capacity loss to the sustainable rate $R_{\text{s}}(k)$ as follows:

$$\begin{aligned} P_{\text{loss}}(k) &= \frac{R_{\text{s}}(k) - \sum_{m=1}^{2^N} \min(R_{\text{s}}(k), \widehat{R}_m(k)) P_m(k)}{R_{\text{s}}(k)} \\ &= \frac{\sum_{m=1}^{2^N} |R_{\text{s}}(k) - \widehat{R}_m(k)| P_m(k)}{2R_{\text{s}}(k)} \end{aligned} \tag{109}$$

APPENDIX D

DERIVATION OF THE CAPACITY VARIATION IN COGNITIVE RADIO NETWORKS

From the capacity state probability, derived in Eq. (108), the variance of the total capacity $\mathbf{R_T(k)}$ can be derived as follows:

$$Var[\mathbf{R_T(k)}] = \sum_{m=1}^{2^N} (\widehat{R}_m(k) - R_{\mathrm{S}}(k))^2 \cdot P_m(k) \tag{110}$$

By comparing the Eq. (109) with Eq. (110), we can see that the variance of the total capacity $Var[\mathbf{R_T(k)}]$ is proportional to the data loss rate $P_{\mathrm{loss}}(k)$. As a result, we can use the capacity variance for the resource allocation, instead of the data loss rate. To apply the variance in Eq. (110) for the optimization, we need another form of the variance expressed in terms of the bandwidth $w_i(k)$ and the normalized capacity $c_i(k)$ of each spectrum. Since spectrum bands are independent of each other, the variance of the total capacity in the selected spectrums can be expressed as follows:

$$\begin{aligned} Var[\mathbf{R_T(k)}] &= Var[\sum_{i \in \mathcal{S}} \mathbf{C_i(k)} \cdot w_i(k)] \\ &= \sum_{i \in \mathcal{S}} Var[\mathbf{C_i(k)} \cdot w_i(k)] \\ &= \sum_{i \in \mathcal{S}} (E[(\mathbf{C_i(k)} \cdot w_i(k))^2] - E[\mathbf{C_i(k)} \cdot w_i(k)]^2) \\ &= \sum_{i \in \mathcal{S}} ((c_i(k)^2 \cdot w_i(k)^2 \cdot \frac{T_i^{\mathrm{off}}}{T_i^{\mathrm{off}} + \tau} \cdot \eta_i - (c_i(k) \cdot w_i(k) \cdot \frac{T_i^{\mathrm{off}}}{T_i^{\mathrm{off}} + \tau} \cdot \eta_i)^2) \\ &= \sum_{i \in \mathcal{S}} \frac{T_i^{\mathrm{off}} \eta_i (T_i^{\mathrm{off}} + \tau - T_i^{\mathrm{off}} \eta_i)}{(T_i^{\mathrm{off}} + \tau)^2} c_i(k)^2 w_i(k)^2 \end{aligned} \tag{111}$$

where $\mathbf{C_i(k)}$ is the random variable to represent the capacity of spectrum i for user k. $\mathcal{S}$ is the set of the selected spectrum bands.

APPENDIX E

DERIVATION OF THE RESOURCE OUTAGE PROBABILITY

To model PU activities in the spectrum, we can use a 2 state Markov chain with the transition probabilities from idle to idle $x_i^{00} = 1 - e^{-\beta_i \Delta t}$, from idle to busy $x_i^{01} = e^{-\beta_i \Delta t}$, from busy to idle $x_i^{10} = e^{-\alpha_i \Delta t}$, and from busy to busy $x_i^{11} = 1 - e^{-\alpha_i \Delta t}$, where Δt is a sensing period. Then, the idle probability of spectrum i after $r\Delta t$, $P_i^{\mathrm{idle}}(r)$, can be expressed as either one of the following probabilities:

$$\begin{aligned} P_i^{\mathrm{i2i}}(r) &= \frac{x_i^{10}}{x_i^{01} + x_i^{10}} + (1 - x_i^{01} - x_i^{10})^r \cdot \frac{x^{01}}{x^{01} + x_i^{10}} \\ P_i^{\mathrm{i2b}}(r) &= \frac{x_i^{10}}{x_i^{01} + x_i^{10}} - (1 - x_i^{01} - x_i^{10})^r \cdot \frac{x^{10}}{x^{01} + x_i^{10}} \end{aligned} \tag{112}$$

where $P_i^{\mathrm{i2i}}(r)$ and $P_i^{\mathrm{i2b}}(r)$ are the expected idle probabilities after $r\Delta t$ when current spectrum states are idle and busy, respectively. If a false alarm probability P_i^{f} is considered, the idle probability of spectrum i can be expressed as either $(1-P_i^{\mathrm{f}})P_i^{\mathrm{i2i}}(r)$ or $(1 - P_i^{\mathrm{f}})P_i^{\mathrm{i2b}}(r)$.

Based on these probabilities, we derive the expected resource outage probability that $W^{\mathrm{av}} < W_{\mathrm{min}}$ as follows: Since the network has M spectrum bands, it has 2^M state according to the status of spectrum bands. Let $\mathcal{L}$ be a set of states that experience resource outage, i.e., available bandwidth W^{av} less than W_{min}. $\mathcal{I}_n$ represents a set of idle spectrum bands at state n. Then, resource outage happens when all spectrum bands in $\mathcal{I}_n, n \in \mathcal{L}$ are idle and the rest of spectrums $i \notin \mathcal{I}_n, n \in \mathcal{L}$ are busy. From this, the resource outage probability after $r\Delta t$, $P_{\mathrm{out}}(r)$ can be derived as follows:

$$P_{\mathrm{out}}(r) = \sum_{n \in \mathcal{L}} \prod_{i \in \mathcal{I}_n} P_i^{\mathrm{idle}}(r) \prod_{i \notin \mathcal{I}_n} (1 - P_i^{\mathrm{idle}}(r)) \tag{113}$$

Based on this probability, we can obtain the expected resource outage probability during $r\Delta t$, P_{out} as $\sum_{r'=1}^{r} P_{\text{out}}(r')/r$.

REFERENCES

[1] Akyildiz, I. F., Lee, W.-Y., and Chowdhury, K., "CRAHNs: Cognitive radio ad hoc networks," *Ad Hoc Networks Journal (Elsevier)*, July 2009.

[2] Akyildiz, I. F., Lee, W.-Y., and Chowdhury, K., "Spectrum management in cognitive radio ad hoc networks," *IEEE Network*, vol. 23, pp. 6–12, July 2009.

[3] Akyildiz, I. F., Lee, W.-Y., Vuran, M. C., and Mohanty, S., "Next generation / dynamic spectrum access / cognitive radio wireless networks: a survey," *Computer Networks Journal (Elsevier)*, vol. 50, pp. 2127–2159, September 2006.

[4] Akyildiz, I. F., Lee, W.-Y., Vuran, M. C., and Mohanty, S., "A survey on spectrum management in cognitive radio networks," *IEEE Communications Magazine*, vol. 46, pp. 40–48, April 2008.

[5] Akyildiz, I. F., McNair, J., HO, J. S. M., Uzunalioglu, H., and Wang, W., "Mobility management in next-generation wireless systems," *Proceedings of the IEEE*, vol. 87, no. 8, pp. 1347–1384, 1999.

[6] Amzallag, D., Bar-Yehuda, R., Raz, D., and Scalosub, G., "Cell selection in 4G cellular networks," in *Proc. IEEE Infocom 2008, Pheonix, AZ, USA*, pp. 1373–1381, April 2008.

[7] Buddhikot, M., Kolody, P., Miller, S., Ryan, K., and Evans, J., "DIMSUMNet: new directions in wireless networking using coordinated dynamic spectrum access," in *Proc. IEEE WoWMoM 2005*, p. 7885., June 2005.

[8] Cabric, D., Mishra, S. M., and Brodersen, R. W., "Implementation issues in spectrum sensing for cognitive radios," in *Proc. 38th Asilomar Conference on Signals, Systems and Computers 2004*, pp. 772–776, November 2004.

[9] Cabric, D., Mishra, S. M., Willkomm, D., Brodersen, R., and Wolisz, A., "A cognitive radio approach for usage of virtual unlicensed spectrum," in *Proc. 14th IST Mobile and Wireless Communications Summit*, June 2005.

[10] Cao, L. and Zheng, H., "Distributed spectrum allocation via local bargaining," in *Proc. IEEE Sensor and Ad Hoc Communications and Networks (SECON)*, pp. 475–486, September 2005.

[11] Cao, L. and Zheng, H., "Distributed rule-regulated spectrum sharing," *IEEE Journal on Selected Areas In Communications*, vol. 26, pp. 130–145, January 2008.

[12] Chari, M. R., Ling, F., Mantravadi, A., Krishnamoorthi, R., Vijayan, R., Walker, G. K., and Chandhok, R., "FLO physical layer: an overview," *IEEE Transactions on Broadcasting*, vol. 53, pp. 45–159, march 2007.

[13] Chen, T., Zhang, H., Maggio, G. M., and Chlamtac, I., "CogMesh: a cluster-based cognitive radio network," in *Proc. IEEE DySPAN 2007*, pp. 168–178, April 2007.

[14] Chen, Y., Zhao, Q., and Swami, A., "Joint design and separation principle for opportunistic spectrum access," in *Proc. IEEE Asilomar Conference on Signals, Systems and Computers 2006*, October 2006.

[15] Chou, C., Shankar, S., Kim, H., and Shin, K. G., "What and how much to gain by spectrum agility?," *IEEE Journal on Selected Areas in Communications*, vol. 25, pp. 576–588, April 2007.

[16] Cordeiro, C. and Challapali, K., "C-MAC: a cognitive MAC protocol for multi-channel wireless networks," in *Proc. IEEE DySPAN 2007*, p. 147157, April 2007.

[17] Cox, D. R., *Renewal Theory.* New York: John Wiley & Sons Inc, 1962.

[18] Digham, F. F., Alouini, M. S., and Simon, M. K., "On the energy detection of unknown signals over fading channels," in *Proc. IEEE ICC 2003*, pp. 3575–3579, May 2003.

[19] Etkin, R., Parekh, A., and Tse, D., "Spectrum sharing for unlicensed bands," *IEEE Journal on Selected Areas in Communications*, vol. 25, pp. 517–526, April 2007.

[20] Evans, J. R. and Minieka, E., *Optimization algorithms for networks and graphs.* CRC Press, second ed., 1992.

[21] FCC, "ET docket no 02-135," *Spectrum Poicy Take Force (SPTF) Report*, November 2002.

[22] FCC, "ET docket no 03-222," *Notice of Proposed Rule Making and Order*, December 2003.

[23] Gandetto, M. and Regazzoni, C., "Spectrum sensing: A distributed approach for cognitive terminals," *IEEE Journal on Selected Areas in Communications*, vol. 25, pp. 546–557, April 2007.

[24] Gandhi, S., Buragohain, C., Cao, L., Zheng, H., and Suri, S., "A general framework for wireless spectrum auctions," in *Proc. IEEE DySPAN 2007*, April 2007.

[25] Geirhofer, S., Tong, L., and Sadler, B., "Dynamic spectrum access in the time domain: modeling and exploiting white space," *IEEE Communications Magazine*, vol. 45, pp. 66–72, May 2007.

[26] Ghasemi, A. and Sousa, E. S., "Collaborative spectrum sensing for opportunistic access in fading environment," in *Proc. IEEE DySPAN 2005*, pp. 131–136, November 2005.

[27] Ghasemi, A. and Sousa, E. S., "Optimization of spectrum sensing for opportunistic spectrum access in cognitive radio networks," in *Proc. IEEE Consumer Communications and Networking Conference (CCNC) 2007, Las Vegas, NV, USA*, pp. 1022–1026, January 2007.

[28] Hanly, S. V., "An algorithm for combined cell-site selection and power control to maximize cellular spread spectrum capacity," *IEEE Journal on Selected Areas in Communications*, vol. 13, pp. 1332–1340, September 1995.

[29] Haykin, S., "Cognitive radio: Brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, pp. 210–220, February 2005.

[30] Huang, J., Berry, R. A., and Honig, M. L., "Spectrum sharing with distributed interference compensation," in *Proc. IEEE DySPAN 2005*, pp. 88–93, November 2005.

[31] Hur, Y., Park, J., Woo, W., Lee, J. S., Lim, K., Lee, C.-H., Kim, H. S., and Laskar, J., "A cognitive radio (CR) system employing a dual-stage spectrum sensing technique: A multi-resolution spectrum sensing (MRSS) and a temporal signature detection (TSD) technique," in *Proc. IEEE Globecom 2006*, November 2006.

[32] IEEE 802.22 WRAN WG, "Draft standard for wireless regional area networks part 22: Cognitive wireless RAN medium access control (MAC) and physical layer (PHY) specifications: Policies and procedures for operation in the TV bands," Tech. Rep. IEEE P802.22/D0.3.8.1, IEEE, September 2007.

[33] Ileri, O., Samardzija, D., and Mandayam, N. B., "Dynamic property rights spectrum access: Flexible ownership based spectrum management," in *Proc. IEEE DySPAN 2007*, April 2007.

[34] Ji, Z. and Liu, K. J. R., "Dynamic spectrum sharing: A game theoretical overview," *IEEE Communications Magazine*, vol. 45, pp. 88–94, May 2007.

[35] Jia, J., Zhang, Q., and Shen, X., "HC-MAC: a hardware-constrained cognitive MAC for efficient spectrum management," *IEEE Journal on Selected Areas in Communications*, vol. 26, p. 106117, January 2008.

[36] Jondral, F. K., "Software-defined radio - basic and evolution to cognitive radio," *EURASIP Journal on Wireless Communication and Networking*, 2005.

[37] Jung, E. and Liu, X., "Opportunistic spectrum access in heterogeneous user environments," in *Proc. IEEE DySPAN 2008, Chicago, IL, USA*, October 2008.

[38] K. Xu, M. G. and Bae, S., "How effective is the IEEE 802.11 RTS/CTS handshake in ad hoc networks?," in *Proc. IEEE Global Telecommunications Conferebce (Globecom) 2002*, vol. 1, pp. 72–76, November 2002.

[39] Kang, X., Liang, Y., Nallanathan, A., Garg, H., and Zhiang, R., "Optimal power allocation for fading channels in cr networks: Ergodic capacity and outage capacity," *IEEE Transactions on Wireless Communications*, vol. 8, pp. 940–950, February 2009.

[40] Kannan, D., *An introduction to stochastic Processes.* North Holland, 1979.

[41] Kim, H. and Shin, K. G., "Efficient discovery of spectrum opportunities with mac-layer sensing in cognitive radio networks," *IEEE Transactions on Mobile Computing*, vol. 7, pp. 533–545, May 2008.

[42] Kim, H. and Shin, K. G., "Fast discovery of spectrum opportunities in cognitive radio networks," in *Proc. IEEE DySPAN 2008, Chicago, IL, USA*, October 2008.

[43] Lee, W.-Y. and Akyildiz, I. F., "Joint spectrum and power allocation for inter-cell spectrum sharing in cognitive radio networks," in *in Proc. IEEE DySPAN 2008, Chicago, IL, USA*, October 2008.

[44] Lee, W.-Y. and Akyildiz, I. F., "Optimal spectrum sensing framework for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 3845–3857, October 2008.

[45] Lee, W.-Y. and Akyildiz, I. F., "Spectrum-aware mobility management for cognitive radio cellular networks," *submitted for publication*, 2009.

[46] Lee, W.-Y. and Akyildiz, I. F., "A spectrum decision framework for cognitive radio networks," *submitted for publication, revised*, July 2009.

[47] Lee, W.-Y. and Akyildiz, I. F., "Spectrum sharing framework for intrastructure-based cr networks," *submitted for publication*, 2009.

[48] Liang, D. and Haas, Z. J., "Predictive distance-based mobility management for pcs networks," in *Proc. of IEEE Infocom, New York, NY, USA*, vol. 3, pp. 1377–1384, March 1999.

[49] Liang, Y. C., Zeng, Y., Peh, E., and Hoang, A. T., "Sensing-throughput tradeoff for cognitive radio networks," *IEEE Transactions on Wireless Communications*, vol. 7, pp. 1326–1337, April 2008.

[50] Luo, L. and Roy, S., "Analysis of search schemes in cognitive radio," in *Proc. IEEE SECON 2007, San Diego, CA, USA*, pp. 647–654, June 2007.

[51] Ma, H., Zheng, L., Ma, X., and Luo, Y., "Spectrum-aware routing for multihop cognitive radio networks with a single transceiver," in *Proc. Cognitive Radio Oriented Wireless Networks and Communications (CrownCom) 2008*, May 2008.

[52] Mishra, S. M., Sahai, A., and Brodersen, R. W., "Cooperative sensing among cognitive radios," in *Proc. IEEE ICC 2006*, vol. 4, pp. 1658–1663, June 2006.

[53] Mitola, J., "Cognitive radio for flexible mobile multimedia communication," in *Proc. IEEE Int'l Workshop on Mobile Multimedia Communications (MoMuC) 1999*, pp. 3–10, November 1999.

[54] Oner, M. and Jondral, F., "On the extraction of the channel allocation information in spectrum pooling systems," *IEEE Journal on Selected Areas in Communications*, vol. 25, p. 558565, April 2007.

[55] Pei, Y., Hoang, A. T., and Liang, Y.-C., "Sensing throughput tradeoff in cognitive radio networks: how frequently should spectrum sensing be carried out?," in *Proc. IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC) 2007, Athens, Greece*, July 2007.

[56] Peng, C., Zheng, H., and Zhao, B. Y., "Utilization and fairness in spectrum assignment for opportunistic spectrum access," *ACM Mobile Networks and Applications (MONET)*, 2006.

[57] Proakis, J. G., *Digital communications*. McGRAW-Hill, fourth ed., 2001.

[58] Raman, C., Yates, R. D., and Mandayam, N. B., "Scheduling variable rate links via a spectrum server," in *Proc. IEEE DySPAN 2005*, pp. 110–118, November 2005.

[59] Ramsdale, P. A. and Harrold, W. B., "Techniques for cellular networks incorporating microcells," in *Proc. IEEE PIMRC 1992, Boston, MA, USA*, pp. 169–173, October 1992.

[60] Rappaport, T., *Wireless communications: principles and practice*. Prentice Hall, 2 ed., 2001.

[61] Sang, A., Wang, X., Madihian, M., and Gitlin, R. D., "Coordinated load balancing, handoff/cell-site selection, and scheduling in multi-cell packet data systems," in *Proc. ACM MobiCom 2004, Philadelphia, PA, USA*, pp. 302–314, September 2004.

[62] Schrijver, A., *Theory of linear and integer programming*. John Wiley and Sons, 1998.

[63] Shankar, S., "Spectrum agile radios: Utilization and sensing architecture," in *Proc. IEEE DySPAN 2005*, pp. 160–169, November 2005.

[64] Shannon, C. E., "Communication in the presence of noise," *Proceedings of the IEEE*, vol. 72, pp. 1192–1201, September 1984.

[65] Song, G. and Li, Y. G., “Utility-based resource allocation and scheduling in ofdm-based wireless broadband networks,” *IEEE Communications Magazine*, pp. 127–134, December 2005.

[66] Sriram, K. and Whitt, W., “Characterizing superposition arrival processes in packet multiplexers for voice and data,” *IEEE Journal on Selected Areas in Communications*, vol. SAC-4, pp. 833–846, September 1986.

[67] Svedman, P., Wilson, S. K., and Ottersten, B., “A QoS-aware proportional fair scheduler for opportunistic OFDM,” in *Proc. IEEE Vehicular Technology Conference 2004*, pp. 558–562, September 2004.

[68] Tang, H., “Some physical layer issues of wide-band cognitive radio system,” in *Proc. IEEE DySPAN 2005*, pp. 151–159, November 2005.

[69] Tanga, J., Misrab, S., and Xueb, G., “Joint spectrum allocation and scheduling for fair spectrum sharing in cognitive radio wireless networks,” *Computer Networks Journal (Elsevier)*, vol. 52, pp. 2148–2158, August 2008.

[70] Wang, P., Xiao, L., Zhou, S., and Wang, J., “Optimization of detection time for channel efficiency in cognitive radio systems,” in *Proc. IEEE Wireless Communications and Networking Conference (WCNC) 2007, Hong Kong, China*, pp. 111–115, March 2007.

[71] Wang, Q. and Zheng, H., “Route and spectrum selection in dynamic spectrum networkstions,” in *Proc. IEEE Consumer Communica Communications and Networking Conference (CNCC) 2006*, January 2006.

[72] Wang, W. and Zhao, M., “Joint effects of radio channels and node mobility on link dynamics in wireless networks,” in *Proc. IEEE Infocom 2008, Pheonix, AZ, USA*, pp. 933–941, April 2008.

[73] Weiss, T. A. and Jondral, F. K., “Spectrum pooling: an innovative strategy for the enhancement of spectrum efficiency,” *IEEE Communications Magazine*, vol. 42, pp. 8–14, March 2004.

[74] Wild, B. and Ramchandran, K., “Detecting primary receivers for cognitive radio applications,” in *Proc. IEEE DySPAN 2005*, pp. 124–130, November 2005.

[75] Willkomm, D., Machiraju, S., Bolot, J., and Wolisz, A., “Primary users in cellular networks: a large-scale measurement study,” in *Proc. IEEE DySPAN 2008, Chicago, IL, USA, October 2008*, 2008.

[76] Yang, L., Cao, L., and Zheng, H., “Proactive channel access in dynamic spectrum network,” in *Proc. Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM) 2007, Orlando, FL, USA*, p. 482486, ,August 2007.

[77] Zhang, L., Liang, Y., and Xin, Y., "Joint beamforming and power allocation for multiple access channels in cognitive radio networks," *IEEE Journal on Selected Areas in Communications*, vol. 26, pp. 38–51, january 2008.

[78] Zhao, Q., Tong, L., Swami, A., and Chen, Y., "Decentrallized cognitive mac opportunistic spectrum access in ad hoc networks: A pomdp framework," *IEEE Journal on Selected Areas in Communications*, vol. 25, pp. 589–600, April 2007.

[79] Zhou, X., Gandhi, S., Suri, S., and Zheng, H., "eBay in the Sky- strategy-proof wireless spectrum auctions," in *Proc. ACM Mobicom 2008*, pp. 2–13, September 2008.

VITA

Won Yeol Lee was born in Daegu, Korea, November 1971. He received his B.S. and M.S. degrees from Department of Electronics Engineering, Yonsei University, Seoul, Korea in 1997 and 1999, respectively. From 1999 to 2004, he was a research engineer of Network R&D Center and Wireless Multimedia Service Division at LG Telecom, Seoul, Korea. He received Ph.D. degree in Electrical and Computer Engineering from Georgia Institute of Technology, Atlanta, GA. in 2009, under the supervision of Prof. Ian F. Akyildiz. In 2008, he received the BWN Lab Researcher of the Year 2008 Award. His current research interests include cognitive radio networks, next generation wireless communications, and wireless sensor networks.