博士論文

A Study of Efficient Spectrum Utilization by Cooperation of Cognitive and Heterogeneous Wireless Networks

(コグニティブネットワークとヘテロジニアスネットワークの協調によるスペクトルの効率的利用に関する研究)

東京大学大学院　情報理工学系研究科

48-087422　李　云

A Study of Efficient Spectrum Utilization by Cooperation of Cognitive and Heterogeneous Wireless Networks

by
YUN LI

DISSERTATION
Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Department of Information and Communication Engineering, Graduate School of Information Science and Technology,
The University of Tokyo

Tokyo, Japan

Preface

The cognitive cell network and the boosting growth of 5G cell network need to explore more spectrum access opportunities to meet ever-growing traffic demand, which necessitates the use of more spectrum data to identify radio environment. This thesis highlights how to encourage Cooperation / Interaction between users in cognitive radio network (CRN) and heterogeneous cell network (HetNet) to recognize radio environment, and proposes a model to optimize the spectrum usages of the mobile network (HetNet) as well as CRN by letting HetNet act as the common control channel for cooperative spectrum sensing of CRN. The throughput of both CRN and HetNet can be improved by designing a proper incentive scheme based on coalitional game theory.

The current advancement of wireless services has resulted in higher requirements for radio frequency resources in wireless networks. The optimum spectrum usage is difficult because of two reasons. The first is that spectrum management has been a politically as well as a technically sensitive issue and it is usually difficult to access the open data of real spectrum usage. The second is that the spectrum usage optimization has been investigated within each network architecture independently from the other networks. CRN is the first architecture for the second user to share the same frequency bands as the primary user if they are not used. Although setting up stable common control channels among nodes has been recognized as difficult within the framework of CRN, no research works have been done to solve this problem using other networks. Firstly this thesis proposes a general framework to solve this problem for an environment that CRN and HetNet are cooperative with each other, where HetNet acts as the common control channel for cooperative spectrum sensing of CRN. The dynamics of CRN and HetNet can be determined by coalitional game theory, and the throughput of these networks can be improved by giving nodes in cooperation proper incentives.

For the difficulty to access the open data of real spectrum usage, this thesis proposes grass-root based spectrum data collections, called SpectrumMap, through large-scale users' participation. Frequency awareness is shown to be important to improve the performance of dynamic spectrum access. It demonstrates the SpectrumMap database implemented in the laboratory environment and discusses in

details the challenge encountered during the system development introduced by the computation complexity of the big spectrum data. Then, the spectrum measurement campaign is conducted in Tokyo area, Japan during several years to show the effectiveness of SpectrumMap. The duty cycle based spectrum utilization for typical wireless bands is shown, with the collected raw data spread from $75MHz$ to $3GHz$. Specifically, as a study case, this part reveals spatio-temporal characteristic of Wi-Fi occupancy in the downtown Tokyo with a carefully arranged measurement schedule. It is shown that spectrum is not fully utilized; even in the crowded Wi-Fi channels observed in downtown areas.

To show how inefficient it is without any cooperation among users, this thesis analyzes multi-hop cognitive network with three restrictions. The first is no cooperation among users. The second is no common control channel between users. The third is no synchronization between users. The existing research works did not consider all these restrictions at a time. Different from the existing studies, this thesis proposes a new method to effectively reduce the decision space of users, thus can substantially reduce the complexity of decision making by the network. However, the conducted result shows that cognitive network can only obtain limited data rate when considering energy constrain.

Considering the crowdsourcing strength of public users, a theoretical model is proposed to incentivize building of the SpectrumMap database through user participation based spectrum data collections. Different from the above case, users are assumed to have a common control channel to a network center, and spectrum data collections are realized by users' cooperation. The validity of the collected data is verified by the simulation with a scenario of cognitive access to white space channels for a cell network. The dynamics of cooperation is determined by coalitional game theory, and the credits to buy white space channel resources when necessary can be improved by giving proper incentives to nodes in cooperation. The proposed data collection model is expected to work as a supplement for the conventional commercial spectrum database to improve the awareness of radio environment.

The above framework is extended to the framework for the collaboration between HetNet and CRN. CRN and HetNet are assumed to have incentive to cooperate with each other, i.e., cognitive users can use small cells in HetNet acting as relay and report the sensed data to SpectrumMap database, while small cells can find good channels by utilizing the sensing capability of cognitive radio users. This thesis formulates the cooperative behavior by utilizing coalitional game theory. Results show that, to achieve the same throughput performance for small cell, the proposed approach can reduce channel resources from 160 to 30 channels to achieve success channel access probability of 0.8, and cognitive users can improve throughput about 10% .

Finally this thesis tries to improve the downlink transmission of HetNet since

it is degraded when HetNet acts as a common control channel for CRN. This part investigates spectrum efficiency for multiple-input multiple-output (MIMO) wireless network, and proposes precoding strategy to improve spectrum efficiency of MIMO. The proposed precoding design strategy for MIMO Relaying Broadcast Channel with coordinated users can overcome the relay's half-duplex constraint and achieve the full degree of freedom (DoF) to improve the frequency efficiency. Furthermore, it evaluates a weighted minimum mean square error (WMMSE) precoding design method to jointly design the precoder of base station and beamforming matrix of relay station according to throughput and fairness criterions. Given base station equipped with N_b antennas, it has been shown that the existing scheme without considering the direct users [1] suffers from the degrees of freedom loss, and the maximum degrees of freedom is N_b, while maximum degrees of freedom of the proposed scheme with the designed precoding strategy is 2 N_b. The conducted simulation has confirmed the effectiveness of the proposed scheme.

[1] the base station keep silence during the phase that relay station transmits data streams to users.

Contents

List of Figures

List of Tables

Acknowledgement

This thesis could not be completed without the support of others. I am indebted for this support and I would like to thank the following people in particular.

First of all, my sincere gratitude goes to my supervisor Prof. Tohru Asami of Department of Information and Communication Engineering, Graduate School of Information Science and Technology, The University of Tokyo, Japan. Prof. Asami has provided me thoughtful supervision and insightful comments, which greatly helps me to achieve the research targets during my whole studies.

Also, I would like to thank Prof. Hitoshi Aida, Prof. Hiroshi Esaki, Prof. Hiroyuki Morikawa, Prof. Kaoru Sezaki, Prof. Yoshihiro Kawahara, from Department of Information and Communication Engineering, Graduate School of Information Science and Technology, The University of Tokyo, Japan, for their valuable comments and encouragement, as well as research method.

In addition, I would like to express my gratitude to Prof. Wen Chen, the Department of Electronic Engineering, Shanghai Jiao Tong University, China, and Prof. Haibin Wan, the College of Computer Science and Electronic Information, Guangxi University, China, Prof. Honggang Zhang, the College of Information Science and Electronic Engineering, Zhejiang University, China, for their patience and guidance throughout my studies.

Furthermore, I want to specially thank Prof. Tatsuro Takahashi and Prof. Ryoichi Shinkuma, from Department of Communications and Computer Engineering, Kyoto University, Japan, for their encouragement during my study in Japan.

I also appreciate research support from other members of Asami & Kawahara Laboratory: Takuya Kato, Trung Tran Quang, Kei Kusunoki, and foreign visiting student, Philip Düe, and other members. All of you have been there to support me for my Ph.D. thesis. I am also very grateful for the help from Meiko Fujita who is at the office of Asami & Kawahara Laboratory and always assist me the final formatting and printing of my work.

Finally I want to thank my family for the support to perform my study.

Chapter 1

Introduction

1.1 Background

The recent improvement of wireless services has resulted in higher requirements for radio frequency resources in wireless networks, and accordingly, spectrum management has been a politically as well as a technically sensitive issue. The cognitive cell network and the boosting growth of 5G cell network need to explore more spectrum access opportunities to meet ever-growing traffic demand, which necessitates the use of more spectrum data to identify radio environment. However, It is usually difficult to access such open data of real spectrum usage.

1.2 Current Situations and Approach of Research

1.2.1 Spectrum database for efficient frequency resource usages

Building a spectrum database in a relatively open access model can promote network users to understand about radio environments. Currently, several standards toward the secondary access under TV white space channels has been worked by the research community, i.e., the 802.22 [12] and 802.11af [13] of IEEE. Thus the existed work associated with exploiting spectrum diversity in TV white space channels mainly advocate the centralized spectrum database approach dominated by authorized organization to derive the distribution of spectrum utilization, based on propagation models from real-world data logs [29, 31–35].

1.2.2 Energy efficient spectrum sensing

One of the most important targets in Cognitive Radio Network protocol design is how to efficiently use limited energy to increase the network throughput. With the energy constraint, the secondary user will choose to work or sleep in order to conserve energy. A number of studies have examined energy-efficient strategies for Cognitive Radio Networks. In [59], Zhao *et al.* proposed decentralized cognitive protocols that search for spectrum opportunities dynamically. Owing to the hardware and energy constraints, their protocols assumed that each secondary user can sense only a limited number of channels in each slot. [60] used the same method and developed optimal distributed MAC protocols for an energy constrained Cognitive Radio Network. The authors established that the optimal sensing and access policies have threshold structures, i.e., a secondary user will sense a channel when the conditional probability of the channel idle state exceeds some threshold value and will access a channel when the sensing outcome shows that the channel is not occupied and its fading condition is better than some threshold value. [61] attempted to improve both spectrum and energy efficiencies and proposed a diversity technology to use channel resources and energy. The above studies focus primarily on a single-hop wireless access network.

1.2.3 Cooperative spectrum sensing

In cooperative spectrum sensing, common control channel is required to realize efficient signaling protocols to identify spectrum opportunity. It is however difficult to construct common control channel in large-scale cognitive radio networks due to dynamic environments of primary users as well as frequency usage efficiency. Most works about cooperative spectrum sensing mainly rely on cluster-based, sequence-based and dedicated common control channel schemes [75]. With these techniques, at the convergent time, cognitive radio users in centralized cooperative spectrum sensing tend to choose one common control channel, while cognitive radio users in distributed cooperative spectrum sensing consume a different common control channel for each available cognitive radio user pair [4]. Thus the spectrum usage is inefficient, and as for the case of centralized cooperative spectrum sensing, this results in a bottleneck, or a single point of failure.

1.2.4 Existing Works for Self-Organization of Femtocell

Self-organization is a method for each femtocell to by itself configure the physical cell, transmission frequency and power, using the acquired information on the wireless environments measured by cellphones in the cell. Existing self-organization for femtocell includes dealing with opportunistic spectrum access

[77], power management [78] and interference mitigation [79], etc. Some techniques depend on the cognitive capability of femtocell itself. However the data collectors are cellphones within each cell and no information on the wireless environments outside the cell is gathered. Thus the optimized cells are difficult to configure.

1.2.5 Existing Works for Spectrum Efficiency for MIMO Relaying Broadcast Channel

Wireless cooperative communication can improve network capacity, service coverage, and reducing energy consumption by using relays [91]. However, half duplex constraint, the inability of relay node to transmit and receive signals simultaneously at the same frequency, is a major potential weakness for relay techniques, which introduces inefficient use of system bandwidth resources due to the necessity of extra dedicated bandwidth for relay retransmissions [92]. The existing works have studied multiple-input multiple-output (MIMO) non-regenerative relaying broadcast channel to design effective source precoding and relay processing matrices. However, their studies have only considered that the source is inactive in the retransmission phase of the relay, which causes half loss of the degrees of freedom (DoF) in the two transmission phases.

1.3 Organization of This Thesis

To solve the issues listed in the previous sections, the goal of this research is to consolidate a configuration methodology of a cooperation system between Femtocell and Cognitive Users, forming Femtocell-aided Cooperative Spectrum Sensing (CSS) (left) and CU-aided self-organization (right), as shown in Fig. 1.1. CSS relies on users in femtocells or heterogeneous network (HetNet) acting as no common control channel. Based on coalitional game theory, the collaboration between two different service users provides them large benefit from the view point of self-organization by Femtocell as well as cooperative spectrum sensing by Cognitive Radio users. To create common control channels, cognitive radio network will ask for the cooperation with HetNet, so as to utilizing the channel resources of HetNet to act the common control channel. Accordingly, it is necessary to promote HetNet user to construct the cooperation with cognitive radio network by useful incentive scheme. Thus, HetNet agrees with cognitive radio network for the cooperation and, as reward, HetNet users can improve their self-organization capabilities by joining the game with cognitive radio users and choosing optimal channels via the sensed channel usage information from cognitive users. As a result, both cognitive radio users and HetNet users can achieve improved throughput.

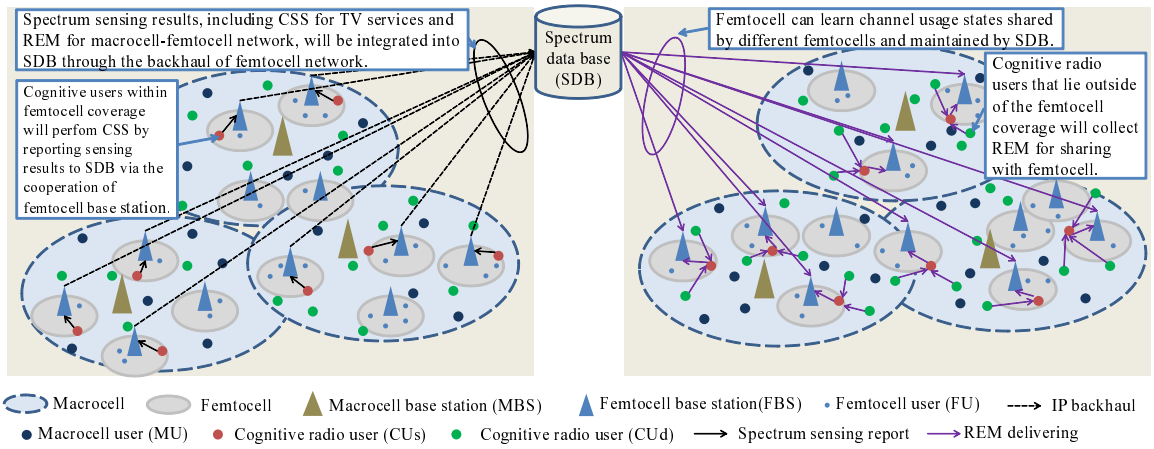

Figure 1.1: Femtocell-aided CSS (left) and CU-aided self-organization (right). CSS is jointly performed by multiple $CU_s s$ and the distributed sensing results are fused into the FC of SDB, lied at the gateway of femtocell network, through the relay of FBSs; Multiple FBSs share the same SDB. CU_d is responsible for collecting REM which will be shared with femtocell for self-organization. The SDB enables large-scale sharing of spectrum sensing and REM exchanging.

The organization of this thesis is as follows.

Chapter 2 proposes SpectrumMap, a grass-rooted spectrum data collection and visualization system for users to know the real spectrum utilization. One of the results of construction of this system shows that spectrum is not fully utilized even for the crowded Wi-Fi channels.

Chapter 3 shows the theoretical limit for the throughput of spectrum sensing policy by independent sensing, and based on Zhao's single-hop sensing /transmission model, proposes a new method to effectively reduce the decision space of users, thus can improve computational efficiency. This method achieves a good approximation of the true optimal schedule, and is extendable to multi-hop transmission.

Chapter 4 investigates how to encourage users to join the spectrum data collections since cooperative spectrum data collections can improve network sensing performance. Based on coalitional game theory, the proposed mechanism is proved to be very effective in the case of small uploading probability of collected data. However, the proposed cooperative spectrum sensing uses a cellular-phone network as the common control channels.

To encourage cellular phone or HetNet users to join the spectrum data collections by Cognitive Radio users, Chapter 5 proposes a coalitional game theory-based model with two incentive schemes specialized for HetNet and Cognitive Ra-

dio users. It is proved there exists a cooperation to maximize the self-organization network (SON) performance as well as the Cognitive Radio network throughput.

To fulfil HetNet users' desire to get more channel resources for common control channel, Chapter 6 proposes a new MIMO precoding scheme to let the base station repeat the same data transmission during the relay-to-user transmission. This is an extension of Wan' s half-duplex relaying model between base station and user [101]. The precoding scheme outperforms the Wan' s method from the points of throughput and fairness conditions.

Chapter 7 summarizes this thesis.

1.4 Terminology

This thesis frequently uses the following terms.

Table 1.1: Key Terms-1

N	total number of nodes (excluding the sink node).
s_i	the ith sensor.
W	a set of nodes in the network ($W = \{s_1, s_2, \ldots, s_N\}$).
x_0	the initial state consisting of all initial states (idle/busy) of channels and the battery levels.
x_t	states of all channels (idle/busy) and battery levels of sensors at time t.
B	battery capacity of each node.
$\mathcal{B}$	a set used to describe the energy state space of sensors ($\mathcal{B} = \{0, 1, 2, \ldots, B\}$).
$b_{s_i,t}$	energy level of sensor s_i at the beginning of time slot t .
e	permissible average energy usage of a sensor denoted by B/T, where the battery capacity is B and the operating time is T.
$\mathcal{A}$	a set of decisions of sensors ($\mathcal{A} = \{0, 1\}$, where '1' represents sensing a channel, and '0' represents not sensing a channel).
z_0, z_1	z_0 indicates the value of '0', and z_1 indicates the value of '1'.
$a_{s_i,t}$	decision of sensor s_i at time slot t ($a_{s_i,t} \in \mathcal{A}$).
a_t	optimal action of a node at time slot t ($a_t \in \mathcal{A}$).
$c_{s_i,s_j,t}$	channel state between s_i and s_j at time slot t.
$\mathcal{S}$	a set used to describe the channel state space of sensors, i.e., busy or idle, at a specified time slot.
x'_t	channel's state ($x'_t \in \mathcal{S}$).
$o_{s_i,s_j,t}$	packet transmission state between s_i and s_j .
$r_{s_i,s_j,t}$	packet reception state between s_i and s_j .
π_T	a set of decisions by every node in every slot, where the operating time is T.
$G_T(\pi)$	average reception rate of a node under policy π, where the operating time is T.
$U_{s_i,T}(\pi)$	average power consumption of node s_i, where the operating time is T.
π_∞	a set of decisions within an infinite operating time (Assume that the electrical energy of a node is infinite. The difference between π_T and π_∞ is: the π_T should take the channel state space and energy state space into the consideration, i.e., $\mathcal{S} \times \mathcal{B}$; while the π_∞ only concentrates on channel state space $\mathcal{S}$.).
G_∞	average reception rate of a node within an infinite operating time (Assume that the electrical energy of the node is infinite.).
U_∞	average power consumption of a node within an infinite operating time (Assume that the electrical energy of the node is infinite.).

Table 1.2: Key Terms-2

$G_T(\pi_T)$	average packet reception rate within time T, under the constraint that the total energy storage of a node is B.
$U_T(\pi_T)$	average power consumption within time T, under the constraint that the total energy storage of the node is B.
$G_\infty(\pi_\infty)$	average packet reception rate within an infinite operating time (Assume that the electrical energy of a node is infinite.).
$U_\infty(\pi_\infty)$	average power consumption within an infinite operating time (Assume that the electrical energy of a node is infinite.).
$U_T(\pi_\infty), G_T(\pi_\infty)$	the values obtained when the policy π_∞ is only executed within the time slot of $[0, T]$.
$U_\infty(\pi_T), G_\infty(\pi_T)$	the values obtained when the policy π_T is executed with infinite operating time (Apply the policy π_T within the time slot of $[0, T]$, then reuse the same policy π_T within the next time slot of $[T+1, 2T]$, and repeat this process by applying the π_T during each of the next period.).
$T^o_{s_1,s_2}(T^o_{s_2,sink})$	the total time of data transmission from node s_1 (node s_2) to node s_2 (node *sink*).
$B^o_{s_1,s_2}(B^o_{s_2,sink})$	the total energy consumption of node s_1 (node s_2) for transmission.
$T^r_{s_1,s_2}$	the total time of data reception from node s_1 to node s_2.
$B^r_{s_1,s_2}$	the total energy consumption of node s_2 for reception.
$\pi^*_{T^o_{s_1,s_2}}$	the optimal transmission strategy of node s_1.
$\pi^*_{T^r_{s_1,s_2}}$	the optimal receipt strategy of node s_2.
$\pi^*_{T^o_{s_2,sink}}$	the optimal transmission strategy of node s_2.
CRN	cognitive radio network.
$HetNet$	heterogeneous cell network.
CSS	cooperative spectrum sensing.
CU_s	cognitive user that lies outside the coverage of Femto cell.
CU_d	cognitive user that lies inside the coverage of Femto cell.
BS	base station.
RS	relay station.
MBS	Macrocell base station.
FBS	Femtocell base station.
REM	radio environment data.
DU	direct user.
RU	relay user.
$CSIT$	imperfect channel state information at the transmitter.
$WMMSE$	weighted minimum mean square error.

Chapter 2

Grassroots Spectrum Measurement

2.1 Introduction

Wireless networks are usually required to perform self-organization due to radio resources constraint, i.e., to wisely determine frequency allocations and adaptively improve the system performance according to the network environment. Cognitive radio (CR) based technology is a significant potential to improve spectrum access and, currently, geolocation database based approach has been approved by FCC for TV white space access. Radio Environment Map (REM) has been considered as a kind of improved geolocation database which holds relatively static information about signal power temperature map collected from spectrum monitors and can be used for improving spectrum utilization for wireless networks. Recently, heterogeneous cell networks has also attracted considerable attention to deal with the coverage of next-generation wireless cellular networks and enhance network capacity to meet the increasing wireless data traffics. Due to the large density of deployment of small cell base stations, interference management becomes the main challenge. To this end, improving spectrum access efficiency is a significant issue for cell networks, such as cognitive radio network (CRN), Wi-Fi [1] as well as heterogeneous cell network, etc.

Due to the spatio-temporal fleetingness of the user behavior in primary (licensed) networks, it is usually difficult for cognitive network users to achieve the expected network performance. However, the empirical time and frequency domain knowledge regarding primary channel activities can allow the cognitive network to learn the radio environment and choose the candidate channels that is more likely in better quality. In this context, the more spectrum raw data, the better performance can be expectable. Hence, it is necessary to build a large-scale

[1]The Wi-Fi includes two kinds : conventional Wi-Fi, such as 2.4GHz based, and the super Wi-Fi which utilizes TV white space channels.

spectrum database. Currently, several standards toward the secondary access under TV white space channels has been worked by the research community, i.e., the 802.22 [12] and 802.11af [13] of IEEE. The existing rules apply a spectrum database approach to regulate the spectrum access of secondary users.

However, observing radio environment data by deploying large-scale dedicated spectrum measurements is subject to the constraint of spectrum monitor cost, and accordingly, it is usually difficult for public users to access the open data of real spectrum usage. Building a spectrum database in a relatively open access model can promote network users to understand radio environments. In this sense, the presented work in this chapter believes grass-root based spectrum data collections is important and proposes a SpectrumMap database to enable public users to collect and share the radio environment data, and consequently to promote network users to aware radio environment.

In this work, there are several challenges associated with the SpectrumMap:

1. Challenges due to spectrum measurement: A wide-area spectrum monitoring should be performed in accordance with a well defined measurement schedule, since spectrum data will reflect the typical life model of human activities.

2. Challenges due to spectrum data visualization: the system needs to provide a flexible and user-friendly interface, which is however is a non-trivial task due to the heavy computation delay among computer servers introduced by huge size of spectrum datasets.

3. Challenges due to user participation: the SpectrumMap database is designed to utilizing crowdsourcing strength by attracting a large number of users' participation for collecting data, which necessitates the designing of an effective incentive mechanism.

To this end, this part provides the following main contributions:

- This work has designed a grass-root based SpectrumMap database, which has been implemented in the laboratory environment. The user interface of the SpectrumMap supports spectrum data uploading and visualization. The framework is expected to improve the radio environment awareness and spectrum usage efficiency for cognitive network.

- This work has implemented a distributed Hadoop based system for SpectrumMap to improve the computing performance due to the collected big size of the data, and it can assist the rapid spectrum usage visualization in accordance with flexible query conditions specified by end users.

- This work has performed large-scale spectrum measurement campaigns continued in three years in Tokyo. A brief summary of the typical bandwidths utilization formulated by duty cycle is reported, and the results indicate that most channels are not well utilized.

The rest of this work is organized as follows: In section 2, this work shows some related works. In section 3, this work describes the SpectrumMap Database architecture, including the data visualizing function. The work also presents a solution to solve the big spectrum data processing via Hadoop system. In subsection 4, this work gives a brief summary about typical bandwidth utilization of Tokyo city as well as a study case of Wi-Fi occupancy.

2.2 Related Works

Concerned with efficient spectrum access, previous works mainly focus on spectrum sensing techniques, such as the widely studied cooperative spectrum sensing [4] by the research community. Besides spectrum sensing based approach, learning spectrum usage in terms of signal strength propagated by primary networks through empirical spectrum measurements has also gained attention, such as the works of [5], [6], [7]. Additionally, there are some researches have made efforts on interference improvement for heterogeneous cell network, i.e., the work of [8], [9], [10]. The work of [11] has concentrated the attention on heterogeneous cloud radio access networks for cooperative radio resource management.

The existing work associated with exploiting spectrum diversity in TV white space channels mainly advocate the centralized spectrum database approach dominated by authorized organization to derive the distribution of spectrum utilization, based on propagation models from real-world data logs [29, 31–35]. To be distinguished from this approach, the SpectrumMap proposed in this work is more close to network users, since it leverages the crowdsourced spectrum monitor, i.e., the grass-root based data collection, to facility data contribution and sharing in large-scale space and time dimensions. More specifically, the spectrum measurements can be dedicated, i.e., the long-term spectrum observation with better accuracy by spectrum analyzers arranged in a well predefined measurement schedule, or relatively opportunistic, i.e., mobile devices equipped with sensors recognize surrounding radio environment and report local spectrum measurements at distributed locations. It should be noted that conventional spectrum database can be attributed to the dedicated spectrum data collection, which is inevitably subject to the cost constraint in the case of mass deployments. The proposed SpectrumMap database tends to exploit the strength of public users to realize wide-area of spectrum sensing. Furthermore, besides assisting service coexisting in TV white space

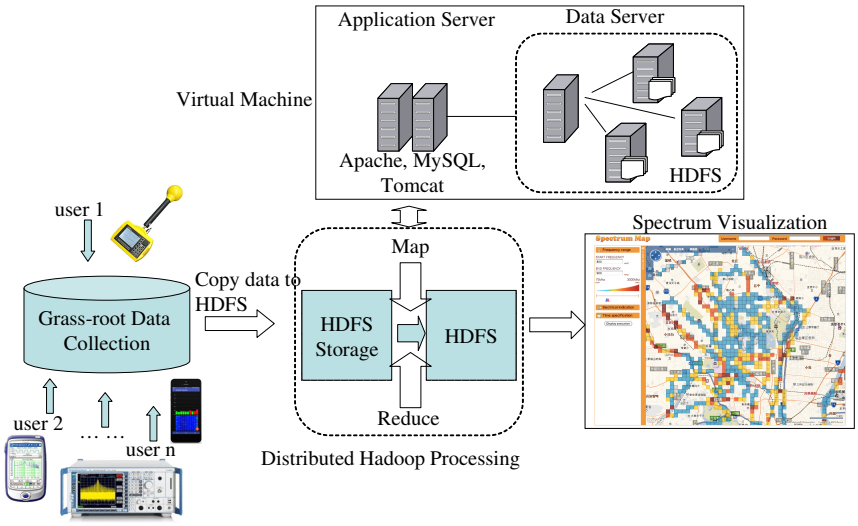

Figure 2.1: System architecture of SpectrumMap

channels, the proposed SpectrumMap framework is applicable to other kinds of radio resource management, such as optimizing self-organization of heterogeneous LTE (Long-Term Evolution) network.

2.3 SpectrumMap Database Architecture

This section describes the system architecture and important functions of SpectrumMap.

2.3.1 Prototype of SpectrumMap

SpectrumMap is designed in a web-based access form to be flexibly interacted with the public with the authority permission. The system architecture is shown in Fig. 2.1, which consists of three components: data collection, data processing, and data visualization. Spectrum data contributors, i.e., specified or volunteer organization, public network users such as Wi-Fi users, HetNet users, cognitive radio users, etc., observe spectrum occupancy at predetermined frequency bands [2], then upload the data into the data center of the *SpectrumMap*. Accordingly, the data center can perform data computing and analyze the channel usage characteristic of the target band.

[2]Since no prior information about primary networks is available, the energy detection is a good candidate.

2.3.2 Requirements of SpectrumMap

SpectrumMap is a mashup database to visualize the spectrum utilization at each geographical location, and the data is collected in a grass-roots manner due to its shortage, which requires the scalability of the data description format on account of large volume data sets. The data include the average frequency occupancy over all the locations, the maximum/minimum frequency occupancy, and signals which cross a power threshold in given locations. In addition, an overview of wide area spectrum utilization requires a large number of data formats at multiple locations, such as longitude, latitude, frequency value and power, which requires SpectrumMap to have high scalability and flexibility for data storage and analysis. SpectrumMap should satisfy the following functions:

1. Fix the measurement cite and capture the varied radio information in three dimensions, such as time (second), frequency(GHz) and spectrum intensity (V/m).

2. Changing measurement cite and obtain the spectrum in three dimensions for each position, with variable time (second), frequency (GHz) and the corresponded spectrum intensity (V/m).

3. Explore all the the satisfied radio information of location, time (second) and frequency band (GHz) by a predetermined threshold level of spectrum intensity (V/m).

4. Given a specified date-time range and threshold level (V/m), then obtain all the satisfied radio information indicated by other elements, i.e. location, frequency (GHz).

There are other kinds of requirements regarding to the measurement activities for SpectrumMap. First, it needs location identification, which requires spectrum monitors to have GPS function. It should be noted that the effect of GPS accuracy on channel usage mining is application dependent, i.e., almost $50m$ location identification error or even more is acceptable for determining TV white space channels, due to its propagation characteristic. Next, the heterogeneity of radiation meter should be addressed, since a variety of sensorized devices may be utilized by users for spectrum measurements [3]. Still, it needs to guarantee the time synchronization of location information and spectrum data, since in most environments, the measurement components of SpectrumMap is separable, which necessitates the time synchronization among the multiple devices.

[3]For this, a significant consideration is to apply metadata with XML schema to describe spectrum data formats, which enables diverse data exchanges from public users to enlarge the spectrum database.

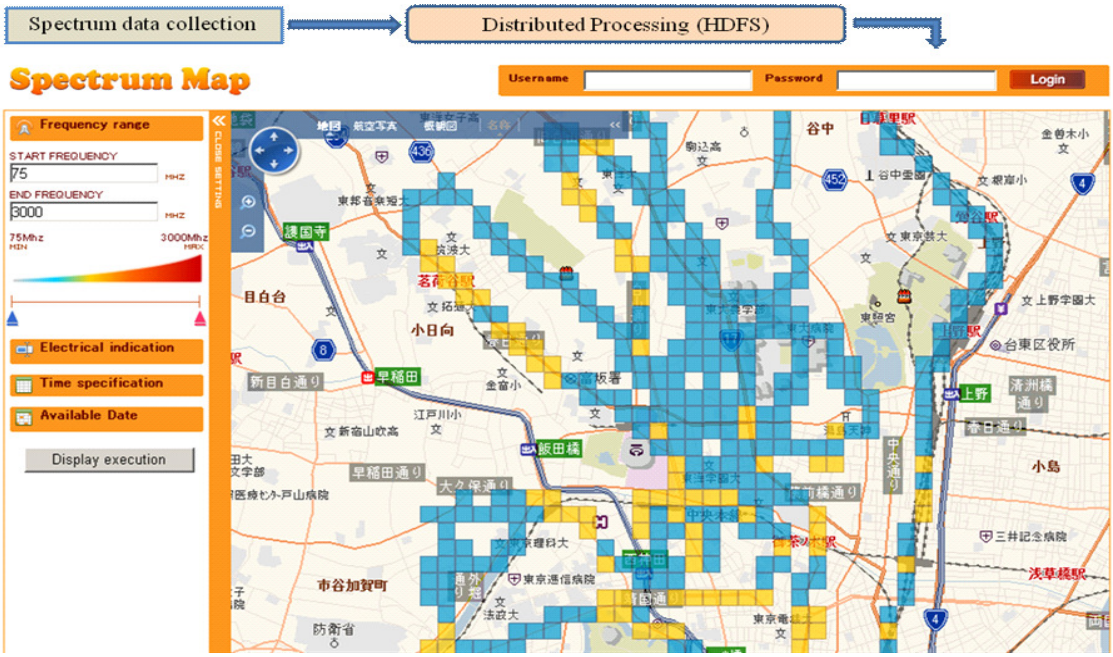

Figure 2.2: User interference of SpectrumMap (The geographical granularity of the map can be adjusted and accordingly, the spectrum intensity level of each grid will be recalculated [36–38].)

2.3.3 Demonstration of SpectrumMap

This part demonstrates the SpectrumMap implemented in laboratory environment. To the best of the author's knowledge, this work is the first time to successfully visualized the spectrum usage in central Tokyo. The accumulated log data is about $10GB$, with $69.8M$ records including $60K$ unique GPS points in 3 years (2009 – 2012). The user interface of SpectrumMap is shown in Fig. 2.2.

Principle Functions: SpectrumMap has multiple optional functions to characterize the radio information . Denoted by Y_{max}, Y_{ave-i}, and Y_{ave-u} the maximum value, average E-field intensity and average E-field utilization, the following methods can be applied to formulate the spectrum utilization:

- Maximum E-field intensity

$$\begin{aligned} Y_{ij} &= max(Y_{ijk}), i \in (1,n), j \in (1,m), k \in (1,l), \\ Y_i &= max(Y_{ij}), i \in (1,n), \quad j \in (1,m), \\ Y_{max} &= max(Y_i), \quad i \in (1,n). \end{aligned}$$

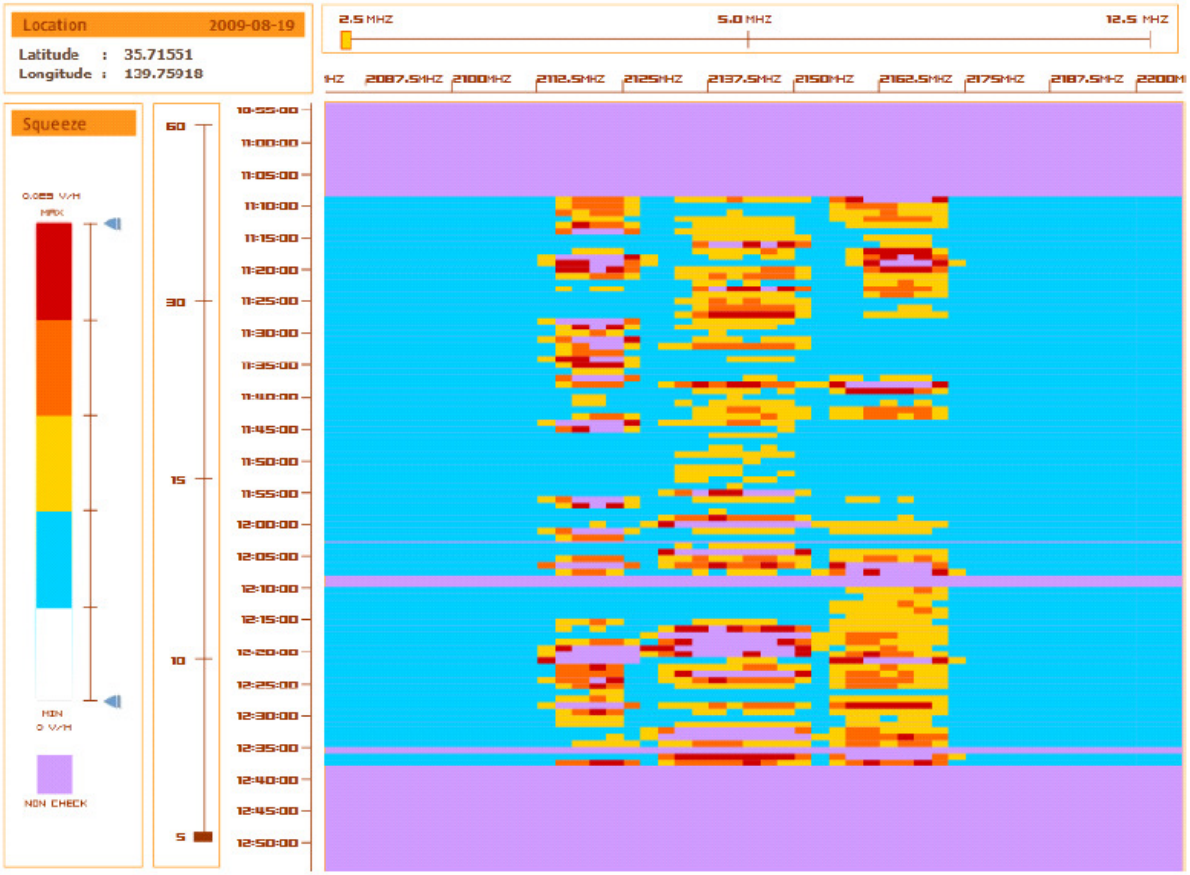

Figure 2.3: Spectrum usage for a certain location on the specified date, with changeable time and frequency granularity [38].

- Average E-field intensity

$$
\begin{aligned}
Y_{ij} &= \frac{1}{l}\sum_{k=1}^{l} Y_{ijk}, i \in (1,n), j \in (1,m), k \in (1,l), \\
Y_{i} &= \frac{1}{m}\sum_{j=1}^{m} Y_{ij}, \quad i \in (1,n), j \in (1,m), \\
Y_{ave-i} &= \frac{1}{n}\sum_{i=1}^{n} Y_{ij}, \quad i \in (1,n).
\end{aligned}
$$

- Average E-field utilization

$$
\begin{aligned}
Y_{ij} &= \frac{1}{l}\sum_{k=1}^{l}(\delta_{Y_{ijk}}), i \in (1,n), j \in (1,m), k \in (1,l), \\
&\quad (\delta_{Y_{ijk}} = 1 \; if \; Y_{ijk} > Threshold) \\
Y_i &= \frac{1}{m}\sum_{j=1}^{m} Y_{ij}, \quad i \in (1,n), j \in (1,m), \\
Y_{ave-u} &= \frac{1}{n}\sum_{i=1}^{n} Y_{ij}, \quad i \in (1,n).
\end{aligned}
$$

In the above, Y_{ijk} indicates the observed data, where i is the number of measurement points in one grid on the map (from 1 to n), j is the number of scan for one point(from 1 to m), and k is the number of data selected from start-frequency to end-frequency for one scan(from 1 to l).

User Interface of SpectrumMap: The user access interface consists of two main panels: control panel and visualized (Fig. 2.2). The visualized panel is constructed by Visual Earth of Microsoft and overlaid by a transparent layer of color tier to show the spectrum intensity. Each grid indicates a target area and the spectrum utilization is marked with different colors. The control panel includes the following elements: (1)The method of calculation: Avg: the average value of spectrum intensity; Max: the max value of the spectrum intensity; Sum: the sum value of the spectrum intensity. (2) Frequency Range: indicated by start frequency as well as end frequency, which can be freely determined by users; (3)Electrical indication: use a bar with different colors to indicate the different spectrum intensity level; (4) Time specification: a pop menu consisting of a calendar panel inside and user can choose the desired daytime; (5) Available Date: a pop menu including series of date information, on which spectrum measurement has been performed; (5) Display execution: perform data calculation according to user's settings and display the result in the visualized panel. In Fig. 2.2, a scenario executed by SpemctrumMap will be given and the real spectrum usage is derived near hongo-sanchome of Tokyo. Besides the above user interface, SpectrumMap is capable to demonstrate more detail radio environment information for each location. Fig. 2.3 has shown the spectrum usage for a certain location ($latitude$, $longitude$) in the specified date, with different time and frequency granularity, i.e., the spectrum intensity every $5 seconds$ with the frequency resolution of $2.5 MHz$.

2.3.4 Big Data Processing for SpectrumMap

SpectrumMap is designed in a web-based system for flexibility and easy access by the general public users. The web system is constructed by Bing Maps of Microsoft, which utilizes 256 x 256 pixel picture, also called tiles, to display a

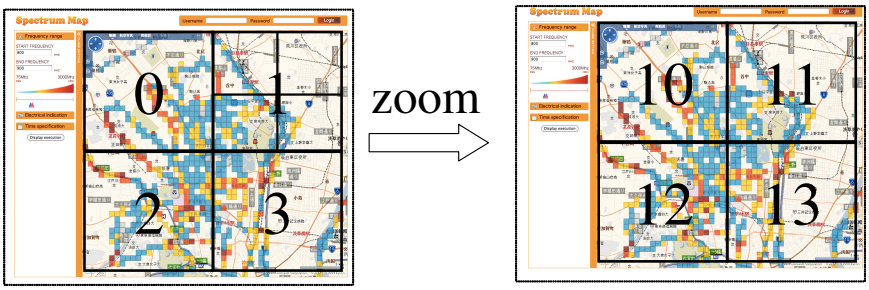

Figure 2.4: Tile and position ID of the SpectrumMap (i.e., the world is divided into 4 parts, each with a unique position ID of 1, 2, 3 and 4; the position ID of 1 can be further divided into 10, 11, 12 and 13.) [36–38].

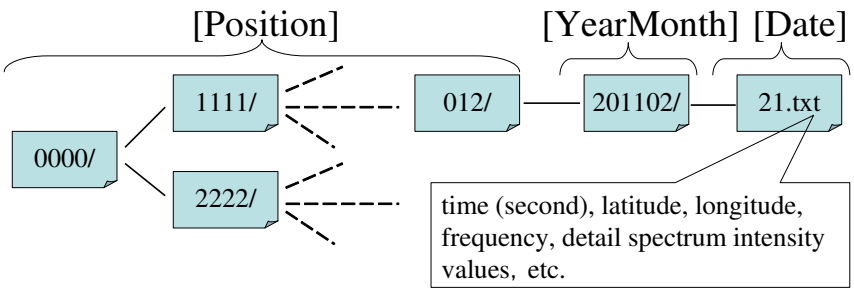

Figure 2.5: Position directory construction of Hadoop Distributed File System (The 23*bits* of position ID is mapped to a tree directory, with bit segmentation like the format of 4/4/4/4/4/3.) [37].

city map with different sizes and zoom levels (Fig. 2.4) [36–38]. Each single tile at a single zoom level is identified by a quadtree key. To realize the visualization function, this work has faced an obstacle that the delay occurred among computer servers due to the heavy computation burden introduced by large size of raw data, since calculate and display the visualized result is based on the dynamic search specified by time, location and frequency. In addition, with the gradual increase in the amount of data, scalability and fault-tolerance of the data server should be treated carefully. Hence, Hadoop Distributed File System (HDFS) is adopted for data processing, known for its high scalability and applicability to large set of distributed applications.

In HDFS, a straightforward solution is to store data file by means of position directory construction indicated by [Position] / [Year Month] / [Date].txt (Fig. 2.5) and each data record has the size of 19.5KB, including time (second), latitude, longitude, minimum frequency band, frequency interval, create time and detail spectrum intensity values of each frequency band [37]. Since one layer of the user interface of SpectrumMap is consisted by 16 pieces of tile regions, each

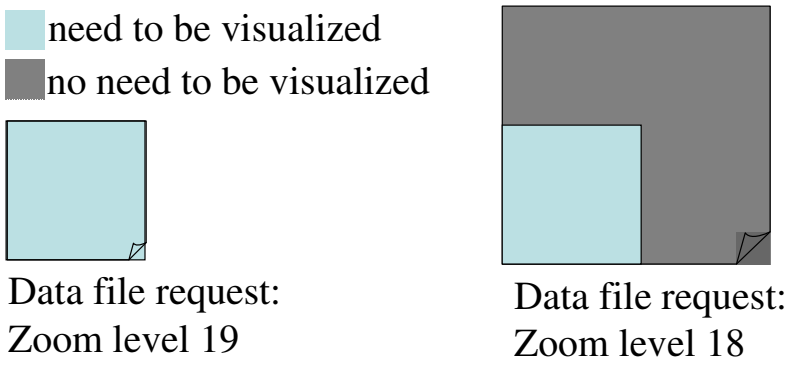

Figure 2.6: A scenario of data mixing (If the integration is based on level 18, the visualized part and non-visualized part will be mixed together when the map only requires the data of level 19.) [37].

request of data visualization has to perform 16 times of data processing. As an example, according to the test, the average time cost of 59.3$seconds$ (with variance of 1.41$seconds$) is necessary to successfully visualize the raw data collected within one-day period with the size of 147MB. To improve the data processing performance, the time cost for each process has been investigated. A complete approach for data processing, $Hdfs_{baseline}$, is given as follows, and is referred as the baseline scheme:

1. Search from HDFS position directory and obtain all the directory lists which meets the query conditions specified by location information.

2. Ready for opening spectrum data file and loading data: (a) attach the search condition of date at the end of the position directory and form a new directory path such as [Position] / [Year Month] / [Date].txt, which is actually the desired spectrum data file. Then, open this file (If no file is existed, indicated as $open0$.). (b) Find the data contents from the spectrum data file (If the data contents which meets the desired beginning time does not exist in the file, indicated as $open1$.). (c) Seek the data which lies behind the beginning time specified by search conditions (indicated as $open2$.).

3. Read the spectrum intensity value for each frequency row by row.

4. Close the file.

In the above $open1$ and $open2$, the loading of meta data has been occurred and data transfer time is necessary. It can be observed that the data processing time of $open0$ (finding the target file from the directory tree of HDFS), $open1$ (finding the meta data from the file obtained by $open0$) and $open2$ (finding the data

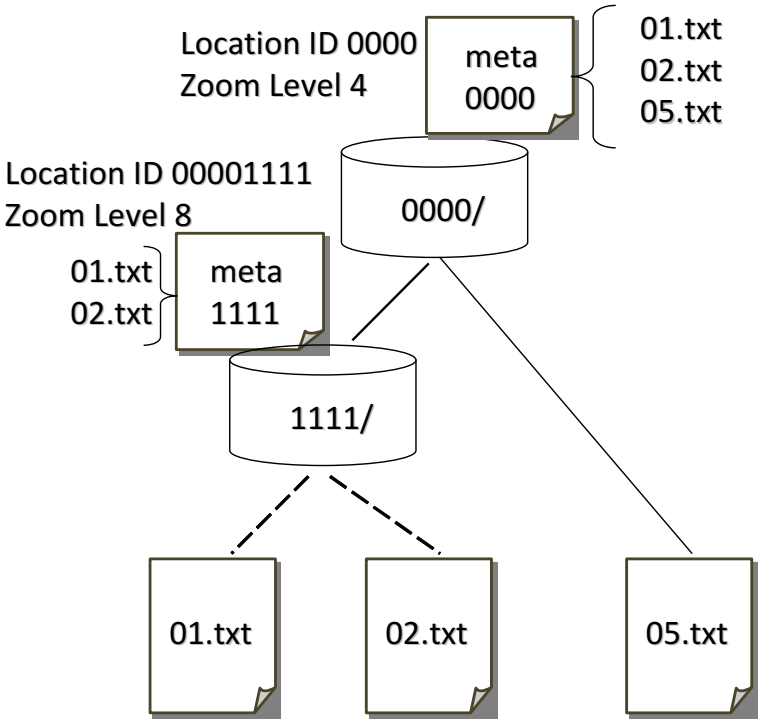

Figure 2.7: Meta data [37].

Table 2.1: Server Configuration

Application server	$(Java(Ver1.6.0), MySQL(Ver5.1.41) \times 1,$ Tomcat (Ver6.0))
Data server	$(Java(Ver1.6.0), HDFS(Ver0.20.2)) \times 6$
Physical machine	$(DellPowerEdgeR610(CPU : IntelXeonE5540x2,$ $memo : 24GB, LinuxKernel : 2.6.32)) \times 1$
Virtual Machine	$KVM(version0.12.4)(CPU : QEMUVirtualCPU \times 2,$ $memo : 2GB, Linux2.6.32)$
Throughput between servers	200 Mbps by iperf
Zoom level of title	14
Zoom level of meta data	16

segment from the meta data obtained by $open1$, based on time range) becomes the main burden, almost 80% time consumption compared with the overall time cost. Particularly, 60% of the total data processing time has been occupied by

$open0$ and $open1$ including the operations for the files outside the search scope, which consequently requires more efficient arrangement of spectrum data files to improve the performance of opening files. Obviously, all the data sets can be stored into one file, yet it is not reasonable, since the data sets to be visualized is more likely a small portion of the overall, and accordingly, all data processing are focused on one file and maintenance becomes difficult with the gradual increase of the data. To avoid this, it has been designed to merge the data files belonging to adjacent geographical locations, i.e., making one file to contain the data of several neighboring title regions on the map, because it is high possibility to display the data simultaneously which has neighboring geographical relationship. Hence, the proposed approach in this work has decided to integrate the data file based on the maximum zoom level of Bing Map, i.e., level 19. Accordingly, visualized data can be separated with non-visualized data (Fig. 2.6) [37]. The new designed data file is called as mete data file, which differs from the traditional spectrum data file and contains the directory path ([Position] / [Year Month] / [Date]) and measured spectrum intensity values inside the text contents. Consequently, searching by directory path of a tree construction in the traditional scheme has been replaced by reading the text contents and obtaining the matched data sets according to time / location information which meets the query conditions, thus the considered approach can only open the target file and relieve the unnecessary burden of time consumption wasted for opening useless files. Moreover, each meta file has been configured with zoom level and position identification, such as Fig. 2.7, and this can enable the system to find the most suitable meta file as soon as possible [37]. From the Fig. 2.7, it can be observed that the position identification of 0000 and 00001 indicate the meta data file of $meta0000$ and $meta1111$, respectively. Accordingly, the improved data processing scheme by adopting meta data, named $Hdfs_{improved}$, is shown as follows:

1. Search from HDFS directory construction and obtain all the meta files according to the most suitable zoom levels (ls).
2. Read meta data files and get meta data ($meta$).
3. Open the target spectrum data files and seek the data which lies behind the beginning time valued by search conditions ($open2$).
4. Read the spectrum intensity value for each frequency row by row.
5. Close the spectrum data files.

It should be noted that, different from the previous $Hdfs_{baseline}$, the meta data has been obtained from the meta data file, and therefor the data loading of $Hdfs_{baseline}$ is no necessary and seek can be performed directly.

Evaluating Spectrum Data Processing

Now the measured data processing performance of SpectrumMap will be shown. The system configuration is given in table 2.1, which includes 7 virtual machines (1 application sever and 6 data servers). Two kinds of data sets has been applied, the one is the real spectrum data obtained through the measurement campaigns in Tokyo city, and another is the dummy spectrum data created by computer program for comparison purpose. Each kind of data is kept working for 20 rounds and each round contains a one-day data set with the size of $147Mb$. The data set of the dummy data is configured in such a way way that it consumes as much processing time as possible. The time cost of data visualization for both $Hdfs_{baseline}$ and $Hdfs_{improved}$ is examined for comparison. To confirm the impact introduced by the size of visualized as well as non-visualized data, the data size is varied and two kinds of configurations has been prepared: (1) Visualized data ($0MB$ - $1GB$) and Non-visualized data ($0MB$); (2) Visualized data ($100MB$) and Non-visualized data ($0MB$ - $1GB$) .

Based on the above configurations, the average overall data processing time of $Hdfs_{baseline}$ is $59.3sec$ (with variation of $1.41sec$), while that of the scheme of $Hdfs_{improved}$ is $7.98sec$ (with variation of $0.132sec$) (Fig. 2.8) [37]. This indicates that the improved scheme by considering the neighboring geographical relationships of titles and integrating data files outperforms the conventional one that relies on HDFS tree constructions of spectrum data files, sine the time consumption has been reduced to 13.5%. Additionally, the time performance impacted by the data size is illustrated in Fig. 2.9 and Fig. 2.10 [37]. It can be observed from Fig. 2.10 that the time cost for $open0$ and $open1$ becomes 0, respectively. Also, since the value of ls becomes small, the useless data communication has been cut down. Note that the process of $open2$ and $read$ in Fig. 2.10 has consumed almost 80% of time consumption and hence the display time of the SpectrumMap depends on visualized spectrum data volume, which can be further verified by Fig. 2.9. For the experiment of Fig.2.9, it has been observed that the time performance for both conventional $Hdfs_{baseline}$ and the improved $Hdfs_{improved}$ remains the same when the data size is no more than 312MB, while the performance difference is obvious when the data size beyonds this value. The main reason behind this result is as follows: The number of tile regions of the zoom level 14 is 16, and each tile region can be further derived to 4^5 new tile regions when the Bing Map is operated to zoom in from the level 14 to 19 [4]. Thus the maximum data files in the improved scheme is $4^5 * 16 = 16384$. In addition, since one row of the data record has the size of $19.5KB$, this indicates that number of the files used in the experiment has just reached the maximum value expected. Accordingly, when the the data size is larger than $312MB$, the newly added data has been continuously recorded in the

[4]The zoom level of 19 is the reference value used to integrate data files in the improved scheme.

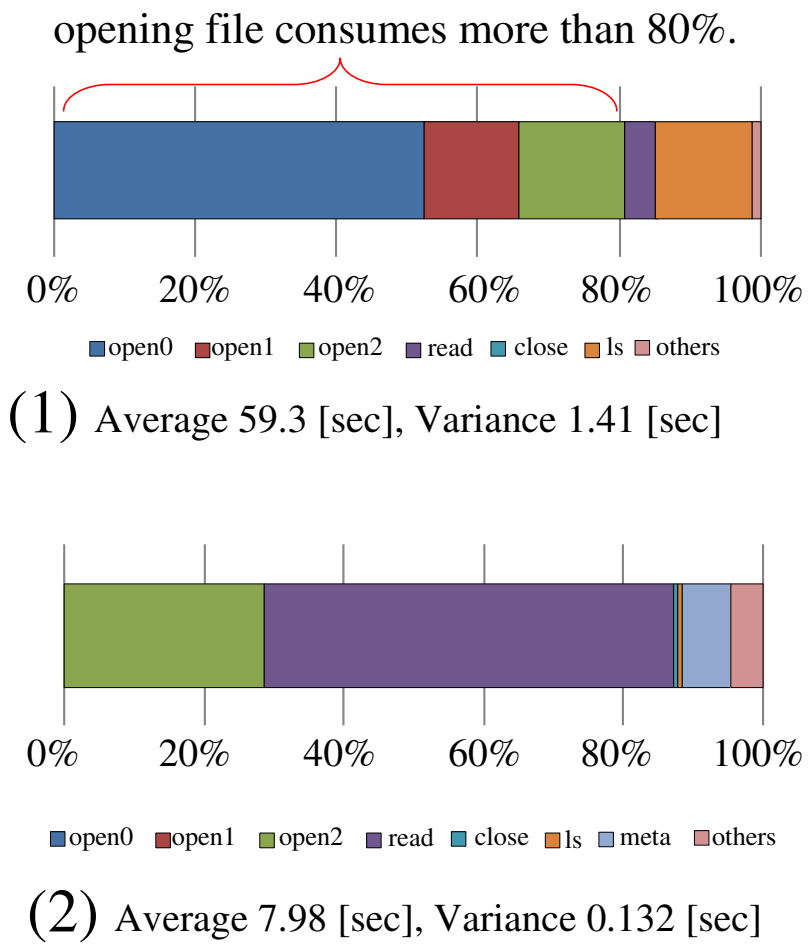

Figure 2.8: Performance comparison for data processing time [37].

same file and thus no additional operations required for opening data file. Fig. 2.10 shows that varying the size of non-visualized data has no effect on the time performance. The performance difference can be observed when non-visualized data is 0MB, this is due to the time consumption for the operation of *ls*, since the meta data has been ready for use in the improved scheme compared to the conventional scheme which needs extra efforts to locate position directories through multi layer of tree construction.

2.4 Spectrum Measurement Campaign in Tokyo

This section presents series of spectrum measurements performed in Tokyo area.

Measurement Construction: A hand-held radiometer SRM-3000 Spectrum An-

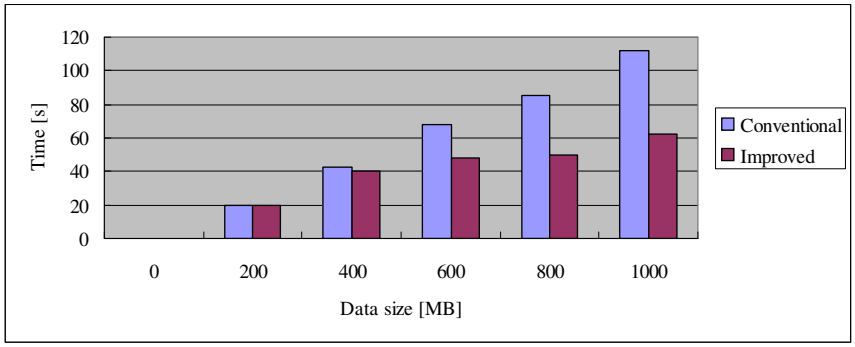

Figure 2.9: Impact of visualized data size on data processing time [37].

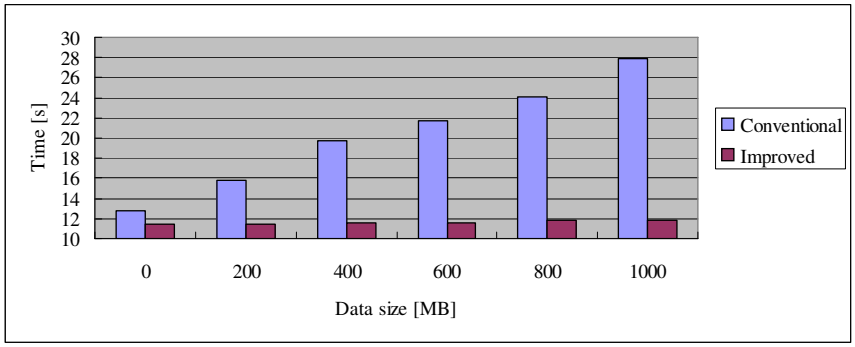

Figure 2.10: Impact of non-visualized data size on data processing time [37].

alyzer (SRM-3000 Selective Radiation Meter) is applied to monitor frequency with the target channels from $75MHz$ to $3GHz$. The location information is identified by a GPS USB (LS23068 Dongle), which consists of time functions to support synchronization between location and radio data. The resolution bandwidth of SRM-3000 is configured as $5MHz$ and the sweep time is from $200ms$ to $1sec$, with the result type of ACT (the current actual spectrum value) for each channel. The antenna is an E-field tri-axial antenna with the range from $0.25mV/m$ to $200V/m$ connected with the device through an N connector (50Ω).

Measurement Schedule: To quantify the impact of environment diversity on spectrum behaviors, the radio data is investigated according to the criteria below: (a) Spectrum data should be observed in business districts, the main terminal stations and resident districts; (b) Spectrum data should reflect typical human life patterns. In view of these reasons, a refined time-space measurement schedule is worked out to meet the above statements as shown in Table 2. Akihabara and Shibuya are selected as the downtown areas due to high congestion of people and potential requirements of wireless services. Roppongi Hills is treated as a busi-

ness district because of its large quantities of office buildings. Ikebukuro area near Kami-Ikebukuro is selected as a residence site as it is a crowded district with many tenements. Three different days (weekday / Friday / weekend) per week and three distinctive time sessions (morning / afternoon / evening; each lasts 15 minutes) per day is considered for spectrum data observations. Namely, spectrum dataset collected at each location contains abundant information used for analysis due to the space-time diversity.

Choose Noise Channel: Duty cycle has been introduced to SpectrumMap to analyze the radio data due to the ability that it can statistically describe channel usage based on a predefined threshold level. The duty cycle is modulated by Eq. (2.1) defined as the ratio of the number of samples which have larger spectrum power than a threshold value ($N_{power>threshold}$) over the number of total datasets (N_{total}).

$$DutyCycle = \frac{N_{power \geq threshold}}{N_{total}} \tag{2.1}$$

It is worthing noting that the spectrum occupancy is quantified by duty cycle according to a well defined threshold level. Some criteria to determine the threshold value have been discussed due to its sensibility to impact spectrum statistics, and most of them assume using fixed threshold value by ignoring signal noise varies with channels [7, 31]. These kinds of techniques are highly questionable since signal noise varies with frequency bands. According to the empirical data, noise distribution is also location dependent (Fig. 2.11), which differs from the viewpoint stated by the work of [16] that applies the average value of one noise channel only while ignoring its diversity with varied frequency and locations. In the proposed approach, the threshold level is obtained based on identifying the signal characteristic of non-communication spectrum bands, chosen from Inter-Satellite Communication services that are allocated at 2075.5 $\sim$ 2110 MHz and 2255.5 $\sim$ 2290 MHz. These frequency bands are selected because they are assigned to satellite-to-satellite communication and the signal usually becomes weaker under the atmospheric boundary layer (ABL), and accordingly, its strength at the ground can be considered as 0. Hence, it is called as a noise channel. Additionally, for the same reason, another noise channel selected is for Mobile Satellite Communication, which has the frequency allocated at 2500 $\sim$ 2535 MHz and 2655 $\sim$ 2690 MHz . In the subsequent parts, it will utilize the aforementioned noise channel to assist the evaluation of the duty cycle. Specifically, this part first analyzes the Wi-Fi occupancy by a linear interpolation model. Then, it proceeds to investigate the other typical bands.

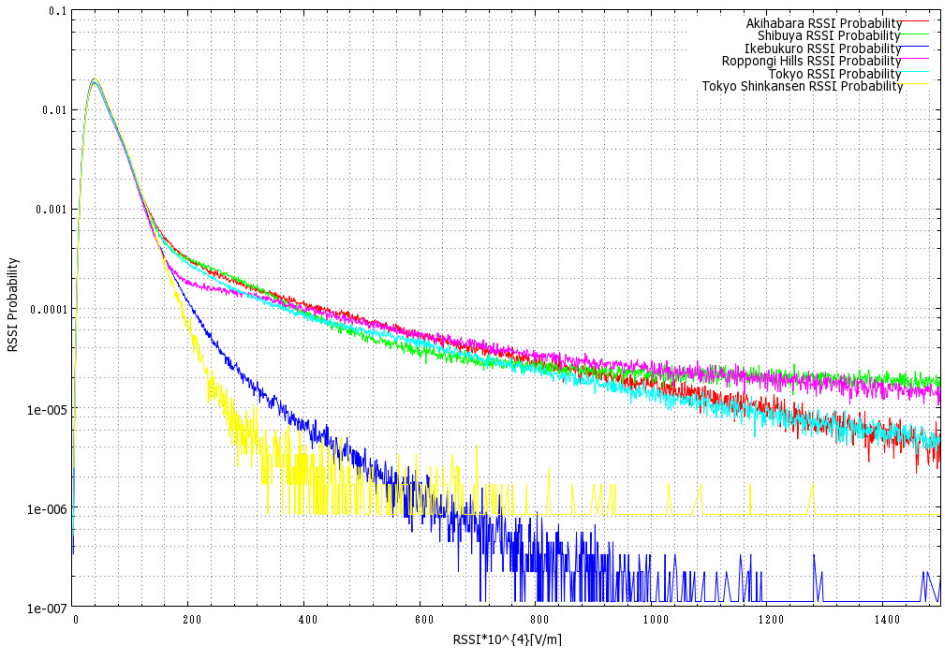

Figure 2.11: RSSI probability of noise channel obtained from 6 locations

2.4.1 Characterize Wi-Fi Occupancy by a Linear Interpolation Model

This part proposes a linear threshold model derived from the signal distribution of the noise channel mentioned above to conduct the accurate threshold value of Wi-Fi. Then, based on this model, Wi-Fi frequency utilization in realistic scenarios will be evaluated.

Following the energy detection principle, it is known that deciding the threshold value is affected by the value of False Alarm Rate (FAR) [5]. The noise distribution for the above satellite channels by a Gaussian distribution including its average as well as standard deviation has been analyzed, shown in Figs. 2.13 and 2.14, respectively. A noteworthy observation is obtainable that both of the average as well as standard deviation can be well fitted by a linear regression function, i.e., $y = k * x + b$, where y denotes average or standard deviation, x is the frequency value, k and b are parameters to be decided. Accordingly, the linear expressions as follows can be obtained:

$$Average : y = 0.0354 * x - 19.534, \quad (2.2)$$

$$Standard\ deviation : y = 0.011 * x - 7.129. \quad (2.3)$$

[5]False Alarm Rate indicates the ratio that noise power is wrongly recognized as target signal, since the noise power may be higher than the decided threshold.

Table 2.2: Measurement Schedule

Location	Weekdays	Friday	Weekends
Akiahabara (bustling area)	Tue.(08:30, 12:00,18:00)	Fri.(08:30, 12:00,18:00)	Sat.(08:30, 12:00,18:00)
Roppongi Hills (business district)	Tue.(09:30, 14:00,19:00)	Fri.(09:30, 14:00,19:00)	Sat.(09:30, 14:00,19:00)
Tokyo Station (Yaesu Side) (business district)	Wed.(09:00, 14:00,19:00)	Fri.(09:00, 14:00,19:00)	Sat.(09:00, 14:00,19:00)
Tokyo Station (Shinkansen Waitting Room)	Wed.(09:30, 15:30,19:30)		
Shibuya Station (bustling area)	Tue.(10:30, 15:00,20:00)	Fri.(10:30, 15:00,20:00)	Sat.(10:30, 15:00,20:00)
Ikebukuro Station (residence)	Wed.(08:00, 15:00,23:00)	Fri.(08:00, 15:00,23:00)	Sat.(08:00, 15:00,23:00)

Consequently, it is capable to acquire the average and standard deviation of noise for Wi-Fi based on the deduction above, since the Wi-Fi bands lies in the center of the noise channels. Another reason to utilize this model is that directly obtaining the Wi-Fi's noise distribution is difficult as it is a linear combination of signal and noise's distribution. By approximating the average and standard deviation of noise in Wi-Fi band using the channel 6, a center channel of Wi-Fi, the typical average and standard deviation value of $\mu = 66.878 \times 10^{-4} V/m$ and $\sigma = 19.882 \times 10^{-4} V/m$ has been obtained, respectively. After transforming Gaussian distribution into Standard Normal Distribution ($Z \sim distribution$) with Eq. 2.4, the desired threshold level can be defined as Eq. 2.5, shown as follows:

$$Z = \frac{X - \mu}{\sigma}. \tag{2.4}$$

$$X = Z\sigma + \mu = Threshold. \tag{2.5}$$

The threshold value can be derived, given as Table 2.3, by considering the relationship between the false alarm rate FAR and Z according to the Standard Normal Distribution.

The "good" Wi-Fi channels is expected to enable opportunistic accesses by other networks systems, and duty cycle of these channels should exhibit the behavior with spatio-temporal independence. The maximum RSSI is referred as the evaluation index of AP activation, because a high RSSI suggests that there exist some APs which are transmitting beacons to wireless devices. Figs. 2.24 - 2.35 show detail results of duty cycle (lower curves) and maximum RSSI (upper curves) obtained from Akihabara, Ikebukuro, Shibuya and Roppongi.

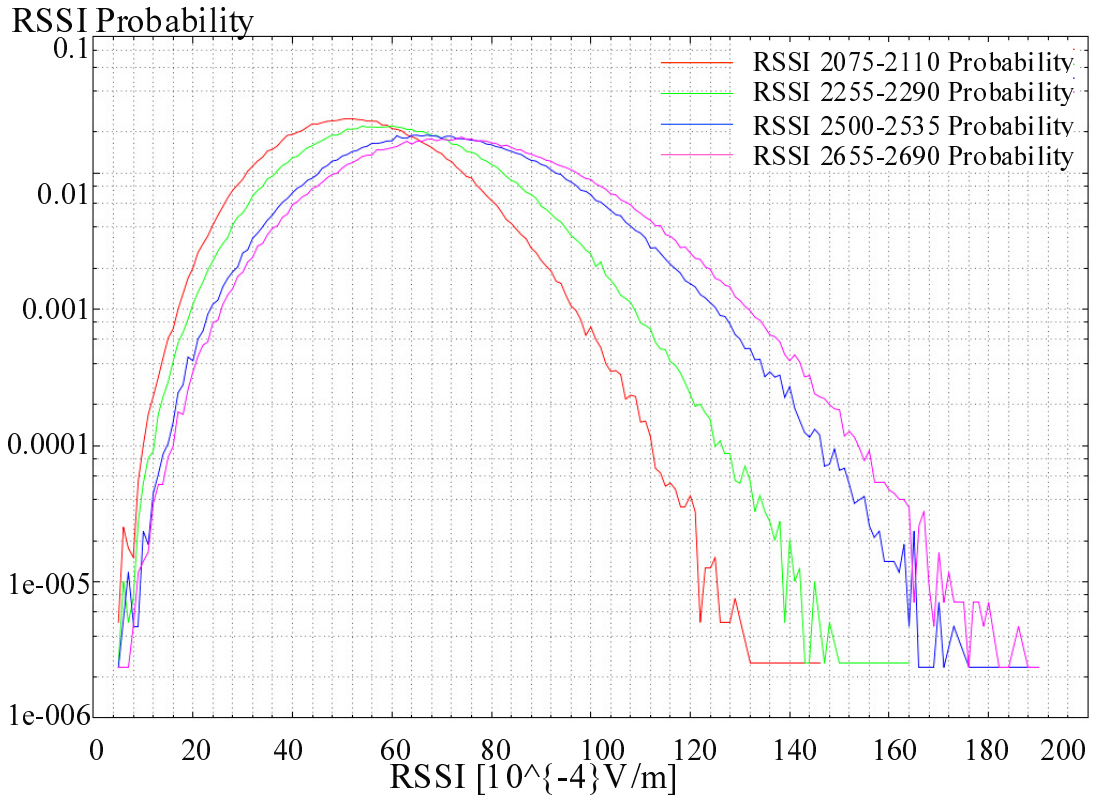

Figure 2.12: Distribution of 4 Noise Channels [39, 40].

Spatial Domain Statistics: This part investigates the duty cycle of target channel (e.g. 1, 6, 11, 14) within the same time classification (Weekday or Friday or Weekend) but at different locations. Figs. 2.24, 2.27, 2.30 and 2.33 exemplify duty cycle on Weekday at four different locations. It worth noting that duty cycle of non-overlapping channels, e.g. channel 1, 6 and 11, are all with little diversity except channel 14, which has a peak during night time in Akihabara (Fig. 2.24) while keeps in low level at other locations (Figs. 2.27, 2.30 and 2.33). For example, the duty cycle of channels 1, 6 and 11 in morning time at four locations varies with 0.03 $\sim$ 0.1, 0.02 $\sim$ 0.06 and 0.07 $\sim$ 0.13, respectively, and they also perform similarly with slight differences in afternoon and night time. The maximum duty cycle variance for these three channels (Weekday; four locations) is around 0.1. Similar conclusions is obtained by comparing duty cycle of all locations on Friday (Figs. 2.25, 2.28 and 2.31, 2.34) or Weekend (Figs. 2.26, 2.29, 2.32 and 2.35), respectively. For instance, channel 1 on Weekend days at four locations has a somewhat large difference in channel occupancy rate, e.g., the maximum with 0.3 in Fig. 2.29 and the minimum with 0.06 in Fig. 2.35, but they could

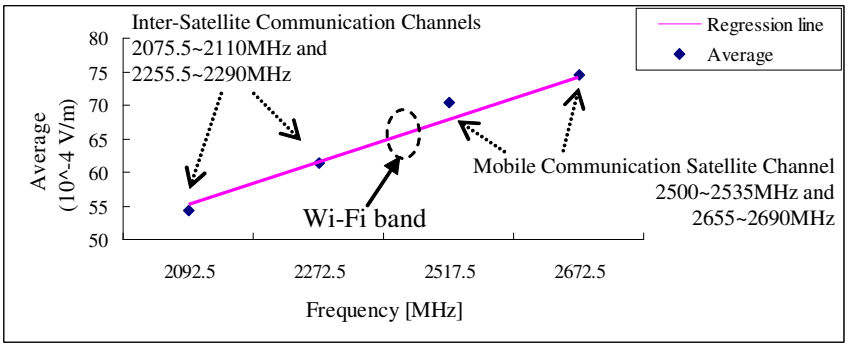

Figure 2.13: Average vs Frequency for Noise Channels [39, 40].

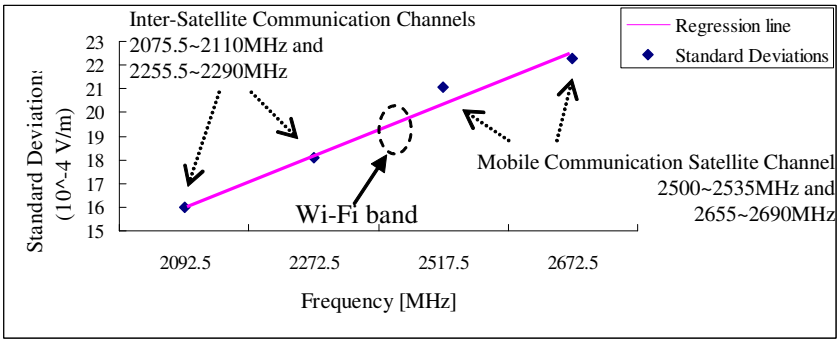

Figure 2.14: Standard Deviation vs Frequency for Noise Channels [39, 40].

still be accounted as low spectrum occupancy. The previous discussion suggests that channels 1, 6 and 11 are in good quality in view of spatial domain. As for channel 14, although duty cycle for morning or afternoon lies steadily at low levels in spite of different locations, the utilization for the evening, however, is very large (Figs. 2.24, 2.25 and 2.26). The duty cycle is more than 90% (Fig. 2.24) in Akihabara, which reveals big difference compared with duty cycle for the evening at other three locations. This phenomenon is more likely due to sales promotion events from commercial enterprises at evening time because the measurement cite is situated in a commerce center. This part is concluded that channels 1, 6, and 11 perform good statics in spatial domain since they show low and stable channel usage when considering location diversity.

Time domain statistics: It is necessary to understand the channel usage behavior over time at a location. By Figs. 2.24, 2.25 and 2.26, it can be observed that channels 1, 6 and 11 perform slight usage diversity at different hours within one day (Weekday, Friday or Weekend), while channel 14 follows a different trend that it has significant change during the evening time. As an example from Fig. 2.24,

Table 2.3: Threshold with Different FAR of Wi-Fi Channel 6

FAR	μ	σ	Z	Threshold $[10^{-4}$V/m]
0.01	66.8783208	19.8824606	2.33	113.204454
0.02	66.8783208	19.8824606	2.05	107.637365
0.03	66.8783208	19.8824606	1.88	104.257347
0.05	66.8783208	19.8824606	1.64	99.4855562

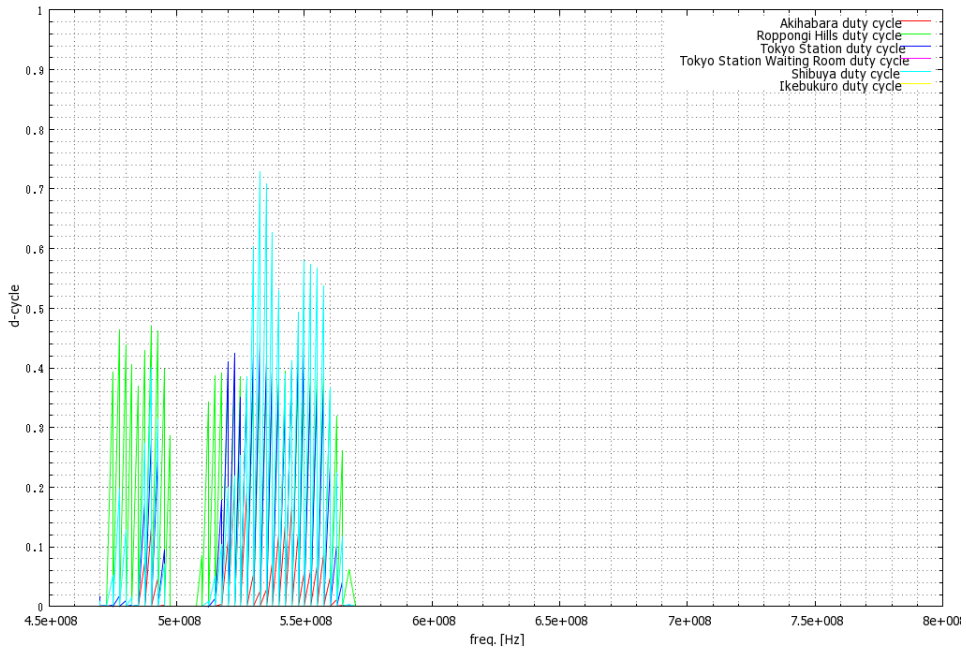

Figure 2.15: Duty Cycle of Japan UHF TV Band [470-770][MHz]

duty cycles for channels 1, 6 and 11 lie in the range of 0.1 ~ 0.14, 0.02 ~ 0.06 and 0.08 ~ 0.1, respectively, except that channel 14 experiences severely variance with 0.04 ~ 0.9. Furthermore, by comparing the duty cycle within the same hour but on different days, such as Figs. 2.24, 2.25 and 2.26, it can be confirmed that spectrum usage for channels 1, 6 and 11 do not vary daily. The results could also be obtained from other locations. Channel 14 observed at Akihabara also has the same behaviors (Figs. 2.24, 2.25 and 2.26), but the collected point at this area seems always under high data traffics at the evening time as mentioned above. In addition, spectrum usage exhibits another behavior along time that the slight ascending of duty cycle could be observed from weekday to Friday or Weekend. This variation is a reflection of human life styles, due to the improved meeting

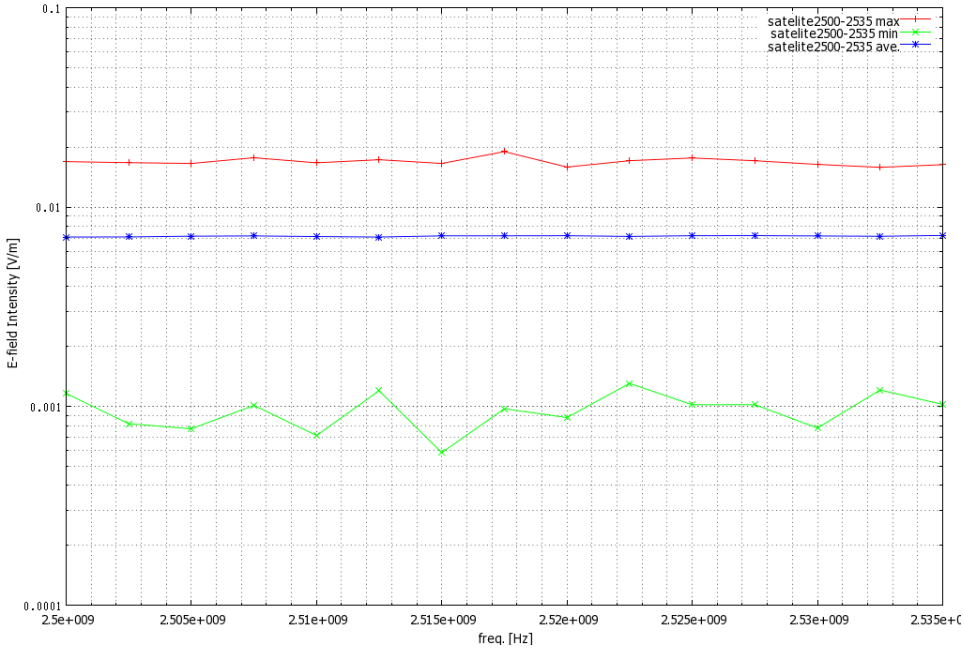

Figure 2.16: Maximum, Minimum and Average Value of Noise Channel at 2500 ~ 2535 MHz

and party during the period of off hours. On the other hand, most of the duty cycle resemble each other in the same time segment on different days at the same location, which indicates that primary transmitters work with habitual tendencies. This section is concluded that channel 1, 6 and 11 are in low channel usage pattern with time independence.

Summarized Evaluation about Wi-Fi Channels: It can observed from the above that good candidates for Wi-Fi channels exist in downtown city environment. A further evaluation is listed in Figs. 2.36 - 2.39, where AP activity is concluded according to the maximum RSSI. The good activation of AP is indicated by *Y*, otherwise by *N*, and the possible throughput of 802.11b is estimated based on duty cycle. Compared the maximum RSSI of all figures, it can be realized that channel 14 for all the cases are in small RSSI, which suggests that AP is inactive for this channel. Therefore, channel 14 is evaluated as the last rank. Channel 6 performs good quantity with low spectrum usage and good AP activation, except that, on Weekday in Shibuya, the maximum RSSI in the afternoon falls around 0.2, which means AP is inactive (Fig. 2.38). Satisfactory results in AP activity and spectrum occupancy are obtained from channel 1 and 11. For example, the average duty cycles reported from channel 1 and 11 are 0.22 and 0.18, respectively. Based on the previous discussion, the Wi-Fi channels of 1 and 11 are good media to serve for other services.

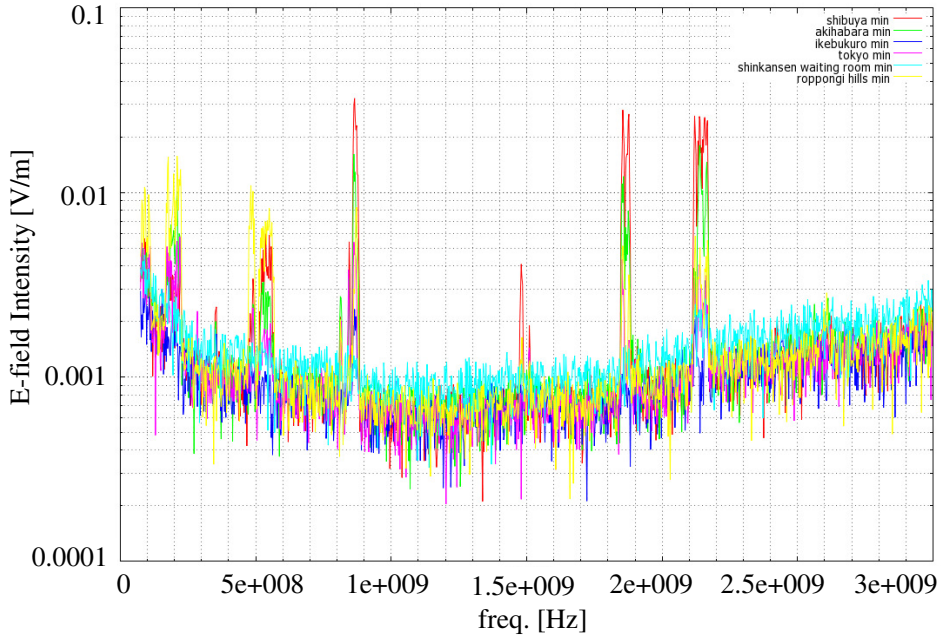

Figure 2.17: Minimum spectrum intensity of all channels observed from 6 locations in Tokyo (75 ~ 3000MHz).

2.4.2 Characterize Other Typical Bands' Occupancy

To evaluate the other typical bands of 75 ~ 3000MHz, this part still utilizes the aforementioned noise channels to determine the threshold value for each channel. In the work of [16], the author obtains a gap between minimum and maximum value in noise channel, then the gap (in dB) is utilized to evaluate the other channels by assuming a constant noise level at each frequency band. Similarly, the approach is adopted to utilize the gap to assist the determination of threshold level for other channels. Fig. 2.16 shows the maximum, minimum and average value for the selected noise channel, i.e., the satellite with 2500 ~ 2535 MHz. It is worth noting that the gap, the spectrum intensity between the maximum and minimum value, is always kept at a constant level, that the maximum is about 24.8dB (with 5MHz RBW) higher than the minimum, and the same result has been obtained for other noise channels, such as 2500 ~ 2535MHz and 2655 ~ 2690MHz assigned to Mobile Satellite Communications. The conducted gap is then taken as the criteria value to evaluate other channels. Accordingly, denoted by $p_min(i)$ the minimum spectrum intensity of the channel i, the determined threshold will be assigned as the value which is 24.8dB higher than $p_min(i)$. It has been found that the approach by utilizing the max-min based gap will get a threshold of 0.017V/m for Wi-Fi band and, due to the low power level of the channels in 802.11b, it is too

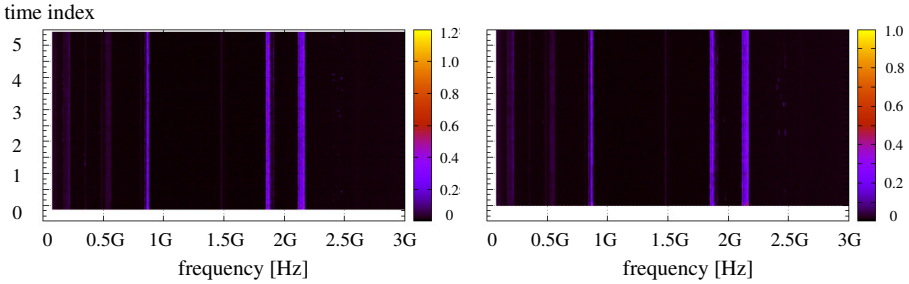

Figure 2.18: Spectrum Histogram at Shibuya (75 ~ 3000MHz).

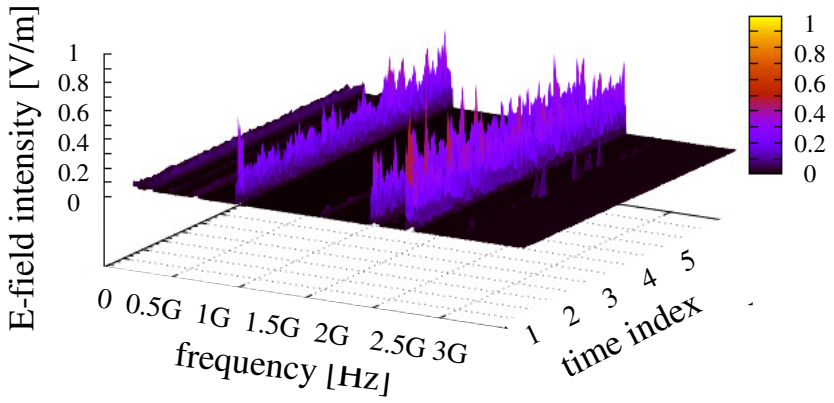

Figure 2.19: Spectrum Histogram on Friday night at Shibuya (75 ~ 3000MHz).

high to capture the Wi-Fi signal. Note that the approach presented in this work is still different with that of [16], since they roughly consider the noise level for each frequency is the same, which contradicts with the real measurement. Further, besides Fig. 2.11 which can illustrate the noise level is location dependent, this work has provided the observed minimum signal for all channels at 6 locations, shown in Fig. 2.17.

Spatial-Temporal Spectrum Usage: Based on the empirical measurement, the data collected at Shibuya station on a week day is selected and the temporal spectrum histogram is plotted by Fig. 2.18. Additionally, a 3-D spectrum intensity is shown with the observation time from 20 : 00 to 20 : 05 on the night at the same location of Shibuya, given by Fig. 2.19. The presented results suggests that, in most cases, spectrum is not well utilized, and most bands is more likely belongs to

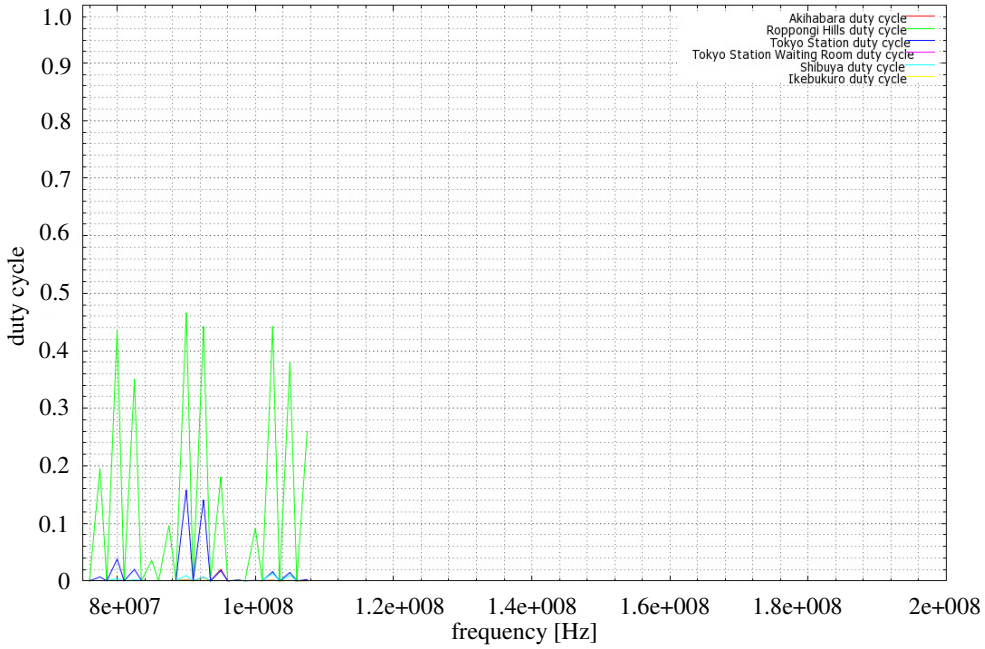

Figure 2.20: Duty cycle (FM and VHF low Band 76 ~ 108 MHz).

the licensed bands. Specifically, it is noted from the figures that the large portion of TV channels is idle, and the measured data in other places besides Shibuya has revealed the same results, which suggests that there might be potential for network operators to improve network performance by considering opportunistic spectrum access in downtown areas. A brief summary about the duty cycle of the main bands is as follows: FM and VHF[76-108MHz]: 3.15%, VHF/cable TV[176-468MHz]: 1.71%, UHF TV[470-770MHz]: 3.84%, cell phone[815-890MHz]: 8.04%, Fixed/Mobile[1759-2170MHz]: 7.91%, WiMAX[2500-2600MHz]: 1.08%.

2.5 Summary

This chapter focused on real spectrum measurement and constructing the visualized application for SpectrumMap database, a kind of crowdsourcing database. The obtained measurement results through practical data collections has shown that spectrum is not fully utilized, even in the crowded Wi-Fi channels observed in downtown Tokyo areas. Spectrum measurements have also been performed by other countries. A significant work comes from [17] which investigates the spectrum usage of Czech, supported by the government organization of France. Results has indicated that the overall spectrum utilization is about 8.3% for the target bands of 400 ~ 3000 MHz obtained from 3 locations. In Singapore, a mea-

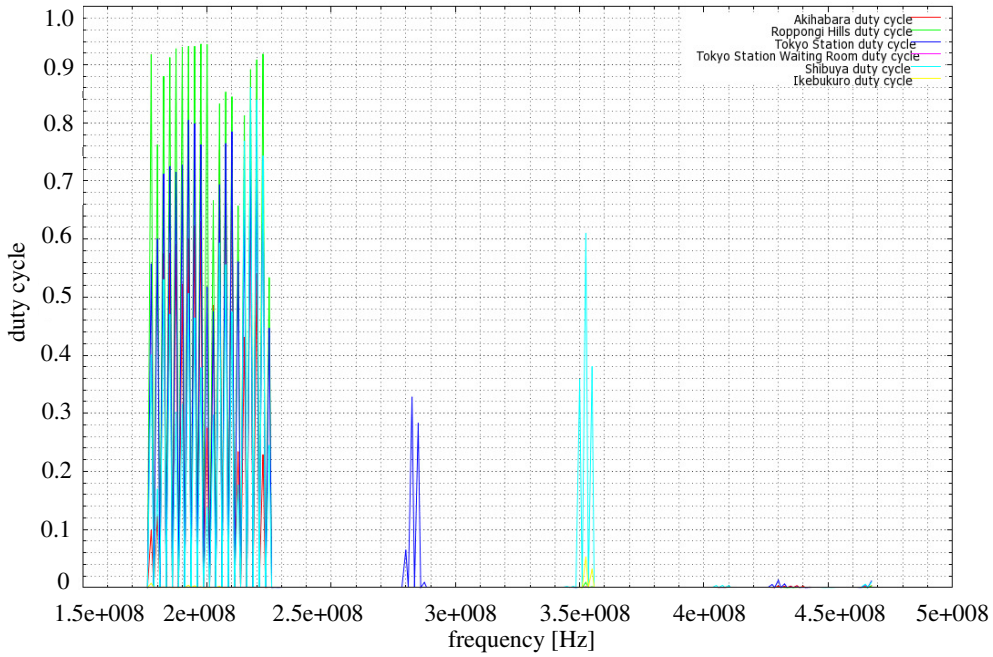

Figure 2.21: Duty cycle (VHF and cable TV Band 176 ~ 468 MHz).

surement campaign of [18] has been performed for the frequency range from 80 MHz to 5.85 GHz and the conducted duty cycle shows that the average spectrum occupancy in Singapore is as low as 4.54%. The study of [19] has also conducted a survey about spectrum usage in Beijing city, and the results suggest that the spectrum occupancy rate is about 13.5% with the target frequency of 450 ~ 2700 MHz. The measured results in this work show that the overall spectrum utilization in Tokyo is lower than 5.11%. Needless to say, the slight difference might be affected by many factors, such as the different settings of threshold values, the diverse of resolution bandwidth of the spectrum analyzer, etc. However, in most cases, spectrum usage efficiency is very low and more effective channel assignments by dynamic spectrum access is necessary.

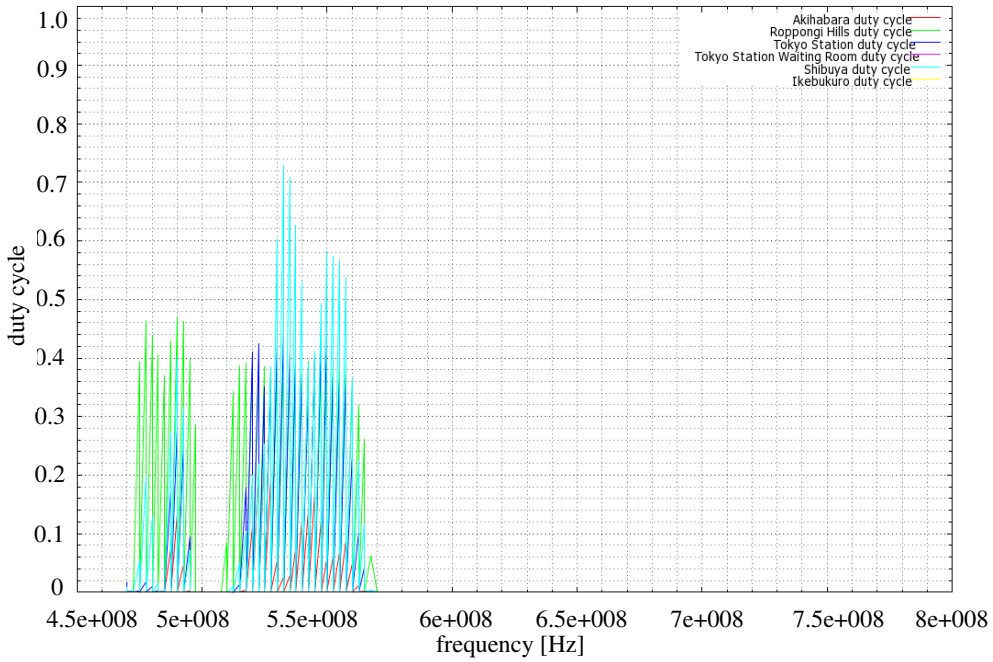

Figure 2.22: Duty cycle (UHF TV Band 470 ~ 770 MHz).

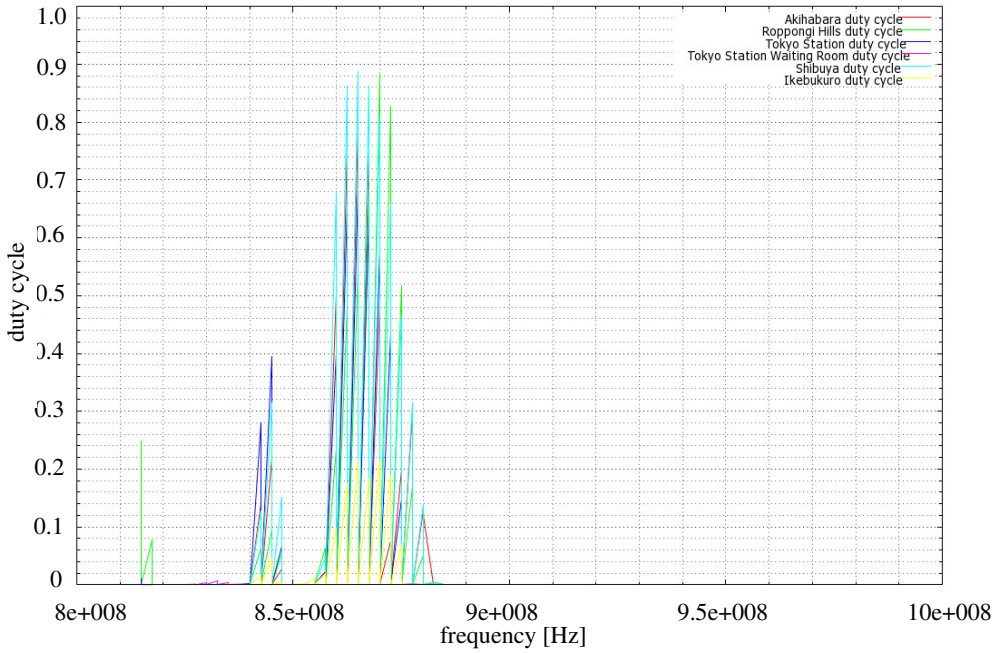

Figure 2.23: Duty cycle (Cell Phone Band 815 ~ 890 MHz).

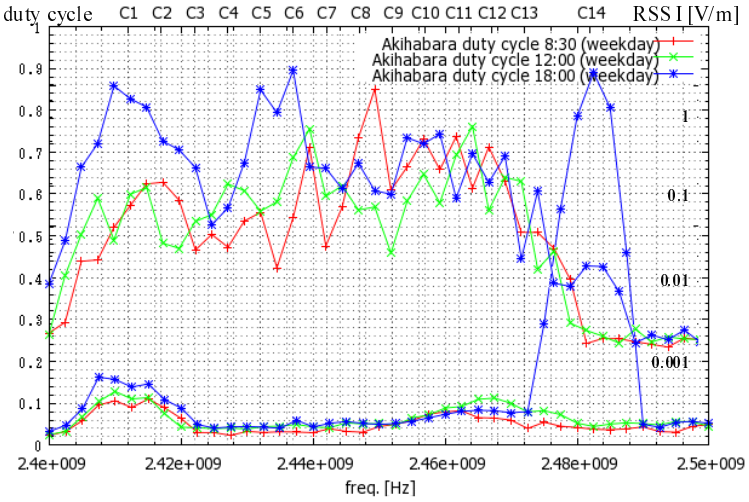

Figure 2.24: Duty Cycle of Wi-Fi at Akihabara - Weekday

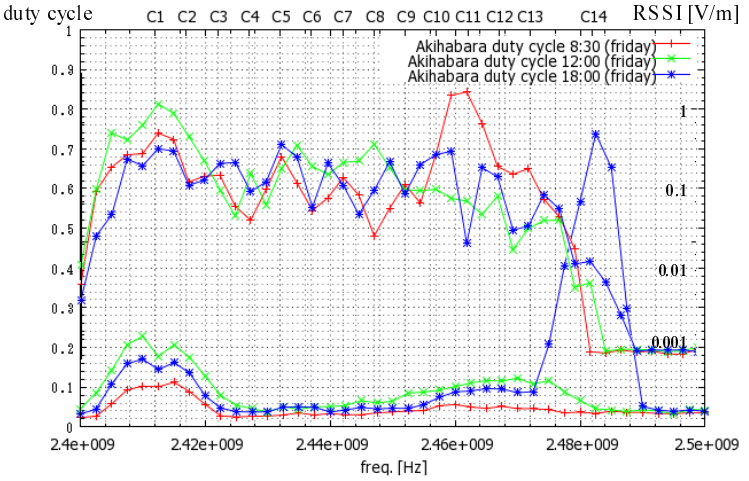

Figure 2.25: Duty Cycle of Wi-Fi at Akihabara - Friday

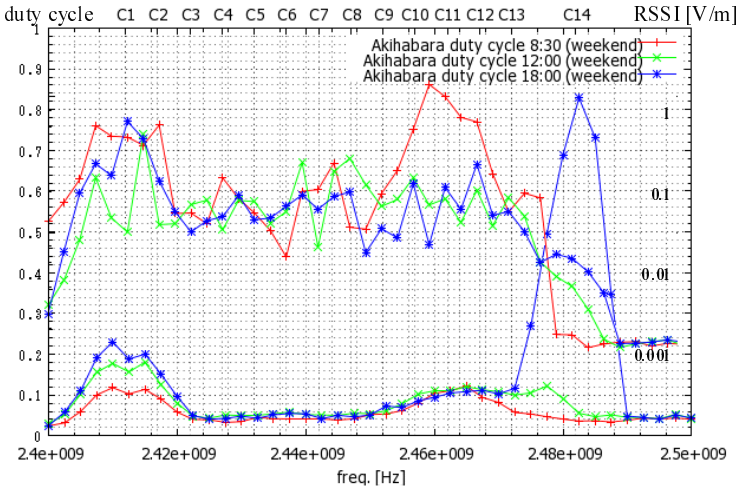

Figure 2.26: Duty Cycle of Wi-Fi at Akihabara - Weekend

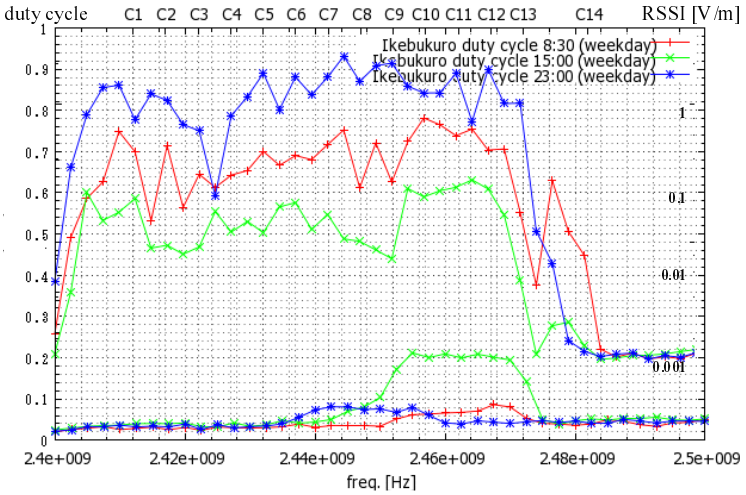

Figure 2.27: Duty Cycle of Wi-Fi at Ikebukuro - Weekday

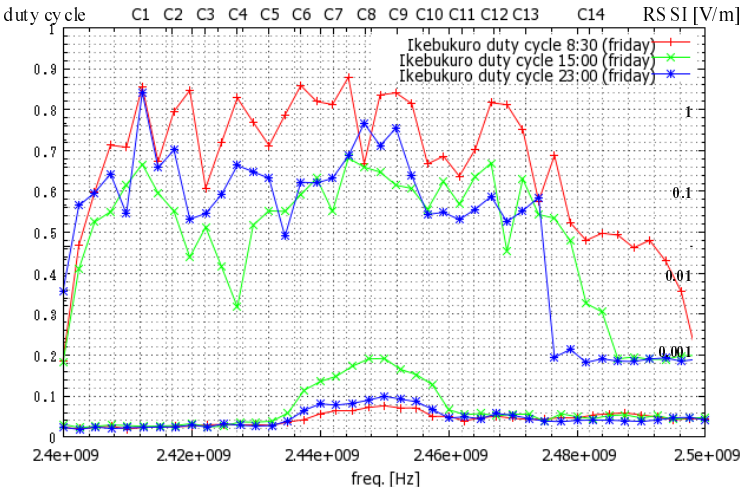

Figure 2.28: Duty Cycle of Wi-Fi at Ikebukuro - Friday

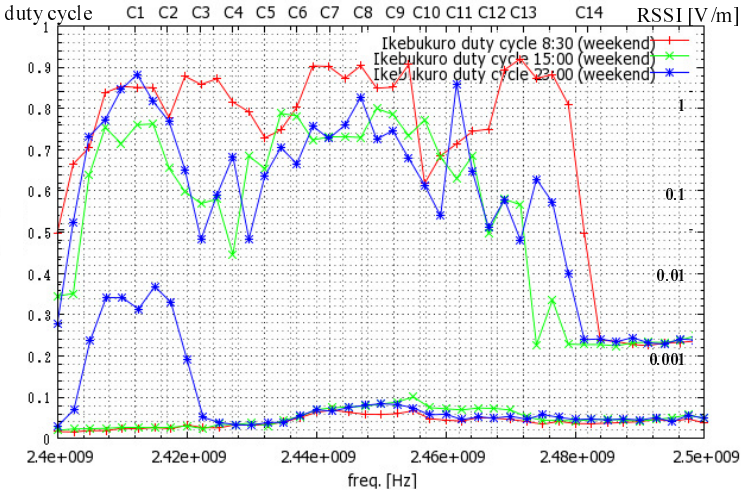

Figure 2.29: Duty Cycle of Wi-Fi at Ikebukuro - Weekend

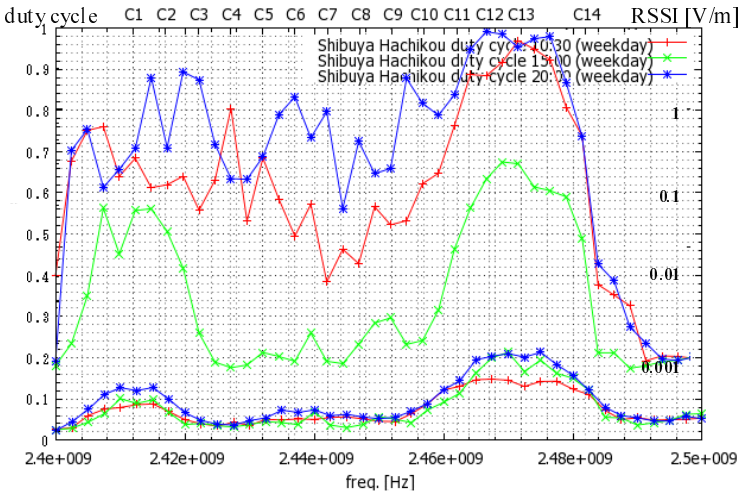

Figure 2.30: Duty Cycle of Wi-Fi at Shibuya - Weekday

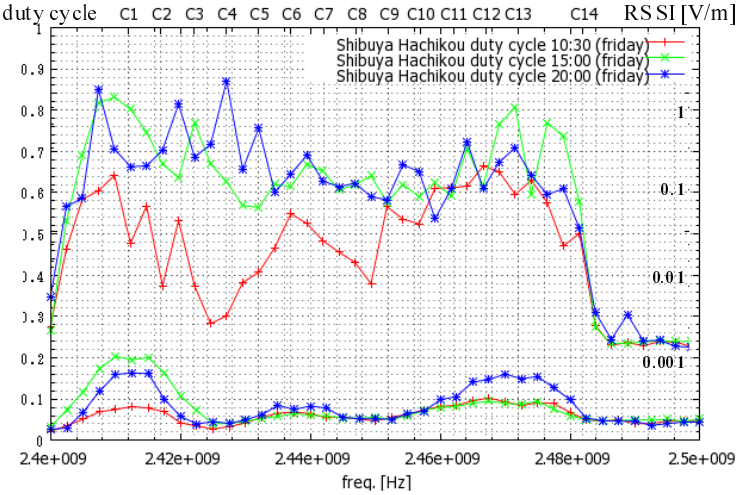

Figure 2.31: Duty Cycle of Wi-Fi at Shibuya - Friday

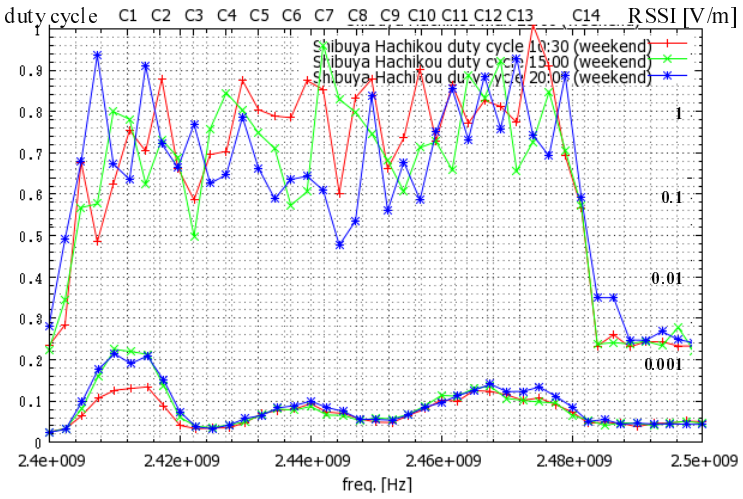

Figure 2.32: Duty Cycle of Wi-Fi at Shibuya - Weekend

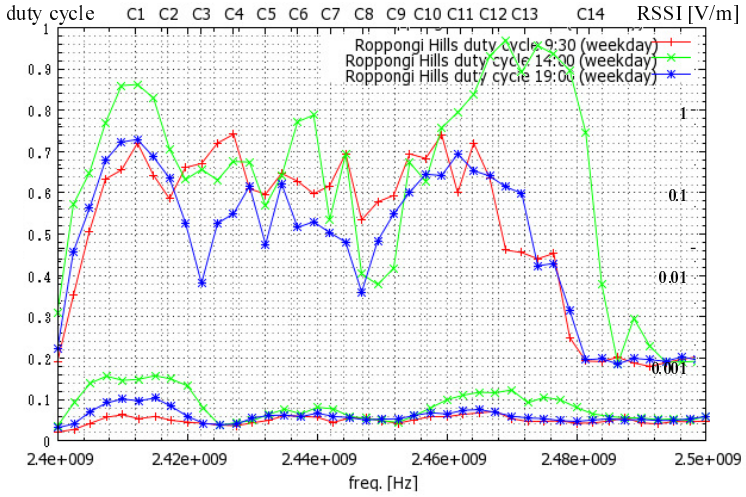

Figure 2.33: Duty Cycle of Wi-Fi at Roppongi - Weekday

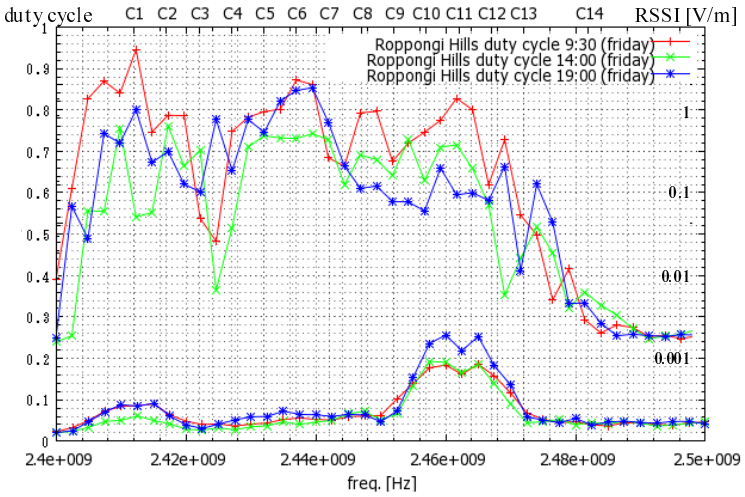

Figure 2.34: Duty Cycle of Wi-Fi at Roppongi - Friday

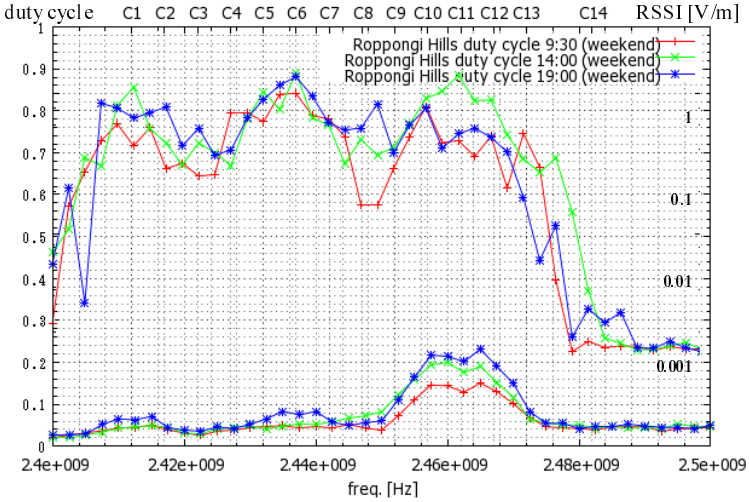

Figure 2.35: Duty Cycle of Wi-Fi at Roppongi - Weekend

Channel	Activation	Duty Cycle	Evaluation	Possible Throughput 802.11b (Mbps)
1	Y	0.19	3	8.1
6	Y	0.05	1	9.5
11	Y	0.11	2	8.9
14	N	0.05	4	9.5

Figure 2.36: Rough Estimation of Throughput Resources for Cognitive Radio at Akihabara - Estimated based on 802.11b

Channel	Activation	Duty Cycle	Evaluation	Possible Throughput 802.11b (Mbps)
1	Y	0.31	3	6.9
6	Y	0.11	1	8.9
11	Y	0.2	2	8
14	N	0.05	4	9.5

Figure 2.37: Rough Estimation of Throughput Resources for Cognitive Radio at Ikebukuro - Estimated based on 802.11b

Channel	Activation	Duty Cycle	Evaluation	Possible Throughput 802.11b (Mbps)
1	Y	0.21	2	7.9
6	N	0.08	3	9.2
11	Y	0.14	1	8.6
14	N	0.04	4	9.6

Figure 2.38: Rough Estimation of Throughput Resources for Cognitive Radio at Shibuya - Estimated based on 802.11b

Channel	Activation	Duty Cycle	Evaluation	Possible Throughput 802.11b (Mbps)
1	Y	0.15	2	8.5
6	Y	0.08	1	9.2
11	Y	0.26	3	7.4
14	N	0.04	4	9.6

Figure 2.39: Rough Estimation of Throughput Resources for Cognitive Radio at Roppongi - Estimated based on 802.11b

Chapter 3

A Near-optimal Sensing Schedule for Multi-hop Cognitive Radio Network

This chapter proposes a dynamic spectrum access policy for multi-hop cognitive radio networks (CRNs), where the transmission in each hop suffers a delay waiting for the communication channel to become available. Recognizing the energy constraints, this work assumes that each secondary user (SU) in the network is powered by a battery with finite initial energy. It develops an energy-efficient policy for CRNs using the Markov decision process, which searches for spectrum opportunities without a common communication channel and assigns each sensor's decision to every time slot. A single-sensor scenario is first considered. Due to the *intermittent activation* of the sensor, achieving the optimal sensing schedule requires excessive complexity and is computationally intractable, owing to the fact that the state space of the Markov decision process evolves exponentially with time variance. In order to overcome this difficulty, a state-reduced suboptimal policy is proposed by relaxing the constrained state space, i.e., assuming that the electrical energy of a node is *infinite*, because this state-reduced suboptimal approach can substantially reduce the complexity of decision-making for CRNs. The performance of the proposed policy is then analyzed and compared it with the optimal solution. Furthermore, the performance of this spectrum access policy under real conditions in which the electrical energy of a node is *finite* is verified. The proposed spectrum access policy uses the dynamic information of each channel. It has been proved that this schedule is a good approximation for the true optimal schedule, which is impractical to obtain. According to the theoretical analysis, the proposed policy has less complexity but comparable performance. It is proved that when the operating time of the CRN is sufficiently long, the data reception rate on the sink node side will converge to the optimal rate with probability *1*. Based

on the results for the single-sensor scenario, the proposed schedule is extended to a multi-hop CRN. The proposed schedule can achieve *synchronization* between transmitter and receiver without relying on a common control channel, and also has near-optimal performance. The performance of the proposed spectrum access policy is confirmed through simulation.

3.1 Introduction

The cognitive radio network (CRN) has emerged as a paradigm to enable efficient spectrum utilization by wireless devices in order to meet the increasing demand of spectrum access. There are two types of users of the CRN: secondary users (SUs) and primary users. SUs are unlicensed users but can use the licensed channel resources of primary users under the constraint that the SUs cause no interference with primary users. Since the available spectrum of a primary user fluctuates widely with location and time, it is usually difficult for the CRN to achieve an optimal design for SUs to support the diverse requirements of quality of service [48, 49].

Recently, the cognitive-radio-enabled wireless sensor network has become a promising candidate for improving network efficiency, i.e., a sensor can find potential vacant channels and determine the proper channel to be moved. By designing the optimal sensing schedule for a sensor-equipped CRN, the CRN can adapt the transceiver to recognize the surrounding radio environment and continuously maintain active communication links. Based on the capability of dynamic spectrum access, the CRN can greatly improve the performance of multimedia networks, heterogeneous sensing, and application of real-time surveillance [50, 51]. A number of studies have considered single-hop-based data communication in CRNs, and the multi-hop CRN has potential advantages for unexplored services. Furthermore, by extending network coverage, the multi-hop CRN can improve a wide range of multimedia applications [52].

Although the sensor-equipped CRN is efficient in terms of spectrum usage and is flexible with respect to deployment, it usually has cost and energy constraints. Accordingly, the stochastic behavior of network users, the channel heterogeneity, and the synchronization problem when neighboring nodes operate over different available channels will likely affect the data transmission performance of a CRN [53]. Some initial steps toward performance optimization regarding the quality of service of a CRN, such as improving the network capacity of a CRN ([54, 55]) and transmission delay optimization ([56, 57]), have been investigated. However, research on energy-efficient design in a CRN remains limited, because assigning the proper sensing actions and decisions for the sensors of a CRN under the complex and dynamic channel environments that coexist with primary users

is very challenging [58].

The above-mentioned fundamental challenges and many other problems must be precisely arranged and effectively addressed in order to exploit the potential advantages of the CRN. The present work focuses on the spectrum-sensing schedule for a multi-hop CRN, where the transmission in each hop is delayed while waiting for the communication channel to become available. The objective is to develop an energy-efficient sensing policy for the CRN in order to maximize the average reception rate of the sink node. This work defines the average reception rate of the sink node as the number of data packets that have been successfully delivered from the source node to the sink node in one time slot.

3.1.1 Primary Results

The present work has three primary contributions.

- First, it assumes that each sensor in the CRN is powered by a battery with finite initial energy and design a dynamic schedule to maximize the energy efficiency by avoiding unnecessary sensing actions. Sensors will dynamically choose whether to sense and access at maximum throughput. Since, under realistic conditions, sensors are always equipped with a battery, the designed schedule will be practical.

- Second, the proposed schedule will use the dynamic information of each channel. It has been proved that this schedule is a good approximation for the true optimal schedule, which is impractical to obtain. The average data reception rate under the proposed policy can be proved to converge to the optimal rate with probability 1 when the operating time of the network is sufficiently long.

- Third, the proposed schedule can be extended to multi-hop CRNs, which will provide an opportunistic relay forwarding scheme and achieve synchronization between the transmitter and receiver without a common control channel. The near-optimal property of the proposed schedule is also proved.

3.1.2 Related Research

One of the most important targets in CRN protocol design is increasing network throughput while better utilizing limited energy resources. If there is no energy constraint, the SU should always sense the availability of all channels and monitor the spectrum, even when there is no data to transmit. With an energy constraint, however, the SU will choose to work or sleep intermittently in order to conserve

energy. This significantly complicates problem solving. Thus, a number of studies have examined energy-efficient strategies for CRNs. In [59], Zhao *et al.* proposed decentralized cognitive protocols that search for spectrum opportunities dynamically. Owing to hardware and energy constraints, their protocols assumed that each SU can sense only a limited number of channels in each slot. [60] used the same method and developed optimal distributed MAC protocols for an energy-constrained CRN. The authors established that the optimal sensing and access policies have threshold structures, i.e., an SU will sense a channel when the conditional probability of the channel idle state exceeds some threshold value and will access a channel when the sensing outcome shows that the channel is not occupied and its fading condition is better than some threshold value. [61] attempted to improve both spectrum and energy efficiencies and proposed a diversity technology to use channel resources and energy.

The above studies focus primarily on a single-hop wireless access network. For multi-hop networking problems in a CRN, different approaches will be used because more issues must be considered. In [62], the authors investigated the problem of designing effective routing solutions for multi-hop CRNs and categorized the routing solutions in CRNs into two main classes depending on the spectrum awareness, namely, full or local spectrum knowledge. Xin *et al.* proposed a graph-based channel assignment framework to model multi-hop CRNs and perform routing assignment from a network perspective [63, 64]. Hou *et al.* developed a mixed-integer non-linear program for multi-hop CRNs [65]. This formulation ensures the existence of a multi-hop path between any source-destination pair. Liu *et al.* proposed a distributed hop relay algorithm that adapts transmission to the dynamic spectrum access opportunities in the neighborhood of the secondary user [52]. They assumed that SUs can exchange messages over a common control channel. Ren *et al.* analyzed the multi-hop delay of ad hoc CRNs under propagation delay and temporal traffic dynamics [66, 67].

Moreover, in the work of [58], the energy efficiency for cluster based cognitive radio sensor network has been addressed. The authors assumed sensor nodes can form multiple clusters and periodically deliver data to the sink node by hierarchical routing. Based on the model, they investigated the energy consumption during channel sensing and switching and showed comparable results of energy consumption for sensors in the cases of utilizing a licensed channel only as well as a default license-free channel only. There are two major differences between the presented work and [58] : (1) In [58], the data communication within each cluster and among different clusters is coordinated in a Time Division Multiple Access (TDMA) manner; while in the presented work, no coordination is assumed between nodes. (2) [58] has not considered the Markov property of channels; while the model in the presented work has applied the Markov property into the channels.

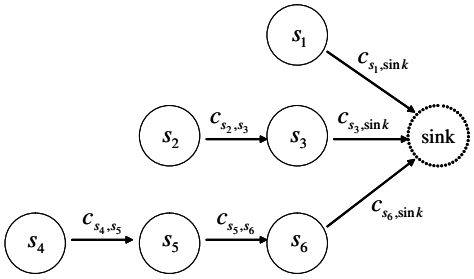

Figure 3.1: Network Configuration Analyzed in the Present Study.

The present work addresses the challenges of developing a sensing schedule for a CRN, where data packets are transmitted to the endpoint through multi-hop links. In contrast to previous studies, instead of considering only a single-node decision, this work considers the decision-making of all sensor nodes in the network. In addition, this work focuses on designing a new sensing schedule that relies on the Markov decision process, and the formulated problem is attributed to *partial observability*. In this sense, the performance optimization for the considered sensing schedule suffers from excessive computational complexity [68], which arises from the fact that the state space of the Markov decision process evolves exponentially with time. In contrast to previous studies such as [59], this work herein proposes a new method that can reduce the system state space, and, without finding the optimal solution directly, the presented work has obtained an approximate solution, which can effectively reduce the decision space of sensors and thus can substantially reduce the complexity of decision-making by the network. Although the designed sensing policy has a simple structure, it can ensure near-optimal performance and is a good approximation for the true optimal schedule.

The remainder of the present work is organized as follows. In Section II, the considered network model and the problem definition is presented. In Section III, data transmission and synchronization for the single-sensor scenario is discussed. Based on the results for the single-sensor scenario, Section IV investigates the performance of a multi-sensor scenario. Section V presents simulation results that validate the proposed sensing schedule. Finally, the present study is summarized in Section VI.

3.2 Network and Problem Definition

3.2.1 Network Model

In this section, a network is designed, which is assumed to contain N nodes and a sink node. The variable $W = \{s_1, s_2, \ldots, s_N\}$ represents all of the nodes in the

network, where s_i represents an individual node. It is also assumed that there are a number of channels between the nodes. The state (busy or idle) of a channel at time slot t shared by nodes s_i and s_j is denoted by $c_{s_i,s_j,t}$. The channel is licensed to primary users, and nodes s_i and s_j can use cognitive radio technology to share this channel.

Next, the working mechanism is described. The time is divided into slots. At the beginning of each slot t, nodes s_i and s_j should sense the state of the channel $c_{s_i,s_j,t}$ simultaneously. If the channel is idle, in other words, if the channel is not used by the primary user, then $c_{s_i,s_j,t}$ is defined as 1, and node s_i transmits one packet to node s_j. If the channel is used by the primary user, then $c_{s_i,s_j,t}$ is 0 and node s_i does not transmit packets in order to save energy. Here, the follows have been assumed: (1) the nodes can be synchronized, (2) each sensing can correctly obtain the channel state, (3) the channels are often occupied randomly by primary users. As for (1), this work describes how synchronization is achieved later herein. As for (2), this work desires to derive the maximum transmission rate at successful data transmissions. As for (3), a Markov chain is used to represent this randomness.

Based on the channel state of $c_{s_i,s_j,t}$ mentioned above, the state transition dynamics can be represented using a transition probability matrix, as follows:

$$\mathbb{P} = \begin{pmatrix} a_{s_i,s_j} & 1 - a_{s_i,s_j} \\ 1 - b_{s_i,s_j} & b_{s_i,s_j} \end{pmatrix}, \tag{3.1}$$

where element $\mathbb{P}_{1,1} = a_{s_i,s_j}$ denotes the probability that channel c_{s_i,s_j} remains in the idle state, $\mathbb{P}_{1,2} = 1 - a_{s_i,s_j}$ denotes the probability that channel c_{s_i,s_j} changes from the idle state to the busy state.

In Fig. 3.1, the packets from s_2, s_4 can only be sent to the sink node through several nodes by multiple hops. For example, the routing $s_4 - s_5 - s_6 - sink$ can be used to achieve this goal. Moreover, each packet transmission should pass through spectrum sensing and use cognitive radio transmission technology. It is assumed that T time slots are needed to operate this network. Since the energy of every node is finite and the channels will be occupied randomly, it is necessary to design a sensor scheduling of all nodes for the average reception rate by sink node to be maximized.

3.2.2 Problem Definition

The primary notation in the present work is listed in Tables 1.1 and 1.2. It has been assumed that all nodes have a battery with capacity B. At the beginning of slot t, the energy level of sensor s_i is $b_{s_i,t}$ (in energy units), and its decision is $a_{s_i,t} \in \mathcal{A}$. The set $\mathcal{A} = \{1, 0\}$, where '1' indicates that sensor s_i will sense the

channel with an energy cost of b_1. If the channel is idle, node s_i will send out one packet with an additional energy cost of b_2. Then, at the beginning of time $t+1$, the residual energy in node s_i is $b_{s_i,t+1} = b_{s_i,t} - b_1 - b_2$. If the channel is busy, then no packet will be sent out, and $b_{s_i,t+1} = b_{s_i,t} - b_1$. Analogously, '0' indicates that the sensor will not sense the channel. Here, it is assumed that the nodes are always aware of their energy levels. The initial state is defined as x_0, which indicates all of the channels' initial states and all of the sensors' battery levels, i.e., $x_0 = \{c_{s_i,s_j,0}, b_{s_i,0}, s_i, s_j \in W\}$. The decision of every node in every slot is represented by π_T, which can be obtained by $\pi_T \triangleq \{a_{s,t}, s \in W, t = 1, 2, \ldots, T\}$. A variable $U_{s_i,T}$ is defined to represent the average power consumption by node s_i. Accordingly, it follows that

$$U_{s_i,T}(\pi_T) = \frac{1}{T}\mathbb{E}\{\sum_{t=1}^{T}\sum_{s\in W} a_{s_i,t}(b_1 + b_2 c_{s_i,s,t})|x_0\}. \tag{3.2}$$

The battery power of the sensor is finite, and the total running time, T, is relatively large and so cannot support continuous operation. In other words, the sensor cannot sense and transmit continuously. Therefore, it is necessary to design a strategy that can guide the operation of all nodes. More specifically, it needs to design a strategy that can make a decision for every node in every slot. The goal is to maximize the average reception rate of the sink node.

Next, two variables are specified to represent the transmission and reception of a packet between s_i and s_j. $o_{s_i,s_j,t} = c_{s_i,s_j,t} \times a_{s_i,t}$ is defined. When $o_{s_i,s_j,t} = 0$, no packets are sent by node s_i to node s_j, whereas $o_{s_i,s_j,t} = 1$ indicates that node s_i transmits one packet to node s_j. Similarly, $r_{s_i,s_j,t} = c_{s_i,s_j,t} \times a_{s_j,t}$ is defined. When $r_{s_i,s_j,t} = 0$, no packet is received by node s_j from node s_i, and when $r_{s_i,s_j,t} = 1$, node s_j will receive one packet from node s_i. Note that, in the case of receiver s_j, decision variable $a_{s_j,t} = 1$ (or 0) indicates that s_j will sense the channel and receive a packet when the channel is idle (or not sense the channel).

A variable G_T is specified as the average reception rate, which is defined as follows:

$$G_T(\pi_T) \triangleq \frac{1}{T}\mathbb{E}\{\sum_{t=1}^{T}\sum_{s\in W} r_{s,sink,t}|x_0\}. \tag{3.3}$$

Next, the primary problem considered herein is proposed:

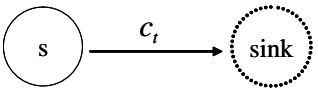

Figure 3.2: Network for a single sensor.

Problem 1

$$
\begin{aligned}
max_{\pi_T} \quad & G_T(\pi_T), \\
s.t. \quad & U_{s_i,T}(\pi_T) \leq e, && (3.4) \\
& \sum_{t=1}^{T'} \sum_{s_j \in W} o_{s_i,s_j,t} = \sum_{t=1}^{T'} \sum_{s_m \in W} r_{s_m,s_i,t}, && (3.5) \\
& \forall s_i \in W, T' \in [1, T],
\end{aligned}
$$

where $e = B/T$ denotes the permissible average energy usage, constraint (3.4) indicates that the total power consumption of each node should be smaller than its total battery capacity, and constraint (3.5) is the flow conservation for relay nodes. The goal is to find the optimal policy π_T^*.

There are a number of differences between the problem considered herein and those of previous studies. First, it is necessary to consider the decision-making of all nodes in the network, for the data packets to be sent to the endpoint via multi-hop CRN. Generally speaking, previous studies analyzed only single-node decisions [59]. Second, the energy of each node should also be considered because the energy is one of the most important factors in a network, and decisions that are not based on energy cannot be applied to practical applications.

This problem can be cast as a Markov decision process problem. However, since the problem involves the stochastic processes of channels and the energy level of each node, the optimal decision by each node will depend simultaneously on the states of the channels and the nodes at each time slot. Finding the optimal result will be complicated and difficult for the nodes to implement. Thus, the present work finds a suboptimal policy, which will be easy for the network to implement. The performance can be theoretically proved to be slightly different from the optimal performance. Furthermore, when the time T increases, the performance of the suboptimal strategy will converge to the optimal performance with probability (w.p.) 1. The following two sections discuss the single-node and multi-node problems. In particular, it finds the best approximate solution for the single-node problem, which is then extended to the multi-node problem.

3.3 Single-sensor Scenario

In this section, the single-sensor problem is analyzed, as shown in Fig. 3.2. Since only a single node is involved, the variable notation can be simplified. For ex-

ample, the channel shared by node s and the sink node is c_t, and its transition probability matrix is

$$\mathbb{P} = \begin{pmatrix} a & 1-a \\ 1-b & b \end{pmatrix}. \tag{3.6}$$

In addition, the battery length of node s is B. Because of the limited battery power, the sensor will sense the channel in some slot in order to determine the channel state. Since the sensor power will sometimes be off, the current state of the channel cannot be obtained. In this way, this problem is one of partial observability. The channel state space can be expressed by the following set:

$$\mathcal{S} = \{\varsigma_{0,0}, \varsigma_{0,1}, \varsigma_{0,2}, \ldots, \varsigma_{1,0}, \varsigma_{1,1}, \varsigma_{1,2}, \ldots\}, \tag{3.7}$$

where state $\varsigma_{0,0}$ indicates that the channel is occupied, $\varsigma_{0,k}(k > 0)$ indicates that the channel is busy k slots before, $\varsigma_{1,0}$ indicates that the channel is idle, and $\varsigma_{1,k}$ indicates that the channel is idle k slots before. Illustrative examples of channel states are given by Figs. 3.3 and 3.4. For this single-node case, **Problem** 1 will turn into

Problem 2

$$\begin{aligned} max_{\pi_T} \quad & G_T(\pi_T), \\ s.t. \quad & U_T(\pi_T) \leq e. \end{aligned} \tag{3.8}$$

Since the sensor can send data directly to the sink node, there are no constraints such as flow conservation (3.5). However, the optimal solution of **Problem** 2, $\pi_T^* = \arg\max_{U_T(\pi_T) \leq e} G_T(\pi_T)$, is also not easy to obtain. Next, the following set is defined: $\mathcal{B} = \{0, 1, 2, \ldots, B\}$. From the Markov decision process theory, the optimal action $a_t \in \mathcal{A}$ of node s in slot t will be based on the state of $x_t \in \mathcal{S} \times \mathcal{B}$. The state space $\mathcal{S} \times \mathcal{B}$ is very large and will continue to expand as time T increases. Solving $\mathcal{S} \times \mathcal{B}$ exactly will be computationally intractable. In particular, the value B is often very large, which leads to difficulty in storing all of the states. In order to overcome these difficulties, this work is currently developing a near-optimal policy that has a simple structure while ensuring near-optimal performance.

3.3.1 State-reduced Suboptimal Policy π_∞^*

In this section, a new problem is defined, namely, **Problem** 3, to provide a suboptimal policy.

Assumption 1 *Assume that the electrical energy of node s is infinite.*

Since under this assumption, the sensor's battery level is always infinite, its decisions will be only related to the channel's states. Define a new state variable $x'_t \in \mathcal{S}$ to represent the channel's state. When the node takes action $a_t =$'1', the state transition probability can be calculated according to the following equation:

$$p(x'_{t+1}|x'_t, a_t) = \begin{cases} f^k_{1,1} & x'_t = \varsigma_{1,k} \ \ and \ \ x'_{t+1} = \varsigma_{1,0}, \\ 1 - f^k_{1,1} & x'_t = \varsigma_{1,k} \ \ and \ \ x'_{t+1} = \varsigma_{0,0}, \end{cases} \tag{3.9}$$

where $f^k_{1,1} = \frac{1-b}{2-a-b} + [a + (1-b)/(a+b-2)](a+b-1)^{k-1}$ denotes the probability that the channel state changes from $\varsigma_{1,k}$ to $\varsigma_{1,0}$, which can be calculated based on the results of the Markov chain [69]. Analogously, the result of $x'_t = \varsigma_{0,k}$ can be obtained. This only requires the introduction of the probability $f^k_{1,0} = \frac{1-b}{2-a-b} + (1 - b + (1-b)/(a+b-2))(a+b-1)^k$ [69], which is the probability that channel state changes from $\varsigma_{0,k}$ to $\varsigma_{1,0}$.

When $a_t =$'0', the state transition probability is

$$p(x'_{t+1}|x'_t, a_t) = \begin{cases} 1 & x'_t = \varsigma_{1,k} \ \ and \ \ x'_{t+1} = \varsigma_{1,k+1} \\ \quad or & x'_t = \varsigma_{0,k} \ \ and \ \ x'_{t+1} = \varsigma_{0,k+1}, \\ 0 & otherwise. \end{cases} \tag{3.10}$$

An illustrative case for $x'_t = \varsigma_{1,k}$ and $x'_{t+1} = \varsigma_{1,k+1}$ is shown in Figure 3.5.

Accordingly, the system will receive a reward, which can be derived using the following formula:

$$r(x'_{t+1}|x'_t, a_t) = \begin{cases} 1 & x'_{t+1} = \varsigma_{1,0}, \\ 0 & otherwise. \end{cases} \tag{3.11}$$

This work proposes the stationary policy $\pi_\infty = \{d_{i,j} = \mathbf{Pr}(a_t =\text{'1'}|\varsigma_{i,j}), \varsigma_{i,j} \in \mathcal{S}\}$. In this case, the average power consumption and average reward can be defined as follows:

$$U_\infty(\pi_\infty) = \lim_{T\to\infty} \frac{1}{T} \mathbb{E} \sum_{t=1}^{T} a_t(b_1 + b_2 p(\varsigma_{1,0}|x'_t, a_t)), \tag{3.12}$$

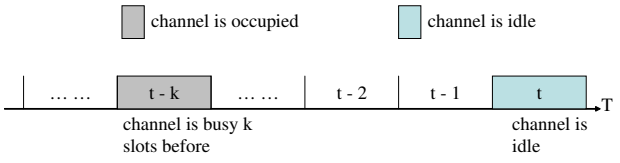

Figure 3.3: Illustrative example of channel states (time slot of t: $\varsigma_{1,0}$; time slot of $t-k$: $\varsigma_{0,k}$).

$$G_\infty(\pi_\infty) = \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\sum_{t=1}^{T} r(x'_{t+1}|x'_t, a_t). \tag{3.13}$$

Next, the following problem is considered.

Problem 3

$$\begin{aligned} max_{\pi_\infty} \quad & G_\infty(\pi_\infty), \\ s.t. \quad & U_\infty(\pi_\infty) \le e. \end{aligned} \tag{3.14}$$

Note that the variables $U_\infty(\pi_\infty)$ and $G_\infty(\pi_\infty)$ represent the average power consumption and the packet reception rate within an infinite operating time under **Assumption** 1, respectively, whereas in **Problem** 2, variables $U_T(\pi_T)$ and $G_T(\pi_T)$ represent the average values within time T and under the constraint that the total energy storage of the node is B, respectively. Furthermore, the distinction between policy π_∞ and policy π_T is that the decision of the former is determined only by the state space $\mathcal{S}$, whereas the latter is determined by $\mathcal{S}\times\mathcal{B}$. Therefore, this approach can substantially reduce the complexity of decision-making. Moreover, $\pi^*_\infty \triangleq \arg\max_{U_\infty(\pi_\infty)\le e} G_\infty(\pi_\infty)$ can be calculated as shown in [70]. Policy π^*_∞ is the suboptimal policy and it is exactly what is needed. Therefore, in the following subsection, the performance of this suboptimal policy will be analyzed and compared with the optimal solution $\pi^*_T = \arg\max_{U_T(\pi_T)\le e} G_T(\pi_T)$.

3.3.2 Near-optimal Performance of Suboptimal Policy π^*_∞

By using the concepts of superior limit and policy space, the following theorem can be obtained.

Theorem 1 *For $\forall\epsilon_1 > 0$, there exists a sufficiently large $T(\epsilon_1)$ such that, for any network run time $T > T(\epsilon_1)$,*

$$max_{U_T(\pi_T)\le e} G_T(\pi_T) \le max_{U_\infty(\pi_\infty)\le e} G_\infty(\pi_\infty) + \epsilon_1. \tag{3.15}$$

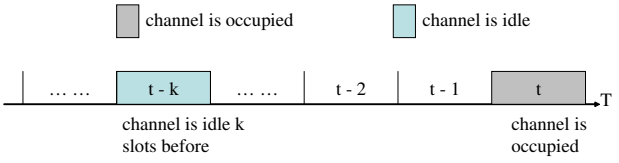

Figure 3.4: Illustrative example of channel states (time slot of t: $\varsigma_{0,0}$; time slot of $t-k$: $\varsigma_{1,k}$).

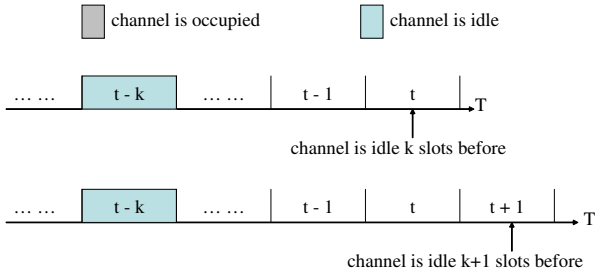

Figure 3.5: Illustrative case for the state transition probability when the node takes action $a_t =$'0' ($x'_t = \varsigma_{1,k}$ and $x'_{t+1} = \varsigma_{1,k+1}$).

or

$$G_T(\pi_T^*) \leq G_\infty(\pi_\infty^*) + \epsilon_1. \tag{3.16}$$

Proof: See 8.1.

In this theorem, $\max_{U_T(\pi_T)\leq e} G_T(\pi_T)$ represents the highest rate of data reception of **Problem** 2, whereas $\max_{U_\infty(\pi_\infty)\leq e} G_\infty(\pi_\infty)$ denotes the maximum reception rate of **Problem** 3. When T is sufficiently large, the performance of policy π_∞^* can be infinitely close to the performance of π_T^*. This implies that policy π_∞^* can be used as an approximate solution of π_T^*. In the hypothesis of this work, however, $G_\infty(\pi_\infty^*)$ is the performance of the system running for an infinitely long time (under **Assumption** 1). Therefore, it is also necessary to analyze its performance under real conditions (whereby the total energy storage of the node is limited and is defined as B), i.e., $G_T(\pi_\infty^*)$.

Theorem 2 *Suppose that nodes always have energy. Then, there exist $\epsilon_0 > 0$ and $M \in \mathbb{Z}^+$ such that the network performance difference after working time T from the average performance can be rewritten as follows:*

$$|G_\infty(\pi_\infty^*) - G_T(\pi_\infty^*)| \leq \frac{M-1}{T\epsilon_0}. \tag{3.17}$$

Proof: From the definition of $G_\infty(\pi_\infty^*)$, there exist $\epsilon_0 > 0$ (e.g., $\epsilon_0 = G_\infty(\pi_\infty^*)/2$) and positive integer M such that for arbitrary initial state x'_0, it can be obtained that

$$G_M(\pi_\infty^*) > \epsilon_0. \tag{3.18}$$

Let $\mathbb{P}_{\pi^*_\infty}$ be the probability transition matrix with state space $\mathcal{S}$ under policy π^*_∞. Assume $\gamma = (\gamma_s), s \in \mathcal{S}$ is the stationary distribution of $\mathbb{P}_{\pi^*_\infty}$. [1] Then, the component for state $\varsigma_{1,0}$ is $\gamma_{\varsigma_{1,0}} = \lim_{T\to\infty} \frac{1}{T}\mathbb{E}\{\sum_{t=1}^{T} 1_{\{x'_t=\varsigma_{1,0}\}}\}$. [2] From the previous definition, $\gamma_{\varsigma_{1,0}} = G_\infty(\pi^*_\infty)$, and this equality is independent of the initial state.

From the theorem of Doeblin [71], for any distribution μ, it can be obtained that

$$\begin{aligned} & |G_\infty(\pi^*_\infty) - G_T(\pi^*_\infty)| \\ \leq \; & \|\gamma - \frac{\mu}{T}\sum_{t=1}^{T} \mathbb{P}^t_{\pi^*_\infty}\|_\nu \\ \leq \; & \frac{M-1}{T\epsilon_0}, \end{aligned} \tag{3.19}$$

where the notation $\|\cdot\|_\nu$ denotes the norm $\|(x_1, x_2, \cdots)\|_\nu = \sum |x_i|$.

The above theorem illustrates that the average performance of policy π^*_∞ will converge to the mean performance under infinite time when the working time increases. On the other hand, when the energy usage is considered to be a random variable, based on the law of large numbers, it is obtainable that

$$\mathbf{Pr}\{\lim_{T\to\infty} U_T(\pi^*_\infty) = e\} = 1. \tag{3.20}$$

Thus, when T is sufficiently large, the total energy cost of policy π^*_∞ is approximately $eT = B$ w.p. 1. Furthermore, for any arbitrarily small $\epsilon_2 > 0$, when T is sufficiently large, set $\hat{T} = \lfloor \frac{B}{e+\epsilon_2} \rfloor = \lfloor T(1 - \frac{\epsilon_2}{e+\epsilon_2}) \rfloor$. [3] It can be seen that

$$\begin{aligned} & \mathbf{Pr}\{U_T(\pi^*_\infty) < e + \epsilon_2\} = 1, \\ & \mathbf{Pr}\{\textit{the total energy cost in } [1, \hat{T}] \leq \widehat{T}(e + \epsilon_2)\} = 1, \\ & \mathbf{Pr}\{\textit{the total energy cost in } [1, \widehat{T}] \leq B\} = 1. \end{aligned} \tag{3.21}$$

With this clarification, policy π^*_∞ can work for at least time $\widehat{T} = \left\lfloor T(1 - \frac{\epsilon_2}{e+\epsilon_2}) \right\rfloor$ w.p. 1 for nodes having a total energy storage of B. Moreover, since ϵ_2 can be infinitely small, while T is sufficiently large, $\widehat{T}$ will converge to the required time T.

By combination with Theorem 1, the following conclusion can be obtained.

[1] γ_s is the probability of occurrence for state s based on probability transition matrix $\mathbb{P}_{\pi^*_\infty}$.

[2] Function $1_{x=a}$ is equal to 1 when $x = a$ and is equal to 0 when $x \neq a$.

[3] Function $\lfloor x \rfloor$ denotes the largest integer that is not more than x.

Theorem 3 *For any $\epsilon > 0$, ϵ_0, and integer M given above, when the network operating time T is sufficiently large, the performance gap between the real performance of the node and that of the optimal solution can be rewritten as follows:*

$$\mathbf{Pr}\{G_T(\pi_T^*) - G_T(\pi_\infty^*) \leq \epsilon\} = 1. \tag{3.22}$$

Proof: If policy π_∞^* is conducted by node s for $\hat{T}$ slots, it can be known that

$$\begin{aligned} & G_T(\pi_T^*) - G_T(\pi_\infty^*) \\ \leq\ & G_\infty(\pi_\infty^*) + \epsilon_1 - G_T(\pi_\infty^*) \\ =\ & \epsilon_1 + G_\infty(\pi_\infty^*) - G_{\widehat{T}}(\pi_\infty^*) + G_{\widehat{T}}(\pi_\infty^*) - G_T(\pi_\infty^*) \\ \leq\ & \epsilon_1 + \frac{M-1}{\widehat{T}\epsilon_0} + \frac{T-\widehat{T}}{T\widehat{T}}. \end{aligned} \tag{3.23}$$

Here, the first inequality is derived from (3.16) in Theorem 1, and the second inequality is derived from Theorem 2. Since ϵ_1 can be arbitrarily small and $\lim_{T\to\infty}(\frac{M-1}{\widehat{T}\epsilon_0} + \frac{T-\widehat{T}}{T\widehat{T}}) = 0$, when the run time T is sufficiently large $\mathbf{Pr}\{G_T(\pi_T^*) - G_T(\pi_\infty^*) \leq \epsilon\} = 1$.

The theorem shows that when the network operating time T is sufficiently large, the performance obtained using suboptimal policy π_∞^* can approach the optimal performance w.p. 1. In practice, T is usually very large, so policy π_∞^* can be used as a good approximation of the optimal solution.

Note 1 In **Problem** *3, since only one constraint is involved, there is at most one randomization element in the near-optimal strategy $\pi_\infty^* = (d_{i,j})$ (see Section* 16.6*, page* 223 *of [70]), where $d_{i,j} = \mathbf{Pr}(a_t = \text{'1'} | \varsigma_{i,j}), \varsigma_{i,j} \in \mathcal{S}$. Assuming that $0 < d_{i_0,j_0} = \frac{m}{n} < 1$, $m, n \in \mathbb{Z}^+$, and other elements are 0 or 1. This shows that, in the special case, when the policy π_∞^* is known by node s and the sink node (which can be calculated in advance), let node s and the sink node run respectively. At time t, when the system condition is $x_t' \neq \varsigma_{i_0,j_0}$, node s and the sink node will operate according to the same policy and operate synchronously. Thus, at the next time, node s and the sink node will be in the same state. If $x_t' = \varsigma_{i_0,j_0}$, it can be supposed that in every n slots, the channel is sensed for m slots and the power is off for $n - m$ slots. Under this mechanism, the effect of d_{i_0,j_0} can be achieved and, at the same time, the network can be synchronized. As an example, consider the elements in the near-optimal strategy to be* $\{0, 1, 0, 0, 1, 0, 1, 1, ...\}$*, which can surely make the two nodes operate synchronously, i.e., "0" indicates sleep, while "1" indicates sensing. However, there is a special case in which the elements*

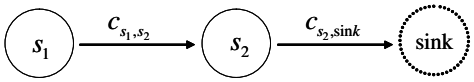

Figure 3.6: Network for two sensors.

may be {..., 0.4, ...}*, and the two nodes can still be made to operate synchronously. Suppose in a period of* 10 *time slots, both of the nodes are allowed to operate at the first four time slots (*1*,* 2*,* 3*, and* 4*). Then, let the two nodes sleep during the remaining six time slots (*5*,* 6*,* 7*,* 8*,* 9*, and* 10*), which can yield the same result.*

3.4 Multi-sensor Scenario

In the previous section, it has been discussed how to realize the communication and synchronization technology of a single node, and in this section, the results will be extended to a multi-node network.

For Fig. 3.6, sensor s_1 can only send data to the sensor sink in the manner of the s_1-s_2-sink. Since the data must be moved from node s_1 to node s_2 before they can be sent to the sink node, nodes s_1 and s_2 need to communicate first and store the data. Under this condition, the variable $T^o_{s_1,s_2}(T^o_{s_2,sink})$ is defined as the total time of data transmission from node s_1 (node s_2) to node s_2 (node *sink*) and the variable $B^o_{s_1,s_2}$ $(B^o_{s_2,sink})$ is defined as the energy consumption of node s_1 (node s_2) accordingly. Similarly, the variable $T^r_{s_1,s_2}$ represents the total time of data reception from node s_1 to node s_2, and the variable $B^r_{s_1,s_2}$ represents the energy consumption of node s_2 accordingly. Note that when sensor s_1 is sending data, node s_2 must be receiving the data at the same time so as not to waste energy. Only in this way can nodes s_1 and s_2 remain synchronized. For each node, the following inequality gives an estimate of the total energy consumption:

$$B^o_{s_1,s_2} \leq B, B^o_{s_1,s_2} = B^r_{s_1,s_2}, B^r_{s_1,s_2} + B^o_{s_2,sink} \leq B, \tag{3.24}$$

where it can be obtained that $T^o_{s_1,s_2} = T^r_{s_1,s_2}$, which means that nodes s_1 and s_2 have the same working condition and working time. [4] When node s_2 communicates with the *sink* node, it must maintain the same synchronization. Furthermore, when node s_2 is sending data, data packets to be transmitted must be stored in the buffer of node s_2. This requires that every time node s_2 makes a decision as to whether the stored data must be checked dynamically. In order to solve the problem simply, it is assumed that nodes s_1 and s_2 communicated for time $T^o_{s_1,s_2}$ before node s_2 communicated with the *sink* node. This means that, in this case, the network operating time is divided into two parts, one is for data transmission and

[4]Here, it has been assumed that $B^o_{s_1,s_2} = B^r_{s_1,s_2}$ in the energy consumption, which can be modified in the practical situation.

Problem 4

$$
\begin{aligned}
max_{\pi_{T^o_{s_1,s_2}},\pi_{T^o_{s_2,sink}}} \quad & \frac{T^o_{s_2,sink}}{T} G_{T^o_{s_2,sink}}(\pi_{T^o_{s_2,sink}}) \\
s.t. \quad & T^o_{s_1,s_2} = T^r_{s_1,s_2} \\
& T^o_{s_1,s_2} + T^o_{s_2,sink} \leq T \\
& B^o_{s_1,s_2} \leq B, B^o_{s_1,s_2} = B^r_{s_1,s_2} \\
& B^r_{s_1,s_2} + B^o_{s_2,sink} \leq B \\
& T^r_{s_1,s_2} G_{T^r_{s_1,s_2}}(\pi^*_{T^r_{s_1,s_2}}) = T^o_{s_2,sink} G_{T^o_{s_2,sink}}(\pi^*_{T^o_{s_2,sink}})
\end{aligned} \tag{3.26}
$$

reception between s_1 and s_2, and another is for data transmission and reception between s_2 and sink node. Each part consists of multiple time slots, and each time slot is further divided into multiple sub-processes, i.e., spectrum sensing, data transmission and reception. As for the channel allocation of each node in more complicated configurations, a method like [72] will be required, which is out of the scope of this work and will be left for future work. On the same link, the total time consumption should meet the following condition:

$$
T^o_{s_1,s_2} = T^r_{s_1,s_2}, T^o_{s_1,s_2} + T^o_{s_2,sink} \leq T. \tag{3.25}
$$

In this manner, the average energy consumption for sending data from node s_1 to node s_2 is $e^o_{s_1,s_2} \triangleq B^o_{s_1,s_2}/T^o_{s_1,s_2}$. For node s_2, the power consumed in sending data to the *sink* node is $e^o_{s_2,sink} \triangleq B^o_{s_2,sink}/T^o_{s_2,sink}$. Thus, the problem can be transformed into **Problem** 4, where the last equation indicates that the total number of data packets sent by node s_2 is equal to that received by node s_2. In **Problem** 4, it is necessary to find the optimal $(T^{o*}_{s_1,s_2}, B^{o*}_{s_1,s_2})$ and $(T^{o*}_{s_2,sink}, B^{o*}_{s_2,sink})$, as well as the optimal transmission strategy of node s_1, $\pi^*_{T^o_{s_1,s_2}}$ (or the optimal receipt strategy of node s_2, $\pi^*_{T^r_{s_1,s_2}}$), and the optimal transmission strategy of node s_2, $\pi^*_{T^o_{s_2,sink}}$. Note that owing to the synchronization, it can be verified that $\pi^*_{T^o_{s_1,s_2}} = \pi^*_{T^r_{s_1,s_2}}$. Since the total network run time T is quite large, it can be made that $T^o_{s_1,s_2} = T^r_{s_1,s_2} = T^o_{s_2,sink} = T/2$. In this way, the problem is transformed, and it is only necessary to find the power consumption of each node $B^{o*}_{s_1,s_2}, B^{o*}_{s_2,sink}$ and its strategy. Now, the optimal reception rate obtained through **Problem** 4 is defined as G^*_{sink}. Similar to Section 3.3, when calculating $\pi^*_{T^o_{s_1,s_2}}$ and $\pi^*_{T^o_{s_2,sink}}$, complexity may be problematic, and the real optimal solution cannot be obtained actually. In this case, the approach described in the previous section is used, with an approximate solution and the algorithm presented in the following table (Algorithm 1).

Algorithm 1 Algorithm for the Multi-sensor Scenario

Step 1	Define $Y = T/2$, $B_1 = 1$, and $B_2 = B - B_1$.
Step 2	Take B_1/Y and B_2/Y as the average energy usages for nodes s_1 and s_2, respectively, and use Section 3.3.2 to calculate the optimal strategies π^1_∞ and π^2_∞.
Step 3	If $G_Y(\pi^1_\infty) = G_Y(\pi^2_\infty)$, it can be obtained that $B^r_{s_1,s_2} = B_1$ and $B^o_{s_2,sink} = B_2$, the transmission and reception strategy of nodes s_1 and s_2 is π^1_∞, and the transmission strategy from node s_2 to the node sink is π^2_∞. Then, proceed to Step 4. Otherwise, define $B_1 = B_1 + 1$ and $B_2 = B - B_1$, and repeat Step 2.
Step 4	The total number of data packets for node s_2 received from node s_1 is $YG_{T^o_{s_1,s_2}}(\pi^1_\infty)$, and the average reception rate of the sink node is $G'_{sink} \triangleq \frac{1}{T} min\{YG_{T^o_{s_1,s_2}}(\pi^1_\infty), YG_{T^o_{s_2,sink}}(\pi^2_\infty)\}$.

In Step 2 of Algorithm 1, the optimal strategies π^1_∞ and π^2_∞ within the infinite range is used to replace the optimal strategy within the finite range. Thus, the total power consumptions of node s_2 for reception, i.e., $B^r_{s_1,s_2}$, and transmission, i.e., $B^o_{s_2,sink}$, are B_1 and B_2, respectively, according to the condition whereby the sum of data reception through node s_2 is equal to that of data transmission in Step 3. Here, the resultant gap between the power consumption and that of the optimal solution, as well as the difference in the average data reception rate of the sink node $G^*_{sink} - G'_{sink}$, where $G_Y(\pi^*_{T^o_{s_2,sink}}) = G^*_{sink}$ and $G_Y(\pi^2_\infty) = G'_{sink}$, is analyzed, as shown in 8.2.

Theorem 4 *For a random $\epsilon > 0$, ϵ_0, and integer M given above, when the network working time T is sufficiently large, the performance gap between the real performance G'_{sink} of the node and the optimal solution G^*_{sink} can be rewritten as follows:*

$$\mathbf{Pr}\{G^*_{sink} - G'_{sink} \leq \epsilon\} = 1. \tag{3.27}$$

Proof: See 8.2.

The above theorem shows that the data reception by the algorithm for the multi-sensor scenario and that by the optimal algorithm do not differ significantly.

In fact, their values will tend to be equal with probability 1. Thus, the solution of the original problem is obtained. When the nodes need to transmit data to the sink node through two intermediate nodes, it needs only to make $Y = T/3$, and, after each node takes Y time, the data can be forwarded to the end. [5]

3.5 Simulation Results

In this section, the numerical results is presented for evaluating the performance of the proposed scheduling and compare the proposed scheduling with other schedulings.

The simulation parameters are chosen from common hardware settings (i.e., MICA2 [73]). In this case, the energy capacity of a node, two batteries with $1000mAh$ is $10.8KJ$. Since the current in active mode is $8mA(3V)$ and that for packet transmission is $25mA(3V)$, $b_1 = 1$ is choosed and accordingly $b_2 = 3$. It is assumed that a slot duration is $1s$, and B=450,000. Moreover, the operating time of the network is T = 2,000,000. The simulation is conducted as follows. Since the proposed strategy can specify how each sensor operates at each time slot, in the settings of this work, the sensors are allowed to continue operating in accordance with this strategy for T rounds, where each round indicates one time slot. Then the total data packets received is added up. Next, based on the cumulative value of data packets that have arrived, the data reception rate can be defined as the ratio of the total received packets to the time T, which actually represents the value of the ordinate in each figure. Note that the data reception rate in the presented results can be converted into other units. [6]

Figures 3.7 and 3.8 show the performance of the near-optimal schedule π^*_∞ for the single-sensor scenario under different spectrum occupancy statistics and battery capacities, respectively. As shown in Section 3.3, since the true optimal policy π^*_T is difficult to obtain, the upper bound is used to verify the proposed policy. The upper bound is derived by the dual linear program according to [74] and has been described in 8.3. As shown in Fig. 3.7, under different probability matrices $\mathbb{P}$ (two sets: $a = 0.8$ and b varies from 0.1 to 0.9; and $a = 0.6$ and b varies from 0.1 to 0.9), the data reception rate achieved by the schedule π^*_∞ matches the upper bound. This confirms the results in Theorem 3, verifying that the proposed

[5]The network operating time and energy of nodes can be averaged and divided into multiple segments. When the energy of a node is exhausted before the end of a certain time segment, operation cannot continue during the remainder of this time segment. The system can be configured to begin operation from a new time segment by adopting battery charging technology.

[6]Suppose the bit length for one data packet is $\hat{b}$. Then one time slot includes $\hat{s}$ seconds, and the average data reception rate is G_T. By multiplying G_T by $\hat{b}/\hat{s}$, i.e., $G_T \times (\hat{b}/\hat{s})$, the data reception rate can be converted to the unit of bits per second ($bits/second$).

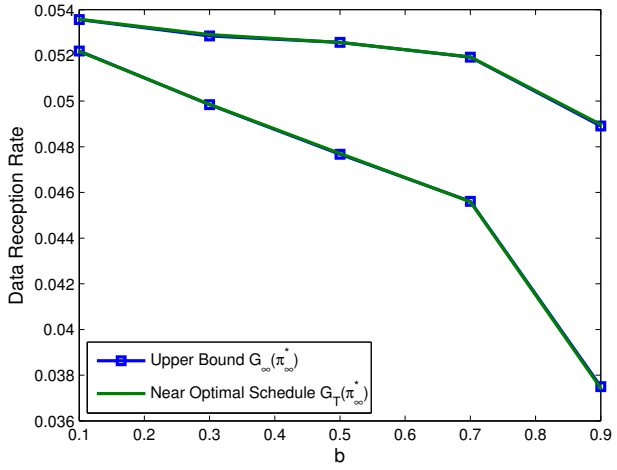

Figure 3.7: Performance of near-optimal scheduling π^*_∞ under different spectrum occupancy conditions. (The above is for $a = 0.8$, while the below is for $a = 0.6$. The unit used to express the data reception rate is packets/slot, which is the same as in the other figures.)

schedule is a good approximation of the optimal solution. Similarly, in Fig. 3.8, given the spectrum occupancy statistics of $a = 0.6$ and $b = 0.7$, the achieved reception rate is also close to the upper bound. When the initial energy budget of the sensor becomes large, it increases with the upper bound because more data packets will be received by the sink node during the same operating time. It is evident that, in both cases, the packet reception simulation matches the analytical expression.

In Fig. 3.9, packets are routed through three sensor nodes, in a manner similar to that shown in Fig. 3.6. In order to simplify the analysis, it is assumed that the channels between the sensors ($c_{s_1,s_2}, c_{s_2,s_3}, c_{s_3,sink}$) follow identical Markovian dynamics, as specified by the following matrix:

$$\mathbb{P} = \begin{pmatrix} 0.6 & 0.4 \\ 0.3 & 0.7 \end{pmatrix}, \tag{3.28}$$

where a and b can be chosen arbitrarily. In this setting, the proposed policy π^*_∞ is compared with two other policies, i.e., the periodic policy π_P and the aggressive policy π_A [69]. For the periodic policy, in each period (the period length is 100 slots in the simulation), sensors will be activated for a fixed duration. The duty cycle is used to denote the fraction of the active times of the sensors. In the aggressive policy, the sensors will sense the channels and transmit or receive data

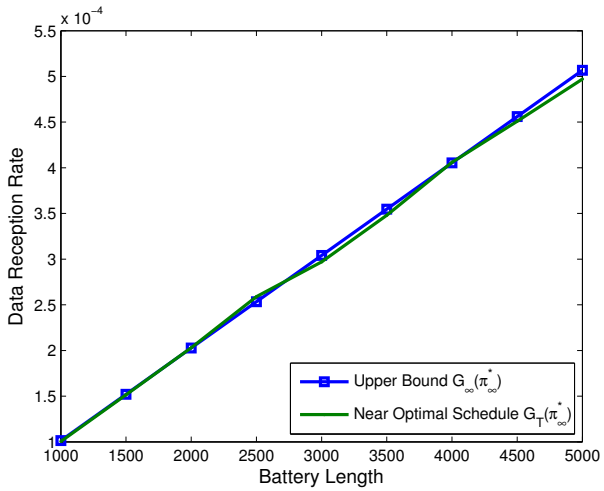

Figure 3.8: Performance of near-optimal scheduling π^*_∞ under different battery capacities ($a = 0.6, b = 0.7$).

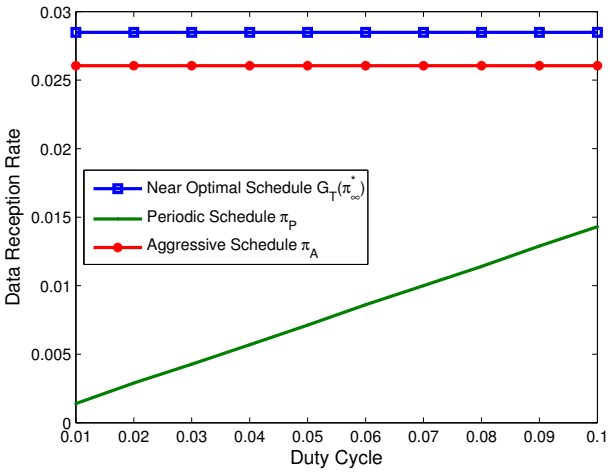

Figure 3.9: Performances of the periodic, aggressive approach and the near-optimal strategy π^*_∞.

whenever they have energy. Moreover, synchronization is assumed to be maintained for π_P and π_A. As shown in Fig. 3.9, π^*_∞ is better than π_P and π_A, because policy π^*_∞ will use the information from the channels and its decision will change with the channel's dynamics. Then, sensors tend to send packets when channels

are of good quality. Thus, energy can be saved. However, for policies π_P and π_A, the decision rules of the sensor are constructed offline. The policies cannot follow changes in the environment and so will have poorer performance.

3.6 Summary

This work has presented a CRN sensing policy that uses the channel's dynamics in the routing. This policy will search for spectrum opportunities without using a common communication channel and will assign the decision of each sensor in every slot under an energy constraint. Unlike existing techniques, which attempt to find the optimal solution directly, this work finds approximate solutions by effectively reducing the system state space, thereby reducing the energy sensors used to make decisions. Although the designed sensing policy has a simple structure, it can ensure near-optimality and is a good approximation for the true optimal schedule. According to the theoretical analysis, the proposed near-optimal policy is less complex but offers comparable performance. As the operating time of the CRN approaches infinity, the data reception rate on the sink node side will converge to the optimal rate w.p. 1.

Although the analysis of the present study is concentrated on a sensor network, the proposed policy possesses good generality and can be fit into other kinds of network scenarios, such as delay-tolerant networks with energy constraints and cognitive-radio-technology-enabled ad-hoc networks. In the future, the optimal cognitive sensing and access policies for energy harvesting sensors based on the proposed policy will be investigated.

Chapter 4

Augment SpectrumMap with User Crowdsourcing - A Novel User Participation Model for Data Collection

4.1 Introduction

Prior work usually assumes dedicated spectrum measurements by authorized organization to derive the distribution of spectrum utilization according to propagation models based on real-world data logs [29, 31–35]. These work is extremely significant and has valid contribution from the aspect of understanding the overall white space channels of networks. However, according to the recent study of [41], the database approach solely relying on propagations models is more likely to introduce errors in estimating the channel usage. Result has shown that the commercial database has caused non-negligible wastage of whitespace channels at almost 71% measured locations according to the observed data within 120 square-km area. It has also been verified the large-scale spectrum data collections by leveraging spectrum monitors in large-scale areas can greatly reclaim this spectrum wastage. Besides dedicated spectrum measurements, SpectrumMap database is designed to include opportunistic spectrum measurement by public users (Fig. 4.1), since large-scale distributed local datasets can assist the database to better deal with white space channels [41]. This section proposes a incentive based user participation model to encourage network users to actively contribute and share their local spectrum observations with SpectrumMap.

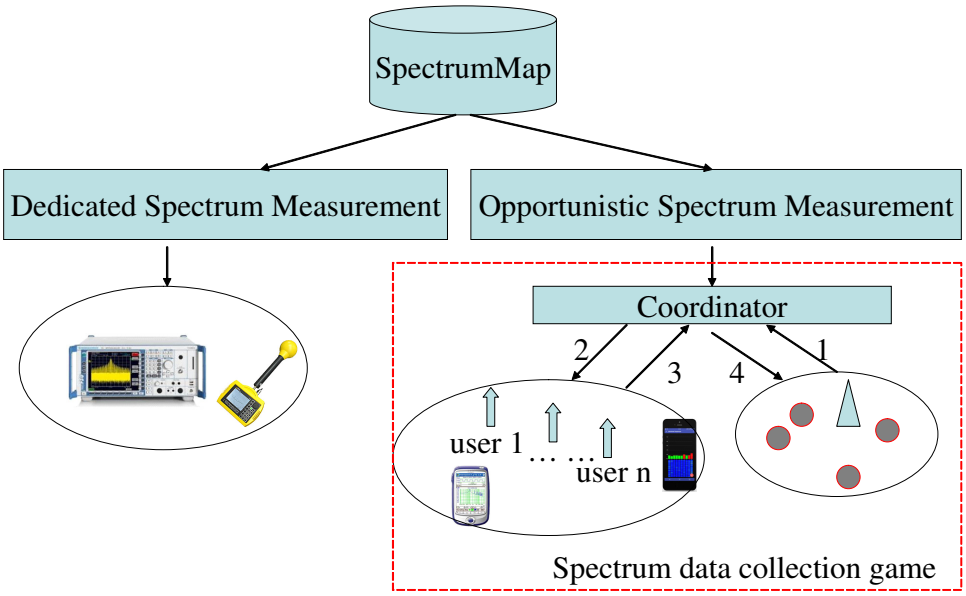

Figure 4.1: User participation model (1: requesting radio data for performing self-organization; 2: broadcast sensing tasks to pubic users; 3: accept the task and collect data; 4: delivery collected data to the required users.).

4.2 User Participation Model and Problem Formulation

A typical spectrum data format can be defined by the dataset of $< Time, Latitude, Longitude, Channel, Power\ intensity >$. The proposed user participation model is actually a spectrum data collection game relying on incentivising public users to provide more available radio data in the case of necessary through user's cooperation. One user might not be able to provide spectrum data to SpectrumMap due to selfishness as well as energy constrain, while multiple users can yield more available spectrum data if they form coalitions through effective cooperation. Fig. 4.2 shows the proposed data collection model. The empowerment of the model can yield high availability of spectrum datasets through the cooperation among network users.

Each user is assumed to have two types of behavior, the contribution and consumption. The contribution behavior, i.e., uploading and sharing spectrum data, is rewarded by *token*, while the consumption behavior, i.e., requesting SpectrumMap for spectrum sensing information, is conditioned on paying token to SpectrumMap. Hence, a trading market for the SpectrumMap can be constructed with token. Meanwhile, SpectrumMap can invite any user to jointly participate spectrum data collection when necessary. Note that cooperative users can be het-

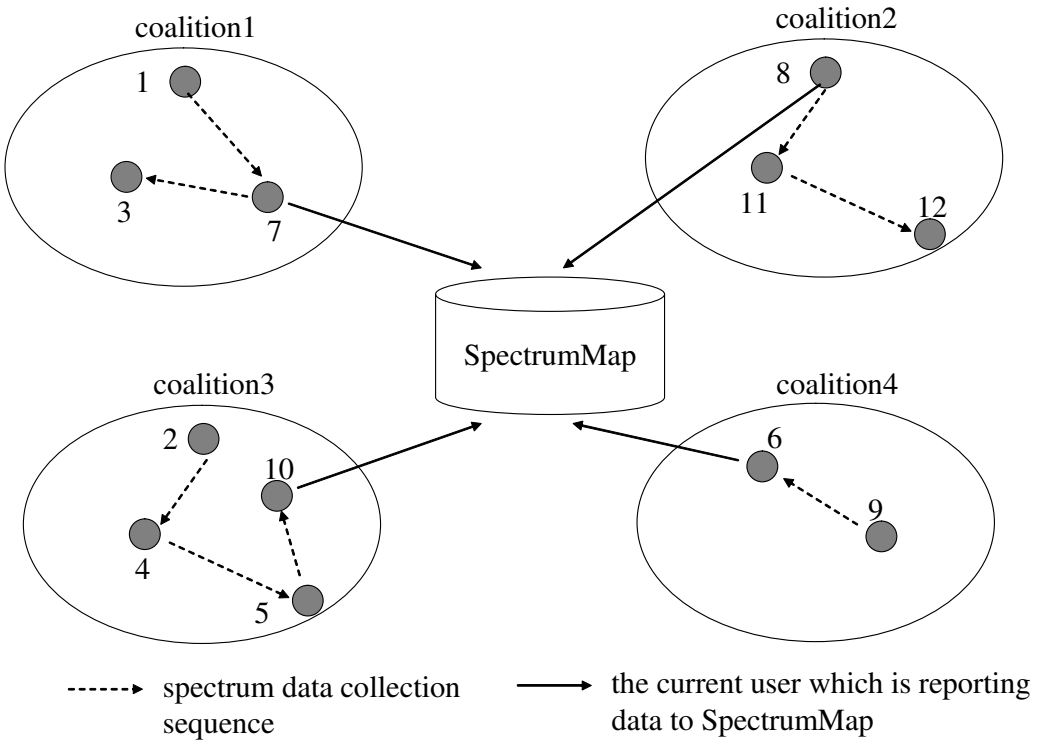

Figure 4.2: User cooperation based data collection (Within each coalition and at one time slot, only one user performs data collection and other users keep idle and wait for their turn by a predefined collection sequence to collect data, i.e., in the coalition1, the data providing sequence is (user 1) $\longrightarrow$ (user 7) $\longrightarrow$ (user 3), which indicates that user 1 collects data at first, then will be user 7, and finally will be user 3. The current data collector is user 7.).

erogeneous, i.e., the members of a coalition can be cognitive users of IEEE 802.22 network, mobile users as well as small cell users of 5G network, etc., since they have common objective to optimize frequency combinations by learning the radio environment. This work assumes a low level of participation due to network users are strategy and autonomous. Accordingly, this work casts the following problem to be solved: *Considering the energy constrain and selfish degree of each autonomous user, to what extent that an cooperative mechanism could be designed for network users to output more available spectrum raw data.* Some definitions are given as follows for simplicity of description:

1. P_{SD}: a value of $(0, 1)$ to formulate the participation capability of a user, that is how much possibility of a user that can contribute its spectrum data. The value can be derived according to the total times of data collection request

from SectrumMap and total times of data uploading agreement from users.

2. R: the rewarded remuneration to a user and formulated by token. The value of the R is jointly determined by the participation capability and the allocation strategy about the token. The allocation strategy determines the amount of rewarded token, which can valuate the activity of each user. More positive action a user performs in spectrum data collection, the more token is obtainable.

3. C: the consumed energy cost for a user. The energy is mainly consumed by spectrum measurement and data delivery to the information center of SpectrumMap.

4. V: the value of a user, which is jointly decided by its rewarded token and cost, such as

$$V = R - C \tag{4.1}$$

4.3 Introduction of Coalitional Game

A spectrum data collection game based on coalitional game approach is proposed to formulate the user participation model. The coalitional game is originated from cooperative game theory and usually expressed by (N, ν), where $N = \{1, 2, 3, ..., n\}$ indicates a set of players, and ν denotes the utility of players. The output of the game is a coalition partition, which is indicated by $\Pi = \{\mathcal{S}_1, ..., \mathcal{S}_K\}$. It can divide the total players N into multiple coalitions, i.e., $\forall k, k \in \{1, ..., K\}$, $\mathcal{S}_k \subset N$ are disjoint coalitions with the constrain $\cup_{k=1}^{K} \mathcal{S}_k = N$. To decide possible coalition partitions for a game, each player has its own choice to join a coalition, and such behavior is denoted by $\succeq_i$. For example, $\mathcal{S}_1 \succeq_i \mathcal{S}_2$ means that game player i desires to join coalition $\mathcal{S}_1$, or choose coalition $\mathcal{S}_1$ or $\mathcal{S}_2$ equally. Further detail concepts about the coalitional game, such as merge and split rule, etc., can be found from [83].

4.4 Evaluating Spectrum Data Collection by Coalitional Game

Multiple users can improve the data collection if they cooperate with each other. That is, within a group consisted of different users, by reaching the agreement with each member, when a user is unable to contribute the data, another candidate user

can be available to collect data. Such cooperation is defined as the sequential-participation in this work. It can be observed that the sequential-participation based coalition can yield high availability of spectrum data collections. It will be shown in the subsequent part that the cooperation among users by the sequential-participation is affected by a data providing sequence (DPS), which is defined as follows:

4.4.1 Sequential-participation based Coalition formed by Users

Definition 1: For a coalition $\mathcal{S}$ consisted by total number of M users, i.e., $\mathcal{S} \in \{U(1), U(2), ..., U(M)\}$, DPS is a kind of *alignment* used to describe the data collection order for coalition members, that is in which sequence users in a coalition should follow to collect spectrum data. Denote π as the DPS, then

$$\pi \Leftrightarrow [\pi[1], \pi[2], ..., \pi[m], ..., \pi[M]], \tag{4.2}$$

where $\pi[m]$ is a integer ($1 \leq m \leq M$) to indicate the sequence index for a user in the coalition. For easy of description, denote by $U_{\pi[m]}$ the user of a coalition with the sequence index $\pi[m]$. SpectrumMap will first ask for the $U_{\pi[1]}$ to monitor the local data and share it with the database; when $U_{\pi[1]}$ agrees with the request, spectrum monitor will be performed by $U_{\pi[1]}$; if $U_{\pi[1]}$ refuse to cooperate with SpectrumMap, SpectrumMap will negotiate with $\pi[2]$ to report its spectrum data. With the similar operation, if $\pi[2]$ ignore the request, SpectrumMap will continue to negotiate with other users until finding some available $U_{\pi[m]}$ according to the DPS. The data contribution possibility of a coalition including M users is defined as follows:

$$D_S = 1 - \Pi_{m=1}^{M}(1 - P_{SD}(\pi[m])), \tag{4.3}$$

where $P_{SD}(i)$ is the spectrum data participation capability of $U(i)$ of the coalition $\mathcal{S}$.

Accordingly, the cost of a coalition can be defined as

$$C_S^{\pi} = \begin{cases} C(\pi[1]) & M = 1 \\ C(\pi[1]) + \\ \quad \sum_{m=2}^{M} \Pi_{l=1}^{m-1}(1 - P_{SD}(\pi[l]))C(\pi[m]) & M > 1 \end{cases}$$

where $P_{SD}(\pi[l])$ is the participation capability of the user $U_{\pi l}$, and $C(\pi[m])$ is the cost of the user $U_{\pi m}$.

Denote $U^{unit}(\pi[m])$ the rewarded unit token to a user $\pi[m]$. Before giving the coalition value for the considered game formulation, two kinds of token allocation strategies are defined to motivate users to join the game.

1. Non-uniform unit token allocation strategy: The amount of unit token, i.e., the rewarded number of token each unit of data contribution, depends on users' participation level. More spectrum data participation capability will yield more number of unit token. The rewarded unit token for a user given by

$$non-uniform: \quad U^{unit}(\pi[m]) = \mathcal{U}^{unit} \times \frac{P_{SD}(\pi[m])}{1}, \tag{4.4}$$

   where $\mathcal{U}^{unit}$ is the unit token to a user when its spectrum data participation capability is 1 and the value can be predetermined by the SpectrumMap.

2. Uniform unit token allocation strategy: Users is treated equally by ignoring the difference in participation capability and assigned the same level of unit token, such as

$$uniform: \quad U^{unit}(\pi[m]) = \mathcal{U}^{unit}, \tag{4.5}$$

The allocation strategy can be summarized as $U^{unit}(\pi[m]) = \mathcal{U}^{unit} \times P_{SD}(\pi[m])^{\alpha} (\alpha = 0 : uniform; \alpha = 1 : non-uniform)$. The coalition value of $\mathcal{S}$ which has a data providing sequence of π can be expressed by

$$v_S^{\pi} = U_S^{unit} \times D_S - C_S^{\pi}, \tag{4.6}$$

where U_S^{unit} is the rewarded unit token to the coalition S according to the participation capability of the total coalition, i.e., $U_S^{unit} = \mathcal{U}^{unit} \times D_S^{\alpha} (\alpha \in \{0, 1\})$. In the following, the game property of non-uniform allocation strategy by assuming $\alpha = 1$ is addressed at first, then the game property of uniform allocation strategy of $\alpha = 0$ is conducted.

4.4.2 Game Property

The proposed spectrum data collection model can be evaluated by a coalitional game (N, ν), where N is total players of users, and ν is a coalitional value defined by Eq. (4.6). By assuming the payoff of a user i in any two coalitions is $x_i(\mathcal{S}_1)$ and $x_i(\mathcal{S}_2)$, respectively, the preference of the user can be defined as

$$\mathcal{S}_1 \succeq_i \mathcal{S}_2 \Leftrightarrow x_i(\mathcal{S}_1) \geq x_i(\mathcal{S}_2) \tag{4.7}$$

Some important properties are presented below for the considered data collection in the view of game theory approach:

Property 1: The proposed game model for spectrum data collection is a coalitional graph game with transferable utility [83](The graph attribute here indicates

that the coalitional value depends on how users are connected with each other based on data providing sequence; the transferable utility means the obtained token of a user can be transfered to another user by distributing the value of (Eq. (4.6)) to coalitional members.).

Property 2: The designed game is not cohesive(This is due to the reason that, from Eq. (4.4.1), forming a grand coalition introduces heavy cost.).

For the proposed game approach, determining optimal interaction relationship between users is important. Hence, the following theorem is proposed for yielding the maximum payoff for each user.

Theorem 1 *The maximum value of a coalition is achievable when users in the coalition is coordinated as follows:*

$$C(\pi[m+1]) \times P_{SD}(\pi[m]) - C(\pi[m]) \times P_{SD}(\pi[m+1]) \qquad \geq 0 \tag{4.8}$$

where $1 \leq m \leq (M-1)$.

Accordingly, with the above result, the optimal cooperation for spectrum data collection among users can be formed.

4.4.3 Payoff Allocation for Each User

The one of *tegalitarian* [84] is applied to determine the payoff allocation for users. The division rule adopted here has lower computation complexity compared with another one of *Shapley value*, and the extra value of a coalition will be distributed equally to the users which join the coalition. Hence, the payoff for the user with alignment sequence of $\pi[m]$ in a coalition $\mathcal{S}$ can be obtained as follows:

$$x_S(\pi[m]) = \frac{1}{|\mathcal{S}|}(v_S^{\pi} - \sum_{k=1}^{M} v(\pi[k])) + v(\pi[m]), \tag{4.9}$$

where $v(\pi[k]))$ and $v(\pi[m]))$ can be obtained similarly according to Eq. 4.6, i.e., $v(\pi[k])) = U^{unit}(\pi[k]) \times P_{SD}(\pi[k]) - C(\pi[k])$.

4.5 Distributed Coalition Formation Algorithm

To form optimal cooperations among users, a coalition formation algorithm is proposed for users to choose their cooperative partners and consequently, each user can achieve a satisfied payoff. The following switch rule [82] for coalition formation has been adopted:

Definition 2: With a coalition partition defined as $\Pi = \{\mathcal{S}_1, ..., \mathcal{S}_K\}$ of the total players of users, the ith user decides to leave $\mathcal{S}_k$ $(1 \leq k \leq K)$, the current coalition

Table 4.1: System Parameters used for Validating Spectrum Access Performance

Channel gain	$\sim exp(1)$	Path loss exponent	3.7
a small cell (γ)			
Transmit power (CBS)	43dBm	Transmit power (SBS)	23dBm
Intensity of CBS	$\frac{1*10^{-6}}{m^2}$	Intensity of small cell	$\frac{8*10^{-5}}{m^2}$
Total channels for a smll cell	k	Total channels for a CBS	$20-200$
Number of users for a small cell	4	Radius of a CBS	500m
Radius of smll cell	40m	QoS threshold of	10,15dB

it belongs to, and join another $\mathcal{S}_m$, where $\mathcal{S}_m \in \Pi \cup \{\phi\}$ and $\mathcal{S}_m \neq \mathcal{S}_k$, then a new partition $\Pi' = \{\Pi \setminus \{\mathcal{S}_k, \mathcal{S}_m\}\} \cup \{\mathcal{S}_k \setminus \{i\}, \mathcal{S}_m \cup \{i\}\}$ is formed with the condition that $\mathcal{S}_m \cup \{i\} \succeq_i \mathcal{S}_k$. As a result, $\{\mathcal{S}_k, \mathcal{S}_m\} \rightarrow \{\mathcal{S}_k \setminus \{i\}, \mathcal{S}_m \cup \{i\}\}$ and $\Pi \rightarrow \Pi'$.

Follow the above switch rule, a user can leave current coalition $\mathcal{S}_k$ and choose to join another coalition $\mathcal{S}_m$ to join, and a distributed coalition formation algorithm can be developed accordingly.

Property 4: The proposed coalition formation algorithm always converges to a final network partition consisted by multiple disjoint coalitions of users. Because no user desires to switch from its current coalition to another coalition or to deviate and act alone, the partition is stable.

4.6 Uniform Token Allocation Strategy

The non-uniform token allocation based game formulation has been discussed in the above, i.e., $\alpha = 1$. Now this section evaluates the allocation strategy by applying uniform token allocation, such as $\alpha = 0$. The following theorem is proposed.

Theorem 2 *By* uniform *token allocation strategy, each user will collect spectrum data separately and the final coalition partition is* $\Pi = \{\{1\}, \{2\}, ..., \{n\}\}$.

4.7 Performance Evaluation for Spectrum Data Collection

The proposed spectrum data collection is evaluated by Monte Carlo simulation with Matlab. two kinds of mechanism have been considered, one is the pro-

posed cooperative game based spectrum data collection (C-RED) and another is the conventional mechanism, i.e., the non-cooperative spectrum data collection (NC-RED) which assumes no cooperation among users and each user acts alone. The data participation capability of each user $P_{SD}(\pi[m])$ is uniformly distributed in (0.2,0.4), (0.7,0.9), and (0,1), respectively. By the value of any single user according to Eq. 4.6, i.e., $v(\pi[m])) = U^{unit}(\pi[m]) \times P_{SD}(\pi[m]) - C(\pi[m])$, the normalized energy cost $C(\pi[m])$, used for spectrum measurement and data transmission, is uniformly distributed in $(0, U^{unit}(\pi[m]) \times P_{SD}(\pi[m])$, since it is necessary to make the value of each user to be always positive.

For fixed data participation capability and cost of users, the proposed game can construct optimal cooperations among users. Fig. 4.3 shows the snapshot of a final formed coalitions yielded by the proposed model. There are altogether six coalitions have been formed, i.e., S1, S2, S3, S4, S5 and S6. Each user includes two parameters, such as its spectrum data participation capability and cost, shown by $(P_{SD}(i), Ci)$ in the figure. Note that the cost is normalized, which can be considered as the consumed energy for data collection. Within each coalition, DPS is plotted based on the order for users to collect spectrum data within a coalition. According to the formed final partitions, the available spectrum data for one coalition and payoff for each user can be maximized.

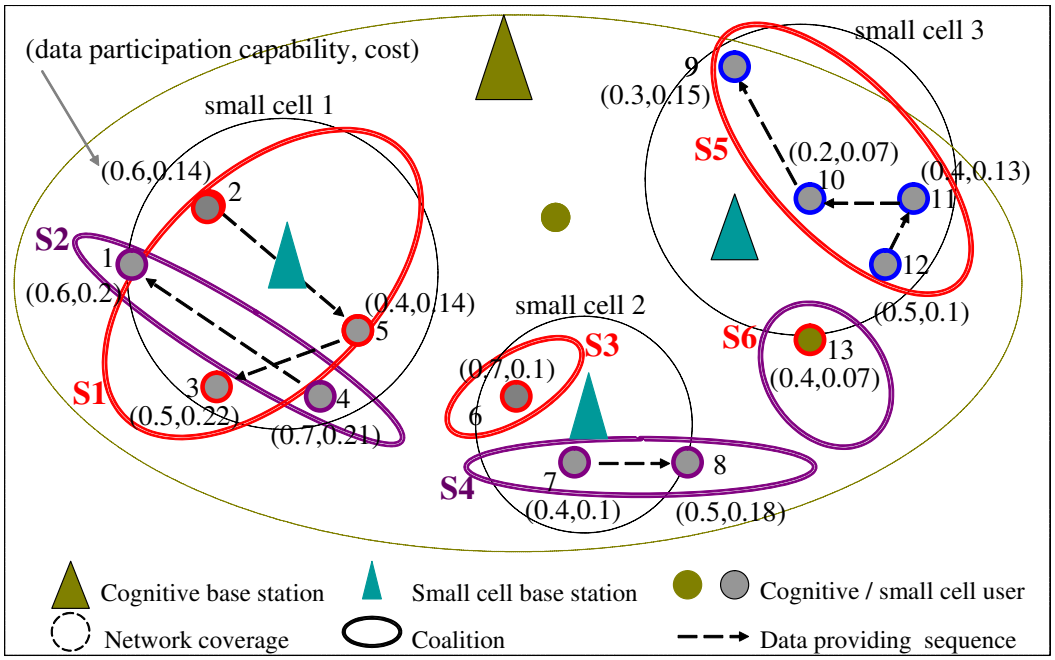

Figure 4.3: A snapshot of final coalition formation (The scenario consists of six stable coalitions. For example, coalition $S1$ includes three users: 2, 3 and 5; and data providing sequence is $\{ 2 \longrightarrow 5 \longrightarrow 3 \}$.

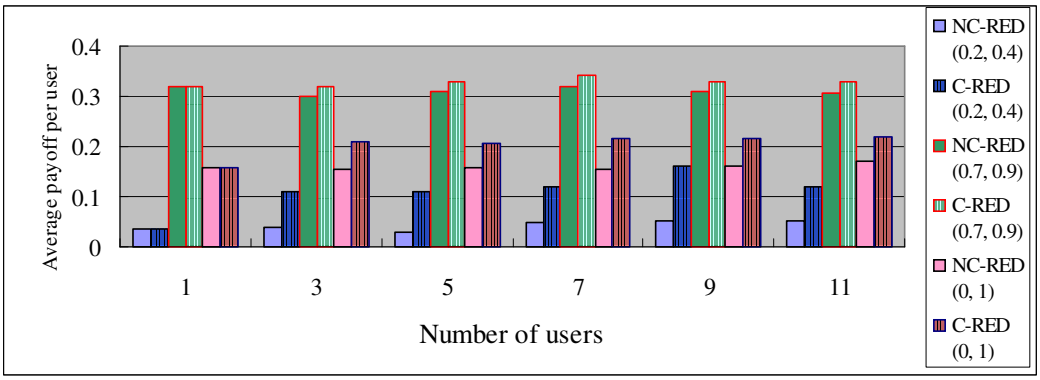

Figure 4.4: Average payoff for each user.

The average payoff of each user is shown in Fig. 4.4, based on various of data participation capability and cost. It can be observed that more average payoff is obtainable. For the participation capability uniformly distributed in (0.2, 0.4), (0.7, 0.9) as well as (0, 1), the proposed C-RED can achieve the improvement about 140%, 5%, 30%, respectively, compared with the conventional N-RED. The proposed algorithm is more beneficial to the users which are in lower spectrum data participation capability. It is observed that the average payoff remains nearly a same level with the increased number of users, which indicates that C-RED can maintain the coalition size to a proper value. This can be further confirmed by Fig. 4.5. It can be observed that more average payoff is obtainable. For the participation capability uniformly distributed in (0.2, 0.4), (0.7, 0.9) as well as (0, 1), the proposed C-RED can achieve the improvement about 140%, 5%, 30%, respectively, compared with the conventional N-RED. The proposed algorithm is more beneficial to the users which are in lower spectrum data participation capability. It is observed the average payoff remains nearly a same level with the increased number of users, which indicates that C-RED can maintain the coalition size to a proper value. This can be further confirmed by Fig. 4.5. Fig. 4.5 presents the average coalition size for the final coalition partition. The average coalition size for C-RED keeps fairly constant when the number of users is increased from 3 to 11, i.e., the average size keeps at the level between 2 and 2.5 when P_{SD} is distributed over (0, 1).

4.8 Validate Spectrum Access Performance Using the Proposed Data Collection Model

The previous parts have demonstrated to incentivize users to participate the SpectrumMap and contribute spectrum datasets in the above. Now this section pro-

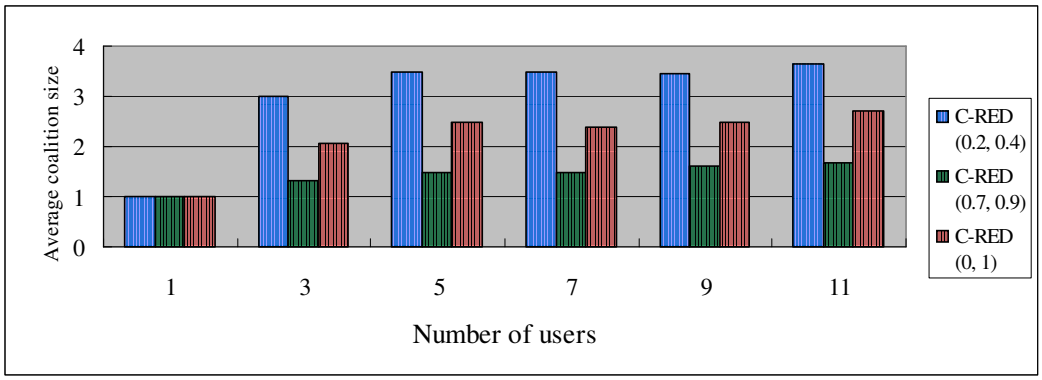

Figure 4.5: Average number of users in a coalition.

ceeds to validate the effectiveness of the user contributed datasets to spectrum access performance by utilizing the proposed data collection model, assisted by a experiment conducted by Monte Carlo simulation of Matlab. A interference scenario caused by heterogeneous coexistence [45] is considered, and specifically, the network scenario consisting of cognitive heterogeneous cell network is investigated, where independently-operated cognitive cells within the same geographical area operate under the same portion of TV white space channels [46, 47] (The standardization effort regarding to 802.11 has tried to address the cognitive Wi-Fi by using TV bands, and LTE operator is also looking forward to this issue.) Without loss of generality, there exists a network area in a $2,000m \times 2,000$ m square existed, where several cognitive cells has been deployed to share the same TV white space channels. Within each cognitive cell, similar to Fig. 4.3, multiple small cells has been deployed densely and underlay with the cognitive cell, also operated in the same TV band with cognitive cell. Since the conventional commercial spectrum database can only manage vertical coexistence interference, i.e., the coexistence between primary TV transmitters and cognitive cell users, while heterogeneous coexistence, such as interference among secondary networks, is still missing. This part will demonstrate the proposed user participation model can assist the performance improvement for heterogeneous coexistence. Accordingly, it has been assumed that a cognitive base station first send a request message to the commercial spectrum database (The one that likes Geo-location database regulated by FCC.) [43] and then with a permission to operate on a set of vacant channels, denoted by $\mathbb{W} = \{1, 2, ..., W\}$, which can be shared among multiple cognitive cells. For a single cognitive cell, the used topology is similar to Fig. 4.3.

All the positions of cognitive cell base stations as well as small cell base stations follow a poison point distribution. Downlink transmission of a target small cell is considered, whose data communication is interfered by multiple surround-

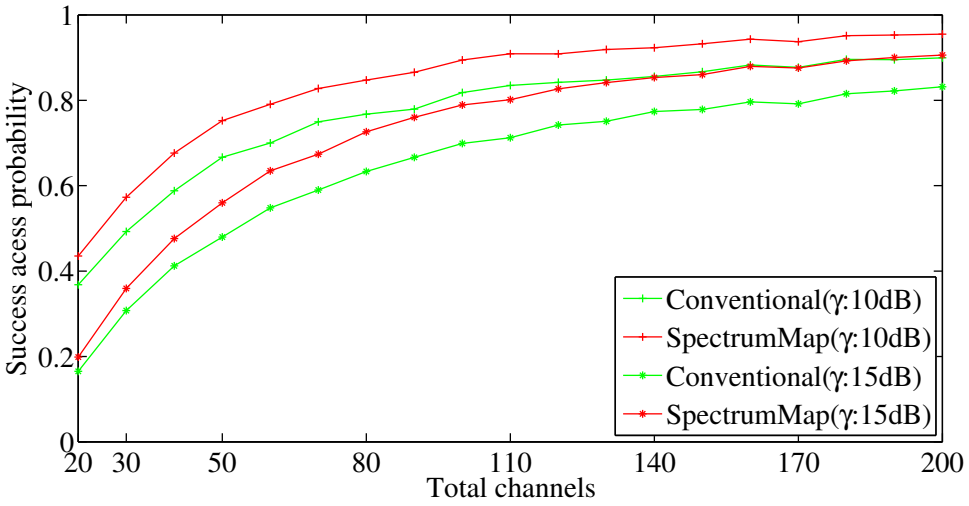

Figure 4.6: Assisted by coalition $S1$ ($D_S = 0.85$) (1 to 5 fixed)).

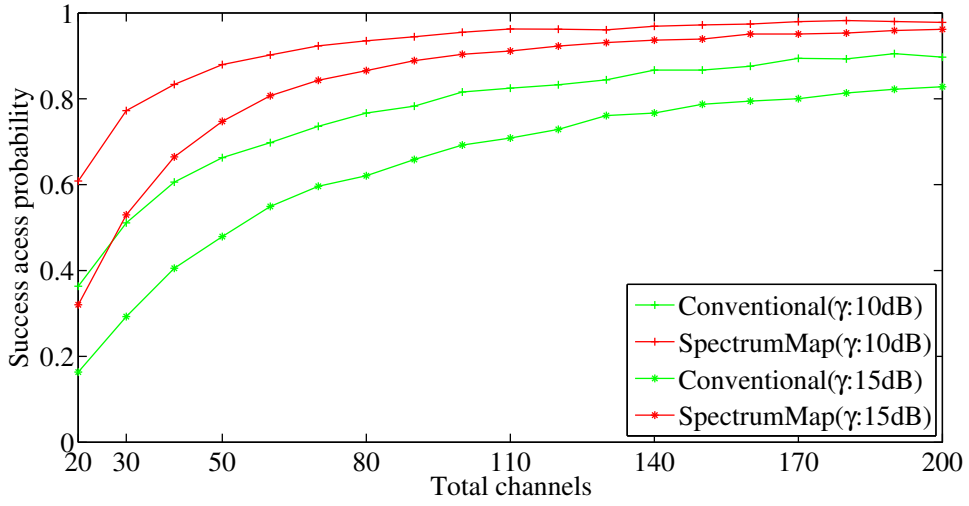

Figure 4.7: Assisted by coalition $S1$ ($D_S = 0.85$) (1 to 10 fixed)).

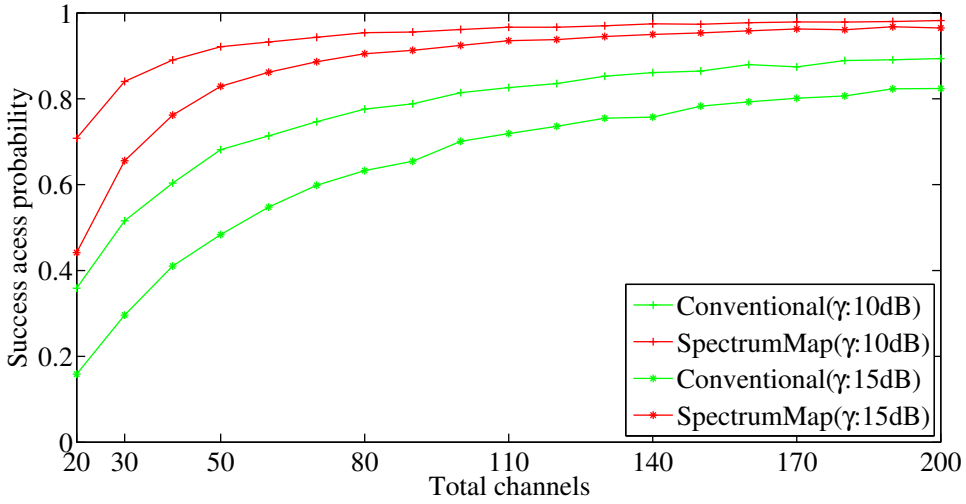

Figure 4.8: Assisted by coalition $S1$ ($D_S = 0.85$) (1 to 20 fixed))

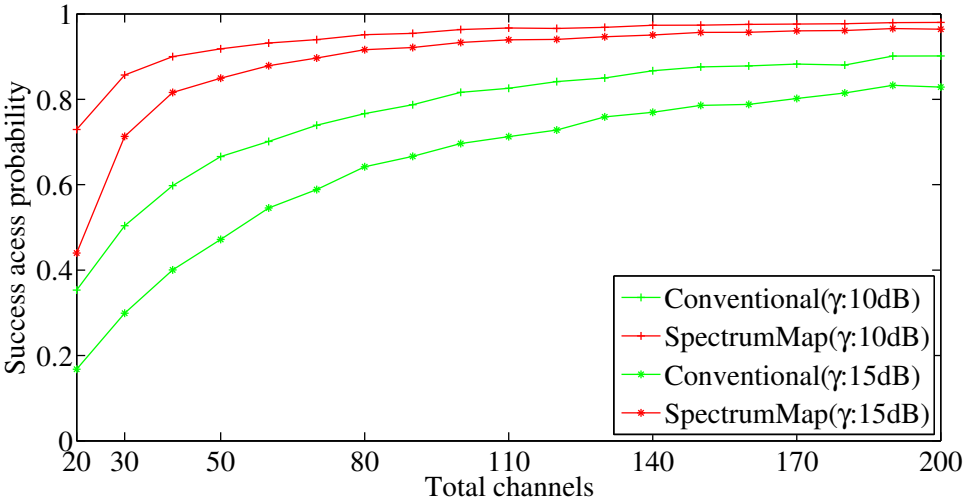

Figure 4.9: Assisted by coalition $S1$ ($D_S = 0.85$) (1 to 5 random)).

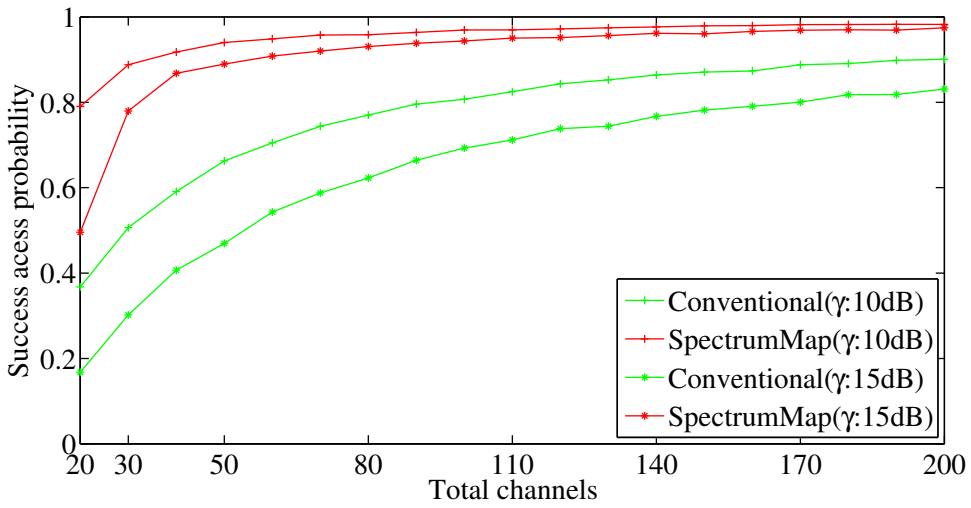

Figure 4.10: Assisted by coalition $S1$ ($D_S = 0.85$) (1 to 10 random)).

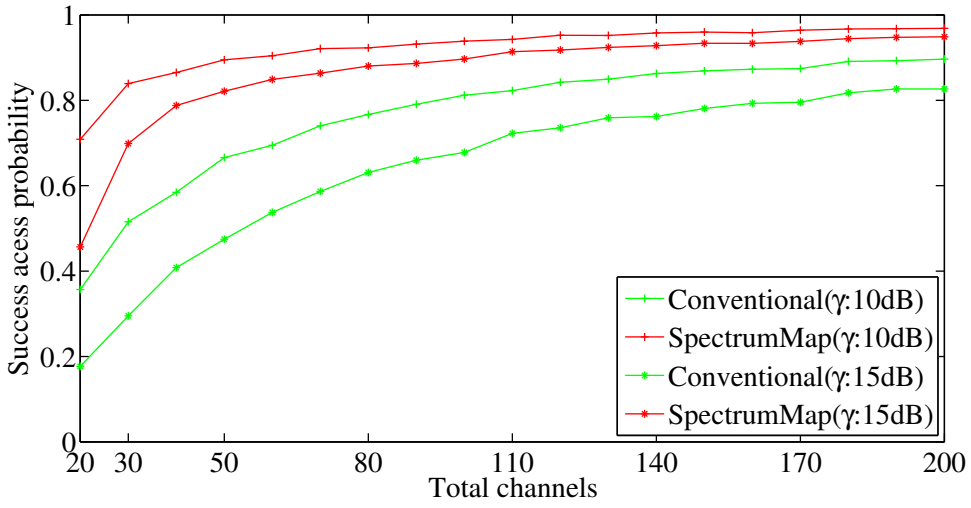

Figure 4.11: Assisted by coalition $S4$ ($D_S = 0.7$) (1 to 10 random))

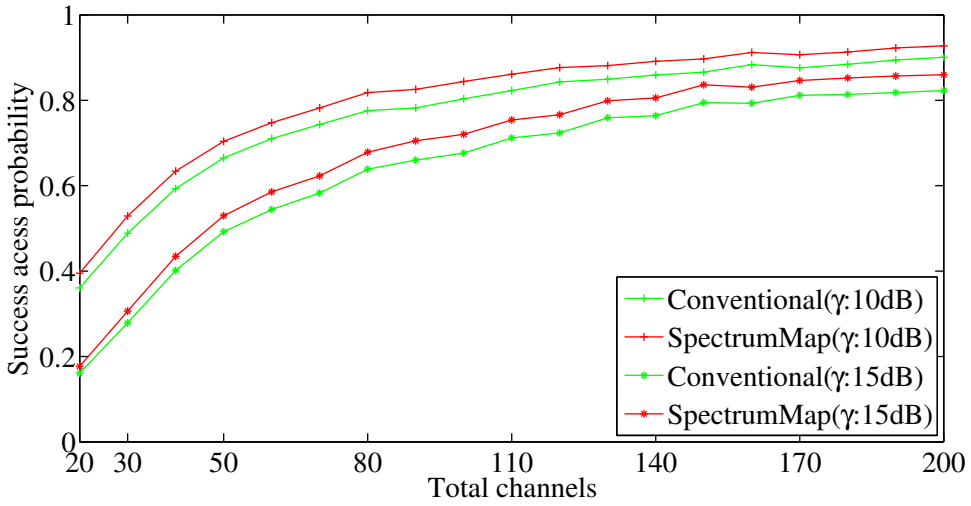

Figure 4.12: Assisted by coalition $S6$ ($D_S = 0.4$) (1 to 5 fixed)).

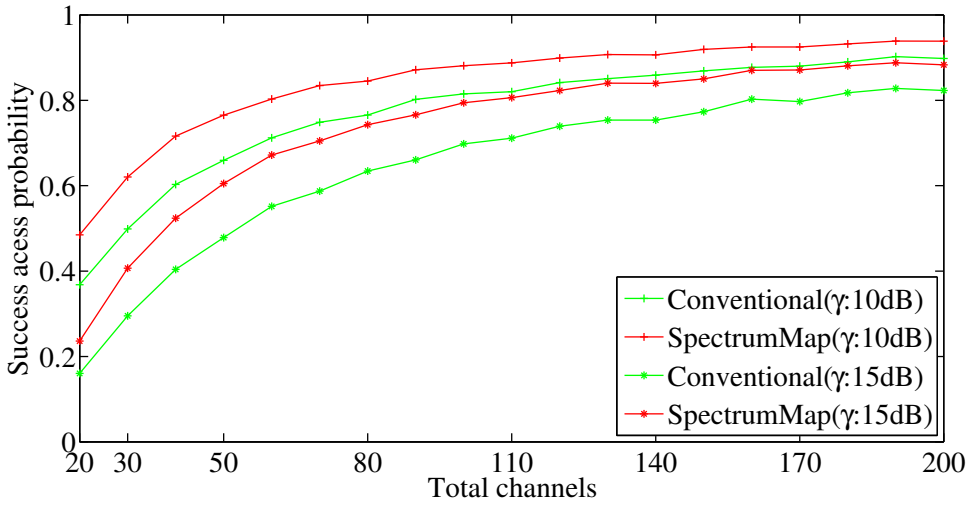

Figure 4.13: Assisted by coalition $S6$ ($D_S = 0.4$) (1 to 10 fixed)).

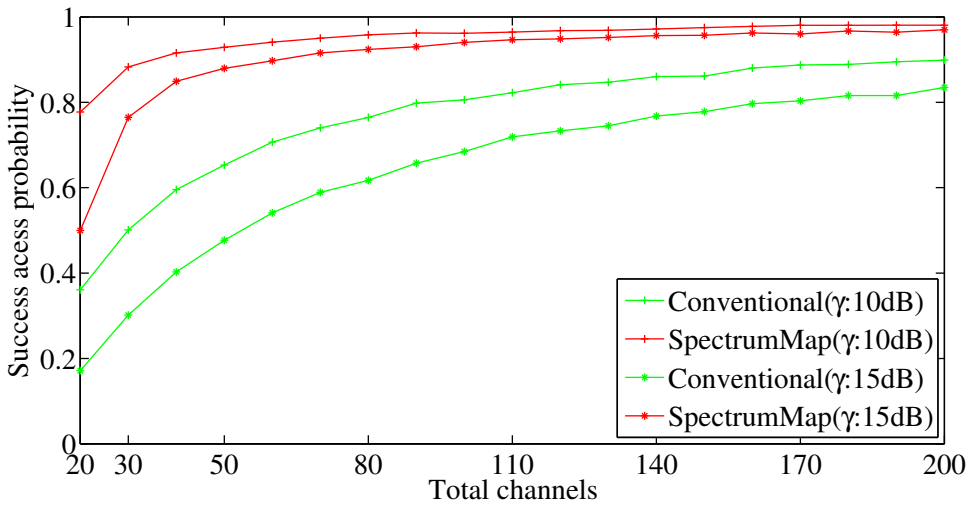

Figure 4.14: Assisted by coalition $S5$ ($D_S = 0.83$) (1 to 10 random))

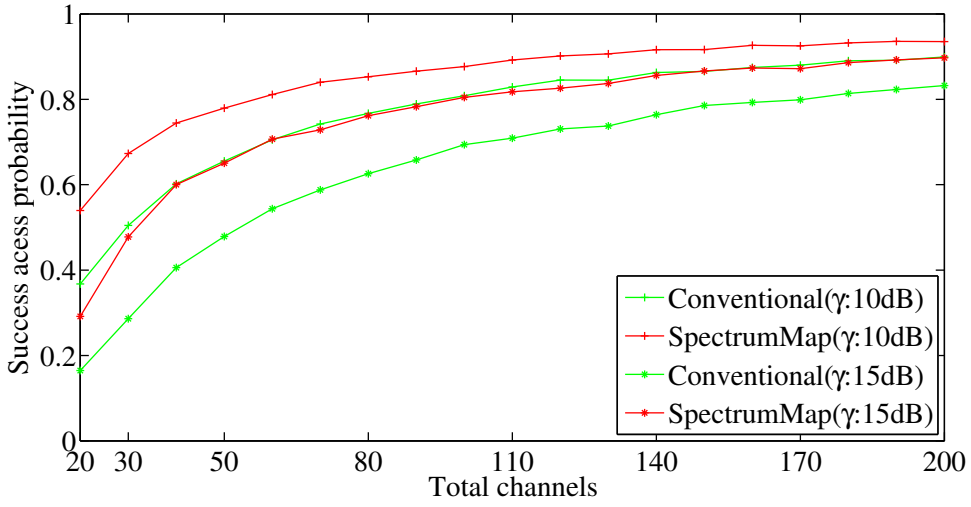

Figure 4.15: Assisted by coalition $S6$ ($D_S = 0.4$) (1 to 5 random)).

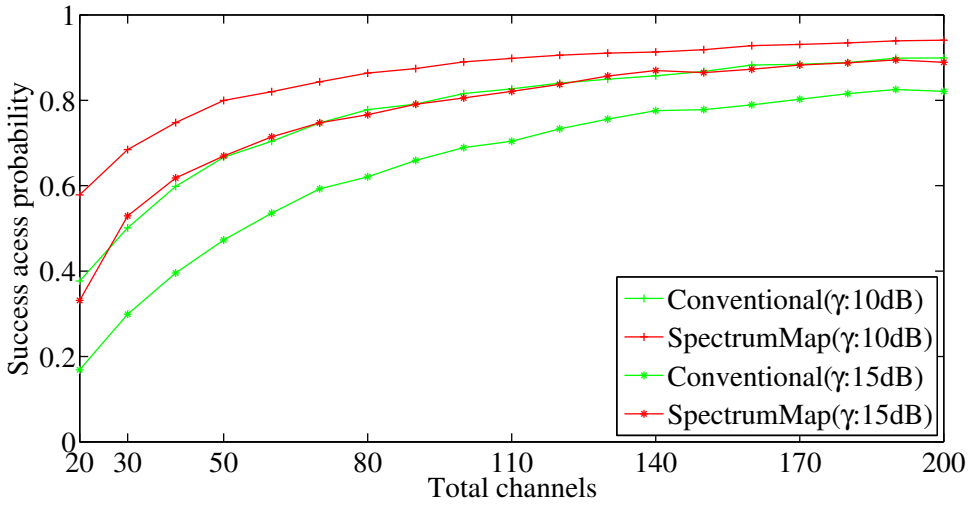

Figure 4.16: Assisted by coalition $S6$ ($D_S = 0.4$) (1 to 10 random)).

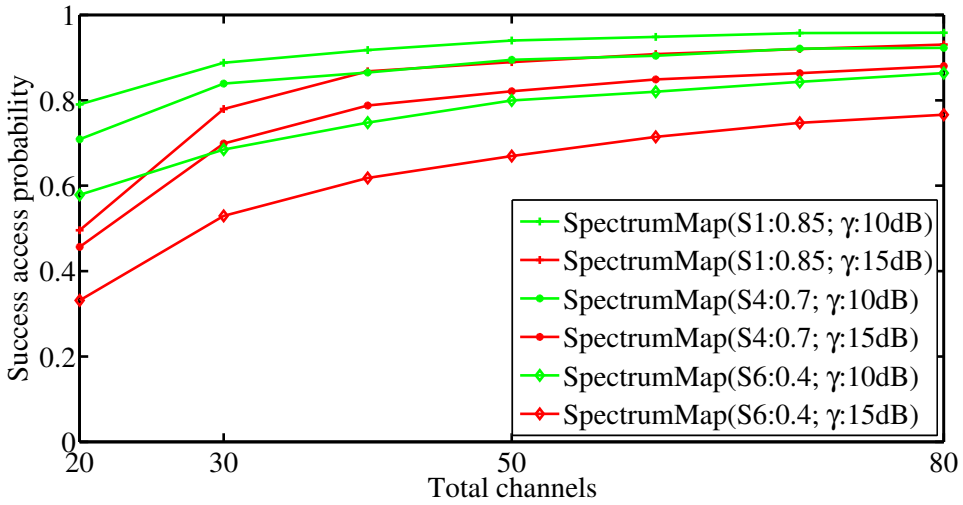

Figure 4.17: Comparison with different participation capability.

ing signals, including all the downlink data traffics between cognitive base stations and their users, as well as that between small cell base stations and small cell users. The formed stable coalition of Fig. 4.3 is used to show the performance improvement by utilizing the user contributed spectrum datasets.

Considering the intra-tier interference (between small cell and small cell) and cross-tier interference (between cognitive cell and small cell), SINR at a reference receiver of small cell can be defined as

$$SINR_{sc,su}^{f,c} = \frac{P_{sc,su}^{f,c} G_{sc,su}^{f,c}}{\sum_{1 \le l \le F, l \ne f} P_{sc,su}^{l,c} G_{sc,su}^{l,c} + \sum_{1 \le m \le M} P_{mbs,fu}^{m,c} G_{mbs,fu}^{m,c} + \sigma^2}. \tag{4.10}$$

In the above, $SINR_{sc,su}^{f,c}$ indicates the achieved SINR at the reference small cell user (su) in a target small cell (sc) f over the subchannel c, and the transmit power for the base station of small cell f is denoted by $P_{sc,su}^{f,c}$; $P_{sc,su}^{l,c}$ and $P_{mbs,su}^{m,c}$ indicates the interference power from the small cell l and cognitive cell m that have occupied the same subchannel c; denote by $G_{sc,su}^{f,c}$, $G_{sc,su}^{l,c}$ and $G_{sc,su}^{m,c}$ the channel gain, respectively; total number of small cell and cognitive cell is represented by F and M, respectively; σ^2 indicates the variance of Additive White Gaussian Noise. In addition, the path loss is represented by $G = h||d||^{-\beta}$, where h denotes fading coefficient, $||d||$ and β indicates the Euclidean distance and path loss exponent, respectively. Accordingly, the success probability of downlink channel access in a target small can be defined as follows:

$$P_{sc} = Prob\{SINR_{sc,su}^{f,c} > \gamma\}, \tag{4.11}$$

where γ is threshold value of QoS (quality of service).

Based on Fig. 4.3, it can be assumed that each coalition performs spectrum measurement at a corresponding time point. In a small cell, at a time portion, there is one coalition which has been coordinated by the base station to collect data for some target channels. Within each coalition, there exists a reference receiver which lies R_f away from the center of small cell. SpectrumMap will analyze the collected channel information from each coalition and recommend the best candidate channels to the reference receiver. The used parameter is given in Table 5.1. The total number of W channels is varied from 20 to 200, which are fully shared by cognitive cell users as well as small cell users. It is assumed that at any time slot, there is only one downlink data traffic for each cognitive cell as well as each small cell. Without coordination, each small cell can choose k channels from the total number of W channels ($k < W$), and different small cell can share the same range of channels. Within each small cell, 4 channels will be further selected from the k channels and allocated to small cell users. In the first kind of experiment, the

target k channels for each coalition to observe is fixed; while in the second kind of experiment, the k channels is not fixed and each coalition can randomly choose k channels from total channels to observe. The result is given by Fig. 4.6 ~ Fig. 4.17, which can be summarized as follows: (1) SpectrumMap assisted spectrum access performance outperforms the conventional one which lacks of the support from user's contributed dataset, and this can be verified by all figures; (2) Increasing observed target channels leads to better performance. This can be confirmed by Fig. 4.6 ~ Fig. 4.8, where target channels is fixed and the number is 5, 10, 20, respectively. Also, by Fig. 4.9 ~ Fig. 4.10, the same result can be obtained, as the target channels is randomly selected with the number of 5 and 10, respectively. The reason behind this is that observing more target channels can increase the chance of finding good channels; (3) Higher spectrum data participation capability achieves better performance than lower spectrum data participation capability, i.e., compare Fig. 4.6 and Fig. 4.12, Fig. 4.7 and Fig. 4.13, respectively. Fig. 4.17 further summarizes the cases which choose target channels randomly. The reason is clear, as the more data contributed from users, the better understanding to channel environment; (4) Choose the observed target channels randomly from total channels yields better performance that that by fixed, i.e., comparing Fig. 4.7 and Fig. 4.10, Fig. 4.13 and Fig. 4.16, respectively. The reason is similar to (2) that the collected spectrum data of randomly selected channels includes more information of the channels which has better SINR.

4.9 Summary

This work proposes a incentive model to encourage public users to contribute spectrum data to the SpectrumMap to improve channel access performance. The result and effectiveness has been confirmed by experiments. An algorithm for forming optimal coalition participations by users has been designed, such that more available spectrum data of each coalition can be yielded and user's payoff can be maximized.

Chapter 5

On the Cooperation Between Cognitive Radio and Femto Cell Networks for Cooperative Spectrum Sensing and Self-organization

This work proposes a single framework for cooperative spectrum sensing of cognitive radio network as well as for the self-organization of femtocells. It will construct a large-scale spectrum database through cooperative spectrum sensing for cognitive radio network, which exploits the existing infrastructure of the two-tier (macrocell-femtocell) network and encourages a small cellular network, i.e., femtocell, to act as the rendezvous point between individual local spectrum monitors and the fusion center of spectrum database. Distributed sensing results reported from many cognitive radio users can be integrated into spectrum database via these rendezvous points. Thus the problem of the dependency of conventional cooperative spectrum sensing on common control channel is improved. Meanwhile, the interaction between cognitive radio users and femtocells can offer an intriguing possibility to improve the self-organization of femtocells by absorbing radio environment maps constructed by cognitive radio users. Game theory can be applied to formulate the mutual benefits in the cooperation between cognitive radio user and femtocell. The numerical effects of the coalitions between cognitive radio users and femtocells are illustrated.

5.1 Introduction

In CSS (cooperative spectrum sensing), CCC (common control channel) is required to realize efficient signaling protocols to identify spectrum opportunity. It

is however difficult to construct CCC in large-scale CR (cognitive radio) networks due to dynamic environments of PUs (primary users) as well as frequency usage efficiency. Most works about CSS mainly rely on cluster-based, sequence-based and dedicated CCC schemes [75]. With these techniques, at the convergent time, CRs in centralized CSS tend to choose one CCC, and CRs in distributed CSS consume different CCCs for each available CR pair [76]. Femtocell is designed to benefit end-users' data rates as well as network operators' costs and capacities, but interference may occur between neighboring FUs (femtocell users) and MUs (macrocell users). Existing self-organization for femtocell includes dealing with opportunistic spectrum access [77], power management [78] and interference mitigation [79], etc. Some techniques depend on the cognitive capability of femtocell itself.

This work presents a single framework to solve CSS for CUs and also self-organization of femtocells by encouraging the cooperation between CU and femtocell. The cooperation is categorized into two parts, Femtocell-aided CSS and CU-aided self-organization, respectively.

By Femtocell-aided CSS, the advantage of the existing infrastructures of small cellular network, i.e., femtocell, can be taken to serve as rendezvous point between CUs and FC (fusion center). Compared with the conventional CSS which always discards the cooperation between CU and PU during CCC construction, reacting to the activity of PU and rebuilding data report channel could be better achieved if the available data report channel of CSS is coordinated with the assistance of femtocell. In addition, since data report channels are distributed into multiple CU-femtocell pairs, the problems of the conventional CSS that are susceptible to report channel saturation and security attacks could be improved.

Unlike the existing approaches, the designed CU-aided self-organization introduces the cognitive capabilities of multiple unsubscribing CUs to strengthen self-organization of femtocell such that femtocell can better grasp the surrounding radio environment and ensure higher throughput. The motivation aims at achieving mutual benefits for CR and femtocell networks. In this regard, CUs and femtocells can achieve desired gain by cooperation while overcome their own limitations. The framework can be further extended to build up large-scale REM to facilitate efficient spectrum utilizations. This work tracks the cooperation by a coalitional game approach in which CU and femtocell are players. The rest of the work is: Section II shows the system model. Section III proposes a distributed algorithm of coalition construction for the framework. Section IV gives simulation results and the work is concluded in Section V.

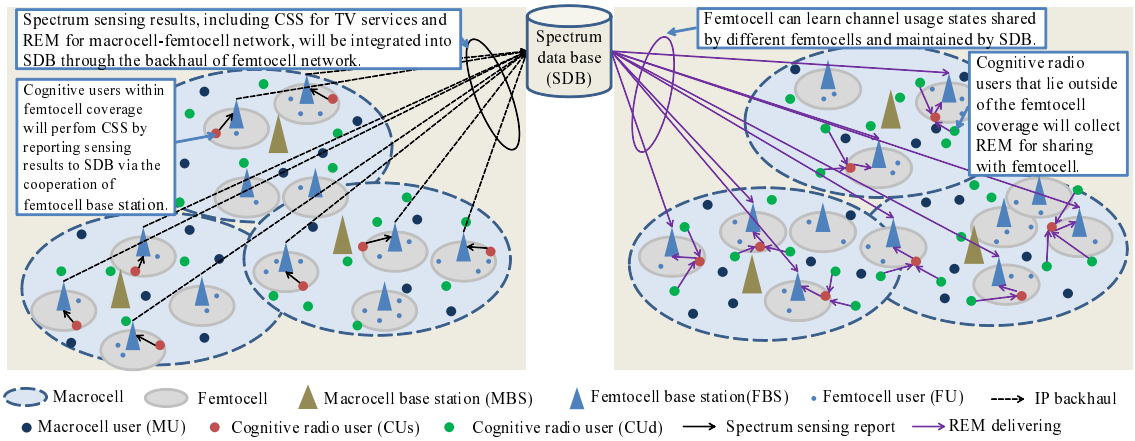

Figure 5.1: Femtocell-aided CSS (left) and CU-aided self-organization (right). CSS is jointly performed by multiple $CU_s s$ and the distributed sensing results are fused into the FC of SDB, lied at the gateway of femtocell network, through the relay of FBSs; Multiple FBSs share the same SDB. CU_d is responsible for collecting REM which will be shared with femtocell for self-organization. The SDB enables large-scale sharing of spectrum sensing and REM exchanging.

5.2 System Model

The designed model of Fig. 5.1 consists of MBSs (Macrocell Base Stations), MUs, FBSs (Femtocell Base Stations), FUs, and CUs. There are two kinds of CUs: CU_s joins CSS to detect white space of TV service, and CU_d collects REM of macrocell-femtocell network. MUs, FUs and $CU_s s$ share the same bandwidths. CU_s uses as many subchannels of femtocell, if they are available, for data fusion of CSS. The TV white space obtained from CSS can be shared among all CUs, i.e. $CU_s s$ and $CU_d s$. CU_d does not share the bandwidths with MUs and FUs, and only uses white space of TV channels for REM transmission and its data communication. The available frequency resources are assumed to be equally divided into orthogonal subchannels and the two tiers share the total spectrum. MBS is arranged according to PPP (poisson point process) and the intensity is λ_m. FBS is distributed by a PPP and the intensity is λ_f. Based on spectrum sensing, denoting ϵ_i as the threshold value of the energy detection for each CU i, the probabilities of detection and false alarm are [80]

$$P_{d,i} = Q((\frac{\epsilon_i}{\sigma^2} - \gamma_i - 1)\sqrt{\frac{N_{TW}}{2\gamma_i + 1}}), \tag{5.1}$$

$$P_{f,i} = Q((\frac{\epsilon_i}{\sigma^2} - 1)\sqrt{N_{TW}}), \tag{5.2}$$

where Q denotes the tail probability for the standard normal distribution, γ_i is received SNR (signal-to-noise ratio) of PU, σ^2 is the variance of AWGN (Additive White Gaussian Noise), and N_{TW} is time-bandwidth product. $P_{f,i}$ is evaluated by a target detection probability, i.e., $\bar{P}_d$, given as

$$P_{f,i} = Q(\sqrt{2\gamma_i + 1}Q^{-1}(\bar{P}_d) + \sqrt{N_{TW}}\gamma_i). \tag{5.3}$$

With CSS, the individual target detection probability is [81]

$$\bar{P}_{d,i}^{css} = 1 - (1 - \bar{P}_d)^{\gamma_i/(\gamma_1+...+\gamma_n)}, \tag{5.4}$$

where n is the number of total participators. SNR diversity of different CUs has been considered, and the obtained value is a weighted target probability of detection. Accordingly, the individual and cooperative false alarm probability are

$$P_{f,i}^{css} = Q(\sqrt{2\gamma_i + 1}Q^{-1}(\bar{P}_{d,i}^{css}) + \sqrt{N_{TW}}\gamma_i), \tag{5.5}$$

$$P_f^{\mathcal{S}} = 1 - \prod_{i=1}^{n}(1 - P_{f,i}^{css}). \tag{5.6}$$

A framework is established between CUs and femtocells shown in Fig. 5.2. The game players include FBSs and CUs. CU_d could decide to join the game with CU_s according to the utility function. If CU_d achieves more throughput by joining the game, it will collect REM, otherwise it will stay alone to perform single spectrum sensing.

5.3 Proposed Coalition Formation Algorithm

5.3.1 Coalitional Game Formation

Multiple players desire to form cooperative groups to strengthen individual gains. They can be formulated as coalitional game and each coalition denotes a single entity, in which agreement can be achieved between players.

Definition 1: Given the set of players $\mathcal{N}$, a coalition partition is expressed by $\Pi = \{\mathcal{S}_1, ..., \mathcal{S}_k\}$ which divides the players set $\mathcal{N}$. For any $l \in \{1, ..., k\}$, $\mathcal{S}_l \subseteq \mathcal{N}$ are disjoint coalitions with the constrain $\cup_{l=1}^{k} S_l = \mathcal{N}$.

Definition 2: Players are distinguished by their functions:

1) FBS: receives spectrum sensing data from nearby CU_s which lies inside its communication range. FBS wishes to cooperate with CU_s by relaying sensing

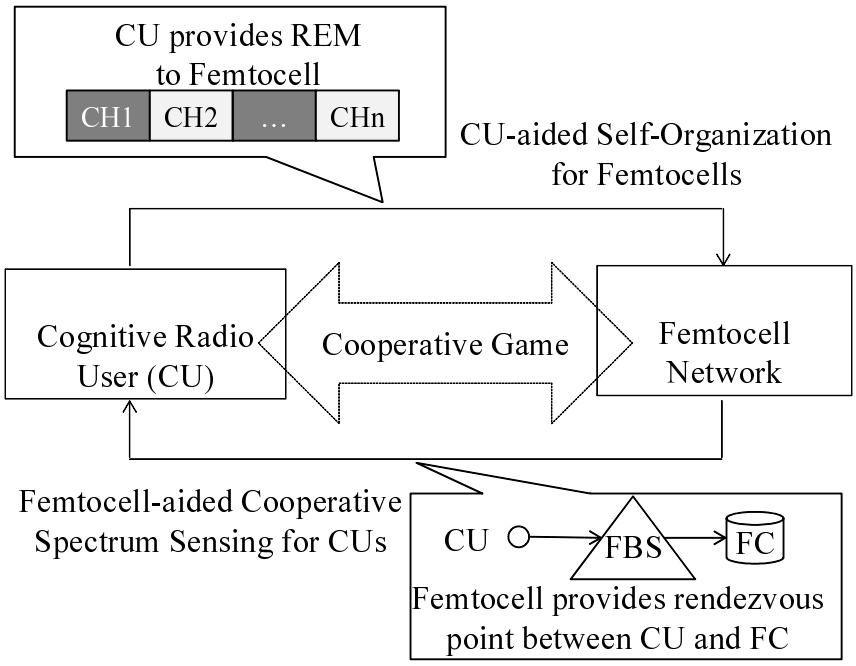

Fig. 5.2: Cooperation based framework between cognitive radio network and heterogeneous cell network.

data to SDB. 2) CU_s: reports spectrum sensing data to FBS. $CU_s(l) \in \mathcal{S}_l$ performs CSS periodically with the help of $FBS(l)$. 3) CU_d: located around CU_s. CU_d lies outside the communication range of FBS and performs spectrum sensing for REM collection.

Definition 3: Dividing $\mathcal{N}$ into three sets $\mathcal{N}_{FBS}$, $\mathcal{N}_{CU_s}$ and $\mathcal{N}_{CU_d}$, where $\mathcal{N}_{FBS} = \{FBS(1), ..., FBS(LF)\}$, $\mathcal{N}_{CU_s} = \{CU_s(1), ..., CU_s(M)\}$ and $\mathcal{N}_{CU_d} = \{CU_d(1), ..., CU_d(N)\}$. The designed one coalition is expressed by $\mathcal{S}_l = \{FBS(l)\} \cup \{CU_s(l)\} \cup \mathcal{N}^l_{CU_d}$, where $\mathcal{N}^l_{CU_d} \subseteq \mathcal{N}_{CU_d}$ assuming each $FBS(l)$ has only one $CU_s(l)$ for each coalition.

Definition 4: The coalitional game based on partition form with nontransferable utility is defined as pair $(\mathcal{N}, \mathcal{V})$. $\mathcal{N}$ is the set of players and $\mathcal{V}(\mathcal{S})$ is a mapping that contains payoff vectors that players in $\mathcal{S}$ can achieve where for every coalition $\mathcal{S} \subseteq \mathcal{N}$, $\mathcal{S}$ is a closed convex subset of R^s [82] [83].

For constructing optimal coalition structures, the CU_s, CU_d and FBS can autonomously choose to join or leave a coalition in terms of well defined individual preference relations, which are based on a coalition value, i.e., the defined utility functions Eqs. (5.11) and (5.20), by considering benefit and cost, such as effective throughput as well as power cost. The coalition value in the proposed cooperative game is defined by

$$\begin{aligned}\mathcal{V}(\mathcal{S}_l) &= V(\{FBS(l)\} \cup \{CU_s(l)\} \cup \mathcal{N}^l_{CU_d}) \\ &= \{v_{FBS(l)}, v_{CU_s(l)}, v_{CU^l_d(1)}, ..., v_{CU^l_d(|\mathcal{S}_l|-2)}\},\end{aligned} \tag{5.7}$$

where l is a coalition index, $v_{FBS}(l)$ is denoted by Eq. (5.20), and $v_{CU_s}(l)$, $v_{CU^l_d}(1)$, ..., $v_{CU^l_d(|\mathcal{S}_l|-2)}$ can be derived from Eq. (5.11).

Definition 5: For any player $CU_d(i)$, its preferences is denoted by an order $\succeq_i$ (a complete, reflexive, and transitive binary relation) over the set (S_l). That is

$$S_1 \succeq_i S_2 \Leftrightarrow v_{CU^1_d(i)} \geq v_{CU^2_d(i)}, \tag{5.8}$$

where $v_{CU^1_d(i)}$ and $v_{CU^2_d(i)}$ are utilities that $CU_d(i)$ received in the coalitions S_1 and S_2, respectively. The definition implies that $CU_d(i)$ prefers to join S_1 rather than S_2 [82]. The main rationale behind the above is that CU and Femtocoell can make distributed decisions to join any coalition they prefer. Algorithms 1, 2 and 3 have been proposed to enable CSS for CU and self-organization for femtocell.

5.3.2 Utility Definition

Utility of CU_s ***and*** CU_d

When the ith CU, $CU(i)$, is acting in non-cooperative manner, the average throughput is [85]

$$R_{CU,H_0}(i) = P_0 \frac{T - T_s}{T}(1 - P_{f,i})R_0(i), \tag{5.9}$$

where H_0 means PU is absent, P_0 is the probability of H_0, T represents the sensing interval and T_s is sensing time, R^i_0 represents the data rate under H_0.

Different $CU_s s$ will participate in CSS assisted by femtocells. CU_d will not always join the cooperation with CU_s, but if it joins the cooperation, it can obtain the gain of white space decision from Femtocell-aided CSS. By considering the cost including sensing time delay, REM transmission and spectrum data fusion, the throughput $R^S_{CU,H_0}(i)$ of $CU(i)$, assuming the data fusion returns the result at every sensing interval T, can be expressed by

$$R^S_{CU,H_0}(i) = \begin{cases} P_0 \frac{T - T^{REM}_s - n^S * T_d}{T}(1 - P^S_f)R_0(i), \\ \qquad\qquad \text{if } CU(i) \in \mathcal{N}_{CU_d}; \\ P_0 \frac{T - T^{CSS}_s - \tau_c * M - n^S * T_d}{T}(1 - P^S_f)R_0(i), \\ \qquad\qquad \text{if } CU(i) \in \mathcal{N}_{CU_s}, \end{cases} \tag{5.10}$$

where T^{REM}_s is time cost for individual CU_d to capture local REM. Note that the task of CU_d is to collect REM for the macrocell-femtocell network, while that of

CU_s is to detect PU signals through CSS, i.e., the TV services. T_d is the time cost for a CU_d to transmit REM to CU_s. n^S is the number of CU_ds within one coalition, thus $n^S * T_d$ denotes the cost for all links corresponded to a common CU_s. T_s^{CSS} is individual sensing time for CU_s to perform CSS. The overhead in combining sensing reports from different CU_s, executed at FC, is also taken into account, given as $\tau_c * M$ which is the required time for the data fusion at FC which is assumed to be proportional to M. Specifically, the probability of false alarm for each CU_d follows the value of CSS performed by CU_s, given by P_f^S. The cost that CU pays for femtocell in Femtocell-aided CSS increases with the number of CU_d within one coalition due to more REM needed to be shared with femtocells.

For benefiting from the data rate gains by forming coalition with femtocell, there is additional power penalty for CU due to REM exchange and sensing data report. In consequence, besides the attainable data rate, i.e., Eq. (5.10), power cost is also an attractive notion to evaluate the utility of a CU. Accordingly, the utility of $CU(i)$ in the proposed cooperative strategy is formulated as

$$v_{CU(i)} = R^S_{CU,H_0}(i) - \hat{P}^S_{CU}(i), \tag{5.11}$$

which has been assumed to be linear functions for those defined in [86].

The power consumption for $CU_d(i)$ in the formed coalition, $\hat{P}^S_{CU_d}(i)$, is approximated by [87]

$$\hat{P}^S_{CU_d}(i) = (\gamma^{th}_{CU} * \sigma^2_{CU})/G_{CU_d}(i), \tag{5.12}$$

where $\hat{P}^S_{CU_d}(i)$ should satisfy power constrain, $\hat{P}^S_{CU_d}(i) \leq P_{max}$ (the power consumption for CU_s is not considered), γ^{th}_{CU} is a defined target SNR for REM exchange, i.e. for delivering REM, σ^2_{CU}represents the variance of AWGN, $G_{CU_d}(i)$ is the path loss of $CU_d(i)$.

Utility of FBS In the considered scenario, intra-tier interference and cross-tier interference will occur. Suppose the available spectrum is organized in orthogonal frequency subchannels. With downlink transmission, the SINR at FU can be defined as Eqs. (5.13) and (5.14),

$$SINR^c_{FBS}(i) = \frac{P^c_{FBS}(i)G^c_{FBS}(i)}{x + \sigma^2} \tag{5.13}$$

$$x = \sum_{l=1,l\neq i}^{LF} P^c_{FBS}(l)G^c_{FBS}(l) + \sum_{m=1}^{LM} P^c_{MBS}(m)G^c_{MBS}(m) \tag{5.14}$$

where $SINR^c_{FBS}(i)$ denotes the received SINR at the FU served by FBS i over subchannel c with the transmit power $P^c_{FBS}(i)$; $P^c_{FBS}(l)$ and $P^c_{MBS}(m)$ represent the interference powers from the FBS l and MBS m that use the same subchannel c;

channel gain is denoted by $G^c_{FBS}(i)$, $G^c_{FBS}(l)$ and $G^c_{FBS}(m)$, respectively; total number of FBS and MBS is LF and LM; σ^2 represents the variance of AWGN. The number of all the subchannels randomly selected from all the available frequencies in the macrocell-femtocell networks is $C(1 \leq c \leq C)$, Channel gain function is given as $G = h\|d\|^{-\beta}$, where h is fading coefficient in a fixed value, $\|d\|$ is the the Euclidean distance between the target femtocell user and its serving FBS, interfering FBSs and MBSs, respectively, and β is the path loss exponent. The success probability of femtocell Tx-Rx link is given by

$$P_{FBS} = Prob\{SINR^c_{FBS}(i) > \gamma_{FU,th}\}, \tag{5.15}$$

where $\gamma_{FU,th}$ is QoS (quality of service) threshold of FU.

In the non-cooperative model, the self-organization of 3GPP is not assumed, and each femtocell is assumed to use subchannels randomly in non-cooperative strategy. The downlink throughput of a subchannel c is [88]

$$\varphi^c_{FBS,non} = p_c N_{FU} \lambda_c P_{FBS,non} log(1 + \gamma_{FU,th}), \tag{5.16}$$

where $P_{FBS,non}$ is channel access success probability defined by Eq. (5.15), N_{FU} is the number of FUs served by a FBS, λ_c is FBS intensity, p_c is subchannel access probability. Each CU_d could maintain a database to store REM datasets used for tracking surrounding radio environment. In the defined REM, a dataset format is shown as below:

$$REM^c_{CU}(i) = \{T, < Lat, Long >, CH, IL\}, \tag{5.17}$$

where time (T), location ($< Lat, Long >$), channel index (CH) and interference level (IL) is stored into local REM database of each CU_d, and would be further integrated into SDB through the cooperation between CU_s and femtocell. The IL for a given channel CH observed by $CU_d(i)$ is defined as aggregate interference value given as

$$IL^{CH} = \sum_{\Phi \in \{MBS \cup FBS\}} P^{\Phi,CH} G^{\Phi,CH}, \tag{5.18}$$

where $P^{\Phi,CH}$ is the transmit power of all $MBS\,s$ and $FBS\,s$ that serve in channel CH, $G^{\Phi,CH}$ is channel gain. Each CU_d chooses a random set of total channels of the two tiers to capture IL, and the target channels observed by each CU_d can be updated periodically. As a result, the combined IL from different $CU_d s$ within one coalition could cover the whole bandwidths from a long-term view. Furthermore, CU_d is more likely to roam around the femtocell, thus the integrated REM from distinct $CU_d s$ will form a rich database for understanding channel usage. Available throughput of femtocell under cooperative strategy is defined as

$$\varphi^c_{FBS,coop} = (1 - \eta) p_c N_{FU} \lambda_c P_{FBS,coop} log(1 + \gamma_{FU,th}), \tag{5.19}$$

where η is the channel access ratio by CU_s used for reporting sensing data, and the success channel access probability, $P_{FBS,coop}$ referred to Eq. (5.15), is supposed to be improved by REM database aided by CU. The utility for the ith FBS is defined as

$$v_{FBS(i)} = \varphi^{c}_{FBS,coop}. \tag{5.20}$$

Algorithm 1 Distributed Coalition Formation

Initial Network Structure :
The whole network contains multiple individual CU, FBS, FU, MBS and MU, no partition formation.
CoalitionFormationProcess :
CU performs local spectrum sensing (targeting bands include TV and macrocell-femtocell services).
Each CU, via measuring RSSI, explores the most nearby FBS which covers the CU in the communication range.
if (CU finds the potential FBS) **then**
  if (Eq. (5.11) is satisfied) **then**
    CU asks FBS for pairwise negotiation.
    if (FBS admits negotiation based on utility Eq. (5.20)) **then**
      1) CU $\rightarrow CU_s$.
      2) CU_s and FBS form an initial coalition S ($S \in \{CU_s \cup FBS\}$).
      3) Call Femtocell-aided CSS (Algorithm 2).
      loop
        4) CU_s asks surrounding $CU_d s$ to join S, and $CU_d s$ decide to join S if payoff determined by Eqs. (5.11) and (5.8) is larger than that by Eq. (5.9)(Information exchange uses white space channels).
        if (CU_d joins S) **then**
          ($S = S \cup \{CU_d\}$)
          Perform CU-aided Self-organization (Algorithm 3).
        end if
      end loop
    end if
  end if
end if

Algorithm 2 Femtocell-aided CSS (new CU_s)

Require: $\bar{P}_d$: target detection probability; { $\gamma_1, \gamma_2, ..., \gamma_{n-1}$ }: observed SINRs by existing $n-1$ CU_ss; { γ_n }: observed SINR by a new CU_s;
Ensure: $R^{\mathcal{S}}_{CU,H_0}$: available throughput for each CU ($CU \in \{CU_s, CU_d\}$)
CU_s reports local sensing result to the corresponding FBS within the same coalition.
FBS redirects the data from CU_s to the SDB (constructed at the gateway by femtocell networks).
The FC performs spectrum data fusion by Eq. (5.6).
The FC informs all FBSs to fetch the new fusion result.

1. Each FBS fetches the new fusion result (i.e., $P^{\mathcal{S}}_f$) to the corresponding CU_s.
2. When CU_d exists, CU_s informs the fusion result to the existing $CU_d s$ within the same coalition.
3. CU_s and $CU_d s$ obtain throughputs $R^{\mathcal{S}}_{CU,H_0}$ by Eq. (5.10).

return $R^{\mathcal{S}}_{CU,H_0}$

Algorithm 3 CU-aided Self-organization

Require: REM^c_{CU}: radio environment map observed by CU_s and CU_d (Eq. (5.17)) ($1 \leq c \leq C$);
Ensure: $\varphi^c_{FBS,coop}$: the available throughput of femtocell

1: Each CU_d performs local observation to get REM^c_{CU} (c is one of the subchannels used by the macrocell-femtocell network).
2: CU_s collects REM^1_{CU}, ..., REM^C_{CU} from multiple $CU_d s$.
3: CU_s shares REM^1_{CU}, ..., REM^C_{CU} with FBS within the same coalition.
4: FBS performs self-organization by assigning each FU the channels with minimum IL in terms of REM , and calculates achievable throughput via Eq. (5.19).
5: FBS records REM^1_{CU}, ..., REM^C_{CU} into SDB for information sharing with other $FBS\,s$ which are in different coalitions.
6: **return** $\varphi^c_{FBS,coop}$

5.4 Simulation Results on the Coalition Effects

By MATLAB with the parameters in Table 5.1, a preliminary simulation is made to evaluate the coalition effects.

A. Performance of Femtocell-aided CSS It is assumed that: 1) each CU_s acts as sensing node and reports its local observation to FC through FBS; 2) FC coordinates CU_ss to report sensing information; 3) FC informs the sensing result to each CU_s and CU_d through FBS. The simulation is done according to: (1) within a square area of 200 m x 200 m, 4 FBSs are created first; within each FBS, a CU_s is randomly deployed; then, 15 CU_ds are created outside the FBSs and no distance limitations between CU_d and CU_s; all the positions of FBSs, CU_ss and CU_ds are in random distributions; (2) all CU_ds should join the coalition which can increase its utility according to Eq. (5.8). A partition example is shown as Fig. 5.3 which has 4 coalitions, each coalition contains one FBS ($\triangle$), one CU_s ($+$) and multiple $CU_d s$ ($\odot$). The number of CU_ss is assigned from 1 to 10, and each CU_s is assumed to have successfully found a FBS which will relay the sensing data to the FC to enable CSS. In addition, it has been assumed that for each CU_s, there are multiple CU_ds that will form a coalition with the CU_s. Each CU_s or CU_d is assigned the same value of γ_i, i.e., -15dB, and their throughputs are obtained by Eqs. (5.9) (for Non-cooperative spectrum sensing) and (5.10) (for CSS), respectively. All the results are based on the average value of 50,000 rounds of calculations. The time cost for delivering REM message by each CU_d and combining sensing results at the FC for each report are set as 0.001 ms, Due to the decreased false alarm probability, CU_ss and CU_ds can achieve higher throughputs after forming coalitions. In the case of large time cost required to integrate a report, i.e, τ_c , CU_s will bear more time cost than CU_d, which may benefit CU_d more throughput gain (see Eq. (5.10)). Fig. 5.4 plots throughput gain of CU_s and CU_d in terms of varied time cost normalized by the case of non-cooperative approach, with random number of CU_d within each coalition ($1 < n^{\mathcal{S}} < 10$). Note that the coalition number of 1 for all the cases belongs to non-cooperative approach actually by reason that the number of CSS participators, i.e., CU_s, is 1.

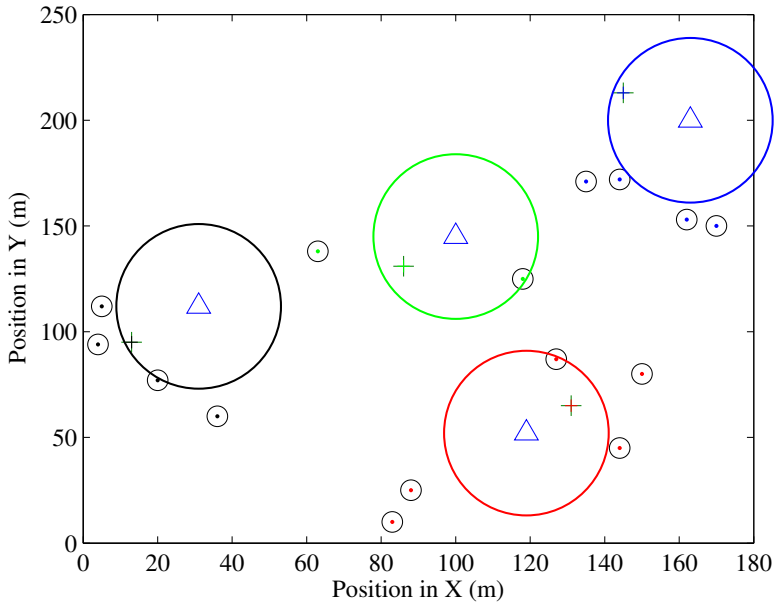

Fig. 5.3: A example scenario of coalition formation

Thus, if CU_d collects REM, the decreased throughput gain will be observed due to extra REM delivering cost, which indicates that CU_d fells no interest to join the cooperation whenever the total coalitions is 1. It can be observed that the time cost is an influence on obtainable throughput gain, i.e., with decreased time cost, better performance can be acquired. Specifically, when the required time cost is large, such as 1 ms, CU_s will lose interest in the cooperation when the coalition number goes beyond 7. Fig. 5.5 depicts the obtained throughput gain of CU_s and CU_d with a series of values of the coalition size (n^S). The time costs for exchanging REM and combining sensing reports are randomly assigned from 0 to 0.3 ms. The reason for the decreased throughput gain for coalition number of 1 is the same as the previous analysis. With the increased number of CU_d, due to the fact that CU has to suspend its data transmission and keep silent until the end of the REM transmissions by all the other CUs within the same coalition, it can be observed that the throughput gain is decreased with the increase of coalition size.

B. Performance of CU-aided Self-organization The macrocell-femtocell network in a 10,000 m x 10,000 m square area is assumed, and the positions of all MBSs and FBSs follow PPP. The scenario is evaluated according to: (1) there is a CU_s at the position of the reference receiver, which is in the coordinate of $(0, 0)$; (2) the CU_s will monitor a series of REM information; (3) these REM information

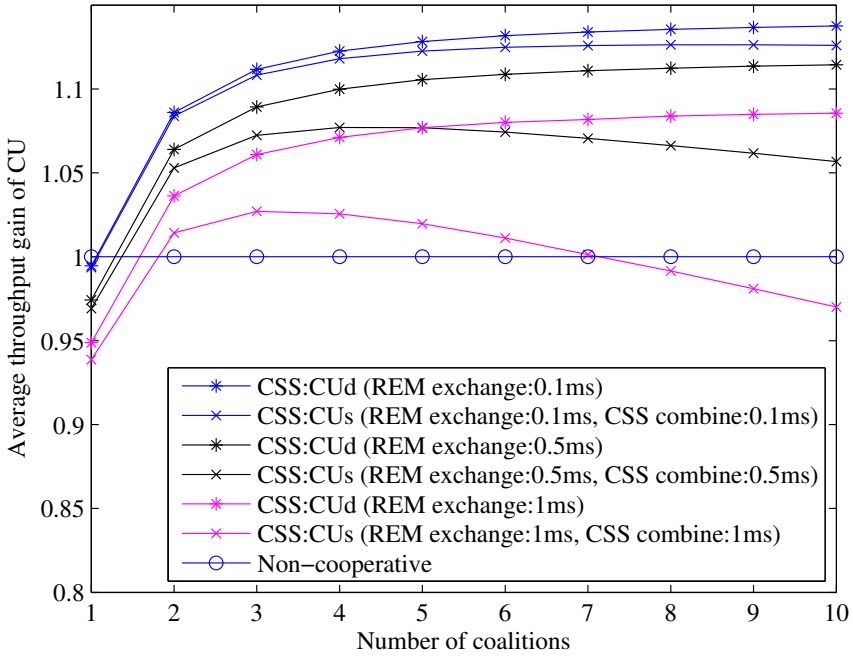

Fig. 5.4: Throughput gain achieved for CU_s and CU_d (with varying time cost)

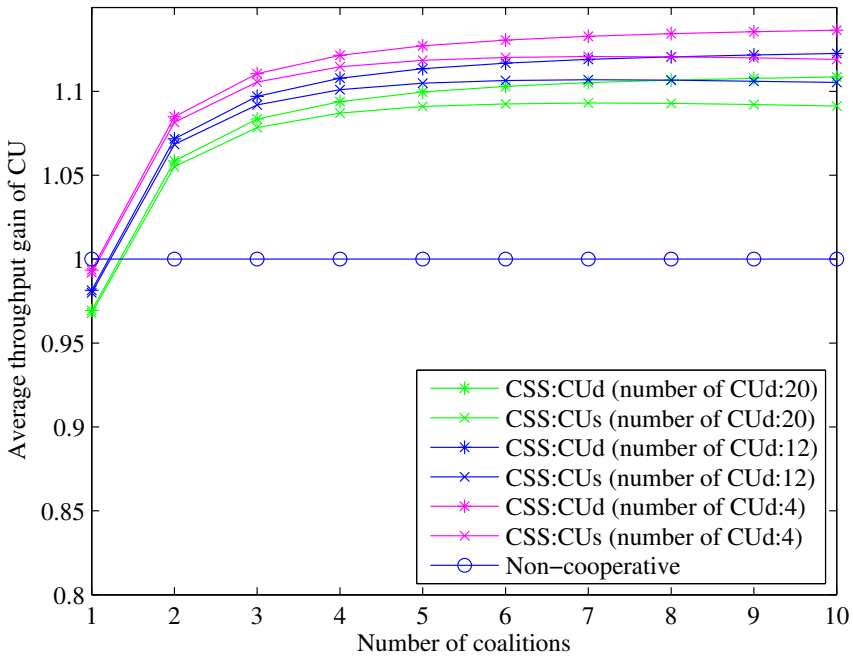

Figure 5.5: Throughput gain achieved for CU_s and CU_d (with varying coalition size)

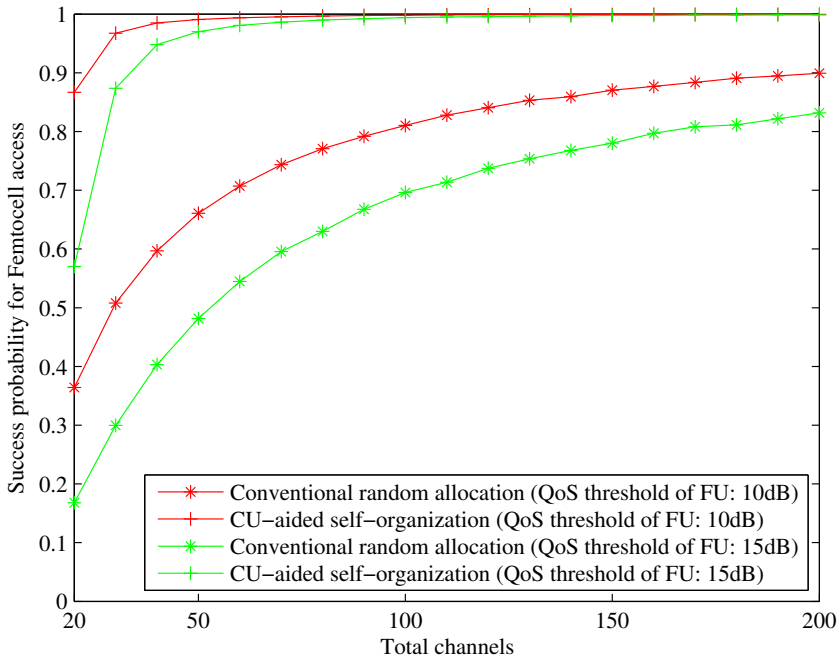

Figure 5.6: Femtocell success channel access probability

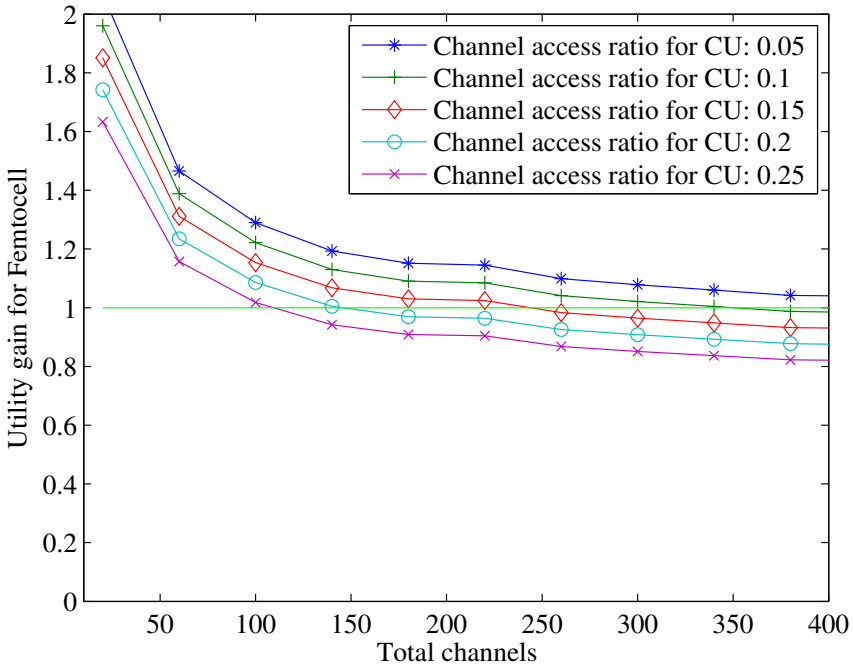

Figure 5.7: Femtocell utility gain

Table 5.1: System Parameters used for Evaluating CSS and Self-organization

Sensing time(T_s)	1ms	femtocell radius	20,40m
Time-bandwidth product(N_{TW})	6000	Number of FUs for one FBS (N_{FU})	4
Frame duration(T)	100ms	Indoor path loss exponent(β)	3.7
Target detection probability($\bar{P}_d$)	0.9	Outdoor path loss exponent(β)	4
QoS threshold of FU($\gamma_{FU,th}$)	10,15dB	Transmit power(MBS)	43dBm
received SNR of PU(γ_i)	-15dB	Transmit power(FBS)	23dBm
SNR target for REM exchange (γ_{CU}^{th})	10dB	FBSs intensity(λ_f)	$0.00008m^{-2}$
Noise level(σ^2)	-90dBm	MBSs intensity(λ_m)	$0.000001m^{-2}$
PU absent probability(P_0)	0.8		

will be shared with the femtocell at the coordinate of $(0, 0)$; (4) the femtocell will obtain success channel access probabilities and achievable utility gain (defined as $\varphi^c_{FBS,coop}/\varphi^c_{FBS,non}$) according to these REM information. The REM information at the position of CU_s is applied to characterize the performance of femtocell approximately. 3GPP is standardizing SON with the similar functions [89], and the conducted simulation scenario can be extended to improve the performance of 3GPP macrocell networks. In Fig. 5.6, the success channel access probabilities of femtocells are compared between CU-aided self-organization and random channel allocation, with varying total available channels. In the following discussion, random channel allocation is referred as the conventional method. Each femtocell serves 4 FUs and the number of assigned channels for one femotocell is set to 10. The outcome is an average value based on $50,000$ rounds of calculations. Since more channels can reduce the chance that MUs and FUs occupy the same frequency resources, it has been observed that the success probability is generally increased with the improvement of the number of total available channels. The success probability is effected by the QoS value of FU, i.e. the $\gamma_{FU,th}$. On the other hand, it can be seen that CU-aided self-organization outperforms the conventional method. As an example that $\gamma_{FU,th}$ is 15 dB, CU-aided self-organization achieves 0.9 of successful probability when the total channels is 35, while conventional method obtains around 0.3. To keep the success channel access probability larger than 0.8, the proposed approach could result in no more than 30 of total channels jointly used by the two-tier network, while the conventional method requires more available channel resources, i.e., about 170, to be jointly assigned to the two-tier network. The reason is clear because the femtocell with CU-aided self-organization can avoid using significant interference channels in its vicinity and allocates the orthogonal channels to FUs in terms of REM provided by CU, which outperforms the conventional method that has no radio environment knowledge. Fig. 5.7 shows the utility gain with CU-aided self-organization. The results

enable femtocell to judge the obtained utility level in terms of lost channel opportunity due to cooperation with CU. When total channel is less than 100, utility gains for all listed channel access ratio to relay sensing data of CU can enable femtocell to form coalition with CU, because the cooperation improves the payoff of femtocell since the utility gain, $\varphi^c_{FBS,coop}/\varphi^c_{FBS,non}$, is above 1. Given total channels, i.e., 300, femtocell has to grant CU no more than 10% of channel access time in order to avoid unwillingness of utility lost.

5.5 Summary

This work extends the cooperative spectrum sensing problem to heterogeneous cell network and proposes a cooperation framework to achieve mutual benefits for both cognitive radio and cellular users. The results has shown that both cognitive radio network and small cell network have incentive to cooperate with each other. The work uses the probability functions to define performance gain to show the convergence of the proposed algorithms, and evaluates the performance by coalitional game based approach. The results indicate that, to achieve the same throughput performance for Femto cell, the proposed approach can reduce channel resources from 160 to 30 channels with achieved success channel access probability of 0.8, and cognitive users can obtain throughput gain about 10%.

Chapter 6

Improve Spectrum Efficiency for General MIMO Relaying Broadcast Channel with Imperfect Channel State Information

Effective communication strategies with a properly designed source precoding matrix (PM) and a properly designed relay beamforming matrix (BM) can significantly improve the spectral efficiency of multiple-input multiple-output (MIMO) relaying broadcast channels (RBCs). In the present paper, this work first proposes a general communication scheme with non-regenerative relay that can overcome the half-duplex relay constraint of the general MIMO-RBC. Based on the proposed scheme, the robust source PM and relay BM are designed for imperfect channel state information at the transmitter (CSIT). In contrast to the conventional non-regenerative relaying communication scheme for the MIMO-RBC, in the proposed scheme, the source can send information continuously to the relay and users during two phases. Furthermore, in conjunction with the advanced precoding strategy, the proposed scheme can achieve a full-degree-of-freedom (DoF) MIMO-RBC with that each entry in the related channel matrix is considered to an i.i.d. complex Gaussian variable. The robust source PM and relay BM designs were investigated based on both throughput and fairness criteria with imperfect CSIT. However, solving the problems associated with throughput and fairness criteria for the robust source PM and relay BM designs is computationally intractable because these criteria are non-linear and non-convex. In order to address these difficulties, this work first sets up equivalent optimization problems based on a tight lower bound of the achievable rate. It then decomposes the equivalent throughput problem into several decoupled subproblems with tractable solutions. Finally, it obtains the suboptimal solution for the throughput problem by an alternating op-

timization approach. The fairness problem is solved by introducing an adjusted algorithm according to the throughput problem. Finally, it demonstrates that, in both cases of throughput and fairness criteria, the proposed relaying communication scheme with precoding algorithms outperforms existing methods.

6.1 Introduction

Wireless cooperative communication can improve network capacity and service coverage and reduce energy consumption by using relays [91]. However, the half-duplex constraint (HDC), i.e., the inability of relay nodes to transmit and receive signals simultaneously at the same frequency, is a major potential weakness because system bandwidth resources are used inefficiently due to the necessity of extra dedicated bandwidth for relay retransmissions [92]. In order to overcome the HDC, recent studies [93–95] have focused on the full-duplex relaying approach, which assumes that the relay node can transmit and receive signals simultaneously at the same frequency. However, for a relay to receive and transmit on the same frequency band at the same time is generally recognized as impossible because doing so will result in strong self-interference from the transmitter to the receiver of the relay node [95]. Moreover, most studies [93–95] considered only the three-terminal model (i.e., one source, one relay, and one destination). The present study focuses on a transparent half-duplex non-generative relaying scheme for multiple-input multiple-output (MIMO) relaying broadcast channels (RBCs), where two communication phases are required for one transmission due to the relay's HDC. In the first phase, the source transmits the signal to both the relay and the destinations, and in the second phase, the relay forwards the obtained signal to the destinations, while the source also simultaneously transmits the signal to the destinations. Note that the protocol is inefficient if the source remains silent during the second phase.

Existing studies have investigated the multiple-input multiple-output (MIMO) non-regenerative relaying broadcast channel to design effective source precoding and relay processing matrices. In [96], the authors examined the scenario of MIMO fixed relays by applying linear processing to improve link capacity in a cellular system. In addition, an implementable architecture integrating Tomlinson-Harashima precoding and adaptive modulation, which can adjust transmit streams adaptively, was proposed. In [97], considering the quality-of-service constraint, the authors jointly investigated linear beamforming and power saving for wireless cellular networks with multi-antenna relays. In [98], the authors studied the joint optimization of precoding design for source and relay to improve system capacity under the constraint of transmit power, and the formulated non-convex problem is solved by quadratic programming approaches. Moreover, in [99], the capacity

optimization for an amplify-and-forward relay network was considered, and they demonstrated that the duality relationships hold when a node is equipped with multiple antennas. However, the above-mentioned studies ignored the receiver's direct links (DLs) and each receiver was assumed to be a single antenna. In [100], a precoding scheme was proposed to optimize power allocation under the constraint of quality-of-service and transmit power budgets by applying a joint zero forcing (ZF) strategy and the DLs' contribution of receivers. However, this study considered only the ZF scheme with single-antenna receivers. In [101] and [102], the authors proposed a two-phase relaying model for a MIMO relaying broadcast network that considers the DLs' contribution of receivers when the source transmits a signal to both the relay and destinations. However, their studies only considered a source to be inactive during the retransmission phase of the relay, which causes the loss of half the degrees of freedom (DoFs) in the two transmission phases. In [103], the authors considered the source to be active during the retransmission phase of the relay, but, considered only the ZF strategy with a regenerative relaying scheme and ignored the receiver's direct links. In [104–107], in the presence of imperfect CSI, robust precoding algorithms were developed in order to deal with CSI quantization or estimation errors, but only an inefficient protocol in which the source remained silent during the second phase was considered. In practice, DLs can provide valuable spatial diversity to the MIMO relay system and should not be ignored. In addition, the source can be active during relay retransmissions in order to overcome the relay's HDC and achieve full DoF.

In the present study, in order to overcome the relay's HDC by allowing the source to be active during the second phase, this work considers a general MIMO-RBC with two types of users, i.e., direct users and relay users. The direct user is defined as a user having better DL (*source-destination*) gain than relay link (RL) (*source-relay-destination*) gain, whereas the relay user is defined as a user having better RL gain than DL gain. In the present study, this work focuses on a general design strategy to deal with the precoding matrix (PM) for the source and the beamforming matrix (BM) for the relay to serve direct users and relay users simultaneously by allowing the source to be active during relay retransmissions within the proposed scheme, which can overcome the relay's HDC and achieve full DoF. Hence, it is no necessary to distinguish direct users from relay users in the remainder of the paper. To the best of the author's knowledge, the present study is the first analytical study on source and relay matrices designed for general MIMO-RBC with coordinated users (direct users and relay users) to overcome the relay's HDC and achieve full DoF during two phases. This work first introduces a theorem to prove that the proposed scheme can achieve the maximum DoF of a MIMO-RBC with full rank, and then consider two-system performance criteria for designing the source PM and relay BM: the summed throughput maximization and the minimum user's rate maximization subjected to transmitted power constraints at both

the source and relay with imperfect channel state information at the transmitter (CSIT). However, this problem is computationally intractable due to the fact that the throughput and fairness criteria for the robust source PM and relay BM designs are non-linear and non-convex [118]. In order to address these difficulties, this work first sets up the equivalent optimization problems based on a tight lower bound of the achievable rate. It then decomposes the equivalent throughput problem into several decoupled subproblems with tractable solutions. Finally, it can achieve a suboptimal solution for the throughput problem using an alternating optimization approach. Accordingly, the fairness problem can be resolved using an adjusted algorithm according to the throughput problem.

The remainder of the present paper is organized as follows. Section II specifies the system model. The full DoF of the MIMO-RBC with the proposed scheme and a general description of the optimization problems are given in Section III. The equivalent optimization problems are presented in Section IV. In Section V, this work describes the design of the base station and relay matrices using the WMMSE method for the throughput optimization problem. Section VI provides the design method for the base station and relay matrices for the fairness optimization problem based on the obtained method by dealing with the throughput optimization problem. Finally, numerical results and conclusions are presented in Sections VII and VIII, respectively.

Notation: The following notational conventions are followed throughout the paper. Vectors and matrices are presented in boldface lowercase and uppercase letters, respectively. The trace, expectation, inverse, transpose, conjugate transpose, determinant, rank, and pseudoinverse of a matrix are denoted by $\mathrm{Tr}(\cdot)$, $\mathsf{E}(\cdot)$, $(\cdot)^{-1}$, $(\cdot)^T$, $(\cdot)^\dagger$, $\det|\cdot|$, $\mathrm{Rank}(\cdot)$, and $\mathrm{Pinv}(\cdot)$, respectively. $\mathrm{diag}(a_1, \ldots, a_N)$ is a diagonal matrix, for which a_i is the ith diagonal entry. $\mathrm{diag}[a]_N$ is an $N \times N$ diagonal matrix with diagonal entry a. A set of $M \times N$ matrices over a complex field is expressed as $C^{M \times N}$, and $\mathbf{I}$ is the identity matrix with appropriate dimensions. $\mathcal{CN}(x, y)$ is used to indicate a circularly symmetric complex Gaussian distribution with mean x and covariance y. Moreover, i.i.d. indicates independent and identically distributed. The set of mobile-users is expressed as $\mathcal{U} = \{1, 2, \cdots, K\}$, where K is the number of mobile-users.

6.2 System Model

The system model considered in the present study, which is a general MIMO-RBC consisting of one source (i.e., base station (BS)), one relay station (RS), and K multi-antenna mobile users (MUs), is shown in Fig 6.1. The numbers of antennas for the BS and RS are N_b and N_r, respectively, and the number of receiver antennas for the kth MU is N_k. There are two kinds of MUs in the MIMO-RBC

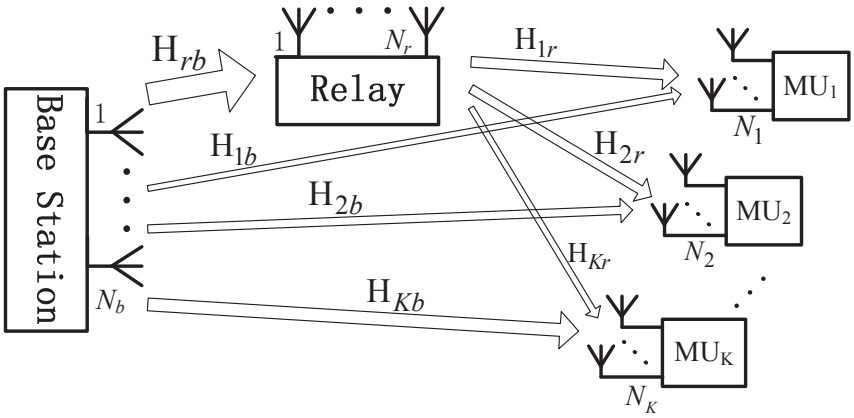

Fig. 6.1: The system model of MIMO-RBC consists of three parts: the base station which acts as source, relay station, and mobile users. The number of mobile users (direct-users and relay-users) is K .

considered herein: direct users and relay users. For the direct users, the DLs gains are better than the RLs gains, and the opposite is true for the relay users. The sets $\mathcal{B}$ and $\mathcal{R}$ contain the indices of the direct users and relay users, respectively. The total number of MUs is $|\mathcal{B} \cup \mathcal{R}| = K$. However, it is not necessary to distinguish direct users from relay users in the remainder of the present paper. The RS is used to aid the data communication between the BS and the relay users. Hence, the number of antennas for the RS must satisfy $N_r \geq \sum_{k\in\mathcal{R}} N_k$. Note that this system model is used for the relay architectures of the third-generation partnership project long-term evolution advanced (3GPP LTE-A) [108]. This work considers the BS and RS to be in a flat fading channel. The BS (or RS) and MUs are also in flat fading channels in the case of a non-line-of-sight (NLOS) scenario[1]. Note that, in the proposed scheme, it is only necessary to satisfy $\sum_{k=1}^{K} N_k \leq 2N_b$ in order to simultaneously support N_k independent substreams for the kth user, which is remarkably different from the conventional schemes for MIMO-RBC [96–102] that require $\sum_{k=1}^{K} N_k \leq N_b$. The considered relaying scheme can be categorized as a non-regenerative, half-duplex scheme [109], and the data transmission scheme has two phases.

[1]In fact, the line-of-sight (LOS) scenario can be seen as a special case of the NLOS, and the results can be extended directly to the LOS scenario.

6.2.1 First Phase:

Suppose that $\mathbf{s}_k$ is the symbol intended for the kth MU, and the signal vector of one MU is independent from those of the other MUs. In the first phase, by applying a linear PM $\mathbf{P} = [\mathbf{P}_1, \mathbf{P}_2, \cdots, \mathbf{P}_K]$ to the data vector $\mathbf{s}$, where $\mathbf{s} = [\mathbf{s}_1^T, \mathbf{s}_2^T, \ldots, \mathbf{s}_K^T]^T \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$, and $\mathbf{P}_k \in C^{N_b \times N_k}$ is a PM acting on signal vector $\mathbf{s}_k$ for the kth MU, the BS broadcasts the precoded data streams to the RS and MUs. Accordingly, the obtained signal vector for the kth MU and RS can be formulated as follows:

$$\mathbf{y}_{1k} = \mathbf{H}_{kb}\mathbf{P}_k\mathbf{s}_k + \sum_{i=1, i\neq k}^{K} \mathbf{H}_{kb}\mathbf{P}_i\mathbf{s}_i + \mathbf{n}_k, \tag{6.1}$$

$$\mathbf{y}_r = \mathbf{H}_{rb}\mathbf{P}\mathbf{s} + \mathbf{n}_r. \tag{6.2}$$

In the above expressions, the channel matrix between the BS and the kth MU (RS) is denoted by $\mathbf{H}_{kb}$ ($\mathbf{H}_{rb}$), and each entry in $\mathbf{H}_{kb}$ ($\mathbf{H}_{rb}$) is an i.i.d. complex Gaussian variable[2] mean zero and variance σ_{kb}^2 (σ_{rb}^2). Moreover, $\mathbf{n}_x \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ ($x = k$ and r) represents the Gaussian noise vectors for the kth MU and RS, respectively.

6.2.2 Second Phase:

By applying a linear BM $\mathbf{G}$ to the received signal vector, the received signals are forwarded to MUs by the RS during the second phase. At the same time, the BS applies a new linear PM $\mathbf{F} = [\mathbf{F}_1, \mathbf{F}_2, \cdots, \mathbf{F}_K]$ to the data vector $\mathbf{s}$ [3], and then broadcasts the precoded data streams to MUs, where $\mathbf{F}_k \in C^{N_b \times N_k}$ is a PM acting on signal vector $\mathbf{s}_k$ for the kth MU. This is another remarkable difference from conventional MIMO-RBC schemes [96–102], in which the BS remains silent during the second phase. Note that the proposed scheme will definitely increase the diversity gain by allowing the BS to transmit continuously during the second phase. By assuming that the BS-receiver channels experience slow fading and remain unchanged during the second phase [111], the received signal vector at the kth MU during the second phase can be written as follows:

$$\begin{aligned}\mathbf{y}_{2k} = {} & \mathbf{H}_{kb}\mathbf{F}_k\mathbf{s}_k + \mathbf{H}_{kr}\mathbf{G}\mathbf{H}_{rb}\mathbf{P}_k\mathbf{s}_k \\ & + \sum_{i=1, i\neq k}^{K} \left(\mathbf{H}_{kb}\mathbf{F}_i + \mathbf{H}_{kr}\mathbf{G}\mathbf{H}_{rb}\mathbf{P}_i\right)\mathbf{s}_i + \mathbf{H}_{kr}\mathbf{G}\mathbf{n}_r + \mathbf{z}_k,\end{aligned} \tag{6.3}$$

[2] The same assumption can be found in [90, 110], etc.

[3] The reason for the BS transmitting the same data vector $\mathbf{s}$ to the MUs during the second phase is that the RS needs the BS to cooperatively transmit the $\mathbf{s}$ to MUs. In fact, the data vector $\mathbf{s}$ is transmitted primarily to the RS by the BS during the first phase and is transmitted primarily to MUs by the BS during the second phase.

where $\mathbf{H}_{kr}$ is the channel matrix from the RS to the kth MU, in which the entry is assumed to be an i.i.d. complex Gaussian variable with mean zero and variance σ_{kr}^2, and $\mathbf{z}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I})$ is the Gaussian noise vector observed at the kth MU.

Suppose that P_b and P_r are power constraints for the BS and the RS, respectively, in the two phases. Then, it can be obtained that

$$\begin{aligned} \mathrm{Tr}(\mathbf{P}^\dagger\mathbf{P}) \le P_b, \text{ and } \mathrm{Tr}(\mathbf{F}^\dagger\mathbf{F}) &\le P_b, && \text{(6.4a)} \\ \mathrm{Tr}(\mathbf{G}\mathbf{H}_{rb}\mathbf{P}\mathbf{P}^\dagger\mathbf{H}_{rb}^\dagger\mathbf{G}^\dagger + \mathbf{G}\mathbf{G}^\dagger) &\le P_r. && \text{(6.4b)} \end{aligned}$$

For simplicity, (6.1) and (6.3) can be rewritten in matrix form as

$$\begin{aligned} \underbrace{\begin{bmatrix} \mathbf{y}_{1k} \\ \mathbf{y}_{2k} \end{bmatrix}}_{\mathbf{y}_k} &= \underbrace{\begin{bmatrix} \mathbf{H}_{kb} & \mathbf{0} \\ \mathbf{H}_{krb} & \mathbf{H}_{kb} \end{bmatrix}}_{\mathbf{H}_k} \underbrace{\begin{bmatrix} \mathbf{P}_k \\ \mathbf{F}_k \end{bmatrix}}_{\widetilde{\mathbf{P}}_k} \mathbf{s}_k + \underbrace{\begin{bmatrix} \mathbf{n}_k \\ \mathbf{H}_{kr}\mathbf{G}\mathbf{n}_r + \mathbf{z}_k \end{bmatrix}}_{\mathbf{N}_k} \\ &\quad + \sum_{i=1, i\neq k}^{K} \begin{bmatrix} \mathbf{H}_{kb} & \mathbf{0} \\ \mathbf{H}_{krb} & \mathbf{H}_{kb} \end{bmatrix} \underbrace{\begin{bmatrix} \mathbf{P}_i \\ \mathbf{F}_i \end{bmatrix}}_{\widetilde{\mathbf{P}}_i} \mathbf{s}_i, \end{aligned} \tag{6.5}$$

where $\mathbf{H}_{krb} \triangleq \mathbf{H}_{kr}\mathbf{G}\mathbf{H}_{rb}$.

Due to the errors introduced by channel estimation, quantization, reciprocity mismatch, and delay, for example, the CSIT is an estimate of the true channel response $\mathbf{H}$ [90, 112]. Then, the channel can be formulated as

$$\begin{aligned} \mathbf{H}_{rb} &= \widehat{\mathbf{H}}_{rb} + e_{rb}\mathbf{\Omega}_{rb}, && \text{(6.6a)} \\ \mathbf{H}_{kb} &= \widehat{\mathbf{H}}_{kb} + e_{kb}\mathbf{\Omega}_{kb}, && \text{(6.6b)} \\ \mathbf{H}_{kr} &= \widehat{\mathbf{H}}_{kr} + e_{kr}\mathbf{\Omega}_{kr}, && \text{(6.6c)} \end{aligned}$$

where $\widehat{\mathbf{H}}_x$ ($x \triangleq rb, kb$ and kr) is the estimated CSIT, e_x is the estimation error, $\mathbf{\Omega}_x$ is independent of $\mathbf{H}_x$, and each entry in $\mathbf{\Omega}_x$ is an i.i.d. complex Gaussian variable with mean zero and variance σ_x^2. The entries of $\widehat{\mathbf{H}}_x$ are also i.i.d. complex Gaussian variables with mean zero and variance $(1 - e_x^2)\sigma_x^2$ [90]. Assume that $e_{kb} = e_{kr}$ for $k \in \mathcal{U}$, and $e_x\sigma_x^2$ is known to the BS [90]. Accordingly, it is necessary to design the BS PM and RS BM using imperfect CSIT.

Considering Gaussian signaling at the source, the achievable rate for the kth MU during the two phases can be obtained as follows:

$$\mathcal{R}_k = \log\det\left|\mathbf{I} + \widetilde{\mathbf{P}}_k^\dagger\mathbf{H}_k^\dagger\mathbf{R}_{\mathbf{I}_k}^{-1}\mathbf{H}_k\widetilde{\mathbf{P}}_k\right|, \tag{6.7}$$

where $\mathbf{R}_{\mathbf{I}_k} = \mathbf{N}_k\mathbf{N}_k^\dagger + \sum_{i=1,i\neq k}^{K}\mathbf{H}_k\widetilde{\mathbf{P}}_i\widetilde{\mathbf{P}}_i^\dagger\mathbf{H}_k^\dagger$, and $\mathbf{N}_k$ and $\mathbf{H}_k$ are given in Eq.(6.5). A concise proof of (6.7) is given in the Appendix.

6.3 Problem Formulation

Before formulating the objective problems, this part first introduces a theorem that demonstrates a significantly different feature of the proposed scheme as compared with the conventional MIMO-RBC [96–102]. This theorem indicates that the proposed scheme can support a maximum of $\sum_{k=1}^{K} N_k$ (if $\sum_{k=1}^{K} N_k \leq 2N_b$) substreams, while existing schemes can only support a maximum of N_b substreams.

Theorem 1 *Let* $\widetilde{\mathbf{H}}_{kb} \triangleq [\mathbf{H}_{krb} \ \ \mathbf{H}_{kb}]$, *and* $\widetilde{\mathbf{H}}_{\Sigma j} \triangleq \left[\widetilde{\mathbf{H}}_{1b}^{T}, \cdots, \widetilde{\mathbf{H}}_{jb}^{T}\right]^{T}$, $(k = 1, \cdots, K;\ 1 \leq j \leq K)$. *It is feasible for the considered scheme to support a maximum of* $\sum_{k=1}^{K} N_k$ *substreams if* $\text{Rank}\left(\widetilde{\mathbf{H}}_{\Sigma K}\right) = \sum_{k=1}^{K} N_k \leq 2N_b$, *i.e., the rank of the correlated channel is full.*

Proof: The multiple-antenna user can be seen as a simple combination of multiple single antenna users in full rank channels, and it will not affect the maximum of substreams which can be supported for the considered scheme, since the post-coding of multiple-antenna user only affects the rate of gain and Theorem 1 focuses on the maximum of supported substreams. Therefore without loss of generality [97], this work can only consider the case that each user is only equipped with single antenna to simplify the proof. Due to the fact that each user is with single antenna, it can be obtained that $\text{Rank}\left(\widetilde{\mathbf{H}}_{kb}\right) = 1$, $(k = 1, \ldots, K)$, and $\text{Rank}\left(\widetilde{\mathbf{H}}_{\Sigma K}\right) = K \leq 2N_b$. Let $\mathbf{G} = \alpha_r \mathbf{I}$ (α_r is a power control factor for satisfying the RS power control), and based on the zero-forcing precoding theory [100], it can be obtained that $\widetilde{\mathbf{P}} \triangleq \begin{bmatrix} \mathbf{P} \\ \mathbf{F} \end{bmatrix} = \text{Pinv}\left(\widetilde{\mathbf{H}}_{\Sigma K}\right)(\text{diag}\,(\gamma_1, \cdots, \gamma_K))^{\frac{1}{2}}$, where γ_i is a power control factor for satisfying the BS power control. Then, substituting $\mathbf{G} = \alpha_r \mathbf{I}$, and $\widetilde{\mathbf{P}}$ into (6.7), it can be directly obtained that it is feasible to support $\sum_{k=1}^{K} N_k$ substreams with appropriate power control factors α_r and diagonal matrix $\text{diag}\,(\gamma_1, \cdots, \gamma_K)$, since the $\mathcal{R}_k$ in (6.7) can be represented as $\log\det\left(1 + \frac{\gamma_k}{\zeta_k}\right)$, where ζ_k can be seen as an invariant factor which is mainly related to the channel matrix $\widetilde{\mathbf{H}}_{\Sigma K}$ and noise [100, 110].

Based on *Theorem 1*, the following corollary is obtainable.

Corollary 1 *The maximum DoF for the relaying scheme considered herein is* $\min\{2N_b, \sum_{k=1}^{K} N_k\}$. *In other words, the maximal spatial multiplexing gain of the MIMO-RBC is* $\min\{2N_b, \sum_{k=1}^{K} N_k\}$

Note that the maximum DoF in the conventional relaying scheme is $\min\{N_b, \sum_{k=1}^{K} N_k\}$ [96–102]. In other words, the maximal spatial multiplexing gain of the existing MIMO-RBC scheme is $\min\{N_b, \sum_{k=1}^{K} N_k\}$. Hence, the proposed scheme significantly increases the DoF of the MIMO-RBC.

Next, two optimization problems is formulated according to the throughput and fairness criteria, respectively. The throughput-based problem is referred as *Problem Tp*, and the fairness-based problem is referred as *Problem Fp*.

6.3.1 Problem Tp

Under the constraint of the BS transmission power budget P_b for the two phases and the constraint of the RS transmission power budget P_r, jointly design BS PM $\mathbf{P}$ and $\mathbf{F}$ and RS BM $\mathbf{G}$ so that the throughput of the system can be maximized according to the imperfect CSIT, i.e.,

$$[\mathbf{P}, \mathbf{F}, \mathbf{G}] = \arg \max_{\{\mathbf{P},\mathbf{F},\mathbf{G}\}} \mathcal{R}_{sum}, \tag{6.8a}$$

$$\text{s.t. :} \quad (6.4), \tag{6.8b}$$

where $\mathcal{R}_{sum} = \sum_{k=1}^{K} \left(\widehat{\mathcal{R}}_k = \mathsf{E}\left[\mathcal{R}_k \left| \widehat{\mathbf{H}}_x, e_x \right. \right] \right)$, $x \triangleq rb, kb$, and kr.

6.3.2 Problem Fp

Under the constraint of the BS transmission power budget P_b for the two phases and the constraint of the RS transmission power budget P_r, jointly design BS PM $\mathbf{P}$ and $\mathbf{F}$ and RS BM $\mathbf{G}$ so that the minimum achievable rate among all MUs can be maximized according to the imperfect CSIT, i.e.,

$$[\mathbf{P}, \mathbf{F}, \mathbf{G}] = \arg \max_{\{\mathbf{P},\mathbf{F},\mathbf{G}\}} \mathcal{R}_{min}, \tag{6.9a}$$

$$\text{s.t. :} \quad (6.4), \tag{6.9b}$$

where $\mathcal{R}_{min} = \min_{k \in \mathcal{U}} \widehat{\mathcal{R}}_k$, and $\widehat{\mathcal{R}}_k = \mathsf{E}\left[\mathcal{R}_k \left| \widehat{\mathbf{H}}_x, e_x \right. \right]$, $x \triangleq rb, kb$ and kr.

It can be verified that both optimization problems are non-linear and non-convex, and the optimal closed-form solutions are difficult to obtain directly [113] [4].

6.4 Equivalent Optimization Problems

Since the above-mentioned optimization problems are non-linear and non-convex, this part is finding suboptimal solutions rather than optimal solutions for the two optimization problems. Hence, The expressions of the achievable data rate for MUs is first expanded according to the imperfect CSIT. Then two equivalent optimization problems are constructed based on a tight lower bound of the achievable

[4] This can be proven by convex optimization based on [113].

rate to replace the original optimization problems. Finally, a general linear iterative design algorithm by alternating optimization is proposed in order to solve the equivalent optimization problems. Let

$$
\begin{aligned}
&\xi_x \triangleq e_x\sigma_x \ (x \triangleq rb,\ kb \text{ and } kr),\\
&\widehat{\mathbf{H}}_{krb} \triangleq \widehat{\mathbf{H}}_{kr}\mathbf{G}\widehat{\mathbf{H}}_{rb},\\
&\mathbb{A}_k \triangleq \widehat{\mathbf{H}}_{kb}\mathbf{P}_k\mathbf{P}_k^{\dagger}\widehat{\mathbf{H}}_{kb}^{\dagger} + \mathrm{diag}\left[\xi_{kb}^2\mathrm{Tr}\left(\mathbf{P}_k\mathbf{P}_k^{\dagger}\right)\right]_{N_k},\\
&\mathbb{B}_k \triangleq \widehat{\mathbf{H}}_{kb}\mathbf{P}_k\mathbf{P}_k^{\dagger}\widehat{\mathbf{H}}_{krb}^{\dagger} + \widehat{\mathbf{H}}_{kb}\mathbf{P}_k\mathbf{F}_k^{\dagger}\widehat{\mathbf{H}}_{kb}^{\dagger}\\
&\quad + \mathrm{diag}\left[\xi_{kb}^2\mathrm{Tr}\left(\mathbf{P}_k\mathbf{F}_k^{\dagger}\right)\right]_{N_k},\\
&\mathbb{C}_k \triangleq \mathbb{B}_k^{\dagger},\\
&\mathbb{E}_k \triangleq \widehat{\mathbf{H}}_{kr}\mathbf{G}\mathbf{G}^{\dagger}\widehat{\mathbf{H}}_{kr}^{\dagger} + \mathrm{diag}\left[\xi_{kr}^2\mathrm{Tr}\left(\mathbf{G}\mathbf{G}^{\dagger}\right)\right]_{N_k} + \mathbf{I}_{N_k},\\
&\mathbb{D}_k \triangleq \widehat{\mathbf{H}}_{kr}\mathbf{G}\mathrm{diag}\left[\xi_{rb}^2\mathrm{Tr}\left(\mathbf{P}_k\mathbf{P}_k^{\dagger}\right)\right]_{N_r}\mathbf{G}^{\dagger}\widehat{\mathbf{H}}_{kr}^{\dagger}\\
&\quad + \widehat{\mathbf{H}}_{krb}\mathbf{P}_k\mathbf{P}_k^{\dagger}\widehat{\mathbf{H}}_{krb}^{\dagger}\\
&\quad + \mathrm{diag}\left[\xi_{kr}^2\mathrm{Tr}\left(\mathbf{G}\widehat{\mathbf{H}}_{rb}\mathbf{P}_k\mathbf{P}_k^{\dagger}\widehat{\mathbf{H}}_{rb}^{\dagger}\mathbf{G}^{\dagger}\right)\right]_{N_k}\\
&\quad + \widehat{\mathbf{H}}_{kb}\mathbf{F}_k\mathbf{P}_k^{\dagger}\widehat{\mathbf{H}}_{krb}^{\dagger} + \widehat{\mathbf{H}}_{krb}\mathbf{P}_k\mathbf{F}_k^{\dagger}\widehat{\mathbf{H}}_{kb}^{\dagger}\\
&\quad + \widehat{\mathbf{H}}_{kb}\mathbf{F}_k\mathbf{F}_k^{\dagger}\widehat{\mathbf{H}}_{kb}^{\dagger} + \mathrm{diag}\left[\xi_{kb}^2\mathrm{Tr}\left(\mathbf{F}_k\mathbf{F}_k^{\dagger}\right)\right]_{N_k}.
\end{aligned}
$$

It can be obtained that

$$
\begin{aligned}
\widehat{\mathcal{R}}_k &= \mathsf{E}\left[\mathcal{R}_k\left|\widehat{\mathbf{H}}_x, e_x\right.\right],\ (x \triangleq rb,\ kb \text{ and } kr)\\
&= \log\det\left|\sum_{k=1}^{K}\begin{bmatrix}\mathbb{A}_k & \mathbb{B}_k\\ \mathbb{C}_k & \mathbb{D}_k\end{bmatrix} + \begin{bmatrix}\mathbf{I}_{N_k} & \mathbf{0}\\ \mathbf{0} & \mathbb{E}_k\end{bmatrix}\right| -\\
&\quad \log\det\left|\sum_{i=1,i\neq k}^{K}\begin{bmatrix}\mathbb{A}_i & \mathbb{B}_i\\ \mathbb{C}_i & \mathbb{D}_i\end{bmatrix} + \begin{bmatrix}\mathbf{I}_{N_k} & \mathbf{0}\\ \mathbf{0} & \mathbb{E}_k\end{bmatrix}\right|, \qquad (6.10)
\end{aligned}
$$

where the following property has been used:

$$\mathsf{E}\left[\mathbf{\Omega}_x\mathbf{X}\mathbf{\Omega}_x^{\dagger}\right] = \mathrm{diag}\left[\sigma_x^2\mathrm{Tr}(\mathbf{X})\right]_{N_k},\ (x = rb,\ kb, \text{and } kr).$$

According to Eq. (6.10), problem (6.8) is a very difficult optimization problem. In order to obtain an efficient solution for the original problem, this work first focuses on finding the tight lower bound for $\widehat{\mathcal{R}}_k$. It then sets up another optimization problem in terms of this lower bound, which enables the solution to proceed. Let

$$\overline{\mathbb{A}}_k \triangleq \mathrm{diag}\left[\xi_{kb}^2\mathrm{Tr}\left(\mathbf{P}_k\mathbf{P}_k^{\dagger}\right)\right]_{N_k},$$

$$\overline{\mathbb{B}}_k \triangleq \operatorname{diag}\left[\xi_{kb}^2 \operatorname{Tr}\left(\mathbf{P}_k \mathbf{F}_k^{\dagger}\right)\right]_{N_k},$$

$$\overline{\mathbb{C}}_k = \overline{\mathbb{B}}_k^{\dagger},$$

$$\begin{aligned}\overline{\mathbb{D}}_k \triangleq\ & \widehat{\mathbf{H}}_{kr}\mathbf{G}\operatorname{diag}\left[\xi_{rb}^2 \operatorname{Tr}\left(\mathbf{P}_k \mathbf{P}_k^{\dagger}\right)\right]_{N_r} \mathbf{G}^{\dagger}\widehat{\mathbf{H}}_{kr}^{\dagger} \\ &+ \operatorname{diag}\left[\xi_{kr}^2 \operatorname{Tr}\left(\mathbf{G}\widehat{\mathbf{H}}_{rb}\mathbf{P}_k\mathbf{P}_k^{\dagger}\widehat{\mathbf{H}}_{rb}^{\dagger}\mathbf{G}^{\dagger}\right)\right]_{N_k} \\ &+ \operatorname{diag}\left[\xi_{kb}^2 \operatorname{Tr}\left(\mathbf{F}_k \mathbf{F}_k^{\dagger}\right)\right]_{N_k},\end{aligned}$$

$$\widehat{\mathbf{H}}_k \triangleq \begin{bmatrix} \widehat{\mathbf{H}}_{kb} & \mathbf{0} \\ \widehat{\mathbf{H}}_{krb} & \widehat{\mathbf{H}}_{kb} \end{bmatrix},$$

$$\widehat{\mathbf{R}}_{\mathbf{I}_k} \triangleq \sum_{i=1, i\neq k}^{K} \begin{bmatrix} \mathbb{A}_i & \mathbb{B}_i \\ \mathbb{C}_i & \mathbb{D}_i \end{bmatrix} + \begin{bmatrix} \mathbf{I}_{N_k} & \mathbf{0} \\ \mathbf{0} & \mathbb{E}_k \end{bmatrix},$$

$$\mathbf{Z}_k \triangleq \begin{bmatrix} \overline{\mathbb{A}}_k & \overline{\mathbb{B}}_k \\ \overline{\mathbb{C}}_k & \overline{\mathbb{D}}_k \end{bmatrix},$$

$$\widehat{\mathbf{R}}_{\Sigma_k} \triangleq \widehat{\mathbf{R}}_{\mathbf{I}_k} + \widehat{\mathbf{H}}_k \widetilde{\mathbf{P}}_k \widetilde{\mathbf{P}}_k^{\dagger} \widehat{\mathbf{H}}_k^{\dagger} + \mathbf{Z}_k.$$

Then, it has

$$\begin{aligned}\widehat{\mathcal{R}}_k &= \log\det\left|\mathbf{I}_{N_k} + \left(\widehat{\mathbf{H}}_k\widetilde{\mathbf{P}}_k\widetilde{\mathbf{P}}_k^{\dagger}\widehat{\mathbf{H}}_k^{\dagger} + \overline{\mathbb{A}}_k\right)\widehat{\mathbf{R}}_{\mathbf{I}_k}^{-1}\right| \\ &\overset{a}{\geq} \log\det\left|\mathbf{I}_{N_k} + \widehat{\mathbf{H}}_k\widetilde{\mathbf{P}}_k\widetilde{\mathbf{P}}_k^{\dagger}\widehat{\mathbf{H}}_k^{\dagger}\widehat{\mathbf{R}}_{\mathbf{I}_k}^{-1}\right| \\ &\overset{b}{\geq} \log\det\left|\mathbf{I}_{N_k} + \widehat{\mathbf{H}}_k\widetilde{\mathbf{P}}_k\widetilde{\mathbf{P}}_k^{\dagger}\widehat{\mathbf{H}}_k^{\dagger}\left(\widehat{\mathbf{R}}_{\mathbf{I}_k} + \mathbf{Z}_k\right)^{-1}\right| \\ &\overset{c}{=} \log\det\left|\mathbf{I}_{N_k} + \widetilde{\mathbf{P}}_k^{\dagger}\widehat{\mathbf{H}}_k^{\dagger}\left(\widehat{\mathbf{R}}_{\mathbf{I}_k} + \mathbf{Z}_k\right)^{-1}\widehat{\mathbf{H}}_k\widetilde{\mathbf{P}}_k\right| \\ &\overset{d}{=} -\log\det\left|\mathbf{I}_{N_k} - \widetilde{\mathbf{P}}_k^{\dagger}\widehat{\mathbf{H}}_k^{\dagger}\Big(\widehat{\mathbf{R}}_{\mathbf{I}_k} + \mathbf{Z}_k \right. \\ &\qquad\qquad \left. + \widehat{\mathbf{H}}_k\widetilde{\mathbf{P}}_k\widetilde{\mathbf{P}}_k^{\dagger}\widehat{\mathbf{H}}_k^{\dagger}\Big)^{-1}\widehat{\mathbf{H}}_k\widetilde{\mathbf{P}}_k\right| \\ &= -\log\det\left|\mathbf{I}_{N_k} - \widetilde{\mathbf{P}}_k^{\dagger}\widehat{\mathbf{H}}_k^{\dagger}\widehat{\mathbf{R}}_{\Sigma_k}^{-1}\widehat{\mathbf{H}}_k\widetilde{\mathbf{P}}_k\right| \\ &\triangleq -\log\det|\mathbf{E}_k|, \end{aligned} \tag{6.11}$$

where (a) is obtained by ignoring $\overline{\mathbb{A}}_k$ caused by errors, (b) is obtained because $\mathbf{Z}_k \geq 0$, (c) is based on the property whereby $\det|\mathbf{I} + \mathbf{AB}| = \det|\mathbf{I} + \mathbf{BA}|$, and (d) originates from the Woodbury matrix identity, i.e., $(\mathbf{B} + \mathbf{UAV})^{-1} = \mathbf{B}^{-1} -$

$\mathbf{B}^{-1}\mathbf{U}\left(\mathbf{A}^{-1}+\mathbf{V}\mathbf{B}^{-1}\mathbf{U}\right)^{-1}\mathbf{V}\mathbf{B}^{-1}$ [114]. $\mathbf{E}_k$ has been used to indicate $\mathbf{I}_{N_k}-\widetilde{\mathbf{P}}_k^{\dagger}\widehat{\mathbf{H}}_k^{\dagger}\widehat{\mathbf{R}}_{\Sigma_k}^{-1}\widehat{\mathbf{H}}_k\widetilde{\mathbf{P}}_k$. Based on the above equation, the lower bound is tight, because $\overline{\mathbb{A}}_k$ and $\mathbf{Z}_k$, which determine the differences produced in the expressions of (a) and (b), respectively, are both small for the case in which $e_x^2 \ll 1$ ($x = kb,\ rb$ and kr). On the other hand, the equation also reveals that the lower bound is tight, i.e., when $e_x = 0$, it can be obtained that $R_k = \widehat{R}_k = -\log\det|\mathbf{E}_k|$.

6.4.1 Equivalent Problem Tp

Therefore, the optimization problem *Tp* based on the lower bound (6.11) can be shown as follows:

$$[\mathbf{P},\mathbf{F},\mathbf{G}] = \arg\max_{\{\mathbf{P},\mathbf{F},\mathbf{G}\}} \sum_{k=1}^{K} -\log\det|\mathbf{E}_k|\,, \tag{6.12a}$$

$$\text{s.t. :}\quad (6.4). \tag{6.12b}$$

In order to obtain the solution for the relaxed problem in (6.12), the following theorem is necessary.

Theorem 2 *Let*

$$\mathbf{M}_k \triangleq \mathbf{I} - \mathbf{A}_k\widehat{\mathbf{H}}_k\widetilde{\mathbf{P}}_k - \widetilde{\mathbf{P}}_k^{\dagger}\widehat{\mathbf{H}}_k^{\dagger}\mathbf{A}_k^{\dagger} + \mathbf{A}_k\widehat{\mathbf{R}}_{\Sigma_k}\mathbf{A}_k^{\dagger}, \tag{6.13}$$

where $\mathbf{A}_k$ is a matrix variable. Then, the optimal solution for (6.14), as formulated below, is also the solution for the relaxed problem given in (6.12):

$$[\mathbf{P},\mathbf{F},\mathbf{G},\mathbf{A}] = \arg\max_{\{\mathbf{P},\mathbf{F},\mathbf{G},\mathbf{A}\}} \sum_{k=1}^{K} -\log\det|\mathbf{M}_k|\,, \tag{6.14a}$$

$$\text{s.t. : (6.4)}, \tag{6.14b}$$

where $\mathbf{A} \triangleq \{\mathbf{A}_k\}_{k=1}^{K}$.

Proof: For any $\mathbf{P}$, $\mathbf{F}$, and $\mathbf{G}$, the optimal $\mathbf{A}_k$ can be readily found, which is equal to $\widetilde{\mathbf{P}}_k^{\dagger}\widehat{\mathbf{H}}_k^{\dagger}\widehat{\mathbf{R}}_{\Sigma_k}^{-1}$. Then, substituting $\mathbf{A}_k = \widetilde{\mathbf{P}}_k^{\dagger}\widehat{\mathbf{H}}_k^{\dagger}\widehat{\mathbf{R}}_{\Sigma_k}^{-1}$ into $\mathbf{M}_k$, $\mathbf{E}_k = \mathbf{M}_k$ is obtainable.

The closed-form solution of problem *Tp*, formulated in (6.14), is still intractable, because it is also a non-linear problem. Actually, according to the proof for Theorem 2, when $\mathbf{P}$, $\mathbf{F}$, and $\mathbf{G}$ are given, it can be obtained that the optimal $\mathbf{A}_k$ as

$$\mathbf{A}_k = \widetilde{\mathbf{P}}_k^{\dagger}\widehat{\mathbf{H}}_k^{\dagger}\widehat{\mathbf{R}}_{\Sigma_k}^{-1}. \tag{6.15}$$

Second, in order to solve problem *Tp* formulated in (6.14), this part first sets up the following optimization problem [102]:

$$\begin{aligned}
&[\mathbf{P}, \mathbf{F}, \mathbf{G}, \mathbf{W}, \mathbf{A}] \\
&= \arg \min_{\{\mathbf{P},\mathbf{F},\mathbf{G},\mathbf{W},\mathbf{A}\}} \sum_{k=1}^{K} w_k \left(\mathrm{Tr}\left(\mathbf{W}_k \mathbf{M}_k\right) - \log\det\left(\mathbf{W}_k\right)\right), \\
&\text{s.t. :(6.4)},
\end{aligned} \tag{6.16}$$

where $\mathbf{A} \triangleq \{\mathbf{A}_k\}_{k=1}^{K}$, $\mathbf{W} \triangleq \{\mathbf{W}_k\}_{k=1}^{K}$, and $\mathbf{W}_k \succeq \mathbf{0}$ is a weight matrix for the kth MU. Then, the following lemma [115] is obtainable.

Lemma 1 *Let the weighting factors* $w_1 = \cdots = w_K = 1$. *Then, the optimal solution for problem (6.16) is also the solution for problem* Tp *formulated in (6.14).*

Proof: Suppose that $\mathbf{P}^{\mathrm{opt}}$, $\mathbf{F}^{\mathrm{opt}}$, and $\mathbf{G}^{\mathrm{opt}}$ are the optimal matrices for the new optimization problem in (6.16) with $w_1 = \cdots = w_K = 1$. Then, it needs only confirm that they are also the optimal matrices for the optimization problem *Tp* in (6.14). For the given optimal matrices $\mathbf{P}^{\mathrm{opt}}$, $\mathbf{F}^{\mathrm{opt}}$, and $\mathbf{G}^{\mathrm{opt}}$, the optimal $\mathbf{A}_k$ for the kth MU to minimize the problem (6.16) can be verified to also be the optimal $\mathbf{A}_k$ for the kth MU to minimize the problem (6.15). It can also be verified that the optimum weight matrix $\mathbf{W}_k$ for the kth MU with the given $\mathbf{P}^{\mathrm{opt}}$, $\mathbf{G}^{\mathrm{opt}}$, $\mathbf{F}^{\mathrm{opt}}$, and $\mathbf{A}$ can be expressed as

$$\mathbf{W}_k = \mathbf{M}_k^{-1}. \tag{6.17}$$

Therefore, substituting $\mathbf{W}_k$ into (6.16) with $w_1 = \cdots = w_K = 1$, the following equations can be obtained:

$$\begin{aligned}
\left[\mathbf{P}^{\mathrm{opt}}, \mathbf{G}^{\mathrm{opt}}, \mathbf{F}^{\mathrm{opt}}\right] &= \arg \min - \sum_{k=1}^{K} \log\det\left(\mathbf{M}_k^{-1}\right) \\
&= \arg \max \sum_{k=1}^{K} - \log\det\left(\mathbf{M}_k\right).
\end{aligned} \tag{6.18}$$

Combining (6.18) and (6.14), the proof of Lemma 1 is completed [115].

6.4.2 Equivalent Problem Fp

Optimization problem *Fp* based on lower bound (6.11) can be rewritten as follows:

$$[\mathbf{P}, \mathbf{F}, \mathbf{G}] = \arg \min_{\{\mathbf{P},\mathbf{F},\mathbf{G}\}} \max_{k \in \mathcal{U}} \log\det\left|\mathbf{E}_k\right|, \tag{6.19a}$$

$$\text{s.t. :} \quad (6.4). \tag{6.19b}$$

In order to obtain the solution for this min-max optimization problem, the following theorem is necessary.

Theorem 3 *The optimal solution for (6.20), as formulated below, is also the solution for the min-max problem given in (6.19):*

$$
\begin{aligned}
[\mathbf{P}, \mathbf{F}, \mathbf{G}] \;=\; & \arg \min_{\{\mathbf{P},\mathbf{F},\mathbf{G}\}} \sum_{k=1}^{K} \log\det|\mathbf{E}_k|\,, \; \textit{and} && \text{(6.20a)}\\
& \text{satisfy}: \; \det|\mathbf{E}_1| = \cdots = \det|\mathbf{E}_K|\,, && \text{(6.20b)}\\
& \text{s.t.}: (6.4). && \text{(6.20c)}
\end{aligned}
$$

Proof: Since the power P_b can be distributed on-demand at the BS to find the optimal solution, if $\mathbf{P}^{\text{opt}}$, $\mathbf{F}^{\text{opt}}$, and $\mathbf{G}^{\text{opt}}$ are the optimal matrices for problem *Fp* in (6.19), they must make $\det|\mathbf{E}_1| = \cdots = \det|\mathbf{E}_K|$ the smallest achievable value, because the logarithm function is a monotonically increasing function. Therefore, $\mathbf{P}^{\text{opt}}$, $\mathbf{F}^{\text{opt}}$, and $\mathbf{G}^{\text{opt}}$ are also the optimal matrices for problem (6.20), and vice versa.

Although both equivalent optimization problems in (6.16) and (6.20) are non-linear and non-convex, this work can first solve the equivalent optimization problem in (6.16) by an alternating optimization method because the equivalent problem can be decoupled into three subproblems and each of them can be solved by the alternating optimization approach [115].

6.5 Base Station Precoding Matrices Design by the Weighted MMSE Method

This section considers the BS PM $\mathbf{P}$ and $\mathbf{F}$ design by an alternating optimization approach, which has been effectively used for solving optimization problems in signal processing and information theory, because this approach is iterative in nature and simple [115, 116]. Based on *Theorem 3*, the optimization problem *Tp* formulated in (6.16) is the basic optimization problem for problem *Fp*. Accordingly, this part first solves optimization problem *Tp* formulated in (6.16). Since this basic optimization problem is also a non-linear and non-convex optimization problem, it is difficult to directly obtain the optimal solution, especially the closed-form solution. However, the problem can be decoupled into three subproblems and solved using an alternating optimization approach. The solution obtained by this method is suboptimal.

6.5.1 BS PM P Design for the First Phase

For given $\mathbf{F}$, $\mathbf{G}$, $\mathbf{W}$, and $\mathbf{A}$, the problem in (6.16) w.r.t. $\mathbf{P}$ can be reformulated as the following weighted minimum problem:

$$\min_{\mathbf{P}} \quad \sum_{k=1}^{K} \left(\mathrm{Tr}(\mathbf{W}_k\mathbf{M}_k)\right), \tag{6.21a}$$

$$\text{s.t. :} \quad \mathrm{Tr}(\mathbf{P}^{\dagger}\mathbf{P}) \leq P_b, \tag{6.21b}$$

$$\mathrm{Tr}(\mathbf{G}\mathbf{H}_{rb}\mathbf{P}\mathbf{P}^{\dagger}\mathbf{H}_{rb}^{\dagger}\mathbf{G}^{\dagger} + \mathbf{G}\mathbf{G}^{\dagger}) \leq P_r. \tag{6.21c}$$

Since an iterative design is considered herein, the relay power constraint in (6.21c) can be ignored in order to simplify the BS PM $\mathbf{P}$ design in the first phase, which does not affect the final result. Next, this part presents the following lemma to illustrate that subproblem (6.21) with respect to PM $\mathbf{P}$ is convex without considering the relay power constraint.

Lemma 2 *For given BS PM* $\mathbf{F}$*, relay BM* $\mathbf{G}$*,* $\mathbf{A}$*, and weight matrices* $\mathbf{W}_k$*, the subproblem of BS PM* $\mathbf{P}$ *with fixed relay power to minimize the matrix-weighted sum-MSE in the considered MIMO-RBC formulated in (6.21) is convex.*

Proof: This part first proves the objective function is convex. Since

$$\mathcal{J}\left(\{\mathbf{P}_k\}_{k=1}^{K}\right) = \sum_{k=1}^{K}\left(\mathrm{Tr}(\mathbf{W}_k\mathbf{M}_k)\right) = \sum_{k=1}^{K}\mathcal{J}_k\left(\mathbf{P}_k\right), \tag{6.22}$$

where

$$\begin{aligned}
\mathcal{J}_k\left(\mathbf{P}_k\right) = \mathrm{Tr}\Big(\mathbf{W}_k\Big(\mathbf{I} - \Big[&\mathbf{A}_{k1}\widehat{\mathbf{H}}_{kb}\mathbf{P}_k + \mathbf{A}_{k2}\widehat{\mathbf{H}}_{krb}\mathbf{P}_k \\
&+\mathbf{A}_{k2}\widehat{\mathbf{H}}_{kb}\mathbf{F}_k\Big] \\
-\Big[&\mathbf{P}_k^{\dagger}\widehat{\mathbf{H}}_{kb}^{\dagger}\mathbf{A}_{k1}^{\dagger} + \mathbf{P}_k^{\dagger}\widehat{\mathbf{H}}_{krb}^{\dagger}\mathbf{A}_{k2}^{\dagger} + \mathbf{F}_k^{\dagger}\widehat{\mathbf{H}}_{kb}^{\dagger}\mathbf{A}_{k2}^{\dagger}\Big] \\
&+\mathbf{A}_{k1}\mathbf{A}_{k1}^{\dagger} + \mathbf{A}_{k2}\mathbb{E}_k\mathbf{A}_{k2}^{\dagger} + \mathbf{A}_k\mathbf{Z}_k\mathbf{A}_k^{\dagger}\Big)\Big) \\
+\sum_{i=1}^{K}\mathrm{Tr}\Big(\mathbf{W}_i\Big[&\mathbf{A}_{i1}\widehat{\mathbf{H}}_{ib}\mathbf{P}_k + \mathbf{A}_{i2}\widehat{\mathbf{H}}_{irb}\mathbf{P}_k \\
+\mathbf{A}_{i2}\widehat{\mathbf{H}}_{ib}\mathbf{F}_k\Big]&\Big[\mathbf{P}_k^{\dagger}\widehat{\mathbf{H}}_{ib}^{\dagger}\mathbf{A}_{i1}^{\dagger} + \mathbf{P}_k^{\dagger}\widehat{\mathbf{H}}_{irb}^{\dagger}\mathbf{A}_{i2}^{\dagger} + \mathbf{F}_k^{\dagger}\widehat{\mathbf{H}}_{ib}^{\dagger}\mathbf{A}_{i2}^{\dagger}\Big]\Big),
\end{aligned}$$

in which the $\mathbf{A}_k$ is needed to be split into two matrices, i.e., $\mathbf{A}_k = [\mathbf{A}_{k1},\ \mathbf{A}_{k2}]$. Hence, it needs only verify that $\mathcal{J}_k\left(\mathbf{P}_k\right)$ is convex because the sum of two convex functions is also a convex function. According to the differential rule for vectors

and the definition of Hessian matrices [117] [113], the following matrix can be obtained:

$$\mathcal{H}(\mathcal{J}_k) = \begin{bmatrix} \mathcal{H}_{\mathbf{p}_k\mathbf{p}_k^*} & \mathcal{H}_{\mathbf{p}_k\mathbf{p}_k} \\ \mathcal{H}_{\mathbf{p}_k^*\mathbf{p}_k^*} & \mathcal{H}_{\mathbf{p}_k^*\mathbf{p}_k} \end{bmatrix} \geq 0, \tag{6.23}$$

where $\mathcal{H}(\mathcal{J}_k)$ is a Hessian matrix, $\mathbf{p}_k = \text{Vec}(\mathbf{P}_k)$, $\text{Vec}(\cdot)$ signifies the matrix vectorization operator, $\mathbf{p}_k^*$ is the conjugate of $\mathbf{p}_k$, and $\mathcal{H}_{\mathbf{p}_k\mathbf{p}_k^*}$ is the partial derivative of $\mathbf{p}_k\mathbf{p}_k^*$.

Hence, the objective function in (6.21a) is convex due to $\mathcal{H}(\mathcal{J}_k) \geq 0$. Similarly, it can be verified that the feasible region of the BS power constraint in (6.21b) is also convex.

Thus, the Lagrangian function of (6.21) for $\mathbf{P}$ is given as

$$\mathcal{L}_{\mathbf{P}} = \sum_{k=1}^{K} (\text{Tr}(\mathbf{W}_k\mathbf{M}_k)) + \lambda_p \left(\text{Tr}\left(\mathbf{P}\mathbf{P}^\dagger\right) - P_b\right).$$

Accordingly, the first-order necessary condition of $\mathcal{L}_{\mathbf{P}}$ w.r.t. each $\mathbf{P}_k$ yields the KKT conditions, as follows:

$$\begin{aligned} \frac{\partial \mathcal{L}_{\mathbf{P}}}{\partial \mathbf{P}_k^*} &= -\widetilde{\mathbf{H}}_{kb}^\dagger \mathbf{A}_k^\dagger \mathbf{W}_k + \widetilde{\mathbf{H}}_{kb}^\dagger \mathbf{A}_k^\dagger \mathbf{W}_k \mathbf{A}_{k2} \widehat{\mathbf{H}}_{kb} \mathbf{F}_k \\ &\quad + \left(\sum_{i=1}^{K} \widetilde{\mathbf{H}}_{ib}^\dagger \mathbf{A}_i^\dagger \mathbf{W}_i \mathbf{A}_i \widetilde{\mathbf{H}}_{ib} + \lambda_p \mathbf{I}\right) \mathbf{P}_k = 0, \end{aligned} \tag{6.24}$$

$$\lambda_p \left(\sum_{k=1}^{K} \text{Tr}\left(\mathbf{P}_k\mathbf{P}_k^\dagger\right) - P_b\right) = 0, \tag{6.25}$$

$$\sum_{k=1}^{K} \text{Tr}\left(\mathbf{P}_k\mathbf{P}_k^\dagger\right) \leq P_b. \tag{6.26}$$

Based on the above KKT conditions, each $\mathbf{P}_k$ can be obtained as:

$$\begin{aligned} \mathbf{P}_k = &\left(\sum_{i=1}^{K} \left(\widetilde{\mathbf{H}}_{ib}^\dagger \mathbf{A}_i^\dagger \mathbf{W}_i \mathbf{A}_i \widetilde{\mathbf{H}}_{ib}\right)\right. \\ &\left. + \lambda_p \mathbf{I}\right)^{-1} \widetilde{\mathbf{H}}_{kb}^\dagger \mathbf{A}_k^\dagger \mathbf{W}_k \left(\mathbf{I} - \mathbf{A}_{k2} \widehat{\mathbf{H}}_{kb} \mathbf{F}_k\right), \end{aligned} \tag{6.27}$$

where $\widetilde{\mathbf{H}}_{ib} = \left[\widehat{\mathbf{H}}_{ib}^T \ \ \widehat{\mathbf{H}}_{irb}^T\right]^T$, and λ_p is the Lagrangian multiplier. The value of λ_p can be obtained through a 1-D search mechanism because $\text{Tr}(\mathbf{P}(\lambda_p)\mathbf{P}(\lambda_p)^\dagger)$ [5] decreases monotonically with λ_p.

[5] $\mathbf{P}(\lambda_p)$ means $\mathbf{P}$ is the function of λ_p.

6.5.2 BS PM F Design for the Second Phase

For given $\mathbf{P}$, $\mathbf{G}$, $\mathbf{W}$, and $\mathbf{A}$, the problem in (6.8) w.r.t. $\mathbf{F}$ can be reformulated as

$$\min_{\mathbf{F}} \sum_{k=1}^{K} \left(\mathrm{Tr}(\mathbf{W}_k\mathbf{M}_k)\right), \text{ s.t. : } \mathrm{Tr}(\mathbf{F}^\dagger\mathbf{F}) \leq P_b. \tag{6.28}$$

Lemma 3 *With given BS PM* $\mathbf{P}$*, relay BM* $\mathbf{G}$*,* $\mathbf{A}$*, and weight matrices* $\mathbf{W}_k$*, the subproblem of the BS PM* $\mathbf{F}$ *design to minimize the matrix-weighted sum-MSE in the MIMO-RBC considered here and formulated in (6.28) is convex.*

Proof: The proof is similar to that of Lemma 2 and is omitted here.

Thus, the Lagrangian function of (6.28) for $\mathbf{F}$ is shown by

$$\mathcal{L}_{\mathbf{F}} = \sum_{k=1}^{K} \left(\mathrm{Tr}(\mathbf{W}_k\mathbf{M}_k)\right) + \lambda_f \left(\mathrm{Tr}\left(\mathbf{F}\mathbf{F}^\dagger\right) - P_b\right).$$

Then, the first-order necessary condition of $\mathcal{L}$ w.r.t. each $\mathbf{F}_k$ yields the KKT conditions as follows:

$$\begin{aligned}\frac{\partial \mathcal{L}_{\mathbf{F}}}{\partial \mathbf{F}_k^*} = &-\widehat{\mathbf{H}}_{kb}^\dagger \mathbf{A}_{k2}^\dagger \mathbf{W}_k + \widehat{\mathbf{H}}_{kb}^\dagger \mathbf{A}_{k2}^\dagger \mathbf{W}_k \left(\mathbf{A}_{k1}\widehat{\mathbf{H}}_{kb} \right. \\ &\left. + \mathbf{A}_{k2}\widehat{\mathbf{H}}_{krb}\right)\mathbf{P}_k \\ &+ \left(\sum_{i=1}^{K} \widehat{\mathbf{H}}_{ib}^\dagger \mathbf{A}_{i2}^\dagger \mathbf{W}_i \mathbf{A}_{i2} \widehat{\mathbf{H}}_{ib} + \lambda_f \mathbf{I}\right)\mathbf{F}_k^\dagger = 0, \end{aligned} \tag{6.29}$$

$$\lambda_f \left(\sum_{k=1}^{K} \mathrm{Tr}\left(\mathbf{F}_k\mathbf{F}_k^\dagger\right) - P_b\right) = 0, \tag{6.30}$$

$$\sum_{k=1}^{K} \mathrm{Tr}\left(\mathbf{F}_k\mathbf{F}_k^\dagger\right) \leq P_b. \tag{6.31}$$

Based on the above KKT conditions, each $\mathbf{F}_k$ can be obtained as:

$$\begin{aligned}\mathbf{F}_k = &\left(\sum_{i=1}^{K} \widehat{\mathbf{H}}_{ib}^\dagger \mathbf{A}_{i2}^\dagger \mathbf{W}_i \mathbf{A}_{i2} \widehat{\mathbf{H}}_{ib} + \lambda_f \mathbf{I}\right)^{-1} \left(\widehat{\mathbf{H}}_{kb}^\dagger \mathbf{A}_{k2}^\dagger \mathbf{W}_k \big(\mathbf{I} - \right. \\ &\left. \left(\mathbf{A}_{k1}\widehat{\mathbf{H}}_{kb} + \mathbf{A}_{k2}\widehat{\mathbf{H}}_{krb}\right)\mathbf{P}_k\big)\right), \end{aligned} \tag{6.32}$$

where λ_f is the Lagrangian multiplier, which can be obtained through a 1-D search mechanism, because $\mathrm{Tr}(\mathbf{F}(\lambda_f)\mathbf{F}(\lambda_f)^\dagger)$ monotonically decreases with λ_f.

6.5.3 RS BM G Design

For given $\mathbf{P}$, $\mathbf{F}$, $\mathbf{W}$, and $\mathbf{A}$, the problem in (6.8) w.r.t. $\mathbf{G}$ can be reformulated as follows:

$$\min_{\mathbf{G}} \sum_{k=1}^{K} \left(\mathrm{Tr}(\mathbf{W}_k \mathbf{M}_k)\right), \quad \text{s.t. : (6.4}b\text{)}. \tag{6.33}$$

Lemma 4 *Given matrices* $\mathbf{P}$*,* $\mathbf{F}$*, and* $\mathbf{A}$ *and weight matrices* $\mathbf{W}_k$*, the subproblem of the RS BM* $\mathbf{G}$ *design to minimize the matrix-weighted sum-MSE in the MIMO-RBC considered here and formulated in (6.33) is convex.*

Proof: The proof is similar to that of the Lemma 2, so it is omitted here.
Therefore, the Lagrangian function of (6.33) for $\mathbf{G}$ is shown by

$$\begin{aligned}\mathcal{L}_{\mathbf{g}} = \sum_{k=1}^{K} \left(\mathrm{Tr}(\mathbf{W}_k \mathbf{M}_k)\right) + \lambda_{\mathbf{g}} \Big(\mathrm{Tr}(\mathbf{G}\widehat{\mathbf{H}}_{rb}\mathbf{P}\mathbf{P}^{\dagger}\widehat{\mathbf{H}}_{rb}^{\dagger}\mathbf{G}^{\dagger} \\ + e_{rb}^2 \sigma_{rb}^2 \mathbf{G}\mathrm{diag}\left[P_b\right]_{N_r} \mathbf{G}^{\dagger} + \mathbf{G}\mathbf{G}^{\dagger}) - P_r\Big).\end{aligned}$$

Accordingly, based on $\mathcal{L}$'s first-order necessary condition with respect to $\mathbf{G}$, the KKT conditions can be derived as follows:

$$\begin{aligned}\frac{\partial \mathcal{L}_{\mathbf{g}}}{\partial \mathbf{G}^{\dagger}} = \sum_{k=1}^{K} \Big(&- \widehat{\mathbf{H}}_{kr}^{\dagger}\mathbf{A}_{k2}^{\dagger}\mathbf{W}_k\mathbf{P}_k^{\dagger}\widehat{\mathbf{H}}_{rb}^{\dagger} \\ &+ \widehat{\mathbf{H}}_{kr}^{\dagger}\mathbf{A}_{k2}^{\dagger}\mathbf{W}_k\mathbf{A}_{k1}\widehat{\mathbf{H}}_{kb}\mathbf{P}\mathbf{P}^{\dagger}\widehat{\mathbf{H}}_{rb}^{\dagger}\Big) \\ &+ \Bigg(\sum_{k=1}^{K} \widehat{\mathbf{H}}_{kr}^{\dagger}\mathbf{A}_{k2}^{\dagger}\mathbf{W}_k\mathbf{A}_{k2}\widehat{\mathbf{H}}_{kr} \\ &+ \lambda_{\mathbf{g}}\mathbf{I}\Bigg)\mathbf{G}\left(\mathbf{I} + \widehat{\mathbf{H}}_{rb}\mathbf{P}\mathbf{P}^{\dagger}\widehat{\mathbf{H}}_{rb}^{\dagger}\right) \\ &+ \sum_{k=1}^{K}\sum_{i=1}^{K}\left(\widehat{\mathbf{H}}_{kr}^{\dagger}\mathbf{A}_{k2}^{\dagger}\mathbf{W}_k\mathbf{A}_{k2}\widehat{\mathbf{H}}_{kb}\mathbf{F}_i\mathbf{P}_i^{\dagger}\widehat{\mathbf{H}}_{rb}^{\dagger}\right) = 0.\end{aligned} \tag{6.34}$$

$$\begin{aligned}\lambda_{\mathbf{g}}\Big(\mathrm{Tr}\Big(\mathbf{G}\widehat{\mathbf{H}}_{rb}\mathbf{P}\mathbf{P}^{\dagger}\widehat{\mathbf{H}}_{rb}^{\dagger}\mathbf{G}^{\dagger} + e_{rb}^2\sigma_{rb}^2\mathbf{G}\mathrm{diag}\left[P_b\right]_{N_r}\mathbf{G}^{\dagger} \\ + \mathbf{G}\mathbf{G}^{\dagger}\Big) - P_r\Big) = 0.\end{aligned} \tag{6.35}$$

$$\begin{aligned}\mathrm{Tr}\Big(\mathbf{G}\widehat{\mathbf{H}}_{rb}\mathbf{P}\mathbf{P}^{\dagger}\widehat{\mathbf{H}}_{rb}^{\dagger}\mathbf{G}^{\dagger} + e_{rb}^2\sigma_{rb}^2\mathbf{G}\mathrm{diag}\left[P_b\right]_{N_r}\mathbf{G}^{\dagger} \\ + \mathbf{G}\mathbf{G}^{\dagger}\Big) \leq P_r.\end{aligned} \tag{6.36}$$

Based on the above KKT conditions, the relay matrix $\mathbf{G}$ can be obtained as:

$$\begin{aligned}\mathbf{G} = &\left(\sum_{k=1}^{K}\mathcal{D}_k + \lambda_{\mathbf{g}}\mathbf{I}\right)^{-1}\left(\sum_{k=1}^{K}\left(O_k - Q_k\right.\right.\\ &\left.\left.-\sum_{i=1}^{K}\mathcal{T}_{ki}\right)\right)\left(\mathbf{I} + \widehat{\mathbf{H}}_{rb}\mathbf{P}\mathbf{P}^{\dagger}\widehat{\mathbf{H}}_{rb}^{\dagger}\right)^{-1},\end{aligned} \tag{6.37}$$

where

$$\begin{aligned}\mathcal{D}_k &\triangleq \widehat{\mathbf{H}}_{kr}^{\dagger}\mathbf{A}_{k2}^{\dagger}\mathbf{W}_k\mathbf{A}_{k2}\widehat{\mathbf{H}}_{kr},\\ O_k &\triangleq \widehat{\mathbf{H}}_{kr}^{\dagger}\mathbf{A}_{k2}^{\dagger}\mathbf{W}_k\mathbf{P}_k^{\dagger}\widehat{\mathbf{H}}_{rb}^{\dagger},\\ Q_k &\triangleq \widehat{\mathbf{H}}_{kr}^{\dagger}\mathbf{A}_{k2}^{\dagger}\mathbf{W}_k\mathbf{A}_{k1}\widehat{\mathbf{H}}_{kb}\mathbf{P}\mathbf{P}^{\dagger}\widehat{\mathbf{H}}_{rb}^{\dagger},\\ \mathcal{T}_{ki} &\triangleq \widehat{\mathbf{H}}_{kr}^{\dagger}\mathbf{A}_{k2}^{\dagger}\mathbf{W}_k\mathbf{A}_{k2}\widehat{\mathbf{H}}_{kb}\mathbf{F}_i\mathbf{P}_i^{\dagger}\widehat{\mathbf{H}}_{rb}^{\dagger},\end{aligned}$$

and the Lagrangian multiplier of $\lambda_{\mathbf{g}}$ is obtained through a 1-D search mechanism.

6.5.4 General Iterative Joint Design Algorithm for Problem Tp

In the aforementioned discussion, by fixing three of the four matrices ($\mathbf{P}$, $\mathbf{F}$, $\mathbf{G}$ and $\mathbf{A}/\mathbf{W}$), the remaining matrix can be optimized. Therefore, this work proposes a joint design algorithm to jointly optimize $\mathbf{P}$, $\mathbf{F}$, $\mathbf{G}$, and $\mathbf{A}/\mathbf{W}$ based on alternating optimization. The joint design algorithm is outlined in Algorithm 1 (In the algorithm, feedback of the communication state from the receiver to the base station is necessary. This feedback is only performed one time and is not necessary for each loop of the algorithm. It is assumed that this information can be transmitted from the receiver to the base station through a dedicated control channel allocated by the base station.).

This algorithm is always convergent to a stationary point. The convergence analysis of this algorithm is also referred to as the block coordinate method [118]. The numerical results of the present study will demonstrate the convergence. In addition, the computational complexity of this algorithm is $O(N_b^3)$ times the number of iterations [118], where N_b is the number of antennas at the BS.

6.6 Matrices Design for Problem Fp

The constrained forms of the newly formulated problem of (6.20) will be relaxed to obtain the optimal matrices design for the above problem *Fp* at first. Next, the results based on the problem *Tp* can be used to solve the relaxed problem. Finally, a jointly designed general algorithm (Algorithm 2), which is similar to that based

Algorithm 4 : Joint Design Algorithm for Problem Tp

1: **Initialize:** $\mathbf{P} = \mathbf{F} = \sqrt{\frac{P_b}{N_b}}\mathbf{I}$, $\mathbf{G} = \rho\mathbf{I}$, and ρ is set under the relay's power constraints, $\mathbf{A}_k/\mathbf{W}_k \longleftarrow$ (6.15)/(6.17) with $\mathbf{P} = \mathbf{F} = \sqrt{\frac{P_b}{N_b}}\mathbf{I}$ and $\mathbf{G} = \rho\mathbf{I}$, $\forall k \in \mathcal{U}$.
2: **For:**
3: $\mathbf{P}_k \longleftarrow$ (6.27) for fixed $\mathbf{F}$, $\mathbf{G}$, $\mathbf{W}$, and $\mathbf{A}$, $\forall k \in \mathcal{U}$;
4: $\mathbf{F}_k \longleftarrow$ (6.32) for fixed $\mathbf{P}$, $\mathbf{G}$, $\mathbf{W}$, and $\mathbf{A}$, $\forall k \in \mathcal{U}$;
5: $\mathbf{G} \longleftarrow$ (6.37) for fixed $\mathbf{P}$, $\mathbf{F}$, $\mathbf{W}$, and $\mathbf{A}$;
6: $\mathbf{A}_k \longleftarrow$ (6.15) for fixed $\mathbf{P}$, $\mathbf{F}$, and $\mathbf{G}$ $\forall k \in \mathcal{U}$;
7: $\mathbf{W}_k \longleftarrow$ (6.17) for fixed $\mathbf{P}$, $\mathbf{F}$, and $\mathbf{G}$, $\forall k \in \mathcal{U}$;
8: **End:** The convergence criterion is satisfied.

on problem *Tp*, can also be obtained for the source PM and the relay BM based on problem *Fp*.

First, the constrained forms of the (6.20) can be rewritten in a relaxed manner as

$$[\mathbf{P}, \mathbf{F}, \mathbf{G}] = \arg \min_{\{\mathbf{P},\mathbf{F},\mathbf{G}\}} \sum_{k=1}^{K} (\zeta_k = \log\det|\mathbf{E}_k|), \tag{6.38a}$$

$$\text{s.t. :} \quad \mathrm{Tr}(\mathbf{P}_k\mathbf{P}_k^{\dagger}) = p_k, \text{ and } \mathrm{Tr}(\mathbf{F}_k\mathbf{F}_k^{\dagger}) = c_k,$$

$$k \in \{1, 2, \cdots, K\}, \tag{6.38b}$$

$$\mathrm{Tr}(\mathbf{G}\mathbf{H}_{rb}\mathbf{P}\mathbf{P}^{\dagger}\mathbf{H}_{rb}^{\dagger}\mathbf{G}^{\dagger} + \mathbf{G}\mathbf{G}^{\dagger}) \leq P_r, \tag{6.38c}$$

where $\mathbf{p} \triangleq \{p_k\}_{k=1}^K$ and $\mathbf{c} \triangleq \{c_k\}_{k=1}^K$ are constant and are determined in advance under the constraint of the power budget, i.e., $\sum_{k=1}^K p_k = P_b$ and $\sum_{k=1}^K c_k = P_b$.

Remark 1 *Due to the fact that $\zeta_k(c_k') < \zeta_k(c_k)$ for all $c_k' > c_k$ and fixed* $\mathbf{p}$ *(or $\zeta_k(p_k') < \zeta_k(p_k)$ for all $p_k' > p_k$ and fixed* $\mathbf{c}$*) with the optimal beamforming structure of the problem in (6.38), the predetermined constant vectors* $\mathbf{p}$ *and* $\mathbf{c}$ *can be adjusted to meet the equivalent condition of (6.20b) using an iterative method.*

Based on the relaxed problem in (6.38) and the results of problem *Tp*, the general iterative algorithm for problem *Fp* can be directly obtained, as shown in Algorithm 2.

This algorithm is also always convergent to a stationary point. The convergence analysis of this algorithm is also referred to as the block coordinate method [118]. The numerical results presented herein will demonstrate the convergence. In addition, it can be obtained that the computational complexity for this algorithm is $O(N_b^3)$ times the number of iterations [118].

Algorithm 5 : Joint Design Algorithm for Problem Fp

1: **Initialize:** $p_k = c_k = \frac{P_b}{K}$, $\mathbf{P} = \mathbf{F} = \sqrt{\frac{P_b}{N_b}}\mathbf{I}$, $\mathbf{G} = \rho\mathbf{I}$, and ρ is set under the relay's power constraints, $\mathbf{A}_k/\mathbf{W}_k \longleftarrow$ (6.15)/(6.17) with $\mathbf{P} = \mathbf{F} = \sqrt{\frac{P_b}{N_b}}\mathbf{I}$ and $\mathbf{G} = \rho\mathbf{I}$, $\forall k \in \mathcal{U}$.
2: **For:**
3: $\mathbf{P}_k \longleftarrow$ (6.27) for fixed $\mathbf{F}$, $\mathbf{G}$, $\mathbf{W}$, and $\mathbf{A}$ based on (6.38), $\forall k \in \mathcal{U}$;
4: $\mathbf{F}_k \longleftarrow$ (6.32) for fixed $\mathbf{P}$, $\mathbf{G}$, $\mathbf{W}$, and $\mathbf{A}$ based on (6.38), $\forall k \in \mathcal{U}$;
5: $\mathbf{G} \longleftarrow$ (6.37) for fixed $\mathbf{P}$, $\mathbf{F}$, $\mathbf{W}$, and $\mathbf{A}$;
6: $\mathbf{A}_k \longleftarrow$ (6.15) for fixed $\mathbf{P}$, $\mathbf{F}$, and $\mathbf{G}$ $\forall k \in \mathcal{U}$;
7: $\mathbf{W}_k \longleftarrow$ (6.17) for fixed $\mathbf{P}$, $\mathbf{F}$, and $\mathbf{G}$, $\forall k \in \mathcal{U}$;
8: Update $\mathbf{p}$ and $\mathbf{c}$ by the following steps: $p_i \longleftarrow (p_i - \triangle)$,$c_i \longleftarrow (c_i - \triangle)$, and $p_j \longleftarrow (p_j + \triangle)$, $c_j = c_j + \triangle$, where $i = \arg\min_{\{i=1,\ldots,K\}} \zeta_i$, $j = \arg\max_{\{j=1,\ldots,K\}} \zeta_j$, and $\triangle = (1 - \frac{K\zeta_i}{\sum_{k=1}^{K}\zeta_k})c_i$.
9: **End:** The convergence criterion is satisfied.

6.7 Numerical Results

In this section, numerical results are obtained in order to verify the performance superiority of the proposed design scheme over 2,000 random channel realizations for an MIMO-RBC with coordinated users (WCU). The proposed design scheme is referred to as WMMSE-WCU and is compared to the following schemes:

1. *WMMSE-MRC&MRT-WCU*. This scheme is derived from that presented in [102]. In [102], only P is designed for the BS. Here, this work has extended the scheme to include F for fair comparison. As a result, the extended scheme from [102] also considers direct users and relay users. However, the BM design for the RS is based on the principles of the maximum ratio combining (MRC) and maximum ratio transmission (MRT).

2. *GCI-I-WCU*. This scheme also considers direct users and relay users. However, the PM design for the BS is based on the GCI scheme [110], and the BM design for the RS is fixed to be scaled by the identity matrix $\mathbf{I}$.

3. *WMMSE-NCU* [102]. This scheme considers only relay users. However, this scheme is near-optimal for scenarios that do not consider coordinated users (direct users).

All schemes are compared under the same condition of various network parameters with the perfect CSIT and imperfect CSIT to verify the effectiveness of the proposed scheme in both cases. This work jointly considers the configuration

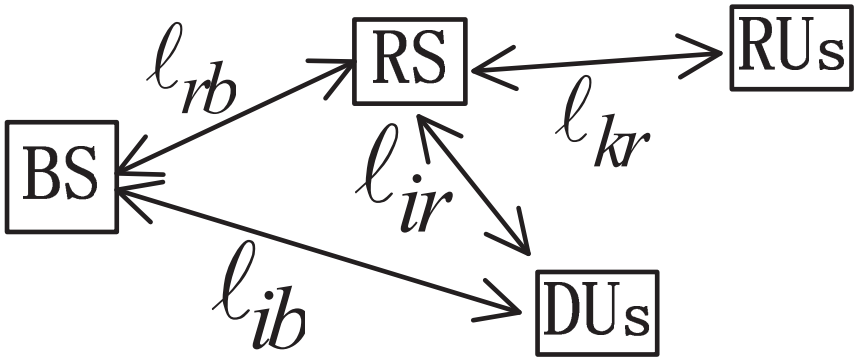

Figure 6.2: Distances between the BS and the RS, between the BS and (RUs), and between the BS and DUs, i.e., $\ell_{ib} = \frac{1}{2}(\ell_{rb} + \ell_{kr})$.

of both large-scale and small-scale fading wireless channels, and the matrices of the channel follow i.i.d. $\mathcal{CN}(0, \frac{1}{\ell^\tau})$ entries [110], where ℓ and $\tau = 3$ indicate the normalized distance between two nodes and the path loss exponent, respectively. In the network settings, the BS, the RS, and the relay users (RUs) are arranged along a line, with all of the RUs located at the same position. All of the direct users (DUs) are also deployed at the same position, and distance (ℓ_{ib}) between the BS and the DUs is equal to half of the distance ($\ell_{rb} + \ell_{kr}$) between the BS and the RUs, as depicted in Fig. 6.2 (In the settings, this work chooses the same position in order to obtain the same distance. For the same distance, different users may experience the same "long-term fading", such as path loss, for example. For the same distance, different users experience different "short-term fading". In other words, users at the same position have different multi-path fading. The proposed scheme is based on precoding, which can help us to reduce the interference.). Note that the RUs (or DUs) are deployed at the same position in these simulations only for simulation convenience, and the results can be extended to other deployment setups for RUs (or DUs).

For problem *Fp*, to provide a fair comparison, all of these schemes are adjusted to be suitable for a max-min achievable rate among all users (For the purposes of performance comparison, this work should use the same standard. For each scheme, this work has conducted a simulation to maximize the minimum-value of each user, which enables fair comparison.).

6.7.1 Convergence Property

This part first shows the convergence properties of the proposed precoding strategy in Figs. 6.3 and 6.4 for problems Tp and Fp, respectively. Figure 6.4 shows

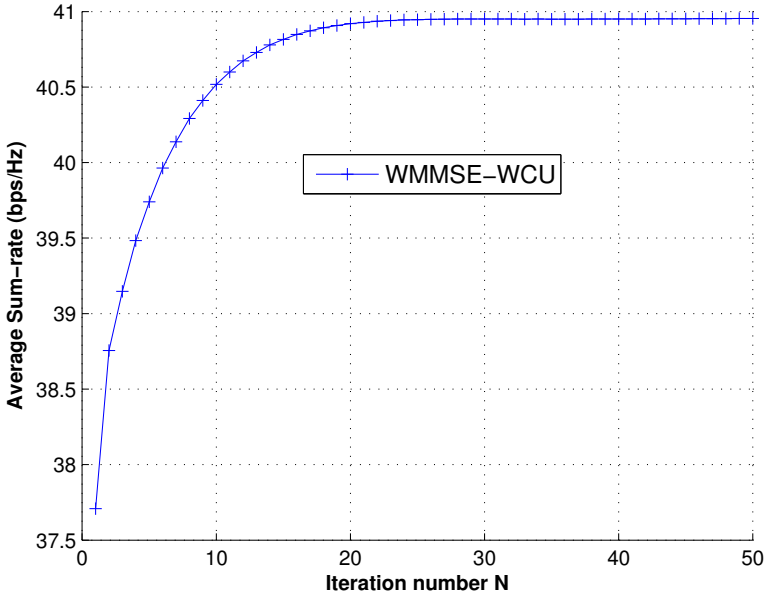

Figure 6.3: Convergence properties for one (randomly selected) channel realization with $P_b = P_r$ (signal-to-noise ratio (SNR) = 24 dB, where $N_b = N_r = 4$, $K = 4$, and $e_x = 0.1$. The BS is located at (0, 0), and the RS is located at (0, 0.5). Moreover, all relay users are located at (0, 1.0), and all direct users are located at (0.25, - $\sqrt{3}/4$) on a two-dimensional surface (i.e., $\ell_{ib} = \frac{1}{2}(\ell_{rb} + \ell_{kr})$ and $\ell_{kr} = \ell_{ir}$), and $w_k = 1$, $\forall k \in \mathcal{U}$ (Here, the SNR indicates the value of the transmitter side [90]. In other words, the SNR is equal to P_b or P_r, which has been normalized based on a noise level of 1.).

that, in the initial rounds, each user has a different minimum rate, and the minimum rates for all of the users eventually converge after 30 iterations (Please note that the convergence property observed in Figs. 6.3 and 6.4 is one kind of realization for Algorithms 1 and 2, respectively. Other topologies, such as different channel-gains, can also be used to demonstrate the convergence property of the proposed algorithms. In addition, the input parameters of Algorithm 1 (Algorithm 2) and Fig. 6.3 (Fig. 6.4) are accordance with each other.).

6.7.2 Rate Comparison

Figures 6.5 through 6.8 show the results for problem *Tp*. Figure 6.5 shows the average sum rate of the network versus the transmit power with perfect CSIT and imperfect CSIT, when the positions of all nodes are fixed. Note that, for fair com-

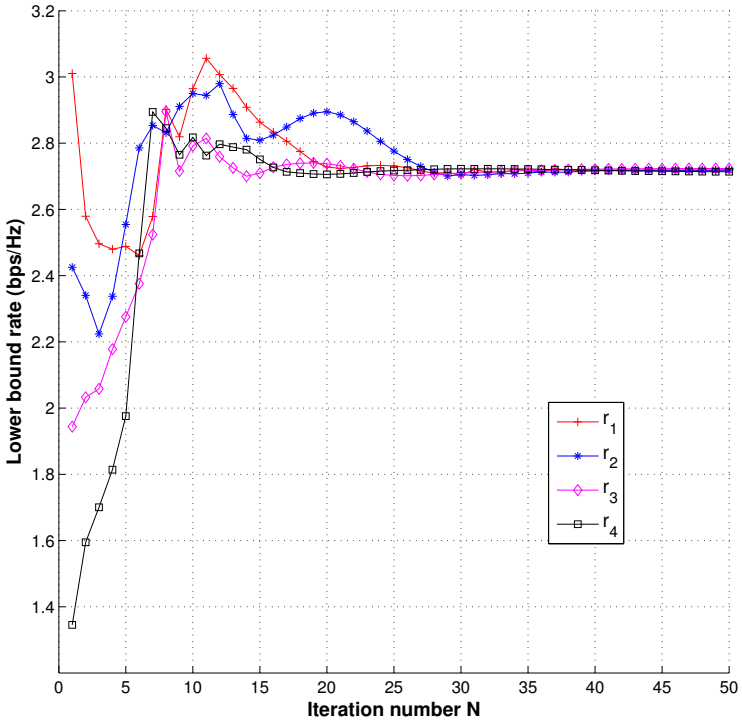

Figure 6.4: Convergence properties for one (randomly selected) channel realization with $P_b = P_r$ (SNR = 20 dB) and $r_k = -\log\det|\mathbf{E}_k|$ (k = 1, 2, 3, 4), where $N_b = N_r = 4$, $K = 4$, and $e_x = 0.2$. The BS is located at (0, 0), and the RS is located at (0, 0.5). Moreover, all of the relay users are located at (0, 1.0), and all of the direct users are located at (0.25, - $\sqrt{3}/4$) on a two-dimensional surface (i.e., $\ell_{ib} = \frac{1}{2}(\ell_{rb} + \ell_{kr})$ and $\ell_{kr} = \ell_{ir}$), and $w_k = 1$, $\forall k \in \mathcal{U}$.

parison, the total transmit powers of the base station during two phases are the same for all schemes. Thus the transmitted power in the first phase of WMMSE-NCU is boosted twice as large as that of WMMSE-WCU since WMMSE-WCU allows BS to transmit signal in the second phase. All of the distances are normalized by that between the BS and the RUs. In Fig 6.5, e_x is constant. This may not be satisfied under the SNR varying condition as e_x has inverse correlation to SNR if the main occurrence factor of CSIT error is thermal noise on the receiver. To reflect the influence of the estimation error, Fig. 6.6 shows the average sum-rate of the network versus estimation error e_x, when the powers at BS and RS are fixed. The proposed scheme with the designed precoding strategy outperforms the other

schemes with different estimation error. Figure 6.7 shows the average sum rate of the network versus the relay position for the case in which the powers at the BS and the RS are fixed. Figure 6.8 shows the average sum rate of the network versus K and versus the numbers of antennas of the BS and the RS, where the number of antennas of the BS is the same as that of the RS.

Figures 6.5 through 6.8 indicate that the proposed scheme with the designed precoding strategy outperforms the other schemes, especially the scheme that does not consider DUs, in the case of either perfect CSIT or imperfect CSIT. This is because schemes that do not consider DUs (i.e., the BS remains silent during the second phase) experience DOF loss, and the maximum number of degrees of freedom is N_b. However, the maximum number of degrees of freedom in the proposed scheme with the designed precoding strategy is $2N_b$. Furthermore, all of the schemes that consider direct users and relay users, i.e., the proposed WMMSE-WCU, WMMSE-MRC&MRT-WCU, and GCI-I-WCU, were applied to the same number of MUs. Figures 6.5 through 6.8 indicate that the proposed WMMSE-WCU scheme performs better than WMMSE-MRC&MRT-WCU or GCI-I-WCU. Schemes that consider non-coordinated users (NCU), such as WMMSE-NCU, only support half of mobile-users. Accordingly, based on Eq. (6.11), it can be concluded that the proposed scheme can better deal with multi-user interferences caused by the RUs and DUs.

6.7.3 Sum Rate and Minimum Rate

The results in Fig. 6.9 show the achievable sum rates and the achievable minimum rate among all users for different schemes with the perfect CSIT and imperfect CSIT. The sum rate and minimum rate of the proposed scheme are higher than those of the other schemes, for the cases of either perfect CSIT or imperfect CSIT, except for the minimum rate of the WMMSE-NCU scheme, because the number of users in the WMMSE-NCU scheme is half that of other schemes, less inter-user interference is introduced.

6.8 Summary

The half-duplex constraint is a major potential weakness for relay techniques because system bandwidth resources are used inefficiently due to the necessity of dedicated bandwidth for relay retransmissions. In the present study, this work proposed a general communication scheme with a precoding design strategy for a MIMO-RBC with coordinated users in order to overcome the relay's HDC and achieve full DoF so as to improve the frequency efficiency. Furthermore, since the problems associated with the throughput and fairness criteria for the robust source

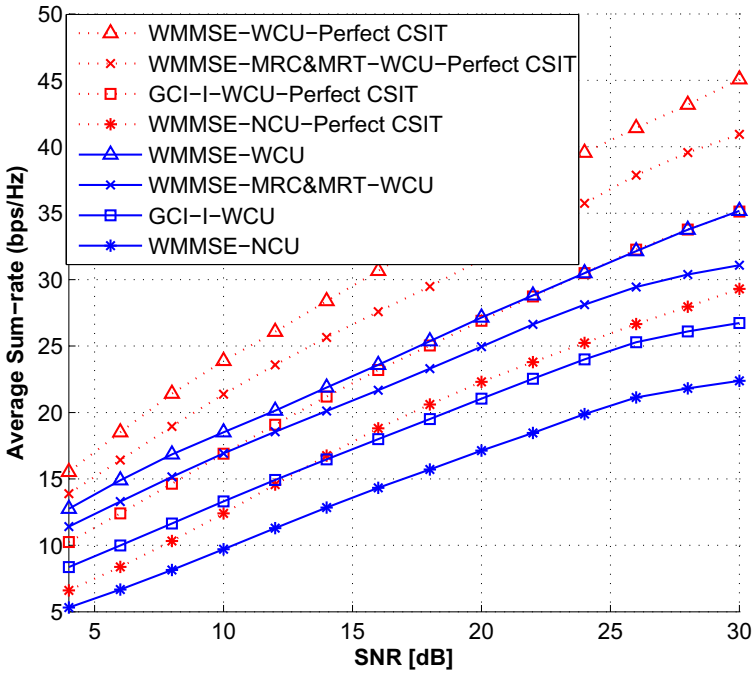

Figure 6.5: Average sum rate versus transmit power with $P_b = P_r$ (SNR dB), where $N_b = N_r = 6$, $K = 6$, $e_x = 0.1$, and $e_x = 0$ (i.e., perfect CSIT). The BS is located at (0, 0), and the RS is located at (0, 0.5). Moreover, all relay users are located at (0, 1.0), and all direct users are located at (0.25, - $\sqrt{3}/4$) on a two-dimensional surface (i.e., $\ell_{ib} = \frac{1}{2}(\ell_{rb} + \ell_{kr})$ and $\ell_{kr} = \ell_{ir}$), and $w_k = 1$, $\forall k \in \mathcal{U}$.

PM and relay BM designs are non-linear and non-convex, this work considered a WMMSE precoding design method by which to jointly design the source PM and relay BM based on an alternating optimization approach. The proposed scheme with the advanced precoding strategy can significantly increase the diversity gain of the MIMO-RBC, and its performance has been verified numerically. As an extension, in the future, this part can intend to investigate a regenerative, full-duplex scheme that considers coordinated users.

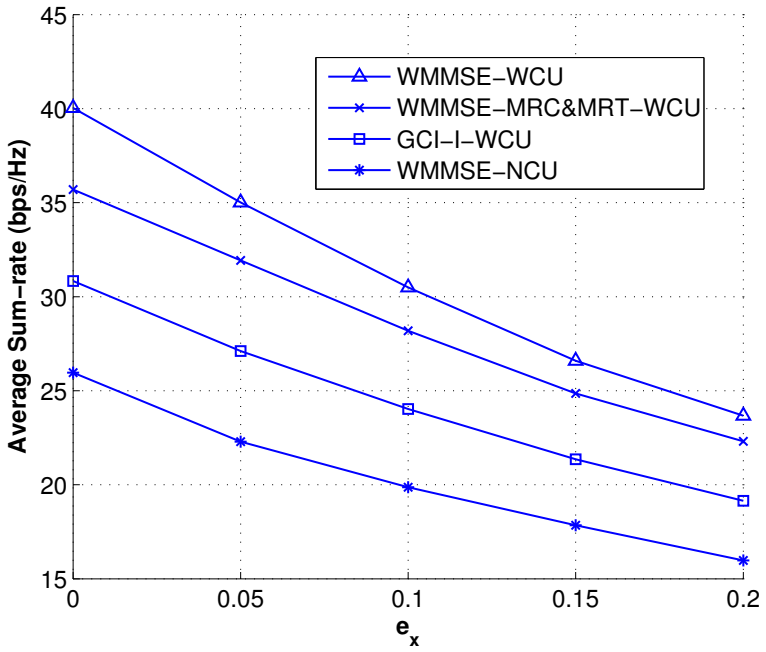

Figure 6.6: Average sum-rate versus the estimation error e_x with $P_b = P_r = 24$ (SNR dB), where $N_b = N_r = 6$, $K = 6$. BS is located at (0, 0), RS is at (0, 0.5), all relay-users are at (0, 1.0) and all direct-users are at (0.25, -$\sqrt{3}$/4) on a two-dimensional surface (i.e., $\ell_{ib} = \frac{1}{2}(\ell_{rb} + \ell_{kr})$ and $\ell_{kr} = \ell_{ir}$), and $w_k = 1$, $\forall k \in \mathcal{U}$.

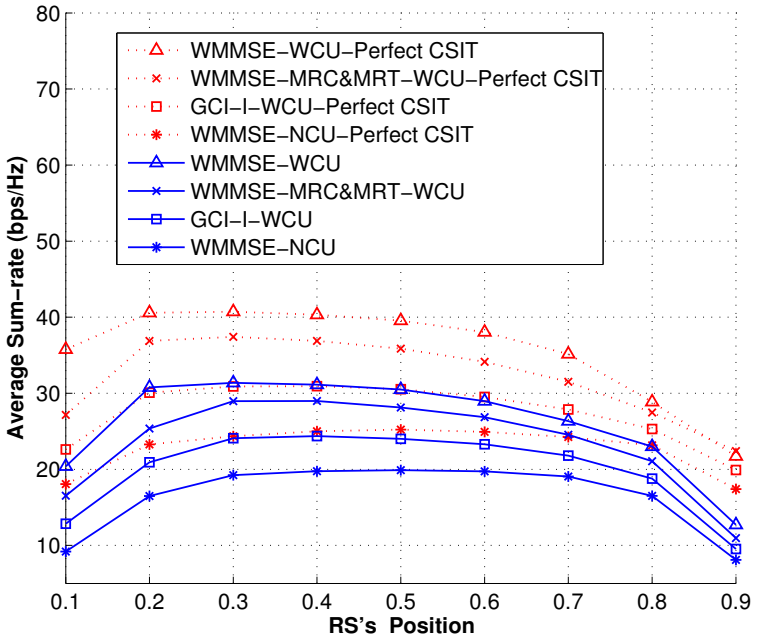

Figure 6.7: Average sum rate versus relay position with $P_b = P_r = 24$ (SNR dB), where $N_b = N_r = 6$, $K = 6$, $e_x = 0.1$ and $e_x = 0$ (i.e., perfect CSIT). The BS is located at (0, 0), and the RS is located at (0, x). Moreover, all relay users are located at (0, 1.0), and all direct users are located at (0.25, - $\sqrt{3}/4$) on a two-dimensional surface (i.e.,$\ell_{ib} = \frac{1}{2}(\ell_{rb} + \ell_{kr})$, where x is in [0.1, 0.9], and $w_k = 1$, $\forall k \in \mathcal{U}$.

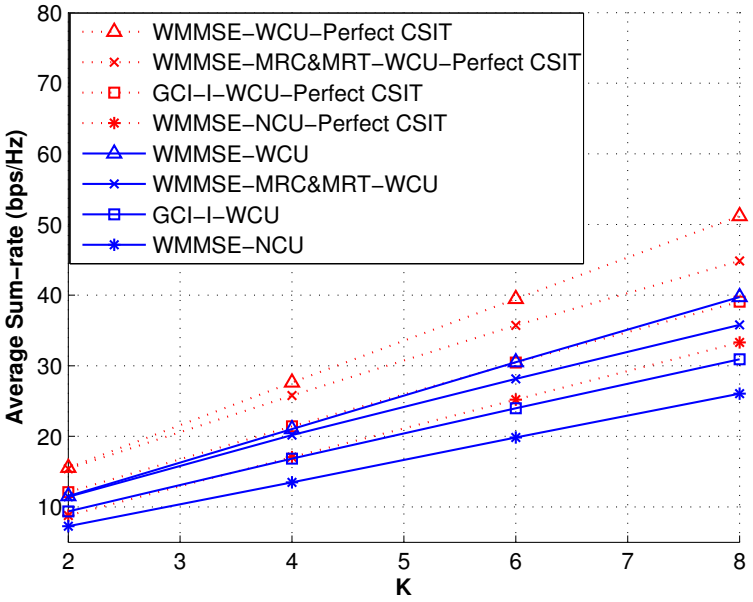

Figure 6.8: Average sum rate versus number of BS antennas, where $N_b = N_r = K$, $N_1 = N_2 = \cdots = N_K = 2$, $e_x = 0.1$, and $e_x = 0$ (i.e., perfect CSIT), and $P_b = P_r = 24$ (SNR dB). The BS is located at (0,0), and the RS is located at (0, 0.5). All relay users are located at (0, 1.0), and all direct users are located at (0.25, - $\sqrt{3}/4$) on a two-dimensional surface (i.e., $\ell_{ib} = \frac{1}{2}(\ell_{rb} + \ell_{kr})$ and $\ell_{kr} = \ell_{ir}$), and $w_k = 1$, $\forall k \in \mathcal{U}$.

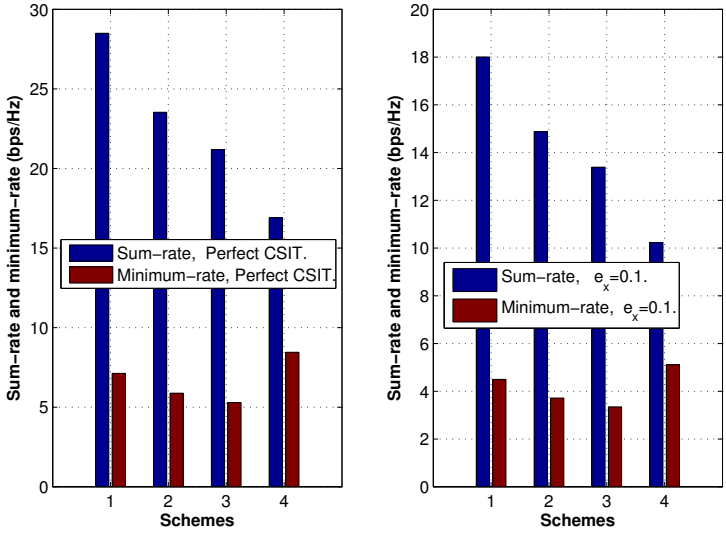

Figure 6.9: Sum rate and minimum rate of various schemes for one (randomly selected) channel realization with $e_{rb} = e_{kb} = e_{kr} = 0$ (i.e., perfect CSIT) and $e_{rb} = e_{kb} = e_{kr} = 0.1$, where $P_b = P_r = 20$ (SNR dB), $N_b = N_r = 4$, and $K = 4$ for the $1^{st}, 2^{nd}, and 3^{rd}$ schemes, and $K = 2$ for the 4^{th} scheme. The BS is located at (0,0), and the RS is located at (0, 0.5). All relay users are located at (0, 1.0), and all direct users are located at (0.25, - $\sqrt{3}/4$) on a two-dimensional surface (i.e., $\ell_{ib} = \frac{1}{2}(\ell_{rb} + \ell_{kr})$ and $\ell_{kr} = \ell_{ir}$), and $w_k = 1$, $\forall k \in \mathcal{U}$. (1 $\triangleq$ WMMSE-WCU, 2 $\triangleq$ WMMSE-MRC&MRT-WCU, 3 $\triangleq$ GCI-I-WCU, and 4 $\triangleq$ WMMSE-NCU.)

Chapter 7

Conclusions and Future Works

7.1 Conclusions

This thesis has mainly addressed the challenges for improving spectrum and energy utilization for wireless networks, and its focus is to promote recognizing radio environment by collecting large-scale and crowdsourcing based spectrum data through service cooperation / interaction between cognitive radio network and heterogeneous cell network. The problems, such as how to construct common control channel between user terminals and data center, how to encourage users to collect data, and how much cooperation gain is obtainable, etc., have been discussed, and system performance has been illustrated by simulation results.

Chapter 2 has focused on real spectrum measurement and constructing the visualized application for SpectrumMap database. Spectrum resource is very limited with rapidly increased wireless mobile devices. Also, spectrum resource is managed and allocated by government organizations, and it is usually difficult to know the real spectrum utilization for public users. In this sense, it is desired to construct the SpectrumMap database, which is a kind of crowdsourcing database and expected to provide spectrum utilization information to public users, and accordingly to incentivize spectrum awareness of public users. The obtained measurement results through practical data collections has shown that spectrum is not fully utilized; even in the crowded Wi-Fi channels observed in downtown areas, spectrum utilization can be further improved.

Chapter 3 has considered spectrum sensing and energy efficient policy for cognitive radio network. The case, where each user performs independent spectrum sensing without any cooperation, has been addressed by the works of [59], [60] and [61] , etc. However these focus primarily on a single-hop wireless access network. Some existing work, i.e., [52], has considered distributed hop relay algorithm which can adapt data transmissions to dynamic spectrum access op-

portunities; however, they have assumed synchronization existed in the network and neighboring secondary users can exchange messages over a common control channel. In addition, these existed studies have obtained the final solutions through direct calculation and the designed algorithms are very complex, not computationally efficient, which is not expected for cognitive radio network when considering energy constrain. Different from the existing studies, this work proposes a new method to effectively reduce the decision space of users, thus can substantially reduce the complexity of decision making by the network. Also, it can achieve very good approximation of the optimal schedule. Furthermore, by adjusting the length of sub-slot operating time for each network user, the proposed algorithm can be extended to multi-hops, and simulation results has confirmed the performance.

Chapter 4 has focused on designing spectrum data collection model in user participation. Network sensing performance of SpectrumMap can be improved if spectrum environment is cooperatively monitored by multiple users. Thus it is expected that more and more users can join the SpectrumMap database by contributing their local data collection through common control channels, and accordingly, this raises the problem how to encourage network users to join the spectrum data collections by designing proper incentive scheme. Therefore, this work has utilized cooperative game theory to formulate users' behavior and proposed effective incentive scheme to promote users' contribution to upload spectrum data to the database center. Especially, it has been proved that the proposed mechanism is more effective in the case that each user has little probability to collect data (i.e., the possibility for users to contribute their data is small). It has also been observed that the incentive effect will be decreased when more users has more probability to collect data.

Chapter 5 has addressed proposing a cooperative spectrum sensing model relying on HetNet infrastructure based on coalitional game theory. Different from the case of Chapter 4, it is supposed that no common control channel existed between cognitive user terminals and the SpectrumMap data center. Thus, to create common control channels, it has been considered that cognitive radio network will ask for the cooperation with HetNet, so as to utilizing the channel resources of HetNet to act the common control channel. Accordingly, it is necessary to promote HetNet user to construct the cooperation with cognitive radio network by useful incentive scheme. Thus, HetNet agrees with cognitive radio network for the cooperation and, as reward, HetNet users can improve their self-organization capabilities by joining the game with cognitive radio users and choosing optimal channels via the sensed channel usage information from cognitive users. As a result, both cognitive radio users and HetNet users can achieve improved throughput, which has been demonstrated by simulation results.

Finally, Chapter 6 tries to improve spectrum efficiency with multiple antenna

technology. HetNet desires to more channel resources since the frequency resources are partly provided to transfer the data for cognitive radio network and HetNet acts as common control channel, in Chapter 5. To improve downlink throughput performance, the existing works, such as [102], has considered to use relay technologies. Compared with the conventional studies with half-duplex constrain that allows no transmission between base station and user terminals during relaying, this work allows the same data streams can be transmitted repeatedly between base station and users, thus can improve the throughput performance. The conducted simulation has confirmed the effectiveness of the proposed scheme.

7.2 Future Works

Currently, spectrum allocation and its utilization is mainly managed by government and official organizations. Throughput the presented work in the thesis, in order to promote spectrum awareness for general public users and let more users to understand spectrum utilization, the grass-root SpectrumMap database has been proposed for providing general information about spectrum utilizations. The SpectrmMap relies on data contributions from various wireless mobile devices at many desired location. To facilitate the large-scale SpectrumMap for cooperative spectrum sensing, this thesis shows a design of incentives based on game theory. It is effective in the cooperation of HetNet and CRN. This framework requires the communications between CRN users and HetNet users while the provider of HetNet is usually different from that of CRN. Without any business cooperation between CRN and HetNet providers, it is difficult for CRN nodes to find this collaboration service and also to set up secure access between a node in HetNet and that in CRN. Thus the collaboration protocol between CRN and HetNet must be designed and standardized.

Chapter 8

Appendix

8.1 Proof for Theorem 1

Proof: Let $A = \max_{U_T(\pi_T) \leq e} G_T(\pi_T)$ and $B = \max_{U_\infty(\pi_T) \leq e} G_\infty(\pi_T)$. Since the policy space

$$\{\pi_T : U_T(\pi_T) \leq e\} \subseteq \{\pi_T : U_\infty(\pi_T) \leq e\},$$

it returns that $A \leq B$. Let $C = \limsup_{T \to \infty} max_{U_\infty(\pi_T) \leq e} G_\infty(\pi_T)$. For $\forall \epsilon_1 > 0$, there exists one $T(\epsilon_1)$ such that when $T > T(\epsilon_1)$,

$$B \leq C + \epsilon_1.$$

There exists one policy π' such that $U_\infty(\pi') \leq e$ and $G_\infty(\pi') = C$. This can be proved as follows: Let $C = \limsup_{T \to \infty} max_{U_\infty(\pi_T) \leq e} G_\infty(\pi_T)$. A subsequence $\{T_k\}$ can be found such that $C = \lim_{T_k \to \infty} max_{U_\infty(\pi_{T_k}) \leq e} G_\infty(\pi_{T_k})$. Within $\{T_k\}$, Further, a new subsequence $\{T_{k'}\}$ can be found such that $\lim_{k' \to \infty} \pi_{T_{k'}} = \pi'$. It can be seen

$$G_\infty(\pi') = \lim_{k' \to \infty} G_\infty(\pi_{T_{k'}}) = C.$$

Thus, $C \leq max_{U_\infty(\pi) \leq e} G_\infty(\pi)$. Therefore, the proof is completed.

8.2 Proof for Theorem 4

Proof: Proof by contradiction. If $\mathbf{Pr}\{G^*_{sink} - G'_{sink} \leq \epsilon\} \neq 1$, then it will take a positive probability as $G^*_{sink} - G'_{sink} > \alpha$. Now observe the part $T^o_{s_2,sink}$. According to the construction of π^2_∞, it can be seen by the Theorem 3, for a random small ϵ_0,

when Y is large enough,

$$\begin{aligned} & \mathbf{Pr}\{G_Y(\pi'_{T^o_{s_2,sink}}) - G_Y(\pi^2_\infty) \le \epsilon_0\} \\ = \ & \mathbf{Pr}\{G_Y(\pi'_{T^o_{s_1,s_2}}) - G'_{sink} \le \epsilon_0\} \\ = \ & 1, \end{aligned}$$

where $\pi'_{T^o_{s_2,sink}}$ represents the optimal solution under the power of B_2. Hence, it can be obtained that

$$\begin{aligned} & G^*_{sink} - G_Y(\pi'_{T^o_{s_2,sink}}) \\ = \ & G_Y(\pi^*_{T^o_{s_2,sink}}) - G_Y(\pi'_{T^o_{s_2,sink}}) \\ > \ & \alpha - \epsilon_0. \end{aligned}$$

Since $\pi^*_{T^o_{s_2,sink}}$ and $\pi'_{T^o_{s_2,sink}}$ represents the optimal solution under the power of $B_{T^o_{s_2,sink}}$ and B_2 respectively, it is easy to know $B_{T^o_{s_2,sink}} > B_2$ according to the previous theorem.

Since node s_2's total power is constant, such that $B_{T^r_{s_1,s_2}} < B_1$. Similarly, it can be obtained that $\beta > 0$, $G_Y(\pi'_{T^o_{s_1,s_2}}) - G_Y(\pi^*_{T^o_{s_1,s_2}}) > \beta$ by using the optimality. In this way,

$$G_Y(\pi^1_\infty) - G_Y(\pi^*_{T^o_{s_1,s_2}}) > \beta - \epsilon_0.$$

The data volume in the node s_2 is balanced, so that there is a contradiction in

$$\begin{aligned} & G_Y(\pi^2_\infty) - G_Y(\pi^*_{T^o_{s_2,sink}}) \\ = \ & G'_{sink} - G^*{}_{sink} \\ > \ & \beta - \epsilon_0. \end{aligned}$$

The proof is completed.

8.3 Obtaining the Upper Bound

According to the section 8.8 of [74], the upper bound is associated with the following problem to maximize the total rewards:

$$\begin{aligned} max \quad & \sum_{s\in\mathcal{S}}\sum_{a\in\mathcal{A}} \widehat{r}(s,a)\widehat{x}(s,a) \\ s.t. \quad & \sum_{a\in\mathcal{A}} \widehat{x}(j,a) - \sum_{s\in\mathcal{S}}\sum_{a\in\mathcal{A}} \widehat{p}(j|s,a)\widehat{x}(s,a) = 0 \ \ (j\in\mathcal{S}), \\ & \sum_{s\in\mathcal{S}}\sum_{a\in\mathcal{A}} \widehat{x}(s,a) = 1, \end{aligned}$$

where $\widehat{x}(s,a) \geq 0, a \in \mathcal{A}, s \in \mathcal{S}, \mathcal{A}$ is decision space, and $\mathcal{S}$ is state space. In the above expression, the $\widehat{p}(j|s,a)$ and $\widehat{r}(s,a)$ is the state transition probability and reward, respectively, which has been given in the section 3.1 of this work and can be calculated according to spectrum occupancy statistics, i.e., the value of a and b. Accordingly, by using linear programming following the instruction of the section 8.8 of the work [74], the value of $\widehat{x}(s,a)$, which indicates the probability of decision of a sensor at each state, can be obtained. That is, the strategy of each sensor can be known now. Then, following the strategy, let sensor to work for T rounds and count the total number of received data packets. Later, the average value of received data packets within one round can be obtained, which is actually the desired value.

8.4 Obtaining the Achievable Rate for MU during Two Phases

The achievable rate of a MIMO channel is an extension of the mutual information formula for a SISO channel. Specifically, the achievable rate is given in terms of the mutual information between the channel input vector $\mathbf{s}_k$ and output vector $\mathbf{y}_k$ as

$$\mathcal{R}_k \leq \mathsf{I}(\mathbf{s}_k;\mathbf{y}_k) = \mathsf{H}(\mathbf{y}_k) - \mathsf{H}(\mathbf{y}_k|\mathbf{s}_k) \tag{8.1}$$

Since this noise $\mathbf{N}_k$ has fixed entropy independent of the channel input, maximizing mutual information is equivalent to maximizing the entropy in $\mathbf{y}_k$. The entropy of $\mathbf{y}_k$ is maximized when $\mathbf{y}_k$ is a zero-mean circularly-symmetric complex Gaussian (ZMCSCG) random vector [119], But $\mathbf{y}_k$ is only ZMCSCG if the input $\mathbf{s}$ is ZMCSCG. This yields

$$\mathsf{H}(\mathbf{y}_k) = \log\det\left|\pi\mathbf{e}(\sum_{i=1}^{K}\mathbf{H}_k\widetilde{\mathbf{P}}_i\widetilde{\mathbf{P}}_i^{\dagger}\mathbf{H}_k^{\dagger} + \mathbf{N}_k\mathbf{N}_k^{\dagger})\right|, \tag{8.2}$$

$$\mathsf{H}(\mathbf{y}_k|\mathbf{s}_k) = \log\det\left|\pi\mathbf{e}(\sum_{i=1,i\neq k}^{K}\mathbf{H}_k\widetilde{\mathbf{P}}_i\widetilde{\mathbf{P}}_i^{\dagger}\mathbf{H}_k^{\dagger} + \mathbf{N}_k\mathbf{N}_k^{\dagger})\right|, \tag{8.3}$$

Then, assuming Gaussian signaling for the source, the kth MU can achieve the rate during two phases as

$$\mathcal{R}_k = \mathsf{H}(\mathbf{y}_k) - \mathsf{H}(\mathbf{y}_k|\mathbf{s}_k) = \log\det\left|\mathbf{I} + \widetilde{\mathbf{P}}_k^{\dagger}\mathbf{H}_k^{\dagger}\mathbf{R}_{\mathbf{I}_k}^{-1}\mathbf{H}_k\widetilde{\mathbf{P}}_k\right|, \tag{8.4}$$

where $\mathbf{R}_{\mathbf{I}_k} = \mathbf{N}_k\mathbf{N}_k^{\dagger} + \sum_{i=1,i\neq k}^{K}\mathbf{H}_k\widetilde{\mathbf{P}}_i\widetilde{\mathbf{P}}_i^{\dagger}\mathbf{H}_k^{\dagger}$.

Bibliography

[1] FEDERAL COMMUNICATION COMMISSION, et al, "Second Report and Order and Memorandum Opinion and Order In the Matter of Unlicensed Operation in the TV Broadcast Bands," Additional Spectrum for Unlicensed Devices Below, 900.

[2] C.-S. Sum, M.-T. Zhou, L. Lu, F. Kojima, and H. Harada, "Performance and coexistence analysis of multiple ieee 802 wpan/wlan/wran systems operating in tv white space," IEEE Dynamic Spectrum Access Networks (DySpan), April 2014.

[3] Y. Zhao, L. Morales, J. Gaeddert, K. K. Bae, J. S. Um; J. H. Reed, "Applying Radio Environment Maps to Cognitive Wireless Regional Area Networks," IEEE Dynamic Spectrum Access Networks (DySpan), April 2007, pp. 115-118.

[4] Ian F. Akyildiz, Brandon F. Lo and Ravikumar Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," Physical Communication 4(1), pp. 40-62, 2011.

[5] M. Wellens and P. Mahonen, "Lessons Learned from an Extensive Spectrum Occupancy Measurement Campaign and a Stochastic Duty Cycle Model," Springer Mobile Networks and Applications, August 2009.

[6] Jones SD, Jung E, Liu X, Merheb N, and Wang IJ, "Characterization of spectrum activities in the U.S. public safety band for opportunistic spectrum access," in Proc. of IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN), pp. 137-146, 2007.

[7] Islam MH, Koh CL, Oh SW, Qing X, Lai YY, Wang C, Liang YC, Toh BE, Chin F, and Tan GL,"Spectrum survey in Singapore: Occupancy measurements and analyses," in Proc. of International Conference on Cognitive Radio Oriented Wireless Networks and Communications (Crown-Com), pp. 1-7, 2008

[8] A. Damnjanovic et al., "A survey on 3GPP heterogeneous networks," IEEE Wireless Commun. Mag., vol. 18, no. 3, pp. 10-21, June 2011.

[9] R. Madan et al., "Cell association and interference coordination in heterogeneous LTE-a cellular networks," IEEE J. Sel. Areas Commun., vol. 28, no. 9, pp. 1479-1489, December 2010.

[10] D. Fooladivanda and C. Rosenberg, "Joint resource allocation and user association for heterogeneous wireless cellular networks," IEEE Trans. Wireless Commun., vol. 12, no. 1, pp. 248-257, 2013.

[11] Gerasimenko M, Moltchanov D, Florea R, et al, "Cooperative Radio Resource Management in Heterogeneous Cloud Radio Access Networks," IEEE Access, 2015, 3, 397-406.

[12] 802.22 Working Group on 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," IEEE, Active Standard, 2011.

[13] 802.11 Wireless Local Area Network Working Group, "802.11af-2013: Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 5: Television White Spaces (TVWS) Operation," IEEE, Active Standard, December 2013.

[14] J. Riihijarvi, P. Mahonen, M. Petrova, and V. Kolar, "Enhancing cognitive radios with spatial statistics: From radio environment maps to topology engine," June 2009, pp. 1-6.

[15] Y. Zhao, B. Le, and J. H. Reed, "Network Support The Radio Environment Map, Cognitive Radio Technology," B. Fette, ed., Elsevier, 2006.

[16] D. Chen, S. Yin, Q. Zhang, M. Liu, and S. Li, "Mining spectrum usage data: a large-scale spectrum measurement study," Proceedings of the 15th annual international conference on Mobile computing and networking, MobiCom '09, New York, NY, USA, pp.13-24, ACM, 2009.

[17] V. Valenta, R. Marsandaandlek, G. Baudoin, M. Villegas, M. Suarez, and F. Robert, "Survey on spectrum utilization in Europe: Measurements, analyses and observations," 2010 Proceedings of the Fifth International Conference on Cognitive Radio Oriented Wireless Networks and Communications, pp.1-5, June 2010.

[18] M. H. Islam et al., “Spectrum survey in Singapore: Occupancy measurements and analyses,” Proc. CrownCom’08, pp. 1-7, May 2008.

[19] Jiantao Xue, Zhiyong Feng,Ping Zhang, “Occupancy Measurements and Analysis in Beijing,” pp. 295-302, International Conference on Electronic Engineering and Computer Science, 2013.

[20] T. Zhang and S. Banerjee, “Inaccurate spectrum databases?: Public transit to its rescue!” In Proc. ACM HotNets, 2013.

[21] T. Zhang and S. Banerjee, “A Vehicle-based Measurement Framework for Enhancing Whitespace Spectrum Databases,” In Proc. ACM MobiCom, pp. 17-28, 2014.

[22] K. Harrison, S. Mishra, and A. Sahai, “How Much White-Space Capacity Is There?” in IEEE Symposium on New Frontiers in Dynamic Spectrum, 2010, April 2010, pp. 1-10.

[23] R. Murty, R. Chandra, T. Moscibroda, and P. Bahl, “SenseLess: A Database-Driven White Spaces Network,” IEEE Transactions on Mobile Computing, 11(2):189-203, Feb. 2012.

[24] Google Spectrum Database. http://www.google.com/get/spectrumdatabase/.

[25] Spectrum Bridge website. http://spectrumbridge.com.

[26] Microsoft Corporation, “White Space Finder,” http://whitespaces.msresearch.us/WSWebGUI/whitespaces.aspx, Mar. 2010.

[27] Zhenjiang Zhang, Wenyu Zhang, Sherali Zeadally, Yanan Wang, and Yun Liu, “Cognitive radio spectrum sensing framework based on multi-agent architecture for 5G networks”. Wireless Communications IEEE, 2015, 22(6), pp. 34-39.

[28] Yao Sun, Yanmin Zhu, Zhenni Feng and Jiadi Yu,“ Sensing processes participation game of smart phones in participatory sensing systems”, 2014 Eleventh Annual IEEE International Conference on Sensing, Communication, and Networking (SECON), IEEE, 2014, pp. 239-247.

[29] J. Riihijarvi and P. Mahonen, “Exploiting spatial statistics of primary and secondary users towards improved cognitive radio networks”, in Proc. 3rd Int. Conf. Cognit. Radio Oriented Wireless Netw. Commun. (CrownCom), May 2008, pp. 1-7.

[30] D. Chen, S. Yin, Q. Zhang, M. Liu, and S. Li, "Mining Spectrum Usage Data: a Large-scale Spectrum Measurement Study", Mobicom 2009.

[31] J. Riihijarvi, P.Mahonen, M.Wellens, andM. Gordziel, "Characterization and modeling of spectrum for dynamic spectrum access with spatial statistics and random fields," in Proc. IEEE 19th Int. Symp. Pers. Indoor Mobile Radio Commun. (PIMRC), Sep. 2008, pp. 1-6.

[32] M. A. McHenry, D. McCloskey, D. Roberson, and J. T. MacDonald, "Spectrum occupancy measurements-Chicago, Illinois," Shared Spectrum Company, Tech. Rep., 2005.

[33] T. Harrold, R. Cepeda, andM. Beach, "Long-term measurements of spectrum occupancy characteristics," in Proc. IEEE Int. Symp. New Front. Dyn. Spectr. Access Netw. (DySPAN), May 2011, pp. 83-89.

[34] V. Atanasovski et al., "Constructing radio environment maps with heterogeneous spectrum sensors," in Proc. IEEE Int. Symp. New Front. Dyn. Spectr. Access Netw. (DySPAN), May 2011, pp. 660-661.

[35] B. Jayawickrama, E. Dutkiewicz, I. Oppermann, G. Fang, and J. Ding, "Improved performance of spectrum cartography based on compressive sensing in cognitive radio networks," in Proc. IEEE Int. Conf. Commun., 2013, pp. 5657-5661.

[36] Kato Takuya, Kusunoki Kei, Kawahara Yoshihiro, Asami Tohru, "An Implementation and Evaluation for Performance Improvement of Spectrum Visualization System," IEICE General Conference, B-7-13, Japan, 2011, pp. 177.

[37] Kato Takuya, Kusunoki Kei, Kawahara Yoshihiro, Asami Tohru, "Performance Improvement of Spectrum Visualization System by Reducing Communications Traffic," IEICE technical report. Information networks, 2011, pp. 85-90.

[38] Y. Li, T. T. Quang, Y. Kawahara, and T. Asami, "Fast White Space Decision for a Tiled Region of a SpectrumMap Based on Spectrum Data Fusion," IEICE General Conference, BS-7-16, Japan, Sept. 2010.

[39] T. T. Quang, Y. Li, Y. Kawahara, and T. Asami, "Statistical Analysis of Spectrum Utilization using Long Term Measurement RF Data," IEICE General Conference, Japan, Sept. 2010, pp.1-2.

[40] Yun Li, Trung Tran Quang, Yoshihiro Kawahara, Tohru Asami, "WiFi Frequency Utilization Analysis by SpectrumMap Database," IEICE Technical Report, Vol. 111, no. 67, IN2011-28, pp. 57-62, Japan, May 2011.

[41] Tan Zhang, Ning Leng, and Suman Banerjee, "A vehicle-based measurement framework for enhancing whitespace spectrum databases," Proceedings of the 20th annual international conference on Mobile computing and networking, ACM, 2014, pp.17-28.

[42] L. Yang, H. Kim, J. Zhang, M. Chiang, and C. W. Tan, "Pricing-based decentralized spectrum access control in cognitive radio networks," IEEE/ACM Transactions on Networking, vol. 21, no. 2, pp. 522-535, 2013.

[43] Xu Chen, Xiaowen Gong, Lei Yang, and Junshan Zhang, "A social group utility maximization framework with applications in database assisted spectrum access," INFOCOM, 2014 Proceedings IEEE, 2014, pp. 1959-1967.

[44] X. Chen and J. Huang, "Database-assisted distributed spectrum sharing," IEEE Journal on Selected Areas in Communications, vol. 31, no. 11, pp. 2349-2361, 2013.

[45] Angela Sara Cacciapuoti, Marcello Caleffi, and Luigi Paura, "Optimal Strategy Design for Enabling the Coexistence of Heterogeneous Networks in TV White Space," IEEE Transactions on Vehicular Technology, 2016, pp. 7361-7363.

[46] Beluri M, Bala E, Dai Y, et al, "Mechanisms for LTE coexistence in TV white space," Dynamic Spectrum Access Networks, IEEE International Symposium on, 2012, pp. 317-326.

[47] IEEE P802.11af, Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications Amendment 2: TV White Spaces Operation, Version D1.0, January 2011.

[48] I.F. Akyildiz, W.Y. Lee, and K.R. Chowdhury, "Crahns: Cognitive radio ad hoc networks," Ad Hoc Networks, vol.7, pp. 810-836, 2009.

[49] Z. Zhang, K. Long, and J. Wang, "Self-organization paradigms and optimization approaches for cognitive radio technologies: a survey," IEEE Wireless Communications, vol.20, pp. 36-42, 2013.

[50] M.A. Alsheikh, D.T. Hoang, D. Niyato, H.P. Tan, and S. Lin, "Markov decision processes with applications in wireless sensor networks: A survey," IEEE Communications Surveys & Tutorials, vol.17, pp. 1239-1267, 2015.

[51] O.B. Akan, O.B. Karli, and O. Ergul, "Cognitive radio sensor networks," IEEE Network, vol.23, pp. 34-40, 2009.

[52] L. Yongkang, L.X. Cai, and S. Xuemin, "Spectrum-aware opportunistic routing in multi-hop cognitive radio networks," IEEE Journal on Selected Areas in Communications, vol.30, no.10, pp. 1958-1968, 2012.

[53] M. Youssef, M. Ibrahim, M. Abdelatif, L. Chen, and A.V. Vasilakos, "Routing metrics of cognitive radio networks: A survey," IEEE Communications Surveys & Tutorials, vol.16, no.1, pp. 92-109, 2015.

[54] P. Spachos and D. Hantzinakos, "Scalable dynamic routing protocol for cognitive radio sensor networks," IEEE Sensors Journal, vol.14, pp. 2257-2266, 2014.

[55] P. Quang and D. Kim, "Throughput-aware routing for industrial sensor networks: Application to isa100. 11a," IEEE Transactions on Industrial Informatics, vol.10, pp. 351-363, 2014.

[56] Z. Liang, S. Feng, D. Zhao, and X. Shen, "Delay performance analysis for supporting real-time traffic in a cognitive radio sensor network," IEEE Transactions on Wireless Communications, vol.10, pp. 325-335, 2011.

[57] L.S. C and C.K. C, "Improving spectrum efficiency via in-network computations in cognitive radio sensor networks," IEEE Transactions on Wireless Communications, vol.13, pp. 1222-1234, 2014.

[58] Ju Ren, Yaoxue Zhang, Ning Zhang, Deyu Zhang and Xuemin (Sherman) Shen, "Dynamic Channel Access to Improve Energy Efficiency in Cognitive Radio Sensor Networks," IEEE Transactions on Wireless Communications, vol. 15, no. 5, pp. 3143-3156, 2016.

[59] Q. Zhao, L. Tong, A. Swami, and Y. Chen, "Decentralized cognitive mac for opportunistic spectrum access in ad hoc networks: A pomdp framework," IEEE Journal on Selected Areas in Communications, vol.25, no.3, pp. 589-600, 2007.

[60] Y.X. Chen, Q. Zhao, and A. Swami, "Distributed spectrum sensing and access in cognitive radio networks with energy constraint," IEEE Transactions on Signal Processing, vol.57, no.2, pp. 783-797, 2009.

[61] P.Y. Ren, Y.C. Wang, and Q.H. Du, "Cad-mac: A channel-aggregation diversity based mac protocol for spectrum and energy efficient cognitive ad hoc networks," IEEE Journal on Selected Areas In Communications, vol.32, no.2, pp. 237-250, 2014.

[62] M. Cesana, F. Cuomo, and E. Ekici, "Routing in cognitive radio networks: Challenges and solutions," Ad Hoc Networks, vol.9, no.3, pp.228-248, 2011.

[63] C. Xin, B. Xie, and C. Shen, "A novel layered graph model for topology formation and routing in dynamic spectrum access networks," IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks (DySPAN), pp. 308-317, 2005.

[64] C. Xin, L. Ma, and C.C. Shen, "A path-centric channel assignment framework for cognitive radio wireless networks," Mobile Networks and Applications, vol.13, no.5, pp. 463-476, 2008.

[65] Y.T. Hou, S. Yi, and H.D. Sherali, "Spectrum sharing for multi-hop networking with cognitive radios," IEEE Journal on Selected Areas in Communications, vol.26, no.1, pp.146-155, 2008.

[66] W. Ren, Q. Zhao, and A. Swami, "Temporal traffic dynamics improve the connectivity of ad hoc cognitive radio networks," IEEE/ACM Transactions on Networking, vol.22, no.1, pp. 124-136, 2014.

[67] W. Ren, Q. Zhao, and A. Swami, "On the connectivity and multihop delay of ad hoc cognitive radio networks," IEEE Journal on Selected Areas In Communications, vol.29, no.4, pp. 805-818, 2011.

[68] C.H. Papadimitriou and J. Tsitsiklis, "The complexity of markov decision processes," IEEE Journal on Selected Areas In Communications, vol.12, pp. 441-450, 1987.

[69] N. Jaggi, K. Kar, and A. Krishnamurthy, "Rechargeable sensor activation under temporally correlated events," Wireless Networks, vol.15, no.5, pp. 619-635, 2009.

[70] E. Altman, Constrained Markov Decision Processes, Chapman & Hall/CRC, 1999.

[71] D.W. Stroock, An Introduction to Markov Processes, Springer, 2005.

[72] J. Ren, Y. Zhang, R. Deng, N. Zhang, D. Zhang, and X. Shen, "Joint channel access and sampling rate control in energy harvesting cognitive radio sensor networks," IEEE Transactions on Emerging Topics in Computing, 2016, to appear, DOI: 10.1109/TETC.2016.2555806.

[73] MICA2 https://www.eol.ucar.edu/isf/facilities/isa/ internal/CrossBow/DataSheets/mica2.pdf.

[74] M. Puterman, Markov Decision Processes: Discrete Stochastic Dynamic Programming, Wiley, 1994.

[75] Kaushik R. Chowdhury, and Ian F. Akyildiz, "OFDM-Based Common Control Channel Design for Cognitive Radio Ad Hoc Networks," IEEE Transactions on Mobile Computing, vol. 10, no. 2, Feburary 2011.

[76] Ian F. Akyildiz, Brandon F. Lo, and Ravikumar Balakrishnan, "Cooperative spectrum sensing in cognitive radio networks: A survey," Physical Communication (Elsevier) Journal, vol. 4, no. 1, pp. 40-62, March 2011.

[77] G. Guandr, S. Bayhan, and F. Alagoandz, "Cognitive femtocell networks: an overlay architecture for localized dynamic spectrum access," IEEE Wireless Commun., vol. 17, no. 4, pp. 62-70, Aug. 2010.

[78] V. Chandrasekhar, J. Andrews, Z. Shen, T. Muharemovic, and A. Gatherer, "Power control in two-tier femtocell networks," IEEE Trans. Wireless Commun., vol. 8, no. 8, pp. 4316-4328, Aug. 2009.

[79] C. D. T. Thai and P. Popovski, "Interference cancellation schemes for uplink transmission in femtocells," in Proceedings of IEEE GLOBECOM 2010, Miami, FL, USA, 30 Nov. 2010.

[80] E. Peh and Y. C. Liang, "Optimization for cooperative sensing in cognitive radio networks," in Proceedings of IEEE Wireless Communications and Networking Conference (WCNC) 2007, pp. 27-32, March 2007.

[81] Z. Khan, J. Lehtomaki, M. Latva-Aho, and L. A. DaSilva, "On Selfish and Altruistic Coalition Formation in Cognitive Radio Networks," 5th Intl. Conf. on Cognitive Radio Oriented Wireless Networks and Communications (CROWNCOM), Cannes, France, June 9-11, 2010.

[82] W. Saad, Z. Han, T. Basar, Hjϕrungnes, and J. B. Song, "Hedonic coalition formation game for secondary base station cooperation in cognitive radio networks," in Proceedings of IEEE Wireless Communications and Networking Conference (WCNC) 2010, pp. 1-6, April 2010.

[83] W. Saad, Z. Han, M. Debbah, A. Hjϕrungnes, and T. Basar, "Coalitional game theory for communication networks: A tutorial," IEEE Signal Processing Magazine, vol. 26, no. 5, pp. 77-97, September 2009.

[84] A. Mas-Colell, M. D. Whinston, and J. R. Green, "Microeconomic Theory," Oxford University Press, 1995.

[85] Y. C. Liang, Y. Zeng, E. C. Y. Peh, and A. T. Hoang, "Sensing-throughput tradeoff for cognitive radio networks," IEEE Transactions on Wireless Communications, vol. 7, no. 4, Article ID 4489760, pp. 1326-1337, 2008.

[86] O. Simeone, I. Stanojev, S. Savazzi, Y. Bar-Ness, U. Spagnolini, and R. Pickholtz, "Spectrum leasing to cooperating secondary ad hoc networks," IEEE J. Sel. Areas Commun., vol. 26, no. 1, pp. 203-213, Jan. 2008.

[87] Walid Saad, Zhu Han, Merouane Debbah, Are Hjϕrungnes, and Tamer Basar, "A distributed coalition formation framework for fair user cooperation in wireless networks," IEEE Trans. on Wireless Commun., vol. 8, no. 9, pp. 4580-4593, Sep, 2009.

[88] W. C. Cheung, T. Q. S. Quek, and M. Kountouris, "Throughput Optimization, Spectrum Allocation, and Access Control in Two-Tier Femtocell Networks," IEEE J. Select. Areas Commun., vol. 30, no. 3, pp. 561-574, Apr. 2012.

[89] S. Hämäläinen, H. Sanneck, and C. Sartori, "LTE Self-Organizing Networks (SON): Network Management Automation for Operational Efficiency," John Wiley & Sons. Ltd., 2012.

[90] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," IEEE Trans. Inf. Theory, vol.52, no.5, pp. 2203-2214, May 2006.

[91] R. Pabst, B.H. Walke, D.C. Schultz, S. Mukherjee, H. Viswanathan, D.D. Falconer, and G.P. Fettweis, "Relay-based deployment concepts for wireless and mobile broadband radio," IEEE Commun. Mag., vol.42, no.9, pp. 80-89, Sept. 2004.

[92] Z. Ding, I. Krikidis, B. Rong, J.S. Thompson, C. Wang, and S. Yang, "On combating the half-duplex constraint in modern cooperative networks: Protocols and techniques," IEEE Wireless Commun., vol.12, no.12, pp. 20-27, Dec. 2012.

[93] I. Krikidis, H.A. Suraweera, S. Yang, and K. Berberidis, "Full-duplex relaying over block fading channel: A diversity perspective," IEEE Trans. Wireless Commun., vol.11, no.12, pp. 4524-4535, Dec. 2012.

[94] B. Yu, L. Yang, X. Cheng, and R. Cao, "Power and location optimization for full-duplex decode-and-forward relaying," IEEE Trans. Commun., vol.63, no.12, pp. 4743-4753, Dec. 2015.

[95] G. Liu, F.R. Yu, H. Ji, V.C.M. Leung, and X. Li, "In-band full-duplex relaying: A survey,research issues and challenges," IEEE Communication Surveys& Turorials, vol.17, no.2, pp. 500-524, sec. 2015.

[96] C.B. Chae, T. Tang, J. Robert W. Heath, and S. Cho, "MIMO relaying with linear processing for multiuser transmission in fixed relay networks," IEEE Trans. Sig. Proc., vol.56, no.2, Feb. 2008.

[97] R. Zhang, C.C. Chai, and Y.C. Liang, "Joint beamforming and power control for multiantenna relay broadcast channel with QoS constraints," IEEE Trans. Sig. Proc., vol.57, no.2, pp. 726-737, Feb. 2009.

[98] W. Xu, X. Dong, and W.S. Lu, "Joint precoding optimization for multiuser multi-antenna relaying downlinks using quadratic programming," IEEE Trans. Commun., vol.59, no.5, pp. 1228-1235, May 2011.

[99] K.S. Gomadam and S.A. Jafar, "Duality of MIMO multiple access channel and broadcast channel with amplify-and-forward relays," IEEE Trans. Commun., vol.58, no.1, pp. 211-217, Jan. 2010.

[100] U. Phuyal, S.C. Jha, and V.K. Bhargava, "Joint zero-forcing based precoder design for QoS-aware power allocation in MIMO cooperative cellular network," IEEE J. Sel. Areas Commun., vol.30, no.2, pp. 350-358, Feb. 2012.

[101] H. Wan, W. Chen, and J. Ji, "Efficient linear transmission strategy for MIMO relaying broadcast channels with direct links," IEEE Wireless Commun. Lett., vol.1, pp. 14-17, Feb. 2012.

[102] H. Wan, W. Chen, and X. Wang, "Joint source and relay design for MIMO relaying broadcast channels," IEEE Commun. Lett., vol.17, no.2, pp. 345-348, Feb. 2013.

[103] J. Kaleva, A. Tolli, G. Venkatraman, and M. Juntti, "Downlink precoder design for coordinated regenerative multi-user relaying," IEEE Trans. Sig. Proc., vol.61, no.5, pp. 1215-1229, March 2013.

[104] B. Zhang, Z. He, K. Niu, and L. Zhang, "Robust linear beamforming for MIMO relay broadcast channel with limited feedback," IEEE Sign. Proc. Lett, vol.17, no.2, pp. 209-212, Oct. 2010.

[105] W. Xu, X. Dong, and W.S. Lu, "MIMO relaying broadcast channels with linear precoding and quantized channel state information feedback," IEEE Trans. Sig. Proc., vol.58, no.10, pp. 5233-5245, Otc. 2010.

[106] Z. Wang, W. Chen, and J. Li, "Efficient beamforming in MIMO relaying broadcast channels with imperfect channel estimations," IEEE Trans. Veh. Tech., vol.61, no.1, pp. 419-429, Jan. 2012.

[107] Y. Cai, R.C. de Lamare, L.L. Yang, and M. Zhao, "Robust mmse precoding based on switched relaying and side information for multiuser MIMO relay systems," IEEE Trans. Veh. Tech., vol.60, no.6, pp. 2608-2619, July 2011.

[108] S.W. Peters, A.Y. Panah, K.T. Truong, and J. R. W. Heath, "Relay architectures for 3GPP LTE-Advanced," EURASIP J.Wireless Commun. Netw., vol.2009.

[109] J.N. Laneman, D. Tse, and G.W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," IEEE Trans. Inf. Theory, vol.50, no.12, pp. 3062-3080, Dec. 2004.

[110] H. Sung, S.R. Lee, and I. Lee, "Generalized channel inversion methods for multiuser MIMO systems," IEEE Trans. Commun., vol.57, no.11, pp. 3489-3499, Nov. 2009.

[111] H.A. Suraweera, I. Krikidis, G. Zheng, C. Yuen, and P.J. Smith, "Low-complexity end-to-end performance optimization in MIMO full-duplex relay systems," IEEE Trans. Wireless Commun., vol.13, no.2, pp. 913-927, Feb. 2014.

[112] A.D. Dabbagh and D.J. Love, "Multiple antenna MMSE based downlink precoding with quantized feedback or channel mismatch," IEEE Trans. Commun., vol.56, no.11, pp. 1859-1868, Nov. 2008.

[113] A. Hjørungens and D. Gesbert, "Complex-valued matrix differentiation: Techniques and key results," IEEE Trans. Sig. Proc., vol.55, no.6, pp. 2740-2746, Jun. 2007.

[114] N.J. Higham, "Accuracy and stability of numerical algorithms, second ed.," Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, MR 1927606, 2002.

[115] Q. Shi, M. Razaviyayn, Z.Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," IEEE Trans. Sig. Proc., vol.59, no.9, pp. 4331-4340, Sep. 2011.

[116] J.C. Bezdek and R.J. Hathaway, "Some notes on alternating optimization," in Advances in Soft Computing-AFSS of Lecture Notes In Computer Science, Springer, Berlin, Germany, vol.2275, no.1, pp. 1-17, Spring 2002.

[117] A. Hjørungens and D. Gesbert, "Hessians of scalar functions of complex-valued matrices: A systematic computational approach," Proc. 9th Int. Sygnal Process. Appl. (ISSPA-2007), 2007.

[118] D. Bertsekas, Nonlinear Programming, Athena Scientific, 1999.

[119] T.M. Cover and J.A. Thomas, Elements of Information Theory, John Wiley & Sons, Inc., 2006.

Publications

Journal

[1] **Yun Li**, Tohru Asami, "A Near-optimal Sensing Schedule for Spectrum Access in Multi-hop Cognitive Radio Network," IEICE Transactions on Communications, vol.E100-B, no.7, Jul. 2017 (to appear).

[2] **Yun Li**, Haibin Wan, Wen Chen, Tohru Asami , "Joint Source and Relay Beamformer Design for General MIMO Relaying Broadcast Channel with Imperfect Channel State Information," IEICE Transactions on Communications, vol.E100-B, no.5, May. 2017 (to appear).

International Conference

[3] **Yun Li**, "Grass-root based SpectrumMap Database for Self-Organized Cognitive Radio and Heterogeneous Networks: Spectrum Measurement, Data Visualization, and User Participating Model," WCNC, IEEE, New Orleans, USA, Mar., 2015.

[4] **Yun Li**, Haibin Wan, Wen Chen, "Robust Linear Beamformer Designs for MIMO Relaying Broadcast Channel with Max-Min Fairness," WCNC, IEEE, New Orleans, USA, Mar., 2015.

[5] **Yun Li**, "A Grass-root based Radio Environment Data Collections in User Participation for improving Frequency Utilization of Cognitive Radio and Heterogeneous Networks," Computing, Networking and Communications (ICNC), IEEE, Anaheim, USA, Feb., 2015.

[6] **Yun Li**, Haibin Wan, "Joint Source and Relay Beamforming Design for General MIMO Relaying Broadcast Channels," Computing, Networking and Communications (ICNC), CNC Workshop, IEEE, Anaheim, USA, Feb., 2015.

[7] **Yun Li**, Honggang Zhang, Tohru Asami, "On the cooperation between cognitive radio users and femtocell networks for cooperative spectrum sensing and self-organization," WCNC, IEEE, Shanghai, China, 2013.

[8] **Yun Li**, T. Tran Quang, Y. Kawahara, T. Asami and M. Kusunoki: "Building a Spectrum Map for Future Cognitive Radio Technology," In Proceedings of ACM Workshop on Cognitive Radio Networks (CoRoNet), ISBN: 978-1-60558-738-7, Beijing, China, Sep. 2009.

IEICE Technical Report

[9] **Yun Li**, Trung Tran Quang,Yoshihiro Kawahara,Tohru Asami, "WiFi Frequency Utilization Analysis by SpectrumMap Database," IEICE Technical Report, Vol. 111, no. 67, IN2011-28, pp. 57-62, May 2011.

IEICE Society / General Conference

[10] **Yun Li**, T. T. Quang, Y. Kawahara, T Asami, : "Fast White Space Decision for a Tiled Region of a SpectrumMap Based on Spectrum Data Fusion," IEICE General Conference, BS-7-16, Osaka, Japan, Sept. 2010.

[11] **Yun Li**, T. T. Quang, Y. Kawahara, T. Asami, M. Kusunoki: "Visualizing Spectrum Utilisation of Tokyo Area," IEICE Society Conference, BS-10-24, Niigata, Japan, Sep. 2009.

[12] **Yun Li**, Yoshihiro Kawahara, Tohru Asami: "Issues for a Cooperative Cognitive Radio Network Construction," IEICE General Conference, BS-4-11, Matsuyama, Japan, Mar. 2009.

[13] **Yun Li**, Ryoichi Shinkuma, and Tatsuro Takahashi: "IFS Control for Improving Fairness of EDCA with Hidden Stations," IEICE Society Conference, B-15-5, Tottori, Japan, Sep. 2007.