Studies on Resource Allocation for OFDMA-based Cellular Cognitive Radio Networks

Hailan Peng

Department of Electronic Engineering

The University of Electro-Communications

A dissertation submitted for the degree of

Doctor of Philosophy

September 2013

Studies on Resource Allocation for OFDMA-based Cellular Cognitive Radio Networks

Approved by the supervisory committee

Chairperson: Associate Professor Takeo Fujii

Member: Professor Yoshio Karasawa

Member: Professor Takeshi Hashimoto

Member: Professor Eiji Oki

Member: Professor Yasushi Yamao

ＯＦＤＭＡに基づくセルラーコグニティブ無線ネットワークのためのリソース割り当てに関する研究

論文概要

無線通信の急速な発展に伴い、周波数不足の問題が深刻になってきている。無線ネットワークの増え続けるサービスとアプリケーションに対応するため、セカンダリユーザ（SU: Secondary User）とプライマリーユーザ（PU: Primary User）との周波数共有を可能にするコグニティブ無線(CR: Cognitive Radio)は、深刻な周波数不足を緩和する有望な技術の１つとして期待されている。CR の端末はダイナミックな無線環境を認知し、その結果に適応することで、無線アクセスの柔軟性を提供できると同時に、限られた周波数資源の利用効率を最大化することができる。CR 端末の主な特徴は、無線ネットワーク環境を認知して、伝送パワー、搬送周波数、変調策略や上位レイヤーのプロトコルパラメータなどのシステムパラメータを適当に調整することである。

CR の基本的な考え方は単純であるが、従来の無線システムと比べて、効率の高い CR ネットワーク（CRN: Cognitive Radio Network）の設計には新たな課題がある。CRN には、システムのアーキテクチャを設計する際の基本課題だけではなく、システムに対する情報理論に基づく解析及び各層でコグニティブ・プロトコルなどの課題も考慮する必要がある。この技術に対して、スペクトルポリシーも含めて、学界と産業界とともに関心が高まっている。PU と SU の共存問題、適応可能な物理層パラメータ設計、リンク適応技術、直交周波数分割多重接続（OFDMA: Orthogonal Frequency Division Multiple Access）コグニティブ無線、超広帯域コグニティブ無線、コグニティブ媒体アクセス制御（MAC: Medium Access Control）プロトコル、スペクトル検出技術、ダイナミックスペクトルアクセス方式などのさまざまのコグニティブ無線通信技術の分野に対する研究は、世界中に広がっている。

従来の無線通信システムに比べて、CRN の資源配分には二つ新たな課題がある。一つは、PU と SU の間の干渉管理である。SU から PU への干渉電力が干渉温度限界以下に保つ必要がある。もう一つは、スペクトルの可用性によるダイナミックな変化への対応である。SU の利用可能なスペクトルは、PU のスペクトルの利用に依存しており、時間的に変化するものとなる。CRN 内の SU の無線リソース割り当ては、システム間の干渉を回避しながらダイナ

ミックに周波数資源を配分する必要があることから難しい課題となっており、PU の性能を保証しつつ、SU の通信品質（QoS: Quality of Service）を高めることが求められる。PU と SU 間のスペクトル共有と周波数資源を効率的に複数 SU に割り当てることを柔軟に実現するため、本研究では、OFDMA に基づくセルラーCRN 上の効率的なリソース割り当てを検討対象とする。OFDMA は、マルチユーザ CRN のような柔軟なネットワークのアクセスのための魅力的な技術である。このネットワークでは、CR 端末である SU はスペクトルオーバレイ（非アクティブの PU 帯域共有）あるいはスペクトルアンダーレイ（全 PU 帯域共有）のどちらかを使用して、適応通信（opportunistic communication）により、PU のプライマリー帯域にアクセスすることができる。

本論文の研究は、セルラープライマリーネットワーク（PN: Primary Network）と共存する単一セル・マルチユーザの CRN から検討始め、前記２つのスペクトル共有方法（スペクトルオーバレイとスペクトルアンダーレイ）により CRN を実現する。この共存システムには、CR セルの SU とプライマリーセルの PU の間の距離が近い可能性があるため、同じサブチャネルを使用している場合、高い同一チャネル干渉が生じることになる。この場合、同一チャネル干渉を回避することが難しい。そこで、効果的にスペクトルを共有するため、コグニティブ基地局（CBS: Cognitive Base Station）は、同一チャネル干渉回避が困難である場合に同じサブチャネルを使用しないようにするか、同一チャネル干渉を制限するために、送信電力を制御する必要がある。しかし、いずれの場合でも、CBS 側は、プライマリースペクトル配分状況や PU の活動状況などのプライマリー情報を保持する必要がある。ここでは、プライマリーからの補助情報を基にした共存アーキテクチャを構築し、プライマリー基地局（PBS: Primary Base Station）は目標信号に対する干渉と雑音の比率（SINR: Signal to Interference plus Noise Ratio）およびシステム中断確率による PU 干渉閾値を決め、電力制御を実施するため、その情報をセカンダリシステムにブロードキャストする。PBS と CBS の距離と PU の干渉閾値情報によって、スペクトルオーバレイあるいはスペクトルアンダーレイのどちらかに適応したスペクトル共有を実現する。さらに、マルチユーザ CRN では、SU のサービル品質を保持しながら、マルチユーザ CRN の総スループットを最大化するために、制約付きの２変数非線形最適化問題（OP: Optimization Problem）を定式化する。この最適化問題を解決するために、(1) クロスレイヤー近似に基づいて MAC 層から物理層までの QoS 制約を変化して、OP を簡略化する。(2) ラグランジュ双対性に基づく技術を使用して、前述の簡略化された OP を解決し、最適な電力とサブチャネルの配分法を見つける。この OP を解決した上で、ダイナミックなサブチャネルと電力配分を実現するためにクロスレイヤーの資源配分と干渉回避を同時に実現するアルゴリズムを提案し、数値解析とコンピュータシミュレーションで有効性を検証する。結果、従来の設計と比較して、本提案アルゴリズムは、スループットを大幅に高めることが可能であり、期待されている PU が必要とする SINR と SU の QoS を同時に保証できる。また、単純なスペクトルオーバレイ共有法と比較して、ハイブリッ

ドスペクトルオーバレイ・アンダーレイ共有法は、極めて高いスペクトル効率が達成できる。

次に、本論文では、マルチセル PN をオーバレイしたマルチセル・マルチユーザシステムの CRN 上の資源配分を研究する。マルチセルの場合、同一チャネル干渉およびセル内干渉の原因のより、単一セルより複雑な適応制御が必要である。そのため、マルチセルの場合、共存アーキテクチャとスペクトル共有方法が非常に重要となる。また、CBS は全ての干渉チャネル情報を取得するのが難しいため、マルチセルの環境では、干渉回避のために分散的な手法が望ましい。そこで、プライマリーからの補助情報を基にした共存アーキテクチャと、セル間のスペクトルオーバレイおよびセル内のスペクトルアンダーレイによる共有方法を提案する。これらのアーキテクチャには、干渉チャネルの評価と電力制御のため、PBS はパイロット信号と干渉閾値を CRN にブロードキャストする。提案するスペクトル共有方法によって、セカンダリー基地局（CBS）はセル内干渉と同一チャネル干渉を容易に回避することが可能となる。本研究の目標は、プライマリーシステムの性能を確保する同時に、スペクトル効率とコグニティブ無線ネットワーク性能（総スループットと SUs の QoS）を最大化することである。PN での干渉閾値を制限条件として、セカンダリセルの間で同一チャネルの干渉を考えるため、セカンダリシステム性能を表すユーティリティ（ペイオフ）関数を定義する。提案する分散型資源配分手法は、プライマリーの性能を保証し、CBS の協力がない状況でも干渉量を考慮しながらセカンダリのスループットを最大化することができる。また、各 SU の瞬時データレートは定められた最小レートよりも大きいことが保証されている。この資源配分問題は、サブチャネル配分と分散型パワー配分ゲーム（DPAG: Distributed Power Allocation Game）の二つのサブ問題に分割することができる。本研究では、この DPAG 問題に対して、唯一のナッシュ均衡が存在することを証明した。さらに、限られた環境で、この DPAG 問題はパレート最適となる。つまり、CBS に対して、ほかの CBS の性能を損なわない前提で、これ以上の性能向上はできないということを確かめた。さらに、シミュレーションによって、提案アルゴリズムは、数回の反復計算で均衡に収束することが確認されている。これらのことから、マルチセルの CRN における提案方式は、大きなオーバーヘッドがなくても効果的に分散リソース分配が可能となる。

このように本論文では、OFDMA に基づくセルラーCRN に適用できる有効なリソース割り当てアルゴリズムと、プライマリーネットワークとセカンダリネットワーク上に実用できる共存アーキテクチャを提案した。また、次世代無線ネットワークにとって、より高いネットワークスループットと QoS の達成は欠かせない要件であり、本論文ではこれら二つの要件を考慮しながら、良好な性能を達成できるアルゴリズムを提案した。本論文の研究では、セルラーコグニティブ無線ネットワークを中心とした検討となっているが、今後の課題として、様々な共存システムにおける有効な資源配分スキームについての検討を行う必要がある。

Abstract

Rapid growth of wireless communications has been worsening the spectrum shortage problem. Cognitive radio (CR) has emerged as a promising technology that can alleviate the severe spectrum shortage problem by making it possible for secondary (unlicensed) users (SUs) to share frequency bands with primary (licensed) users (PUs). A CR transceiver is able to intelligently recognize and adapt itself to the dynamic radio environment to maximize the utilization of the limited radio resources and provide flexibility in wireless access. The key features required to a CR transceiver are awareness of the radio environment and adaptation of system parameters such as transmit power, carrier frequency, modulation strategy, and higher-layer protocol parameters.

Even though the basic idea of CR is simple, efficient design of CR networks (CRNs) imposes new challenges compared to the conventional wireless networks. It is necessary to consider not only the problems in designing network architectures but also the information-theoretic analysis of cognitive radio networks and cognitive protocols. There is an increasing interest on CR technology among the researchers in academia, industry and the spectrum policy makers. A rich set of research works on cognitive radio wireless communication networks includes co-existence issues, adaptive physical layer protocols, link adaptation techniques, orthogonal frequency division multiple access (OFDMA) and ultra wide band (UWB) based cognitive radio, cognitive medium access control (MAC) protocols, spectrum sensing, dynamic spectrum access methods and so on.

In wireless communication networks, the field of resource allocation is a versatile area that covers a broad range of issues. For example, resource allocation across various network layers encounters different design constraints and parameters; different networking scenarios have different performance goals and service objective; and different formulations of resource allocations need to employ different optimization tools. The basics of resource allocation involve power control, rate adaptation, multiple access and spectrum access, and cross-layer optimization. Compared to the conventional wireless communication networks, two new issues emerged for resource allocation in CRNs. The first one is the interference management between PUs and SUs. The interference power to the PUs should be kept below primary interference temperature limits. Another one is the dynamic spectrum availability. The available spectrum for SUs depends on the spectrum utilization of PUs and is time-varying. Moreover, to flexibly implement spectrum sharing between PUs and SUs, coexistent architectures and spectrum sharing methods of these two networks also have significant importance.

In this dissertation, we study coexistent architectures of primary & secondary networks, different spectrum sharing methods and resource allocation for OFDMA-based cellular CRNs. OFDMA is an attractive access technology for such flexible networks because it is possible for CRNs to implement dynamic spectrum allocation, which is considered in the proposals proposed in this dissertation. Usually, the SUs with CR capability can access the primary bands using either spectrum overlay sharing (non-active PU bands sharing) or spectrum underlay sharing (whole PU bands sharing) for communications. We compare these two sharing methods in Chapter 4, and it shows that the latter one can obtain better performance. Hence, it is better to utilize the spectrum underlay or hybrid spectrum overlay/underlay sharing methods in the CRNs. The studied resource allocation problems in this dissertation mainly focus on interference management between PUs and SUs. Since we assume the primary spectrum utilization is static during one scheduling time (i.e., one resource allocation). Therefore, in our studies, dynamic spectrum availability is not considered. However, it is not difficult to add this dynamic feature if we assume arbitrary distributions for PUs' activities.

First, we consider a single-cell multi-user CRN, which coexists with a cellular primary network (PN). In the coexistent system, the SUs in the CR cell may be near to the PUs in the primary cell. In this situation, it is difficult to avoid co-channel interference because it is too high when SUs and PUs use the same subchannels. In order to share the spectrum effectively, the cognitive base station (CBS) needs to avoid using the same subchannels if the co-channel interference is hard to be avoided, or to control the transmission power on these subchannels to limit the co-channel interference. No matter which situation it is, primary spectrum information is necessary for the CBS. Therefore, we develop a primary-assistance based coexistent architecture, where the primary base station (PBS) broadcasts the interference margins at PUs according to its target signal to interference plus noise ratio (SINR) and outage probability, to the secondary network for power control. Here, the sharing method is either spectrum underlay or spectrum overlay based on the distance between the PBS and the CBS and the interference margins at PUs. In the multi-user CRN, to provide the SUs with satisfactory quality of service (QoS), and to optimize the sum rate of the CRN as well, a constrained two-variable nonlinear optimization problem (OP) is formulated. We solve this OP by (i) transforming the QoS constraints from MAC layer to physical-layer based on a cross-layer approximation to simplify the OP and (ii) using the Lagrangian duality based technique to solve the simplified OP and find the optimal power and subchannel allocation. A joint cross-layer resource allocation and interference avoidance algorithm is proposed for dynamic subchannel and power allocation. The effectiveness of the proposed algorithm is verified by numerical analysis and computer simulations. Simulation results show that, compared to the conventional designs, our algorithm achieves significantly higher throughput and can guarantee the required SINR for PUs and the QoS for SUs. Moreover, compared to the spectrum overlay sharing method, the hybrid spectrum underlay & overlay sharing can provide substantially higher spectrum efficiency.

Next, we study resource allocation for a multi-cell multi-user CRN, which coexists with a multi-cell PN. Due to the co-channel interference and the inter-cell interference, the multi-cell case is more complicated than single-cell. It is difficult for the CBSs to obtain all the interference channel information. Therefore, in the multi-cell environment, distributed operation is preferred for interference avoidance. To manage the coexistence, a primary-willingness based coexistent architecture and a novel intra-cell spectrum overlay and inter-cell spectrum underlay sharing method are proposed. In this architecture, the PBSs are assumed to broadcast pilot signals and interference margins to assist the CRN for interference channel evaluation and power control. Our objective in this study is to guarantee primary performance (i.e., received SINR) and, at the same time, maximize both spectrum efficiency and network performance (i.e., total throughput, and QoS of SUs). Subject to the interference margins imposed by the PN, we introduce a utility (payoff) function that can represent the secondary system performance while taking into account the co-channel interference among the secondary cells. A distributed resource allocation scheme is devised to guarantee the primary performance and, at the same time, maximize the secondary utility function without cooperation among CBSs. QoS among SUs is considered in this proposal so that the instantaneous data rate for each SU is larger than a given minimum rate. The resource allocation problem is decomposed into two subproblems: subchannel allocation and distributed power allocation game (DPAG). We prove that there exists a Nash equilibrium in the DPAG and the equilibrium is unique. Moreover, we prove that the DPAG is also Pareto optimal if the transmission power from CBSs to SUs is limited, that is, no CBS can further improve its performance without impairing others. Through simulations, the effectiveness of our algorithm is shown. The solution turns out to converge within a small number of iterations. The QoS for SUs also can be satisfied well. Compared to the centralized algorithm, the proposed distributed algorithm has great advantages, i.e., good system performance without large signaling overhead and without coordination among CBSs.

This dissertation provides effective subchannel and power allocation algorithms that are suitable for OFDMA-based cellular CRNs, and devises practical coexistent architectures for the overlaid/mixed primary and secondary networks. For the next generation wireless networks, high network throughput and QoS are important requirements. Our algorithms consider these two aspects and can achieve good performance. Moreover, our researches provide some new directions for the coexistence of primary and secondary networks. However, in this dissertation, only cellular CRNs are focused. There still exist many other coexistent scenarios and other directions on design of effective resource allocation schemes and coexistent architectures for the next generation wireless networks.

Contents

List of Figures **vii**

List of Tables **ix**

Acronyms **xi**

1 Introduction **1**

1.1 Background 1

1.2 Scope 3

1.3 Resource Allocation Problems in Cognitive Radio Networks 4

1.4 Motivation and Objectives 6

1.5 Overview of the Dissertation 8

2 Resource Allocation for Wireless Networks **11**

2.1 Introduction 11

2.2 Basics of Resource Allocation 12

2.2.1 Power Control 14

2.2.2 Rate Adaptation 17

2.2.3 Multiple Access and Spectrum Access 19

2.2.4 Cross-Layer Designs 21

2.3 Optimization Formulation of Resource Allocation Problems 22

2.3.1 Constrained Optimization 23

2.4 Optimization Techniques 25

2.4.1 Mathematical Programming 26

2.4.2 Integer/Combinatorial Optimization 27

2.4.3 Game Theory 28

2.5 Previous Resource Allocation Algorithms for the Traditional OFDM-based Wireless Networks 30
2.6 Chapter Summary 32

3 Cognitive Radio Networks 35
3.1 Cognitive Radio Technology 35
3.2 Cognitive Radio Networks and its Applications 37
3.3 Fundamental Issues in Cognitive Radio Networks 40
3.4 Coexistence and Spectrum Sharing 42
3.5 Previous Resource Allocation Algorithms for OFDM-based CRNs 44
3.6 Chapter Summary 46

4 Cross-layer Resource Allocation with QoS Support for Multi-user Cognitive Radio Networks 49
4.1 Introduction 50
4.2 Related Works 51
4.3 System Model and Interference to SUs 52
4.3.1 OFDMA-based Multiuser Cognitive Radio Network 52
4.3.2 Physical Layer and Propagation Model 53
4.3.3 Interference to SUs 54
4.4 Constraints and Problem Definition 55
4.4.1 Power Limits for SUs 55
4.4.2 QoS Constraints for SUs 58
4.4.3 Optimization Problem 59
4.5 Joint Cross-layer Resource Allocation and Interference Avoidance with QoS Support 60
4.5.1 Transform the MAC-layer QoS Constraints to PHY-layer . 61
4.5.2 QoS-guaranteed Resource Allocation Algorithm 64
4.6 Simulation Results 66
4.6.1 Effectiveness of the Proposed Design 68
4.6.2 Performance Comparision between Spectrum Underlay and Overlay 70
4.7 Chapter Summary 73

5 Distributed Resource Allocation for Multi-cell Cognitive Radio Networks **77**
5.1 Introduction 78
5.2 Related Works 79
5.3 System Model and Problem Description 80
5.3.1 System Model and Assumptions 80
5.3.2 Problem Description 83
5.4 Power Control and Interference Avoidance 85
5.5 Distributed Resource Allocation 87
5.5.1 Subchannel Allocation 88
5.5.2 Distributed Power Allocation Game 90
5.5.3 Uniqueness of Nash Equilibrium Point 93
5.5.4 Pareto Optimality 95
5.6 Performance Evaluation 97
5.6.1 Simulation Results of Three-cell Case 98
5.6.2 Simulation Results of Seven-cell Case 103
5.7 Chapter Summary 104

6 Conclusions **109**
6.1 Contributions and Discussions 109
6.2 Future Work 111

A Proof of Proposition 2 in Chapter 4 **115**

B Proof of Solution S^* in Chapter 4 **121**

References **127**

Publications **135**

List of Figures

1.1 Active subbands of PUs, spectrum holes, and OFDM subcarriers of SUs. 3
1.2 Two kinds of spectrum sharing methods: (a) Spectrum overlay; (b) Spectrum underlay. 4

2.1 Wireless standards comparison. 13
2.2 Example of multi-cell case. 15
2.3 Example of single-cell case. 16

3.1 Spectrum sharing process. 36
3.2 An example of centralized cognitive radio networks. 37
3.3 An example of distributed cognitive radio networks. 38
3.4 Basic cognitive cycle. 41
3.5 Dynamic spectrum access. 42

4.1 An example of system model. 52
4.2 Active subbands of PUs and OFDM subcarriers of SUs 53
4.3 The minimum QoS-guaranteed power per time slot when K=3, τ =1, 3, 5, 8 62
4.4 Cumulative probability of SINR of PU 1, R^{RT}=600kbps 69
4.5 Cumulative probability of the reciprocal of relative interference of PU 1 in dB: $10\log_{10}(I_n^{max}/I_{siml})$, R^{RT}=600kbps 70
4.6 QoS support for RT SUs 71
4.7 Average throughput for NRT SUs 72
4.8 Sum rate of CRN 73

4.9 Cumulative probability of the value exceeded the QoS requirements λ_k, R^{RT}=600kbps . . . 74
4.10 Maximum number of SUs with different sharing schemes, R^{RT}=600kbps 75
4.11 Sum rate of CRN with different sharing schemes, R^{RT}=600kbps . 75

5.1 System model. . . . 81
5.2 Available subchannels for P-cells and CR-cells. . . . 82
5.3 Sum data rate per CR-cell vs. the number of iterations. . . . 99
5.4 Transmission power allocation for the same subchannels in CR-cells when P_{PU}=10dBm. . . . 100
5.5 Transmission power allocation for the same subchannels in CR-cells when P_{PU}=25dBm. . . . 101
5.6 Instantaneous data rate of cell-edge SU in CR-cells vs. scheduling time. . . . 102
5.7 Average sum data rate per CR-cell vs. R^k_{min}. . . . 103
5.8 Average sum data rate per CR-cell vs. P^m_{max}. . . . 104
5.9 Average sum data rate of CRN vs. R^k_{min}. . . . 105
5.10 Average sum data rate of CRN vs. P^m_{max}. . . . 106
5.11 Outage probability of PN. . . . 107

List of Tables

4.1 SIMULATION PARAMETERS 67

5.1 SIMULATION PARAMETERS 98

ACRONYMS

Acronyms

3GPP 3rd Generation Partnership Project

AWGN Additive white Gaussian noise

BER Bit error rate

CBS Cognitive base station

CCI Co-channel interference

CDMA Code-division multiple access

CR Cognitive Radio

CR-cell Cognitive radio cell

CRN Cognitive radio network

DCA Dynamic channel allocation

DPAG Distributed power allocation game

DRA Dynamic resource allocation

DSA Dynamic spectrum access

DSL Digital subscriber line

FCC Federal Communications Commission

FCA Fixed channel allocation

FDD Frequency-division duplexing

FDMA Frequency-division multiple access

ACRONYMS

FHMA Frequency-hopped multiple access

HCA Hybrid channel allocation

ISM Industrial, scientific and medical

IWF Iterative water-filling

KKT Karush-Kuhn-Tucker

LP Linear programming

LTE-Advanced Long Term Evolution Advanced

MAC Medium Access Control

MCR Multiuser Cognitive Radio

MI Mutual interference

NE Nash equilibrium

NLP Nonlinear programming

NP-hard Nondeterministic-polynomial-hard

NRT Non-real-time

OFDM Orthogonal frequency division multiplexing

OFDMA Orthogonal frequency division multiple access

OP Optimization problem

OSI Open system interconnection

PBS Primary base station

P-cell Primary cell

PDF Probability density function

PF Proportional fairness

PHY Physical

PN Primary network

PSD Power spectral density

PU Primary user

QoS Quality of service

RA Resource allocation

RF Radio frequency

RT Real-time

SDMA Space-division multiple access

SDR Software-defined radios

SINR Signal to interference plus noise ratio

SU Secondary user

TCP Transmission Control Protocol

TDD Time-division duplexing

TDMA Time-division multiple access

U-NII Unlicensed National Information Infrastructure

UWB Ultra wide band

WiMAX Worldwide interoperability for microwave access

WLAN Wireless local area networks

WPAN Wireless personal area networks

WMAN Wireless metropolitan area networks

ACRONYMS

Chapter 1

Introduction

This dissertation represents studies on resource allocation algorithms for OFDMA-based cellular cognitive radio networks (CRNs). In this chapter, the research background and scope are introduced first. Then, the problems existed in resource allocation of CRNs are reviewed. Research motivation and objectives are also presented. Finally, the overview of the dissertation is given with the description of each chapter.

1.1 Background

Recently, advances in reconfigurable hardware have paved the way to the flexible radios (or software-defined radios (SDR) [1]) that can adapt their air interface and communication protocol to use existing standards or access technologies. The merits for users are twofold. First, it is now possible to use lots of applications, relying on different wireless communication techniques, in a single portable device. Second, users can now seamlessly and opportunistically roam across various wireless access networks in the search for more throughput or cheaper bandwidth.

However, the realization of true seamless handover requires a tight coupling of the hardware flexibility with the protocol layers. Intelligent schemes for environment awareness, hand-off and distributed Quality-of-Service (QoS) control are preferable. The combination of flexibility and increased protocol intelligence has recently led to the novel concept of cognitive radio (CR) that adapts terminals to the current environment and spectrum use [1]. It was first presented by Mitola [2] as a novel wireless communications approach with the ability to sense the external radio environment, to learn from its history, and to make intelligent decisions in adjusting its transmission parameters based on the current environment. It is an innovative technology to exploit the available flexibility leveraging reconfigurability, increased awareness and intelligent control. In a more restricted definition, cognitive devices use their increased flexibility and awareness to control channel access dynamically, i.e., to achieve a dynamic spectrum access (DSA).

With the ever increasing demand for mobile and wireless applications, the static assignment of radio resources to licensed holders has become a limiting factor in efficient spectrum utilization. In many countries, there is little spectrum left for exclusive allocation [3]. However, studies have shown that a large portion of the assigned spectrum is used only sporadically, and the report from the Federal Communications Commission (FCC) has shown that most of the licensed spectrum is currently under-utilized [4].

Recognizing the special ability of CR, to alleviate the looming spectrum shortage problem, the FCC has suggested the use of CR technology to allow secondary users (i.e., unlicensed users) to share radio resources with primary users (i.e., licensed users) while not unduly interfering with them [5]. It is an excellent candidate for improving spectrum utilization.

The underutilized frequency bands of the radio spectrum, legally owned by primary users (PUs), are referred to as *spectrum holes*, which are formally defined as [6]:

"A spectrum hole is a band of frequencies owned by a primary network, but at a particular time and specific geographic location, the band is not being utilized by any primary user."

Orthogonal frequency division multiplexing (OFDM) is a frequency division multiplexing scheme that uses a large number of closely spaced orthogonal subcarriers to carry data. It has been considered as an appropriate modulation technology for cognitive radio networks [7] because of its high spectral efficiency and low interference between adjacent subcarriers. Furthermore, the orthogonal frequency division multiple access (OFDMA) technique is also an attractive

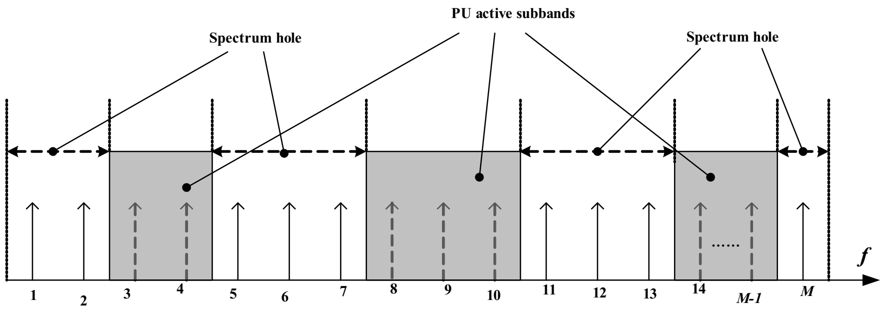

Figure 1.1: Active subbands of PUs, spectrum holes, and OFDM subcarriers of SUs.

access scheme for CRNs because it enables CRNs to allocate radio resources to multiple users dynamically and efficiently [7].

Figure 1.1 shows an example of the spectrum in a typical OFDM-based CRN. The frequency bands that are currently used by PUs are the shadowed areas. The remaining areas, not occupied by the PUs, are *spectrum holes* at this time. When CRNs only utilize the spectrum holes of primary networks (PNs), we refer this sharing method to *spectrum overlay sharing* or *protective sharing*. When both spectrum holes and PU active subbands are shared by CRNs, it is referred to as *spectrum underlay sharing* or *aggressive sharing*.

In this dissertation, the studies are about cognitive radio technology and OFDMA-base cognitive radio networks.

1.2 Scope

To implement CR, a wide range of tasks are involved [6], [8], which include the challenges in radio-scene analysis (information-theoretic analysis of the systems, radio environment estimation and spectrum detection), channel identification (channel-state information estimation and channel capacity prediction), dynamic spectrum management and transmission power control, and system architecture design. The fundamental issues in CRNs are discussed in detail in Sect. 3.3. In this dissertation, we focus on the last two tasks and aim to design practical coexistent architectures and efficient resource allocation (RA) algorithms for OFDMA-based cellular CRNs.

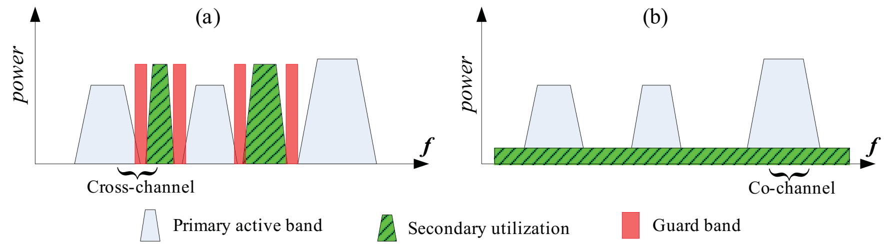

Figure 1.2: Two kinds of spectrum sharing methods: (a) Spectrum overlay; (b) Spectrum underlay.

1.3 Resource Allocation Problems in Cognitive Radio Networks

Although a CRN is allowed to share primary bands to maximize the spectrum efficiency, the utilization by secondary users (SUs) should not cause degradation of services for the PUs. Thus, SUs should monitor and keep the generated interference to PUs to an acceptable level. This level is referred to as the interference temperature limit by the FCC Spectrum Policy Task Force [4]. The definition of the interference temperature limit for a PU is a maximum allowed level of interference power. SUs can use PU bands as long as the total interference power to the PUs is kept below this limit.

Compared to the conventional wireless communication systems, for the RA in CRNs, two new issues arise, namely,

- The interference power to the PU bands should be kept below the interference temperature limit. So, the interference management is more complicated in the coexistent primary and secondary networks, compared to the traditional wireless communications.

- The available spectrum, depending on the spectrum utilization of PNs, is time-varying in CRNs. Good quality-of-service (QoS) should be provided to SUs in spite of the time-varying available spectrum. So, effective resource allocation algorithms are necessary.

At first, we analyze the interferences in the coexistent system. As we have introduced before, there are two kinds of spectrum sharing methods: spectrum overlay sharing and spectrum underlay sharing, as shown in Figure 1.2. In spectrum sharing, two types of interferences will be generated by SUs to PUs. One is the cross-channel interference from the adjacent channels used by SUs (Figure 1.2 (a)), and the other is the co-channel interference generated by SUs using the PU active frequency bands (Figure 1.2 (b)).

For an OFDM based CRN, due to the orthogonality, inter-carrier interference among SUs can be ignored. However, since PNs may not be using OFDM, there could be cross-channel interference between PU bands and SU subcarriers. To avoid the cross-channel interference, appropriate guard bands can be utilized between adjacent channels, as shown in Figure 1.2 (a). To avoid large co-channel interference, power control is necessary at secondary transmitters on PU active subbands. Hence, when designing RA algorithms for CRNs, if the spectrum underlay sharing method is implemented, not only the cross-channel interference but also the co-channel interference need to be considered. For the spectrum overlay sharing, the interference from SUs to PUs is unnecessary to be considered. So, the overlay sharing model can greatly simplify RA design in CRNs.

Even though interference to PUs does not need to be considered in spectrum overlay sharing, it does not mean that RA schemes designed for the conventional OFDM systems can be applied directly to OFDM-based CRNs. In a CRN, besides the fading characteristics of transmission channels, the available transmission spectrum also changes over time. RA schemes designed for the conventional OFDM or OFDMA systems assumed fixed available spectrum, which is not the case in CRNs. Thus, new RA algorithms that take into account both the fading characteristics of the transmission channels and the time-varying spectrum sharing are needed.

In OFDMA based CRNs, power and spectrum allocation results are different in each time slot. So the resource allocation can be implemented one time slot by one time slot. Moreover, although the available spectrum is time-varying in CRNs during a long time, it can be assumed to be fixed during a very short time, i.e., one time slot. In this dissertation, we assume that the PUs' activities are static during one resource allocation for simplicity. Therefore, even though the available spectrum is time-varying for a long period, it can be assumed to be fixed in a short time. It is also not difficult to consider this dynamics in resource allocation problems of CRNs if we assume arbitrary distributions for PUs' activities. Generally, the PUs' activities can be characterized by the Markov chain model.

When SUs share both spectrum holes and PU active subbands, the achievable capacity of CRNs is higher than the use of spectrum hole only. So, it can achieve higher spectrum utilization by spectrum underlay sharing. However, in some practical situations, underlay sharing may not be possible. This can happen, for example, when the CRN is co-located with a broadcast PN, in which there are so many primary receivers that it is impossible to keep the interference power below the specified interference limit at every primary receiver. In such situations, PU active bands may not be shared in order to avoid excessive co-channel interference. In a fading environment, however, a SU signal may undergo deep fading and be received with very little power at the primary receiver. So, for SUs in fading environments, it may be possible to share both the spectrum holes and PU active frequency bands opportunistically if the received interference power at the PU is below the specified interference limit.

Note that RA algorithms designed for spectrum underlay sharing systems also can be applied to spectrum overlay sharing systems by setting the PU interference power threshold at each active primary receiver equals to 0. However, due to the simplicity of the overlay sharing, many studies have been done for spectrum overlay sharing CRNs, and only few researches focused on the spectrum underlay sharing or hybrid spectrum overlay/underlay sharing. Thus, new RA algorithms for spectrum underlay or hybrid spectrum overlay/underlay sharing systems are needed.

The QoS for SUs is also important. In a CRN, the available resources are limited and depend on the utilization of PNs. To provide good QoS for SUs, effective RA algorithms should be devised.

1.4 Motivation and Objectives

The overall objective of this research is to enable spectrum sharing between primary users and secondary users and design effective RA algorithms by using both spectrum overlay and spectrum underlay sharing methods for OFDMA-based cellular CRNs.

Until now, there are many works based on the two spectrum sharing methods separately. They have not been considered jointly. For the spectrum overlay sharing, it is unnecessary to consider the co-channel interferences. However, for the spectrum underlay sharing, both cross-channel and co-channel interferences need to be considered. To ensure the PUs' normal operation, in the hybrid spectrum overlay/underlay sharing systems, the total interference power to each PU has to be kept below a specified interference power threshold. Therefore, transmission power control has to be taken into account in RA, especially when the PUs do not use OFDM.

In this dissertation, our objectives are as followings:

For the single-cell CRN:

- Different coexistent architectures for CR-cell and primary cell are considered. The performance of the two sharing methods is evaluted for different architectures in order to find the optimal sharing method.
- To devise efficient RA algorithms to allocate subchannels and power for multiple SUs in OFDMA-based single-cell CRN, which share both the spectrum holes and PU active frequency bands with PUs, while guaranteeing that the total generated interference power at PUs does not exceed the specified interference margins.
- The QoS support for different SUs is guaranteed in a fading environment with time-varying spectrum.

For the multi-cell CRN:

- A novel intra-cell spectrum overlay and inter-cell spectrum underlay sharing method is proposed for the coexistence of CRN/PN.
- A QoS-guaranteed distributed resource allocation algorithm for a multi-cell CRN is designed by using the proposed intra-cell spectrum overlay and inter-cell spectrum underlay sharing method.

1.5 Overview of the Dissertation

This dissertation includes our research works on resource allocation algorithms for cellular cognitive radio networks. The dissertation consists of six chapters as described below.

Chapter 1 introduces the research background, scope, resource allocation problems in CRNs, research motivation and objectives.

Chapter 2 This chapter presents the basics of resource allocation methods, optimization techniques and related previous resource allocation algorithms for the traditional wireless networks.

Chapter 3 The basics of the CR technology and CR networks with applications are described. Moreover, the fundamental challenges and issues in CRNs and the previous RA algorithms for OFDM-based CRNs are also introduced.

Chapter 4 A joint cross-layer resource allocation and interference avoidance algorithm with QoS support for a multi-user cognitive radio network is presented. In this chapter, the performances of both primary networks and secondary networks are considered jointly. A constrained two-variable nonlinear optimization problem (OP) is formulated. A cross-layer design and convex optimization methodology are utilized to achieve the objectives in this chapter.

Chapter 5 A distributed resource allocation algorithm is proposed for a multi-cell cognitive radio network. In the algorithm, the resource allocation problem is decomposed into subchannel allocation and distributed power allocation game (DPAG). The DPAG is explained in details in this chapter. The existence and uniqueness of nash equilibrium point, and pareto optimality of DPAG are proved.

Chapter 6 The proposed RA algorithms designed for cellular CRNs are summarized. The main contributions of the dissertation and suggestions for future research work are also presented.

Chapter 2

Resource Allocation for Wireless Networks

This chapter presents an overview of resource allocation for wireless networks. Starting from an introduction of wireless networks and the basic principles of resource allocation, such as power control, rate adaptation, multiple access and cross-layer design, to the optimization techniques, then, related works, i.e. previous resource allocation algorithms for the traditional OFDM-based wireless networks are also discussed. This chapter will help us to understand the basics and the importance of wireless resource allocation.

2.1 Introduction

Over the past decade, there has been a significant advance in the design of wireless networks, ranging from physical (PHY) layer algorithm development and MAC layer protocol design to network and system level optimization. Many wireless standards have been proposed to suit the demands of various applications. For wireless networks, because of fading channels, user mobility, energy/power resources, and many other factors, one cannot optimize wireless networks as has been traditionally done for wired networks, in which one can simply focus on and optimize each networking layer without paying much attention to the effects of other layers. For wireless networks, cross-layer optimization is a central issue to ensure overall system performance. Resource allocation is one of the most important issues for cross-layer optimization of wireless networks.

Resource allocation is used to assign the available resources in an economic way. The purpose of resource allocation for wireless networks is to intelligently allocate the limited resources, e.g. total transmission power and available frequency bandwidth, among users to meet users' service requirements.

Due to the number of degrees of freedom and many different parameters, resource allocation is the issue covering a wide range of problems. Therefore, the optimization tools that can be employed vary a lot. Besides the commonly used convex optimization in communication system design, many resource allocation problems are nonlinear and nonconvex. When it comes to channel allocation and scheduling, sometimes the problems become integer, combinatorial, or both. If one takes into account time-varying conditions, then the problem evolves into one of dynamic optimization. When allocation among distributed and autonomous users is considered, game theory can be employed to find the optimal strategy and solution. It is fair to say that there is no single optimization tool available to solve all resource allocation problems at once.

What makes resource allocation more challenging is that, in fact, when it comes to the applications, different wireless networks aim at different service goals, and therefore have different design specifications. One network can be severely energy sensitive and power constrained, whereas the other can be bandwidth limited and throughput hungry. In some situations, a network may have a high degree of mobility with opportunistic access, whereas in other cases a network has an ultrawide bandwidth to share with others but little mobility.

As such, different networks face different resource allocation problems; different characteristics of problems employ different optimization techniques; and joint considerations of different layers encounter different constrained optimization issues.

2.2 Basics of Resource Allocation

Wireless network refers to a telecommunications network whose interconnections between nodes are implemented without the use of wires. There is great growth

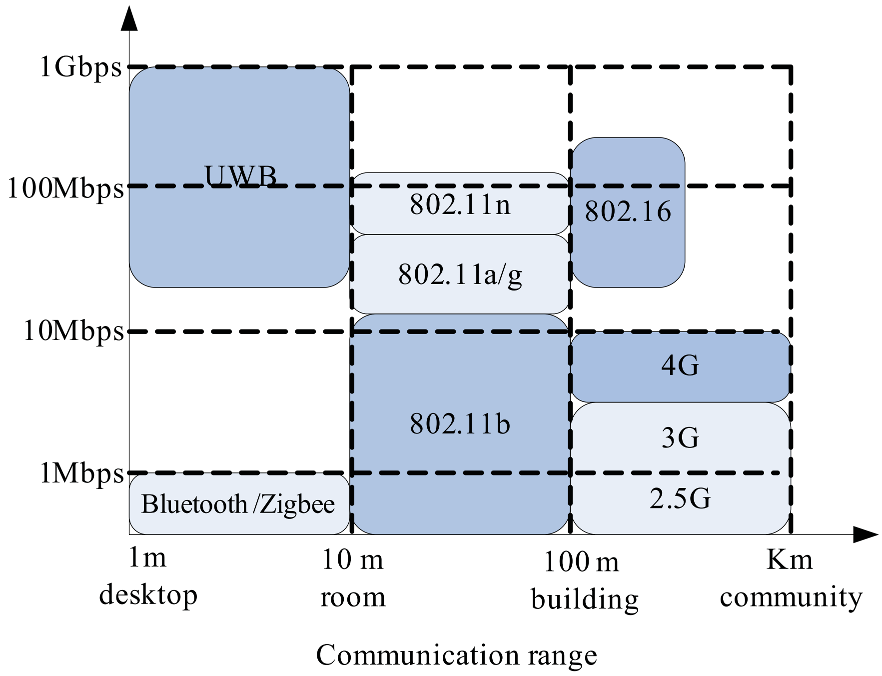

Figure 2.1: Wireless standards comparison.

during the past few decades and it will continuously evolve in the future.

There are many existing wireless standards. Figure 2.1 shows different standards for different communication rates and different communication ranges. These standards will fit different needs of various applications, and the same techniques can be utilized by multiple standards in different situations. According to the decreasing order of the coverage areas, there are four types of wireless networks: cellular networks, wireless metropolitan area networks (WMAN), wireless local area networks (WLAN), and wireless personal area networks (WPAN).

Besides the above wireless networks, there are some wireless networks without specified standards, such as wireless ad hoc networks, wireless sensor networks, and cognitive radio networks. Different networks can be applied to different environments. The details of cognitive radio networks will be explained in the next chapter.

For different wireless networks, there are different resource allocation algorithms. However, the basics of resource allocation are the same. In wireless networks, the resources allocated to multiple users always consist of power resource and space/time/spectrum resource. Multiple users in wireless networks will share the limited resources under different practical constraints. We will explain the basic principles of resource allocation following the four aspects: power control, rate adaptation, multiple access and spectrum access, and cross-layer designs [9].

2.2.1 Power Control

In wireless communications, transmission power is an important resource. Power control is a significant design problem in modern wireless networks. It serves several purposes, including combating fading channel, reducing co-channel interference (CCI), managing data quality, maximizing cell capacity, minimizing handset mean transmission power and so on. In wireless networks, two important detrimental effects that decrease network performance are the time-varying nature of the channels and CCI. The average channel gain is primarily determined by large-scale path-loss factors such as propagation loss and shadowing. The instant channel gain is also affected by small-scale fading factors such as multipath fading. Because the available bandwidth is limited, the channels are shared for different transmissions. The channel sharing increases the network capacity but causes CCI. Power control is an effective resource-allocation method to combat these detrimental effects. The transmission power is adjusted according to the channel condition so as to maintain the received signal quality. Power control is not a single user's problem because a user's transmission power causes other users' interferences. The objective of power control is to control the transmission power to guarantee a certain link quality and reduce CCI.

In wireless networks, there are the following difficulties in power control:

- There is a trade-off for each link's power. The increase in transmission power will increase the link's signal to interference plus noise ratio (SINR), but, on the other hand, the increased power will interfere with other links and cause degradation of other links.

- For the uplink or multi-cell case, to reduce network overhead, it is better to implement power control in a distributed way. That is to say, all

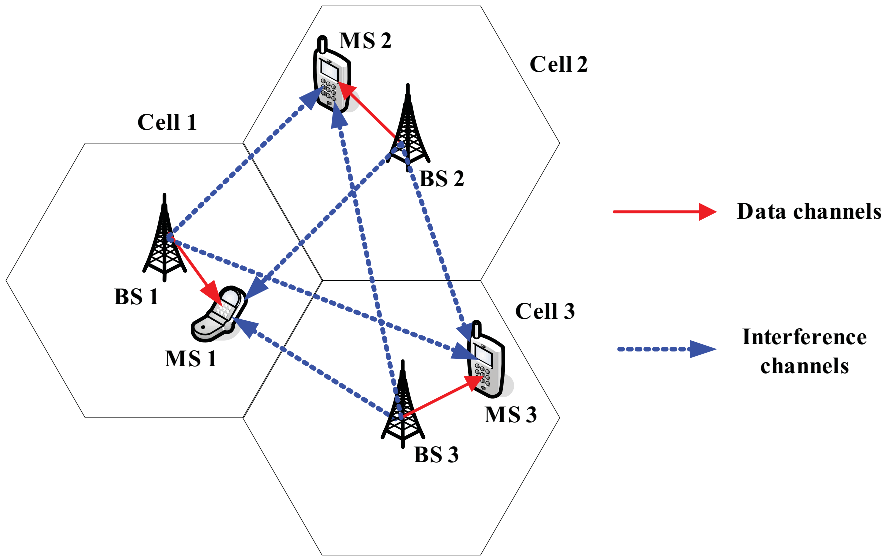

Figure 2.2: Example of multi-cell case.

users should use their local information to control the power so that the limited power resources can be effectively utilized to improve the system performance while maintaining the users' QoS.

- It is necessary to have a simple implementation of power control without causing too much communication overhead and burden.
- To combat fading channels, the convergence speed for a power-control algorithm should be fast enough compared with the changing speed of the fading channels.
- The power control scheme should be able to accommodate heterogeneous QoS requirements.

Hence, the efficient management of the power resource has become an important research issue in the recent years.

For different scenarios, such as single-cell case and multi-cell case, uplink and downlink, power control schemes are different. Figure 2.2 illustrates an example of a multi-cell communication network. A set of transmitter-receiver pairs share the same channel in different cells. The channels between different cells are referred as interference channels, i.e., the dotted lines shown in Figure 2.2. The interferences

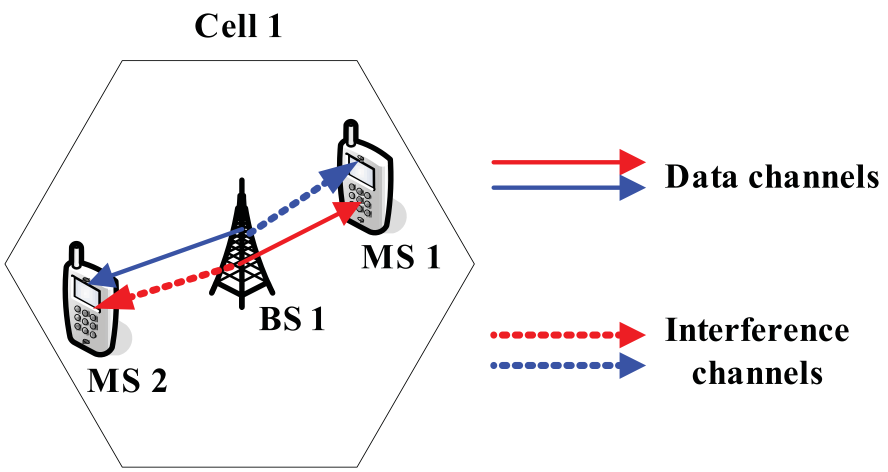

Figure 2.3: Example of single-cell case.

are referred to as co-channel interferences. If the cells sharing the same channels are separated far enough, the interferences may be much less than the thermal noise. Under this condition, the power control problem becomes a single-user optimization problem and can be easily solved. However, if the channel-sharing cells are close to each other, we need to control the users' transmission power in the different cells. Moreover, the estimation of interference channels is difficult. So distributed power control using only local information is preferred to multi-cell power control implementation.

In Figure 2.3, it shows an example of a single-cell communication network. In contrast to the multi-cell case, there is only one base station in the network. Multiple users are distributed in one cell. When the multiple users share the same spectrum, the interferences might be much larger compared with those in the multi-cell case. So power control is always necessary for all users every time. Since the users access to the same base station, the interference channel information can be obtained easily. Consequently, centralized power control can be possibly implemented in the single-cell case.

Power control schemes can be classified according to how to measure the power, what the available measurements are, what the constraints are, and how much time delay can be accepted [9]. Based on the directions of communications, the power control schemes can be classified as uplink and downlink power control. According to what is measured to determine power, power control techniques can be classified into the strength of arriving signal based, received SINR based and bit-error-rate (BER) based power control. Depending on whether feedbacks exist, power control techniques can be classified as closed-loop, open-loop and combined closed/open-loop power control.

In addition, according to network infrastructures, there exist two main classifications: centralized and distributed power control. The centralized power control is limited by the overhead of channel estimation. So, it can work well in a system with a small number of users and a centralized topology. On the other hand, the distributed power control can provide practical implementation. The disadvantages are low convergent speed and possible infeasibility.

2.2.2 Rate Adaptation

Rate adaptation is one of the most important resource allocation issues. Wireless networks can adapt the users' rates so that the limited resources can be efficiently utilized. Compared with power control, rate adaptation gives a new dimension of freedom to change the transmission rate over time, i.e., power control maintains the desired link quality, whereas rate adaptation adjusts this link quality.

According to the different layers of the OSI (Open System Interconnection) model, rate adaptation can be classified into three different types: source rate adaptation in the application layer, rate control for data communication in the network/MAC layer, and channel protection adaptation in the PHY layer. They are briefly summarized in the following:

Source Rate Adaptation

This type of adaptation adjusts the quality of transmitting information at source nodes. Because the structures of coders for different services (such as voice and video) are different, the design concerns for source adaptation are different. That is, different types of services have different source encoders to control the source rates. For example, the voice encoder can change the information rate according to the talking period and the silence period, as it is useless to have a high data rate for the silence period. For video transmission, the data is very bursty over time, because of the different video scenarios and different frames.

Because the capacity to deliver the information is limited by the communication systems, the wireless networks can utilize the limited system resources effectively by the source rate adaptation.

Rate control for network/MAC layer

The network/MAC layer utilizes buffers to accommodate the rate differences between rates of source coders and rates that channel can provide. Rate control is critical to optimize the buffer behaviors and maintain the QoS.

For data transmission over wireless networks, one of the important QoS targets is the average delay required for delivering a packet from the source to the destination. The delay is influenced by the packet arrival rate, service rate and others. Moreover, the maximum buffer size is also an important issue for practical implementation. Therefore, there are the following two kinds of rate control problems:

- The first one is delay-constrained. For different applications, the delay constraints are different. For example, for the voice packet, the delay requirement is very strict, because the delayed packets can significantly reduce voice quality, whereas for services like e-mail, the delay can be arbitrarily long. So the problem formulation is usually constrained by the maximum delays for the specific types of applications. The optimization is performed by adapting the service rate from the link or the incoming packet rate from the source.

- The second one is maximum-buffer-size-constrained. The buffer size of the communication system is limited. So this kind of problem has the constraint on the maximum buffer size. The resource allocation scheme tries to prevent buffer overflow. If the buffer does overflow, some packets will be selectively dropped.

For the multiple-user case, the multiple-access nature of the wireless channels requires rate control for different users.

Rate Adaptation for PHY layer

The wireless channel gains and phases fluctuate over time, which causes fading in the communication links. To combat the fading, channel adaptation schemes, such as adaptive channel coding and adaptive modulation schemes, should be considered. In channel adaptation schemes, rate adaptation can be implemented to protect the channel from transmission errors according to the channel state.

In specific wireless networks, different types of rate adaptation can be applied or combined together. In addition, rate control can be combined with power control to further improve the system performance. With regard to implementation, both rate adaptation and power control impose practical challenges. For rate adaptation, different modulation and coding schemes have to be implemented. There shall be a reliable channel to feed back the selected rate without any delay. For power control, the link quality shall also be monitored and fed back. In the most of current standards, the power control is utilized much more frequently than the rate adaptation.

2.2.3 Multiple Access and Spectrum Access

Multiple access considers the problem of allocating limited radio resources, such as spectrum, time and space, to multiple users. Spectrum access decides whether an indicidual user can access the spectrum.

Multiple Access

Multiple access is a general accessing strategy to allocate the limited resources, such as space, bandwidth and time, to guarantee the basic QoS, improve the system performances, and reduce the cost for the network infrastructures.

The basic idea of the multiple access is to combine several signals at the transmitter by a multiplexor and split up at the receiver by a de-multiplexor. Based on how to divide the limited radio resources to multiple users, the multiple-access schemes can be classified as time-division multiple access (TDMA), frequency-division multiple access (FDMA), code-division multiple access (CDMA), space-division multiple access (SDMA), and others. For multiple users' communication, TDMA, FDMA, CDMA, SDMA, frequency-hopped multiple access (FHMA) and orthogonal frequency-division multiple access (OFDMA) are major access techniques [10].

Based on how to coordinate access for multiple users, multiple-access schemes can be classified into scheduling and random access. In scheduling, there is a centralized control, i.e., the base station, to control which user can transmit by using specific resources such as the bandwidth at different times. In random access, there is no such centralized control. Users access and utilize the resources in a distributed way. If conflicts of resource usage occur, certain mechanisms are employed to avoid conflicting in the future. These two types of schemes are employed in different scenarios depending on the networks' situations. For example, in cellular networks, it is possible for centralized control where scheduling can be employed. On the other hand, in the WLAN, mobile users distributively share the limited bandwidth. As a result, radom-access schemes are widely deployed.

Spectrum Access

For spectrum access, there are two kinds of methods: channel allocation and opportunistic spectrum access.

Because the radio spectrum is limited, a given radio spectrum is to be divided into a set of disjointed channels that can be used simultaneously while minimizing interference in adjacent channels by allocating channels appropriately (especially for traffic channels). Frequency allocation should be carefully planned to avoid degradation caused by CCI.

With the control of a central processor, the channel allocation schemes can be coordination-based. Coordination-based channel allocation schemes can be divided in general into fixed channel allocation schemes (FCA), dynamic channel allocation schemes (DCA), and hybrid channel allocation schemes (HCA, combining both FCA and DCA techniques).

Besides the coordination-based approach, there are also distributed measurement-based methods, which can alleviate the processing time of the central processor greatly.

Opportunistic spectrum access enables the dynamic management of radio resources within a single-user access system or between different radio-access systems. The conventional fixed spectrum allocation results in low utilization of the allocated spectrum. Opportunistic spectrum access can improve spectral efficiency, increase capacity and improve ease of access to the spectrum. In addition, the FCC has investigated a novel cognitive radio technology to improve spectrum utilization by allowing SUs to borrow unused radio spectrum from PUs. The cognitive radio technology will be introduced in detail in the next chapter.

2.2.4 Cross-Layer Designs

In wireless networks, the different layers of the OSI model interact in a nontrivial manner in order to support information transfer. Due to the time varying nature of a wireless network (either fading channels or user mobility), it is difficult to capture the network nature and control the network in real-time only by single-layer information. Cross-layer designs from the PHY to transport layer in wireless networks are with significant importance.

In cross-layer designs of wireless networks, a number of PHY and MAC layer parameters are jointly controlled and in synergy with higher layer functions like transport and routing. In general, cross-layer operation terminology refers to "any violation of the layered architecture" [11], in order to adapt to the dynamics of the wireless environment in terms of the traffic pattern, the service demand, mobility and other variations in environment. Considering cross-layer design proposals in the references, they can be summarized into the following categories:

- *Air interface-centric*: In this category of cross-layer optimization, the main focus is on efficient utilization of the scarce wireless radio resources through adaptation to the time-variant channel, based on throughput efficiency, fairness and QoS. Therefore, the problem is generally modeled between the PHY and MAC layers while the upper layers of protocol stack (Application/Transport) contribute mainly to the traffic pattern. This imposes additional constraints on the original problem. Two main examples of this approach are opportunistic scheduling in cellular single-hop networks [12] and joint congestion-control and scheduling in multi-hop wireless networks [13].

- *User-centric*: In this approach, the main concern is the adaptation of upper layer protocols to the time-variant, unreliable channel in wireless environment to achieve a certain level of user satisfaction in terms of the end-to-end throughput, delay, and power consumption. Hence, the emphasis is on the upper layer protocols whereas channel characteristics are modeled using simplified assumptions with less consideration of the air interface technology, radio resource management policy or interference modeling. The proposals for improving the performance of Transmission Control Protocol (TCP) [14] in wireless networks can be categorized as user-centric cross-layer operation.

- *Route-centric*: Recently, there has been much activity on cross-layer routing in multi-hop wireless networks. The cross-layer routing designs effectively couple the network layer and the PHY or MAC layer to select the best routing. Some examples of this approach can be found in [15], [16].

In thisdissertation, the main focus is on air interface-centric cross-layer resource allocation designs. Typical PHY and MAC layer functions include power control and channel allocation.

2.3 Optimization Formulation of Resource Allocation Problems

A standard problem in network design deals with the question of how the available resources should be shared between competing users to meet some share objectives. One possible objective is to allocate resources to a set of users so as to maximize the total throughput. The main drawback of this objective is that it may be quite unfair in the sense that some users may be denied access to the links. For this reason, another objective that must be considered is the issue of fairness. However, the “perfect fairness” is usually at the expense of a considerable drop in efficiency (i.e., total throughput). Therefore, there seems to be a fundamental trade-off between throughput and fairness, with the throughput-optimal policy and perfect fair policy being two extremes of this trade-off [17].

Except for throughput and fairness, QoS (e.g., service delay)is another important objective. In order to achieve different objectives, there exist different resource allocation problems in wireless networks. A common approach to balance these objectives is to maximize the aggregate (overall) utility of resource allocations subject to different constraints.

In this section, we discuss how to formulate a resource allocation optimization issue. Specifically, we need to know what the resources are, what the parameters are, what the practical constraints are, and what the optimized performances across the different layers are. For multi-user scenarios, the trade-offs between the different optimization goals and different users' interests are also needed to be considered.

2.3.1 Constrained Optimization

Many wireless resource allocation problems can be formulated as constrained optimization problems, which can be optimized from the network point of view or from the individual point of view. The general formulation can be written as

$$\min_{\mathbf{x}\in\Omega} f(\mathbf{x}), \tag{2.1}$$

$$\Omega: \quad \begin{cases} g_i(\mathbf{x}) \leq 0, & for \;\; i = 1, \ldots, I, \\ h_j(\mathbf{x}) = 0, & for \;\; j = 1, \ldots, J. \end{cases} \tag{2.2}$$

where $\mathbf{x}$ is the parameter vector for resource allocation, Ω is the feasible range of the parameter vector, and $f(\mathbf{x})$ is utility function that represents the performance or cost. Here, $g_i(\mathbf{x})$ and $h_j(\mathbf{x})$ are the inequality and equality constraints for the parameter vector, respectively. The optimization process finds the solution $\mathbf{x}^* \in \Omega$ that satisfies all inequality and equality constraints.

If the optimization goal, the inequality constraints, and the equality constraints are all linear functions of the parameter $\mathbf{x}$, then the problem in (2.1)-(2.2) is called a linear program. There is a global optimal point that is very easy to obtain by linear programming. However, most of the practical problems in wireless networking and resource allocation are nonlinear. If either the optimization goal or the constraint functions are nonlinear, the problem is called a nonlinear programming. In general, there are multiple local optima in a nonlinear program and to find the global optimum is not an easy task. Furthermore, if the feasible set Ω consists of integers, the problem is an integer programming. Most integer programs are nondeterministic-polynomial-hard (NP-hard) problems that cannot be solved in polynomial time.

One special kind of nonlinear programming is the convex optimization problem in which the feasible set Ω is a convex set [18], and the optimization goal and the constraints are convex/concave/linear functions. The advantages of convex optimization for wireless networking and resource allocation problems are shown as follows [9]:

- Computation time is usually quadratic. Problems can then be solved efficiently, using interior-point methods or other special methods for convex optimization.

- Solution methods are reliable enough to be embedded in a computer-aided design or analysis tool, or even a real-time reactive or automatic control system.

The challenges of convex optimization are to recognize and model the problem as a convex optimization. There are many tricks for transforming problems into convex forms.

The basics of constrained optimization problems are discussed above. In wireless networking and resource allocation, the parameters, utility functions, and constrains in (2.1)-(2.2) can have the following physical meanings:

- **Parameters:**

  1. PHY layer: transmission power, modulation level, channel-coding rate, channel/code selection, and others.

  2. MAC layer: transmission time/frequency, service rate, priorities for transmission, and others.

  3. Network layer: route selection, routing cost, and others.

4. Application layer: source-coding rate, buffer priority, packet arrival rate, and others.

- **Optimization goals:**

  1. PHY layer: minimize overall power, maximize throughput, maximize rate per link, minimize BER, and others.

  2. MAC layer: maximize overall throughput, minimize buffer overflow probability, minimize delay, and others.

  3. Network layer: minimize cost, maximize profit, and others.

  4. Application layer: minimize distortion, minimize delay, and others.

- **Constraints:**

  1. PHY layer: maximum user transmission power, available modulation constellation, available channel-coding rate, limited energy, and others.

  2. MAC layer: contentions, limited time/frequency slot, limited information about other users, and others.

  3. Network layer: maximum hops, security concerns, and others.

  4. Application layer: the base-layer transmission, limited source rate, strict delay requirement, security, and others.

After formulating the constrained optimization problem for resource allocation, we need to get the solution by using some optimization techniques, which will be introduced in the next section.

2.4 Optimization Techniques

Usually, the resource allocation issues in wireless networks can be formulated as optimization problems with different objective goals, different resources, parameters, and several constraints. Various optimization techniques can be applied to wireless resource allocation problems.

2.4.1 Mathematical Programming

If the optimization problem is to find the best objective function within a constrained feasible region, such a formulation is sometimes called a mathematical programming, which has the following forms:

- given: a function $f : \mathbf{A} \rightarrow \mathbf{R}$ from a certain set $\mathbf{A}$ to the real numbers;
- find: an element x_0 in $\mathbf{A}$ such that $f(x_0) \leq f(x), \forall x \in \mathbf{A}$ (minimization), or such that $f(x_0) \geq f(x), \forall x \in \mathbf{A}$ (maximization).

Typically, $\mathbf{A}$ is a certain subset of the Euclidean space $\Re^n$, often specified by a set of constraints, equalities, or inequalities that the members of $\mathbf{A}$ have to satisfy. The domain $\mathbf{A}$ of f is called the search space. The elements of $\mathbf{A}$ are called feasible solutions. The function f is called an objective (utility) function, or cost function. A feasible solution that minimizes or maximizes the objective function is called an optimal solution.

There are four major subfields of mathematical programming: linear programming, convex programming, nonlinear programming and dynamic programming.

- Linear programming (LP) studies the case in which the objective function f is linear and the set $\mathbf{A}$ is specified using only linear equalities and inequalities.
- Convex programming studies the case in which the constraints and the optimization goals are all convex or linear.
- Nonlinear programming (NLP) studies the general case in which the objective function or the constraints or both contain nonlinear parts.
- Dynamic programming studies case in which the optimization strategy is based on splitting the problem into smaller subproblems or considers the optimization problems over time.

In general, mathematical programming can be solved using the following approaches:

For twice-differentiable functions, unconstrained problems can be solved by finding the points where the gradient of the objective function is 0 (i.e., the stationary points) and then using the Hessian matrix to decide whether the point is local optimal. If the objective function is convex over the region of interest, the local optimal point will also be a global optimal point.

There exist robust and fast numerical techniques for optimizing twice differentiable convex functions. Besides, constrained problems can often be transformed into unconstrained problems with the help of Lagrangian multipliers.

The details of the mathematical programming can be found in [18] and [19]. In Chapter 4, Lagrangian method-based convex optimization problem is investigated for resource allocation in multi-user CRNs.

2.4.2 Integer/Combinatorial Optimization

Combinatorial optimization problems are problems of choosing the best combination out of all possible combinations. Most combinatorial problems can be formulated as integer programming. In wireless resource allocation, many variables, such as channel allocation, have a combinatorial nature. Integer/combinatorial optimization is the process of finding one or more best (optimal) solutions in a well-defined discrete problem space. For example, in a WLAN, the time slots are occupied by different users. The allocation of time is restricted to a discrete nature. In WiMAX or any OFDM systems, the distinct time-frequency slot is also allocated to the admitted users. Moreover, for practical implementation, the coding rate and adaptive modulation can have only discrete values. To design future wireless networks, it is with significant importance to study integer optimizations.

The general problem formulation can be given by:

$$\min_{\mathbf{x},\mathbf{y},\mathbf{z}\in\Omega} f(\mathbf{x},\mathbf{y},\mathbf{z}), \tag{2.3}$$

$$\Omega: \quad \begin{cases} g_i(\mathbf{x},\mathbf{y},\mathbf{z}) \leq 0, \quad for \;\; i = 1, \ldots, I, \\ h_j(\mathbf{x},\mathbf{y},\mathbf{z}) = 0, \quad for \;\; j = 1, \ldots, J, \\ \mathbf{x} \in \Re, \mathbf{y} \in \{0,1\}, \mathbf{z} \in \mathbf{I}. \end{cases} \tag{2.4}$$

where, function f is the objective function, g_i is the inequality constraint, h_i is the equality constraint, the component of vector $\mathbf{x}$ is a real value variable, $\mathbf{y}$ is a variable of either 0 or 1, and $\mathbf{z}$ is a integer value in a space $\mathbf{I}$. If $\mathbf{y} = 0$ and $\mathbf{z} = 0$, it becomes a nonlinear optimization. If $\mathbf{z} = 0$, the problem is referred to a pure 0–1 integer programming problem. If $\mathbf{y} = 0$, it is called a pure integer programming problem. Otherwise, the problem is a mixed integer programming problem.

For wireless networking and resource allocation, there are many potential applications of integer optimization, such as routing and network graph problems, scheduling problems, assignment problems (i.e., a number of tasks should be allocated to different agents). Knapsack problem [20] is one special case of this kind of optimization problem.

There are at least three different approaches for solving integer programming. They include relaxation and decomposition techniques, enumeration techniques, and cutting-plane approachs based on polyhedral combinatorics [21], [22]. The three approaches also can be combined into a “hybrid” method in computational practice. In Chapter 4, relaxation method is utilized to simplify the optimization problem, and transfer it as a convex optimization.

2.4.3 Game Theory

Game theory [23], [24] is a branch of applied mathematics that uses models to study interactions with formalized incentive structures. It studies the mathematical models of conflict and cooperation among intelligent and rational decision makers. “Intelligent” means that each individual understands everything about the structure of the situation, including the fact that others are intelligent, rational decision makers. “Rational”means that each individual’s decision-making behavior is consistent with the maximization of subjective expected utility.

In wireless networks, to obtain information such as channel conditions, signaling is performed so that resource allocation can be conducted in an optimal way. However, signaling has considerable overhead for communications. Most of the current wireless networks have more than 50% of overhead. Reducing the overhead can increase the spectrum utilization, increase the number of users, and improve the network performance. One way to reduce overhead is to do resource optimization by using only local information. This is very important, especially if the system topology is distributive.

In some wireless network scenarios, it is hard for an individual user to know the channel conditions of the other users. The users cannot cooperate with each other. They can act selfishly to maximize their own performances in a distributive way. Such a fact motivates us to adopt game theory. Resouce allocation can be modeled as a game that deals with how rational and intelligent individuals interact with each other in an effort to achieve their own goals. In the game, each user is self-interested and trying to optimize its utility function, in which the utility function represents the user's performance and controls the outcomes of the game. There are many advantages of applying game theory to wireless networking and resource allocations:

- Reducing network overhead: The individual user observes the outcome of the game and adjusts only its own resources in response to optimize its benefit. So, there is no need to collect all the information and conduct constrained optimization in a centralized way.

- More robust outcome: If the information for optimization is not quite accurately obtained, the optimized results can be far from optimality. In contrast, local information is more accurate, so the outcome of distributed game approaches is more robust.

- Combinatorial nature: The traditional optimization techniques are hard to handle combinatorial problems. For game theory, it is natural to discuss the problem in a discrete form. In problems such as discrete modulation levels and channel-coding rates, analyzing the combinatorial problems by game theory is considerably convenient.

- Rich mathematics for optimization: There are many mathematics tools to analyze the outcome of the game. Specifically, if the game is played noncooperatively, the static game can be studied. If the game is played multiple times, dynamic game theory is employed. If some contracts and mutual benefits can be obtained, cooperative game explains how to divide the profits. Auction theory studies the behaviors of both seller an bidder.

In general, there are four types of games, namely, the noncooperative game, repeated game, cooperative game and auction theory. In Chapter 5 of the theis, a noncooperative game based dynamic resource allocation algorithm for multi-cell CRNs is studied.

2.5 Previous Resource Allocation Algorithms for the Traditional OFDM-based Wireless Networks

In thisdissertation, resource allocation for OFDMA-based Cellular Cognitive Radio Networks is studied. Therefore, related previous works on resurce allocation algorithms for traditional OFDM-based wireless networks are presented here.

Centralized Physical Layer Approach

The bit and power loading problem for single-user OFDM systems can be solved by using the well-known water-filling [25] algorithm if we assume that the number of bits to be loaded is a real number, or implement a greedy approach that assigns one bit at a time to the subcarrier that requires the least additional power for the integer bit case. To reduce computational complexity for the integer bit case, various low complexity algorithms have been proposed, for both optimal (e.g. [26]) and suboptimal solutions (e.g., [27]).

In the case of the downlink transmission from a BS to multiple users, the subchannels need to be assigned to users exclusively [28]. Therefore, RA involves subchannel assignment in addition to power and bit allocation. When the goal is to maximize system throughput, the problem can be solved in two separate steps [28], namely, assigning each subchannel to the user with the best channel condition, followed by power and bit allocation.

When there are QoS or fairness requirements, subchannel, bit, and power allocation becomes more complicated. Since optimal solutions are generally computationally complex, various sub-optimal solutions have been proposed. In [29], suboptimal solutions are proposed to minimize the total transmit power while satisfying rate and BER requirements for real-time (RT) services. For non-real-time (NRT) services, maximizing system throughput while guaranteeing a certain level of fairness among users is a reasonable goal [30].

Most of these suboptimal solutions use the divide-and-conquer approach, in which the subcarrier, power, and bit allocation problem is broken down into two steps, i.e., allocate subcarriers to users and load appropriate power and bits to each subcarrier. During the first step, power is often assumed to be the same across all subcarriers so as to simplify the problem.

Centralized MAC Layer Approach

RA also occurs in the MAC layer, which is responsible for packet scheduling. Almost all existing studies extend opportunistic scheduling strategies from the single carrier case to the multiuser OFDM case with multiple subcarriers. For NRT services, some schemes (e.g., [31]) extend the proportional fair (PF) rule, while others (e.g., [32]) extend the modified-largest weighted delay first rule [33] for RT traffic. An urgency and efficiency based packet scheduling algorithm is proposed in [34] for both RT and NRT services using an urgency factor that reflects the urgency of meeting QoS requirements combined with the PF rule to maximize system throughput.

Centralized Cross-Layer Approach

Some researchers have adopted a cross-layer design approach in allocating system resources. In [35] and [36], sub-optimal algorithms for NRT services are proposed and algorithms for both RT and NRT services are studied in [37] and [38]. In [37], the QoS for RT applications is improved by giving high priority to users whose head-of-line packet deadlines are approaching. In [38], the MAC layer QoS requirement for each user is converted to a PHY layer fixed rate requirement based on the average user packet arrival rate and delay constraint. An optimal subchannel and power allocation strategy is proposed that maximizes system throughput subject to a total transmit power limit and user delay requirements.

Distributed Approach

While centralized RA is suitable for single-cell systems, distributed algorithms may be more appropriate for multi-cell cellular systems or ad hoc systems. Although distributed dynamic channel allocation (DCA) has been studied for multiple cell cellular networks for voice services, it cannot be easily ported to multiuser OFDM systems. This is because traditional DCA schemes assume homogeneous applications with a pre-determined SINR threshold, and may not efficiently support services with different requirements.

To dynamically allocate resources in a multi-cell system or an ad hoc system, subcarriers may be simultaneously shared among served users in order to improve system performance. In this case, co-channel interference has to be considered. In [39], other users' signals are treated as noise, and the power allocation problem is viewed as a non-cooperative game. A distributed iterative water-filling (IWF) algorithm is proposed for digital subscriber line (DSL) systems in [39]. To achieve the optimal power allocation solution, the achievable target rates must be known. This is not a big problem for DSL systems, but is unrealistic for time-varying wireless channels. To make IWF suitable for wireless systems, a scheme is proposed in [40] for multi-cell wireless systems in which a virtual referee is introduced to displace some users out of certain subchannels when necessary, to allow ITWF to converge to good solutions.

Power and bit allocation for multiuser OFDM systems with co-channel interference have been formulated as a constrained nonlinear programming problem in [41]. To reduce the complexity of finding a solution, a distributed algorithm is proposed that allocates one bit per iteration.

2.6 Chapter Summary

We introduce the basics of resource allocation for wireless networks in this chapter. The optimization formulation and techniques used in resource allocation are also presented. Related previous works on resource allocation for traditional OFDM-based wireless networks are also discussed.

Different wireless network scenarios always have different design objectives, so, the overall utility functions, which need to be optimized in resource allocation, are always distinctive. Due to different features of the optimization problems, different optimization techniques (i.e., solving methods) will be applied to obtain the optimal/sub-optimal solution. In thisdissertation, two different resource allocation problems using different optimization techniques are investigated for single-cell and multi-cell multi-user CRNs in chapter 4 and chapter 5, respectively.

Chapter 3

Cognitive Radio Networks

Cognitive Radio (CR) technology provides a new and promising solution to improve the spectrum utilization. Some basic concepts about the CR technology, cognitive radio networks (CRNs), and its applications are intoduced in this chapter. The fundamental challenges and issues in designing CRNs are also presented. Researches on CRNs are mainly focused on technologies for PHY and MAC layers. Resource allocation is one of the most important topics for dynamic spectrum access in CRNs. The related previous works on resource allocation for CRNs are also reviewed in this chapter.

3.1 Cognitive Radio Technology

Cognitive radio [1] has emerged as a promising technology for alleviating the severe spectrum shortage problem while accommodating the increasing amount of services and applications in wireless networks. Haykin [6] defines cognitive radio as:

"Cognitive radio is an intelligent wireless communication system that is aware of its surrounding environment (i.e., outside world), and uses the methodology of understanding-by-building to learn from the environment and adapt its internal states to statistical variations in the incoming RF (radio frequency) stimuli by making corresponding changes in certain operating parameters (e.g., transmit-power, carrier frequency, and modulation strategy) in real-time, with two primary

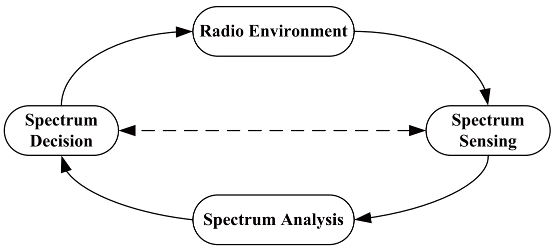

Figure 3.1: Spectrum sharing process.

objectives in mind: (1) highly reliable communications whenever and wherever needed; (2) efficient utilization of the radio spectrum."

A CR transceiver is able to adapt itself to the dynamic radio environment to maximize the utilization of the limited radio resources while providing flexibility in wireless access. A CR transceiver is able to aware of the radio environment, in terms of spectrum usage, power spectral density of transmitted/received signals, wireless protocol signaling, etc. and intelligently configure its system parameters, such as transmit-power, carrier frequency, physical-layer modulation strategy, and higher-layer protocol parameters. This intelligence is achieved through its learning ability.

Due to this ability of CR, the use of CR technology to allow SUs (i.e., unlicensed users) to share radio resources with PUs (i.e., licensed users) is suggested by FCC [5] to improve spectrum efficiency. Therefore, the SUs equipped with CR have the following functionalities in a spectrum sharing process [8], as shown in Figure 3.1:

- Spectrum sensing: monitoring the available spectrum bands to detect spectrum holes. It is also useful to monitor other information in the wireless communication environment, such as, the activities of PUs, and higher layer information in the transmission protocol stack. The main challenges are energy-limit and hardware cost.
- Spectrum analysis: Based on the measurements obtained through spectrum sensing, it is required to determine which portion of the spectrum is available for SUs and build a spectrum available model. Due to hardware

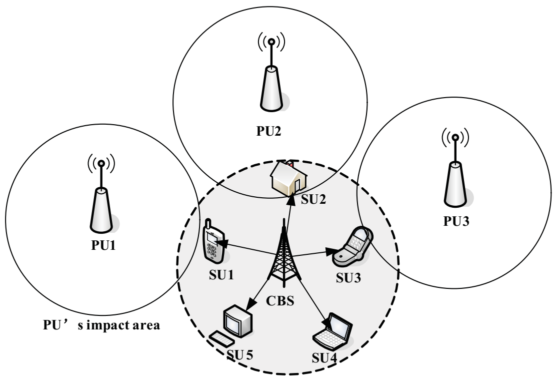

Figure 3.2: An example of centralized cognitive radio networks.

and energy limitations, it is difficult to build a full spectrum scene and an accurate primary communication network. To improve the accuracy, cooperation between nodes can be implemented at the cost of increased communication overhead. Also, techniques for local spectrum analysis are important to build the utilization model of primary spectrum.

- Spectrum decision: It is about how to access the spectrum, including selecting the best available channel and coordinating access to this channel with PUs (e.g., power control to avoid interferences to PUs). The spectrum decision is an optimization problem using the model built during spectrum analysis. This optimization problem can have a local or a global optimization goal.

Even though the SUs have these functionalities, it is unnecessary for each SU to perform all the functionalities at the same time. This depends on the type of cognitive radio networks and the coexistent architectures of CRNs and primary networks (PNs).

3.2 Cognitive Radio Networks and its Applications

Devices with cognitive abilities can create cognitive radio networks. The general examples of CRNs distinguish two types of users sharing the common spectrum

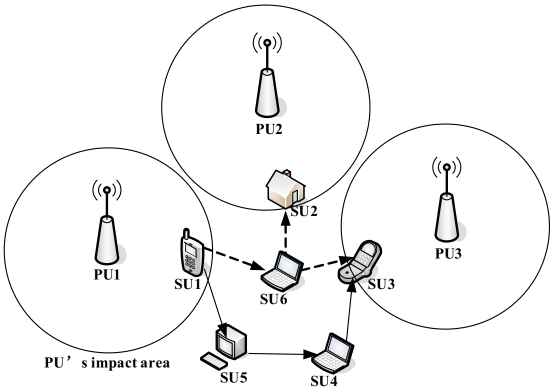

Figure 3.3: An example of distributed cognitive radio networks.

portion with different rules: PUs and SUs. PUs have priority in spectrum utilization within the band they have licensed, and SUs must access the spectrum in a non-intrusive manner. PUs use the traditional wireless communication systems with static spectrum allocation. SUs are equipped with CR tranceivers and exploit spectrum opportunities to sustain their communication activities without interfering with PU transmissions. Figures 3.2 and 3.3 illustrate examples of centralized and distributed CRNs, respectively, where the SUs share spectrum bands with PUs.

Cognitive radio networks changes their configurations based on the spectral environment. This capability opens up the possibility of designing flexible and dynamic spectrum access strategies with the purpose of opportunistically reusing portions of the spectrum temporarily vacated by licensed primary users. On the other hand, the flexibility in the spectrum access phase comes with an increased complexity in the design of communication protocols at different layers. Most of the researches on CRNs focuses on tackling PHY layer and/or MAC layer issues, including the definition of effective spectrum sensing, spectrum decision and spectrum sharing techniques.

With the ability to learn from and adapt users’ need to their surrounding environment, cognitive radio networks offer a great number of benefits in all kinds of application markets: military, government, public safety, and commercial [42]. In general, the applications can be summarized as the following cases:

- **Leased network [8]:** The primary network can provide a leased network by allowing opportunistic access to its licensed spectrum with the agreement with a third party without sacrificing the service quality of the primary users. For example, the primary network can lease its spectrum access right to a mobile virtual network operator. Also the primary network can provide its spectrum access rights to a regional community for the purpose of broadband access. To implement the leased network, devices with cognitive abilities are important.

- **Emergency network [42]:** Public safety and emergency networks are another area in which cognitive radio networks can be implemented. In the case of natural disasters, which may temporarily disable or destroy existing communication infrastructure, emergency personnel working in the disaster areas need to establish emergency networks. Since emergency networks deal with the critical information, reliable communication should be guaranteed with minimum latency. In addition, emergency communication requires a significant amount of radio spectrum for handling huge volume of traffic including voice, video and data. Cognitive radio networks can enable the usage of the existing spectrum without the need for an infrastructure and by maintaining communication priority and response time.

- **Military Network [42]:** One of the most interesting potential applications of an xG network is in a military radio environment.Cognitive radio networks can enable the military radios to choose arbitrary, intermediate frequency bandwidth, modulation schemes, and coding schemes, adapting themselves to the variable radio environment of battlefield. Military networks also have a strong need for security and protection of the communication in hostile environment. Cognitive radio networks could allow military personnel to perform spectrum handoff to find secure spectrum band for themselves and their allies.

- **Cognitive ad-hoc/mesh network [43]:** Wireless ad-hoc/mesh networks are emerging as a cost-effective technology for providing broadband connectivity. However, as the network density increases and the applications require higher throughput, ad-hoc/mesh networks require higher capacity to meet the requirements of the applications. Cognitive ad-hoc/mesh networks can further improve the network throughput by enabling the access to larger amount of spectrum. Cognitive ad-hoc/mesh networks can be used to deploy in dense urban areas, where there is the possibility of significant contention.

3.3 Fundamental Issues in Cognitive Radio Networks

Cognitive radio offers a novel way of solving spectrum underutilization problems. To implement CR technology, a wide range of research issues are involved. For a single-user case (i.e., a transmitter linked to a receiver), the main technical challenges rooted in signal-processing and communication technologies. These technical challenges are further compounded by the fact that the spectrum holes come and go in a stochastic manner. For a multi-user case (i.e., multiple transmitter-receiver links), multi-user CRNs would have to be flexible enough. So, except for the radio-scene issues, the fundamental challenges in designing system architectures also need to be studied. Moreover, there would have to be a paradigm shift from traditional transmitter-centric wireless communications to a new receiver-centric mode in CRNs, so as to control the interference from SUs.

The fundamental issues in CRNs are the followings [8]:

- Radio-scene analysis: this encompasses the information-theoretic analysis of the systems; estimation of interference temperature of the radio environment around a primary receiver; detection of spectrum holes.

- Channel identification: this encompasses the estimation of channel-state information and prediction of channel capacity.

- Dynamic spectrum management and transmission power control: this refers to decision-making and action taken by the secondary transmitter in re-

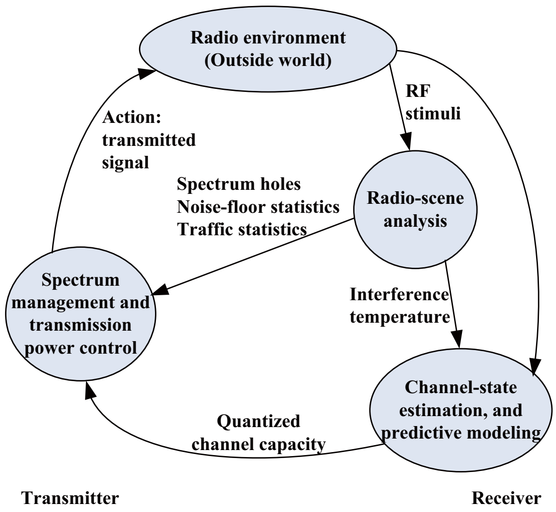

Figure 3.4: Basic cognitive cycle.

sponse to the analysis of RF stimuli picked up by the secondary receiver and includes spectrum access, interference avoidance, and so on.

- System architecture design: coexistent architectures decide how the SUs should share spectrum with PUs and which spectrum sharing method is better.

The first two issues are solved in the receiver, and the third one is performed in the transmitter, as depicted in the *basic cognitive cycle* in Figure 3.4 [8]. Every node of the network is equipped with a transceiver (i.e., transmitter/receiver combination), and each user can either be a transmitter or a receiver. The cognitive module in the transmitter must work in a harmonious manner with the cognitive module in the receiver. In order to maintain this harmony between the CR transmitter and receiver at all times, a feedback channel is necessary to connect the receiver to the transmitter as shown in Figure 3.4. Through the feedback channel, the receiver is enabled to convey the following two information to the transmitter:

- Information on the performance of the link for adaptive modulation;
- Information on the spectral state of the radio environment in the area the receiver located.

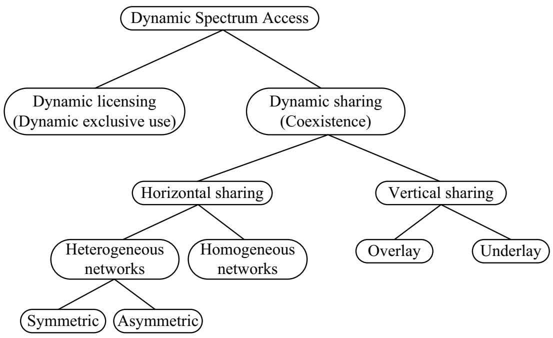

Figure 3.5: Dynamic spectrum access.

The cognitive radio is therefore, an example of a feedback control system.

Cognitive radio technology has different degrees of cognition. On one hand, the user may simply pick one spectrum hole and build its cognitive cycle around that hole. On the other hand, the user may employ multiple implementation technologies to build its cognitive cycle around a wideband spectrum hole or a set of narrowband spectrum holes to provide the best expected performance in terms of spectrum management, transmit-power control, data rate and reliable communication and to do all this in a secure manner.

The last issue is a system-level coexistent problem, which should be studied by considering the network structure of PNs to enable the coexistence, which is discussed in detail in the following section.

3.4 Coexistence and Spectrum Sharing

Various dynamic spectrum access approaches are possible to make the spectrum management more adaptive, as shown in Figure 3.5 [8]. The dynamic spectrum sharing or coexistence is one of them. Coexistence or dynamic sharing allows such sharing, in theory, on a packet-per-packet basis, since it licenses spectrum and networks simultaneously while relying on in-network spectrum sharing techniques to avoid conflicts.

Coexistence scenarios can be of two types: horizontal coexistence and vertical coexistence. In the former case, all of the users have equal regulatory status to access the radio spectrum, where the nodes may use similar wireless access technology (i.e., homogeneous networks, e.g., 802.11a based unlicensed ISM and the U-NII 5GHz frequency band) or different wireless access techniques (i.e., heterogeneous networks, e.g., 802.11b and 802.15.1, 802.11b and 802.16a networks).

In case of vertical coexistence, the radio spectrum is licensed to the PUs only, while the SUs can access the spectrum opportunistically without affecting the PUs' performance. Two approaches for spectrum access to minimize the interference caused to the PUs by the SUs' communication are spectrum overlay and spectrum underlay (as shown in Figure 1.2).

For the vertical coexistence, considering how primary and secondary users cooperate with each other, on one extreme, the primary and secondary users may have completely isolated networks (i.e., without cooperation), and on the other extreme, primary and secondary users may fully cooperate to form a single network where access rights between PUs and SUs transfer with packets. Based on this consideration, there are the following spectrum sharing scenarios:

- Isolated networks (noncooperative networks): A network built to support municipal services such as police or emergency dispatch may allow the same spectrum to be used on a secondary basis when there is no demand for the primary user. In this case, the primary network should be completely isolated from secondary networks.

- Secondary Market(cooperative networks): The primary user may sell secondary access rights to his spectrum and may even allow the secondary user to access to his infrastructure through the cooperative networking approach. In this way, primary users and secondary users can support the development and maintenance costs of the infrastructure together, without sacrificing PUs' access rights.

- Broadband Access Development (opportunistic secondary utilization or fee-based primary utilization): Since it may not be financially viable for a provider to build infrastructure and support access in some regions, the same equipments for primary networks can be used by local communities or neighborhoods to build their own networks and wireless broadband access on a secondary basis or fee-based primary use. Even though the service provider decides to develop infrastructure and provide services in this region, users still can continue using the network on a secondary basis or pay for the primary use and its associated services.

In spite of the type of coexistent scenarios, three functionalities, namely, spectrum sensing, spectrum analysis, and spectrum decision, are fundamental for any CRNs to share spectrum with other PNs. For a multi-user CRN, a centralized or a distributed network architecture can be used for spectrum analysis and spectrum decision, as shown in Figure 3.2 and Figure 3.3. That is, spectrum decision can be taken in a cooperative way (i.e., global optimization) or a non-cooperative way (i.e., local optimization). By cooperation (either through a distributed or a centralized way), the cognitive nodes can share network information among each other or at a cognitive base station to achieve a coordinated and efficient spectrum management. However, the exchanges of information may cause high overhead. Moreover, in a distributed network (without central controller), it may require synchronization among the nodes, which results in a more complex network design. In contrast, a competitive or non-cooperative approach may simplify the distributed network design at the expense of network performance.

3.5 Previous Resource Allocation Algorithms for OFDM-based CRNs

Related previous works on resurce allocation algorithms for OFDM-based CRNs are reviewed below.

Algorithms Dealing with Cross-Channel Interference

Cross-channel interference is considered in [44] and [45]. In [44], the bit and power loading problem is studied for the downlink of an OFDM-based CR system, in which the PU channel located in the middle of a frequency band is available to SUs. An optimal scheme based on a Lagrange formulation and two suboptimal schemes are proposed assuming that there is only one SU in the system. A similar model is used in [45] to study subcarrier, power, and bit allocation for multiple SUs. Greedy algorithms are proposed based on minimum SU power and minimum PU interference considerations.

Algorithms Dealing with Co-Channel Interference

Different optimization problems are formulated and solved based on various interference temperature limit considerations. In [46], to simplify the problem, this limit is converted to a power constraint in each PU band by defining a protection area for the PUs. The power constraint is calculated based on a path loss factor and the distance between the protection area and the SU transmitter. The optimization problems, formulated in [47] for a multiple-SU and multiple-PU system, use two interference temperature models proposed in [48]. The first model, which assumes a unified interference temperature limit on each subchannel, is translated into an average interference power threshold at the measurement point. The second model, which assumes different interference temperature limits on different PU active frequency bands, is translated into an average interference power threshold at each PU receiver. The authors in [49] proposed an uplink RA algorithm to maximize the system throughput in a centralized manner.

Instead of the interference temperature limit, some other means of protection for PU signals are considered in [50], [51] and [52]. Minimum average rate is guaranteed in [50], by assuming that PUs are willing to be cooperative in RA. PU outage probability is ensured in [51]. In [52], the average PU transmission rate is maintained using SU cooperation.

The above-mentioned algorithms, designed for multiple SUs, assume that each subchannel can only be used by at most one SU at any given time. In some situations, e.g., in an ad hoc system or a multicell cellular network, allowing multiple SUs to share each subchannel can result in a higher spectrum utilization. In [53], CRNs with one channel are considered in which all SUs access the channel at the same time, while keeping the total generated interference below the predefined interference temperature limit at a single measurement point. Two co-located cellular systems, consisting of one PU system and one SU system, are studied in [54], in which the average interference from the SUs to the PUs is ensured to be below the interference temperature limit. In [55], the interference to the PUs is limited by a per channel power mask, which specifies the highest power that can be used by a SU on each channel. In [56], two distributed fair subcarrier and power allocation schemes for both the downlink and uplink of CRNs have been proposed.

Algorithms Making Use of Spectrum Holes

Studies assuming the use of spectrum holes appear in [57],[58], where a game theoretic approach is utilized to solve the channel allocation problem based on the observation that users in CRNs may not be willing to cooperate with others but rather may selfishly try to maximize their own performance. A dynamic channel allocation scheme based on a potential game [59] approach is proposed for ad hoc networks in [57]. In [58], a non-cooperative game formulation is used to model the multi-channel allocation problem.

In [60], cross-layer based MAC protocols are proposed to allow SUs to share the spectrum holes, which are detected by integrated PHY layer spectrum-sensing policies. The authors in [61] minimized SU throughput variance in a single-user CRN. A spectrum overlay sharing method based distributed subchannels, bits, and power allocation has been proposed in [62] for an OFDM-based CRN. Subchannels adjacent to PU bands are assumed not to be used by SUs. As a result, the authors in [62] do not consider cross-channel or co-channel interference. A non-active PU bands access based spectrum overlay sharing was considered in [63], [64] for resource allocation in OFDM-based CR systems.

3.6 Chapter Summary

The cognitive radio networks is a new emerging fields. The applications and fundamental issues of the Cognitive radio networks are presented in this chapter. Then, an overview of the various types of coexistence and spectrum sharing scenarios is also given. In addition, related works, i.e., previous resource allocation algorithms for OFDM-based CRNs are also discussed. Few studies focused on multi-cell multi-user CRNs, especially with the utilization of hybrid overlay/underlay spectrum sharing method.

In this dissertation, considering both spectrum overlay and spectrum underlay sharing methods, two new subchannel allocation and power control algorithms are proposed for single-cell and multi-cell multi-user CRNs, respectively.

Chapter 4

Cross-layer Resource Allocation with QoS Support for Multi-user Cognitive Radio Networks

A cross-layer resource allocation algorithm for a single-cell multi-user cognitive radio network is presented in this chapter. To implement the coexistence of a primary cell and the cognitive radio cell, a primary-assistance based coexistent architecture is proposed. In the coexistent system, primary base station (PBS) will provide interference margins of PUs to cognitive base station (CBS) for secondary power control and interference management. In the proposed algorithm, joint resource allocation and interference avoidance with QoS support for multiple SUs is considered. The objectives are formulated as a constrained two-variable nonlinear optimization problem (OP), which is solved by using convex optimization.

This chapter is organized as follows. Section 4.1 is a brief introduction of background and the proposed algorithm. Related works are presented in Section 4.2. In Section 4.3, system models and related assumptions are described, which include the system architecture, the wireless propagation model, and the interference to SUs. The constrained resource allocation problem is defined in Section 4.4. A joint cross-layer optimization is elaborately considered in Section 4.5. The efficiency and accuracy of the proposed method is verified by computer simulations, which are presented in Section 4.6. The results show that the joint cross-layer design has significant improvement compared to two conventional designs. Compared to the spectrum overlay sharing, the spectrum underlay sharing can provide a substantial performance improvement due to the higher spectrum efficiency. Finally, Sect. 4.7 concludes this chapter.

4.1 Introduction

Cognitive devices like SUs can use their increased flexibility and awareness to control channel access dynamically, i.e., to achieve a dynamic spectrum access. Even though the basic idea of CR is simple, the efficient design of CRNs imposes new challenges compared to the conventional wireless networks.

Considering the two new issues mentioned in Sect. 1.3, the resource allocation problem in CRNs involves mutual interference (MI) [65] between PUs and SUs, QoS, fairness, and so on. In this chapter, to flexibly implement spectrum sharing between PUs and SUs, the coexistence of an OFDMA-based single-cell multiuser CRN and a single-cell PN is studied. To enhance the spectrum efficiency, a dynamic resource allocation (DRA) algorithm for multiple SUs in the CRN is proposed.

So far, the coexistence and optimization of a multiuser cognitive radio (MCR) network considering the MI, QoS support and the different spectrum sharing schemes still have not been well studied. Several technical difficulties are involved. First, SU-to-PU interference as well as PU-to-SU interference exist in the CRN/PN. Furthermore, the interference information should be obtained using very limited information. The CRN has to maximize the sum rate of all SUs and, at the same time, make sure that the SU-to-PU interference at each PU receiver does not exceed a limit. Second, to account for the MI, limited transmission power and satisfactory QoS, a large number of constraints are involved in the optimization procedure. Simplified and fast update algorithms are needed.

For the CRN/PN coexistence, we consider a novel infrastructure-based dynamic system architecture, in which the CRN can be either independent of or overlapped with the primary cell. Moreover, a primary-assistance based joint spectrum underlay/overlay method is proposed for the spectrum sharing and real-time SU-to-PU interference control. The PBS determines the interference limits at each PU receiver based on its target performance, such as predefined signal to interference plus noise ratio (SINR), system outage probability, and so on. Then, the PBS broadcasts the interference limits and pilot signals for SU-to-PU interference channel estimation. According to the interference limits and geographic location of the CRN, the CBS decides available spectrum resources in the CRN and utilizes adaptive power control to limit the SU-to-PU interference.

For DRA in the CRN, we propose a Lagrangian duality-based optimization framework under transmit power and QoS constraints for downlink transmissions. Our considered scenario can be modelled as a constrained two-variable non-linear optimization problem. In order to solve the problem and achieve our objectives, we develop near-optimal and low-complexity problems. Based on the transmission power of the CBS and the interference limits of PU receivers, a joint power control and interference avoidance method is analyzed to simplify the constraints and guarantee the performance in the PS with priority. Then, a cross-layer design and the Lagrangian duality based technique are considered to transform the QoS requirements in MAC layer to PHY-layer, so as to provide QoS support for the SUs during each scheduling time. Finally, a modified iterative water-filling (IWF) algorithm is implemented to solve the problem.

4.2 Related Works

As we have introduced in Sect. 3.5, various resource allocation methods have been developed for OFDM-based CRNs. In [45], [54], the fair resource allocation of subcarrier, bit, and power in PHY layer to maximize the system throughput while guaranteeing the interference power limited is studied for OFDM-based CRNs. However, these algorithms cannot dynamically adjust their rate requirement to different SUs. Moreover, in [54], the authors assume that the PN and the CRN are both OFDM-based systems. It is impractical that the PN always uses OFDM. Currently, to the best of our knowledge, there are few studies on QoS support in OFDMA-based CRNs. The QoS designs in [60] and [61] for CRNs only considered non-real-time applications.

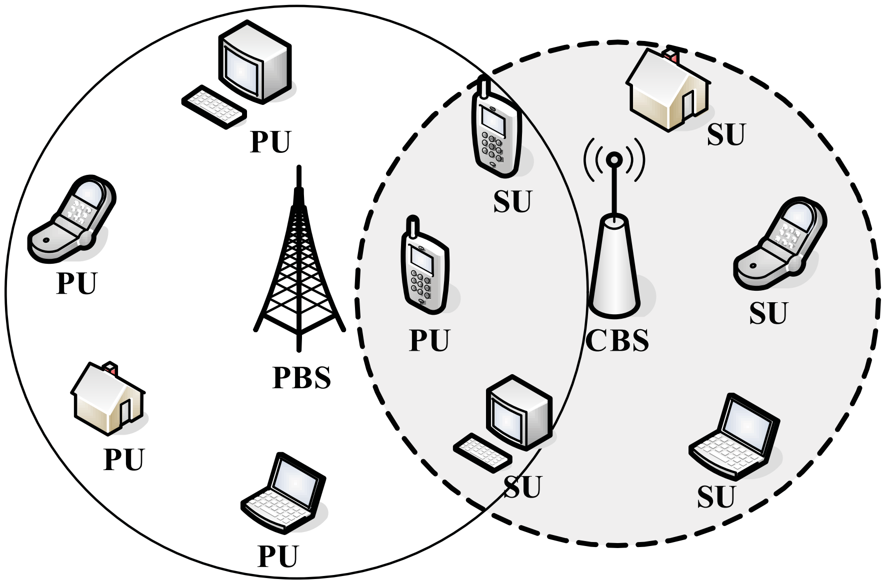

Figure 4.1: An example of system model.

4.3 System Model and Interference to SUs

4.3.1 OFDMA-based Multiuser Cognitive Radio Network

We consider an OFDMA-based MCR network with K SUs and a CBS as the controller to share the spectrum with a PN, which is also an infrastructure-based cellular system with one PBS and N PUs, see Figure 4.1.

In practical applications, it is possible that uplink transmission power of PUs is too small. Then, the SUs cannot access primary bands in order to protect the PUs. Therefore, with considering the feasibility of the coexistent architecture, we assume that, in the PN, uplink and downlink transmissions use the time-division duplex (TDD) mode, meanwhile, in the MCR network, the frequency-division duplexing (FDD) is employed. Hence, in the worst case, both uplinks and downlinks of the MCR network can access primary bands when the PN is on downlink transmissions.

In this study, we consider downlink transmissions in both the CR cell and primary cell. The whole spectrum, which is originally licensed to PUs, is divided into M subcarriers in the CRN. Active subbands allocated to the PUs and OFDM subcarriers of the SUs are shown in Figure 4.2, where the bandwidth of subband

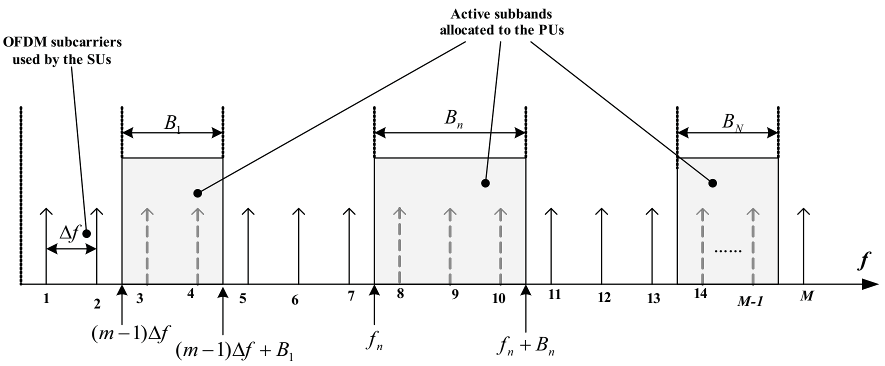

Figure 4.2: Active subbands of PUs and OFDM subcarriers of SUs

n is denoted by B_n, which is allocated to PU n, and Δf is the subcarrier spacing in the CRN. The time slot duration in the CRN is equal to one OFDM symbol period T_s and the subcarriers are modelled in discrete time with the time-varying gain. The set of the SUs, the PUs, the subcarriers and the time slots are denoted by $\mathcal{K} = \{1, \ldots, K\}$, $\mathcal{N} = \{1, \ldots, N\}$, $\mathcal{M} = \{1, \ldots, M\}$, and $\mathcal{T} = \{1, \ldots, T\}$, respectively.

4.3.2 Physical Layer and Propagation Model

For all the links in the MCR network, the channels are subject to frequency selective fading. The channel gain is given by:

$$G_{k,m,t}{=}g_{k,m,t} \cdot \mu(d_k/d_0)^{-\alpha}, \quad \forall k \in \mathcal{K}, \forall m \in \mathcal{M}, \forall t \in \mathcal{T}, \tag{4.1}$$

where $\sqrt{g_{k,m,t}}$ is the small-scale fading being modelled as a Rayleigh distribution. μ is the free-space factor of the channel gain, which can be calculated from $\mu{=} G_s(\frac{\lambda}{4\pi d_0})^2$, where G_s denotes the transmit and receive antenna gain, λ is the wave length, d_0 is a reference distance set to be $d_0 = 100$ m [66]. d_k denotes the distance between the CBS and SU k. α is the path loss exponent. For any link, the power gains $G_{k,m,t}$ are independent and identically distributed (i.i.d.) random variables. Furthermore, we assume the channel is block fading, i.e., $g_{k,m,t}$ is fixed during each time slot, which is much longer than the total duration of information collecting and reporting.

4.3.3 Interference to SUs

In the coexistent system, we consider two kinds of MI: the interference from the PBS to SUs and that from the CBS to PUs. The first interference is discussed here. The second one will be introduced in Sect. 4.4.1.

In the CRN, we assume that the interference from primary transmitters to SUs can be measured properly. Therefore, after the interference limits and the pilot signals from the PBS have been collected, the channel information between two systems can be known at the CBS.

According to Figure 4.2, the interference power from the primary transmitters (i.e., the PBS on downlink, or PUs on uplink) to SU k at subcarrier m and time slot t can be given by

$$I_{k,m,t} = \begin{cases} P_p^{(d)} G_{k,m,t}^{ps} \int_{(m-1)\Delta f}^{m\Delta f} \Phi_p(f)df, & \texttt{if downlinks in the PN} \\ \sum_{n=1}^{N} P_n^{(u)} G_{k,m,t}^{n} \int_{(m-1)\Delta f}^{m\Delta f} \Phi_n(f)df, & \texttt{if uplinks in the PN}, \end{cases} \tag{4.2}$$

where $P_p^{(d)}$ is the downlink transmit power at the PBS, $G_{k,m,t}^{ps}$ is the power gain from the PBS to SU k, $\Phi_p(f)$ is the equivalent baseband power spectral density (PSD) of the PBS's signal when the transmission power is normalized to one watt. $P_n^{(u)}$ is the uplink transmit power at PU n, $G_{k,m,t}^{n}$ is the power gain from PU n to SU k, $\Phi_n(f)$ is the normalized equivalent baseband PSD of the PU n's signal.

We now make further assumptions about channel gain information. We assume that the transmission power and PSD $\Phi_p(f)$ and $\Phi_n(f)$ are known to SUs. Based on these, the SUs can estimate the mean channel gains from the primary transmitters to themselves. Due to the reciprocal characteristic of the wireless channel, the mean channel gains from the SUs to the primary transmitters would be equal to these values. Similarly, the mean channel gains from the CBS to primary transmitters also can be estimated.

4.4 Constraints and Problem Definition

4.4.1 Power Limits for SUs

In the study of this chapter, we use a predefined SINR value γ_p and an interference violation probability $\delta^{(p)}$ together as the primary targets. First, the received SINR at each PU must be no less than the predefined value γ_p. Let $\gamma_{n,t}$ denote the SINR experienced by PU n at subband n. Therefore, we must have:

$$\gamma_{n,t} = \frac{P_{n,t}^{PBS} G_{n,t}^{p}}{N_0 + I_n} \geq \gamma_p, \quad \forall n \in \mathcal{N}, \tag{4.3}$$

where $P_{n,t}^{PBS}$ is the transmission power from the PBS to PU n. $G_{n,t}^{p}$ is the channel gain between the PBS and PU n, which is frequency selective over subband n. N_0 is the complex Gaussian noise power. I_n is the interference from the CBS to PU n. According to (4.3), the interference limit I_n^{max} of PU n can be obtained as follows:

$$I_n^{max} = \frac{P_{n,t}^{PBS} G_{n,t}^{p}}{\gamma_p} - N_0, \tag{4.4}$$

where the set $\{I_n^{max}, n \in \mathcal{N}\}$ is the interference power limits at the PUs.

Assume that B_n is a multiple of Δf, which is from frequency f_n to frequency $f_n + B_n$, as shown in Figure 4.2. Let x be the beginning subcarrier index of subband n, so $f_n = (x-1)\Delta f$. The SU-to-PU interference I_n, which should be no larger than the interference limit I_n^{max}, can be given by:

$$I_n = \sum_{m=x}^{S+x} P_{k,m,t} G_{n,m,t}^{sp} \int_{(m-1)\Delta f}^{m\Delta f} \Phi_s(f) df \leq I_n^{max}, \tag{4.5}$$

where S is the total secondary subcarrier number by which PU subband n is affected, $P_{k,m,t}$ is the allocated transmission power from the CBS to SU k on subcarrier m and time slot t, $G_{n,m,t}^{sp}$ is the channel gain from the CBS to PU n on subcarrier m. $\Phi_s(f)$ is the normalized equivalent baseband PSD of the secondary OFDM signal. Here, we only consider the main lobe power of the OFDM signal, because the interference power caused by the side lobes of the OFDM signal is very low, only 4.922% of the transmit power without multiplying the path loss [65]. Hence, we can assume $P_{k,m,t} \int_{(m-1)\Delta f}^{m\Delta f} \Phi_s(f) df \approx P_{k,m,t}$ for simplicity. Moreover, as long as the interference power on any subcarrier m has been controlled no larger than the average limit I_n^{max}/S, the predefined SINR at each PU can be guaranteed. Let $I_m^{lim} = I_n^{max}/S$ denote the interference limit for subcarrier m. Therefore, Eq. (4.5) can be rewritten as:

$$I_m = P_{k,m,t} G_{n,m,t}^{sp} \leq I_m^{lim}, \tag{4.6}$$

where $\{I_m^{lim} = (I_n^{max}/S)^+, n \in \mathcal{N}, m \in \mathcal{M}\}$ is a mapping of the interference limits from primary subband n to S secondary subcarriers, where, $(x)^+ = \mathtt{max}(0, x)$.

In the CRN, the adaptive power control is used to manage the interference from the CBS to PUs. Since the instantaneous interference level I_n may exceed the tolerable limit I_n^{max}, and violate the absolute interference constraint, we define the interference violation probability as $\mathbf{Pr}\{I_n > I_n^{max}\}$, where $\mathbf{Pr}\{A\}$ denotes the probability of event A, which should be no larger than the value $\delta^{(p)}$. Therefore, considering the simplification from (4.5) to (4.6), there is a constraint on the interference violation probability $\delta^{(p)}$ as follows:

$$\mathbf{Pr}\{P_{k,m,t} G_{n,m,t}^{sp} > I_m^{lim}\} \leq \delta^{(p)}, \quad \forall m \in \mathcal{M}. \tag{4.7}$$

Similar to Eq.(1), $G_{n,m,t}^{sp} = g_{n,m,t}^{sp} \cdot \mu(d_{n,t}^{sp}/d_0)^{-\alpha}$, where $g_{n,m,t}^{sp}$ is the small scale fading between the CBS and PU n on subcarrier m and has been characterized as a Rayleigh distribution with the probability density function (PDF) $f(x;\sigma) = \frac{x}{\sigma^2} e^{(-x^2/2\sigma^2)}$ and the cumulative distribution function (CDF) $F(x) = 1 - e^{(-x^2/2\sigma^2)}$, where $x \in [0, \infty), \sigma > 0$. $d_{n,t}^{sp}$ is the distance between the CBS and PU n.

From (4.7), we have the following proposition for power control at the CBS.

Propositon 1: At the CBS, the allocated power for each subcarrier should be controlled no larger than the following value:

$$P_{m,t}^{max} = \frac{I_m^{lim}}{\mu(d_{n,t}^{sp}/d_0)^{-\alpha} F^{-1}(1 - \delta^{(p)})}, \quad \forall m \in \mathcal{M}, \tag{4.8}$$

where $F^{-1}(\cdot)$ is the inverse function of the CDF of the Rayleigh distribution. Assume that location information of the PUs is available to the CBS, where a resource allocation algorithm is executed. A variety of location-awareness techniques are introduced in [67] and the references therein. The location-based primary protection and resource allocation methods can be found in [68],[70].

However, the methods proposed in [68] and [70] cannot guarantee the QoS for each SU, and they ignored the small-scale fading totally.

Proof: In (4.7), the CDF of $g^{sp}_{n,m,t}$ is given by $F(x) = \mathbf{Pr}\{X \leq x\} = 1 - e^{(-x^2/2\sigma^2)}$, $x \in [0, \infty), \sigma > 0$. Then, for Eq. (4.7), there is the following derivation:

$$\begin{aligned}
&\mathbf{Pr}\{P_{k,m,t} g^{sp}_{n,m,t} \cdot \mu(d^{sp}_{n,t}/d_0)^{-\alpha} > I^{lim}_m\} \\
&= \mathbf{Pr}\{g^{sp}_{n,m,t} > \frac{I^{lim}_m}{P_{k,m,t}\mu(d^{sp}_{n,t}/d_0)^{-\alpha}}\} \\
&= 1 - \mathbf{Pr}\{g^{sp}_{n,m,t} \leq \frac{I^{lim}_m}{P_{k,m,t}\mu(d^{sp}_{n,t}/d_0)^{-\alpha}}\} \\
&= 1 - F(\frac{I^{lim}_m}{P_{k,m,t}\mu(d^{sp}_{n,t}/d_0)^{-\alpha}}) \leq \delta^{(p)}. \qquad (4.9)
\end{aligned}$$

Then, from (4.9), we has the following inequality:

$$F(\frac{I^{lim}_m}{P_{k,m,t}\mu(d^{sp}_{n,t}/d_0)^{-\alpha}}) \geq 1 - \delta^{(p)}. \qquad (4.10)$$

Due to the monotone non-decreasing property of cumulative distribution functions, from (4.10), the following power constraint on subcarrier m can be obtained:

$$\frac{I^{lim}_m}{P_{k,m,t}\mu(d^{sp}_{n,t}/d_0)^{-\alpha}} \geq F^{-1}(1 - \delta^{(p)}). \qquad (4.11)$$

From (4.11), the power limits for secondary subcarriers can be derived based on the primary performance targets and the channel characteristics:

$$P_{k,m,t} \leq \frac{I^{lim}_m}{\mu(d^{sp}_{n,t}/d_0)^{-\alpha} F^{-1}(1 - \delta^{(p)})}, \qquad (4.12)$$

where $P_{k,m,t}$ can be written as $P_{m,t}$ instead.

Then, the interference constraint (4.6) for PUs can be replaced by the following power limit for SUs:

$$P_{k,m,t} \leq P^{max}_{m,t}, \quad \forall m \in \mathcal{M}, \qquad (4.13)$$

where $P^{max}_{m,t}$ is the maximum power that can be allocated to subcarrier m. This power-limited access control method is based on the assumptions and system models in this chapter.

4.4.2 QoS Constraints for SUs

At MAC layer, a cross-layer resource allocation algorithm is proposed to support both real-time (RT) and non-real-time (NRT) services in the CRN. Assume that the first i users from K are with RT service, denoted by $\mathcal{I} = \{1, \ldots, i\}$; and the other K-i users are with NRT service, denoted by $\mathcal{J} = \{i+1, \ldots, K\}$.

RT Service is the services such as MPEG or streaming video or audio. It provides guarantees on throughput and latency, that is, each packet, which has a length of l_i^{RT}, needs to be received by SU i within d_i^{RT} time slots after the packet has been transmitted.

NRT Service provides guarantees on throughput, can tolerate longer delays, and is insensitive to delay jitter. So it is suitable for FTP applications. Its average data rate that the system needs to provide is R_j^{NRT}.

The instantaneous rate for SU k at subcarrier m and time slot t can be given by:

$$R_{k,m,t} = \Delta f \log_2(1 + \beta_{k,m,t} P_{k,m,t}), \quad \forall k \in \mathcal{K}, \tag{4.14}$$

where $\beta_{k,m,t} = \frac{G_{k,m,t}}{N_0 + I_{k,m,t}}$, $I_{k,m,t}$ is defined in Section 2.3, $P_{k,m,t}$ is limited by Eq. (4.13).

In order to provide satisfactory QoS for the SUs, there are following constraints for different services:

$$t_i^D - t_i^S \leq d_i^{RT}, \quad \forall i \in \mathcal{I}, \tag{4.15}$$

$$\sum_{t=1}^{d_i^{RT}} \sum_{m=1}^{M} R_{i,m,t} \geq l_i^{RT}, \quad \forall i \in \mathcal{I}, m \in \mathcal{M}, t \in \mathcal{T}, \tag{4.16}$$

$$\bar{R}_j = \frac{1}{T} \sum_{t=1}^{T} \sum_{m=1}^{M} R_{j,m,t} \geq R_j^{NRT}, \quad \forall j \in \mathcal{J}, m \in \mathcal{M}, t \in \mathcal{T}, \tag{4.17}$$

where for RT SUs, t_i^D is the arriving time slot when the data arrived at the destination (i.e. SU i), and t_i^S is the sending time slot from the source node (i.e. the CBS). Inequality (4.15) indicates that the duration time slots from the CBS to SU i should be no larger than the delay of RT SUs d_i^{RT}. In (4.16), the packet length l_i^{RT} needs to be received by SU i within d_i^{RT} time slots, that is , Equation (4.16) is another form of (4.15). Equation (4.17) is the MAC-layer QoS constraint for NRT SUs, where, $\bar{R}_j$ is the average rate of SU j from time slot 1 to time slot T, and should be no smaller than R_j^{NRT}.

4.4.3 Optimization Problem

In this study, our objective is to maximize the system throughput under several constraints while ensuring that the RT services can be provided within their specified deadlines, as well as the average data rates for NRT can satisfy the requirements. Let $\mathcal{L} = \{L_{k,m,t}, k \in \mathcal{K}, m \in \mathcal{M}, t \in \mathcal{T}\}$ denote the allocation results of the continuous instantaneous M subcarriers in the time slot t. Therefore, based on the above system models, the optimization problem (OP) can be formulated as follows:

OP-1:

$$max \quad \sum_{t=1}^{T}\sum_{m=1}^{M}\sum_{k=1}^{K} R_{k,m,t}L_{k,m,t} \tag{4.18}$$

s.t.

$$t_i^D - t_i^S \leq d_i^{RT}, \tag{4.19}$$

$$\sum_{t=1}^{d_i^{RT}}\sum_{m=1}^{M} R_{i,m,t} \geq l_i^{RT}, \tag{4.20}$$

$$\bar{R}_j = \frac{1}{T}\sum_{t=1}^{T}\sum_{m=1}^{M} R_{j,m,t} \geq R_j^{NRT}, \tag{4.21}$$

$$P_{k,m,t} \leq P_{m,t}^{max}, \tag{4.22}$$

$$\sum_{k=1}^{K}\sum_{m=1}^{M} P_{k,m,t} \leq P_0, \tag{4.23}$$

$$\sum_{k=1}^{K} L_{k,m,t} \leq 1, \quad L_{k,m,t} = \{0, 1\}, \tag{4.24}$$

$$0 \leq R_{k,m,t} \leq R_{max}. \tag{4.25}$$

Constraints (4.19)–(4.22) are introduced before. In (4.23), P_0 is the total allowed power in the CBS per time slot. Equations (4.24) are the constraints to guarantee that there is only one connection on one subcarrier. To protect the PS better, besides the interference limits from the PBS, we consider a maximum link rate R_{max} to limit the transmit power from the CBS further. The (4.25) indicates that the actual link capacity on each subcarrier should be no more than R_{max}, where $R_{max} = r_{max}\Delta f$ and r_{max} is the maximum number of bits that can be allowed per subcarrier.

4.5 Joint Cross-layer Resource Allocation and Interference Avoidance with QoS Support

The OP-1 in Section 4.4.3 is difficult to be solved directly both in mathematically and in practical consideration. Since it involves an optimization over T time slots, M subcarriers, K users and with MAC-layer QoS requirements, there may be many local optimal points and the problem feasibility cannot be guaranteed. In addition, the CBS should be able to allocate the subcarriers and power to K SUs at the beginning of T time slots based on the solution of OP-1, therefore, in OP-1, the knowledge of future channel gains (i.e., the channel gain at future time slots) is required. However, it is impossible for the CBS to obtain future channel information. Therefore, we need to simplify the OP1 and find reasonable approximation, so as to solve the multi-variable resource allocation problem optimally and accurately. In the following section, we formulate a problem OP-2, which is a power minimization problem to find the minimum transmit power that can guarantee the QoS requirements for SUs during τ time slots ($\tau < T$), so as to transform the MAC-layer QoS constraints over T time slots to less time slots.

4.5.1 Transform the MAC-layer QoS Constraints to PHY-layer

Consider the QoS requirements in τ time slots, Eqs. (4.15) and (4.16) can be rewritten as:

$$\sum_{t=1}^{\tau}\sum_{m=1}^{M} R_{i,m,t}L_{i,m,t} \geq r_{i,\tau}^{Req}, \quad \forall i \in \mathcal{I}, \tau \in 1,2,\ldots,T, \tau < d_i^{RT}, \tag{4.26}$$

where τ is the number of time slots considered in the resource allocation algorithm, which is less than T. $r_{i,\tau}^{Req}$ is the packet length that needs to be transmitted in τ time slots for RT SU i.

Eq. (4.17) can be rewritten as:

$$\sum_{t=1}^{\tau}\sum_{m=1}^{M} R_{j,m,t}L_{j,m,t} \geq r_{j,\tau}^{Req}, \quad \forall j \in \mathcal{J}, \tau \in 1,2,\ldots,T, \tag{4.27}$$

where $r_{i,\tau}^{Req}$ is the total bit rate that needs to be transmitted in τ time slots for NRT SU j.

Combining (4.26) and (4.27) together, the QoS requirements in (4.15)–(4.17) can be rewritten as a rate constraint during τ time slots, that is:

$$\sum_{t=1}^{\tau}\sum_{m=1}^{M} R_{k,m,t}L_{k,m,t} \geq r_{k,\tau}^{Req}, \quad \tau \in 1,2,\ldots,T, \tau < d_i^{RT}, \tag{4.28}$$

where $r_{k,\tau}^{Req}$ is the required number of bits that needs to be transmitted in τ time slots.

We consider the minimum required power used for QoS support. There is the following optimization problem:

OP-2:

$$min \quad \sum_{t=1}^{\tau}\sum_{m=1}^{M}\sum_{k=1}^{K} (2^{\frac{R_{k,m,t}}{\Delta f}} - 1)\frac{L_{k,m,t}}{\beta_{k,m,t}} \tag{4.29}$$

s.t.

$$\sum_{t=1}^{\tau}\sum_{m=1}^{M} R_{k,m,t}L_{k,m,t} = r_{k,\tau}^{Req}, \quad \forall k \in \mathcal{K}, \tag{4.30}$$

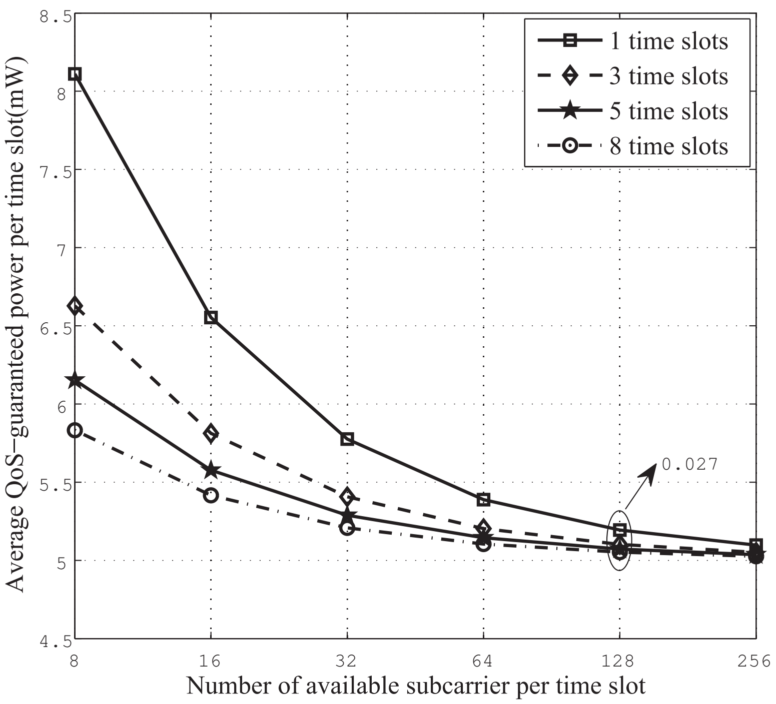

Figure 4.3: The minimum QoS-guaranteed power per time slot when K=3, τ =1, 3, 5, 8

$$\sum_{k=1}^{K} L_{k,m,t} \leq 1, \quad L_{k,m,t} = \{0, 1\}, \tag{4.31}$$

$$0 \leq R_{k,m,t} \leq R_{max}. \tag{4.32}$$

The objective of OP-2 is to minimize the required power and satisfy the QoS requirements at the same time. For the solution of OP-2, there is the following proposition:

Propositon 2: As $M \to \infty$, for $\forall t \in \mathfrak{T}$, the optimal solution $\{R^*_{k,m,t}, L^*_{k,m,t}\}$ for OP-2 satisfies

$$\sum_{m=1}^{M} R^*_{k,m,t} L^*_{k,m,t} = r^{Req}_{k,t}, \quad \forall t \in \mathfrak{T}, \tag{4.33}$$

where $r^{Req}_{k,t} = \begin{cases} l_i^{RT}/d_i^{RT}, & \forall i \in \mathfrak{I} \\ R_j^{NRT} T_s, & \forall j \in \mathfrak{J} \end{cases}$, is the required bit rate at each time slot.

Proof: Details are provided in Appendix A.

In Proposition 2, we assumed that $M \to \infty$, and Equation (4.33) shows that the required rate needed to be transmitted at each time slot is the same for any allocation time. However, in this paper, M is finite, in order to utilize the Proposition 2 to simplify the OP-1, we evaluated the average minimum QoS-guaranteed power per time slot with different τ, which is shown in Figure 4.3.

In this simulation, we assume that all the SUs are with RT service, the subcarrier spacing is 16 KHz, and the required data rate is $R_k = 800$kbps, so the packet length is $r_{k,\tau}^{Req} = R_k T_s \tau$. Let the SU number be 3, and the value of τ is 1, 3, 5, 8 respectively. From Figure 4.3, we can see that for different value of τ , the average QoS-guaranteed power is different. When the number of available subcarriers per time slot M is increased, the effect of τ value over the QoS-guaranteed power is smaller. When $M = 64$, the difference between the values of ordinate is less than 0.05; when M increased to 128, the difference is only 0.027. It means that when M becomes large, enough selectivity and diversity can be obtained for the system to achieve nearly the same minimum power for different τ values. Thus, based on Proposition 2, a reasonable approximation can be provided to transform the QoS constraints during τ time slots to one time slot with a reasonable number of subcarriers M [71].

Therefore, the MAC-layer QoS constraints (4.15)–(4.17) in OP-1 can be replaced by the following PHY-layer constraint in one time slot:

$$\sum_{m=1}^{M} R_{k,m,t} L_{k,m,t} \geq r_k^{Req}. \tag{4.34}$$

Assume that there are enough subcarriers to satisfy the requirements of multiple services at each time slot. In order to derive the solutions, we introduced the surplus variable λ_k [18] into the constraint (4.34), then, (4.34) is replaced by:

$$\sum_{m=1}^{M} R_{k,m,t} L_{k,m,t} = r_k^{Req} + \lambda_k, \quad \lambda_k \geq 0, \tag{4.35}$$

where λ_k is the surplus variable for SU k and represents the amount by which the total allocated rate is exceeded. Moreover, the larger the value of λ_k, the higher the value of the system throughput; thus, λ_k should be maximized in each time slot, then, the system utility function (4.18) in OP-1 can be replaced by the maximization of λ_k.

4.5.2 QoS-guaranteed Resource Allocation Algorithm

Due to the analysis and the simplifications above, we can describe the algorithm as following. Since the problem only depends on the parameters in the current time slot, the time index t can be removed for simplicity. Therefore, the OP of joint cross-layer resource allocation (RA) can be described as:

OP-3:

$$max \quad \sum_{k=1}^{K} \lambda_k \tag{4.36}$$

s.t.

$$0 \leq P_{k,m} \leq P_m^{max}, \quad \sum_{k=1}^{K}\sum_{m=1}^{M} P_{k,m} \leq P_0, \tag{4.37}$$

$$\sum_{k=1}^{K} L_{k,m} \leq 1, \quad L_{k,m} = \{0,1\}, \tag{4.38}$$

$$0 \leq R_{k,m} \leq R_{max}, \tag{4.39}$$

$$\sum_{m=1}^{M} R_{k,m} L_{k,m} = r_k^{Req} + \lambda_k, \quad \lambda_k \geq 0. \tag{4.40}$$

OP-3 is a constrained nonlinear programming problem, and in general, is intractable. It is shown in Appendix B that Problem OP-3 can be decomposed into three subproblems. The following optimal solution can be obatined using Lagrangian duality based technique [18], [72].

Let $R^*_{k,m}, P^*_{k,m}, L^*_{k,m}, k^*, \lambda^*_k$, $\xi^*_{k,m}$,φ^*, η^*_k, ς^*_k be an optimal solution set, where $\xi_{k,m}, \varphi, \eta_k$, ς_k, $k \in \mathcal{K}$, $m \in \mathcal{M}$ are the non-negative Lagrange multipliers [18] (see details in Appendix B). Here, $R_{k,m}$ is replaced by $\widehat{R}_{k,m}\Delta f$, so, $\widehat{R}_{max} = R_{max}/\Delta f$ in the following equations.

Solution S^****:*** The optimal solution $S^* = \{P^*_{k,m}, L^*_{k,m}, k^*\}$ for Problem OP-3 has the following properties.

1)For a given SU k, if the subcarrier allocation $L^*_{k,m} = 1$, the optimal power allocation strategy is:

$$P^*_{k,m} = \begin{cases} 0, & \omega_k < \frac{\ln(2)}{\Delta f \beta_{k,m}} \\ \frac{\omega_k \Delta f}{\ln(2)} - \frac{1}{\beta_{k,m}}, & \frac{\ln(2)}{\Delta f \beta_{k,m}} \le \omega_k \le \frac{2^{\widehat{R}_{max}} \ln(2)}{\Delta f \beta_{k,m}} \\ \frac{2^{\widehat{R}_{max}} - 1}{\beta_{k,m}}, & \omega_k > \frac{2^{\widehat{R}_{max}} \ln(2)}{\Delta f \beta_{k,m}} \end{cases} \tag{4.41}$$

where $\omega_k = \frac{\eta^*_k L^*_{k,m}}{\xi^*_{k,m} + \varphi^*}$, represents the update of the multipliers and can be viewed as the iterative water-filling (IWF) level for SU k and will be discussed later. From (4.41), the rate allocation $R^*_{k,m}$ and $\widehat{R}^*_{k,m}$ can be obtained by using (4.14).

2)The subcarrier allocation strategy for subcarrier m is:

$$L^*_{k,m} = \begin{cases} 1, & k = k^* \ \texttt{and} \ m \in \mathcal{M} \\ 0, & \texttt{otherwise} \end{cases} \tag{4.42}$$

$$k^* = \texttt{argmax} \quad \eta^*_k \widehat{R}^*_{k,m} \tag{4.43}$$

For a given t and m, if the values of $\eta^*_k \widehat{R}^*_{k,m}$ are the same for multiple SUs, we will choose one SU arbitrarily.

3)IWF level ω_k is:

$$\omega_k = \begin{cases} \omega_B, & \lambda^*_k > 0 \\ (1 + \varsigma^*_k)\omega_B, & \lambda^*_k = 0 \end{cases} \tag{4.44}$$

where $\omega_B = 1/(\xi^*_{k,m} + \varphi^*)$, $\xi^*_{k,m} > 0$, $\varphi^* > 0$.

Proof: Details are provided in Appendix B.

We can know from the above optimal values, if the water-filling levels $\{\omega_k\}$ for all SUs can be found, the optimal power and subcarrier allocation are obtained from (4.41) and (4.42).

In the system, we assume the primary bandwidth is large enough for SUs to guarantee the QoS, that is, according to Figure 4.3, $M > 64$. There are two phases in our algorithm. The first one is to provide QoS support for all SUs. Then, it is the system throughput maximization. According to (4.41) and (4.14), the higher the value of water-filling level ω_k, the higher the number of allocated bit rate for SU k. So, at the starting point, ω_B should be set to be the lowest water-filling level among all SUs. At each iteration, ω_B is increased to the next level that is higher than before. Based on the water-filling levels, the allocation algorithm is performed according to (4.41), (4.42) and (4.43). The process will stop when $\sum_{k=1}^{K}\sum_{m=1}^{M} P_{k,m}^{*} \geq P_0$ or the subcarriers have been used out. If $\sum_{k=1}^{K}\sum_{m=1}^{M} P_{k,m}^{*} > P_0$, the algorithm will use (4.41) to find an appropriate value of ω_k that satisfies the KKT condition (B.9), see Appendix B.

A brief description of the procedure is given as follows:

(i) Initialize the water-filling level ω_k, ω_B.

(ii) In Phase I, for $m = 1, ..., M$, do the following:

- For $k = 1, ..., K$, do subcarrier and power allocation using (4.41), (4.42) and (4.43).
- Check if the QoS requirements have been satisfied. If for all SUs, $\sum_{m=1}^{M} R_{k,m}^{*} L_{k,m}^{*} \geq r_k^{Req}$, go to step (iii). If no, find the SUs whose bit rates less than r_k^{Req} as set K^-, and find the SUs whose bit rates greater than r_k^{Req} as set K^+. Adjust ω_k for SUs in set K^- and K^+, and reset $L_{k,m}^{*} = 0, k \in K^- \cup K^+$, reallocate the left subcarriers and power to SUs in K^- and K^+, until $\sum_{m=1}^{M} R_{k,m}^{*} L_{k,m}^{*} = r_k^{Req}$ or $m = M$, then go to step (iii).

(iii) Check if the transmit power and the power limits have been fulfilled. First, check if $P_{k,m}^{*} \leq P_m^{max}$. If no, update ω_k until $P_{k,m}^{*} \leq P_m^{max}$. Next, compare $\sum_{k=1}^{K}\sum_{m=1}^{M} P_{k,m}^{*}$ with P_0. If $\sum_{k=1}^{K}\sum_{m=1}^{M} P_{k,m}^{*} = P_0$, the algorithm will be finished. If $\sum_{k=1}^{K}\sum_{m=1}^{M} P_{k,m}^{*} < P_0$, go to step (iv). If $\sum_{k=1}^{K}\sum_{m=1}^{M} P_{k,m}^{*} > P_0$, adjust the base water-filling level ω_B to a smaller one, then back to step (ii).

(iv) In Phase II, the power has not been used out. We need to adjust ω_B to a higher value, and allocate all the resource to the SUs using (4.41), (4.42) and (4.43) to maximize the system throughput.

4.6 Simulation Results

In this section, simulations are performed for the downlink OFDMA-based MCR network to evaluate the effectiveness of the proposed algorithm. The simulation

Table 4.1: SIMULATION PARAMETERS

Parameters	**Value**
Number of PUs N	8
Number of SUs M	10
System Bandwidth B_w	2 MHz
System center frequency f_c	1.9 GHz
Total power at the PBS	10 W
Total power at the CBS P_0	13 W
Predefined SINR of PUs in dB γ_p^{dB}	10 dB
Interference violation probability $\sigma^{(p)}$	0.01
antenna gain G_s	8 dBi
Number of subcarriers M	128
Path loss exponent α	4
Number of RT SUs I	5
Number of NRT SUs J	5
Delay of RT service $d_i^{RT}(time\ slots)$	90
Symbol period T_s	40 μs
Real-time Data Rate $R^{RT}(kbps)$	[100,600]
Required average data rate R_j^{NRT}	300 kbps
Max.number of bits per subcarrier r_{max}	8 bits

parameters are summarized in Table 4.1. The simulation area is 2km*2km with the CBS located at the positions d_{sp} far away from the PBS, where d_{sp} is the distance between two BSs and varies from 0 to 1km. It is assumed that the channel gain is constant during 1ms periods, thus resource allocation is performed once every 1ms, which is also called one scheduling time. The WINNER Phase II channel model [69] is utilized to implement the channels in the simulations. We assume the wireless propagation environment is urban area, and the cell radius is 500m. PUs and SUs are randomly located in its own cell area at each scheduling time. All performance results are obtained by over 1000 simulation runs.

Moreover, we assume the PS is always downlink transmissions in the simulation. So, the impact of TDD frame ratio will not be considered here, since the performance with considering the frame ratio is proportional to the evaluate value, and not difficult to deduce.

4.6.1 Effectiveness of the Proposed Design

Firstly, we consider the worst case that all the primary bandwidth has been uniformly allocated to PUs, and each with 250 kHz bandwidth, that means there is no non-active PU bands in the PS. The transmit power at the PBS is uniformly allocated to 8 PUs. In this case, if the CRN is overlapped with primary system, the interference power limits I_m^{lim} may be very small, even zero. In order to share the bandwidth with the PS, the distance between two BSs d_{sp} should be large enough, then, the SU-to-PU interference can be controlled. Here, we set d_{sp} =1000 m at first.

For comparison, we study two conventional resource allocation schemes: channel greedy with power control and proportional fairness (PF) with equal power. For traditional OFDMA systems, channel greedy scheduling with water-filling/equal power allocation in [73], and PF scheduling with water-fill/equal power allocation in [74], have been proposed. However, OFDMA-based CRNs are different from traditional ones. Here, we would like to compare the proposed design to these two schemes to indicate that our proposed scheme is more suitable for CRNs. The first conventional scheme allocates the subcarriers to the SU who has the largest SINR on the considered subcarrier, and allocates the power based on the power limits to control the interference. The second one assigns the subcarriers uniformly to all SUs for fairness, and equal power to all subcarriers.

Figures 4.4 and 4.5 are cumulative probability of achieved SINR and relative interference power of PU 1, which is located at the cell edge of the PS, respectively. The relative interference power is defined as I_{siml}/I_n^{max}, where I_{siml} is the simulated interference power. $10\log_{10}(I_n^{max}/I_{siml})$ (dB) is the dB-value of the reciprocal of relative interference. From Figures 4.4 and 4.5, we can see that the PF scheme cannot provide the PUs with predefined SINR because of the high interference. Compared to Greedy scheme and PF scheme, the proposed method

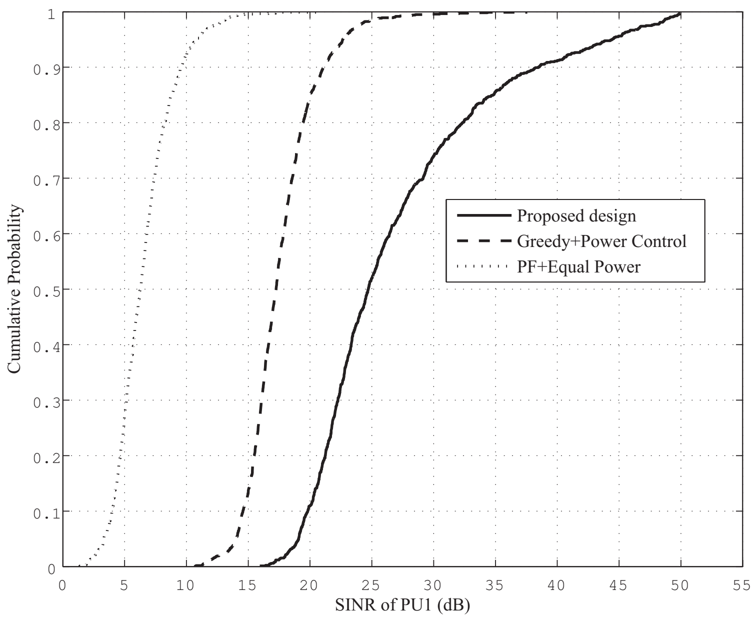

Figure 4.4: Cumulative probability of SINR of PU 1, R^{RT}=600kbps

not only achieves the highest SINR, which is much higher than predefined value, but also controls the SU-to-PU interference well.

In Figures 4.6 and 4.7, the simulation results of QoS support for RT and NRT SUs are shown respectively. In Figure 4.6, we set d_2 = 500m, d_5 = 100m, where, d_2 and d_5 are the distances from SU 2 and SU 5 to the CBS, respectively. We can see from this figure that SU 2, which is the cell edge user with great channel fading, can only obtain the basic QoS-guaranteed data rate 600kbps; however, SU 5, which is near to the CBS with good channel state, can achieve much higher rate so as to maximize the system throughput. Figure 4.7 shows the average data rate of NRT SUs versus different RT data rate. Compared to conventinal schemes, the proposed design can achieve much higher average data rate.

Figure 4.8 shows the sum rate of the CRN. Compared to the other two resource allocation schemes, it yields a significant higher sum rate. For different RT data rate, the sum rate of CRN is almost the same due to the same available CBS power and spectrum resource. For the Greedy scheme, even though it can control the SU-to-PU interference, the spectrum efficiency is really low. From Figure 4.8,

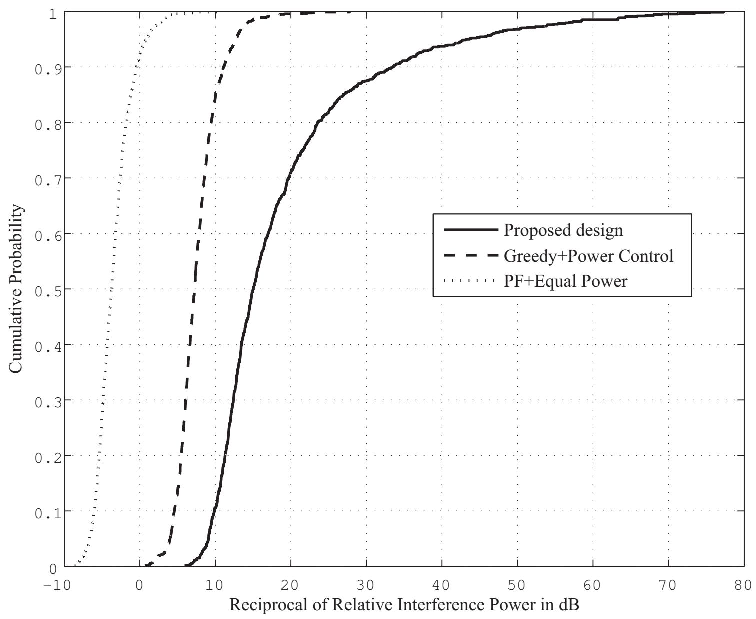

Figure 4.5: Cumulative probability of the reciprocal of relative interference of PU 1 in dB: $10\log_{10}(I_n^{max}/I_{siml})$, R^{RT}=600kbps

we can see the proposed design has the best performance.

4.6.2 Performance Comparision between Spectrum Underlay and Overlay

In practical application, two networks may be overlapped with each other, and the primary bandwidth may not be used out, that is, the probability of non-active PU bands satisfies $P_{non\text{–}active}\geq 0$.

When CBS and PBS are too close to each other and the primary bands are all in use, it is difficult to control the mutual interference with power control, or to provide satisfactory QoS to all SUs. Figure 4.9 is the CDF of λ_k at different CBS-PBS distance d_{sp}, which has been changed from 350m to 650m. Here, the simulation parameters are the same with Section 4.6.1 and $P_{non\text{–}active} = 0$. We define $\mathbf{Pr}\{\lambda_k < 0\} = \delta^{(s)}$, which is the secondary QoS-unsatisfactory probability. When $\delta^{(s)} > 0.01$, we consider that the CRN cannot guarantee the QoS for all SUs. From Figure 4.9, we can see that in order to limit the interference at the PUs and

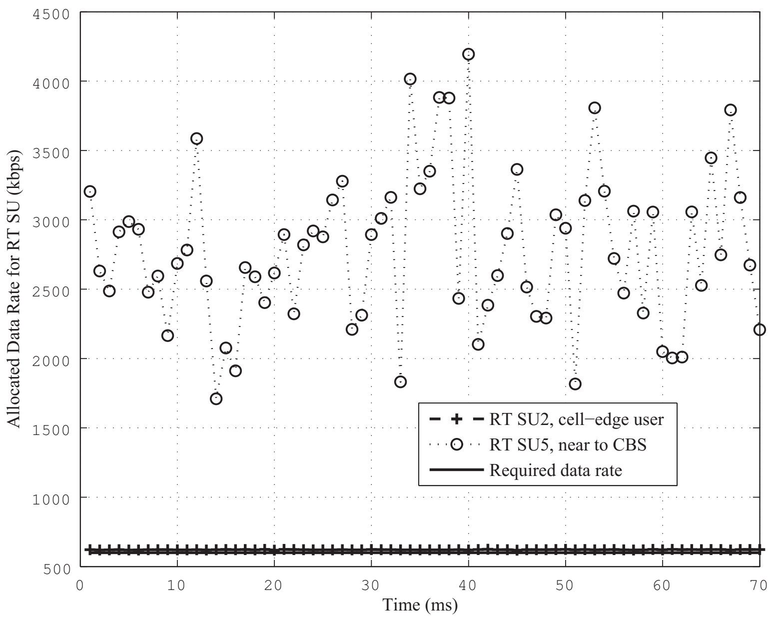

Figure 4.6: QoS support for RT SUs

provide the QoS for 10 SUs as well, d_{sp} cannot be very small. When $d_{sp}=650$m, $\delta^{(s)}$ is only 0.005; while d_{sp} decreased to 550m, $\delta^{(s)}$ increased to 0.029. Then, we consider that when $d_{sp}=550$m the CRN cannot satisfy the requirements for all SUs at the same time. In this situation, we should choose other better sharing methods for the coexistence.

Therefore, in order to improve the effectiveness of the coexistent system, we compared the performance between spectrum underlay and spectrum overlay sharing methods.

Figures 4.10 and 4.11 are the average maximum number of SUs and the sum rate of the CRN, respectively. In these two figures, we consider two different sharing schemes: spectrum overlay (only available non-active subbands can be utilized), and spectrum underlay & overlay (the whole primary bands can be utilized, and control the interference level to the active subbands). We set the probability of non-active PU bands to be $P_{non\text{–}active}=0$ or 20%. When $P_{non\text{–}active}=0$, only the spectrum underlay sharing can be utilized for the CRN; while, when $P_{non\text{–}active}=20\%$, both the spectrum underlay and spectrum overlay are available

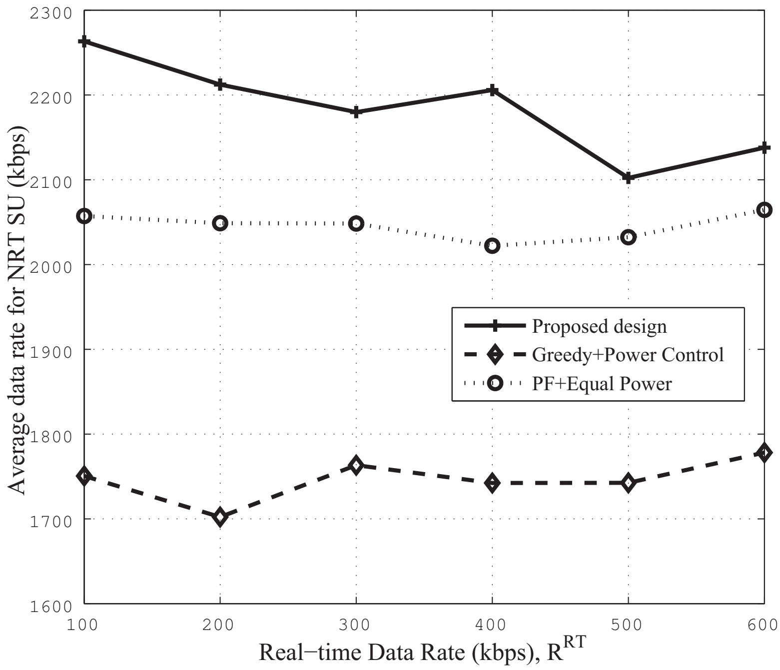

Figure 4.7: Average throughput for NRT SUs

for the CRN to share the spectrum with PS.

From Figures 4.10 and 4.11, we can see that the maximum RT/NRT SU number and the sum rate of the CRN are increasing with the distance between the CBS and the PBS when utilizing the spectrum underlay scheme for both $P_{non\text{-}active}$=0 and 20%. However, the performance of the spectrum overlay sharing scheme is almost the same, and only related to $P_{non\text{-}active}$, which is the available spectrum resource. Compared the performance of these sharing schemes in Figures 4.10 and 4.11, the spectrum underlay has substantial higher spectrum efficiency and can be utilized for both $P_{non\text{-}active}$=0 and 20%. When $P_{non\text{-}active}>0$, the spectrum underlay & overlay sharing method is the best candidate for CR systems to access unlicensed spectrum.

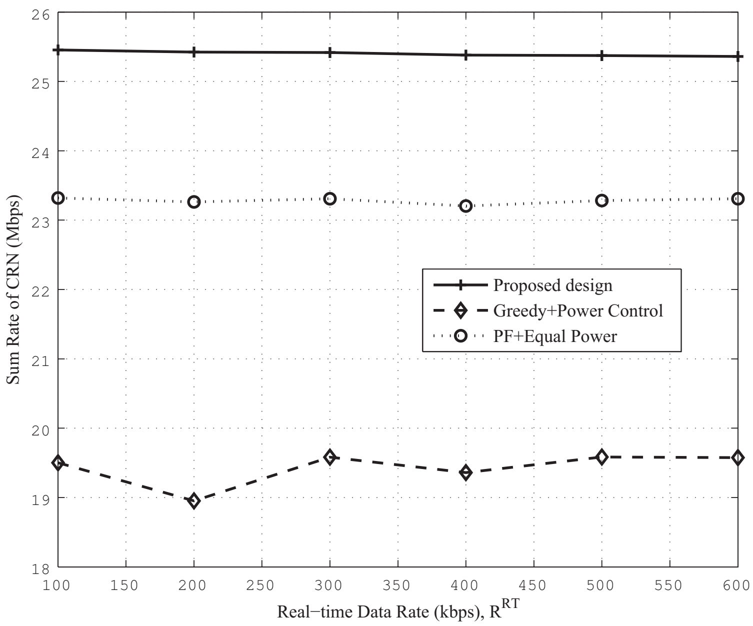

Figure 4.8: Sum rate of CRN

4.7 Chapter Summary

A combined cross-layer resource allocation and interference avoidance optimization design for downlink OFDMA-based MCR networks has been proposed. We utilize a predefined SINR and an interference violation probability at the primary receivers for the power allocation and interference control. QoS constraints transformation and IWF method have been analyzed to maximize the system throughput for the CRN and provide satisfactory QoS support for different services of the SUs. To obtain optimal solution, we developed a convex optimization problem to solve the system utility function.

Compared to the conventional resource allocation schemes, the proposed cross-layer design with the spectrum underlay sharing method could share the spectrum with the PUs more effectively. Simulation results illustrate that our proposed design has a significant performance gain. On the other hand, the comparision between the spectrum underlay and overlay has shown that if there are available subbands in the PN, it has the best spectrum efficiency with both the spectrum underlay and overlay sharing methods.

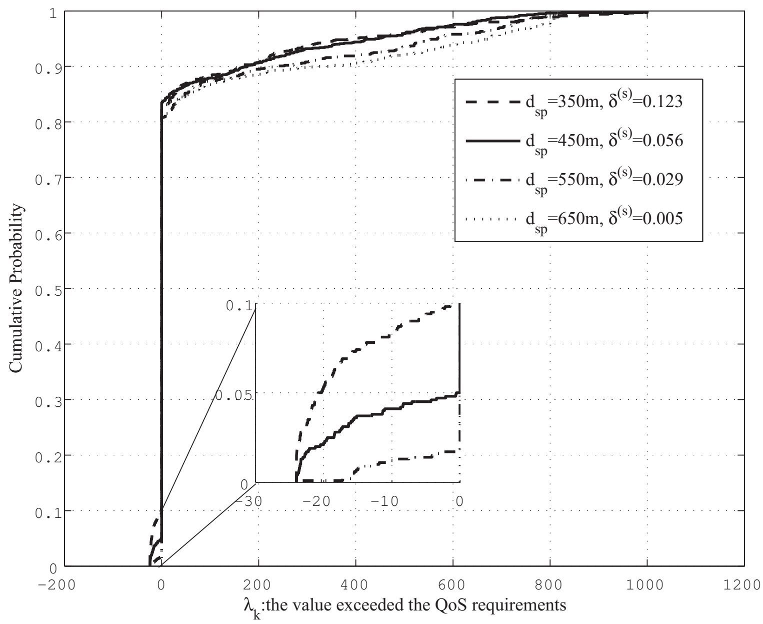

Figure 4.9: Cumulative probability of the value exceeded the QoS requirements λ_k, R^{RT}=600kbps

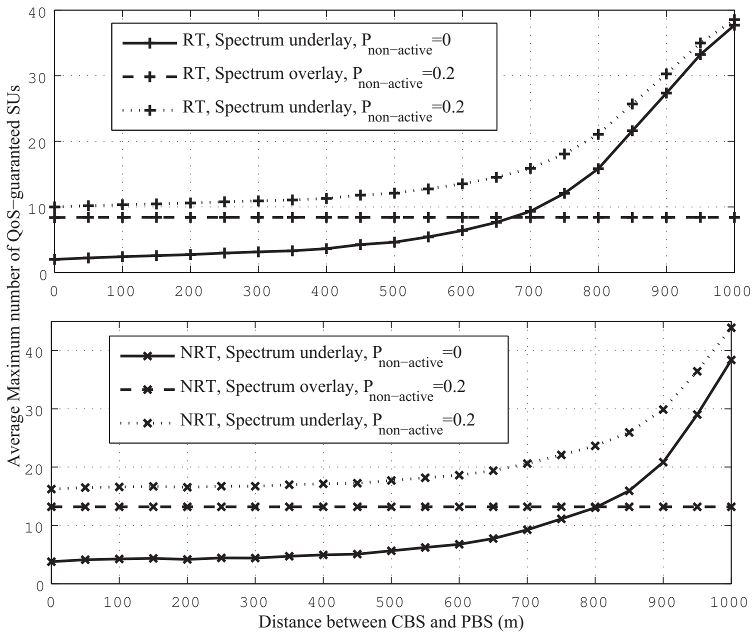

Figure 4.10: Maximum number of SUs with different sharing schemes, R^{RT}=600kbps

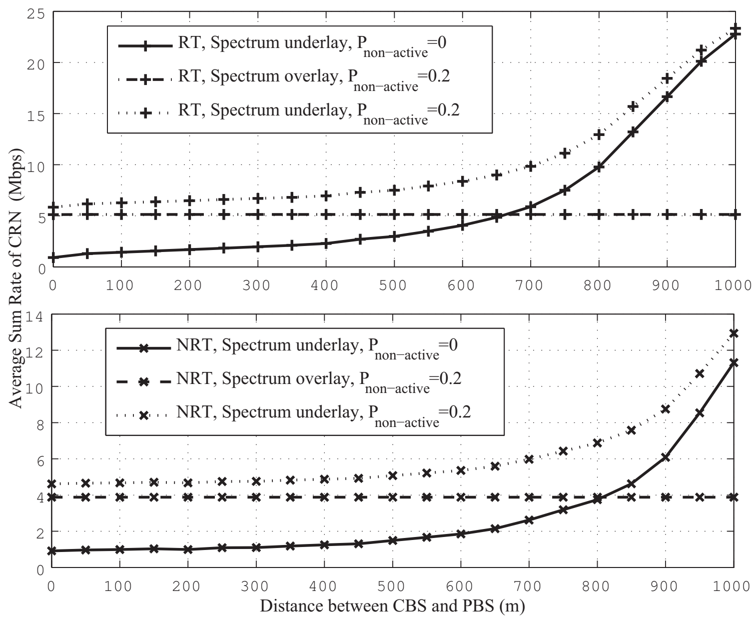

Figure 4.11: Sum rate of CRN with different sharing schemes, R^{RT}=600kbps

Chapter 5

Distributed Resource Allocation for Multi-cell Cognitive Radio Networks

In this chapter, we consider a multi-cell cognitive radio network, which overlays a multi-cell primary network. To manage the coexistence, a primary-willingness based coexistent architecture and a novel intra-cell spectrum overlay and inter-cell spectrum underlay sharing method are proposed. In the system, primary base stations will broadcast pilot signals and interference margins to assist the CRN for interference channel evaluation and power control. Subject to the interference margins imposed by the primary network, we define a utility (payoff) function that can represent the secondary system performance while taking into account the co-channel interference among secondary cells. A distributed resource allocation scheme is devised to guarantee the primary performance, and at the same time, maximize the secondary utility without any cooperation among cognitive base stations. QoS among SUs is also considered by the scheme such that the instantaneous data rate for each secondary user is larger than a given minimum rate. The resource allocation problem can be decomposed into two subproblems: subchannel allocation and distributed power allocation game (DPAG). We prove that there exists a Nash equilibrium in the DPAG and the equilibrium is unique. Moreover, the DPAG is also Pareto optimal in some constrained environments, that is, no CBS can further improve its performance without impairing others.

The proposed algorithm turns out to converge to an equilibrium within a small number of iterations.

This chapter is organized as follows. Related works are described in section 5.2. System model and problem description are defined in section 5.3. Section 5.4 explains the power control and interference avoidance in the CRN. In section 5.5, a distributed resource allocation is defined. The existance, uniqueness, and pareto optimality of DPAG are proved, and a distributed resource allocation algorithm is also described in this section. Simulation results are shown in sect. 5.6. Finally, section 5.7 summarizes the chapter.

5.1 Introduction

As we have introduced before, cognitive radio is utilized to alleviate the severe spectrum shortage problem by making it possible for SUs to share frequency bands with PUs in some geographical locations. The SUs can access the licensed frequency using either spectrum overlay or spectrum underlay sharing methods.

Efficient designs for CRNs imposes new challenges compared to conventional wireless systems. To implement a practical multi-cell CRN, one of the major issues is that the secondary utilization should not degrade the service in PN. The other one is how to maximize the performance by controlling the co-channel interference among the neighboring cells.

Here, we consider an OFDMA-based multi-cell CRN, which overlays with a multi-cell PN. Both spectrum overlay and underlay sharing methods are utilized to encourage CRN/PN coexistence. Each CR cell utilizes different subchannels from the colocated primary cell (i.e., intra-cell spectrum overlay), but the same subchannels allocated to the neighboring primary cells (i.e., inter-cell spectrum underlay). To obtain necessary information from PNs, a primary-willingness based intra-cell overlay and inter-cell underlay method is proposed. First, PBSs determine the interference margins that can be accepted at each primary receiver based on its targets, such as predefined SINR, outage probability, etc. Then, the PBSs broadcast the interference margins on occupied subchannels and pilot signals for secondary-to-primary interference estimation. According to the information collected from the PBSs and its own SUs, each CBS decides subchannel and power allocation for its own SUs distributedly.

In this chapter, a noncooperative scheme for the downlink resource allocation in multi-cell CRNs is presented. We first define a utility (payoff) function that represents the sum data rate and the power consumption in a cell. Then, the problem is decomposed into two subproblems: subchannel allocation and distributed power allocation game. The defined utility can be maximized under the total power constraint at the CBS, interference margins at primary receivers, and QoS constraint for each SU, in a distributed manner. In this scheme, a CBS in each cell individually controls the assignment of subchannels to the SUs and the power allocation to each subchannel to maximize its utility. For the power allocation game, the existence and uniqueness of Nash Equilibrium (NE) point are investigated. Moreover, there is no CBS can further improve its performance without impairing others in some condition (i.e., Pareto optimality). The proposed algorithm turns out to converge to an equilibrium within a small number of iterations, and can substantially improve the total utility without requiring any coordination among CBSs.

5.2 Related Works

In the related works introduced in Sect. 3.5, many previous studies on resource allocation methods for CRNs have assumed the single-cell or ad hoc model. The interference limited underlay spectrum sharing resource allocation techniques for OFDMA-based single-cell CRNs have been presented in [47],[49],[54],[56]. In [47], the weighted sum rate for several SU links was optimized using Lagrangian duality tool and both centralized and distributed algorithms were designed. The authors in [49],[54] focus on throughput maximization in a centralized manner. In [56], two distributed fair subcarrier and power allocation schemes by using Colonel Blotto game for both the downlink and uplink of CRNs have been proposed. A white space access based distributed subchannels, bits, and power allocation was proposed in [62] for an ad hoc or multi-cell CRN. Hybrid overlay/underlay systems were considered in [75],[76], where SUs were classified into several sets to determine whether overlay or underlay should be applied, and optimal power allocation and access control for ad hoc CRNs was studied. However, these schemes are not applicable for multi-cell multi-user utilization. Moreover, the hybrid overlay/underlay in [75],[76] is different from the intra-cell overlay and inter-cell underlay introduced above, which is proposed specially for the multi-cell CRN/PN coexistence in this paper. In [77], a multiple access channel is assumed for a two-cell CRN model. Based on the spectrum overlay sharing, a Nash equilibrium based sum rate maximization resource allocation algorithm has been proposed. The authors in [78] considered the weighted sum rate optimization for multi-cell CRNs in a centralized way. The inter-cell iterative water-filling method has been implemented to control the inter-cell interference. However, such an approach induces signaling overhead and also requires efforts for cell planning. Therefore, in multi-cell environment, distributed operation is a preferred approach for the co-channel interference management.

5.3 System Model and Problem Description

5.3.1 System Model and Assumptions

We consider a cellular CRN in this chapter, which is supposed to be deployed over cellular PNs in the following ways. First, the future broadband cellular technology, such as 3rd Generation Partnership Project (3GPP) Long Term Evolution Advanced (LTE-Advanced), requires up to 100 MHz per channel [79], but the amount of spectrum is limited. The CRN can enable bandwidth aggregation by sharing spectrum owned by other cellular operators to solve this data explosion problem. Second, in the current cellular networks, the BS has only a RF unit. A digital unit for all communication functionalities is implemented in a separate central server [80]. As a result, the cost of BSs will be cheap enough for anybody to install anywhere. This allows a new type of a mobile virtual operator based on CR, which can operate its own BSs in a current primary cell without spectrum licenses. Therefore, an example of CRN/PN coexistent model is shown in Figure 5.1. The similar network architectures also can be found in [54] and [56].

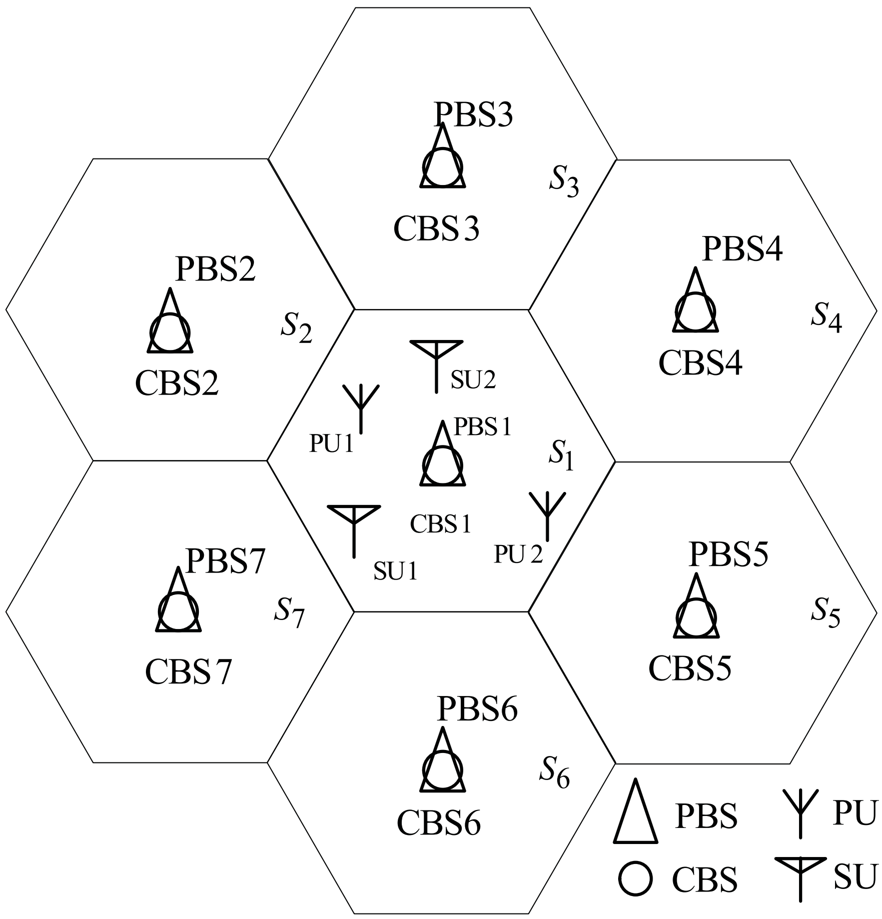

Figure 5.1: System model.

Assume that a CRN can recognize some essential information about a PN in advance, e.g., range of frequency bandwidth, frequency reuse factor, frequency allocation, etc. A primary-willingness based coexistent architecture is considered, so, PBSs will report interference margins and pilot signals to CBSs periodically for power control. This can be easily implemented among wired infrastructure-based base stations, especially when CBSs are colocated with PBSs. The detail of how to use pilot signals are out of the scope of this paper, which has been introduced in [81].

Both CRN and PN are assumed that the time-division duplexing (TDD) mode is employed for uplink and downlink transmissions. We consider uplink transmissions in primary cells (P-cells) and downlink transmissions in cognitive radio cells

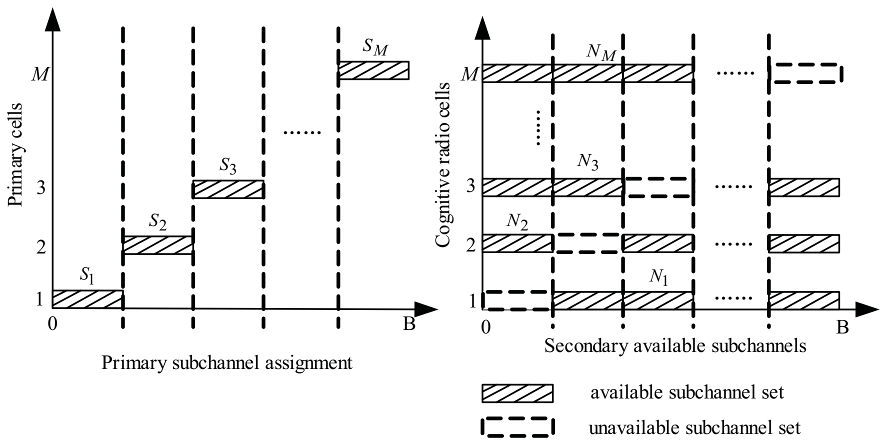

Figure 5.2: Available subchannels for P-cells and CR-cells.

(CR-cells), because it is much easier to evaluate the interference level when PBSs as primary receivers and CBSs as secondary transmitters due to their fixed locations. We assume that all CR-cells are in a quasi-synchronous mode and either all in uplink or all in downlink.

A M-cell system model is given in Figure 5.1, $M = 7$. Each cell has one CBS, one PBS, K SUs, and J PUs. The OFDMA technique is used for both CRN and PN. Therefore, the intra-cell interference in the network can be neglected by using different subchannels, but the inter-cell interferences (CBS-to-PBS, CBS-to-SU and PU-to-SU interferences) need to be considered. Besides, only the interferences from first-tier neighboring cells are taken into account and that from far-away neighbors will be viewed as an additive white Gaussian noise (AWGN) [66]. Therefore, this one-tier system model can be easily extended to multiple tiers situation.

Here, we assume that the total bandwidth B is divided into N subchannels, which are assigned to M P-cells fixedly. The non-overlapping subchannel sets for M P-cells are $S_1, S_2, \cdots, S_M$, respectively. As introduced in Sect. 1, the available subchannel set for CR-cell m can be denoted by $N_m = \sum_{m'=1, m' \neq m}^{M} S_{m'}$. The primary subchannel assignment and the secondary available subchannels for each CR-cell are shown in Figure 5.2. From this figure, we can see that CR-cells can avoid generating intra-cell interference to the colocated P-cell by using subchannels allocated to other P-cells.

We denote by $\mathbf{p}^m = (p_1^m, \cdots, p_{N_m}^m)$ the transmission power vector of CBS m, with p_n^m denoting the transmission power on subchannel n at CBS m. $\mathbf{P}=[\mathbf{p}^1\mathbf{p}^2\cdots\mathbf{p}^M]$ is the network power vector, a concatenation of the transmission power matrices of the M CBSs. We let $\mathbf{P}^{-m}$ denote the interference power vector of CBS m. We assume that the total transmission power ($i.e.$, $\sum_{n=1}^{N_m} p_n^m$) of each CBS is constrained to be less than P_{max}^m, which is the maximum power of CBS m. Since each subchannel can only be assigned to one SU, let $\mathbf{A}^m=[a_{k,n}^m]_{K\times N}$ denote the subchannel allocation matrix, where $a_{k,n}^m$ is 1 if subchannel n is assigned to SU k and 0 otherwise.

Let $g_{k,n}^m$ denote the channel gain between SU k and CBS m on subchannel n. Then the SINR of SU k in cell m for a given power vector $\mathbf{P}$ can be expressed by

$$\gamma_{k,n}^m(\mathbf{P}) = \frac{g_{k,n}^m p_n^m}{I_n^{m'}(\mathbf{P}^{-m}) + N_0}, \tag{5.1}$$

where $I_n^{m'}(\mathbf{P}^{-m})=\sum_{m'=1,m'\neq m}^{M}\left(g_n^{j\to k}p_{j,n}^{m'}+g_n^{m'\to k}p_n^{m'}\right)$ is the interference on subchannel n from primary and secondary transmitters in cell m', $g_n^{j\to k}$ and $g_n^{m'\to k}$ are interference channel gains from PU j and CBS m'. $p_{j,n}^{m'}$ and $p_n^{m'}$ are the primary and secondary transmission power. N_0 is the background complex Gaussian noise power on subchannel n.

The achievable data rate of SU k is given by

$$R_{k,n}^m(\mathbf{P}) = B_n \log_2\left(1+\gamma_{k,n}^m(\mathbf{P})\right), \tag{5.2}$$

where B_n is the bandwidth of each subchannel.

The total data rate allocated to SU k, which not only depends on the power allocation in other cells $\mathbf{P}^{-m}$, but also the subchannel allocation $\mathbf{A}^m$ and the power allocation in the corresponding cell $\mathbf{p}^m$, is given as follows:

$$R_k(\mathbf{P}, \mathbf{A}^m) = \sum_{n=1}^{N_m} a_{k,n}^m R_{k,n}^m(\mathbf{P}). \tag{5.3}$$

5.3.2 Problem Description

The authors in [54],[49]-[78] only considered the sum data rate maximization. If the transmission power of the CBS increases, the sum data rate increases, but it causes an increase of co-channel interference in the neighboring cells too. Consequently, the transmission power increase leads to conflicting interests among multiple CR-cells. Therefore, it is not enough for considering the sum data rate as utility function only. Here, the sum data rate is viewed as the reward obtained by consuming power resource. The allocated total power is as the cost. The utility function is defined to be the reward minus the cost. In order to protect the primary performance and guarantee the QoS of SUs, the utility function is subject to several constraints. The allocated power on each subchannel should be less than the power limit. QoS in this paper is defined as achieving data rate for each SU no less than R^k_{min}, the minimum data rate. Therefore, we can formulate the problem as follows:

$$\max_{\mathbf{P},\mathbf{A}^m} \; u_m\left(\mathbf{P},\mathbf{A}^m\right) = \sum_{k=1}^{K} R_k(\mathbf{P},\mathbf{A}^m) - c\sum_{n=1}^{N_m} P_n^m, \tag{5.4}$$

subject to

$$\sum_{n=1}^{N_m} a^m_{k,n} R^m_{k,n}(\mathbf{P}) \geq R^k_{min}, \tag{5.5}$$

$$\sum_{k=1}^{K} a^m_{k,n} - 1 \leq 0, \;\; a^m_{k,n} = \{0,1\}, \tag{5.6}$$

$$p^m_n \leq p^{lim}_{n,m}, \tag{5.7}$$

$$\sum_{n=1}^{N_m} p^m_n \leq P^m_{max}, \tag{5.8}$$

where c is the price per unit power, having the unit bps/W. The power price represents the cost imposed on each CBS for the co-channel interference generated by it. Therefore, the utility function (5.4) encourages each CBS to maximize the sum data rate but using minimum power, i.e., causing minimum co-channel interference to other cells. Equation (5.5) guarantees a minimum level of QoS requested by SUs. Note that the allocated data rate for each SU depends on the network power vector $\mathbf{P}$. Constraints in (5.6) are to guarantee that there is only one connection on one subchannel. In Eq. (5.7), $p_{n,m}^{lim}$ is the power limit on subchannel n, which will be described in Sect. 3. The allocated transmission power on subchannel n should be less than this limit. The total transmission power of CBS is limited as constraint (5.8).

To solve (5.4)-(5.8) by centralized algorithms, all the channel information is required, especially the interference channel information shown in (5.1), which causes computational complexity and large amount of channel estimation overhead when the number of cells and number of users are large. In a multi-cell CRN, noncooperation between CBSs can be a realistic situation. In the subsequent sections, we are going to solve this problem in a distributed manner by using a game theoretical approach.

5.4 Power Control and Interference Avoidance

In this section, we will introduce the location-based [68] power control and interference avoidance method, to avoid CBS-to-PBS interference and guarantee the performance in primary networks.

As we have mentioned before, a primary-willingness based coexistent architecture is considered. So, firstly, each PBS will dertermine its interference margins based on its targets. In this paper, we consider a predefined SINR value γ_p and an interference violation probability $\delta^{(p)}$(i.e., primary SINR outage probability), as primary targets. Considering the uplink transmissions in P-cell $m^{'}$. The received SINR experienced by PBS $m^{'}$ on subchannel n should be no less than the predefined SINR γ_p. That is:

$$\gamma_{j,n}^{m^{'}} = \frac{g_{j,n}^{m^{'}} p_{j,n}^{m^{'}}}{\sum_{m=1,m\neq m^{'}}^{M} p_n^m g_n^{m\to m^{'}} + N_0} \geq \gamma_p, \tag{5.9}$$

where $g_{j,n}^{m^{'}}$, $p_{j,n}^{m^{'}}$ are the channel gain and transmission power from PU j to PBS $m^{'}$. p_n^m is the transmission power on subchannel n at CBS m, $g_n^{m\to m^{'}}$ is the interference channel gain between CBS m and PBS $m^{'}$.

We define the interference margin level on subchannel n at PBS $m^{'}$ as $I^{th}_{n,m^{'}}$. From (5.9), we have the following equation:

$$\sum_{m=1,m\neq m^{'}}^{M} p_n^m g_n^{m\to m^{'}} \leq I^{th}_{n,m^{'}} = \frac{g_{j,n}^{m^{'}} p_{j,n}^{m^{'}}}{\gamma_p} - N_0. \tag{5.10}$$

From (5.10), the PBSs can estimate the interference margin on each subchannel to guarantee the predefined SINR γ_p. Then, the interference margins $\{I^{th}_{n,m^{'}}\}_{n,m^{'}}$ on all subchannels at each PBS will be broadcast to the CRN for power control at CBSs.

Equation (5.10) is a sum interference constraint for all neighboring CR-cells around PBS $m^{'}$. To obatin individual power limits for each neighboring CBS, we assume that the power limit is the same for all neighboring cells. Therefore, from (5.10), for CR-cell m, on subchannel n, we have:

$$p_n^m g_n^{m\to m^{'}} \leq \frac{I^{th}_{n,m^{'}}}{M-1}, \quad m \in [1, M], m \neq m^{'}. \tag{5.11}$$

If the above constraint has been guaranteed in any neighboring CR-cell m, the interference margin at PBS $m^{'}$ will not be exceeded. However, it is difficult to satisfy (5.11) perfectly. Therefore, to avoid the instantaneous interference level exceeding the interference margin, we need to control the interference violation probability no larger than $\delta^{(p)}$. We formulate as follows:

$$\mathbf{Pr}\left\{ p_n^m g_n^{m\to m^{'}} > \frac{I^{th}_{n,m^{'}}}{M-1} \right\} \leq \delta^{(p)}, \tag{5.12}$$

where $\mathbf{Pr}\{\mathrm{A}\}$ denotes the probability of event A.

Here, we consider the average interference channel gain to neglect the small-scale fading in $g_n^{m\to m^{'}}$, therefore, the interference channel gain can be written as $g_n^{m\to m^{'}} = \mathrm{PL} \cdot \varphi^{-1} = \mu(d^{m\to m^{'}}/d_0)^{-\chi}\varphi^{-1}$ [66], where, PL denotes the distance-dependent path loss part, μ is a constant, depends on the antenna characteristics and the average channel attenuation; $d^{m\to m^{'}}$ is the distance between secondary transmitter CBS and primary receiver PBS; d_0 is a reference distance; χ is the path loss exponent; φ is a log-normal distributed random variable with mean zero and variance σ_φ^2, denotes the shadowing fading part. So, $p_n^m g_n^{m\to m^{'}}$ in Eq. (5.12) can be considered as a log-normal distribution due to shadowing effect. The left part of Eq. (5.12) can be written as:

$$\mathbf{Pr}\left\{p_n^m g_n^{m\to m'} > \frac{I_{n,m'}^{th}}{M-1}\right\}$$

$$= 1 - Q\left(\frac{p_{n\,dB}^m + \mathrm{PL}_{dB} - I_{n,m'\,dB}^{th} + 10\log_{10}(M-1)}{\sigma_{\varphi_{dB}}}\right), \tag{5.13}$$

where $Q(x) = \frac{1}{\sqrt{2\pi}}\int_x^{\infty} exp\left(\frac{-t^2}{2}\right) dt$ is the Gaussian-Q function and a monotone decreasing function. Therefore, from (5.12) and (5.13), the dB-valued power limit on subchannel n of CR-cell m can be calculated as:

$$p_{n\ dB}^{m,lim} = I_{n,m'\,dB}^{th} - 10\log_{10}(M-1) - \mathrm{PL}_{dB} + \sigma_{\varphi_{dB}} Q^{-1}(1-\delta^{(p)}), \tag{5.14}$$

where $I_{n,m'}^{th}$ is the interference margin received from PBSs. The distance $d^{m\to m'}$ in PL can be evaluated from pilot signals. $Q^{-1}(\cdot)$ is the inverse-Q function.

Here, from the dB-valued power limit $p_{n\ dB}^{m,lim}$, we can obtain the power limit $p_{n,m}^{lim}$ mentioned in (5.7).

5.5 Distributed Resource Allocation

From Sect. 5.3.2, the resource allocation problem is a constrained nonlinear programming problem, and is intractable. To simplify the problem, it can be decomposed into two subproblems in each CR-cell: subchannel allocation and power allocation. In (5.4), each CBS maximizes its own utility function subject to the constraints (5.5)-(5.8) for a given interference power from all the other cells. Here, we consider the interferences in a distributed manner, without any cooperation among CBSs, that is, the interference power is obtained according to the feedback information from its SUs in each cell.

5.5.1 Subchannel Allocation

First, we consider the problem of optimizing the subchannel assignment for a given network power vector $\mathbf{P}_0$, which is given by

$$\max_{\mathbf{A}^m} \; u_m\left(\mathbf{P}_0,\mathbf{A}^m\right)=\sum_{k=1}^{K} R_k(\mathbf{P}_0,\mathbf{A}^m), \tag{5.15}$$

which is subject to (5.5) and (5.6). Compared to (5.4), note that the cost term is suppressed as it depends only on the transmission power vector, which is the same for all SUs in the same CR-cell. Moreover, the constraints (5.7) and (5.8) are also unnecessary for the given transmission power vector $\mathbf{p}_0^m$.

The problem in (5.15) is a multi-user nonlinear optimization problem with equality and inequality constraints. In order to utilize the method of Lagrangian multipliers and Karush-Kuhn-Tucker (KKT) conditions [18] to find the optimal solutions, the constraints $a_{k,n}^m = \{0,1\}$ in (5.6) should be relaxed as $0 \leq a_{k,n}^m \leq 1$. So, the Lagrangian function associated with the optimization problem (5.15) can be written as:

$$\begin{aligned}
& L_m\left(\mathbf{P}_0, \mathbf{A}^m, \phi_k^m, \psi_n^m, \zeta_{k,n}^m, \upsilon_{k,n}^m\right) \\
&= \sum_{k=1}^{K}\sum_{n=1}^{N_m} a_{k,n}^m R_{k,n}^m(\mathbf{P}_0) + \sum_{k=1}^{K}\phi_k^m\left(\sum_{n=1}^{N_m} a_{k,n}^m R_{k,n}^m(\mathbf{P}_0) - R_{min}^k\right) \\
&- \sum_{n=1}^{N_m}\psi_n^m\left(\sum_{k=1}^{K} a_{k,n}^m - 1\right) - \sum_{k=1}^{K}\sum_{n=1}^{N_m}\zeta_{k,n}^m\left(a_{k,n}^m - 1\right) + \sum_{k=1}^{K}\sum_{n=1}^{N_m}\upsilon_{k,n}^m a_{k,n}^m \\
&= \sum_{k=1}^{K}\sum_{n=1}^{N_m} a_{k,n}^m\left((1+\phi_k^m)R_{k,n}^m(\mathbf{P}_0) - \psi_n^m - \zeta_{k,n}^m + \upsilon_{k,n}^m\right) \\
&- \sum_{k=1}^{K}\phi_k^m R_{min}^k + \sum_{n=1}^{N_m}\psi_n^m + \sum_{k=1}^{K}\sum_{n=1}^{N_m}\zeta_{k,n}^m,
\end{aligned} \tag{5.16}$$

where $\phi_k^m, \psi_n^m, \zeta_{k,n}^m, \upsilon_{k,n}^m$ are non-negative Lagrangian multipliers in CR-cell m for the QoS constraint, subchannel allocation constraints, respectively.

Let $a_{k,n}^{m*}, \phi_k^{m*}, \psi_n^{m*}, \zeta_{k,n}^{m*}, \upsilon_{k,n}^{m*}$ be an optimal solution set, then the KKT conditions [18] state that:

$$\phi_k^{m*} \geq 0, \psi_n^{m*} \geq 0, \zeta_{k,n}^{m*} \geq 0, \upsilon_{k,n}^{m*} \geq 0, \tag{5.17}$$

$$\phi_k^{m*} \left(\sum_{n=1}^{N_m} a_{k,n}^{m*} R_{k,n}^m(\mathbf{P}_0) - R_{min}^k \right) = 0, \tag{5.18}$$

$$-\psi_n^{m*} \left(\sum_{k=1}^{K} a_{k,n}^{m*} - 1 \right) = 0, \tag{5.19}$$

$$-\zeta_{k,n}^{m*} \left(a_{k,n}^{m*} - 1 \right) = 0,\ \upsilon_{k,n}^{m*} a_{k,n}^{m*} = 0, \tag{5.20}$$

$$(1 + \phi_k^{m*}) R_{k,n}^m(\mathbf{P}_0) - \psi_n^{m*} - \zeta_{k,n}^{m*} + \upsilon_{k,n}^{m*} = 0, \tag{5.21}$$

where equation (5.21) is obtained by setting $\partial L_m / \partial a_{k,n}^m = 0$.

According to (5.6), for any subchannel n, there is only one SU k^* with a nonzero value of $a_{k^*,n}^{m*} = 1$. Hence, for SU k^*, according to (5.20), we have $\zeta_{k^*,n}^{m*} \geq 0$, and $\upsilon_{k^*,n}^{m*} = 0$, so, from (5.21), we have:

$$\psi_n^{m*} = (1 + \phi_{k^*}^{m*}) R_{k^*,n}^m(\mathbf{P}_0) - \zeta_{k^*,n}^{m*}. \tag{5.22}$$

For any SU $k \neq k^*$, $a_{k,n}^{m*} = 0$, so, according to (5.20), we have $\zeta_{k,n}^{m*} = 0$, and $\upsilon_{k,n}^{m*} \geq 0$, combined with (5.21), we have:

$$\psi_n^{m*} = (1 + \phi_k^{m*}) R_{k,n}^m(\mathbf{P}_0) + \upsilon_{k,n}^{m*}. \tag{5.23}$$

Compare (5.22) and (5.23), we have that $(1+\phi_{k^*}^{m*}) R_{k^*,n}^m(\mathbf{P}_0) \geq (1+\phi_k^{m*}) R_{k,n}^m(\mathbf{P}_0)$, so, for the given network power vector $\mathbf{P}_0$, we can obtain the optimal subchannel allocation strategy by assigning subchannel n to the SU who yields the maximum weighted data rate. That is, for each subchannel

$$a_{k,n}^{m*} = \begin{cases} 1, & \text{if } k^* = \text{argmax}_k (1+\phi_k^{m*}) R_{k,n}^m(\mathbf{P}_0), \\ 0, & \text{otherwise.} \end{cases} \tag{5.24}$$

where ϕ_k^{m*} is the optimal Lagrangian multiplier (i.e., the weight) for the QoS constraint, which can be determined by sub-gradient search method.

5.5.2 Distributed Power Allocation Game

According to (5.24), when the network power vector $\mathbf{P}$ is determined, the optimal subchannel allocation matrix $\mathbf{A}^{*m}(\mathbf{P})$ can be determined. Therefore, for a decided $\mathbf{A}^{*m}(\mathbf{P})$, the distributed resource allocation problem turns into a distributed power allocation, which can be constructed as a distributed power allocation game (DPAG) $G=[\mathcal{M},\{\mathbb{P}^m\},\{u_m(\mathbf{P},\mathbf{A}^{*m}(\mathbf{P}))\}]$, where, $\mathcal{M}=\{1,2,\cdots,M\}$ is the index set of CBSs, i.e., the players, and $\mathbb{P}^m$ is the power allocation strategy space of CBS m, defined by $\mathbb{P}^m=\left\{\mathbf{p}^m|0\le\sum_{n=1}^{N_m}p_n^m\le\min\{\sum_{n=1}^{N_m}p_{n,m}^{lim},P_{max}^m\}\right\}$. So, the DPAG is as follows:

$$\max_{\mathbf{p}^m\in\mathbb{P}^m} u_m\left(\mathbf{p}^m,\mathbf{P}^{-m},\mathbf{A}^{*m}(\mathbf{P})\right)=\sum_{n=1}^{N_m}\max_k\left[(1+\phi_k^{m*})R_{k,n}^m(\mathbf{P})\right]-c\sum_{n=1}^{N_m}p_n^m$$

$$=\sum_{n=1}^{N_m}(1+\phi_{k^*}^{m*})R_{k^*,n}^m(\mathbf{P})-c\sum_{n=1}^{N_m}p_n^m, \tag{5.25}$$

which is subject to (5.7) and (5.8). k^* is the optimal SU on subchannel n.

The game DPAG involving M players based on utility function (5.25) is expected to achieve a Nash Equilibrium (NE), which is defined as:

Definition 1: A strategy vector $\mathbf{P}^*\in\mathbb{P}$ is a Nash equilibrium, if for all players $m\in\mathcal{M}$ and each alternate strategy $\mathbf{q}^m\in\mathbb{P}^m$, we have that

$$u_m\left(\mathbf{p}^m,\mathbf{P}^{*-m},\mathbf{A}^{*m}(\mathbf{P}^*)\right)\ge u_m\left(\mathbf{q}^m,\mathbf{P}^{*-m},\mathbf{A}^{*m}(\mathbf{Q})\right), \tag{5.26}$$

where $\mathbf{Q}=[\mathbf{p}^1\cdots\mathbf{p}^{m-1}\mathbf{q}^m\mathbf{p}^{m+1}\cdots\mathbf{p}^M]$.

In other words, assuming that all other players stick to the strategies they have chosen in $\mathbf{P}^*$, no player m can change its chosen strategy from $\mathbf{p}^m$ to $\mathbf{q}^m$ and thereby improve its utility.

Theorem 1: For the DPAG in (5.25), a Nash equilibrium exists.

Proof: In [23], it is established that if for all $m\in\mathcal{M}$, the following two conditions are satisfied, a Nash equilibrium exists in game $G=[\mathcal{M},\{\mathbb{P}^m\},\{u_m(\mathbf{P})\}]$:

- $\mathbb{P}^m$ is a nonempty, convex and compact subset of some Euclidean space $\Omega^{K\times N_m}$;

- $u_m(\mathbf{P})$ is continuous in $\mathbf{P}$ and quasi-concave in $\mathbf{p}^m$.

For the first condition, the strategy space $\mathbb{P}^m$ of each CBS is defined by a set of power vectors $\mathbf{p}^m$, and all the power value elements in $\mathbf{p}^m$ is between zero and the maximum power limit on each subchannel. Thus, it is obvious that the first condition is satisfied.

From (5.25), the utility function $u_m(\mathbf{P})$ is obviously a continuous function of $\mathbf{P}$. From (5.1) and (5.2), $R^m_{k^*,n}\left(p^m_n, \mathbf{P}^{-m}\right)$ is a monotonically increasing function of p^m_n on subchannel n. So, for any given $\alpha > 0$, the sub-level power set $S^m_n \equiv \{x | R^m_{k^*,n}(x) \geq \alpha\}$ is given by $\{x | x \geq \left(R^m_{k^*,n}\right)^{-1}(\alpha)\}$. Since S^m_n is a convex set, $R^m_{k^*,n}$ is a quasi-concave function of p^m_n. Then, the utility function is a sum of quasi-concave functions in the corresponding p^m_n. Therefore, we can prove that the utility function $u_m(\mathbf{P})$ is quasi-concave in $\mathbf{p}^m$.

The transmission power vector of a CBS that maximizes the utility function in the strategy space is called the *best response* (or, *optimal solution*) to the transmission power chosen by other CBSs. Let $p^{m*}_n(\mathbf{P}^{-m})$ denote the best response of CBS m on subchannel n to a given interference power allocation $\mathbf{P}^{-m}$. As we have proved above that, in the DPAG, the utility is a convex function of $\mathbf{p}^m$, so the problem is also a convex optimization problem. The Lagrangian function associated with the optimization problem (5.25) can be written as:

$$\begin{aligned}
&L_m\left(\mathbf{P}, \mathbf{A}^{m*}, \eta^m_n, \lambda^m\right) \\
&= \sum_{n=1}^{N_m} (1+\phi^{m*}_{k^*}) R^m_{k^*,n}(\mathbf{P}) - c\sum_{n=1}^{N_m} p^m_n - \sum_{n=1}^{N_m} \eta^m_n \left(p^m_n - p^{lim}_{n,m}\right) - \lambda^m \left(\sum_{n=1}^{N_m} p^m_n - P^m_{max}\right) \\
&\qquad = \sum_{n=1}^{N_m} \left((1+\phi^{m*}_{k^*}) R^m_{k^*,n}(\mathbf{P}) - \omega^m_n p^m_n\right) + \sum_{n=1}^{N_m} \eta^m_n p^{lim}_{n,m} + \lambda^m P^m_{max},
\end{aligned} \tag{5.27}$$

where $\omega^m_n = c + \eta^m_n + \lambda^m$. η^m_n, λ^m are non-negative Lagrangian multipliers in CR-cell m for the power limit, and total power constraint.

Therefore, for a given subchannel allocation $\mathbf{A}^m_0$, we can obtain the best response of CBS m as follows, by using the Lagrangian function and KKT conditions [18]. The details are omitted here, please refer to [18] and Sect. 5.5.1.

$$p^{m*}_n(\mathbf{P}^{-m}) = \left[\frac{(1+\phi^{m*}_{k^*}) B_n}{\omega^{m*}_n \ln 2} - \frac{I^{m'}_n(\mathbf{P}^{-m}) + N_0}{g^m_{k^*,n}}\right]^+, \tag{5.28}$$

$$-\eta_n^{m*}\left(p_n^{m*}-p_{n,m}^{lim}\right)=0, \tag{5.29}$$

$$-\lambda^{m*}\left(\sum_{n=1}^{N_m} p_n^{m*}-P_{max}^m\right)=0, \tag{5.30}$$

where $[x]^+ = x$ if $x \geq 0$ and 0 otherwise. Equation (5.28) is the best power response, and $\omega_n^{m*}=c+\eta_n^{m*}+\lambda^{m*}$, η_n^{m*} and λ^{m*} are the optimal Lagrangian multipliers for the best response $p_n^{m*}(\mathbf{P}^{-m})$. Equations (5.29) and (5.30) are the KKT conditions. In the algorithm, the sub-gradient search method is implemented to update the Lagrangian multipliers to achieve best response. From (5.28), we can see the best power response is a modified water-filling allocation with the water level determined by the power price and the multipliers of power limit and total power.

Based on (5.24) and (5.28), the proposed distributed resource allocation (DRA) algorithm is devised as follows. It determines the subchannel and power allocation iteratively:

1. Initially, assume the initialized power allocation as: $\mathbf{P}_0 = \{\frac{P_{max}^m}{N_m}\}$, which means each CBS distributed the total power P_{max}^m to each subchannel equally. The initialized Lagrangian multipliers are set to zero.

2. Each SU measures the SINR for all the available subchannels for the given transmit power allocation of other CBSs in the previous iteration. Note that, at the beginning, the given transmission power is $\mathbf{P}_0$.

3. Each SU feeds back the measured values to the CBS associated with it. Note that, only the SUs feed back SINR information to its corresponding CBS. There is no any cooperation among CBSs.

4. Each CBS performs subchannel assignment according to (5.24), and obtains the optimal weight value ϕ_k^{m*} according to (5.18).

5. Then, each CBS performs power allocation according to (5.28) on each subchannel. The multiplier updation is also implemented.

6. Iterate step 2 – 5 until the resource allocation converges to an equilibrium.

In the above DRA algorithm, not only the interference from other CR-cells, but also the power and subchannel allocation are updated through iterations. Each CBS will maximize its own utility function by using only local information (i.e., SINR) received from the SUs and does not need the transmission power information in other cells. Therefore, the DRA algorithm operates in a distributed way.

We can see from the above process that the DRA algorithm includes both subchannel allocation performed in step 4 and power allocation performed in step 5. In each iteration, CBSs obtain the best power response $p_n^{m*}(\mathbf{P}^{-m})$ according to (5.28), and then allocate subchannels to maximize the utility based on the best power response until the DRA algorithm converges. We proved that a Nash equilibrium exists in the DPAG. Once all the players (i.e., CBSs) are in a Nash equilibrium, it means that the power strategy $\mathbf{p}^m$ is stable and will not change to another alternate one. Hence, for this given NE power stategy, the subchannel allocation is also stable. Therefore, it is obvious that if the DRA algorithm converges, it will converge to a NE point.

The uniqueness of NE point and Pareto optimality of DPAG will be discussed in the following subsections.

5.5.3 Uniqueness of Nash Equilibrium Point

From Sect. 5.5.2, the optimal solution $\mathbf{S}^*(\mathbf{P})$ of the DPAG can be written as $\mathbf{S}^*(\mathbf{P}) = [\mathbf{p}^{1*}(\mathbf{P})\mathbf{p}^{2*}(\mathbf{P}) \cdots \mathbf{p}^{M*}(\mathbf{P})]$, where $\mathbf{p}^{m*}(\mathbf{P}) = [p_n^{m*}(\mathbf{P}^{-m})]$. Let $\mathbf{S}^{*2}(\mathbf{P}) \equiv \mathbf{S}^*\left(\mathbf{S}^*(\mathbf{P})\right)$, that is, using solution $\mathbf{S}^*(\mathbf{P})$ as the initial power to achieve the power allocation at next iteration. For the uniqueness of the NE point in the DPAG, we have the following theorem:

Theorem 2: The DPAG has a unique Nash equilibrium if the best response $\mathbf{S}^*(\mathbf{P})$ satisfies the following properties [23],[24]:

- Positivity. There exist m and n such that $p_n^{m*}(\mathbf{P}) > 0$ for any network power vector $\mathbf{P}$, where $p_n^{m*}(\mathbf{P})$ is the optimal power allocated to subchannel n in CR-cell m.

- Monotonicity. If $\mathbf{P} \geq \mathbf{Q}$, then $\mathbf{S}^{*2}(\mathbf{P}) \geq \mathbf{S}^{*2}(\mathbf{Q})$.
- Scalability. If $p_n^{m*}(\mathbf{P}) > 0$, then for all $\alpha > 1$, there is $\alpha p_n^{m*2}(\mathbf{P}) > p_n^{m*2}(\alpha\mathbf{P})$.

*The following proof is to explain that the best response of the DPAG can satisfy the properties described in **Theorem 2**.*

Proof: Let $\beta_{k,n}^m(\mathbf{P}^{-m}) = \frac{g_{k,n}^m}{I_n^{m'}\left(\mathbf{P}^{-m}\right)+N_0}$ be the ratio of channel gain to interference plus noise power. Thus, the power solution on each subchannel in Eq. (5.28) can be rewritten as:

$$p_n^{m*}(\mathbf{P}^{-m}) = \left[\frac{(1+\phi_{k^*}^{m*})B_n}{\omega_n^{m*}\ln 2} - \frac{1}{\beta_{k^*,n}^m(\mathbf{P}^{-m})}\right]^+. \tag{5.31}$$

- Positivity. According to (5.31), we can know that when $\beta_{k^*,n}^m(\mathbf{P}^{-m})$ is very high on some subchannel, the allocated power on this subchannel will definitely be larger than 0. Therefore, it is realistic that there exist m and n that $p_n^{m*}(\mathbf{P}) > 0$ for any network power vector $\mathbf{P}$.
- Monotonicity. Let $y = \beta_{k^*,n}^m(\mathbf{P}^{-m})$. It is obvious that function y is monotonically decreasing with $\mathbf{P}^{-m}$, but $p_n^{m*}(y) = \left[\frac{(1+\phi_{k^*}^{m*})B_n}{\omega_n^{m*}\ln 2} - \frac{1}{y}\right]^+$ is monotonically increasing with y. Hence, $p_n^{m*}(\mathbf{P}^{-m})$ is monotonically decreasing with $\mathbf{P}^{-m}$. Therefore, if $\mathbf{P} \geq \mathbf{Q}$, we have $\mathbf{p}^{m*}(\mathbf{P}) \leq \mathbf{p}^{m*}(\mathbf{Q})$, so, $\mathbf{S}^*(\mathbf{P}) \leq \mathbf{S}^*(\mathbf{Q})$. Therefore, we have $\mathbf{S}^*\left(\mathbf{S}^*(\mathbf{P})\right) \geq \mathbf{S}^*\left(\mathbf{S}^*(\mathbf{Q})\right)$, that is $\mathbf{S}^{*2}(\mathbf{P}) \geq \mathbf{S}^{*2}(\mathbf{Q})$.
- Scalability. Assume that the noise power N_0 is much smaller than the co-channel interference $I_n^{m'}\left(\mathbf{P}^{-m}\right)$, this is reasonable in interference-limit cellular systems. Thus, from Eq. (5.31) and its monotonicity proved above, we have:

$$\alpha p_n^{m*}(\alpha\mathbf{P}) > p_n^{m*}(\mathbf{P}), \tag{5.32}$$

$$p_n^{m*}(\frac{1}{\alpha}\mathbf{P}) < \alpha p_n^{m*}(\mathbf{P}). \tag{5.33}$$

From (5.32), there is:

$$p_n^{m*}\left(p_n^{m*}(\alpha\mathbf{P})\right) < p_n^{m*}\left(\frac{1}{\alpha}p_n^{m*}(\mathbf{P})\right). \tag{5.34}$$

The right side of (5.34) satisfies $p_n^{m*}\left(\frac{1}{\alpha}p_n^{m*}(\mathbf{P})\right) < \alpha p_n^{m*}\left(p_n^{m*}(\mathbf{P})\right)$ by using (5.33). Therefore, we have $p_n^{m*2}(\alpha\mathbf{P}) < \alpha p_n^{m*2}(\mathbf{P})$.

5.5.4 Pareto Optimality

Pareto optimality [24] is defined as an allocation upon which no player can be made better off in utility without making any other player worse off. Define $\mathbf{P}^* = [\mathbf{p}^1 \cdots \mathbf{p}^m \cdots \mathbf{p}^j \cdots \mathbf{p}^M]$ is the Nash Equilibrium; $\mathbf{Q}=[\mathbf{p}^1 \cdots \mathbf{p}^{m-1}\mathbf{q}^m\mathbf{p}^{m+1} \cdots \mathbf{p}^j \cdots \mathbf{p}^M]$ is another power allocation strategy that can improve the utility of player m.

The mathematical definition of Pareto optimality is given as following:

Definition 2: A game is Pareto optimal, if the Nash Equilibrium $\mathbf{P}^* \in \mathbb{P}$ satisfies,

$\exists \mathbf{Q} \neq \mathbf{P}^*, u_m(\mathbf{Q}) > u_m(\mathbf{P}^*) \Rightarrow \exists j \in \mathcal{M}, u_j(\mathbf{Q}) < u_j(\mathbf{P}^*)$.

Namely, if player m can get better utility by changing power allocation from $\mathbf{p}^m$ to $\mathbf{q}^m$, and there exists player j whose utility is made worse due to the changed power allocation of player m, we say that the Nash Equilibrium $\mathbf{P}^*$ is Pareto optimal.

Theorem 3: The DPAG in (5.25) is Pareto optimal if the transmission power from CBSs to SUs is limited smaller than $\frac{B_n^m}{c\ln 2} - \frac{1}{\beta_{k^*,n}^m}$, where $\beta_{k^*,n}^m$ is the ratio of channel gain to interference plus noise.

*The following proof is to explain that the condition in **Theorem 3** can always be satisfied in our CRN/PN system.*

Proof: From the **Definition 2**, if $\exists \mathbf{Q} \neq \mathbf{P}^*$, we have $u_m(\mathbf{Q}) > u_m(\mathbf{P}^*)$.

Suppose 1: There is no CBS will be impaired. That is, $\forall j \in \mathcal{M} - m$, we have $u_j(\mathbf{Q}) \geq u_j(\mathbf{P}^*)$.

From (5.1), (5.2) and (5.25), we have

$$u_m(\mathbf{P}^*) = \sum_{n=1}^{N_m} B_n^m \log_2\left(1 + \frac{g_{k^*,n}^m p_n^m}{I_n^{m'}(\mathbf{P}^{*-m}) + N_0}\right) - c\sum_{n=1}^{N_m} p_n^m, \tag{5.35}$$

$$u_m(\mathbf{Q}) = \sum_{n=1}^{N_m} B_n^m \log_2\left(1 + \frac{g_{k^*,n}^m q_n^m}{I_n^{m'}(\mathbf{Q}^{-m}) + N_0}\right) - c\sum_{n=1}^{N_m} q_n^m, \tag{5.36}$$

where $B_n^m = (1 + \phi_{k^*}^{m*})B_n$.

Similarly, for CBS j, we have

$$u_j(\mathbf{P}^*)=\sum_{n=1}^{N_j}B_n^j\log_2\left(1+\frac{g_{k^*,n}^j p_n^j}{I_n^{j'}(\mathbf{P}^{*-j})+N_0}\right)-c\sum_{n=1}^{N_j}p_n^j, \tag{5.37}$$

$$u_j(\mathbf{Q})=\sum_{n=1}^{N_j}B_n^j\log_2\left(1+\frac{g_{k^*,n}^j q_n^j}{I_n^{j'}(\mathbf{Q}^{-j})+N_0}\right)-c\sum_{n=1}^{N_j}q_n^j, \tag{5.38}$$

where $B_n^j=(1+\phi_{k^*}^{j*})B_n$.

From the definition of $\mathbf{Q}$ and $\mathbf{P}^*$, we know that $\mathbf{Q}^{-m}=\mathbf{P}^{*-m}$. Moreover, the network condition and channel information are the same, therefore, we have $I_n^{m'}(\mathbf{Q}^{-m})=I_n^{m'}(\mathbf{P}^{*-m})$ on subchannel n. Let $\beta_{k^*,n}^m=\frac{g_{k^*,n}^m}{I_n^{m'}(\cdot)+N_0}$ denote the ratio of channel gain to interference plus noise. Due to $u_m(\mathbf{Q})>u_m(\mathbf{P}^*)$, for any subchannel n in CR-cell m, there is the following inequality:

$$B_n^m\log_2\left(\frac{1+\beta_{k^*,n}^m q_n^m}{1+\beta_{k^*,n}^m p_n^m}\right)>c\left(q_n^m-p_n^m\right), \tag{5.39}$$

where the parameters $B_n^m>0$, $c>0$, $\beta_{k^*,n}^m>0$, and variables $q_n^m\geq 0$, $p_n^m\geq 0$. Moreover, the function $\log_2(\cdot)$ is monotonically increasing, therefore, from (B.11), we have

$$\frac{1+\beta_{k^*,n}^m q_n^m}{2^{\frac{c}{B_n^m}q_n^m}}>\frac{1+\beta_{k^*,n}^m p_n^m}{2^{\frac{c}{B_n^m}p_n^m}}. \tag{5.40}$$

Let $f(x)=\frac{1+\beta_{k^*,n}^m x}{2^{\frac{c}{B_n^m}x}}$, $x\geq 0$, the derivative of function $f(x)$ is

$$\frac{\partial f(x)}{\partial x}=\frac{\beta_{k^*,n}^m-\frac{c\ln 2}{B_n^m}(1+\beta_{k^*,n}^m x)}{2^{\frac{c}{B_n^m}x}}. \tag{5.41}$$

If $0\leq x<\frac{B_n^m}{c\ln 2}-\frac{1}{\beta_{k^*,n}^m}$, we have $\frac{\partial f(x)}{\partial x}>0$, so, $f(x)$ is a monotonically increasing function. From (5.40), we have $q_n^m>p_n^m$ for any subchannel n in CR-cell m, i.e., $\mathbf{q}^m>\mathbf{p}^m$. Compare $I_n^{j'}(\mathbf{P}^{*-j})=I_n^{j'}(\mathbf{p}^1\cdots\mathbf{p}^m\cdots\mathbf{p}^{j-1}\mathbf{p}^{j+1}\cdots\mathbf{p}^M)$ and $I_n^{j'}(\mathbf{Q}^{-j})=I_n^{j'}(\mathbf{p}^1\cdots\mathbf{q}^m\cdots\mathbf{p}^{j-1}\mathbf{p}^{j+1}\cdots\mathbf{p}^M)$ in (5.37) and (5.38), the interferences on subchannel n from CR-cell m are increased due to $\mathbf{q}^m>\mathbf{p}^m$, that is, $I_n^{j'}(\mathbf{Q}^{-j})>I_n^{j'}(\mathbf{P}^{*-j})$. Moreover, from the definition of $\mathbf{Q}$ and $\mathbf{P}^*$, we know $\mathbf{p}^j=\mathbf{q}^j$. Hence, there exists player j whose utility is made worse, i.e., $u_j(\mathbf{Q})<u_j(\mathbf{P}^*)$. This is contradictive

to **Suppose 1**. Therefore, if the transmit power from CBSs to SUs is limited smaller than $\frac{B_n^m}{c\ln 2} - \frac{1}{\beta_{k^*,n}^m}$, no CBS can further improve its utility without impairing other players, and the DPAG is Pareto optimal.

If $x > \frac{B_n^m}{c\ln 2} - \frac{1}{\beta_{k^*,n}^m}$, we have $\frac{\partial f(x)}{\partial x} < 0$, so, $f(x)$ is a monotonically decreasing function. The similar as stated before, if x is larger than or equal to $\frac{B_n^m}{c\ln 2} - \frac{1}{\beta_{k^*,n}^m}$, we cannot say that the DPAG is Pareto optimal.

Consider the CRN/PN system in our paper. When the carrier to interference and noise ratio $\beta_{k^*,n}^m \approx 1$, the value $\frac{B_n^m}{c\ln 2} - \frac{1}{\beta_{k^*,n}^m}$ in our system is approximate 8W (this is impossible for CBSs to transmit on one subchannel by using 8W power). Moreover, from Sect. 3, we know that the transmit power on each subchannel is limited by primary targets. Since the transmit power of PUs is only 10mW in our coexistent system, from (13), the power limit $p_{n,m}^{lim}$ at CBSs on each subchannel is no larger than 15mW under the parameters defined in next Section. Therefore, the variable x in function $f(x)$ can always satisfy $0 \leq x < \frac{B_n^m}{c\ln 2} - \frac{1}{\beta_{k^*,n}^m}$, and we can summarize that the DPAG is Pareto optimal in some constrained environment, and in our system, it can satisfy this condition with high probability.

5.6 Performance Evaluation

In this section, simulations are performed for the downlink OFDMA-based multi-cell CRN/PN to evaluate the effectiveness of the proposed algorithm. The simulation parameters are summarized in Table 5.1. It is assumed that the channel gain is constant during $1ms$ period, thus the algorithm is performed once every $1ms$, which is also called one scheduling time. The WINNER Phase II channel model [69] is utilized to implement the channels in the simulations. We assume the wireless propagation environment is urban area, and both CRN and PN have the same cell radius R, with reference distance d_0 and center frequency f_c.

Consider K SUs and J PUs in each cell. SUs are randomly located in its cell area at each scheduling time, but the distance from SU to CBS should be larger than the reference distance. In each P-cell, the subchannels are uniformly allocated to J PUs. Consider a pessimistic assumption that all PUs are located at cell-edge area (i.e., 900m–1000m) from its PBS, then, the tolerable interference power margins at PBSs will not be very large. We assume all CR-cells are in a

Table 5.1: SIMULATION PARAMETERS

Parameters	Value
Number of cells M	7
Cell radius R	1000 m
Number of PUs per P-cell J	8
Number of SUs per CR-cell K	10
Number of subchannels per P-cell	16
Subchannel Bandwidth B_n	17 KHz
System center frequency f_c	2.4 GHz
Reference distance d_0	100 m
Total power at each CBS P^m_{max}	20–40 dBm
Predefined SINR of PUs in dB γ_p^{dB}	10 dB
Interference violation probability $\sigma^{(p)}$	0.01
antenna gain	8 dB
Path loss exponent	4
Shadowing standard deviation $\sigma_{\varphi_{dB}}$	6 dB
Noise power spectral density	-174 dBm/Hz
Power price c	3 kbps/W
Required minimum data rate for each SU R^k_{min}	100–600 kbps/W

quasi-synchronous mode, and there is no coordination among CBSs in resource allocation.

5.6.1 Simulation Results of Three-cell Case

In this subsection, in order to evaluate the performance of the algorithm, we consider a simple three-cell case at first, i.e., the cell 1,2,3 shown in Figure 5.1. So, the available subchannels for each CR-cell is $N_m = 32$.

In the coexistant system, the transmission power of primary transmitters has significant importance for the spectrum utilization in the CRN, since the primary interference power limit on each subchannel cannot be exceeded in order to guar-

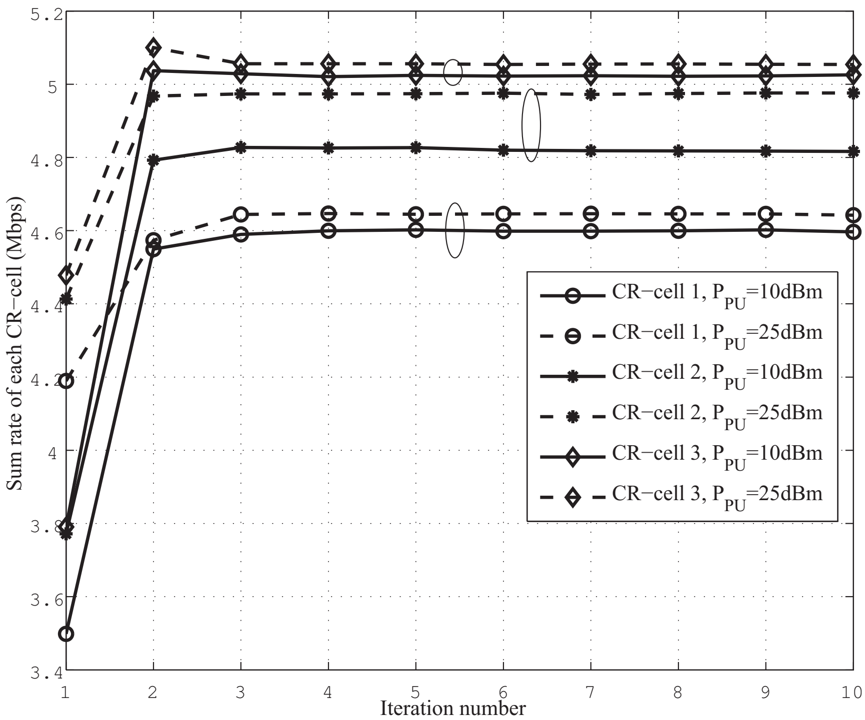

Figure 5.3: Sum data rate per CR-cell vs. the number of iterations.

antee the primary performance targets. Thus, when the targets are defined, if primary transmission power is higher, the allowable interference power limit will also be higher, so, the throughput of each secondary cell should be higher. However, the higher transmission power will also cause larger inter-cell interference. In order to evaluate the influence of primary transmission power, we have the results in Figures 5.3–5.5.

Firstly, the convergence behavior of the proposed algorithm is demonstrated in Figure 5.3, which is the sum rate of each CR-cell vs. the number of iterations for different PU transmission power P_{PU}=10dBm and 25dBm, respectively. It shows that for different P_{PU}, the sum rate converges in about three to four iterations. We can know from the result that the algorithm is effective and can get the NE point in each CR-cell fast. Moreover, from Figure 5.3, when P_{PU} increased from 10dBm to 25dBm, the sum rate only increased by an insignificant amount compared to the consumed transmission power, i.e., 0.04 Mbps in CR-cell 1, 0.15 Mbps in CR-cell 2, 0.03 in CR-cell 3. Furthermore, Figures 5.4 and 5.5 show the resulting power

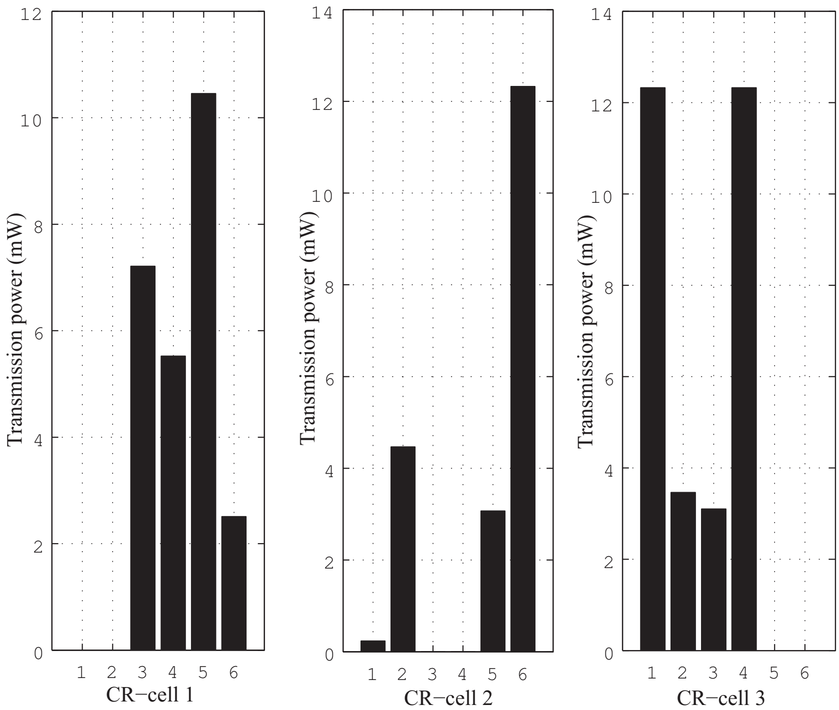

Figure 5.4: Transmission power allocation for the same subchannels in CR-cells when P_{PU}=10dBm.

allocation at the Nash equilibrium in each CR-cell when P_{PU}=10dBm and 25dBm, respectively. We can see the maximum allocated power in Figure 5.5 is more than 30 times than that in Figure 5.4. Higher transmission power results in higher interference, such that the sum rate cannot be increased so much. Therefore, it is not practical and economical for P_{PU}=25dBm. So, for the following evaluations, the primary transmission power is set to be P_{PU}=10dBm. The sum rate result in fourth iterations will be utilized as optimal resource allocation solution.

In Figure 5.6, the instantaneous data rate of cell-edge SU in each CR-cell is evaluated. Here, the cell-edge user is also defined as the user who is in the range from 900m to 1000m. The required minimum data rate for each SU is R^k_{min}=100kbps. From Figure 5.6, we can see that the cell-edge SU in each CR-cell can obtain its minimum data rate when R^k_{min} is proper. However, when R^k_{min} is too large, the system resources may not be enough for real-time services.

Then, we evaluate the performance of the proposed DRA algorithm in com-

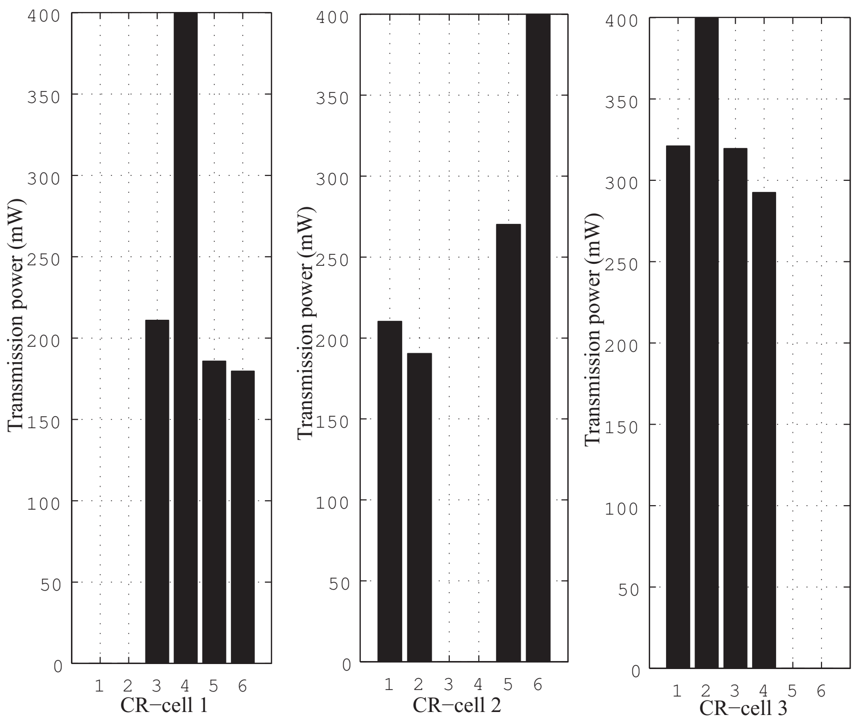

Figure 5.5: Transmission power allocation for the same subchannels in CR-cells when P_{PU}=25dBm.

parison with the centralized resource allocation algorithm, that is, each CBS knows the power allocation in other CR-cells and interference channel information well. Figures 5.7 and 5.8 show the average sum data rate with different QoS requirement and different CBS total power, respectively. Each average data rate is obtained by over 1000 simulation runs. From Figures 5.7 and 5.8, we can see that the average sum rate of centralized resource allocation algorithm is higher than that of distributed one, but the difference is not so significant. In Figure 5.7, the total power P^m_{max} at CBSs is 40dBm. When the required minimum data rate increases, the average data rate decreases, because the system needs more resources to guarantee the QoS for each SU, even the cell-edge SUs, when R^k_{min} increases. However, when R^k_{min} is large, the difference between two algorithms becomes larger. Since then, in order to guarantee the QoS, the allocated data rate to each subchannel increases, the interference increases as well, the centralized algorithm can manage the interference better than the distributed algorithm due

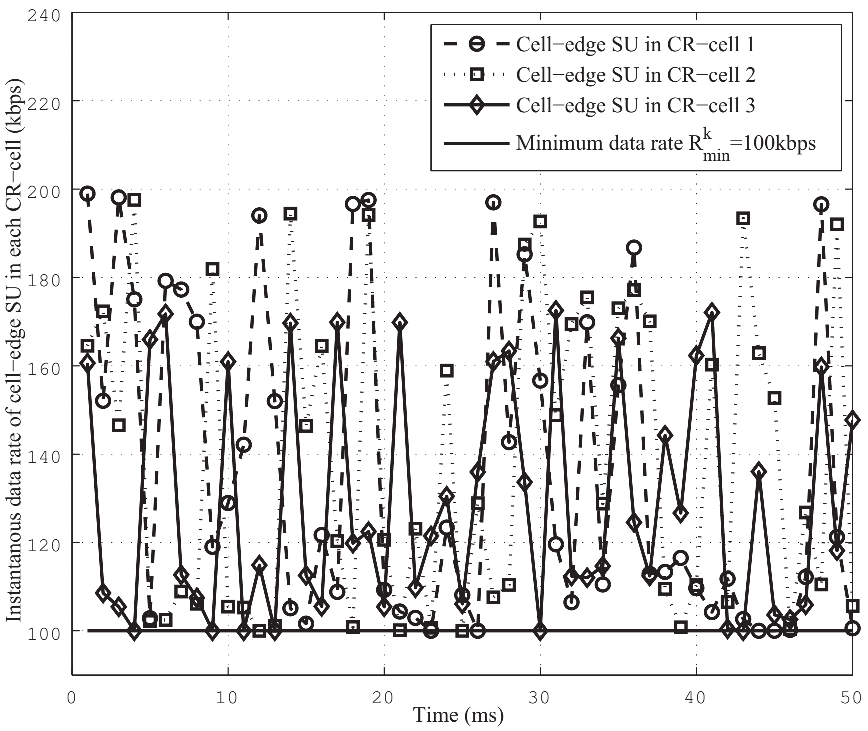

Figure 5.6: Instantaneous data rate of cell-edge SU in CR-cells vs. scheduling time.

to signaling exchanging and cooperation.

The similar result is shown in Figure 5.8, where, R^k_{min} is 100kbps. When P^m_{max} increases, the difference between two algorithms becomes larger. Moreover, we can see that if P^m_{max} is too large, i.e., more than 35dBm, in order to control large inter-cell interference, the sum rate in each cell is much more different from each other by using the proposed algorithm. This is because, they try to get an equlibrium to avoid large interference. On the other hand, for the centralized algorithm, there is not so much difference. Eventhough the performance of the proposed distributed method is a little worse than that of the centralized method, it is still a good choice due to its distribution and flexibility when the number of cells and users becomes very large, especially for cognitive radio networks.

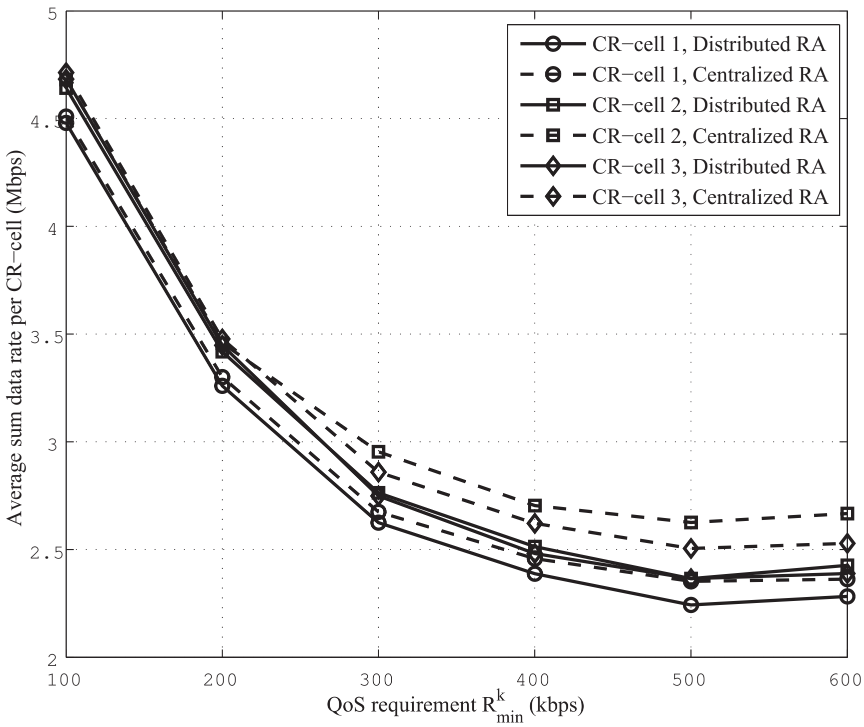

Figure 5.7: Average sum data rate per CR-cell vs. R^k_{min}.

5.6.2 Simulation Results of Seven-cell Case

The performance of seven-cell case, i.e., one-tier model, is evaluated in this subsection. The available subchannels for each CR-cell is $N_m = 96$. The primary transmission power is set to be $P_{PU} = 10$dBm for the following results.

In Figures 5.9 and 5.10, the sum rate of seven cells and three cells compared with the centralized algorithm is evaluated. Figures 5.9 and 5.10 show the average sum data rate with different R^k_{min} and different P^m_{max}, respectively. In Figure 5.9, P^m_{max} at CBSs is 40dBm, and in Figure 5.10, R^k_{min} is 100kbps. We can see that centralized algorithm outperforms the distributed one, but the difference is not significant. Compared to the three-cell case, the sum rate of the seven-cell case is much higher, because the available subchannels for each CR-cell is much more than three-cell case. Moreover, the variation tendency of these two figures is similar to figures 5.7 and 5.8 respectively, and the reasons of the variation tendency are described in Section 5.1, which will be omitted here.

Figure 5.11 shows that the outage probability of primary networks is always

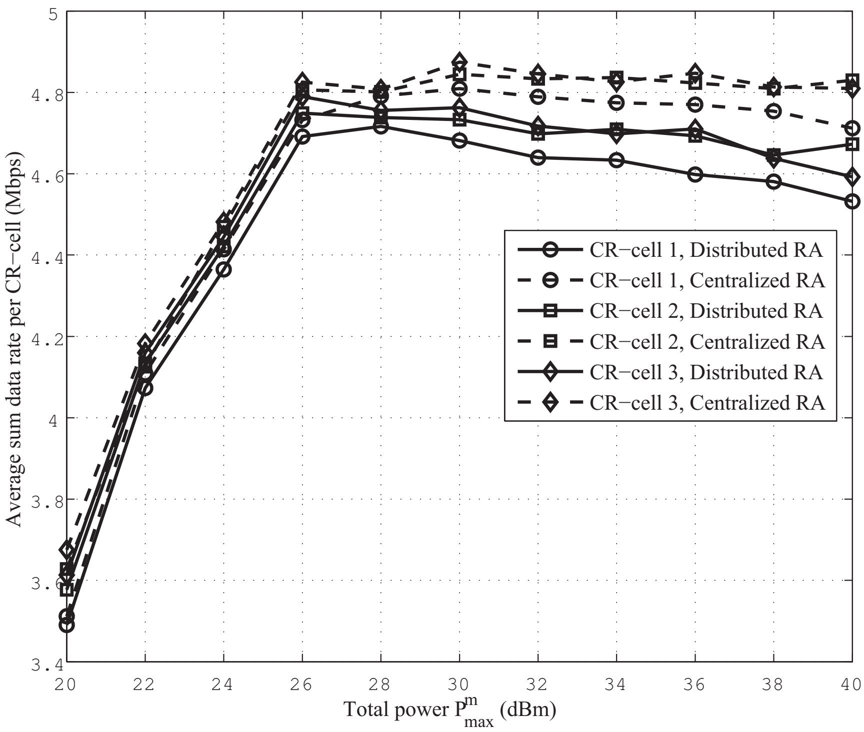

Figure 5.8: Average sum data rate per CR-cell vs. P^m_{max}.

zero even though the QoS requirement and the total power at CBSs are different. It also indicates that the interference constraints are satisfied, and the performance of PN can be guaranteed well.

5.7 Chapter Summary

In this chapter, we proposed a novel distributed resource allocation scheme for downlink transmission in OFDMA-based multi-cell cognitive radio networks. In the DRA algorithm, each CBS tries to maximize its sum rate while minimizing the co-channel interference to other CR-cells and without causing unacceptable interference to primary receivers. QoS for each SU is also considered. A primary-willingness based coexistent architecture is devised for CRN/PN. A novel intra-cell spectrum overlay and inter-cell spectrum underlay sharing method is utilized for primary and secondary spectrum share. According to the analysis of the DRA problem, it can be solved by two steps: subchannel allocation and power

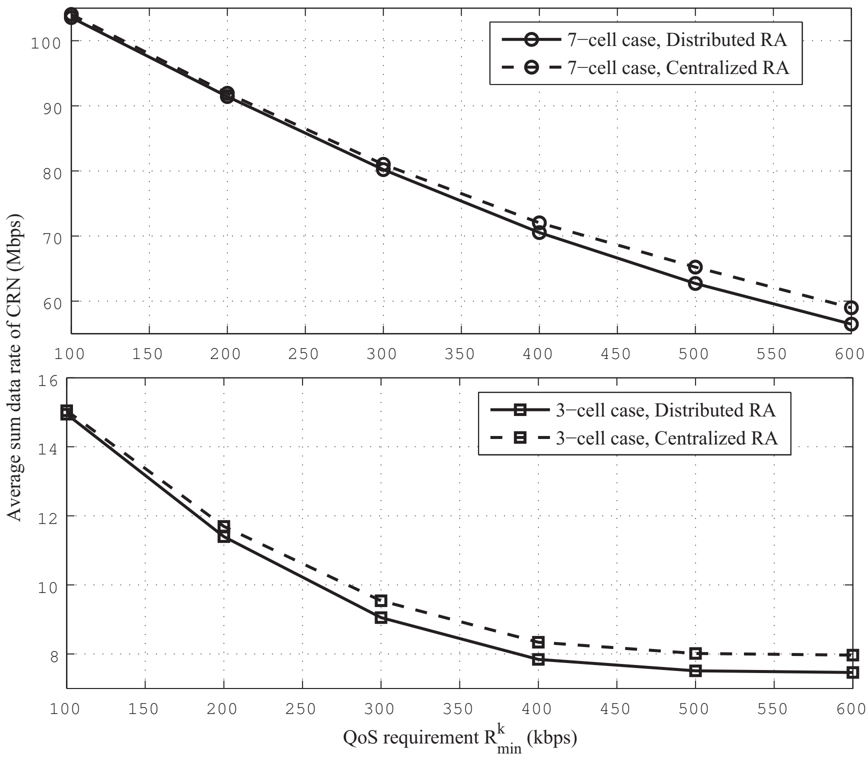

Figure 5.9: Average sum data rate of CRN vs. R^k_{min}.

allocation, and finally reduced into a distributed power allocation game. The Nash equilibrium point of the DPAG is obtained by using the Lagrangian duality based technique and KKT conditions. It was proven that the Nash equilibrium exists in the power allocation game, and is unique and Pareto optimal with high probability in our system. Through the simulation, the efficiency of our algorithm is shown, which has good convergent performance. Moreover, the QoS for cell-edge SU can be satisfied well. Compared to the centralized algorithm, the proposed algorithm shows its advantages, i.e., good system performance without large signaling overhead and without any coordination among CBSs.

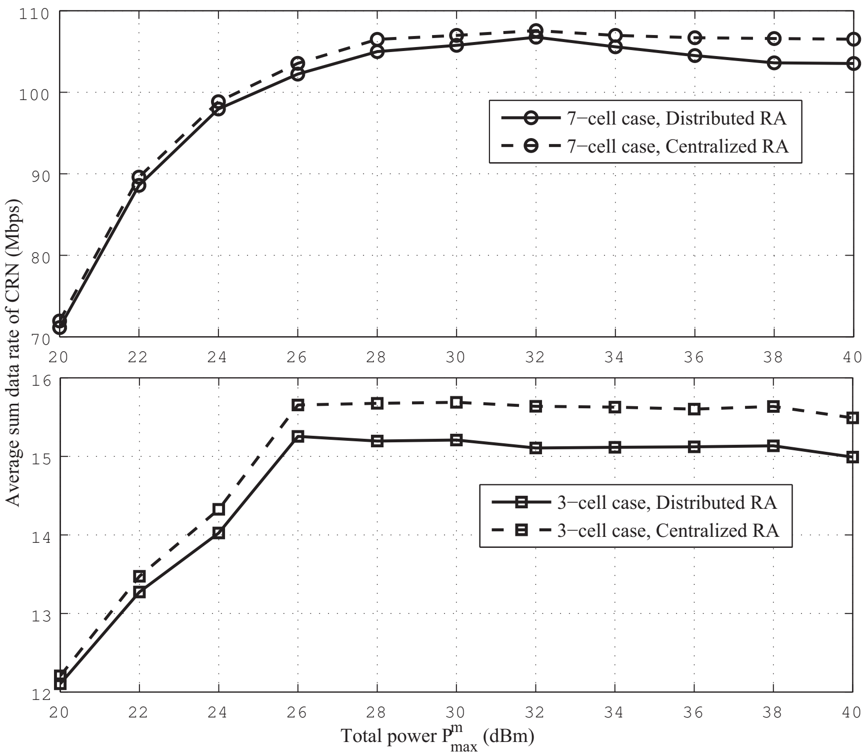

Figure 5.10: Average sum data rate of CRN vs. P^m_{max}.

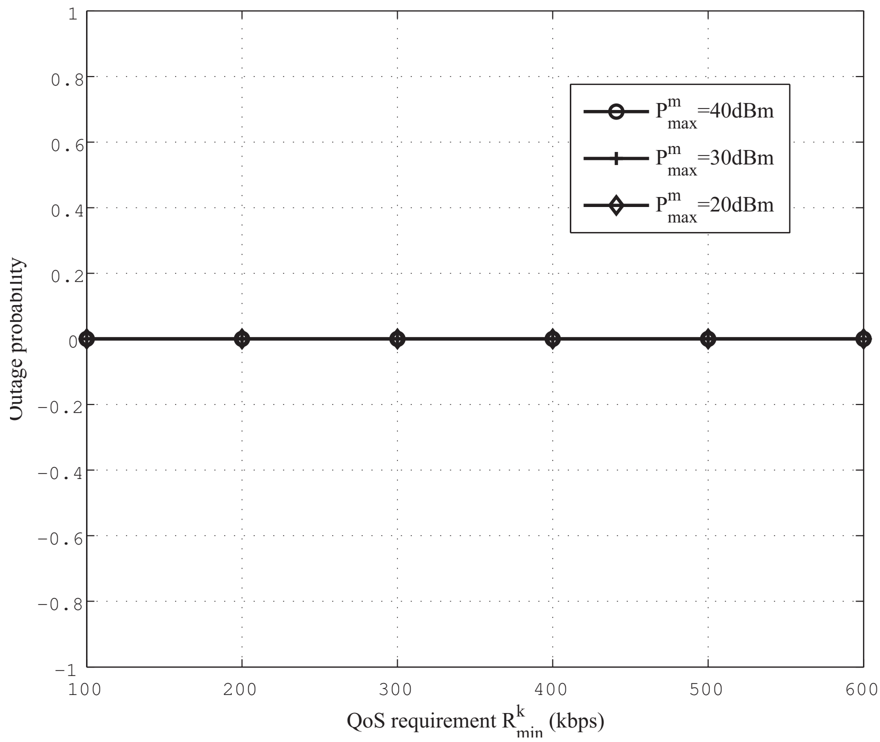

Figure 5.11: Outage probability of PN.

Chapter 6

Conclusions

This chapter concludes our research work based on the study of resource allocation schemes for cellular cognitive radio networks, which is a key technology to alleviate the severe spectrum shortage problem for the next generation wireless networks. First, we describe the advantages and contributions of the proposed resource allocation schemes. Secondly, the potential future research direction is discussed.

6.1 Contributions and Discussions

Cognitive radio is a promising technology for alleviating the severe spectrum shortage problem by allowing secondary users to share spectrum with primary users. Devices equipped with CR can be networked to create cognitive radio networks. With the ability to learn from and adapt to both their surrounding environment and user needs, cognitive radio networks have a great number of benefits in all kinds of applications, such as, military, government, public safety, and commercial areas. Considering how secondary users share spectrum with primary users, there are different spectrum sharing scenarios, such as, noncooperative networks, cooperative networks, and opportunistic utilization.

When implementing cognitive radio, many challenges occur throughout all layers of networks, especially PHY and MAC layers. Spectrum sensing, spectrun analysis and spectrum decision are the core functionalities of SUs equipped with CR. The spectrum decision is about whether and how to access the spectrum.

The optimal spectrum access option is the one that maximizes the application or user requirements with the given radio environment or spectrum constraints. It is different from spectrum access in traditional wireless networks due to the inter-system interferences and time-varying available spectrum. The interference power from SUs to PUs should be kept below a certain threshold. The interference management is more complicated during spectrum sharing. Moreover, The available spectrum is time-varying in CRNs, because it depends on the spectrum utilization in PNs. Nevertheless, good QoS still should be provided to SUs. Considering these problems, in this thesis, we focus on spectrum decision process and aim at designing operable spectrum sharing architectures and efficient resource allocation algorithms for cellular cognitive radio networks.

First, we carry out our study on single-cell multi-user CRNs, which coexist with a cellular PN. We develop a primary-assistance based coexistent architecture, where, the PBS determines the interference margins at PUs according to its target performance; then, the interference margins on occupied subchannels and pilot signals will be broadcast to secondary network for power control. Two different spectrum sharing methods, i.e., the spectrum underlay and spectrum overlay, are implemented in the study. The sharing method can be adapted to one of them based on the distance between the PBS and the CBS and the interference margins at PUs. Furthermore, a joint cross-layer resource allocation and interference avoidance algorithm is proposed for dynamic resource allocation in multi-user CRNs, based on the primary-assistant sharing architecture. The effectiveness of our proposed algorithm is verified by numerical analysis and computer simulations. Two conventional resource allocation schemes have also been studied : channel greedy access with power control and proportional fairness access with equal power. Compared to the conventional schemes, our algorithm achieves significant higher throughput and can guarantee the required SINR of the PUs and the QoS of the SUs well. Moreover, compared to the spectrum overlay sharing method, the spectrum underlay sharing could share the spectrum with the PUs more effectively. In addition, if there are non-active subbands in the PN, the hybrid spectrum underlay & overlay sharing can provide substantial higher spectrum efficiency.

Our next research work is for the multi-cell multi-user cognitive radio networks. Due to the co-channel interference and the inter-cell interference, the multi-cell case is much more complicated than single-cell. To manage the coexistence, a primary-willingness based coexistent architecture and a novel intra-cell spectrum overlay and inter-cell spectrum underlay sharing method are proposed. Then, for this spectrum sharing scenario, a distributed resource allocation scheme is devised to guarantee the primary performance, and at the same time, maximize the secondary utility without any cooperation among CBSs. Through the simulation, the efficiency of our algorithm is shown. The proposed algorithm turns out to converge to the equilibrium only within a small number of iterations. Compared to the centralized algorithm, the proposed distributed algorithm shows its advantages, i.e., good system performance without large signaling overhead and without any coordination among CBSs. Moreover, QoS among SUs is also considered by the scheme such that the instantaneous data rate for each secondary user is larger than a given minimum rate. The QoS for SUs can be satisfied well in the scheme.

Our investigation of resource allocation algorithms for CRNs provides some new research directions and practical applications for next generation networks. Although many researches have been done on resource allocation for CRNs, our research works focus on the hybride overlay/underlay spectrum sharing method and interference avoidance to primary networks. These research works significantly improve the spectrum efficiency and system throughput of CRNs.

6.2 Future Work

So far, very few researches have been done on multi-cell case for the cognitive radio networks. Therefore, in this thesis, coexistent architectures and resource allocation schemes on cellular cognitive radio networks are considered. However, there still exist lots of research scopes for future works not only about the centralized CRNs but also distributed CRNs.

First, to achieve the dynamic characteristics and implement CR technology, the information from only one layer is far from enough, so the information exchange and information fusion of multiple layers have significant importance. A very prominent example of such information exchange is the spectrum sensing information that requires cooperation of PHY and MAC layers to obtain. Also, in dynamic spectrum management and transmission power control, information from multiple layers is necessary not only in centralized CRNs but also distributed ones. Hence, cross-layer designs are nontrivial for CRNs ranging from the PHY layer to transport layer.

Second, in previous works, spectrum availability in CRNs all depends on the activities of PUs, which are assumed as simple structures (such as ON/OFF models). However, the mobility of secondary users is rarely considered in spectrum management of CRNs. It is still a new topic, and also should be considered in the future RA algorithm designs, especially for the vehicular cognitive radio networks (vehicular ad-hoc cognitive radio networks, or central-controlled vehicular CRNs), where the moving routes of the SUs are fixed and can be known beforehand.

As for future work, designing an effective cross-layer RA scheme for CRNs in mobility environment is an interesting research topic. This could lead to interesting and useful applications in the future wireless broadband access.

Appendix

Appendix A

Proof of Proposition 2 in Chapter 4

From OP-2, we have the following optimization problem:

OP-A:

$$min \quad \sum_{t=1}^{\tau}\sum_{m=1}^{M}\sum_{k=1}^{K}(2^{\frac{R_{k,m,t}}{\Delta f}}-1)\frac{L_{k,m,t}}{\beta_{k,m,t}} \tag{A.1}$$

s.t.

$$\sum_{k=1}^{K} L_{k,m,t}-1 \leq 0,\ L_{k,m,t}-1 \leq 0,\ -L_{k,m,t} \leq 0 \tag{A.2}$$

$$R_{k,m,t} - R_{max} \leq 0, \quad -R_{k,m,t} \leq 0 \tag{A.3}$$

$$r_{k,\tau}^{Req} - \sum_{t=1}^{\tau}\sum_{m=1}^{M} R_{k,m,t}L_{k,m,t} = 0 \tag{A.4}$$

The functions (A.1) in OP-A are convex functions in convex set $\mathcal{C} = \{L_{k,m,t} = \{0,1\}, R_{k,m,t} \in [0, R_{max}]\}$. The largrangian function [18] of the above convex optimization OP-A is:

$$L = \sum_{t=1}^{\tau}\sum_{m=1}^{M}\sum_{k=1}^{K}\frac{(2^{\frac{R_{k,m,t}}{\Delta f}}-1)L_{k,m,t}}{\beta_{k,m,t}} + \sum_{t=1}^{\tau}\sum_{m=1}^{M} a_{m,t}(\sum_{k=1}^{K} L_{k,m,t}-1) + \sum_{t=1}^{\tau}\sum_{m=1}^{M}\sum_{k=1}^{K} b_{k,m,t}(L_{k,m,t}-1)$$

$$-\sum_{t=1}^{\tau}\sum_{m=1}^{M}\sum_{k=1}^{K}c_{k,m,t}L_{k,m,t}+\sum_{t=1}^{\tau}\sum_{m=1}^{M}\sum_{k=1}^{K}e_{k,m,t}(R_{k,m,t}-R_{max})-\sum_{t=1}^{\tau}\sum_{m=1}^{M}\sum_{k=1}^{K}f_{k,m,t}R_{k,m,t}$$

$$+\sum_{k=1}^{K}h_k(r_{k,\tau}^{Req}-\sum_{t=1}^{\tau}\sum_{m=1}^{M}R_{k,m,t}L_{k,m,t}) \tag{A.5}$$

where $\{a_{m,t}, b_{k,m,t}, c_{k,m,t}, e_{k,m,t}, f_{k,m,t}, h_k\}$,$t$=1, 2, ..., τ, $k\in\mathcal{K}$,$m\in\mathcal{M}$ are the Lagrange multipliers, and each multiplier should be no less than zero.

Let $\{R^*_{k,m,t}, L^*_{k,m,t}, a^*_{m,t}, b^*_{k,m,t}, c^*_{k,m,t}, e^*_{k,m,t}, f^*_{k,m,t}, h^*_k\}$ be an optimal solution. Then the Karush-Kuhn-Tucker (KKT) conditions are as following:

$$a^*_{m,t}\geq 0, b^*_{k,m,t}\geq 0, c^*_{k,m,t}\geq 0, e^*_{k,m,t}\geq 0, f^*_{k,m,t}\geq 0, h^*_k\geq 0 \tag{A.6}$$

$$a^*_{m,t}(\sum_{k=1}^{K}L^*_{k,m,t}-1)=0 \tag{A.7}$$

$$b^*_{k,m,t}(L^*_{k,m,t}-1)=0 \tag{A.8}$$

$$c^*_{k,m,t}L^*_{k,m,t}=0 \tag{A.9}$$

$$e^*_{k,m,t}(R^*_{k,m,t}-R_{max})=0 \tag{A.10}$$

$$f^*_{k,m,t}R^*_{k,m,t}=0 \tag{A.11}$$

$$h^*_k\left(r_{k,\tau}^{Req}-\sum_{t=1}^{\tau}\sum_{m=1}^{M}R^*_{k,m,t}L^*_{k,m,t}\right)=0 \tag{A.12}$$

$$\frac{L^*_{k,m,t}\ln(2)2^{\frac{R^*_{k,m,t}}{\Delta f}}}{\beta_{k,m,t}\Delta f}+e^*_{k,m,t}-f^*_{k,m,t}-h^*_kL^*_{k,m,t}=0 \tag{A.13}$$

$$\frac{(2^{\frac{R^*_{k,m,t}}{\Delta f}}-1)}{\beta_{k,m,t}}+a^*_{m,t}+b^*_{k,m,t}-c^*_{k,m,t}+h^*_kR^*_{k,m,t}=0 \tag{A.14}$$

Equations (A.13) and (A.14) are obtained by setting $\partial L/\partial R_{k,m,t} = 0$ and $\partial L/\partial L_{k,m,t}=0$ respectively.

From (A.13), when $L^*_{k,m,t}\neq 0$, we have

$$R^*_{k,m,t}=\Delta f \log_2\{\frac{(h^*_k L^*_{k,m,t}+f^*_{k,m,t}-e^*_{k,m,t})\beta_{k,m,t}\Delta f}{\ln(2)L^*_{k,m,t}}\} \tag{A.15}$$

According to (A.10) and (A.11), $e^*_{k,m,t}$ and $f^*_{k,m,t}$ cannot be positive at the same time. Therefore, when $R^*_{k,m,t}=0$, $e^*_{k,m,t}=0, f^*_{k,m,t}\geq 0$, then, according to (A.15), $h^*_k \leq \frac{\ln(2)}{\beta_{k,m,t}\Delta f}$; when $R^*_{k,m,t}=R_{max}$, $e^*_{k,m,t}\geq 0, f^*_{k,m,t}=0$, then, according to (A.15), $h^*_k \geq \frac{2^{\frac{R_{max}}{\Delta f}}\ln(2)}{\beta_{k,m,t}\Delta f}$; when $0 < R^*_{k,m,t} < R_{max}$, $e^*_{k,m,t}=0, f^*_{k,m,t}=0$, then, according to (A.15), $\frac{\ln(2)}{\beta_{k,m,t}\Delta f} < h^*_k < \frac{2^{\frac{R_{max}}{\Delta f}}\ln(2)}{\beta_{k,m,t}\Delta f}$. To summarize, the optimal rate allocation $R^*_{k,m,t}$ at time slot t can be:

$$R^*_{k,m,t}=\begin{cases} 0, & h^*_k \leq \frac{\ln(2)}{\beta_{k,m,t}\Delta f} \\ \Delta f \log_2(\frac{h^*_k\beta_{k,m,t}\Delta f}{\ln(2)}), & \frac{\ln(2)}{\beta_{k,m,t}\Delta f} < h^*_k < \frac{2^{\frac{R_{max}}{\Delta f}}\ln(2)}{\beta_{k,m,t}\Delta f} \\ R_{max}, & h^*_k \geq \frac{2^{\frac{R_{max}}{\Delta f}}\ln(2)}{\beta_{k,m,t}\Delta f} \end{cases} \tag{A.16}$$

Suppose that subcarrier m has been allocated to more than one SU, that is, there exists $0 < L^*_{k,m,t} < 1$ for SUs $k_1, k_2, ..., k_B$, $B > 1$. From (A.8) and (A.9), we have $b^*_{k,m,t}=0$ and $c^*_{k,m,t}=0$. Then, from (A.14), we have

$$\frac{(2^{\frac{R^*_{k,m,t}}{\Delta f}}-1)}{\beta_{k,m,t}}+h^*_k R^*_{k,m,t}=-a^*_{m,t} \tag{A.17}$$

We define $h_{k,m,t}=(2^{\frac{R^*_{k,m,t}}{\Delta f}}-1)/\beta_{k,m,t}+h^*_k R^*_{k,m,t}$, therefore, we have $h_{k,m,t}=-a^*_{m,t}$ for all $k=k_1, k_2, ..., k_B$, that is $h_{k_1,m,t}=h_{k_2,m,t}=...=h_{k_B,m,t}$.

However, for the left side in (A.17), unless $\beta_{k,m,t}$ is equal for SUs $k_1, k_2, ..., k_B$, it is highly impossible that any of the two $h_{k,m,t}$ values will be equal. Since $\beta_{k,m,t}$ are channel state information, modeled as independent and random variables. Therefore, we conclude that for any time slot t and subcarrier m, there is only one SU k^*, that is $L^*_{k^*,m,t}=1$ if subcarrier m has been allocated. The method how to find this SU, which is similar to that analyzed in Appendix B, will be omitted here.

Now, the proof of Proposition 2 will continue. We consider $L^*_{k,m,t}{=}1$ for SU k. The allocated rates for SU k during τ time slots r_k can be calculated as following by using (A.16):

$$r_k{=}\sum_{t=1}^{\tau}\sum_{m=1}^{M} R^*_{k,m,t} \tag{A.18}$$

We define some probabilities:

Pr{a given subcarrier is allocated to SU k}$= p_k$;

Pr{the power gain of the allocated subcarrier satisfies $\frac{\ln(2)}{\beta_{k,m,t}\Delta f} < h^*_k < \frac{2^{\frac{R_{max}}{\Delta f}}\ln(2)}{\beta_{k,m,t}\Delta f}$}$=$ p^a_k;

Pr{the power gain of the allocated subcarrier satisfies $h^*_k \geq \frac{2^{\frac{R_{max}}{\Delta f}}\ln(2)}{\beta_{k,m,t}\Delta f}$}$= p^b_k$.

Equation (A.18) can be calculated as:

$$\begin{aligned} r_k{=} & \sum_{\frac{\ln(2)}{\beta_{k,m,t}\Delta f}<h^*_k<\frac{2^{\frac{R_{max}}{\Delta f}}\ln(2)}{\beta_{k,m,t}\Delta f}} \Delta f \log_2\left(\frac{h^*_k\beta_{k,m,t}\Delta f}{\ln(2)}\right) + \sum_{h^*_k\geq\frac{2^{\frac{R_{max}}{\Delta f}}\ln(2)}{\beta_{k,m,t}\Delta f}} R_{max} \\ = & \sum_{t=1}^{\tau} m(t)p_k p^a_k \Delta f \log_2\left(\frac{h^*_k\beta_{k,m,t}\Delta f}{\ln(2)}\right) + \sum_{t=1}^{\tau} m(t)p_k p^b_k R_{max} \end{aligned} \tag{A.19}$$

where $m(t)$ is the available subcarrier number at time slot t.

Similarly, at time slot t, the optimal number of bit rate allocated to SU k is obtained:

$$\sum_{m=1}^{M} R^*_{k,m,t}L^*_{k,m,t} = m(t)p_k\{p^a_k \Delta f \log_2\left(\frac{h^*_k\beta_{k,m,t}\Delta f}{\ln(2)}\right) + p^b_k R_{max}\} \tag{A.20}$$

From (A.19) and (A.20), we have

$$\sum_{m=1}^{M} R^*_{k,m,t}L^*_{k,m,t}{=}\frac{m(t)}{\sum_{t=1}^{\tau} m(t)} r_k \tag{A.21}$$

Suppose that the allocation can achieve the QoS requirements for SU k, therefore, the allocated bit rates during τ time slots satisfied $r_k = r^{Req}_{k,t}\tau$, where, $r^{Req}_{k,t}$ is the required bit rate at each time slot. So, equation (A.21) can be rewritten as:

$$\sum_{m=1}^{M} R^*_{k,m,t}L^*_{k,m,t}{=}\frac{m(t)\tau}{\sum_{t=1}^{\tau} m(t)} r^{Req}_{k,t} \tag{A.22}$$

Due to primary interference limits, not all the subcarriers are available for SUs. If the available bandwidth $m(t)$ is considerably small, only few bit rates can be allocated to SU k at time slot t. Then, it is difficult to satisfy the required bit rate $r_{k,t}^{Req}$ at each time slot, and the required bit rate during τ time slots also cannot be satisfied, since $r_k < r_{k,t}^{Req}\tau$ due to limited primary bandwidth. To guarantee the QoS requirments, M needs to be large enough. Assume $M \to \infty$, so that at each time slot, there are $m(t) \geq m_{req}$ that can achieve SUs' requirements, where, m_{req} is the minimum number of required subcarrier. Therefore, for OP-A, to achieve the QoS and minimize the transmit power, $m(t)$ should be equal to m_{req} at each time slot. If $M \to \infty$, $\sum_{t=1}^{\tau} m(t) = \tau m_{req}$, and we can have:

$$\sum_{m=1}^{M} R_{k,m,t}^{*} L_{k,m,t}^{*} = r_{k,t}^{Req} \tag{A.23}$$

Appendix B

Proof of Solution S^* in Chapter 4

In Problem OP-3, if we replace $R_{k,m}$ by $R_{k,m} = \widehat{R}_{k,m}\Delta f$, which can be obtained from (4.10-refeq:5). The following duality optimization problem is obtained:

OP-B:

$$min \quad \sum_{k=1}^{K} -\lambda_k \tag{B.1}$$

s.t.

$$P_{k,m} - P_m^{max} \leq 0, \ \sum_{k=1}^{K}\sum_{m=1}^{M} P_{k,m} - P_0 \leq 0 \tag{B.2}$$

$$\sum_{k=1}^{K} L_{k,m} - 1 \leq 0, \ L_{k,m} - 1 \leq 0, \ -L_{k,m} \leq 0 \tag{B.3}$$

$$\widehat{R}_{k,m}\Delta f - R_{max} \leq 0, \ \ -\widehat{R}_{k,m}\Delta f \leq 0 \tag{B.4}$$

$$r_k^{Req} + \lambda_k - \sum_{m=1}^{M} \widehat{R}_{k,m}\Delta f L_{k,m} = 0, \ -\lambda_k \leq 0 \tag{B.5}$$

where $P_{k,m} = \frac{2^{\widehat{R}_{k,m}} - 1}{\beta_{k,m}}$.

The Lagrangian function [18], [72] associated with the above duality problem OP-B can be written as:

$$L(\widehat{R}_{k,m}, L_{k,m}, \lambda_k, \xi_{k,m}, \varphi, \psi_m, \phi_{k,m}, \upsilon_{k,m}, \zeta_{k,m}, \epsilon_{k,m}, \eta_k, \varsigma_k)$$

$$
\begin{aligned}
&= \sum_{k=1}^{K}(-\lambda_k) + \sum_{k=1}^{K}\sum_{m=1}^{M}\xi_{k,m}\left(\frac{2^{\widehat{R}_{k,m}}-1}{\beta_{k,m}} - P_{max}\right) + \varphi\left(\sum_{k=1}^{K}\sum_{m=1}^{M}\frac{2^{\widehat{R}_{k,m}}-1}{\beta_{k,m}} - P_0\right)\\
&+\sum_{m=1}^{M}\psi_m\left(\sum_{k=1}^{K}L_{k,m}-1\right) + \sum_{k=1}^{K}\sum_{m=1}^{M}\phi_{k,m}(L_{k,m}-1) - \sum_{k=1}^{K}\sum_{m=1}^{M}\upsilon_{k,m}L_{k,m}\\
&+\sum_{k=1}^{K}\sum_{m=1}^{M}\zeta_{k,m}(\widehat{R}_{k,m}\Delta f - R_{max}) - \sum_{k=1}^{K}\sum_{m=1}^{M}\epsilon_{k,m}\widehat{R}_{k,m}\Delta f\\
&+\sum_{k=1}^{K}\eta_k\left(r_k^{Req} + \lambda_k - \sum_{m=1}^{M}\widehat{R}_{k,m}\Delta f L_{k,m}\right) - \sum_{k=1}^{K}\varsigma_k\lambda_k
\end{aligned}
\tag{B.6}
$$

where $\xi_{k,m}, \varphi, \psi_m, \phi_{k,m}, \upsilon_{k,m}, \zeta_{k,m}, \epsilon_{k,m}, \eta_k, \varsigma_k, k \in \mathcal{K}, m \in \mathcal{M}$ are the Lagrange multipliers.

Let $\widehat{R}^*_{k,m}, L^*_{k,m}, \lambda^*_k, \xi^*_{k,m}, \varphi^*, \psi^*_m, \phi^*_{k,m}, \upsilon^*_{k,m}, \zeta^*_{k,m}, \epsilon^*_{k,m}, \eta^*_k, \varsigma^*_k$, be an optimal solution set, then the Karush-Kuhn-Tucker (KKT) conditions state that [18]:

$$
\begin{gathered}
\xi^*_{k,m} \geq 0,\ \varphi^* \geq 0,\ \psi^*_m \geq 0,\ \phi^*_{k,m} \geq 0,\ \upsilon^*_{k,m} \geq 0,\\
\zeta^*_{k,m} \geq 0,\ \epsilon^*_{k,m} \geq 0,\ \eta^*_k \geq 0,\ \varsigma^*_k \geq 0
\end{gathered}
\tag{B.7}
$$

$$
\xi^*_{k,m}\left(\frac{2^{\widehat{R}^*_{k,m}}-1}{\beta_{k,m}} - P_{max}\right) = 0 \tag{B.8}
$$

$$
\varphi^*\left(\sum_{k=1}^{K}\sum_{m=1}^{M}\frac{2^{\widehat{R}^*_{k,m}}-1}{\beta_{k,m}} - P_0\right) = 0 \tag{B.9}
$$

$$
\psi^*_m\left(\sum_{k=1}^{K}L^*_{k,m} - 1\right) = 0 \tag{B.10}
$$

$$
\phi^*_{k,m}(L^*_{k,m} - 1) = 0, \quad -\upsilon^*_{k,m}L^*_{k,m} = 0 \tag{B.11}
$$

$$
\zeta^*_{k,m}(\widehat{R}^*_{k,m}\Delta f - R_{max}) = 0, \epsilon^*_{k,m}\widehat{R}^*_{k,m}\Delta f = 0 \tag{B.12}
$$

$$\eta_k^*(r_k^{Req}+\lambda_k^*-\sum_{m=1}^{M}\widehat{R}_{k,m}^*\Delta f L_{k,m}^*)=0 \tag{B.13}$$

$$\varsigma_k^*\lambda_k^*=0 \tag{B.14}$$

$$(\xi_{k,m}^*+\varphi^*)\frac{2^{\widehat{R}_{k,m}^*}\ln(2)}{\Delta f\beta_{k,m}}-\eta_k^*L_{k,m}^*=\epsilon_{k,m}^*-\zeta_{k,m}^* \tag{B.15}$$

$$\psi_m^*+\phi_{k,m}^*-\upsilon_{k,m}^*-\eta_k^*\widehat{R}_{k,m}^*\Delta f=0 \tag{B.16}$$

$$-1+\eta_k^*-\varsigma_k^*=0 \tag{B.17}$$

Equations (B.15)-(B.17) are obtained by setting $\partial L/\partial \widehat{R}_{k,m}$=0, $\partial L/\partial L_{k,m}$=0, and $\partial L/\partial \lambda_k$=0 respectively.

In order to analyze the KKT consitions and get the optimal solution, we have the following steps to solve Problem OP-B:

1) ***Step 1****: Power Allocation*

From (B.15), the following equation can be obtained:

$$2^{\widehat{R}_{k,m}^*}\frac{\ln(2)}{\Delta f\beta_{k,m}}-\frac{\eta_k^*L_{k,m}^*}{\xi_{k,m}^*+\varphi^*}=\frac{\epsilon_{k,m}^*-\zeta_{k,m}^*}{\xi_{k,m}^*+\varphi^*} \tag{B.18}$$

According to (B.7) and (B.12), we know that $\zeta_{k,m}^*$ and $\epsilon_{k,m}^*$ cannot be both positive and they are all nonnegative. So the optimal values of $\zeta_{k,m}^*$ and $\epsilon_{k,m}^*$ can only be one of the following cases:

$$\epsilon_{k,m}^*>0,\ \zeta_{k,m}^*=0$$

$$\epsilon_{k,m}^*=0,\ \zeta_{k,m}^*>0$$

$$\epsilon_{k,m}^*=0,\ \zeta_{k,m}^*=0$$

Here, we set $\omega_k=\frac{\eta_k^*L_{k,m}^*}{\xi_{k,m}^*+\varphi^*}$, therefore, if $\omega_k<2^{\widehat{R}_{k,m}^*}\frac{\ln(2)}{\Delta f\beta_{k,m}}$, that means we must have $\epsilon_{k,m}^*$>0, $\zeta_{k,m}^*$=0. So, according to (B.12), we must have $\widehat{R}_{k,m}^*=0$.

If $2^{\widehat{R}_{k,m}^*}\frac{\ln(2)}{\Delta f\beta_{k,m}}\leq\omega_k\leq 2^{\widehat{R}_{max}}\frac{\ln(2)}{\Delta f\beta_{k,m}}$, that means $0\leq\widehat{R}_{k,m}^*\leq\widehat{R}_{max}$, where, $\widehat{R}_{max}=\frac{R_{max}}{\Delta f}$, we must have $\epsilon_{k,m}^*$=0, $\zeta_{k,m}^*$=0. So, according to (B.18), we can get that $\widehat{R}_{k,m}^*=\log_2(\frac{\omega_k\Delta f\beta_{k,m}}{\ln(2)})$.

If $\omega_k > 2^{\widehat{R}_{max}} \frac{\ln(2)}{\Delta f \beta_{k,m}}$, we must have $\epsilon^*_{k,m}=0$, $\zeta^*_{k,m}>0$. According to (B.15), we must have $\widehat{R}^*_{k,m}=\widehat{R}_{max}$.

Therefore, to summarize, the optimal value of $\widehat{R}^*_{k,m}$ is:

$$\widehat{R}^*_{k,m}=\begin{cases} 0, & \omega_k < \frac{\ln(2)}{\Delta f \beta_{k,m}} \\ \log_2(\frac{\omega_k \Delta f \beta_{k,m}}{\ln(2)}), & \frac{\ln(2)}{\Delta f \beta_{k,m}} \leq \omega_k \leq \frac{2^{\widehat{R}_{max}} \ln(2)}{\Delta f \beta_{k,m}} \\ \widehat{R}_{max}, & \omega_k > \frac{2^{\widehat{R}_{max}} \ln(2)}{\Delta f \beta_{k,m}} \end{cases} \tag{B.19}$$

The following optimal power allocation $P^*_{k,m} = \frac{2^{\widehat{R}^*_{max}-1}}{\beta_{k,m}}$ can be obtained:

$$P^*_{k,m}=\begin{cases} 0, & \omega_k < \frac{\ln(2)}{\Delta f \beta_{k,m}} \\ \frac{\omega_k \Delta f}{\ln(2)} - \frac{1}{\beta_{k,m}}, & \frac{\ln(2)}{\Delta f \beta_{k,m}} \leq \omega_k \leq \frac{2^{\widehat{R}_{max}} \ln(2)}{\Delta f \beta_{k,m}} \\ \frac{2^{\widehat{R}_{max}-1}}{\beta_{k,m}}, & \omega_k > \frac{2^{\widehat{R}_{max}} \ln(2)}{\Delta f \beta_{k,m}} \end{cases} \tag{B.20}$$

where $\omega_k = \frac{\eta^*_k L^*_{k,m}}{\xi^*_{k,m}+\varphi^*}$, represents the update of the multipliers and is viewed as the iterative water-filling level for the SU k and will be discussed later.

2) ***Step 2****: Subcarrier Allocation*

For simplicity and in order to maximize the system throughput, $L^*_{k,m}$ is set to be either 0 or 1, and one subcarrier must be allocated to any user. So, for any given time slot t and subcarrier m, there is only one SU k^* with a nonzero value of $L^*_{k^*,m}$ and $L^*_{k^*,m}=1$ according to (B.3). Now, we'll discuss how to determine the SU k^*.

For SU k^*, according to (B.11), we have $\phi^*_{k^*,m}\geq 0$, and $\upsilon^*_{k^*,m}=0$, and it follows from (B.16) that

$$\psi^*_m = \eta^*_{k^*} \widehat{R}^*_{k^*,m} \Delta f - \phi^*_{k^*,m} \tag{B.21}$$

For any other SU $k \neq k^*$, $L^*_{k,m}=0$, according to (B.11), we have $\phi^*_{k,m}=0$, and $\upsilon^*_{k,m} \geq 0$, combined with (B.16), we have

$$\psi^*_m = \upsilon^*_{k,m} + \eta^*_k \widehat{R}^*_{k,m} \Delta f \tag{B.22}$$

Compare (B.21) with (B.22), we have

$$\eta^*_{k^*} \widehat{R}^*_{k^*,m} \geq \eta^*_k \widehat{R}^*_{k,m} \tag{B.23}$$

The subcarrier allocation strategy for any subcarrier m is

$$L^*_{k,m} = \begin{cases} 1, & k = k^* \ \ and \ \ m \in \mathcal{M} \\ 0, & otherwise \end{cases} \tag{B.24}$$

$$k^* = argmax \qquad \eta^*_k \widehat{R}^*_{k,m} \tag{B.25}$$

where $\widehat{R}^*_{k,m}$ is the optimal value from (B.19). Suppose that for a given t and m, the values of $\eta^*_k \widehat{R}^*_{k,m}$ are the same for several users, we will choose one SU arbitrarily.

*3) **Step 3**: Iterative Water-filling Level*

For all SUs, according to KKT conditions, when $\lambda^*_k > 0$, according to (B.14), $\varsigma^*_k = 0$ and according to (B.17), $\eta^*_k = 1$, so $\omega_k = \frac{L^*_{k,m}}{\xi^*_{k,m} + \varphi^*}$. When $\lambda^*_k = 0$, then $\varsigma^*_k \geq 0$ and $\eta^*_k = 1 + \varsigma^*_k$, so $\omega_k = \frac{(1+\varsigma^*_k)L^*_{k,m}}{\xi^*_{k,m} + \varphi^*}$. The initial value of $L^*_{k,m}$ in ω_k can be set to 1. Let $\omega_B = \frac{1}{\xi^*_{k,m} + \varphi^*}$ be the base water-level for all SUs.

From (B.7)- (B.9), if the optimal value $\varphi^* > 0$, that means all the power has been used to optimize the system throughput. Therefore, the initial value of φ^* and $\xi^*_{k,m}$should be set to $\varphi^* > 0$ and $\xi^*_{k,m} > 0$ to obtain the maximum system throughput. Therefore, to summarize, the water-filling level ω_k is:

$$\omega_k = \begin{cases} \omega_B, & \lambda^*_k > 0 \\ (1 + \varsigma^*_k)\omega_B, & \lambda^*_k = 0 \end{cases} \tag{B.26}$$

where $\omega_B = 1/(\xi^*_{k,m} + \varphi^*)$, $\xi^*_{k,m} > 0$, $\varphi^* > 0$.

References

[1] J. Mitola III, ***Cognitive radio: An integrated agent architecture for software defined radio,*** Ph.D. dissertation, Royal Institute of Technology (KTH), Stockholm, Swedem, 2000.

[2] J. Mitola III and G. Q. Maguire Jr, **"Cognitive radio: making software radios more personal,"** IEEE Personal Communications, vol. 6, no. 4, pp. 13–18, Aug. 1999.

[3] National Telecommunications and Information Administration, **United states frequency allocation chart,** [Online]. Available: http://www.ntia.doc.gov/osmhome/allochrt.html.

[4] Federal Communications Commission, **"FCC Spectrum policy task force report, Report of the Spectrum Efficiency Working Group,"** ET Docket no. 02–135, Nov. 2002.

[5] Federal Communications Commission, **"Facilitating opportunities for flexible, efficient, and reliable spectrum use employing cognitive radio technologies,"** notice of proposed rulemaking and order, FCC 03–322, 2003.

[6] S. Haykin, **"Cognitive radios: Brain-empowered wireless communications,"** IEEE J. Sel. Areas Commun., vol. 23, no. 2, pp. 13–18, Feb. 2005.

[7] T. A. Weiss and F. K. Jondral, **"Spectrum pooling: An innovative strategy for the enhancement of spectrum efficiency,"** IEEE Commun. Mag., vol. 42, no. 3, pp. S8–14, Mar. 2004.

[8] E. Hossain (Editor) and V. K. Bhargava (Editor), ***Cognitive Wireless Communication Networks,*** Springer, Nov. 2007.

[9] Zhu Han and K.J. Ray Liu, ***Resource Allocation for Wireless Networks: Basics, Techniques, and Applications,*** Cambridge University Press, April 2008.

[10] T.S. Rappaport, ***Wireless Communications: Principle and Practice (2nd ed.),*** Englewood Cliffs, NJ: Prentice-Hall, 2002.

[11] V. Srivastava and M. Motani, **"Cross-layer design: a survey and the road ahead,"** IEEE Commun. Mag., vol. 43, no. 12, pp. 112–119, Dec. 2005.

[12] X. Lin, N. Shroff and R. Srikant, **"A tutorial on cross-layer optimization in wireless networks,"** IEEE Journal Sel. Areas Commun., vol. 24, no. 8, pp.1452–1463, Aug. 2006.

[13] W. Huang and K. Letaief, **"Cross-layer scheduling and power control combined with adaptive modulation for wireless ad hoc networks,"** IEEE Trans. Commun., vol. 57, no. 4, pp. 728–739, Apr. 2007.

[14] V. T. Raisinghani and S. Iyer, **"Cross-layer feedback architecture for mobile device protocol stacks,"** IEEE Commun. Mag., vol. 44, no. 1, pp. 85–92, Jan. 2006.

[15] M. Chiang, **"To layer or not to layer: balancing transport and physical layers in wireless multihop networks,"** in Proc. IEEE INFOCOM Conf., vol. 4, pp. 2525-2536, Mar. 2004.

[16] L. Iannone and S. Fdida, **"Evaluating a cross-layer approach for routing in wireless mesh networks,"** Telecommun. Systems Journal (Springer), Special issue: Next Generation Networks- Architectures, Protocols, Performance, vol. 31, no. 2-3, pp. 173–193, Mar. 2006.

[17] A. Tang, J. Wang and S. Low, **"Is fair allocation always inefficient,"** in Proc. 23rd IEEE Conf. on Computer and Commun. (INFOCOM), vol. 1, pp. 35–45, Mar. 2004.

[18] S. Boyd and L. Vandenberghe, ***Convex Optimization,*** Cambridge University Press, 2004.

[19] M. S. Bazaraa, H. D. Sherali and C. M. Shetty, ***Nonlinear Programming: Theory and Algorithms* (2nd ed.),** New York: Wiley, 1993.

[20] S. Martello and P. Toth, ***Knapsack Problems***, New York: Wiley, 1990.

[21] M. Cooper and K.Farhangian, **"Multicriteria optimization for nonlinear integer-variable programming,"** Large Scale Systems 9, pp. 73–78, 1985.

[22] M. L. Fisher, **"The Lagrangian relaxation method for solving integer programming problems,"** ACM Journal of Management Science, vol. 50, no. 12, pp. 1861–1871, Dec. 2004.

[23] D. Fudenberg and J. Tirole, ***Game Theory***, MIT Press, Cambridge, MA, 1991.

[24] N. Nisan, T. Roughgarden, E. Tardos and V.V. Vazirani, ***Algorithmic Game Theory***, Cambridge University Press, 2007.

[25] R. G. Gallager, ***Information Theory and Reliable Communication***, New York: John Wiley & Sons, 1968.

[26] N. Papandreou and T. Antonakopoulo, **"A new computationally efficient discrete bit-loading algorithm for DMT applications,"** IEEE Trans. Commun., vol. 53, no. 5, pp. 785–789, May 2005.

[27] L. Piazzo, **"Fast algorithm for power and bit allocation in OFDM systems,"** IEEE Electronics Letters, vol. 35, no. 25, pp. 2173–2174, Dec. 1999.

[28] J. Jang and K. B. Lee, **"Transmit power adaptation for multiuser OFDM systems,"** IEEE Journal Sel. Areas in Commun., vol. 21, no. 2, pp. 171–178, Feb. 2003.

[29] G. Yu, Z. Zhang, Y. Chen, J. Shi and P.Qiu, **"A novel resource allocation algorithm for real-time services in multiuser OFDM systems,"** in Proc. IEEE 63rd Vehicular Tech. Conf. (VTC 2006-Spring), vol. 3, pp. 1156–1160, May 2006.

[30] Z. Shen, J. G. Andrews and B. L. Evans, **"Adaptive resource allocation in multiuser OFDM systems with proportional rate constraints,"** IEEE Trans. Wireless Commun., vol. 4, no. 6, pp. 2726–2737, Nov. 2005.

[31] P. Svedman, S. K. Wilson and B. Ottersten, **"A QoS-aware proportional fair scheduler for opportunistic OFDM,"** in Proc. IEEE 60th Vehicular Tech. Conf.(VTC 2004-Fall), vol. 1, pp. 558–562., Sept. 2004.

[32] P. Parag, S. Bhashyam and R. Aravind, **"A subcarrier allocation algorithm for OFDMA using buffer and channel state information,"** in Proc. IEEE 62nd Vehicular Tech. Conf. (VTC 2005-Fall), vol. 1, pp. 622–625, Sept. 2005.

[33] M. Andrews, K. Kumaran, K. Ramanan, etc., **"Providing quality of service over a shared wireless link,"** IEEE Commun. Magazine, vol. 39, no. 2, pp. 150–154, Feb. 2001.

[34] S. Ryu, B. H. Ryu, H. Seo, etc., **"Wireless packet scheduling algorithm for OFDMA system based on time-utility and channel state,"** ETRI Journal, vol. 27, no. 6, pp. 777–787, 2005.

[35] G. Song and Y. Li, **"Cross-layer optimization for OFDM wireless networks-part I: theoretical framework,"** IEEE Trans. Wireless Commun., vol. 4, no. 2, pp. 614–624, Mar. 2005.

[36] C. Mohanram and S. Bhashyam, **"A sub-optimal joint subcarrier and power allocation algorithm for multiuser OFDM,"** IEEE Commun. Letters, vol. 9, no. 8, pp. 685–687, Aug. 2005.

[37] S. S. Jeong, D. G. Jeong and W. S. Jeon, **"Cross-layer design of packet scheduling and resource allocation in OFDMA wireless multimedia networks,"** in Proc. IEEE 63rd Vehicular Tech. Conf. (VTC 2006-Spring), vol. 1, pp. 309–313, May 2006.

[38] D. S. W. Hui, V. K. N. Lau and W. H. Lam, **"Cross-layer design for OFDMA wireless systems with heterogeneous delay requirements,"** IEEE Trans. Wireless Commun., vol. 6, no. 8, pp. 2872–2880, Aug. 2007.

[39] W. Yu, G. Ginis and J. M. Cio, **"Distributed multiuser power control for digital subscriber lines,"** IEEE Journal Sel. Areas in Commun., vol. 20, no. 5, pp. 1105–1115, Jun. 2002.

[40] Z. Han, Z. Ji and K. J. R. Liu, **"Non-cooperative resource competition game by virtual referee in multi-cell OFDMA networks,"** IEEE Journal Sel. Areas in Commun., vol. 25, no. 6, pp. 1079–1090, Aug. 2007.

[41] G. Kulkarni, S. Adlakha and M. Srivastava, **"Subcarrier allocation and bit loading algorithms for OFDMA-based wireless networks,"** IEEE Trans. Mobile Computing, vol. 4, no. 6, pp. 652–662, Nov.-Dec. 2005.

[42] D. Maldonado, B. Lie, A. Hugine, etc., **"Cognitive radio applications to dynamic spectrum allocation,"** in Proc. IEEE Dyspan'05, pp. 597-600, Nov. 2005.

[43] J. A. Stine, **"Spectrum management: the killer application of ad hoc and mesh networking,"** in Proc. IEEE Dyspan'05, pp. 184-193, Nov. 2005.

[44] G. Bansal, M. J. Hossain and V. K. Bhargava, **"Adaptive power loading for OFDM-based cognitive radio systems,"** in Proc. IEEE Int. Conf. on Commun. (ICC '07), pp. 5137–5142, Jun. 2007.

[45] T. Qin and C. Leung, **"Fair adaptive resource allocation for multiuser OFDM cognitive radio systems,"** in Proc. 2nd Int. Conf. CHINACOM, pp. 115–119, Aug. 2007.

[46] P. Wang, M. Zhao, L. Xiao, etc., **"Power allocation in OFDM-based cognitive radio systems,"** in Proc. IEEE Global Telecommun. Conf. (GLOBECOM'07), pp. 4061–4065, Nov. 2007.

[47] P. Cheng, Z. Zhang, H.-H. Chen, etc., **"Optimal distributed joint frequency, rate and power allocation in cognitive OFDMA systems,"** IET Commun., vol. 2, no. 6, pp. 815–826, Jul. 2008.

[48] T. C. Clancy, **"Formalizing the interference temperature model,"** Wiley Wireless Commun. Mobile Computing, vol. 7, no. 9, pp. 1077–1086, May 2007.

[49] W. Wang, W. Wang, Q. Lu and T. Peng, **"An uplink resource allocation scheme for OFDMA-based cognitive radio networks,"** Int. J. Commun. Syst., vol. 22, no. 5, pp. 603-623, 2009.

[50] J.-A. Bazerque and G. B. Giannakis, **"Distributed scheduling and resource allocation for cognitive OFDMA radios,"** in Proc. 2nd Int. Conf. on Cognitive Radio Oriented Wireless Networks and Commun. (CrownCom 2007), pp. 343–350, Aug. 2007.

[51] Y. Chen, G. Yu, Z. Zhang, etc., **"On cognitive radio networks with opportunistic power control strategies in fading channels,"** IEEE Trans. Wireless Commun., vol. 7, no. 7, pp. 2752–2761, Jul. 2008.

[52] H. Luo, G. Yu and Z. Zhang, **"Optimal bit and power allocation in broadband cognitive radio system,"** in Proc. IEEE 67th Vehicular Tech. Conf. (VTC 2008-Spring), pp. 1569–1573, May 2008.

[53] Y. Xing, C. N. Mathur, M. A. Haleem, etc., **"Dynamic spectrum access with QoS and interference temperature constraints,"** IEEE Trans. Mobile Computing, vol. 6, no. 4, pp. 423–433, Apr. 2007.

[54] A. Attar, O. Holland, M. R. Nakhai, etc., **"Interference limited resource allocation for cognitive radio in orthogonal frequency-division multiplexing networks,"** IET Commun., vol. 2, no. 6, pp. 806–814, Jul. 2008.

[55] F. Wang, M. Krunz, and S. Cui, **"Price-based spectrum management in cognitive radio networks,"** IEEE Journal Sel. Topics in Signal Processing, vol. 2, no. 1, pp. 74–87, Feb. 2008.

[56] C. K. Tan, T.C. Chuah and S.W. Tan, **"Fair subcarrier and power allocation for multiuser orthogonal frequency-division multiple access cognitive radio networks using a Clonel Blotto game,"** IET Communs. vol.5, no. 11, pp. 1607-1618, 2011.

[57] N. Nie and C. Comaniciu, **"Adaptive channel allocation spectrum etiquette for cognitive radio networks,"** Springer Mobile Networks and Applications, vol. 11, no. 6, pp. 779–797, Dec. 2006.

[58] M. Felegyhazi, M. Cagalj, S. S. Bidokhti, etc., **"Non-cooperative multi-radio channel allocation in wireless networks,"** in Proc. 26th IEEE Int. Conf. on Computer Commun. (INFOCOM 2007), pp. 1442–1450, May 2007.

[59] D. Monderer and L. S. Shapley, **"Potential games,"** Elsevier Games and Economic Behavior, vol. 14, no. 1, pp. 124–143, May 1996.

[60] H. Su and X. Zhang, **"Cross-layer-based opportunistic MAC protocols for QoS provisionings over cognitive radio wireless networks,"** IEEE Journal Sel. Areas Commun., vol. 26, no. 1, pp. 118–129, Jan. 2008.

[61] J. W. Mwangoka, K. B. Letaief and Z. Cao, **"Robust end-to-end QoS maintenance in non-contiguous OFDM based cognitive radios,"** in Proc. IEEE Int. Conf. on Commun. (ICC '08), pp. 2905–2909, May 2008.

[62] Y.H. Zhang and C. Leung, **"A distributed algorithm for resource allocation in OFDM cognitive radio systems,"** IEEE Trans. Veh. Tech., vol. 60, no. 2, pp. 546-554, 2011.

[63] R. Wang, V. K. N. Lau, L. Lv and B. Chen, **"Joint cross-layer scheduling and spectrum sensing for OFDMA cognitive radio systems,"** IEEE Trans. Wireless Commun., vol. 8, no. 5, pp. 2410–2416, May 2009.

[64] Y. Zhang and C. Leung, **"Cross-layer resource allocation for mixed services in multiuser OFDM-based cognitive radio systems,"** IEEE Trans. Veh. Tech., vol. 58, no. 8, pp. 4605–4619, Oct. 2009.

[65] T. A. Weiss, J. Hillenbrand, A. Krohn and F. K. Jondral, **"Mutual interference in OFDM-based Spectrum Pooling Systems,"** in Proc. IEEE 59th Vehicular Technology Conference (VTC 2004-Spring), vol. 4, pp. 1873-1877, 2004.

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

[67] H. Celebi and H. Arslan, **"Cognitive Positioning Systems,"** IEEE Trans. Wireless Commun., vol. 6, no. 12, pp. 4475-4483, Dec. 2007.

[68] H. Nam, M. B. Ghorbel and M.-S. Alouin, **"Location-based resource allocation for OFDMA cognitive radio systems,"** in Proc. CrownCom2010, pp. 1-5, Jun. 2010.

[69] P. Kyosti, et al., **"WINNER II Channel Models,"** IST-WINNER D1.1.2 v 1.1, Sept. 2007. Available: https://www.ist-winner.org/WINNER2-Deliverables/D1.1.2v1.1.pdf

[70] M. Ben Ghorbel, H. Nam and M.-S. Alouini, **"Discrete rate resource allocation for OFDMA cognitive radio systems with location information,"** in PIMRC workshops 2010, pp. 247-251, Sept. 2010.

[71] Y. Zhang, ***Resource allocation for OFDM-based cognitive radio systems*,** Ph.D. dissertation, Univ. British Columbia, Vancouver, BC, Canada, 2008.

[72] D. G. Luenberger, ***Optimization by Vector Space Methods*,** NewYork: Wiley, 1969.

[73] K. Kim, Y. Han and S. L. Kim, **"Joint Subcarrier and Power Allocation in Uplink OFDMA Systems,"** IEEE Commun. Lett., Vol. 9, pp.526-528, Jun. 2005.

[74] Y. Ma, **"Rate Maximization for Downlink OFDMA With Proportional Fairness,"** IEEE Trans. Veh. Technol., vol.57, pp.3267-3274, Sep.2008.

[75] H.P. Yao, Z. Zhou, H. Liu and L.Y. Zhang, **"Optimal power allocation in joint spectrum underlay and overlay cognitive radio networks,"** in Proc. CrownCom2009, pp. 1-5, 2009.

[76] M. Rong and S.H. Zhu, **"Hybrid overlay/underlay spectrum sharing in cognitive radio networks,"** IEICE Trans. Commun., vol. E94-B, no. 9, pp. 2672-2676, 2011.

[77] H. Zhang, D. L. Ruyet, D.Roviras and H. Sun, **"Resource allocation of non-cooperative multi-cell for cognitive radio networks,"** Int. Symposium on Wireless Commun. Systems (ISWCS 2009), pp. 101-105, 2009.

[78] Y. Ma, D. I. Kim and Z.Q. Wu, **"Optimization of OFDMA-Based cellular cognitive radio networks,"** IEEE Trans. Commun., vol. 58, no. 8, pp. 2265-2276, 2010.

[79] 3GPP TSG RAN, **"Requirements for further advancements for Evolved Universal Terrestrial Radio Access (E-UTRA),"** 3GPP TR 36.913 V9.0.0, Dec. 2008.

[80] 3GPP TSG RAN WG4, **"RF requirements for multicarrier and multiple radio access technology BS (Release 9),"** 3GPP TR 37.900 V1.0.0, Sept. 2009.

[81] T.H. Kim and T.J. Lee, **"Frequency sharing mechanism using pilot sensing in OFDMA-based cognitive radio networks,"** IEICE Trans. Commun., vol. E94-B, no. 4, pp. 986-996, 2011.

Publications

List of Publications Directly Related to The Dissertation

Journal Papers

1. Hailan Peng, Takeo Fujii, **"Joint cross-layer resource allocation and interference avoidance with QoS support for cognitive radio systems,"** EURASIP (European Association for Signal Processing) Journal on Wireless Communications and Networking, 2012:41, Feb. 2012.

2. Hailan Peng, Takeo Fujii, **"Distributed Resource Allocation for Multi-cell Cognitive Radio Networks based on Intra-cell Overlay and Inter-cell Underlay Spectrum Sharing,"** IEICE Transactions on Communications, vol.E96-B, no.06, pp. 1566-1576, Jun. 2013.

International Conference Papers

3. Hailan Peng, Takeo Fujii, **"Joint Resource Allocation and Interference Avoidance with Fairness Consideration for Multi-cell Cognitive Radio Networks,"** 2012 IEEE Wireless Communications and Networking Conference (WCNC 2012), Paris, France, Apr. 1-4, 2012.

4. Hailan Peng, Takeo Fujii, **"Joint Resource Allocation and Interference Avoidance for Multi-cell Cognitive Radio Networks with Primary Assistance,"** International Triangle Symposium on Advanced ICT 2011 (TriSAI 2011), Daejeon, Korea, Aug. 25-26, 2011.

5. Hailan Peng, Takeo Fujii, **"Cross-Layer Resource Allocation for Downlink OFDMA-based Cognitive Radio Systems using Convex Optimization,"** International Triangle Symposium on Advanced ICT 2010 (TriSAI 2010), Beijing, China, Oct. 25-27, 2010.

6. Hailan Peng, Takeo Fujii, **"Joint Cross-Layer Resource Allocation and Interference Avoidance with QoS Support for Multiuser Cognitive Radio Systems,"** in the 7th International Symposium on Wireless Communication Systems (ISWCS 2010), York, UK, Sept. 19-22, 2010.

Domestic Conference Papers

7. Hailan Peng, Takeo Fujii, **"Joint Resource Allocation and Interference Avoidance for Multi-cell Cognitive Radio Networks with Primary Assistance,"** IEICE Society Conference 2011, B-17-18, Sapporo, Japan, Sept. 13-16, 2011.

Lists of Publications from Masters (M.S.)

1. Hailan Peng, Jie Cai, Lei You, Yong Zhang, Mei Song, **"Cost Performance Analysis of Multi-hop Relay based Broadband Access Network,"** Domestic Journal of Application Research of Computers, vol. 26 No. 2, pp. 677-679, Feb. 2009.

2. Hailan Peng, Lei You, Yong Zhang, Mei Song, **"Joint Routing and Scheduling Scheme in Wireless Mesh Backhaul Networks,"** 2008 China-Ireland International Conference on Information and Communications Technologies (CIICT 2008), Beijing, China, Sept. 26-28, 2008.

3. Hailan Peng, Wei Luo, Lei You, Mei Song, **"Mesh QoS Routing: A Novel QoS Routing Protocol for WMNs,"** Future Telecommunication Conference 2007 (FTC 2007), Beijing, China, Oct. 11-12, 2007.

Hailan Peng

She has received the B.E. and M.E. degrees in Electrical Engineering from Beijing University of Posts and Telecommunications, China, in 2006 and 2009, respectively. Currently, she is pursuing her Ph.D. degree in Advanced Wireless Communication research Center, The University of Electro-Communications, Tokyo, Japan. This dissertation is submitted for the degree of Doctor of Philosophy in Engineering in The University of Electro-Communications. Her current research interests include cross-layer optimization, resource allocation, multi-cell cognitive radio networks, game theory and graph theory.