Resource Allocation with Carrier Aggregation for Spectrum Sharing in Cellular Networks

Haya Shajaiah

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Electrical Engineering

Charles Clancy, Chair
Jeffrey H. Reed
Lamine Mili
Anil Vullikanti
Ravi Tandon

March 14, 2016
Arlington, Virginia

Keywords: Optimal Resource Allocation, Carrier Aggregation, Utility Proportional Fairness, User Discrimination, Resource Block Scheduling, Multi-Tier Secure Spectrum Auction

©Copyright 2016, Haya Shajaiah

Resource Allocation with Carrier Aggregation for Spectrum Sharing in Cellular Networks

Haya Shajaiah

(ABSTRACT)

Recently, there has been a massive growth in the number of mobile users and their traffic. The data traffic volume almost doubles every year. Mobile users are currently running multiple applications that require higher bandwidth which makes users so limited to the service providers' resources. Increasing the utilization of the existing spectrum can significantly improve network capacity, data rates and user experience. Spectrum sharing enables wireless systems to harvest under-utilized swathes of spectrum, which would vastly increase the efficiency of spectrum usage. Making more spectrum available can provide significant gain in mobile broadband capacity only if those resources can be aggregated efficiently with the existing commercial mobile system resources. Carrier aggregation (CA) is one of the most distinct features of 4G systems including Long Term Evolution Advanced (LTE-Advanced). In this dissertation, a resource allocation with carrier aggregation framework is proposed to allocate multiple carriers resources optimally among users with elastic and inelastic traffic in cellular networks. We use utility proportional fairness allocation policy, where the fairness among users is in utility percentage of the application running on the user equipment (UE). A resource allocation (RA) with CA is proposed to allocate single or multiple carriers resources optimally among users subscribing for mobile services. Each user is guaranteed a minimum quality of service (QoS) that varies based on the user's application type. In addition, a resource allocation with user discrimination framework is proposed to allocate single or multiple carriers resources among users running multiple applications. Furthermore, an application-aware resource block (RB) scheduling with CA is proposed to assign RBs of multiple component carriers to users' applications based on a utility proportional fairness scheduling policy.

We believe that secure spectrum auctions can revolutionize the spectrum utilization of cellular networks and satisfy the ever increasing demand for resources. Therefore, a framework for multi-tier dynamic spectrum sharing system is proposed to provide an efficient sharing of spectrum with commercial wireless system providers (WSPs) with an emphasis on federal spectrum sharing. The proposed spectrum sharing system (SSS) provides an efficient usage of spectrum resources, manages intra-WSP and inter-WSP interference and provides essential level of security, privacy, and obfuscation to enable the most efficient and reliable usage of the shared spectrum. It features an intermediate spectrum auctioneer responsible for allocating resources to commercial WSPs' base stations (BS)s by running secure spectrum auctions. In order to insure truthfulness in the proposed spectrum auction, an optimal bidding mechanism is proposed to enable BSs (bidders) to determine their true bidding values. We also present a resource allocation based on CA approach to determine the BS's optimal aggregated rate allocated to each UE from both the BS's permanent resources and winning auctioned spectrum resources.

Resource Allocation with Carrier Aggregation for Spectrum Sharing in Cellular Networks

Haya Shajaiah

GENERAL AUDIENCE ABSTRACT

In recent years, the number of mobile users and their traffic volume have increased rapidly. The data traffic volume almost doubles every year. Mobile users are currently running multiple applications that require higher bandwidth which makes users so limited to the service providers' resources. The volume of data traffic is expected to continue growing up and reaches 1000 times its value in 2010 by 2020 which is referred to as 1000x data challenge. With the increasing volume of data traffic, more spectrum is required. Federal agencies are now willing to share their spectrum with commercial users due to the high demand for spectrum by commercial operators. Making more spectrum available can provide significant gain in mobile broadband capacity only if those resources can be aggregated efficiently with the existing commercial mobile system resources. In this dissertation, we introduced new resource allocation methods for future wireless systems that takes into consideration aggregating multiple wireless providers' resources and showed the efficiency of the proposed methods compared to other existing methods in improving mobile users' quality of experience.

Preface

In the name of God, the Most Gracious, the Most Merciful
"Are those who know and those who do not know alike? Only the men of understanding are mindful"
[The Holy Quran, 39:9]

Acknowledgments

All praise and glory to Almighty Allah SWT, the Lord of the worlds. Peace and blessing of Allah be upon Prophet Muhammad PBUH. The Messenger of Allah Muhammad PBUH said in the Hadith that: "The best things that a man can leave behind are three: A righteous son who will pray for him, ongoing charity whose reward will reach him, and knowledge which is acted upon after his death" [Sunan Ibn e Majah].

First of all, I thank Allah for giving me health, wisdom and strength to pursue PhD studies. Next, I would like to express my sincere appreciation to my advisor Dr. Charles Clancy for all his support, advice, help, and guidance. I would not have accomplished this work without his vision and support. I also thank my PhD advisory committee members, Prof. Jeffrey H. Reed, Prof. Lamine Mili, Prof. Anil Vullikanti, and Prof. Ravi Tandon for their valuable suggestions and feedback on my research and dissertation.

Next, I would like to thank Dr. Ahmed Abdelhadi, currently a Research Scientist at Hume Center, for being a very helpful mentor during my PhD studies and for co-authoring many papers with me. I also thank my lab-mates Ms. Jasmin Mahal, Dr. Awais Khawar, Dr. Mo Ghorbanzadeh and Dr. Chowdhury Shahriar for the great experience I had while working with them.

Finally, I thank all my family members for their support and encouragement. I would like to express my gratitude to my beloved husband, Dr. Islam Younis, for his constant love, support and encouragement throughout my PhD studies over the past three and a half year. I would have to admit that this dissertation would not have been possible without Islam's support and love. I would also like to express my gratitude to my parents, Dr. Jamal Nouh Shajaiah and Amal Okab, for their encouragement, support, unconditional love and care. My dad has always been my role model in life. I admire and regard him with the utmost respect. My mom has done a lot for me, I specially thank her for all the help she has provided while I am pursuing PhD. I love my parents so much and I would not have made it this far without them. I also thank my sisters, Dr. Hiba and Hala, my brother, Hussein, my aunt, Nawar Okab, and my wonderful grandmother, Om Rasem whom I unfortunately lost few months before the completion of this dissertation, for their continuous encouragement and support. Moreover, I thank my dear friend Hamideh Bitaraf for always being with me in good and bad times during my wonderful days of PhD.

Dedication

To my parents, my beloved husband Islam, and our beloved children Rasem and Tayma.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

In recent years, the number of mobile subscribers and their traffic have increased rapidly. Mobile subscribers are currently running multiple applications, simultaneously, on their smart phones that require a higher bandwidth and make users so limited to the carrier resources. Multiple services are now offered by network providers such as mobile-TV and multimedia telephony [3]. According to the Cisco Visual Networking Index (VNI) [4], the volume of data traffic is expected to continue growing up and reaches 1000 times its value in 2010 by 2020 which is referred to as 1000x data challenge. With the increasing volume of data traffic, more spectrum is required [5]. However, due to spectrum scarcity and fragmentation, it is difficult to provide the required resources with a single frequency band. Therefore, aggregating frequency bands, that belong to different carriers, is needed to utilize the radio resources across multiple carriers and expand the effective bandwidth delivered to user terminals, leading to interband non-contiguous carrier aggregation [6].

1.1 Motivation and Background

Carrier aggregation is one of the most distinct features of 4G systems including LTE-Advanced. Given the fact that LTE requires wide carrier bandwidths to utilize such as 10 and 20 MHz, CA needs to be taken into consideration when designing the system to overcome the spectrum scarcity challenges. With the CA being defined in [7], two or more component carriers (CCs) of the same or different bandwidths can be aggregated to achieve wider transmission bandwidths between the evolve node B (eNodeB) and the UE. This feature allows LTE-Advanced to meet the International Mobile Telecommunications (IMT) requirements for the fourth-generation standards defined by the International Telecommunications Union (ITU) [8]. An overview of CA framework and cases is presented in [5]. Many operators are willing to add the CA feature to their plans across a mixture of macro cells and small cells. This will provide capacity and performance benefits in areas where small cell coverage is available while enabling network operators to provide robust mobility management on their macro cell networks.

The non-contiguous carrier aggregation task is a challenging. The challenges are both in hardware implementation and joint optimal resource allocation. Hardware implementation challenges are in the need for multiple oscillators, multiple RF chains, more powerful signal processing, and longer battery life [9].

Increasing the utilization of the existing spectrum can significantly improve network capacity, data rates and user experience. Some spectrum holders such as government users do not use their entire allocated spectrum in every part of their geographic boundaries most of the time. Therefore, the National Broadband Plan (NBP) and the findings of the President's Council of Advisors on Science and Technology (PCAST) spectrum study have recommended making the under-utilized federal spectrum available for secondary use [10]. Spectrum sharing enables wireless systems to use the underutilized spectrum efficiently. Making more

spectrum available can provide significant gain in mobile broadband capacity only if those resources can be aggregated efficiently with the existing commercial mobile system resources. As a result of the high demand for spectrum by commercial wireless operators, federal agencies are now willing to share their spectrum with commercial users. This has led to proposals to share spectrum allocated for federal radar operations with commercial users. The 3550-3650 MHz band, currently used for military radar operations, is identified for spectrum sharing between military radars and communication systems, according to the NTIA's 2010 Fast Track Report [11]. This band is very favorable for commercial cellular systems such as LTE-Advanced systems. Therefore, innovative methods are required to make spectrum sharing between radars and cellular systems a reality.

Beside CA capability, next-generation wireless networks need to support diverse QoS requirements of multiple applications since different applications require different application's performance. Furthermore, certain types of users may require to be given priority when allocating the network resources (i.e. such as public safety users) which needs to be taken into consideration when designing the resource allocation framework.

The public safety wide area wireless communication system is currently separate from the commercial cellular networks. Industries are willing to support both communities by providing a common technology. Release 12 of 3GPP LTE standards has enhanced LTE to support public safety requirements. Advanced standards such as LTE provide multimedia capabilities and voice and messages services at multi-megabit per second. The services that public safety networks provide such as communications for police, fire and ambulance require systems development to meet the communication needs of emergency services. A common technical standard for commercial and public safety users provides advantages for both. The public safety systems market is much smaller than the commercial cellular market which makes it unable to attract the level of investment that goes in to commercial cellular networks and this makes a common technical standards for both the best solution. The public safety

community gains access to the technical advantages provided by the commercial cellular networks whereas the commercial cellular community gains enhancement in their systems and make it more attractive to consumers. The USA has reserved spectrum in the 700MHz band for an LTE based public safety network. The current public safety standards support medium speed data which drives the need of new technology. An efficient resource allocation framework is needed for cellular networks that support both of commercial and public safety communities and takes into consideration that users' applications should not be treated evenly for both communities.

1.2 Carrier Aggregation

1.2.1 Motivation for Developing Carrier Aggregation

The idea of using multi-carrier has been driven by the rapid data user growth and the increasing demand for resources. Operators are facing operational challenges in terms of data capacity. The carrier aggregation feature has been added to Release 10 of the 3GPP LTE-Advance standard to allow single users to employ multiple carriers in order to achieve higher bandwidth [12]. With the increasing number of applications and their required bandwidth, smart phones are now require large bandwidth allocations which makes them limited to the network resources. The peak data rates required by IMT-Advanced can be satisfied LTE-Advanced as it support wider bandwidth by using the carrier aggregation feature.

Carrier aggregation is also needed because of the fact that the current frequency spectrum is highly segmented [13]. Figure 1.1 [14] shows the current frequency allocation table for the US and how segmented the spectrum is. Fragmented spectrum can be utilized more efficiently by aggregating non contiguous carriers.

The overall goal of carrier aggregation is to provide an enhanced QoS for mobile users

Figure 1.1: US Frequency Spectrum Allocation.

throughout the cell by combining peak capacities available at different frequencies, providing more consistent QoS to users by utilizing unused capacity available at other frequencies, improving mobility and enabling interference management.

1.2.2 Deployment Scenarios for Carrier Aggregation

Different deployment scenarios have been considered for the design of LTE-Advanced carrier aggregation [2]. Figure 1.2 shows five different deployment scenarios with two component carriers F1 and F2. The five scenarios are described below.

Scenario 1: Cells with the two carrier frequencies are collocated and overlaid in the same band. Both frequencies F1 and F2 almost have the same coverage area. Carrier aggregation enables a higher achievable data rates throughput the cell.

Scenario 2: Cells with the two carriers are collocated and overlaid in different bands. Different carriers have different coverage because higher frequency bands have larger path loss. Higher frequency bands carriers are used to improve data rates.

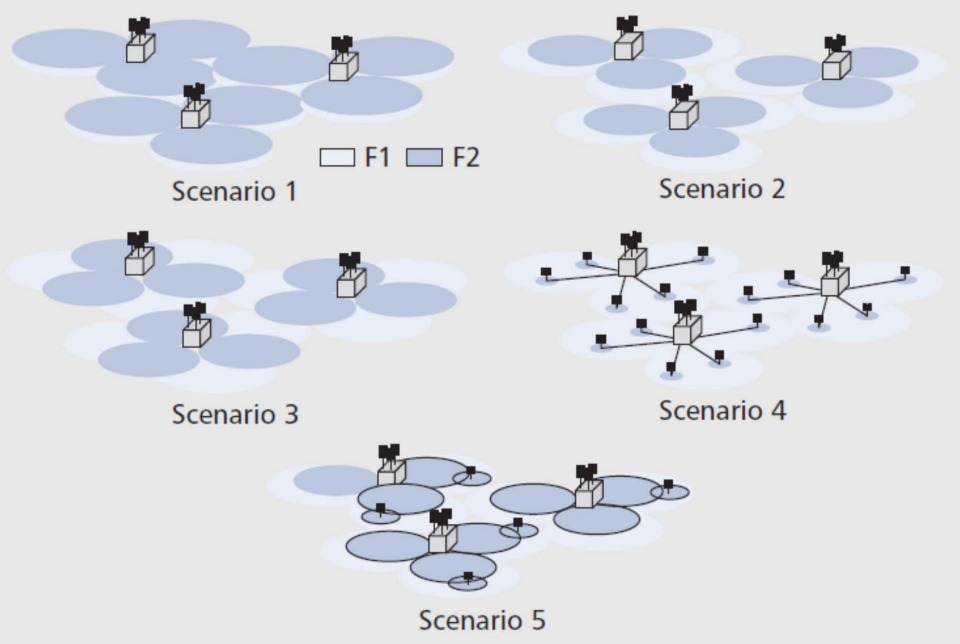


Figure 1.2: Carrier aggregation deployment scenarios with F2 > F1 [2].

Scenario 3: Cells with the two carriers are co-located in different bands. To improve the throughput of cell edge, the antennas for cells of F2 are directed to the cell boundaries of F1. Carrier aggregation is applied for areas with overlapping coverage.

Scenario 4: Remote radio heads (RRHs) of carrier F2 are used in hot spots to improve the throughput and cells of carrier F1 are the macro cells. There are usually different bands for frequencies F1 and F2. Carrier aggregation is applied for users under the coverage area of both the RRHs and the macro cells.

Scenario 5: Similar to scenario 2 except that in order to extend one of the frequencies coverage frequency selective repeaters are deployed.

1.2.3 Types of Carrier Aggregation

Three types of carrier aggregation have been defined in 3GPP in order to meet operators spectrum scenarios. These types are intra-band contiguous, intra-band non contiguous and inter-band non contiguous [5] The uplink and downlink can be configured independently.

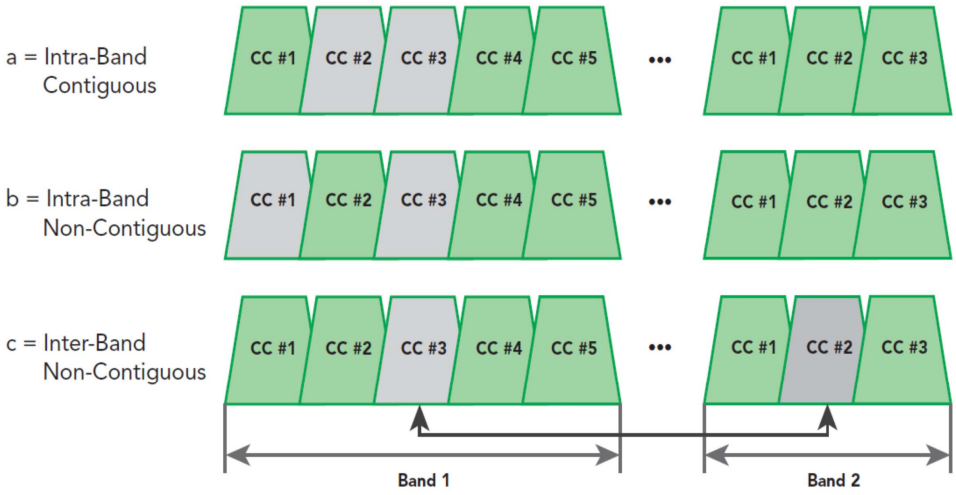


Figure 1.3: Types of carrier aggregation in LTE-Advanced.

However, the number of uplink carriers needs not to exceed the number of downlink carriers. The three types of CA are illustrated in Figure 1.3 and discussed below.

Intra-band continuous: This type refers to the situation where all carriers on the uplink or the downlink are adjacent in frequency [5]. The hardware implementation of this type of CA is not complicated since this type of CA can be achieved by a single RF chain. However this type of CA is unlikely since the current spectrum is highly segmented.

Intra-band non contiguous: In this type of CA, the combined carriers fall within the same band but are not adjacent in frequency [5]. This type is more realistic since the frequency bands are highly segmented. The hardware implementation of this type can simply be achieved through a single RF chain given that carriers are in the same frequency band.

Inter-band non contiguous: In this type of CA, the two carriers are within different bands [5]. The user hardware implementation for this type is the most complex since a single RF chain has limitation in terms of a certain band of interest for practical reasons.

1.3 Related Work

In this section, we discuss several related research work in the area of radio resource allocation for spectrum sharing and spectrum auction mechanisms.

1.3.1 Previous Studies in Resource Allocation for Spectrum Sharing

There has been several works in the area of resource allocation optimization to utilize the scarce radio spectrum efficiently. The authors in [**?**, 15–17] have used a strictly concave utility function to represent each user's elastic traffic and proposed distributed algorithms at the sources and the links to interpret the congestion control of communication networks. Their work have only focussed on elastic traffic and did not consider real-time applications as it have non-concave utility functions as shown in [18]. The authors in [19] and [20] have argued that the utility function, which represents the user application performance, is the one that needs to be shared fairly rather than the bandwidth. In this research work, we consider using resource allocation to achieve a utility proportional fairness that maximizes the user satisfaction. If a bandwidth proportional fairness is applied through a max-min bandwidth allocation, users running delay-tolerant applications receive larger utilities than users running real-time applications as real-time applications require minimum encoding rates and their utilities are equal to zero if they do not receive their minimum encoding rates.

The proportional fairness framework of Kelly introduced in [15] does not guarantee a minimum QoS for each user application. To overcome this issue, a resource allocation algorithm that uses utility proportional fairness policy is introduced in [21]. We believe that this approach is more appropriate as it respects the inelastic behavior of real-time applications. The

utility proportional fairness approach in [21] gives real-time applications priority over delay tolerant applications when allocating resources and guarantees that no user is allocated zero rate. In [21,22] and [23], the authors have presented optimal resource allocation algorithms to allocate single carrier resources optimally among mobile users. However, their algorithms do not support multi-carrier resource allocation. To incorporate the carrier aggregation feature, we have introduced a multi-stage resource allocation using carrier aggregation in [24]. In [25] and [26], we present resource allocation with users discrimination algorithms to allocate the eNodeB resources optimally among mobile users with elastic and inelastic traffic. In [27], the authors have presented a radio resource block allocation optimization problem using a utility proportional fairness approach. The authors in [28] have presented an application-aware resource block scheduling approach for elastic and inelastic adaptive real-time traffic where users are assigned to resource blocks.

On the other hand, resource allocation for single cell multi-carrier systems have been given extensive attention in recent years [29–31]. In [32–35], the authors have represented this challenge in optimization problems. Their objective is to maximize the overall cell throughput with some constraints such as fairness and transmission power. However, transforming the problem into a utility maximization framework can achieve better users satisfaction rather than better system-centric throughput. Also, in practical systems, the challenge is to perform multi-carrier radio resource allocation for multiple cells. The authors in [36, 37] suggested using a distributed resource allocation rather than a centralized one to reduce the implementation complexity. In [38], the authors propose a collaborative scheme in a multiple base stations (BSs) environment, where each user is served by the BS that has the best channel gain with that user. The authors in [39] have addressed the problem of spectrum resource allocation in carrier aggregation based LTE-Advanced systems, with the consideration of UEs MIMO capability and the modulation and coding schemes (MCSs) selection.

Most of the previous research work have focused on finding resource allocation approaches

for intra-system and intra-operator of a single network operator. However, current research on resource allocation are for more complex network topologies [40, 41]. Carrier aggregation in networks that involve multiple network operators in HetNets need to be further investigated. In [42], the authors have analyzed the performance of their proposed carrier aggregation framework that combines a statically assigned spectrum with spectrum resources from a shared spectrum pool. A tractable multi-band multi-tier CA models for HetNets are proposed in [43]. Two models are considered: multi-flow CA and single-flow CA, each UE performs cell selection based on the reference signal's maximum received power. A major concern about deploying small cells is their small coverage areas and low transmit power. The authors in [44, 45] have addressed this issue and suggested biasing to allow small cells to expand their coverage areas.

In the past, wireless systems were able to share government bands by operating on a low power to prevent the interference with the incumbent systems such as wireless local area network (WLAN) in the 5.25-5.35 and 5.47-5.725 GHz radar bands [46]. Small cells operating in a low power have been proposed recently to operate in the 3.5 GHz radar band [47].

To mitigate radar interference to LTE-Advanced systems, a spatial approach for spectrum sharing between a MIMO radar and LTE cellular system with N_{BS} base stations was proposed in [48]. Radar signals are manipulated such that they are not a source of interference to the LTE-Advanced BSs. Because there exist many interference channels between the two systems, the interference channel with the maximum null space dimension is chosen based on the algorithm proposed by the authors, the radar signal is then projected onto the null space of that interference channel to mitigate interference to the LTE-Advanced BS. This spatial approach results in small degradation in the radar performance [49].

In [49], the authors proposed a technique to project radar waveforms onto the null space of an interference channel matrix between the radar and the communication system. In their proposed approach, the cognitive radar is assumed to have full knowledge of the interference

channel and modifies its signal vectors in a way such that they are in the null space of the channel matrix. In order to avoid interference to the communication system, a radar signal projection onto the null space of interference channel between radar and communication systems is presented in [50]. In [51], a novel signal processing approach is developed for coherent MIMO radar to minimize the arbitrary interferences generated by wireless systems from any direction while operating at the same frequency using cognitive radio technology.

As 4G wireless mobile systems including LTE and LTE-Advanced continue to evolve, higher data rates and improved QoS, even for cell edge users, are promised to be guaranteed for end users. The capacity promised by MIMO systems may not be fully realizable without having a sufficient control of inter-cell interference which limits throughput for cell-edge users [52]. In order to mitigate inter-cell interference, three major frequency reuse patterns can be used: hard frequency reuse, fractional frequency reuse (FFR) and soft frequency reuse. Hard frequency reuse divides the system bandwidth into a number of sub-bands according to certain reuse factor such that neighboring cells transmit on different sub-bands. FFR divides the system bandwidth into an inner and an outer part. The inner part is only allocated to the near users with reduced power while applying a frequency reuse factor of 1 such that the inner part is reused by all other BSs. On the other hand, the outer part of bandwidth is allocated to far users (cell edge users) with a frequency reuse factor greater than one. Soft frequency reuse allows the overall bandwidth to be shared by all BSs with a reuse factor of 1 while the BSs are restricted to certain power bound for the transmission on each sub-carrier. Hard frequency reuse suffers from a reduced spectral efficiency whereas soft frequency reuse [53, 54] has full spectral efficiency but requires centralized coordination of resource allocation which becomes impractical for a large number of BSs. Unless otherwise specified, we consider using FFR as it compromises between hard and soft frequency reuse and therefore will be a proficient option for future wireless systems. However, we do not intend to address inter-cell interference throughout this dissertation.

1.3.2 Previous Studies in Spectrum Auctions

Traditionally, radio spectrum management is controlled by a central government agency such as the Federal Communications Commission (FCC) in the United States. Such a centralized spectrum assignment mechanism predetermines static bands for specific usage without taking into consideration the service requirements and the dynamic nature of the radio spectrum. This results in an under-utilized pre-assigned spectrum bands while many of the commercial bands are overcrowded due to the rapid growth of wireless services. To address this limitation in the spectrum utilization, the FCC has legalized secondary markets for spectrum such that a primary spectrum licensee can lease its under-utilized spectrum to secondary incumbents [55]. Inspired by microeconomics mechanisms [56–58], spectrum auction seems to be a promising solution to release the under-utilized spectrum to potential secondary users [59–61]. There has been some previous work to deal with security issues in auction design. These works have focused on adding some new features to the auction design, such as confidentiality, fairness [62, 63] and anonymity.

Because of the reusability feature of the radio spectrum, traditional auctions can not be directly used in a spectrum auction design. Spectrum auctions should allow bidders, that are not within the interference radius of each other, to use the same frequency simultaneously. Therefore, the optimal spectrum allocation is considered NP-complete [64,65] whereas conventional auctions are based on optimal allocations [59].

Most early works in spectrum auctions, such as [59,61], have focused on single-seller multi-buyer auctions that deal with homogeneous channels. In [59], the authors have proposed **VERITAS**, a truthful mechanism that supports an eBay-like dynamic spectrum market. It is a good fit for short term and small regions based spectrum auction which is not the case in FCC required spectrum auction which is for long term and large geographical regions. To deal with interference between neighboring bidders, a conflict graph and a wireless spectrum auction framework have been proposed in [59]. Based on these concepts, a conflict graph is

used to represent the interference relationship in **VERITAS** [59]. In a sealed secondary price and VCG auctions, the dominant strategy for certain bidder, when he has no information about other bidders' bids, is to bid with his true evaluation values [66]. The authors in [67] have showed that it is not always right to allocate spectrum bands to the bidder with the highest bid, as proposed in [59], if the sum of the neighbors bids is much higher than the highest bid. Their proposed solution is based on grouping nodes such that nodes with no interference are grouped together. However, their group partition approach is NP-complete under interference constraints [65].

The authors in [60] have proposed **TRUST**, a spectrum trading approach that satisfies some good properties. However, it achieves truthfulness while sacrificing one group of bidders, as it takes the group's bid as the clearing price. In [68], the authors have improved the idea of **TRUST** as they succeeded to achieve truthfulness by only sacrificing one buyer in each group. But, both works [60,68] have inherited **McAfee** mechanism [69] which requires homogeneous channels. In [70], the proposed **TASC** mechanism was the first to consider heterogeneous channels. However, it can reduce the system efficiency as all channels are restricted to a unique clearing price. In [71, 72], **TASC** mechanism has been extended to consider spectrum reusability and diversity of channel characteristics. In [73], the authors have proposed a privacy preserving auction for spectrum trading. In [74, 75], an auction based framework is proposed. A third party leases its unused resources to service providers to provide dynamic cellular offloading.

In [59, 72], the authors have exploited frequency interference property. They used interference graph model that makes spectrum allocation, allows spectrum reuse and avoids interference. In [60,68], the authors have utilized the reusability property by dividing buyers into groups such that buyers in the same group do not interfere with each other. Each group of buyers either wins or loses the same channel.

Most of existing works have focused on the case of identical spectrum bands. Spectrum

reusability in an auction design has been first addressed in [60]. In [76], the authors have modeled a spectrum auction based on spectrum reusability in a time-frequency division manner. The authors in [77] have also considered spectrum reusability in their auction design by assuming that each spectrum buyer is allowed to have multiple radios. The proposed **MTSSA** scheme also supports the frequency reusability property.

Beside the properties of secondary price auctions that are beneficial to have in a spectrum auction, i.e. such as incentive compatibility, individual rationality and no positive transfers, it is important to secure the spectrum auction to avoid potential back room dealing. An ideal spectrum auction design would allow the auctioneer to find the best allocation of the frequency bands, determine the winners and their payments while the bidders keep their actual bidding values secret and unknown to the auctioneer. This can prevent frauds made by insincere auctioneers and bid rigging between the auctioneer and the bidders. There has been some previous works in secure spectrum auctions. The authors in [78–80], have used homomorphic encryption to secure traditional auction designs. In [81], the authors have considered frequency reuse in their secure spectrum auction design, and propose **THEMIS**. However, **THEMIS** does not support multi-tier spectrum sharing systems where spectrum reuse is possible among multiple service providers. In these systems a dynamic spectrum sharing approach is required to provide an efficient sharing of the spectrum among multiple service providers. Furthermore, the computational and communication complexity of **THEMIS** is closely related to the number of available frequency bands. Therefore **THEMIS** may incur a heavy cost for a large number of frequency bands and bidding values.

1.4 Contributions

An outline of the contributions of this research study is as follows:

1.4.1 Multi-Stage Resource Allocation with Carrier Aggregation

In this area, the following contributions are made:

- We formulate the resource allocation with CA problem into a convex optimization framework and develop a utility proportional fairness RA with CA optimization problem that gives priority to real-time application due to the nature of the sigmoidal-like utility functions.
- We develop a distributed multi-stage resource allocation algorithm to solve the optimization problem and allocate the eNodeBs' resources optimally among users, with a minimum QoS guaranteed for each user, and present its corresponding simulation results.
- We develop a price selective centralized resource allocation with CA scheme to allocate multiple carriers resources optimally among users located under one carrier or multiple carriers coverage area and present the corresponding algorithm that allows each UE to choose its primary carrier and the order of its secondary carriers based on the price offered by all in range carriers.
- We present simulation results for the performance of the proposed price selective centralized algorithm and show how it converges to the optimal rates whether the eNodeBs' available resources are abundant or scarce.
- We develop a spectrum sharing approach, based on multi-stage resource allocation with CA, for sharing the Federal under-utilized 3.5 GHz spectrum with commercial users and present its corresponding simulation results.

1.4.2 Resource Allocation with Joint Carrier Aggregation

In this area, the following contributions are made:

- We formulate a RA optimization problem with joint CA to allocate multi-carrier resources and use a utility proportional fairness approach to solve for logarithmic and sigmoidal-like utility functions representing delay-tolerant and real-time applications, respectively.
- We prove that the RA with joint CA optimization problem is convex and therefore the global optimal solution is tractable. In addition, we present a robust distributed resource allocation algorithm to solve the optimization problem and provide optimal rates in high-traffic and low-traffic situations. We present simulation results for the proposed algorithm and compare its performance with the multi-stage RA with CA algorithm.

1.4.3 Resource Allocation with User Discrimination

In this area, the following contributions are made:

- We develop a spectrum sharing scheme for public safety and commercial users running elastic or inelastic traffic and formulate a resource allocation optimization problem to allocate the eNodeB resources optimally among public safety and commercial users. In addition, we present a resource allocation algorithm to allocate an optimal rate to each UE with a priority given to public safety users. Within the same group of users, a priority is given to real time applications presented by sigmoidal-like utility functions.
- We develop a resource allocation with user discrimination framework to allocate a single carrier resources optimally among different types of users running multiple applications.

- We propose a two-stage rate allocation method for the single carrier RA with user discrimination optimization problem and present its corresponding algorithms. First, the eNodeB and the UE collaborate to allocate an optimal rate to each UE. Each UE then allocates its assigned rate optimally among its applications.

- We develop a multi-stage resource allocation with user discrimination optimization problem to allocate multi-carrier resources optimally among different classes of users. In addition, we prove that the resource allocation optimization problem is convex and therefore the global optimal solution is tractable.

- We present a resource allocation algorithm to solve the multi-stage RA with user discrimination optimization problem and allocate each user an aggregated final rate from its in range carriers. We present simulation results for the performance of the proposed algorithm.

1.4.4 Resource Allocation for Spectrum Sharing between Radar and Communication Systems

In this area, the following contributions are made:

- We present a spectrum sharing scenario between a MIMO radar and a LTE system with multiple base stations and propose a channel-selection algorithm to select the best channel for radar's signal projection that maintains a minimum degradation in the radar performance while causing no interference to the LTE BS. We also present a null-space projection (NSP) algorithm that performs the null space computation.

- We present a multi-stage RA with CA algorithm for the proposed spectrum sharing approach to allocate both of the radar and the LTE-Advanced carriers' resources optimally among users running real-time or delay-tolerant applications. We show through

simulation results that the proposed algorithm is a robust algorithm that converges to the optimal rates for high available resources and scarce resources cases.

1.4.5 Resource Block Scheduling with Carrier Aggregation based on Utility Proportional Fairness

In this area, the following contributions are made:

- We develop a framework for the problem of utility proportional fairness RB scheduling with CA for multi-carrier cellular networks.
- We prove that the proposed resource scheduling policy, that is based on CA, exists and that the optimal solution is tractable. We show through simulation results the performance of the proposed resource scheduling with CA approach and compare it with other resource scheduling policies.

1.4.6 Resource Management for a Multi-Tier Wireless Spectrum Sharing System Leveraging Secure Spectrum Auctions

In this area, the following contributions are made:

- We develop a secure spectrum auction, **MTSSA**, that considers spectrum reusability and the case of heterogeneous frequency bands, e.g. commercial and federal bands. The proposed spectrum auction **MTSSA** has the following properties: it optimizes the usage of spectrum resources by managing intra-WSP and inter-WSP interference, it provides a truthful auction that is achieved when each bidder submits its true evaluation value, it uses a payment method that satisfies essential economic properties, and it provides a secure spectrum auction that prevents frauds of insincere auctioneers

and bid-rigging. Simulation results show that **MTSSA** achieves an efficient spectrum utilization, revenue and bidders' satisfaction.

- We propose an optimal bidding mechanism for determining true bidding values to be used in secure spectrum auctions by BSs, that belong to different WSPs, participating in a spectrum auction.

- We present a resource allocation based on carrier aggregation approach to determine the BS's optimal aggregated rate allocated to each UE, under its coverage area, from both the BS's permanent resources and the BS's winning auctioned spectrum resources. We show through simulation results the performance of the proposed optimal bidding strategy

1.5 Organization of Dissertation

The rest of this proposal is organized as follows. Chapter 2 discusses the users applications utility functions used and their properties. Chapter 3 presents a multi-stage distributed and centralized resource allocation with CA framework. Chapter 4 develops a robust resource allocation with joint carrier aggregation for multi-carrier cellular networks and compare the performance of the proposed algorithms with the multi-stage resource allocation with CA approach. Chapter 5 develops a spectrum sharing architecture between commercial and public safety cellular systems and provides a resource allocation with user discrimination framework for multi-carrier cellular networks. Chapter 6 presents a spectrum sharing approach between radar and communication systems and provides a resource allocation with CA approach for a LTE-Advanced Cellular System Sharing Spectrum with S-band Radar. Chapter 7 presents a utility proportional RB scheduling with CA approach and compares the proposed scheduling policy with other resource scheduling policies. Chapter 8 develops a multi-tier dynamic spectrum sharing system, proposes a secure spectrum auction mechanism

and introduces an optimal bidding strategy for determining true bidding values. Chapter 9 points out some future extensions of this work.

Chapter 2

Preliminaries

2.1 User Applications Utility Functions

The user satisfaction with the provided service can be expressed using utility functions that represent the degree of satisfaction of the user function of the rate allocated by the cellular network [18, 82, 83]. We assume that the applications utility functions $U(r)$ are strictly concave or sigmoidal-like functions.

These applications utility functions have the following properties:

- $U(0) = 0$ and $U(r)$ is an increasing function of r.
- $U(r)$ is twice continuously differentiable in r and bounded above.

We use the normalized sigmoidal-like utility function, same as the one presented in [82], that is

$$U(r) = c\left(\frac{1}{1+e^{-a(r-b)}} - d\right), \tag{2.1}$$

where $c = \frac{1+e^{ab}}{e^{ab}}$ and $d = \frac{1}{1+e^{ab}}$ so it satisfies $U(0) = 0$ and $U(\infty) = 1$. The normalized sigmoidal-like function has an inflection point at $r^{\inf} = b$. In addition, we use the normalized

logarithmic utility function, used in [21], that can be expressed as

$$U(r) = \frac{\log(1 + kr)}{\log(1 + kr_{\max})}, \tag{2.2}$$

where $r_{\max}$ gives 100% utilization and k is the slope of the curve that varies based on the user application. So, it satisfies $U(0) = 0$ and $U(r_{\max}) = 1$.

Figure 2.1 shows an example of sigmoidal and logarithmic utility functions. It shows three normalized sigmoidal-like utility functions that are expressed by equation 2.1 with different parameters $a = 5$, $b = 10$ which is an approximation to a step function at rate $r = 10$ (e.g. VoIP), $a = 3$, $b = 20$ which is an approximation of an adaptive real-time application with inflection point at rate $r = 20$ (e.g. standard definition video streaming), $a = 1$, $b = 30$ which is also an approximation of an adaptive real-time application with inflection point at rate $r = 30$ (e.g. high definition video streaming). In addition Figure 2.1 shows three logarithmic functions that are expressed by equation 2.2 with $r_{\max} = 100$ and different k parameters which are approximations for delay tolerant applications (e.g. FTP). We use $k = \{15, 3, 0.5\}$. It is noticeable that real-time applications require a minimum rate, i.e. the inflection point, after that rate the application QoS is fulfilled to a large extent whereas logarithmic utility functions provide some QoS at low rates which is suitable for the delay-tolerant applications nature.

2.2 Utility Proportional Fairness Resource Allocation

In proportional fairness resource allocation model, each user must be allocated some rate. This is guaranteed as allocating zero rate to any user will set the efficiency of the network

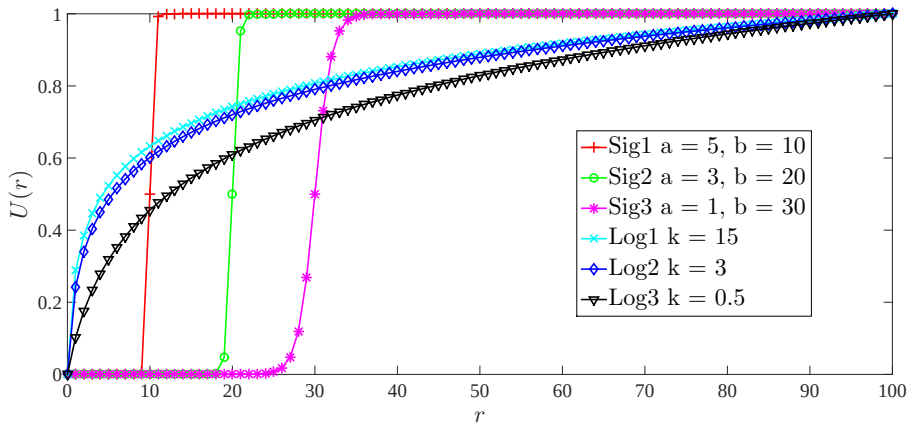


Figure 2.1: Logarithmic and sigmoidal utility functions $U(r)$ representing delay-tolerant and real-time applications, respectively.

to zero. The proportional fairness model is presented in the following equation:

$$r_i = \arg\max_{r_i} \prod_{i=1}^{N} U_i(r_i) \tag{2.3}$$

where $U_i(r_i)$ is the utility function of the i^{th} user allocated resource r_i and N is the number of users. The objective function in equation (2.3) ensures non-zero resource allocation for all users which guarantees minimum QoS for all users. Frank Kelly algorithm [15] can be used to achieve rate allocation with the fairness model. It achieves Pareto optimal resource allocation across the network while using a proportional fairness approach to distribute all network resources where Pareto optimal or Pareto efficient solutions are those solutions that distribute all of the network resources; i.e. also referred to as the Pareto front [84]. Frank Kelly algorithm uses an iterative process to determine the rate that needs to be allocated to each user as well as the price the network should charge each user for the allocated resources. In the next chapters we will be using methods that are based on the Frank Kelly algorithm to solve different proportional fairness resource allocation formulations.

2.3 Utility Proportional Fairness Resource Allocation with Carrier Aggregation

In this dissertation, a utility proportional fairness (UPF) resource allocation optimization framework is proposed to allocate multi-carrier resources optimally among active mobile users from their all in range carriers based on carrier aggregation scenario. Throughout the next chapters we present different resource allocation methods for multi-carrier wireless systems. First, we present a multi-stage RA approach which uses a utility proportional fairness RA optimization problem to allocate each carrier resources separately in a multi-stage basis while taking into consideration the resources allocated to each user from other carriers every time the RA optimization problem is executed. The UPF resource allocation optimization problem that we use in the multi-stage RA with CA approach is given by

$$\begin{aligned} \max_{\mathbf{r}^j} \quad & \prod_{i=1}^{M_j} U_i(r_i^j + c_i^j) \\ \text{subject to} \quad & \sum_{i=1}^{M_j} r_i^j \leq R_j, \\ & r_i^j \geq 0, \qquad i = 1, 2, ..., M_j, \\ & c_i^j = \sum_{n=1, n\neq j}^{K} v_i^n r_i^{n,opt}, \\ & v_i^n = \begin{cases} 1, & \text{the } i^{th} \text{ UE} \in \mathcal{M}_n \\ 0, & \text{the } i^{th} \text{ UE} \notin \mathcal{M}_n \end{cases} \end{aligned} \tag{2.4}$$

where optimization problem (2.4) is carrier j RA optimization problem, $\mathcal{M}_j$ is the set of users located under the coverage area of the j^{th} eNodeB and $M_j = |\mathcal{M}_j|$ is the number of users in the set $\mathcal{M}_j$, $\mathbf{r}^j = \{r_1^j, r_2^j, ..., r_{M_j}^j\}$, R_j is the j^{th} carrier available resources, c_i^j is equivalent to the total rates allocated to the i^{th} user by other carriers in its range, v_i^n is

equivalent to 1 if the i^{th} UE $\in \mathcal{M}_n$ and is equivalent to 0 if the i^{th} UE $\notin \mathcal{M}_n$ and $r_i^{n,opt}$ is the optimal rate allocated to the i^{th} user by the n^{th} carrier.

Then, we present a RA optimization approach based on joint carrier aggregation such that the multi-carrier resource assignment process is performed jointly from all carriers and not in a multi-stage basis. This approach guarantees that each user receives an optimal minimum price for the aggregated resources. The UPF resource allocation optimization problem with joint CA is given by

$$\begin{aligned} \max_{\mathbf{r}} \quad & \prod_{i=1}^{M} U_i(r_{1i} + r_{2i} + ... + r_{Ki}) \\ \text{subject to} \quad & \sum_{i=1}^{M} r_{1i} \leq R_1, \sum_{i=1}^{M} r_{2i} \leq R_2, ... \\ & ... \ , \ \sum_{i=1}^{M} r_{Ki} \leq R_K, \\ & r_{li} \geq 0, \ l = 1, 2, ..., K, \ i = 1, 2, ..., M \end{aligned}$$

where $\mathbf{r} = \{\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_M\}$ and $\mathbf{r}_i = \{r_{1i}, r_{2i}, ..., r_{Ki}\}$, K is the number of carriers eNodeBs with K cells, M is the number of UEs distributed in these cells. The rate allocated by the l^{th} carrier eNodeB to i^{th} UE is given by r_{li} where $l = \{1, 2, ..., K\}$ and $i = \{1, 2, ..., M\}$ and R_l is the total available rate at the l^{th} carrier eNodeB.

In order to consider the case when it is required to treat users differently when assigning the network resources, we introduced a user discrimination feature to the resource allocation framework such that certain group of users (e.g. public safety users in systems that consider spectrum sharing between public safety and commercial users) are given priority when allocating the network resources. Furthermore, we developed resource allocation with CA methods to allocate multi-carrier resources based on user discrimination and used UPF optimization problem to calculate the allocated resources.

Chapter 3

Multi-Stage Resource Allocation with CA in Cellular Networks

In this chapter we present a resource allocation with carrier aggregation optimization problem to allocate the eNodeB's carrier resources optimally among users in its coverage area while taking into consideration the rates allocated to each user from other carriers. We propose two multi-stage resource allocation with carrier aggregation algorithms. The first algorithm is a distributed (decentralized) multi-stage algorithm that allocates users, under the coverage area of a primary carrier and a secondary carrier, resources from both carriers. The second algorithm is a centralized multi-stage algorithm that allocates users resources optimally from all in band carriers and gives each user the ability to select its primary and secondary carriers based on their offered prices in order to provide a minimum price for the allocated resources.

3.1 Multi-Stage Distributed Resource Allocation with Carrier Aggregation

In this section, we focus on finding an optimal solution for the carrier aggregation resource allocation problem for a group of users running two types of applications presented by logarithmic utility functions or sigmoidal-like utility functions. These utility functions are concave and non-concave utility functions, respectively. The RA optimization problem assigns part of the bandwidth from two carriers to each user subscribing for a mobile service taking into consideration that each user is getting a minimum QoS. Our objective is to allocate the resources from two carriers to each user based on its application that is represented by a utility function.

Our contributions in this section are summarized as:

- We present a resource allocation optimization problem with carrier aggregation that gives priority to real-time application users when allocating resources.
- We prove that the optimal rate allocated by the two carriers to each user when using carrier aggregation is equivalent to the optimal rate allocated to the same user by one carrier that has resources equivalent to the total resources in the two carriers. We present a two-stage carrier aggregation rate allocation algorithm to solve the optimization problem and present its corresponding simulation results.

3.1.1 Problem Formulation

We consider two eNodeBs that have the same coverage area and M UEs. One of the eNodeBs is considered to be the primary carrier and the other one is the secondary carrier. Each user is allocated certain bandwidth r_i based on the type of application the UE is

running. Our goal is to determine the optimal bandwidth that needs to be allocated to each user by the two eNodeBs.

We assume the utility functions $U_i(r_i)$ to be a strictly concave or a sigmoidal-like functions. Logarithmic utility functions expressed by equation (2.2) and sigmoidal-like utility functions expressed by equation (2.1) are used to represent delay tolerant and real-time applications, respectively.

3.1.1.1 Single Carrier Optimization Problem

The basic formulation of a single carrier resource allocation problem is given by the following optimization problem:

$$\begin{aligned} \max_{\mathbf{r_{single}}} \quad & \prod_{i=1}^{M} U_i(r_{i,single}) \\ \text{subject to} \quad & \sum_{i=1}^{M} r_{i,single} \leq R, \\ & r_{i,single} \geq 0, \qquad i = 1, 2, ..., M. \end{aligned} \tag{3.1}$$

where R is the maximum achievable rate of the eNodeB, $r_{i,single}$ is the rate for user i and M is the number of UEs.

The optimization problem (3.1) is a convex optimization problem and there exists a unique tractable global optimal solution [21]. The objective function in the optimization problem (3.1) is equivalent to $\max_{\mathbf{r_{single}}} \sum_{i=1}^{M} \log U_i(r_{i,single})$. The solution of this optimization problem is the global optimal solution for the resource allocation problem when resources are allocated by one eNodeB.

For the carrier aggregation resource allocation case, the optimization problem is divided into two stages as shown in section 3.1.2.

3.1.2 Two Carriers Optimization Problem

3.1.2.1 Primary Carrier

The two carriers optimization problem is done in two stages, primary and secondary stages.

The optimization problem for the first carrier can be written as:

$$\begin{aligned} \max_{\mathbf{r_p}} \quad & \prod_{i=1}^{M} U_i(r_{i,p}) \\ \text{subject to} \quad & \sum_{i=1}^{M} r_{i,p} \leq R_p, \\ & r_{i,p} \geq 0, \qquad i = 1, 2, ..., M. \end{aligned} \tag{3.2}$$

where $\mathbf{r_p} = \{r_{1,p}, r_{2,p}, ..., r_{M,p}\}$ and M is the number of UEs in the coverage area of primary user eNodeB and R_p is the maximum achievable rate of the primary carrier. The resource allocation objective function is to maximize the total system utility when allocating resources to each user. Furthermore, it provides proportional fairness among utilities. Users running real-time applications are allocated more resources in this approach.

The optimization problem (3.2) is a convex optimization problem and there exists a unique tractable global optimal solution [21]. The objective function in the optimization problem (3.2) is equivalent to $\max_{\mathbf{r_p}} \sum_{i=1}^{M} \log U_i(r_{i,p})$. The solution of this optimization problem is the first optimal solution that gives each of the M users the optimal rate $r_{i,p}^{\text{opt}}$ only from the primary carrier and not yet the final optimal rate.

3.1.2.2 Secondary Carrier

As mentioned before, we consider a secondary carrier eNodeB located in the same coverage area of the same mobile system. Again, M is the number of mobile users in the coverage area. Once the primary carrier finishes allocating its resources to the M users, the secondary

carrier starts to allocate its resources to the same users while ensuring a minimum user QoS. Therefore, we assume again that the secondary carrier will allocate the resources based on utility proportional fairness.

The optimization problem for the secondary carrier can be written as:

$$\begin{aligned} \max_{\mathbf{r_s}} \quad & \prod_{i=1}^{M} U_i(r_{i,s} + r_{i,p}^{\text{opt}}) \\ \text{subject to} \quad & \sum_{i=1}^{M} r_{i,s} \leq R_s, \\ & r_i \geq 0, \qquad i = 1, 2, ..., M. \end{aligned} \tag{3.3}$$

where $\mathbf{r_s} = \{r_{1,s}, r_{2,s}, ..., r_{M,s}\}$ is the rate for user i, R_s is the maximum achievable rate by the secondary carrier and $r_{i,p}^{\text{opt}}$ is the first optimal rate allocated to user i by the primary carrier and estimated in (3.2). The optimization problem here gives priority to the real-time application users and ensures a minimum rate for each user equals to the first optimal rate $r_{i,p}^{\text{opt}}$ estimated in (3.2).

The optimization problem (3.3) is a convex optimization problem and there exists a unique tractable global optimal solution [21]. The objective function in the optimization problem (3.3) is equivalent to $\max_{\mathbf{r_s}} \sum_{i=1}^{M} \log U_i(r_{i,s} + r_{i,p}^{\text{opt}})$. The global optimal rate for each user is obtained by the sum of the solution given by (3.2) $r_{i,p}^{\text{opt}}$ and the solution given by (3.3) $r_{i,s}^{\text{opt}}$ for user i and is equal $r_{i,agg}^{\text{opt}} = r_{i,s}^{\text{opt}} + r_{i,p}^{\text{opt}}$, such that $r_{i,agg}^{\text{opt}}$ is the global optimal solution that gives each of the M users the optimal rate from both the primary and secondary carriers and considered the final optimal rate.

3.1.2.3 Equivalence

In this section, we show the equivalence of the optimal rate $r_{i,agg}^{\text{opt}}$ given to each user by the primary and secondary eNodeBs to the optimal rate given to the same user by a single

eNodeB, given by the single carrier optimization problem (3.1), when its available resources are equivalent to the resources available in both the primary and secondary eNodeBs in the carrier aggregation case.

Theorem 3.1.1. *The optimal rate $r_{i,agg}^{opt}$ allocated to user i by the two carriers from optimization problem (3.2) and optimization problem (3.3) is equivalent to the optimal rate allocated to the same user by the single carrier optimization problem (3.1) when $R = R_p + R_s$.*

Proof. From the optimization problem (3.2), we have the Lagrangian:

$$L_p(r_{i,p}) = (\sum_{i=1}^{M} \log U_i(r_{i,p})) - P_p(\sum_{i=1}^{M} r_{i,p} - R_p - z_p) \tag{3.4}$$

where $z_p \geq 0$ is the slack variable and P_p is the Lagrange multiplier which is equivalent to the shadow price that corresponds to the total price per bandwidth for the M channels as in [21]. So we have

$$\frac{\partial L_p(r_{i,p})}{\partial r_{i,p}} = \frac{U_i^{'}(r_{i,p})}{U_i(r_{i,p})} - P_p = 0 \tag{3.5}$$

solving for $r_{i,p}$ we obtain $r_{i,p}^{\text{opt}}$.

From optimization problem (3.3), we have the Lagrangian:

$$L_s(r_{i,s}) = (\sum_{i=1}^{M} \log U_i(r_{i,s} + r_{i,p}^{\text{opt}})) - P_s(\sum_{i=1}^{M} r_{i,s} - R_s - z_s) \tag{3.6}$$

where $z_s \geq 0$ is the slack variable and P_s is the Lagrange multiplier. So we have

$$\frac{\partial L_s(r_{i,s})}{\partial r_{i,s}} = \frac{U_i^{'}(r_{i,s} + r_{i,p}^{\text{opt}})}{U_i(r_{i,s} + r_{i,p}^{\text{opt}})} - P_s = 0 \tag{3.7}$$

solving for $r_{i,s}$ we obtain $r_{i,s}^{\text{opt}}$. Replacing $r_{i,s} + r_{i,p}^{\text{opt}}$ in equation (3.6) by a new variable $r_{i,agg}$

such that $r_{i,agg} = r_{i,s} + r_{i,p}^{\text{opt}}$ and rewrite the Lagrangian in terms of $r_{i,agg}$ we obtain

$$\begin{aligned} L_{agg}(r_{i,agg}) =& (\sum_{i=1}^{M} \log U_i(r_{i,agg})) \\ & - P_s(\sum_{i=1}^{M}(r_{i,agg} - r_{i,p}^{\text{opt}}) - R_s - z_s) \end{aligned} \tag{3.8}$$

where $r_{i,agg} \geq r_{i,p}^{\text{opt}}$. From the primary carrier we have $\sum_{i=1}^{M} r_{i,p}^{\text{opt}} = R_p$. So equation (3.8) is equivalent to

$$L(r_{i,agg}) = \sum_{i=1}^{M} \log U_i(r_{i,agg}) - P_s(\sum_{i=1}^{M} r_{i,agg} - R - z_s) \tag{3.9}$$

From problem (3.1) we have

$$\begin{aligned} L_{single}(r_{i,single}) =& (\sum_{i=1}^{M} \log U_i(r_{i,single})) \\ & - P(\sum_{i=1}^{M}(r_{i,single} - R - z)) \end{aligned} \tag{3.10}$$

equivalent to (3.8) for $r_i \geq r_{i,p}^{\text{opt}}$. Therefore, the optimal solution $r_{i,agg}^{\text{opt}}$ given by (3.8) is equivalent to the optimal solution $r_{i,single}^{\text{opt}}$ given by (3.10) when $R = R_p + R_s$. □

3.1.3 Algorithm

We use the same approach used in [21] for utility proportional fairness. Our algorithm is divided into two stages. In first stage (stage1), algorithm 1 and algorithm 2 are the UE and the eNodeB algorithms, respectively. In stage 1, each UE transmits an initial bid $w_{i,p}(1)$ to the primary eNodeB. The eNodeB checks whether the difference between the current received bid and the previous one is less than a threshold δ, if so it exits. Otherwise, if the difference is greater than δ, eNodeB calculates the shadow price $P_p(n) = \frac{\sum_{i=1}^{M} w_{i,p}(n)}{R_p}$. The shadow price does not depend on the number of users competing for some resources, it only

depends on the users bids and the eNodeB's available resources. The estimated $P_p(n)$ is then sent to the UE where it is used to calculate the rate $r_{i,p}(n)$ which is the solution of the optimization problem $r_{i,p}(n) = \arg\max_{r_{i,p}}(\log U_i(r_{i,p}) - P_p(n)r_{i,p})$. A new bid $w_{i,p}(n)$ is calculated using $r_{i,p}(n)$ where $w_{i,p}(n) = P_p(n)r_{i,p}(n)$. All UEs send their new bids $w_{i,p}(n)$ to the primary eNodeB. Stage 1 of the Algorithm is finalized by the primary eNodeB. Each UE then calculates its allocated rate $r_{i,p}^{\text{opt}} = \frac{w_{i,p}(n)}{P_p(n)}$.

Algorithm 1 UE Stage 1 of Carrier Aggregation

Send initial bid $w_{i,p}(1)$ to eNodeB
loop
  Receive shadow price $P_p(n)$ from eNodeB
  if STOP from eNodeB **then**
    Calculate allocated rate $r_{i,p}^{\text{opt}} = \frac{w_{i,p}(n)}{P_p(n)}$
  else
    Solve $r_{i,p}(n) = \arg\max_{r_{i,p}}\Big(\log U_i(r_{i,p}) - P_p(n)r_{i,p}\Big)$
    Send new bid $w_{i,p}(n) = P_p(n)r_{i,p}(n)$ to eNodeB
  end if
end loop

Algorithm 2 eNodeB Stage 1 of Carrier Aggregation

loop
  Receive bids $w_{i,p}(n)$ from UEs {Let $w_{i,p}(0) = 0 \ \forall i$}
  if $|w_{i,p}(n) - w_{i,p}(n-1)| < \delta \ \forall i$ **then**
    STOP and allocate rates (i.e $r_{i,p}^{\text{opt}}$ to user i)
  else
    Calculate $P_p(n) = \frac{\sum_{i=1}^{M} w_{i,p}(n)}{R_p}$
    Send new shadow price $P_p(n)$ to all UEs
  end if
end loop

After allocating rates from primary carrier, stage 2 starts performing. Each UE transmits an initial bid $w_{i,s}(1)$ to the secondary eNodeB. The secondary eNodeB checks whether the difference between the current received bid and the previous one is less than a threshold δ, if so it exits. Otherwise, if the difference is greater than δ, the secondary eNodeB calculates the shadow price $P_s(n) = \frac{\sum_{i=1}^{M} w_{i,s}(n)}{R_s}$. The estimated $P_s(n)$ is then sent to the UE where it is used to calculate the rate $r_{i,s}(n)$ which is the solution of the optimization problem

$r_{i,s}(n) = \arg\max_{r_{i,s}}(\log U_i(r_{i,s} + r_{i,p}^{\text{opt}}) - P_s(n)r_{i,s})$. A new bid $w_{i,s}(n)$ is calculated using $r_{i,s}(n)$ where $w_{i,s}(n) = P_s(n)r_{i,s}(n)$. All UEs send their new bids $w_{i,s}(n)$ to the secondary eNodeB. Stage 2 of the Algorithm is finalized by the secondary eNodeB. Each UE then calculates its allocated rate $r_{i,s}^{\text{opt}} = \frac{w_{i,s}(n)}{P_s(n)}$.

Algorithm 3 UE Stage 2 of Carrier Aggregation

Send initial bid $w_{i,s}(1)$ to eNodeB
loop
  Receive shadow price $P_s(n)$ from eNodeB
  if STOP from eNodeB **then**
    Calculate allocated rate $r_{i,s}^{\text{opt}} = \frac{w_{i,s}(n)}{P_s(n)}$
  else
    Solve $r_{i,s}(n) = \arg\max_{r_{i,s}}\Big(\log U_i(r_{i,s} + r_{i,p}^{\text{opt}}) - P_s(n)r_i\Big)$
    Send new bid $w_{i,s}(n) = P_s(n)r_{i,s}(n)$ to eNodeB
  end if
end loop

Algorithm 4 eNodeB Stage 2 of Carrier Aggregation

loop
  Receive bids $w_{i,s}(n)$ from UEs {Let $w_{i,s}(0) = 0 \ \forall i$}
  if $|w_{i,s}(n) - w_{i,s}(n-1)| < \delta \ \forall i$ **then**
    STOP and allocate rates (i.e $r_{i,s}^{\text{opt}}$ to user i)
  else
    Calculate $P_s(n) = \frac{\sum_{i=1}^{M} w_{i,s}(n)}{R_s}$
    Send new shadow price $P_s(n)$ to all UEs
  end if
end loop

3.1.4 Simulation Results

As shown in Figure 3.1, we consider two eNodeBs with the same coverage area and six UEs. One of the eNodeBs is the primary carrier and the other one is the secondary carrier with a coverage area that is almost the same for the two carriers. In Figure 3.2, we show three normalized sigmoidal-like utility functions expressed in equation (2.1), each one is corresponding to one user. We use different parameters a and b for each one where $a = 5$,

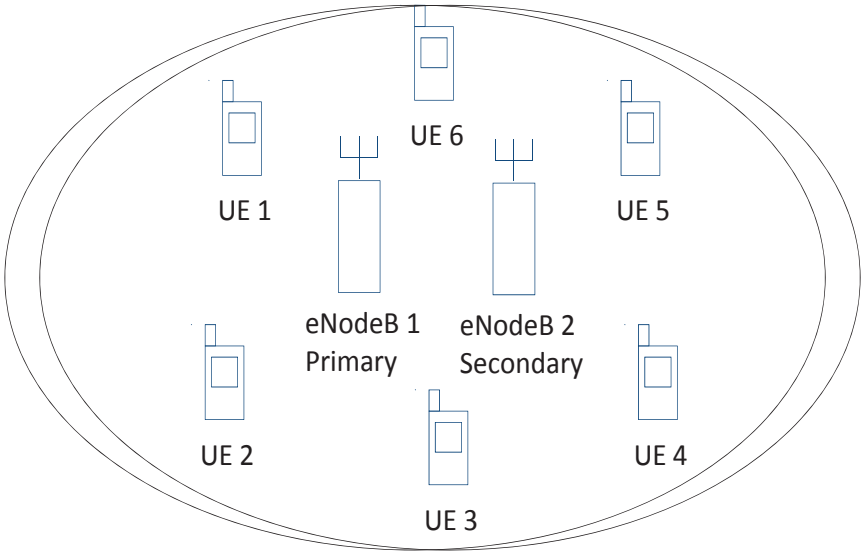


Figure 3.1: System Model.

$b = 10$ for the first user, $a = 3$, $b = 20$ for the second user and $a = 1$, $b = 30$ for the third user. Each sigmoidal-like function is an approximation to a step function at rate b. We also show three logarithmic functions expressed in equation (2.2), which represent delay tolerant applications, with $k = \{15, 3, 0.5\}$ for user four, five and six, respectively. We set $r_{max} = 120$.

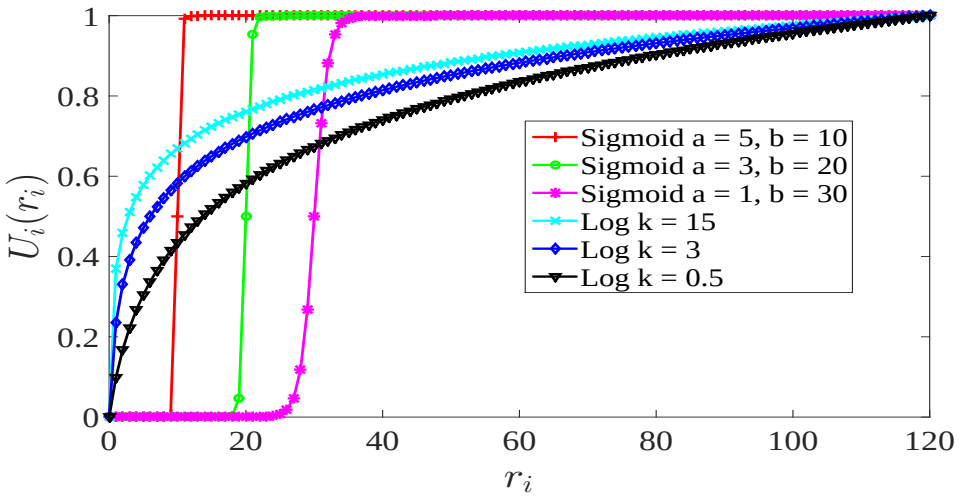


Figure 3.2: The users utility functions $U_i(r_i)$.

3.1.4.1 Convergence Dynamics for $R_p = 70$ in stage 1 of the Algorithm

We applied algorithm 1 and 2 of stage 1 in C++ to the sigmoidal-like and logarithmic utility functions shown in Figure 3.2. We set $R_p = 70$ and $\delta = 10^{-2}$. In Figure 3.3, we show the simulation results for the rate of different users and the number of iterations. As mentioned before the sigmoidal-like utility functions are given priority over the logarithmic utility functions for rate allocation and this explain the results we got in Figure 3.3 where the steady state rate of each sigmoidal-like function exceeds the inflection point b_i. In Figure 3.4, we show the bids of the six users with the number of iterations. As expected, the higher the user bids the higher the allocated rate is for that user. The algorithm allows users with real-time applications, presented in sigmoidal-like utility functions, to bid higher than the other users until each one of them reaches its inflection point then the elastic traffic starts dividing the remaining resources among them based on their parameters. The first optimal rates for the six users $r_{i,p}^{\text{opt}} = \{10.64, 20.88, 31.41, 1.54, 2.19, 3.26\}$ are obtained at the end when running Algorithm 1 and 2 of stage 1. The first optimal rates are used in the next simulation that is performed for the secondary eNodeB and the same six UEs.

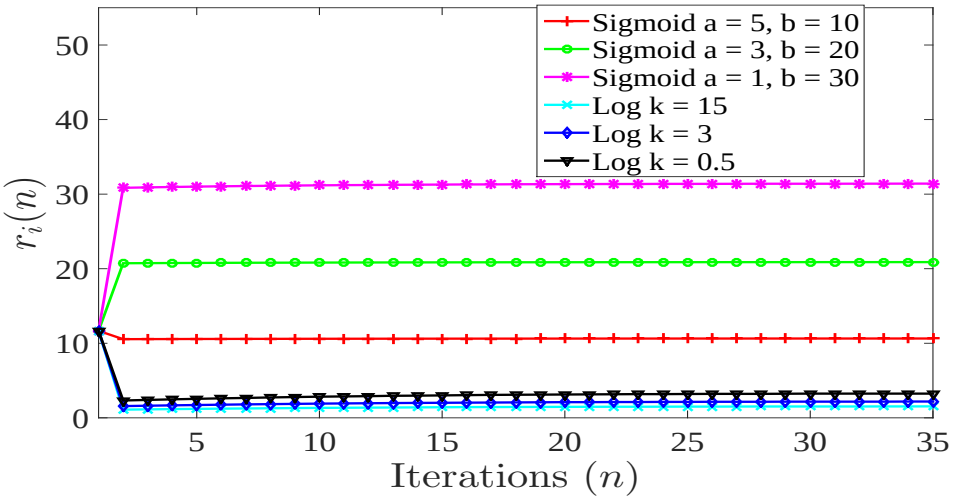


Figure 3.3: The rates $r_{i,p}(n)$ with the number of iterations n for different users and $R_p = 70$.

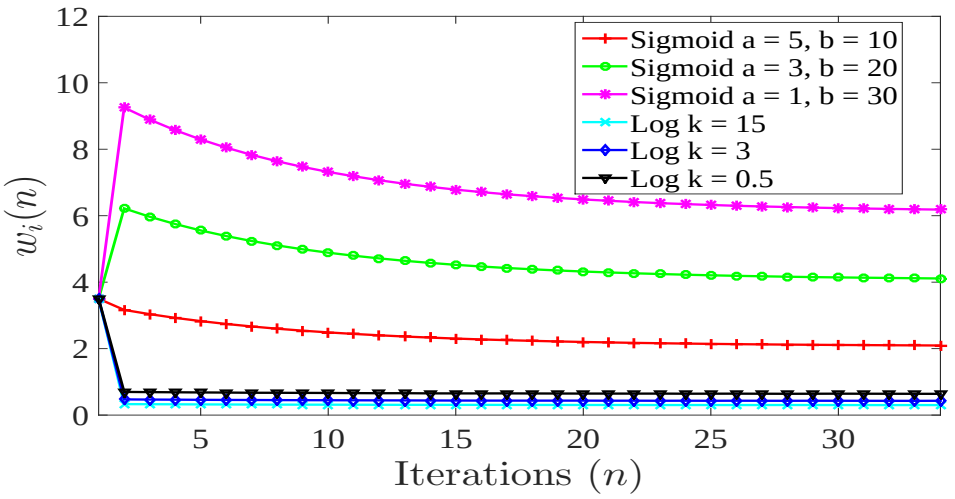


Figure 3.4: The bids convergence $w_{i,p}(n)$ with the number of iterations n for different users and $R_p = 70$.

3.1.4.2 Convergence Dynamics for the carrier aggregation $R_s = 50$ in stage 2 of the Algorithm

We applied algorithm 3 and 4 of stage 2 in C++ to the sigmoidal-like and logarithmic utility functions. We set $R_s = 50$ and $\delta = 10^{-2}$.

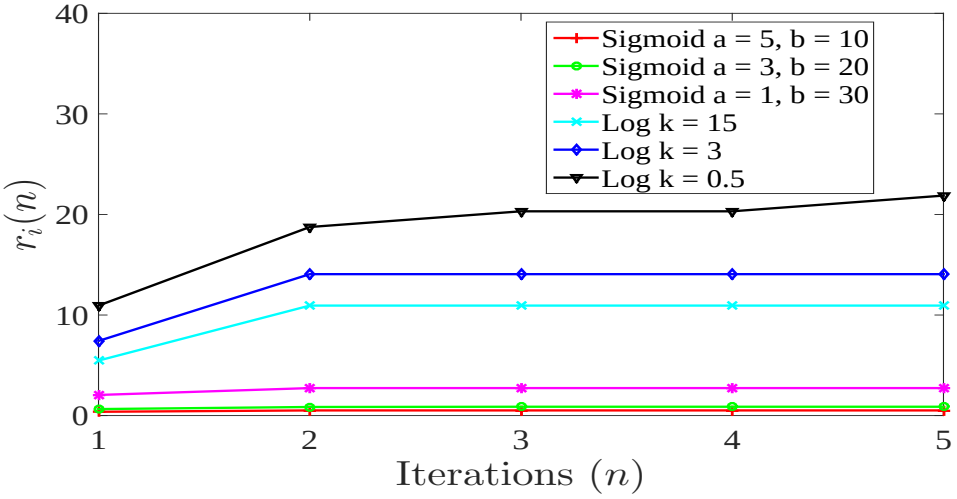


Figure 3.5: The rates $r_{i,s}(n)$ with the number of iterations n for different users and $R_s = 50$.

In Figure 3.5, we show the simulation results for the rate of the six users and the number of iterations. Again, the sigmoidal-like utility functions are given priority over the logarithmic

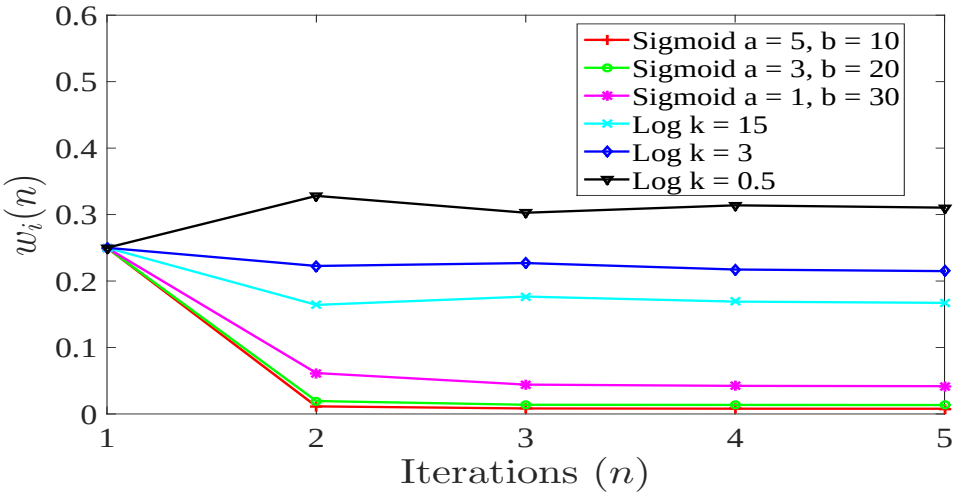


Figure 3.6: The bids convergence $w_{i,s}(n)$ with the number of iterations n for different users and $R_s = 50$.

utility functions for rate allocation, but since each sigmoidal-like function reached its steady state in stage 1 of the Algorithm most of R_s is distributed among the logarithmic functions. In stage 2 the optimal rates for the real time applications users $r_{i,s}^{\text{opt}}$ slightly increased from the first optimal rate $r_{i,p}^{\text{opt}}$ as they were given priory to reach their optimal rates in stage 1 by the primary eNodeB, whereas the elastic traffic divided the remaining resources among them and showed a high increase in their second optimal rate $r_{i,s}^{\text{opt}}$ from their first optimal rates obtained in stage 1. The optimal rates obtained at the end of stage 2 are $r_{i,s}^{\text{opt}} = \{0.51, 0.88, 2.735, 10.94, 14.06, 21.87\}$.

In Figure 3.6, we show the bids of the six users with the number of iterations. As expected the higher the user bids the higher the allocated rate is for that user. The algorithm allows users with real-time applications, presented in sigmoidal-like utility functions, to bid higher than the other users until each one of them reaches its inflection point, but since these users reached their steady states in stage 1 of the Algorithm the elastic traffic users bid higher than the inelastic traffic users and share the secondary carrier's resources among them based on their parameters.

The final optimal rate for each user $r_{i,agg}^{\text{opt}}$ is the sum of $r_{i,p}^{\text{opt}}$ obtained at the end of stage 1

of the Algorithm and $r_{i,s}^{\text{opt}}$ obtained at the end of stage 2 of the Algorithm. As expected the final optimal rates for the six users sum up to 120 which is the total rate of the primary and secondary maximum rates.

3.1.4.3 Equivalence of Optimal rate $r_{i,single}^{\text{opt}}$ with $r_{i,p}^{\text{opt}} + r_{i,s}^{\text{opt}}$ when $R = R_p + R_s$

Figure 3.7 shows the optimal rates obtained when we run Algorithm 1 and 2 of stage 1 for the same six users sharing resources of a single carrier with $R = 120$. We made $R_p = R$, $r_{i,p}(n) = r_{i,single}(n)$, $w_{i,p}(n) = w_i(n)$ and $P_p(n) = P(n)$ when running Algorithm 1 and 2 of stage 1 for the single carrier case. The optimal rates obtained in this case are $r_{i,single}^{\text{opt}} = \{11.16, 21.74, 34.22, 13.12, 16.87, 22.50\}$, they are almost similar to the final optimal rates $r_{i,agg}^{\text{opt}}$ in the carrier aggregation case when the same users share the resources of two carriers one being the primary and the other being the secondary with a total R_p and R_s of 120.

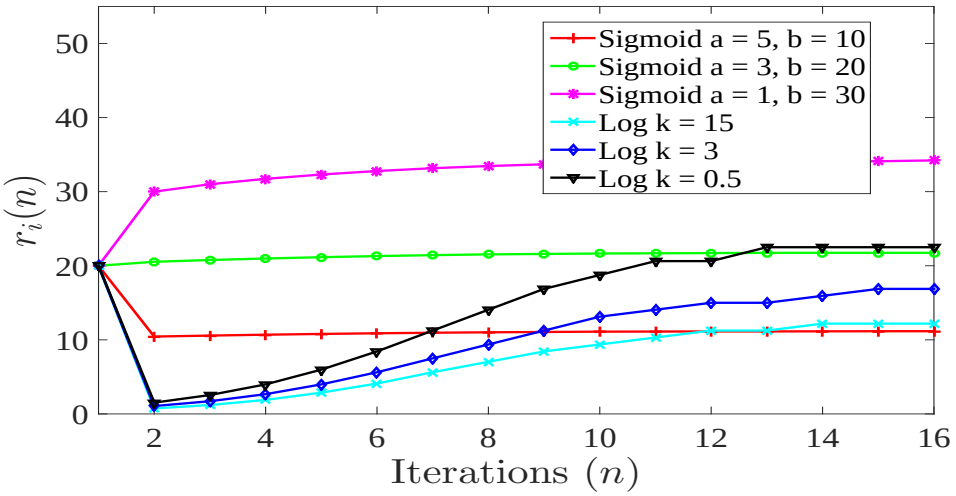


Figure 3.7: The rates $r_{i,single}(n)$ with the number of iterations n for different users for the single carrier case with $R = 120$.

3.1.4.4 Impact of Dynamic User Activities in the Convergence of the RA algorithms

We investigate the sensitivity of the proposed resource allocation algorithms to users arrival, departures and user utility changes due to application activity changes. Let $\mathcal{M}$ represents the set of original users that the eNodeB has started calculating their optimal rates by running the RA algorithms before users arrival or departure occur, and let w_i^{opt} represents the optimal bidding values that correspond to the optimal rates calculated by the RA algorithms. Let $\mathcal{M}^{'}$ represents the new set of users after users departure or arrival. We compare the number of iterations it takes the algorithms to converge when changing users activities for the two cases described below:

- Case 1: The eNodeB uses the optimal bidding values w_i^{opt}, determined after the convergence of the RA algorithms for the original users, as initial bidding values (i.e. $w_i^{'} = w_i^{\text{opt}}$) for common bidders in $\mathcal{M}^{'}$ and $\mathcal{M}$ when it starts running the RA algorithms to determine the optimal rates for users in $\mathcal{M}^{'}$ after the changes in users' activities.

- Case 2: Cold start, the eNodeB and all active UEs start running the RA algorithms without taking into consideration using the optimal bids determined by the algorithms, before the changes in users activities, for common users in $\mathcal{M}^{'}$ and $\mathcal{M}$.

We considered the same six UEs ($|\mathcal{M}| = 6$) with the same simulation setup described above. We ran the resource allocation with CA algorithms for the six users and observed the number of iterations that takes the algorithms to converge to the optimal rates allocated from the eNodeB's primary resources as well as the eNodeB's secondary resources. On the other hand, we considered the arrival of additional two users where the number of users subscribing for mobile services changed from 6 users to 8 users ($|\mathcal{M}^{'}| = 8$) with user 7 running real-time application represented by sigmoidal utility function with $a = 5$, $b = 10$,

and user 8 is running delay tolerant application represented by logarithmic utility function with parameter $k = 15$. For the two cases described above, we compared the number of iterations that took the algorithms to converge when the number of users changed to 8. The observed number of iterations in case 1 (when considering the optimal bidding values of the first 6 users) and case 2 (cold start) are almost the same. In addition, we consider users departure where after the convergence of the algorithms for the original 6 users, two users (user 5 and user 6) departed and are no longer active for the eNodeB. Again, for the two cases described above, we compared the number of iterations it took the algorithms to converge. We noticed that the number of iterations are almost the same for the two cases.

For common users in $\mathcal{M}$ and $\mathcal{M}^{'}$, when using the users bidding values w_i^{opt}, that correspond to the optimal rates r_i^{opt} calculated by the eNodeB and UE algorithms when considering all users in $\mathcal{M}$, as initial bidding values $w_i^{'}$ for determining new optimal allocated rates by the RA algorithms, the number of iterations required for the convergence of the algorithms during the process of calculating the optimal rates for the updated users set $\mathcal{M}^{'}$ are not necessarly less than the number of iterations required for the convergence of the algorithms when common UEs in $\mathcal{M}$ and $\mathcal{M}^{'}$ send new bidding values $w_i^{'} \neq w_i^{\text{opt}}$. This is because the optimal rates calculated by the algorithms before users departure or arrival are no longer optimal and new optimal rates will be calculated by the algorithms for users in $\mathcal{M}^{'}$. Therefore, in situations of dynamic users activities, it does not matter whether the algorithms use the latest calculated bidding values or new ones (cold start) for users who did not change their running applications and are still active; i.e. the algorithms will not necessarily converge faster when using the latest calculated bidding values before the changes in users activities.

3.2 Multi-Stage Centralized Resource Allocation with Carrier Aggregation Based on a Price Selective Algorithm

In this section, we formulate the RA with CA problem into a convex optimization framework. We use logarithmic utility functions to represent delay-tolerant applications and sigmoidal-like utility functions to represent real-time applications running on the UEs subscribing for a mobile service. The primary and secondary carriers optimization problems assign part of the bandwidth from the multiple carriers to each user. A minimum QoS is guaranteed for each user by using a proportional fairness approach. Our objective is to allocate multiple carriers resources optimally among users in their coverage area while giving the user the ability to select the carrier with the lowest price to be its primary carrier and the others to be its secondary carriers. This mechanism allows users to improve their allocated rates by using the CA feature while maintaining the lowest possible price for their allocated aggregated rates. Additionally, our centralized algorithm is performed mostly in the eNodeBs which reduces the transmission overhead created by the distributed algorithm introduced in [24].

Our contributions in this section are summarized as:

- We present a resource allocation optimization problem with carrier aggregation that solves for logarithmic and sigmoidal-like utility functions.
- We propose a price selective centralized RA with CA algorithm to allocate multiple carriers resources optimally among users.
- We show that our algorithm is a robust one that converges to the optimal rates whether the eNodeBs available resources are abundant or scarce. We present simulation results

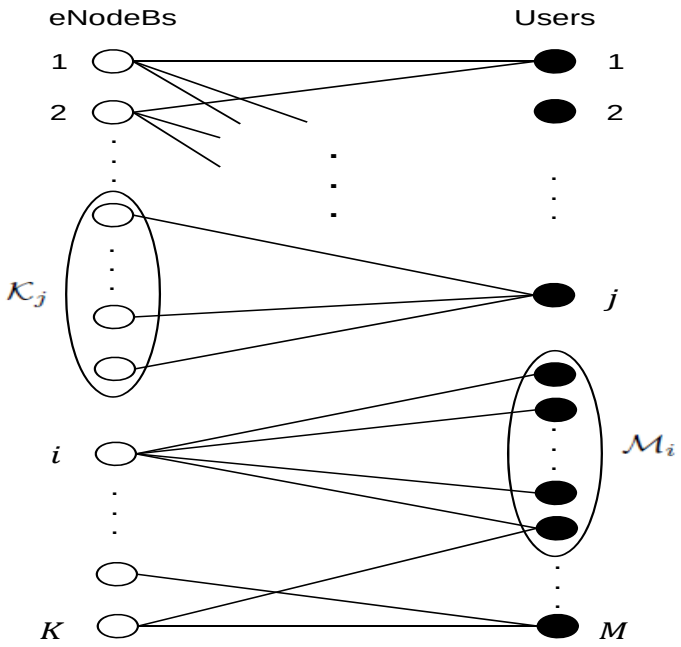


Figure 3.8: System model for a LTE mobile system with M users and K carriers eNodeBs. $\mathcal{M}_i$ represents the set of users located under the coverage area of the i^{th} eNodeB and $\mathcal{K}_j$ represents the set of all in range eNodeBs for the j^{th} user.

for the performance of our resource allocation algorithm.

3.2.1 Problem Formulation

We consider a LTE mobile system with M users and K carriers eNodeBs, one eNodeB in each cell, as illustrated in Figure 3.8. The users located under the coverage area of the i^{th} eNodeB are forming a set of users $\mathcal{M}_i$ where $\mathcal{M}_i \in \{\mathcal{M}_1, \mathcal{M}_2, ..., \mathcal{M}_K\}$ and $M_i = |\mathcal{M}_i|$ is the number of users in the users set $\mathcal{M}_i$ under the coverage area of the i^{th} eNodeB. Each joint user j is located under the coverage area of a set of eNodeBs, as shown in Figure 3.8, that is given by $\mathcal{K}_j$ where $\mathcal{K}_j \in \{\mathcal{K}_1, \mathcal{K}_2, ..., \mathcal{K}_M\}$ and $K_j = |\mathcal{K}_j|$ is the number of eNodeBs in the set $\mathcal{K}_j$ of all in range eNodeBs for user j.

Each eNodeB calculates its offered price per unit bandwidth (assuming it is the primary carrier for all users under its coverage area) and provides each user under its coverage area

with its offered price. Each joint user selects the carrier with the least offered price to be its primary carrier and the rest of all in range carriers to be its secondary carriers. The eNodeB with the least offered price first allocates its resources to all users under its coverage area based on the applications running on their UEs. The remaining eNodeBs then start allocating their resources in the order of their offered prices to all users under their coverage area based on the users applications and the rates that are allocated to the joint users from other eNodeBs (with lower offered prices).

We express the user satisfaction with its provided service using utility functions [18, 82, 83]. We assume that the j^{th} user' application utility function $U_j(r_j)$ is strictly concave or sigmoidal-like function where r_j is the rate allocated to the j^{th} user. Delay tolerant applications are represented by logarithmic utility functions expressed by equation (2.2) whereas real-time applications are represented by sigmoidal-like utility functions expressed by equation (2.1).

3.2.2 Multiple Carriers Optimization Problem

In this section we formulate the RA problem for allocating the primary and secondary carriers resources optimally among users under their coverage areas. Each carrier first calculates its offered price per unit bandwidth assuming that it is the primary carrier for all UEs under its coverage area. Then, each carrier starts allocating its available resources optimally among all users in its coverage area in the order of the carrier's offered price, such that the carrier with a lower offered price performs the RA prior to the one with a higher offered price.

3.2.2.1 The Price Selection Problem and enodeB Sorting

As mentioned earlier, each carrier calculates its offered price assuming it is the primary carrier for all users under its coverage area. The carrier's offered price is obtained from the following RA optimization problem:

$$\begin{aligned} \max_{\mathbf{r_i}} \quad & \prod_{j=1}^{M_i} U_j(r_{i,j}) \\ \text{subject to} \quad & \sum_{j=1}^{M_i} r_{i,j} \leq R_i, \\ & r_{i,j} \geq 0, \qquad j = 1, 2, ..., M_i. \end{aligned} \tag{3.11}$$

where $\mathbf{r_i} = \{r_{i,1}, r_{i,2}, ..., r_{i,M_i}\}$, M_i is the number of UEs under the coverage area of the i^{th} eNodeB and R_i is the maximum achievable rate of the i^{th} eNodeB. The resource allocation objective function is to maximize the total system utility when allocating the eNodeB resources. Furthermore, it provides proportional fairness among utilities. Therefore, no user is allocated zero resources and a minimum QoS is provided to each user. Real-time applications are given priory when allocating the eNodeB resources using this approach. Optimization problem (3.11) is a convex optimization problem and there exists a unique tractable global optimal solution [21]. The objective function in optimization problem (3.11) is equivalent to $\max_{\mathbf{r_i}} \sum_{j=1}^{M_i} \log U_j(r_{i,j})$.

From optimization problem (3.11), we have the Lagrangian:

$$\begin{aligned} L_i(r_{i,j}) = & (\sum_{j=1}^{M_i} \log U_j(r_{i,j})) \\ & - p_i^{\text{offered}} (\sum_{j=1}^{M_i} r_{i,j} - R_i - z_i) \end{aligned} \tag{3.12}$$

where $z_i \geq 0$ is the slack variable and p_i^{offered} is the Lagrange multiplier which is equivalent to

the shadow price that corresponds to the i^{th} carrier price per unit bandwidth for the M_i channels as in [21]. The set of all carriers in the LTE mobile system is given by $\mathcal{K} = \{1, 2, ..., K\}$ and their corresponding offered prices are given by $\mathcal{P}^{\text{offered}} = \{p_1^{\text{offered}}, p_2^{\text{offered}}, ..., p_K^{\text{offered}}\}$. The j^{th} user set of all in range carriers $\mathcal{K}_j$ (i.e. $\mathcal{K}_j = \{1, 2, ..., K_j\}$) corresponding offered prices are given by $\mathcal{P}^j = \{p_1^j, p_2^j, ..., p_{K_j}^j\}$.

All in range carriers $\mathcal{K}_j$ of the j^{th} user are arranged based on their offered prices as follows:

$$
\begin{aligned}
l_1^j &= \arg\min_{\mathcal{K}_j} \{p_1^j, p_2^j, ..., p_{K_j}^j\} \\
l_2^j &= \arg\min_{\mathcal{K}_j - \{l_1^j\}} \{p_1^j, p_2^j, ..., p_{K_j}^j\} \\
&\vdots \\
l_{K_j}^j &= \arg\min_{\mathcal{K}_j - \{l_1^j, ..., l_{K_j-1}^j\}} \{p_1^j, p_2^j, ..., p_{K_j}^j\}
\end{aligned}
$$

where l_1^j is the carrier with the lowest offered price and $l_{K_j}^j$ is the carrier with the highest offered price within the j^{th} user set $\mathcal{K}_j$ of all in range carriers and $\mathcal{P}^j = \{p_1^j, p_2^j, ..., p_{K_j}^j\}$ is the set of the offered prices of all in range carriers for the j^{th} user. The j^{th} user sends an assignment of 1 to the i^{th} eNodeB that is corresponding to eNodeB l_1^j (i.e. the eNodeB with the least offered price among the j^{th} user's all in range carriers). On the other hand, the j^{th} user sends an assignment of 0 to each of the remaining eNodeBs in its range. Once the i^{th} eNodeB receives an assignment of 1 from each UE in its coverage area it starts allocating its resources to the M_i UEs in $\mathcal{M}_i$ such that the j^{th} UE is allocated an optimal rate $r_i^{j,opt}$ from the i^{th} eNodeB. Once the j^{th} UE is allocated rate from its primary carrier l_1^j, it then sends an assignment of 1 to the i^{th} eNodeB that is corresponding to eNodeB l_2^j and sends an assignment of 0 to each of the remaining eNodeBs in its range. The process continues until the j^{th} UE sends an assignment of 1 to the i^{th} eNodeB that is corresponding to eNodeB $l_{K_j}^j$

and receives its allocated rate from that eNodeB. The j^{th} UE then calculates its aggregated final optimal rate r_j^{agg}.

3.2.2.2 RA Optimization Problem

Once the carriers offered prices are calculated as discussed in 3.2.2.1, each user j selects eNodeB l_1^j to be its primary carrier and the remaining carriers in its range to be its secondary carriers. The eNodeB with the least offered price is the first one to start allocating its resources among all users in its coverage area. Each of the remaining eNodeBs then starts allocating its available resources after all the users in its coverage area are allocated rates from carriers in their range with lower offered prices. Eventually, each user j is allocated rates from all of the K_j carriers in its range. As discussed before, the i^{th} carrier eNodeB starts allocating its resources among all users in its coverage area once it receives an assignment of 1 from each of the M_i users in $\mathcal{M}_i$. The rate allocated to the j^{th} user from its i^{th} carrier is given by $r_i^{j,opt}$.

The RA optimization problem for the i^{th} carrier eNodeB in $\mathcal{K}$, such that the i^{th} eNodeB received an assignment of 1 from each of the users under its coverage area, can be written as:

$$\begin{aligned}
\max_{\mathbf{r_i}} \quad & \prod_{j=1}^{M_i} U_j(r_i^j + c_i^j) \\
\text{subject to} \quad & \sum_{j=1}^{M_i} r_i^j \leq R_i, \\
& r_i^j \geq 0, \qquad j = 1, 2, ..., M_i, \\
& c_i^j = \sum_{n=1, n\neq i}^{K} v_n^j r_n^{j,opt}, \\
& v_n^j = \begin{cases} 1, & \text{the } j^{th} \text{ UE} \in \mathcal{M}_n, \\ 0, & \text{the } j^{th} \text{ UE} \notin \mathcal{M}_n, \end{cases}
\end{aligned} \tag{3.13}$$

where $\mathbf{r_i} = \{r_i^1, r_i^2, ..., r_i^{M_i}\}$, R_i is the i^{th} eNodeB available resources, c_i^j is equivalent to the total rates allocated to the j^{th} user by the carriers in its range with lower offered prices than the i^{th} carrier offered price, v_n^j is equivalent to 1 if the j^{th} UE $\in \mathcal{M}_n$ and is equivalent to 0 if the j^{th} UE $\notin \mathcal{M}_n$ and $r_n^{j,opt}$ is the optimal rate allocated to the j^{th} user by the n^{th} eNodeB (i.e. the n^{th} carrier $\in \mathcal{K}$). Once the j^{th} user is allocated rate from all the carriers in its range, it then calculates its aggregated final optimal rate $r_j^{agg} = \sum_{i=1}^{K} v_i^j r_i^{j,opt}$.

Optimization problem (3.13) gives priority to the real-time application users and ensures that the minimum rate allocated to each user is c_i^j. Optimization problem (3.13) is a convex optimization problem and there exists a unique tractable global optimal solution [21]. The objective function in optimization problem (3.13) is equivalent to $\max_{\mathbf{r_i}} \sum_{j=1}^{M_i} \log U_j(r_i^j + c_i^j)$.

From optimization problem (3.13), we have the Lagrangian:

$$\begin{aligned} L_i(r_i^j) =& (\sum_{j=1}^{M_i} \log U_j(r_i^j + c_i^j)) \\ & - p_i(\sum_{j=1}^{M_i} r_i^j - R_i - z^i) \end{aligned} \tag{3.14}$$

where $z^i \geq 0$ is the slack variable and p_i is the Lagrange multiplier which is equivalent to the shadow price that corresponds to the i^{th} carrier price per unit bandwidth for the M_i channels as in [21].

3.2.3 Algorithm

In this section, we present our price selective centralized RA with CA algorithm. Each UE is allocated optimal rates from its all in range carriers and the final optimal rate allocated to each UE is the aggregated rate. The algorithm starts when each UE transmits its application parameters to all in range eNodeBs. Each eNodeB assigns initial values $w_{i,j}(0)$ to the users applications. Each eNodeB performs an internal iterative algorithm to calculate its offered

price per unit bandwidth. In each iteration, the eNodeB checks the difference between the current value $w_{i,j}(n)$ and the previous one $w_{i,j}(n-1)$, if the difference is greater than a threshold δ, the shadow price $p_i^{\text{offered}}(n) = \frac{\sum_{j=1}^{M_i} w_{i,j}(n)}{R_i}$ is calculated by the eNodeB. Each eNodeB uses $p_i^{\text{offered}}(n)$ to calculate the rate $r_{i,j}(n)$ that is the solution of the optimization problem $r_{i,j}(n) = \arg\max_{r_{i,j}}(\log U_j(r_{i,j}) - p_i^{\text{offered}}(n) r_{i,j})$. The calculated rate is then used to calculate a new value $w_{i,j}(n)$ where $w_{i,j}(n) = p_i^{\text{offered}}(n) r_{i,j}(n)$. Each eNodeB checks the fluctuation condition as in [22] and calculates a new value $w_{i,j}(n)$. Once the difference between the current $w_{i,j}(n)$ and the previous one is less than δ for all UEs, the i^{th} eNodeB sends its offered price p_i^{offered} to all UEs in its coverage area.

Once the j^{th} UE receives the offered prices p_i^{offered} from all in range carriers, it sends an assignment of 1 to the i^{th} eNodeB with the lowest offered price that is corresponding to eNodeB l_1^j and an assignment of 0 to the remaining eNodeBs in its range. The j^{th} UE then receives its allocated rate $r_i^{j,opt}$ and shadow price p_i from that eNodeB. It then updates the c_i^j value and sends it to the i^{th} eNodeB that is corresponding to eNodeB l_2^j, it also sends an assignment of 1 to that eNodeB and an assignment of 0 to the remaining eNodeBs in its range. The process continues until the j^{th} UE receives its allocated rate $r_i^{j,opt}$ and shadow price p_{K_j}, it then calculates its aggregated final optimal rate r_j^{agg}.

On the other hand, Once the i^{th} eNodeB receives assignments of 1 from all UEs in its coverage area it calculates the optimal rate $r_i^{j,opt}$ and shadow price p_i and sends them to each UE in its coverage area. The process continues until the eNodeB with the highest offered price receives assignment of 1 from all UEs in its coverage area, it then sends each of these UEs its allocated optimal rate $r_i^{j,opt}$ and shadow price p_i.

Algorithm 5 The j^{th} UE Algorithm

Let $c_i^j = 0 \ \forall i \in \{1, 2, ..., K\}$
Send the UE application utility parameters k_j, a_j and b_j to all in range eNodeBs
Receive offered prices that are equivalent to $\mathcal{P}^j = \{p_1^j, p_2^j, ..., p_{K_j}^j\}$ from all in range carriers eNodeBs
loop
  for $m \leftarrow 1$ to K_j **do**
    $l_m^j = \arg \min_{\mathcal{K}_j - \{l_1^j, ..., l_{m-1}^j\}} \{p_1^j, p_2^j, ..., p_{K_j}^j\}$ is carrier l_m^j for the j^{th} UE
  end for
end loop
loop
  for $m \leftarrow 1$ to $K_j - 1$ **do**
    Send Flag assignment of 1 to the i^{th} eNodeB and an assignment of 0 to the remaining carriers in $\mathcal{K}_j$ {eNodeB i = eNodeB l_m^j}
    Send c_i^j to the i^{th} eNodeB {eNodeB i = eNodeB l_m^j}
    Receive the optimal rate $r_i^{j,opt}$ from the i^{th} eNodeB {eNodeB i = eNodeB l_m^j}
    Receive shadow price p_i from the i^{th} eNodeB {eNodeB i = eNodeB l_m^j}
    Send the optimal rate $r_i^{j,opt}$ to the i^{th} eNodeB {the i^{th} eNodeB corresponds to eNodeB l_{m+1}^j}
    Calculate new $c_i^j = \sum_{n=1, n \neq i}^{K} v_n^j r_n^{j,opt}$ for the i^{th} eNodeB that corresponds to eNodeB l_{m+1}^j
  end for
end loop
Send c_i^j to the i^{th} eNodeB {the i^{th} carrier corresponds to carrier $l_{K_j}^j$}
Receive the optimal rate $r_i^{j,opt}$ from the i^{th} eNodeB {the i^{th} carrier corresponds to carrier $l_{K_j}^j$}
Receive shadow price p_i from the i^{th} eNodeB {the i^{th} carrier corresponds to carrier $l_{K_j}^j$}
Calculate the aggregated final optimal rate $r_j^{agg} = c_i^j + r_i^{j,opt}$ {the i^{th} carrier corresponds to carrier $l_{K_j}^j$}

3.2.4 Simulation Results

Algorithm (5) and Algorithm (6) were applied in C++ to different sigmoidal-like and logarithmic utility functions. The simulation results showed convergence to the global optimal rates. In this section, we present the simulation results for two carriers and 9 UEs shown in Figure 3.9. Three UEs {UE1,UE2,UE3} (first group) are under the coverage area of only Carrier 1 eNodeB, another three UEs {UE4,UE5,UE6} (second group) are joint users under

Algorithm 6 The i^{th} eNodeB Algorithm

Let $w_{i,j}(0) = 0 \ \forall j \in \mathcal{M}_i$
Receive application utility parameters k_j, a_j and b_j from all UEs under the coverage area of the i^{th} eNodeB
loop
  while $|w_{i,j}(n) - w_{i,j}(n-1)| > \delta$ for any $j = \{1, ..., M_i\}$ where the j^{th} UE under the coverage area of the i^{th} eNodeB **do**
    Calculate $p_i^{\text{offered}}(n) = \frac{\sum_{j=1}^{M_i} w_{i,j}(n)}{R_i}$
    for $j \leftarrow 1$ to M_i **do**
      Solve $r_{i,j}(n) = \arg\max_{r_{i,j}} \Big(\log U_j(r_{i,j}) - p_i^{\text{offered}}(n) r_{i,j}(n) \Big)$
      Calculate new $w_{i,j}(n) = p_i^{\text{offered}}(n) r_{i,j}(n)$
      if $|w_{i,j}(n) - w_{i,j}(n-1)| > \Delta w$ **then**
        $w_{i,j}(n) = w_{i,j}(n-1) + \text{sign}(w_{i,j}(n) - w_{i,j}(n-1)) \Delta w(n)$
        $\{\Delta w(n) = l_1 e^{-\frac{n}{l_2}}\}$
      end if
    end for
  end while
  Send the i^{th} eNodeB' shadow price $p_i^{\text{offered}} = p_i^{\text{offered}}(n) = \frac{\sum_{j=1}^{M_i} w_{i,j}(n)}{R_i}$ to all UEs in the eNodeB coverage area
end loop
if The i^{th} eNodeB received Flag assignment of 1 from each UE (the j^{th} UE where $j \in \mathcal{M}_i$) in its coverage area **then**
  loop
    Let $w_i^j(0) = 0 \ \forall j, j = \{1, ..., M_i\}$
    while $|w_i^j(n) - w_i^j(n-1)| > \delta$ for any $j = \{1, ..., M_i\}$ **do**
      Calculate $p_i(n) = \frac{\sum_{j=1}^{M_i} w_i^j(n)}{R_i}$
      for $j \leftarrow 1$ to M_i **do**
        Receive c_i^j value from the j^{th} UE
        Solve $r_i^j(n) = \arg\max_{r_i^j} \Big(\log U_j(r_i^j + c_i^j) - p_i(n) r_i^j(n) \Big)$
        Calculate new $w_i^j(n) = p_i(n) r_i^j(n)$
        if $|w_i^j(n) - w_i^j(n-1)| > \Delta w$ **then**
          $w_i^j(n) = w_i^j(n-1) + \text{sign}(w_i^j(n) - w_i^j(n-1)) \Delta w(n)$
          $\{\Delta w(n) = l_1 e^{-\frac{n}{l_2}}\}$
        end if
      end for
    end while
    Send rate $r_i^{j,opt} = \frac{w_i^j(n)}{R_i}$ to all UEs in the eNodeB coverage area
    Send the shadow price $p_i = p_i(n)$ to all UEs in its coverage area
  end loop
end if

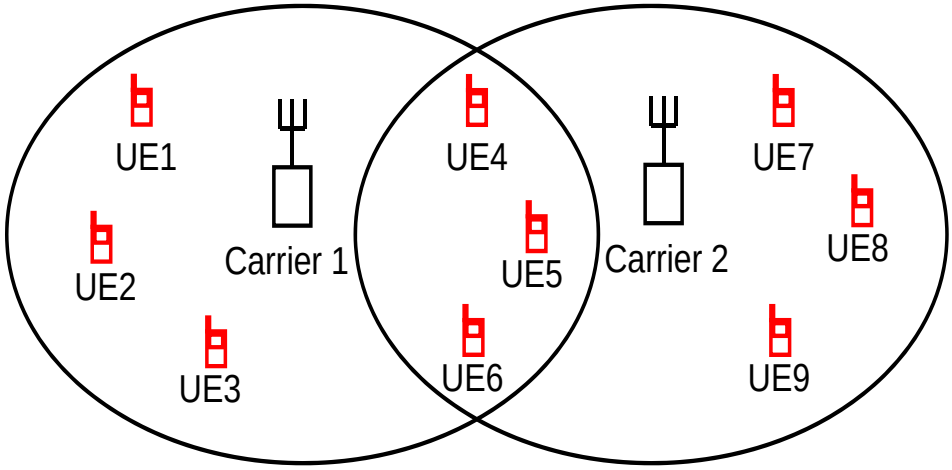


Figure 3.9: System model with two carriers eNodeBs and three groups of users. UE1,UE2 and UE3 under the coverage area of only carrier 1. UE4, UE5 and UE6 under the coverage area of both carriers. UE7, UE8 and UE9 under the coverage area of only carrier 2.

the coverage area of both carrier 1 and carrier 2 eNodeBs and three UEs {UE7,UE8,UE9} (third group) are under the coverage area of only carrier 2 eNodeB. UE1 and UE7 are running the same real-time application that is represented by a normalized sigmoidal-like utility function, that is expressed by equation (2.1), with $a = 5$, $b = 10$ which is an approximation to a step function at rate $r = 10$. UE2 and UE8 are running the same real-time application that is represented by another sigmoidal-like utility function with $a = 3$ and $b = 20$. UE3 and UE9 are running the same delay-tolerant application that is represented by a logarithmic function with $k = 15$. The joint users UE4 and UE5 are running delay tolerant applications that are represented by logarithmic functions with $k = 3$ and $k = 0.5$, respectively. The joint user UE6 is running real-time application that is represented by sigmoidal-like utility function with $a = 1$ and $b = 30$. Additionally, We use $r_{max} = 100$ for all logarithmic functions, $l_1 = 5$ and $l_2 = 10$ in the fluctuation decay function of the algorithm and $\delta = 10^{-3}$. The utility functions corresponding to the nine UEs applications are shown in Figure 3.10.

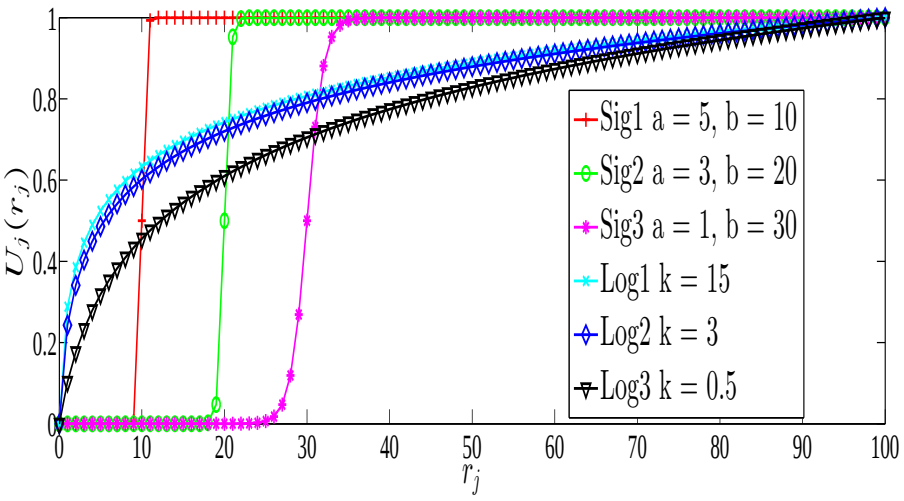


Figure 3.10: The users utility functions $U_j(r_j)$. Sig1 represents UE1 and UE7 applications, Sig2 represents UE2 and UE8 applications, Log1 represents UE3 and UE9 applications, Log2 represents UE4 application, Log3 represents UE5 application and Sig3 represents UE6 application, r_j is the rate allocated to the j^{th} user from all in range eNodeBs.

3.2.4.1 The i^{th} carrier offered Price p_i^{offered} for $50 \leq R_1 \leq 200$ and $R_2 = 100$

In the following simulations, carrier 1 eNodeB available resources R_1 takes values between 50 and 200 with step of 10, and carrier 2 eNodeB available resources is fixed $R_2 = 100$. In Figure 3.11, we consider each carrier to be the primary carrier for all UEs under its coverage area and show that carrier 1 offered price p_1^{offered} is higher than carrier 2 offered price p_2^{offered} when $R_1 \leq R_2$ where $R_2 = 100$. On the other hand, Figure 3.11 shows that $p_2^{\text{offered}} > p_1^{\text{offered}}$ when $R_2 < R_1 \leq 200$. This shows how the carrier's offered price depends on its available resources, the shadow price increases when the carrier's available resources decreases for a fixed number of users. As mentioned before, the joint users select the carrier with the lowest offered price to be their primary carrier. Therefore, in this case the joint users select carrier 2 to be their primary carrier and carrier 1 to be their secondary carrier when $R_1 \leq 100$ whereas they select carrier 1 to be their primary carrier and carrier 2 to be their secondary carrier when $100 < R_1 \leq 200$.

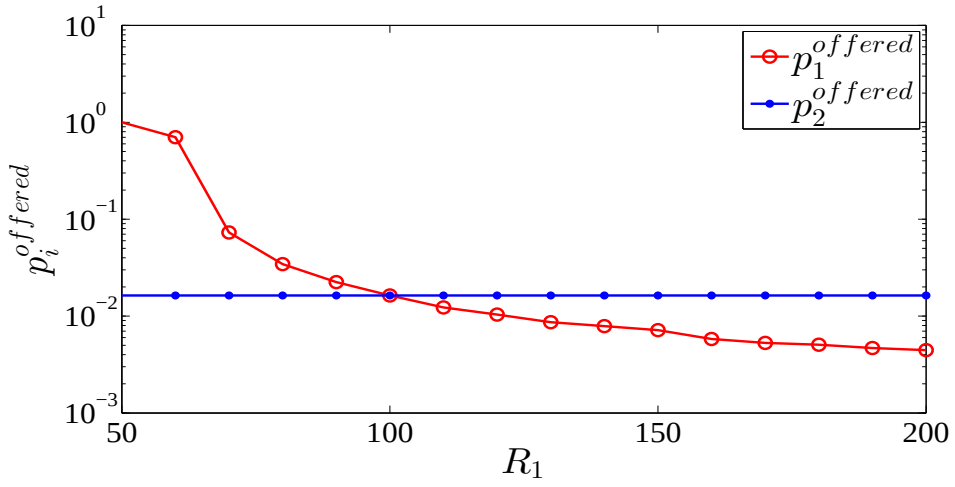


Figure 3.11: Carrier 1 offered price p_1^{offered} for different values of R_1 and fixed number of users and carrier 2 offered price p_2^{offered} for $R_2 = 100$ assuming that each carrier is the primary carrier for all UEs under its coverage area.

3.2.4.2 Aggregated rates r_j^{agg} for $50 \leq R_1 \leq 200$ and $R_2 = 100$

In the following simulations, carrier 1 available resources R_1 takes values between 50 and 200 with step of 10 and carrier 2 eNodeB available resources is fixed $R_2 = 100$. In Figure 3.12, we show the aggregated final optimal rates for the nine users with different available resources R_1 of carrier 1. The final optimal rates r_j^{agg} for the first group of UEs are allocated to them by only carrier 1 as they are under the coverage area of only that carrier and do not have secondary carriers. Similarly, the final optimal rates r_j^{agg} for the third group of UEs are allocated to them by carrier 2 as they are under the coverage area of only that carrier and do not have secondary carriers. On the other hand, the second group of UEs are joint users and are allocated rates from both carriers. The joint users select their primary carrier l_1^j to be the carrier with the lowest shadow price $l_1^j = \arg\min_{\{1,2\}}\{p_1^{\text{offered}}, p_2^{\text{offered}}\}$ and the other carrier with a higher offered price to be their secondary carrier l_2^j. The aggregated final optimal rate allocated to each joint user is the aggregated rate of its primary carrier allocated rate and its secondary carrier allocated rate. Figure 3.12 shows that users running real-time applications are given priority over users running delay tolerant applications and

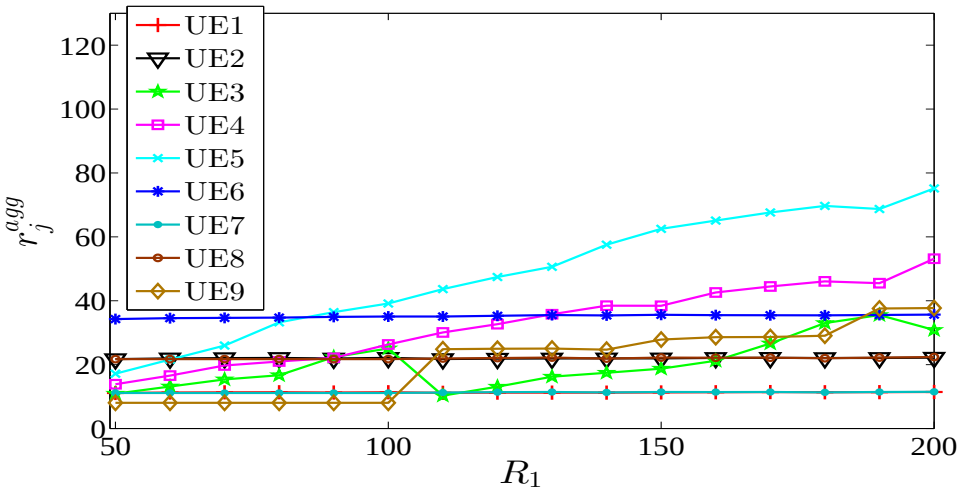


Figure 3.12: The aggregated final optimal allocated rate r_j^{agg} for each user from its all in range carriers versus carrier 1 available resources $50 \leq R_1 \leq 200$ with carrier 2 available resources fixed at $R_2 = 100$.

are allocated higher rates in the case of low carrier's available resources.

3.3 Multi-Stage Resource Allocation with Carrier Aggregation for Commercial Use of 3.5 GHz Spectrum

The Commission and the President have outlined a path to double the available spectrum for wireless broadband use, the PCAST Report identifies two technological advances to increase wireless broadband capabilities. First, increasing the deployment of small cell networks and second using spectrum sharing technology. The 3.5 GHz Band is an ideal band for small cell deployments and shared spectrum use because of its smaller coverage. The National Institute of Standards and Technology (NTIA) Fast Track Report [11] identified the 3.5 GHz Band for potential shared federal and non-federal broadband use. This band is very favorable for commercial cellular systems such as LTE-Advanced systems. Small cells are low-powered wireless base stations designed to play well with macro networks in a heteroge-

neous network (HetNet). Small cells are backed up by a macro cell layer of coverage so that if a small cell shuts down in the 3.5 GHz shared band, operators can pick up coverage again in the macro network.

In this section, we introduce an application-aware spectrum sharing approach for sharing the Federal under-utilized 3.5 GHz spectrum with commercial users. In our model, users are running elastic or inelastic traffic and each application running on the UE is assigned a utility function based on its type. Furthermore, each of the small cells' users has a minimum required target utility for its application. In order for users located under the coverage area of the small cells' eNodeBs, with the 3.5 GHz band resources, to meet their minimum required quality of experience (QoE), the network operator makes a decision regarding the need for sharing the macro cell's resources to obtain additional resources. Our objective is to provide each user with a rate that satisfies its application's minimum required utility through spectrum sharing approach and improve the overall QoE in the network. We present an application-aware spectrum sharing algorithm that is based on resource allocation with carrier aggregation to allocate macro cell permanent resources and small cells' leased resources to UEs and allocate each user's application an aggregated rate that can at minimum achieves the application's minimum required utility.

Our contributions in this section are summarized as:

- We present a spectrum sharing approach for sharing the Federal under-utilized 3.5 GHz spectrum with commercial users.
- We present a spectrum sharing algorithm that is based on resource allocation with CA to allocate the small cells' under-utilized 3.5 GHz resources to small cells' users and allocate the macro cell's resources to both macro cell's users and small cell's users that did not reach their applications minimum required utilities by the small cells allocated rates.

- We present simulation results for the performance of the proposed resource allocation algorithm.

3.3.1 Problem Formulation

We consider LTE-Advanced mobile system consisting of a macro cell, referred to by the index B, with a coverage radius D_B, that is overlaid with S small cells. The macro cell's eNodeB is configured at the LTE-Advanced carrier and the small cell's eNodeB is configured to use the 3.5 GHz under-utilized spectrum band. Let $\mathcal{S}$ denotes the set of small cells located within the coverage area of the macro cell B where $S = |\mathcal{S}|$. All small cells are connected to the core network. The small cells are assumed to have a closed access scheme where only registered UEs, referred to by SUEs, are served by the small cells eNodeBs. On the other hand, all UEs under the coverage area of the macro cell B and not within the coverage of any small cell, referred to by MUEs, are served by the macro cell's eNodeB. The set of all MUEs under the coverage area of macro cell B is referred to by μ. The set of SUEs associated to small cell s is referred to by $\mathcal{Q}_s$. We assume that the association of the UEs with their eNodeBs remains fixed during the runtime of the resource allocation process. We have $\bigcup_{s=1}^{S} \mathcal{Q}_s = \Theta$ and $\bigcap_{s=1}^{S} \mathcal{Q}_s = \emptyset$. Each SUE i has a minimum QoE requirement for its applications that is represented by the utility of the user's application with its allocated rate. Let u_i^{req} denotes the minimum required utility of SUE $i \in \Theta$.

We express the user satisfaction with its application rates using utility functions. We represent the i^{th} user application utility function $U_i(r_i)$ by sigmoidal-like function or logarithmic function where r_i is the rate of the i^{th} user application. Logarithmic utility functions expressed by equation (2.2) and sigmoidal-like utility functions expressed by equation (2.1) are used to represent delay tolerant and real-time applications, respectively.

Figure 3.13 shows a heterogeneous network that consists of one macro cell with one eNodeB

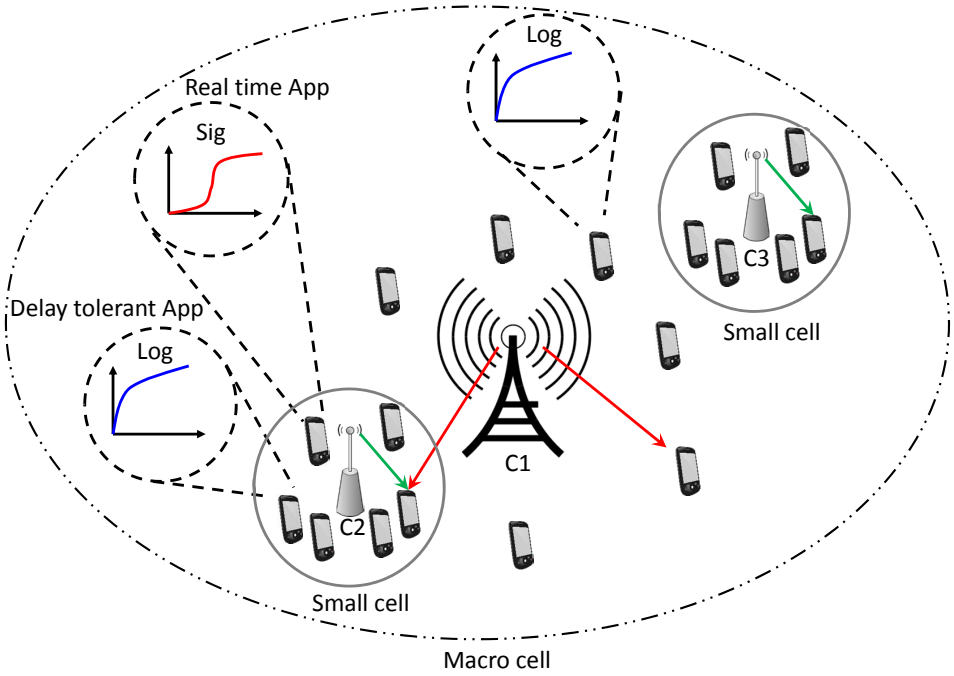


Figure 3.13: System model for a LTE-Advanced mobile system with one macro cell and two small cells within the coverage area of the macro cell. Each of the small cells is configured to use the 3.5 GHz under-utilized spectrum.

and two small cells within the coverage area of the macro cell, each of the small cells has one eNodeB that is configured to use the 3.5 GHz under-utilized spectrum. Mobile users under the coverage of the macro cell and the small cells are running real time or delay tolerant applications that are represented by sigmoidal-like or logarithmic utility functions, respectively.

3.3.2 Resource Allocation Optimization for Spectrum Sharing with the 3.5 GHz Spectrum

In this section, we present a resource allocation framework for cellular networks sharing the federal under-utilized 3.5 GHz spectrum. In our model, SUEs are allocated resources from the leased under-utilized 3.5 GHz resources at the small cells eNodeBs whereas MUEs are allocated resources only by the macro cell's eNodeB. Each of the SUEs has a minimum

required utility u_i^{req} for each of its applications. First the small cell's eNodeB allocates its available leased resources then the network operator decides which SUEs still require additional resources in order to achieve their minimum required utilities and allocate them more resources from the macro cell eNodeB based on a resource allocation with carrier aggregation optimization problem.

The resource allocation process starts by allocating each of the small cells resources to SUEs under it coverage area. We use a utility proportional fairness resource allocation optimization problem to allocate the small cell resources. The RA optimization problem of the small cell s is given by:

$$\begin{aligned} \max_{\mathbf{r}^s} \quad & \prod_{i=1}^{|\mathcal{Q}_s|} U_i(r_i^s) \\ \text{subject to} \quad & \sum_{i=1}^{|\mathcal{Q}_s|} r_i^s \leq R_s \\ & 0 \leq r_i^s \leq R_s, \qquad i = 1, 2, ..., |\mathcal{Q}_s|, \end{aligned} \tag{3.15}$$

where $\mathbf{r}^s = \{r_1^s, r_2^s, ..., r_{|\mathcal{Q}_s|}^s\}$, $|\mathcal{Q}_s|$ is the number of SUEs under the coverage area of the small cell s and R_s is the maximum achievable rate of the under-utilized 3.5 GHz leased spectrum available at the eNodeB of small cell s. The resource allocation objective function is to maximize the entire small cell utility when allocating its resources. It also achieves proportional fairness among utilities such that non of the SUEs will be allocated zero resources. Therefore, a minimum QoS is provided to each SUE. This approach gives real time applications priority when allocating the small cell resources. The objective function in optimization problem (3.15) is equivalent to $\max_{\mathbf{r}^s} \sum_{i=1}^{|\mathcal{Q}_s|} \log U_i(r_i^s)$. Optimization problem (3.15) is a convex optimization problem and there exists a unique tractable global optimal solution [21].

From optimization problem (3.15), we have the Lagrangian:

$$L_s(\mathbf{r}^s, p^s) = (\sum_{i=1}^{|\mathcal{Q}_s|} \log U_i(r_i^s)) - p^s(\sum_{i=1}^{|\mathcal{Q}_s|} r_i^s + z_s - R_s) \tag{3.16}$$

where $z_s \geq 0$ is the slack variable and p^s is the Lagrange multiplier which is equivalent to the shadow price that corresponds to the service provider's price per unit bandwidth for the small cell resources [21].

The solution of equation (3.15) is given by the values r_i^s that solve equation $\frac{\partial \log U_i(r_i^s)}{\partial r_i^s} = p^s$ and are the intersection of the time varying shadow price, horizontal line $y = p^s$, with the curve $y = \frac{\partial \log U_i(r_i^s)}{\partial r_i^s}$ geometrically. Once the RA process is performed by the small cell s, each SUE in $\mathcal{Q}_s$ will be allocated $r_i^{s,\text{all}} = r_i^s$ rate. However, the network operator decides if any of the SUEs requires additional resources in order to reach the minimum required utility u_i^{req} of its application by comparing the utility of the small cell allocated rate that is given by $U_i(r_i^{s,\text{all}})$ with the value u_i^{req}. If the achieved utility for certain SUE is less that the minimum required utility, the network operator requests additional resources from the macro cell for that SUE. The small cell s eNodeB creates a set $\mathcal{Q}_{sB}$ of all SUEs that needs to be allocated additional resources where $\mathcal{Q}_{sB} = \{\text{SUEs} \in \mathcal{Q}_s \;\; s.t. \;\; u_i^{\text{req}} > U_i(r_i^{s,\text{all}})\}$.

Once each small cell s within the coverage area of the macro cell B performs its RA process based on optimization problem (3.15), the macro cell starts allocating its resources to all MUEs within its coverage area as well as the SUEs that were reported, by the network operator, for their need of additional resources. Let $\mathcal{Q}$ be the set of SUEs that will be allocated additional resources by the macro cell where $\mathcal{Q} = \bigcup_{s=1}^{S} \mathcal{Q}_{sB}$. The set of UEs that will be served by the macro cell's eNodeB; i.e. participate in the macro cell RA process, is given by β where $\beta = \mu \bigcup \mathcal{Q}$. The resource allocation optimization problem of the macro cell B is given by:

$$\begin{aligned} \max_{\mathbf{r}} \quad & \prod_{i=1}^{|\beta|} U_i(r_i + C_i) \\ \text{subject to} \quad & \sum_{i=1}^{|\beta|} r_i \leq R_B \\ & C_i = \begin{cases} 0 & \text{if UE } i \notin \mathcal{Q} \\ r_i^{s,\text{all}} & \text{if UE } i \in \mathcal{Q} \end{cases} \\ & 0 \leq r_i \leq R_B, \quad i = 1, 2, ..., |\beta|, \end{aligned} \tag{3.17}$$

where $\mathbf{r} = \{r_1, r_2, ..., r_{|\beta|}\}$, $|\beta|$ is the number of UEs that will be be served by the macro cell's eNodeB and R_B is the maximum achievable rate of the resources available at the macro cell's eNodeB. The resource allocation objective function is to maximize the entire macro cell utility when allocating its resources. The RA optimization problem (3.17) is based on carrier aggregation. It seeks to maximize the multiplication of the utilities of the rates allocated to MUEs by the macro cell's eNodeB and the utilities of the rates allocated to the SUEs in β by small cells' eNodeBs and macro cell's eNodeB. Utility proportional fairness is used to guarantee that non of the UEs will be allocated zero resources. Real time applications are given priority when allocating the macro resources using this approach. The objective function in optimization problem (3.17) is equivalent to $\max_{\mathbf{r}} \sum_{i=1}^{|\beta|} \log U_i(r_i + C_i)$. Optimization problem (3.17) is a convex optimization problem and there exists a unique tractable global optimal solution [21].

From optimization problem (3.17), we have the Lagrangian:

$$L_B(\mathbf{r}, p^B) = (\sum_{i=1}^{|\beta|} \log U_i(r_i + C_i)) - p^B(\sum_{i=1}^{|\beta|} r_i + z_B - R_B) \tag{3.18}$$

where $z_B \geq 0$ is the slack variable and p^B is the Lagrange multiplier which is equivalent to the shadow price that corresponds to the service provider's price per unit bandwidth for the macro cell resources [21].

The solution of equation (3.17) is given by the values r_i that solve equation $\frac{\partial \log U_i(r_i+C_i)}{\partial r_i} = p^B$ and are the intersection of the time varying shadow price, horizontal line $y = p^B$, with the curve $y = \frac{\partial \log U_i(r_i+C_i)}{\partial r_i}$ geometrically. Once the macro cell eNodeB is done performing the RA process based on optimization problem (3.17), each UE in β will be allocated $r_i^{\text{all}} = r_i + C_i$ rate.

3.3.3 The Macro Cell and Small Cells RA Optimization Algorithm

In this section, we present our resource allocation algorithm. The proposed algorithm consists of SUE, MUE, small cell eNodeB and macro cell eNodeB parts shown in Algorithm 7, 8, 9 and 10, respectively. The execution of the algorithm starts by SUEs and MUEs, subscribing for mobile services, transmitting their applications utilities parameters to their corresponding eNodeBs. First, each small cell s eNodeB calculates its allocated rate $r_i^{s,\text{all}}$ to each SUE in $\mathcal{Q}_s$. It then checks whether the achievable utility of that rate is less or greater than the SUE's minimum required utility u_i^{req}. If for any SUE $U_i(r_i^{s,\text{all}}) < u_i^{\text{req}}$, the small cell's eNodeB sends the application parameters and the allocated rate $r_i^{s,\text{all}}$ for that SUE to the macro cell's eNodeB requesting additional resources. Otherwise, it allocates the rate $r_i^{s,\text{all}}$ to that SUE.

Once the macro cell's eNodeB receives the set $\mathcal{Q}_{sB}$ from each small cell in $\mathcal{S}$ within its coverage area. It starts the RA process to allocate its available resources to each UE in β based on a RA with carrier aggregation optimization problem. Once the RA process of the macro cell is performed, the macro cell allocates rate $r_i^{\text{all}} = r_i + C_i$ to the i^{th} UE in β.

Algorithm 7 The i^{th} SUE $\in \mathcal{Q}_s$ Algorithm

loop

Send application utility parameters k_i, a_i, b_i, $r_i^{\max}$ and u_i^{req} to the SUE's in band small cell's eNodeB.

Receive the final allocated rate $r_i^{s,\text{all}}$ from the small cell s eNodeB or from the macro cell's eNodeB.

end loop

Algorithm 8 The i^{th} MUE $\in \mu$ Algorithm

loop
  Send application utility parameters k_i, a_i, b_i and $r_i^{\max}$ to the macro cell's eNodeB.
  Receive the final allocated rate r_i^{all} from the macro cell's eNodeB.
end loop

Algorithm 9 Small Cell s eNodeB Algorithm

loop
  Initialize $\mathcal{Q}_{sB} = \emptyset$; $r_i^{\text{all}} = 0$.
  Receive application utility parameters k_i, a_i, b_i, $r_i^{\max}$ and u_i^{req} from all SUEs in $\mathcal{Q}_s$.
  Solve $\mathbf{r}^s = \arg\max_{\mathbf{r}^s} \sum_{i=1}^{|\mathcal{Q}_s|} \log U_i(r_i^s) - p^s(\sum_{i=1}^{|\mathcal{Q}_s|}(r_i^s) - R_s)$.
  Let $r_i^{s,\text{all}} = r_i^s$ be the rate allocated by the s small cell's eNodeB to each user in $\mathcal{Q}_s$.
  Calculate the SUE utility $U_i(r_i^{s,\text{all}})$ $\forall i \in \mathcal{Q}_s$
  for SUE $i \leftarrow 1$ to $|\mathcal{Q}_s|$ **do**
    if $U_i(r_i^{s,\text{all}}) < u_i^{\text{req}}$ **then**
      $\mathcal{Q}_{sB} = \mathcal{Q}_{sB} \bigcup \text{SUE}\{i\}$
      Send SUE i parameters k_i, a_i, b_i, $r_i^{\max}$ and $r_i^{s,\text{all}}$ to the macro cell's eNodeB
    else
      Allocate rate $r_i^{\text{all}} = r_i^{s,\text{all}}$ to SUE i
    end if
  end for
end loop

Algorithm 10 The Macro Cell's eNodeB Algorithm

loop
  Initialize $C_i = 0$; $r_i^{\text{all}} = 0$.
  for $s \leftarrow 1$ to S **do**
    Receive application utility parameters k_i, a_i, b_i, $r_i^{\max}$ and $r_i^{s,\text{all}}$ for all SUEs in $\mathcal{Q}_{sB}$ from small cell s eNodeB.
    $C_i = r_i^{s,\text{all}}$ $\forall i \in \mathcal{Q}_{sB}$
  end for
  Create user group $\mathcal{Q} = \bigcup_{s=1}^{S} \mathcal{Q}_{sB}$
  Create user group $\beta = \mu \bigcup \mathcal{Q}$
  Solve $\mathbf{r} = \arg\max_{\mathbf{r}} \sum_{i=1}^{|\beta|} \log U_i(r_i + C_i) - p^B(\sum_{i=1}^{|\beta|}(r_i) - R_B)$.
  Allocate $r_i^{\text{all}} = r_i + C_i$ to each UE i in β
end loop

Table 3.1: Users and their applications utilities

User's Index	User's Type	Applications Utilities Parameters
UE1 $i = \{1\}$	SUE	Sig2 $a_i = 3$, $b_i = 20$, $u_i^{\text{req}} = 0.8$
UE2 $i = \{2\}$	SUE	Sig3 $a_i = 1$, $b_i = 30$, $u_i^{\text{req}} = 0.8$
UE3 $i = \{3\}$	SUE	Log2 $k_i = 3$, $r_i^{\max} = 100$, $u_i^{\text{req}} = 0.5$
UE4 $i = \{4\}$	SUE	Log3 $k_i = 0.5$, $r_i^{\max} = 100$, $u_i^{\text{req}} = 0.5$
UE5 $i = \{5\}$	MUE	Sig1 $a_i = 5$, $b_i = 10$
UE6 $i = \{6\}$	MUE	Sig3 $a_i = 1$, $b_i = 30$
UE7 $i = \{7\}$	MUE	Log1 $k_i = 15$, $r_i^{\max} = 100$
UE8 $i = \{8\}$	MUE	Log3 $k_i = 0.5$, $r_i^{\max} = 100$

3.3.4 Simulation Results

Algorithm 7, 8, 9 and 10 were applied in C++ to multiple utility functions with different parameters. Simulation results showed convergence to the global optimal rates. In this section, we consider a macro cell with one eNodeB. Within the coverage area of the macro cell there exists one small cell s. Four SUEs are located under the coverage area of the small cell s with UEs indexes $\{1, 2, 3, 4\}$. The SUEs user group is given by $\mathcal{Q}_s = \{1, 2, 3, 4\}$. Four MUEs are located under the coverage area of the macro cell's eNodeB but not within the small cell. The MUEs user group is given by $\mu = \{5, 6, 7, 8\}$. Each UE whether it is SUE or MUE is running either real time application or delay tolerant application. Each of the SUEs' applications utilities has a minimum required utility that is given by u_i^{req} that is equivalent to the C_i value for that user whereas MUEs do not have minimum required utilities for their applications. The UEs' indexes, types and applications utilities parameters are listed in table 3.1. Figure 3.14 shows the sigmoidal-like utility functions and the logarithmic utility functions used to represent the SUEs and MUEs applications.

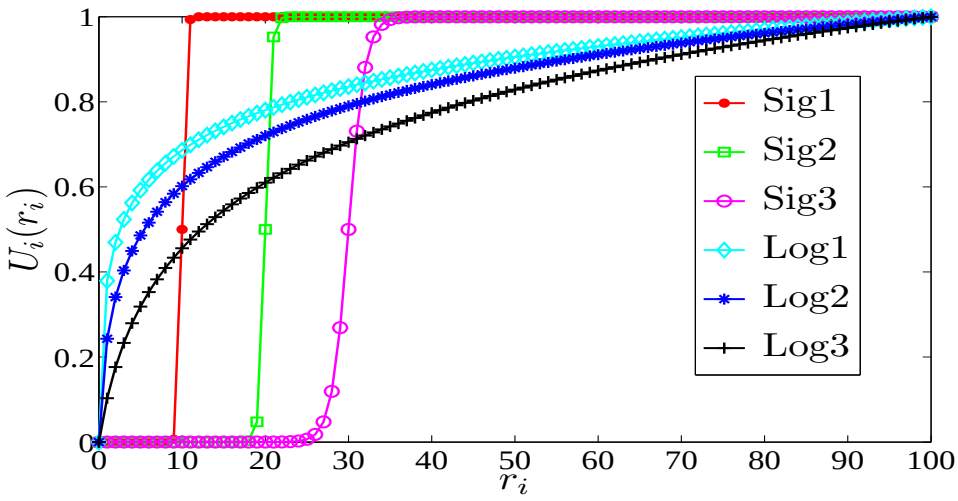
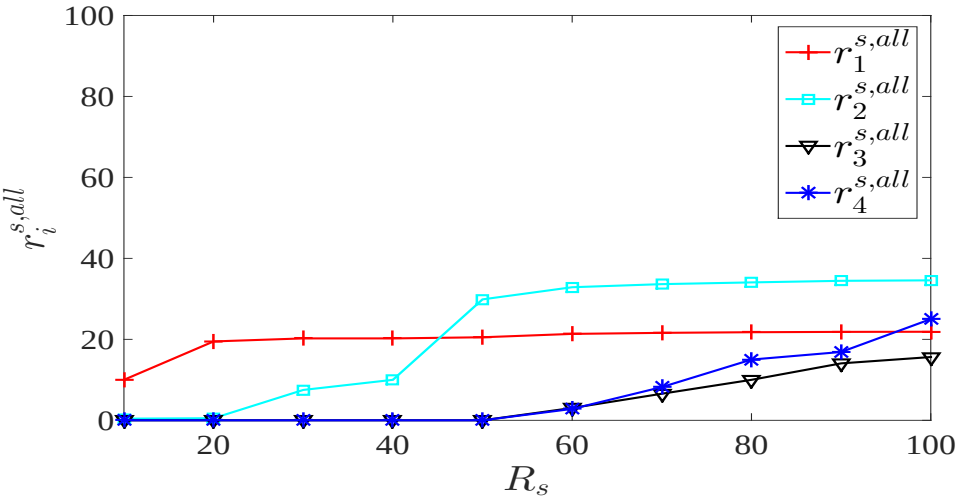


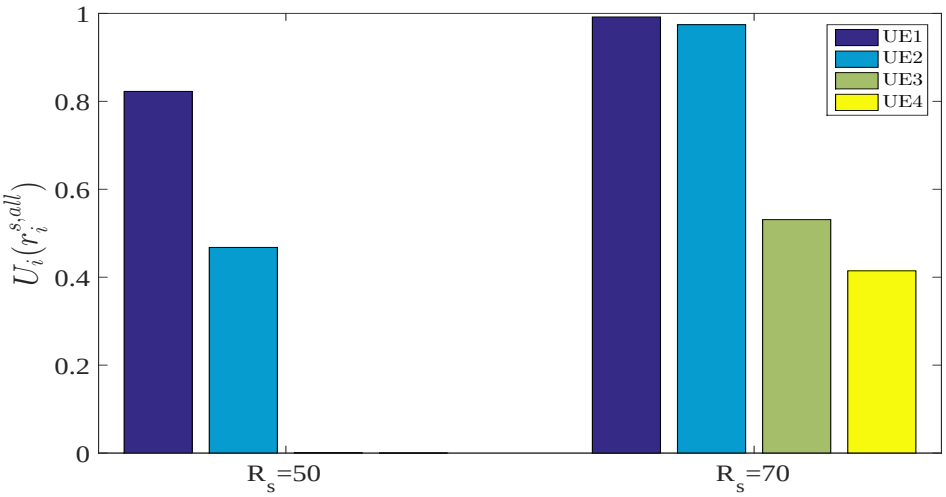
Figure 3.14: The users utility functions $U_i(r_i)$ used in the simulation (three sigmoidal-like functions and three logarithmic functions).

3.3.4.1 Small Cell Allocated Rates and Users QoE

In the following simulations, the small cell's carrier total rate R_s takes values between 10 and 100 with step of 10. In Figure 3.15, we show the small cell's allocated rates $r_i^{s,\text{all}}$ for users in $\mathcal{Q}_s$ with different values of the small cell's carrier total rate R_s and the users QoE with the small cell allocated rates when $R_s = 50$ and $R_s = 70$. In Figure 3.15(a), we show that users running real time applications are given priority when allocating the small cell's resources due to their sigmoidal-like utility function nature. We also observe that non of the UEs is allocated zero resources because we used a utility proportional fairness approach. We also show how the proposed rate allocation algorithm converges for different values of R_s. In Figure 3.15(b), we show the QoE for the four SUEs which is represented by their applications utilities of the small cell allocated rates $U_i(r_i^{s,\text{all}})$ when $R_s = 50$ and $R_s = 70$. We notice that in the case of $R_s = 50$, the utilities of the small cell allocated rates for UE2, UE3 and UE4 did not reach the minimum required utilities for these SUEs whereas in the case of $R_s = 70$ the utility of the small cell allocated rate for UE4 did not reach the minimum required utility for that SUE. Therefore, based on the proposed algorithm the network operator will request additional resources for these UEs from the macro cell's



(a) The rates $r_i^{s,\text{all}}$ allocated by the small cell's eNodeB to users in $\mathcal{Q}_s$ with $10 < R_s < 100$.



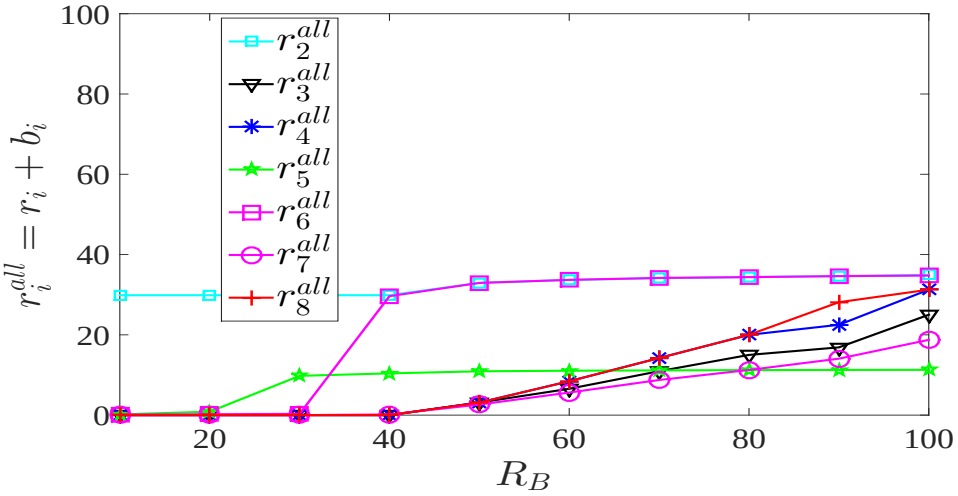
(b) Users' QoE represented by the utility of user's application of its allocated rate $U_i(r_i^{s,\text{all}})$ when $R_s = 50$ and $R_s = 70$.

Figure 3.15: The small cell's eNodeB allocated rates with $10 < R_s < 100$ and users' QoE when $R_s = 50$ and $R_s = 70$.

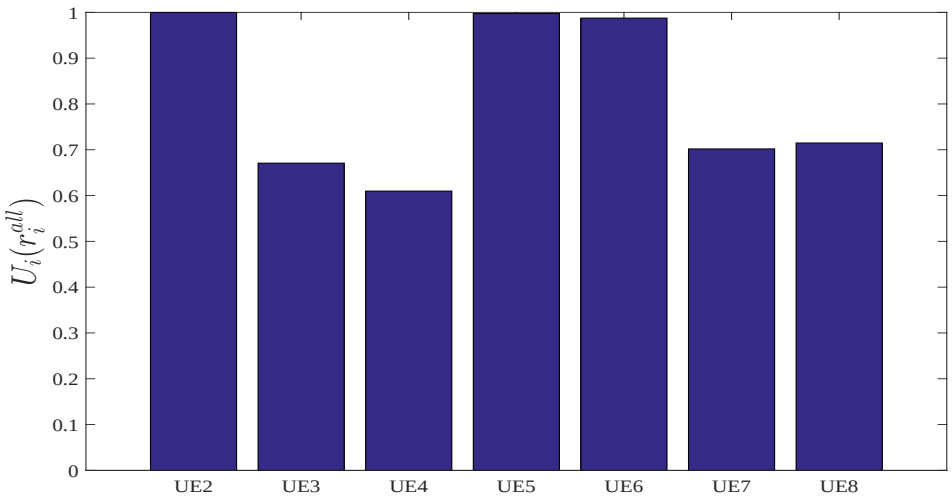
eNodeB and these UEs will be allocated additional resources based on a resource allocation with carrier aggregation scenario.

3.3.4.2 Macro Cell Allocated Rates and Users QoE

In the following simulations, the macro cell's carrier total rate R_B takes values between 10 and 100 with step of 10 and R_s is fixed at 50. As discussed in 3.3.4.1, in the case of $R_s = 50$ the network operator requests additional resources for three SUEs (i.e. UEs in $\mathcal{Q}_{sB} = \{2, 3, 4\}$) as they did not reach their minimum required utilities. Therefore, the macro cell's eNodeB performs a resource allocation with carrier aggregation process to allocate resources to the UEs in user group β where $\beta = \{2, 3, 4, 5, 6, 7, 8\}$. In Figure 3.16, we show the final allocated rates r_i^{all} for the UEs in β and these users QoE with the final allocated rates when $R_B = 80$. In Figure 3.16(a), we show the macro cell's final allocated rates converge for different values of R_B. Again we observe that non of the users is allocated zero resources and that real time applications are given priority when allocating the macro cell's resources. In Figure 3.16(b), we show the QoE for each of the seven UEs in β which is represented by the utility, of the final allocated rate $U_i(r_i^{\text{all}})$, of the user's application when $R_s = 50$ and $R_B = 80$. We notice that the utilities of the final allocated rates for the three SUEs in $\mathcal{Q}_{sB}$ (i.e. UE {2,3,4}) exceed the minimum required utilities for these SUEs because of the additional resources allocated to these users by the macro cell's eNodeB.



(a) The aggregated rates $r_i^{\text{all}} = r_i + C_i$ allocated by the macro cell's eNodeB to users in β when $R_s = 50$.



(b) Users' QoE represented by the utility of user's application of its allocated rate $U_i(r_i^{\text{all}})$ when $R_B = 80$ and $R_s = 50$.

Figure 3.16: The total aggregated rates $r_i^{\text{all}} = r_i + C_i$ allocated by the macro cell's eNodeB to users in β with $10 < R_B < 100$ when $R_s = 50$ and the users' QoE when $R_B = 80$ and $R_s = 50$.

3.4 Summary and Conclusions

In this chapter, we introduced a novel RA with CA optimization problem in cellular networks. We considered mobile users with elastic or inelastic traffic and used utility functions to represent the applications running on the UEs. We presented an iterative decentralized rate allocation with CA algorithm to allocate both the primary and secondary carriers resources optimally among users located under the coverage area of both carriers. We also presented a novel price selective centralized algorithm for allocating resources from different carriers optimally among users. Our price selective algorithm allows each user to select its primary and secondary carriers based on their offered prices in order to guarantee a minimum price for the aggregated final allocated rates. The centralized algorithm is performed mostly in the eNodeBs. Therefore, it requires less transmission overhead and less computations in the UEs. The proposed algorithms use proportional fairness approach to provide a minimum QoS to all users while giving priority to real-time application users. We analyzed the convergence of the algorithms with different carriers available resources and showed through simulations that our algorithms converge to optimal values.

In addition, we proposed a spectrum sharing approach for sharing the Federal underutilized 3.5 GHz spectrum with commercial users and presented multi-stage resource allocation with CA algorithms to allocate the macro cell and small cells resources optimally among users under their coverage area. Users located under the coverage area of the small cells are allocated resources by the small cells' eNodeBs whereas both the macro cell users and the small cells' users that did not reach their minimum required utilities, by their small cells' allocated rates, are allocated resources by the macro cell's eNodeB based on carrier aggregation. We showed through simulations that the proposed algorithm converges to the optimal rates. We also showed that small cells' users can achieve their minimum required QoE by using the proposed spectrum sharing approach.

Chapter 4

Robust RA with Joint CA for Multi-Carrier Cellular Networks

In this chapter, we focus on solving the problem of utility proportional fairness optimal RA with joint CA for multi-carrier cellular networks. The RA with joint CA algorithm presented in [1] fails to converge for high-traffic situations due to the fluctuation in the RA process. In this chapter, we present a robust algorithm that solves the drawbacks in [1] and allocates multiple carriers resources optimally among UEs in their coverage area for both high-traffic and low-traffic situations. Additionally, our proposed distributed algorithm outperforms the multi-stage RA with CA algorithms presented in [24,85,86] as it guarantees that mobile users are assigned optimal (minimum) price for resources. We formulate the multi-carrier RA with CA optimization problem into a convex optimization framework. We use logarithmic and sigmoidal-like utility functions to represent delay-tolerant and real-time applications, respectively, running on the mobile users' smart phones [21]. Our model supports both contiguous and non-contiguous carrier aggregation from one or more network providers. During the resource allocation process, our distributed algorithm allocates optimal resources from one or more carriers to provide the lowest resource price for the mobile

users. In addition, we use a utility proportional fairness approach that ensures non-zero resource allocation for all users and gives real-time applications priority over delay-tolerant applications due to the nature of their applications that require minimum encoding rates.

Our contributions in this chapter are summarized as:

- We consider the RA optimization problem with joint CA presented in [1] that uses utility proportional fairness approach and solves for logarithmic and sigmoidal-like utility functions representing delay-tolerant and real-time applications, respectively.
- We prove that the optimization problem is convex and therefore the global optimal solution is tractable. In addition, we present a robust distributed resource allocation algorithm to solve the optimization problem and provide optimal rates in high-traffic and low-traffic situations.
- Our proposed algorithm outperforms that presented in [1] by preventing the fluctuations in the RA process when the resources are scarce with respect to the number of users. It also outperforms the algorithms presented in [24, 85, 86] as it guarantees that mobile users receive optimal price for resources.
- We present simulation results for the performance of our RA algorithm and compare it with the performance of the multi-stage RA algorithm presented in [24, 85, 86].

4.1 Problem Formulation

We consider LTE mobile system consisting of K carriers eNodeBs with K cells and M UEs distributed in these cells. The rate allocated by the l^{th} carrier eNodeB to i^{th} UE is given by r_{li} where $l = \{1, 2, ..., K\}$ and $i = \{1, 2, ..., M\}$. Each UE has its own utility function $U_i(r_{1i} + r_{2i} + ... + r_{Ki})$ that corresponds to the type of traffic being handled by the i^{th} UE. Our objective is to determine the optimal rates that the l^{th} carrier eNodeB should allocate

to the nearby UEs. We express the user satisfaction with its provided service using utility functions that represent the degree of satisfaction of the user function with the rate allocated by the cellular network [82] [18] [83]. We assume the utility functions $U_i(r_{1i} + r_{2i} + ... + r_{Ki})$ to be a strictly concave or a sigmoidal-like functions. The utility functions have the following properties:

- $U_i(0) = 0$ and $U_i(r_{1i} + r_{2i} + ... + r_{Ki})$ is an increasing function of r_{li} for l.
- $U_i(r_{1i} + r_{2i} + ... + r_{Ki})$ is twice continuously differentiable in r_{li} for all l.

In our model, we use the normalized sigmoidal-like utility function, as in [82], that can be expressed as

$$U_i(r_{1i} + r_{2i} + ... + r_{Ki}) = c_i \left(\frac{1}{1 + e^{-a_i(\sum_{l=1}^{K} r_{li} - b_i)}} - d_i \right) \tag{4.1}$$

where $c_i = \frac{1+e^{a_i b_i}}{e^{a_i b_i}}$ and $d_i = \frac{1}{1+e^{a_i b_i}}$. So, it satisfies $U_i(0) = 0$ and $U_i(\infty) = 1$. We use the normalized logarithmic utility function, as in [83], that can be expressed as

$$U_i(r_{1i} + r_{2i} + ... + r_{Ki}) = \frac{\log(1 + k_i \sum_{l=1}^{K} r_{li})}{\log(1 + k_i r_{max})} \tag{4.2}$$

where r_{max} is the required rate for the user to achieve 100% utility percentage and k_i is the rate of increase of utility percentage with allocated rates. So, it satisfies $U_i(0) = 0$ and $U_i(r_{max}) = 1$. We consider the utility proportional fairness objective function that is given by

$$\max_{\mathbf{r}} \prod_{i=1}^{M} U_i(r_{1i} + r_{2i} + ... + r_{Ki}) \tag{4.3}$$

where $\mathbf{r} = \{\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_M\}$ and $\mathbf{r}_i = \{r_{1i}, r_{2i}, ..., r_{Ki}\}$. The goal of this resource allocation objective function is to maximize the total system utility while ensuring proportional fairness between utilities (i.e., the product of the utilities of all UEs). This resource allocation objective function inherently guarantees:

- non-zero resource allocation for all users. Therefore, the corresponding resource allocation optimization problem provides a minimum QoS for all users.
- priority to users with real-time applications. Therefore, the corresponding resource allocation optimization problem improves the overall QoS for LTE system.

The basic formulation of the utility proportional fairness resource allocation problem is given by the following optimization problem:

$$\begin{aligned} \max_{\mathbf{r}} \quad & \prod_{i=1}^{M} U_i(r_{1i} + r_{2i} + ... + r_{Ki}) \\ \text{subject to} \quad & \sum_{i=1}^{M} r_{1i} \leq R_1, \sum_{i=1}^{M} r_{2i} \leq R_2, ... \\ & ... \ , \ \sum_{i=1}^{M} r_{Ki} \leq R_K, \\ & r_{li} \geq 0, \quad l = 1, 2, ..., K, \ \ i = 1, 2, ..., M \end{aligned} \tag{4.4}$$

where R_l is the total available rate at the l^{th} carrier eNodeB.

We prove in Section 4.2 that the solution of the optimization problem (4.4) is the global optimal solution.

4.2 The Global Optimal Solution

In the optimization problem (4.4), since the objective function $\arg\max_{\mathbf{r}} \prod_{i=1}^{M} U_i(r_{1i} + r_{2i} + ... + r_{Ki})$ is equivalent to $\arg\max_{\mathbf{r}} \sum_{i=1}^{M} \log(U_i(r_{1i} + r_{2i} + ... + r_{Ki}))$, then optimization problem (4.4) can be written as:

$$\begin{aligned} \max_{\mathbf{r}} \quad & \sum_{i=1}^{M} \log\left(U_i(r_{1i} + r_{2i} + ... + r_{Ki})\right) \\ \text{subject to} \quad & \sum_{i=1}^{M} r_{1i} \leq R_1, \sum_{i=1}^{M} r_{2i} \leq R_2, ... \\ & ... \ , \ \sum_{i=1}^{M} r_{Ki} \leq R_K, \\ & r_{li} \geq 0, \quad l = 1, 2, ..., K, \ \ i = 1, 2, ..., M. \end{aligned} \tag{4.5}$$

Lemma 4.2.1. *The utility functions* $\log(U_i(r_{1i}+...+r_{Ki}))$ *in the optimization problem (4.5) are strictly concave functions.*

Proof. In Section 4.1, we assume that all the utility functions of the UEs are strictly concave or sigmoidal-like functions.

In the strictly concave utility function case, recall the utility function properties in Section 4.1, the utility function is positive $U_i(r_{1i} + ... + r_{Ki}) > 0$, increasing and twice differentiable with respect to r_{li}. Then, it follows that $\frac{\partial U_i(r_{1i}+...+r_{Ki})}{\partial r_{li}} > 0$ and $\frac{\partial^2 U_i(r_{1i}+...+r_{Ki})}{\partial r_{li}^2} < 0$. It follows that, the utility function $\log(U_i(r_{1i} + r_{2i} + ... + r_{Ki}))$ in the optimization problem (4.5) have

$$\frac{\partial \log(U_i(r_{1i} + ... + r_{Ki}))}{\partial r_{li}} = \frac{\frac{\partial U_i}{\partial r_{li}}}{U_i} > 0 \tag{4.6}$$

and

$$\frac{\partial^2 \log(U_i(r_{1i} + ... + r_{Ki}))}{\partial r_{li}^2} = \frac{\frac{\partial^2 U_i}{\partial r_{li}^2} U_i - (\frac{\partial U_i}{\partial r_{li}})^2}{U_i^2} < 0. \tag{4.7}$$

Therefore, the strictly concave utility function $U_i(r_{1i} + r_{2i} + ... + r_{Ki})$ natural logarithm $\log(U_i(r_{1i} + r_{2i} + ... + r_{Ki}))$ is also strictly concave. It follows that the natural logarithm of the logarithmic utility function in equation (4.2) is strictly concave.

In the sigmoidal-like utility function case, the utility function of the normalized sigmoidal-like function is given by equation (4.1) as $U_i(r_{1i}+r_{2i}+...+r_{Ki}) = c_i\left(\frac{1}{1+e^{-a_i(\sum_{l=1}^{K} r_{li} - b_i)}} - d_i\right)$.

For $0 < \sum_{l=1}^{K} r_{li} < \sum_{l=1}^{K} R_l$, we have

$$
\begin{aligned}
0 < c_i \left(\frac{1}{1 + e^{-a_i(\sum_{l=1}^{K} r_{li} - b_i)}} - d_i \right) < 1 \\
d_i < \frac{1}{1 + e^{-a_i(\sum_{l=1}^{K} r_{li} - b_i)}} < \frac{1 + c_i d_i}{c_i} \\
\frac{1}{d_i} > 1 + e^{-a_i(\sum_{l=1}^{K} r_{li} - b_i)} > \frac{c_i}{1 + c_i d_i} \\
0 < 1 - d_i(1 + e^{-a_i(\sum_{l=1}^{K} r_{li} - b_i)}) < \frac{1}{1 + c_i d_i}
\end{aligned}
$$

It follows that for $0 < \sum_{l=1}^{K} r_{li} < \sum_{l=1}^{K} R_l$, we have the first and second derivative as

$$
\begin{aligned}
\frac{\partial}{\partial r_{li}} \log U_i(r_{1i} + ... + r_{Ki}) =& \frac{a_i d_i e^{-a_i(\sum_{l=1}^{K} r_{li} - b_i)}}{1 - d_i(1 + e^{-a_i(\sum_{l=1}^{K} r_{li} - b_i)})} \\
&+ \frac{a_i e^{-a_i(\sum_{l=1}^{K} r_{li} - b_i)}}{(1 + e^{-a_i(\sum_{l=1}^{K} r_{li} - b_i)})} > 0 \\
\frac{\partial^2}{\partial r_{li}^2} \log U_i(r_{1i} + ... + r_{Ki}) =& \frac{-a_i^2 d_i e^{-a_i(\sum_{l=1}^{K} r_{li} - b_i)}}{c_i \left(1 - d_i(1 + e^{-a(\sum_{l=1}^{K} r_{li} - b_i)}) \right)^2} \\
&+ \frac{-a_i^2 e^{-a_i(\sum_{l=1}^{K} r_{li} - b_i)}}{(1 + e^{-a_i(\sum_{l=1}^{K} r_{li} - b_i)})^2} < 0
\end{aligned}
$$

Therefore, the sigmoidal-like utility function $U_i(r_{1i} + ... + r_{Ki})$ natural logarithm $\log(U_i(r_{1i} + ... + r_{Ki}))$ is strictly concave function. Therefore, all the utility functions in our model have strictly concave natural logarithm. □

Theorem 4.2.2. *The optimization problem (4.4) is a convex optimization problem and there exists a unique tractable global optimal solution.*

Proof. It follows from Lemma 4.2.1 that for all UEs utility functions are strictly concave. Therefore, the optimization problem (4.5) is a convex optimization problem [87]. The optimization problem (4.5) is equivalent to optimization problem (4.4), therefore it is a convex optimization problem. For a convex optimization problem, there exists a unique tractable

global optimal solution [87]. □

4.3 The Dual Problem

The key to a distributed and decentralized optimal solution of the primal problem in (4.5) is to convert it to the dual problem similar to [21], [15] and [88]. The optimization problem (4.5) can be divided into two simpler problems by using the dual problem. We define the Lagrangian

$$\begin{aligned} L(\mathbf{r},\mathbf{p}) &= \sum_{i=1}^{M} \log\left(U_i(r_{1i}+r_{2i}+...+r_{Ki})\right) \\ &\quad - p_1(\sum_{i=1}^{M} r_{1i} + z_1 - R_1) - ... \\ &\quad - p_K(\sum_{i=1}^{M} r_{Ki} + z_K - R_K) \\ &= \sum_{i=1}^{M}\left(\log(U_i(r_{1i}+r_{2i}+...+r_{Ki})) - \sum_{l=1}^{K} p_l r_{li}\right) \\ &\quad + \sum_{l=1}^{K} p_l(R_l - z_l) \\ &= \sum_{i=1}^{M} L_i(\mathbf{r}_i,\mathbf{p}) + \sum_{l=1}^{K} p_l(R_l - z_l) \end{aligned} \tag{4.8}$$

where $z_l \geq 0$ is the l^{th} slack variable and p_l is Lagrange multiplier or the shadow price of the l^{th} carrier eNodeB (i.e. the total price per unit rate for all the users in the coverage area of the l^{th} carrier eNodeB) and $\mathbf{p} = \{p_1, p_2, ..., p_K\}$. Therefore, the i^{th} UE bid for rate from the l^{th} carrier eNodeB can be written as $w_{li} = p_l r_{li}$ and we have $\sum_{i=1}^{M} w_{li} = p_l \sum_{i=1}^{M} r_{li}$. The first term in equation (4.8) is separable in $\mathbf{r}_i$. So we have $\max_{\mathbf{r}} \sum_{i=1}^{M}(\log(U_i(r_{1i}+r_{2i}+...+r_{Ki})) - \sum_{l=1}^{K} p_l r_{li}) = \sum_{i=1}^{M} \max_{\mathbf{r}_i}\left(\log(U_i(r_{1i}+r_{2i}+...+r_{Ki})) - \sum_{l=1}^{K} p_l r_{li}\right)$. The dual problem objective function

can be written as

$$\begin{aligned} D(\mathbf{p}) =& \max_{\mathbf{r}} L(\mathbf{r}, \mathbf{p}) \\ =& \sum_{i=1}^{M} \max_{\mathbf{r}_i} (L_i(\mathbf{r}_i, \mathbf{p})) + \sum_{l=1}^{K} p_l (R_l - z_l) \end{aligned} \tag{4.9}$$

The dual problem is given by

$$\begin{aligned} \min_{\mathbf{p}} \quad & D(\mathbf{p}) \\ \text{subject to} \quad & p_l \geq 0, \qquad l = 1, 2, ..., K. \end{aligned} \tag{4.10}$$

So we have

$$\frac{\partial D(\mathbf{p})}{\partial p_l} = R_l - \sum_{i=1}^{M} r_{li} - z_l = 0 \tag{4.11}$$

substituting by $\sum_{i=1}^{M} w_{li} = p_l \sum_{i=1}^{M} r_{li}$ we have

$$p_l = \frac{\sum_{i=1}^{M} w_{li}}{R_l - z_l}. \tag{4.12}$$

Now, we divide the primal problem (4.5) into two simpler optimization problems in the UEs and the eNodeBs. The i^{th} UE optimization problem is given by:

$$\begin{aligned} \max_{r_i} \quad & \log(U_i(r_{1i} + r_{2i} + ... + r_{Ki})) - \sum_{l=1}^{K} p_l r_{li} \\ \text{subject to} \quad & p_l \geq 0 \\ & r_{li} \geq 0, \qquad i = 1, 2, ..., M, l = 1, 2, ..., K. \end{aligned} \tag{4.13}$$

The second problem is the l^{th} eNodeB optimization problem for rate proportional fairness that is given by:

$$\begin{aligned} \min_{p_l} \quad & D(\mathbf{p}) \\ \text{subject to} \quad & p_l \geq 0. \end{aligned} \tag{4.14}$$

The minimization of shadow price p_l is achieved by the minimization of the slack variable

$z_l \geq 0$ from equation (4.12). Therefore, the maximum utility percentage of the l^{th} eNodeB rate R_l is achieved by setting the slack variable $z_l = 0$. In this case, we replace the inequality in primal problem (4.5) constraints by equality constraints and so we have $\sum_{i=1}^{M} w_{li} = p_l R_l$. Therefore, we have $p_l = \frac{\sum_{i=1}^{M} w_{li}}{R_l}$ where $w_{li} = p_l r_{li}$ is transmitted by the i^{th} UE to l^{th} eNodeB. The utility proportional fairness in the objective function of the optimization problem (4.4) is guaranteed in the solution of the optimization problems (4.13) and (4.14).

4.4 Distributed Optimization Algorithm

The distributed resource allocation algorithm, in [1], for optimization problems (4.13) and (4.14) is a modified version of the distributed algorithms in [21–23], [15] and [88], which is an iterative solution for allocating the network resources for a single carrier. The algorithm in [1] allocates resources from multiple carriers simultaneously with utility proportional fairness policy. The algorithm is divided into the i^{th} UE algorithm as shown in Algorithm 1 [1] and the l^{th} eNodeB carrier algorithm as shown in Algorithm 2 [1]. In Algorithm 1 and 2 [1], the i^{th} UE starts with an initial bid $w_{li}(1)$ which is transmitted to the l^{th} carrier eNodeB. The l^{th} eNodeB calculates the difference between the received bid $w_{li}(n)$ and the previously received bid $w_{li}(n-1)$ and exits if it is less than a pre-specified threshold δ. We set $w_{li}(0) = 0$. If the value is greater than the threshold, the l^{th} eNodeB calculates the shadow price $p_l(n) = \frac{\sum_{i=1}^{M} w_{li}(n)}{R_l}$ and sends that value to all UEs in its coverage area. The i^{th} UE receives the shadow prices p_l from all in range carriers eNodeBs and compares them to find the first minimum shadow price $p^1_{\min}(n)$ and the corresponding carrier index $l_1 \in L$ where $L = \{1, 2, ..., K\}$. The i^{th} UE solves for the l_1 carrier rate $r_{l_1 i}(n)$ that maximizes $\log U_i(r_{1i} + ... + r_{Ki}) - \sum_{l=1}^{K} p_l(n) r_{li}$ with respect to $r_{l_1 i}$. The rate $r^1_i(n) = r_{l_1 i}(n)$ is used to calculate the new bid $w_{l_1 i}(n) = p^1_{\min}(n) r^1_i(n)$. The i^{th} UE sends the value of its new bid $w_{l_1 i}(n)$ to the l_1 carrier eNodeB. Then, the i^{th} UE selects the second minimum shadow price $p^2_{\min}(n)$ and the corresponding carrier index $l_2 \in L$. The i^{th} UE solves for the l_2

carrier rate $r_{l_2 i}(n)$ that maximizes $\log U_i(r_{1i} + ... + r_{Ki}) - \sum_{l=1}^{K} p_l(n) r_{li}$ with respect to $r_{l_2 i}$. The rate $r_{l_2 i}(n)$ subtracted by the rate from l_1 carrier $r_i^2(n) = r_{l_2 i}(n) - r_i^1(n)$ is used to calculate the new bid $w_{l_2 i}(n) = p_{\min}^2(n) r_i^2(n)$ which is sent to l_2 carrier eNodeB. In general, the i^{th} UE selects the m^{th} minimum shadow price $p_{\min}^m(n)$ with carrier index $l_m \in L$ and solves for the l_m carrier rate $r_{l_m i}(n)$ that maximizes $\log U_i(r_{1i} + ... + r_{Ki}) - \sum_{l=1}^{K} p_l(n) r_{li}$ with respect to $r_{l_m i}$. The rate $r_{l_m i}(n)$ subtracted by $l_1, l_2, ..., l_{m-1}$ carriers rates $r_i^m(n) = r_{l_m i}(n) - (r_i^1(n) + r_i^2(n) + ... + r_i^{m-1}(n))$ is used to calculate the new bid $w_{l_m i}(n) = p_{\min}^m(n) r_i^m(n)$ which is sent to l_m carrier eNodeB. This process is repeated until $|w_{li}(n) - w_{li}(n-1)|$ is less than the threshold δ for all l carriers.

The distributed algorithm in [1] is set to avoid the situation of allocating zero rate to any user (i.e. no user is dropped). This is inherited from the utility proportional fairness policy in the optimization problem, similar to [21], [22] and [23]. In addition, the UE chooses from the nearby carriers eNodeBs the one with the lowest shadow price and starts requesting bandwidth from that carrier eNodeB. If the allocated rate is not enough or the price of the resources increases due to high demand on that carrier eNodeB resources from other UEs, the UE switches to another nearby eNodeB carrier with a lower resource price to be allocated the rest of the required resources. This is done iteratively until an equilibrium between demand and supply of resources is achieved and the optimal rates are allocated in the LTE mobile network. Figure 4.1 shows a block diagram that represents the distributed RA algorithm.

4.5 Convergence Analysis

In this section, we present the convergence analysis of Algorithm 1 and 2 in [1] for different values of carriers eNodeBs rates R_l. This analysis is equivalent to low and high-traffic hours analysis in cellular systems (e.g. change in the number of active users M and their traffic in

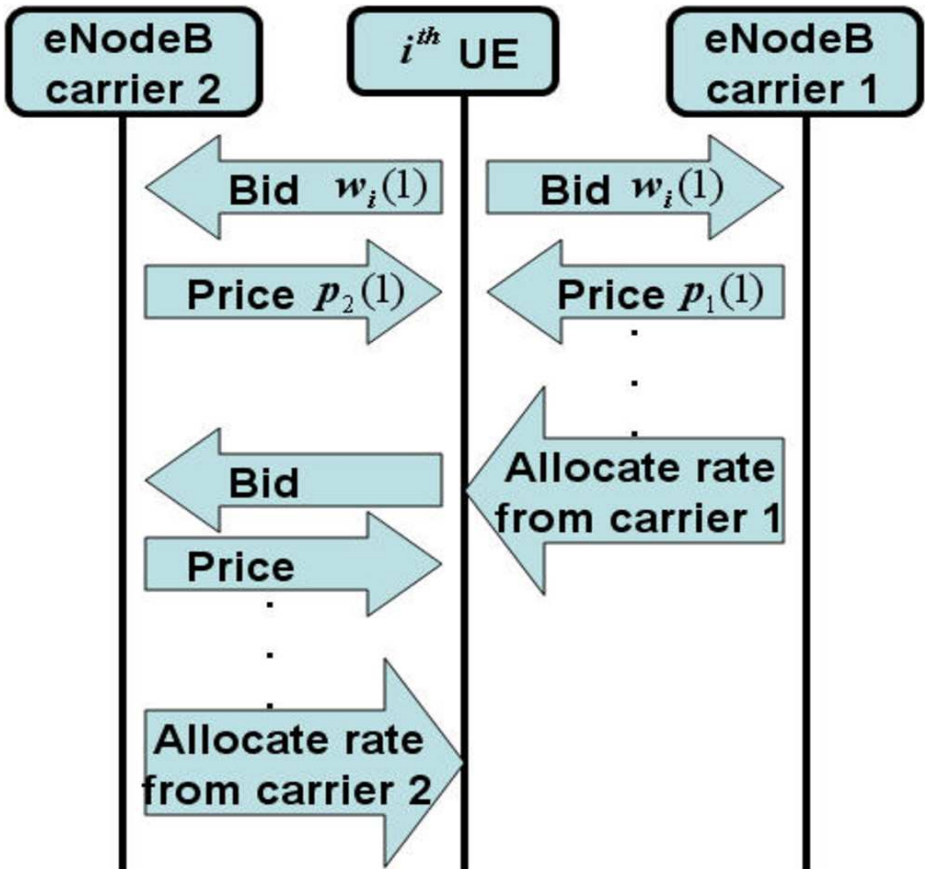


Figure 4.1: Flow Diagram with the assumption that the shadow price from the first carrier eNodeB p_1 is less before the n_1th iteration so rate r_{1i} of the i^{th} user is allocated. After the n_1th iteration, the shadow price from the second carrier eNodeB p_2 is less so rate r_{2i} is allocated.

the cellular system [22]).

4.5.1 Drawback in Algorithm 1 and 2 in [1]

Lemma 4.5.1. *For sigmoidal-like utility function $U_i(r_{1i}+r_{2i}+...+r_{Ki})$, the slope curvature function $\frac{\partial \log U_i(r_{1i}+r_{2i}+...+r_{Ki})}{\partial r_{li}}$ has an inflection point at $\sum_{l=1}^{K} r_{li} = r_i^s \approx b_i$ and is convex for $\sum_{l=1}^{K} r_{li} > r_i^s$.*

Proof. For the sigmoidal-like function $U_i(r_{1i} + r_{2i} + ... + r_{Ki}) = c_i\left(\frac{1}{1+e^{-a_i(\sum_{l=1}^{K} r_{li} - b_i)}} - d_i\right)$,

let $S_i(r_{li}) = \frac{\partial \log U_i(r_{1i}+r_{2i}+...+r_{Ki})}{\partial r_{li}}$ be the slope curvature function. Then, we have that

$$\frac{\partial S_i}{\partial r_{li}} = \frac{-a_i^2 d_i e^{-a_i(\sum_{l=1}^{K} r_{li}-b_i)}}{c_i\left(1-d_i(1+e^{-a_i(\sum_{l=1}^{K} r_{li}-b_i)})\right)^2} - \frac{a_i^2 e^{-a_i(\sum_{l=1}^{K} r_{li}-b_i)}}{\left(1+e^{-a_i(\sum_{l=1}^{K} r_{li}-b_i)}\right)^2}$$

and (4.15)

$$\frac{\partial^2 S_i}{\partial r_{li}^2} = \frac{a_i^3 d_i e^{-a_i(\sum_{l=1}^{K} r_{li}-b_i)}(1-d_i(1-e^{-a_i(\sum_{l=1}^{K} r_{li}-b_i)}))}{c_i\left(1-d_i(1+e^{-a_i(\sum_{l=1}^{K} r_{li}-b_i)})\right)^3} + \frac{a_i^3 e^{-a_i(\sum_{l=1}^{K} r_{li}-b_i)}(1-e^{-a_i(\sum_{l=1}^{K} r_{li}-b_i)})}{\left(1+e^{-a_i(\sum_{l=1}^{K} r_{li}-b_i)}\right)^3}.$$

We analyze the curvature of the slope of the natural logarithm of sigmoidal-like utility function. For the first derivative, we have $\frac{\partial S_i}{\partial r_{li}} < 0 \ \forall r_{li}$. The first term S_i^1 of $\frac{\partial^2 S_i}{\partial r_{li}^2}$ in equation (4.15) can be written as

$$S_i^1 = \frac{a_i^3 e^{a_i b_i}(e^{a_i b_i} + e^{-a_i(\sum_{l=1}^{K} r_{li}-b_i)})}{(e^{a_i b_i} - e^{-a_i(\sum_{l=1}^{K} r_{li}-b_i)})^3} \tag{4.16}$$

and we have the following properties:

$$\begin{cases} \lim_{\sum_{l=1}^{K} r_{li}\to 0} S_i^1 = \infty, \\ \lim_{\sum_{l=1}^{K} r_{li}\to b_i} S_i^1 = 0 \ \text{ for } \ b_i \gg \frac{1}{a_i}. \end{cases} \tag{4.17}$$

For second term S_i^2 of $\frac{\partial^2 S_i}{\partial r_i^2}$ in equation (4.15), we have the following properties:

$$\begin{cases} S_i^2(r_{li} = b_i - \sum_{j \neq l} r_{ji}) = 0, \\ S_i^2(r_{li} > b_i - \sum_{j \neq l} r_{ji}) > 0, \\ S_i^2(r_{li} < b_i - \sum_{j \neq l} r_{ji}) < 0. \end{cases} \tag{4.18}$$

From equation (4.17) and (4.18), S_i has an inflection point at $\sum_{l=1}^{K} r_{li} = r_i^s \approx b_i$. In addition, we have the curvature of S_i changes from a convex function close to origin to a concave function before the inflection point $\sum_{l=1}^{K} r_{li} = r_i^s$ then to a convex function after the inflection point. □

Our rate allocation approach guarantees non-zero rate allocation for all active users in the coverage area of a specific carrier eNodeB. We define the set $\mathcal{M}^l := \{i : r_{li} \neq 0\}$ to be the set of active users covered by the l^{th} eNodeB. Then, we have the following Corollary.

Corollary 4.5.2. *If $\sum_{i \in \mathcal{M}^l} r_i^{inf} \ll R_l \; \forall \, l \in L$ then Algorithm 1 and 2 in [1] converge to the global optimal rates which correspond to the steady state shadow price $p_{ss} < \frac{a_{i_{\max}} d_{i_{\max}}}{1 - d_{i_{\max}}} + \frac{a_{i_{\max}}}{2}$ where $i_{\max} = \arg\max_{i \in \mathcal{M}^l} b_i$.*

Proof. For the sigmoidal-like function $U_i(r_{1i} + r_{2i} + ... + r_{Ki}) = c_i \left(\frac{1}{1 + e^{-a_i(\sum_{l=1}^{K} r_{li} - b_i)}} - d_i \right)$, the optimal solution is achieved by solving the optimization problem (4.5). In Algorithm 1 [1], an important step to reach to the optimal solution is to solve the optimization problem $r_{li}(n) = \arg\max_{r_{li}} \left(\log U_i(r_{1i} + r_{2i} + ... + r_{Ki}) - p_l(n) r_{li} \right)$ for every UE in the l^{th} eNodeB coverage area. The solution of this problem can be written, using Lagrange multipliers method, in the form

$$\frac{\partial \log U_i(r_{1i} + r_{2i} + ... + r_{Ki})}{\partial r_{li}} - p_l = S_i(r_{li}) - p_l = 0. \tag{4.19}$$

From equation (4.17) and (4.18) in Lemma 4.5.1, we have the curvature of $S_i(r_{li})$ is convex

for $\sum_{l=1}^{K} r_{li} > r_i^s \approx b_i$. The algorithm in [1] is guaranteed to converge to the global optimal solution when the slope $S_i(r_{li})$ of all the utility functions natural logarithm $\log U_i(r_{1i} + r_{2i} + ... + r_{Ki})$ are in the convex region of the functions, similar to analysis of logarithmic functions in [15] and [88]. Therefore, the natural logarithm of sigmoidal-like functions $\log U_i(r_{1i} + r_{2i} + ... + r_{Ki})$ converge to the global optimal solution for $\sum_{l=1}^{K} r_{li} > r_i^s \approx b_i$. The inflection point of sigmoidal-like function $U_i(r_{1i}+r_{2i}+...+r_{Ki})$ is at $r_i^{\inf} = b_i$. For $\sum_{i\in\mathcal{M}^l} r_i^{\inf} \ll R_l$, the algorithm in [1] allocates rates $\sum_{l=1}^{K} r_{li} > b_i$ for all users. Since $S_i(r_{li})$ is convex for $\sum_{l=1}^{K} r_{li} > r_i^s \approx b_i$ then the optimal solution can be achieved by Algorithm 1 and 2 in [1]. We have from equation (4.19) and as $S_i(r_{li})$ is convex for $\sum_{l=1}^{K} r_{li} > r_i^s \approx b_i$, that $p_{ss} < S_i(\sum_{l=1}^{K} r_{li} = \max_{i\in\mathcal{M}^l} b_i)$ where $S_i(\sum_{l=1}^{K} r_{li} = \max_{i\in\mathcal{M}^l} b_i) = \frac{a_{i_{\max}} d_{i_{\max}}}{1-d_{i_{\max}}} + \frac{a_{i_{\max}}}{2}$ and $i_{\max} = \arg\max_{i\in\mathcal{M}^l} b_i$. □

We define the set $\mathcal{M}^{\mathcal{L}} := \{i : r_{li} \neq 0 \,\forall\, l \in \mathcal{L}, r_{li} = 0 \,\forall\, l \notin \mathcal{L}\}$ to be the set of active users covered exclusively by the set of carriers eNodeBs $\mathcal{L} \subseteq L$. Then, we have the following Corollary.

Corollary 4.5.3. *For $\sum_{i\in\mathcal{M}^{\mathcal{L}}} r_i^{inf} > \sum_{l\in\mathcal{L}} R_l$ and the global optimal shadow price $p_{ss} \approx \frac{a_i d_i e^{\frac{a_i b_i}{2}}}{1-d_i(1+e^{\frac{a_i b_i}{2}})} + \frac{a_i e^{\frac{a_i b_i}{2}}}{(1+e^{\frac{a_i b_i}{2}})}$ where $i \in \mathcal{M}^{\mathcal{L}}$, then the solution given by Algorithm 1 and 2 in [1] fluctuates about the global optimal rates.*

Proof. For the sigmoidal-like function $U_i(r_{1i} + r_{2i} + ... + r_{Ki}) = c_i\left(\frac{1}{1+e^{-a_i(\sum_{l=1}^{K} r_{li} - b_i)}} - d_i\right)$, it follows from lemma 4.5.1 that for $\sum_{i\in\mathcal{M}^{\mathcal{L}}} r_i^{\inf} > \sum_{l\in\mathcal{L}} R_l \;\exists\; i \in \mathcal{M}^{\mathcal{L}}$ such that the optimal rates $\sum_{l=1}^{K} r_{li}^{\text{opt}} < b_i$. Therefore, if $p_{ss} \approx \frac{a_i d_i e^{\frac{a_i b_i}{2}}}{1-d_i(1+e^{\frac{a_i b_i}{2}})} + \frac{a_i e^{\frac{a_i b_i}{2}}}{(1+e^{\frac{a_i b_i}{2}})}$ is the optimal shadow price for optimization problem (4.5). Then, a small change in the shadow price $p_l(n)$ in the n^{th} iteration can lead the rate $r_{li}(n)$ (root of $S_i(r_{li}) - p_l(n) = 0$) to fluctuate between the concave and convex curvature of the slope curve $S_i(r_{li})$ for the i^{th} user. Therefore, it causes fluctuation in the bid $w_{li}(n)$ sent to the eNodeB and fluctuation in the shadow price $p_l(n)$ set by eNodeB. Therefore, the iterative solution of Algorithm 1 and 2 in [1] fluctuates about the global optimal rates $\sum_{l=1}^{K} r_{li}^{\text{opt}}$. □

Algorithm 11 The i^{th} UE Algorithm

Send initial bid $w_{li}(1)$ to l^{th} carrier eNodeB (where $l \in L = \{1, 2, ..., K\}$)
loop
  Receive shadow prices $p_{l \in L}(n)$ from all in range carriers eNodeBs
  if STOP from all in range carriers eNodeBs **then**
    Calculate allocated rates $r_{li}^{\text{opt}} = \frac{w_{li}(n)}{p_l(n)}$
    STOP
  else
    Set $p_{\min}^0 = \{\}$ and $r_i^0 = 0$
    for $m = 1 \rightarrow K$ **do**
      $p_{\min}^m(n) = \min(\mathbf{p} \setminus \{p_{\min}^0, p_{\min}^1, ..., p_{\min}^{m-1}\})$
      $l_m = \{l \in L : p_l = \min(\mathbf{p} \setminus \{p_{\min}^0, p_{\min}^1, ..., p_{\min}^{m-1}\})\}$ {l_m is the index of the corresponding carrier}
      Solve $r_{l_m i}(n) = \arg\max_{r_{l_m i}} \Big(\log U_i(r_{1i} + ... + r_{Ki}) - \sum_{l=1}^{K} p_l(n) r_{li} \Big)$ for the l_m carrier eNodeB
      $r_i^m(n) = r_{l_m i}(n) - \sum_{j=0}^{m-1} r_i^j(n)$
      if $r_i^m(n) < 0$ **then**
        Set $r_i^m(n) = 0$
      end if
      Calculate new bid $w_{l_m i}(n) = p_{\min}^m(n) r_i^m(n)$
      if $|w_{l_m i}(n) - w_{l_m i}(n-1)| > \Delta w(n)$ **then**
        $w_{l_m i}(n) = w_i(n-1) + \text{sign}(w_{l_m i}(n) - w_{l_m i}(n-1)) \Delta w(n)$ {$\Delta w = h_1 e^{-\frac{n}{h_2}}$ or $\Delta w = \frac{h_3}{n}$}
      end if
      Send new bid $w_{l_m i}(n)$ to l_m carrier eNodeB
    end for
  end if
end loop

Theorem 4.5.4. *Algorithm 1 and 2 in [1] does not converge to the global optimal rates for all values of R_l.*

Proof. It follows from Corollary 4.5.2 and 4.5.3 that Algorithm 1 and 2 in [1] does not converge to the global optimal rates for all values of R_l. □

4.5.2 Solution using Algorithm 11 and 12

For a robust algorithm, we add a fluctuation decay function to the algorithm presented in [1] as shown in Algorithm 11. Our robust algorithm ensures convergence for all values of the carriers eNodeBs maximum rate R_l for all l. Algorithm 11 and 12 allocated rates coincide with Algorithm 1 and 2 in [1] for $\sum_{i\in\mathcal{M}^l} r_i^{\inf} \ll R_l \ \forall \ l \in L$. For $\sum_{i\in\mathcal{M}^{\mathcal{L}}} r_i^{\inf} > \sum_{l\in\mathcal{L}} R_l$, robust algorithm avoids the fluctuation in the non-convergent region discussed in the previous section. This is achieved by adding a convergence measure $\Delta w(n)$ that senses the fluctuation in the bids w_{li}. In case of fluctuation, it decreases the step size between the current and the previous bid $w_{li}(n) - w_{li}(n-1)$ for every user i using *fluctuation decay function*. The fluctuation decay function could be in the following forms:

- *Exponential function*: It takes the form $\Delta w(n) = h_1 e^{-\frac{n}{h_2}}$.
- *Rational function*: It takes the form $\Delta w(n) = \frac{h_3}{n}$.

where h_1, h_2, h_3 can be adjusted to change the rate of decay of the bids w_{li}.

Remark 4.5.5. *The fluctuation decay function can be included in the UE or the eNodeB Algorithm.*

In our model, we add the decay part to the UE Algorithm as shown in Algorithm 11.

4.6 Simulation Results

Algorithm 11 and 12 were applied to various logarithmic and sigmoidal-like utility functions with different parameters in MATLAB. The simulation results showed convergence to the global optimal rates. In this section, we present the simulation results for two carriers in a heterogeneous network (HetNet) that consists of one macro cell, one small cell and 12

Algorithm 12 The l^{th} eNodeB Algorithm

loop
  Receive bids $w_{li}(n)$ from UEs {Let $w_{li}(0) = 0 \ \forall i$}
  if $|w_{li}(n) - w_{li}(n-1)| < \delta \ \forall i$ **then**
    Allocate rates, $r_{li}^{\text{opt}} = \frac{w_{li}(n)}{p_l(n)}$ to i^{th} UE
    STOP
  else
    Calculate $p_l(n) = \frac{\sum_{i=1}^{M} w_{li}(n)}{R_l}$
    Send new shadow price $p_l(n)$ to all UEs
  end if
end loop

active UEs as shown in Figure 4.2. The UEs are divided into two groups. The 1^{st} group of UEs (index $i = \{1, 2, 3, 4, 5, 6\}$) is located in the macro cell under the coverage area of both the 1^{st} carrier (C1) and the 2^{nd} carrier (C2) eNodeBs. We use three normalized sigmoidal-like functions that are expressed by equation (4.1) with different parameters. The used parameters are $a = 5$, $b = 10$ corresponding to a sigmoidal-like function that is an approximation to a step function at rate $r = 10$ (e.g. VoIP) and is the utility of UEs with indexes $i = \{1, 7\}$, $a = 3$, $b = 20$ corresponding to a sigmoidal-like function that is an approximation of an adaptive real-time application with inflection point at rate $r = 20$ (e.g. standard definition video streaming) and is the utility of UEs with indexes $i = \{2, 8\}$, and $a = 1$, $b = 30$ corresponding to a sigmoidal-like function that is also an approximation of an adaptive real-time application with inflection point at rate $r = 30$ (e.g. high definition video streaming) and is the utility of UEs with indexes $i = \{3, 9\}$, as shown in Figure 4.3. We use three logarithmic functions that are expressed by equation (4.2) with $r_{max} = 100$ and different k_i parameters which are approximations for delay-tolerant applications (e.g. FTP). We use $k = 15$ for UEs with indexes $i = \{4, 10\}$, $k = 3$ for UEs with indexes $i = \{5, 11\}$, and $k = 0.5$ for UEs with indexes $i = \{6, 12\}$, as shown in Figure 4.3. A summary is shown in table 4.1. A three dimensional view of the sigmoidal-like utility function $U_i(r_{1i} + r_{2i})$ is show in Figure 4.4.

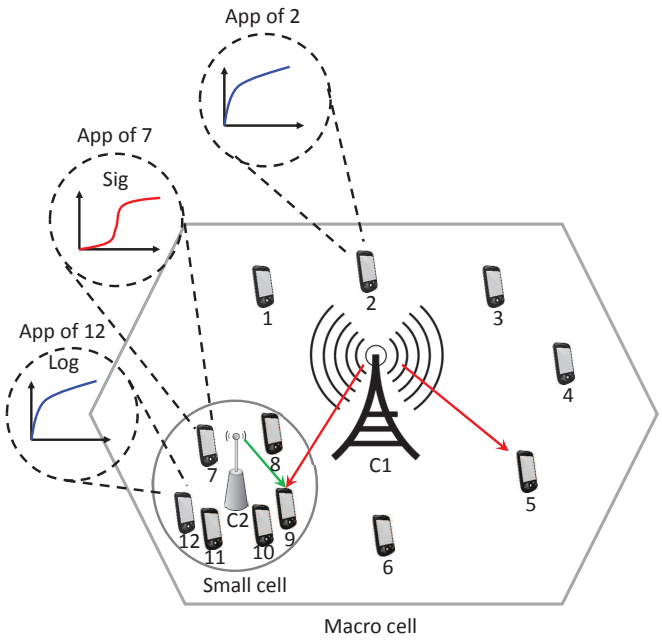


Figure 4.2: System model with two groups of users. The 1^{st} group with UE indexes $i = \{1, 2, 3, 4, 5, 6\}$, 2^{nd} group with UE indexes $i = \{7, 8, 9, 10, 11, 12\}$.

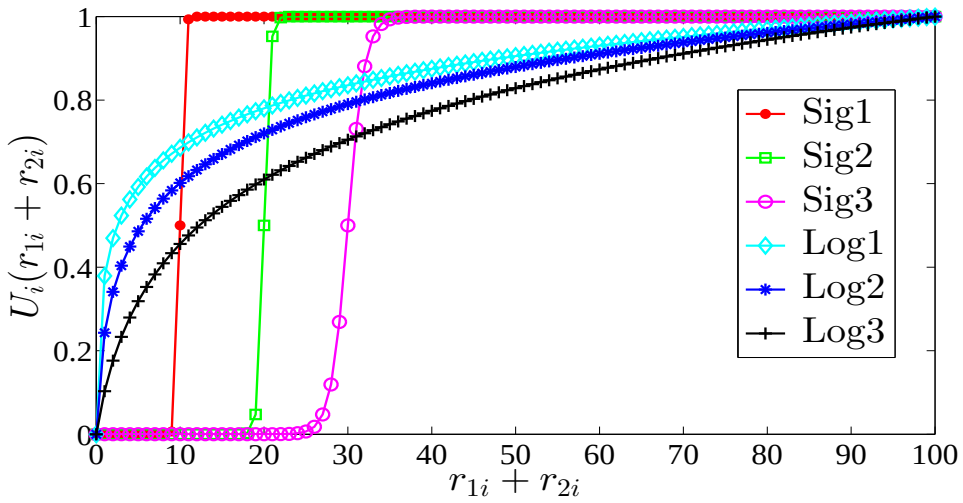


Figure 4.3: The users utility functions $U_i(r_{1i}+r_{2i})$ used in the simulation (three sigmoidal-like functions and three logarithmic functions).

4.6.1 Allocated Rates for $30 \leq R_1 \leq 200$ and $R_2 = 70$

In the following simulations, we set $\delta = 10^{-3}$, the 1^{st} carrier eNodeB rate R_1 takes values between 30 and 200 with step of 10, and the 2^{nd} carrier eNodeB rate is fixed at $R_2 = 70$.

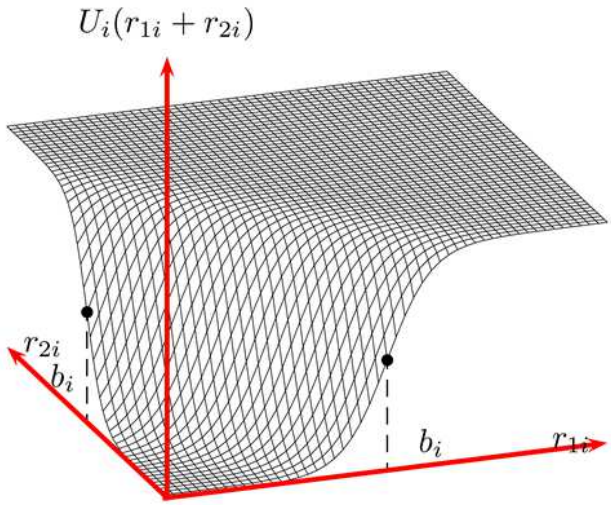


Figure 4.4: The sigmoidal-like utility $U_i(r_{1i}+r_{2i}) = c_i(\frac{1}{1+e^{-a_i(r_{1i}+r_{2i}-b_i)}} - d_i)$ of the i^{th} user, where r_{1i} is the rate allocated by 1^{st} carrier eNodeB and r_{2i} is the rate allocated by 2^{nd} carrier eNodeB.

Table 4.1: Users and their applications utilities

Applications Utilities Parameters		Users Indexes
Sig1	Sig $a = 5,\ b = 10$	$i = \{1, 7\}$
Sig2	Sig $a = 3,\ b = 20$	$i = \{2, 8\}$
Sig3	Sig $a = 1,\ b = 30$	$i = \{3, 9\}$
Log1	Log $k = 15,\ r_{max} = 100$	$i = \{4, 10\}$
Log2	Log $k = 3,\ r_{max} = 100$	$i = \{5, 11\}$
Log3	Log $k = 0.5,\ r_{max} = 100$	$i = \{6, 12\}$

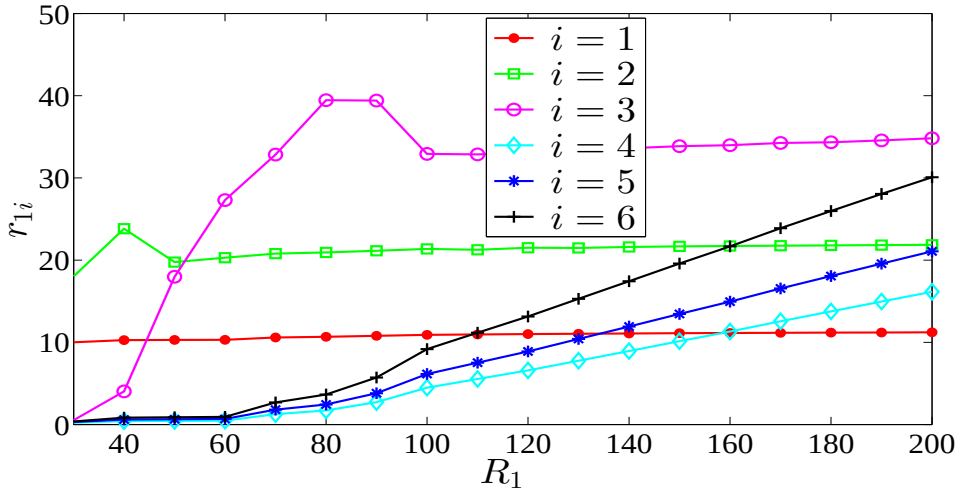
In Figure 4.5, we show the final allocated optimal rates $r_i = r_{1i} + r_{2i}$ of different users with different 1^{st} carrier eNodeB total rate R_1 and observe how the proposed rate allocation algorithm converges when the eNodeBs available resources are abundant or scarce. In Figure 4.5(a), we show the rates allocated to the 1^{st} group of UEs by only C1 eNodeB since C2 eNodeB is not within these users range, we observe the increase in the rate allocated to these users with the increase in R_1. Figure 4.5(b) shows the final allocated rates to the 2^{nd} group of UEs by both C1 and C2 eNodeBs. Since these users located under the coverage

area of both the macro cell and the small cell, they are allocated rates jointly using the proposed RA with joint CA approach. Figure 4.5(a) and 4.5(b) show that by using the RA with joint CA algorithm, no user is allocated zero rate (i.e. no user is dropped). However, the majority of the eNodeBs resources are allocated to the UEs running adaptive real-time applications until they reach their inflection rates the eNodeBs then allocate more resources to the UEs with delay-tolerant applications, as real-time application users bid higher than delay-tolerant application users by using the utility proportional fairness policy.

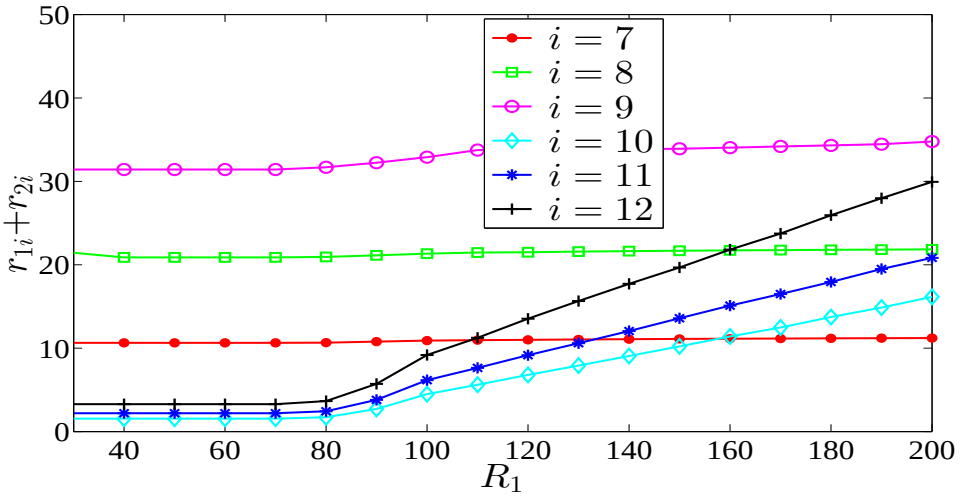
In Figure 4.6, we show the rates allocated to the 2^{nd} group users, located under the coverage area of both the macro cell and small cell eNodeBs, by each of the two carriers' eNodeBs with the increase in the 1^{st} carrier eNodeB resources. In Figure 4.6(a) and 4.6(b), when the resources available at C2 eNodeB (i.e. R_2) is more than that at C1 eNodeB, we observe that most of the 2^{nd} group rates are allocated by C2 eNodeB. However, the delay tolerant applications are not allocated much resources since most of R_2 is allocated to the real-time applications. With the increase in C1 eNodeB resources R_1, we observe a gradual increase in the 2^{nd} group rates allocated to real-time applications from C1 eNodeB and a gradual decrease from C2 eNodeB resources allocated to real-time-applications. This shift in the resource allocation increases the available resources in C2 eNodeB to be allocated to 2^{nd} group delay tolerant applications by C2 eNodeB.

4.6.2 Pricing Analysis and Comparison for $30 \leq R_1 \leq 200$ and $R_2 = 70$

In the following simulations, we set $\delta = 10^{-3}$ and the 1^{st} carrier eNodeB rate R_1 takes values between 30 and 200 with step of 10, and C2 eNodeB total rate is fixed at $R_2 = 70$. As discussed before, the users' allocated rates are proportional to the users' bids. Real-time application users bid higher than delay-tolerant application users due to their applications



(a) The rates allocated r_{1i} from the 1^{st} carrier eNodeB (i.e. the macro cell eNodeB) to users of the 1^{st} group (i.e. $i = 1, 2, 3, 4, 5, 6$).

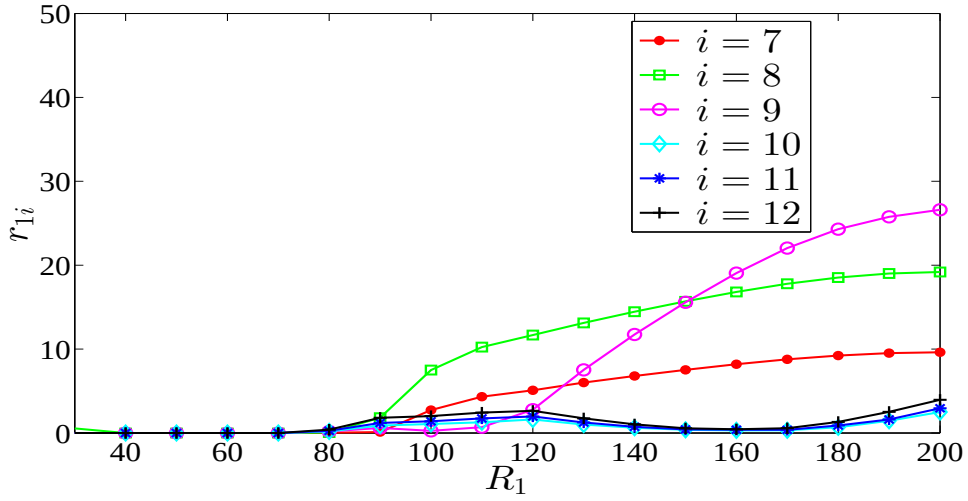


(b) The rates $r_{1i} + r_{2i}$ allocated from 1^{st} and 2^{nd} carriers eNodeBs (i.e. the macro cell and the small cell eNodeBs) to users of the 2^{nd} group (i.e. $i = 7, 8, 9, 10, 11, 12$).

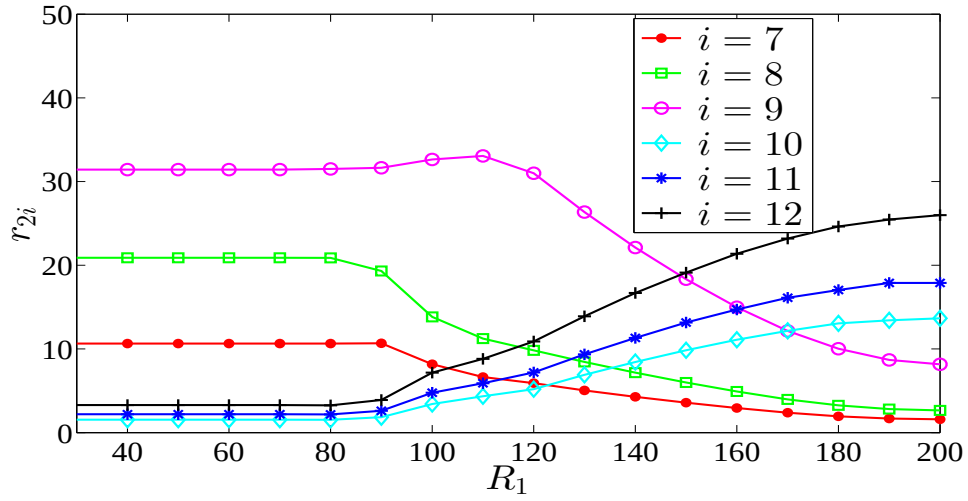
Figure 4.5: The allocated rates $\sum_{l=1}^{K} r_{li}$ of the two groups of users verses 1^{st} carrier rate $30 < R_1 < 200$ with 2^{nd} carrier rate fixed at $R_2 = 70$.

nature and the utility proportional fairness policy. Therefore, the pricing which is proportional to the bids is traffic-dependent, i.e. when the demand by users increases, as a result the price increases and vice versa.

In Figure 4.7, we compare between the shadow price of C1 and C2 eNodeBs when using the proposed RA with joint CA approach with their shadow prices obtained when using



(a) The allocated rates r_{1i} from the 1^{st} carrier eNodeB to the 2^{nd} group of users.



(b) The allocated rates r_{2i} from the 2^{nd} carrier eNodeB to the 2^{nd} group of users.

Figure 4.6: The allocated rates from C1 and C2 eNodeBs to the 2^{nd} group of users with 1^{st} carrier eNodeB rate $30 < R_1 < 200$ and 2^{nd} carrier eNodeB rate fixed at $R_2 = 70$.

the multi-stage RA with CA approach in [24, 85, 86]. For the RA with joint CA case, we observe that the shadow price of C1 eNodeB is higher than that of C2 eNodeB for $R_1 < 80$ and approximately equal for $80 \leq R_1 \leq 200$ which shows how it is very efficient to use the joint CA approach for the pricing of the user. We also show how the prices decrease with the increase in the eNodeBs total rate. By using this traffic-dependent pricing, the network providers can flatten the traffic specially during peak hours by setting traffic-dependent resource price, which gives an incentive for users to use the network during less traffic hours.

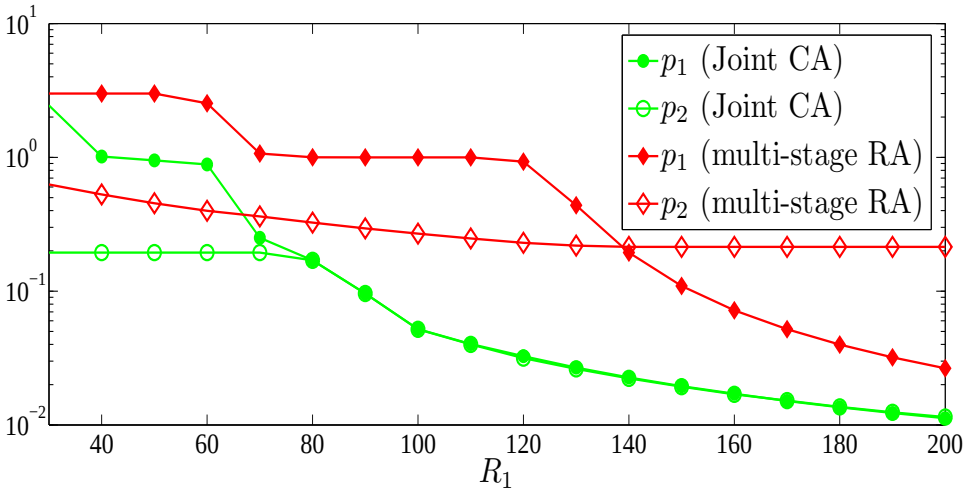


Figure 4.7: The 1^{st} carrier shadow price p_1 and 2^{nd} carrier shadow price p_2 for both multi-stage RA with CA and joint RA methods with C1 eNodeB rate $30 < R_1 < 200$ and C2 eNodeB rate $R_2 = 70$.

On the other hand, for the multi-stage RA with CA approach, we show in Figure 4.7 the changes in C1 and C2 eNodeBs shadow prices with R_1. When using the multi-stage RA with CA approach, all users are first allocated rates by the macro cell eNodeB, once C1 eNodeB is done allocating its resources C2 eNodeB starts allocating its resources only to the 2^{nd} group users as they are located within its coverage area. Since the pricing method in multi-stage RA with CA approach is not optimal, this explains why the shadow prices of C1 and C2 eNodeBs, in Figure 4.7, when using the proposed RA with joint CA approach are less than their corresponding prices when using the multi-stage RA with CA approach. This shows how the proposed algorithm outperforms the algorithms presented in [24, 85, 86] as it guarantees that mobile users receive optimal price (minimum) for resources.

4.7 Summary and Conclusions

In this chapter, we introduced a novel resource allocation optimization problem with joint carrier aggregation in cellular networks. We considered mobile users running real-time and delay-tolerant applications with utility proportional fairness allocation policy. We proved

that the global optimal solution exists and is tractable for mobile stations with logarithmic and sigmoidal-like utility functions. We presented a novel robust distributed algorithm for allocating resources from different carriers optimally among the mobile users. Our algorithm ensures fairness in the utility percentage achieved by the allocated resources for all users. Therefore, the algorithm gives priority to users with adaptive real-time applications while providing a minimum QoS for all users. In addition, the proposed RA with joint CA algorithm guarantees allocating resources from different carriers with the lowest resource price for the user. We analyzed the convergence of the algorithm with different network traffic densities and presented a robust algorithm that overcomes the fluctuation in allocation during peak traffic hours. We showed through simulations that our algorithm converges to the optimal resource allocation and that the proposed algorithm outperforms the multi-stage RA with CA algorithms presented in [24, 85, 86] as it guarantees that mobile users receive optimal price for the allocated resources.

Chapter 5

Resource Allocation with User Discrimination for Spectrum Sharing

In this chapter, we focus on the problem of radio resource allocation with user discrimination for different scenarios in cellular networks. First, we present a resource allocation with user discrimination approach between public safety and commercial users. It is important to have a common technical standard for commercial and public safety users as it provides advantages for both. The public safety systems market is much smaller than the commercial cellular market which makes it unable to attract the level of investment that goes in to commercial cellular networks and this makes a common technical standards for both the best solution. The public safety community gains access to the technical advantages provided by the commercial cellular networks whereas the commercial cellular community gains enhancement in their systems and makes it more attractive to consumers. The National Public Safety Telecommunications Council (NPSTC) and other organizations recognized the desirability of having an inter operable national standard for a next generation public safety network with broadband capabilities. The USA has reserved spectrum in the 700MHz band for an LTE based public safety network. The current public safety standards support medium speed

data which drives the need of new technology to add true mobile broadband capabilities and makes LTE the baseline technology for next generation broadband public safety networks.

Then, we provide a resource allocation with user discrimination optimization framework in cellular networks for different types of users running multiple applications simultaneously. Mobile users are now running multiple applications simultaneously on their smart phones. Operators are moving from single-service to multi-service and new services such as multimedia telephony and mobile-TV are now provided. In addition, different users subscribing for the same service may receive different treatment from the network providers [3] because of the subscriber differentiation provided by the service providers.

In addition, we present an efficient resource allocation with user discrimination framework for 5G Wireless Systems to allocate multiple carriers resources among users with elastic and inelastic traffic. As 5G systems' expected capabilities have started to take shape, CA is expected to be supported by 5G. Therefore, CA needs to be taken into consideration when designing 5G systems. Beside CA capability, 5G wireless network promises to handle diverse QoS requirements of multiple applications since different applications require different application's performance. Furthermore, certain types of users may require to be given priority when allocating the network resources (i.e. such as public safety users) which needs to be taken into consideration when designing the resource allocation framework.

5.1 Spectrum Sharing between Public Safety and Commercial Users in Cellular Networks

In this section, we propose a spectrum sharing approach between two groups of users, public safety and commercial users. We focus on finding an optimal solution for the resource allocation problem for the two groups of users running applications that are presented by

logarithmic utility functions or sigmoidal-like utility functions. These utility functions are concave and non-concave utility functions, respectively. The optimization problem allocates part of the bandwidth from one eNodeB to each user subscribing for a mobile service taking into consideration that each user is getting a minimum QoS. In addition, the public safety users in emergency mode are given priority over the commercial users and within each group the non concave functions that are approximated by sigmoidal-like functions and presenting real-time applications are given priority over the concave functions approximated by logarithmic functions and presenting delay tolerant applications. In our system model, each public safety subscriber has an assigned application target rate that varies based on the application type and assigned to the public safety subscriber by the network.

Our resource allocation algorithm first allocates the application target rate to each public safety UE when that UE is in emergency mode. It then allocates the remaining resources among the commercial UEs subscribing for resources.

Our contributions in this section are summarized as:

- We present a resource allocation optimization problem to allocate the eNodeB resources optimally among public safety and commercial users. The eNodeB and the UE collaborate to allocate an optimal rate to each UE with priority given to public safety users. Within the same group of users, a priority is given to real time applications presented by sigmoidal-like utility functions.
- We show that each of our two cases resource allocation (RA) optimization problems has a unique tractable global optimal solution.

5.1.1 Problem Formulation

We consider a single cell 4G-LTE mobile system with a single eNodeB, N commercial UEs and M public safety UEs. The user i is allocated certain bandwidth r_i based on the type of

application the UE is running. Each user is assigned a utility function $U_i(r_i)$ based on the application running on the UE and whether it is a commercial or public safety user. Our goal is to determine the optimal bandwidth that needs to be allocated to each user by the eNodeB.

Utility functions $U_i(r_i)$ are used to represent the applications running on the UEs. Logarithmic utility functions expressed by equation (2.2) and sigmoidal-like utility functions expressed by equation (2.1) are used to represent delay tolerant and real-time applications, respectively. The basic formulation of the resource allocation problem is given by the following optimization problem:

$$\begin{aligned} \max_{\mathbf{r}} \quad & \prod_{i=1}^{M} U_i(r_{i,s}) \prod_{j=1}^{N} U_j(r_{j,c}) \\ \text{subject to} \quad & \sum_{i=1}^{M} r_{i,s} + \sum_{j=1}^{N} r_{j,c} \leq R, \\ & r_{i,s} \geq r_{i,s}^{t}, \qquad i = 1, 2, ..., M \\ & r_{j,c} \geq 0, \qquad j = 1, 2, ..., N. \end{aligned} \tag{5.1}$$

where R is the maximum achievable rate of the eNodeB, $\mathbf{r} = \{r_{1,s}, ..., r_{M,s}, r_{1,c}, ..., r_{N,c}\}$ where $r_{i,s}$ is the rate for public safety user i, $r_{j,c}$ is the rate for commercial user j, $r_{i,s}^{t}$ is the application target rate for public safety user i which is the mnimum rate that the user wants to achieve, M and N are the numbers of the public safety and commercial UEs, respectively. The resource allocation objective function maximizes the product of users utilities system utility when allocating resources to each user. Therefore, it provides a proportional fairness among utilities. Public safety users that are running real-time applications are given the priority when allocating resources by the eNodeB. The next priority is given to the elastic traffic running by public safety users. Once each public safety user satisfies its application target rate the eNodeB starts allocating resources to commercial users giving priority to users running real time applications. We assume that the public safety users are in an

emergency mode, therefore these users are given a higher priority over the commercial users. The optimization problem (5.1) has a unique tractable global optimal solution [21] that will be discussed in the next section.

We used utility proportional fairness model because non-zero rate allocation is guaranteed to all users. So it is impossible to set a users allocation to zero without setting the efficiency of the network to zero. Because this resource allocation strategy does not disenfranchise any given user, it will be considered as an appropriate fairness model for this problem.

5.1.2 Resource Allocation Optimization Problem

The resource allocation for public safety and commercial users is divided into two cases. The first case is when the maximum available resources R for the eNodeB is less than the sum of the total application target rates of the public safety UEs subscribing for a service from that eNodeB and the second case is when R is greater than that total. The two cases are two different optimization problems that will be solved by our proposed algorithm to obtain the optimal rate for each UE.

5.1.2.1 The First Case RA Optimization Problem when $\sum_{i=1}^{M} r_{i,s}^{t} \geq R$

As mentioned before the first case optimization problem is applied in the case of $\sum_{i=1}^{M} r_{i,s}^{t} \geq R$. In this case the eNodeB only allocates resources to the public safety users because they are considered more important and the eNodeB's available resources doesn't exceed their need. The commercial users will not be given any of the eNodeB resources in this case. This optimization problem can be written as:

$$\begin{aligned} \max_{\mathbf{r}} \quad & \prod_{i=1}^{M} U_i(r_{i,s}) \\ \text{subject to} \quad & \sum_{i=1}^{M} r_{i,s} \leq R, \\ & 0 \leq r_{i,s} \leq r_{i,s}^{t}, \qquad i = 1, 2, ..., M. \end{aligned} \tag{5.2}$$

where U_i is the public safety i^{th} utility function and $\mathbf{r} = \{r_{1,s}, ..., r_{M,s}\}$ and M is the number of public safety UEs in the coverage area of the eNodeB. The solution of the optimization problem (5.2) is the optimal solution when $\sum_{i=1}^{M} r_{i,s}^{t} \geq R$. This solution will guarantee that the public safety users are given priority when allocating the eNodeB resources. The optimal rate for each public safety UE is less than or equal to the application target rate for each public safety UE. The public safety users running real time applications will be given priority over public safety users with elastic traffic.

The objective function in the optimization problem (5.2) is equivalent to $\max_{\mathbf{r}} \sum_{i=1}^{M} \log U_i(r_{i,s})$. The optimization problem (5.2) is a convex optimization problem and there exists a unique tractable global optimal solution as shown in Theorem (III.1) [21]. This optimal solution gives each of the M users an optimal rate $r_{i,s}^{\text{opt}}$.

5.1.2.2 The Second Case RA Optimization Problem when $\sum_{i=1}^{M} r_{i,s}^{t} < R$

The second case optimization problem is applied in the case of $\sum_{i=1}^{M} r_{i,s}^{t} < R$. The eNodeB collaborate with the UEs to solve this optimization problem. The eNodeB allocates resources to both public safety and commercial users because its available resources exceed the minimum need of the public safety UEs expressed by the application target rates. As mentioned before, the eNodeB gives priority to the public safety users and within the public safety group the priority is given to the UEs running inelastic traffic. This optimization problem can be written as:

$$\begin{aligned} \max_{\mathbf{r}} \quad & \prod_{i=1}^{M} U_i(r_{i,s}) \prod_{j=1}^{N} U_j(r_{j,c}) \\ \text{subject to} \quad & \sum_{i=1}^{M} r_{i,s} + \sum_{j=1}^{N} r_{j,c} \leq R, \\ & r_{i,s} \geq r_{i,s}^{t}, \qquad i = 1, 2, ..., M \\ & r_{j,c} \geq 0, \qquad j = 1, 2, ..., N. \end{aligned} \tag{5.3}$$

This optimization problem is same as the one discussed in the problem formulation (section 5.1.1). First, the eNodeB allocates the application target rate to each public safety UE. It then starts allocating its remaining resources both to the public safety and commercial UEs based on utility proportional fairness. The solution of the optimization problem (5.3) is the global optimal solution that gives an optimal rate $r_{i,s}^{\text{opt}}$ to each public safety UE and an optimal rate $r_{i,c}^{\text{opt}}$ to each commercial user UE.

Proposition 5.1.1. *The optimization problem (5.3) is a convex optimization problem and there exists a unique tractable global optimal solution.*

Proof. We introduce a new parameter c_i where c_i is the application target rate for the public safety UE whereas it is 0 for the commercial UE, the optimization problem (5.3) can be rewritten as follows:

$$\begin{aligned} \max_{\mathbf{r}} \quad & \prod_{i=1}^{M+N} U_i(r_i + c_i) \\ \text{subject to} \quad & \sum_{i=1}^{M+N} (r_i + c_i) \leq R, \\ & r_i \geq 0, \qquad i = 1, 2, ..., M + N. \end{aligned} \tag{5.4}$$

$$c_i = \begin{cases} r_{i,s}^{t} & \text{if public safety UE} \\ 0 & \text{if commercial UE} \end{cases}$$

where R is the maximum achievable rate of the eNodeB, $\mathbf{r} = \{r_1, ..., r_M, r_{M+1}, ..., r_{M+N}\}$

where the first M rates are for the M public safety users and the last N rates are for the N commercial users, $U_i(r_i + c_i)$ is the UE utility function, this optimization problem guarantees an optimal rate that is at least equal to the application target rate for the public safety UE. The objective function in the optimization problem (5.4) can be written as $\sum_{i=1}^{M+N} \log U_i(r_i + c_i)$.

The utility function $U_i(r_i + c_i)$ for the UE is strictly concave or sigmoidal-like function as mentioned in section 5.1.1. As shown in Theorem (III.1) [21], $\log U_i(r_i)$ is a strictly concave function for a strictly concave or sigmoidal-like utility function. It follows that the optimization problem 5.4 that is equivalent to (5.3) is convex. Therefore, there exists a tractable global optimal solution for the optimization problem (5.3). □

5.1.3 Algorithm

In our proposed iterative algorithm, the eNodeB and the UEs collaborate to allocate optimal rates for the public safety and commercial users subscribing for a mobile service. Algorithm 1 and algorithm 2 are the public safety UE and the commercial UE algorithms, respectively. Algorithm 3 is the eNodeB algorithm. The algorithm starts when each UE transmits an initial bid $w_i(1)$ to the eNodeB. Additionally, each public safety UE transmits its application target rate to the eNodeB. The eNodeB checks whether the $\sum_{i=1}^{M} r_{i,s}^t$ is less or greater than R and send a flag with this information to each UE. In the case of $\sum_{i=1}^{M} r_{i,s}^t \geq R$, the commercial UEs will not be allocated any of the resources and will not be sending any further bids to the eNodeB unless they receive a flag from the eNodeB with $\sum_{i=1}^{M} r_{i,s}^t < R$.

On the other hand, each public safety UE checks whether the difference between the current received bid and the previous one is less than a threshold δ, if so it exits. Otherwise, if the difference is greater than δ, eNodeB calculates the shadow price $p(n) = \frac{\sum_{i=1}^{M} w_i(n)}{R}$. The estimated $p(n)$ is then sent to the public safety UEs where it is used to calculate the

rate $r_{i,s}(n)$ which is the solution of the optimization problem $r_{i,s}(n) = \arg\max_{r_{i,s}}(\log U_i(r_{i,s}) - p(n)r_{i,s})$. A new bid $w_i(n)$ is calculated using $r_i(n)$ where $w_i(n) = p(n)r_{i,s}(n)$. All public safety UEs send their new bids $w_i(n)$ to the eNodeB. The Algorithm is finalized by the eNodeB. Each public safety UE then calculates its allocated rate $r_{i,s}^{\text{opt}} = \frac{w_i(n)}{p(n)}$.

In the case of $\sum_{i=1}^{M} r_{i,s}^{t} < R$, the eNodeB sends a flag with this information to each UE. Each public safety and commercial UE checks whether the difference between the current received bid and the previous one is less than a threshold δ, if so it exits. Otherwise, if the difference is greater than δ, eNodeB calculates the shadow price $p(n) = \frac{\sum_{i=1}^{M+N} w_i(n)}{R}$. The estimated $p(n)$ is then sent to the public safety and commercial UEs where it is used by the public safety UE to calculate the rate $r_{i,s}(n) = r_i + r_{i,s}^{\text{t}}$ which is the solution of the optimization problem $r_{i,s}(n) = \arg\max_{r_{i,s}}(\log U_i(r_i + c_i) - p(n)(r_i + c_i))$. A new bid $w_i(n)$ is calculated by the public safety UE using $r_i(n)$ where $w_i(n) = p(n)(r_i(n) + c_i)$. All public safety UEs send their new bids $w_i(n)$ to the eNodeB. On the other hand, the commercial UEs receive $p(n)$ and use it to calculate the rate $r_{i,c}(n)$ which is the solution of the optimization problem $r_{i,c}(n) = \arg\max_{r_{i,c}}(\log U_i(r_{i,c}) - p(n)r_{i,c})$. A new bid $w_i(n)$ is calculated by the commercial UE using $r_{i,c}(n)$ where $w_i(n) = p(n)r_i(n)$. All public safety UEs send their new bids $w_i(n)$ to the eNodeB. The Algorithm is finalized by the eNodeB. Each public safety UE then calculates its allocated rate $r_{i,s}^{\text{opt}} = \frac{w_i(n)}{p(n)}$ and each commercial UE calculates its allocated rate $r_{i,c}^{\text{opt}} = \frac{w_i(n)}{p(n)}$.

5.1.4 Simulation Results

We consider one eNodeB with four public safety UEs and another four commercial UEs in its coverage area. We use multiple sigmoidal-like and logarithmic utility functions in our simulations and present two cases, one when the eNodeB resources R is less than the total application target rates of the public safety UEs and the other when R is greater than that total. We applied algorithm 1, 2 and 3 in C++ to the sigmoidal-like and logarithmic utility functions. The simulation results showed convergence to the optimal global point in both

Algorithm 13 Public Safety UE Algorithm

Send initial bid $w_i(1)$ to eNodeB
Send the application target rate $r_{i,s}^{\mathrm{t}}$ to eNodeB
loop
  while Flag $\sum_{i=1}^{M} r_{i,s}^{t} \geq R$ from eNodeB **do**
    Receive shadow price $p(n)$ from eNodeB
    if STOP from eNodeB **then**
      Calculate allocated rate $r_{i,s}^{\mathrm{opt}} = \frac{w_i(n)}{p(n)}$
    else
      Solve $r_{i,s}(n) = \arg\max_{r_{i,s}} \Big(\log U_i(r_{i,s}) - p(n) r_{i,s} \Big)$
      Send new bid $w_i(n) = p(n) r_{i,s}(n)$ to eNodeB
    end if
  end while
  while Flag $\sum_{i=1}^{M} r_{i,s}^{t} < R$ from eNodeB **do**
    Receive shadow price $p(n)$ from eNodeB
    if STOP from eNodeB **then**
      Calculate allocated rate $r_{i,s}^{\mathrm{opt}} = \frac{w_i(n)}{p(n)}$
    else
      Solve $r_{i,s}(n) = r_i + r_{i,s}^{\mathrm{t}} = \arg\max_{r_i} \Big(\log U_i(r_i + c_i) - p(n)(r_i + c_i) \Big)$
      Send new bid $w_i(n) = p(n)(r_i(n) + c_i)$ to eNodeB
    end if
  end while
end loop

Algorithm 14 Commercial UE Algorithm

Send initial bid $w_i(1)$ to eNodeB
loop
  while Flag $\sum_{i=1}^{M} r_{i,s}^{t} \geq R$ from eNodeB **do**
    Allocated rate $r_{i,c}^{\mathrm{opt}} = 0$
  end while
  while Flag $\sum_{i=1}^{M} r_{i,s}^{t} < R$ from eNodeB **do**
    Receive shadow price $p(n)$ from eNodeB
    if STOP from eNodeB **then**
      Calculate allocated rate $r_{i,c}^{\mathrm{opt}} = \frac{w_i(n)}{p(n)}$
    else
      Solve $r_{i,c}(n) = \arg\max_{r_{i,c}} \Big(\log U_i(r_{i,c}) - p(n) r_{i,c} \Big)$
      Send new bid $w_i(n) = p(n) r_{i,c}(n)$ to eNodeB
    end if
  end while
end loop

Algorithm 15 eNodeB Algorithm

loop
  Receive bids $w_i(n)$ from UEs {Let $w_i(0) = 0\ \forall i$}
  Receive application target rates from public safety UES
  while $\sum_{i=1}^{M} r_{i,s}^{t} \geq R$ **do**
    Send flag $\sum_{i=1}^{M} r_{i,s}^{t} \geq R$ to all UEs
    if $|w_i(n) - w_i(n-1)| < \delta$, $i = \{1, ..., M\}$ **then**
      STOP and allocate rates (i.e $r_{i,s}^{\text{opt}}$ to public safety user i)
    else
      Calculate $p(n) = \frac{\sum_{i=1}^{M} w_i(n)}{R}$, $i = \{1, ..., M\}$
      Send new shadow price $p(n)$ to public safety UEs
    end if
  end while
  while $\sum_{i=1}^{M} r_{i,s}^{t} < R$ **do**
    Send flag $\sum_{i=1}^{M} r_{i,s}^{t} < R$ to all UEs
    if $|w_i(n) - w_i(n-1)| < \delta\ \forall i$ **then**
      STOP and allocate rates (i.e $r_{i,s}^{\text{opt}}$ or $r_{i,c}^{\text{opt}}$ to user i)
    else
      Calculate $p(n) = \frac{\sum_{i=1}^{M+N} w_i(n)}{R}$
      Send new shadow price $p(n)$ to all UEs
    end if
  end while
end loop

cases. We present the simulation results for eight utility functions that correspond to public safety and commercial UEs running real time application or delay tolerant applications. We use two normalized utility functions expressed in equation (2.1) with different parameters a and b for each utility function, $a = 3$, $b = 20$ for the first public safety user, $a = 1$, $b = 30$ for the second public safety user. We set the application target rate $r_{i,s}^{t}$ for these two users to equal b that is 20 and 30 respectively. Another two normalized utility functions are used with the same a and b parameters to represent two commercial users running real time applications. Each sigmoidal-like function is an approximation to a step function at rate b. We also use two logarithmic functions expressed in equation (2.2) with different parameters $k = 3$ for one public safety UE and $k = 0.5$ for second public safety UE running delay tolerant application. We set the application target rate $r_{i,s}^{t}$ for each of these two users to equal 15. Another two logarithmic utility functions are used with the same k parameters to

represent two commercial users running delay tolerant applications.

5.1.4.1 Convergence Dynamics for $R = 70$ where $\sum_{i=1}^{M} r_{i,s}^{t} \geq R$

This represents the first case where $\sum_{i=1}^{M} r_{i,s}^{t} \geq R$. We set $R = 70$ and $\delta = 10^{-2}$. As mentioned before, in this case the commercial UEs will not be allocated any of the eNodeB resources because R does not exceed the public safety application target rates which need to be satisfied before the eNodeB starts allocating resources to the commercial users. In Figure 5.1, we show the simulation results for the rate of different public safety users and the number of iterations. The sigmoidal-like utility functions are given priority over the logarithmic utility functions for rate allocation. This explain the results we got in Figure 5.1. In this case the final optimal rate does not exceed the user application target rate. In Figure 5.2, we show the bids of the four public safety users with the number of iterations. As expected, user rates are proportional to the user bids. The algorithm allows users with real-time applications to bid higher than the other users until each one of them reaches its inflection point, which is equivalent to their application target rates, then users with elastic traffic start dividing the remaining resources among them based on their parameters while not exceeding their application target rates. In Figure 5.3, we show the shadow price $p(n)$ with the number of iterations where the convergence behavior of the shadow price with the number of iterations is shown.

5.1.4.2 Convergence Dynamics for $R = 200$ where $\sum_{i=1}^{M} r_{i,s}^{t} < R$

Figure 5.4 shows four public safety normalized sigmoidal-like utility functions expressed in equation (2.1) corresponding to two public safety users and another two commercial users. We also show four logarithmic functions expressed in equation (2.2), which represent delay tolerant applications for two public safety users and another two commercial users. We set $R = 120$ and $\delta = 10^{-2}$. This represents the second case where $\sum_{i=1}^{M} r_{i,s}^{t} < R$. In this case

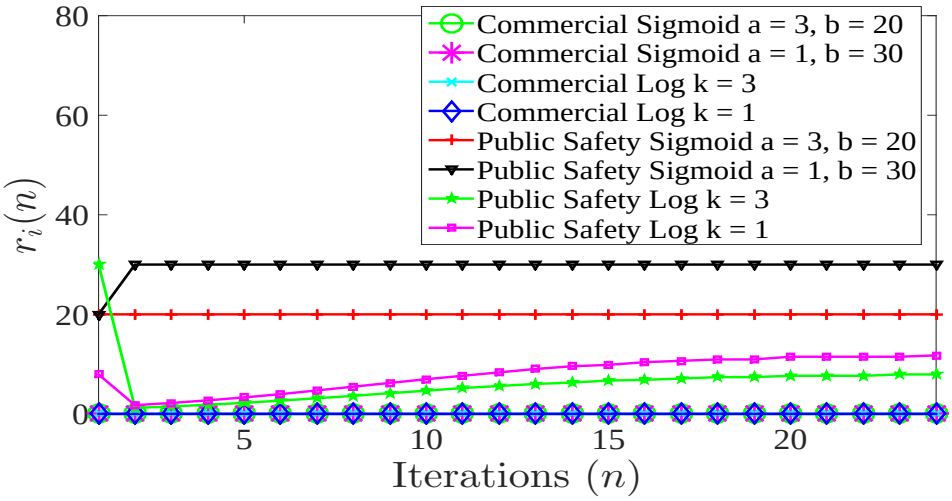


Figure 5.1: The rates $r_i(n)$ with the number of iterations n for different users and $R = 70$.

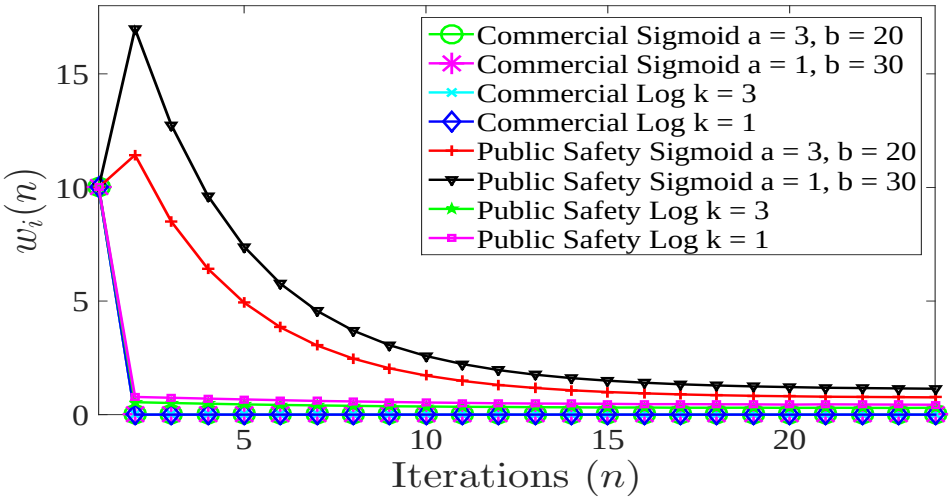


Figure 5.2: The bids convergence $w_i(n)$ with the number of iterations n for different users and $R = 70$.

the public safety UEs are given priority over the commercial UEs. In Figure 5.5, we show the simulation results for the rate of different public safety and commercial users and the number of iterations., first the algorithm allocates an equivalent amount of resources to the application target rate to each public safety user. It then starts allocating resources to each commercial UE with inelastic traffic until it reaches the inflection point of that user utility function. It then starts dividing the remaining resources among all users based on their parameters. In Figure 5.6, we show the bids of the eight users with the number of iterations.

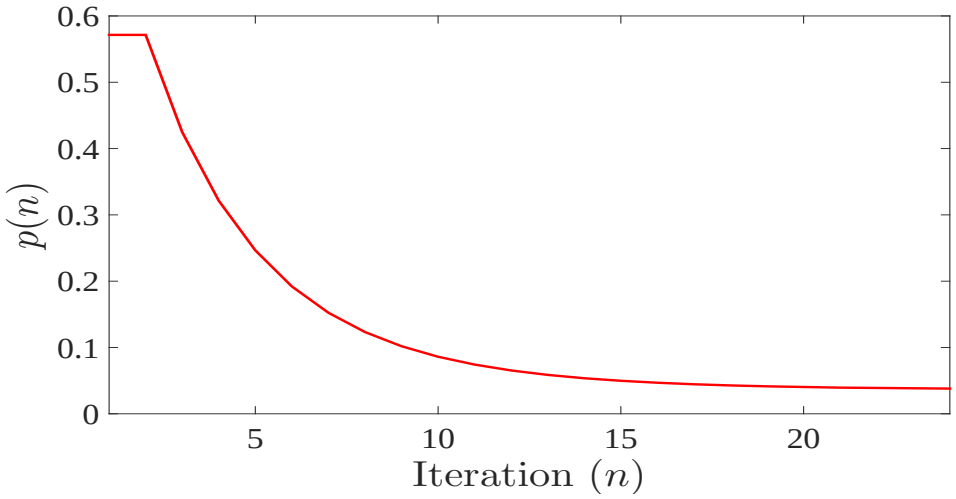


Figure 5.3: The shadow price convergence with the number of iterations n.

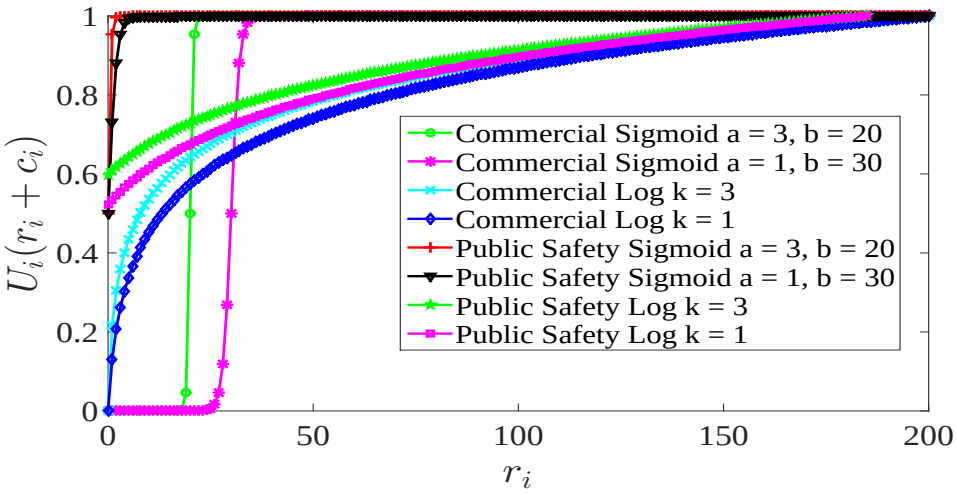


Figure 5.4: The users utility functions $U_i(r_i + c_i)$.

The algorithm allows public safety users to bid higher than the other users until each one of them reaches its application target rate. Commercial users with inelastic traffic then start bidding higher until they each utility function reaches its inflection point. In Figure 5.7, we show the shadow price $p(n)$ with the number of iterations where the convergence behavior of the shadow price with the number of iterations is shown.

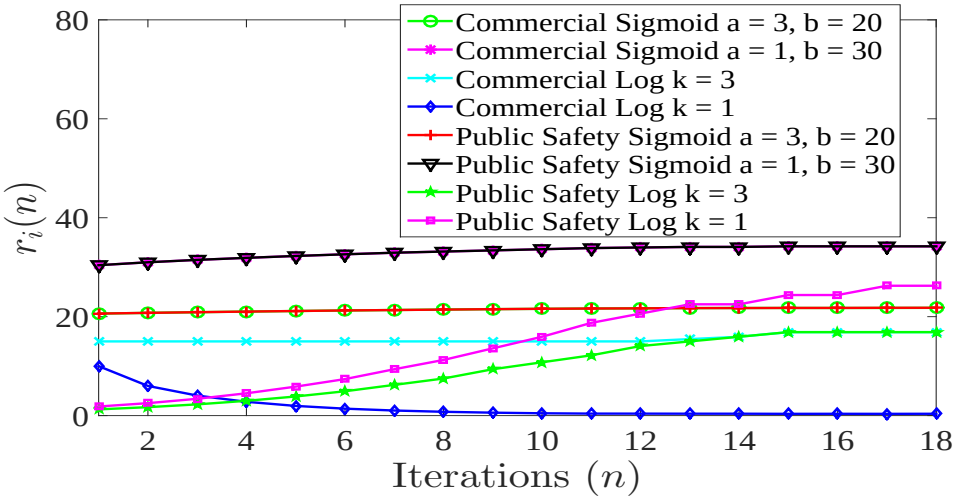


Figure 5.5: The rates $r_i(n)$ with the number of iterations n for different users and $R = 200$.

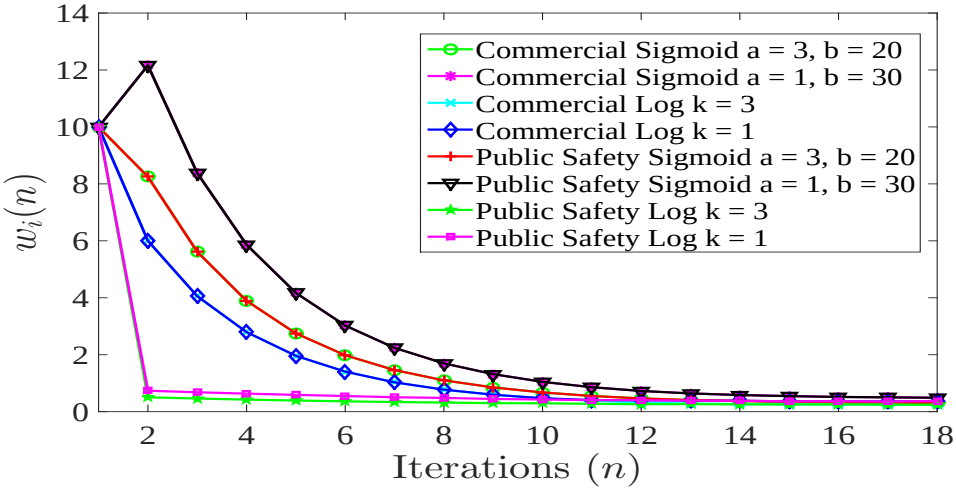


Figure 5.6: The bids convergence $w_i(n)$ with the number of iterations n for different users and $R = 200$.

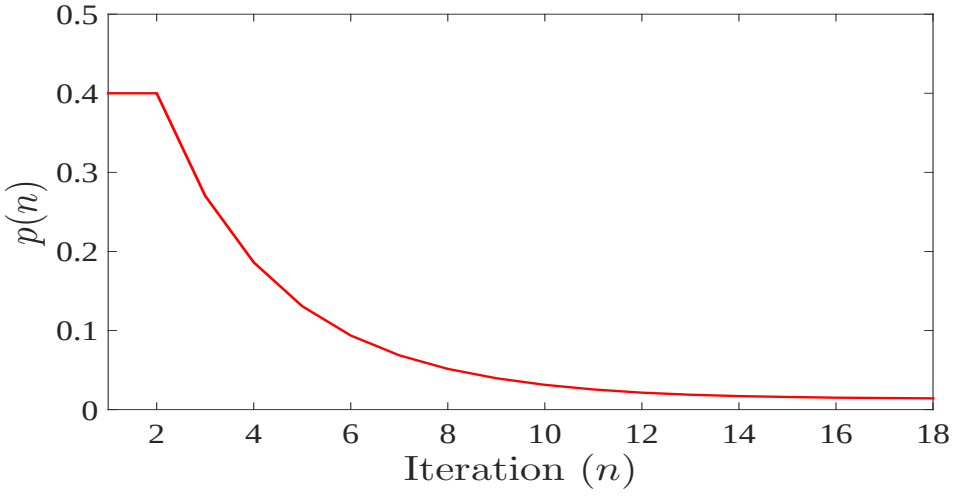


Figure 5.7: The shadow price convergence with the number of iterations n.

5.2 Multi-Application Resource Allocation with User Discrimination in Cellular Networks

In this section, we focus on finding an optimal solution for the resource allocation problem for different types of users running multiple types of applications simultaneously on their UEs. We considered subscriber differentiation, application status differentiation (application weight) and application target rate when formulating the resource allocation optimization problem. In our model, each user subscribing for a service is assigned a subscription weight by the network. Each user can run multiple applications simultaneously and each application is represented by a utility function based on the application type. In addition, each application is assigned an application weight by the UE based on the application instantaneous usage percentage and importance to the UE. Furthermore, certain type of users with higher priority (e.g. VIP users) are assigned applications target rates by the network. Therefore, these VIP UEs' applications are given higher priority by the network when allocating resources. A minimum QoS is guaranteed for each user by using a proportional fairness approach and real-time applications are given priority over delay-tolerant applications. Our objective is to allocate the resources optimally among the UEs and their applications from a single eNodeB based on a utility proportional fairness policy. We propose a two-stage rate allocation algorithm to allocate the eNodeB resources among users and their applications. In the first stage, the eNodeB collaborates with the UEs to allocate user rates. In the second stage, the rates are allocated to user applications internally by the UEs. Our contributions in this section are summarized as:

- We present a resource allocation optimization problem to allocate the eNodeB resources optimally among different types of users running multiple applications.
- We propose a two-stage rate allocation method to allocate rates optimally among users.

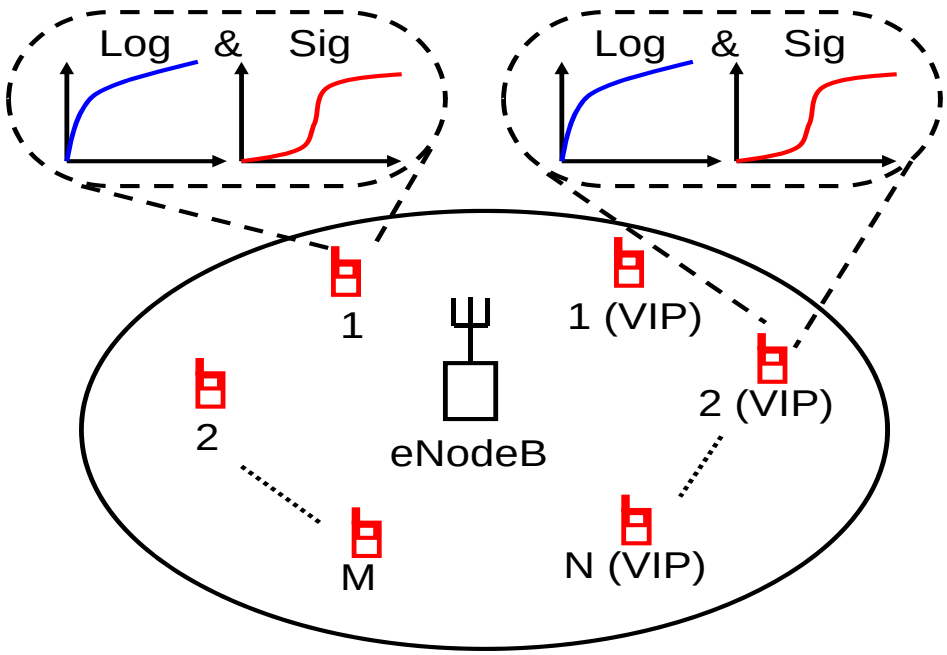


Figure 5.8: System Model, one eNodeB with N VIP UEs and another M regular UEs subscribing for a mobile service in the eNodeB coverage area.

First, the eNodeB and the UE collaborate to allocate an optimal rate to each UE. Each UE then allocates its assigned rate optimally among its applications.

- We show that our resource allocation optimization problems have unique tractable global optimal solutions.

5.2.1 Problem Formulation

We consider a single cell mobile system that consists of a single eNodeB, M regular UEs and another N VIP UEs as shown in Figure 5.8. The rate allocated by the eNodeB to the i^{th} UE is given by r_i. Each UE has its own utility function $X_i(r_i)$ that corresponds to the user satisfaction with its allocated rate r_i. Our objective is first to determine the optimal rates the eNodeB shall allocate to the UEs. We assume that the utility function $X_i(r_i)$ that is assigned to the i^{th} user is given by:

$$X_i(r_i) = \prod_{j=1}^{L_i} U_{ij}^{\alpha_{ij}}(r_{ij} + c_{ij}) \tag{5.5}$$

$$c_{ij} = \begin{cases} r_{ij}^t & \text{if the } j^{th} \text{ application is assigned} \\ & \text{an application target rate} \\ 0 & \text{if the } j^{th} \text{ application is not assigned} \\ & \text{an application target rate} \end{cases}$$

where $U_{ij}(r_{ij})$ is the j^{th} application utility function for user i, r_{ij} is the rate allocated to the j^{th} application running on the i^{th} UE, L_i is the number of applications running on the i^{th} UE, c_{ij} is the application target rate for the j^{th} application of user i if it is assigned one whereas it is 0 if the j^{th} application is not assigned an application target rate by the network, α_{ij} is the j^{th} application usage percentage (application weight) of the i^{th} UE and r_{ij}^t is the application target rate assigned to the j^{th} application of the i^{th} user.

We express the user satisfaction with its provided service using utility functions [18,82,83]. We assume that the j^{th} application utility function for user i is given by $U_{ij}(r_{ij})$ that is strictly concave function expressed by equation (2.2) or sigmoidal-like function expressed by equation (2.1) where r_{ij} is the rate allocated to the j^{th} application of user i. Delay tolerant applications are represented by logarithmic utility functions whereas real-time applications are represented by sigmoidal-like utility functions.

5.2.2 Resource Allocation Optimization Problem

The resource allocation (RA) optimization problem for multi-application users is divided into two cases. The first-case is when the maximum available resources R of the eNodeB is

less than or equal to the total VIP UEs applications target rates. The second-case is when R is greater than the total UEs applications target rates. The RA optimization problems for the two cases will be solved by our proposed algorithm to obtain the optimal rate for each UE as well as the optimal rates for the UE applications.

5.2.2.1 First-Case RA Optimization Problem when $\sum_{i=1}^{M}\sum_{j=1}^{L_i} r_{ij}^{t} \geq R$

In this case, the eNodeB only allocates resources to the M VIP UEs as they are considered more important and regular users will not be allocated any of the eNodeB resources since its available resources are limited. In this case, the optimization problem is divided into two stages. In the first-stage, the eNodeB allocates rates r_i to the M group of users. Both the eNodeB and the M UEs collaborate to achieve the UEs resource allocation. In the second-stage, each one of these M UEs uses the rate allocated to it by the eNodeB to allocate optimal rates r_{ij} to its L_i applications. The second-stage is performed internally in the UE.

5.2.2.1.1 First-Stage of the First-Case Optimization Problem

In this case, the optimization problem for the first-stage can be written as:

$$\begin{aligned} \max_{\mathbf{r}} \quad & \prod_{i=1}^{M} X_i^{\beta_i}(r_i) \\ \text{subject to} \quad & \sum_{i=1}^{M} r_i \leq R \\ & 0 \leq r_i \leq \sum_{j=1}^{L_i} r_{ij}^{t}, \qquad i = 1, 2, ..., M. \end{aligned} \tag{5.6}$$

where $X_i = \prod_{j=1}^{L_i} U_{ij}^{\alpha_{ij}}(r_{ij})$, $\mathbf{r} = \{r_1, r_2, ..., r_M\}$ is the rate allocated by the eNodeB to the i^{th} UE, M is the number of VIP UEs in the coverage area of the eNodeB, R is the maximum achievable rate of the given eNodeB and β_i is the i^{th} user subscription weight assigned by

the network.

The objective function in the optimization problem (5.6) is equivalent to $\sum_{i=1}^{M} \beta_i \log(X_i(r_i))$. Therefore, the optimization problem (5.6) is a convex optimization problem and there exists a unique tractable global optimal solution as shown in Corollary (III.1) [23]. This optimal solution gives each of the M users an optimal rate r_i^{opt} that is less than or equal to the total applications target rates for that UE.

5.2.2.1.2 Second-Stage of the First-Case Optimization Problem

Each one of the M VIP UEs allocates optimal rates r_{ij}^{opt} to its L_i applications. The optimal rate allocated to each application depends on the application differentiation weight and the application type. This optimization problem is solved internally in the UE and can be written for the i^{th} UE as follows:

$$\begin{aligned} \max_{\mathbf{r}_i} \quad & \prod_{j=1}^{L_i} U_{ij}^{\alpha_{ij}}(r_{ij}) \\ \text{subject to} \quad & \sum_{j=1}^{L_i} r_{ij} \leq r_i^{\text{opt}} \\ & 0 \leq r_{ij} \leq r_{ij}^{\text{t}}, \qquad j = 1, 2, ..., L_i. \end{aligned} \tag{5.7}$$

where $\mathbf{r}_i = \{r_{i1}, r_{i2}, ..., r_{iL_i}\}$, r_i^{opt} is the optimal rate allocated by the eNodeB to the i^{th} UE and L_i is number of the UE applications. Since the objective function in the optimization problem (5.7) is equivalent to $\sum_{j=1}^{L_i} \alpha_{ij} \log(U_{ij}(r_{ij}))$, then optimization problem (5.7) is convex and there exists a unique tractable global optimal solution as shown in Corollary (III.2) [23]. This optimal solution represents the optimal rate r_{ij}^{opt} allocated to each of the L_i applications.

5.2.2.2 Second-Case RA Optimization Problem when $\sum_{i=1}^{M}\sum_{j=1}^{L_i} r_{ij}^{t}<R$

In this case, the eNodeB first allocates resources to the M VIP UEs. It then allocates the remaining resources based on the proportional fairness approach. The optimization problem in this case is divided into two stages. In the first-stage, the eNodeB collaborates with the UEs to allocate rates r_i to all UEs. In the second-stage, each one of these $M+N$ UEs allocates optimal rates r_{ij} to its applications. The second-stage is performed internally in the UE. The inelastic traffic are given priority when allocating the resources internally by the UEs.

5.2.2.2.1 First-Stage of the Second-Case Optimization Problem

In this case, the optimization problem of the first-stage can be written as:

$$\begin{aligned} \max_{\mathbf{r}} \quad & \prod_{i=1}^{M+N} X_i^{\beta_i}(r_i) \\ \text{subject to} \quad & \sum_{i=1}^{M+N} r_i \leq R \\ & r_i \geq 0, \qquad i = 1, 2, ..., M+N. \end{aligned} \tag{5.8}$$

where $X_i = \prod_{j=1}^{L_i} U_{ij}^{\alpha_{ij}}(r_{ij} + c_{ij})$ and $\mathbf{r} = \{r_1, r_2, ..., r_{M+N}\}$ and $M+N$ is the number of the VIP and regular UEs subscribing for a service in the coverage area of the eNodeB and β_i is the i^{th} user subscription weight assigned by the network. Each UE is allocated at least the total amount of its applications target rates if it has any.

The objective function in the optimization problem (5.8) is equivalent to $\sum_{i=1}^{M+N} \beta_i \log(X_i(r_i))$. Therefore, optimization problem (5.8) is a convex optimization problem and there exists a unique tractable global optimal solution r_i^{opt} for each of the $M+N$ users as shown in Corollary (III.1) [23].

5.2.2.2.2 Second-Stage of the Second-Case Optimization Problem

Each one of the $M+N$ UEs allocates optimal rates r_{ij}^{opt} to its applications. Each UE first allocates the application target rate to each of its applications if it is assigned one. It then starts allocating the remaining resources among all the applications based on the application differentiation weight and the type of the application. This optimization problem is solved internally in the UE and can be written for the i^{th} UE as follows:

$$\begin{aligned} \max_{\mathbf{r}_i} \quad & \prod_{j=1}^{L_i} U_{ij}^{\alpha_{ij}}(r_{ij}+c_{ij}) \\ \text{subject to} \quad & \sum_{j=1}^{L_i}(r_{ij}+c_{ij}) \leq r_i^{\text{opt}} \\ & r_{ij} \geq 0, \qquad j=1,2,...,L_i. \end{aligned} \tag{5.9}$$

where $\mathbf{r}_i = \{r_{i1}, r_{i2}, ..., r_{iL_i}\}$, r_i^{opt} is the rate allocated by the eNodeB to the i^{th} UE in the first-stage and c_{ij} is same as before. The objective function of the optimization problem (5.9) is equivalent to $\sum_{j=1}^{L_i} \alpha_{ij} \log(U_{ij}(r_{ij}+c_{ij}))$. Therefore, optimization problem (5.9) is a convex optimization problem and there exists a unique tractable global optimal solution as shown in Corollary (III.2) [23]. Each UE allocates an optimal rate $r_{ij}^{\text{opt}} = r_{ij} + c_{ij}$ to each of its applications.

5.2.3 Algorithms

As mentioned before, the RA for the multi-application users with different priorities is achieved in two-stages. In the first-stage, the eNodeB and the UEs collaborate to allocate optimal rates r_i for users as shown in VIP UE Algorithm (16), regular UE Algorithm (17) and eNodeB Algorithm (18). In the second-stage, the UE internal algorithm allocates applications rates r_{ij} to the UE's applications as shown in the internal UE Algorithm (19).

5.2.3.1 First-Stage RA Algorithm

The first-stage of the RA algorithm is presented in this section. The algorithm starts when each UE transmits an initial bid $w_i(1)$ to the eNodeB. Additionally, each VIP UE transmits its applications target rates to the eNodeB. The eNodeB checks whether the $\sum_{i=1}^{M}\sum_{i=1}^{L_i} r_{ij}^{t}$ is less or greater than R and sends a flag with this information to each UE. In the case of $\sum_{i=1}^{M}\sum_{i=1}^{L_i} r_{ij}^{t} \geq R$, the regular UEs will not be allocated any of the resources and will not be sending any further bids to the eNodeB.

Algorithm 16 VIP UE Algorithm

Send initial bid $w_i(1)$ to eNodeB
Send the applications target rates r_{ij}^{t} to eNodeB
loop
  while Flag $\sum_{i=1}^{M}\sum_{j=1}^{L_i} r_{ij}^{t} \geq R$ from eNodeB **do**
    Receive shadow price $p(n)$ from eNodeB
    if STOP from eNodeB **then**
      Calculate allocated rate $r_i^{\mathrm{opt}} = \frac{w_i(n)}{p(n)}$
    else
      Solve $r_i(n) = \arg\max_{r_i}\Big(\beta_i \log X_i(r_i) - p(n) r_i\Big)$
      Send new bid $w_i(n) = p(n) r_i(n)$ to eNodeB
    end if
  end while
  while Flag $\sum_{i=1}^{M}\sum_{j=1}^{L_i} r_{ij}^{t} < R$ from eNodeB **do**
    Receive shadow price $p(n)$ from eNodeB
    if STOP from eNodeB **then**
      Calculate allocated rate $r_i^{\mathrm{opt}} = \frac{w_i(n)}{p(n)}$
    else
      Solve $r_i(n) = \arg\max_{r_i}\Big(\beta_i \log X_i(r_i) - p(n)(r_i + \sum_{j=1}^{L_i} r_{ij}^{t})\Big)$
      Calculate new bid $w_i(n) = p(n)(r_i(n) + \sum_{j=1}^{L_i} r_{ij}^{t})$
      if $|w_i(n) - w_i(n-1)| > \Delta w$ **then**
        $w_i(n) = w_i(n-1) + \mathrm{sign}(w_i(n) - w_i(n-1))\Delta w(n)$
        $\{\Delta w(n) = l_1 e^{-\frac{n}{l_2}}\}$
      end if
      Send new bid $w_i(n)$ to eNodeB
    end if
  end while
end loop

Each VIP UE checks whether the difference between the current received bid and the

previous one is less than a threshold δ, if so it exits. Otherwise, the eNodeB calculates the shadow price $p(n) = \frac{\sum_{i=1}^{M} w_i(n)}{R}$ and sends it to the VIP UEs where it is used to calculate the i^{th} VIP UE rate $r_i(n)$ which is the solution of the optimization problem $r_i(n) = \arg\max_{r_i}(\beta_i \log X_i(r_i) - p(n)r_i)$ where $X_i(r_i) = \prod_{j=1}^{L_i} U_{ij}^{\alpha_{ij}}(r_{ij})$. A new bid $w_i(n) = p(n)r_i(n)$ is then calculated and the VIP UEs check the fluctuation condition as in [22] and send their new bids to the eNodeB. The Algorithm is finalized by the eNodeB. Each VIP UE then calculates its allocated rate $r_i^{\text{opt}} = \frac{w_i(n)}{p(n)}$.

Algorithm 17 Regular UE Algorithm

```
Send initial bid w_i(1) to eNodeB
loop
  while Flag Σ_{i=1}^{M} Σ_{j=1}^{L_i} r_ij^t ≥ R from eNodeB do
    Allocated rate r_i^opt = 0
  end while
  while Flag Σ_{i=1}^{M} Σ_{j=1}^{L_i} r_ij^t < R from eNodeB do
    Receive shadow price p(n) from eNodeB
    if STOP from eNodeB then
      Calculate allocated rate r_i^opt = w_i(n)/p(n)
    else
      Solve r_i(n) = arg max_{r_i} ( β_i log X_i(r_i) − p(n) r_i )
      Calculate new bid w_i(n) = p(n) r_i(n)
      if |w_i(n) − w_i(n−1)| > Δw then
        w_i(n) = w_i(n−1) + sign(w_i(n) − w_i(n−1)) Δw(n)
        {Δw(n) = l_1 e^{−n/l_2}}
      end if
      Send new bid w_i(n) to eNodeB
    end if
  end while
end loop
```

In the case of $\sum_{i=1}^{M} \sum_{i=1}^{L_i} r_i^t < R$, a flag with this information is sent to each UE by the eNodeB. Each UE checks whether the difference between the current received bid and the previous one is less than a threshold δ, if so it exits. Otherwise, the eNodeB calculates the shadow price $p(n) = \frac{\sum_{i=1}^{M+N} w_i(n)}{R}$ and sends it to each UE where it is used by the VIP UE to calculate the rate $r_i = r_i(n) + \sum_{j=1}^{L_i} r_{ij}^{\text{t}}$, $r_i(n)$ is the solution of the optimization problem $r_i(n) = \arg\max_{r_i}(\beta_i \log X_i(r_i) - p(n)(r_i + \sum_{j=1}^{L_i} r_{ij}^{\text{t}}))$ where $X_i(r_i) = \prod_{j=1}^{L_i} U_{ij}^{\alpha_{ij}}(r_{ij} + c_{ij})$. A

new bid $w_i(n) = p(n)(r_i(n) + \sum_{j=1}^{L_i} r_{ij}^{\mathrm{t}})$ is calculated by the VIP UE. All VIP UEs check the fluctuation condition and send their new bids to the eNodeB. On the other hand, the regular UEs receive $p(n)$ and calculate the rate $r_i(n)$ which is the solution of the optimization problem $r_i(n) = \arg\max_{r_i}(\beta_i \log X_i(r_i) - p(n)r_i)$ where $X_i(r_i) = \prod_{j=1}^{L_i} U_{ij}^{\alpha_{ij}}(r_{ij} + c_{ij})$. A new bid $w_i(n) = p(n)r_i(n)$ is calculated by the regular UE. All regular UEs check the fluctuation condition and send their new bids to the eNodeB. The Algorithm is finalized by the eNodeB. Each VIP and regular UE then calculates its allocated rate $r_i^{\mathrm{opt}} = \frac{w_i(n)}{p(n)}$.

Algorithm 18 eNodeB Algorithm

loop
  Receive bids $w_i(n)$ from UEs {Let $w_i(0) = 0\ \forall i$}
  Receive applications target rates from VIP UEs
  while $\sum_{i=1}^{M}\sum_{j=1}^{L_i} r_{ij}^t \geq R$ **do**
    Send flag $\sum_{i=1}^{M}\sum_{j=1}^{L_i} r_{ij}^t \geq R$ to all UEs
    if $|w_i(n) - w_i(n-1)| < \delta$, $i = \{1, ..., M\}$ **then**
      STOP and allocate rates (i.e r_i^{opt} to VIP user i)
    else
      Calculate $p(n) = \frac{\sum_{i=1}^{M} w_i(n)}{R}$, $i = \{1, ..., M\}$
      Send new shadow price $p(n)$ to VIP UEs
    end if
  end while
  while $\sum_{i=1}^{M}\sum_{j=1}^{L_i} r_{ij}^t < R$ **do**
    Send flag $\sum_{i=1}^{M}\sum_{j=1}^{L_i} r_{ij}^t < R$ to all UEs
    if $|w_i(n) - w_i(n-1)| < \delta\ \forall i$ **then**
      STOP and allocate rates (i.e r_i^{opt} to user i)
    else
      Calculate $p(n) = \frac{\sum_{i=1}^{M+N} w_i(n)}{R}$
      Send new shadow price $p(n)$ to all UEs
    end if
  end while
end loop

5.2.3.2 Second-Stage RA Algorithm

The second-stage of RA is presented in this section and shown in Algorithm (19) where the rates r_{ij} are allocated internally by the UE to its applications. Each UE uses its allocated rate

r_i^{opt} in the first-stage to solve the optimization problem $\mathbf{r}_i = \arg\max_{\mathbf{r}_i} \sum_{j=1}^{L_i}(\alpha_{ij} \log U_{ij}(r_{ij} + c_{ij}) - p(r_{ij}+c_{ij})) + pr_i^{\text{opt}}$. The rate $r_{ij}^{opt} = r_{ij}+c_{ij}$ is then allocated to the UE's j^{th} application.

Algorithm 19 Internal UE Algorithm

loop
Receive r_i^{opt} from eNodeB Algorithm (16), (17) and (18)
Solve
$\mathbf{r}_i = \arg\max_{\mathbf{r}_i} \sum_{j=1}^{L_i}(\alpha_{ij} \log U_{ij}(r_{ij} + c_{ij}) - p(r_{ij} + c_{ij})) + pr_i^{\text{opt}}$
$\{\mathbf{r}_i = \{r_{i1}, r_{i2}, ..., r_{iL_i}\}\}$
Allocate $r_{ij}^{opt} = r_{ij} + c_{ij}$ to the j^{th} application
end loop

5.2.4 Simulation Results

In this section, we consider one eNodeB with four UEs in its coverage area subscribing for a mobile service. The first and second UEs are VIP UEs and the third and fourth UEs are regular UEs. Each one of the four UEs is running two applications simultaneously. The first application is a real-time application whereas the second application is a delay-tolerant application.

We applied algorithm (16), (17), (18) and (19) in C++ to the UEs functions. The simulation results showed convergence to the optimal global point in the two stages of the algorithm. We present the simulation results for the four users. The first UE is a VIP UE, we use a normalized sigmoidal-like utility function that is expressed by equation (2.1) to represent its first application with $a = 3$, $b = 20$ which is an approximation to a step function at rate $r = 20$ and we set $r_{11}^t = 20$. Additionally, for the second application of the first user (VIP user) we use a logarithmic function that is expressed by equation (2.2) with $k = 3$ which is an approximation of a delay-tolerant application. The second user is a VIP user, we use a normalized sigmoidal-like utility function to represent its first application with $a = 1$, $b = 30$ and we set $r_{21}^t = 30$. Additionally, for the second application of the second user (VIP user)

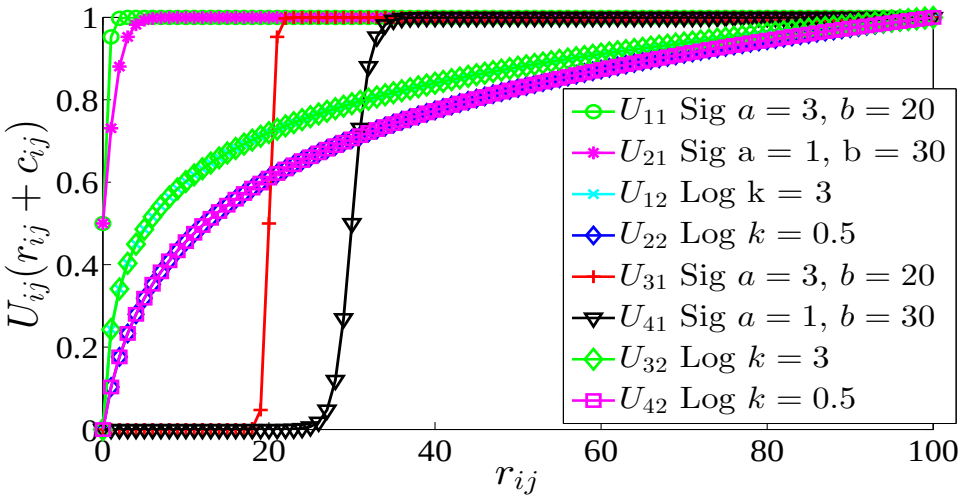


Figure 5.9: The applications utility functions $U_{ij}(r_{ij})$.

we use a logarithmic function with $k = 0.5$ to represent its delay tolerant application. The same parameters of the first user are used for the third user's utility functions except that its applications are not assigned applications target rates. Also, the same parameters of the second user are used for the fourth user's utility functions except that its applications are not assigned applications target rates. Furthermore, we set $\beta_i = 1$ for all UEs. We use $r_{max} = 100$ for all logarithmic functions, $l_1 = 5$ and $l_2 = 10$ in the fluctuation decay function of the algorithm and $\delta = 10^{-3}$. Let the application weight α_{ij} in the set α corresponds to the j^{th} application of user i where α be $\alpha = \{\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}, \alpha_{31}, \alpha_{32}, \alpha_{41}, \alpha_{42}\}$.

Figure 5.9 shows eight applications utility functions corresponding to the four UEs. The real-time applications of the VIP UEs are assigned applications target rates, this explains their shifted utility functions by the amount of r_{ij}^t in Figure 5.9. The other applications do not have applications target rates ($c_{ij} = 0$ for each one). Figure 5.10 shows the aggregated utilities $X_i(r_i)$ for each user.

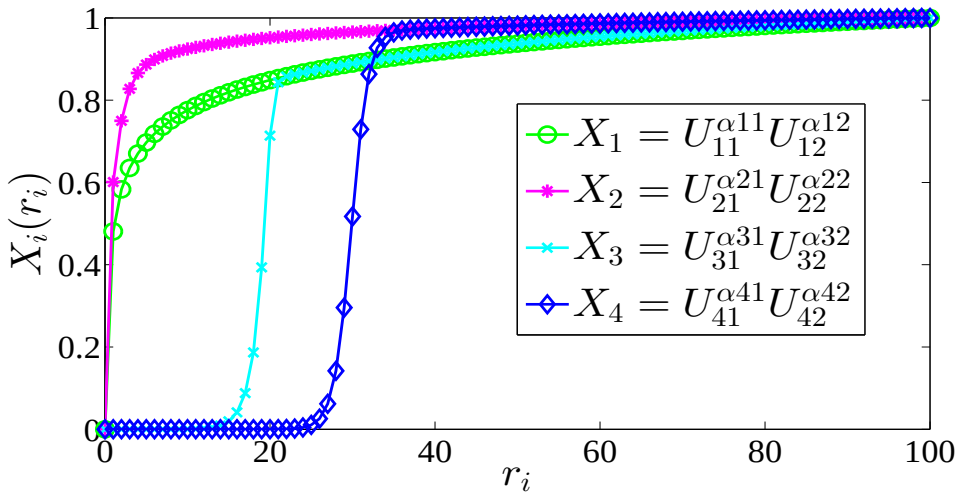


Figure 5.10: The aggregated utility functions $X_i(r_i)$ of the i^{th} user.

5.2.4.1 Convergence Dynamics for $5 \leq R \leq 200$

In the following simulations, we set $\alpha = \{0.5, 0.5, 0.9, 0.1, 0.5, 0.5, 0.9, 0.1\}$ and the eNodeB available resources R takes values between 5 and 200 with step of 5. In Figure 5.11, we show the four users optimal rates r_i^{opt} with different eNodeB resources R. This represents the solution of optimization problem (5.6) when $R \leq 50$ and optimization problem (5.8) when $R > 50$, using the first-stage of the algorithm, where 50 is the total applications target rates for the the two VIP users. Figure 5.11 shows that when $R \leq 50$ the regular UEs are not allocated any of the eNodeB resources. Furthermore, when $R > 50$ each VIP user is first allocated its total applications target rates and the remaining resources are then allocated to all users based on the proportional fairness approach.

In Figure 5.12, we show the final optimal applications rates r_{ij}^{opt} for the four users with different eNodeB resources R. This is the solution of optimization problem (5.7) when $R \leq 50$ and the solution of (5.9) when $R > 50$ using the user internal algorithm. The figure shows that when $R \leq 50$, the real-time applications are given priority over the delay tolerant applications when allocating rates by each VIP UE to its applications whereas when $R > 50$, the VIP UEs first allocate the applications target rates to the applications that are assigned

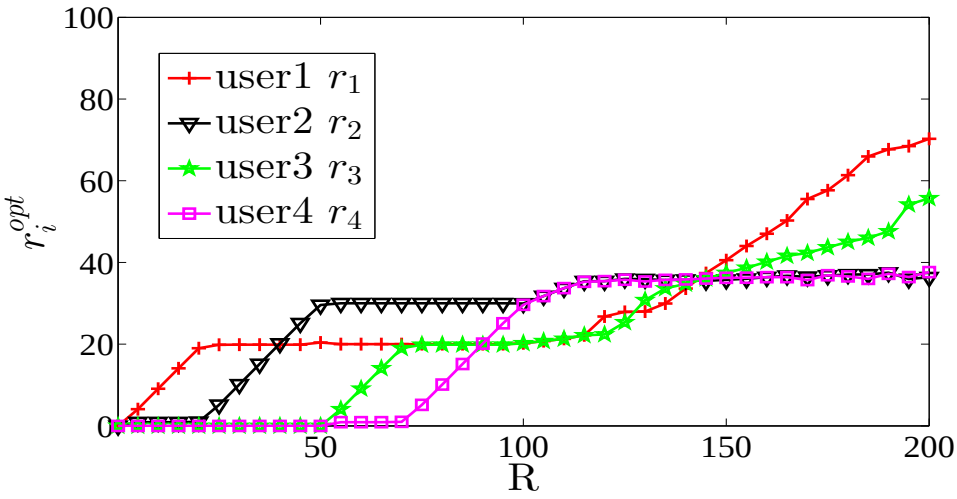


Figure 5.11: The users optimal rates r_i^{opt} for different values of R.

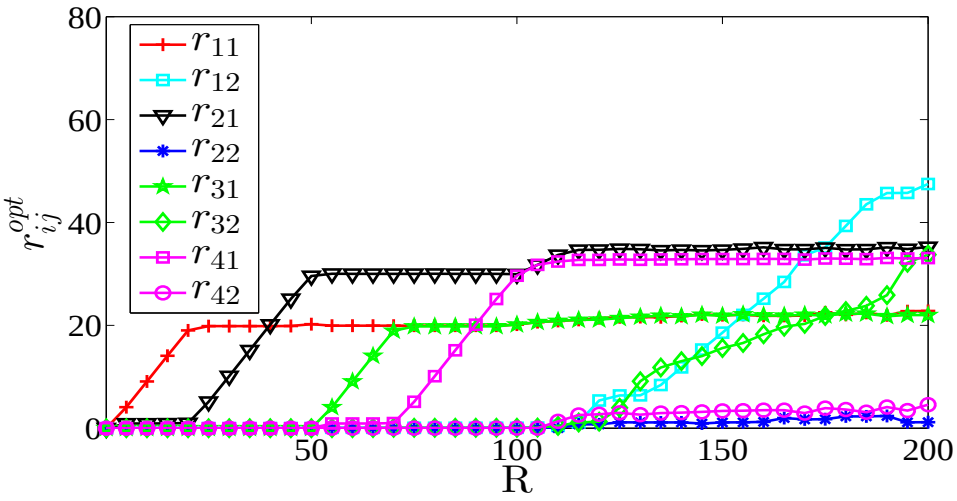


Figure 5.12: The applications optimal rates r_{ij}^{opt} for different values of R.

ones and then allocate the remaining resources among all applications using proportional fairness approach while giving the priority to the real-time applications. The regular users also give the priority to their real-time applications when allocating resources as shown in the same figure.

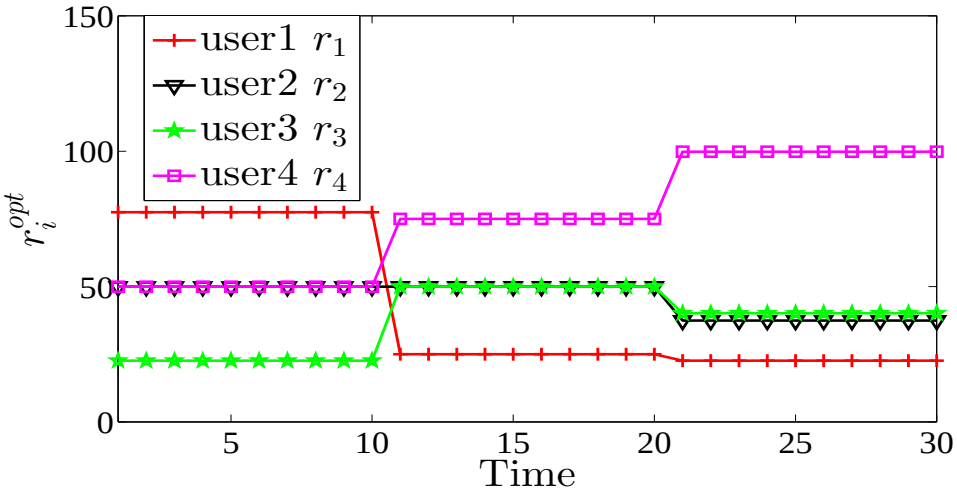


Figure 5.13: The users optimal rates r_i^{opt} with the change in users' applications usage percentages $\alpha(t)$.

5.2.4.2 Rate Allocation Sensitivity to change in α

In the following simulations, we measure the sensitivity of the change in application weight that is corresponding to the application usage percentage in the UE. We use $R = 200$ and the same parameters as before for the four users. The users change their applications usage percentage with time as the following

$$\alpha(t) = \begin{cases} \alpha = \{0.1, 0.9, 0.5, 0.5, 0.9, 0.1, 0.5, 0.5\}; \\ \qquad\qquad for \;\; 0 \leq t \leq 10 \\ \alpha = \{0.5, 0.5, 0.3, 0.7, 0.2, 0.8, 0.1, 0.9\}; \\ \qquad\qquad for \;\; 10 \leq t \leq 20 \\ \alpha = \{1.0, 0.0, 0.9, 0.1, 0.8, 0.2, 0.1, 0.9\}; \\ \qquad\qquad for \;\; 20 \leq t \leq 30 \end{cases} \tag{5.10}$$

Figure 5.13 shows the users optimal rates r_i^{opt} with time for the changing usage percentages given by $\alpha(t)$.

5.3 Resource Allocation with User Discrimination Framework for Multi-Carrier Cellular Networks

In this section, we provide an efficient framework for the resource allocation problem to allocate multi-carrier resources optimally among users that belong to different classes of user groups. In our model, we use utility functions to represent users' applications. Sigmoidal-like utility functions and logarithmic utility functions are used to represent real-time and delay-tolerant applications, respectively, running on the UEs [21]. The resource allocation with user discrimination framework presented in [26] does not consider the case of multi-carrier resources available at the eNodeB. It only solves the problem of resource allocation with user discrimination in the case of single carrier. In this section, we consider the case of multiple carriers' resources available at the eNodeB and multiple classes of users located under the coverage area of these carriers. We use a priority criterion for the resource allocation process that varies based on the user's class and the type of application running on the UE. We consider two classes of users, VIP users (i.e. public safety users or users who require emergency services) and regular users. VIP users are assigned a minimum required application rate for each of their applications whereas regular users' applications are not assigned any.

We formulate the resource allocation with user discrimination problem in a multi-stage resource allocation with carrier aggregation optimization problem to allocate resources to each user from its all in range carriers based on a utility proportional fairness policy. Each application running on the UE is assigned an application minimum required rate by the network that varies based on the type of user's application and the user's class. Furthermore, if the user's in range carriers have enough available resources, the user is allocated at minimum its applications' minimum required rates. VIP users are given priority over regular users by the network when allocating each carrier's resources, and real-time applications are given

priority over delay-tolerant applications.

Our contributions in this section are summarized as:

- We present a multi-stage resource allocation with user discrimination optimization problem to allocate multi-carrier resources optimally among different classes of users.
- We prove that the resource allocation optimization problem is convex and therefore the global optimal solution is tractable.
- We present a resource allocation algorithm to solve the optimization problem and allocate each user an aggregated final rate from its in range carriers. The proposed algorithm outperforms that presented in [26] as it considers allocating each user resources from multiple carriers using a resource allocation with carrier aggregation approach.
- We present simulation results for the performance of the proposed resource allocation algorithm.

5.3.1 Problem Formulation

In this paper, we consider a single cell mobile system with one eNodeB, K carriers (frequency bands) that have resources available at the eNodeB, M regular and VIP UEs. Let $\mathcal{M}$ be the set of all regular and VIP UEs where $M = |\mathcal{M}|$. The set of carriers is given by $\mathcal{K} = \{1, 2, ..., K\}$ with carriers in order from the highest frequency to the lowest frequency. Higher frequency carriers have smaller coverage area than lower frequency carriers. The eNodeB allocates resources from multiple carriers to each UE. Users located under the coverage area of multiple carriers are allocated resources from all in range carriers. The rate allocated by the eNodeB to UE i from all in range carriers is given by r_i. Each application running on the UE is mathematically represented by a utility function $U_i(r_i)$ that corresponds to the application's type and represents the user satisfaction with its allocated rate r_i. Our goal is

to determine the optimal rates that the eNodeB shall allocate from each carrier to each UE in order to maximize the total system utility while ensuring proportional fairness between utilities.

The rate allocated to the i^{th} user in $\mathcal{M}$ by the j^{th} carrier in $\mathcal{K}$ is given by $r_i^{j,all}$. The final allocated rate by the eNodeB to the i^{th} user is given by

$$r_i = \sum_{j \in \mathcal{K}} r_i^{j,all} \tag{5.11}$$

where r_i is equivalent to the sum of rates allocated to the i^{th} user from all carriers in its range. Based on the coverage area of each carrier and the users' classes, a user grouping method is introduced in 5.3.1.1 to partition users into groups. The eNodeB performs resource allocation with user discrimination based on carrier aggregation to allocate each carrier's resources to users located within the coverage area of that carrier.

We express the user satisfaction with its rate using utility functions that represent the degree of satisfaction of the user function with the rate allocated by the cellular network [18, 22, 82, 83]. We represent the i^{th} user application utility function $U_i(r_i)$ by sigmoidal-like function expressed by equation (2.1) or logarithmic function expressed by equation (2.2) where r_i is the rate of the i^{th} user.

5.3.1.1 User Grouping Method

In this section we introduce a user grouping method to create user groups for each carrier $j \in \mathcal{K}$. The eNodeB creates a user group $\mathcal{M}_j$ for each carrier where $\mathcal{M}_j$ is a set of users located under the coverage area of the j^{th} carrier. The number of users in $\mathcal{M}_j$ is given by $M_j = |\mathcal{M}_j|$. Furthermore, users in $\mathcal{M}_j$ are partitioned into two groups of users. A VIP user group $\mathcal{M}_j^{\text{VIP}}$ and a regular user group $\mathcal{M}_j^{\text{Reg}}$, where $\mathcal{M}_j^{\text{VIP}}$ and $\mathcal{M}_j^{\text{Reg}}$ are the sets of all VIP users and regular users, respectively, located under the coverage area of the j^{th} carrier with

$\mathcal{M}_j = \mathcal{M}_j^{\text{VIP}} \cup \mathcal{M}_j^{\text{Reg}}$. The number of users in $\mathcal{M}_j^{\text{VIP}}$ and $\mathcal{M}_j^{\text{Reg}}$ is given by $M_j^{\text{VIP}} = |\mathcal{M}_j^{\text{VIP}}|$ and $M_j^{\text{Reg}} = |\mathcal{M}_j^{\text{Reg}}|$, respectively. The eNodeB allocates the j^{th} carrier resources to users in $\mathcal{M}_j$ with a priority given to VIP users (i.e. users in $\mathcal{M}_j^{\text{VIP}}$). Users located under the coverage area of multiple carriers (i.e. common users in multiple user groups) are allocated resources from these carriers and their final rates are aggregated under a non adjacent inter band aggregation scenario.

The i^{th} user is considered part of user group $\mathcal{M}_j$ if it is located within a distance of D_j from the eNodeB where D_j represents the coverage radius of the j^{th} carrier. Let d_i denotes the distance between the eNodeB and user i. The j^{th} carrier user group $\mathcal{M}_j$ is defined as

$$\mathcal{M}_j = \{i : d_i < D_j, 1 \leq i \leq M\}, 1 \leq j \leq K. \tag{5.12}$$

On the other hand, the eNodeB creates a set of carriers $\mathcal{K}_i$, for each user, that is defined as

$$\mathcal{K}_i = \{j : d_i < D_j, 1 \leq j \leq K\}, 1 \leq i \leq M. \tag{5.13}$$

The number of carriers that the i^{th} user can be allocated resources from is given by $N_i = |\mathcal{K}_i|$. Higher frequency carriers have smaller coverage radius than lower frequency carriers (i.e. $D_1 < D_2 < ... < D_K$). Therefore, user group $\mathcal{M}_1 \subseteq \mathcal{M}_2 \subseteq ... \subseteq \mathcal{M}_K$. Figure 5.14 shows one cellular cell with one eNodeB under non adjacent inter band scenario with K carriers in $\mathcal{K}$ and M users in $\mathcal{M}$ and how users are partitioned into user groups based on their location and their class.

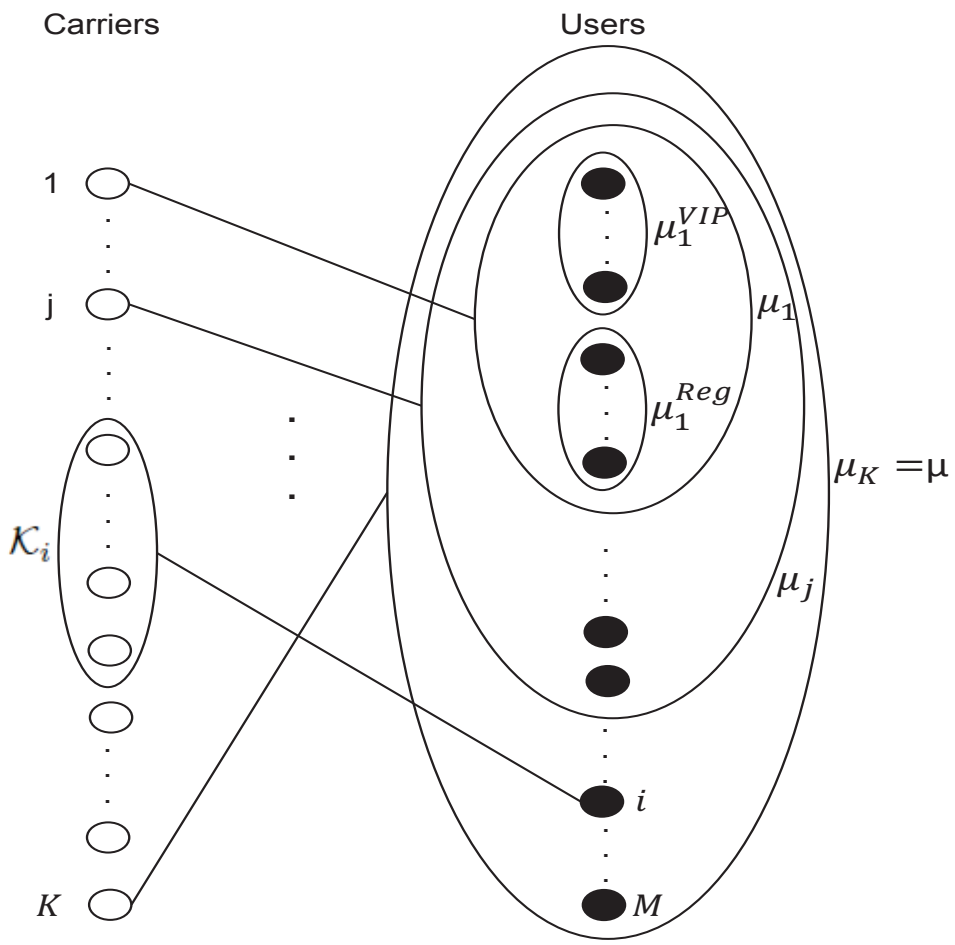


Figure 5.14: User grouping for a LTE mobile system with M users in $\mathcal{M}$ and K carriers in $\mathcal{K}$. $\mathcal{M}_j$ represents the set of users located under the coverage area of the j^{th} carrier with $\mathcal{M}_j = \mathcal{M}_j^{VIP} \cup \mathcal{M}_j^{Reg}$. $\mathcal{K}_i$ represents the set of all in range carriers for the i^{th} user.

5.3.2 Multi-Carrier Resource Allocation with User discrimination Optimization Problem

In this section, we present a multi-stage resource allocation (RA) with user discrimination optimization problem to allocate multi-carrier resources optimally among users in their coverage area. Our objective is to find the final allocated rate to each user from its all in range carriers based on a utility proportional fairness policy. We use utility functions of users rates to represent the type of application running on the UE. Every user subscribing for a mobile service is guaranteed to achieve a minimum QoS with priority criterion. VIP users are given

priority when allocating each carrier's resources and within each user class group, whether it is VIP or regular user group, real time applications are given priority when allocating each carrier's resources. This is due to the nature of sigmoidal-like utility functions that are used to represent real-time applications.

The eNodeB performs the resource allocation process for all carriers one at a time and one after another in ascending order of their coverage radius D_j. Each carrier $j \in \mathcal{K}$ has a limited amount of available resources that is given by R_j and each user's application has a minimum required rate r_i^{req} that is equivalent to zero in the case of regular users and is equivalent to certain value (i.e. rate) in the case of VIP users. The eNodeB starts the RA process by performing a RA for carrier 1 in $\mathcal{K}$ as it has the smallest coverage radius D_1. After allocating its resources to users in $\mathcal{M}_1$, the eNodeB then starts the RA process to allocate carrier 2 resources to users in $\mathcal{M}_2$. In addition, since $\mathcal{M}_1 \subseteq \mathcal{M}_2$ the eNodeB allocates users in $\mathcal{M}_1$ resources from carrier 2 and the rates are aggregated based on a non adjacent inter band aggregation scenario. The eNodeB continues the resource allocation process by allocating the j^{th} carrier resources to users in $\mathcal{M}_j$. Let $r_i^{j,\text{all}}$ represents the rate allocated by the j^{th} carrier to UE i and let C_i represents the total aggregated rate allocated to UE i by carriers $\{1, 2, ..., j-1\}$ where $C_i = \sum_{l=1}^{j-1} r_i^{l,\text{all}}$. Furthermore, let C_i^j be a constant that is always equivalent to zero for regular users whereas for VIP users C_i^j is equivalent to zero or $r_i^{\text{req}} - C_i$ based on some conditions that are discussed later in this section. The resource allocation process is finalized by allocating the K^{th} carrier resources to users in $\mathcal{M}_K$, i.e. all users in the cellular cell as they are all located within its coverage radius. We consider a utility proportional fairness objective function, based on carrier aggregation, that the eNodeB seeks to maximize for each time it allocates a carrier's resources.

The proposed RA optimization problem for multi-carrier cellular systems is divided into three cases. In order for the eNodeB to guarantee that VIP users are given priority when allocating each carrier's resources, each time the eNodeB performs a RA process for a car-

rier it checks the values of 1) the carrier's available resources R_j, 2) the current total rate allocated to each VIP UE $i \in \mathcal{M}_j^{\text{VIP}}$ from other carriers (i.e. $C_i = \sum_{l=1}^{j-1} r_i^{l,\text{all}}$) and 3) the value of $r_i^{\text{req}} - C_i$ for each VIP UE $i \in \mathcal{M}_j^{\text{VIP}}$ if $C_i < r_i^{\text{req}}$. Based on these values, the eNodB performs the RA process that corresponds to the most appropriate case among the three cases. The three cases and their RA optimization framework are presented below.

Case 1. *RA Optimization Problem when* $C_i \geq r_i^{req}$ $\forall i \in \mathcal{M}_j$:

The eNodeB chooses the RA optimization problem of this case in order to allocate the j^{th} carrier resources if the total aggregated rate C_i that is allocated to each UE $i \in \mathcal{M}_j$ from carriers $\{1, 2, ..., j-1\}$ is greater than or equal the minimum required application rate r_i^{req}. In this case, since each UE has already been allocated at least its application minimum required rate from other carriers, the eNodeB performs the RA process among all users under the coverage area of carrier j. The RA optimization problem for the j^{th} carrier in this case is given by:

$$\begin{aligned} \max_{\mathbf{r}^j} \quad & \prod_{i=1}^{M_j} U_i(C_i + C_i^j + r_i^j) \\ \text{subject to} \quad & \sum_{i=1}^{M_j} r_i^{j,\text{all}} \leq R_j, \;\; r_i^{j,\text{all}} \geq 0 \\ & r_i^{j,\text{all}} = r_i^j + C_i^j, \;\; C_i^j = 0 \\ & C_i = \sum_{l=1}^{j-1} r_i^{l,\text{all}}, \;\; C_i \geq r_i^{\text{req}}, \quad i = 1, 2, ..., M_j, \end{aligned} \tag{5.14}$$

where C_i^j is a constant that is equivalent to zero in this case, $U_i(C_i + C_i^j + r_i^j)$ is the utility function of the summation of the rate C_i allocated to the application running on the i^{th} user by carriers $\{1, 2, ..., j-1\}$ and the rate $r_i^{j,\text{all}}$ allocated to the same application by carrier j where $r_i^{j,\text{all}} = C_i^j + r_i^j$, $\mathbf{r}^j = \{r_1^j, r_2^j, ..., r_{M_j}^j\}$ and M_j is the number of users in $\mathcal{M}_j$ (i.e. both VIP and regular users) located under the coverage area of the j^{th} carrier. After the eNodeB performs the RA process for the j^{th} carrier by solving optimization problem

(5.14), the total rate allocated to each user by the eNodeB is equivalent to $C_i + r_i^{j,\text{all}}$. In optimization problem (5.14), we consider a utility proportional fairness objective function, based on carrier aggregation, that the eNodeB seeks to maximize when it performs RA for carrier j.

Case 2. *RA Optimization Problem when* $C_i < r_i^{req}$ *for any user* $i \in \mathcal{M}_j$ *and* $\sum_{i=1}^{M_j^{VIP}} q_i^j \geq R_j$ *where* $q_i^j = 0$ *if* $C_i \geq r_i^{req}$ *and* $q_i^j = r_i^{req} - C_i$ *if* $C_i < r_i^{req}$:

The eNodeB selects the optimization problem of this case to allocate the j^{th} carrier resources if the total aggregated rate C_i for any user i is less than the user's application minimum required rate r_i^{req} and $\sum_{i=1}^{M_j^{\text{VIP}}} q_i^j$ for VIP users in $\mathcal{M}_j^{\text{VIP}}$ is greater than or equal the carrier's available resources R_j. In this case, the eNodeB allocates the j^{th} carrier resources only to VIP UEs in $\mathcal{M}_j^{\text{VIP}}$ as they are considered more important and regular users in $\mathcal{M}_j^{\text{Reg}}$ are not allocated any of the j^{th} carrier resources since the carrier's resources are limited. The RA optimization problem for the j^{th} carrier in this case is given by:

$$
\begin{aligned}
\max_{\mathbf{r}^j} \quad & \prod_{i=1}^{M_j^{\text{VIP}}} U_i(C_i + C_i^j + r_i^j) \\
\text{subject to} \quad & \sum_{i=1}^{M_j^{\text{VIP}}} r_i^{j,\text{all}} \leq R_j, \;\; r_i^{j,\text{all}} \geq 0 \\
& C_i = \sum_{l=1}^{j-1} r_i^{l,\text{all}}, \;\; r_i^{j,\text{all}} = r_i^j + C_i^j \\
& C_i^j = 0 \\
& q_i^j = \begin{cases} 0 & \text{if } \; C_i \geq r_i^{\text{req}} \\ r_i^{\text{req}} - C_i & \text{if } \; C_i < r_i^{\text{req}} \end{cases} \\
& \sum_{i=1}^{M_j^{\text{VIP}}} q_i^j \geq R_j, \quad i = 1, 2, ..., M_j^{\text{VIP}},
\end{aligned}
\tag{5.15}
$$

where $\mathbf{r}^j = \{r_1^j, r_2^j, ..., r_{M_j^{\text{VIP}}}^j\}$, $C_i^j = 0$ and M_j^{VIP} is the number of users in $\mathcal{M}_j^{\text{VIP}}$. After the

eNodeB performs the RA process for the j^{th} carrier by solving optimization problem (5.15), each VIP user in $\mathcal{M}_j^{\text{VIP}}$ is allocated a rate that is equivalent to $r_i^{j,\text{all}}$ by carrier j whereas users in $\mathcal{M}_j^{\text{Reg}}$ are not allocated any of the j^{th} carrier resources. The total rate allocated by the eNodeB to each user is equivalent to $C_i + r_i^{j,\text{all}}$. In optimization problem (5.15), we consider a utility proportional fairness objective function, based on carrier aggregation, that the eNodeB seeks to maximize when it performs RA for carrier j.

Case 3. *RA Optimization Problem when* $C_i < r_i^{req}$ *for any user* $i \in \mathcal{M}_j^{VIP}$ *and* $\sum_{i=1}^{M_j^{VIP}} q_i^j < R_j$ *where* $q_i^j = 0$ *if* $C_i \geq r_i^{req}$ *and* $q_i^j = r_i^{req} - C_i$ *if* $C_i < r_i^{req}$:

The eNodeB selects the optimization problem of this case to allocate the j^{th} carrier resources if the total aggregated rate C_i for any user i is less than the user's application minimum required rate r_i^{req} and the summation $\sum_{i=1}^{M_j^{\text{VIP}}} q_i^j$ for VIP users in $\mathcal{M}_j^{\text{VIP}}$ is less than the carrier's available resources R_j. In this case, the eNodeB allocates the j^{th} carrier resources to all UEs in $\mathcal{M}_j$. The RA optimization problem for the j^{th} carrier in this case is given by:

$$
\begin{aligned}
\max_{\mathbf{r}^j} \quad & \prod_{i=1}^{M_j} U_i(C_i + C_i^j + r_i^j) \\
\text{subject to} \quad & \sum_{i=1}^{M_j} r_i^{j,\text{all}} \leq R_j, \;\; r_i^{j,\text{all}} \geq 0 \\
& C_i = \sum_{l=1}^{j-1} r_i^{l,\text{all}}, \;\; r_i^{j,\text{all}} = r_i^j + C_i^j \\
& C_i^j = \begin{cases} 0 & \text{if } \; C_i \geq r_i^{\text{req}} \\ r_i^{\text{req}} - C_i & \text{if } \; C_i < r_i^{\text{req}} \end{cases} \\
& q_i^j = \begin{cases} 0 & \text{if } \; C_i \geq r_i^{\text{req}} \\ r_i^{\text{req}} - C_i & \text{if } \; C_i < r_i^{\text{req}} \end{cases} \\
& \sum_{i=1}^{M_j^{\text{VIP}}} q_i^j < R_j, \;\; i = 1, 2, ..., M_j,
\end{aligned}
\tag{5.16}
$$

where $\mathbf{r}^j = \{r_1^j, r_2^j, ..., r_{M_j}^j\}$ and M_j is the number of users in $\mathcal{M}_j$. After the eNodeB performs the RA process for the j^{th} carrier by solving optimization problem (5.16), each user in $\mathcal{M}_j$ is allocated a rate that is equivalent to $r_i^{j,\text{all}}$ by carrier j and the total rate allocated by the eNodeB to each user is equivalent to $C_i + r_i^{j,\text{all}}$. In optimization problem (5.16), we consider a utility proportional fairness objective function, based on carrier aggregation, that the eNodeB seeks to maximize when it performs RA for carrier j.

Each of the three RA optimization problems (5.14), (5.15) and (5.16) of the j^{th} carrier can

be expressed by the following generalized optimization problem:

$$\begin{aligned}
\max_{\mathbf{r}^j} \quad & \prod_{i=1}^{|\alpha_j|} U_i(C_i + C_i^j + r_i^j) \\
\text{subject to} \quad & \sum_{i=1}^{|\alpha_j|} r_i^{j,\text{all}} \le R_j, \;\; r_i^{j,\text{all}} \ge 0 \\
& C_i = \sum_{l=1}^{j-1} r_i^{l,\text{all}}, \;\; r_i^{j,\text{all}} = r_i^j + C_i^j \\
& q_i^j = \begin{cases} 0 & \text{if } \; C_i \ge r_i^{\text{req}} \\ r_i^{\text{req}} - C_i & \text{if } \; C_i < r_i^{\text{req}} \end{cases} \\
& i = 1, 2, ..., |\alpha_j|,
\end{aligned} \tag{5.17}$$

where C_i^j and α_j in (5.17) are given by

$$C_i^j = \begin{cases} 0 & \text{if } \; C_i \ge r_i^{\text{req}} \\ r_i^{\text{req}} - C_i & \text{if } \; C_i < r_i^{\text{req}} \;\text{ and }\; \sum_{i=1}^{|\mathcal{M}_j^{\text{VIP}}|} q_i^j < R_j \\ 0 & \text{if } \; C_i < r_i^{\text{req}} \;\text{ and }\; \sum_{i=1}^{|\mathcal{M}_j^{\text{VIP}}|} q_i^j \ge R_j \end{cases}$$

$$\alpha_j = \begin{cases} \mathcal{M}_j & \text{if } \; C_i \ge r_i^{\text{req}} \;\; \forall i \in \mathcal{M}_j \\ \mathcal{M}_j^{\text{VIP}} & \text{if } \; C_i < r_i^{\text{req}} \text{ for any user } i \in \mathcal{M}_j \\ & \;\; \text{and } \sum_{i=1}^{M_j^{\text{VIP}}} q_i^j \ge R_j \\ \mathcal{M}_j & \text{if } \; C_i < r_i^{\text{req}} \text{ for any user } i \in \mathcal{M}_j^{\text{VIP}} \\ & \;\; \text{and } \sum_{i=1}^{M_j^{\text{VIP}}} q_i^j < R_j \end{cases} \tag{5.18}$$

where $\mathbf{r}^j = \{r_1^j, r_2^j, ..., r_{|\alpha_j|}^j\}$, α_j is a set of users located under the coverage area of carrier j that is equivalent to $\mathcal{M}_j$ or $\mathcal{M}_j^{\text{VIP}}$ based on certain conditions as shown in (5.18) and $|\alpha_j|$ is the number of users in α_j.

The objective function in optimization problem (5.17) is equivalent to $\sum_{i=1}^{|\alpha_j|} \log U_i(C_i + C_i^j + r_i^j)$. Later in this section we prove that optimization problem (5.17) is a convex optimization problem and there exists a unique tractable global optimal solution. Once the eNodeB is done performing the RA process, for the j^{th} carrier, by solving optimization problem (5.17), each user in α_j is allocated a rate that is equivalent to $r_i^{j,\text{all}} = r_i^j + C_i^j$ and the user's total aggregated rate allocated by the eNodeB from carriers $\{1, 2, ..., j\}$ is given by $\sum_{l=1}^{j} r_i^{l,\text{all}}$.

Lemma 5.3.1. *The utility functions* $\log U_i(C_i + C_i^j + r_i^j)$ *in optimization problem (5.17) are strictly concave functions.*

Proof. The utility functions are assumed to be logarithmic functions expressed by equation (2.2) or sigmoidal-like functions expressed by equation (2.1). Therefore, $U_i(C_i + C_i^j + r_i^j)$ is a strictly concave (i.e. in the case of logarithmic utility functions) or a sigmoidal-like function of the total aggregated rate $C_i + C_i^j + r_i^j$ allocated to user i application from carriers $\{1, 2, ..., j\}$ after performing the RA process of the j^{th} carrier by the eNodeB.

In the case of logarithmic utility function, recall the utility function properties in Chapter 2 Section 2.1, the utility function of the application rate is positive, increasing and twice differentiable with respect to the application rate. It follows that $U_i'(C_i + C_i^j + r_i^j) = \frac{dU_i(C_i+C_i^j+r_i^j)}{dr_i^j} > 0$ and $U_i''(C_i + C_i^j + r_i^j) = \frac{d^2U_i(C_i+C_i^j+r_i^j)}{{dr_i^j}^2} < 0$, i.e. since $C_i + C_i^j$ is greater or equal zero. Then the function $\log U_i(C_i + C_i^j + r_i^j)$ has $\frac{d\log(U_i(C_i+C_i^j+r_i^j))}{dr_i^j} = \frac{U_i'(C_i+C_i^j+r_i^j)}{U_i(C_i+C_i^j+r_i^j)} > 0$ and $\frac{d^2\log(U_i(C_i+C_i^j+r_i^j))}{{dr_i^j}^2} = \frac{U_i''(C_i+C_i^j+r_i^j)U_i(C_i+C_i^j+r_i^j) - U_i'^2(C_i+C_i^j+r_i^j)}{U_i^2(C_i+C_i^j+r_i^j)} < 0$. Therefore, the natural logarithm of the logarithmic utility function $\log(U_i(C_i + C_i^j + r_i^j))$ is strictly concave.

On the other hand, in the case of sigmoidal-like utility function, the normalized sigmoidal-like function is given by $U_i(C_i + C_i^j + r_i^j) = c_i\left(\frac{1}{1+e^{-a_i(C_i+C_i^j+r_i^j-b_i)}} - d_i\right)$. For $0 < r_i^j < (R_j - C_i^j)$,

we have

$$
\begin{aligned}
0 &< c_i\left(\frac{1}{1+e^{-a_i(C_i+C_i^j+r_i^j-b_i)}}-d_i\right)<1\\
d_i &< \frac{1}{1+e^{-a_i(C_i+C_i^j+r_i^j-b_i)}} < \frac{1+c_id_i}{c_i}\\
\frac{1}{d_i} &> 1+e^{-a_i(C_i+C_i^j+r_i^j-b_i)} > \frac{c_i}{1+c_id_i}\\
0 &< 1-d_i(1+e^{-a_i(C_i+C_i^j+r_i^j-b_i)}) < \frac{1}{1+c_id_i}
\end{aligned}
$$

It follows that for $0 < r_i^j < (R_j - C_i^j)$, we have the first and second derivatives as

$$
\begin{aligned}
&\frac{d}{dr_i^j}\log U_i(C_i+C_i^j+r_i^j) =\\
&\qquad\frac{a_id_ie^{-a_i(C_i+C_i^j+r_i^j-b_i)}}{1-d_i(1+e^{-a_i(C_i+C_i^j+r_i^j-b_i)})}\\
&\qquad+\frac{a_ie^{-a_i(C_i+C_i^j+r_i^j-b_i)}}{(1+e^{-a_i(C_i+C_i^j+r_i^j-b_i)})}>0\\
&\frac{d^2}{dr_i^{j\,2}}\log U_i(C_i+C_i^j+r_i^j) =\\
&\qquad\frac{-a_i^2d_ie^{-a_i(C_i+C_i^j+r_i^j-b_i)}}{c_i\left(1-d_i(1+e^{-a(C_i+C_i^j+r_i^j-b_i)})\right)^2}\\
&\qquad+\frac{-a_i^2e^{-a_i(C_i+C_i^j+r_i^j-b_i)}}{(1+e^{-a_i(C_i+C_i^j+r_i^j-b_i)})^2}<0.
\end{aligned}
$$

Therefore, the natural logarithm of the sigmoidal-like utility function $\log(U_i(C_i + C_i^j + r_i^j)$ is strictly concave function. Therefore, the utility functions natural logarithms have strictly concave natural logarithms in both cases of logarithmic utility functions and sigmoidal-like utility functions. □

Theorem 5.3.2 proves the convexity of optimization problem (5.17).

Theorem 5.3.2. *Optimization problem (5.17) is a convex optimization problem and there exists a unique tractable global optimal solution.*

Proof. It follows from Lemma 5.3.1 that all UEs utility functions of applications rates are strictly concave. Therefore, optimization problem (5.17) is a convex optimization problem. For a convex optimization problem there exists a unique tractable global optimal solution [89]. □

5.3.3 RA Optimization Algorithm

In this section, we present our multi-carrier resource allocation with user discrimination algorithm. The proposed algorithm consists of UE and eNodeB parts shown in Algorithm 20 and Algorithm 21, respectively. The execution of the algorithm starts by UEs, subscribing for mobile services, transmitting their application utility parameters to the eNodeB, which allocates available carriers' resources to UEs based on a proportional fairness policy. First, the eNodeB performs the user grouping method described in Section 5.3.1.1 for each carrier by creating three user group sets $\mathcal{M}_j^{\text{VIP}}$, $\mathcal{M}_j^{\text{Reg}}$ and $\mathcal{M}_j$ for UEs located within the coverage area of the j^{th} carrier. It then starts performing the RA process to allocate the carriers resources starting with carrier 1 in $\mathcal{K}$ (i.e. the carrier with the smallest coverage radius) in ascending order $1 \rightarrow K$. In order to allocate certain carrier's resources, the eNodeB performs the RA process that corresponds to the most appropriate case among the three cases presented in Section 5.3.2. From optimization problem (5.17), we have the following Lagrangian

$$L(\mathbf{r}^j, p^j) = \sum_{i=1}^{|\alpha_j|} \log U_i(C_i + C_i^j + r_i^j) - p^j(\sum_{i=1}^{|\alpha_j|}(C_i^j + r_i^j) + \sum_{i=1}^{|\alpha_j|} z_i - R_j), \tag{5.19}$$

where $z_i \geq 0$ is the slack variable and p^j is Lagrange multiplier that represents the shadow price (price per unit bandwidth for all the $|\alpha_j|$ channels). The rates, solutions to equation (5.17), are the values r_i^j which solve equation $\frac{\partial \log U_i(C_i+C_i^j+r_i^j)}{\partial r_i^j} = p^j$ and are the intersection of the time varying shadow price, horizontal line $y = p^j$, with the curve $y = \frac{\partial \log U_i(C_i+C_i^j+r_i^j)}{\partial r_i^j}$ geometrically. The rate allocated by carrier j to the i^{th} UE is equivalent to $r_i^{j,\text{all}} = r_i^j + C_i^j$. When the eNodeB is done allocating the K^{th} carrier resources, each user is then allocated its final aggregated rate $r_i = \sum_{j=1}^{K} r_i^{j,\text{all}}$.

Algorithm 20 The i^{th} UE Algorithm

loop
Send application utility parameters k_i, a_i, b_i, $r_i^{\max}$ and r_i^{req} to eNodeB.
Receive the final allocated rate r_i from the eNodeB.
end loop

5.3.4 Simulation Results

Algorithm 20 and 21 were applied in C++ to multiple utility functions with different parameters. Simulation results showed convergence to the global optimal rates. In this section, we consider a mobile cell with one eNodeB, two carriers with available resources and 8 active UEs located under the coverage area of the eNodeB as shown in Figure 5.15. The UEs are divided into two groups. The 1^{st} group of UEs (index $i = \{1, 2, 3, 4\}$) represents user group $\mathcal{M}_1$ located within the coverage radius D_1 of carrier 1. Each user in $\mathcal{M}_1$ belongs to one of the two classes of user groups, i.e. VIP user group and Regular user group, where $\mathcal{M}_1^{\text{VIP}} = \{2, 4\}$, $\mathcal{M}_1^{\text{Reg}} = \{1, 3\}$ and $\mathcal{M}_1 = \mathcal{M}_1^{\text{VIP}} \cup \mathcal{M}_1^{\text{Reg}}$. On the other hand, the 2^{nd} group of UEs (index $i = \{1, 2, 3, 4, 5, 6, 7, 8\}$) represents user group $\mathcal{M}_2$ located within the coverage

Algorithm 21 The eNodeB Algorithm

loop
  Initialize $C_i = 0$; $C_i^j = 0$; $r_i^{j,\text{all}} = 0$.
  Receive application utility parameters k_i, a_i, b_i, $r_i^{\max}$ and r_i^{req} from all UEs in $\mathcal{M}$.
  for $j \leftarrow 1$ to K **do**
    Create user groups $\mathcal{M}_j^{\text{VIP}}$, $\mathcal{M}_j^{\text{Reg}}$ and $\mathcal{M}_j$ for UEs located within the coverage area of the j^{th} carrier.
  end for
  for $i \leftarrow 1$ to $|\mathcal{M}_j|$ **do**
    Create carrier group $\mathcal{K}_i$ for the i^{th} UE's all in range carriers.
  end for
  for $j \leftarrow 1$ to K **do**
    if $C_i < r_i^{\text{req}}$ **then**
      $q_i^j = r_i^{\text{req}} - C_i$
    else
      $q_i^j = 0$
    end if
    if $C_i \geq r_i^{\text{req}}\ \forall i \in \mathcal{M}_j$ **then**
      $C_i^j = 0$
      Solve $\mathbf{r}^j = \arg\max_{\mathbf{r}^j} \sum_{i=1}^{|\mathcal{M}_j|} \log U_i(C_i + C_i^j + r_i^j) - p^j(\sum_{i=1}^{|\mathcal{M}_j|}(r_i^j + C_i^j) - R_j)$.
      Allocate rate $r_i^{j,\text{all}} = r_i^j + C_i^j$ by the j^{th} carrier to each user in $\mathcal{M}_j$.
      Calculate new $C_i = C_i + r_i^{j,\text{all}}\ \forall i \in \mathcal{M}_j$
    else if $C_i < r_i^{\text{req}}$ for any user $i \in \mathcal{M}_j$ && $\sum_{i=1}^{M_j^{\text{VIP}}} q_i^j \geq R_j$ **then**
      $C_i^j = 0$
      Solve $\mathbf{r}^j = \arg\max_{\mathbf{r}^j} \sum_{i=1}^{|\mathcal{M}_j^{\text{VIP}}|} \log U_i(C_i + C_i^j + r_i^j) - p^j(\sum_{i=1}^{|\mathcal{M}_j^{\text{VIP}}|}(r_i^j + C_i^j) - R_j)$.
      Allocate rate $r_i^{j,\text{all}} = r_i^j + C_i^j$ by the j^{th} carrier to each user in $\mathcal{M}_j^{\text{VIP}}$.
      Calculate new $C_i = C_i + r_i^{j,\text{all}}\ \forall i \in \mathcal{M}_j^{\text{VIP}}$
    else if $C_i < r_i^{\text{req}}$ for any user $i \in \mathcal{M}_j^{\text{VIP}}$ and $\sum_{i=1}^{|\mathcal{M}_j^{\text{VIP}}|} q_i^j < R_j$ **then**
      if $C_i < r_i^{\text{req}}$ **then**
        $C_i^j = r_i^{\text{req}} - C_i$
      else
        $C_i^j = 0$
      end if
      Solve $\mathbf{r}^j = \arg\max_{\mathbf{r}^j} \sum_{i=1}^{|\mathcal{M}_j|} \log U_i(C_i + C_i^j + r_i^j) - p^j(\sum_{i=1}^{|\mathcal{M}_j|}(r_i^j + C_i^j) - R_j)$.
      Allocate rate $r_i^{j,\text{all}} = r_i^j + C_i^j$ by the j^{th} carrier to each user in $\mathcal{M}_j$.
      Calculate new $C_i = C_i + r_i^{j,\text{all}}\ \forall i \in \mathcal{M}_j$
    end if
  end for
  Allocate total aggregated rate $r_i = \sum_{j=1}^{K} r_i^{j,\text{all}}$ by the eNodeB to each UE i in $\mathcal{M}$
end loop

Table 5.1: Users and their applications utilities

Applications Utilities Parameters		Users Indexes
Sig1	Sig $a_i = 5,\ b_i = 10$	$i = \{5\}$
Sig2	Sig $a_i = 3,\ b_i = 20$	$i = \{1\}$
Sig3	Sig $a_i = 1,\ b_i = 30$	$i = \{2, 6\}$
Log1	Log $k_i = 15,\ r_i^{\max} = 100$	$i = \{7\}$
Log2	Log $k_i = 3,\ r_i^{\max} = 100$	$i = \{3\}$
Log3	Log $k_i = 0.5,\ r_i^{\max} = 100$	$i = \{4, 8\}$

radius D_2 of carrier 2. Each user in $\mathcal{M}_2$ belongs to a VIP user group or a regular user group where $\mathcal{M}_2^{\text{VIP}} = \{2, 4, 6, 8\}$, $\mathcal{M}_2^{\text{Reg}} = \{1, 3, 5, 7\}$ and $\mathcal{M}_2 = \mathcal{M}_2^{\text{VIP}} \cup \mathcal{M}_2^{\text{Reg}}$.

We use sigmoidal-like utility functions and logarithmic utility functions with different parameters to represent each of the users' applications. We use three normalized sigmoidal-like functions that are expressed by equation (2.1) with different parameters. The used parameters are $a_i = 5$, $b_i = 10$ that correspond to a sigmoidal-like function with inflection point $r_i = 10$ which represents the utility of UE with index $i = \{5\}$, $a_i = 3$, $b_i = 20$ that correspond to a sigmoidal-like function with inflection point $r_i = 20$ which represents the utility of UE with index $i = \{1\}$, and $a_i = 1$, $b_i = 30$ that correspond to a sigmoidal-like function with inflection point $r_i = 30$ which represents the utility of UEs with indexes $i = \{2, 6\}$, as shown in Figure 5.16. We use three logarithmic functions expressed by equation (2.2) with $r_i^{\max} = 100$ and different k_i parameters to represent delay-tolerant applications. We use $k_i = 15$ for UE with index $i = \{7\}$, $k_i = 3$ for UE with index $i = \{3\}$, and $k_i = 0.5$ for UEs with indexes $i = \{4, 8\}$, as shown in Figure 5.16. A summary is shown in table 5.1. We use an application minimum required rate that is equivalent to the inflection point of the sigmoidal-like function, i.e. $r_i^{\text{req}} = b_i$, for each VIP user running a real-time application, we use $r_i^{\text{req}} = 15$ for each VIP user running a delay-tolerant application and $r_i^{\text{req}} = 0$ for each regular user whether it is running real-time application or delay-tolerant application.

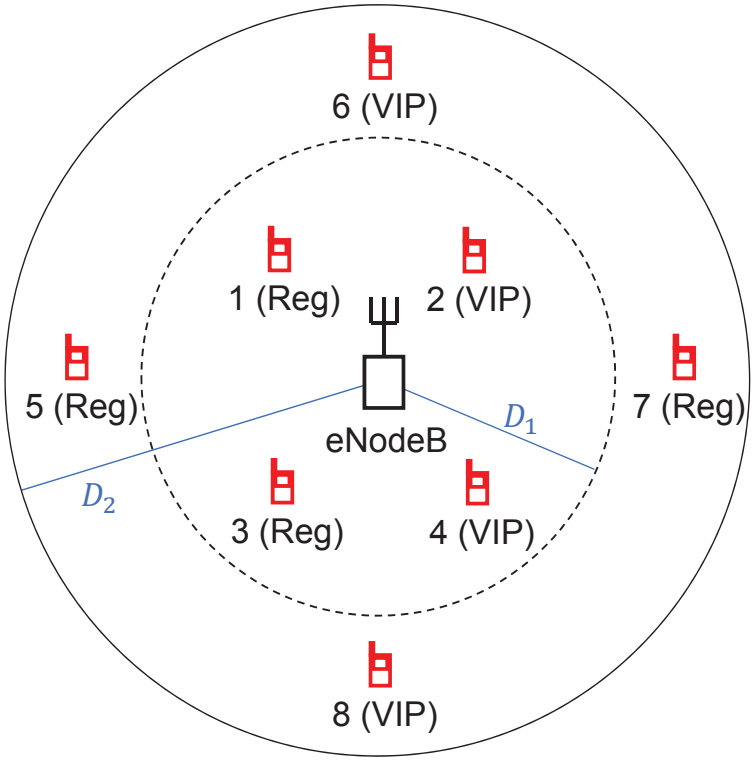


Figure 5.15: System model for a mobile system with $M = 8$ users and $K = 2$ carriers available at the eNodeB. Carrier 1 coverage radius is D_1 and carrier 2 coverage radius is D_2 with $D_1 < D_2$. $\mathcal{M}_1 = \{1, 2, 3, 4\}$ and $\mathcal{M}_2 = \{1, 2, ..., 8\}$ represent the sets of user groups located under the coverage area of carrier 1 and carrier 2, respectively.

5.3.4.1 Carrier 1 Allocated Rates for $60 \leq R_1 \leq 150$

In the following simulations, we set $\delta = 10^{-3}$, carrier 1 rate R_1 takes values between 60 and 150 with step of 10. In Figure 5.17, we show the allocated rates $r_i^{1,\text{all}}$ of different users with different values of carrier 1 total rate R_1 and observe how the proposed rate allocation algorithm converges for different values of R_1. In Figure 5.17, we show that both VIP and regular users in user group $\mathcal{M}_1$ are allocated resources by carrier 1 when $60 \leq R_1 \leq 150$ since carrier 1 available resources R_1 is greater than the total applications minimum required rates for users in $\mathcal{M}_1$. Figure 5.17 also shows that by using the proposed RA with user discrimination algorithm, no user is allocated zero rate (i.e. no user is dropped). However, carrier 1 resources are first allocated to the VIP users until each of their applications reaches

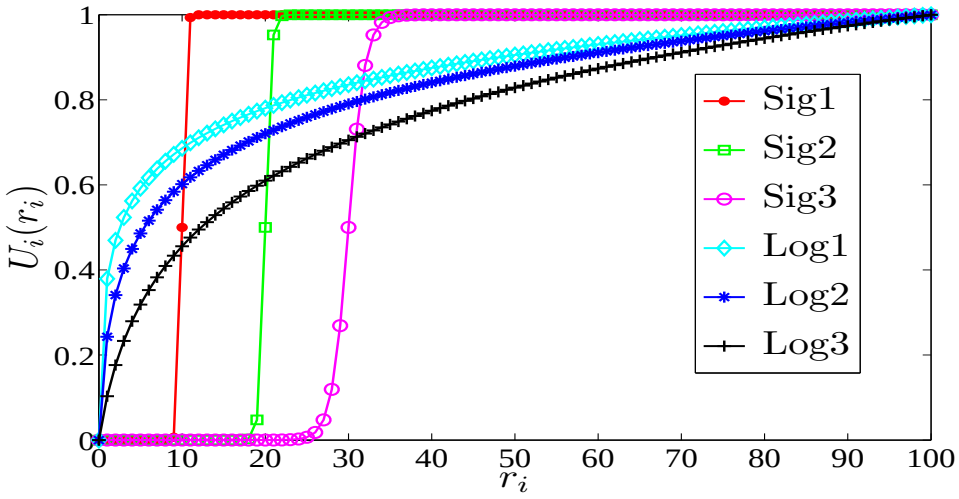


Figure 5.16: The users utility functions $U_i(r_i)$ used in the simulation (three sigmoidal-like functions and three logarithmic functions).

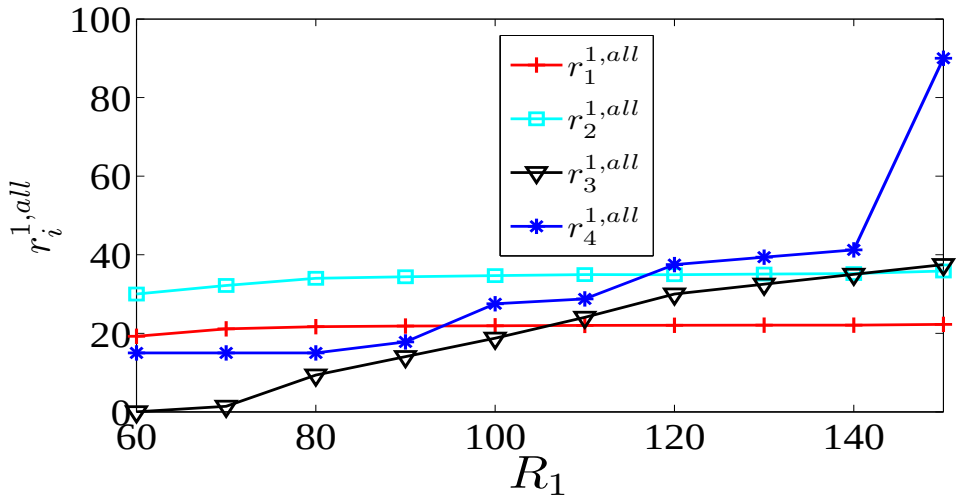


Figure 5.17: The rates $r_i^{1,\text{all}}$ allocated from carrier 1 to $\mathcal{M}_1$ user group with carrier 1 available resources $60 < R_1 < 150$.

the application minimum required rate r_i^{req}. Then the majority of carrier 1 resources are allocated to the UEs running adaptive real-time applications until they reach their inflection rates, the eNodeB then allocates more of carrier 1 resources to UEs with delay-tolerant applications.

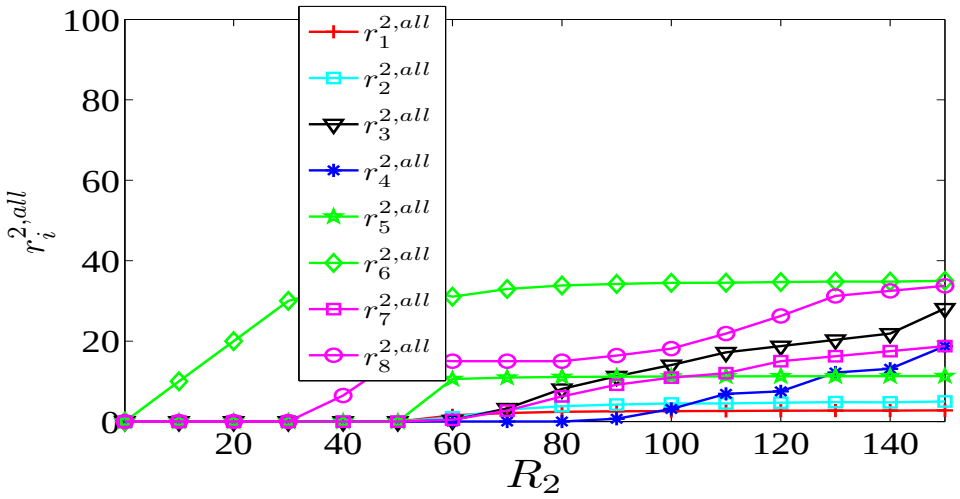
5.3.4.2 Carrier 2 Allocated Rates and the Total Aggregated Rates for $10 \leq R_2 \leq 150$

In the following simulations, we set $\delta = 10^{-3}$, carrier 2 rate R_2 takes values between 10 and 150 with step of 10 and carrier 1 rate is fixed at $R_1 = 60$. In Figure 5.18, we show the allocated rates $r_i^{2,\text{all}}$ and the final aggregated rates r_i of different users with different values of carrier 2 total rate R_2 and observe how the proposed rate allocation algorithm converges for different values of R_2. In Figure 5.18(a), we show that when $10 \leq R_2 \leq 45$ only VIP users in $\mathcal{M}_2$ (i.e. UEs in $\mathcal{M}_2^{\text{VIP}}$) that were not allocated resources by carrier 1 or did not reach their applications minimum required rates are allocated resources by carrier 2. Whereas when $45 < R_2 \leq 150$, both VIP and regular users in $\mathcal{M}_2$ are allocated resources by carrier 2 as carrier 2 total rate R_2 is greater than $\sum_{i=1}^{M_2^{\text{VIP}}} q_i^2$ (i.e. the total required rates for UEs to reach their r_i^{req}). Figure 5.18(a) also shows that by using the proposed RA with user discrimination algorithm that is based on carrier aggregation, the eNodeB takes into consideration the rates allocated to users in $\mathcal{M}_2$ by carrier 1 when allocating carrier 2 resources. Carrier 2 resources are first allocated to VIP users until each of their applications reaches the application minimum required rate r_i^{req}. Then the majority of carrier 2 resources are allocated to the UEs running adaptive real-time applications until they reach their inflection rates, the eNodeB then allocates more of carrier 2 resources to UEs with delay-tolerant applications.

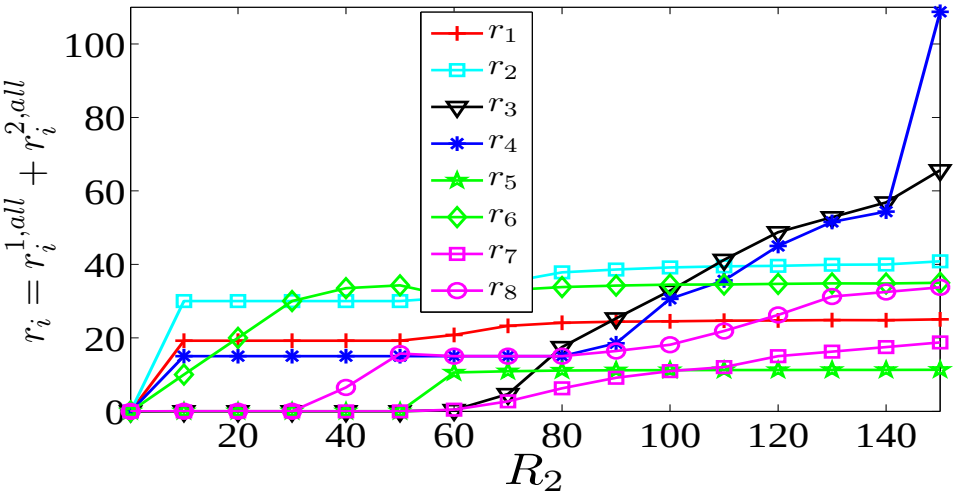
Figure 5.18(b) shows the total aggregated rates $r_i = \sum_{j=1}^{2} r_i^{j,\text{all}}$ for the 8 users.

5.3.4.3 Pricing Analysis for Carrier 1 and Carrier 2

In the following simulations, we set $\delta = 10^{-3}$. In Figure 5.19, we show carrier 1 shadow price with $60 \leq R_1 \leq 150$. We observe that carrier 1 price p^1 is traffic-dependant as it decreases for higher values of R_1. In Figure 5.20, we show the offered price of carrier 2



(a) The rates $r_i^{2,\text{all}}$ allocated from carrier 2 to $\mathcal{M}_2$ user group.



(b) The total aggregated rates r_i allocated by the eNodeB to the 8 users.

Figure 5.18: The rates $r_i^{2,\text{all}}$ allocated from carrier 2 to users in $\mathcal{M}_2$ and the total aggregated rates allocated to the 8 users with carrier 2 available resources $10 < R_2 < 150$ and carrier 1 resources fixed at $R_1 = 60$.

with $10 \leq R_2 \leq 150$ and $R_1 = 60$. We observe that p^2 decreases when R_2 increases for $10 \leq R_2 \leq 45$, only VIP users are allocated rates by carrier 2 when $10 \leq R_2 \leq 45$. However, we observe a jump in the price when $R_2 = 50$ as more users are considered in the rate allocation process (i.e VIP users and regular users in $\mathcal{M}_2$). Figure 5.20 also shows that carrier 2 price p^2 decreases when R_2 increases for $50 \leq R_2 \leq 150$.

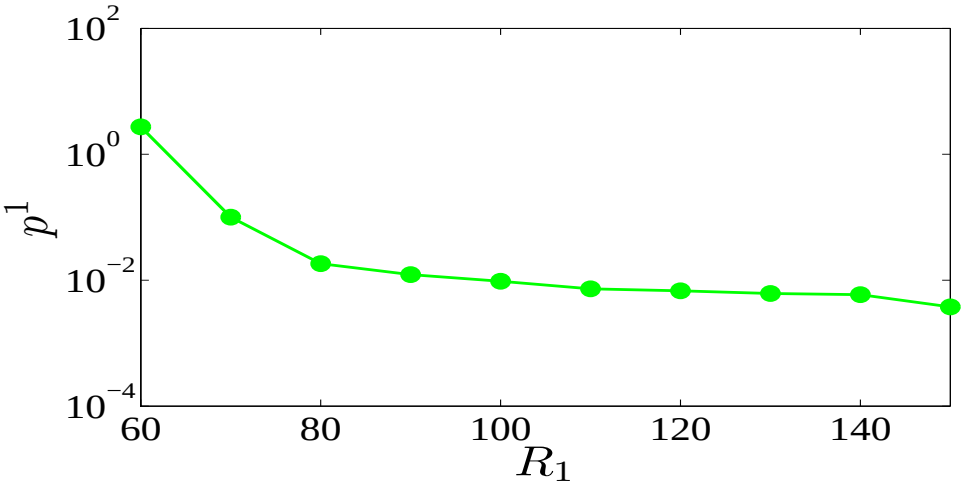


Figure 5.19: Carrier 1 shadow price p^1 with carrier 1 resources $60 < R_1 < 150$.

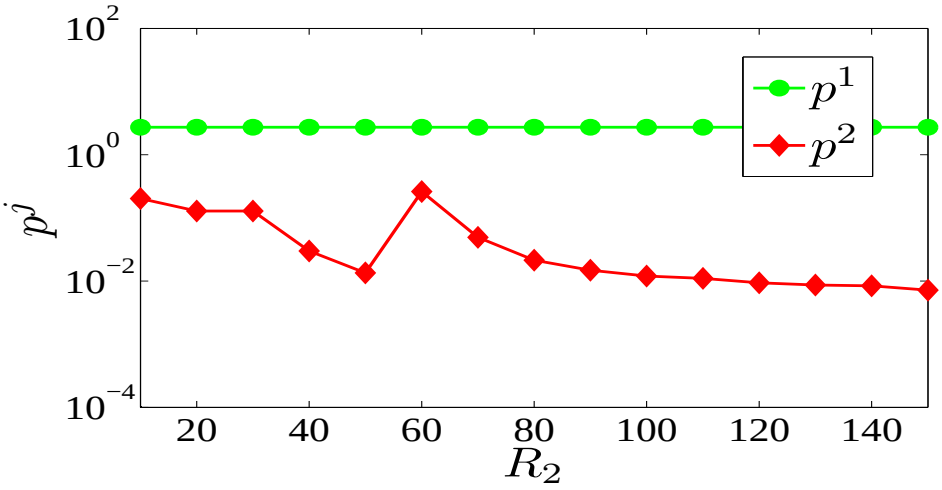


Figure 5.20: Carrier 1 shadow price p^1 and carrier 2 shadow price p^2 with carrier 2 resources $10 < R_2 < 150$ and carrier 1 resources fixed at $R_1 = 60$.

5.4 Summary and Conclusions

In this chapter, we introduced a framework for the problem of resource allocation with user discrimination in cellular systems. In section 5.1, We presented a spectrum sharing approach between two types of users; i.e. public safety users and commercial users, running delay tolerant or real-time applications. We proposed an iterative decentralized RA algorithm for the eNodeB and both the public safety and commercial UEs. The algorithm provides a

utility proportional fair resource allocation which guarantees a minimum QoS based on the public safety UEs application target rates, the group that the UE belongs to and the eNodeB available resources. The public safety users group is given priority over the commercial users group and within each group, users running real time applications are prioritized over those running delay tolerant applications. We showed through simulations that our algorithm converges to the optimal rates.

In section 5.2, we proposed a novel RA approach to allocate a single eNodeB resources optimally among multi-application UEs with different priority. Two cases of RA optimization problems are considered in our approach. The two cases are based on the total applications target rates of the VIP UEs compared to the eNodeB available resources. A two-stage RA algorithm is presented for each case to allocate the eNodeB resources among users and their applications. Different parameters are taken into consideration by our algorithm when allocating resources such as the application type, the application target rate (if the user application has one), the user subscription weight and the application weight. We showed through simulations that our two-stage RA algorithm converges to the optimal rates.

In section 5.3, we proposed an efficient resource allocation with user discrimination approach for 5G systems to allocate multiple carriers resources optimally among UEs that belong to different user groups classes. We used utility functions to represent the applications running on the UEs. Each user is assigned a minimum required application rate based on its class and the type of its application. Users are partitioned into different user groups based on their class and the carriers coverage area. We presented resource allocation optimization problems based on carrier aggregation for different cases. We proved the existence of a tractable global optimal solution. We presented a RA algorithm for allocating resources from different carriers optimally among different classes of mobile users. The proposed algorithm ensures fairness in the utility percentage, gives priority to VIP users and within a VIP or a regular user group it gives priority to adaptive real-time applications while providing a

minimum QoS for all users. We showed through simulations that the proposed resource allocation algorithm converges to the optimal rates. We also showed that the pricing provided by our algorithm depends on the traffic load.

Chapter 6

RA with CA for a Cellular System Sharing Spectrum with S-band Radar

As a result of the high demand for spectrum by commercial wireless operators, federal agencies are now willing to share their spectrum with commercial users. The 3550-3650 MHz band, currently used for military radar operations, is identified for spectrum sharing between military radars and communication systems, according to the NTIA's 2010 Fast Track Report [11]. This band is very favorable for commercial cellular systems such as LTE-Advanced systems. However, radar interference to cellular systems is a cause of concern for commercial operators and thus innovative methods are required to make spectrum sharing between radars and cellular systems a reality.

In this chapter, we consider a LTE-Advance cellular system sharing the $3550 - 3650$ MHz band with a MIMO radar. The LTE-Advanced cellular system has N_{BS} base stations. In order to mitigate radar interference, a spectrum sharing algorithm is proposed. The algorithm selects the best interference channel for radar's signal projection to mitigate radar interference to the i^{th} BS. We consider a MIMO colocated radar mounted on a ship. Colocated radars have improved spatial resolution over widely-spaced radars [90]. The LTE cellular

system operates in its regular licensed band and shares the 3.5 GHz band with a MIMO radar in order to increase its capacity such that the two systems do not cause interference to each other. We focus on finding an optimal solution for the resource allocation with carrier aggregation problem to allocate the LTE-Advanced BS/eNodeB and the available MIMO radar resources optimally among users subscribing for a service in the cellular cell coverage area. Each user is assigned a utility function based on the application running on its UE. Real-time applications are represented by sigmoidal-like utility functions whereas delay-tolerant applications are represented by logarithmic utility functions. Real-time applications are given the priority when allocating resources. A resource allocation with carrier aggregation algorithm is proposed in this chapter to allocate the LTE-Advanced eNodeB and the MIMO radar resources optimally among users. The proposed algorithm is performed in two stages, the LTE-Advanced eNodeB resources are first allocated to users subscribing for a service and then the available MIMO radar resources are allocated to the same users. The algorithm employs a proportional fairness approach in its two stages to guarantee that no user is allocated zero resources and gets dropped.

Our contributions in this chapter are summarized as:

- We present a spectrum sharing scenario between a MIMO radar and LTE system with multiple base stations and propose a channel-selection algorithm to select the best channel for radar's signal projection that maintains a minimum degradation in the radar performance while causing no interference to the LTE BS. We also present our null-space projection (NSP) algorithm that performs the null space computation.

- We present a resource allocation optimization problem with carrier aggregation to allocate the LTE-Advanced and the MIMO radar carriers resources optimally among users running real-time or delay-tolerant applications.

- We propose a two-stage resource allocation algorithm to allocate the two carriers re-

sources optimally among users. First, the LTE-Advanced eNodeB and the UEs collaborate to allocate an optimal rate to each UE. Once the LTE-Advanced eNodeB finishes allocating resources to the UEs, the eNodeB then allocates the MIMO radar's available resources to these UEs.

The remainder of this chapter is organized as follows. Section 6.1 discusses the spectrum sharing scenario between MIMO radar and LTE cellular system. In Section 6.2, we describe colocated MIMO radars. In section 6.3, we present our channel-selection and NSP algorithms and explain the projection of radar signal onto the null space of the selected interference channel. In Section 6.4 we present our resource allocation with carrier aggregation optimization problem using a utility proportional fairness approach. In section 6.5 we present our two-stage distributed robust resource allocation with carrier aggregation algorithm for the optimization problem. Section 6.6 discusses simulation setup and provides quantitative results along with discussion. Section 6.7 concludes the paper.

6.1 System Model

We consider a colocated MIMO radar and a MIMO LTE communication system. The two systems are the primary users of the 3550-3650 MHz band under consideration. The MIMO radar has M_T transmit antennas and M_R receive antennas. The LTE communication system has N_{BS} base stations, each BS is equipped with N_T^{BS} transmit antennas and N_R^{BS} receive antennas, with the i^{th} BS supporting K_i^{UE} user equipments (UE)s. Each UE is equipped with N_T^{UE} transmit antennas and N_R^{UE} receive antennas. The colocated radars give better target parameter identifiability and improved spatial resolution as their antenna spacing is on the order of half the wavelength of the carrier [90]. The MIMO radar projects its signal onto the null space of the interference channel while illuminating a target. The MIMO radar is sharing N_{BS} interference channels $\mathbf{H}_i^{N_R^{\text{BS}} \times M_T}$ with the LTE system. Let $\mathbf{x}_{\text{Radar}}(t)$ and $\mathbf{x}_j^{\text{UE}}(t)$

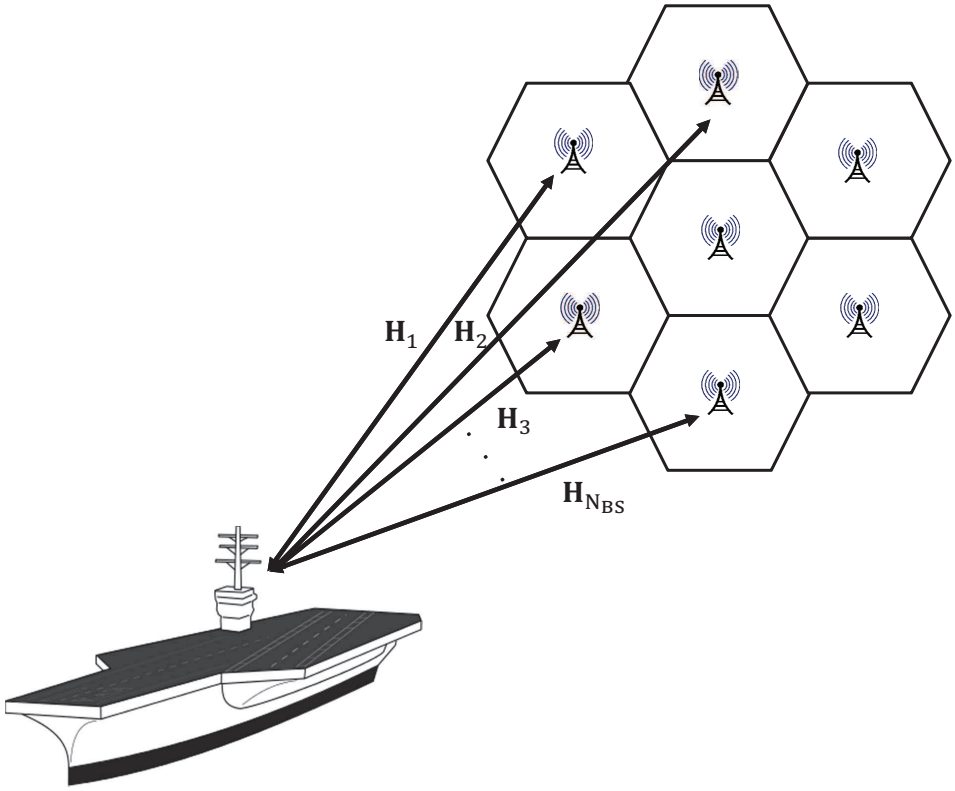


Figure 6.1: Spectrum-sharing scenario between LTE cellular system and a maritime MIMO radar.

be the signals transmitted from the MIMO radar and the j^{th} UE in the i^{th} cell, respectively. The received signal at the i^{th} BS receiver can be written as

$$\mathbf{y}_i(t) = \mathbf{H}_i^{N_R^{\text{BS}} \times M_T} \mathbf{x}_{\text{Radar}}(t) + \sum_j \mathbf{H}_j^{N_R^{\text{BS}} \times N_T^{\text{UE}}} \mathbf{x}_j^{\text{UE}}(t) + \mathbf{w}(t)$$

$$\text{for } 1 \leq i \leq N_{\text{BS}} \text{ and } 1 \leq j \leq K_i^{\text{UE}}$$

where $\mathbf{w}(t)$ is the additive white Gaussian noise. In order to avoid interference to the i^{th} LTE BS, the MIMO radar maps $\mathbf{x}_{\text{Radar}}(t)$ onto the null-space of $\mathbf{H}_i^{N_R^{\text{BS}} \times M_T}$. Figure 6.1 shows a spectrum sharing scenario between a maritime MIMO radar and a LTE cellular system where the MIMO radar is sharing N_{BS} interference channels $\mathbf{H}_i^{N_R^{\text{BS}} \times M_T}$ with the LTE system.

6.2 Radar-LTE Spectrum Sharing Approach

The MIMO radar we consider is a colocated MIMO radar with M_T transmit antennas and M_R receive antennas. Let $\mathbf{x}_{\text{Radar}}(t)$ be the signal transmitted from the MIMO radar, defined as

$$\mathbf{x}_{\text{Radar}}(t) = \begin{bmatrix} x_1(t)e^{j\omega_c t} & x_2(t)e^{j\omega_c t} & \cdots & x_{M_T}e^{j\omega_c t}(t) \end{bmatrix}^T$$

where ω_c is the carrier angular frequency, $x_k(t)$ is the baseband signal from the k^{th} transmit element and $t \in [0, T_o]$ with T_o being the observation time. The radar transmit steering vector is defined as

$$\mathbf{a}_T(\theta) \triangleq \begin{bmatrix} e^{-j\omega_c \tau_{T_1}(\theta)} & e^{-j\omega_c \tau_{T_2}(\theta)} & \cdots & e^{-j\omega_c \tau_{T_{M_T}}(\theta)} \end{bmatrix}^T$$

the radar receive steering vector is defined as

$$\mathbf{a}_R(\theta) \triangleq \begin{bmatrix} e^{-j\omega_c \tau_{R_1}(\theta)} & e^{-j\omega_c \tau_{R_2}(\theta)} & \cdots & e^{-j\omega_c \tau_{R_{M_R}}(\theta)} \end{bmatrix}^T$$

and the transmit-receive steering matrix is defined as

$$\mathbf{A}(\theta) \triangleq \mathbf{a}_R(\theta)\mathbf{a}_T^T(\theta).$$

Then, the signal received from a single point target at an angle θ is given by

$$\mathbf{y}_{\text{Radar}}(t) = \alpha\, e^{-j\omega_D t}\, \mathbf{A}(\theta)\, \mathbf{x}_{\text{Radar}}(t - \tau(t))$$

where $\tau(t) = \tau_r = \tau_{T_k}(t) + \tau_{R_l}(t)$ is the sum of propagation delays between the target and the k^{th} transmit element and between the target and the l^{th} receive element, respectively; and α represents the complex path loss including the propagation loss and the coefficient of reflection.

6.3 Spectrum Sharing Algorithms

In this section, we present a channel-selection algorithm to select the best interference channel on which radar signals are projected. We also present NSP algorithm that performs the null space computation.

6.3.1 Channel-Selection Algorithm

Our channel-selection algorithm, shown in Algorithm (22), selects the best interference channel onto which radar signals are projected. Based on our system model, we assume that there exist N_{BS} interference channels $\mathbf{H}_i, i = 1, 2, \ldots, N_{\mathrm{BS}}$ between the MIMO radar and the LTE system. Our goal is to select the best interference channel defined as

$$\begin{aligned} i_{\min} &\triangleq \underset{1 \le i \le N_{\mathrm{BS}}}{\arg\min} \|\mathbf{x}_{\mathrm{Radar}} - \mathbf{P}_{\mathbf{V}_i}\mathbf{x}_{\mathrm{Radar}}\| \\ \mathbf{H}_{\mathrm{Best}} &\triangleq \mathbf{H}_{i_{\min}} \end{aligned}$$

we also seek to avoid the worst interference channel defined as

$$\begin{aligned} i_{\max} &\triangleq \underset{1 \le i \le N_{\mathrm{BS}}}{\arg\max} \|\mathbf{x}_{\mathrm{Radar}} - \mathbf{P}_{\mathbf{V}_i}\mathbf{x}_{\mathrm{Radar}}\| \\ \mathbf{H}_{\mathrm{Worst}} &\triangleq \mathbf{H}_{i_{\max}} \end{aligned}$$

where $(\mathbf{x}_{\mathrm{Radar}} - \mathbf{P}_{\mathbf{V}_i}\mathbf{x}_{\mathrm{Radar}})$ is the difference between the original radar waveform $\mathbf{x}_{\mathrm{Radar}}$ and the radar waveform projected onto the null space of $\mathbf{H}_i$ and the Euclidean norm of $(\mathbf{x}_{\mathrm{Radar}} - \mathbf{P}_{\mathbf{V}_i}\mathbf{x}_{\mathrm{Radar}})$ is defined as

$$\|\mathbf{x}_{\mathrm{Radar}} - \mathbf{P}_{\mathbf{V}_i}\mathbf{x}_{\mathrm{Radar}}\| = \sqrt{(\mathbf{x}_{\mathrm{Radar}} - \mathbf{P}_{\mathbf{V}_i}\mathbf{x}_{\mathrm{Radar}})^H(\mathbf{x}_{\mathrm{Radar}} - \mathbf{P}_{\mathbf{V}_i}\mathbf{x}_{\mathrm{Radar}})}.$$

We use the blind null space learning algorithm introduced in [91] to estimate the channel state information (CSI) of the N_{BS} interference channels at the MIMO radar. The projection matrix $\mathbf{P}_{\mathbf{V}_i}$ of each of the N_{BS} interference channels is then found using Algorithm (23). Once Algorithm (22) receives the projection matrices of the interference channels, it selects the best interference channel $\breve{\mathbf{H}}$ and sends it to Algorithm (23) for NSP of radar signals. Selecting the best interference channel using our channel-selection algorithm (i.e. Algorithm (22)) guarantees minimum degradation in the performance of the radar while maintaining no interference to the LTE BS.

Algorithm 22 Channel-Selection Algorithm

loop
  for $i = 1 : N_{\text{BS}}$ **do**
    Estimate CSI of $\mathbf{H}_i$.
    Send $\mathbf{H}_i$ to Algorithm (23) for null space computation.
    Receive projection matrix $\mathbf{P}_{\mathbf{V}_i}$ from Algorithm (23).
  end for
  Find $i_{\min} = \arg\min_{1 \le i \le N_{\text{BS}}} \|\mathbf{x}_{\text{Radar}} - \mathbf{P}_{\mathbf{V}_i}\mathbf{x}_{\text{Radar}}\|$.
  Set $\breve{\mathbf{H}} = \mathbf{H}_{i_{\min}}$ as the best interference channel.
  Set $\mathbf{P}_{\breve{\mathbf{V}}} = \mathbf{P}_{\mathbf{V}_{i_{\min}}}$.
  Send $\mathbf{P}_{\breve{\mathbf{V}}}$ to Algorithm (23) to get NSP radar waveform.
end loop

6.3.2 Null-Space Projection (NSP) Algorithm

In this section, we present our proposed null-space projection algorithm. We also explain the projection of radar signals onto null space of the best interference channel selected using Algorithm (22). The CSI of each of the N_{BS} interference channels is first estimated using a blind null space learning algorithm [91]. Algorithm (23) gets the CSI estimates of the interference channels from Algorithm (22) and finds the null space of each $\mathbf{H}_i^{N_R^{\text{BS}} \times M_T}$. This is performed using the singular value decomposition (SVD) theorem as shown in our NSP

algorithm (Algorithm (23)). The SVD for the complex i^{th} interference channel is given by

$$\mathbf{H}_i^{N_R^{\text{BS}} \times M_T} = \mathbf{U}_i \mathbf{\Sigma}_i^{N_R^{\text{BS}} \times M_T} \mathbf{V}_i^H$$

and $\mathbf{\Sigma}_i^{N_R^{\text{BS}} \times M_T}$ is given by

$$\mathbf{\Sigma}_i{}^{N_R^{\text{BS}} \times M_T} = \text{diag}(\sigma_1, ..., \sigma_k, ..., \sigma_l) \in \mathbf{R}^{N_R^{\text{BS}} \times M_T}$$

$$\text{s.t.} \quad l = \min\{N_R^{\text{BS}}, M_T\}\sigma_i \qquad = \begin{cases} \sigma_i, & i \leq k \\ 0\ , & i > k \end{cases}$$

where $\mathbf{U}_i$ is the complex unitary matrix, $\mathbf{\Sigma}_i$ is the matrix of singular values, $\sigma_1 > \sigma_2 > ... > \sigma_k > \sigma_{k+1} = ... = \sigma_l = 0$ and $\mathbf{V}_i^H$ is the complex unitary matrix. Once the null space of all interference channels is determined, $\mathbf{\Sigma}_i'^{M_T \times M_T}$ is then calculated as follows

Algorithm (23) uses $\mathbf{\Sigma}_i'^{M_T \times M_T}$ for the formation of the projection matrix $\mathbf{P}_{\mathbf{V}_i}$ that is given by

$$\mathbf{P}_{\mathbf{V}_i} = \mathbf{V}_i \mathbf{\Sigma}_i'^{M_T \times M_T} \mathbf{V}_i^H$$

where $\mathbf{P}_{\mathbf{V}_i}$ satisfies the following properties:

- $\mathbf{H}_i \mathbf{P}_{\mathbf{V}_i} = 0$.
- $\mathbf{P}_{\mathbf{V}_i}{}^2 = \mathbf{P}_{\mathbf{V}_i}$.

Algorithm (22) receives the projection matrices $\mathbf{P}_{\mathbf{V}_i}$ and uses them to determine the best interference channel $\breve{\mathbf{H}}$ and its corresponding $\mathbf{P}_{\breve{\mathbf{V}}}$, the one with the minimum $\|\mathbf{x}_{\text{Radar}} - \mathbf{P}_{\mathbf{V}_i}\mathbf{x}_{\text{Radar}}\|$, which according to our Algorithm (22) is given by

$$i_{\min} = \underset{1 \leq i \leq N_{\text{BS}}}{\arg\min} \|\mathbf{x}_{\text{Radar}} - \mathbf{P}_{\mathbf{V}_i}\mathbf{x}_{\text{Radar}}\|$$

$$\breve{\mathbf{H}} = \mathbf{H}_{i_{\min}}$$

$$\mathbf{P}_{\breve{\mathbf{V}}} = \mathbf{P}_{\mathbf{V}_{i_{\min}}}.$$

Algorithm (22) sends $\mathbf{P}_{\breve{\mathbf{V}}}$ to Algorithm (23) where it is used for the projection of the radar waveform. The radar waveform projected onto the null space of $\breve{\mathbf{H}}$ can be written as

$$\breve{\mathbf{x}}_{\text{Radar}} = \mathbf{P}_{\breve{\mathbf{V}}}\mathbf{x}_{\text{Radar}}. \tag{6.1}$$

Algorithm 23 Null-Space Projection (NSP) Algorithm

if $\mathbf{H}_i$ received from Algorithm (22) **then**
  Perform SVD on $\mathbf{H}_i$ (i.e. $\mathbf{H}_i = \mathbf{U}_i\mathbf{\Sigma}_i\mathbf{V}_i^H$).
  Find projection matrix $\mathbf{P}_{\mathbf{V}_i} = \mathbf{V}_i\mathbf{\Sigma}_i'^{M_T \times M_T}\mathbf{V}_i^H$.
  Send projection matrix $\mathbf{P}_{\mathbf{V}_i}$ to Algorithm (22).
end if
if $\mathbf{P}_{\breve{\mathbf{V}}}$ received from Algorithm (22) **then**
  Get NSP radar signal via $\breve{\mathbf{x}}_{\text{Radar}} = \mathbf{P}_{\breve{\mathbf{V}}}\mathbf{x}_{\text{Radar}}$.
end if

6.4 RA with CA for Radar-LTE Spectrum Sharing

Each of the N_{BS} LTE-Advanced base stations has L^{UE} UEs/mobiles and two carriers. One of the carriers is the LTE-Advanced carrier that is considered to be the primary carrier and the other one is the MIMO radar carrier considered to be the secondary carrier. Each user is allocated certain bandwidth r_i based on the type of application the UE is running. Our goal is to determine the optimal bandwidth that needs to be allocated to each user by the two carriers.

Each UE has its own utility function $U_i(r_i)$ that corresponds to the application running on the UE. We assume that the utility function assigned to the i^{th} user is a strictly concave utility function if the user is running delay-tolerant application or a sigmoidal-like utility

function if the user is running real-time application.

The first resource allocation optimization problem is the primary carrier (LTE-Advanced carrier) optimization. The primary carrier allocates its resources using a utility proportional fairness approach to guarantee that no user is allocated zero resources.

The LTE-Advanced carrier optimization problem can be written as:

$$\begin{aligned} \max_{\mathbf{r}_{\mathrm{LTE}}} \quad & \prod_{i=1}^{L^{\mathrm{UE}}} U_i(r_{i,\mathrm{LTE}}) \\ \text{subject to} \quad & \sum_{i=1}^{L^{\mathrm{UE}}} r_{i,\mathrm{LTE}} \leq R_{\mathrm{LTE}} \\ & 0 \leq r_i \leq R_{\mathrm{LTE}}, \qquad i = 1, 2, ..., L^{\mathrm{UE}}. \end{aligned} \tag{6.2}$$

where $\mathbf{r}_{\mathrm{LTE}} = \{r_{1,\mathrm{LTE}}, r_{2,\mathrm{LTE}}, ..., r_{L^{\mathrm{UE}},\mathrm{LTE}}\}$ and L^{UE} is the number of mobile users in the coverage area of primary carrier and R_{LTE} is the maximum achievable rate of the primary carrier. This resource allocation objective function is to maximize the total system utility when allocating resources to each user. Furthermore, it provides a proportional fairness among utilities. Users running real-time applications are allocated more resources in this approach.

The objective function in the optimization problem (6.2) is equivalent to $\max_{\mathbf{r}_{\mathrm{LTE}}} \sum_{i=1}^{L^{\mathrm{UE}}} \log U_i(r_{i,\mathrm{LTE}})$, so the optimization problem (6.2) is a convex optimization problem and there exists a unique tractable global optimal solution as shown in [21]. The solution of this optimization problem is the first optimal solution that gives each of the L^{UE} users the optimal rate $r_{i,\mathrm{LTE}}^{\mathrm{opt}}$ only from the primary carrier and not yet the final optimal rate.

As mentioned before, once the LTE-Advanced carrier finishes allocating its resources to the L^{UE} users, the MIMO radar carrier starts to allocate its available resources to the same users using proportional fairness approach to ensure a minimum user QoS.

The optimization problem for the secondary carrier (MIMO radar) can be written as:

$$\begin{aligned} \max_{\mathbf{r}_{\text{radar}}} \quad & \prod_{i=1}^{L^{\text{UE}}} U_i(r_{i,\text{radar}} + r_{i,\text{LTE}}^{\text{opt}}) \\ \text{subject to} \quad & \sum_{i=1}^{L^{\text{UE}}} r_{i,\text{radar}} \leq R_{\text{radar}} \\ & 0 \leq r_{i,\text{radar}} \leq R_{\text{radar}}, \qquad i = 1, 2, ..., L^{\text{UE}}. \end{aligned} \tag{6.3}$$

where $\mathbf{r}_{\text{radar}} = \{r_{1,\text{radar}}, r_{2,\text{radar}}, ..., r_{L_{\text{UE}},\text{radar}}\}$ and L^{UE} is the number of UEs in the coverage area, R_{radar} is the maximum achievable rate by the secondary carrier and $r_{i,\text{LTE}}^{\text{opt}}$ is the optimal rate allocated to user i by the LTE-Advanced carrier in (6.2). Optimization problem (6.3) ensures a minimum rate of $r_{i,\text{LTE}}^{\text{opt}}$ for each user and gives priority for users running real-time applications.

The objective function in the optimization problem (6.3) is equivalent to $\max_{\mathbf{r}_{\text{radar}}} \sum_{i=1}^{L^{\text{UE}}} \log U_i(r_{i,\text{radar}} + r_{i,\text{LTE}}^{\text{opt}})$, so the optimization problem (6.3) is a convex optimization problem and there exists a unique tractable global optimal solution [21].

The final optimal aggregated rate $r_{i,\text{agg}}$ for user i is obtained by the sum of the solution of the optimization problem (6.2) $r_{i,\text{LTE}}^{\text{opt}}$ and the solution of (6.3) $r_{i,\text{radar}}^{\text{opt}}$ and can be written as $r_{i,\text{agg}}^{\text{opt}} = r_{i,\text{radar}}^{\text{opt}} + r_{i,\text{LTE}}^{\text{opt}}$, such that $r_{i,\text{agg}}^{\text{opt}}$ is the global final optimal solution that gives each of the L^{UE} users the optimal rate from both the LTE-Advanced and the MIMO radar carriers. The solution of the optimization problem (6.3) is the global optimal solution that gives each of the M users optimal rates from both the primary and secondary carriers.

6.5 Two-stage Carrier Aggregation Algorithm

Our two-stage algorithm is a modified version of the algorithm proposed in [24]. In the first stage, the UEs and the primary carrier collaborate to allocate an optimal rate to each UE. The first stage of the algorithm starts when each UE transmits an initial bid $w_{i,\text{LTE}}(1)$ to

the LTE-Advanced eNodeB. The eNodeB checks the difference between the current received bid and the previous one, if it is less than a threshold ϵ it exits. Otherwise, if the difference is greater than ϵ, the shadow price $P_{\text{LTE}}(n) = \frac{\sum_{i=1}^{L^{\text{UE}}} w_{i,\text{LTE}}(n)}{R_{\text{LTE}}}$ is calculated by the LTE-Advanced eNodeB. The shadow price represents the total price per unit bandwidth for all users. It depends on the users bids and the eNodeB's available resources. The LTE-Advanced eNodeB sends the calculated $P_{\text{LTE}}(n)$ to each UE where it is used to calculate the rate $r_{i,\text{LTE}}(n)$ that is the solution of the optimization problem $r_{i,\text{LTE}}(n) = \arg\max_{r_{i,\text{LTE}}}(\log U_i(r_{i,\text{LTE}}) - P_{\text{LTE}}(n)r_{i,\text{LTE}})$. The calculated rate is then used to estimate a new bid $w_{i,\text{LTE}}(n)$ where $w_{i,\text{LTE}}(n) = P_{\text{LTE}}(n)r_{i,\text{LTE}}(n)$. All UEs check the fluctuation condition and send their new bids $w_{i,\text{LTE}}(n)$ to the LTE eNodeB. Once the first stage is finalized by the eNodeB, each UE calculates its allocated rate $r_{i,\text{LTE}}^{\text{opt}} = \frac{w_{i,\text{LTE}}(n)}{P_{\text{LTE}}(n)}$.

Algorithm 24 UE First Stage Algorithm

Send initial bid $w_{i,\text{LTE}}(1)$ to LTE-Advanced eNodeB
loop
  Receive shadow price $P_{\text{LTE}}(n)$ from LTE eNodeB
  if STOP from LTE eNodeB **then**
    Calculate allocated rate $r_{i,\text{LTE}}^{\text{opt}} = \frac{w_{i,\text{LTE}}(n)}{P_{\text{LTE}}(n)}$
  else
    Solve $r_{i,\text{LTE}}(n) = \arg\max_{r_{i,\text{LTE}}}\Big(\log U_i(r_{i,\text{LTE}}) - P_{\text{LTE}}(n)r_{i,\text{LTE}}\Big)$
    Calculate new bid $w_{i,\text{LTE}}(n) = P_{\text{LTE}}(n)r_{i,\text{LTE}}(n)$
    if $|w_{i,\text{LTE}}(n) - w_{i,\text{LTE}}(n-1)| > \Delta w$ **then**
      $w_{i,\text{LTE}}(n) = w_{i,\text{LTE}}(n-1) + \text{sign}(w_{i,\text{LTE}}(n) - w_{i,\text{LTE}}(n-1))\Delta w(n)$
      $\{\Delta w(n) = l_1 e^{-\frac{n}{l_2}}\}$
    end if
    Send new bid $w_{i,\text{LTE}}(n)$ to eNodeB
  end if
end loop

After allocating rates $r_{i,\text{LTE}}^{\text{opt}}$ from the LTE carrier, the second-stage of the algorithm starts performing. Each UE transmits its initial bid $w_{i,\text{radar}}(1)$ to the MIMO radar eNodeB. The eNodeB checks the difference between the current received bid and the previous one if it is less than a threshold ϵ it exits. Otherwise, if the difference is greater than ϵ, the MIMO radar eNodeB calculates the shadow price $P_{\text{radar}}(n) = \frac{\sum_{i=1}^{L^{\text{UE}}} w_{i,\text{radar}}(n)}{R_{\text{radar}}}$. The radar eNodeB sends the

Algorithm 25 LTE eNodeB Algorithm

loop
  Receive bids $w_{i,\mathrm{LTE}}(n)$ from UEs {Let $w_{i,\mathrm{LTE}}(0) = 0 \ \forall i$}
  if $|w_{i,\mathrm{LTE}}(n) - w_{i,\mathrm{LTE}}(n-1)| < \epsilon \ \forall i$ **then**
    STOP and allocate rates (i.e $r^{\mathrm{opt}}_{i,\mathrm{LTE}}$ to user i)
  else
    Calculate $P_{\mathrm{LTE}}(n) = \frac{\sum_{i=1}^{L^{\mathrm{UE}}} w_{i,\mathrm{LTE}}(n)}{R_{\mathrm{LTE}}}$
    Send new shadow price $P_{\mathrm{LTE}}(n)$ to all UEs
  end if
end loop

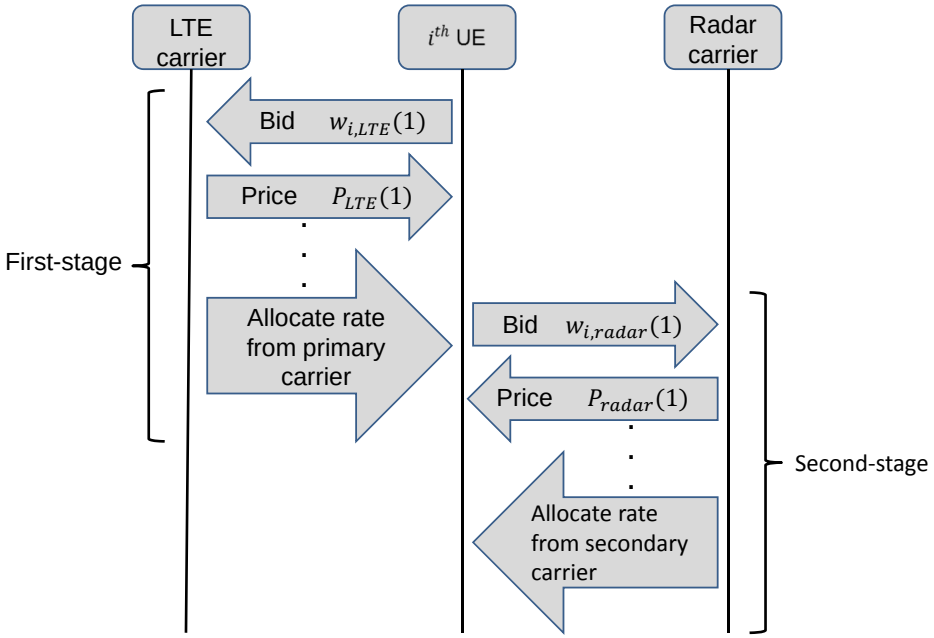


Figure 6.2: Flow Diagram for the two-stage RA with carrier aggregation Algorithm.

calculated $P_{\mathrm{radar}}(n)$ to the UEs. Each UE calculates the rate $r_{i,\mathrm{radar}}(n)$ which is the solution of the optimization problem $r_{i,\mathrm{radar}}(n) = \arg\max_{r_{i,\mathrm{radar}}} (\log U_i(r_{i,\mathrm{radar}} + r^{\mathrm{opt}}_{i,\mathrm{LTE}}) - P_{\mathrm{radar}}(n) r_{i,\mathrm{radar}})$. A new bid $w_{i,\mathrm{radar}}(n)$ is calculated using $r_{i,\mathrm{radar}}(n)$ where $w_{i,\mathrm{radar}}(n) = P_{\mathrm{radar}}(n) r_{i,\mathrm{radar}}(n)$. All UEs check the fluctuation condition and send their new bids $w_{i,\mathrm{radar}}(n)$ to the radar eNodeB. The second-stage of the Algorithm is finalized by the radar eNodeB. Each UE then calculates its allocated rate $r^{\mathrm{opt}}_{i,\mathrm{radar}} = \frac{w_{i,\mathrm{radar}}(n)}{P_{\mathrm{radar}}(n)}$ by the radar eNodeB. The final global optimal rate $r^{\mathrm{opt}}_{i,\mathrm{agg}} = r^{\mathrm{opt}}_{i,\mathrm{radar}} + r^{\mathrm{opt}}_{i,\mathrm{LTE}}$ is then allocated to each UE. Figure 6.2 shows a flow chart of the LTE-Advanced two-stage RA with carrier aggregation Algorithm.

Algorithm 26 UE Second Stage Algorithm

Send initial bid $w_{i,\text{radar}}(1)$ to the radar eNodeB
loop
  Receive shadow price $P_{\text{radar}}(n)$ from the radar eNodeB
  if STOP from the radar eNodeB **then**
    Calculate allocated rate $r_{i,\text{agg}}^{\text{opt}} = \frac{w_{i,\text{radar}}(n)}{P_{\text{radar}}(n)} + r_{i,\text{LTE}}^{\text{opt}}$
  else
    Solve $r_{i,\text{radar}}(n) = \arg\max_{r_{i,\text{radar}}} \Big(\log U_i(r_{i,\text{radar}} + r_{i,\text{LTE}}^{\text{opt}}) - P_{\text{radar}}(n) r_{i,\text{radar}} \Big)$
    Calculate new bid $w_{i,\text{radar}}(n) = P_{\text{radar}}(n) r_{i,\text{radar}}(n)$
    if $|w_{i,\text{radar}}(n) - w_{i,\text{radar}}(n-1)| > \Delta w$ **then**
      $w_{i,\text{radar}}(n) = w_{i,\text{radar}}(n-1) + \text{sign}(w_{i,\text{radar}}(n) - w_{i,\text{radar}}(n-1))\Delta w(n)$
      $\{\Delta w(n) = l_1 e^{-\frac{n}{l_2}}\}$
    end if
    Send new bid $w_{i,\text{radar}}(n)$ to eNodeB
  end if
end loop

Algorithm 27 MIMO Radar eNodeB Algorithm

loop
  Receive bids $w_{i,\text{radar}}(n)$ from UEs $\{\text{Let } w_{i,\text{radar}}(0) = 0 \ \forall i\}$
  if $|w_{i,\text{radar}}(n) - w_{i,\text{radar}}(n-1)| < \epsilon \ \forall i$ **then**
    STOP and allocate rates (i.e $r_{i,\text{radar}}^{\text{opt}}$ to user i)
  else
    Calculate $P_{\text{radar}}(n) = \frac{\sum_{i=1}^{L^{\text{UE}}} w_{i,\text{radar}}(n)}{R_{\text{radar}}}$
    Send new shadow price $P_{\text{radar}}(n)$ to all UEs
  end if
end loop

6.6 Simulation Results

In our spectrum sharing model, the LTE-Advanced system has N_{BS} BS, only the i^{th} BS is under zero interference from the MIMO radar due to the spectrum sharing approach employed by the proposed spectrum sharing model. We consider this BS which has two eNodeBs, one is configured at the LTE-Advanced carrier and the second is configured to use radar carrier when there is no interference from radar. In this BS we consider four UEs in its coverage area subscribing for a mobile service. The first and second UEs are running real-time applications presented by sigmoidal-like utility functions whereas the third and fourth

UEs are running delay-tolerant applications presented by logarithmic utility functions. The four UEs are to be allocated resources from the LTE-Advanced and the MIMO radar carriers.

The proposed RA with CA algorithm is applied in C++ to the sigmoidal-like and logarithmic utility functions. Simulation results showed convergence to the optimal global point in the two stages of the algorithm. Each of the four UEs is allocated a final optimal rate by the two carriers. We use a normalized sigmoidal-like utility function that is expressed by equation (2.1) to represent the first user real-time application with $a = 3$, $b = 20$ which is an approximation to a step function at rate $r = 20$. Additionally, we use another sigmoidal-like utility function to represent the second user real-time application with $a = 1$, $b = 30$. Furthermore, we use logarithmic functions to represent the third and fourth UEs delay-tolerant applications with $k = 3$ and $k = 0.5$, respectively. Additionally, We use $r_{max} = 100$ for all logarithmic functions, $l_1 = 5$ and $l_2 = 10$ in the fluctuation decay function of the algorithm and $\epsilon = 10^{-7}$.

6.6.1 Rate Allocation for $10 \leq R_{\mathbf{LTE}} \leq 70$ in the First-Stage of the RA Algorithm

We apply Algorithm (24) and (25) of the first-stage in C++ to the sigmoidal-like and logarithmic utility functions. The LTE-Advanced eNodeB available resources R_{LTE} takes values between 10 and 70 with step of 10. In Figure 6.3, we show the four users optimal rates $r_{i,\text{LTE}}^{opt}$ allocated by the LTE-Advanced eNodeB with different eNodeB resources R_{LTE}. This represents the solution of optimization problem (6.2). As mentioned before the sigmoidal-like utility functions are given priority over the logarithmic utility functions for rate allocation and this explain the results we got in Figure 6.3 where the algorithm gives priority to real-time applications when allocating the LTE-Advanced eNodeB resources as it uses proportional fairness approach. Users with real-time applications bid higher than the other users until

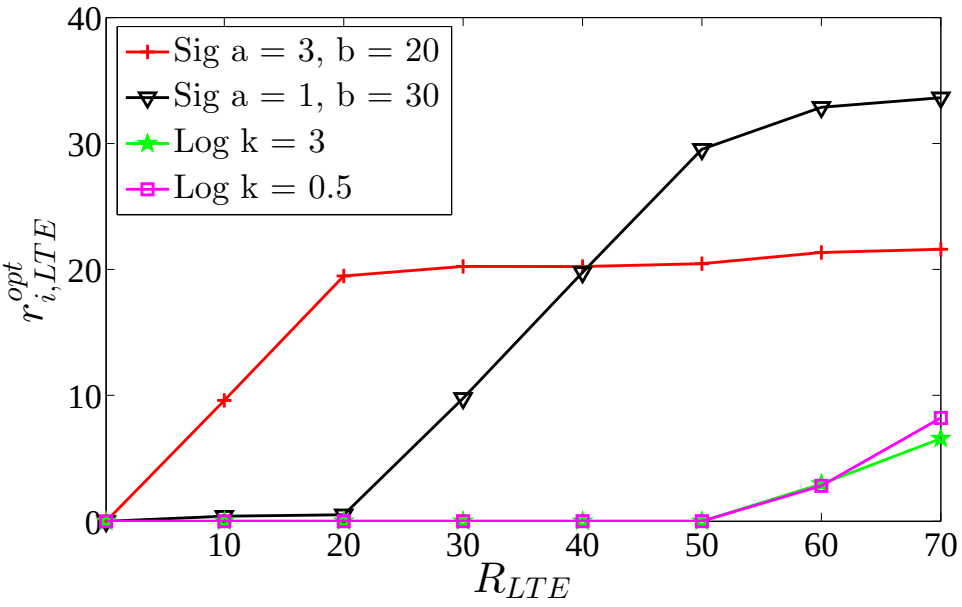


Figure 6.3: The users optimal rates $r^{opt}_{i,\text{LTE}}$ for different values of R_{LTE} for Algorithm (24) and (25).

each one of them reaches its inflection point then the algorithm starts dividing the remaining resources among users running delay-tolerant applications based on their utility functions parameters.

6.6.2 Rate Allocation for $10 \leq R_{\text{radar}} \leq 80$ in the Second-Stage of the RA Algorithm

We apply Algorithm (26) and (27) of the second-stage in C++ to the sigmoidal-like and logarithmic utility functions. The radar carrier available resources R_{radar} takes values between 10 and 80 with step of 10. In Figure 6.4, we show the four users optimal rates $r^{opt}_{i,\text{radar}}$ allocated by the radar eNodeB with different available resources R_{radar}. This represents the solution of optimization problem (6.3). Each user running real-time application is allocated at least its utility inflection rate $r_i = b_i$ by the LTE-Advanced carrier in the first-stage of the Algorithm, this explains the result we got in Figure 6.4 where most of the radar carrier resources are allocated to users running delay-tolerant applications.

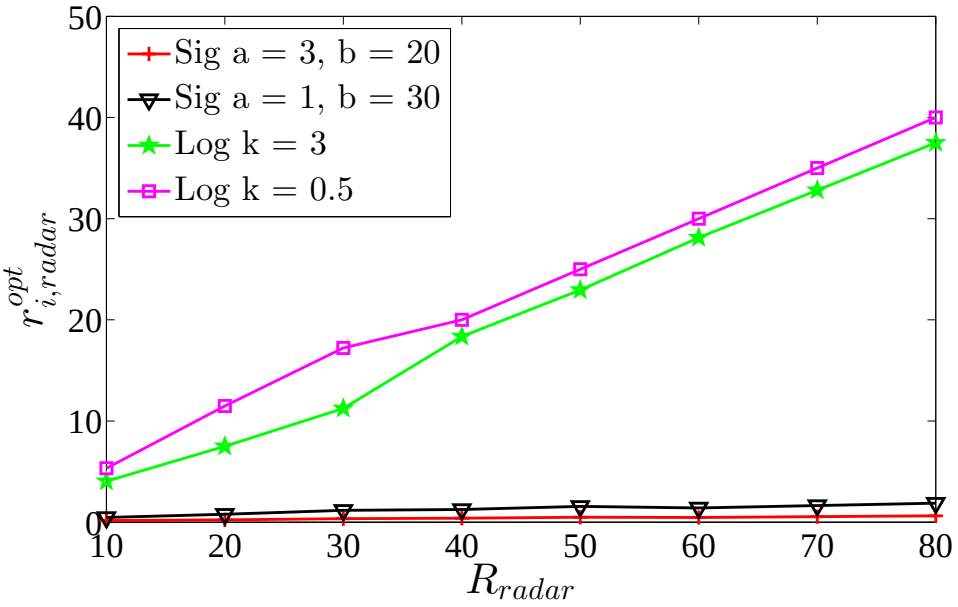


Figure 6.4: The users optimal rates $r_{i,\text{radar}}^{opt}$ for different values of R_{radar} for Algorithm (26) and (27).

6.6.3 RA with Carrier Aggregation for $10 \leq R \leq 150$

In the following simulations, the total rate of the LTE-Advanced carrier takes values between 10 and 70 and the MIMO radar carrier has available resources that takes values between 10 and 80. The two carriers resources are to be allocated to the four users subscribing for a mobile service in the LTE-Advanced cellular cell using RA with carrier aggregation.

In Figure 6.5, we show the optimal rate allocated to each user by the first-stage of the algorithm when $10 \leq R \leq 70$ is the LTE-Advanced carrier available resources. The final optimal rates allocated to each user by the second-stage of the algorithm are also shown in Figure 6.5 for $70<R \leq 150$ where R is the total available resources of $R_{\text{LTE}} = 70$ and $10 \leq R_{\text{radar}} \leq 80$. The LTE carrier allocates the majority of its resources to the UEs running real-time applications until they reach the inflection rate $r_i = b_i$. When the LTE-Advanced resources R_{LTE} exceed the total inflection rates of the users real-time applications, the LTE-Advanced carrier starts allocating resources to the delay-tolerant applications. The aggregated final optimal rate allocated to each user by the LTE-Advanced and the radar

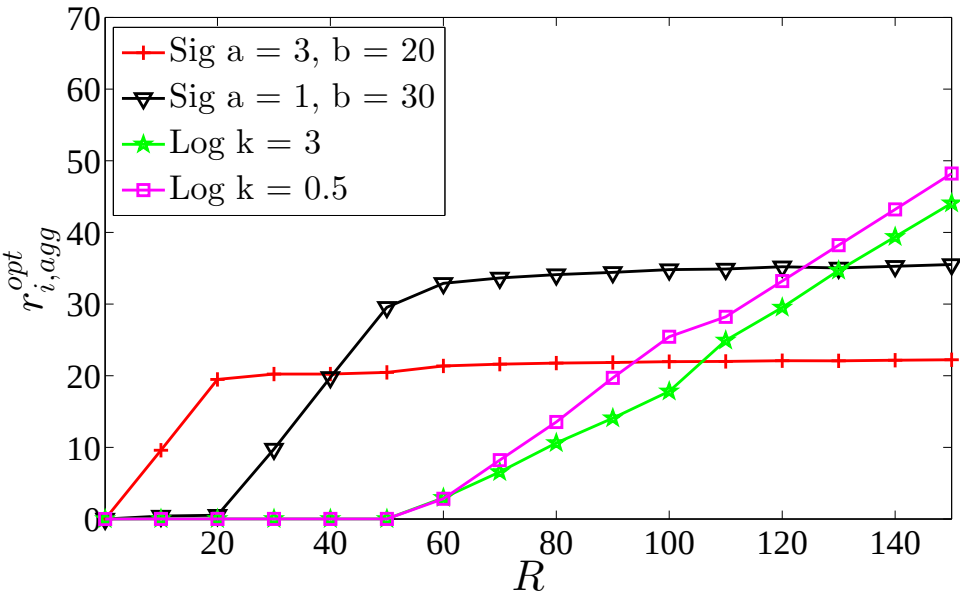


Figure 6.5: The users final optimal rates $r_{i,\text{agg}}^{\text{opt}}$ for different values of R where $10 \leq R \leq 70$ is the LTE-Advanced carrier available resources and $70 < R \leq 150$ is the total available resources of $R_{\text{LTE}} = 70$ and $10 \leq R_{\text{radar}} \leq 80$.

carriers is the total optimal rate allocated to each user by the LTE-Advanced carrier when $R_{\text{LTE}} = 70$ and the optimal rate allocated to the same user by the radar carrier when $10 \leq R_{\text{radar}} \leq 80$. Since users running real-time applications are allocated at least their utilities inflection rates $r_i = b_i$ by the LTE-Advanced carrier, the radar carrier allocates most of its resources to users running delay-tolerant applications.

6.6.4 Price Sensitivity to Change in R

In the following simulations, the total available resources takes different values between 10 and 150 with step of 10. In Figure 6.6, we show the shadow price P, that represents the total price per unit bandwidth for all users, with the total available resources R of the LTE-Advanced and radar carriers. R is the LTE-Advanced carrier available resources for $10 \leq R \leq 70$ whereas when $70 < R \leq 150$ R is the total available resources of $R_{\text{LTE}} = 70$ and $10 \leq R_{\text{radar}} \leq 80$. As expected the price is higher for smaller R when the number of users is

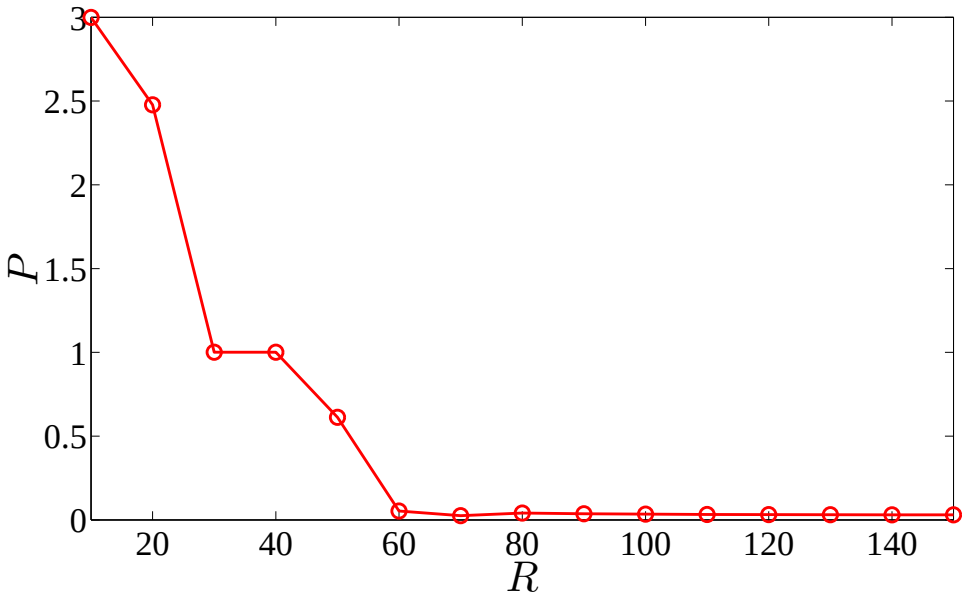


Figure 6.6: The shadow price P for different values of R and fixed number of users (same four users), R is the LTE-Advanced carrier available resources for $10 \leq R \leq 70$ whereas when $70 < R \leq 150$ R is the total available resources of $R_{\text{LTE}} = 70$ and $10 \leq R_{\text{radar}} \leq 80$.

fixed (the same four users).

6.7 Summary and Conclusions

In this chapter, we presented a spectrum sharing scenario between a MIMO radar and LTE cellular system with multiple BSs. We proposed a channel-selection algorithm and NSP algorithm to select the best interference channel and project the radar signal onto it. Our proposed algorithms guarantee a minimum degradation in the radar's performance by selecting the best interference channel for the NSP of the radar signal. In addition, we presented an optimal resource allocation with carrier aggregation approach to allocate LTE-Advanced and MIMO radar carriers' resources optimally among LTE-Advanced users in a cellular cell. We considered two utility functions based on the application type running on the UE, sigmoidal-like utility functions represent real time applications and logarithmic utility functions represent delay tolerant applications. As a result of our analysis, we presented

an iterative distributed RA with carrier aggregation algorithm for the UEs and both LTE-Advanced and radar carriers. The algorithm provides a utility proportional fair resource allocation which guarantees a minimum QoS to each user while giving priority to users running real-time applications. We showed through simulations that our algorithm converges to the optimal rates in its two stages.

Chapter 7

Utility Proportional Fairness Resource Block Scheduling with Carrier Aggregation

Utility proportional fairness (PF) resource allocation for a single carrier in cellular networks have been extensively studied in [21]. The problem of RA for multi-carrier systems in single cell have been given attention in recent years [24, 29–31]. In [24], a multi-stage resource allocation with CA algorithms are presented. However, non of these RA approaches have considered the problem of RB scheduling for multiple component carriers.

In this chapter, we focus on solving the problem of utility PF resource block scheduling with CA for multi-carrier cellular networks. The resource scheduling approach presented in [28, 92] does not consider the case of multi-carrier resources available at the eNodeB. It only solves the problem of RB scheduling in the case of single carrier.

Our contributions in this chapter are summarized as:

- We propose a framework for the problem of utility proportional fairness RB scheduling

with CA for multi-carrier cellular networks.

- We introduce a user grouping method that creates a user group for each component carrier such that each carrier assigns its resources only to users in its user group. Each user subscribing for a mobile service is assigned on multiple component carriers' RBs based on the proposed user grouping method and a utility proportional fairness policy.

- We prove that the proposed resource scheduling policy, that is based on CA, exists and that the optimal solution is tractable.

- We present simulation results for the proposed resource scheduling with CA approach and compare its performance in the case of using the proposed resource scheduling policy and the case of using the scheduling policy presented in [92].

7.1 System Model and Problem Setup

The transmission resources in a LTE downlink have dimensions in frequency, time and space [93]. The frequency is represented by subcarriers. The time is divided into frames and each frame is further divided into subframes. The space is provided by the transmit and receive antennas. One RB consists of 12 continuous subcarriers. In reuse-1 radio systems, that is considered in this chapter, a RB can be allocated to only one user.

We consider a single cell LTE-Advanced mobile system with one eNodeB and M users. Let the number of CCs that the system can aggregate be K. The set of CCs is given by $\mathcal{K} = \{f_1, f_2, ..., f_K\}$ with CCs in order from the highest frequency to the lowest frequency (i.e. $f_1 > f_2 > ... > f_K$). We consider an equal power allocation (EPA) scheme that each frequency component has the same transmitting power. Furthermore, a non adjacent inter band aggregation scenario is considered. Because the channel fading for high frequency is larger than that for low frequency, higher frequency carriers have smaller coverage areas

than lower frequency carriers. Users located under the coverage area of multiple carriers are scheduled resources from all in band carriers. The eNodeB assigns RBs from multiple carriers to each UE. The total allocated rate achieved by assigning RBs to the i^{th} UE is given by r_i. Each UE has its own utility function $U_i(r_i)$ that corresponds to the type of application running on the i^{th} UE. Our goal is to determine which RBs from each CC should be allocated to each UE by the eNodeB in order to maximize the total system utility while ensuring PF between utilities.

We define $\mathcal{Z}_k$, where $1 \leq k \leq K$, to be the set of RBs available by f_k carrier where $z_{k,j}$ denotes a single RB in $\mathcal{Z}_k = \{z_{k,1}, z_{k,2}, ...\}$, $z_{k,j} \in \mathcal{Z}_k$ is the j^{th} RB in CC f_k and $|\mathcal{Z}_k|$ denotes the number of RBs available by f_k carrier. The signal to noise ratio (SNR) of user i on RB $z_{k,j}$ is given by $\gamma_{i,z_{k,j}} = P_{z_{k,j}}|G_{i,z_{k,j}}|^2/N_{i,z_{k,j}}$ where $G_{i,z_{k,j}}$ is the complex channel gain between the eNodeB and the i^{th} UE on RB $z_{k,j}$, $N_{i,z_{k,j}}$ is the noise power experienced by the i^{th} UE on RB $z_{k,j}$ and $P_{z_{k,j}}$ is the transmission power that the eNodeB assigns to RB $z_{k,j}$. Under the EPA, $P_{z_{k,j}} = P_k/|\mathcal{Z}_k|$ where P_k is the transmitting power of CC f_k. Then the achievable data rate of the i^{th} user on RB $z_{k,j}$ is given by

$$H_{i,z_{k,j}} = W \log(1 + \beta_{z_{k,j}} \gamma_{i,z_k}), \tag{7.1}$$

where W is the bandwidth of a RB and $\beta_{z_{k,j}}$ is the SNR gap.

In each frame, the eNodeB schedules each of the frame's RBs to one UE. Let $\phi_{i,z_{k,j}}$ be the proportion of frames that the i^{th} UE is scheduled by the eNodeB on RB $z_{k,j}$. The i^{th} UE rate on all RBs scheduled by carrier f_k is given by

$$r_{i,f_k} = \sum_{z_{k,j} \in \mathcal{Z}_k} \phi_{i,z_{k,j}} H_{i,z_{k,j}}. \tag{7.2}$$

The overall rate of the i^{th} UE, that is the sum of the rates achieved by all carriers RBs assignments, is given by $r_i = \sum_{f_k \in \mathcal{K}} r_{i,f_k}$.

We express the user satisfaction with its application rates using utility functions. We represent the i^{th} user application utility function $U_i(r_i)$ by sigmoidal-like function or logarithmic function where r_i is the rate of the i^{th} user application. Logarithmic utility functions expressed by equation (2.2) and sigmoidal-like utility functions expressed by equation (2.1) are used to represent delay tolerant and real-time applications, respectively.

A user grouping method is introduced in 7.2 to partition users into groups depending on their location in the cell. The eNodeB performs RBs assignments from each CC to the user group located in the coverage area of that carrier.

7.2 User Grouping Method

In this section we introduce a user grouping method to create one user group $\mathcal{M}_{f_k}$ for each CC f_k where $\mathcal{M}_{f_k}$ is a set of users located under the coverage area of carrier f_k. Users in $\mathcal{M}_{f_k}$ are assigned RBs on CC f_k by the eNodeB. Users located under the coverage area of multiple carriers (i.e. common users in multiple user groups) are assigned RBs on these carriers and their final rates are aggregated under a non adjacent inter band aggregation scenario.

The i^{th} user is part of user group $\mathcal{M}_{f_k}$ if it satisfies certain path loss constraints on CC f_k. Assume that the maximum pathloss in a carrier can not exceed a threshold L^{th}. In order for the eNodeB to identify a user group for each CC, it first computes the i^{th} user pathloss on each CC and creates a set α_i that includes all in range carriers such that the i^{th} user is assigned RBs only from carriers in α_i.

Higher frequency carriers have smaller coverage radius R_k than lower frequency carriers (i.e. $R_1 < R_2 < ... < R_K$). Therefore, user group $\mathcal{M}_{f_1} \subseteq \mathcal{M}_{f_2} \subseteq ... \subseteq \mathcal{M}_{f_K}$.

7.3 RB Scheduling with CA Problem

In this section, we present our RB scheduling with CA approach. Our objective is to assign RBs to each user (i.e. the i^{th} user) on all of its in range carriers (i.e. CCs in α_i) based on a utility PF policy. We use utility functions of users' applications rates to represent the type of application running on the UE. Given that different applications may have different QoS requirements, every user subscribing for a mobile service is guaranteed to achieve minimum QoS for each of its applications with a priority criterion. Users running real-time applications are given priority when assigning RBs due to the sigmoidal-like utility functions nature used to represent their applications. In addition, our utility PF approach guarantees that no user is assigned zero RBs.

The eNodeB performs the RBs assignment for each of the CC's RBs in $\mathcal{Z}_k$. It assigns the RBs of each CC f_k one at a time and one after another in ascending order of their coverage radius R_k. It starts with CC f_1 as it has the smallest coverage radius R_1. After assigning all users in $\mathcal{M}_{f_1}$ on f_1 RBs, the eNodeB then assigns users in $\mathcal{M}_{f_2}$ on f_2 RBs. In addition, since $\mathcal{M}_{f_1}$ users are also in $\mathcal{M}_{f_2}$ (i.e. $\mathcal{M}_{f_1} \subseteq \mathcal{M}_{f_2}$), the eNodeB assigns $\mathcal{M}_{f_1}$ users on f_2 RBs and the rates are aggregated based on a non adjacent inter band aggregation scenario. The eNodeB continues the RB assignment process by assigning $\mathcal{M}_{f_k}$ users on CC f_k RBs. Finally, the RB assignment process is finalized by assigning carrier f_K RBs to all users in the cellular cell as they are all located within its coverage radius. We consider a utility PF objective function, based on CA, that the eNodeB seeks to maximize for each time it assigns user on a RB. The utility PF resource scheduling with CA optimization problem for the

eNodeB assignments of $\mathcal{M}_{f_k}$ users on $\mathcal{Z}_k$ RBs is given by

$$\begin{aligned} \max_{\phi_{i,z_k}} \quad & \prod_{i=1}^{M_k} U_i \left(c_{i,f_k} + \sum_{z_{k,j} \in \mathcal{Z}_k} (\phi_{i,z_{k,j}} H_{i,z_{k,j}}) \right) \\ \text{subject to} \quad & \sum_{i=1}^{M_k} \phi_{i,z_{k,j}} = 1, \;\; c_{i,f_1} = 0, \\ & \phi_{i,z_{k,j}} \geq 0, \quad i = 1, 2, ..., M_k \\ & c_{i,f_k} = \sum_{l=1}^{k-1} r_{i,f_l}, \;\; k > 1. \end{aligned} \tag{7.3}$$

where $M_k = |\mathcal{M}_{f_k}|$ is the number of UEs in the coverage area of carrier f_k, $c_{i,f_1} = 0$ and c_{i,f_k} for $k > 1$ is equivalent to $\sum_{l=1}^{k-1} r_{i,f_l}$ that is the i^{th} UE total rate on all RBs scheduled by carriers $\{f_1, ..., f_{k-1}\}$. The eNodeB seeks to maximize the objective function of this resource scheduling optimization problem that is achieved by maximizing the product of all UEs' utilities when assigning the UEs on the carriers' RBs. The goal of this resource scheduling objective function is to allocate the resources to the UE that maximizes the total cellular network objective (i.e. the product of the utilities of all UEs) while ensuring PF between individual utilities. This objective function ensures non-zero RA for all users. Therefore, the resource scheduling optimization problem guarantees minimum QoS for all users. In addition, this approach allocates more resources to real-time applications providing improvement to the QoS of LTE system.

Later in this section we prove that there exists a tractable global optimal solution to optimization problem (7.3). However, the user's final rate, achieved by assigning each user on its in range carriers' RBs, is determined using a multi-stage approach where optimization problem (7.3) is required for each CC f_k. In addition, optimization problem (7.3) needs to be applied in a multi-stage scenario starting from the carrier with the smallest coverage area (i.e. f_1) and ending with the carrier that has the largest coverage area (i.e. f_K). The rate achieved for each user after assigning CC f_k RBs is needed for the next stage optimization problem

(7.3) of carrier f_{k+1}. The objective function in optimization problem (7.3) is equivalent to $\arg\max_{\phi_{i,z_k}} \sum_{i=1}^{M_k} \log(U_i(c_{i,f_k} + \sum_{z_{k,j}\in\mathcal{Z}_k}(\phi_{i,z_{k,j}} H_{i,z_{k,j}})))$. The utility functions $\log(U_i(c_{i,f_k} + \sum_{z\epsilon Z} \phi_{i,b(i),z} H_{i,b(i),z}))$ that are equivalent to $\log(U_i(c_{i,f_k} + r_{i,f_k})$ are strictly concave functions as proved in [21]. As a result, optimization problem (7.3) is a convex optimization problem and there exists a unique tractable global optimal solution [21, 24].

In order to consider the case when the entire input is not available from the beginning, we use an online algorithm as in [92, 94]. The total achieved data rate of each UE when assigning it on different CCs' RBs, i.e. r_i, requires the knowledge of $\phi_{i,z_{k,j}}$ on each RB $z_{k,j}$ the UE is assigned on. We use an online scheduling algorithm to decrease the computation overhead while processing the rate information as in [92].

Let $\phi_{i,z_{k,j}}[n]$ be the proportion of the frames that UE i is scheduled on RB $z_{k,j}$ in the first n frames. Then, the proportion of the frames that UE i is scheduled on RB $z_{k,j}$ in the $[n+1]^{th}$ frame is defined as follows:

$$\phi_{i,z_{k,j}}[n+1] = \begin{cases} \frac{n-1}{n}\phi_{i,z_{k,j}}[n] + \frac{1}{n}, \\ \text{if UE } i \text{ is scheduled on RB } z_{k,j} \\ \text{in the } (n+1)^{th} \text{ frame} \\ \frac{n-1}{n}\phi_{i,z_{k,j}}[n], \text{otherwise.} \end{cases} \tag{7.4}$$

In the proposed scheduling policy, for certain CC's RB $z_{k,j}$, the eNodeB schedules the UE that maximizes $\frac{U_i'(c_{i,f_k}+\sum_{z_{k,j}\in\mathcal{Z}_k}\phi_{i,z_{k,j}}H_{i,z_{k,j}})H_{i,z_{k,j}}}{U_i(c_{i,f_k}+r_{i,f_k})}$ on RB $z_{k,j}$.

Lemma 7.3.1. *Using the scheduling policy in (7.4), we show that* $\liminf_{n\to\infty} \sum_{i=1}^{M_k} \log U_i(c_{i,f_k} + \sum_{z_{k,j}\in\mathcal{Z}_k}(\phi_{i,z_{k,j}}[n]H_{i,z_{k,j}}))$ *exists for optimization problem (7.3).*

Proof. We define $L(\phi) = \sum_{i=1}^{M_k} \log U_i(c_{i,f_k} + \sum_{z_{k,j}\in\mathcal{Z}_k}(\phi_{i,z_{k,j}} H_{i,z_{k,j}}))$ where ϕ, $\phi[n]$ and H are the short terms for $\phi_{i,z_{k,j}}$, $\phi_{i,z_{k,j}}[n]$ and $H_{i,z_{k,j}}$, respectively. Let $r_{i,f_k}[n] = \sum_{z_{k,j}}(\phi_{i,z_{k,j}}[n]H_{i,z_{k,j}})$.

Using Taylor's theorem, for any ϕ and $\Delta\phi$ we have

$L(\phi + \Delta\phi) = L(\phi) + L'(\phi)\Delta\phi + \pi(\phi, \Delta\phi)$

where $|\pi(\phi + \Delta\phi)| < b|\Delta\phi|^2$, for some constant b.

Let $\Delta\phi_{i,z_{k,j}}[n] = \phi_{i,z_{k,j}}[n+1] - \phi_{i,z_{k,j}}[n]$, then

$$\Delta\phi_{i,z_{k,j}}[n] = \begin{cases} \frac{1}{n} - \frac{\phi_{i,z_{k,j}}[n]}{n}, \\ \text{if UE } i \text{ is scheduled on RB } z_{k,j} \\ \text{in the } (n+1)^{th} \text{ frame} \\ \frac{-\phi_{i,z_{k,j}}[n]}{n}, \text{otherwise.} \end{cases}$$

$|\Delta\phi_{i,z_{k,j}}[n]| < \frac{1}{n}$, for all i and $z_{k,j}$. As a result;

$$\begin{aligned} L(\phi[n+1]) &= L(\phi[n] + \Delta\phi[n]), \\ &\geq L(\phi[n]) + \Delta L(\phi[n]) - \frac{b}{n^2}, \\ &= L(\phi[n]) + \left(\sum_i \frac{U_i'(c_{i,f_k} + \sum_{z_{k,j}} \phi H)}{U_i(c_{i,f_k} + r_{i,f_k})} \right. \\ &\left. H\Delta\phi \right) - \frac{b}{n^2} \\ &= L(\phi[n]) + \frac{1}{n}\left(\max_i \frac{U_i'(c_{i,f_k} + \sum_{z_{k,j}} \phi H)}{U_i(c_{i,f_k} + r_{i,f_k})} \right. \\ &\left. H - \sum_i \frac{U_i'(c_{i,f_k} + \sum_{z_{k,j}} \phi H)}{U_i(c_{i,f_k} + r_{i,f_k})} H\phi[n] \right) - \frac{b}{n^2} \\ &\geq L(\phi[n]) - \frac{b}{n^2}, \end{aligned} \tag{7.5}$$

where $\Delta\phi[n]$ is substituted by $(\frac{1}{n} - \frac{\phi_{i,z_{k,j}}[n]}{n})$ (i.e. user i has the largest $\frac{U_i'(c_{i,f_k} + \sum_{z_{k,j}} \phi H)H}{U_i(c_{i,f_k} + r_{i,f_k})}$ among all users) and the last inequality holds since $\sum_i \phi_{i,z_{k,j}[n]} = 1$ for all i and $z_{k,j}$.

Let $\beta := \limsup_{n\to\infty} L(\phi[n])$. For any $\epsilon > 0$, there exists large enough N so that $L(\phi[N]) > \beta - \frac{\epsilon}{2}$ and $\sum_{n=N}^{\infty} \frac{b}{n^2} < \frac{\epsilon}{2}$. For any $\hat{n} > N$, $L(\phi[\hat{n}]) \geq L(\phi[N]) - \sum_{n=N}^{\hat{n}} \frac{b}{n^2} > \beta - \epsilon$. Therefore, $L(\phi[n])$ converges to β, as $n \to \infty$.

Due to the constraint $\sum_{i=1}^{M_k} \phi_{i,z_{k,j}} = 1$ in (7.3), ϕ is a solution to optimization problem (7.3) if and only if

$$\begin{aligned} \frac{dL}{d\phi_{i,z_{k,j}}} &= \frac{U_i'(c_{i,f_k} + \sum_{z_{k,j}} \phi H) H}{U_i(c_{i,f_k} + r_{i,f_k})} \\ &= \max_m \frac{U_m'(c_{m,f_k} + \sum_{z_{k,j}} \phi_{m,z_{k,j}} H_{m,z_{k,j}})}{U_m(c_{m,f_k} + r_{m,f_k})} H_{m,z_{k,j}}, \end{aligned} \tag{7.6}$$

for all i and $z_{k,j}$ such that $\sum_{i=1}^{M_k} \phi_{i,z_{k,j}} = 1$ and $\phi_{i,z_{k,j}} \geq 0$. □

Theorem 7.3.2. *Using the scheduling policy (7.6),* $\lim_{n\to\infty} L(\phi)[n] = \sum_{i=1}^{M_k} \log U_i(c_{i,f_k} + \sum_{z_{k,j}} (\phi_{i,z_{k,j}}[n] H_{i,z_{k,j}}))$ *(i.e.* $\lim_{n\to\infty} L(\phi[n])$*) achieves the maximum of optimization problem (7.3).*

Proof. Suppose $\lim_{n\to\infty} L(\phi[n])$ does not achieve the maximum of the optimization problem. There exists $\delta > 0$, $\lambda > 0$, and positive integer N such that for all $n > N$, there exists some $i^n \in M_k$ and $z_{k,j}^n \in \mathcal{Z}_k$ so that $\phi_{i^n,z_{k,j}^n}[n] > \delta$ and $\frac{U_{i^n}'(c_{i^n,f_k} + \sum_{z_{k,j}} \phi_{i^n,z_{k,j}^n} H_{i^n,z_{k,j}^n}) H_{i^n,z_{k,j}^n}}{U_i(c_{i^n,f_k} + r_{i^n,f_k})} < \max_m \frac{U_m'(c_{m,f_k} + \sum_{z_{k,j}} \phi_{m,z_{k,j}^n} H_{m,z_{k,j}^n}) H_{m,z_{k,j}^n}}{U_m(c_{m,f_k} + r_{m,f_k})} - \lambda$. Now we have:

$$\begin{aligned} &L(\phi[n+1]) - L(\phi[n]) \geq L'(\phi[n]))\Delta\phi[n] - \frac{b}{n^2} \\ &= \sum_{i=1}^{M_k} \frac{U_i'(c_{i,f_k} + \sum_{z_{k,j}} \phi[n] H) H}{U_i(c_{i,f_k} + r_{i,f_k})} \Delta\phi[n] - \frac{b}{n^2} \\ &= \frac{\delta\lambda}{n} - \frac{b}{n^2} \geq \frac{\delta\lambda}{2n}, \end{aligned}$$

for large enough n. Since $\sum_{n=1}^{\infty} \frac{1}{n} = \infty$, which is a contradiction. As a result, $\lim_{n\to\infty} L(\phi[n])$ achieves the maximum of the optimization problem. □

7.4 Simulation Results

In this section we present simulation results for the proposed resource scheduling with CA approach. We consider a LTE-Advanced mobile system with $M = 8$ users and two CCs f_1 and f_2 available at the eNodeB with $f_1 > f_2$ as shown in Figure 7.1. We apply the user grouping method presented in 7.2 and two user groups are obtained, $\mathcal{M}_{f_1} = \{1, 2, 3, 4\}$ and $\mathcal{M}_{f_2} = \{1, 2, ..., 8\}$ where user $i \in \mathcal{M}_{f_k}$ represents the i^{th} user located under the coverage area of carrier f_k. Users $\{1, 2, 5, 6\}$ are running real-time applications that are represented by sigmoidal-like utility functions with parameters $a_i = 5$ and $b_i = 10$ for users $\{1, 5\}$ and $a_i = 1$ and $b_i = 30$ for users $\{2, 6\}$. Users $\{3, 4, 7, 8\}$ are running delay-tolerant applications that are represented by logarithmic utility functions with parameters $k_i = 15$ for users $\{3, 7\}$ and $k_i = 0.5$ for users $\{4, 8\}$. The simulation was run using MATLAB.

We compare the performance of the resource scheduling with CA approach in the case of using the proposed utility proportional fairness (UPF) resource scheduling policy and in the case of using the traditional proportional fairness (traditional-PF) scheduling policy presented in [92]. We assume equal channel gain in our simulation. In Figure 7.2, we show simulation results and compare the performance of different scheduling policies for users in $\mathcal{M}_{f_1}$ that are assigned RBs by carrier f_1 and users in $\mathcal{M}_{f_2}$ that are assigned RBs by carrier f_1 and f_2. Figure 7.2 shows the objective function of carrier f_1 RA optimization problem that is given by the multiplication of all users' applications quality of experience (QoE) for users in $\mathcal{M}_{f_1}$ and the objective function of carrier f_2 RA optimization problem when using the aforementioned scheduling policies. Figure 7.2 shows that the system performance, represented by the objective function value of the RA optimization problem that is given by the multiplication of all users applications' utilities, that represent users' satisfaction with the allocated rates in the case of the proposed UPF scheduling policy is much greater than the objective function value when using the traditional-PF scheduling policy. It also shows

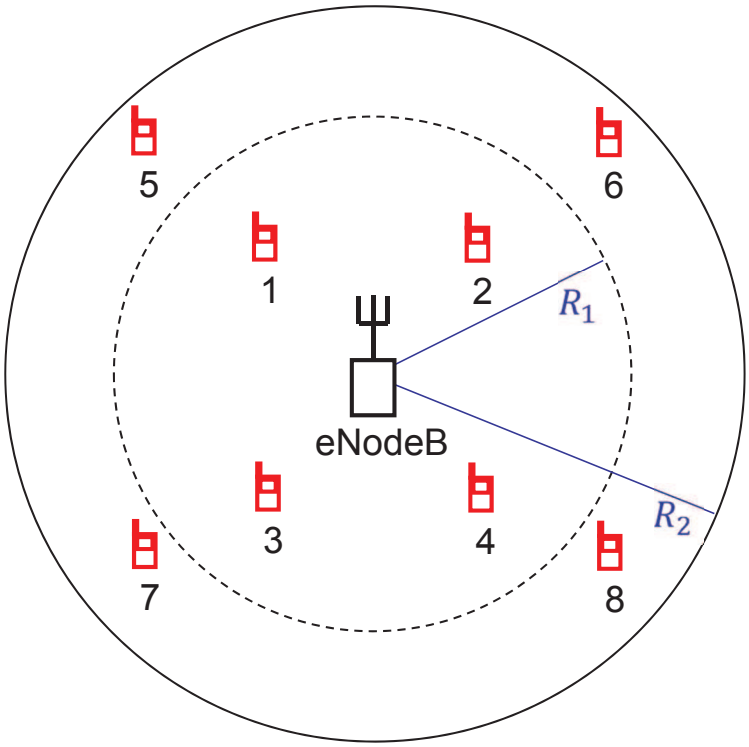


Figure 7.1: LTE-Advanced mobile system with two component carriers (i.e. f_1 and f_2) available at the eNodeB with $f_1 > f_2$ and $R_1 < R_2$.

that the system performance when using the traditional-PF with equal priority weights is worse than the system performance when using the traditional-PF with non equal priority weights.

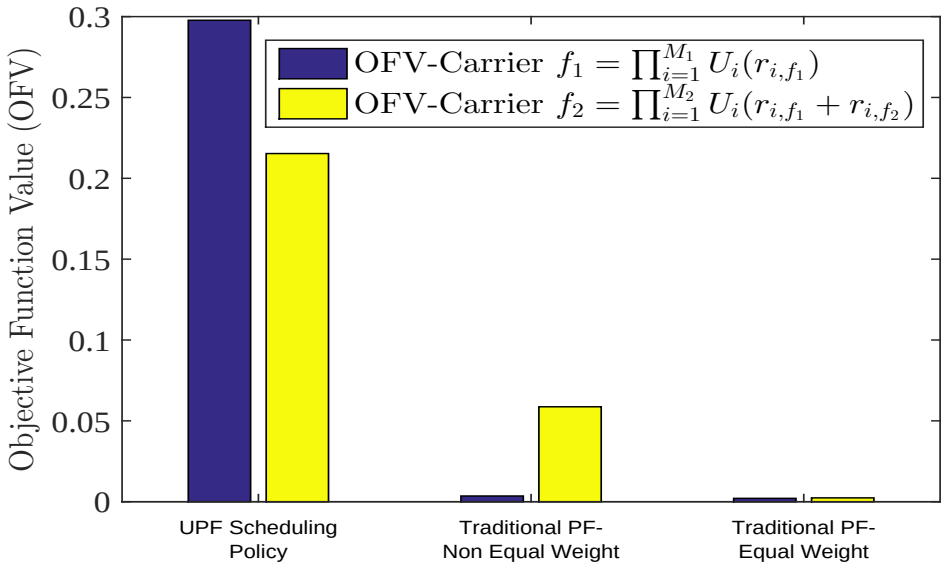


Figure 7.2: Performance comparison for different scheduling policies represented by the objective function of carrier f_1 and f_2 RA optimization problems.

7.5 Summary and Conclusions

In this cchapter, we introduced a RB scheduling with CA approach in LTE-Advanced. Users are partitioned in user groups and each user is assigned on RBs of its corresponding in range carriers. We used utility PF with CA policy and presented users' applications using utility functions. We proved that our scheduling policy exists and therefore the optimal solution is tractable. Simulation results showed that the proposed resource scheduling with CA policy achieves better QoE than the traditional proportional fairness policy.

Chapter 8

Resource Management for a Multi-Tier Wireless Spectrum Sharing System Leveraging Secure Spectrum Auctions

The dilemma between the spectrum scarcity and the booming demand for more spectrum has driven the FCC to pour out new technologies that allow opportunistic access to the under-utilized spectrum bands [95, 96]. The FCC has acknowledged the possibility for spectrum licensees to trade their unused spectrum to secondary spectrum markets by leasing it temporarily or on long term basis [55]. An efficient functionality by these secondary spectrum markets can improve spectrum utilization and therefore increase the spectrum capacity available to wireless service demands. Recently, auction theory has been recognized as a promising tool to solve various spectrum trading problems due to their economically robust allocation efficiency [56–58]. However, traditional auctions can not be directly applied in a spectrum auction because of the reusability property of the radio spectrum.

Secure spectrum auctions can revolutionize the spectrum utilization of cellular networks and satisfy the ever increasing demand for resources. In this research work, we focus on designing a framework for a multi-tier dynamic spectrum sharing system to provide an efficient sharing of spectrum with commercial wireless system providers (WSPs), with an emphasis on federal spectrum sharing. In this chapter we present a spectrum sharing system that provides an efficient usage of spectrum resources, manage intra-WSP and inter-WSP interference and provide essential level of security, privacy, and obfuscation to enable the most efficient and reliable usage of the shared spectrum. The proposed spectrum sharing system features an intermediate spectrum auctioneer responsible for allocating resources to commercial WSPs by running secure spectrum auctions while preventing possible fraud and bid-rigging.

In addition, we propose an optimal bidding mechanism for a truthful secure spectrum auction in which bidders are BSs that seek to obtain additional temporary resources by participating in the spectrum auction. By using the proposed bidding strategy, each bidder first determines its true bidding price for each number of spectrum bands it is bidding for and submits its corresponding encrypted bidding value. The proposed bidding mechanism takes into consideration that BSs participating in the spectrum auction have permanent resources and will be aggregating their winning spectrum bands (temporary resources) with the permanent resources, based on a non adjacent carrier aggregation scenario, when allocating optimal rates to active users under their coverage area. We consider a secure spectrum auction that uses homomorphic encryption through Pailliar cryptosystem to prevent possible fraud and bid-rigging. We also focus on providing an efficient resource management solution to distribute the auctioned resources among end users (i.e. UEs) in order to improve the UEs quality of experience. Furthermore, a resource allocation based on carrier aggregation approach is proposed to allocate the BS's permanent spectrum resources as well as its temporary resources (i.e. the auctioned under-utilized frequency bands) optimally among mobile users.

8.1 A Multi-Tier Wireless Spectrum Sharing System Leveraging Secure Spectrum Auctions

In this section we design a secure spectrum auction framework by considering the spectrum spatial reuse property. We propose **MTSSA**, a secure spectrum auction design that provides framework for a multi-tier dynamic spectrum sharing system to allocate spectrum resources that are managed by a broker (i.e. the auctioneer). Beside supporting multi-tier spectrum sharing systems, the computational and communication complexity for the proposed spectrum auction framework is less than other secure spectrum auctions, e.g. [81], making it a practical and implementable system. **MTSSA** allows the auctioneer to allocate its under-utilized frequency bands that are leased from federal government to commercial WSPs' BSs by running secure spectrum auction. By leveraging Paillier cryptosystem, **MTSSA** can prevent possible frauds and bid-rigging.

The major contributions of the proposed spectrum auction are summarized as:

- **MTSSA** considers spectrum reusability and the case of heterogeneous frequency bands, e.g. commercial and federal bands.
- **MTSSA** provides a framework for a multi-tier dynamic spectrum sharing system that allows an efficient spectrum sharing of the under-utilized spectrum with commercial WSPs. The auctioneer allocates the under-utilized frequency bands to commercial WSPs' BSs by running a secure spectrum auction. **MTSSA** optimizes the usage of spectrum resources by managing intra-WSP and inter-WSP interference. In order to account for frequency reusability, the network is divided into subnets and the auctioneer auctions the frequency bands in each of the subnets one after another. An intermediate federal gateway maintains a conflict-table for each BS that is updated when changes in frequency bands allocation are made.

- **MTSSA** provides a truthful auction that is achieved when each bidder submits its true evaluation value. Truthfulness is a dominant strategy for **MTSSA** as it prevents manipulating the auction.

- **MTSSA** uses a payment method that satisfies some essential economic properties such as incentive compatibility, individual rationality and no positive transfers.

- **MTSSA** leverages Paillier cryptosystem [81, 97, 98] to create a ciphertext for the bidding values. Each BS submits its bidding values through a buffer that creates an encrypted version of the bidding values. While the actual bidding values are kept secret from the auctioneer, the auctioneer is still able to reveal the auction results and charge the bidders securely.

- **MTSSA** provides a secure spectrum auction that prevents frauds of insincere auctioneers and bid-rigging with less computational and communication complexity compared to other secure spectrum auctions, e.g. [81]. Simulation results show that **MTSSA** achieves an efficient spectrum utilization, revenue and bidders' satisfaction.

8.1.1 System Model

8.1.1.1 Spectrum Trading Architecture

We consider a spectrum trading scenario where the spectrum owner is a federal regulatory agency that leases its under-utilized spectrum on a long-term basis to a broker which manages spectrum assets and plays the role of a middleman for the spectrum owner of the under-utilized spectrum, e.g. federal government, and the WSPs. The architecture of this spectrum assignments is represented through a spectrum pyramid as shown in Figure 8.1. At the top of this pyramid, is the spectrum owner that leases the under-utilized frequency bands to a spectrum broker under certain rules [55, 99, 100]. The broker represents a secondary market

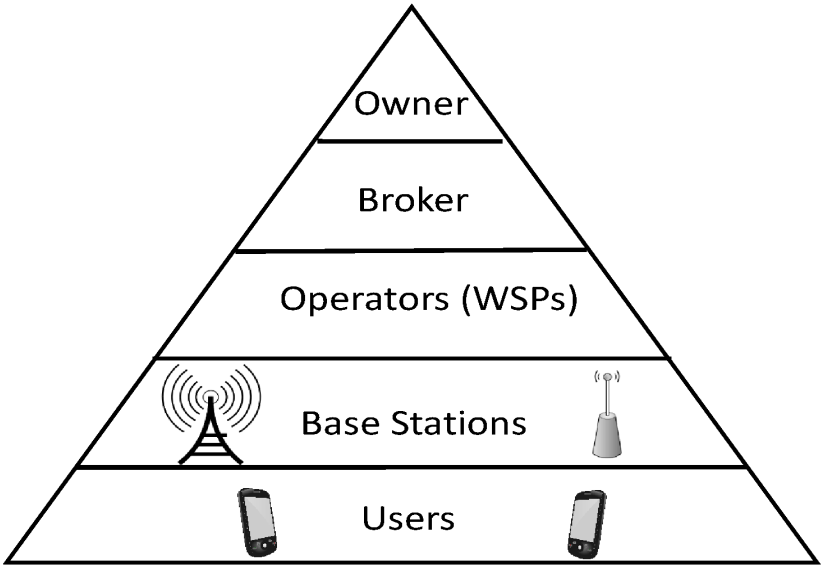


Figure 8.1: A spectrum pyramid that represents an architecture for the under-utilized spectrum assignments.

place that auctions these frequency bands to WSPs' base stations (BS)s. At the bottom of the pyramid, are the end users devices (i.e. users equipments (UE)s) that are assigned spectrum by the WSPs' BSs. We focus on designing a secure spectrum auction between the broker (i.e. the auctioneer) and the WSPs base stations to allocate the under-utilized frequency bands.

8.1.1.2 Spectrum Auction Model

Consider a spectrum auction setting, where one auctioneer (i.e. the broker in Figure 8.1) auctions a set of frequency bands $\mathcal{M} = \{1, 2, ..., M\}$ to $\mathcal{N} = \{1, 2, ..., N\}$ bidders (i.e. nodes representing BSs) located in the same geographical region where $\mathcal{N}$ represents a set of all bidders that belong to different WSPs and $N = |\mathcal{N}|$ represents the number of these bidders. Let L be the number of WSPs where each WSP has a coverage area within the auction's geographical region. Each WSP (i.e. the l^{th} WSP) provides a mobile wireless service over multiple cellular cells. Its cellular network consists of macro cells and small cells. Within the coverage area of some macro cells, there exist one or more small cells with pico/femto BSs,

see Figure 8.2. Let $\mathcal{N}^l$ be the set of all macro cells and small cells BSs that belong to the l^{th} WSP and are interested in bidding for frequency bands, the set of all bidders $\mathcal{N}$ consists of all BSs that belong to the L WSPs where $\mathcal{N} = \mathcal{N}^1 \cup \mathcal{N}^2 \cup ... \cup \mathcal{N}^L$ and $N = |\mathcal{N}|$.

We consider a multi-band spectrum auction where each BS can bid for a single or multiple frequency bands from the set of available frequency bands $\mathcal{M}$ based on its demand. Let $M = |\mathcal{M}|$ denotes the number of the auctioneer's available frequency bands and let J be the number of bidders that is announced by the federal gateway to the BSs where $N < J < N+\delta$ and δ is a positive integer that is greater than or equal 2. Once the broker leases the spectrum owner's unused frequency bands in $\mathcal{M}$ for a time duration T, the broker becomes the owner of the spectrum bands in $\mathcal{M}$. Meanwhile, the interested BSs submit their bids to the broker. In the proposed spectrum auction model, since BSs do not know the right number of bidders, each BS bids for certain number of frequency bands while taking into consideration that the number of bidders is J. Let $\mathcal{K} = \{\alpha^1, \alpha^2, ...\}$ be the allocation set for the M frequency bands that each BS bids for based on its knowledge that the number of all bidders is J. For example, given that $\mathcal{M} = \{A, B\}$ and $\mathcal{N} = \{1, 2\}$, let $J = 3$ then we have $\mathcal{K} = \{\alpha^1 = (2, 0, 0), \alpha^2 = (0, 2, 0), \alpha^3 = (0, 0, 2), \alpha^4 = (1, 1, 0), \alpha^5 = (1, 0, 1), \alpha^6 = (0, 1, 1)\}$. Each BS submits its sealed bids for the allocations in $\mathcal{K}$ where $\mathbf{b_n} = [b_n(\alpha^1), b_n(\alpha^2), ...]$ represents the sealed bids that bidder n in $\mathcal{N}$ submits for the allocations in $\mathcal{K}$, e.g. $\mathbf{b_1} = [4, 0, 0, 2, 2, 0]$ indicates that BS 1 bids 4 for allocation α^1, 0 for allocation α^2, 0 for allocation α^3, 2 for allocation α^4, 2 for allocation α^5 and 0 for allocation α^6. For certain allocation α, the n^{th} BS has a true evaluation value $v_n(\alpha)$. Let $\mathbf{v_n} = [v_n(\alpha^1), v_n(\alpha^2), ...]$ be the true evaluation vector for BS n. Let p_n represents the price that is charged by the auctioneer to BS n for allocating the frequency bands. The utility of BS n, denoted by U_n, is defined as the difference between the BS's true evaluation value and the actual price it pays to the auctioneer p_n, $U_n = v_n(\alpha) - p_n$, for a specific allocation α. The Auctioneer's revenue from the spectrum sales is defined as $R = \sum_{n=1}^{n=N} p_n$. We assume that each BS submits same bids for different allocations in $\mathcal{K}$ if the number of frequency bands that corresponds to that BS (i.e. the number of frequency

bands that the BS bids for) in these allocations is the same; i.e. in the example mentioned above BS 1 submits the same bid for allocations $\alpha^4 = (1, 1, 0)$ and $\alpha^5 = (1, 0, 1)$. We assume that bidders treat different frequency bands similarly[1]. The bidding values of each BS is proportional to the price of the frequency bands it is requesting. This price depends on the BS's demand for spectrum and is traffic-dependent; i.e. the price per unit bandwidth is optimally calculated by each BS and is proportional to the number of active UEs and the type of their applications as shown in [22]. Table 8.1 summarizes some of the notations used in the design.

Table 8.1: Key symbols

$\mathcal{M}$	Frequency bands set
$\mathcal{N}^l$	Set of all BSs that belong to WSP l
$\mathcal{N}$	Set of all BSs that belong to the L WSPs, $\mathcal{N} = \mathcal{N}^1 \cup \mathcal{N}^2 \cup ... \cup \mathcal{N}^L$
$\mathcal{K}$	Allocation set $\mathcal{K} = \{\alpha^1, \alpha^2, ...\}$
$\mathbf{b_n}$	BS n sealed bids vector for the allocation set $\mathcal{K}$
$\mathbf{v_n}$	True evaluation vector of BS n
p_n	Price charged by the auctioneer to BS n
U_n	BS n utility, $U_n = v_n(\alpha) - p_n$
R	Auctioneer's revenue, $R = \sum_{n=1}^{n=N} p_n$

In Figure 8.2, we show two WSPs (i.e. $L = 2$) providing service in the same geographical region where the broker performs its spectrum auction. Both WSPs are interested in the auctioneer's frequency bands $\mathcal{M}$. Therefore, both of them participate in the spectrum

[1]This is a valid assumption since bidders in the system model are BSs and not end users. Once each bidder/BS is allocated frequency bands by the auctioneer, the BS then allocates these bands to UEs such that the channel state is maximized for the link between the BS and each UE in order to maximize the BS's aggregated throughput. Therefore, each BS submits its bidding values based on the number of frequency bands it is bidding for and does not submit different bids for different frequency bands since these frequency bands will be eventually utilized by allocating them to UEs subscribing for the BS's resources. If bidders are end users (i.e. UEs) and not BSs as the model presented in [81], each bidder submits its bidding values based on which frequency bands it is bidding for. In this case, each UE bids for more allocations since the set $\mathcal{K}$ consists of more possible allocations when bidders are end users compared to the number of allocations when bidders are BSs.

auction. In the coverage area of each WSP there exists multiple macro cells and small cells managed by that WSP. BSs requesting additional frequency bands submit sealed bidding vectors to the auctioneer via an intermediate secure gateway to participate in the auction of the under-utilized federal spectrum bands. Considering the frequency reuse property [60,65], each BS has certain coverage radius (i.e. assume it is equivalent to the cell's radius). Within the coverage radius of the n^{th} BS, non of the interfering BSs can simultaneously use any of the frequency bands that the n^{th} BS is using. However, a non-interfering BS can use the same frequency band that is simultaneously used by a BS located outside its coverage radius without causing interference, i.e. frequency reuse is utilized in our model. An interference conflict graph is constructed by the federal gateway for all the BSs that are participating in the auction. In Figure 8.3, we show the frequency conflict graph with all bidders/BSs that belong to the two WSPs, each BS is connected with other BSs located within its coverage radius (i.e. BS n is connected with all BSs that must not simultaneously use same frequency bands due to interference between them) where the edges represent mutual interference between the corresponding BSs. The interference conflict graph can be constructed using physical or protocol channel model [101]. The interference conflict graph is constructed and updated by an intermediate gateway that is operated by federal government which keeps it unknown to the bidders. It is important to keep the interference conflict graph unknown to the bidders so that each bidder is unaware of other BSs connected to it in the interference conflict graph to prevent possible bid-rigging among bidders in case if certain bidder is aware of which BSs are connected to it and attempts to collude with them. It is assumed that there exists a pilot channel, like the one in [85], to exchange information between the federal gateway and the BSs or simply by sending that unsecured information with the bids. Furthermore, the proposed spectrum auction is executed in one subnet after another where a subnet is defined to be a group of BSs that includes one root BS, i.e. BS n, and all other BSs that are connected to it through interference edges (i.e. the BSs that have mutual interference with BS n) but not previously considered root BSs. Figure 8.3 shows two subnets in the

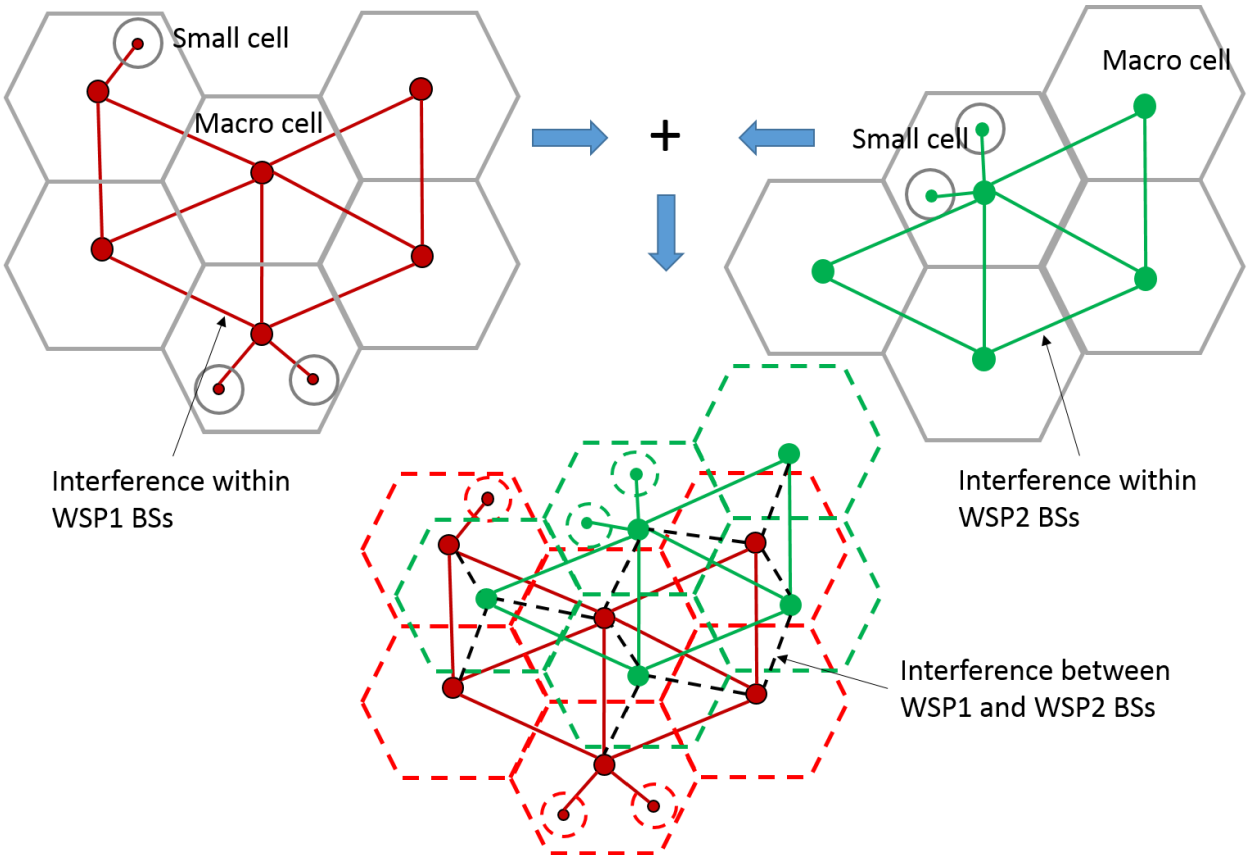


Figure 8.2: Two WSPs with a coverage area within the geographical region where the auction takes place. In each WSP's macro cells and small cells, all the BSs that are interested in the auctioneer's under-utilized frequency bands are part of the interference conflict graph.

frequency conflict graph of the two WSPs. In Figure 8.4, we show the spectrum auction model for **MTSSA** for two WSPs that participate in the spectrum auction. First, all BSs submit their encrypted bidding vectors to the federal gateway. The auctioneer then carries out a secure spectrum auction in one subnet after another. It then allocates the winning BSs frequency bands and charges them for the allocated resources.

8.1.2 Design Considerations

In this section, we present the payment method for the proposed auction. We also discuss some economic properties that need to be considered in the design and prove that by using a VCG based auction approach, when running the auction in each subnet, some desired

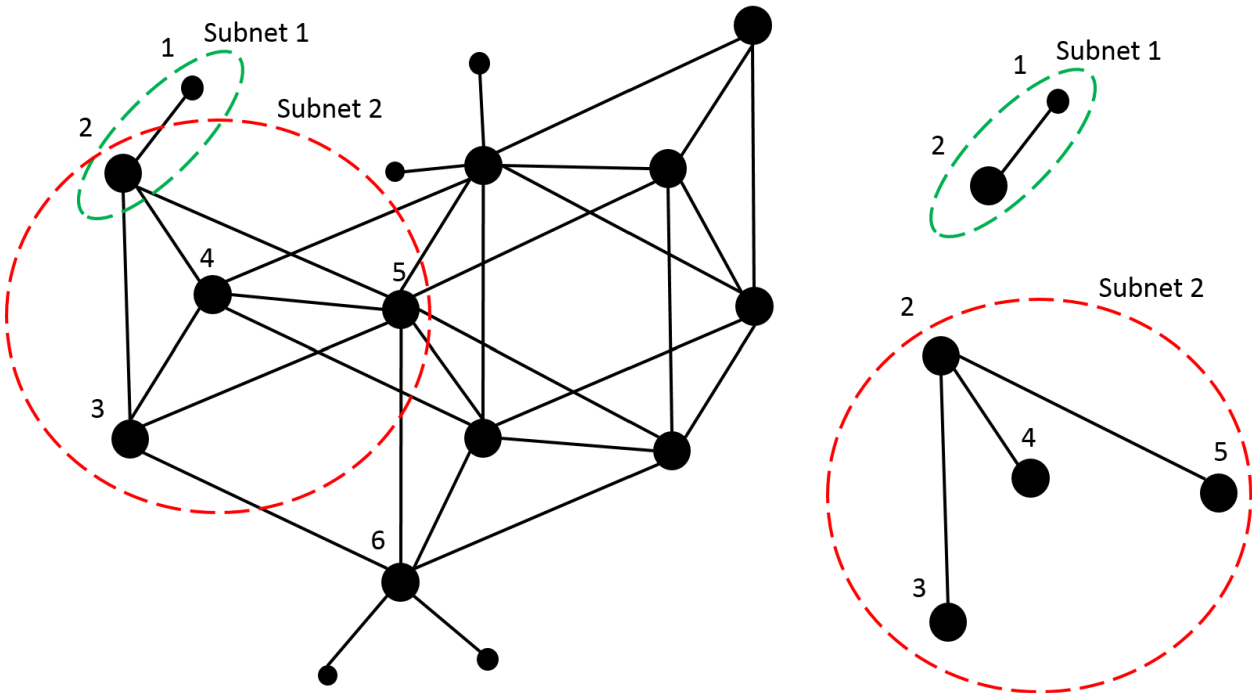


Figure 8.3: Frequency conflict graph for all BSs that belong to the two WSPs shown in Figure 8.2. Each node represents one BS and the edges represent mutual interference between the end points (i.e. BSs). Subnet 1 consists of the small cell's BS (i.e. BS 1), which represents the root BS for the subnet, and the macro cell's BS (i.e. BS 2). Subnet 2 consists of BSs 2, 3, 4 and 5 where BS 2 is the root BS.

economic properties can be satisfied.

8.1.2.1 The Payment Method

Our goal is to use a payment rule that satisfies some of the required economic properties, such as incentive compatibility, individual rationality and no positive transfers. In addition, it is important for the payment rule to provide a satisfactory revenue for the auctioneer. Under certain assumptions, it has been proven that VCG auction satisfies these three economic properties while maximizing the auctioneer's revenue [102]. VCG auction is also proven to be Pareto efficient [103]. In VCG, each bidder submits its true evaluation values regardless of the bidding values that other bidders submit. This is a dominant strategy for the bidder to maximize its utility and win the auction. In our design, we use a payment method that is

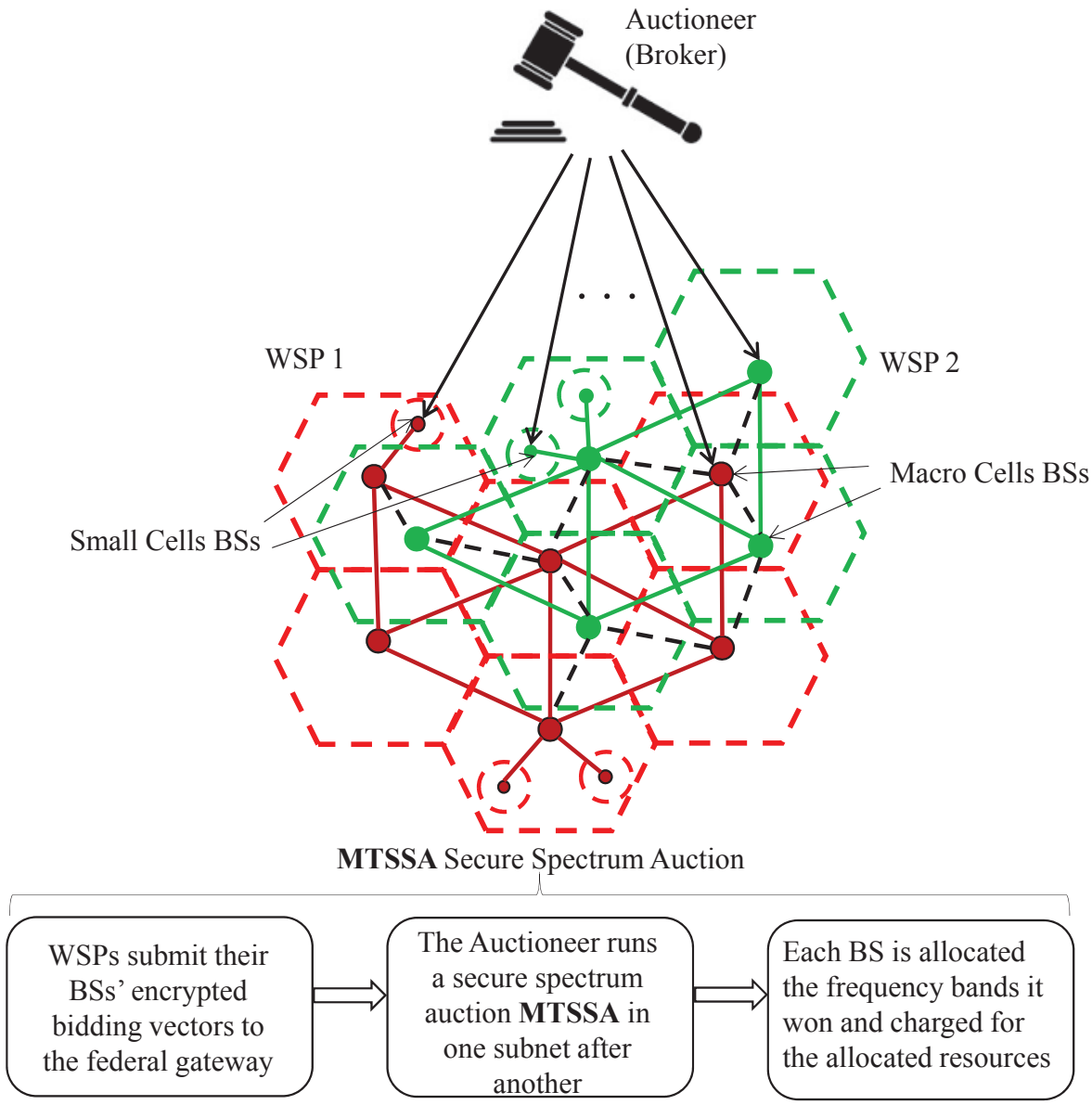


Figure 8.4: Spectrum auction model for the proposed **MTSSA** with two WSPs' BSs participating in the auction.

based on VCG mechanism with Clarke pivot payments [104]. Using this payment rule, each bidder pays the difference between the social welfare with and without his participation (i.e. bidder n pays the externality he causes). Consider the same system setup as described in Section 8.1.1. Each BS n submits its sealed bidding vector $\mathbf{b_n}$ for the allocation set $\mathcal{K}$. The auctioneer selects a Pareto efficient allocation $\alpha^* \in \mathcal{K}$ where α^* is defined as

$$\alpha^* = \arg\max_{\alpha \in \mathcal{K}} \sum_n b_n(\alpha). \tag{8.1}$$

With truthful bidding values, the auctioneer assigns its frequency bands $\mathcal{M}$ to bidders based on α^* allocation. Furthermore, let $\alpha^*_{-n} \in \mathcal{K}$ be an allocation without BS n participating that is defined as

$$\alpha^*_{-n} = \arg\max_{\alpha \in \mathcal{K}} \sum_{k \neq n} b_k(\alpha). \tag{8.2}$$

The auctioneer charges BS n a payment p_n that is equivalent to

$$p_n = \sum_{k \neq n} b_k(\alpha^*_{-n}) - \sum_{k \neq n} b_k(\alpha^*). \tag{8.3}$$

Then, the utility of BS n can be expressed as

$$\begin{aligned} U_n &= v_n(\alpha^*) - p_n \\ &= v_n(\alpha^*) - (\sum_{k \neq n} b_k(\alpha^*_{-n}) - \sum_{k \neq n} b_k(\alpha^*)) \\ &= [v_n(\alpha^*) + \sum_{k \neq n} b_k(\alpha^*)] - \sum_{k \neq n} b_k(\alpha^*_{-n}). \end{aligned} \tag{8.4}$$

8.1.2.2 Desired Economic Auction Properties

It is essential for an auction to have certain economic properties. First, we discuss these economic properties and then we prove that by using a VCG based auction approach these

properties can be satisfied.

1. Incentive Compatibility (truthfulness): An auction is incentive compatible if non of the bidders can get higher utility by not reporting its true evaluation vector. Based on this property, a dominant strategy for any bidder is to declare its true evaluation value regardless of what the other bidders do.

2. Individual Rationality: An auction is individually rational if the utility U_n for each bidder n is greater or equal zero (i.e. $U_n \geq 0$). Meaning that the winning bidders obtain non-negative utility (i.e. bidders do not pay more than their evaluation values) from the auction and no one suffer as a result of participating in the auction.

3. No Positive Transfers: In auctions with no positive transfers, the payment of any bidder n must be greater or equal zero (i.e. $p_n \geq 0$). This prevents situations when the auctioneer has to pay the bidders.

In Lemma 8.1.1, Lemma 8.1.2 and Lemma 8.1.3, we show that by using the payment method and the VCG based auction approach discussed above, the aforementioned desired economic properties can be satisfied.

Lemma 8.1.1. *Let* $\mathbf{v_n}$ *and* $\mathbf{v_n^{'}} \neq \mathbf{v_n}$ *be the* n^{th} *BS bidding vector when it is equivalent to its true evaluation values and any other values, respectively, and let* α^* *and* $\alpha^{*'}$ *be the allocations that maximize the social welfare when* $\mathbf{v_n}$ *and* $\mathbf{v_n^{'}}$ *are declared, respectively. Then, for the* n^{th} *BS, the utility* $U_n \geq U^{'}{}_n$.

Proof. Using the utility definition and payment method in Section 8.1.2.1, the utility of BS n is $U_n = v_n(\alpha^*) + \sum_{k \neq n} v_k(\alpha^*) - \sum_{k \neq n} v_k(\alpha^*_{-n})$ when declaring $\mathbf{v_n}$ whereas the utility of BS n is $U^{'}{}_n = v_n(\alpha^{*'}) + \sum_{k \neq n} v_k(\alpha^{*'}) - \sum_{k \neq n} v_k(\alpha^*_{-n})$ when declaring $\mathbf{v}^{'}{}_n$. Since α^* is the

allocation that maximizes the social welfare, we have the following inequality:

$$\sum_{n} v_n(\alpha^*) \geq \sum_{n} v_n(\alpha^{*'}). \tag{8.5}$$

Now, by subtracting the term $\sum_{k\neq n} v_k(\alpha^*_{-n})$ from both sides of equation (8.5), we get $U_n \geq U^{'}{}_n$ which is the incentive compatibility property. □

Lemma 8.1.2. *Let α^* and ${\alpha^*}_{-n}$ be the allocations that maximize the social welfare with and without BS n's participation, respectively, with the assumption that each BS submits its true evaluation values. Then each BS n do not suffer as a result of participating in the auction and the auction's winners do not pay more than their evaluation values (i.e. $U_n \geq 0$).*

Proof. To show individual rationality, consider the utility of BS n:

$$\begin{aligned} U_n &= v_n(\alpha^*) + \sum_{k\neq n} v_k(\alpha^*) - \sum_{k\neq n} v_k(\alpha^*_{-n}) \\ &\geq \sum_{j} v_j(\alpha^*) - \sum_{j} v_j(\alpha^*_{-n}) \\ &\geq 0. \end{aligned}$$

The first inequality holds since $v_n(\alpha^*)+\sum_{k\neq n} v_k(\alpha^*) = \sum_j v_j(\alpha^*)$, $\sum_j v_j(\alpha^*_{-n}) \geq \sum_{k\neq n} v_k(\alpha^*_{-n})$ and $\sum_j v_j(\alpha^*_{-n}) \geq 0$. The second inequality holds because α^* is the allocation that maximizes the social welfare, $\sum_j v_j(\alpha^*)$. □

Lemma 8.1.3. *As a result of using the payment method in Section 8.1.2.1, the auction has no positive transfers (i.e. $p_n \geq 0\ \forall n \in \mathcal{N}$).*

Proof. From equation (8.3), we have $p_n = \sum_{k\neq n} b_k(\alpha^*_{-n}) - \sum_{k\neq n} b_k(\alpha^*) \geq 0$, since α^*_{-n} is the allocation that maximizes the social welfare without the n^{th} BS participation, $\sum_{k\neq n} b_k(\alpha^*_{-n})$.

□

8.1.2.3 Design Challenges

Truthfullness is one of the important properties that needs to be taking into consideration when designing a spectrum auction. Sealed secondary price auction and VCG auction are very preferable as they guarantee that bidders submit their true evaluation values. As mentioned before, VCG auction has many properties that are essential to have in a spectrum auction. However, VCG requires finding an optimal allocation which is NP-complete because of the spectrum spatial reusability property. In addition, VCG is vulnerable to frauds of the auctioneer and bid-rigging between the insincere auctioneer and greedy bidders [81]. Therefore, VCG auction can not be used in a spectrum auction without countermeasures for fraud and bid-rigging.

Bid-rigging between a greedy bidder and an auctioneer can occur for the benefit of both. Since the auctioneer is aware of all bidders' bidding values, he can collude with a greedy bidder and reveal the winning bid value to him. In Figure 8.5, we show an example of a spectrum auction where the auctioneer auctions one frequency band $|\mathcal{M}| = 1$ to four BSs (i.e. subnet 2 of the frequency conflict graph that is shown in Figure 8.3). The auctioneer runs a VCG auction that is equivalent to a sealed secondary price auction for one frequency band auction. In Figure 8.5(a), we show an example of bid-rigging. Bidder 4 is the winner and bidder 2 is a greedy bidder who colludes with the auctioneer and learns about the highest bid. As a result, bidder 2 bids a value that is higher than his true evaluation but a little bit less than the highest bid. By doing so, the auctioneer considers the bidding value of bidder 2 to be the charging price for the winner (i.e. bidder 4). By such a bid-rigging action, the auctioneer can make more profit and share the spoils with bidder 2. Bid-rigging can also occur among bidders without the auctioneer's participation for the benefit of one or multiple bidders. This is possible if a bidder succeeds to know the bidding values of all bidders in his subnet. However, since each subnet has bidders that belong to different WSPs bid-rigging among bidders is less likely to happen as it is difficult for a bidder to know the bidding values

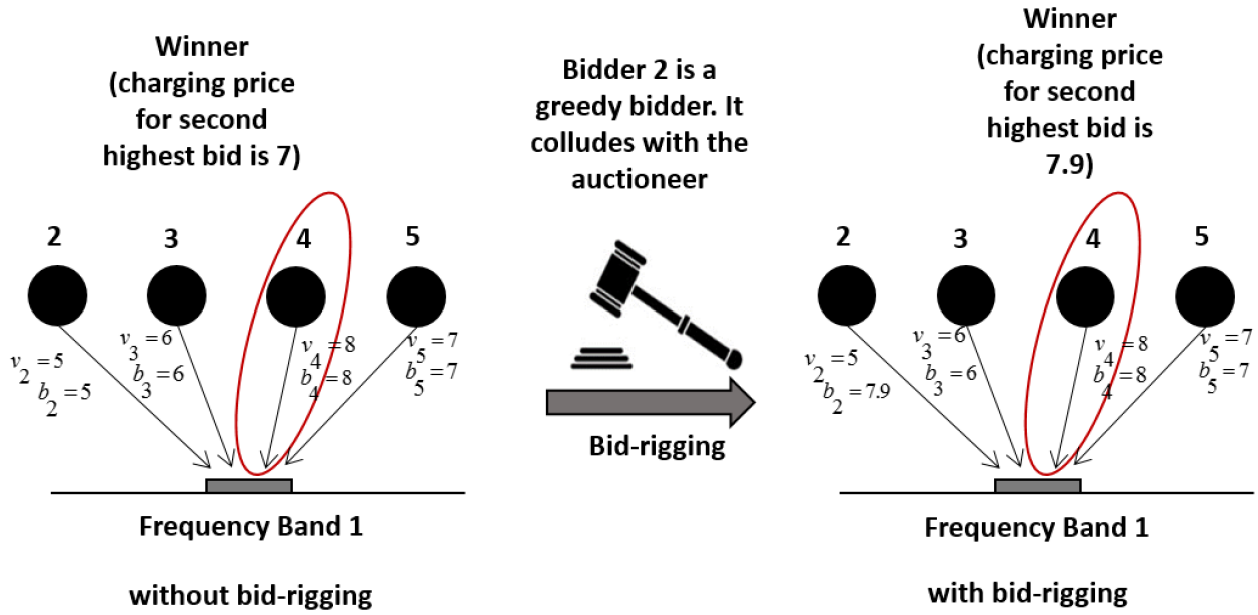
of bidders that are not in his WSP. If all bidders in a subnet belong to the same subnet and a bidder is aware of that and succeeds to learn other bidders' bidding values, which may be possible since all bidders belong to the same WSP, he can collude with other bidders and manipulate the auction.

On the other hand, a fraud occurs when an insincere auctioneer overcharges the winner in order to increase his own profit. This results is an unexpected bad utility for the winner. In Figure 8.5(b), bidder 4 is the winner and the charging price should be 7 which is equivalent to the second highest bid. However, the insincere auctioneer charges bidder 4 at 7.9 to obtain higher revenue. This is possible since all bidding values are sealed and bidders do not know about the bidding values of each other.

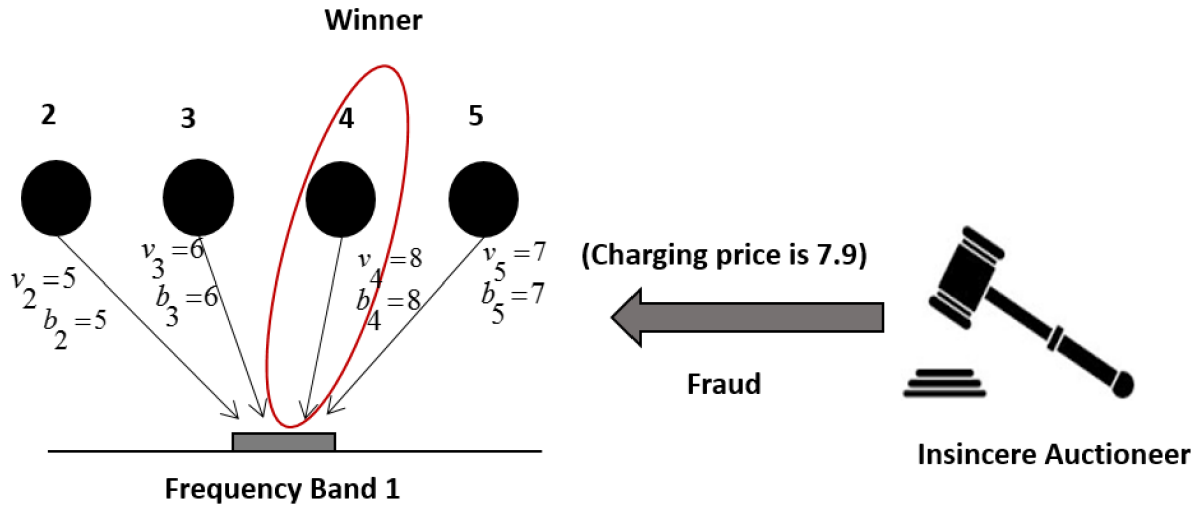
To avoid possible bid-rigging and frauds, a successful spectrum auction design needs to take into consideration securing the auction by making the auctioneer able to decide how to allocate the frequency bands while keeping the bidders actual bidding values unknown to the auctioneer. This is essential to avoid possible back-room dealing and ensure a secure spectrum auction.

8.1.3 MTSSA: Secure Spectrum Auction Design

In order to enable an efficient usage of the under-utilized shared spectrum managed by a broker. It is important to design a secure spectrum auction that allows the broker to provide a sufficient level of security, privacy and obfuscation to enable a reliable and efficient usage of the shared spectrum. In order to thwart back-room dealing, it is essential to have a mechanism that allows the auctioneer to find the maximum bid among all bidders without knowing their actual bids. The proposed **MTSSA** leverages Paillier cryptosystem to avoid possible frauds and bid-rigging between greedy bidders and the auctioneer. On the other hand, in order to avoid bid-rigging between bidders, the federal gateway decides whether



(a) Bid-rigging between an insincere auctioneer and a greedy bidder



(b) Frauds of an insincere auctioneer

Figure 8.5: Examples of bid-rigging and frauds in an unsecured spectrum auction of one frequency band and four BSs.

to provide the bidders with the right number of the current bidders or not based on the number of bidders participating in the spectrum auction and the WSPs they belong to. If all bidders belong to the same WSP or there are only small number of bidders participating

in the auction, the federal gateway provides the BSs with a random number J for the number of bidders where $N < J < N + \delta$ and δ is a positive integer that is greater than or equal 2 whereas it provides the bidders with the exact number of bidders if the bidders belong to multiple WSPs. In this section, we first describe Paillier cryptosystem and point out its special features. We then discuss **MTSSA** frequency bands allocation procedure. Finally, we present the security part of **MTSSA**.

8.1.3.1 Paillier Cryptosystem

Some of Paillier cryptosystem properties are essential for our secure spectrum auction design. Paillier cryptosystem [81,97,98] is a probabilistic public key encryption system, i.e. the term probabilistic encryption indicates that when encrypting the same plaintext for multiple times it yields different ciphertexts, that satisfies special features such as homomorphic addition, indistinguishability and self blinding.

The homomorphic properties of Paillier cryptosystem provide it with a notable feature. As the encryption function of a message m, is given by $C(m)$, is additively homomorphic. i.e. $C(m_1 + m_2) = C(m_1)C(m_2)$. On the other hand, with the indistinguishability property of Paillier cryptosystem, if the plaintext m is encrypted twice, the two created cyphertexts are different from each other and no one can distinguish the original plaintexts, except by random guessing, unless decrypting the original ciphertexts. The self blinding property allows changing the ciphertext publicly without affecting the plaintext. Therefore, from the ciphertext $C(m)$, it is possible to compute a different randomized ciphertext $C^{'}(m)$ without knowing the decryption key or the original plaintext.

8.1.3.2 Frequency Bands Allocation Procedure

All BSs that are interested in the auction and belong to the WSPs within the geographical region of the auction submit their bidding values to participate in the auction. Based on the location of these BSs and which WSPs they belong to, the federal gateway creates an interference conflict graph (i.e. like the one in Figure 8.3). The auctioneer executes the auction in one subnet at a time. For each subnet, the auctioneer selects a random BS $n \in \mathcal{N}$ to be the current root BS and considers its corresponding subnet, i.e. connected nodes/BSs. After solving for the current subnet, the auctioneer selects a new BS, that has not been a root BS before, to be the new root BS and excludes any previous root BS from its subnet along with the allocated frequency bands to these BSs. Following the same procedure, the auctioneer continues to execute the auction in one subnet after another until each BS has participated in the auction. Based on the subnet auction results, the auctioneer allocates the corresponding root BS the frequency bands and charges it for the allocated resources. The federal gateway maintains a conflict table for each BS participating in the auction, i.e. as in [67].

The **MTSSA** procedure is presented in the following steps:

1. The federal gateway sets up a conflict-table for each BS $n \in \mathcal{N}$ and each BS submits its encrypted version of bidding values $\mathbf{b_n}$: The federal gateway creates a set of all BSs $\mathcal{N}^l$ that are interested in bidding for the auctioneer's under-utilized frequency bands for each WSP l within the auctioneer's geographical region. The federal gateway creates a conflict table for each BS $n \in \mathcal{N}^l$ with all the interfering BSs denoted by $\mathcal{I}_n$ (i.e. $\mathcal{I}_n$ is a set of all BSs that are located within the coverage area of BS n). The interfering BSs $\mathcal{I}_n$ for each bidder $n \in \mathcal{N}^l$ is updated by the federal gateway[2]. The federal gateway decides the

[2]It is assumed that the federal gateway is aware of all BSs in the coverage area of each WSP l, within the auction's geographical region, whether they belong to WSP l or to other WSPs. Therefore, the set of interfering BSs $\mathcal{I}_n$ includes all BSs within the coverage area of BS n that belong to WSP l as well as BSs that belong to other WSPs.

value of J and sends out the number of bidders J to each bidder. Each bidder $n \in \mathcal{N}^l$ creates its bidding vector $\mathbf{b_n}$ that will be an input to a buffer that encrypts the bidding values. The encrypted bids are then submitted to federal gateway for randomization, see Section 8.1.3.3, then sent to auctioneer, see Figure 8.4. Neither the auctioneer nor the other BSs know the actual bidding values $\mathbf{b_n}$ that BS n has submitted. We show in Section 8.1.3.3.1 the procedure of encrypting the bidding values using Paillier encryption.

2. **Start with a random BS $n \in \mathcal{N}$ and consider its corresponding subnet:** The auctioneer does not have an optimal choice regarding which subnet it starts the auction from in order to maximize his revenue. Therefore, the auctioneer selects a random bidder n from the set $\mathcal{N}$ and considers its corresponding subnet (i.e. a subnet consists of a root BS n and all other BSs connected to BS n in the interference conflict graph except BSs that have been previously considered root BSs).

3. **The auctioneer carries out a secure spectrum auction in the subnet of the selected BS n:** The auctioneer performs a secure spectrum auction procedure (detailed in Section 8.1.3.3) in the current subnet under consideration.

4. **Allocate frequency bands and charge price:** Based on the subnet auction's results, the auctioneer allocates the root BS frequency bands and charges it for the allocated resources. The allocation and the payment vary based on the location of root BS and its relative bid with respect to neighboring BSs. For each winning BS, the federal gateway stores the allocated frequency bands and the auctioneer's charging price in the conflict table.

5. **Proceed to next root BS:** A new root BS is selected based on a random selection done by the auctioneer and the corresponding subnet secure bids are sent to federal gateway and the process is repeated starting from Step **2**.

Algorithm 28 summarizes **MTSSA** spectrum auction procedure.

Algorithm 28 MTSSA Frequency Bands Allocation

$\mathcal{N} = \mathcal{N}^1 \cup \mathcal{N}^2 ... \cup \mathcal{N}^L$ {i.e. $\mathcal{N}$ is the set of all BSs in the interference conflict graph}
$\mathcal{N}_0 = \mathcal{N}$
Auctioneer generates his private and public keys of Paillier cryptosystem
Auctioneer sends public key and element x to all BSs via pilot channel
while $\mathcal{N} != \phi$ **do**
  $n = random(\mathcal{N})$ {Auctioneer selects a random BS}
  $\mathcal{N}_n = include\_conflict(n)$ {Auctioneer adds BSs that form the n^{th} subnet from conflict table of n^{th} BS as root}
  $\mathcal{N}_- = (\mathcal{N}_0 \setminus \mathcal{N}) \cap \mathcal{N}_n$ {$\mathcal{N}_-$ is set of previous root BSs in the n^{th} subnet}
  $\mathcal{N}_n = \mathcal{N}_n \setminus \mathcal{N}_-$ {Auctioneer removes from $\mathcal{N}_n$ previous root BSs}
  $\mathcal{M}_- = include\_alloc(\mathcal{N}_-)$ {$\mathcal{M}_-$ is set of freq. bands allocated to $\mathcal{N}_-$}
  $\mathcal{M}_n = \mathcal{M} \setminus \mathcal{M}_-$ {Auctioneer removes from $\mathcal{M}$ freq. bands allocated to $\mathcal{N}_-$}
  $\mathcal{K}_n = alloc\_vect(\mathcal{N}_n, \mathcal{M}_n)$ {Auctioneer forms allocation vector $\mathcal{K}_n$ and sends to $\mathcal{N}_n$}
  BSs $\in \mathcal{N}_n$ send encrypted bids to federal gateway
  Federal gateway randomizes bids and forward to auctioneer
  Auctioneer selects the highest allocation $\alpha^\star$
  Auctioneer charges price p_n to BS n
  $\mathcal{N} = \mathcal{N} \setminus \{n\}$
end while

8.1.3.3 Secure Spectrum Auction Using Paillier Cryptosystem

In order for **MTSSA** to ensure a secure auction, it is important to design **MTSSA** such that no way for the auctioneer to manipulate the auction. VCG auction is proven to have the incentive compatibility property from the bidders side which is essential for our design. In order for **MTSSA** to prevent the auctioneer from conducting any frauds or bid-rigging [81], it is important to limit the auctioneer's capability by making him only able to exploit the winners and their payments without knowing the actual bidding values. So, by leveraging Paillier cryptosystem in our design, **MTSSA** can ensure a secure spectrum auction. Next, we discuss in details how both the BSs and the auctioneer need to collaborate in order to carry out a secure spectrum auction.

8.1.3.3.1 Impact of Paillier Cryptosystem on the Bidding Values

Encrypting the Bidding Values: Each BS submits its bidding values through a buffer that uses Paillier cyptosystem to encrypt the bidding values and create a vector of ciphertexts for each bidding value. Let s be a number that any actual bidding value does not exceed and let $z = b(\alpha)$ be the actual bidding value for allocation α such that $1 \leq z \leq s$. Let the vector of ciphertexts for z be $\mathbf{c}(z)$. As in [105], the vector of ciphertexts $\mathbf{c}(z)$ is given by

$$\begin{aligned} \mathbf{c}(z) &= (c^1, ..., c^s) \\ &= (\underbrace{C(x), ..., C(x)}_{z}, \underbrace{C(0), ..., C(0)}_{s-z}), \end{aligned} \tag{8.6}$$

where $C(x)$ is the Paillier encryption of the public element x (i.e. $x \neq 0$) and $C(0)$ is the Paillier encryption of 0. As mentioned before, C has a self blinding property which makes z undeterminable without decrypting the elements in $\mathbf{c}(z)$.

Selecting the Maximum Bidding Value: The auctioneer can determine the BS with the maximum bidding value from the encrypted bidding values without knowing their actual values. Let $\mathbf{c}(z_j) = (c_j^1, ..., c_j^s)$ be the encrypted bidding vector of BS n for certain allocation α. First, consider the product of all encrypted bidding vectors for allocation α,

$$\prod_j \mathbf{c}(z_j) = (Q_1, ..., Q_s) = (\prod_j c_j^1, ..., \prod_j c_j^s). \tag{8.7}$$

Due to the homomorphic addition property of Paillier cryptosystem, Q_i (i.e. $1 \leq i \leq s$ and $i \neq j$) is equivalent to

$$Q_i = \prod_j c_j^i = C^{\gamma(i)}(x) = C(\gamma(i)x), \tag{8.8}$$

where $\gamma(i)$ represents the number of values that are equal to or greater than i, i.e. $\gamma(i) = |\{j : z_j \geq i\}|$. Given that $\gamma(i)$ monotonically decreases when i increases, one way to find the maximum of these bidding values is to decrypt Q_i and check whether the decrypted value $C^{-1}(Q_i)$ equals 0 or not for i changing from $s \rightarrow 1$. Once the largest i with a decrypted value $C^{-1}(Q_i) \neq 0$ is found, then the maximum bidding value for the allocation α is determined to be i (i.e. $i = \max\{z_j\}$).

Randomizing the Encrypted bidding Values: Without knowing z, the federal gateway adds a constant t to the encrypted vector $\mathbf{c}(z)$ and randomizes the rest of its elements. This results in the following vector

$$\mathbf{c}^{'}(z+t) = (\underbrace{C(x), ..., C(x)}_{z}, c_1^{'}, ..., c_{s-z}^{'}), \tag{8.9}$$

where t can not be obtained from either $\mathbf{c}(z)$ or $\mathbf{c}^{'}(z+t)$ because of the self blinding property of Paillier cryptosystem. In addition, t can not be figured out by comparing $\mathbf{c}(z)$ and $\mathbf{c}(z+t)$ during the shifting and randomizing process.

8.1.3.3.2 Securing the MTSSA Subnet Auction

By using Paillier cryptosystem as discussed in Section 8.1.3.3.1, with encrypted bidding values it is still possible to find the maximum bid and the encrypted bidding vectors are randomized without knowing their actual values. This makes it possible to apply a VCG based auction in each subnet. As mentioned before, the proposed **MTSSA** auction is carried out in one subnet after another. In certain subnet, **MTSSA** auction is performed as follows:

1. The auctioneer generates his private and public keys of Paillier cryptosystem and publishes his public key and element x over the pilot channel.

2. Each BS submits its sealed bidding vector $\mathbf{b_n} = \mathbf{v_n}$ (i.e. its true evaluation values since we consider a VCG based auction). The auctioneer creates representing vectors $\mathbf{C}_T = \mathbf{C}(O)$, $\mathbf{C}_1 = \mathbf{C}(O)$,...,$\mathbf{C}_N = \mathbf{C}(O)$ where N is the number of bidders (i.e. BSs), the initial $O(\alpha)$ equals 0 and the size of vector $\mathbf{C}$ equals $|\mathcal{K}|$. In order for BS z to keep his bidding value b_z secret, he adds his encrypted bidding value to all of the representing vectors except $\mathbf{C}_z$. Once all BSs are done performing this addition, the auctioneer obtains

$$\mathbf{C}_T = (\prod_n \mathbf{c}(b_n(\alpha_1)), ..., \prod_n \mathbf{c}(b_n(\alpha_{|\mathcal{K}|}))). \tag{8.10}$$

Due to the homomorphic addition property of Paillier cryptosystem, equation (8.10) is equivalent to

$$\mathbf{C}_T = (\mathbf{c}(\sum_n b_n(\alpha_1)), ..., \mathbf{c}(\sum_n b_n(\alpha_{|\mathcal{K}|}))) = \mathbf{C}(\sum_z b_z), \tag{8.11}$$

and

$$\mathbf{C}_z = \mathbf{C}(\sum_{n \neq z} b_z) \quad 1 \leq z \leq N. \tag{8.12}$$

3. Federal gateway adds $\theta(\alpha) = t$ to $\mathbf{C}_T, \mathbf{C}_1, ..., \mathbf{C}_N$ to obtain $\mathbf{C}(\sum_n b_n + \theta)$ and $\mathbf{C}(\sum_{n \neq z} b_n + \theta) \forall z$. It sends these randomized encrypted bids to auctioneer.

4. In order for the auctioneer to select an allocation for the current subnet and find its corresponding charging price, it finds the maximum sum value

$$\begin{aligned} g &= \arg\max_{\alpha \in \mathcal{K}} (\sum_n b_n(\alpha) + \theta(\alpha)) \\ &= \arg\max_{\alpha \in \mathcal{K}} (\sum_n b_n(\alpha)) + t, \end{aligned} \tag{8.13}$$

which can be determined by the auctioneer by taking the product of all the encrypted elements in $\mathbf{C}_T$, i.e. as discussed in Section 8.1.3.3.1, which is equivalent to $\prod_{i=1}^{|\mathcal{K}|} \mathbf{c}(\sum_{n=1}^{N} b_n(\alpha_i) +$

t). The auctioneer determines the maximum element in that product which is equivalent to g in (8.13).

5. For each allocation α, the auctioneer decrypts the g^{th} element of vector $\mathbf{c}(\sum_n b_n(\alpha) + \theta(\alpha))$ in $\mathbf{C}_T$ and finds whether it equals 0 or x. If it equals x at allocation α^*, then the auctioneer selects α^* to be the allocation that maximizes $\sum_n b_n$ in the current subnet and considers its corresponding BSs.

6. To find the charging price for the root BS, the auctioneer decrypts $\mathbf{c}(\sum_{n\neq z} b_n(\alpha^*) + \theta)$ of $\mathbf{C}_z$ and finds the masked value $(\sum_{n\neq z} b_n(\alpha^*) + \theta)$.

7. The auctioneer then finds the maximum masked bid of the product of the encrypted elements $\max_{\alpha\in\mathcal{K}}(\sum_{n\neq z} b_n(\alpha)+t)$, similar to Step **4**, which is equivalent to $(\sum_{n\neq z} b_n(\alpha^*_{-z})+t)$.

8. The auctioneer then finds the charging price for root BS of allocation α^* that is given by

$$p_z = (\sum_{n\neq z} b_n(\alpha^*_{-z}) + t) - (\sum_{n\neq z} b_n(\alpha^*) + t). \tag{8.14}$$

8.1.4 Simulation and Analysis

In this section, we first evaluate the performance of the proposed **MTSSA** spectrum auction and compare it with the performance of other spectrum auction mechanisms. Three performance metrics are considered: spectrum utilization, auctioneer's revenue and bidders' satisfaction. These are the most important performance metrics that need to be maximized in a successful spectrum auction. Then, we analyze the security strategy of the proposed secure spectrum auction **MTSSA** that makes it able to avoid possible frauds and bid-rigging. In addition, we compare the computational and communication complexity of **MTSSA** with the spectrum auction mechanism in [81].

8.1.4.1 Performance Analysis

We consider a spectrum auction hosted by the auctioneer (the broker) in a $A * A$ m^2 square geographical region with two cellular networks located within the same region where the auction takes place. Each cellular network belongs to different WSP, i.e. there exists two WSPs $L = 2$ that are interested in participating in the spectrum auction. Each WSP has certain number of BSs, located in macro cells or small cells, that are interested in bidding for the auctioneer's under-utilized frequency bands. The BSs are randomly placed in the auction's geographical area. Suppose that the frequency mutual interference between any two BSs is based on the distance between them. Any two macro cells' BSs located within a distance of $0.4A$ can not be allocated the same frequency bands and these BSs are connected together in the frequency conflict graph. Also, any small cell's BS can not be allocated the same frequency bands of any other BS located within a distance of $0.05A$ from it. In our simulation setup, bids are selected randomly with biding per frequency band is monotonically decreasing, i.e BS's bid for first frequency band is higher than second frequency band and BS's bid for second frequency band is higher than third frequency band and so on, see [**?**, 21, 22].

Based on the frequency assignment policy, three spectrum auction mechanisms are considered in our simulation as described in the following three cases:

- *Case* 1 : Conventional spectrum leasing (**CSL**) case where the government directly leases the under-utilized spectrum to the WSP with highest bid. Once the winning WSP is exclusively assigned all frequency bands, it then allocates these resources internally to its BSs. Other WSPs are deprived from these frequency bands.

- *Case* 2 : **MTSSA** where each WSP directly submits all of its BSs' encrypted bids to the federal gateway. The auctioneer decides the frequency bands allocation to each BS whereas the WSP has no control on the resources allocation process. By using **MTSSA** frequency assignment process, each BS can be allocated any number of frequency bands

between zero and all of the auctioneer's under-utilized frequency bands.

- *Case* 3 : **MTSSA** with fixed limit (**MTSSA-FL**) is a special case of the proposed **MTSSA** where the frequency bands allocation policy is similar to the proposed **MTSSA** but is restricted in the number of frequency bands that each BS can bid for. Each BS bids for a fixed number of frequency bands and it can be allocated any number of frequency bands between zero and that fixed number of frequency bands it submitted the bids for.

We ran Monte Carlo Simulation, for the three cases described above, and the results are averaged over 25 independent runs in which the location and the bidding values of the BSs are generated randomly and the performance metrics are evaluated. We consider the network setup described above with different number of macro cells and small cells' BSs that belong to the two WSPs. First, we consider 8 BSs, i.e. 4 macro cells' BSs and 4 small cells' BSs. Second, we consider 12 BSs, i.e. 6 macro cells' BSs and 6 small cells' BSs. Third, we consider 16 BSs, i.e. 8 macro cells' BSs and 8 small cells' BSs.

We consider three performance metrics to compare between **CSL**, **MTSSA**, and **MTSSA-FL**. These performance metrics are:

- Spectrum Utilization: It is represented by the sum of the frequency bands that are allocated by the auctioneer to the winning BSs.
- Auctioneer's Revenue: It is given by the sum of all BSs' payments, i.e $R = \sum_{n=1}^{n=N} p_n$.
- Bidders' Satisfaction: It is represented by the sum of all winning BSs' utilities divided by the sum of all BSs' evaluation values, i.e. $\sum_{n\in\mathcal{A}} U_n / \sum_{n\in\mathcal{N}} v_n$, where $\mathcal{A}$ is the set of BSs that are allocated frequency bands (i.e. winning BSs).

In Figure 8.6, we compare the performance of the proposed **MTSSA** and its special case **MTSSA-FL** with that of a **CSL** based auction. We plot the spectrum utilization,

auctioneers revenue and BSs satisfaction of the three auction designs with different number of BSs, i.e. 8 BSs, 12 BSs and 16 BSs as mentioned before.

Figure 8.6(a) shows the spectrum utilization versus the number of available under-utilized frequency bands. As the number of frequency bands increases, the spectrum utilization, which is represented by the number of allocated frequency bands, also increases for each of the three auction mechanisms. For certain number of frequency bands, each of the three mechanisms shows higher utilization when the number of BSs (bidders) increases. However, it is not surprising that the performance in terms of utilization for **CSL** is lower than that for the other two mechanisms. This is because in **CSL**, the auctioneer assigns each WSP different frequency bands and the frequency bands assigned to each WSP are then auctioned among BSs that belong to that WSP. In the case of **CSL**, the auctioneer considers one frequency conflict graph for each WSP and frequency reusability is not applicable among BSs that belong to different WSPs, i.e. BSs that belong to different WSPs and are not within the interference range of each other are not allowed to use the same frequency bands. Moreover, the utilization in the cases of **MTSSA** and **MTSSA-FL** is almost the same when the number of available frequency bands is low and is slightly higher for **MTSSA** than that for **MTSSA-FL** when the number of available frequency bands is higher.

Figure 8.6(b) shows that for each of the three mechanisms, the auctioneer's revenue increases when the number of BSs increases. This is expected as the auctioneer's revenue increases with more bidders requesting more resources. However, for certain number of BSs, the auctioneer's revenue for **MTSSA-FL** is higher than that for **MTSSA** and **CSL** and as expected the auctioneer's revenue is the lowest in the case of **CSL**. The bump of **MTSSA** over **CSL** is from the payments received from winning BSs that belong to different WSPs, and not located within the interference range of each other, but are allocated similar frequency bands. Figure 8.6(b) shows that the auctioneers revenue for **MTSSA-FL** is higher than that for **MTSSA**. This is due to the frequency reuse that is better achieved in the

case of **MTSSA-FL** than in **MTSSA** which makes frequency utilization higher in the case of **MTSSA-FL** compared to **MTSSA**. Also, the difference in the auctioneer's revenue between **MTSSA** and **MTSSA-FL** is higher when the number of BSs is higher as shown in Figure 8.6(b). This is because the frequency reuse is higher when the number of BSs increases.

We show in Figure 8.6(c) that as the number of frequency bands increases, the bidders' satisfaction also increases until it saturates when each bidder is allocated the number of frequency bands he bids for. On the other hand, the bidders satisfaction for **CSL** is higher than that for **MTSSA** and **MTSSA-FL** due to the less number of BSs competing for resources as all the frequency bands are allocated to one WSP. Therefore, the bidders who belong to the winning WSP get their requested frequency bands while paying less.

The comparison between the three mechanisms in Figure 8.6 shows **MTSSA**'s high performance and superiority over conventional spectrum leasing mechanism as it considers spectrum reusability and in the same time it guarantees a secure spectrum auction.

In addition, we compare the performance of the proposed spectrum auction **MTSSA** and its ability to maximize the auctioneer's revenue with the spectrum auction mechanism in [73] (**SPRING**). We ran simulations for different number of bidders where each bidder is requesting one spectrum at a time by submitting its true evaluation bidding value from the set $\{1, 2, 3, 4, 5\}$. Simulation results showed that **MTSSA** and **SPRING** achieve almost similar spectrum utilization and bidders satisfaction when each bidder requests one spectrum band. However, the auctioneer's revenue is much higher in the case of **MTSSA**. This is due to **MTSSA** payment method which is based on VCG auction as well as the spectrum allocation procedure used in **MTSSA** whereas **SPRING** considers dividing all bidders in groups such that each group contains bidders that do not have mutual frequency interference among them. When running **SPRING**, each group of bidders submits a bidding value for the whole group that is equivalent to the minimum bidding value among bidders' bids in that

group multiplied by the number of bidders in that group. The auctioneer selects the group with the maximum bidding value to be the winner and changes each bidder within that group a price that is equivalent to the minimum bid value in that group. Based on the payment method adopted in **SPRING**, it is obvious that in the case of **MTSSA** spectrum auction, the auctioneer will be gaining more revenue compared to the case when using **SPRING** spectrum auction. Figure 8.7 shows the auctioneer's revenue in the case of **MTSSA** and in the case of **SPRING** for different number of bidders. In addition, **MTSSA** prevents possible insincere behavior of the auctioneer or bidders by limiting the auctioneer's capability and making him only able to exploit the winners and their payments without knowing the actual bidding values whereas in the case of **SPRING**, the auctioneer is able to determine the value of the minimum bid value which makes fraud of insincere auctioneer possible. Furthermore, **MTSSA** considers the possibility of bidders colluding with each other to manipulate the auction (bid-rigging among bidders) and is designed in away to avoid these situations whereas bid-rigging among greedy bidders is possible in the case of **SPRING**.

8.1.4.2 MTSSA Security Analysis

As discussed before, our proposed **MTSSA** leverages Paillier cryptosystem in order to ensure that the BSs' bidding values are kept unknown to the auctioneer while the auctioneer is still able to find the winners and charges them their corresponding payments. This is possible because of the indistinguishability property of Paillier cryptosystem, i.e. it is not possible to know the value of z without decrypting each element in $c(z)$, and the self blinding property that makes it impossible to find a mapping function from $c(t)$ to $c^{'}(z+t)$ [81,97,98]. In order to prevent an insincere auctioneer from performing any frauds, we consider a secure gateway that is operated by federal government. Its main function is to send out the number of bidders J to each of the participating BSs in the auction. The number of bidders J can either be the actual number of bidders or a random number that is decided by the federal

gateway in order to avoid possible bid-rigging between bidders. The federal gateway also randomizes the encrypted bids by the random constant t. So auctioneer can determine the winning allocation and assign secondary price without any knowledge of the original bidding values of BSs. This way **MTSSA** can avoid bid-rigging between an insincere auctioneer and a greedy bidder. This can be guaranteed because even if certain bidder colludes with auctioneer, he can not find out about the bidding values as federal gateway randomized it. Therefore, all BSs that belong to different WSPs are treated equally by the proposed **MTSSA** and their bidding values are kept secret from the auctioneer who is only able to determine the winners and their corresponding charged price.

8.1.4.3 MTSSA Complexity Analysis

In this section, we analyze the computational and communication complexity of **MTSSA** and show its efficiency compared to other secure spectrum auction models such as **THEMIS**. We assume a connected random graph with N nodes, so the size of each subnet is in the order of $\mathcal{O}(\log N)$ [106]. Given the number of spectrum bands available for auction in a subnet to be M, then the number of possible allocations is $\binom{\log N+M-1}{M}$. Using Stirling's formula, we have the number of possible allocations in the order of

$$\mathcal{O}\left(\frac{1}{\sqrt{2\pi M}}\left(1+\frac{M}{\log N-1}\right)^{\log N-\frac{1}{2}}\left(1+\frac{\log N-1}{M}\right)^{M}\right).$$

Table 8.2 shows computational complexity of both **MTSSA** and **THEMIS** where s is the number of possible bidding values. In Figure 8.8, we plot the upper bounds of the number of possible allocations for **MTSSA** and **THEMIS** using the simulation setup given in Section 8.1.4. It can be observed that our model computational complexity is more practical and efficient compared to **THEMIS**. For instant, for the allocation of 15 frequency bands in a subnet, **THEMIS** has 10 million possible allocations compared to 100 for **MTSSA**. This explains why authors in [81] have only considered the case when each bidder requests

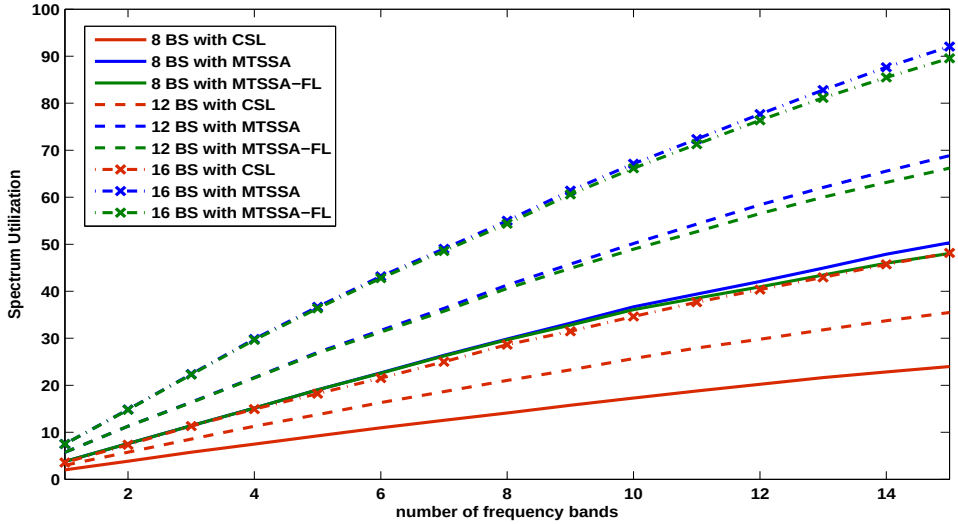
Table 8.2: Computational Complexity Comparison

	Computational Complexity
MTSSA	$\mathcal{O}\left(\frac{N\log N}{\sqrt{2\pi M}}\left(1+\frac{M}{\log N-1}\right)^{\log N-\frac{1}{2}}\left(1+\frac{\log N-1}{M}\right)^{M} s\log N\right)$
THEMIS	$\mathcal{O}\left(N\log N(\log N)^{M} s\log N\right)$

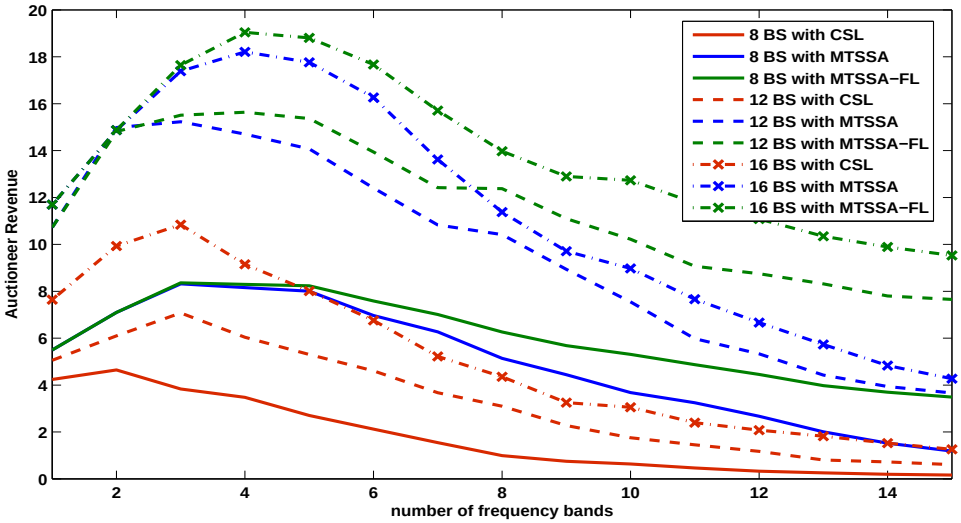
Table 8.3: Communication Complexity Comparison

	Round	**Volume**
MTSSA	$\mathcal{O}(N\log N)$	$\mathcal{O}\left(\frac{N\log N}{\sqrt{2\pi M}}\left(1+\frac{M}{\log N-1}\right)^{\log N-\frac{1}{2}}\left(1+\frac{\log N-1}{M}\right)^{M} s\log N\right)$
THEMIS	$\mathcal{O}(N\log N)$	$\mathcal{O}\left(N\log N(\log N)^{M} s\log N\right)$

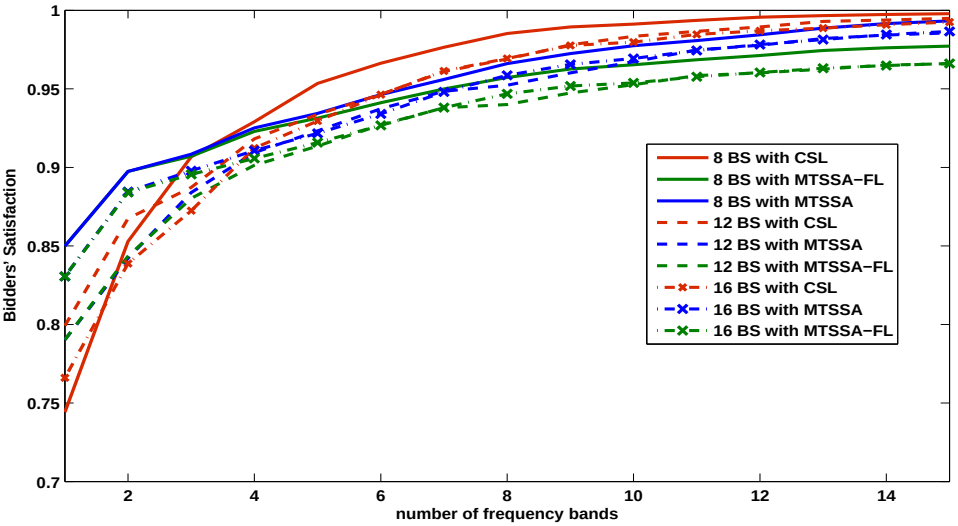
one spectrum band in their simulation results. Similarly, Table 8.3 shows communication complexity of both **MTSSA** and **THEMIS**.



(a) Spectrum Utilization



(b) Auctioneer's Revenue



(c) Bidders' Satisfaction

Figure 8.6: Performance comparison of **MTSSA**, **MTSSA-FL** and **CSL**.

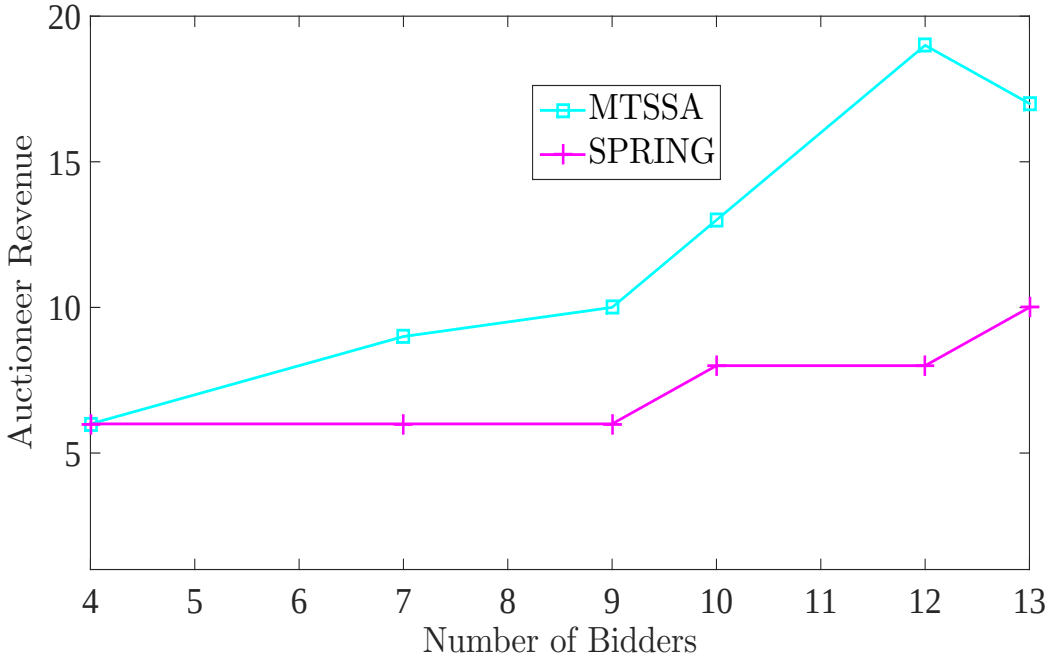


Figure 8.7: Comparison between auctioneer's revenue for **MTSSA** and **SPRING**.

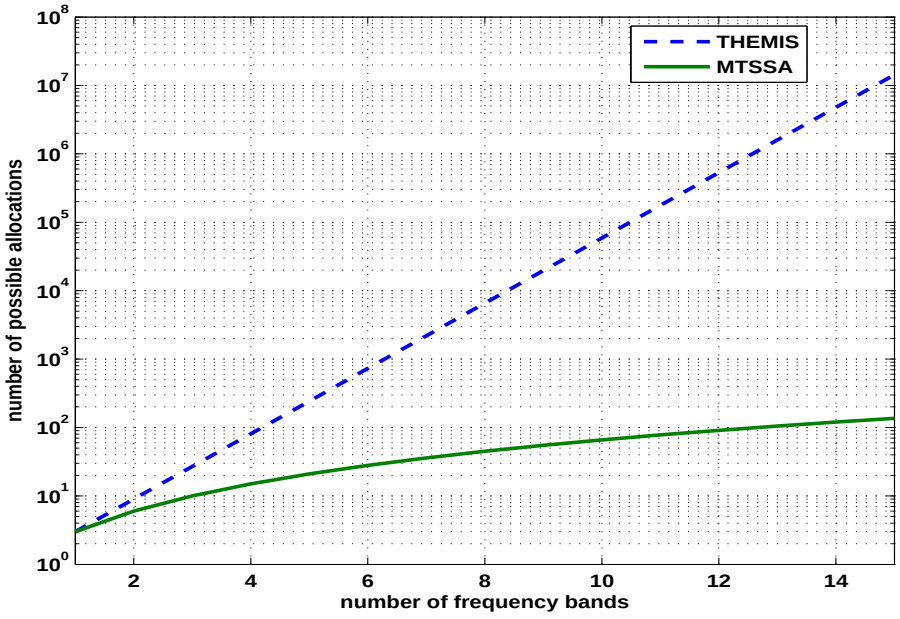


Figure 8.8: Comparison between upper bounds of the number of possible allocations for **MTSSA** and **THEMIS**.

8.2 An Optimal Strategy for Determining True Bidding Values in Secure Spectrum Auctions

Some previous research effort has considered designing secured spectrum auctions. In [78–80], homomorphic encryption is used to provide security in traditional auctions. Our proposed secure spectrum auction **MTSSA** allows spectrum reusability among multiple service providers and provides a truthful secure spectrum auction by assuming that each bidder submits its true bidding value. However, **MTSSA** does not provide a bidding strategy for BSs participating in the spectrum auction. Given that truthfulness is a dominant property in a secure spectrum auction, as it prevents insincere bidders from manipulating the auction, it is important to design a bidding mechanism to be used by bidders in order for them to determine their true bidding values.

In this section, we focus on providing an optimal bidding strategy for BSs participating in a truthful secure spectrum auction. BSs that belong to different WSPs and are interested in additional resources participate in the spectrum auction by submitting theirs bidding prices to the auctioneer. The auctioneer carries out a secure spectrum auction for its available under-utilized spectrum bands. Based on the auction's result, the auctioneer allocates each of the winning bidders the corresponding spectrum bands and charges it for the allocated resources. The proposed bidding strategy is essential in a truthful secure spectrum auction as it guarantees that bidders submit their true bidding values preventing insincere bidders from manipulating the auction. By using the proposed bidding strategy, each BS participating in the spectrum auction calculates its true bidding value for the number of spectrum bands it is bidding for based on its demand for these resources. The BS's decision regarding its demand for additional resources is driven by the traffic of the active UEs, located under its coverage area, that are subscribing for mobile services from that BS. In addition, we provide an efficient resource management solution to enable BSs, participating in the spectrum auction, to

distribute their temporary auctioned resources as well as their permanent resources optimally among UEs based on a carrier aggregation scenario.

Our contributions in this section are summarized as:

- We propose an optimal bidding mechanism for determining true bidding values to be used in secure spectrum auctions by BSs, that belong to different WSPs, participating in a spectrum auction seeking additional resources to improve the QoE for active UEs located under their coverage area. The BS's true bidding values are determined based on the number of spectrum bands it is bidding for, the applications traffic of users under its coverage area and the quantity of bidder's permanent resources.

- We prove that by using the proposed bidding mechanism, when a BS bids for certain number of spectrum bands n, its bidding price of the l^{th} spectrum band ($l \in \{1, 2, .., n-1\}$) is greater than its bidding price of the $(l+1)^{th}$ spectrum band which is expected as the BS's price per spectrum band depends on its demand for that spectrum band.

- We present a resource allocation based on carrier aggregation approach to determine the BS's optimal aggregated rate allocated to each UE, under its coverage area, from both the BS's permanent resources and the BS's winning auctioned spectrum resources.

- We show through simulation results the performance of the proposed optimal bidding strategy when used by BSs participating in a secure spectrum auction and show that the spectrum auction achieves an efficient spectrum utilization, revenue and bidders' satisfaction.

8.2.1 System Model

We consider a spectrum sharing scenario where a federal agency leases its under-utilized spectrum bands on a long term basis to a broker which plays the role of an auctioneer and

auctions its under-utilized spectrum bands to BSs that belong to different WSPs and located in the same geographical area. Let L denotes the number of WSPs with coverage areas within the auction's geographical region. Each WSP provides wireless services to multiple cellular cells through BSs where one BS is located in the middle of each cellular cell. Each BS has permanent resources assigned to it by its WSP and each BS allocates its resources to UEs that belong to its WSP and located within its coverage area. Each BS with a high traffic volume (i.e. during peak hours) and scarce permanent resources can participate in the spectrum auction by submitting its bidding price to the auctioneer. Let $\mathcal{K}^l$ denotes the set of BSs, that belong to WSP l, participating in the spectrum auction and let $\mathcal{K} = \mathcal{K}^1 \cup \mathcal{K}^2 \cup ... \cup \mathcal{K}^L$ denotes the set of the L WSPs' BSs participating in the spectrum auction where $K = |\mathcal{K}|$ denotes the number of BSs in $\mathcal{K}$. Furthermore, let $\mathcal{M}_k$ be the set of UEs subscribing for mobile services from the k^{th} BS in $\mathcal{K}$. The set of auctioneer's spectrum bands is given by $\mathcal{N}$ where $N = |\mathcal{N}|$ denotes the number of the auctioneer's available spectrum bands. Let R_k^p denotes the total achievable rate of the permanent resources of BS k at the time during which the auction takes place.

In this paper, we consider a multi-band spectrum auction, similar to the one presented in [107], where each BS can bid for a single or multiple spectrum bands in the set $\mathcal{N}$ auctioned by the broker (auctioneer). Once the auction takes place, each BS in $\mathcal{K}$ submits its bids to the auctioneer. Let $\boldsymbol{\alpha} = \{\alpha_1, \alpha_2, ...\}$ be the allocation set for the N spectrum bands where $\alpha_j \in \boldsymbol{\alpha}$ is given by $\alpha_j = (\alpha_j(1), ..., \alpha_j(k), ..., \alpha_j(K))$ with $\alpha_j(i)$ represents the number of spectrum bands that the k^{th} BS in $\mathcal{K}$ is bidding for. For example, for $\mathcal{K} = \{1, 2, 3\}$ and $\mathcal{N} = \{a, b\}$ we have $\boldsymbol{\alpha} = \{\alpha_1 = (2, 0, 0), \alpha_2 = (0, 2, 0), \alpha_3 = (0, 0, 2), \alpha_4 = (1, 1, 0), \alpha_5 = (1, 0, 1), \alpha_6 = (0, 1, 1)\}$. Each BS $k \in \mathcal{K}$ creates its bidding vector $\mathbf{b}_k = [b_k(\alpha_1), b_k(\alpha_2), ...]$ and submits an encrypted bidding vector $\hat{\mathbf{b}}_k = [\hat{b}_k(\alpha_1), \hat{b}_k(\alpha_2), ...]$ for the allocations in $\boldsymbol{\alpha}$ where $\hat{b}_k(\alpha_j)$ is the encrypted bidding value for the actual bid $b_k(\alpha_j)$. In the above example, $\mathbf{b}_1 = [6, 0, 0, 3, 3, 0]$ indicates that BS 1 bids 6 for allocation α_1, 0 for allocation α_2, 0 for allocation α_3, 3 for allocation α_4, 3 for allocation α_5 and 0 for allocation α_6. We assume

that the BSs treat different spectrum bands similarly. Let $e_k(\alpha_j)$ represents the k^{th} BS true evaluation value for allocation α_j and let p_k denotes the actual price that the k^{th} BS in $\mathcal{K}$ pays to the auctioneer. The k^{th} BS utility is given by $u_k = e_k(\alpha_j) - p_k$, which is the difference between the k^{th} BS true evaluation value and the actual price the BS pays to the auctioneer for the allocated spectrum bands. The auctioneer's revenue is given by $Rev = \sum_{k=1}^{K} p_k$, which is the sum of all BSs' payments from the spectrum sale.

Once the auctioneer shares with all WSPs' BSs (i.e. BSs located within the auction geographical region) the number of its spectrum bands N, each BS decides whether to participate in the spectrum auction or not based on the current traffic of the UEs located under its coverage area (i.e. UEs in $\mathcal{M}_k$) and its current permanent available resources R_k^p. If the BS's R_k^p value is considered scarce for the current active UEs' demand of resources, the BS decides to participate in the spectrum auction and submit its bidding values for all possible allocations in $\boldsymbol{\alpha}$. We assume that the k^{th} BS's decision on whether to participate in the spectrum auction or not is driven by its system utility; i.e. if the total system Quality of Experience (QoE) when the UEs in $\mathcal{M}_k$ are allocated resources only from the BS's permanent resources is below a minimum expected predefined value then the BS decides to participate in the spectrum auction in order to obtain additional resources and allocates both of its permanent resources as well as its auctioned resources to the UEs in $\mathcal{M}_k$ based on a carrier aggregation scenario. As in [107], all bidders submit their bidding values to the auctioneer through an intermediate secure gateway. In order to consider the spectrum reusability property [60, 65], the secure gateway builds an interference conflict graph which consists of nodes that represent all BSs participating in the auction. Each BS (node) in the interference conflict graph is connected with other nodes located within its coverage area; which indicates that the auctioneer can not allocate the k^{th} BS spectrum bands similar to the ones simultaneously allocated to any other BS connected to BS k in the interference conflict graph. The interference conflict graph can be constructed using a physical or a protocol channel model [101]. Figure 8.9 shows two WSPs located within the auction's geographical

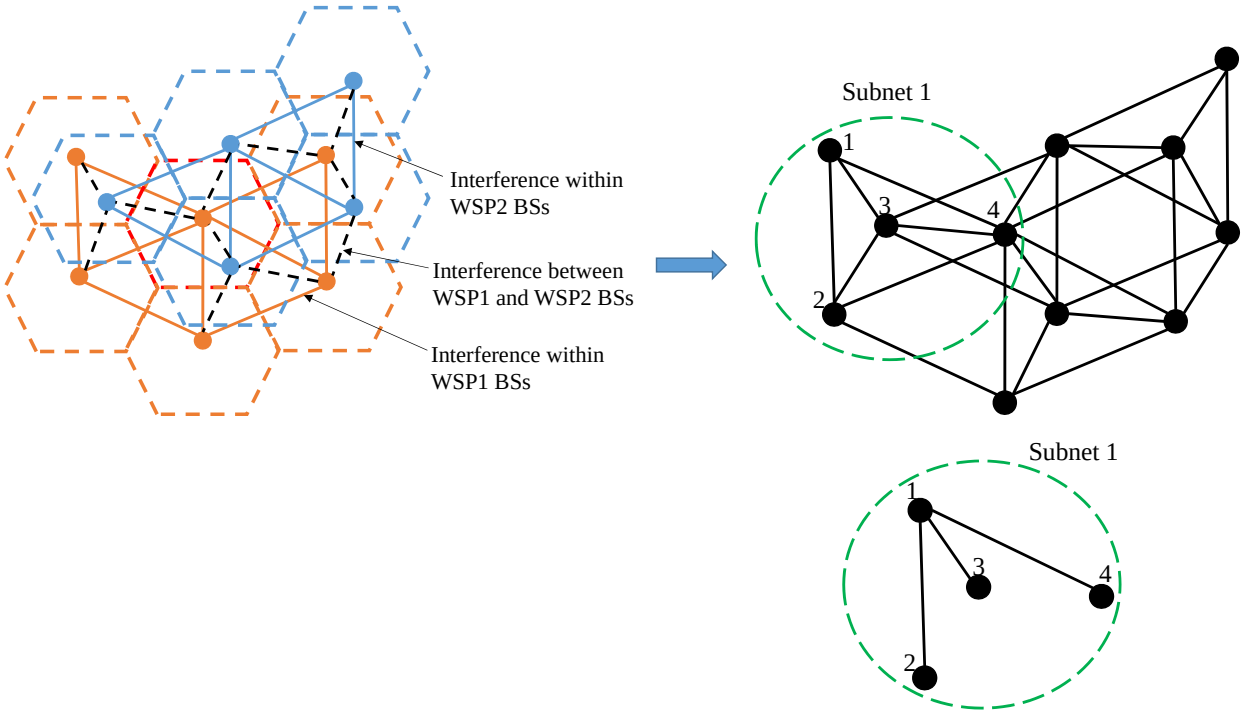


Figure 8.9: Frequency conflict graph for two WSPs's BSs participating in the spectrum auction where nodes represent BSs and edges represent mutual interference between end points (BSs) with an illustration of one subnet; i.e. subnet 1 which consists of BSs 1, 2, 3 and 4 where BS 1 is the root BS.

region with all of their BSs participating in the auction. An interference conflict graph for the two WSPs' BSs is shown in the same figure with an illustration of one subnet. The auctioneer runs the spectrum auction in one subnet after another where a subnet consists of one root BS (i.e. BS1 in Figure 8.9) and all other nodes connected to it in the interference conflict graph that are not previously considered root BSs.

8.2.2 Spectrum Sharing through Secure and Truthful Spectrum Auction

In this section, we present a spectrum sharing system that provides an efficient sharing of spectrum resources with commercial WSPs using a secure truthful spectrum auction. Figure 8.10 shows the proposed spectrum sharing model that is performed in the following steps: 1)

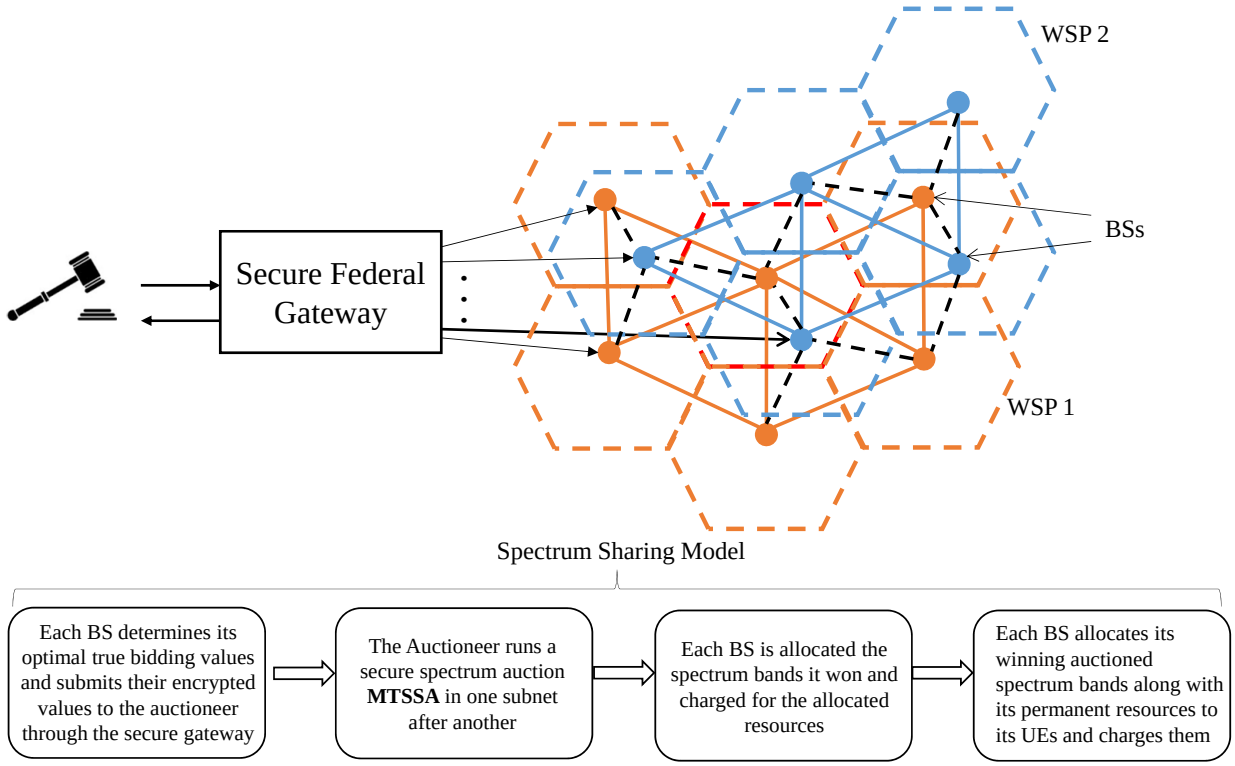


Figure 8.10: Spectrum sharing model through a truthful and secure spectrum auction with BSs that belong to two WSPs participating in the auction.

Each BS that belongs to a WSP within the auction region, and decides to participate in the spectrum auction, determines its optimal true bidding values (by using the proposed optimal bidding mechanism) and submits them to the auctioneer through the secure gateway. 2) The Auctioneer runs a secure spectrum auction **MTSSA** [107] in one subnet after another. 3) Each BS is allocated the spectrum bands it won and charged for the allocated resources. 4) Each BS allocates its winning auctioned spectrum bands (temporary resources) along with its permanent resources to the UEs under its coverage area based on a carrier aggregation scenario and charges the UEs for the allocated resources.

8.2.2.1 An Optimal Mechanism for Determining True Bidding Values

Each BS decides whether its primary resources are sufficient or not based on its available permanent resources and the traffic of the applications running on the UEs subscribing for mobile services from that BS. Each BS in $\mathcal{K}$ located within the auction region informs the auctioneer about its interest in participating in the spectrum auction seeking additional resources. Once the auctioneer runs the spectrum auction, each of the BSs participating in the auction receives the number of the auctioneer's spectrum bands N as well as the number of all bidders K participating in the auction.

Each BS $k \in \mathcal{K}$ performs a resource allocation process using Utility Proportional Fairness (UPF) approach to determine its assigned resources from its permanent resources to each active UE $i \in \mathcal{M}_k$. We express the user satisfaction with its application rates using utility functions. We represent the i^{th} user application utility function $U_i(r_i)$ by sigmoidal-like function or logarithmic function where r_i is the rate of the i^{th} user application. Logarithmic utility functions expressed by equation (2.2) and sigmoidal-like utility functions expressed by equation (2.1) are used to represent delay tolerant and real-time applications, respectively.

We use a utility proportional fairness resource allocation optimization problem to find the k^{th} BS assigned resources to each of the UEs in $\mathcal{M}_k$ and the BS's price per unit bandwidth. The UPF resource allocation optimization problem that is used for assigning the k^{th} BS permanent resources R_k^p is given by:

$$\begin{aligned} \max_{\mathbf{r}_k^p} \quad & \prod_{i=1}^{|\mathcal{M}_k|} U_i(r_{i,k}^p) \\ \text{subject to} \quad & \sum_{i=1}^{|\mathcal{M}_k|} r_{i,k}^p \leq R_k^p \\ & 0 \leq r_{i,k}^p \leq R_k^p, \qquad i = 1, 2, ..., |\mathcal{M}_k|, \end{aligned} \tag{8.15}$$

where $\mathbf{r}_k^p = \{r_{1,k}^p, r_{2,k}^p, ..., r_{|\mathcal{M}_k|,k}^p\}$, $r_{i,k}^p$ is the achievable rate of the i^{th} UE in $\mathcal{M}_k$ from BS

k assignment to that UE from its permanent resources, $|\mathcal{M}_k|$ is the number of active UEs under the coverage area of BS k and R_K^p is the maximum achievable rate of the k^{th} BS from its available permanent resources. The objective function of the above UPF resource allocation is to maximize the entire cell utility when allocating the BS's resources and achieve utility proportional fairness such that non of the UEs is allocated zero resources. Therefore, it guarantees a minimum QoS for each UE. Real time applications are given priority when allocating the BS's resources due to the nature of their utility functions used to represent their applications. The objective function in optimization problem (8.15) is equivalent to $\max_{\mathbf{r}_k^p} \sum_{i=1}^{|\mathcal{M}_k|} \log U_i(r_{i,k}^p)$. Optimization problem (8.15) is a convex optimization problem and there exists a unique tractable global optimal solution as shown in [21,24]. From optimization problem (8.15), we have the Lagrangian:

$$\begin{aligned} L_p(\mathbf{r}_k^p, p_k^p) = &\sum_{i=1}^{|\mathcal{M}_k|} \log U_i(r_{i,k}^p) \\ &- p_k^p \left(\sum_{i=1}^{|\mathcal{M}_k|} r_{i,k}^p + z_k^p - R_k^p \right), \end{aligned} \tag{8.16}$$

where $z_k^p \geq 0$ is the slack variable and p_k^p is the Lagrange multiplier which represents the shadow price (price per unit bandwidth) of BS k for its assigned resources from its permanent available resources. The solution of optimization problem (8.15) is given by the values $r_{i,k}^p$ which are the solution of $\frac{\partial \log U_i(r_{i,k}^p)}{\partial r_{i,k}^p} = p_k^p$ and are the intersection of the time varying shadow price, horizontal line $y = p_k^p$, with $y = \frac{\partial \log U_i(r_{i,k}^p)}{\partial r_{i,k}^p}$ geometrically. Once BS k performs the UPF resource assignments for its permanent resources, each UE in $\mathcal{M}_k$ will be assigned $r_{i,k}^{p,\text{opt}} = r_{i,k}^p$ rate but not yet allocated its assigned resources. As mentioned above, we assume that each BS $k \in \mathcal{K}$ is demanding additional resources and therefore will be participating in the spectrum auction; i.e. the k^{th} BS's cell utility obtained from the assigned resources $r_{i,k}^{p,\text{opt}}$ is less than the minimum required utility.

Each BS k uses the information it receives from the auctioneer regarding the number of all

bidders K and the number of the auctioneer's spectrum bands N and uses them to create the allocation set $\boldsymbol{\alpha}$. For the number of spectrum bands $n = \alpha_j(k)$ (i.e. $n \leq N$) that corresponds to BS k in the allocation α_j, BS k determines the maximum achievable rate $R^t_{k,n}$ if BS k allocates these temporary resources (i.e. auctioned resources) to its UEs. We assume that each BS k calculates its bidding price for certain allocation α_j based on the number of spectrum bands that corresponds to that BS (i.e. $\alpha_j(k)$) regardless of the number of spectrum bands that correspond to other BSs in the same allocation α_j. Therefore, BS k bidding price will be the same for all allocations $\alpha_j \in \boldsymbol{\alpha}$ that have similar $\alpha_j(k)$ values. In order for BS k to determine its true bidding price that needs to be submitted to the auctioneer, it uses a UPF resource allocation with carrier aggregation to find its aggregated assigned resources from both its permanent resources R^p_k and its auctioned resources $R^t_{k,n}$, assuming that this BS will be winning its corresponding number of spectrum bands in that allocation (i.e. $\alpha_j(k)$). Additionally, each BS also determines the price per unit bandwidth for the temporary auctioned resources based on resource allocation with carrier aggregation optimization and calculates its bidding value for each allocation that has $\alpha_j(k) = n$ based on that price. The k^{th} BS's UPF resource allocation optimization problem that is used for assigning the temporary auctioned resources $R^p_{k,n}$ that corresponds to n spectrum bands ($n \leq N$) is given by:

$$\begin{aligned} \max_{\mathbf{r}^t_{k,n}} \quad & \prod_{i=1}^{|\mathcal{M}_k|} U_i(r^t_{i,k,n} + r^{p,\text{opt}}_{i,k}) \\ \text{subject to} \quad & \sum_{i=1}^{|\mathcal{M}_k|} r^t_{i,k,n} \leq R^t_{k,n} \\ & 0 \leq r^t_{i,k,n} \leq R^t_{k,n}, \qquad i = 1, 2, ..., |\mathcal{M}_k|, \end{aligned} \tag{8.17}$$

where $\mathbf{r}^t_{k,n} = \{r^t_{1,k,n}, r^t_{2,k,n}, ..., r^t_{|\mathcal{M}_k|,k,n}\}$, $r^t_{i,k,n}$ is the achievable rate of the i^{th} UE in $\mathcal{M}_k$ from BS k assignment to that UE from its temporary auctioned resources, $|\mathcal{M}_k|$ is the number of UEs that will be assigned temporary auctioned resources by BS k and $R^t_{k,n}$ is the maximum

achievable rate of the k^{th} BS from the auction resources. The objective function of the above optimization problem is to maximize the entire cell utility when allocating the BS's temporary resources. The RA optimization problem (8.17) is based on carrier aggregation. It seeks to maximize the multiplication of the utilities of the rates allocated to UEs in $\mathcal{M}_k$ by BS's permanent and auctioned resources. Utility proportional fairness is used to guarantee that non of the UEs will be allocated zero resources and real time applications are given priority when allocating the BS's auctioned resources. The objective function in optimization problem (8.17) is equivalent to $\max_{\mathbf{r}^t_{k,n}} \sum_{i=1}^{|\mathcal{M}_k|} \log U_i(r^t_{i,k,n} + r^{p,\text{opt}}_{i,k})$. Optimization problem (8.17) is a convex optimization problem and there exists a unique tractable global optimal solution [21, 24]. From optimization problem (8.17), we have the Lagrangian:

$$\begin{aligned} L_t(\mathbf{r}^t_{k,n}, p^t_{k,n}) = &\sum_{i=1}^{|\mathcal{M}_k|} \log U_i(r^t_{i,k,n} + r^{p,\text{opt}}_{i,k}) \\ &- p^t_{k,n}\left(\sum_{i=1}^{|\mathcal{M}_k|} r^t_{i,k,n} + z^t_{k,n} - R^t_{k,n}\right), \end{aligned} \tag{8.18}$$

where $z^t_{k,n} \geq 0$ is the slack variable and $p^t_{k,n}$ is the Lagrange multiplier which represents the price per unit bandwidth of BS k for its assigned resources from its temporary resources $R^t_{k,n}$. The solution of optimization problem (8.17) is given by the values $r^t_{i,k,n}$ that solve equation $\frac{\partial \log U_i(r^t_{i,k,n}+r^{p,\text{opt}}_{i,k})}{\partial r^t_{i,k,n}} = p^t_{k,n}$ and are the intersection of the time varying shadow price, horizontal line $y = p^t_{k,n}$, with the curve $y = \frac{\partial \log U_i(r^t_{i,k,n}+r^{p,\text{opt}}_{i,k})}{\partial r^t_{i,k,n}}$ geometrically.

Once BS k is done performing the resource assignment with carrier aggregation process for each possible number of auctioned spectrum bands n based on optimization problem (8.17), BS k calculates each bidding value $b_k(\alpha_j)$ as the following:

$$\begin{aligned} b_k(\alpha_j) &= \sum_{l=1}^{n} D * p^t_{k,l} \\ \text{where} \quad & n = \alpha_j(k), \end{aligned} \tag{8.19}$$

where D is the number of unit bandwidths in one spectrum band and $p^t_{k,l}$ is the price per unit bandwidth of the l^{th} spectrum band; $l = \{1, 2, ..., n\}$. After calculating all bidding values $b_k(\alpha_j)$, BS k creates a bidding vector $\mathbf{b}_k = [b_k(\alpha_1), b_k(\alpha_2), ...]$ for the allocations in $\boldsymbol{\alpha}$ and submits its corresponding encrypted bidding vector $\hat{\mathbf{b}}_k$ to the secure gateway. Algorithm 29 summarizes the k^{th} BS procedure for determining its optimal bidding values for the allocations in $\boldsymbol{\alpha}$.

Once the secure gateway receives the encrypted bidding values from all bidders, it starts performing spectrum bands allocation as will be discussed next in 8.2.2.2. The auctioneer then decides the winning BSs, allocates them the corresponding spectrum bands and charges them for the allocated resources. Based on the number of spectrum bands that each BS k won, it allocates an aggregated rate $r^{\text{agg}}_{i,k} = r^{t,\text{opt}}_{i,k,n} + r^{p,\text{opt}}_{i,k}$ to each UE $i \in \mathcal{M}_k$; i.e. n is the number of winning spectrum bands for BS k from the spectrum auction.

Lemma 8.2.1. *The price per unit bandwidth of the assigned temporary resources by BS k that is represented by the shadow price $p^t_{k,n}$ in equation (8.18) (i.e. the Lagrangian of optimization problem (8.17)) is a strictly decreasing function with respect to $R^t_{k,n}$ which represents the maximum achievable rate of the k^{th} BS.*

Proof. Let $\sum_{i=1}^{|\mathcal{M}_k|} \log U_i(r^t_{i,k,n} + r^{p,\text{opt}}_{i,k})$, which is equivalent to the objective function in (8.17), be denoted by $f(\mathbf{r})$ where $\mathbf{r} = (r^t_{1,k,n}, r^t_{2,k,n}, ..., r^t_{|\mathcal{M}_k|,k,n})$ are the variables. Consider the problem of finding the maximum of $f(\mathbf{r})$ subject to the constraint $\sum_{i=1}^{|\mathcal{M}_k|} r^t_{i,k,n} + z^t_{k,n} = R^t_{k,n}$ with $\sum_{i=1}^{|\mathcal{M}_k|} r^t_{i,k,n} + z^t_{k,n}$ represented by $g(\mathbf{r}^t_{k,n})$. Let R and p be the short terms used for $R^t_{k,n}$ and $p^t_{k,n}$, respectively. For each choice of the parameter R, let $\mathbf{r}^*(R)$ of $\mathbf{r}$ be the values that maximize f; i.e. $\mathbf{r}^* = (r^*_1, r^*_2, ..., r^*_{|\mathcal{M}_k|})$ and r^*_i is the solution of $p = S_i(r_i) = \frac{d \log U_i(r_i)}{dr_i}$. For logarithmic or sigmoidal-like utility functions $U_i(r_i)$ (recall utility function properties in Section 8.2.2.1), $r_i = S_i^{-1}(p)$ is strictly decreasing function and each solution r^*_i is equivalent to $S_i^{-1}(p^*)$. Then $f(\mathbf{r}^*(R))$ is the maximum of f for fixed value of the inputs R. The

Algorithm 29 The k^{th} BS Optimal Bidding Algorithm

Receive application utility parameters k_i, a_i, b_i and $r_i^{\max}$ from all UEs in $\mathcal{M}_k$.
Receive the number of auctioneer's spectrum bands N and the number of bidders K from the auctioneer.
Find the BS's assigned rates from its permanent resources R_k^p by solving $\mathbf{r}_k^p = \arg\max_{\mathbf{r}_k^p} \sum_{i=1}^{|\mathcal{M}_k|} \log U_i(r_{i,k}^p) - p_k^p(\sum_{i=1}^{|\mathcal{M}_k|}(r_{i,k}^p) - R_k^p)$.
Let $r_{i,k}^{p,\text{opt}} = r_{i,k}^p \ \forall i \in \mathcal{M}_k$.
for $n = 1 \rightarrow N$ **do**
  Associate the number of spectrum bands n with a corresponding achievable rate $R_{k,n}^t$ {i.e. for each number of spectrum bands $n \in \mathcal{N}$ that BS k is bidding for.}
  Find the BS's assigned rates from the temporary resources $R_{k,n}^t$ by solving $\mathbf{r}_{k,n}^t = \arg\max_{\mathbf{r}_{k,n}^t} \sum_{i=1}^{|\mathcal{M}_k|} \log U_i(r_{i,k,n}^t + r_{i,k}^{p,\text{opt}}) - p_{k,n}^t(\sum_{i=1}^{|\mathcal{M}_k|}(r_{i,k,n}^t) - R_{k,n}^t)$.
  Let $r_{i,k,n}^{t,\text{opt}} = r_{i,k,n}^t \ \forall i \in \mathcal{M}_k$.
  for $j = 1 \rightarrow |\boldsymbol{\alpha}|$ **do**
    if $\alpha_j(k) = n$ **then**
      Calculate $b_k(\alpha_j) = \sum_{l=1}^{n} D * p_{k,l}^t$.
    end if
  end for
end for
Create a bidding vector $\mathbf{b}_k = [b_k(\alpha_1), b_k(\alpha_2), ...]$.
Encrypt each bidding value in $\mathbf{b}_k$ and create an encrypted bidding vector $\hat{\mathbf{b}}_k$.
Submit the encrypted bidding vector $\hat{\mathbf{b}}_k$ to the secure gateway.

derivative

$$\frac{d}{dR} f(\mathbf{r}^*(R))$$

represents the rate of change in the optimal output from the change of parameter R.

Corresponding to $\mathbf{r}^*(R)$ there is a value $p = p^*(R)$ such that they are a solution to the Lagrange multiplier problem (8.18), i.e.,

$$\bigtriangledown f(\mathbf{r}^*(R)) = p^*(R) \bigtriangledown g(\mathbf{r}^*(R))$$

$$R = g(\mathbf{r}^*(R))$$

where $r_i^* = R - \sum_{j,j\neq i}(r_j^*) = R - \sum_{j,j\neq i} S_j^{-1}(p^*)$. It follows that the following equation holds true

$$\begin{aligned} p^*(R) &= \frac{d}{dR} f(\mathbf{r}^*(R)) \\ &= \sum_{i=1}^{|\mathcal{M}_k|} \left(\frac{\partial}{\partial r_i^*} \Big(f(\mathbf{r}^*(R)) \Big) \frac{dr_i^*}{dR} \right) \\ &= \sum_{i=1}^{|\mathcal{M}_k|} \frac{U_i'(r_{i,k}^{p,\text{opt}} + r_i^*)}{U_i(r_{i,k}^{p,\text{opt}} + r_i^*)}, \end{aligned} \tag{8.20}$$

where the derivative in (8.20) is obtained by applying chain rule and $\frac{dr_i^*}{dR} = 1$. Therefore, the Lagrange multiplier is equivalent to the rate of the change in the optimal output $f(\mathbf{r}^*)$ resulting from the change of the parameter R. Next, we show that $p^*(R)$ is a decreasing function with R. For $p^*(R)$, we have the first derivative as

$$\begin{aligned} \frac{d}{dR} p^*(R) &= \sum_{i=1}^{|\mathcal{M}_k|} \left(\frac{\partial}{\partial r_i} \Big(\sum_{i=1}^{|\mathcal{M}_k|} \frac{U_i'(r_{i,k}^{p,\text{opt}} + r_i^*)}{U_i(r_{i,k}^{p,\text{opt}} + r_i^*)} \Big) \frac{dr_i^*}{dR} \right) \\ &= \sum_{i=1}^{|\mathcal{M}_k|} \left(\frac{U_i''(r_{i,k}^{p,\text{opt}} + r_i^*) U_i(r_{i,k}^{p,\text{opt}} + r_i^*)}{U_i^2(r_{i,k}^{p,\text{opt}} + r_i^*)} \right. \\ &\quad \left. - \frac{U_i'^2(r_{i,k}^{p,\text{opt}} + r_i^*)}{U_i^2(r_{i,k}^{p,\text{opt}} + r_i^*)} \right). \end{aligned} \tag{8.21}$$

where the derivative in (8.21) is obtained by applying chain rule, $U_i''(r_{i,k}^{p,\text{opt}} + r_i^*) = \frac{d^2}{dr_i{}^2} \log U_i(r_{i,k}^{p,\text{opt}} + r_i^*)$. In the case of logarithmic utility function $U_i(r_{i,k}^{p,\text{opt}} + r_i^*)$, recall the utility function properties in Section 8.2.2.1, the utility function of the application rate is positive, increasing and twice differentiable with respect to the application rate. It follows that $U_i'(r_{i,k}^{p,\text{opt}} + r_i^*) = \frac{dU_i(r_{i,k}^{p,\text{opt}} + r_i^*)}{dr_i} > 0$ and $U_i''(r_{i,k}^{p,\text{opt}} + r_i^*) = \frac{d^2 U_i(r_{i,k}^{p,\text{opt}} + r_i^*)}{dr_i} < 0$, i.e. since $r_{i,k}^{p,\text{opt}} + r_i^*$ is greater than zero. Then the function $\log U_i(r_{i,k}^{p,\text{opt}} + r_i^*)$ has $\frac{d \log(U_i(r_{i,k}^{p,\text{opt}} + r_i^*)}{dr_i} = \frac{U_i'(r_{i,k}^{p,\text{opt}} + r_i^*)}{U_i(r_{i,k}^{p,\text{opt}} + r_i^*)} > 0$ and $\frac{d^2 \log(U_i(r_{i,k}^{p,\text{opt}} + r_i^*)}{dr_i{}^2} = \frac{U_i''(r_{i,k}^{p,\text{opt}} + r_i^*) U_i(r_{i,k}^{p,\text{opt}} + r_i^*) - U_i'^2(r_{i,k}^{p,\text{opt}} + r_i^*)}{U_i^2(r_{i,k}^{p,\text{opt}} + r_i^*)} < 0$. On the other hand, in the

case of sigmoidal-like utility function, the normalized sigmoidal-like function is given by $U_i(r_{i,k}^{p,\text{opt}} + r_i^*) = c_i\left(\frac{1}{1+e^{-a_i(r_{i,k}^{p,\text{opt}}+r_i^*-b_i)}} - d_i\right)$. For $R > 0$, we have

$$0 < c_i\left(\frac{1}{1+e^{-a_i(r_{i,k}^{p,\text{opt}}+r_i^*-b_i)}} - d_i\right) < 1$$

then,

$$0 < 1 - d_i(1+e^{-a_i(r_{i,k}^{p,\text{opt}}+r_i^*-b_i)}) < \frac{1}{1+c_i d_i}$$

It follows that for $R > 0$, we have

$$\frac{d}{dr_i}\log U_i(r_{i,k}^{p,\text{opt}} + r_i^*) = \frac{a_i d_i e^{-a_i(r_{i,k}^{p,\text{opt}}+r_i^*-b_i)}}{1-d_i(1+e^{-a_i(r_{i,k}^{p,\text{opt}}+r_i^*-b_i)})} + \frac{a_i e^{-a_i(r_{i,k}^{p,\text{opt}}+r_i^*-b_i)}}{(1+e^{-a_i(r_{i,k}^{p,\text{opt}}+r_i^*-b_i)})} > 0$$

$$\frac{d^2}{dr_i{}^2}\log U_i(r_{i,k}^{p,\text{opt}} + r_i^*) = \frac{-a_i^2 d_i e^{-a_i(r_{i,k}^{p,\text{opt}}+r_i^*-b_i)}}{c_i\left(1-d_i(1+e^{-a(r_{i,k}^{p,\text{opt}}+r_i^*-b_i)})\right)^2} + \frac{-a_i^2 e^{-a_i(r_{i,k}^{p,\text{opt}}+r_i^*-b_i)}}{(1+e^{-a_i(r_{i,k}^{p,\text{opt}}+r_i^*-b_i)})^2} < 0.$$

So, we have $\frac{d}{dR}p^*(R) = \frac{d^2}{dR^2}f(\mathbf{r}^*(R)) < 0$ where $f(\mathbf{r}^*(R)) = \sum_{i=1}^{|\mathcal{M}_k|}\log U_i(r_{i,k}^{p,\text{opt}} + r_i^*)$ and each U_i is a logarithmic or sigmoidal utility function. Therefore, $p^*(R)$ is strictly decreasing function with R. □

Theorem 8.2.2. *By using the proposed bidding mechanism to determine the true bidding price of BS k for any allocation $\alpha_j \in \boldsymbol{\alpha}$, for the bidding price $b_k(\alpha_j)$ of the $n = \alpha_j(k)$*

spectrum bands where that price is determined by the summation of the price of each of the n spectrum bands (equation (8.19)), the price of the l^{th} spectrum band ($l \in \{1, 2, ..., n-1\}$) is greater than the price of the $(l+1)^{th}$ spectrum band; i.e. $D * p^t_{k,1} > D * p^t_{k,2} > ... > D * p^t_{k,n}$.

Proof. It follows from Lemma 8.2.1 that the price per unit bandwidth $p^t_{k,l}$ is a strictly decreasing function with $R^t_{k,l}$, where $l \in \{1, 2, ..., n\}$ and $R^t_{k,1} > R^t_{k,2} > ... > R^t_{k,n}$. Therefore, for certain number of spectrum bands n that BS k is bidding for, the price of the l^{th} spectrum band which is determined by $D * p^t_{k,l}$ is lower for larger l; i.e. $D * p^t_{k,1} > D * p^t_{k,2} > ... > D * p^t_{k,n}$. □

8.2.2.2 Spectrum Bands Allocation

The spectrum bands allocation procedure is similar to the one introduced in [107]. The auctioneer generates a private and public keys of Paillier cryptosystem and publishes his public key and element x (i.e. $x \neq 0$) over the pilot channel. Each bidder determines its true bidding values as shown in 8.2.2.1 and sends their encrypted values to the secure gateway which creates an interference conflict graph for the bidders. The auctioneer selects a random BS to be the root BS and executes the spectrum auction in its subnet. The auctioneer then carries out the spectrum auction in the next subnet while excluding BSs that are previously considered root BSs along with the spectrum bands allocated to these BSs. The auctioneer continues to carry out the spectrum auction in one subnet after another until all BSs in $\mathcal{K}$ have participated in the auction. The secure gateway maintains a conflict table for each bidder as in [67]. After the execution of the spectrum auction in any subnet, the secure gateway updates the conflict tables of all bidders in that subnet.

In order to enable an efficient secure spectrum bands allocation, it is important to prevent possible frauds of an insincere auctioneer and bid-rigging between greedy bidders and an insincere auctioneer. This is possible if the auction mechanism allows the auctioneer to determine the winning BSs and their payments without knowing their actual bidding values.

In order to achieve this, we use a spectrum auction mechanism that leverages homomorphic encryption through Pailliar cryptosystem as in [107].

On the other hand, we adopt a payment method similar to the one used in VCG auction with Clarke pivot payments [104] as it satisfies the following desired economic properties: incentive compatibility, individual rationality and no positive transfers as proved in [107]. By using this payment method, each winning bidder (i.e. bidder k is a winning bidder) is charged a price p_k that is equivalent to the difference between the social welfare with and without that bidder's participation. The charging price p_k is given by

$$p_k = \sum_{i \neq k} \hat{b}_i(\alpha^{*-k}) - \sum_{i \neq k} \hat{b}_i(\alpha^*), \tag{8.22}$$

where allocation $\alpha^* = \arg\max_{\alpha \in \boldsymbol{\alpha}} \sum_i \hat{b}_i(\alpha)$ and allocation $\alpha^{*-k} = \arg\max_{\alpha \in \boldsymbol{\alpha}} \sum_{i \neq k} \hat{b}_i(\alpha)$.

The spectrum bands allocation procedure is presented as follows:

1. The secure gateway receives a sealed bidding vector $\hat{\mathbf{b}}_k$ from each BS k for its true bidding values of the allocations in $\boldsymbol{\alpha}$. We assume that the secure gateway is aware of which WSP each bidder belongs to. The secure gateway creates a conflict table, for each bidder, which contains all BSs within the interference range of that bidder with their allocated spectrum bands and charging price. The procedure of encrypting the bidding values using Paillier encryption is similar to the one presented in [107]. The secure gateway randomizes the bidding values and sends them to the auctioneer. For any bidder k, neither the auctioneer nor the other bidders know the actual bidding values of BS k.

2. The auctioneer selects a random bidder k as a root BS and considers its corresponding subnet. The auctioneer does not have an optimal choice for which subnet it starts running the auction from in order to maximize its profit.

3. The auctioneer runs a secure spectrum auction in the subnet of the current root BS k

as presented in [107]. Once the auction results are revealed, the auctioneer determines the winning bidders and their charging prices.

4. Each winning BS is allocated its corresponding spectrum bands and is charged for the allocated resources. The secure gateway updates the conflict table of each winning BS with the allocated spectrum bands and their charging price. If all BSs in $\mathcal{K}$ have already been considered root BSs, the spectrum auction is considered complete and there is no need to proceed to the next step.

5. The auctioneer selects a new root BS k, that has not been considered as a root BS before, based on a random selection and considers its corresponding subnet which consists of all BSs within the interference range of the root BS k that have not been previously considered as root BSs. The process is then repeated starting from Step **2**.

The spectrum bands allocation process is summarized in Algorithm 30

8.2.3 Simulation Results

In this section, we evaluate the performance of the proposed bidding mechanism for determining true bidding values that are used in a truthful and secure spectrum auction. We evaluate three performance metrics that need to be maximized in a successful spectrum auction. The considered performance metrics are: spectrum utilization, auctioneer's revenue and bidders' satisfaction.

We consider two cellular networks that belong to different WSPs located within the auctioneer's geographical region where a secure spectrum auction is hosted in $X * X$ m^2 area by a broker who plays the role of an auctioneer. Each WSP has a number of BSs that decide to participate in the spectrum auction to obtain additional resources. We assume that each BS is located in the middle of the cellular cell such that UEs located within the BS's coverage area and belong to that BS's WSP are only allocated resources form that BS. Each active

Algorithm 30 Spectrum Bands Allocation Algorithm

Auctioneer creates a set of all bidders $\mathcal{K} = \mathcal{K}^1 \cup \mathcal{K}^2 ... \cup \mathcal{K}^L$
The secure gateway creates the interference conflict graph for BSs in $\mathcal{K}$ and a conflict table for each bidder
Auctioneer generates his private and public keys of Paillier cryptosystem and sends his public key along with element x to all BSs in $\mathcal{K}$
$\mathcal{K}_0 = \mathcal{K}$
Initialize $\beta = \phi$
while $\mathcal{K} \,! = \phi$ **do**
  $k = random(\mathcal{K})$ {Auctioneer selects a random BS}
  $\mathcal{K}_k = include\_conflict(k)$ {Auctioneer creates a new subnet for root BS k by adding BSs that form its subnet from conflict table of that BS}
  $\beta_k = \mathcal{K}_k \cap \beta$ {β_k is set of previous root BSs in the subnet of root BS k}
  $\mathcal{K}_k = \mathcal{K}_k \setminus \beta_k$ {Auctioneer removes from $\mathcal{K}_k$ previous root BSs}
  $\mathcal{N}_k = \mathcal{N} \setminus \mathcal{N}_{\beta_k}$ {Auctioneer removes from $\mathcal{N}$ spectrum bands allocated to BSs in β_k}
  $\boldsymbol{\alpha}_k = alloc\_vect(\mathcal{K}_k, \mathcal{N}_k)$ {Auctioneer forms allocation vector $\boldsymbol{\alpha}_k$ and sends to $\mathcal{K}_k$}
  BSs $\in \mathcal{K}_k$ send encrypted bids to the secure gateway
  The secure gateway randomizes bids and sends them to auctioneer
  Auctioneer selects the highest allocation $\alpha^\star$
  Auctioneer charges price p_k to BS k
  $\mathcal{K} = \mathcal{K} \setminus \{k\}$
  $\beta = \beta \cap \{k\}$
  The secure gateway updates the conflict table for the winning BSs
end while

UE is running either real-time or delay tolerant application. We assume that each BS k has a permanent resources R_k^p that does not change while the auctioneer is running the spectrum auction. The BSs of the two WSPs are randomly placed in the auction's geographical area. We assume that any two BSs located within the interference range of each other can not be allocated the same spectrum bands by the auctioneer. We assume that the interference range is determined based on the distance. The interference range for each BS used in the simulations is equivalent to $0.6X$ m. The auctioneer announces to all bidders the number of its spectrum bands N. We assume that one auctioned spectrum band consists of $D = 10$ units of bandwidth and that one spectrum band corresponds to an achievable rate that is equivalent to 10; i.e. $R_{k,1}^t = 10$ corresponds to $n = 1$ auctioned spectrum band, $R_{k,2}^t = 20$ corresponds to $n = 2$ auctioned spectrum bands and so on. Each application running on

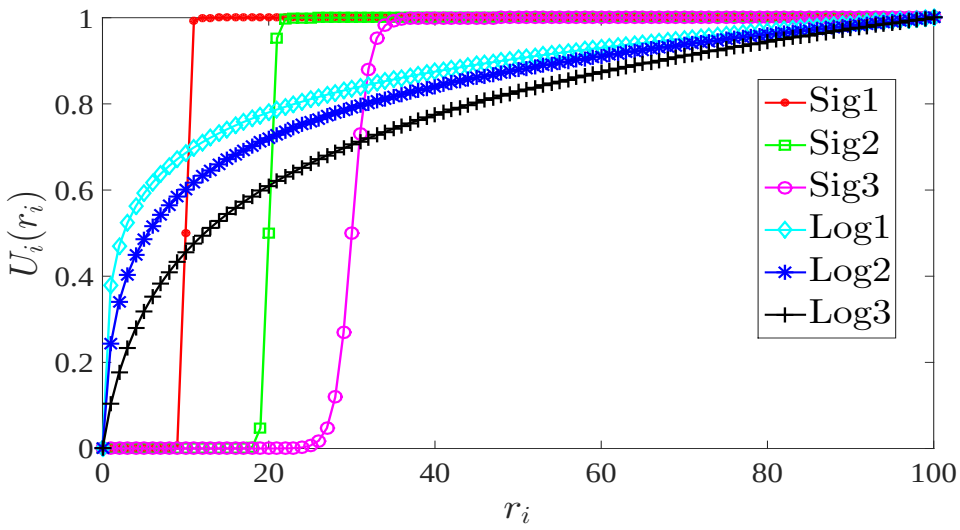


Figure 8.11: The users utility functions $U_i(r_i)$ used in the simulation (three sigmoidal-like functions and three logarithmic functions).

the UE is represented by a sigmoidal or logarithmic utility function based on the application type. We use three normalized sigmoidal utility functions to represent real time applications. The used parameters are $a_i = 5$, $b_i = 10$ with a corresponding utility function denoted by Sig1, $a_i = 3$, $b_i = 20$ with a corresponding utility function denoted by Sig2, and $a_i = 1$, $b_i = 30$ with a corresponding utility function denoted by Sig3 as shown in Figure 8.11. We also use three logarithmic utility functions with $r_i^{\max} = 100$ and different k_i parameters to represent delay tolerant applications. The used parameters are $k_i = 15$, $k_i = 3$ and $k_i = 0.5$ with corresponding utility functions denoted by Log1, Log2 and Log3, respectively, as shown in Figure 8.11.

8.2.3.1 BSs' Bidding Prices and The Final Allocated Rates

In the following simulations, we consider 4 BSs (bidders) $\mathcal{K} = \{1, 2, 3, 4\}$ participating in a secure and truthful spectrum auction. BS1 and BS2 belong to WSP1 (i.e. $\mathcal{K}^1 = \{1, 2\}$) and BS3 and BS4 belong to WSP2 (i.e. $\mathcal{K}^2 = \{3, 4\}$). The 4 BSs are demanding additional resources and therefore they decide to participate in the spectrum auction. The permanent

resources of the k^{th} BS in $\mathcal{K} = \{1, 2, 3, 4\}$ is as follows: $R_1^p = 10$, $R_2^p = 20$, $R_3^p = 30$ and $R_4^p = 40$. Each BS has 4 UEs under its coverage area subscribing for mobile services; i.e. $\mathcal{M}^k = \{1, 2, 3, 4\}$ is the set of active UEs for BS k where 1^{st} UE in $\mathcal{M}^k$ is running real time application represented by Sig2, the 2^{nd} UE is running real time application represented by Sig3, the 3^{rd} UE is running delay tolerant application represented by Log2 and the 4^{th} UE is running delay tolerant application represented by Log3. Algorithm 29 and Algorithm 30 were applied in Matlab for different number of auctioneer's spectrum bands $N \in \{1, 2, ..., 15\}$. Optimal bidding prices for all allocations in $\boldsymbol{\alpha}$ are determined by each BS (bidder) after the execution of Algorithm 29. Additionally, after the execution of Algorithm 30 each BS is allocated its winning spectrum bands by the auctioneer and is charged the corresponding price.

In Figure 8.12(a), we show each of the BS's shadow price (price per unit bandwidth) $p_{k,n}^t$ with the BS's temporary resources vary as $R_{k,n}^t = \{10, 20, 30, ..., 150\}$ which corresponds to $n = \{1, 2, 3, ..., 15\}$ auctioned spectrum bands. We notice that the BS's price per unit bandwidth is monotonically decreasing with $R_{k,n}^t$, i.e. BS's shadow price for first spectrum band is higher than second spectrum band and its price for the second spectrum band is higher than third spectrum band and so on, which is expected as the BS's demand for the l^{th} spectrum band is higher than its demand for the $(l+1)^{th}$ spectrum band. In Figure **??**, we show each BS's calculated bidding price $b_k(\alpha_j)$ for each number of spectrum bands n (i.e $1 \leq n \leq 15$ and $\alpha_j(k) = n$) the BS is bidding for where $b_k(\alpha_j) = \sum_{l=1}^{n} 10 * p_{k,l}^t$. We observe that as the number of spectrum bands n increases, the BS's true bidding price increases. However, the increase in the bidding price is so minimal for higher number of spectrum bands n since the BS's demand for additional resources decreases.

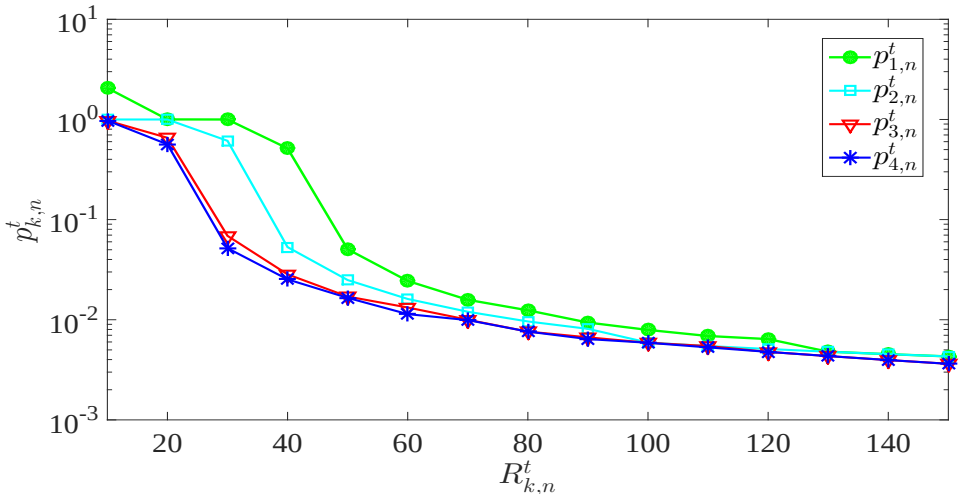
We ran Algorithm 29 and Algorithm 30 in Matlab for $N = 7$ spectrum bands and the number of each BS's winning spectrum bands is obtained as follows: BS1, BS2, BS3 and BS4 won 4, 4, 3 and 3 spectrum bands, respectively. Figure 8.13(a) shows BS3 allocated

rates from its primary resources $R_3^p = 30$, temporary auctioned resources $R_{3,3}^t = 30$ and the aggregated final resources to each of the 4 active UEs under its coverage area. Figure 8.13(b) shows the UEs applications' utilities of their allocated rates from BS3. We notice that real time applications are given priority when allocating the BS's resources due to the nature of their applications and the utility functions used to represent them, and non of the applications is allocated zero resources.

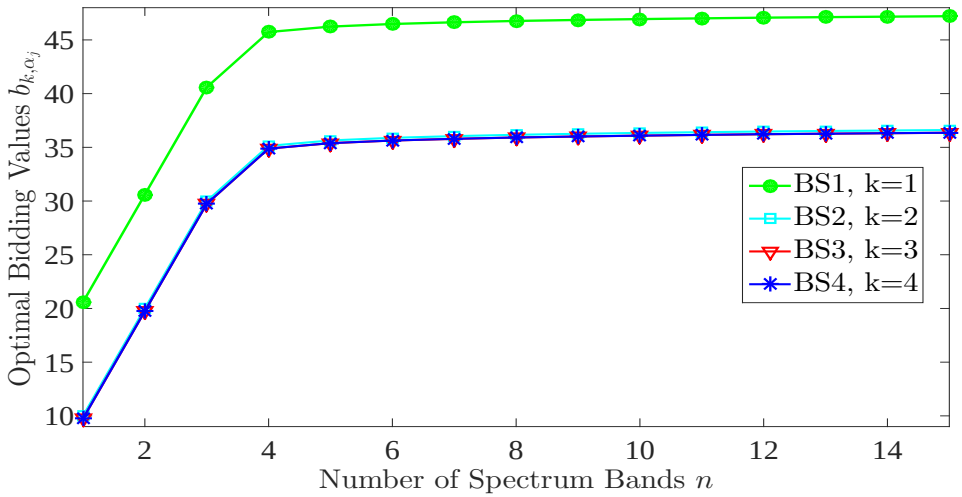
8.2.3.2 Performance Analysis

In the following simulations, we consider the network setup described above for two cases. In the first case, we considered 5 BSs (i.e. $\mathcal{K} = \{1, 2, 3, 4, 5\}$) 3 of them belong to the first WSP and the other 2 belong to the second WSP with the k^{th} BS in $\mathcal{K}$ has a permanent resources $R_k^p = 10 * k$. In the second case, we considered 6 BSs (i.e. $\mathcal{K} = \{1, 2, 3, 4, 5, 6\}$) 3 of them belong to the first WSP and the other 3 belong to the second WSP with the k^{th} BS in $\mathcal{K}$ has a permanent resources $R_k^p = 10 * k$. We ran Monte Carlo Simulation for the two cases and the results are averaged over 5 independent runs. In the 1^{st} run, we consider 4 active UEs under the coverage area of each BS with each UE running one application that is selected randomly with a corresponding utility function equivalent to one of the utility functions shown in Figure 8.11. Similarly, in the 2^{nd}, 3^{rd}, 4^{th} and 5^{th} run, we consider 6, 8, 10 and 12 active UEs, respectively, under the coverage area of each BS with each UE running one application that is selected randomly with a corresponding utility function equivalent to one of the utility functions shown in Figure 8.11. Because of the difference in the number of active UEs and their applications, each run generated different bidding values. We observe the performance of the secure and truthful spectrum auction when using the proposed bidding strategy for the two cases with different number of spectrum bands $N = \{1, 2, 3, ..., 15\}$.

We consider the three performance metrics listed and described below:



(a) The shadow price $p^t_{k,n}$ for each of the 4 BSs (i.e BS $k \in \{1, 2, 3, 4\}$) with the BSs' temporary resources $10 \leq R^t_{k,n} \leq 150$.

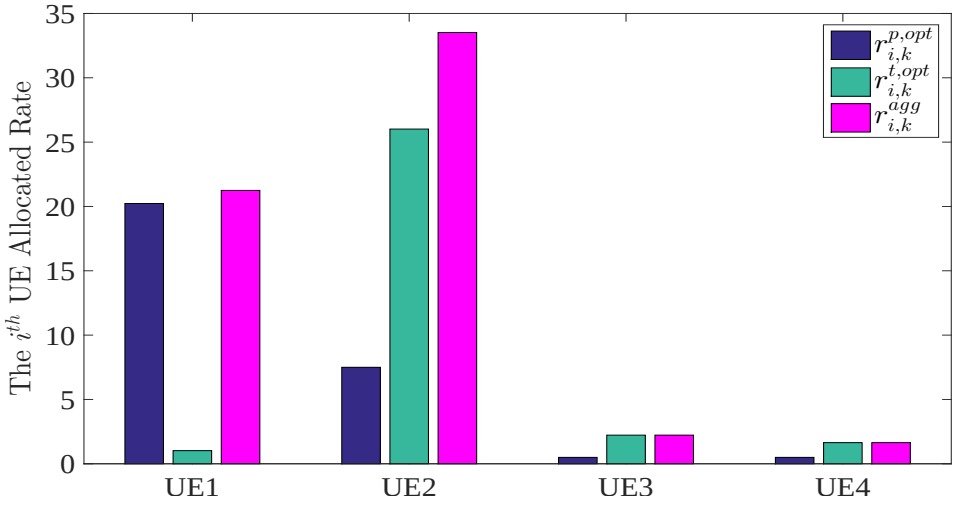


(b) The 4 BSs' (bidders') optimal bidding values b_{k,α_j} with the number of spectrum bands each BS is bidding for, i.e $1 \leq n \leq 15$, for each of the 4 BSs (i.e BS $k \in \{1, 2, 3, 4\}$).

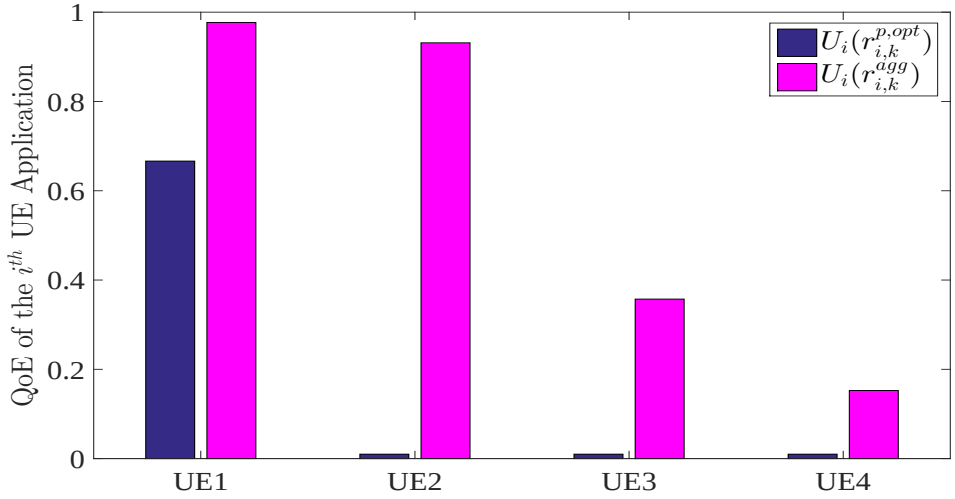
Figure 8.12: The 4 BSs (bidders) calculated shadow price with their temporary resources $10 \leq R^t_{k,n} \leq 150$ and the BSs optimal bidding values with the number of spectrum bands n each BS is bidding for; when the permanent resources of BS1, BS2, BS3 and BS4 are $R^p_1 = 10$, $R^p_2 = 20$, $R^p_3 = 30$ and $R^p_4 = 40$, respectively.

- Spectrum Utilization: It is represented by the sum of the spectrum bands that are allocated by the auctioneer to the winning Bidders.

- Auctioneer's Revenue: It is given by the sum of all Bidders' payments, i.e $Rev = \sum_{k=1}^{k=K} p_k$.



(a) BS3 optimal allocated rates $r_{i,3}^{p,opt}$, $r_{i,3}^{t,opt}$, $r_{i,3}^{agg}$ to each of the 4 UEs when BS3 permanent resources $R_3^p = 30$ and its winning auctioned resources is 3 spectrum bands (n=4) which is equivalent to $R_{3,3}^t = 30$.



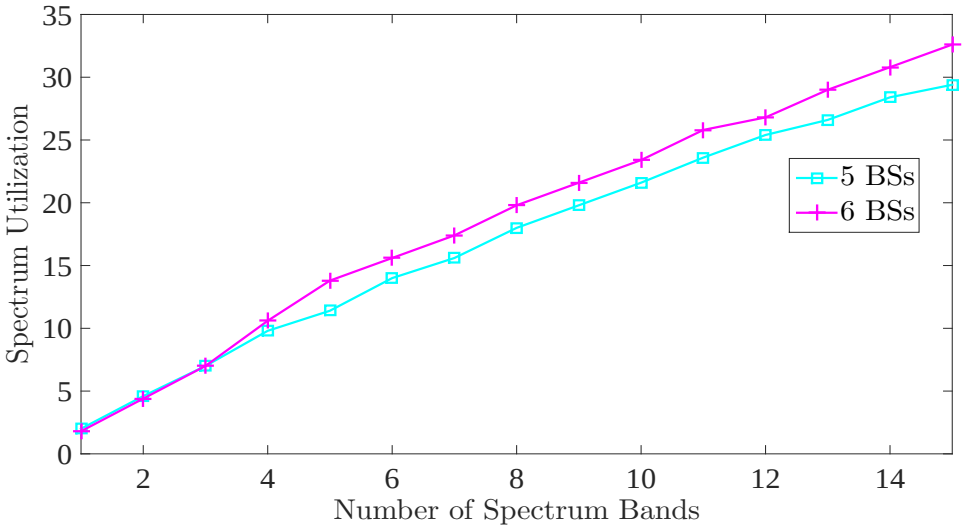
(b) BS3 users' QoE represented by the utility of user's application of its allocated rate $U_i(r_{i,k}^{p,opt})$, $U_i(r_{i,k}^{agg})$ when $R_3^p = 30$ and $R_{3,3}^t = 30$.

Figure 8.13: BS3 allocated rates to users under its coverage area and its users' QoE when $R_3^p = 30$ and $R_{3,3}^t = 30$.

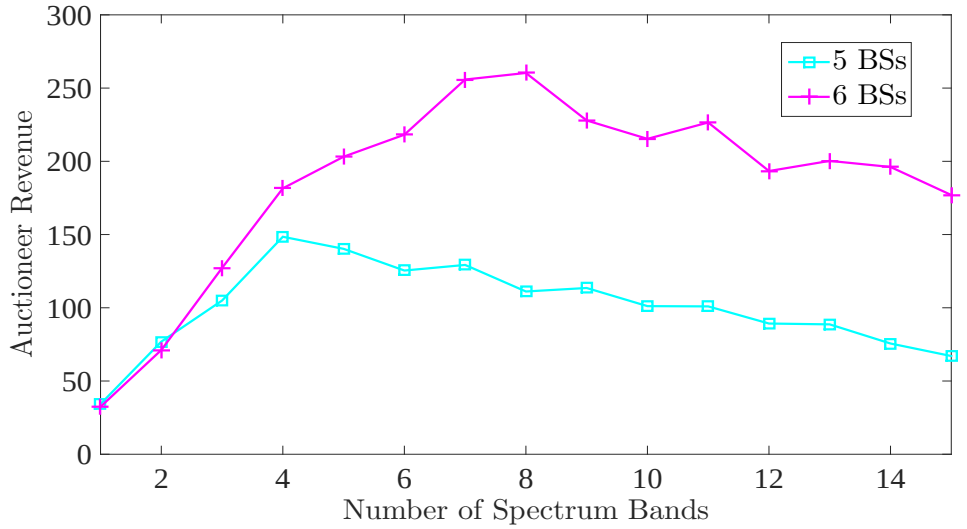
- Bidders' Satisfaction: It is represented by the sum of the utilities of all winning BSs divided by the sum of all bidders' evaluation values, i.e. $\sum_{k\in\mathcal{W}} u_k / \sum_{k\in\mathcal{K}} e_k$, where $\mathcal{W}$ is the set of all winning BSs.

In Figure 8.14(a), we show the spectrum utilization versus the number of the auctioneer's available under-utilized spectrum bands. We observe that as the number of spectrum bands

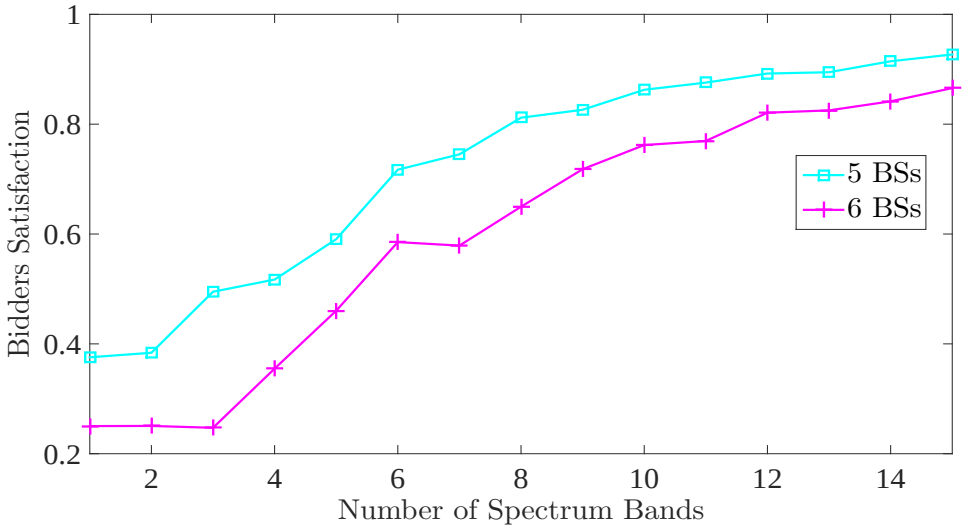
increases, the spectrum utilization, that is represented by the number of allocated spectrum bands, also increases. We also observe that for certain number of spectrum bands, the spectrum utilization is higher when the number of bidders increases. Figure 8.14(b) shows that the auctioneer's revenue increases when the number of BSs increases which is expected as the auctioneer's revenue increases with more bidders requesting more spectrum bands. We show in Figure 8.14(c) that as the number of spectrum bands increases, the bidders' satisfaction also increases until it saturates when each bidder is allocated the number of spectrum bands he bids for.



(a) Spectrum Utilization



(b) Auctioneer's Revenue



(c) Bidders' Satisfaction

Figure 8.14: Performance of the secure and truthful spectrum auction when using the proposed bidding mechanism.

8.3 Summary and Conclusions

In this chapter, we proposed a secure spectrum auction **MTSSA** for a multi-tier dynamic spectrum sharing system. By considering the spectrum reusability property, **MTSSA** enables an efficient sharing of the under-utilized frequency bands with commercial WSPs. In order to allow spectrum reuse among multiple WSPs, the frequency conflict graph that is considered by **MTSSA** includes all BSs that belong to multiple WSPs and the auction is carried out in one subnet after another. **MTSSA** leverages Paillier cryptosystem and a federal gateway to keep the BSs' bidding values unknown to the auctioneer. The auctioneer uses the additive homomorphic property of Paillier cryptosystem to find the winning BSs and their charging prices. This prevents possible frauds and bid-rigging between an insincere auctioneer and greedy BSs. Compared with conventional spectrum leasing mechanism, **MTSSA** has shown better performance while providing a secure spectrum auction against possible back room dealings. Computational and communication complexity analysis showed that the proposed **MTSSA** is more realistic and efficient compared to other spectrum auction mechanisms.

In addition, we proposed an optimal bidding mechanism to be used by BSs participating in a secure spectrum auction. The proposed bidding mechanism allows each BS located within the auction's region and has a demand for additional resources, in order to improve the QoE for its end users' applications, to calculate its true bidding value for each number of spectrum bands it is bidding for. The proposed bidding strategy is ideal for a secure and truthful spectrum auction to guarantee truthfulness which is a dominant strategy to prevent insincere behaviors of greedy bidders and an insincere auctioneer. We presented the BS's bidding algorithm which is used to determine the BS's true bidding value that is based on the number of spectrum bands it is bidding for, the applications traffic of users under its coverage area and the quantity of bidder's permanent resources. In addition, we proved that

by using the proposed bidding mechanism, when a BS bids for certain number of spectrum bands the biding price per each spectrum band is monotonically decreasing as the demand for each additional spectrum band decreases as the number increases. We also presented a resource allocation based on carrier aggregation approach to determine the BS's optimal aggregated rate allocated to each UE from both the BS's permanent resources and the BS's winning auctioned spectrum resources. Simulation results showed the performance of the proposed bidding mechanism when used in a secure spectrum auction.

Chapter 9

Future Research Directions

In this chapter, we discuss some possible research directions in the future to improve and expand the proposed methods presented in this dissertation. An outline of future research direction is as follows:

- Develop resource allocation with carrier aggregation frame work for heterogeneous radio access technologies such as LTE-A and WiFi. Such approach is a challenging one because the two technologies have differences is spectrum access and physical layer. However, the coexistence of LTE and WiFi provides benefit for both, it increases the rate capacity for LTE end users and in the same time introduces reliability for WiFi users. This requires resource management algorithms that are based on CA where UEs are connected to LTE eNodeB as well as WiFi access point and are allocated resources from both, based on carrier aggregation.

- Develop efficient component carriers assignment algorithms for networks incorporating LTE-A with Carrier Aggregation. Component carriers involved in the resource assignment process should not be assigned based on the signal received power level, estimated by each UE, as it is insufficient in case of carrier aggregation. It is more important

to consider the load on each component carrier which can be estimated from counting users on each CC or by measuring the interference on each CC. For example, in heterogeneous networks, the coexistence of macro cells and pico cells makes the resource management task challenging. This is due to the differences in output power between macro BS and pico BS; i.e. 43 to 46 dBm in the case of macro cell and 30 dBm in the case of pico cell, which causes under-utilization in small cells' BS resources due to their lower power levels that make most of users select macro BS.

- Consider providing methods to mitigate the interference caused by carrier aggregation. For example, in LTE-A network, when considering femto cells that are installed by customers, it is important to develop a scheme that optimizes the carrier aggregation selection while avoiding interference among eNodeBs and femoto cells. This can be achieved by selecting an optimal group of component carriers for carrier aggregation in a LTE-A network.

- The proposed spectrum auction mechanisms presented in this dissertation has only considered spectrum heterogeneity in terms of the spectrum being commercial or federal. However, spectrum heterogeneity of the auctioneer's spectrum bands (i.e. spectrum bands with different central frequencies) has not been taken into consideration. Spectrums with different frequencies have different path losses and therefore different transmission ranges. For example: in a spectrum auction, when considering BSs (bidders) in cellular networks that have macro cells and micro/pico/femto cells, each bidder has different targeted cell coverage when bidding for spectrum bands; i.e. macro cells' BSs may favor low frequency spectrum because of its long transmission range whereas femto cells' BSs would prefer high frequency spectrum with transmission range that is enough for an indoor area as it generates less cross-tier interference. In addition, when constructing a conflict graph for bidders participating in a spectrum auction, bidders that are bidding for high frequency spectrum bands and are considered interference free

(i.e. not connected directly to each other in the conflict graph) could cause interference to each other when bidding for low frequency spectrum bands. Therefore, the conflict graph in a spectrum auction with heterogeneous spectrum bands (have different central frequencies) is not fixed during the auction process. Designing a robust spectrum auction, with heterogeneous spectrum bands, that achieves high spectrum utilization while satisfying the essential economic properties, required in a successful auction, is a challenging task. Future research work in spectrum auctions needs to address this concern.

References

[1] A. Abdel-Hadi and C. Clancy, "An Optimal Resource Allocation with Joint Carrier Aggregation in 4G-LTE," arXiv:1405.6448v1, Accepted in ICNC, 2015.

[2] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2 (Release 10)," TSG RAN. TS 36.300 v10.3.0.

[3] H. Ekstrom, "QoS control in the 3GPP evolved packet system," *Communications Magazine, IEEE*, vol. 47, pp. 76 –83, february 2009.

[4] Cisco, *Visual Networking Index*. White paper at Cisco.com, Feb. 2014.

[5] M. Iwamura, K. Etemad, M.-H. Fong, R. Nory, and R. Love, "Carrier aggregation framework in 3GPP LTE-advanced [WiMAX/LTE Update]," *IEEE Communications Magazine*, vol. 48, no. 8, pp. 60–67, 2010.

[6] Y. Wang, K. I. Pedersen, T. B. Sørensen, and P. E. Mogensen, "Utility Maximization in LTE-Advanced Systems with Carrier Aggregation," in *VTC Spring*, pp. 1–5, 2011.

[7] G. RP-091440, "Work Item Description: CarrierAggregation for LTE," December 2009.

[8] S. Parkvall, A. Furuskar, and E. Dahlman, "Evolution of LTE toward IMT-advanced," *Communications Magazine, IEEE*, vol. 49, pp. 84–91, February 2011.

[9] R. L. Kurrle, "Resource Allocation for Smart Phones in 4G LTE Advanced Carrier Aggregation," November 2012.

[10] PCAST, "Final PCAST Spectrum Report ," July 2012.

[11] National Telecommunications and Information Administration (NTIA), "An assessment of the near-term viability of accommodating wireless broadband systems in the 1675-1710 MHz, 1755-1780 MHz, 3500-3650 MHz, 4200-4220 MHz, and 4380-4400 MHz bands (Fast Track Report)." Online, October 2010.

[12] G. Yuan, X. Zhang, W. Wang, and Y. Yang, "Carrier aggregation for LTE-advanced mobile communication systems," *Communications Magazine, IEEE*, vol. 48, pp. 88–93, February 2010.

[13] Z. Shen, A. Papasakellariou, J. Montojo, D. Gerstenberger, and F. Xu, "Overview of 3GPP LTE-advanced carrier aggregation for 4G wireless communications," *Communications Magazine, IEEE*, vol. 50, pp. 122–130, February 2012.

[14] "Frequency spectrum wall chart."

[15] F. Kelly, A. Maulloo, and D. Tan, "Rate control in communication networks: shadow prices, proportional fairness and stability," in *Journal of the Operational Research Society*, vol. 49, 1998.

[16] S. Low, F. Paganini, and J. Doyle, "Internet congestion control," *Control Systems, IEEE*, vol. 22, pp. 28–43, Feb 2002.

[17] J. Mo and J. Walrand, "Fair end-to-end window-based congestion control," *Networking, IEEE/ACM Transactions on*, vol. 8, pp. 556–567, Oct 2000.

[18] S. Shenker, "Fundamental design issues for the future internet," *Selected Areas in Communications, IEEE Journal on*, vol. 13, pp. 1176–1188, Sept 1995.

[19] Z. Cao and E. Zegura, "Utility max-min: an application-oriented bandwidth allocation scheme," in *INFOCOM '99. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 2, pp. 793–801 vol.2, Mar 1999.

[20] S. Sarkar and L. Tassiulas, "Fair allocation of utilities in multirate multicast networks: a framework for unifying diverse fairness objectives," *Automatic Control, IEEE Transactions on*, vol. 47, pp. 931–944, Jun 2002.

[21] A. Abdel-Hadi and C. Clancy, "A utility proportional fairness approach for resource allocation in 4G-LTE," in *Computing, Networking and Communications (ICNC), 2014 International Conference on*, pp. 1034–1040, Feb 2014.

[22] A. Abdel-Hadi and C. Clancy, "A robust optimal rate allocation algorithm and pricing policy for hybrid traffic in 4G-LTE," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2013 IEEE 24th International Symposium on*, pp. 2185–2190, Sept 2013.

[23] A. Abdel-Hadi, C. Clancy, and J. Mitola, "A Resource Allocation Algorithm for Multi-Application Users in 4G-LTE," in *MobiCom Workshop*, 2013.

[24] H. Shajaiah, A. Abdel-Hadi, and C. Clancy, "Utility Proportional Fairness Resource Allocation with Carrier Aggregation in 4G-LTE," in *Military Communications Conference, MILCOM 2013 - 2013 IEEE*, pp. 412–417, Nov 2013.

[25] H. Shajaiah, A. Abdel-Hadi, and C. Clancy, "Spectrum sharing between public safety and commercial users in 4G-LTE," in *Computing, Networking and Communications (ICNC), 2014 International Conference on*, pp. 674–679, Feb 2014.

[26] H. Shajaiah, A. Abdelhadi, and C. Clancy, "Multi-Application Resource Allocation with Users Discrimination in Cellular Networks," in *PIMRC*, 2014.

[27] M. Ghorbanzadeh, A. Abdelhadi, and C. Clancy, "A utility proportional fairness radio resource block allocation in cellular networks," *arXiv:1406.2630v1*.

[28] T. Erpek, A. Abdelhadi, and T. Clancy, "An optimal application-aware resource block scheduling in lte," in *Computing, Networking and Communications (ICNC), 2015 International Conference on*, pp. 275–279, Feb 2015.

[29] M. Awad, V. Mahinthan, M. Mehrjoo, X. Shen, and J. W. Mark, "A Dual-Decomposition-Based Resource Allocation for OFDMA Networks With Imperfect CSI," *Vehicular Technology, IEEE Transactions on*, vol. 59, pp. 2394–2403, Jun 2010.

[30] M. Mehrjoo, S. Moazeni, and X. S. Shen, "Resource allocation in OFDMA networks based on interior point methods," *Wireless Communications and Mobile Computing*, vol. 10, no. 11, pp. 1493–1508, 2010.

[31] P. Tejera, W. Utschick, J. Nossek, and G. Bauch, "Rate Balancing in Multiuser MIMO OFDM Systems," *Communications, IEEE Transactions on*, vol. 57, pp. 1370–1380, May 2009.

[32] L. Xu, X. Shen, and J. W. Mark, "Fair resource allocation with guaranteed statistical QoS for multimedia traffic in wideband CDMA cellular network," *Mobile Computing, IEEE Transactions on*, vol. 4, pp. 166–177, March 2005.

[33] M. Mehrjoo, M. Awad, M. Dianati, and X. Shen, "Design of fair weights for heterogeneous traffic scheduling in multichannel wireless networks," *Communications, IEEE Transactions on*, vol. 58, pp. 2892–2902, October 2010.

[34] R. Madan, S. Boyd, and S. Lall, "Fast algorithms for resource allocation in wireless cellular networks," *Networking, IEEE/ACM Transactions on*, vol. 18, pp. 973–984, June 2010.

[35] Y.-B. Lin, T.-H. Chiu, and Y.-T. Su, "Optimal and near-optimal resource allocation algorithms for OFDMA networks," *Wireless Communications, IEEE Transactions on*, vol. 8, pp. 4066–4077, August 2009.

[36] G. Li and H. Liu, "Downlink dynamic resource allocation for multi-cell OFDMA system," in *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on*, vol. 1, pp. 517–521 Vol.1, Nov 2003.

[37] S. Cicalo, V. Tralli, and A. Perez-Neira, "Centralized vs Distributed Resource Allocation in Multi-Cell OFDMA Systems," in *Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd*, pp. 1–6, May 2011.

[38] M. Dianati, X. Shen, and K. Naik, "Cooperative Fair Scheduling for the Downlink of CDMA Cellular Networks," *Vehicular Technology, IEEE Transactions on*, vol. 56, pp. 1749–1760, July 2007.

[39] P.-L. Tsai, K.-J. Lin, and W.-T. Chen, "Downlink radio resource allocation with Carrier Aggregation in MIMO LTE-advanced systems," in *Communications (ICC), 2014 IEEE International Conference on*, pp. 2332–2337, June 2014.

[40] T. Yang, L. Zhang, and L. Yang, "Cognitive-based distributed interference management for home-enb systems with single or multiple antennas," in *Personal Indoor and Mobile Radio Communications (PIMRC), 2010 IEEE 21st International Symposium on*, pp. 1260–1264, Sept 2010.

[41] A. Attar, V. Krishnamurthy, and O. Gharehshiran, "Interference management using cognitive base-stations for UMTS LTE," *Communications Magazine, IEEE*, vol. 49, pp. 152–159, August 2011.

[42] J. McMenamy, I. Macaluso, N. Marchetti, and L. Doyle, "A methodology to help operators share the spectrum through an enhanced form of carrier aggregation," in *Dynamic Spectrum Access Networks (DYSPAN), 2014 IEEE International Symposium on*, pp. 334–345, April 2014.

[43] X. Lin, J. Andrews, R. Ratasuk, B. Mondal, and A. Ghosh, "Carrier aggregation in heterogeneous cellular networks," in *Communications (ICC), 2013 IEEE International Conference on*, pp. 5199–5203, June 2013.

[44] A. Ghosh, N. Mangalvedhe, R. Ratasuk, B. Mondal, M. Cudak, E. Visotsky, T. Thomas, J. Andrews, P. Xia, H. Jo, H. Dhillon, and T. Novlan, "Heterogeneous cellular networks: From theory to practice," *Communications Magazine, IEEE*, vol. 50, pp. 54–64, June 2012.

[45] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, "A survey on 3GPP heterogeneous networks," *Wireless Communications, IEEE*, vol. 18, pp. 10–21, June 2011.

[46] Federal Communications Commission (FCC), "In the matter of revision of parts 2 and 15 of the commissions rules to permit unlicensed national information infrastructure (U-NII) devices in the 5 GHz band." MO&O, ET Docket No. 03-122, June 2006.

[47] Federal Communications Commission (FCC), "FCC proposes innovative small cell use in 3.5 GHz band." Online: http://www.fcc.gov/document/fcc-proposes-innovative-small-cell-use-35-ghz-band, December 12, 2012.

[48] A. Khawar, A. Abdel-Hadi, and C. Clancy, "Spectrum Sharing between S-band Radar and LTE Cellular System: A Spatial Approach," 2012. IEEE DySPAN-SSPARC Workshop.

[49] S. Sodagari, A. Khawar, T. C. Clancy, and R. McGwier, "A projection based approach for radar and telecommunication systems coexistence," in *IEEE Global Communications Conference (GLOBECOM)*, 2012.

[50] A. Khawar, A. Abdel-Hadi, T. Clancy, and R. McGwier, "Beampattern analysis for mimo radar and telecommunication system coexistence," in *Computing, Networking and Communications (ICNC), 2014 International Conference on*, pp. 534–539, Feb 2014.

[51] H. Deng and B. Himed, "Interference mitigation processing for spectrum-sharing between radar and wireless communications systems," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 49, pp. 1911–1919, July 2013.

[52] H. Huang, M. Trivellato, A. Hottinen, M. Shafi, P. Smith, and R. Valenzuela, "Increasing downlink cellular throughput with limited network MIMO coordination," *Wireless Communications, IEEE Transactions on*, vol. 8, pp. 2983–2989, June 2009.

[53] A. Gjendemsjo, D. Gesbert, G. Oien, and S. Kiani, "Optimal Power Allocation and Scheduling for Two-Cell Capacity Maximization," in *Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks, 2006 4th International Symposium on*, pp. 1–6, April 2006.

[54] S. Das, H. Viswanathan, and G. Rittenhouse, "Dynamic load balancing through coordinated scheduling in packet data systems," in *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, vol. 1, pp. 786–796 vol.1, March 2003.

[55] FCC (2004), "Second report and order: Promoting efficient use of spectrum through elimination of barrier to the development of secondary markets," tech. rep., Tech. rep., FCC 04-167.

[56] S. Sengupta and M. Chatterjee, "An Economic Framework for Dynamic Spectrum Access and Service Pricing," *Networking, IEEE/ACM Transactions on*, vol. 17, pp. 1200–1213, Aug 2009.

[57] M. Pan, F. Chen, X. Yin, and Y. Fang, "Fair Profit Allocation in the Spectrum Auction Using the Shapley Value," in *Global Telecommunications Conference, 2009. GLOBECOM 2009. IEEE*, pp. 1–6, Nov 2009.

[58] J. Zhang and Q. Zhang, "Stackelberg Game for Utility-based Cooperative Cognitive radio Networks," MobiHoc '09, (New York, NY, USA), pp. 23–32, ACM, 2009.

[59] X. Zhou, S. Gandhi, S. Suri, and H. Zheng, "eBay in the Sky: Strategy-proof Wireless Spectrum Auctions," MobiCom '08, (New York, NY, USA), pp. 2–13, ACM, 2008.

[60] X. Zhou and H. Zheng, "TRUST: A General Framework for Truthful Double Spectrum Auctions," in *INFOCOM 2009, IEEE*, pp. 999–1007, April 2009.

[61] J. Jia, Q. Zhang, Q. Zhang, and M. Liu, "Revenue Generation for Truthful Spectrum Auction in Dynamic Spectrum Access," MobiHoc '09, (New York, NY, USA), pp. 3–12, ACM, 2009.

[62] C.-C. Wu, C.-C. Chang, and I.-C. Lin, "New Sealed-Bid Electronic Auction with Fairness, Security and Efficiency," *J. Comput. Sci. Technol.*, vol. 23, no. 2, pp. 253–264, 2008.

[63] K. Peng, C. Boyd, and E. Dawson, "Batch Verification of Validity of Bids in Homomorphic e-Auction," *Comput. Commun.*, vol. 29, pp. 2798–2805, Sept. 2006.

[64] S. Gandhi, C. Buragohain, L. Cao, H. Zheng, and S. Suri, "A General Framework for Wireless Spectrum Auctions," in *DySPAN 2007*, pp. 22–33, April 2007.

[65] K. Jain, J. Padhye, V. N. Padmanabhan, and L. Qiu, "Impact of Interference on Multihop Wireless Network Performance," MobiCom '03, (New York, NY, USA), pp. 66–80, ACM, 2003.

[66] R. Weber, "Auction Theory: By Vijay Krishna. Academic Press, 2002," *Games and Economic Behavior*, vol. 45, no. 2, pp. 488–497, 2003.

[67] Y. Wu, B. Wang, K. Liu, and T. Clancy, "A multi-winner cognitive spectrum auction framework with collusion-resistant mechanisms," in *DySPAN 2008.*, pp. 1–9, Oct 2008.

[68] F. Wu and N. Vaidya, "A Strategy-Proof Radio Spectrum Auction Mechanism in Noncooperative Wireless Networks," *Mobile Computing, IEEE Transactions on*, vol. 12, pp. 885–894, May 2013.

[69] R. McAfee, "A dominant strategy double auction," *Journal of Economic Theory*, vol. 56, no. 2, pp. 434 – 450, 1992.

[70] D. Yang, X. Fang, and G. Xue, "Truthful Auction for Cooperative Communications," MobiHoc '11, (New York, NY, USA), pp. 9:1–9:10, ACM, 2011.

[71] X. Feng, Y. Chen, J. Zhang, Q. Zhang, and B. Li, "TAHES: A Truthful Double Auction Mechanism for Heterogeneous Spectrums," *Wireless Communications, IEEE Transactions on*, vol. 11, pp. 4038–4047, November 2012.

[72] Y. Chen, J. Zhang, K. Wu, and Q. Zhang, "TAMES: A Truthful Auction Mechanism for heterogeneous spectrum allocation," in *INFOCOM, 2013 Proceedings IEEE*, pp. 180–184, April 2013.

[73] Q. Huang, Y. Tao, and F. Wu, "SPRING: A Strategy-proof and Privacy preserving spectrum auction mechanism," in *INFOCOM, 2013 Proceedings IEEE*, pp. 827–835, April 2013.

[74] W. Dong, S. Rallapalli, R. Jana, L. Qiu, K. Ramakrishnan, L. Razoumov, Y. Zhang, and T. W. Cho, "iDEAL: Incentivized dynamic cellular offloading via auctions," in *INFOCOM, 2013 Proceedings IEEE*, pp. 755–763, April 2013.

[75] X. Zhuo, W. Gao, G. Cao, and Y. Dai, "Win-Coupon: An incentive framework for 3G traffic offloading," in *Network Protocols (ICNP), 2011 19th IEEE International Conference on*, pp. 206–215, Oct 2011.

[76] M. Dong, G. Sun, X. Wang, and Q. Zhang, "Combinatorial auction with time-frequency flexibility in cognitive radio networks," in *INFOCOM, 2012 Proceedings IEEE*, pp. 2282–2290, March 2012.

[77] F. Wu and N. Vaidya, "SMALL: A Strategy-proof Mechanism for radio spectrum allocation," in *INFOCOM, 2011 Proceedings IEEE*, pp. 81–85, April 2011.

[78] M. Yokoo and K. Suzuki, "Secure Multi-agent Dynamic Programming Based on Homomorphic Encryption and Its Application to Combinatorial Auctions," in *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems: Part 1*, AAMAS '02, (New York, NY, USA), pp. 112–119, ACM, 2002.

[79] K. Suzuki and M. Yokoo, "Secure Generalized Vickrey Auction Using Homomorphic Encryption," in *Financial Cryptography* (R. N. Wright, ed.), vol. 2742 of *Lecture Notes in Computer Science*, pp. 239–249, Springer, 2003.

[80] M. Yokoo and K. Suzuki in *Financial Cryptography* (A. Juels, ed.), vol. 3110 of *Lecture Notes in Computer Science*, pp. 132–146, Springer, 2004.

[81] M. Pan, J. Sun, and Y. Fang, "Purging the Back-Room Dealing: Secure Spectrum Auction Leveraging Paillier Cryptosystem," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 4, pp. 866–876, 2011.

[82] J.-W. Lee, R. R. Mazumdar, and N. B. Shroff, "Downlink power allocation for multi-class wireless systems," *IEEE/ACM Trans. Netw.*, vol. 13, pp. 854–867, Aug. 2005.

[83] G. Tychogiorgos, A. Gkelias, and K. K. Leung, "Utility-proportional fairness in wireless networks.," in *PIMRC*, pp. 839–844, IEEE, 2012.

[84] J. B. Taylor, *Principles of microeconomics*. Microeconomics Series, Houghton Mifflin, 1998.

[85] H. Shajaiah, A. Khawar, A. Abdel-Hadi, and T. Clancy, "Resource allocation with carrier aggregation in LTE Advanced cellular system sharing spectrum with S-band radar," in *Dynamic Spectrum Access Networks (DYSPAN), 2014 IEEE International Symposium on*, pp. 34–37, April 2014.

[86] H. Shajaiah, A. Abdelhadi, and T. C. Clancy, "A price selective centralized algorithm for resource allocation with carrier aggregation in LTE cellular networks," *arXiv:1408.4151, Accepted in WCNC, 2015*.

[87] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[88] S. H. Low and D. E. Lapsley, "Optimization flow control, i: Basic algorithm and convergence," *IEEE/ACM Transactions on Networking*, vol. 7, no. 6, pp. 861–874, 1999.

[89] S. Boyd and L. Vandenberghe, *Introduction to convex optimization with engineering applications*. Course Reader, 2001.

[90] J. Li and P. Stoica, "MIMO radar with colocated antennas," *Signal Processing Magazine, IEEE*, vol. 24, no. 5, pp. 106–114, 2007.

[91] Y. Noam and A. Goldsmith, "Blind null-space learning for mimo underlay cognitive radio with primary user interference adaptation," *Wireless Communications, IEEE Transactions on*, vol. 12, no. 4, pp. 1722–1734, 2013.

[92] I.-H. Hou and C. S. Chen, "Self-organized resource allocation in LTE systems with weighted proportional fairness," in *Communications (ICC), 2012 IEEE International Conference on*, pp. 5348–5353, June 2012.

[93] I. T. S. Sesia and M. Baker, *LTE - The UMTS Long Term Evolution: From Theory to Practice*. John Wiley Son, 2011.

[94] T. Eprek, A. Abdelhadi, and T. Clancy, "An Optimal Application-Aware Resource Block Scheduling in LTE.," in *Accepted in ICNC, 2014*, 2014.

[95] FCC, *Spectrum Policy Task Force Report*. Federal Communications Commission, Spectrum Policy Task Force, 2002.

[96] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty, "Next generation/dynamic spectrum access/cognitive radio wireless networks: A survey," *Comput. Netw.*, vol. 50, pp. 2127–2159, Sept. 2006.

[97] P. Paillier, "Public-Key Cryptosystems Based on Composite Degree Residuosity Classes," *j-LECT-NOTES-COMP-SCI*, vol. 1592, pp. 223–238, 1999.

[98] P. Paillier and D. Pointcheval, "Efficient Public-Key Cryptosystems Provably Secure Against Active Adversaries," in *Proceedings of the International Conference on the Theory and Applications of Cryptology and Information Security: Advances in Cryptology*, ASIACRYPT '99, (London, UK, UK), pp. 165–179, Springer-Verlag, 1999.

[99] FCC (2003), "The development of secondary markets-Report and Order and further notice of proposed rule marketing ," tech. rep., Tech. rep., FCC 03-113.

[100] FCC (2010), "Unlicensed Operations in the TV Broadcast Bands, Second Memorandum Opinion and Order," tech. rep., Tech. rep., FCC 10-174.

[101] J. Jose, A. Abdel-Hadi, P. Gupta, and S. Vishwanath, "On the impact of mobility on multicast capacity of wireless networks," in *INFOCOM, 2010 Proceedings IEEE*, pp. 1–5, March 2010.

[102] T. Groves, "Incentives in teams," *Econometrica*, vol. 41, p. 617631, 1973.

[103] W. Vickrey, "Counterspeculation, Auctions and Competitive Sealed Tenders," *Journal of Finance*, pp. 8–37, 1961.

[104] N. Nisan, T. Roughgarden, É. Tardos, and V. V. Vazirani, *Algorithmic Game Theory*. New York, NY, USA: Cambridge University Press, 2007.

[105] M. Abe and K. Suzuki, "M + 1-st price auction using homomorphic encryption," in *Public Key Cryptography* (D. Naccache and P. Paillier, eds.), vol. 2274 of *Lecture Notes in Computer Science*, pp. 115–124, Springer Berlin Heidelberg, 2002.

[106] F. Xue and P. R. Kumar, "The number of neighbors needed for connectivity of wireless networks," *Wirel. Netw.*, vol. 10, pp. 169–181, Mar. 2004.

[107] A. Abdelhadi, H. Shajaiah, and C. Clancy, "A Multi-Tier Wireless Spectrum Sharing System Leveraging Secure Spectrum Auctions," *Cognitive Communications and Networking, IEEE Transactions on*, 2015.