

UNIVERSITY OF OKLAHOMA
GRADUATE COLLEGE

DESIGN AND DEVELOPMENT OF CARRIER ASSIGNMENT AND PACKET
SCHEDULING IN LTE-A AND Wi-Fi

A DISSERTATION
SUBMITTED TO THE GRADUATE FACULTY
in partial fulfillment of the requirements for the
Degree of
DOCTOR OF PHILOSOPHY

By
HUSNU SANER NARMAN
Norman, Oklahoma
2016

DESIGN AND DEVELOPMENT OF CARRIER ASSIGNMENT AND PACKET
SCHEDULING IN LTE-A AND Wi-Fi

A DISSERTATION APPROVED FOR THE
SCHOOL OF COMPUTER SCIENCE

BY

Dr. Mohammed Atiquzzaman, Chair

Dr. Changwook Kim

Dr. Dean Hougen

Dr. Qi Cheng

Dr. Lucy Lifschitz

© Copyright by HUSNU SANER NARMAN 2016
All Rights Reserved.

This dissertation is dedicated to my family...

Acknowledgments

This dissertation has been possible because of the guidance of my adviser, the helps from friends, and the supports from my sponsor, my family, wife and children. First of all, I would like to express my deepest gratitude to my advisor, Dr. Mohammed Atiquzzaman, for his expert advices, continuous supervision throughout my PhD studies and providing me with an excellent atmosphere for doing research.

I am truly grateful to my committee members: Dr. Changwook Kim, Dr. Dean Hougen, Dr. Qi Cheng and Dr. Lucy Lifschitz for their valuable advices and time to review the dissertation. Their comments and suggestions have been very helpful to improve the quality of this dissertation.

I would like to thank Dr. Turgay Korkmaz and Serkan Ozturk for their spiritual supports and they were always willing to help and give their best suggestions during my PhD. I am also thankful to all the faculty and staff members of the School of Computer Science at the University of Oklahoma for providing a supportive environment for my study.

I appreciate the financial support of The Ministry of National Education, Republic of Turkey. Without their supports, I could not have the opportunity to work on my higher education in the United States. Their valuable supports made this work happen.

I am thankful to my parents, sisters and brothers. They have always been supporting me and encouraging me with their best wishes. I would also like to thank my wife, Fatma Narman. She has always stood by me through the good and bad times. Finally, I would like to thank my children, Esat and Elif for their warm smiles which have motivated me to study.

Table of Contents

Acknowledgments	iv
List of Tables	x
List of Figures	xi
Abstract	xiv
1 Introduction	1
1.1 Carrier Aggregation	2
1.2 Multi-Band Wi-Fi Routers	3
1.3 Objective and Contributions of the Research	5
1.4 Organization of the Dissertation	6
2 Component Carrier Assignment: Past and Future	7
2.1 Introduction	7
2.2 Carrier Assignment Challenges	11
2.2.1 Primary and Secondary Component Carrier Assignment	11
2.2.2 Frequency Division Duplex (FDD) - Time Division Duplex (TDD) CA	12
2.2.3 Dual Connectivity	13
2.2.4 Licensed Assisted Access (LAA) / LTE in Unlicensed Band (LTE-U)	13
2.3 Resource Management: PCC and SCCs Assignment	14
2.3.1 Carrier Deployment Scenarios	15
2.3.2 Component Carrier Assignment Strategies	16
2.3.2.1 Random	18
2.3.2.2 Load Balancing	18
2.3.2.3 Channel Quality Indicator	19
2.3.2.4 Fairness	20
2.3.2.5 Data Traffic Type and Usage	21
2.3.2.6 Energy Efficiency and Number of CCs	21
2.3.2.7 Mobility	22

2.3.3	The Carrier Assignment Methods	22
2.3.3.1	Packet Scheduling and CQI Measurement	23
2.3.3.2	Comparison of the Methods	23
2.4	Summary	25
3	Joint and Selective Periodic Component Carrier Assignment for LTE and LTE-A	26
3.1	Introduction	27
3.2	System Model with Joint and Selective Techniques	28
3.2.1	Joint Periodic Component Carrier Assignment (j-pCCA)	29
3.2.2	Proposed Selective Periodic Component Carrier Assignment (s-pCCA)	30
3.2.3	Methods	31
3.2.3.1	Random (R)	32
3.2.3.2	Least Load (LL)	32
3.2.3.3	Channel Quality (CQ)	32
3.2.3.4	Least Load Rate (LR)	32
3.3	Analysis	33
3.3.1	Notations	34
3.3.2	Disjoint Queue Scheduler of j-pCCA and s-pCCA for Downlink	34
3.3.2.1	Assumptions	35
3.3.2.2	Performance Metrics	36
3.3.3	Joint Queue Scheduler of j-pCCA and s-pCCA for Downlink	38
3.3.3.1	Assumptions	38
3.3.3.2	State Probability	39
3.3.3.3	Drop Probability	41
3.3.3.4	Average Queue Length and Average Delay	42
3.4	Simulation of the System	43
3.4.1	Assumptions for eNBs	43
3.4.2	Assumptions for UEs	44
3.4.3	Packet Scheduling	45
3.4.4	Observation Methodology	45
3.5	Results	47
3.5.1	Average Delay and Drop Ratio during Carrier Assignment Operations	47
3.5.2	Overall Performance of the System	49
3.5.2.1	Average Delay	49
3.5.2.2	Throughput Ratio	50
3.5.3	Average Delay and Throughput Ratio According to Equipment Type	52
3.5.3.1	Average Delay	52
3.5.3.2	Throughput Ratio	54
3.5.4	Summary of Results	55

3.6	Summary	56
4	Component Carriers Assignment Method Based on User Profile in LTE and LTE-A	57
4.1	Introduction	58
4.2	User Profile	60
4.2.1	The Technique to Collect User Information	61
4.2.2	User Profile Detection	62
4.2.3	Number of Required CCs for Each User	63
4.2.4	Component Carrier Assignment Process Based on User Profile .	64
4.2.5	Carrier Assignment Process	65
4.3	Simulation of the System	65
4.3.1	Assumptions for UEs	66
4.3.2	Observation Methodology	66
4.4	Results	66
4.4.1	Utilization	67
4.4.2	Average Delay	68
4.4.3	Throughput Ratio	71
4.4.4	Summary of Results	73
4.5	Summary	74
5	Analysis of Multi Class Traffic of Single and Multi-Band Routers	75
5.1	Introduction	76
5.2	The Communication Models of Single and Multi-Band Systems	79
5.3	Single Band Router Architecture	80
5.4	Current Multi-Band Router Architecture	81
5.5	Proposed Multi-Shared-Band Router Architecture	82
5.5.1	Time and Space Priority	83
5.5.1.1	Q_B Queue	83
5.5.1.2	Q_R Queue	84
5.5.1.3	Q_N Queue	84
5.5.2	Scheduling Algorithm	85
5.6	Analysis	86
5.6.1	Assumptions	86
5.6.2	Notations	86
5.6.3	Priority	86
5.6.4	Performance Metrics of Current Multi-Band Systems	88
5.6.5	Performance Metrics of the Proposed Multi-Shared-Band System	89
5.6.5.1	Total Arrival Rates in Each Queue	89
5.6.5.2	Computing Drop Probability	90
5.6.5.3	Average Queue Length	91
5.6.5.4	Throughput	92
5.6.5.5	Average Delay	93

5.6.6	Overall System Performance	93
5.6.6.1	Average Occupancy of the System	93
5.6.6.2	Drop Probability of the System	94
5.6.6.3	Throughput of the System	94
5.6.6.4	Average Delay of the System	94
5.7	Simulation of the System	94
5.8	Results	95
5.8.1	Validation of Developed Analytical Formulas	95
5.8.1.1	Queue Throughput	96
5.8.1.2	Average Class Occupancy	96
5.8.1.3	Class Throughput	97
5.8.2	Comparison of Single, Current Multi-Band and Multi-Shared-Band Systems	98
5.8.2.1	Band Utilization	98
5.8.2.2	Average Class Occupancy	99
5.8.2.3	Class Throughput	101
5.8.3	Discussion on Comparison of Systems	102
5.8.4	Comparison of Single and Multi-Shared-Band Architectures with FSF, LUF, and SSF Allocations	103
5.8.4.1	Band Utilization, Packet Delay, and Drop Ratio of Single and Multi-Shared-Band Systems	103
5.8.4.2	Average Class Delay and Throughput	106
5.8.5	Discussion on Comparison of Single Band and Multi-Shared-Band Allocations	108
5.9	Summary	109
6	Energy Aware Scheduling and Queue Management for Multi-Shared-Band Router	110
6.1	Introduction	110
6.2	Single and Multi-Band Routers	112
6.3	Energy Aware Scheduling Algorithm (e-ASA)	113
6.3.1	Notations	114
6.3.2	Wake up and Sleep Modes of the Bands	114
6.3.3	The Scheduling Procedure	116
6.4	Simulation of the System	117
6.4.1	Assumptions for the System	117
6.4.2	Assumptions for Users	118
6.4.3	Band Selection and Packet Scheduling	118
6.4.4	Observation Methodology	119
6.5	Results	120
6.5.1	Band Utilization	120
6.5.2	Energy	121

6.5.3 Throughput Ratio	122
6.5.4 Discussion on Results	123
6.6 Summary	123
7 Conclusions and Future Works	124
Bibliography	127
Appendix A	
Author's List of Publications	141
A.1 Peer Reviewed Publications	141
A.2 Papers under Review	142
Appendix B	
Acronyms	143

List of Tables

2.1	The comparison of the methods.	24
3.1	The notations for Chapter 3.	35
3.2	The simulation parameters.	43
4.1	Mobile user profiles.	58
4.2	User profile detection based on eNBs.	62
5.1	The notations for Chapter 5.	87
6.1	The notations for Chapter 6.	114
6.2	The parameters of the systems.	117

List of Figures

1.1	eNodeB (eNB) with multiple bands and several UEs.	2
1.2	A current multi-band router communication scenario.	4
2.1	Types of Carrier Aggregation.	8
2.2	Radio Resource Management framework in LTE-A [29].	9
2.3	Component carrier allocations for LTE and LTE-A.	11
2.4	Carrier deployment scenarios for multi-band [42].	15
2.5	Symmetric and asymmetric allocations.	17
3.1	A carrier assignment model with n users and m available CCs.	29
3.2	Downlink Disjoint Queue Model with n users and <i>one</i> available CC.	35
3.3	Downlink Disjoint Queue Model with one user and primary and secondary carrier queues.	36
3.4	Downlink Joint Queue Model with n users and m available CCs.	38
3.5	Downlink Joint Queue Model with one user and m available CCs.	39
3.6	State transaction diagram for the Joint Queue Scheduler model.	40
3.7	Average delay during the periodic carrier assignment process for joint and selective techniques.	47
3.8	Drop probability during the periodic carrier assignment process for joint and selective techniques.	48
3.9	Average delay per packet for joint and selective techniques.	50
3.10	Throughput ratio for joint and selective techniques.	51
3.11	Average delay per packet of LTE type devices for joint and selective techniques.	52
3.12	Average delay per packet of LTE-A type devices for joint and selective techniques.	53
3.13	Throughput ratio of LTE type devices for joint and selective techniques. .	54
3.14	Throughput ratio of LTE-A type devices for joint and selective techniques.	55
4.1	Utilization of CCs for user profile and the other methods.	67
4.2	Average delay per packet of LTE type equipment for user profile and the other methods.	68
4.3	Average delay per packet of LTE-A type equipment for user profile and the other methods.	69

4.4	Average delay per packet for user profile and the other methods.	70
4.5	Throughput ratio of LTE type equipment for user profile and the other methods.	71
4.6	Throughput ratio of LTE-A type equipment for user profile and the other methods.	72
4.7	Throughput ratio for user profile and the other methods.	73
5.1	A single band router communication scenario.	79
5.2	A current multi-band router communication scenario.	80
5.3	Single band mobile router architecture.	80
5.4	Architecture of a current (simultaneous) multi-band mobile router.	81
5.5	Proposed architecture of a simultaneous multi-shared-band mobile router.	82
5.6	A multi-shared-band router communication scenario.	83
5.7	Queue corresponding to Q_B band.	84
5.8	Queue corresponding to Q_R band.	84
5.9	Queue corresponding to Q_N band.	85
5.10	Queue throughput of the multi-shared-band system obtained through simulations and analytical model.	96
5.11	Average class occupancy of the multi-shared-band system obtained through simulations and analytical model.	97
5.12	Class throughput of the multi-shared-band system obtained through simulations and analytical model.	97
5.13	Band utilization of the single and current multi-band systems.	98
5.14	Band utilization of the current multi-band and multi-shared-band systems.	99
5.15	Average class occupancies of the single and current multi-band systems. .	100
5.16	Average class occupancies of the current multi-band and multi-shared-band systems.	100
5.17	Class throughput of the single and current multi-band systems.	101
5.18	Class throughput of the current multi-band and multi-shared-band systems.	102
5.19	Band Utilization of Single, FSF, LUF, and SSF systems.	103
5.20	Average delays of Single, FSF, LUF, and SSF systems.	104
5.21	Drop probabilities of Single, FSF, LUF, and SSF systems.	105
5.22	Average class delays of FSF and Single.	105
5.23	Average class delays of LUF, SSF and Single.	106
5.24	Average class delays of FSF, LUF, and SSF.	106
5.25	Class throughput of FSF and Single.	107
5.26	Class throughput of LUF, SSF, and Single.	107
5.27	Class throughput of FSF, LUF, and SSF.	108
6.1	Single band queuing system.	112
6.2	Current multi-band queuing system.	113
6.3	Multi-shared-band queuing system.	113
6.4	Average band utilization.	120

6.5	Energy consumptions of the systems.	121
6.6	Throughput ratios of the systems.	122

Abstract

The highly competitive environment in today's wireless and cellular network industries is making the management of systems seek for better and more advance techniques to keep masses of data, complexity of systems and deadline constrains under control with a lower cost and higher efficiency. Therefore, the management is getting significant attentions by researchers in order to increase the efficiency of the resource usage to provide high quality services. Two of the cornerstones of the management system in wireless and cellular network are carrier assignment and packet scheduling. Therefore, this work focuses on analysis and development of carrier assignment and packet scheduling methods in multi-band Wi-Fi and LTE-A networks. First, several existing carrier assignment methods which are developed by considering different strategists in LTE and LTE-A are analyzed. Secondly, a new technique for the carrier assignment methods for LTE and LTE-A is developed to improve the efficiency of carrier assignment methods. Thirdly, a novel carrier assignment method is proposed by considering the behaviors of mobile users for LTE and LTE-A. Then, a novel architecture with packet scheduling scheme is proposed for next generation mobile routers in multi-band Wi-Fi environment as similar to LTE-A. Finally, the scheme is improved based on energy awareness. Results show that the developed methods improve the performance of the systems in comparison to existing methods. The proposed methods and related analysis should help network engineers and service providers build next generation carrier assignment and packet scheduling methods to satisfy users in LTE, LTE-A and Wi-Fi.

Chapter 1

Introduction

Mobile devices (such as tablet, smartphones, etc.) are being an essential part of human life [1–3]. This necessity results in an enormous growth in the number of mobile devices. According to Gsma Intelligence report [4], the number of active mobile devices passed human population in the world. Currently, there are 7.6 billion mobile devices with 3.7 billion unique mobile subscribers [4]. In 2013, the number of purchased smartphones passed one billion and in 2017, two billion smartphones are expected to be sold [5]. The most notable reason for the increase in the number of such devices is that the users can reach wide range of applications under different platforms (e.g., GooglePlay, AppStore) by cutting cross time and place restrictions [6–8]. For example, more than 100 billion applications have been downloaded in 2013 and more than 250 billion applications are expected to be downloaded in 2017 [5].

The number of mobile device also affects the Internet usage because an increasing number of mobile users access larger data by using wireless connectivity [9–13] in mobile environment [14–16]. Therefore, the bandwidth demand for mobile Internet access is increasing. To answer users' needs, Carrier Aggregation¹ in LTE-A and multi-band routes² in Wi-Fi, have been developed.

¹Detailed information about carrier aggregation is given in Chapter 2

²Detailed information about multi-band routers is given in Chapter 5

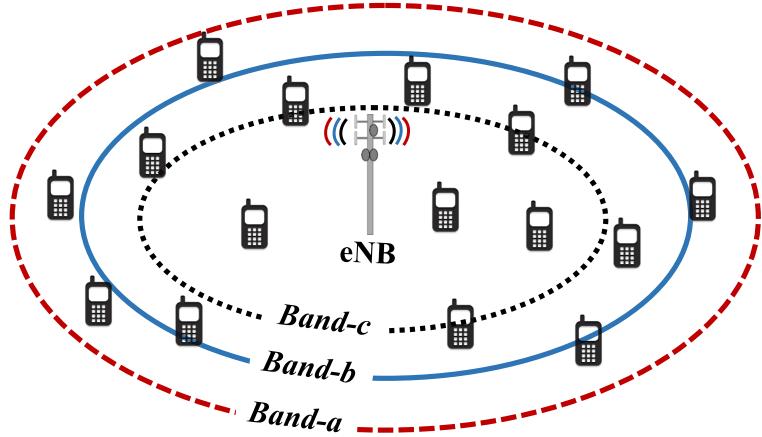


Figure 1.1: eNodeB (eNB) with multiple bands and several UEs.

1.1 Carrier Aggregation

Figure 1.1 shows a deployment scenario for the multi-band system in mobile networks. In this system, each band may have different communication coverages and each band is sub-divided into Component Carriers (CC). Each bandwidth of CC can be 1.4, 3, 5, 10, 15 and 20 MHz. By Carrier Aggregation, User Equipment (UE) can simultaneously connect one or multiple CCs from the same or different bands according to the capacities of equipment³ if users are in the coverages of the bands as illustrated in Figure 1.1.

CCs are classified as primary and secondary component carriers. Primary CC is the main carrier for each user and Secondary CCs are the auxiliary carriers to boost data rates. By Carrier Aggregation and multi-input multi-output (MIMO) technologies⁴, LTE-A supports 1.5 Gbps for uploading and 3 Gbps for downloading peak data rates [17].

Due to multiple bands and CCs connections, there are several challenges which need to be solved. One of the main challenges is how to assign primary and secondary carriers to each user to serve better. This is the carrier assignment problem which is mostly focused in this research. Some of the other challenges are: How should be carrier assignment (i) if both time and frequency full duplex based carriers exist in the system, (ii) if a user can

³For example, LTE equipment cannot connect multiple CCs.

⁴Multi-antenna multi-path

receive service from two different eNBs at the same time, and (iii) if the same band is used by other network such as Wi-Fi. Therefore, assignment of multiple CCs to UEs must be carefully designed because an inefficient carrier assignment method can decrease system performances [18–21].

Several methods have been proposed for primary and secondary carrier assignment [19] by following different strategies such as load balancing, fairness, energy etc. to increase the efficiency of the system. Some of the methods assign carriers to users as a result of the mandatory changes such as path loss, low channel quality, etc. However, it has been shown that if the carrier assignment occurs periodically in addition to mandatory changes, the resource management becomes more efficient [22].

Although the known methods have improved the efficiency of the carrier assignment, there are several limitations which need to be improved. One of the limitation of the previous works is that the methods are evaluated based on the overall performance by ignoring the behaviors of the methods during the carrier assignment operations. Especially, the periodic carrier assignment interrupts packet transfers during the carrier assignment process due to the fact that all carriers are updated together. Therefore, in Chapter 3, the performance of the system is analyzed during the carrier assignment operations and a new technique is proposed for the periodic carrier assignment. Moreover, a new strategy, user profile, based on the historical information of the users is proposed in Chapter 4 to make the carrier assignment more efficient because continuous increase in bandwidth demand of users forces the operators to manage the resource allocation more intelligently.

1.2 Multi-Band Wi-Fi Routers

Figure 1.2 shows the multi-band architecture for a Wi-Fi router. Similar to LTE-A, not only some of the current router operate but also next generation Wi-Fi routers will simultaneously operate on multiple bands to provide services to users. The benefit of using a multi-band in the routers is less interference, higher capacity and better reliability. For

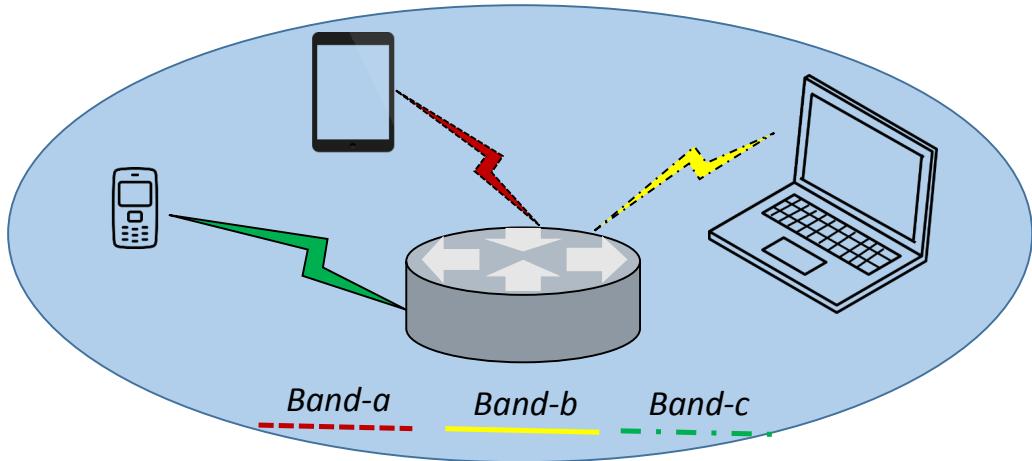


Figure 1.2: A current multi-band router communication scenario.

example, it is expected that speed of future IEEE 802.11ad (WiGig) tri-band routers be up to 7 Gbps [23]. However, the current architecture does not allow multiple users to receive services over multiple bands simultaneously which wastes resource due to unutilized bands. There have been some previous works which explain the possible Wi-Fi architecture with multiple physical and link layers to support multiple frequency bands simultaneously [24, 25]. Yet, none of them proposes any band scheduling algorithm for multi-band system considering multiple traffic types and a band-sharing architecture. Therefore, a band-sharing architecture for multi-band Wi-Fi routers with an appropriate scheduling scheme by considering multiple data traffic types is developed in Chapter 5.

The other problem in the multi-band routers is the energy consumption because even the standby energy billing cost of a single band router is \$27 per year, which is the highest standby cost in the home gadgets, according to Ecotricity [26]. The more active bands result in the more energy consumptions. Therefore, in Chapter 6, the proposed scheduler for the band sharing architecture in Chapter 5 is improved by adding the energy awareness.

1.3 Objective and Contributions of the Research

The *objective* of this research is to satisfy end users and increase Quality of Service (QoS)⁵ by developing carrier assignment and scheduling methods for LTE-A and Wi-Fi.

The key *contributions* of the dissertation are summarized as follows:

- The carrier assignment procedure for LTE and LTE-A is explained with technical challenges.
- The strategies which are required to provide QoS in carrier assignment methods in LTE and LTE-A are discussed. Then according to the strategies, several current methods are classified to point out the advantages and disadvantages of the methods.
- Selective periodic carrier component assignment in LTE and LTE-A is proposed to increase QoS by considering performances of the methods during the periodic carrier assignment operations in addition to overall system performance.
- A new user profile strategy for carrier assignment methods in LTE and LTE-A is proposed to increase QoS. Then, the performance of user profile strategy is compared to some of the other strategies.
- The carrier aggregation architecture is applied to multi-band Wi-Fi routers, and a band-sharing architecture with a band scheduler which utilizes available resources is developed for next generation Wi-Fi routers.
- The proposed multi-band scheduler is improved based on energy awareness to decrease the energy consumption of the next generation multi-band Wi-Fi routers.

As results of the above mentioned contributions, several research papers have been published and several new ones were submitted for publication. The list of the author's publication can be found in Appendix A. Each chapter also has the related paper references at the first page footnote section.

⁵QoS refers to bandwidth, throughput, delay and energy efficiency

1.4 Organization of the Dissertation

The rest of the dissertation is organized as follows: In Chapter 2, the current developments of the carrier assignment in LTE and LTE-A are explained. In Chapter 3, selective technique for periodic carrier assignment methods for LTE and LTE-A is presented and followed by user profile strategy for carrier assignment methods in Chapter 4. In Chapter 5, a novel multi-band scheduler for multi-band Wi-Fi router is explained and the scheduler is improved based on energy awareness in Chapter 6. Finally, Chapter 7 has concluding remarks, future works and the related paper list.

Chapter 2

Component Carrier Assignment: Past and Future

The bandwidth demand for mobile Internet access is increasing with the number of mobile users. To satisfy users, Carrier Aggregation has been proposed. In Carrier Aggregation, the best available one or more component carriers of each band are assigned to each user to provide efficient services. Several methods have been reported in the literature on the component carrier assignment and have improved the performance of LTE and LTE-A systems. However, many technical challenges still exist. Therefore, in this chapter, we firstly explain the technical challenges for the carrier assignment to provide better services to users. Secondly, we review the strategies for component carrier assignment methods and then compare the current carrier assignment methods according to the strategies. Our explanations and discussions in this chapter should help researchers from all fields to understand the carrier assignment procedures with the future directions for further research in LTE-A.

2.1 Introduction

The bandwidth demand for the mobile Internet access is increasing with the number of mobile users. To satisfy users, Carrier Aggregation (CA) has been developed [17]. In CA, the system operates on multiple bands and each band is sub-divided into Component Carriers (CC). Users can get services over multiple CC from different bands. Therefore,

there are three types of CA, *Intra-band contiguous*, *Intra-band non-contiguous* and *Inter-band non-contiguous* [27], as shown in Figure 2.1.

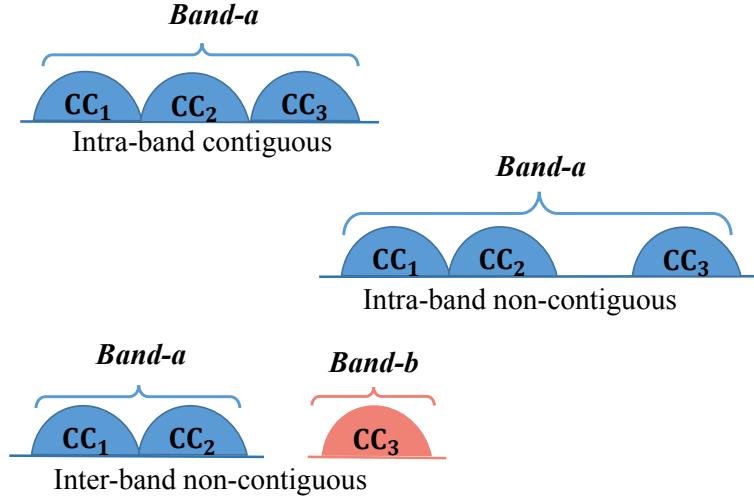


Figure 2.1: Types of Carrier Aggregation.

- **Intra-Band Contiguous:** In this type, the bandwidth is wider than 20MHz. However, it is unlikely to allocate more than 20MHz bandwidth due to the interference of carriers in lower frequencies. It can be possible in the future with new larger bands like 3.5 GHz [19, 27].
- **Intra-Band Non-Contiguous:** In this type, the multiple CCs from the same band non-contiguously assigned to users when the contiguous CCs are not available for CA [27, 28].
- **Inter-Band Non-Contiguous:** In this type, the multiple CCs from different bands are used for the communication. With this type of aggregation, performances can be improved by considering the different characteristics of the bands [19, 27].

Radio Resource Management (RRM) framework of a multi-component carrier LTE-A system is presented in Figure 2.2. Each CC has a RRM block and RRM independently operates to maintain backward compatibility so that LTE and LTE-A equipment users can

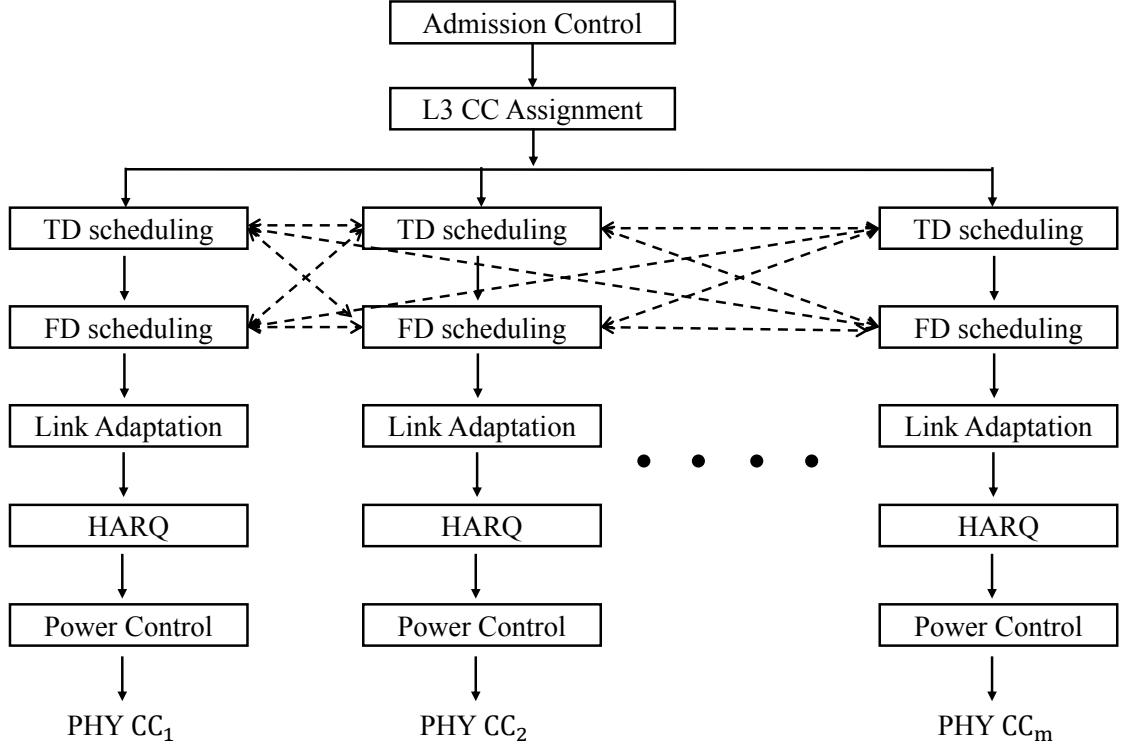


Figure 2.2: Radio Resource Management framework in LTE-A [29].

co-exist. Link Adaption (LA) and Hybrid Automatic Repeat Request (HARQ) are different for each CC. However, packets are jointly scheduled on different CCs by exchanging required information such as throughput to improve cell coverage and fairness [29].

- **Link Adaption:** One of the key issues in wireless systems is to maintain the connections of users to service stations based on channel status (such as path loss, interference). To maintain the connection and increase the service quality in terms of throughput, the link adaptation techniques are required [30–32].
- **Hybrid Automatic Repeat Request:** One of the other issue in wireless system is to guarantee the correctness of the transferred data in receiver sides. To do so, the error checking and correction methods are used. If a receiver cannot correct the errors of the received data, the same data is requested again by sending NACK (Non-acknowledgment) by ignoring any received parts of the data. This request

for retransmission is called Automatic Repeat Request (ARQ) [33]. However, in Hybrid ARQ, if the receiver cannot correct the errors of the received data, the same data is requested again by sending NACK but keeping the received parts of the data in a buffer [34, 35].

- **CC Assignment:** To use the system resources efficiently and provide better service to users, the base stations or eNodeBs (eNB) must assign one or more CCs to UEs by considering the specifications of each UE, the number of UEs in the system, CQI of carriers, etc. However, if the carrier assignment methods and techniques are not carefully designed, CCs from one band can be overloaded while CCs from the other band can be idle. Hence, the quality of services for end users can be dropped significantly. For this reason, the carrier assignment methods and preceding techniques during the carrier assignment process significantly affect systems performance [18–21]. *My research focus is mostly on the carrier assignment in RRM.*

In LTE-A, there are several proposed carrier assignment methods and their analysis [19–21]. However, further improvements are necessary because continuously increasing users' requests force the operators to manage the data traffic more intelligently [22]. Therefore, the *objective* of this chapter is to review the current developments on the carrier assignment and the challenges which arise with the new enhancements in LTE and LTE-A systems. The key *contributions* of this chapter are: (i) The current research trends with technical challenges for LTE and LTE-A are explained; (ii) The strategies which are required to provide QoS in the component carrier assignment methods are discussed; (iii) The current carrier assignment methods are reviewed and classified according to the strategies which are considered by the methods while assigning carriers to users.

The rest of this chapter is organized as follows: In Section 2.2, the current trends and technical challenges of the carrier assignment are discussed. In Section 2.3, we explain

the carrier assignment procedure and the used strategies to provide QoS. Then the current carrier assignment methods are reviewed and classified according to the strategies. Finally, Section 2.4 has the summary for this chapter.

2.2 Carrier Assignment Challenges

In this section, the current research trends for LTE and LTE-A are explained with the technical challenges.

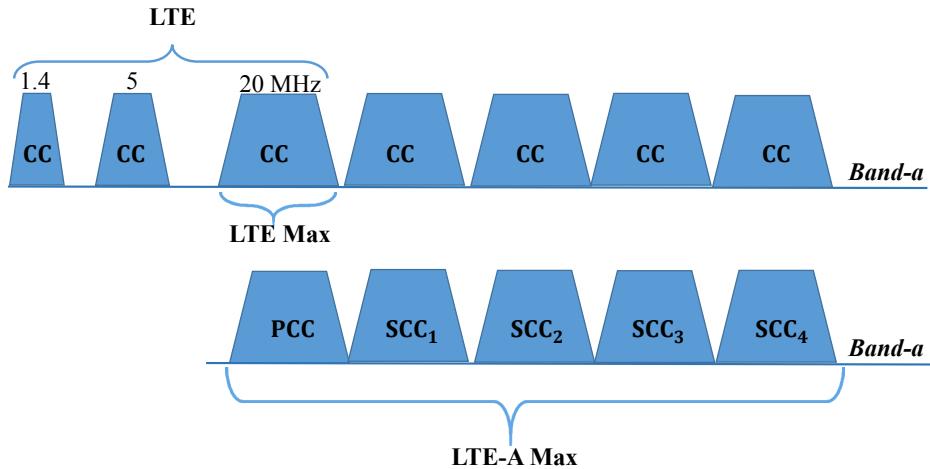


Figure 2.3: Component carrier allocations for LTE and LTE-A.

2.2.1 Primary and Secondary Component Carrier Assignment

In CA, CCs are classified as primary and secondary carriers. However, firstly Primary Component Carrier (PCC) is determined and then Secondary Component Carriers (SCCs) are activated based on QoS requirements, service expectations of users, etc. It is important to note that a PCC of a user can be different than a PCC of another user [27, 36].

Figure 2.3 shows the bandwidth capacities of LTE (Rev. 8/9) and LTE-A (Rev. 10 and above) type users. While LTE type (Rev. 8/9) users can only connect one CC, LTE-A type (Rev. 10 and above) users can connect up to five CCs and one of them must be PCC.

PCC can only be updated during handover or cell reselection for each user while SCCs can be activated or deactivated at any time. If a SCC or all SCCs need to be updated due to path loss, user requirements, channel conditions, etc. for a user, PCC continues serving the user. However, while PCC is updated, SCCs may or may not also be updated because the RRConnectionReconfiguration IE may contain a list of new SCCs which are same or different sets of carriers. Therefore, packet transfer can be interrupted during the PCC reassignment. To overcome the packet interruption caused by PCC updates, PCC is generally selected from the band which supports the highest coverage area and has the highest CQI [27]. However, there are some periodic carrier assignment methods do not distinguish between PCC and SCCs while updating them [22].

PCCs of users are the main carriers for the communication and SCCs are the carriers which boost the data rate. Therefore, PCC and each SCC should carefully be selected for each user. However, the coexistence of heterogeneous and small cell networks increases the frequency of PCC reselection and activation/deactivation of SCCs thus, the frequency of data transfer interruptions. Therefore, more investigations are needed to improve PCC assignment and reselection process, and activation/deactivation of SCCs in different networks to overcome the interruption problem. *In my research, I focus on general PCC and SCCs assignment challenge.*

2.2.2 Frequency Division Duplex (FDD) - Time Division Duplex (TDD) CA

One of the enhancement in LTE-A is that FDD and TDD carriers can be aggregated together by selecting PCC from FDD and SCCs from TDD or vice versa [27, 37–39]. The typical benefits of FDD-TDD CA are that more resources will be available by combination of TDD and FDD resources, load balancing between FDD and TDD can be handled

efficiently and configuration of downlink and uplink¹ is more flexible on TDD [40] because one carrier can be used as downlink and uplink in TDD but it is not possible in FDD. QoS is improved further by FDD-TDD CA. However, because of different natures of FDD and TDD, it is hard to fully utilize the resources of TDD and FDD [41].

2.2.3 Dual Connectivity

With release 12 of LTE-A by 3GPP² [42], dual connectivity is supported by LTE-A. Dual connectivity means that a user simultaneously receives services by using multiple base stations (such as multiple eNBs which are called as Master eNB and Secondary eNB) [27,43,44]. The benefits of using dual connectivity are to increase throughput especially for cell edge UEs³, enhance the mobility robustness and decrease overheads because of frequent handovers [45]. Such improvements take QoS a huge step ahead. Despite such improvements, several challenges should be addressed regarding CCs management, buffer status report, user transmission power management, Radio Resource Control (RRC) signaling from/to eNBs and user power saving operations depending on deployment scenarios [45].

2.2.4 Licensed Assisted Access (LAA) / LTE in Unlicensed Band (LTE-U)

With 3GPP Revision 13 in LTE-A, LTE-A is enhanced to use unlicensed bands (e.g. 5GHz) [46–48]. By this enhancement, PCC is selected from licensed bands and SCCs are selected from unlicensed bands to boost downlink or uplink data rate. The main

¹Shortly, uplink carriers are the ones which used to upload data from UEs to eNBs and downlink carries are the ones which used to download data from eNBs to UEs.

²The 3rd Generation Partnership Project unites [Seven] telecommunications standard development organizations (ARIB, ATIS, CCSA, ETSI, TSDSI, TTA, TTC), known as "Organizational Partners" and provides their members with a stable environment to produce the Reports and Specifications that define 3GPP technologies.

³Edge users are located near the boundary of the base station.

challenges in such environment is that unlicensed bands may be used by different networks [49]. Therefore, a band must be shared. However, sharing a band can be unfair for the users of other networks because like Wi-Fi uses Orthogonal Frequency Division Multiplexing (OFDM) unlike LTE because LTE uses Orthogonal Frequency Division Multiple Access (OFDMA)⁴ for data transfers and if any collision is detected by Wi-Fi, it stops the data transfers for a certain time. Therefore, Wi-Fi users cannot receive services because of the collision and that is unfair for Wi-Fi users [50]. To overcome this problem, Listen-Before-Talk (LBT)⁵ is integrated to 3GPP Revision 13 [27]. There are also other solutions like Dynamic Frequency Selection and Transmission Power Control for a fair and flexible coexistence of LTE-A with the other networks. However, there are still arguments regarding the effects of such control mechanism on the other network users.

2.3 Resource Management: PCC and SCCs Assignment

As explained in Section 2.2.1, my research is related to PCC and SCCs assignment. Therefore, in this section, the resource management is explained based on PCC and SCCs assignment.

The resource management is the most important factor which affects QoS in the mobile network because users have to share the resources (here resources are primary and secondary component carriers) to receive services. Therefore, in this section, the carrier deployment scenarios [42] and the strategies which have been used for the carrier assignment methods are discussed. Then the current carrier assignment methods are reviewed and classified according to the strategies to investigate the limitations of the methods.

⁴OFDMA is multi-user version of OFDM.

⁵LBT is designed and enforced by European Union regulations and enabling a flexible and fair coexistence between different systems. LBT is not enforced by some countries such as US, South Korea etc. Therefore, operators in US and South Korea can deploy LTE-U before actual LBT integration into LTE-A.

2.3.1 Carrier Deployment Scenarios

Figure 2.4 shows the carrier deployment scenarios for a multi-band architecture in mobile networks [42]. The behaviors of the carrier deployment scenarios can be summarized as follows:

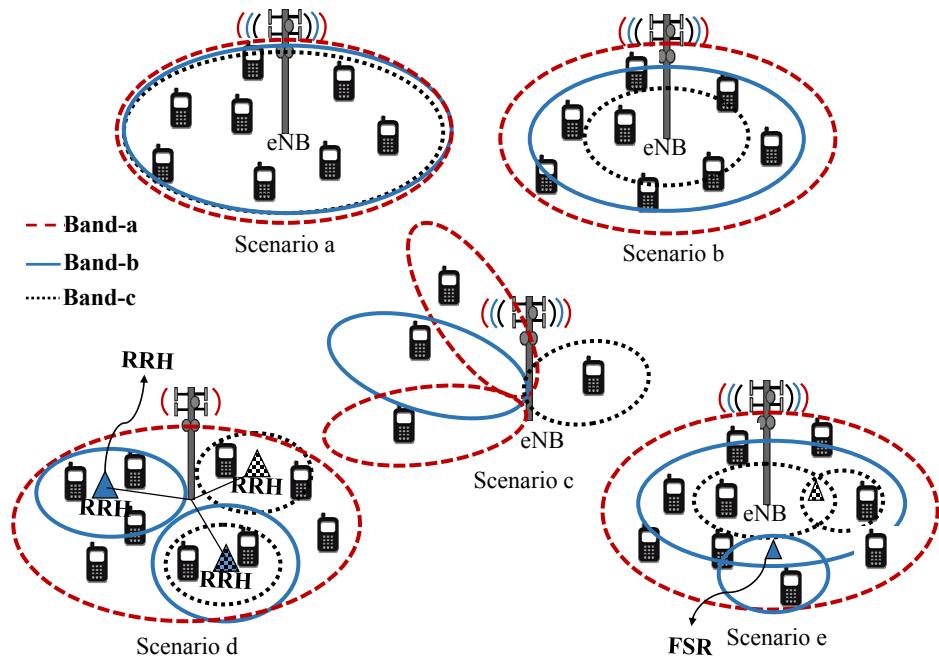


Figure 2.4: Carrier deployment scenarios for multi-band [42].

- **Scenario a:** All bands have same coverage and any of the carriers from any band can be assigned to users.
- **Scenario b:** An eNB serves users by using multi-band and each band has different coverage. Therefore, carriers from different bands can be assigned to only the users which are in their coverages.
- **Scenario c:** All bands can have same or different coverages but the antenna of one band is directed to the boundaries of the other band to increase the coverage and service quality.

- **Scenario d:** An eNB provides service in the macro coverage by using one of the multiple bands. However, there are Remote Radio Heads (RRH) which uses different bands (higher frequency such as 60GHz) than the eNB to increase the throughput at hot spots.
- **Scenario e:** All bands can have same or different coverages but the coverages of the bands are extended by using Frequency Selective Repeaters (FSR). Therefore, users can get service by connecting the bands which are extended by FSR.

It is also possible to use hybrid models of the above scenarios to increase the efficiency. For example, Scenario *c* and Scenario *d* or Scenario *c* and Scenario *e* can be combined, and the bands (a higher frequency such as 60GHz) can serve users with beamforming to increase the directional coverage while the other bands (lower frequency such as 800MHz) can manage the mobility by providing services in all covered areas. In fact, the best scenario can be deployed by considering the number of different factors such cost, efficiency, population, etc. in the area.

2.3.2 Component Carrier Assignment Strategies

The component carrier assignment is one of the crucial parts in the mobile network and is determined after the admission control in Layer 3 [19, 51] (see Figure 2.2). According to the direction of data traffic flow, the carrier assignment can be grouped as *downlink* or *uplink* [52, 53]:

- **Downlink (DL):** Downlink carrier assignment methods assign carriers to users when users download data to their equipment. The aim of the methods is to optimize bandwidth usage.
- **Uplink (UP):** Uplink carrier assignment methods assign carriers to users when users upload data from their equipment to eNB. The main purpose of the methods is to optimize not only bandwidth usages but also energy consumptions of equipment because of battery capacities of devices.

According to the number of CCs in DL and UP, carrier aggregation is called Symmetric or Asymmetric allocations as shown in Figure 2.5 [36].

There are several channel formats in LTE and four of them are shown in Figure 2.5. Physical Uplink Control Channel (PUCCH) carries Uplink Control Information (UCI) which consists of CQI, ACK/NACK for downlink information and Scheduling Request (SR) [54–56]. CQI informs the eNB about the current channel condition according to a UE. CQI can also include MIMO related feedback if MIMO transmission is used. ACK/NACK is the part of HARQ. SR requests resources from an eNB to transmit data.

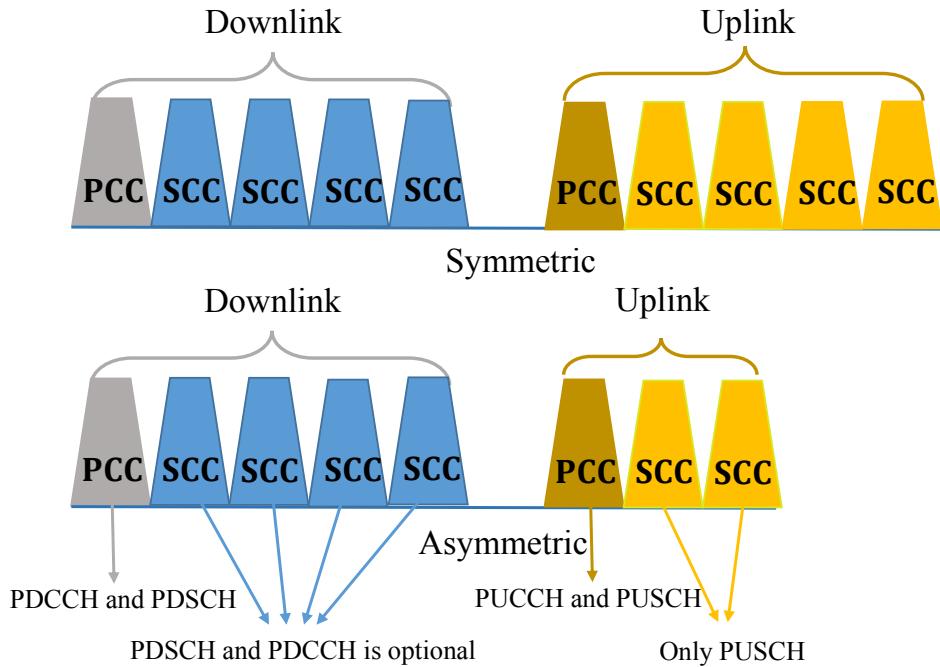


Figure 2.5: Symmetric and asymmetric allocations.

Physical Uplink Shared Channel (PUSCH) carries the data traffic of users and signaling messages in the uplink [57]. Physical Downlink Control Channel (PDCCH) carries Downlink Control Information (DCI) which consists of channel quality requests, power control commands, resource allocations with the corresponding modulation and coding schemes and some other commands [58]. Physical Downlink Shared Channel (PDSCH)

carries user data traffic, signaling messages, page messages and system information in the downlink [59, 60].

According to the reasons of the carrier reassignment, the carrier assignment methods can be grouped as *mandatory or periodic carrier assignments* [22]. The mandatory carrier assignment is because of the mandatory reasons such as path loss, CQI changes, etc. and the periodic carrier assignment depends on time and periodically update carriers of each user. In both types, the methods can be categorized under seven groups based on the following strategies; *Random, Load Balancing, Channel Quality Indicator, Fairness, Traffic Type, Energy Efficiency and Mobility*.

2.3.2.1 Random

Randomness or Random methods [61] are the basic and easily implementable carrier assignment methods. Random methods randomly assign some of the available carriers to users by ignoring user requirements, QoS requirements, CQI and load balancing. Hence, R is considered as the simplest strategy and cannot guarantee users' satisfaction.

2.3.2.2 Load Balancing

Load Balancing (LB) [51] methods assign users to the carriers according to their loads. If there are multiple available carriers, some of them are randomly selected from the available carriers. Therefore, LB methods well balance user loads across carriers in short and long terms. It is important to note that load balancing is also part of the bandwidth arrangement [19]. Load balancing techniques can be grouped as active and idle modes [62]:

- **Active Mode:** The system has information about the carrier loads and active users. Therefore, the load balancing is decided based on the number of the active users. Currently, the proposed methods are mostly based on Active Mode because of its simplicity.

- **Idle Mode:** This model is harder than Active Mode because load balancing is decided based on not only the number of the active users but also the idle users. However, the system has only information about the active users or the users that send control messages. In order to consider idle users, various adaptation functions are used for load balancing optimization.

However, only considering load balancing fully optimizes neither the system performance nor user Quality of Experience (QoE)⁶ because the characteristic of the bands (e.g., comparing 2.4GHz to 60GHz, 2.4GHz serves larger coverage areas), the carrier conditions according to position of users and the expectation of the users (e.g., battery state information or data traffic types) are not taken into account.

2.3.2.3 Channel Quality Indicator

The carriers are assigned to the users after gathering the feedback about the channel conditions from the users [64]. Therefore, QoE is expected to be higher in this type of the methods because of dynamic carrier assignment. There are two type feedback obtaining techniques, *full and partial* [64]:

- **Full Feedback:** The system requires to have all information about channel conditions including all sub-bands (sub-bands mean not only CCs but all Radio Resource Block). According to full feedback, the carriers are assigned to the users.
- **Partial Feedback:** The system requires only limited information about channel conditions. According to the partial feedback, the conditions of some channels are obtained by feedback from the selected users and then, the other channel conditions are estimated by using several algorithms to decrease signaling overheads. According to the obtained and estimated channel conditions, the carriers are assigned to users.

⁶QoS is generally defined in terms of the network delivery capacity and resource availability, but QoE is generally defined in terms of the satisfaction of users. Although it is well-known that there is a strong relation between QoS and QoE, some of QoS measurements such as delay does not affect user decisions as much as the other measurements of QoS such as bandwidth for some services [63].

The full feedback system is expected to have a better performance because of the awareness of instant conditions of the channels. However, as shown in [64], the partial feedback system can be as efficient as the full feedback system because of lower feedback overheads.

2.3.2.4 Fairness

Fairness is generally accepted as a part of the packet scheduler rather than the carrier assignment [51]. However, we have also mentioned fairness because existence of LTE and LTE-A user types and their CCs capacities affect QoS and QoE. For example, without fairness, the experienced service quality by the users who have LTE-A type devices can be much higher than the experienced service quality by the users who have LTE type devices. This is because LTE-A type equipment can connect more than one CC to receive services while LTE type equipment connects only one CC to receive services.

We can also group fairness as full and partial fairness according to Independent-Packet Scheduling and Cross CCs-Packet Scheduling [51]:

- **Partial Fairness:** The aim of this type of fairness is to provide equally likely services to same types of users. Therefore, the experienced performance such as delay and throughput of the same type users are similar. This type of fairness is simple because all same types of users are expected to have similar properties in terms of capacity, devices types, etc.
- **Full Fairness:** The aim of this type of fairness is to provide equally likely services to all types of users. Therefore, the experienced performance such as delay and throughput of all type users are similar. This type of fairness is more complex because of the existence of a various number of types of equipment and users.

2.3.2.5 Data Traffic Type and Usage

The data traffic types in wireless networks are generally classified as real-time, non-real-time [65] and signaling traffic such as binding update for mobile IP management in the mobile environment [66,67]. Some of the traffic types (such as real-time) have strict delay constraints [68, 69]; the other signaling traffic is required for the mobility management of the mobile users [70, 71]. In addition, each user has different expectations from the different traffic types. For example, while user X waits ten seconds for video loading, user Y waits only five seconds. By considering the traffic requirements and the expectations of the users, some carrier assignment methods [72] have been developed. In this type of methods, each carrier (or carrier groups) with its bandwidth is determined for each application such as video application according to instant data rates. Then, users which use the same type of application are assigned to the reserved carriers for those types of applications. According to [72], this model improves the QoS and QoE. However, such arrangement may cause data congestion on some carriers because many users can desire to use the same types of applications at the same time.

2.3.2.6 Energy Efficiency and Number of CCs

In LTE and LTE-A systems, the LTE-A type equipment can connect up to five CCs while LTE type device can connect only one CC. For example, there are several methods which assign all CCs for each LTE-A type equipment [73] and it has been shown that using all CCs can improve performance for LTE-A type equipment. However, assigning all available CCs increases power consumption [19]. Therefore, using all CCs can also decrease user QoE because of short battery lives of UEs. To provide better services, not only the number of required CCs needs to be carefully determined but also the battery state information of UEs should be considered to increase the efficiency of energy and resource usages [74].

2.3.2.7 Mobility

User mobility is one of the important factors in LTE and LTE-A because user movements may cause path loss, cell reselection, handover, etc. Therefore, the mobility strategy is used to assign carriers to eliminate the frequent handover and cell reselection by guessing the user directional movement according to the vectorial directions to assign carriers [75]. However, such calculation is extremely costly if hundreds of users are in the system.

2.3.3 The Carrier Assignment Methods

Several carrier assignment methods have been proposed and analyzed in the literature [22, 51, 72, 73, 75–84]. In [51, 76], Round Robin, which is based on the load balancing strategy, and Mobile Hashing, which is based on the randomness strategy, methods have been investigated. In [77], firstly, Channel Quality Indicator (CQI) rates from all users for each component carrier are measured, then according to the highest rate, the carriers are assigned to users. In [72], a service-based method is proposed by giving the priorities for some of the traffic types while assigning the carriers to users. In [78], absolute and relative carrier assignment methods are proposed according to a predetermined CQI threshold and CQI of PCC, respectively. In [79], G-factor carrier assignment method is proposed by considering load balancing for non-edge users and better coverages (based on CQI) for edge users. Edge users are located near the cell edge in LTE. In [80], firstly, bands of pico and macro cells are decided according to interference, then beamforming is used to provide services to each user. In [81], a self-organized method, which assumes availability of CQI for each resource block to avoid interference, is proposed. A resource block is the smallest unit of resources that can be allocated to a user. In [82], the least user loaded carriers with highest CQI are considered to assign carriers to users. In [75], the mobility of users is estimated in real time while assigning carriers to users in order to decrease carrier reselection and handover. In [73, 83, 84], uplink carrier assignment methods have been proposed by considering a ratio function, traffic type and CQI to increase throughput while sending data from users to eNBs. While the aim of uplink carrier assignment

is to optimize the bandwidth and power limitations, downlink carrier assignment aims to optimize only bandwidth.

2.3.3.1 Packet Scheduling and CQI Measurement

In addition to carrier assignment methods, methods to measure CQI and methods for packet scheduling have been proposed for LTE and LTE-A [64, 85–89]. In [64, 85–87], the methods are proposed to measure CQI. In [88, 89], full or partial feedbacks related to CQI are used to find the best available resource blocks in carriers for each user. In [90], the service-based methods are proposed by giving priority for some services while assigning resource blocks to users. In [91, 92], multiple resource blocks are assigned to users in such a way that the delay is decreased. In [93, 94], the uplink resource block scheduling has been proposed by considering a ratio function of CQI and data traffic.

2.3.3.2 Comparison of the Methods

The current component carrier assignment methods are categorized under R, LB, CQI, Device Types (Backward Compatible), Number of CC (Number of used carriers for each type of users), Scenario (specifically for a given scenario type), CA (specifically for given aggregation type), Fairness, Downlink/Uplink and Mobility as shown in Table 2.1. *The current methods mostly focus on load balancing, channel conditions and characteristic of bands by generally assuming the availability of the full channel information. The methods have also been evaluated according to the overall throughput, drop rate, delay, and energy efficiency parameters to show QoS levels. However, partial CQI, full fairness and "e" scenario cases are generally ignored.*

Table 2.1: The comparison of the methods.

Methods	R ¹	LB ²	CQI ³	Device Types	Number of CC	Scenario	CA ⁴	Fairness	Down/Up	Mobility
[22]		✓	Full	LTE-A	At most all	a,b	All	✗	Downlink	✗
[51]	✓	✓	Full	LTE- A/LTE	All and 1	a,b	NE ⁵	Full	Downlink	✗
[72]		✗	LTE-A		At most all	a	All	✗	Downlink	✗
[73]	✓	✗	LTE- A/LTE		All and 1	a	All	✗	Uplink	✗
[75]	✓	✗	LTE-A		At most all	a,b	IANC ⁶ /IENC ⁷	✗	Downlink	✓
[76]	✓	✓	Full	LTE- A/LTE	All and 1	a,b	IANC/IENC	Partial	Downlink	✗
[77]	✓	✓	Full	LTE-A	At least 2	a,b	IANC/IENC	Partial	Downlink	✗
[78]		✓	Full	LTE- A/LTE	At most all and 1	a,c	NE	Full	Downlink	✗
[79]	✓	✓	Full	LTE- A/LTE	All and 1	b	IANC	✗	Downlink	✗
[80]		✓	Full	LTE-A	At most all	c,d,e	NE	Full	Downlink	✗
[81]	✓	✓	Full	LTE-A	At most all	c,d	All	Partial	Downlink	✗
[82]		✓	Full	LTE-A	All	a	IANC/IENC	Partial	Downlink	✗
[83]		✓	Partial	LTE-A	At most all	b	All	✗	Uplink	✗
[84]		Full	LTE-A	At most all	a,b	IAC ⁸	✗	Uplink	✗	

¹ Random; ² Load Balancing; ³ Channel Quality Indicator; ⁴ Carrier Aggregation;
⁶ Intra-band non-contiguous; ⁷ Inter-band non-contiguous; ⁸ Intra-band contiguous.

⁵ Not Explained;

2.4 Summary

In this chapter, the challenges of PCC and SCC assignment, Frequency Division Duplex - Time Division Duplex carrier aggregation, dual connectivity and LTE-U which significantly improve the quality of services were explained. Moreover, the strategies (randomness, load balancing, channel quality indicator, fairness, traffic type, energy efficiency and mobility) for primary and secondary component carrier assignment methods were discussed. Additionally, the current methods were reviewed and compared based on the strategies. The current methods mostly focus on load balancing, channel conditions and characteristic of carriers by assuming availability of full channel information but partial channel information and full fairness are generally ignored. Moreover, the current methods are analyzed and evaluated according to the overall system performances.

In the next chapter, rather than focusing on the overall performance of the methods, the behaviors of the methods during the carrier assignment operations in addition to their overall performances are considered to improve the performances of the periodic carrier assignment methods.

Chapter 3

Joint and Selective Periodic Component Carrier Assignment for LTE and LTE-A

In the previous chapter, the carrier assignment strategies and methods with their analysis were briefly explained. In this chapter, firstly, the packet drops and delay which are experienced by users during the carrier assignment operations are investigated because the previous works only analyze the overall performance of the system by neglecting their behaviors during the carrier assignment operations. The other limitation of the previous works is that the data transfer is interrupted during the periodic component carrier assignment operations which can decrease the performance of the system. Therefore, selective periodic component carrier assignment technique, which allows continuous data transfer during the periodic carrier assignment operations, is proposed by considering the behaviors of the carrier assignment methods during the operations. Selective technique is integrated into four component carrier assignment methods, Least Load, Least Load Rate, Random, and Channel Quality, to observe the performance improvements. Results show that the proposed technique increases throughput ratio up to 18% and decreases average delay up to 50%.

Results presented in this chapter have appeared as two research papers in IEEE Global Communications Conference (GLOBECOM) [95, 96].

3.1 Introduction

The previous works on the carrier assignment improved the performance of LTE-A. However, one of their limitations is that the overall system performance is analyzed in order to evaluate the performance of the carrier assignment methods, yet the behavior of the system such as packet drops and delay metrics during the carrier assignment process are ignored. However, delays and packet drops can occur during the carrier assignment operations because the carrier assignment operations could consume considerable amount of time based on the selected carrier assignment method due to the required time for CQI feedback, QoS measurement, queue migration, process time, etc. For example, if a method is based on CQI feedback, then it increases delay and packet drops of packets waiting for services during the carrier assignment operations, thus packet retransfer rate increases.

In [22], a Periodic Component Carrier Assignment (pCCA) method is proposed and the carriers are periodically assigned to each user in specified time interval. In the periodic carrier assignment, CCs of all users are updated periodically in addition to the mandatory carrier assignment. As presented in [22], the periodic carrier assignment method significantly improves the performance of LTE-A systems because the carriers are frequently updated for all users according to CQI, etc. However, one known limitation of such system is the interruption of the data transfer during the periodic carrier assignment operations. This interruption is due to reassigning all carriers of users at the same time in the periodic carrier assignment [22]. For example, while UE_i is leaving from *Band-c* coverage to enter *Band-b* coverage, simultaneously reassigning all CCs to UE_i causes delay for packets of UE_i which are waiting for service. However, reassigning all CCs may increase the performance in a long term if CQI of the new CCs are higher than the CQI of the previous CCs (We called the policy of the reassignment of all CCs as Joint Periodic Component Carrier Assignment Technique (j-pCCA).). On the other hand, updating the CCs in *Band-c* by allowing the CCs in *Band-b* or *Band-a* to continue serving UE_i , prevents packets experiencing delay or drop. Therefore, the *aim* of this chapter is to improve the

performance of j-pCCA by eliminating the frequent packet transfer interruptions during the carrier assignment operations.

The *objective* of this chapter is to consider the packet drops and delay, which are experienced by users during the periodic carrier assignment process, and to propose selective periodic carrier assignment technique (s-pCCA) to increase the performance of the periodic carrier assignment methods in LTE and LTE-A systems. The key *contributions* of this work are as follows: (i) Selective periodic carrier assignment technique is proposed; (ii) The system models for joint and selective techniques are explained by using Joint (JQS) and Disjoint (DQS) Queue Scheduler [97]; (iii) Analytical expression for joint and selective techniques are derived by using M/M/m/N queue model according to JQS and DQS; (vi) Joint and selective techniques are compared by using four carrier assignment methods, Least Load (LL), Least Load Rate(LR), Random (R), and Channel quality (CQ)¹ within an extensive simulation. *Results* show that the proposed selective technique increases the throughput ratio up to 18% and decreases the average delay up to 50% in comparison to joint technique. Our proposed technique and related analysis will help service providers build efficient periodic component carrier assignment methods in order to increase throughput and decrease average delay time.

The rest of the chapter is organized as follows: In Section 3.2, the system model of the carrier assignment procedure for joint and selective techniques are explained and followed by the queuing analysis of the both techniques in Section 3.3. Simulation environments with parameters are described in Section 3.4. In Section 3.5, the simulation results are presented and analyzed. Finally, Section 3.6 has the summary for this chapter.

3.2 System Model with Joint and Selective Techniques

Figure 3.1 demonstrates a simple model of the carrier assignment and packet scheduler. There are n number of users and each user can only connect up to m number of CCs.

¹Detailed information about the methods is given in Section 3.2.3

Currently, LTE-A system can only support up to five simultaneous CCs connection for each user providing IMT-A level service [17]. One of the CCs is PCC for uplink and downlink, and can only be updated during handover or cell reselection [17], and the rest of carriers are SCCs which are updated for each user based on CQI of channels, path loss, etc. However, as stated in [22], the periodic carrier assignment is a new method trying reassign all CCs periodically in addition to mandatory carrier assignments. Therefore, both PCC and SCCs are updated during the periodic carrier assignment operations for all users [22]. After the carrier assignment process finishes, Packed Scheduler transfers

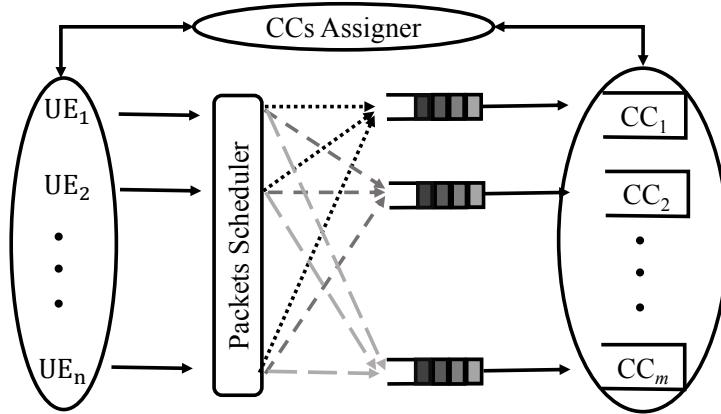


Figure 3.1: A carrier assignment model with n users and m available CCs.

packets over the selected carriers in time and frequency domains. Currently Proportional Fairness and max-min are common packet schedulers in LTE systems [22, 51].

3.2.1 Joint Periodic Component Carrier Assignment (j-pCCA)

Mandatory carrier assignment methods allocate users to carriers based on the mobility of users (including path loss, connection problems, low CQI, etc.). Therefore, when UE_i moves from one position to another position, uplink and downlink carriers are updated to maintain UE_i connection. On the other hand, the periodic carrier assignment allocates users to carriers based on time and updates the carriers in pre-specified time intervals [22]

regardless of the mandatory changes. During j-pCCA, all carriers are simultaneously updated for all users; packet transfers are thus interrupted. After j-pCCA is completed, the packet transfers are restarted.

3.2.2 Proposed Selective Periodic Component Carrier Assignment (s-pCCA)

As explained in Section 3.2.1, the disadvantage of joint technique is that simultaneous reassignments of all carriers for all users result in the packet transfer interruptions. In order to provide better service, we have proposed a novel s-pCCA to solve the disadvantage of joint technique. In selective technique, only the selected carriers of users are periodically updated. However, it is possible to update all carriers if it is required.

For selective technique, the selection algorithm decides which CCs must be updated for users. In order to make decision on CCs, CQI threshold is used in the selection algorithm. Therefore, selective technique depends on three crucial factors which are the selection algorithm, time and the strategy of the carrier assignment method. For example, LL method with selective technique is processed as follows for each periodic time:

- The threshold of CQI is predetermined for the selection algorithm. Here, the highest possible CQI is selected as a predetermined threshold for s-pCCA but the threshold can be dynamic according to user profile².
- Partially or fully CQI feedback is obtained to measure the quality of the carrier for each user. Note that, although CQI is low, the channel can transfer only a limited number of packets.
- The carriers, which have lower CQI than the predetermined threshold, are selected to be updated for each user.

²Detailed information related to user profile is given in Chapter 4

- Selective technique firstly finds the available carriers for each user according to CQI of carriers and their loads in terms of the active number of users. The new carriers should currently load by a lower number of users comparing to the other carriers and the qualities of the new carriers should be equal or higher than the threshold. It is very important to note that the number of the new carriers may not be equal to the number of the previous assigned carriers for each user. To make them equal, more carriers, which have a lower number of active users, are assigned. For example, assume that UE_i receives data by using C_1 , C_2 , and C_3 component carriers and CQI of C_1 and C_2 are lower than the threshold. Therefore, selective technique chooses C_1 and C_2 to update for UE_i. However, selective technique only finds CQI of C_4 is equal or higher than the threshold from all available CCs for UE_i. Therefore, LL method with selective technique assigns C_4 and the CC, which is loaded by the least number of active users.
- To increase the efficiency and QoS, the packet transfer priority is given to the CC, which has the highest CQI.

Similarly, LR, R and CQ methods with selective technique are processed as above except that the strategies of carrier assignment methods. The method details are explained in Section 3.2.3.

3.2.3 Methods

To analyze the impacts of joint and selective techniques on the carrier assignment, four different carrier assignment methods are used. The methods are based on random, load balancing and CQI and they are Random (R)³, Least Load (LL), Least Load Rate (LR) and Channel Quality (CQ). Those methods are selected for test cases because of common usage in the literature and the different properties are considered while assigning the carriers to UEs.

³In Chapter 2, Random is also shown as a strategy. However, here we called the method name as Random to make each method name be easily understandable by readers.

3.2.3.1 Random (R)

R method is one of the well-known methods in the literature [51, 61]. However, R method ignores QoS requirements of each user and CQI of channels. In this work, R method assigns carriers to users according to Java Random Generator and Java Random Generator is based on Uniform Distribution. Therefore, R randomly selects available carriers for each user but it only well balances user loads across carriers in long term.

3.2.3.2 Least Load (LL)

LL method is also one of the well-known methods in the literature [51]. LL assigns the carriers to users according to load balancing strategy by selecting the least loaded carriers thus, it well balances user loads across the carriers in short and long terms [51]. LL method also ignores QoS requirements of each user and CQI of the carriers. It is important to note that ignoring CQI does not mean the performance of LL method is lower than other methods.

3.2.3.3 Channel Quality (CQ)

CQI can vary according to the positions of UEs because of obstacles and distances. Therefore, there are several versions of CQ methods like [78]. In this chapter, CQ method assigns the carriers to users by selecting the carriers which have the highest CQI [88] and it is similar to Relative method in [78]. Because of only considering CQI, user loads and QoS requirements of users are ignored.

3.2.3.4 Least Load Rate (LR)

LR method assigns the carriers to users by selecting the highest rate which is measured by using the total capacity in terms of the bandwidth, the number of users and CQI for each

carrier. The rate is measured as in [77] but instead of considering the queue length⁴, we have considered the number of users in each carrier as follows:

$$\text{Rate} = \frac{\text{CQI of carrier} * \text{Bandwidth of carrier}}{\text{The number of users on carrier}} \quad (3.1)$$

3.3 Analysis

In this section, the analytical expressions of the performance metrics such as the drop probability, throughput, and average delay will be derived for joint and selective techniques during the periodic carrier assignment operations by using queuing theory for Joint and Disjoint Queue Scheduler [97]. However, Disjoint Queue Scheduler is used for the simulation because Disjoint Queue Scheduler is more realistic than Joint Queue Scheduler [19].

Packet schedulers enqueue an arrived packet which is requested by a user to one of the assigned CCs. During the joint periodic carrier assignment operations, packet transfers of UE_i are terminated all the time. However, during the selective periodic carrier assignment operations, packet transfers of UE_i are terminated if all the carriers are updated or PCC needs to be updated (if PCC is updated then all the carriers may be updated). Therefore, there are three cases in the system for joint and selective techniques:

- **Case 1:** PCC is updated, therefore SCCs are updated.
- **Case 2:** All carriers are updated.
- **Case 3:** SCCs are updated but no needs to update PCC.

It is worth noting that if it were possible to change one of SCCs as PCC when PCC is required to be updated in LTE-A, there would be four cases. Simply, Case 1 would be divided to two cases: Case 1-a: There is at least one SCC, which is not required to be

⁴The queue length is considered in packet scheduling rather than carrier assignment for all methods.

updated, can be altered as PCC. Case 1-b: There is no such SCC, therefore all the carriers are updated. It is important to note that while PCC is updated, SCCs may not be updated because the RRConnectionReconfiguration IE may contain a list of new SCCs which are same or different sets of carriers.

The performance metrics of joint and selective techniques are same for Case 1 and Case 2. Hence, only Case 3 is explained to distinguish differences between joint and selective techniques. During the periodic assignment operations (Case 3) in joint technique for UE_i, the packet transfer operation is as follows: (i) Packet transfer is interrupted for the user; (ii) All CCs of the user are updated; (iii) Packet transfer is restarted for the user over the new carriers. On the other hand, during the periodic carrier assignment operations (Case 3) in selective technique, the process is as follows: (i) For all users, some carriers (CCs) are selected to be updated according to the selective algorithm (here, it is based on CQI); (ii) Packet transfer is only interrupted on the carriers which are needed to be updated for each user; (iii) The new carriers are assigned to users; (iv) Packet transfer is started on the new carriers for the users.

In joint periodic carrier assignments, all the carriers are updated for UE_i. Therefore, the system is not in steady state because the service rate is zero during the carrier assignment operations. Hence, we only mention the possibilities of the performance of joint technique. On the other hand, we approximately derive the performance metrics of selective technique.

3.3.1 Notations

The notations used for the analysis in the rest of this chapter are listed in Table 3.1.

3.3.2 Disjoint Queue Scheduler of j-pCCA and s-pCCA for Downlink

Figure 3.2 illustrates the downlink process for n users with one CC. The queuing scheduler is Disjoint Queue Scheduler [97]. Disjoint Queue Scheduler allows all CCs to have disjoint buffers for each user as shown in Figure 3.2.

Table 3.1: The notations for Chapter 3.

i	$\in \{1, 2, \dots, n\}$
j	$\in \{1, 2, \dots, m\}$
$Q_{CC_{ij}}$	Queue of UE $_i$ for CC $_j$ in Disjoint Queue Scheduler
Q_j	Queue of CC $_j$ in Joint Queue Scheduler
N	Size of Queues
p_k	Probability of k number of packets in the system
μ_j	Service rate of CC $_j$
λ_j	Packet arrival rate to j^{th} queue
λ_i	Packet arrival rate of UE $_i$
λ_{ij}	Packet arrival rate of UE $_i$ to j^{th} queue
δ	Average delay during the periodic carrier assignment operations
n	Average queue length during the periodic carrier assignment operations
D	Drop probability during the periodic carrier assignment operations

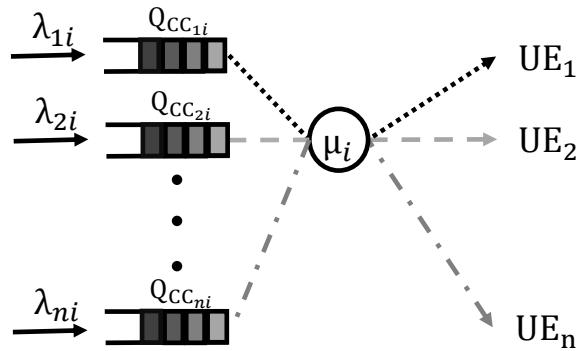


Figure 3.2: Downlink Disjoint Queue Model with n users and *one* available CC.

Downlink packet arrival rate for UE $_i$ is λ_i , each CC represented by a server and service rates of CCs are μ_j where $j \in \{1, 2, \dots, m\}$ and each buffer, Q_j , can hold at most N packets.

3.3.2.1 Assumptions

To make the model analytically tractable, it is assumed that there is only one UE in the system as demonstrated in Figure 3.3. All carriers are capable of transferring all type of packets, the queuing system is under heavy traffic flows, packet arrivals follow Poisson

Distribution, and service times for packets are exponentially distributed. Type of queue

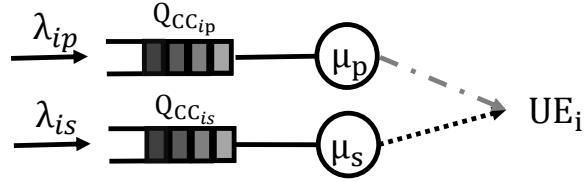


Figure 3.3: Downlink Disjoint Queue Model with one user and primary and secondary carrier queues.

discipline used in the analysis is FIFO. Bandwidth and CQI of carriers can be different, thus service rate of all servers can also be different. The assumption of one user in the system makes the derivation of analytical expressions of performance metrics simpler.

The model can be more realistic if priority based packet arrivals can be considered. In such system, packets are classified according to their priorities, then the priority queue system will be used to derive of analytical expressions of performance metrics. In Chapter 5, priority based packet arrivals are already considered while deriving the analytical expressions. Therefore, it is assumed that the system has one user without packet classification in this model. Moreover, assuming existence of more than one user in the system will not affect realism of the system model while deriving of analytical approximations because the arrival rate (λ) can be considered to represent arrival rates of multiple users rather than one user.

3.3.2.2 Performance Metrics

In this subsection, we approximately derive the drop probability, the average queue length and the average delay for joint and selective technique for Case 3 because the performance metrics of joint and selective techniques are same for Case 1 and Case 2. In both joint and selective techniques, min-delay scheduler is used and the system is under heavy traffic flows. Therefore, the total service rate ($\mu_p + \mu_s$) and the overall arrival rate (λ_i) can be used instead of separate analysis for the queues to approximate the performance metrics.

The drop probability of packets in the system for UE_i can be approximated using the standard M/M/1/N formula as follows [98]:

$$D_i = \begin{cases} \frac{\rho_i^N(1-\rho_i)}{1-\rho_i^{N+1}}, & \rho_i \neq 1 \\ \frac{1}{N+1}, & \rho_i(t) = 1 \end{cases} \quad (3.2)$$

ρ_i is a general term which is used in queue theory and represent division of the arrival rate to service rate. Therefore,

$$\rho_i = \frac{\lambda_i}{\mu_p + \mu_s} \quad (3.3)$$

since the arrival rate is λ_i and total service rate is $\mu_p + \mu_s$.

The average queue length for UE_i in selective technique can also be approximated by using the standard M/M/1/N formula as follows [98]:

$$n_i = \begin{cases} \frac{\rho_i - (N+1)\rho_i^{N+1} + N\rho_i^{(N+2)}}{(1-\rho_i)(1-\rho_i^{N+1})}, & \rho_i \neq 1 \\ \frac{N}{2}, & \rho_i = 1 \end{cases} \quad (3.4)$$

By Little's Law [99] and using Equations (3.2) and (3.4), average delay (δ_i) for UE_i can be written as

$$\delta_i = \frac{n_i}{\lambda_i(1 - D_i)} \quad (3.5)$$

Similarly, the drop probability (D_i) and average queue length (n_i) of selective technique can be represented by using same equations (3.2), (3.4), and (3.5) during the periodic carrier assignment process. However, selective technique may or may not interrupt the packet transfers for UE_i, the service rate will be at least μ_p and at most $\mu_p + \mu_s$. Therefore, ρ_i will be

$$\rho_i = \begin{cases} \frac{\lambda_i}{\mu_p + \mu_s}, & \mu_s \neq 0 \\ \frac{\lambda_i}{\mu_p}, & \mu_s = 0 \end{cases} \quad (3.6)$$

In other words, if there are m CCs of which v are not updated (assuming $v \leq m$ and $\text{CC}_1, \text{CC}_2, \dots, \text{CC}_v$ are not updated during the periodic carrier assignment process and CC_1 is PCC), then total service rate is $\sum_{k=1}^v \mu_k$. Therefore,

$$\rho_i = \frac{\lambda_i}{\sum_{k=1}^v \mu_k} \quad (3.7)$$

On the other hand, for joint technique, the average queue length (n_i) will be $n_i \approx N$ and $D_i \approx 1$. Therefore, the average delay (δ_i) will be $\delta_i \approx \infty$. However, because the periodic carrier assignment time duration is limited (assume τ), $\delta_i = \tau$

3.3.3 Joint Queue Scheduler of j-pCCA and s-pCCA for Downlink

Figure 3.4 illustrates a downlink process for n_i users in LTE-A in Joint Queue Scheduler. Joint Queue Scheduler allows all users to have disjoint buffers.

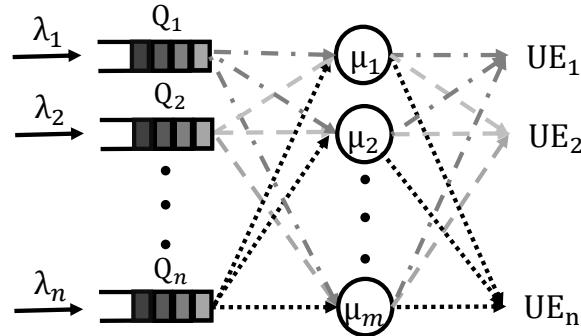


Figure 3.4: Downlink Joint Queue Model with n users and m available CCs.

The downlink packet arrival rate for UE_i is λ_i , each CC represented by a server and service rates of CCs are μ_j where $j \in \{1, 2, \dots, m\}$ and each buffer, Q_j , can hold at most N packets.

3.3.3.1 Assumptions

To make the model analytically tractable, it is assumed that there is only one UE in the system as demonstrated in Figure 3.5, all the servers are capable of serving all types of

packets, the queuing system is under heavy traffic flows, packet arrivals follow Poisson Distribution, and service times of packets are exponentially distributed. Type of queue

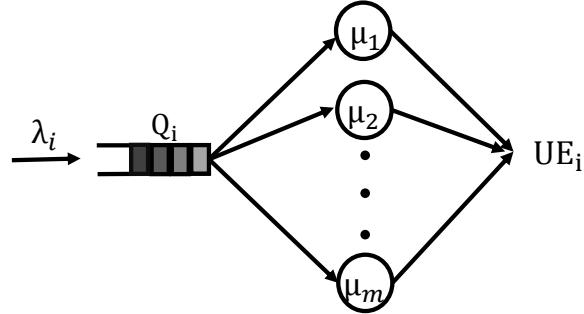


Figure 3.5: Downlink Joint Queue Model with one user and m available CCs.

discipline used in the analysis is FIFO. Bandwidth and CQI of carriers can be different, so can service rates of all servers. The model can also be more realistic if priority based packet arrivals can be considered.

3.3.3.2 State Probability

In this section, we will approximately derive the state probability in order to derive the drop probability, average queue length and average delay during the periodic carrier assignment operations for Case 3 in joint and selective techniques according to Joint Queue Scheduler [97]. The service rate of the system is state-dependent. When one packet is in the system, the service rate is $\mu_{t1} = \mu_1$ and when two packets are in the system, the service rate is $\mu_{t2} = \mu_1 + \mu_2$. The service rate of the system increases until all carriers are utilized (c carriers for UE_i). Then the total server rate of the system is fixed at μ_{tc} . It is important to note that there is at least one carrier which serves incoming traffic in selective technique which means $1 \leq c \leq m$. By using the above approach, the state transaction diagram for selective techniques can be obtained as in Figure 3.6.

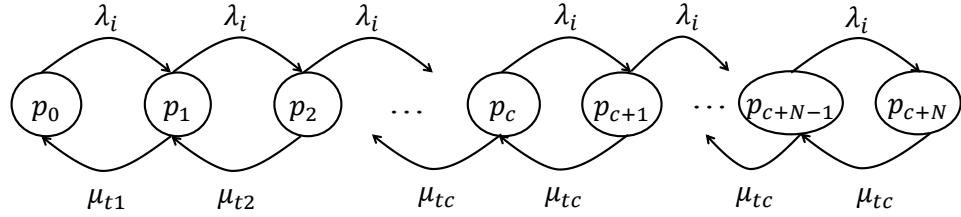


Figure 3.6: State transaction diagram for the Joint Queue Scheduler model.

The state probability equations until m_{tc} state can be written by using the state transition diagram [98, 100, 101] in Figure 3.6 as follows:

$$\begin{aligned}
 \lambda_i p_0 &= \mu_{t1} p_1 \Rightarrow p_1 = p_0 \frac{\lambda_i}{\mu_{t1}}, \\
 \lambda_i p_1 &= \mu_{t2} p_2 \Rightarrow p_2 = p_0 \frac{\lambda_i^2}{\mu_{t1} \mu_{t2}}, \\
 &\vdots \\
 \lambda_i p_{c-1} &= \mu_{tc} p_c \Rightarrow p_c = p_0 \frac{\lambda_i^c}{\prod_{v=1}^c \mu_{tv}}.
 \end{aligned} \tag{3.8}$$

The state probability equations after m_{tc} state are different because the system has only c servers for requests. Thus, the state probability equations can be written as follows:

$$\lambda_i p_{c-1} = \mu_{tc} p_c \Rightarrow p_k = p_0 \frac{\lambda_i^k}{\mu_{tc}^{k-c} \prod_{v=1}^c \mu_{tv}} \tag{3.9}$$

where $c < k < c + N$. Shortly,

$$p_k = \begin{cases} p_0 \frac{\lambda_i^k}{\prod_{v=1}^k \mu_{tv}}, & k \leq c \\ p_0 \frac{\mu_{tc}^c \rho_i^k}{\prod_{v=1}^c \mu_{tv}}, & c < k \leq c + N \end{cases} \tag{3.10}$$

where $\rho_i = \lambda_i / \mu_{tc}$. The sum of the state probabilities is equal to one. Therefore,

$$\sum_{k=0}^{c+N} p_k = 1 \tag{3.11}$$

Equation (3.11) can be divided to two parts as follows:

$$\begin{aligned} 1 &= \sum_{k=0}^{c+N} p_k \\ &= p_0 + \sum_{k=1}^c p_k + \sum_{k=c+1}^{c+N} p_k \end{aligned} \quad (3.12)$$

In order to find the state probabilities, we need to measure p_0 by substituting equation (3.10) into equation (3.12) as follows:

$$\begin{aligned} 1 &= p_0 + \sum_{k=1}^c p_k + \sum_{k=c+1}^{c+N} p_k \\ &= p_0 + \sum_{k=1}^c p_0 \frac{\lambda_i^k}{\prod_{v=1}^k \mu_{tv}} + \sum_{k=c+1}^{c+N} p_0 \frac{\mu_{tc}^c \rho_i^k}{\prod_{v=1}^c \mu_{tv}}. \end{aligned} \quad (3.13)$$

Therefore, from equation (3.14), p_0 can be simplified as

$$p_0^{-1} = \begin{cases} 1 + \sum_{k=1}^c \frac{\lambda_i^k}{\prod_{v=1}^k \mu_{tv}} + \frac{\mu_{tc}^c}{\prod_{v=1}^c \mu_{tv}} \sum_{k=c+1}^{c+N} \rho_i^k, & \rho \neq 1 \\ 1 + \sum_{k=1}^c \frac{\lambda_i^k}{\prod_{v=1}^k \mu_{tv}} + N \frac{\mu_{tc}^c}{\prod_{v=1}^c \mu_{tv}}, & \rho_i = 1 \end{cases} \quad (3.14)$$

where $\rho_i = \lambda_i / \mu_{tc}$

3.3.3.3 Drop Probability

The drop probability of the model is the final state probability which is p_{c+N} because after all servers and the queue are filled, the new arrived packet would be dropped. Therefore, by substituting $c + N$ instead of k in equation (3.10) will give the drop probability for selective technique as follows:

$$D_i = p_0 \frac{\mu_{tc}^c \rho_i^{c+N}}{\prod_{v=1}^c \mu_{tv}} \quad (3.15)$$

On the other hand, the drop probability for joint technique cannot be obtained because of unsteady state. It may be obtained by using $\lim_{\rho_i \rightarrow \infty} D_i$ or $\lim_{\mu \rightarrow 0} D_i$. Therefore, $D_i \approx 1$.

3.3.3.4 Average Queue Length and Average Delay

The average queue length and average delay can be formulated by using the state probabilities. The average queue length for $M/M/1/N$ will be as follows:

$$n_i = \sum_{k=1}^N kp_k \quad (3.16)$$

However, in $M/M_i/c/N$, the system has c servers and from the above state probabilities (equation (3.10)), n_i will be

$$n_i = \sum_{k=c+1}^{c+N} (k - c)p_k \quad (3.17)$$

After substituting equation (3.10) into equation (3.17) and simplifying by using geometric series, the following expressions for n_i is obtained as follows:

$$n_i = \begin{cases} p_0 \frac{\mu_{te}^c}{\prod_{v=1}^c \mu_{tv}} \rho_i^{c+1} \left(\frac{1-(N+1)\rho_i^N + N\rho_i^{N+1}}{(1-\rho_i)^2} \right), & \rho_i \neq 1 \\ p_0 \frac{\mu_{te}^c}{\prod_{v=1}^c \mu_{tv}} \left(\frac{N(N+1)}{2} \right), & \rho_i = 1 \end{cases} \quad (3.18)$$

By using Little's law [99] and equations (3.15) and (3.18), the average delay can be obtained as follows:

$$\delta_i = \frac{n_i}{\lambda_i(1 - D_i)} \quad (3.19)$$

In Section 3.3, we approximately derive the analytical performance metrics for selective technique and the possible performance values of joint technique for Case 3. Based on the obtained performance metrics, we can definitely sure that selective technique has improved the performance of the system during the periodic carrier assignment operations (The simulation results on delay and drop ratio during the periodic carrier assignment process in Figures 3.7 and 3.8 also verify the correctness of the improvements.). However, the overall system performance metrics can be different because the service rates of the carriers for each user depend on user positions. Therefore, we have implemented an extensive simulation to observe the overall system performances of joint and selective techniques.

3.4 Simulation of the System

A discrete event simulation has been implemented in Matlab and Java by considering the carrier assignment methods which are mentioned in Sections 3.2 and 3.2.3. The simulation parameters with additional assumptions for downlink process are explained in the following subsections.

3.4.1 Assumptions for eNBs

It is assumed that there is only one eNB which has three bands to provide service to users. The parameters of eNB and the simulation are given in Table 3.2. In the simulation, Sce-

Table 3.2: The simulation parameters.

Scenario [42]	b
Number of eNB	1
Used Bands	800MHz, 1.8GHz, 2.6GHz
Number of CCs in Each Band	4
Total Number of CCs	12
Queue Length of Each Queue	50 packets [102]
Bandwidth of CCs	10MHz
Modulations	BPSK, QPSK, 16QAM, and 64QAM
CQI	3, 5, 7, and 11
Transmission Time Interval	10ms (10ms is average, it can be more or less)
Time for CCA	20ms (at most 20ms)
CQI Threshold	The highest possible
Simulation Model	Finite buffer [103]

nario b is used to represent the general macro model. Only one eNB is considered not to deal with the handover process in case users change base stations. However, assuming one eNB does not affect the obtained results in terms of performance comparison between methods. The eNB provides service to users by using three bands similar to real case scenario and each band can have four CCs with 10MHz bandwidth. The number of CCs in each band is selected as four because LTE-A type equipment can connect at most four CCs to download data. Therefore, even if a LTE-A type user in the coverage

of *Band-a* can connect four CCs to get services similar to real case scenario. To simulate saturation of the system, a higher number of CCs are not selected. 10MHz and 20MHz bandwidths are used in LTE-A to provide IMT-A level speed [42]. BPSK, QPSK, 16QAM and 64QAM are the modulations techniques to transfer bits according to CQI in LTE systems. Therefore, to simulate those modulations, four CQI levels are used and each CQI level is modulation changing point. The average Transmission Time Interval (TTI) is 10ms for a packet (TTI can be less or more according to different packet sizes) to simulate the low and high latency requirements because the accepted TTI in LTE is 1ms to meet the low latency requirements [42]. In order to show the lowest improvements of selective technique comparing to joint technique, time for CCA is kept as 20ms and lower because the carrier assignment operations can consume considerable amount of time according to carrier assignment methods. As simulation model, finite buffer is used because finite buffer simulation well presents the reality comparing to full buffer simulation [103].

3.4.2 Assumptions for UEs

There are two types of equipment, LTE and LTE-A types in the system. Half number of equipment is LTE type and can only use one carrier and the other half is LTE-A type and can use multiple carriers (up to five). In simulation, four CCs can be simultaneously used by LTE-A type equipment because maximum five CCs can be used by LTE-A type equipment, and one of them must be used for upload primary component carriers (see Section 3.2). Users are initially non-uniformly distributed in area which is arranged as the most users are located nearby to eNB. 50% of users can move around of the eNB in specified time interval.

Each user can only download one type of traffic. Packet arrivals follow Pareto Distribution with shape parameter 2.5 and different packet arrival rates. Pareto Distribution is selected for simulation because Pareto based traffic models well simulate the high speed networks with unexpected demand on packet transfers by considering the long-term correlation in packet arrival times [104]. If there is one user in the network, the total packet

arrival rate is 250. If there are two users in the network, total packet arrival rate is 500 (different users can have distinct or same packet arrival rates). The arrival rate is simulated in such way to be suitable to finite buffer simulation. Therefore, the total arrival rates of traffic are enlarged when the number of users is increased.

3.4.3 Packet Scheduling

In the simulation, we have used a min-delay packet scheduling method in order to compare joint and selective techniques. Proportional Fairness is not preferred because Proportional Fairness packet scheduling can block packet transfer [51]. Therefore, the performance of the carrier assignment methods could not be observed correctly.

Packet arrival traffics are kept same for all test cases. Because of UEs and eNB positions, CQI for all carriers can be one of four options which are given in Table 3.2. Each packet is transferred by using one of the assigned carriers. To increase the efficiency and QoS, the packet transferring priority is given to the CC, which belongs to the highest frequency and minimizes packet delay if multiple carriers are available. If there are no available carriers to serve arrived packets, the packets are enqueued to corresponding user queues. Queue lengths are kept equal (of 50 packets) for each user queue. Buffer lengths are kept small [102], similar to real system to reduce packet delay. If there are not any empty spaces in queues, the arrived packets are dropped.

3.4.4 Observation Methodology

The results in Section 3.5 are the average of 200 realizations for different size of users. The impacts of light and heavy user loads on joint and selective techniques are investigated by using four different methods which are explained in Section 3.2.3. In each figure, the method name is given on the title and the labels are used to distinguish joint and selective techniques.

We present the performances of joint and selective techniques by comparing the throughput ratio and average delay. The throughput ratio shows how much data is successfully

transferred out of generated packets and is measured by dividing the transferred packets to all the processed packets (dropped packets + successfully transferred packets). Therefore, while the number of users is increased, the throughput ratio decreases because of the carrier capacities. Drop probability is not given because it is just inverse of the throughput ratio (Throughput ratio = 1 - drop probability). However, drop probability due to carrier assignment operations is presented to verify the correctness of our approach. The drop probability is measured by summing all the dropped packets during the carrier assignment operations and then divided by all the processed packets.

The average delay per packet shows how much time a packet waits to transfer. Here waiting times of the dropped packets are ignored and only the delays of transferred packets are considered. It is determined based on the waiting time in the queues and service. Additionally, the average delay which is experienced by packets during the periodic carrier assignment process is shown to verify the analytical expressions in Section 3.3. To measure the average delay during the carrier assignment process, we consider the time of packet arrival, the beginning time of the carrier assignment process and the finishing time of the carrier assignment process for each packet. After summing delays experienced by all packets during the carrier assignment process, the sum is divided to the number of processed packets (transferred and dropped packets) to find the average delay per packet. Some packets may or may not experience delay because of the carrier assignment process but the overall average delay is affected by any delay. Furthermore, the performances of joint and periodic techniques are evaluated in terms of equipment types (LTE and LTE-A type equipment) by using the explained performance metrics.

As a result of the average delay and throughput ratio comparison between joint and selective techniques, tradeoff between their resource usages and managed QoS are compared.

3.5 Results

In this section, delay experienced by users during carrier assignment process, overall system performance and experienced performance by each device type are presented for joint and selective techniques.

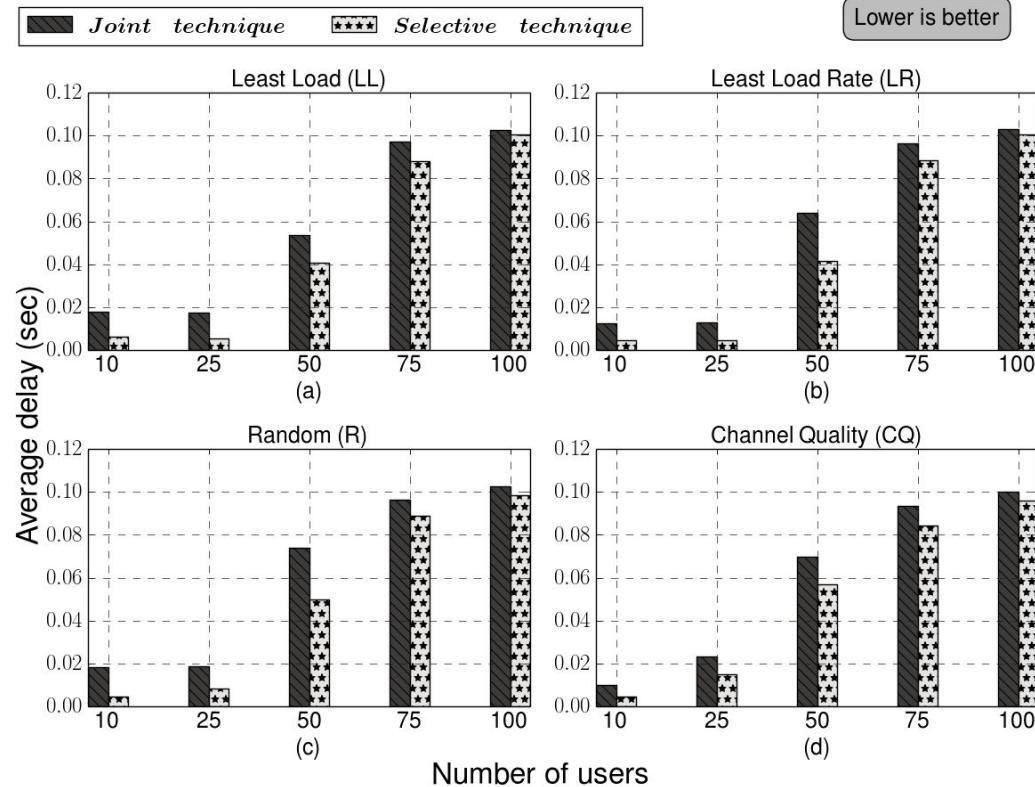


Figure 3.7: Average delay during the periodic carrier assignment process for joint and selective techniques.

3.5.1 Average Delay and Drop Ratio during Carrier Assignment Operations

In this section, the average delay and drop ratio due to carrier assignment operations are presented to show how different methods are effected by joint and selective techniques during the carrier assignment operations.

Figure 3.7 demonstrates the average delay due to the carrier assignment operations for joint and selective techniques. When the number of users is 10 and 25, the average delay is lower than 0.03 second for all the methods and the average delay is significantly lower in selective technique. When the number of users is 50 and more, the average delay also gradually increases for all cases but the average delay of joint technique is again higher than the average delay of selective technique for the methods due to a lower num-

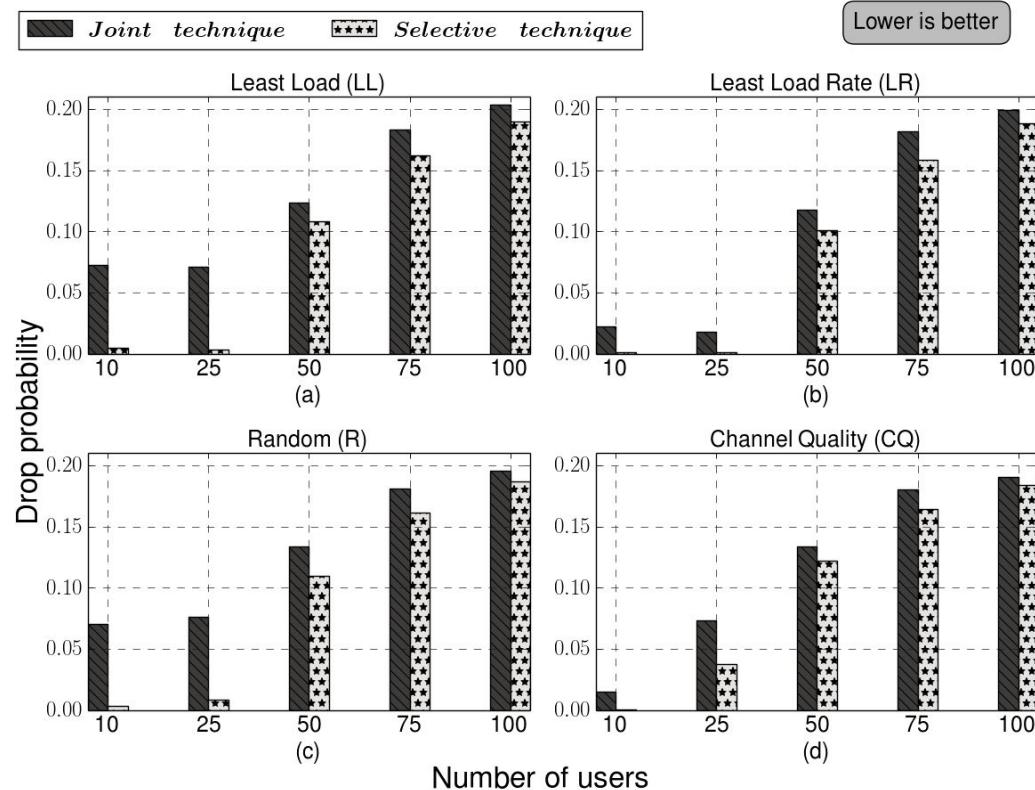


Figure 3.8: Drop probability during the periodic carrier assignment process for joint and selective techniques.

ber of packet interruptions in selective technique. However, the average delay difference between joint and selective techniques is decreasing while the number of users is raising.

Figure 3.8 demonstrates the drop probability due to the carrier assignment operations for joint and selective techniques. When the number of users is 10 and 25, the drop probability is remarkably lower than in selective technique especially for LL and R methods. When the number of users is 50 and more, the drop probability also gradually increases

for all cases but the drop probability of selective technique is again lower than the drop probability of joint technique for the methods due to a lower number of packet interruption in selective technique. However, similar to the average delay difference, the drop probability difference between joint and selective techniques is decreasing while the number of users is raising (see Figure 3.7). Figures 3.7 and 3.8 clearly show that selective technique significantly decreases the average delay and drop probability experienced by users during the carrier assignment operations. However, the overall system performance can be different. Therefore, the following section includes the results for the overall performance analysis.

3.5.2 Overall Performance of the System

In this subsection, the overall system performances of the methods for joint and selective techniques are presented by using the average delay and throughput ratio.

3.5.2.1 Average Delay

Figure 3.9 demonstrates the average delays of the methods for joint and selective techniques. When the number of users is increasing, the average delay is regularly getting higher for all cases due to a high number of packet arrival rates. In all cases, selective technique is better than joint technique as shown in Figure 3.9. For instance, the average delay of joint technique is between 0.06 and 0.47 seconds for all the methods when the number of users is 50 and fewer. However, the average delay of selective technique is between 0.03 and 0.22 seconds for the same number of users. Therefore, selective technique decreases the average delay up to 50%.

When the number of users is 75 and more, the average delay is changing between 0.93 and 1.25 seconds for joint technique. However, the average delay is between 0.80 and 1.08 seconds for selective technique. Therefore, selective technique improves the average delay up to 15% while the system is under heavy data traffic loads. It is worth mentioning that while the number of users is increasing (after 50 users), the average delay

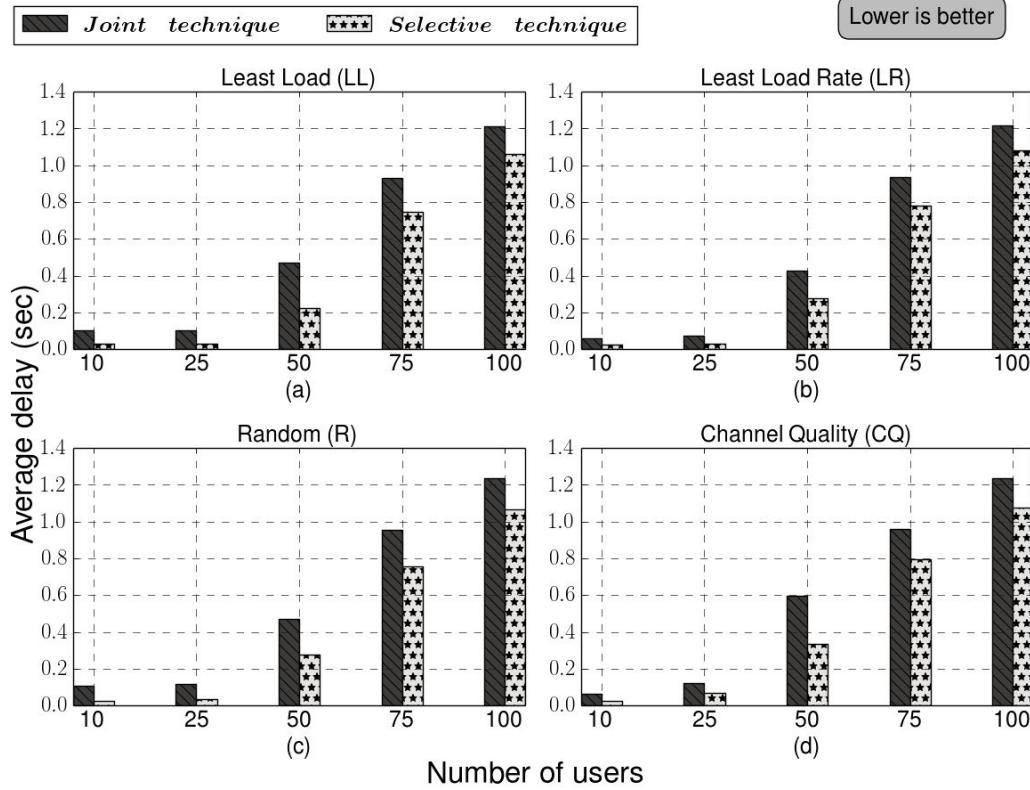


Figure 3.9: Average delay per packet for joint and selective techniques.

gap between selective and joint is decreasing for all the methods as expected. This is due to the capacity limitation of the system.

3.5.2.2 Throughput Ratio

Figure 3.10 shows the throughput ratios for joint and selective techniques. The throughput ratios are gradually decreasing for all cases while the number of users is increasing. For all cases, selective technique has a higher throughput ratio than joint technique. While the number of users is 25 or lower, selective technique improves the throughput ratio up to 14% (almost 0.87 to 0.99) in LL and R methods comparing to joint technique. Selective technique also increases the throughput ratios of LR and CQ but the improvement is not as significant as LL and R methods for the same number of the users. When the number of users is 50, selective method improves even more (up to 18%). However, the throughput

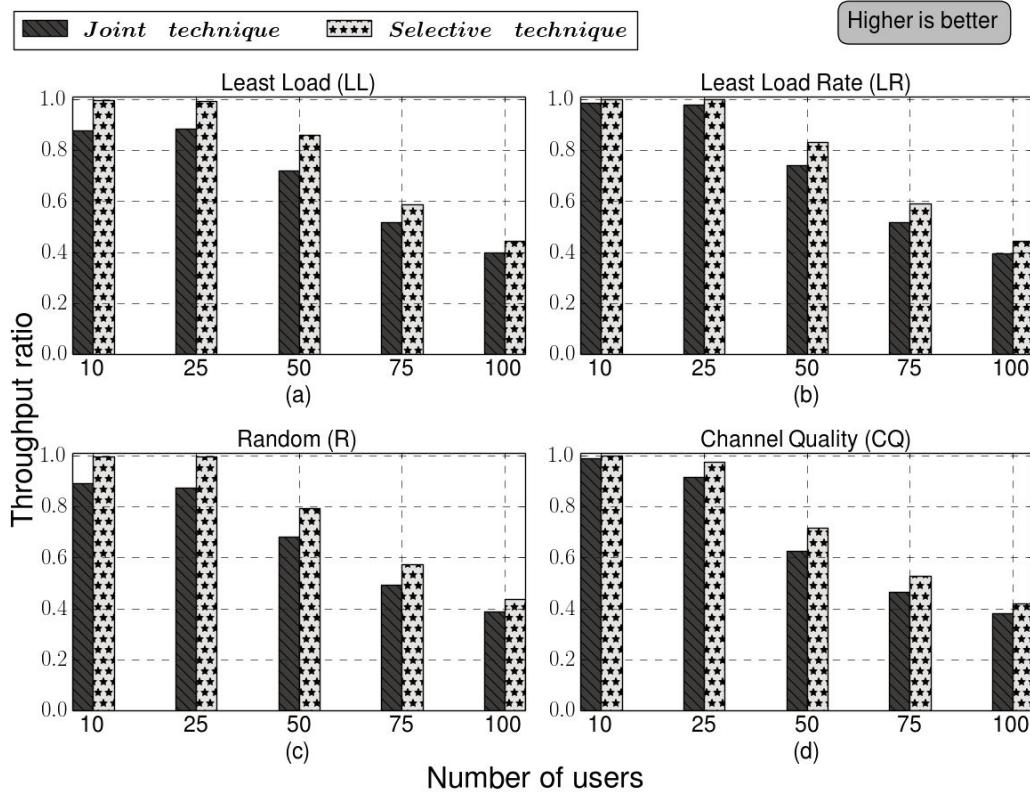


Figure 3.10: Throughput ratio for joint and selective techniques.

ratio improvement for a higher number of users (more than 50) is not as much as for a lower number of users in LL and R methods due to carrier capacity and packet arrival rates.

Moreover, all the methods with selective technique have almost the optimum (=1.0) throughput ratio when the number of users is 25 or lower. However, only LR method with joint technique has almost the optimum throughput ratio for the same number of users. It is worth mentioning that LL and LR methods have almost the same and the highest throughput ratio in selective technique and LR method has the highest throughput in joint technique.

3.5.3 Average Delay and Throughput Ratio According to Equipment Type

In the following subsections, the experienced performance by each equipment type (LTE and LTE-A equipment types) for four methods with selective and joint techniques is pre-

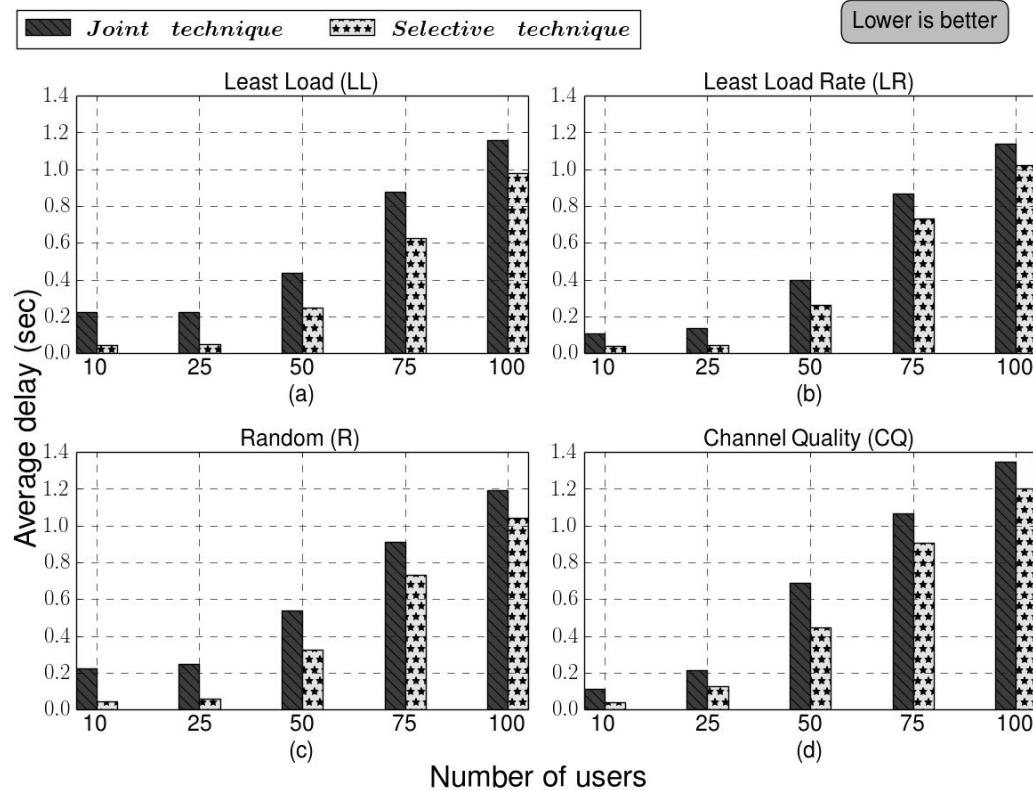


Figure 3.11: Average delay per packet of LTE type devices for joint and selective techniques.

sented according to the average delay and throughput ratio. The equipment based comparison is shown to investigate how the users of different types of equipment will be affected by joint and selective techniques if there are multiple types of equipment in the system.

3.5.3.1 Average Delay

Figures 3.11 and 3.12 show the average delays per packet which are experienced by LTE and LTE-A type equipment, respectively. When the number of users is 25 or lower, the

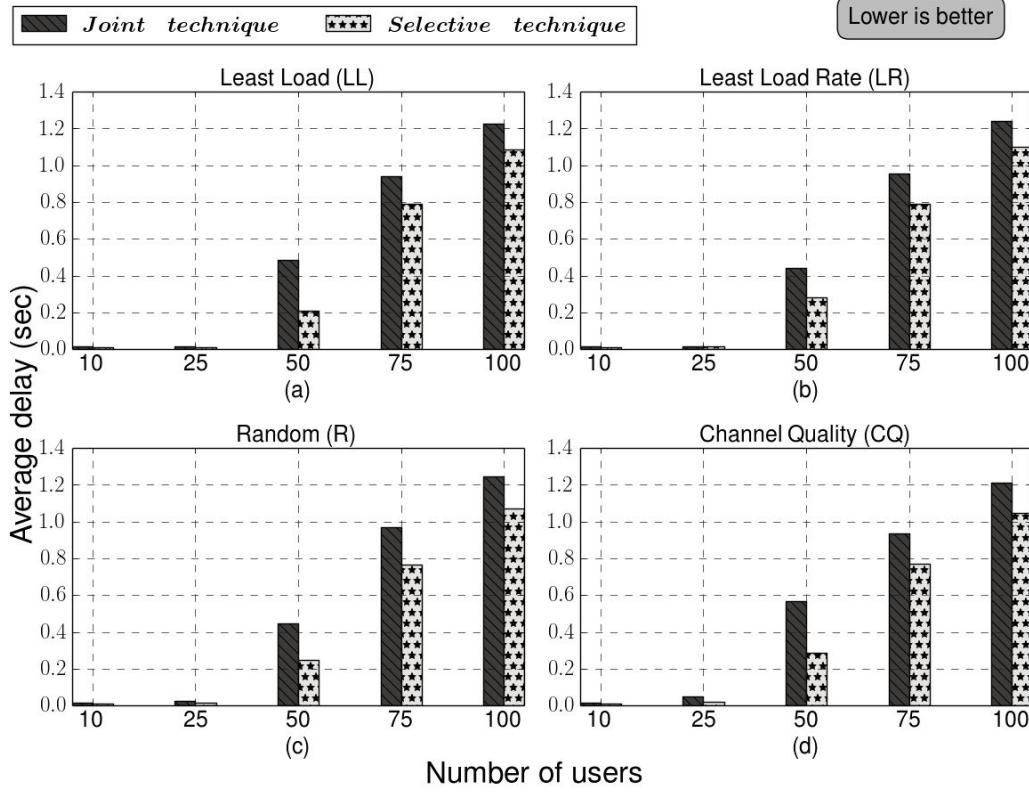


Figure 3.12: Average delay per packet of LTE-A type devices for joint and selective techniques.

average delay of LTE type equipment is higher than the average delay of LTE-A type equipment for all the methods because there is only one assigned CC to serve for LTE type equipment and multiple assigned CCs for LTE-A type equipment. Due to light packet arrival loads, the carriers are not busy all the time. Thus, the packets of LTE-A type equipment do not experience much delay. For the same number of users, selective technique remarkably decreases the average delay of LTE type equipment and slightly improves the average delay of LTE-A type equipment comparing to joint technique for all the methods. This shows that joint technique frequently interrupts the packet transfers for LTE type devices.

When the number of users increases to 50 and more, there are slightly differences between the average delays of LTE and LTE-A type equipment because LTE type equipment makes carriers busier due to higher packet arrival rates. However, all the methods with

selective technique have up to 50% lower average delays than joint technique for both LTE and LTE-A type equipment because the number of the packet transfer interruptions is lower in selective technique. Additionally, the average delay difference between joint and selective techniques for LTE type equipment is decreasing while the number of users is increasing. This is also true for LTE-A type equipment when the number of users is 50 and more.

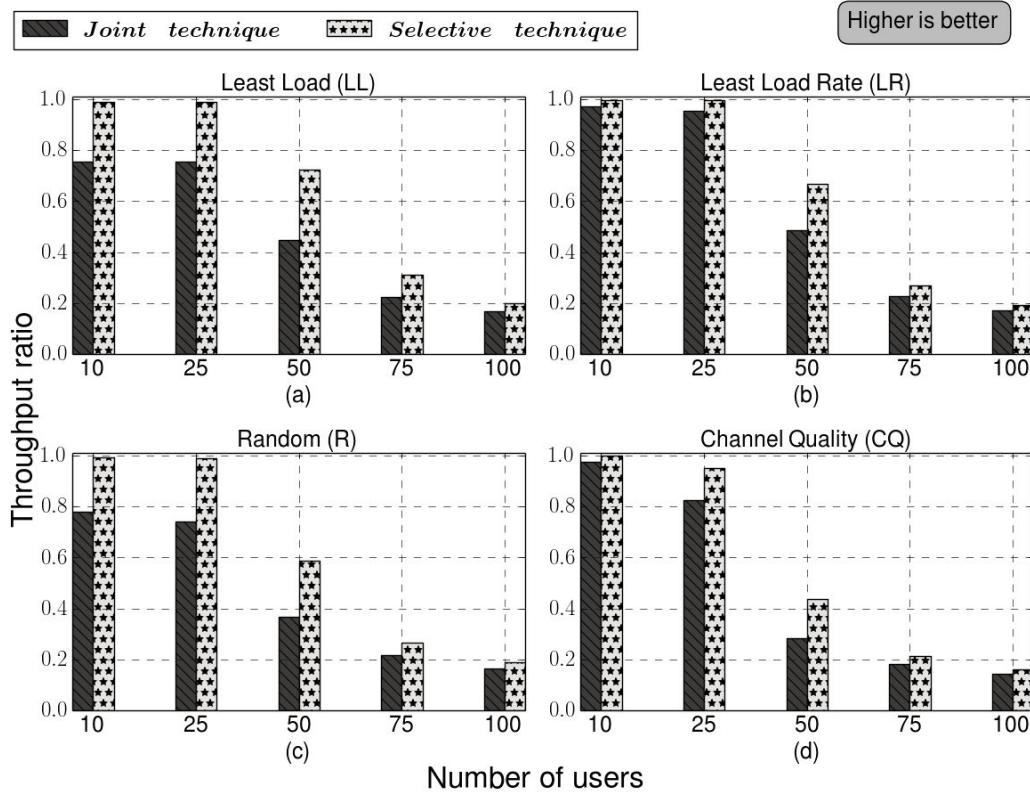


Figure 3.13: Throughput ratio of LTE type devices for joint and selective techniques.

3.5.3.2 Throughput Ratio

Figures 3.13 and 3.14 demonstrate the throughput ratios which are experienced by LTE and LTE-A type equipment for joint and selective techniques. The throughput ratio of LTE type equipment is lower than the throughput ratio of LTE-A type equipment for all the methods because of different capacities of the equipment. The throughput ratio of

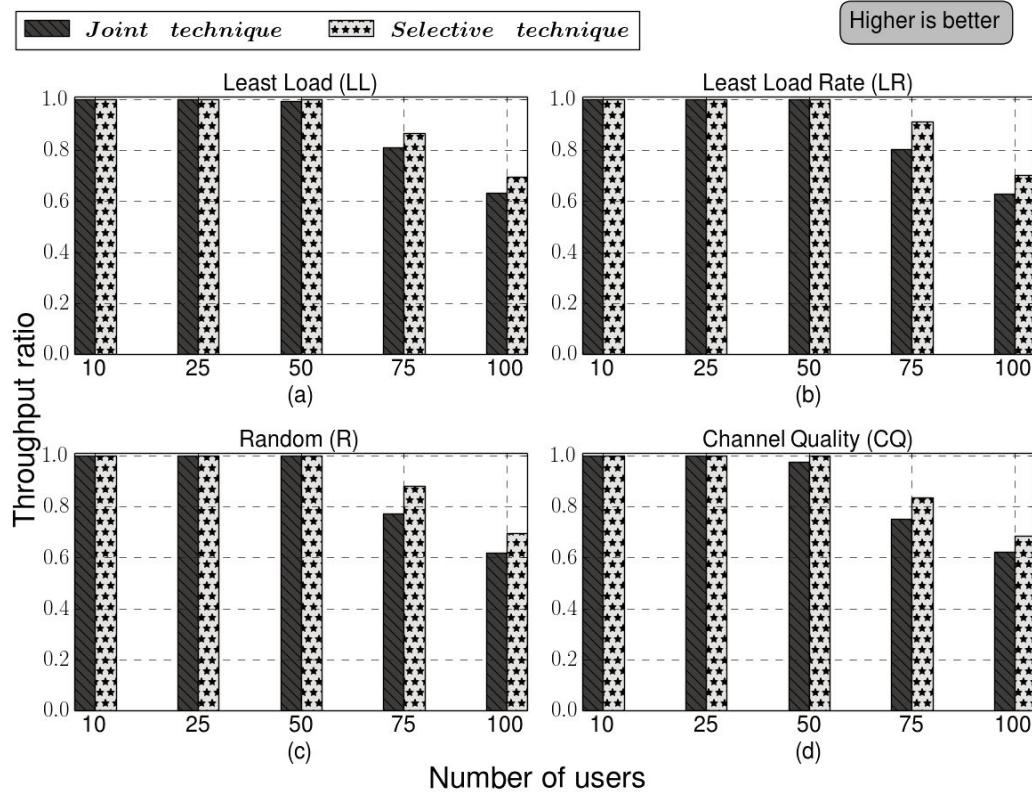


Figure 3.14: Throughput ratio of LTE-A type devices for joint and selective techniques.

LTE-A type equipment is 1.0 for both joint and selective techniques when the number of users is 50 and lower. However, only LR with joint technique and all the methods with selective technique have almost 1.0 throughput ratio for LTE type equipment when the number of user is 25 and fewer. This actually shows that selective technique significantly increases throughput ratio of LTE type equipment (almost up to 35%). Additionally, selective technique also improves throughput ratio of LTE-A type equipment for all the methods when the number of users is 75 and more.

3.5.4 Summary of Results

Based on the results, we make the following observations: (i) Joint technique shows that LTE type equipment traffic suffers higher delay than LTE-A type equipment traffic due to the interruptions of packet transfer; (ii) Selective technique significantly enhances the

performance of LTE and LTE-A. However, the improvement in LTE type equipment is higher than the improvement in LTE-A type equipment because of capacity of LTE type equipment; (iii) Selective technique remarkably decreases the overall (up to 50%) average delay and improve (up to 18%) throughput ratio comparing to joint technique.

3.6 Summary

In this chapter, selective periodic component carrier assignment technique was proposed by considering the behavior of the system during the component carrier assignment operations. The performances of current joint and proposed selective component carrier assignment techniques were compared by using queuing theory and an extensive simulation. Both techniques were analyzed according to not only the overall system performance but also the device-based performances. Results show that the proposed technique efficiently uses system resources and improves the overall throughput ratio up to 18% and average delay up to 50% in LTE and LTE-A systems. Our proposed selective technique and related analysis should help service providers build efficient periodic component carrier assignment methods in order to improve performances metrics such as throughput ratio and delay.

In the next chapter, a novel carrier assignment method is proposed by considering behaviors of users to provide better services.

Chapter 4

Component Carrier Assignment Method Based on User Profile in LTE and LTE-A

In the previous chapter, selective technique was developed for the periodic carrier assignment methods to eliminate the interruption of packet transfers because of joint technique. In this chapter, behaviors of users are investigated while assigning carriers to users because continuous increase in bandwidth demand of users forces the operators to manage the resource allocation more intelligently. Therefore, a novel component carrier assignment method is proposed by considering user profiles (new strategy), which is a tracking technique to identify the mobility and data usage of users, to increase quality of services and experiences getting by mobile users. Results show that the proposed method uses the system resources efficiently and can improve the overall throughput ratio up to 15% and the average delay up to 13% in comparison to other methods. Our method will help service providers build efficient carrier components assignment methods through considering user profile to improve performance metrics, such as throughput ratio and delay.

Results presented in this chapter have appeared in IEEE GLOBECOM Workshop on Broadband Wireless Access and IEEE Wireless Communications and Networking Conference (WCNC) [105, 106].

4.1 Introduction

Continuously increasing the bandwidth demand of users forces the operators to manage data traffic more intelligently because economical and physical limitations do not allow the operators to extend network capacity [107]. Although Load balancing, CQI, etc., as summarized in Chapter 2, have been used to manage the carrier assignments process, more advance techniques [107] in addition to the developed methods will be needed to answer users' demands in the future. Therefore, we have developed a user profile carrier assignment method to manage the carrier assignment process more intelligently in LTE and LTE-A because not only the mobility of each user profile is different but also each user profile needs various QoS from distinct types of data traffic [107]. Five common user

Table 4.1: Mobile user profiles.

		User Profile					
		Teen.	H. Wife	B. Man	Grad. Stu.	G. Parent	
Traffic Types	RT	Video	V. High	Middle	Low	Medium	Low
	RT	Onl. Game	V. High	Low	Low	Medium	Low
	RT	Movie	V. High	V. High	Low	Medium	Low
	RT	Talk	Low	Medium	V. High	Medium	Medium
NRT	NRT	Web	High	Low	V. High	Medium	Low
	NRT	Mail	High	Low	V. High	Medium	Low
	NRT	SMS	V. High	Medium	Low	Medium	Low
	NRT	Mobility	Low	Medium	V. High	Low	Low
Con.	Location	Location	Low	Medium	High	Medium	Low

profiles are considered in Table 4.1. They are Teenager (Teen.), Housewife (H. Wife), Businessman (B. Man), Graduate Student (Grad. Stu.) and Grandparent (G. Parent). As illustrated in Table 4.1, the bandwidth demand of each user varies depending on applications (Real-Time (RT) and Non-Real-Time (NRT) services) and the mobility of each user is different (see Table 4.1 for Teenager and Businessman). Therefore, user profile, in addition to CQI, can be considered to increase QoS and QoE. It is important to note

that the actual traffic types of LTE standard are referred to as Quality Class of Identifier (QCI) and can be grouped under nine categories [42]. However, we make them simple to be understandable by all readers from different fields.

None of the listed previous works in Chapter 2 considers user profiles for the carrier assignment process although load balancing, CQI and traffic loads are considered. However, in [75], only the mobility of users is estimated in real time while assigning carriers to users in order to decrease the number of carrier reselection and handover, yet the real-time cost for mobility measurements, the throughput and average delay effects of the method are not demonstrated. Therefore, the *aim* of this chapter is to improve the performance of LTE and LTE-A by proposing a novel carrier assignment method.

In this chapter, the benefits of user profile strategy have been shown by comparing the user profile based carrier assignment method to four different carrier assignment methods; Least Load (LL), Random (R), Channel Quality (CQ) and Least Load Rate (LR)¹. In the user profile based method, firstly we show how to estimate user profiles. Then obtained user profile information, CQI and load balancing are used to improve the efficiency of the carrier assignment in LTE and LTE-A.

The *objective* of this chapter is to propose a technique to identify user profile and propose a novel carrier assignment method by considering user profile, CQI and load balance properties in LTE and LTE-A. The key *contributions* of this chapter are as follows: (i) Defining user profiles with respect to data traffic types and mobility, (ii) proposing a carrier assignment method based on user profile, CQI and load balancing, and (iii) evaluating the performance of the proposed method by comparing with four different carrier assignment methods according to the overall and equipment type performances.

Results show that the proposed method uses system resources efficiently and can improve the overall throughput ratio up to 15% and the average delay up to 13% in comparison the other methods. Therefore, the proposed method and related analysis should help

¹Details about the methods are given in Section 3.2.3

service providers build component carrier assignment methods through considering user profile to improve performance metrics, such as throughput ratio and delay.

The rest of the chapter is organized as follows: User profile with its properties and the proposed method are presented in Section 4.2. The simulation environment with its parameters are briefly explained in Section 4.3. In Section 4.4, simulation results are analyzed. Finally, Section 4.5 has the concluding remarks.

4.2 User Profile

The historical information about each user plays a crucial role to identify user profiles because the data consumption and mobility of users, which are the main factors for user profiles, are measured according to the behaviors of users. Before explaining how user profile is determined according to the data usages of users and their mobility, the effects of data usage and the mobility on carrier assignment are illustrated with some of the scenarios as follows by considering user movement from *Band-b* coverage to *Band-c* coverage in Figure 1.1: (i) The user has a higher data usage and the user mobility is low, therefore updating its CCs by selecting CCs from *Band-c* will increase the service quality because *Band-c* has more non-interference carriers; (ii) The mobility of the user is high and the data usage of the user is low. Hence, assigning new CCs from *Band-c* may cause service interruptions because frequent position changes of the user cause the user to frequently update CCs; (iii) The user does not consume much data, thus no need to update CCs; (iv) The mobility of the user is high and the user consumes more data. Therefore, the user can use multiple CCs from the multiple bands.

In addition to the carrier assignment, determining the number of required component carriers for each user is important because of power usages and QoS efficiency. For example, when a user can enter an eNB coverage, some of the scenarios to determine the number of carriers for the user can be as follows: (i) The data traffic of the user is low, thus only one carrier (one CC) will be enough; (ii) The data traffic of the user is high,

hence assigning multiple CCs can increase QoS and QoE; (iii) One CC can be assigned to the user because the device type of the user does not allow assigning more than one CC. Above-mentioned scenarios are only a few examples which show the importance of the determination of the bands and the number of the required carriers during the carrier assignment process.

4.2.1 The Technique to Collect User Information

To find the data consumption of users and their mobility, we have proposed an eNB-based data collection technique. Each eNB collects the past activities of users as follows: (i) Each eNB has a table similar to Table 4.2. However, instead of column *eNB-ID*, each eNB table has *UE-ID* column for each user. The other columns are the same as Table 4.2; (ii) Each user activity information (such as the amount of downloaded and uploaded data according to traffic types and how much time the user is in the service of each CC or each band) is collected by eNB (the information in Table 4.2); (iii) If the user leaves from one eNB boundary to enter another eNB boundary, the collected information about the user is transferred from old eNB to the new eNB and also to UE during the handover process according to *UE-ID* (this requires each user must have a unique identification ID). In the meantime, UE also sends the previous recorded information from all eNBs to the new eNB; (iv) The new eNB uses *UE-ID* to locate the user information to update the activities of the user. It is important to note that the data is not redundant at different eNBs but each UE holds only the information of its own activity.

The above process is only one strategy to collect the user information in each eNB and report back to users. There would be a numerous amount of strategies. For example, the more detailed information about users can be collected by eNB according to time frames rather than one record.

4.2.2 User Profile Detection

As shown in Table 4.2, each UE holds *Times*, *Connection Time (Con. T)* and *Idle Time (Idle T.)*, *RT* and *NRT services* data sizes for each eNB. In Table 4.2, *Times* illustrates how often a user connects to eNBs, *Con. T* represents how long a user keeps connected eNBs and *Idle T.* gives how long a user keeps connected but does not receive any services from the previous sessions for each band.

Table 4.2: User profile detection based on eNBs.

	<i>Band-a/Band-b/Band-c</i>			RT-Services		NRT-Services	
eNB-ID	Times	Con. T.	Idle T.	Video	Game	Web	Mail
<i>ID</i> ₁	<i>f</i> ₁	<i>c</i> ₁	<i>t</i> ₁	<i>v</i> ₁	<i>g</i> ₁	<i>w</i> ₁	<i>m</i> ₁
<i>ID</i> ₂	<i>f</i> ₂	<i>c</i> ₂	<i>t</i> ₂	<i>v</i> ₂	<i>g</i> ₂	<i>w</i> ₂	<i>m</i> ₂
<i>ID</i> ₃	<i>f</i> ₃	<i>c</i> ₃	<i>t</i> ₃	<i>v</i> ₃	<i>g</i> ₃	<i>w</i> ₃	<i>m</i> ₃
<i>ID</i> ₄	<i>f</i> ₄	<i>c</i> ₄	<i>t</i> ₄	<i>v</i> ₄	<i>g</i> ₄	<i>w</i> ₄	<i>m</i> ₄
<i>ID</i> ₅	<i>f</i> ₅	<i>c</i> ₅	<i>t</i> ₅	<i>v</i> ₅	<i>g</i> ₅	<i>w</i> ₅	<i>m</i> ₅
<i>ID</i> ₆	<i>f</i> ₆	<i>c</i> ₆	<i>t</i> ₆	<i>v</i> ₆	<i>g</i> ₆	<i>w</i> ₆	<i>m</i> ₆
<i>ID</i> ₇	<i>f</i> ₇	<i>c</i> ₇	<i>t</i> ₇	<i>v</i> ₇	<i>g</i> ₇	<i>w</i> ₇	<i>m</i> ₇
<i>ID</i> ₈	<i>f</i> ₈	<i>c</i> ₈	<i>t</i> ₈	<i>v</i> ₈	<i>g</i> ₈	<i>w</i> ₈	<i>m</i> ₈

In order to identify user profile from Table 4.2, some statistical analysis such as rate, ratio, percentage, etc., can be used. For example, the percentage of *Connection Time* of UE *i* to eNB *j* according to the total connection time (C_j^i) and the percentage of *Times* of UE *i* to eNB *j* according to the total connection frequency (T_j^i) can be simply calculated as follows:

$$C_j^i = 100 \times \frac{c_j}{\sum_{s=1}^k c_s} \quad \text{and} \quad T_j^i = 100 \times \frac{f_j}{\sum_{s=1}^k f_s} \quad (4.1)$$

where *k* is the number of eNBs. Lower T_j^i and higher C_j^i indicate that UE *i* spends its more time around eNB *j* with specified carrier band. On the other hand, higher T_j^i and lower C_j^i indicate that UE *i* temporarily requests service from eNB *j*. For example, UE *i* just receives services from eNB *j* while driving home, to work or school.

The data usage of a UE can also be estimated from Table 4.2. For example, the percentage of RT (RT is real-time data and NRT is non-real-time data) data usage according to the total data usage of a data traffic type for UE i in eNB j can be simply measured as

$$RT_j^i = 100 \times \frac{v_j + g_j}{\sum_{s=1}^k (v_s + g_s)} \quad (4.2)$$

Like RT_j^i , NRT_j^i can be obtained. Furthermore, the percentage of the active time connection of UE i in eNB j (AT_j^i) can be measured as

$$AT_j^i = 100 \times \frac{c_j - t_j}{\sum_{s=1}^k c_s - \sum_{s=1}^k t_s} \quad (4.3)$$

Similarly, the data usage analysis can be obtained according to service types such as mail, game, etc. for any eNB as above.

In addition to the percentage analysis, the average can be used. For example, the average connection time (AC_j^i), RT average (ART_j^i), NRT average ($ANRT_j^i$) and the average of all type (Θ_j^i) data usage of UE i in eNB j can be measured per connection as follows:

$$AC_j^i = \frac{c_j}{f_j}, \quad ART_j^i = \frac{v_j + g_j}{f_j}, \quad ANRT_j^i = \frac{w_j + m_j}{f_j}, \quad \Theta_j^i = \frac{v_j + g_j + w_j + m_j}{f_j} \quad (4.4)$$

The average analysis can be used by an eNB to identify a user profile although no information is available for the eNB in the user profile table. It is worth mentioning that Table 4.2 is prepared by considering only eNB rather than considering each CC for clarity. Therefore, Equations (4.1) - (4.4) are some of statistical analysis examples which can be used to identify user profiles based on eNB.

4.2.3 Number of Required CCs for Each User

In order to estimate the number of required CCs for UE i in eNB j , the average data consumption during the active connection time, which is obtained from Table 4.2, is used. Therefore,

$$\sigma = \frac{v_j^i + g_j^i + w_j^i + m_j^i}{c_j^i - t_j^i} \quad (4.5)$$

The number of the required CCs for UE i in eNB j can be obtained by using σ as follows:

$${}^{num}CC = \begin{cases} 1, & if \frac{\sigma}{\varphi_x^i} \leq 1 \\ \frac{\sigma}{\varphi_x^i}, & if \frac{\sigma}{\varphi_x^i} \geq 1 \quad and \quad \frac{\sigma}{\varphi_x^i} \leq 4 \\ 4, & if \frac{\sigma}{\varphi_x^i} \geq 4 \end{cases} \quad (4.6)$$

where φ_x^i is the maximum service rate of CC_x for UE_i and it is obtained by using the bandwidth and CQI of CC_x . $\sigma/\varphi_x^i \leq 4$ because only five CCs will be aggregated in LTE-A and one of five CCs is for uplink.

4.2.4 Component Carrier Assignment Process Based on User Profile

The proposed method considers three crucial parameters that enable the dynamic carrier assignment: (i) User equipment types in terms of LTE and LTE-A. (LTE device can only connect one CC while LTE-A device can connect up to five CCs [42]), (ii) CQI of CCs [86–89], and (iii) user profiles.

In order to assign the most suitable CCs to UE_i in eNB_j, we consider the total data usage of UE_i, the active connection time of UE_i per connection, and the number of active users with their previous data usages for each CC. The active connection time of UE_i to eNB_j per connection is measured for each band as follows:

$$\alpha_{jl}^i = \frac{c_{jl}^i - t_{jl}^i}{f_{jl}^i} \quad (4.7)$$

where i represents UE, j represents eNB and l represents band. c_{jl}^i is the total connection time of UE_i in eNB_j by using the carriers on band_l, t_{jl}^i is the idle time of UE_i in eNB_j on the carriers of band_l and f_{jl}^i is the number of connections of UE_i to eNB_j by using the carriers of band- l . After calculating the active connection time per connection (α_{jl}^i), the service rates of the carriers for UE_i are estimated by considering active users and their data usage rates, CQI and the data usage per connection of UE_i in eNB_j on each carrier in band- l as follows:

$$\beta_{jlx}^i = \frac{\varphi_x^i}{\eta_j} \quad (4.8)$$

where η_j is measured by summing all packet arrival rates of active users per connection ($\eta_j = \frac{\sum_{i=1}^n(v_j^i+g_j^i+w_j^i+m_j^i)}{f_j^i}$ and n is the number of active users on each CC). φ_x^i is the maximum service rate of CC_x for UE_i and it is obtained by using the bandwidth and CQI of CC_x . Then, CC which has the highest $\alpha_{jlx}^i * \beta_{jlx}^i$ as a result of equations (4.7) and (4.8) is selected for the user.

4.2.5 Carrier Assignment Process

As explained in the previous section, CC which has the highest $\alpha_{jlx}^i * \beta_{jlx}^i$ as a result of equations (4.7) and (4.8) is selected for the user. However, the carrier assignment process is as follows: (i) The device capacity of information of UE_i is transferred to eNB_j; (ii) Partially or fully CQI feedback is obtained from UE_i, the eNB lists all available CCs from the resources and measures the number of UEs waiting for services; (iii) The number of the required CCs is found according to Equation 4.6 and the carriers are selected by using equations (4.7) and (4.8), respectively; (iv) CC assignment is finished and buffers for each carrier are created for UE_i because DQS [97] is used as a queue scheduler because of its realistic approach [19]; (v) Packet transfer is started on the assigned CCs. To increase the efficiency and QoS, packet transferring priority is given to the CC, which is the closest to the eNB; (vi) Repeating process until all users are allocated.

4.3 Simulation of the System

The discrete event simulation environment, parameters and interested performance metrics such as utilization throughput ratio and average delay are similar to the previous chapter (see Section 3.4). However, several parameters of UEs have been changed because of user profile definitions.

4.3.1 Assumptions for UEs

Initially users are non-uniformly distributed in area (more users are located nearby to eNB). Half of users can randomly move around of the eNB in specified time interval to create high mobility condition. Each user can only download one type of data traffic. Packet arrivals follow Pareto Distribution and packet arrival rates are enlarged when the number of users is increased. Those assumptions are similar to the assumptions in Chapter 3. However, as a different from previous assumptions, the packet arrival rates are completely different for all users. For example, if there is one user in the system, the arrival rate is ten. If there are two users in the system, the packet arrival rate is ten for first user and twenty for second user. The packet arrival rates are chosen such a way to implement the model by distinguishing behaviors of users.

4.3.2 Observation Methodology

Three versions of user profile method; user profile based on perfect estimation, user profile based on the estimation with 10% error (UP^{10}) and user profile based on the estimation with 25% error (UP^{25}) are analyzed. For example, if the data usage rate of a user is 100MB then the estimated data usage of the user can be 125MB or 75MB for UP^{25} and 110MB or 90MB for UP^{10} . Therefore, the proposed method is evaluated under a more realistic scenario.

We present the performances of the methods by comparing CC utilization, throughput ratio and average delay.

4.4 Results

In this section, utilization, overall average delay and overall throughput ratio, and experienced average delay and throughput ratio by each device type are presented to compare the methods.

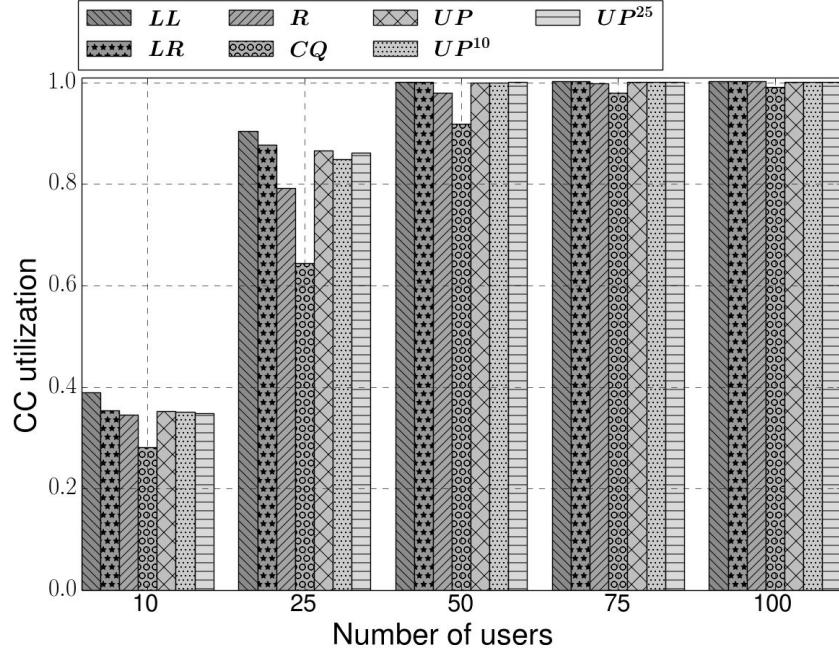


Figure 4.1: Utilization of CCs for user profile and the other methods.

4.4.1 Utilization

Figure 4.1 shows CC utilization for the methods. When the number of users increases, the utilization also raises for all cases. Though a larger number of users, the carrier utilization of CQ does not reach peak rate (=1) because CQ method does not assign lower quality carriers to users even if there are no active users on the carriers. This also results in that CQ has the lowest utilization. When the number of users is 25 and fewer, LL has the highest utilization because LL well balances user loads. UPs (UPs means UP, UP¹⁰ and UP²⁵) have almost the same utilization although UP with perfect estimation has the highest utilization.

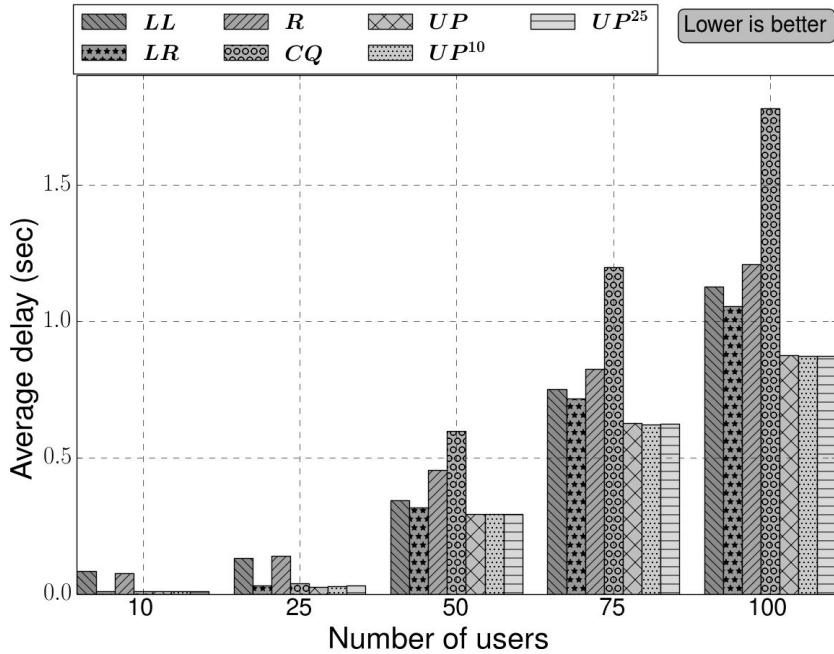


Figure 4.2: Average delay per packet of LTE type equipment for user profile and the other methods.

4.4.2 Average Delay

Figure 4.2 demonstrates the average delay experienced by LTE type equipment for all methods. The average delay regularly increases with the number of users for all cases due to enlarging requests from eNB. When the number of users is 25 and fewer, the average delay is lower than 0.2 second for all methods because the methods can efficiently handle the incoming data traffic and the assigned carriers are enough to provide services to users. For the same number of users, R and LL have higher average delays than other methods. When the number of users is 50 and more, the average delay is noticeably higher for all methods but especially the growth in CQ is significant. The delay of CQ is almost double of the delay of UPs since CQ does not handle the higher number of LTE type users due to taking only CQI into consideration. It is important to note that UPs have almost similar delays and delays of UPs are lower than all other methods. Quantitatively, UPs improves

the average delay up to 20% in comparison to LR which is the best method after UPs for LTE type equipment.

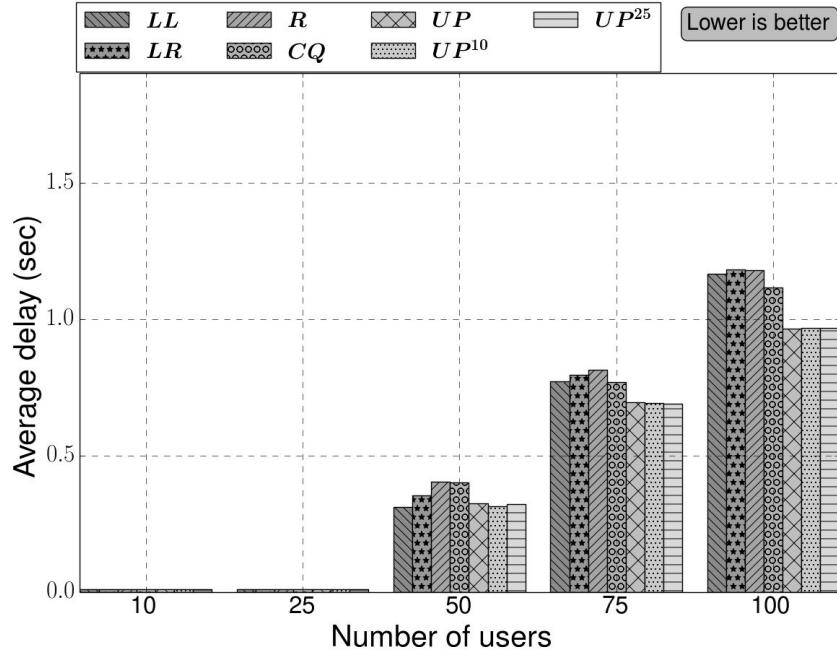


Figure 4.3: Average delay per packet of LTE-A type equipment for user profile and the other methods.

Figure 4.3 shows the average delay experienced by LTE-A type equipment for all methods. The average delays of all methods are extremely low when the number of users is 25 and fewer. Indeed, this is expected because assigning multiple component carriers to LTE-A type equipment causes the packets not to experience much delay during the packet transfer process. The other expected results by cause of multiple CCs assignment is that LTE-A equipment experiences much low delay than LTE equipment when the number of user is 25 users or fewer. When the number of users increases (50 and higher), the average delay also raises because of heavy traffic loads. Although there are small differences between the average delays of all methods in the network, UPs have lower delays than other methods. Quantitatively, UPs decrease average delay up to 18% in comparison to LR and R, 16% in comparison to LL and 14% in comparison to CQ for LTE-A type equipment.

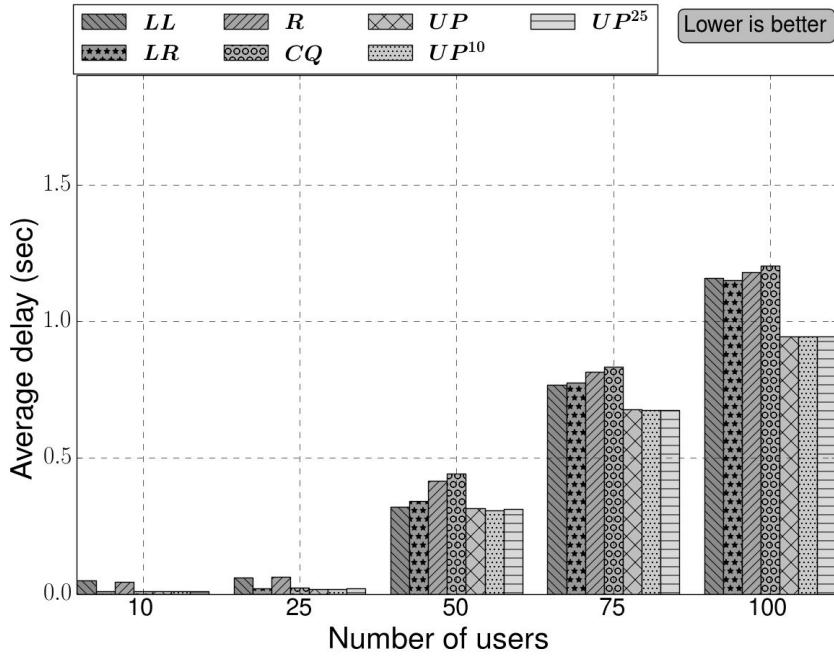


Figure 4.4: Average delay per packet for user profile and the other methods.

Figure 4.4 shows the average delay which is measured by considering all types of equipment for all methods. The average delay of all methods are below 0.1 second when the number of users is 25 and fewer. For a higher number of users, the average delays of all methods increase because of heavy traffic loads. After 50 users, UPs have lower delay than the other methods and decrease the average delay up to 20%.

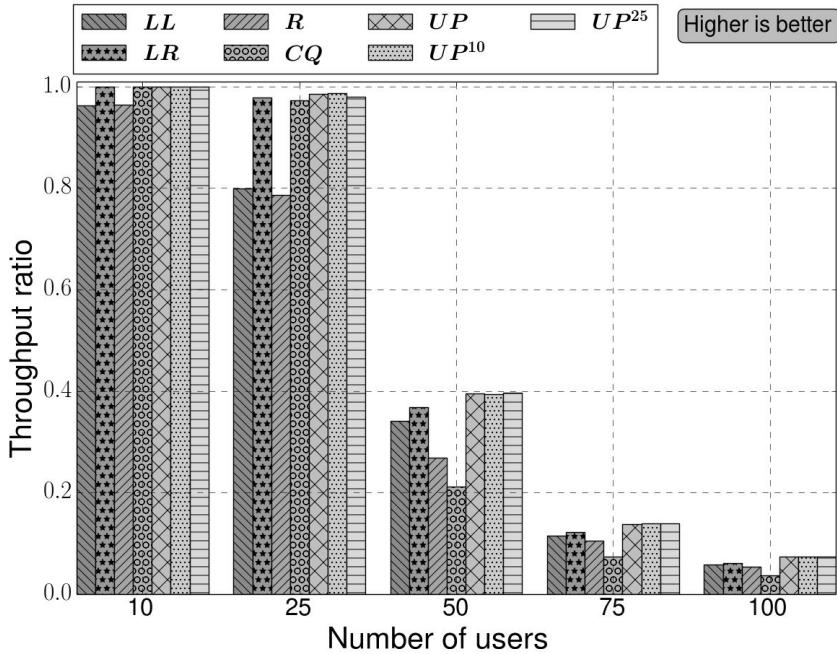


Figure 4.5: Throughput ratio of LTE type equipment for user profile and the other methods.

4.4.3 Throughput Ratio

Figure 4.5 demonstrates throughput ratios of all methods for LTE type equipment. An increment in the number of users gradually reduces throughput ratios of all cases because of growth of packet arrival rates. However, when the number of users is increased from 10 to 25, the throughput ratios of LL and R methods decrease almost up to 20% although there are insignificant changes of the throughput ratios of the other methods. Indeed, this shows the importance of CQI for the system which has light traffic loads because CQ, LR and UPs include CQI for carrier assignment. Moreover, increasing the number of users to 50 decreases the throughput ratios up to 60% for UPs, more than 60% for LR and up to 80% for CQ. This shows the importance of load balancing for the system which is under heavy traffic loads. Therefore, CQ shows the worst performance for LTE type equipment when there are a higher number of users in the system. In all cases, UPs have higher

throughput ratios than the other methods and improve the throughput ratio up to 13% in comparison to LR which is the best method after UPs for LTE type equipment.

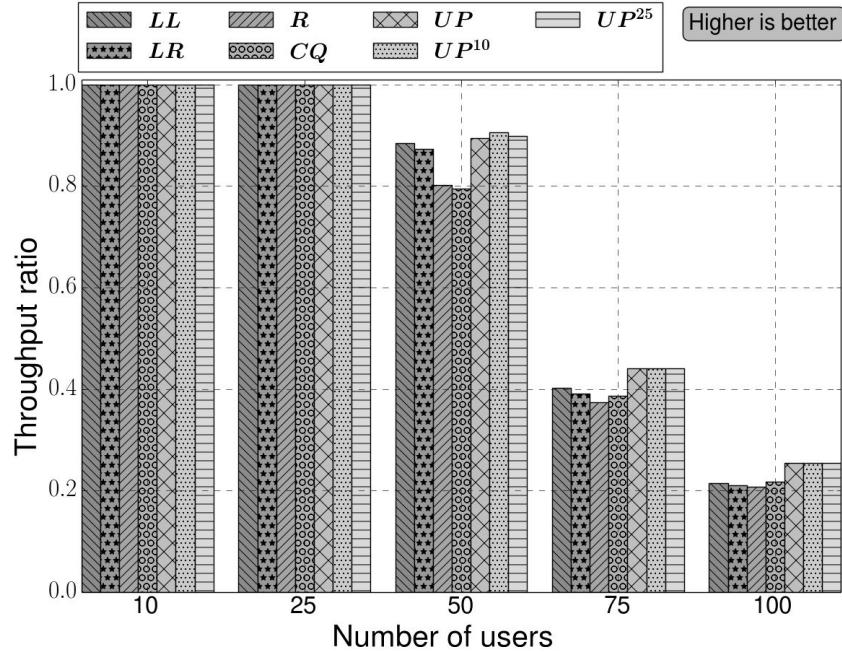


Figure 4.6: Throughput ratio of LTE-A type equipment for user profile and the other methods.

Figure 4.6 shows the throughput ratios of all methods for LTE-A type equipment. When the number of users is 25 and fewer, all methods have the optimum throughput ratio (=1) because assigning multiple carriers to users is enough to serve the incoming data traffic. However, when the number of users increases, the throughput ratios of LTE-A type equipment also decrease but the decrements are not as much as decrements in the throughput ratios of LTE type equipment. Indeed, LTE-A type equipment has more than two times a higher throughput ratio than LTE type equipment for all methods. It is important to note that UPs also improve throughput ratio up to 17% for LTE-A type equipment.

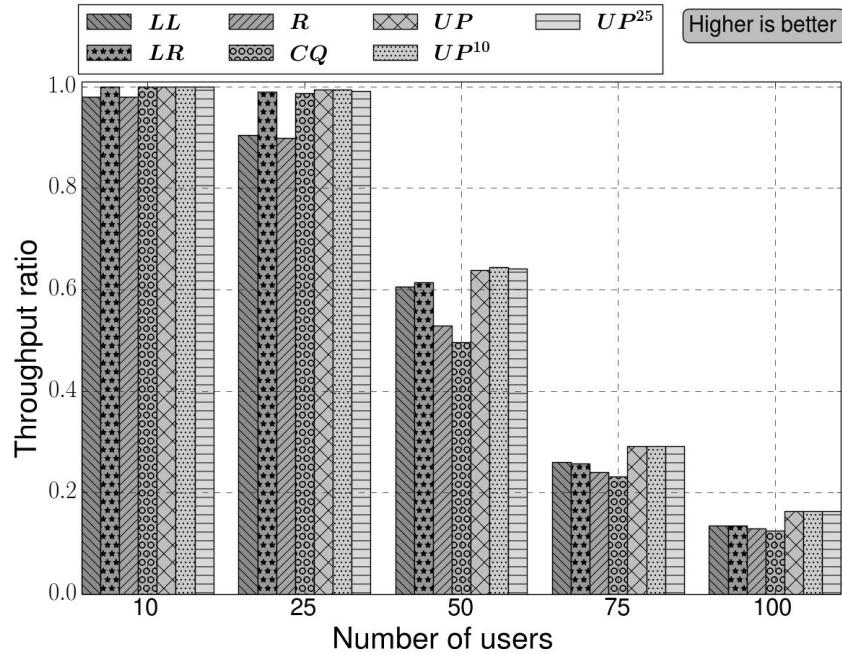


Figure 4.7: Throughput ratio for user profile and the other methods.

Figure 4.7 shows overall throughput ratios which are measured by considering all types of equipment in the system. UPs have higher overall throughput ratios than other methods and improve the throughput ratio up to 15% in comparison to LR method which is the best after UPs.

4.4.4 Summary of Results

Based on the results, we make the following observations: (i) UPs improve the overall throughput ratio up to 15% and the average delay up to 20% in comparison the other methods. However, the improvements can be more according to equipment type; (ii) Increasing error percentage of user profile does not affect overall performance of UPs (see Figures 4.2-4.7 for UP, UP¹⁰, and UP²⁵); (iii) Strategies for carrier assignment methods behaves differently according to equipment type and traffic loads (see LL and CQ in figures).

4.5 Summary

In this chapter, a user profile table was used to track user behaviors in each eNodeB. In order to show the benefits of, a carrier component assignment method was proposed for LTE and LTE-A based on user profiles. Utilization, average delay and throughput ratio were presented for the proposed method according to overall and device-based performances through an extensive simulation. Results show that the proposed method uses system resources efficiently and improves overall throughput ratio up to 15% and overall average delay up to 20% in comparison the other methods in LTE and LTE-A systems. Our proposed method and related analysis should help service providers build efficient carrier assignment methods by considering user behaviors.

Until this chapter, LTE and LTE-A have been considered. In the next chapter, multi-band in Wi-Fi will be considered for carrier and band selection with scheduling algorithms.

Chapter 5

Analysis of Multi Class Traffic of Single and Multi-Band Routers

In the previous chapters, the carrier assignment was investigated in LTE-A. In this chapter, integration of Carrier Aggregation into next generation Wi-Fi routers is investigated because modern mobile Wi-Fi routers are capable of supporting simultaneous multi-band, leading to less interference, higher capacity and better reliability to facilitate higher bandwidth. However, there exists neither previous works that attempts to maximize utilization of available bandwidth through the sharing of traffic classes among different frequency bands of the mobile router. In this chapter, a novel multi-shared-band architecture with a scheduling algorithm is proposed for multi-band mobile routers which transmits different classes of traffic through different frequency bands to achieve improved performance. An analytical model is developed to perform queuing analysis of the multi-shared-band system and various performance metrics are derived and validated by extensive simulations. Results are shown by comparing multi-shared-band model with existing single and multi-band models. The results show that the proposed architecture and the scheduler can ensure maximum possible utilization through sharing of capacities among the bands. Additionally, it is evident from the results that multi-band systems are not always

Results presented in this chapter have appeared in IEEE International Conference on Communication (ICC), IEEE Global Communications Conference (GLOBECOM) and Springer Wireless Personal Communication (WPC) Journal [100, 108, 109].

better than the single band systems although multi-band systems are expected to have better performance. Based on the results, recommendations are listed for choosing single or multi-band systems and allocation policies based on traffic conditions, and their priorities.

5.1 Introduction

To satisfy a higher bandwidth demands of users, currently wireless routers which are available commercially with simultaneous multi-band support of 2.4 and 5 GHz. Future IEEE 802.11ad (WiGig) tri-band enabled devices, operating in the 2.4, 5 and 60 GHz bands, are expected to deliver data transfer rates up to 7 Gbps [23, 110]. The benefit of using a multi-band in the Mobile Routers (MR) is *less interference, higher capacity and better reliability*. Exploitation of rarely-used frequency bands in wireless networks reduces interference in heavily-used frequency band, e.g., 2.4 GHz, thereby increasing total capacity of the wireless network.

Current simultaneous multi-band Mobile Router (MR) uses two different frequencies (2.4 GHz and 5 GHz) for different types of devices in a home network. However, they do not attempt to exploit the under-utilized frequency band when other one is flooded with data. Therefore, the current multi-band architecture does not efficiently use bands to make systems more productive.

Moreover, the multi-band router system is a heterogeneous multi-server system which means each server's service rate is different than the other. Heterogeneity of a system raises a problem: Which (arrived) packet should be distributed to which server (namely allocation policy) [111]? The problem becomes more complex when different classes of packets are considered since some of the traffic types (such as, real-time) have strict delay constraints [112–114]; some other signaling traffic (required for mobility management) is crucial for maintaining Internet connectivity of the mobile users. Therefore, *flexibility of each class* (i.e., which class can be served by which server), and *priority of class* (i.e., which class can be served first if a server can serve more than one class) can be taken into

account. Therefore, it is essential to propose an appropriate multi-band architecture with a band scheduling and queue management scheme for the multi-class traffic to ensure the maximum possible utilization of the system resources in multi-band MRs [18].

The *aim* of this chapter is to propose a multi-shared-band architecture with a scheduling algorithm for multi-band routers, then compare single and multi-band systems with different allocation policies to investigate that under which circumstances single or multi-band performs better through the use of different router service rates.

There have been several research works [18, 23, 24, 110, 115–119] reported in the literature that attempt to extend current single band technology through the use of multiple frequency bands, leading to increased bandwidth while reducing interference. Even though multi-band usage has been widely investigated in cell networks such GSM [117–119] and LTE, it is a relatively new concept in wireless networks. Authors [24, 25] explain possible Wi-Fi architecture with multiple physical and link layers to support multiple frequency bands simultaneously. Singh et al. [115] proposed a method to assign different frequency bands to end-devices based on their distances from the access router. In [23, 110, 116], authors proposed the use of 60 GHz frequency band (having low range) to attain faster data transfer rate in wireless networks. However, none of these works [18, 23, 24, 115–118] propose any band scheduling algorithm for multi-band system considering multi-class traffic, compare between single and multi-band, or perform any queuing analysis to measure difference performance metrics.

To the best of my knowledge, there have been no earlier works on band scheduling and queue management for multi-band MRs that attempts to maximize utilization of available bands by using band-sharing architecture. On the other hand, there are other type approaches; dynamic bandwidth size arrangement, priority base traffic management and dynamic band allocation [120–122] to increase the performance of the current single and multi-band systems. However, dynamic bandwidth size arrangement increases interference, priority base traffic management leads low priority traffic to get low service

and dynamic band allocation cannot utilize all bands simultaneously. Besides these disadvantages, the multi-shared-band architecture and scheduling algorithm can co-exists with aforementioned approaches without any limitations. Therefore, this is a *novel work* that aims at attaining maximum possible band utilization using different allocation policies while comparing the performance of single, current multi-band, and multi-shared-band systems. The *objective* of this work is to analyze the performance of multi-band MRs while ensuring maximum possible utilization through sharing of bands among different classes of traffic and determine whether the multi-shared-band architecture performs better than single and current multi-band architectures. The key *contributions* of this work are: (i) Proposing a band-sharing router architecture and a novel scheduling algorithm that aims at improved utilization of the system and testing performance of different allocation policies such as fastest server first, low utilization first, and slowest server first, (ii) developing an analytical model to evaluate the performance (utilization of bands, average class occupancy, packet drop rate, average delay, and throughput) of the multi-shared-band system, (iii) validating the analytical model by extensive simulations, (iv) comparing the performance of multi-band router with single band router by a developed realistic simulation, and (v) analyzing the results to make recommendations for choosing single or multi-band architecture with allocation policies based on traffic conditions and class priority.

The proposed algorithm considers multi-class Internet traffic and schedules them through alternate under-utilized frequency bands, thereby reducing packet loss and delay. The *results* of this work are: (i) Drop probability and throughput are significantly improved through the proposed band-sharing architecture; (ii) The simulation results validate the analytical model; (iii) Low priority classes of traffics in single band systems can suffer long delay; (iv) Multi-band routers can suffer low band utilization under light traffic.

The rest of the chapter is organized as follows: In Section 5.2, communication models of single and current multi-band systems are presented. In Section 5.3, the architecture of single band MRs is explained, followed by the current multi-band architecture in Section 5.4. In Section 5.5, the proposed architecture is explained and followed by scheduling

algorithm and analysis of the proposed system to derive different performance metrics in Section 5.6. In Section 5.8, validation of the developed formulas is presented and the performances of single, current multi-band and proposed multi-shared-band architectures are compared. Finally, Section 5.9 has the summary for this chapter.

5.2 The Communication Models of Single and Multi-Band Systems

Figure 5.1 shows the communication model for single band systems. All devices such as

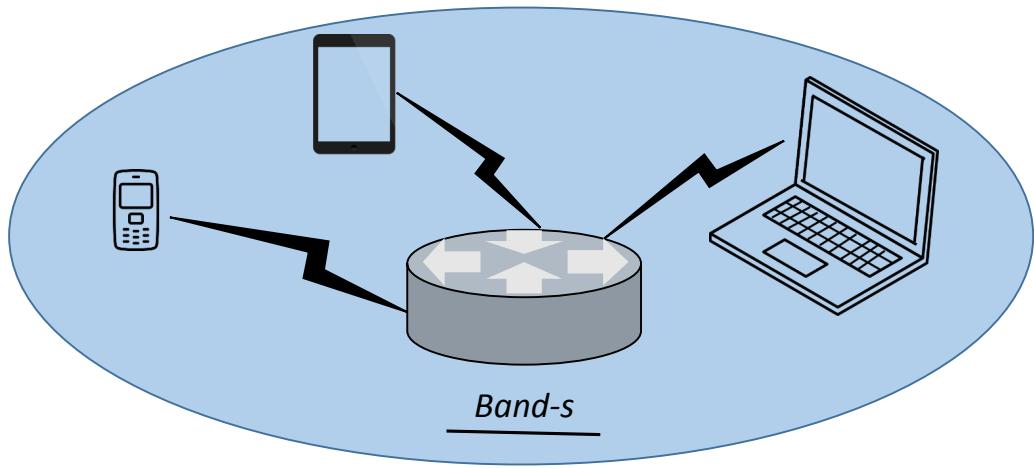


Figure 5.1: A single band router communication scenario.

tablet, TV and laptop in the network are connected to the router by using only *Band-s*. Therefore, simultaneously connecting several devices to the router over one band dramatically decreases the quality of the service.

On the other hand, the current multi-band system uses several bands to send and receive data as presented in Figure 5.2. Each device can connect different bands. For example, while the laptop is connecting router over *Band-b*, the tablet is using *Band-a*. Although, simultaneously connecting several devices to the router does not decrease the quality of the service, one band can be idle during communications.

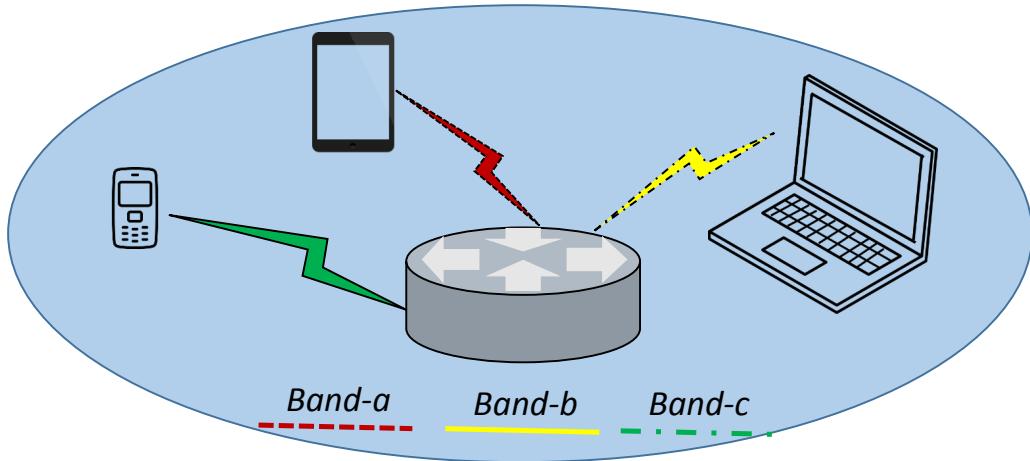


Figure 5.2: A current multi-band router communication scenario.

5.3 Single Band Router Architecture

Traditional single band MRs use only one frequency band for all types of traffic. Fig-

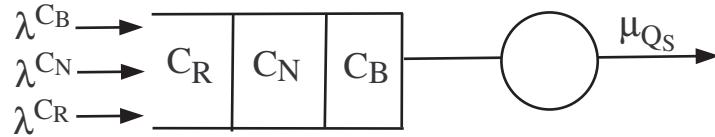


Figure 5.3: Single band mobile router architecture.

ure 5.3 shows the architecture of a single band MR with arrival rates of different class of traffic: signaling traffic or Binding Update (C_B), non-real-time (C_N), and real-time (C_R) traffic with λ^{C_B} , λ^{C_N} and λ^{C_R} arrival rates. All the traffic is queued and served by the single band with rate (μ_{Q_S}) based on the priority level of each class. Generally, C_B packets are given the highest priority, then C_N and C_R are served [123–125]. A problem of priority scheduling (of different traffic classes) of single band architectures is that one type of packet may be served continuously while others may suffer starvation. To prevent such starvation, a threshold is used for each class. However, identifying an optimum threshold is another problem. In the model, absolutely non-preemptive priority is used for each class.

5.4 Current Multi-Band Router Architecture

Commercial (simultaneous) multi-band MRs available today uses two different bands (2.4 GHz and 5 GHz) for different types of devices in a home network. Laptops may connect to 2.4 GHz network while Wi-Fi-enabled TV and gaming devices may connect to 5 GHz network. This reduces interference with the heavily-used 2.4 GHz network (as cordless phones, microwave uses the similar band). In addition, video streaming can be done through the high frequency band. The main principle of today's simultaneous multi-band

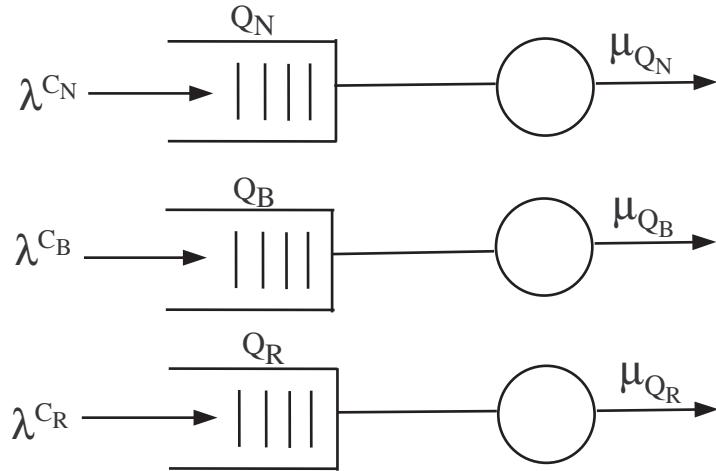


Figure 5.4: Architecture of a current (simultaneous) multi-band mobile router.

MR architecture is the non-sharing of bands among different flows of traffic. Moreover, some of the devices today (such as, IPTV) mostly deal with real-time traffic. Based on this fact, it is assumed that each of the band of current simultaneous multi-band MR only deals with one type of traffic. This might be a slight deviation from the real MR used today. However, this is assumed to compare the proposed architecture with current simultaneous multi-band MR architecture.

Figure 5.4 shows the current architecture of a simultaneous tri-band MR. Here, three bands are assumed to be used for three different classes of traffic: signaling traffic or Binding Update (C_B), real-time (C_R) and non-real-time (C_N) traffic. Each class of traffic is solely assigned to each designated frequency band as shown in Figure 5.4 and the

corresponding queues are named as Q_B , Q_R and Q_N . There will be absolutely no sharing of traffic among different bands even if one (or more) bands are under-utilized due to low traffic arrivals to those queues.

5.5 Proposed Multi-Shared-Band Router Architecture

In this section, the proposed architecture of multi-band MRs, that promotes sharing of bands to maximize system utilization, is explained. Three different queues are considered

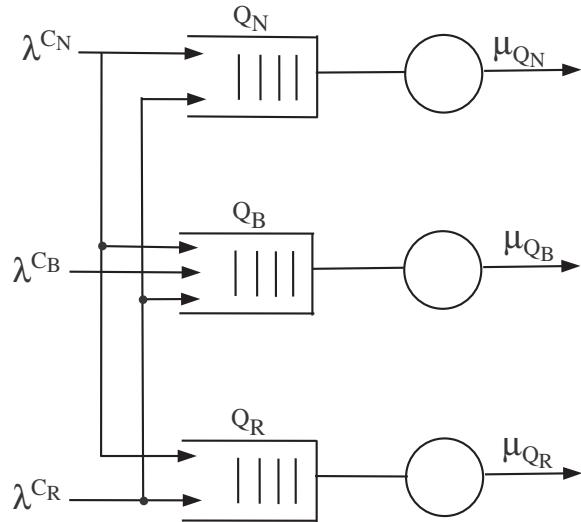


Figure 5.5: Proposed architecture of a simultaneous multi-shared-band mobile router.

(shown in Figure 5.5), each of which corresponds to a frequency band of a simultaneous tri-band MR. As illustrated in Figure 5.4, three classes of traffic are also considered and each queue is designated for each class of traffic. However, unlike the current multi-band architecture, in this multi-shared-band architecture (see Figure 5.5), data traffic of a class can flow through other queues which have empty slots, thereby ensuring better utilization of buffer spaces available. For example, if the Q_B has some empty spaces available and bursty C_R traffic comes in, the overflowed C_R traffic can be queued in the Q_B and subsequently served (or sent) through the Q_B -server (transmitter). Therefore, the communication model can be presented as shown in Figure 5.6.

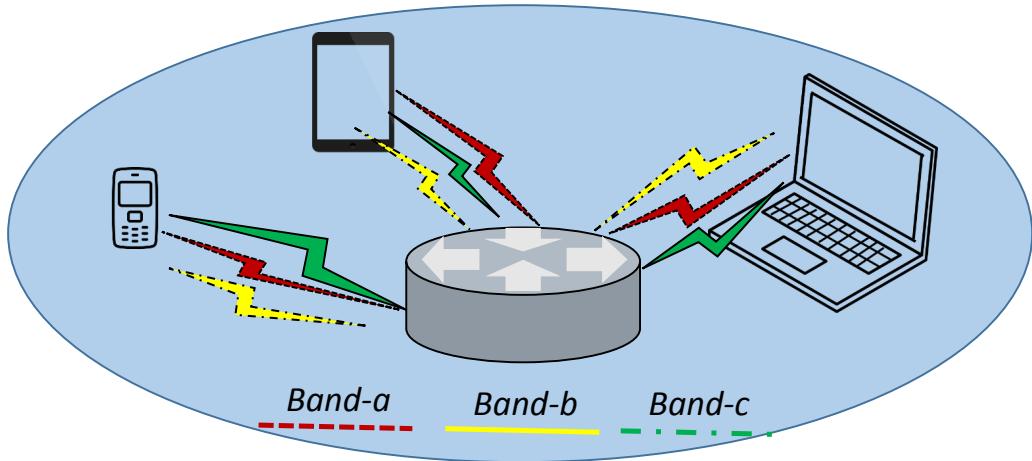


Figure 5.6: A multi-shared-band router communication scenario.

Similar to the current multi-band, the multi-shared-band routers can also use a number of bands to communicate with devices as shown in Figure 5.6. However, the devices can simultaneously send and receive data over different bands. Therefore, all bands are utilized and the system can have higher throughput.

5.5.1 Time and Space Priority

The time and space priority for the three queues of the scheduling for the proposed multi-shared-band architecture are explained in Figures 5.7, 5.8 and 5.9.

5.5.1.1 Q_B Queue

For Q_B , C_B packets have the highest priority; C_R and C_N packets have dynamic priority based on arrival rates (see equations(5.1) and (5.2)). Regarding space priority, C_B packets are queued in front of Q_B and if there are empty spaces available, other types (C_R and C_N) can be accommodated as shown in Figure 5.7.

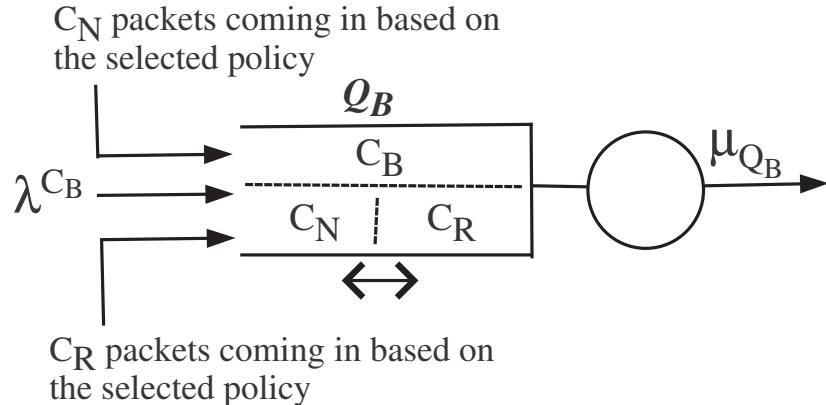


Figure 5.7: Queue corresponding to Q_B band.

5.5.1.2 Q_R Queue

Q_R can have only C_R and C_N packets as shown in Figure 5.8. C_R traffic has higher priority over C_N traffic. Therefore, Q_R can have C_N packets only if C_R packets cannot fill Q_R at any instant and there are C_N packets overflowed from Q_B or Q_N .

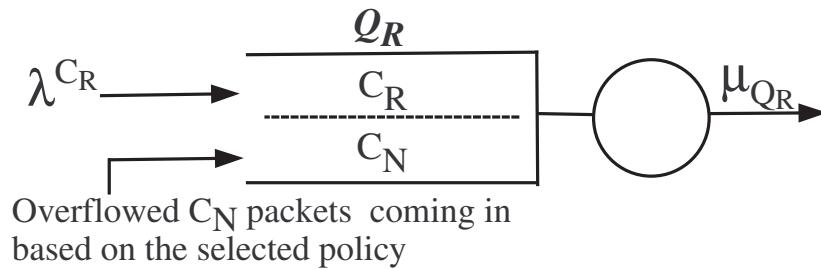


Figure 5.8: Queue corresponding to Q_R band.

5.5.1.3 Q_N Queue

Finally, Figure 5.9 shows Q_N which is designated for C_N traffic. However, if there are empty spaces in this queue, overflowed C_R traffic out of Q_B or Q_R can be enqueued in Q_N (see Figure 5.5).

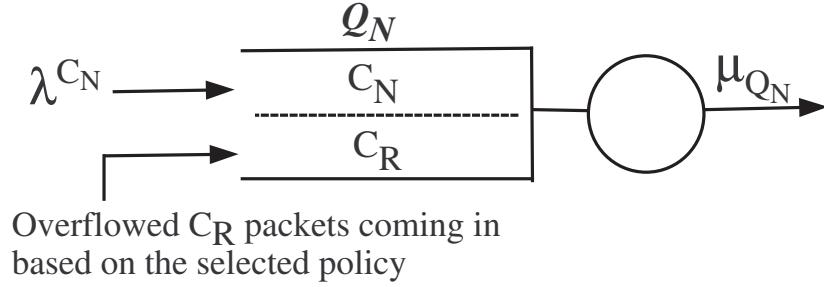


Figure 5.9: Queue corresponding to Q_N band.

5.5.2 Scheduling Algorithm

The following two crucial factors are considered to ensure improved performance of the multi-band MR:

- The unused buffer space of one queue (or band) can be used for other traffic types, thereby reducing the idle time of the system.
- Priorities of different traffic classes are also considered while selecting a particular type of packet over others.

Three types of allocation policies are used in the proposed multi-shared-band architecture:

- Fastest server first (FSF),
- Least utilization first (LUF),
- Slowest server first (SSF).

Queue allocation policies are explained as follows: (i) Attempts are first made to send different class of traffic through the designated frequency band; (ii) If there is overflow of C_R or C_N packets from Q_R or Q_N , they are forwarded to other servers based on the two principles: Faster server first and slowest server first (computed by comparing μ_{Q_B} , μ_{Q_N} , and μ_{Q_R}) or lower utilization server (computed by $\lambda^{C_B} / \mu_{Q_B}$ for Q_B , $\lambda^{C_N} / \mu_{Q_N}$ for Q_N ,

and $\lambda^{C_R} / \mu_{Q_R}$ for Q_R); (iii) If there is no space available in the chosen queue, the packets are queued in the third queue (if there is space in it). Otherwise, packets are dropped from the system; (iv) The race between different classes of traffics is resolved through the use of priority explained in Section 5.6.3; (v) A similar policy is enforced while dealing with each class of traffic.

5.6 Analysis

In this section, explanation for computation of various metrics in single band, current multi-band and the proposed multi-shared-band systems is given.

5.6.1 Assumptions

To make the model analytically tractable, the following assumptions have been made:

- (i) Packet arrival follows Poisson distribution; (ii) Type of queue discipline is FIFO with non-preemptive priority among various traffic classes.

5.6.2 Notations

The notations used in the analysis are listed in Table 5.1. To simplify the notation, $C \in \{C_B, C_N, C_R\}$ is used as the common notation for different traffic class types and $Q \in \{Q_B, Q_N, Q_R\}$ as the common notation for different queue types.

5.6.3 Priority

Priorities of different classes are taken into account while allowing traffic into Q_B . Priority of C_B packets in Q_B is $\sigma_{Q_B}^{C_B} = 1$. Priorities of other classes of traffic in Q_B are measured as follows:

$$\sigma_{Q_B}^{C_N} = \frac{\lambda^{C_R}}{\lambda^{C_B} + \lambda^{C_R} + \lambda^{C_N}} \quad (5.1)$$

Table 5.1: The notations for Chapter 5.

N_Q	Queue size of Q
λ^C	Total packet arrival rate of class C
λ_{C_B}	λ^{C_B}
μ_Q	Service rate at Q
$\sigma_{Q_B}^C$	Priority of class C traffic in Q_B
δ^C	Average delay of class C
n^C	Average occupancy of class C
D^C	Drop probability of class C
γ^C	Throughput of class C
δ_Q	Average delay of packets in Q
n_Q	Average occupancy of packets in Q
D_Q	Drop probability of packets from Q
γ_Q	Throughput of Q
δ_Q^C	Average delay of class C in Q
n_Q^C	Average occupancy of class C in Q
D_Q^C	Drop probability of class C from Q
χ_Q^C	Total dropped packets of class C from Q
γ_Q^C	Throughput of class C in Q
δ	Average delay of packets in the system
n	Average occupancy of packets in the system
D	Drop probability of the system
γ	Throughput of the system

$$\sigma_{Q_B}^{C_R} = \frac{\lambda^{C_N}}{\lambda^{C_B} + \lambda^{C_R} + \lambda^{C_N}} \quad (5.2)$$

If there are space in Q_B and both C_N and C_R types packets are overflowed from Q_N and Q_R , respectively, they will be enqueued to Q_B based on the priority equations (5.1) and (5.2). If $\sigma_{Q_B}^{C_N} \geq \sigma_{Q_B}^{C_R}$, C_N packets will be enqueued after C_B packets. Otherwise, C_R packets will be enqueued after C_B packets. While Q_B is full and a C_B packet or higher priority packet is arrived, the lowest priority packet is dropped. Similarly, C_N packets are queued to Q_R if there is a space in Q_R . While Q_R is full and there are C_N packets in Q_R , arrived C_R packets are enqueued by dropping C_N packets except C_N packet in service. Similar policy is used for Q_N .

5.6.4 Performance Metrics of Current Multi-Band Systems

The class performance metrics can be obtained for current multi-band system by using M/M/1/N [126] standard formula. Standard equations of M/M/1/N [126] are listed as follows:

$$n_{MM1N} = \begin{cases} \frac{\rho - (N+1)\rho^{N+1} + N\rho^{(N+2)}}{(1-\rho)(1-\rho^{N+1})}, & \text{if } \rho \neq 1 \\ \frac{N}{2}, & \text{if } \rho = 1 \end{cases} \quad (5.3)$$

$$D_{MM1N} = \begin{cases} \frac{\rho^N(1-\rho)}{1-\rho^{N+1}}, & \text{if } \rho \neq 1 \\ \frac{1}{N+1}, & \text{if } \rho = 1 \end{cases} \quad (5.4)$$

where the service rate, arrival rate and buffer size denoted by μ , λ and N , respectively, and $\rho = \lambda/\mu$. Therefore, the performance metrics for each class can easily be obtained because of non-shared bands. Moreover, the average occupancy, delay, drop rate and throughput of each priority class can be measured by using some previous works [127, 128] in the literature. The drop probability of each class [127], average class occupancy and delay [128] have been analytically formalized for non-preemptive priority classes by considering randomized push out mechanism. However, verification of analytical formulas is not presented. Therefore, the verification of the developed formulas for the proposed architecture will be presented in next sections.

5.6.5 Performance Metrics of the Proposed Multi-Shared-Band System

In this section, drop probability, average queue length, average delay and throughput are derived for the proposed multi-shared-band system.

5.6.5.1 Total Arrival Rates in Each Queue

For queuing analysis of the proposed system, we need to determine the total arrival rate of all class of traffic in each queue. In the proposed system, the total arrival rate of each queue not only depends on the arrival rate of that particular queue, but also depends on the overflows of packets from the other queues. In general, overflow in a queue can happen when there is no buffer space left.

Q_B -queue: For Q_B , the number of packets overflowed from Q_N and Q_R queues goes to Q_B . Thus, the arrival rates of NRT and RT packets to the Q_B (denoted by $\lambda_{Q_B}^{C_N}$ and $\lambda_{Q_B}^{C_R}$) can be obtained as follows:

$$\lambda_{Q_B}^{C_N} = \lambda_{Q_N}^{C_N} D_{Q_N} \quad (5.5)$$

$$\lambda_{Q_B}^{C_R} = \lambda_{Q_R}^{C_R} D_{Q_R} \quad (5.6)$$

where $D_{Q_N}^{C_N}$ and $D_{Q_R}^{C_R}$ are packet drop probabilities of NRT packets from Q_N and RT packets from Q_R . Now, the total (effective) arrival rate of all class of traffic in Q_B can be obtained as follows:

$$\lambda_{Q_B} = \lambda^{C_B} + \lambda_{Q_B}^{C_N} + \lambda_{Q_B}^{C_R} \quad (5.7)$$

N -queue: Q_N is designated for NRT traffic. However, if there are empty buffer spaces available in Q_N (due to low NRT arrival rate), this queue can be used to transmit RT traffic that is overflowed from the Q_B . Let $\lambda_{Q_N}^{C_R}$ denotes the arrival rate of RT packets in Q_N . Therefore, the total arrival rate (of both NRT and RT packets) in the Q_N is as follows:

$$\lambda_{Q_N} = \lambda^{C_N} + \lambda_{Q_N}^{C_R} \quad (5.8)$$

If all RT packets cannot be accommodated in Q_N , the rest of RT packets are dropped from the system.

Q_R -queue: Q_R is designated for RT traffic. However, if there is empty buffer space available in the Q_R (due to low RT arrival rate), this queue can be used to transmit NRT traffic which has been overflowed from the Q_B . Therefore, the total arrival rate (of both RT and NRT) in Q_R is as follows:

$$\lambda_{Q_R} = \lambda^{C_R} + \lambda_{Q_R}^{C_N} \quad (5.9)$$

If the empty buffer space in Q_R is higher than the overflowed NRT packets, then all NRT packets can be transmitted through Q_R . Otherwise, NRT packets are dropped from Q_R .

5.6.5.2 Computing Drop Probability

The packet drop probability of RT packets in Q_R can be obtained using standard M/M/1/N drop probability equation (5.4) as follows [126]:

$$D_{Q_R}^{C_R} = \frac{\rho_{Q_R}^{C_R}(1 - \rho_{Q_R}^{C_R})}{1 - (\rho_{Q_R}^{C_R})^{N_{Q_R}} + 1} \quad (5.10)$$

where $\rho_{Q_R}^{C_R} = \frac{\lambda_{Q_R}^{C_R}}{\mu_{Q_R}}$. Similarly, the packet drop probability of NRT packets in Q_N can be obtained as follows:

$$D_{Q_N}^{C_N} = \frac{\rho_{Q_N}^{C_N}(1 - \rho_{Q_N}^{C_N})}{1 - (\rho_{Q_N}^{C_N})^{N_{Q_N}} + 1} \quad (5.11)$$

where $\rho_{Q_N}^{C_N} = \frac{\lambda_{Q_N}^{C_N}}{\mu_{Q_N}}$. Let us assume that the priority of RT packets is higher than that of NR packets in Q_B . Therefore, while computing RT packet drop probability in Q_B , we can consider only BU and RT packets in Q_B by ignoring NRT packets for simplicity. Let us define utilization in Q_B considering only BU and RT packets be $\rho_{Q_B}^{C_{BR}} = \frac{\lambda_{Q_B}^{C_B} + \lambda_{Q_B}^{C_R}}{\mu_{Q_B}}$. Thus, the packet drop probability of BU packets in Q_B , denoted by $D_{Q_B}^{C_B}$, can be obtained by using derived formulas in [129] as follows:

$$D_{Q_B}^{C_B} = \frac{\rho_{Q_B}^{C_{BR}} \rho_1^{N_{Q_B}} (1 - \rho_1) (1 - (\rho_{Q_B}^{C_{BR}})^{N_{Q_B}+1})}{(1 - \rho_1^{N_{Q_B}+1}) (1 - (\rho_{Q_B}^{C_{BR}})^{N_{Q_B}+2})} \quad (5.12)$$

where $\rho_1 = \lambda_{Q_B}^{C_B} / \mu_{Q_B}$. Using equation (5.12), the packet drop probability of RT packets in Q_B can be obtained according to [129] as follows:

$$\begin{aligned} D_{Q_B}^{C_R} &= \frac{(1 - \rho_{Q_B}^{C_{BR}})}{(1 - (\rho_{Q_B}^{C_{BR}})^{N_{Q_B}+2})} (\rho_{Q_B}^{C_{BR}})^{N_B+1} \\ &+ \frac{\lambda_{Q_B}^{C_B}}{\lambda_{Q_B}^{C_R}} \left(\frac{(1 - \rho_{Q_B}^{C_{BR}})}{(1 - (\rho_{Q_B}^{C_{BR}})^{N_{Q_B}+2})} (\rho_{Q_B}^{C_{BR}})^{N_{Q_B}+1} - D_{Q_B}^{C_B} \right) \end{aligned} \quad (5.13)$$

Hence, RT packet arrival rate in Q_N is the dropped packets from Q_R , Therefore, it can be obtained as follows:

$$\lambda_{Q_N}^{C_R} = \lambda_{Q_B}^{C_R} D_{Q_B}^{C_R} \quad (5.14)$$

The total arrival to Q_N is the sum of two arrival rates λ_{Q_N} and $\lambda_{Q_N}^{C_R}$ (see equation (5.8)); the former has the higher priority than the latter. Therefore, following a similar approach as in Equation (5.12), we can compute $D_{Q_N}^{C_N}$. Then we can follow similar approach in Equation (5.13) to compute $D_{Q_N}^{C_R}$ which is the final drop of R-type packets from the system. That is, $D_{Q_N}^{C_R} = D_{Q_N}^{C_R}$. The computation of NRT packet drop probability follows similar steps as followed for RT packets and can be represented by $D_{Q_R}^{C_N} = D_{Q_R}^{C_N}$.

5.6.5.3 Average Queue Length

Each queue behaves as M/M/1/N queue when total packet arrival rates are calculated. Therefore, the estimated queue length can be obtained by using standard M/M/1/N queue length formulas as follows [126]:

$$n_{Q_T} = \begin{cases} \frac{\rho_{Q_T} - (N_{Q_T} + 1)(\rho_{Q_T})^{N_{Q_T}+1} + N_{Q_T}(\rho_{Q_T})^{(N_{Q_T}+2)}}{\left(1 - \rho_{Q_T}\right)\left(1 - (\rho_{Q_T})^{N_{Q_T}+1}\right)}, & \text{if } \rho_{Q_T} \neq 1 \\ \frac{N_{Q_T}}{2}, & \text{if } \rho_{Q_T} = 1 \end{cases} \quad (5.15)$$

The average occupancy of RT class depends on the number of RT packets in all queues (Q_R , Q_B and Q_N). Therefore, sum of the RT packet occupancies in all queues gives the occupancy of RT class in the system and this is computed as follows:

$$\begin{aligned} n^{C_R} &= n_{Q_R}^{C_R} + n_{Q_B}^{C_R} + n_{Q_N}^{C_R} \\ &= n_{Q_R}^{C_R} + \left(n_{Q_B}^{C_{BR}} - n_{Q_B}^{C_B} \right) + \left(n_{Q_N}^{C_{NR}} - n_{Q_N}^{C_N} \right) \end{aligned} \quad (5.16)$$

For computing $n_{Q_R}^{C_R}$ which is the RT class occupancy in Q_R , we need to put $N_{Q_T} = N_{Q_R}$, $\rho_{Q_T} = \rho_{Q_R} = \lambda_{Q_R}/\mu_R$ in equation (5.15). Similar approach can be used for the rest of the terms in equation (5.16).

To find BU class occupancy, only Q_B should be considered because BU packets are only queued in Q_B , therefore,

$$n^{C_B} = n_{Q_B}^{C_B} \quad (5.17)$$

To compute average occupancy of NRT packets in the system, a similar approach as in equation (5.16) can be used. Therefore,

$$n^{C_N} = n_{Q_N}^{C_N} + n^{C_N Q_B} + n^{C_N Q_R} \quad (5.18)$$

5.6.5.4 Throughput

The throughput of T class of traffic can be obtained according to M/M/1/N standard throughput formula [126] as follows:

$$\gamma^{C_T} = (1 - D^{C_T}) \lambda^{C_T} \quad (5.19)$$

5.6.5.5 Average Delay

Thus, the average packet delay of BU class depends on only Q_B because BU packets are served by only Q_B . Therefore, by Little's law,

$$\delta^{C_B} = \frac{n^{C_B}}{\gamma^{C_B}} \quad (5.20)$$

However, the average delays of RT and NRT class depend on all queues. Therefore, all queues must be considered for to measure average delay. Hence, the average delay of RT is measured by considering Q_R , Q_N and Q_B . However, because of the priority orders of different class traffic in different queues, RT packets experience delay of BU packets in Q_B , delay of NRT packets in Q_N and delay of RT packets in Q_R . Therefore, the overall average delay can be obtained as follows:

$$\delta^{C_R} = \frac{\gamma_{Q_B}^{C_R}(\delta_{Q_B}^{C_B} + \frac{n_{Q_B}^{C_R}}{\gamma_{Q_B}^{C_R}}) + \gamma_{Q_N}^{C_R}(\delta_{Q_N}^{C_N} + \frac{n_{Q_N}^{C_R}}{\gamma_{Q_N}^{C_R}}) + \gamma_{Q_R}^{C_R}(\frac{n_{Q_R}}{\gamma_{Q_R}^{C_R}})}{\gamma^{C_R}} \quad (5.21)$$

and by considering that RT has higher priority than NRT packets in Q_B , we have to use $\delta_{Q_B}^{C_R}$ rather than $\delta_{Q_B}^{C_B}$. NRT packets experience delay of BU and RT packets in Q_B , delay of NRT packets in Q_N and delay of RT packets in Q_R . Therefore, the overall average delay can be obtained as follows:

$$\delta^{C_N} = \frac{\gamma_{Q_B}^{C_N}(\delta_{Q_B}^{C_{BR}} + \frac{n_{Q_B}^{C_N}}{\gamma_{Q_B}^{C_N}}) + \gamma_{Q_R}^{C_N}(\delta_{Q_R}^{C_R} + \frac{n_{Q_R}^{C_N}}{\gamma_{Q_R}^{C_N}}) + \gamma_{Q_N}^{C_N}(\frac{n_{Q_N}}{\gamma_{Q_N}^{C_N}})}{\gamma^{C_N}} \quad (5.22)$$

5.6.6 Overall System Performance

In order to compare overall performances of the single band, current multi-band and proposed multi-shared-band, the following calculations are used.

5.6.6.1 Average Occupancy of the System

The total average occupancy of multi-band architectures (in all queues) can be computed by summing all occupancies of classes as follows:

$$n = n^{C_B} + n^{C_N} + n^{C_R} \quad (5.23)$$

5.6.6.2 Drop Probability of the System

After finding the drop probability of each class, the overall drop probability of multi-band systems is computed by averaging from all arrivals to dropped packets as follows:

$$D = \frac{\lambda^{C_B} D^{C_B} + \lambda^{C_N} D^{C_N} + \lambda^{C_R} D^{C_R}}{\lambda^{C_B} + \lambda^{C_N} + \lambda^{C_R}} \quad (5.24)$$

5.6.6.3 Throughput of the System

The total throughput of multi-band systems can be obtained by summing of each class throughput as follows:

$$\gamma = \gamma^{C_B} + \gamma^{C_N} + \gamma^{C_R} \quad (5.25)$$

5.6.6.4 Average Delay of the System

The average delay in multi-band systems can be obtained by averaging delays of three classes as follows:

$$\delta = \frac{\gamma^{C_B} \delta^{C_B} + \gamma^{C_N} \delta^{C_N} + \gamma^{C_R} \delta^{C_R}}{\gamma} \quad (5.26)$$

5.7 Simulation of the System

A discrete event simulation is implemented in Matlab environment by taking into account all the assumptions and scheduling policies mentioned in Sections 5.3, 5.4 and 5.5. M/M/1/N and M/M/3/N [126] procedures are followed for the implementation of the simulation. Buffer lengths are kept equal (of 50 packets) for each multi-band queue. Buffer lengths are kept small [102], similar to real routers to reduce packet delay. However, to

have fair comparison with the single band architecture, the total buffer length for single band is used to hold 150 packets which is three times of the length of a multi-band buffer. C_R and C_N packets are assumed to be 512 bytes [23, 130] whereas C_B packets are assumed to be 64 bytes. The service rates of Q_B , Q_N and Q_R are kept 27, 75 and 132 packets/sec which is proportional to service rates of multi-band routers [23] according to extended bandwidth. The single band router can only have one band; therefore, the highest service rate in multi-band systems (i.e., 132 packets/sec) is used for the service rate of the single band. Each simulation is run with 10000 samples for 20 trials having different traffic class arrival rates as follows:

$$\lambda^{C_B} = (i) = \{ i \}, \lambda^{C_N} = \{ 3\lambda^{C_B} \}, \lambda^{C_R} = \{ 10\lambda^{C_B} \} \text{ where } i = 1, 2, 3, \dots, 20.$$

Arrival rates of all types of traffic are increased in the simulation to observe the impact of heavy traffic on the systems. The arrival rates of C_B and C_N type packets are increased slowly in each trial whereas C_R traffic arrival rate is increased at a much higher rate. This will saturate Q_R and the impact of this overflow on different performance metrics of the proposed system and the current existing systems is explained.

5.8 Results

In this section, validation of analytical approximations is presented and the single, current multi-band and multi-shared-band systems are compared.

5.8.1 Validation of Developed Analytical Formulas

In this subsection, the simulation and analytical results are presented for the proposed system to compare and validate the correctness of the derived analytical approximations in Section 5.6 ¹.

¹Both simulation and analytical expressions are results of SSF allocation policy.

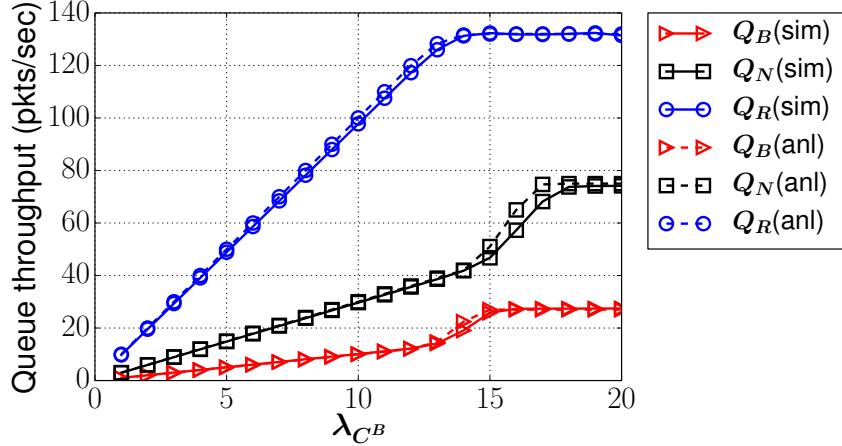


Figure 5.10: Queue throughput of the multi-shared-band system obtained through simulations and analytical model.

5.8.1.1 Queue Throughput

Figure 5.10 represents the throughput of queues for the multi-shared-band system. The simulation results closely match the analytical ones. The throughput of queues slowly increases when the packet arrival rates are raised. After the system reaches the maximum throughput capacity, they are fixed. The throughput of Q_B and Q_N is lower than Q_R throughput due to the lower service rates of Q_B and Q_N .

5.8.1.2 Average Class Occupancy

Figure 5.11 shows the average class occupancy of the multi-shared-band system obtained through simulations and analytical formulas. The simulation and analytical results are close to each other. The class occupancies of C_N and C_B are almost zero as their service rates are much higher than their arrival rates. However, this is not the case for C_R class where its arrival rate is higher than Q_R service rate. Hence, the excessive C_R packets are enqueued in other two queues. Due to priorities of C_B and C_N in Q_B and Q_N , the average occupancy of C_R is much higher than other classes of data traffic.

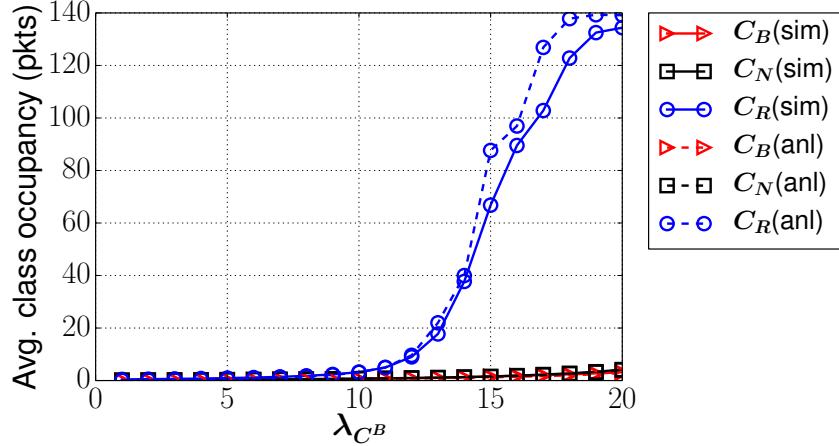


Figure 5.11: Average class occupancy of the multi-shared-band system obtained through simulations and analytical model.

5.8.1.3 Class Throughput

Figure 5.12 shows the class throughput for the multi-shared-band system. Again the simulation results closely match the analytical ones. The throughput of C_B and C_N is lower

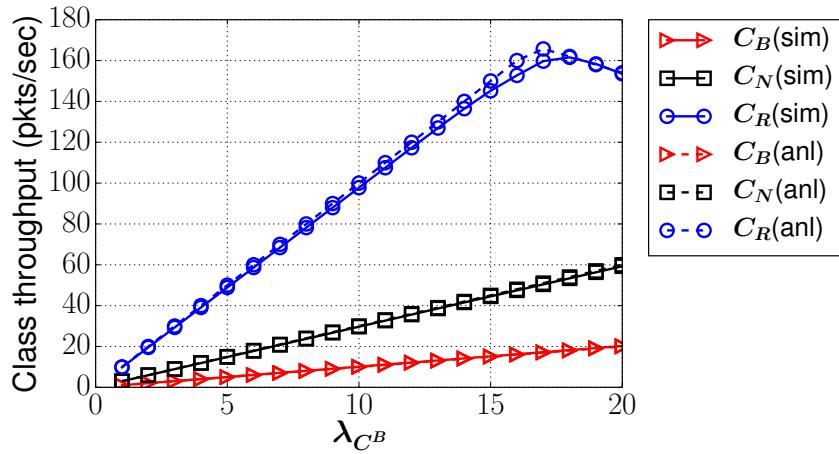


Figure 5.12: Class throughput of the multi-shared-band system obtained through simulations and analytical model.

than the throughput of C_R due to the low arrival rates of C_B and C_N . However, the throughput of C_R rises as high arrival rates in subsequent trials are used.

5.8.2 Comparison of Single, Current Multi-Band and Multi-Shared-Band Systems

In this subsection, SSF allocation policy is used in the proposed scheduling algorithm to compare the multi-shared-band system with the single and current multi-band systems because SSF allocation in the multi-shared-band system represents the worst case scenario.

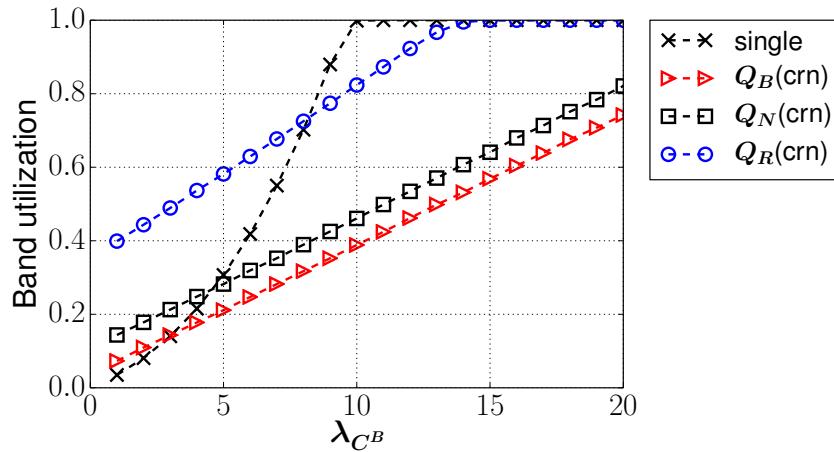


Figure 5.13: Band utilization of the single and current multi-band systems.

5.8.2.1 Band Utilization

Utilization is a performance measurement that indicates how efficiently bands are used and whether there are any unused resources of the system. Figure 5.13 shows the band utilization for single and current multi-band systems. The single band utilization is lower than utilization of all the bands of the current multi-band system for low arrival rates. Gradually increasing arrival rates saturates the single band and makes it reach its maximum capacity. Then all arrived new packets are dropped. However, the current multi-band system uses multiple queues to serve different packet types. In spite of some improvement comparing to the single band by considering high arrival rates, Q_R of the current multi-band system can be saturated by high C_R packets arrival. Therefore, Q_B and Q_N utilization is lower than Q_R utilization because of lower arrival rates of C_B and C_N comparing

to the arrival rate of C_R . On the other hand, the multi-shared-band system distributes

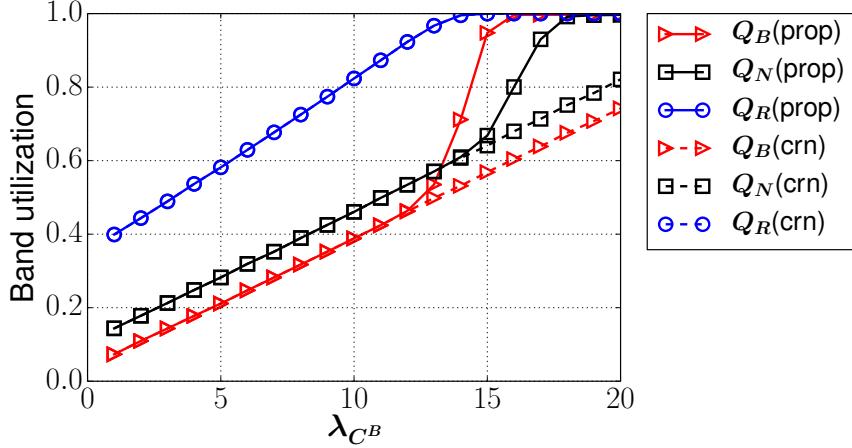


Figure 5.14: Band utilization of the current multi-band and multi-shared-band systems.

C_R packets to Q_B and Q_N as shown in Figure 5.14 (the proposed (shared) and current (non-shared) multi-band systems). When the arrival rates are low ($\lambda^{C_B} = 1, \dots, 13$), all the queues have similar utilization for the current and proposed multi-band systems. However, for $\lambda^{C_B} = 14, \dots, 20$, the utilization of Q_B and Q_N of the multi-shared-band systems is higher than the current multi-band ones. This is because an increasing number of C_R packets are dropped in the current system whereas in the proposed one, they are accommodated in Q_B and Q_N , thereby improving their utilization and maximizing the system performance.

5.8.2.2 Average Class Occupancy

Figure 5.15 shows the average packet occupancy of each class of traffic for the single and current multi-band systems. C_B and C_N occupancies are lower for the single and current multi-band systems because of low arrival rates and the priorities of C_B and C_N in the single band system. When packet arrival rates increase, C_R occupancy sharply increases in the single band. However, it slowly increases in the current multi-band system because of higher service rate of the multiple bands and using a different queue to process only C_R traffic rather than processing C_B and C_N then C_R type packets as in the single band.

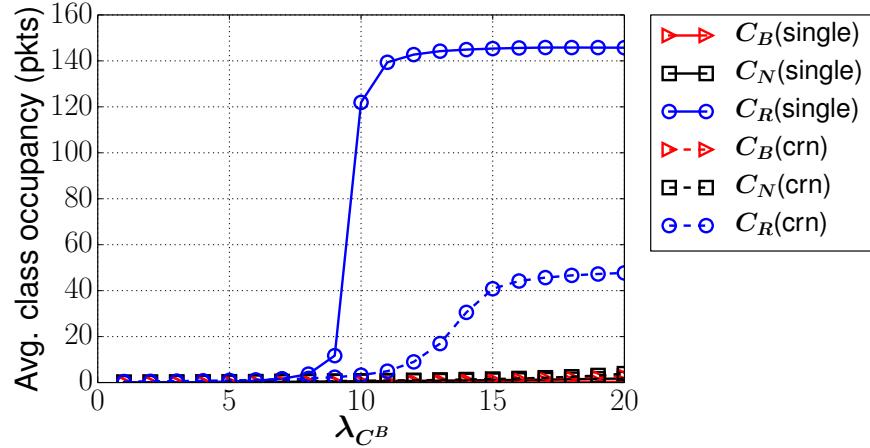


Figure 5.15: Average class occupancies of the single and current multi-band systems.

Figure 5.16 shows the average packet occupancy of each class of traffic for the proposed and current multi-band systems. The occupancy of C_R ($\lambda^{C_B} = 14, \dots, 20$) in the

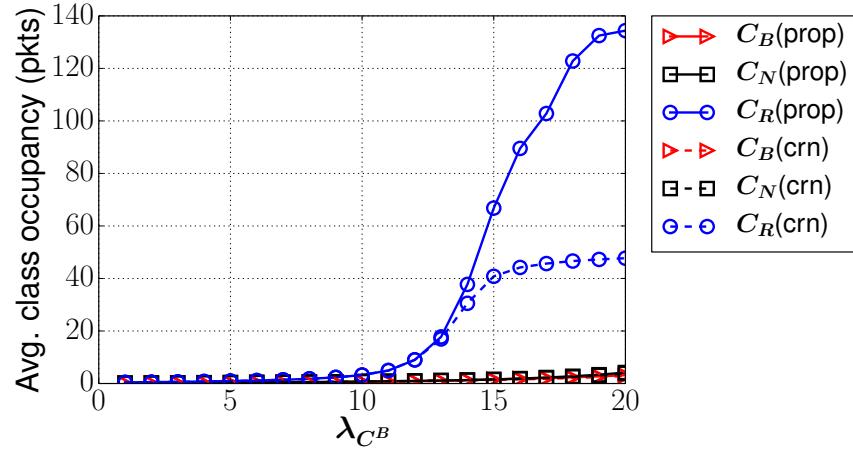


Figure 5.16: Average class occupancies of the current multi-band and multi-shared-band systems.

proposed system is higher than the current one. This is because excessive C_R packets are immediately dropped from the system in the current multi-band system and these lost packets do not come into account in occupancy calculations. On the contrary, in the proposed system the overflowed C_R packets get chances to be enqueued in Q_B and Q_N before being dropped. C_R packets are the third priority packets in Q_B and the second priority in

Q_N and they have to wait for C_B and C_N packets, respectively before being scheduled for service. Hence, it increases the occupancy of C_R packets.

5.8.2.3 Class Throughput

Figure 5.17 shows the throughput of each class for the single and current multi-band systems. C_B and C_N throughput is increasing in the single and current multi-band systems while the arrival rates increase until $\lambda^{C_B} = 10$. When C_N and C_B arrival rates increase,

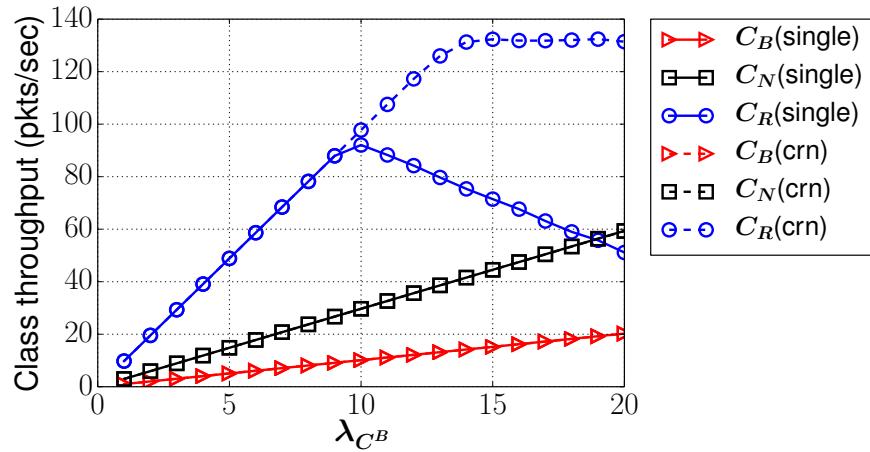


Figure 5.17: Class throughput of the single and current multi-band systems.

C_R throughput is getting lower for the single band due to the fact that the priority level of C_B and C_N type packets are higher than C_R type packets. However, C_R throughput in the current multi-band is higher than single band one and increasing until C_R traffic reaches the service rate of Q_R while it does not affect the throughput of C_B and C_N type packets because of distinct queues and servers. In the single band, it is expected to have lower throughput for C_N and C_R type packets while increasing C_B arrival rate because C_B has the highest priority in the single band.

Figure 5.18 shows the throughput of each class for the proposed and current multi-band systems. The throughput of C_N and C_B is increasing with the arrival rates for the both systems. However, for C_R class in the current system, the throughput is saturated at μ_{Q_R} ($= 132$ pkts/sec) when the C_R arrival rate reaches this value. However, C_R throughput

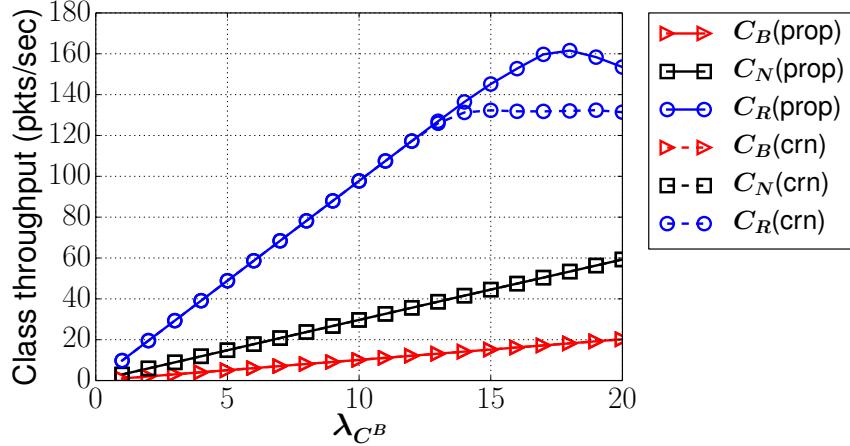


Figure 5.18: Class throughput of the current multi-band and multi-shared-band systems.

for the proposed system is growing much higher (due to sharing of other under-utilized bands) and reaches its peak value at $\lambda^{C_B}=18$. After that, it starts to decrease slowly due to the impact of the increased arrival rates of other classes (C_B and C_N) that results in less available space for the overflowed C_R packets.

5.8.3 Discussion on Comparison of Systems

According to above results, the following observations are obtained: (i) The performance of the proposed multi-band system (multi-shared-band) is better than the single band and current multi-band systems; (ii) Though improved band utilization in the multi-shared-band system, the both types of multi-band systems (proposed and current) do not use band efficiently as the single band while system is loaded with light data traffic; (iii) The highest priority class in single band can have lower delay than same class in both types of multi-band systems; (vi) Under heavy traffic loads, the lower priority class in single band has longer waiting time (in queue) than same class in both types of multi-band systems; (v) The multi-shared-band system significantly improves the throughput of the system while causing small amount of delay for the packets which are overflowed from their reserved queues.

5.8.4 Comparison of Single and Multi-Shared-Band Architectures with FSF, LUF, and SSF Allocations

In this subsection, the performances of FSF, LUF, and SSF allocation policies are also observed for the proposed multi-band (multi-shared-band) system and their performances are compared to the performance of the single band system. Because of the input parameters of the simulation, the performance of SSF is exactly matched with LUF performance. It is realized that overflow C_R packets are firstly enqueued to Q_B for LUF as it is in SSF because λ^{C_B}/μ_{Q_B} is always smaller than λ^{C_N}/μ_{Q_N} for the service rates. LUF is tested with other inputs when the overflowed C_R packets are firstly enqueued to interchangeable Q_B or Q_N . However, these service rates do not represent the real case scenario that was discussed at the beginning of Section 5.8. Therefore, input parameters are kept as realistic as possible. LUF/SSF is used to represent both LUF and SSF allocations in the following figures because their performances are equal under these circumstances.

5.8.4.1 Band Utilization, Packet Delay, and Drop Ratio of Single and Multi-Shared-Band Systems

The band utilization results for each trial are shown in Figure 5.19. The single band

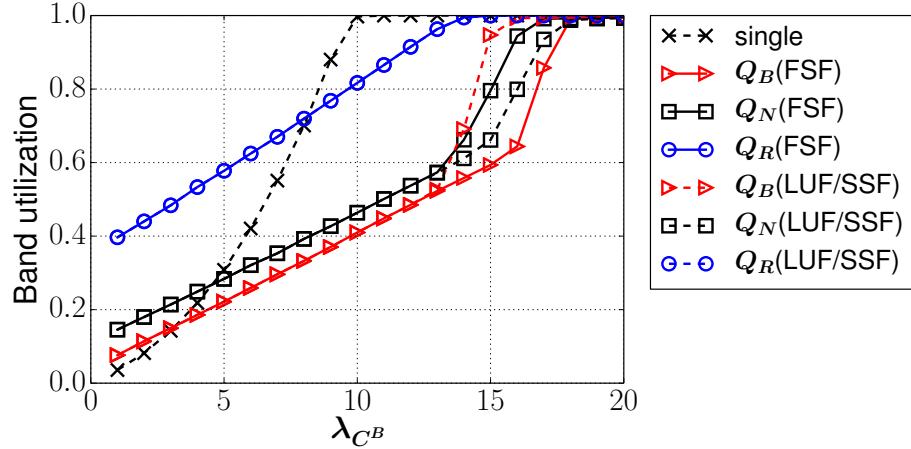


Figure 5.19: Band Utilization of Single, FSF, LUF, and SSF systems.

utilization is lower than the utilization of FSF and LUF/SSF until λ_{C_B} is equal to four. However, under heavy traffic loads, the multi-shared-band utilization increases gradually but not as fast as the single band utilization. Q_B utilization of FSF is the lowest because of the low arrival rates of C_B packets and forwarding of the overflowed C_R packets to Q_N first, then Q_B when the system overwhelms with C_R packets. The band utilization of LUF can be vary because forwarding C_R class packet to other queues depends on the rate of λ/μ (here every time overflowed C_R packets are forwarded to first Q_B , then Q_N as in SSF). Therefore, Q_B utilization of FSF is lower than Q_N utilization and it is reverse for LUF/SSF after $\lambda^{C_B} = 15$. When the system reaches the maximum capacity in the multi-shared-band, the utilization of bands is similar for both FSF and LUF/SSF allocations.

The average delay and drop probability results² for each trial are given in Figures 5.20 and 5.21, respectively. The average delay and drop probability of the single and multi-

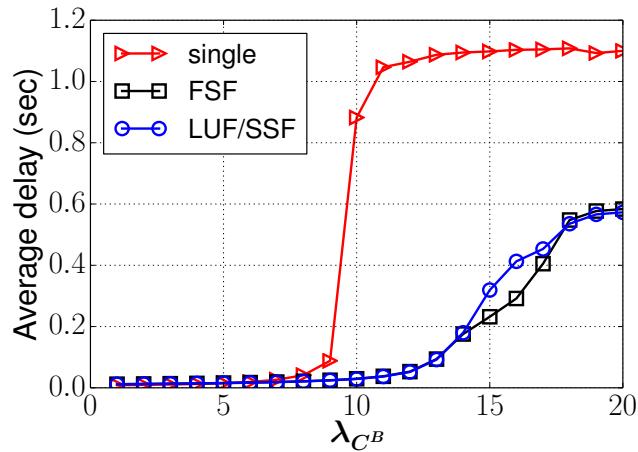


Figure 5.20: Average delays of Single, FSF, LUF, and SSF systems.

shared-band are significantly low while the systems are under light traffic loads (see Figures 5.20 and 5.21). However, under heavy traffic loads, the delay of the single band sharply increases and saturates at its maximum capacity. While there are no significant differences between the average delays of FSF and LUF/SSF in the multi-shared-band

²Here, they are measured according to explanations in Section 5.6.6

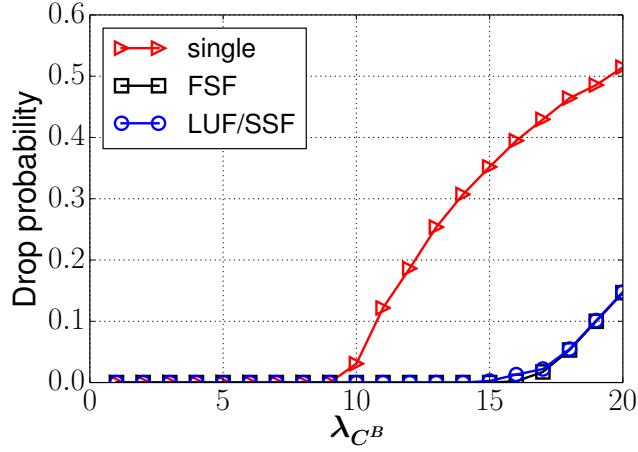


Figure 5.21: Drop probabilities of Single, FSF, LUF, and SSF systems.

system, their average delays are two times better than the average delay of the single band.

In Figure 5.21, the drop probabilities of FSF and LUF/SSF in the multi-shared-band are lower than the single band drop probability because the total service rate of the multi-shared-band is almost two times of that of the single band. Moreover, the drop probability of FSF is the lowest although it has similar drop probability with LUF/SSF.

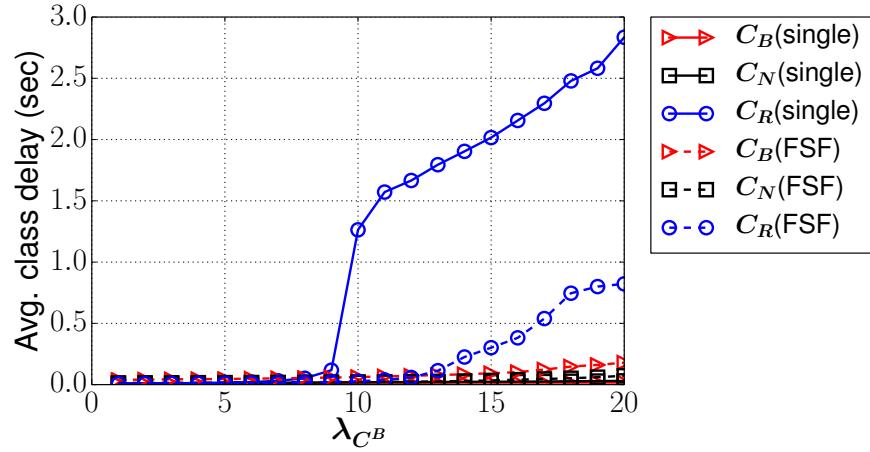


Figure 5.22: Average class delays of FSF and Single.

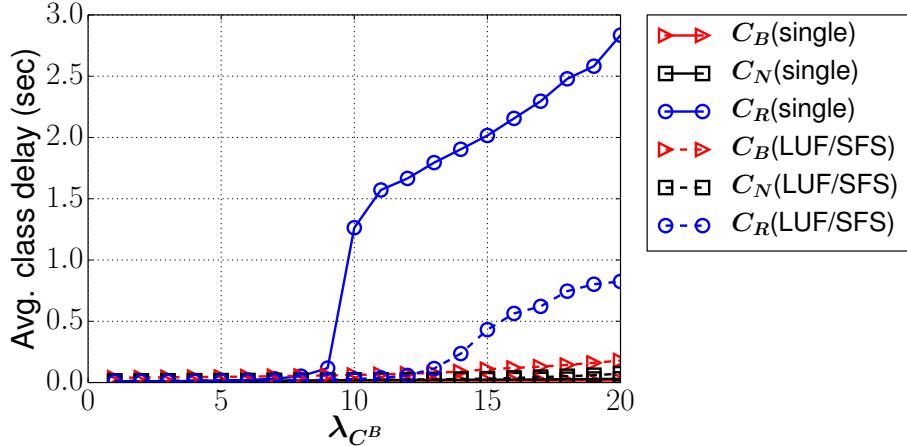


Figure 5.23: Average class delays of LUF, SSF and Single.

5.8.4.2 Average Class Delay and Throughput

The average delays for each class traffic are given in Figures 5.22, 5.23, and 5.24. C_B and C_N class delays in the single band are low because of their priorities (see Figure 5.22 and 5.23). Under heavy C_R traffic, C_R class delay sharply increases in the single band because of priority order of C_B , C_N , and C_R of the single band. Interestingly, the total service rate of the multi-shared-band is almost two times higher than the service rate of the single band, C_R class delay of the single band is at least three times higher than C_R class delay of the multi-shared-band (see Figures 5.22 and 5.23).

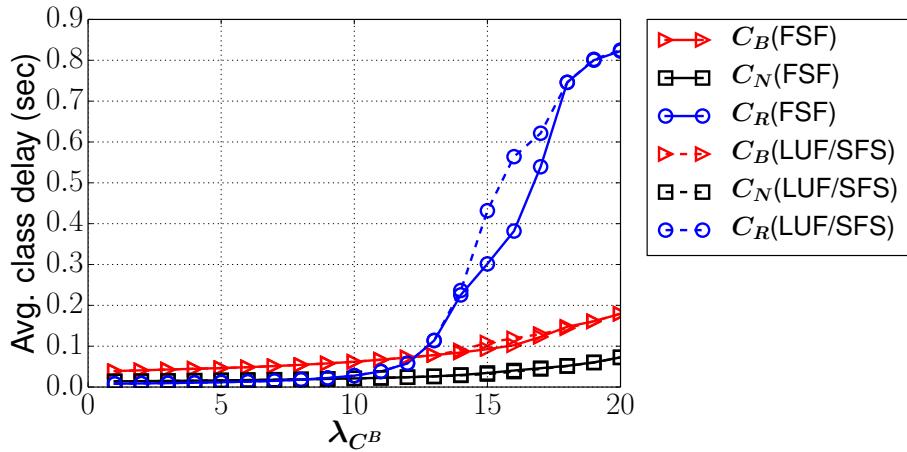


Figure 5.24: Average class delays of FSF, LUF, and SSF.

Although LUF/SSF allocation of the multi-shared-band system shows notable performance for C_R traffic, FSF is better than LUF/SSF (see Figure 5.24).

The class throughput results are given in Figures 5.25, 5.26, and 5.27. C_B and C_N

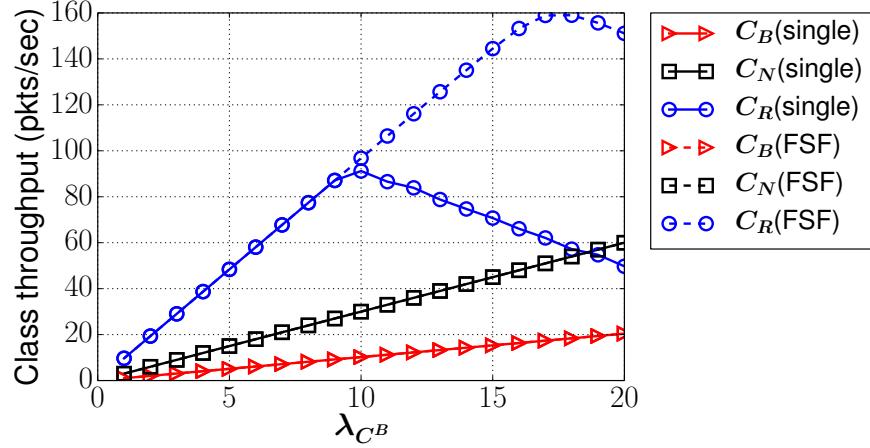


Figure 5.25: Class throughput of FSF and Single.

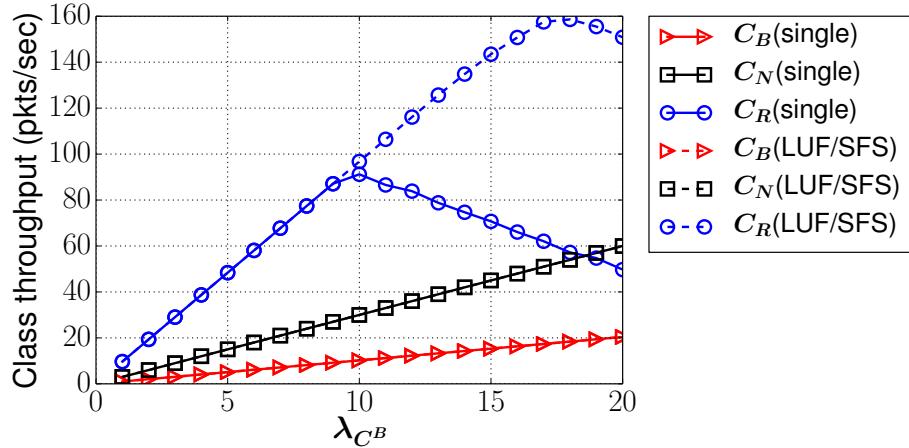


Figure 5.26: Class throughput of LUF, SSF, and Single.

throughput is increasing in the single and multi-shared-band system as arrival rates get higher as shown in Figures 5.26 and 5.27. When C_R , C_B and C_N traffic arrival rates increase, C_R traffic throughput is lower in the single band system because C_B and C_N traffics make the single band busy and C_R cannot get service. However, C_R throughput in the multi-shared-band system has a higher throughput and it is increasing until Q_N and Q_B

have spaces to serve C_R traffics. After Q_B and Q_N cannot serve, C_R throughput decreases and will be a constant after some point. It is also interesting that the performances of FSF and LUF/SSF are slightly different (see Figure 5.27).

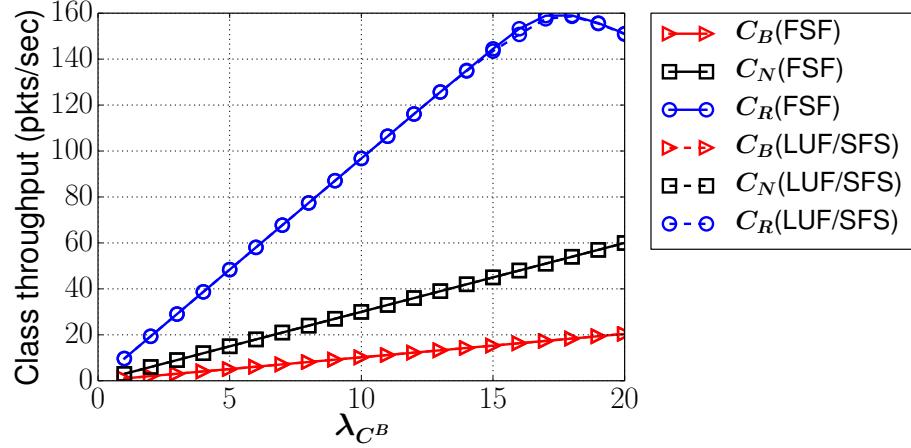


Figure 5.27: Class throughput of FSF, LUF, and SSF.

5.8.5 Discussion on Comparison of Single Band and Multi-Shared-Band Allocations

According to overall and class-wise analysis results, the following observations are obtained: (i) The multi-shared-band system (for three allocation policies) is better than the single band system for heavy traffic loads; (ii) FSF, LUF and SSF allocations do not use bands efficiently as the single band while system is under low traffic loads; (iii) LUF allocation shows similar performance with SSF or FSF; (iv) FSF allocation policy in multi-shared-band system has the best performance; (v) Although FSF has less delay than LUF and SSF for C_R class, there is no significant difference between throughput of FSF, LUF, and SSF allocation policies.

5.9 Summary

In this chapter, a novel multi-shared-band architecture with a scheduling algorithm was proposed for multi-band mobile routers that exploits band sharing. Analytical formulas of the multi-shared-band system were presented that were validated by extensive simulations. The performances of the single and multi-band systems under different scenarios were compared. After analyzing the systems according to class-wise, queue-wise and system-wise performances, the following results were obtained: (i) The proposed architecture can ensure maximum possible utilization through the sharing of capacities among the bands in the multi-band systems; (ii) The single band system with priority is recommended for light traffic loaded systems; (iii) The multi-band system is recommended for heavy traffic loaded systems; (iv) The priorities in single band plays a crucial role for efficient services; (v) Different allocation policies in the multi-shared-band system show similar effects on the system performances; (vi) The priority in the multi-band systems does not have impact on the system performance as much as the single band systems. The results obtained in this chapter should help network engineers to develop efficient routers, and also end-users to identify suitable routers to fulfill their needs.

In the next chapter, an energy efficient scheduling algorithm will be discussed for the proposed multi-share-band architecture.

Chapter 6

Energy Aware Scheduling and Queue Management for Multi-Shared-Band Router

In the previous chapter, the multi-shared-band architecture with several allocation policies was discussed and, the multi-shared-band, current multi-band and single band systems were compared. However, the energy consumptions of the systems have not been considered. Therefore, in this chapter, an energy aware scheduling algorithm for the multi-shared-band system is proposed to decrease the energy consumption of the proposed multi-shared-band system. Results show that the proposed energy aware scheduling algorithm uses system resources efficiently and decreases the energy consumption of the multi-shared-band system up to 60%.

6.1 Introduction

In the previous chapter, the benefits of the multi-shared-band system architecture with different allocation policies over the single-band system architecture were demonstrated through analytical analysis and an extensive simulation. However, using multiple bands increases the energy consumption because of active bands. Therefore, it is essential to consider the energy efficiency in the next generation systems because even energy billing

Results presented in this chapter have appeared in IEEE Wireless Communication and Networking Conference (WCNC) Workshop [131].

cost of a single band idle router system is \$27, (which is the highest cost for the idle gadgets at home) per year according to Ecotricity [26]. Using multiple bands increase the energy consumption more because of active multiple bands, thus the number of active bands and their energy consumptions must be taken into account not to waste energy. Therefore, it is necessary to propose an appropriate scheduling and queue management for multi-band MRs to ensure maximum possible utilization of the system resources [18,132] by decreasing the energy consumption. Hence, the *aim* of this chapter is to improve the energy efficiency of the proposed multi-shared-band system without decreasing the throughput ratio.

There have been several types of energy saving algorithm in the literature [133–136]. In [133], Wake-on-Line is proposed by following wake up and sleep procedure of remote devices to decrease the energy consumption of the connected devices. In [134], a single channel selection based on the energy consumption of the channel and quality of service requirements of users is used while providing services. In [135], an energy aware path selection method for MIMO capable wireless systems is discussed. In [136], the sizes of packets are decreased to save energy. Mostly, the above methods are developed when the system is actively transferring packets. However, when the system is idle, even a single band system consumes significant amount energy [26]. Therefore, in this chapter, the energy aware scheduling algorithm is developed for multi-shared-band systems when not only the system is active but also idle.

The *objective* of this work is to ensure maximize utilization of MRs by reducing the energy consumption through the band-sharing algorithm. The *contributions* of this work are: (i) Proposing an Energy Aware Scheduling Algorithm (e-ASA) to improve the utilization of the system while decreasing the energy consumption in MRs, (ii) developing an energy consumption model for MRs, and (iii) comparing the performance of the single, current multi-band system, and multi-shared-band system to multi-shared-band system with e-ASA in terms of resource usages, throughput ratios and energy consumptions. Results show that the proposed e-ASA uses system resources efficiently and decrease the

energy consumption of the multi-shared-band architecture up to 60% without reducing the system throughput ratio. The proposed energy aware scheduling algorithm and related analysis should help network engineers build next generation wireless routers by considering the energy usage of the systems.

The rest of the chapter is organized as follows: In Section 6.2, single and multi-band systems are discussed. e-ASA for multi-shared-band systems is explained in Section 6.3. Section 6.4 presents the simulation environment to analyze the performances of the systems. In Section 6.5, the simulation results are presented by showing band utilization, energy consumptions and throughput ratios. Finally, Section 6.6 has the summary of this chapter.

6.2 Single and Multi-Band Routers

In this section, the queuing models of single band and multi-band (non-shared and shared) systems are explained slightly different from what explained in the previous chapter. In the previous chapter, the priority based queue model is analyzed. On the contrary, here, systems are operated without the priorities of different data traffic classes as shown in Figures 6.1, 6.2 and 6.3. Therefore, the brief descriptions of the queue systems are given below.

Figure 6.1 shows a queuing model sample for the single band system. There is a

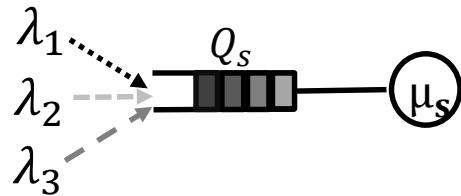


Figure 6.1: Single band queuing system.

server (μ_s) and a buffer (Q_s). The incoming data traffic ($\lambda_1, \lambda_2, \lambda_3$) is served by one server. When the server is busy and if there is an empty space in the buffer, a new arrived packet is enqueued to the buffer, Q_s .

However, the multi-band system has three servers (μ_a, μ_b, μ_c) and three buffers (Q_a, Q_b, Q_c) as shown in Figures 6.2 and 6.3. In current multi-band (see Figure 6.2), each user

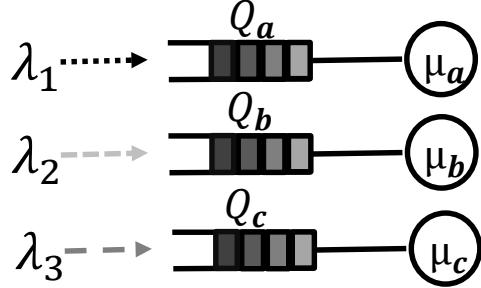


Figure 6.2: Current multi-band queuing system.

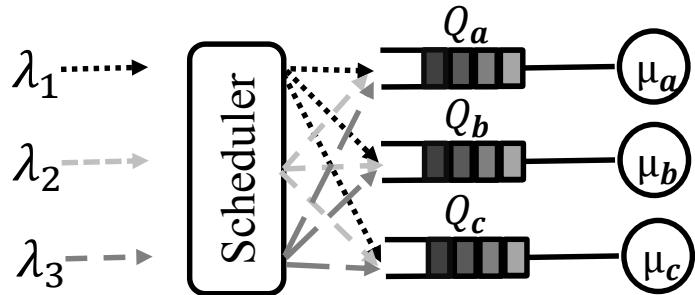


Figure 6.3: Multi-shared-band queuing system.

is served by only one server. However, multiple servers can simultaneously serve to each user in the multi-shared-band system (see Figure 6.3). Therefore, the band sharing model improves the system performance as explained in Chapter 5. However, because of multiple active bands, the energy consumption is significantly increasing. To decrease the energy consumption, Energy Aware Scheduling Algorithm (e-ASA) is developed for multi-shared-band system.

6.3 Energy Aware Scheduling Algorithm (e-ASA)

The energy aware scheduling algorithm is developed to decrease the energy consumption of the multi-shared-band system by following wake up and sleep procedures of the bands

according to incoming data traffic. For example, if users can be satisfied by one band, the system uses only one band for data transfer. To have such a system, three crucial parameters which are the total arrival rate, the total service rate of the bands and a predetermined threshold, are considered to developed e-ASA.

6.3.1 Notations

Notations in Table 6.1 will be used in the rest of the chapter to explain wake up sleep procedures.

Table 6.1: The notations for Chapter 6.

i	$\in \{s,a,b,c\}$
Q_i	Queue of μ_i
N_i	Size of Q_i
ρ_i	λ_i/μ_i
λ	$\sum_{i=1}^3 \lambda_i$
μ	$\mu_a + \mu_b + \mu_c$
θ	Throughput ratio
D	Drop probability
γ	Throughput
E	Overall energy consumption of the bands
α	Energy consumption of a band while data transfer
β	Energy consumption of a band during the idle time

6.3.2 Wake up and Sleep Modes of the Bands

The wake up and sleep procedures of the bands are used to save energy. However, it is crucial to decide when the bands change their status to wake up or sleep modes. To make decision of the band statuses, the data traffic load and desired QoS by users are used. For example, if one band is flooded with data, another band is activated. Here, it is very important to decide when one, two or three bands are needed according to data traffic

loads because the performance of the system depends on the band operations. Therefore, to efficiently determine the required number of bands, the throughput ratio of the system is measured for three possible cases (one, two, or three bands are active) by using queuing theory. The predetermined threshold for the throughput ratio (θ) (Throughput ratio is 1 - drop probability) and the expected throughput ratios¹ of three cases are used while determining how many bands need to be active. For example, assume that the expected throughput ratios are 0.7, 0.8, and 0.9 by using one, two and three bands, respectively. This means that if the threshold is 0.8 then only two bands become active.

To find the system throughput, M/M/1/N queuing model [98] is used. However, the more complex measurement can also be used but the complex measurement can increase the calculation cost of the system throughput. To find the system throughput, the drop probability (D) must be obtained. From M/M/1/N standard drop probability formula [126], drop probability can be written as

$$D = \begin{cases} \rho^N \frac{1-\rho}{1-\rho^{N+1}}, & \text{if } \rho \neq 1 \\ \frac{1}{N+1}, & \text{if } \rho = 1 \end{cases} \quad (6.1)$$

where $\rho = \lambda/\mu$ and N is the queue length. μ is determined based on service time and bandwidth. In this chapter, Transmission Time Interval (TTI) is 1ms and bandwidth is 20MHz in all three bands. Therefore, service rates of all three bands are same ($\mu_a = \mu_b = \mu_c$) and μ can be μ_a , $\mu_a + \mu_b$ or $\mu_a + \mu_b + \mu_c$. N can be 150, 100, or 50 according to number of active bands. Therefore, by using drop rate (D), the throughput (γ) will be

$$\gamma = \lambda(1 - D) \quad (6.2)$$

and throughput ratio is

$$\theta = (1 - D) \quad (6.3)$$

¹The expected throughput ratio term is used for the throughput ratio which is approximately calculated based on the number of active bands.

Therefore, the number of the required bands can be measured according to the predetermined thresholds and estimated throughput ratio as follows:

$$\text{The number of the required bands} = \begin{cases} 1, & \text{if } \theta \leq th_1 \\ 2, & \text{if } th_1 < \theta \leq th_2 \\ 3, & \text{if } th_2 < \theta \end{cases} \quad (6.4)$$

where th_1 and th_2 are the predetermined thresholds, and $th_1 < th_2$.

6.3.3 The Scheduling Procedure

The scheduling algorithm e-ASA in the multi-shared-band system is processed as follows:

- The traffic arrival rate information is determined by the router according to experiences.
- Partially or fully CQI feedback is obtained to measure the quality of band links.
- The number of the required bands is determined based on Equations (6.3) and (6.4). It is very important to note that the active bands are also selected based on the coverage. For example, if two bands are enough to serve all users, and all users in *Band-a* coverage and *Band-b* coverage but not in *Band-c* coverage, then *Band-a* and *Band-b* are activated.
- If the communication between users and the router is ensured by non-active bands (which is recently decided to become non-active), firstly, the communication links between users and the router are recreated by using new active bands. Then, deactivation of the selected bands is processed.
- To increase the efficiency and QoS, packet transferring priority is given to the band, which has the least number of active users (Least Load) and which has the highest CQI. If the band is not enough to data traffic, the other bands can also be used.

6.4 Simulation of the System

A discrete event simulation has been implemented in Matlab by considering the scheduling algorithm and the system architectures which are mentioned in Sections 5.3, 5.4 and 5.5

6.4.1 Assumptions for the System

It is assumed that there are three systems which are the single band, current multi-band and the multi-shared-band. While the single band system has only one band, the multi-band systems have three bands to provide service to users. In addition, the systems can only operate over one channel on each band. Some parameters of the systems are summarized in Table 6.2.

Table 6.2: The parameters of the systems.

Used Bands	2.4GHz, 3.6GHz, 5GHz
Length of Q_s	150 packets
Length of Q_a , Q_b and Q_c	50 packets
Bandwidth	20MHz
Modulations	BPSK, QPSK, 16QAM, and 64QAM
Channel Quality Indicator	3, 5, 7, and 11
Transmission Time Interval	1ms
Threshold for one band	0.8
Threshold for two bands	0.9
α and β	10 and 3, respectively

In the simulation, three bands are used similar to IEEE 802.11ad (WiGig) tri-band. Queue length is kept small to decrease delay time similar to the previous simulations. Due to Wi-Fi, the bandwidth is 20MHz and Transmission Time Interval (TTI) for a packet is 1ms similar to default IEEE 802.11n setting. BPSK, QPSK, 16QAM and 64QAM are the modulations techniques to transfer bits according to CQI. Therefore, to simulate those modulations, four CQI levels are used and each CQI level is modulation changing point. To simulate the energy consumption during the idle and data transfer, α and β are 10

and 3, respectively according to rates of receiving and sending data energy consumption results with equal packet size in [137].

6.4.2 Assumptions for Users

It is assumed that all users can simultaneously connect multiple bands in the systems. However, according to the system architecture, users can transfer data over one or multiple bands. Initially, users are uniformly distributed in the coverage area and they can move around of the routers in specified time interval. Therefore, CQI can change for each user. Each user can only download or upload one type of traffic. Packet arrivals follow Poisson Distribution and packet arrival rates are enlarged when the number of users is increased.

6.4.3 Band Selection and Packet Scheduling

In the simulation, the routers can only operate over one channel from each band. When users arrive, the system assigns a band to each user in the current multi-band system. In the single band system, the band selection is not a problem because there is only one band and all users must use same band. However, the current multi-band system allows each user to transfer data over only one band and the band is automatically assigned to users by selecting the least loaded band with the highest CQI. For example, if there are nine users around a current tri-band Wi-Fi router, three users are allocated to each band if by assuming CQIs of carriers are same. On the other hand, in the multi-shared-band system, the users can transfer data over all active bands.

In addition to the band selection, the packet scheduling is required in the simulation. Each packet is transferred by using the assigned band which minimizes packet delay in the multi-shared-band systems. Packet scheduling is first come first serve in single and current multi-band systems because a user only uses one band to send and receive data. If there are no available bands to serve arrived packets, packets are enqueued to corresponding band queues in single and current multi-band systems. However, packets are enqueued to

corresponding band queues based on minimum delay measurement in the multi-shared-band system. If there are no spaces in queues, arrived packets are dropped.

6.4.4 Observation Methodology

The results in Section 6.5 are average of 100 realizations for different number of users. The effects of light and heavy user loads on the energy consumption and the performance of the single band, current multi-band, multi-shared-band and e-ASA based multi-shared-band systems are investigated.

The performances of the single band and multi-band (current and multi-shared-band) systems are presented by comparing average band utilization (for three bands, after summing utilization of three bands, the sum is divided by three), the throughput ratio and the energy consumption. The throughput ratio is measured by dividing transferred packets to all processed packets as similar to previous chapters. Therefore, while the number of users is increased, the throughput ratio decreases because the bands are shared by multiple users. The energy consumption of each band (E_a, E_b, E_c for *Band-a, Band-b, Band-c*) is obtained by using utilization of the bands (ρ_a, ρ_b, ρ_c for *Band-a, Band-b, Band-c*) and simulation time (T). For instance, the energy consumption of *Band-a* is obtained as follows:

$$E_a = T * (\rho_a * \alpha + (1 - \rho_a) * \beta) \quad (6.5)$$

where α presents the energy consumption of a band during the data transfer and β is the energy consumption of a band during the standby. After finding E_b and E_c similar to E_a , sum of them gives the overall energy consumption of the bands. Therefore, the overall energy consumption of the bands (E) will be

$$E = E_a + E_b + E_c \quad (6.6)$$

As a result of these evaluations, tradeoff between resource usage, energy efficiency and managed QoS are compared for single band, multi-band, multi-share-band and multi-share-band with e-ASA.

6.5 Results

In this section, the overall performances of the single band and multi-band systems are presented and the effects of e-ASA on the multi-shared-band system are shown. In Figures, *Single* represents the single band system, *Current* represents the current multi-band system, *Shared* represents the multi-shared-band system without any restriction (all bands are active any time), and *Shared(e-ASA)* represents e-ASA based multi-shared-band system.

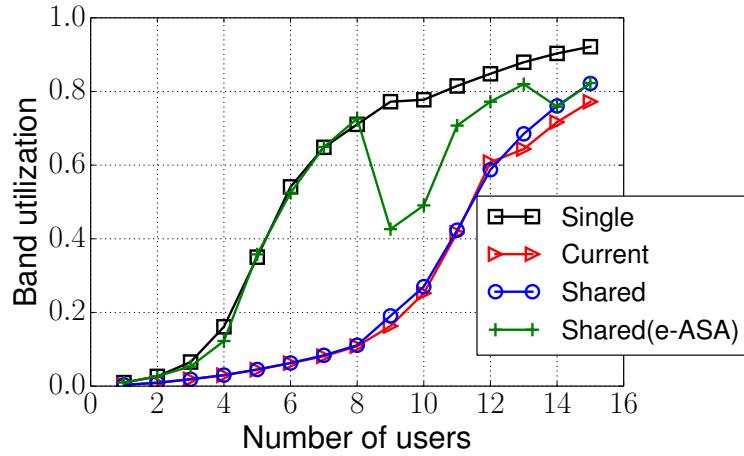


Figure 6.4: Average band utilization.

6.5.1 Band Utilization

Figure 6.4 shows the band utilization of the single band and multi-band systems (current, multi-shared-band and e-ASA based multi-shared-band). The band utilization of all cases is lower than 0.2 when the number of the users in the system is lower than four. However, the band utilization of the current and multi-shared-band systems slowly increases when the number of users is lower than eight. In contrast to multi-band systems, the band utilization of the single band and e-ASA based multi-shared-band is sharply increasing while the number of users is greater than four. This increase is because the single band system has only one band and e-ASA only uses one band. On the other hand, the band utilization

of e-ASA based multi-shared-band system is dropped from around 0.8 to 0.4 when the number of users in the system is nine because e-ASA activates another band. After that, the band utilization of e-ASA based multi-shared-band system gradually increases until the number of users reaches 14. e-ASA activates the last band when the number of users reaches 14. Therefore, the band utilization of e-ASA drops. It is important to note that the single band system has the highest band utilization because of one band.

6.5.2 Energy

Figure 6.5 demonstrates the energy consumptions of single and multi-band systems. While

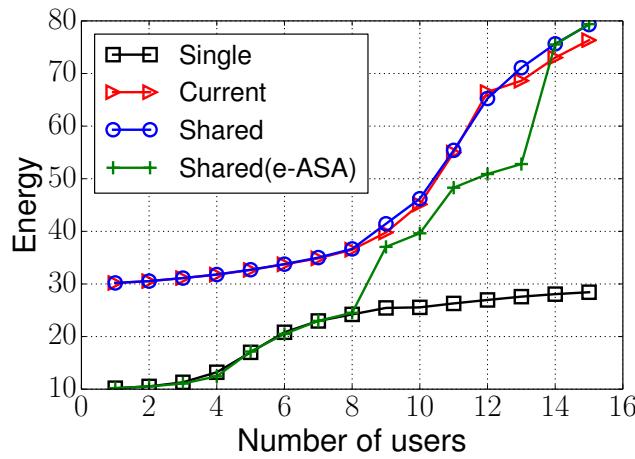


Figure 6.5: Energy consumptions of the systems.

the energy consumptions of the current and multi-shared-band systems are close to each other for the same number of users, the energy consumptions of single and e-ASA based multi-shared-band systems are equal until the number of the user is eight. Moreover, the energy consumptions of the current and multi-shared-band systems are 2 to 3 times higher than the energy consumptions of single and e-ASA based multi-shared-band systems for the same number of users. When the number of users is larger than eight, the energy consumption of e-ASA based multi-shared-band systems is sharply increasing and almost reaches to the energy consumptions of the current and multi-shared-band systems. This sharp increase is because of the dynamic activation of the bands in e-ASA. Furthermore,

the active number of the bands in e-ASA based multi-shared-band system is three when the number of users reaches 14. Therefore, the energy consumption of e-ASA based multi-shared-band is passing the energy consumption of the current multi-band system and reaches the energy consumption of the multi-shared-band system.

6.5.3 Throughput Ratio

Figure 6.6 depicts the throughput ratios of the single and multi-band systems. Until the

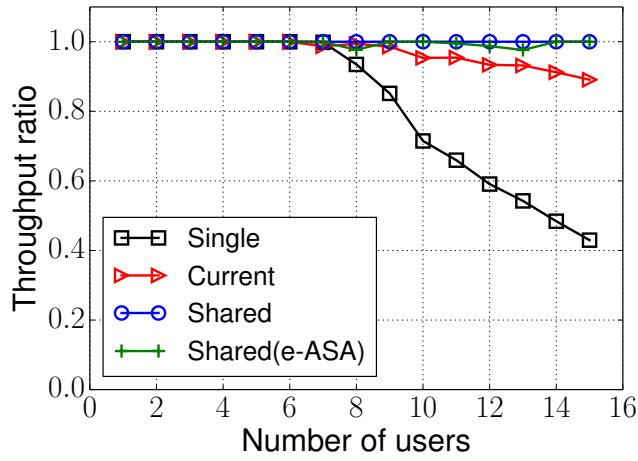


Figure 6.6: Throughput ratios of the systems.

number of users reaches eight, all systems are capable of serving to all traffic of users. However, the throughput ratio of the single band system is regularly decreasing while the number of users is getting higher. On the other hand, the throughput ratio of the current multi-band system slowly decreases after the number of users is ten. In contrast to single and current multi-band systems, the throughput ratios of multi-shared-band and e-ASA based multi-shared-band systems are still optimum while the number of users is 15. However, when the number of users is raising, the throughput ratios of the multi-shared-band and e-ASA based multi-shared-band systems will also decrease.

It is important to note that when the number of the users are 8 and 13, there is insignificant amount of decrease in the throughput ratio of e-ASA based multi-shared-band system because e-ASA allows only one band to be active until the number of the users is

eight and two bands when the number of the users is between 9 and 13. After that, all bands are activated by e-ASA.

6.5.4 Discussion on Results

Based on the results, the following observations are made: (i) The throughput ratios of the multi-band systems are significantly higher than the throughput ratio of the single band system (the same result is obtained in the previous chapter); (ii) The energy consumptions of the current and multi-shared-band systems are close to each other; (iii) e-ASA can decrease the energy consumption of the multi-shared-band system up to 60% without decreasing the throughput ratio.

6.6 Summary

In this chapter, an energy aware scheduling algorithm was proposed for the multi-shared-band system by considering traffic arrivals, channel quality indicator of bands, and service rates. The band utilization, energy consumptions and throughput ratios of the single band, current multi-band and multi-shared-band systems were demonstrated through an extensive simulation. Results show that Energy Aware Scheduling Algorithm uses system resources efficiently and decreases the energy consumption of the multi-band system up to 60%. The proposed energy aware scheduling algorithm and related analysis should help network engineers build next generation wireless systems by considering the energy usage of the systems.

The next chapter includes the conclusions and future directions of this dissertation.

Chapter 7

Conclusions and Future Works

In this research, carrier assignment for LTE and LTE-A, and packet scheduling for multi-band Wi-Fi routers have been studied and their performances are improved by filling gaps of the limitations of the previous works. The key contributions of this dissertation can be summarized as follows:

- **Selective Technique:** The selective periodic component carrier assignment technique (Chapter 3) has been proposed by considering the behavior of the system during the carrier assignment operations. The performances of the current joint and the proposed selective periodic component carrier assignment techniques have been compared by using analytical analysis based on queuing theory and an extensive simulation. Results show that the proposed selective technique efficiently uses system resources and improves the overall throughput ratio up to 18% and average delay up to 50% in LTE and LTE-A systems.
- **User Profile Strategy:** A novel carrier component assignment method for LTE and LTE-A based on user profiles has been developed (Chapter 4). Utilization, delay and throughput ratio have been presented for the proposed method according to overall and equipment types through an extensive simulation. Results show that the proposed method uses system resources efficiently and improve overall throughput ratio up to 15% and overall delay up to 20% comparing the other methods in LTE and LTE-A systems.

- **Architecture and Scheduler for Multi-Band Wi-Fi Routers:** Similar to Carrier Aggregation in LTE-A, a novel multi-shared-band architecture with a packet scheduling algorithm has been proposed for multi-band Wi-Fi routers that exploits band sharing (Chapter 5). The performance of the multi-shared-band system has been presented by using priority based queuing system and the analytical expressions have been validated by an extensive simulation. Then, the performances of the single and multi-band systems under different scenarios have been compared according to class-wise, queue-wise and system-wise cases and according to the results, the recommendations for end-users have been listed to identify suitable routers to fulfill their needs.
- **Energy Aware Scheduler for Multi-Band Wi-Fi Routers:** The proposed multi-shared-band scheduling algorithm has been improved by adding of energy awareness (Chapter 6). Band utilization, energy consumptions and throughput ratios of single band, current multi-band and multi-shared-band Wi-Fi routers have been demonstrated through an extensive simulation. Results show that Energy Aware Scheduling Algorithm uses system resources efficiently and decrease energy consumption of the multi-band routers up to 60% without decreasing the system throughput ratio.

The proposed resource management models and related analysis should help service providers and network engineers build efficient carrier assignment and packet scheduling methods to improve quality services such as throughput, delay and energy efficiency. However, the proposed methods could be further improved in a number of ways. Some of the future works are listed below:

- **Time-Based Analysis for Periodic Carrier Assignment Methods in LTE-A:** In Chapter 3, selective technique has been developed. In our research, the carrier assignment occurs periodically with in a given time (e.g., sixteen times periodic update in between each mandatory carrier assignment) and it has been evaluated

according to the given time. In [22], there is no information related to the frequency of periodic carrier assignment. However, the frequencies of periodic carrier assignments can vary. Therefore, the performance of selective and joint technique can also be evaluated according to distinct carrier reassignment frequencies to investigate performance of periodic carrier assignment.

- **User Profile for LTE-U:** LTE-Unlicensed is one of the crucial development which allows LTE equipment to receive services over unlicensed bands. However, unlicensed bands such as 2.4 GHz are used by users of Wi-Fi. Therefore, it is challenging to manage carrier assignment by considering the co-existence of users of LTE-A and Wi-Fi. In order to overcome the challenge, Listen-Before-Talk has been integrated into LTE-A. On the other hand, in our proposed user profile strategy (Chapter 4), the benefits of user profile for LTE and LTE-A have been presented by ignoring Wi-Fi users. However, user profile strategy can be applied to such environment by not only considering user profile but also map profile of Wi-Fi network activities over time by eNBs to increase service efficiency because knowing possible activities in Wi-Fi will help eNBs assign suitable and non-interference carriers.
- **Mismatch of Data and Signaling in Multi-Shared-Band System:** In Chapters 5, a novel multi-shared-band architecture with a packet scheduling algorithm has been proposed for multi-band Wi-Fi routers. However, using multiple bands increases the complexity of the system especially in terms of signal and data messages because such shared algorithm changes bands frequently to send and receive data and may cause the sequence mismatch between signal message and data message. To overcome such problems, non-preemptive priority for different data traffic classes is used. However, signaling and data messages should be investigated more in multi-band Wi-Fi routers to overcome such mismatch by reserving specific channels for signaling messages as done in LTE-A.

There are clearly many works to be done in the area of carrier assignment and packet scheduling. However, the most direct extension of this work is to extend the proposed methods according to coexistence of different heterogeneous networks and analyze the effects of not only coexistence of the networks but also their heterogeneity levels on the performance.

All in all, this research guides researchers along illuminating journey through current developments on the resource management in terms of the carrier assignment and packet scheduling methods in LTE-A and Wi-Fi, and shedding some lights on the some of the challenges to satisfy end users and increase Quality of Service.

Bibliography

- [1] M. L. Smith, R. Spence, and A. T. Rashid, “Mobile phones and expanding human capabilities,” *Information Technologies & International Development*, vol. 7, no. 3, pp. 77–88, 2011.
- [2] D. Pareit, B. Lannoo, I. Moerman, and P. Demeester, “The history of WiMAX: A complete survey of the evolution in certification and standardization for IEEE 802.16 and WiMAX,” *IEEE Communications Surveys Tutorials*, vol. 14, no. 4, pp. 1183–1211, Fourth quarter 2012.
- [3] R. Baldemair, E. Dahlman, G. Fodor, G. Mildh, S. Parkvall, Y. Selen, H. Tullberg, and K. Balachandran, “Evolving wireless communications: Addressing the challenges and expectations of the future,” *IEEE Vehicular Technology Magazine*, vol. 8, no. 1, pp. 24–30, Mar 2013.
- [4] Gsma-Intelligence. Definitive data and analysis for the mobile industry. Accessed: Feb 10, 2016. [Online]. Available: <https://gsmaintelligence.com/>
- [5] F. Richter. (2013, Sep) Smartphone sales break the billion barrier. Accessed: June 12, 2014. [Online]. Available: <http://www.statista.com/chart/777/global-connected-device-shipments/>
- [6] Y. Xu and S. Mao, “A survey of mobile cloud computing for rich media applications,” *IEEE Wireless Communications*, vol. 20, no. 3, pp. 46–53, June 2013.
- [7] D. Kovachev, Y. Cao, and R. Klamma, “Mobile cloud computing: A comparison of application models,” *CoRR*, vol. abs/1107.4940, 2011.
- [8] A. Khan, M. Othman, S. Madani, and S. Khan, “A survey of mobile cloud computing application models,” *IEEE Communications Surveys Tutorials*, vol. PP, no. 99, July 2013.
- [9] J. Pereira, P. Silva, P. Lima, and A. Martinoli, “An experimental study in wireless connectivity maintenance using up to 40 robots coordinated by an institutional robotics approach,” in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Tokyo, Nov 3-8, 2013, pp. 5073–5079.

- [10] O. Karimi, J. Liu, and C. Wang, “Seamless wireless connectivity for multimedia services in high speed trains,” *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 4, pp. 729–739, May 2012.
- [11] S. Dasgupta and G. Mao, “On the quality of wireless network connectivity,” in *IEEE Global Communications Conference (GLOBECOM)*, Anaheim, CA, Dec 3-7, 2012, pp. 500–505.
- [12] W. Ren, Q. Zhao, and A. Swami, “Connectivity of heterogeneous wireless networks,” *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4315–4332, 2011.
- [13] J. Reich, V. Misra, D. Rubenstein, and G. Zussman, “Connectivity maintenance in mobile wireless networks via constrained mobility,” *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 5, pp. 935–950, 2012.
- [14] X. Wang, X. Lin, Q. Wang, and W. Luan, “Mobility increases the connectivity of wireless networks,” *IEEE/ACM Transactions on Networking*, vol. 21, no. 2, pp. 440–454, Apr 2013.
- [15] D. Shila and Y. Cheng, “Ad hoc wireless networks meet the infrastructure: Mobility, capacity and delay,” in *IEEE INFOCOM*, Orlando, FL, Mar 25-30, 2012, pp. 3031–3035.
- [16] J. Xie and X. Wang, “A survey of mobility management in hybrid wireless mesh networks,” *IEEE Network*, vol. 22, no. 6, pp. 34–40, Nov 2008.
- [17] J. Wannstrom. (2013, June) LTE-Advanced. Accessed: Mar 18, 2015. [Online]. Available: <http://www.3gpp.org/technologies/keywords-acronyms/97-lte-advanced>
- [18] I. F. Akyildiz, D. M. Gutierrez-Estevez, and E. C. Reyes, “The evolution to 4G cellular systems: LTE-Advanced,” *Physical Communication*, vol. 3, pp. 217–244, Mar 2010.
- [19] H. Lee, S. Vahid, and K. Moessner, “A survey of radio resource management for spectrum aggregation in LTE-Advanced,” *IEEE Communications Surveys Tutorials*, vol. 16, no. 2, pp. 745–760, Second quarter 2014.
- [20] Y. L. Lee, T. C. Chuah, J. Loo, and A. Vinel, “Recent advances in radio resource management for heterogeneous LTE/LTE-A networks,” *IEEE Communications Surveys & Tutorials*, vol. 16, no. 4, pp. 2142–2180, June 2014.
- [21] E. Ahmed, A. Gani, S. Abolfazli, L. Yao, and S. Khan, “Channel assignment algorithms in cognitive radio networks: Taxonomy, open issues, and challenges,” *IEEE Communications Surveys Tutorials*, vol. 18, no. 1, pp. 795–823, First quarter 2016.

- [22] X. Cheng, G. Gupta, and P. Mohapatra, “Joint carrier aggregation and packet scheduling in LTE-Advanced networks,” in *Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*, New Orleans, LA, June 24-27, 2013, pp. 469–477.
- [23] E. Perahia, C. Cordeiro, M. Park, and L. L. Yang, “IEEE 802.11ad: Defining the next generation multi-Gbps Wi-Fi,” in *7th IEEE Consumer Communications and Networking Conference (CCNC)*, Las Vegas, NV, Jan 9-12, 2010.
- [24] L. Verma and S. S. Lee, “Multi-band Wi-Fi systems: A new direction in personal and community connectivity,” in *IEEE International Conference on Consumer Electronics (ICCE)*, Las Vegas, NV, Jan 9-12, 2011, pp. 665–666.
- [25] M. N. Alqahtani, W. N. L. Mahadi, and T. F. Tengkumohmed, “Design of multi band antenna for wireless communication,” in *IEEE International Conference on Semiconductor Electronics*, Melaka, June 28-30, 2010, pp. 338–343.
- [26] N. Blackmore. (2014, July) Wireless Internet router is expensive on standby but is it wise to turn off at night? Accessed: Dec 4, 2014. [Online]. Available: <http://www.telegraph.co.uk/finance/personalfinance/money-saving-tips/10953517/Wireless-internet-router-is-expensive-on-standby-but-is-it-wise-to-turn-off-at-night.html>
- [27] 3GPP. (2015, Oct) LTE; evolved universal terrestrial radio access (E-UTRA) and evolved universal terrestrial radio access network (E-UTRAN); overall description; stage 2 (3GPP TS 36.300 version 12.7.0 Release 12) . Accessed: Nov 18, 2015. [Online]. Available: <http://www.3gpp.org/dynareport/36300.htm>
- [28] 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, Aug 2010.
- [29] 4G Americas. (2014, Oct) LTE carrier aggregation technology development and deployment worldwide. Accessed: Feb 13, 2016. [Online]. Available: http://www.4gamericas.org/files/8414/1471/2230/4G_Americas_Carrier_Aggregation_FINALv1_0_3.pdf
- [30] G. Ku and J. Walsh, “Resource allocation and link adaptation in LTE and LTE Advanced: A tutorial,” *IEEE Communications Surveys Tutorials*, vol. 17, no. 3, pp. 1605–1633, Dec 2014.
- [31] H. Sahlin, “Channel prediction for link adaptation in LTE uplink,” in *IEEE Vehicular Technology Conference (VTC Fall)*, Quebec City, QC, Sep 3-6, 2012.
- [32] G. Caire, R. Muller, and R. Knopp, “Hard fairness versus proportional fairness in wireless communications: The single-cell case,” *IEEE Transactions on Information Theory*, vol. 53, no. 4, pp. 1366–1385, Apr 2007.

- [33] H. A. Ngo and L. Hanzo, “Hybrid automatic-repeat-request systems for cooperative wireless communications,” *IEEE Communications Surveys Tutorials*, vol. 16, no. 1, pp. 25–45, Aug 2014.
- [34] Y. Lifeng and L. Zhifang, “A hybrid automatic repeat request (HARQ) with turbo codes in OFDM system,” in *International Conference on Computational Intelligence and Software Engineering (CiSE)*, Wuhan, Dec 10-12 2010.
- [35] N. Prasad and X. Wang, “Efficient combining techniques for multi-input multi-output multi-user systems employing hybrid automatic repeat request,” *IET Communications*, vol. 5, no. 13, pp. 1785–1796, Sep 2011.
- [36] R. Schwarz. (2012, Nov) Carrier aggregation ? (one) key enabler for LTE-Advanced. Accessed: Feb 10, 2016. [Online]. Available: https://cdn.rohde-schwarz.com/pws/dl_downloads/dl_common_library/dl_brochures_and_datasheets/pdf_1/Article_Carrier-aggregation.pdf
- [37] Y. Li, Q. Mu, L. Liu, L. Chen, M. Peng, and W. Wang, “Control channel design for carrier aggregation between LTE FDD and LTE TDD systems,” in *IEEE 75th Vehicular Technology Conference (VTC Spring)*, Yokohama, May 6-9, 2012.
- [38] R. Shrivastava, S. Costanzo, K. Samdanis, D. Xenakis, D. Grace, and L. Merakos, “An SDN-based framework for elastic resource sharing in integrated FDD/TDD LTE-A HetNets,” in *IEEE 3rd International Conference on Cloud Networking (CloudNet)*, Luxembourg, Oct 8 -10, 2014, pp. 126–131.
- [39] A. Niknejad, S. Thyagarajan, E. Alon, Y. Wang, and C. Hull, “A circuit designer’s guide to 5G mm-wave,” in *IEEE Custom Integrated Circuits Conference (CICC)*, San Jose, CA, Sep 28-30, 2015.
- [40] F. Boccardi, J. Andrews, H. Elshaer, M. Dohler, S. Parkvall, P. Popovski, and S. Singh, “Why to decouple the uplink and downlink in cellular networks and how to do it,” *arXiv preprint arXiv:1503.06746*, 2015.
- [41] H. Xie, F. Gao, S. Zhang, and S. Jin, “A simple DFT-aided spatial basis expansion model and channel estimation strategy for massive MIMO systems,” *CoRR*, vol. abs/1511.04841, 2015.
- [42] 3GPP. (2015, Feb) LTE; evolved universal terrestrial radio access (E-UTRA) and evolved universal terrestrial radio access network (E-UTRAN); overall description; stage 2 (3GPP TS 36.300 version 12.4.0 Release 12) . Accessed: Mar 18, 2015. [Online]. Available: http://www.etsi.org/deliver/etsi_ts/136300_136399/136300/12.04.00_60/ts_136300v120400p.pdf
- [43] M. Pan, T. Lin, C. Chiu, and C. Wang, “Downlink traffic scheduling for LTE-A small cell networks with dual connectivity enhancement,” *IEEE Communications Letters*, no. 99, Jan 2016.

- [44] M.-S. Woo, S.-M. Kim, S.-G. Min, and S.-E. Hong, “Micro mobility management for dual connectivity in LTE HetNets,” in *IEEE International Conference on Communication Software and Networks (ICCSN)*, Chengdu, June 6-7, 2015, pp. 395–398.
- [45] S. Ahmad and L. DaSilva, “Power control and soft topology adaptations in multi-hop cellular networks with multi-point connectivity,” *IEEE Transactions on Communications*, vol. 63, no. 3, pp. 683–694, Mar 2015.
- [46] F. Liu, E. Bala, E. Erkip, M. Beluri, and R. Yang, “Small-cell traffic balancing over licensed and unlicensed bands,” *IEEE Transactions on Vehicular Technology*, vol. 64, no. 12, pp. 5850–5865, Dec 2015.
- [47] A. Al-Dulaimi, S. Al-Rubaye, Q. Ni, and E. Sousa, “5G communications race: Pursuit of more capacity triggers LTE in unlicensed band,” *IEEE Vehicular Technology Magazine*, vol. 10, no. 1, pp. 43–51, Mar 2015.
- [48] R. Zhang, M. Wang, L. Cai, Z. Zheng, and X. Shen, “LTE-unlicensed: the future of spectrum aggregation for cellular networks,” *IEEE Wireless Communications*, vol. 22, no. 3, pp. 150–159, June 2015.
- [49] M. Khawer, J. Tang, and F. Han, “usicic - a proactive small cell interference mitigation strategy for improving spectral efficiency of LTE networks in the unlicensed spectrum,” *IEEE Transactions on Wireless Communications*, no. 99, Nov 2015.
- [50] F. Abinader, E. Almeida, F. Chaves, A. Cavalcante, R. Vieira, R. Paiva, A. So-brinho, S. Choudhury, E. Tuomaala, K. Doppler, and V. Sousa, “Enabling the coexistence of LTE and Wi-Fi in unlicensed bands,” *IEEE Communications Magazine*, vol. 52, no. 11, pp. 54–61, Nov 2014.
- [51] Y. Wang, K. Pedersen, T. Sorensen, and P. Mogensen, “Carrier load balancing and packet scheduling for multi-carrier systems,” *IEEE Transactions on Wireless Communications*, vol. 9, no. 5, pp. 1780–1789, May 2010.
- [52] G. Berardinelli, T. Sorensen, P. Mogensen, and K. Pajukoski, “Transmission over multiple component carriers in LTE-A uplink,” *IEEE Wireless Communications*, vol. 18, no. 4, pp. 67–73, Aug 2011.
- [53] Y. Yu and D. Gu, “Enhanced MU-MIMO downlink transmission in the FDD-based distributed antennas system,” *IEEE Communications Letters*, vol. 16, no. 1, pp. 37–39, Jan 2012.
- [54] M. Lichtman, T. Czuski, S. Ha, P. David, and J. Reed, “Detection and mitigation of uplink control channel jamming in LTE,” in *IEEE Military Communications Conference (MILCOM)*, Baltimore, MD, Oct 6-8, 2014, pp. 1187–1194.

- [55] T. Chaitanya and E. Larsson, “Improving 3GPP-LTE uplink control signaling performance using complex-field coding,” *IEEE Transactions on Vehicular Technology*, vol. 62, no. 1, pp. 161–171, Jan 2013.
- [56] D. Wang, S. Yang, Y. Liao, and Y. Liu, “Efficient receiver scheme for LTE PUCCH,” *IEEE Communications Letters*, vol. 16, no. 3, pp. 352–355, Mar 2012.
- [57] N. Hou, K. Niu, Z. He, and S. Sun, “Test and performance analysis of PUSCH channel of LTE system,” in *IEEE International Symposium on Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications (MAPE)*, Chengdu, Oct 29-31, 2013, pp. 110–114.
- [58] P. Osti, P. Lassila, S. Aalto, A. Larmo, and T. Tirronen, “Analysis of PDCCH performance for M2M traffic in LTE,” *IEEE Transactions on Vehicular Technology*, vol. 63, no. 9, pp. 4357–4371, Nov 2014.
- [59] K. Digish and R. Thilagavathy, “ASIC implementation of physical downlink shared channel for LTE,” in *International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*, Kanyakumari, July 10-11, 2014, pp. 370–376.
- [60] J. Zhu and H. Li, “On the performance of LTE physical downlink shared channel,” in *International Conference on Computer Science and Network Technology (ICCSNT)*, vol. 2, Harbin, Dec 24-26, 2011, pp. 983–986.
- [61] T. Dean and P. Fleming, “Trunking efficiency in multi-carrier CDMA systems,” in *56th Vehicular Technology Conference*, Vancouver, Canada, Sep 24-28, 2002, pp. 156–160.
- [62] N. Zia and A. Mitschele-Thiel, “Self-organized neighborhood mobility load balancing for LTE networks,” in *IFIP Wireless Days (WD)*, Valencia, Nov 13-15, 2013.
- [63] H. Rifai, S. Mohammed, and A. Mellouk, “A brief synthesis of QoS-QoE methodologies,” in *International Symposium on Programming and Systems (ISPS)*, Algiers, Apr 2011, pp. 32–38.
- [64] R. Nordin and M. Ismail, “Partial feedback scheme with an interference-aware sub-carrier allocation scheme in a correlated LTE downlink,” in *19th Asia-Pacific Conference on Communications*, Denpasar, Aug 29-31, 2013, pp. 643–648.
- [65] P. Ferguson and G. Huston, *Quality of service: delivering QoS on the Internet and in corporate networks.* Wiley, Jan 1998.
- [66] S. O. Yang, S. Kim, and C.-S. Hwang, “Low-cost binding update schemes based on mobility type in HMIPv6,” *Canadian Journal of Electrical and Computer Engineering*, vol. 31, no. 4, pp. 195–202, Dec 2006.

- [67] R.-C. Wang and R. Chang, “Cross-layer binding update for TCP performance enhancement over mobile IPv6 networks,” *IET Communications*, vol. 1, no. 5, pp. 924–932, Oct 2007.
- [68] B. Blaszczyszyn, M. Jovanovic, and M. Karray, “Quality of real-time streaming in wireless cellular networks - stochastic modeling and analysis,” *IEEE Transactions on Wireless Communications*, vol. 13, no. 6, pp. 3124–3136, June 2014.
- [69] H. Thomsen, N. Pratas, C. Stefanovic, and P. Popovski, “Analysis of the LTE access reservation protocol for real-time traffic,” *IEEE Communications Letters*, vol. 17, no. 8, pp. 1616–1619, Aug 2013.
- [70] D. Xenakis, N. Passas, L. Merakos, and C. Verikoukis, “Mobility management for femtocells in LTE-Advanced: Key aspects and survey of handover decision algorithms,” *IEEE Communications Surveys Tutorials*, vol. 16, no. 1, pp. 64–91, July 2014.
- [71] R.-H. Liou, Y.-B. Lin, and S.-C. Tsai, “An investigation on LTE mobility management,” *IEEE Transactions on Mobile Computing*, vol. 12, no. 1, pp. 166–176, Jan 2013.
- [72] F. Liu, W. Xiang, Y. Zhang, K. Zheng, and H. Zhao, “A novel QoE-based carrier scheduling scheme in LTE-Advanced networks with multi-service,” in *Vehicular Technology Conference*, Quebec City, Canada, Sep 3-6, 2012.
- [73] H. Wang, C. Rosa, and K. Pedersen, “Uplink component carrier selection for LTE-Advanced systems with carrier aggregation,” in *IEEE International Conference on Communications*, Kyoto, June 5-9, 2011.
- [74] G. Yu, Q. Chen, R. Yin, H. Zhang, and G. Ye Li, “Joint downlink and uplink resource allocation for energy-efficient carrier aggregation,” *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3207–3218, June 2015.
- [75] Z. Chen, G. Cui, C. Zhai, W. Wang, Y. Zhang, and X. Li, “Component carrier selection based on user mobility for LTE-Advanced systems,” in *IEEE 78th Vehicular Technology Conference (VTC Fall)*, Las Vegas, NV, Sep 2-5, 2013.
- [76] Y. Wang, K. Pedersen, P. Mogensen, and T. Sorensen, “Resource allocation considerations for multi-carrier LTE-Advanced systems operating in backward compatible mode,” in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, Tokyo, Sep 13-16, 2009, pp. 370–374.
- [77] H. Tian, S. Gao, J. Zhu, and L. Chen, “Improved component carrier selection method for non-continuous carrier aggregation in LTE-Advanced systems,” in *IEEE Vehicular Technology Conference (VTC Fall)*, San Francisco, CA, Sep 5-8, 2011.

- [78] L. Liu, M. Li, J. Zhou, X. She, L. Chen, Y. Sagae, and M. Iwamura, “Component carrier management for carrier aggregation in LTE-Advanced system,” in *IEEE Vehicular Technology Conference*, Budapest, May 15-18, 2011.
- [79] H. Wang, C. Rosa, and K. Pedersen, “Performance analysis of downlink inter-band carrier aggregation in LTE-Advanced,” in *IEEE Vehicular Technology Conference*, San Francisco, CA, Sep 5-8, 2011.
- [80] C. Sun, H. Qing, S. Wang, and G. Lu, “Component carrier selection and beamforming on carrier aggregated channels in heterogeneous networks,” *Communications and Network*, vol. 5, no. 3B, pp. 211–216, Sep 2013.
- [81] A. Shahid, S. Aslam, S. Sohaib, H. S. Kim, , and K.-G. Lee, “A self-organized metaheuristic approach towards inter-cell interference management for LTE-Advanced,” *EURASIP Journal on Wireless Communications and Networking*, Oct 2014.
- [82] H. Tang, Y. Tian, H. Wang, and R. Huang, “A component carrier selection algorithm based on channel quality for LTE-Advanced system with carrier aggregation,” *Journal of Computational Information Systems*, pp. 8953–8962, Oct 2014.
- [83] R. Sivaraj, A. Pande, K. Zeng, K. Govindan, and P. Mohapatra, “Edge-prioritized channel- and traffic-aware uplink carrier aggregation in LTE-Advanced systems,” in *International Symposium on a World of Wireless, Mobile and Multimedia Networks*, San Francisco, CA, June 25-28, 2012.
- [84] S. N. K. Marwat, Y. Dong, X. Li, Y. Zaki, and C. Goerg, “Novel schemes for component carrier selection and radio resource allocation in lte-advanced uplink,” in *Mobile Networks and Management*, ser. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, R. Agero, T. Zinner, R. Goleva, A. Timm-Giel, and P. Tran-Gia, Eds. Springer International Publishing, 2015, vol. 141, pp. 32–46.
- [85] X. Chen, H. Yi, H. Luo, H. Yu, and H. Wang, “A novel CQI calculation scheme in LTE\LTE-A systems,” in *International Conference on Wireless Communications and Signal Processing*, Nanjing, Nov 9-11, 2011.
- [86] S. Donthi and N. Mehta, “Performance analysis of subband-level channel quality indicator feedback scheme of LTE,” in *National Conference on Communications*, Chennai, Jan 29-31, 2010.
- [87] N. Kolehmainen, J. Puttonen, P. Kela, T. Ristaniemi, T. Henttonen, and M. Moisio, “Channel quality indication reporting schemes for UTRAN long term evolution downlink,” in *IEEE Vehicular Technology Conference*, Singapore, May 11-14, 2008, pp. 2522–2526.

- [88] L. xiang Lin, Y. an Liu, F. Liu, G. Xie, K. ming Liu, and X. yang Ge, “Resource scheduling in downlink LTE-Advanced system with carrier aggregation,” *The Journal of China Universities of Posts and Telecommunications*, vol. 19, no. 1, pp. 44 – 49, Feb 2012.
- [89] S.-B. Lee, S. Choudhury, A. Khoshnevis, S. Xu, and S. Lu, “Downlink MIMO with frequency-domain packet scheduling for 3GPP LTE,” in *IEEE INFOCOM*, Rio de Janeiro, Apr 19-25, 2009, pp. 1269–1277.
- [90] W. Fu, Q. Kong, W. Tian, C. Wang, and L. Ma, “A QoS-aware scheduling algorithm based on service type for LTE downlink,” in *International Conference on Computer Science and Electronics Engineering*, Hangzhou, China, Mar 22-23, 2013, pp. 2468–2474.
- [91] S. Bodas, S. Shakkottai, L. Ying, and R. Srikant, “Scheduling for small delay in multi-rate multi-channel wireless networks,” in *IEEE INFOCOM*, Shanghai, China, Apr 10-15, 2011.
- [92] B. Sadiq, S. J. Baek, and G. de Veciana, “Delay-optimal opportunistic scheduling and approximations: The log rule,” *IEEE/ACM Transactions on Networking*, vol. 19, no. 2, pp. 405–418, Apr 2011.
- [93] H. K. Rath, M. Sengupta, and A. Simha, “Novel transport layer aware uplink scheduling scheme for LTE-based networks,” in *National Conference on Communications*, New Delhi, India, Feb 15-17, 2013.
- [94] H. Yang, F. Ren, C. Lin, and J. Zhang, “Frequency-domain packet scheduling for 3GPP LTE uplink,” in *IEEE INFOCOM*, San Diego, CA, Mar 14-19, 2010.
- [95] H. S. Narman and M. Atiquzzaman, “Analysis of joint and partial component carrier assignment techniques in LTE and LTE-A,” in *IEEE Global Communications Conference (GLOBECOM)*, San Diego, CA, Dec 6-10, 2015.
- [96] ———, “Selective periodic component carrier assignment technique in LTE and LTE-A systems,” in *IEEE Global Communications Conference (GLOBECOM)*, San Diego, CA, Dec 6-10, 2015.
- [97] L. Chen, W. Chen, X. Zhang, and D. Yang, “Analysis and simulation for spectrum aggregation in LTE-Advanced system,” in *70th Vehicular Technology Conference*, Anchorage, AK, Sep 20-23, 2009.
- [98] D. Gross and C. M. Harris, *Fundamentals of Queueing Theory (Wiley Series in Probability and Statistics)*. Wiley-Interscience, Feb 1998.
- [99] J. D. Little and S. C. Graves, “Little’s law,” in *Building intuition*. Springer, 2008, pp. 81–100.

- [100] H. S. Narman, M. S. Hossain, and M. Atiquzzaman, “Multi class traffic analysis of single and multi-band queuing system,” in *IEEE Global Communications Conference (GLOBECOM)*, Atlanta, GA, Dec 9-13, 2013, pp. 1422 – 1427.
- [101] F. S. Q. Alves, H. C. Yehia, L. A. C. Pedrosa, F. R. B. Cruz, and L. Kerbache, “Upper bounds on performance measures of heterogeneous M/M/c queues,” *Mathematical Problems in Engineering*, vol. 2011, May 2011.
- [102] G. Appenzeller, I. Keslassy, and N. McKeown, “Sizing router buffers,” *Computer Communication Review*, vol. 34, pp. 281–292, Oct 2004.
- [103] P. Ameigeiras, Y. Wang, J. Navarro-Ortiz, P. Mogensen, and J. Lopez-Soler, “Traffic models impact on OFDMA scheduling design,” *EURASIP Journal on Wireless Communications and Networking*, vol. 2012, no. 1, pp. 1–13, Feb 2012.
- [104] A. Adas, “Traffic models in broadband networks,” *IEEE Communications Magazine*, vol. 35, no. 7, pp. 82–89, July 1997.
- [105] H. S. Narman and M. Atiquzzaman, “Carrier components assignment method for LTE and LTE-A systems based on user profile and application,” in *IEEE GLOBECOM Workshop on Broadband Wireless Access*, Austin, TX, Dec 12, 2014.
- [106] H. Narman and M. Atiquzzaman, “Joint and partial carrier components assignment techniques based on user profile in LTE systems,” in *IEEE Wireless Communications and Networking Conference (WCNC)*, New Orleans, LA, Mar 9-12, 2015, pp. 983–988.
- [107] Ixia. (2013, Dec) Quality of service (QoS) and policy management in mobile data networks. White Paper, Accessed: July 10, 2014. [Online]. Available: https://www.ixiacom.com/sites/default/files/resources/whitepaper/policy_management.pdf
- [108] M. S. Hossain, H. S. Narman, and M. Atiquzzaman, “A novel scheduling and queue management scheme for multi-band mobile routers,” in *IEEE International Conference on Communications (ICC)*, Budapest, Hungary, June 9-13, 2013, pp. 3787 – 3791.
- [109] H. S. Narman, M. Hossain, and M. Atiquzzaman, “Management and analysis of multi class traffic in single and multi-band systems,” *Wireless Personal Communications*, Feb 2015.
- [110] R. Daniels and R. Heath, “60 ghz wireless communications: emerging requirements and design recommendations,” *IEEE Vehicular Technology Magazine*, vol. 2, no. 3, pp. 41–50, Sep 2007.
- [111] T. Heath, B. Diniz, E. V. Carrera, W. M. Jr., and R. Bianchini, “Energy conservation in heterogeneous server clusters,” in *Principles and Practice of Parallel Programming*, Chicago, IL, June 15-17, 2005, pp. 186–195.

- [112] N. Kumar, N. Chilamkurti, J. Park, and D. Park, “Load balancing with fair scheduling for multiclass priority traffic in wireless mesh networks,” in *Future Information Technology*, ser. Communications in Computer and Information Science, J. Park, L. Yang, and C. Lee, Eds. Springer Berlin Heidelberg, 2011, vol. 184, pp. 101–109.
- [113] N. Lu and J. Bigham, “Utility-maximization bandwidth adaptation for multi-class traffic QoS provisioning in wireless networks,” in *1st ACM International Workshop on Quality of Service & Security in Wireless and Mobile Networks*, Montreal, Canada, Oct 10 - 13, 2005, pp. 136–143.
- [114] A. Ghosh, S. Ha, E. Crabbe, and J. Rexford, “Scalable multi-class traffic management in data center backbone networks,” *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 12, Dec 2013.
- [115] H. Singh, J. Hsu, L. Verma, S. S. Lee, and C. Ngo, “Green operation of multi-band wireless LAN in 60 GHz and 2.4/5 GHz,” in *Consumer Communications and Networking Conference (CCNC)*, Las Vegas, NV, Jan 9-12, 2011, pp. 787–792.
- [116] S. Singh, R. Mudumbai, and U. Madhow, “Distributed coordination with deaf neighbors: Efficient medium access for 60 GHz mesh networks,” in *IEEE INFOCOM*, San Diego, CA, Mar 14-19, 2010.
- [117] Y. Lin, W. Lai, and R. Chen, “Performance analysis for dual band PCS networks,” *IEEE Transactions on Computers*, vol. 49, pp. 148–159, Feb 2000.
- [118] K. Doppler, C. Wijting, T. Henttonen, and K. Valkealahti, “Multiband scheduler for future communication systems,” *I. J. Communications, Network and System Sciences*, vol. 1, no. 1, pp. 6744 – 6748, Feb 2008.
- [119] R. Samanta, G. Sanyal, and P. Bhattacharjee, “Study and analysis of cellular wireless networks with multiclass traffic,” in *IEEE International Advance Computing Conference*, Patiala, Mar 6-7, 2009, pp. 1081–1086.
- [120] Cisco. Congestion management overview. Accessed: May 10, 2015. [Online]. Available: http://www.cisco.com/c/en/us/td/docs/ios/12_2/qos/configuration/guide/fqos_c/qcfconmg.html
- [121] D. Medhi, *Network Routing Algorithms, Protocols, and Architectures*. Morgan Kaufmann, Mar 2007.
- [122] D. Gong, M. Zhao, and Y. Yang, “Distributed channel assignment algorithms for 802.11n WLANs with heterogeneous clients,” *Journal of Parallel and Distributed Computing*, vol. 74, no. 5, pp. 2365 – 2379, May 2014.

- [123] M. S. Hossain, M. Atiquzzaman, and W. Ivancic, “Scheduling and queue management for multi-class traffic in access router of mobility protocol,” in *IEEE HPCC*, Melbourne, Australia, Sep 1-3, 2010, pp. 653–658.
- [124] N. Nasser, L. Karim, and T. Taleb, “Dynamic multilevel priority packet scheduling scheme for wireless sensor network,” *IEEE Transactions on Wireless Communications*, vol. 12, no. 4, pp. 1536–1276, Feb 2013.
- [125] M. Al-Sanabani, S. Shamala, M. Othman, and Z. Zukarnain, “Multi-class bandwidth reservation scheme based on mobility prediction for handoff in multimedia wireless/mobile cellular networks,” *Wireless Personal Communications*, vol. 46, no. 2, pp. 143–163, 2008.
- [126] D. Gross, J. Shortle, J. Thompson, and C. M. Harris, *Fundamentals of Queueing Theory*. Wiley-Interscience, Aug 2008.
- [127] K. E. Avrachenkov, N. O. Vilchevsky, and G. L. Shevlyakov, “Priority queueing with finite buffer size and randomized push-out mechanism,” *Performance Evaluation*, vol. 61, Oct 2005.
- [128] V. Zaborovsky, O. Zayats, and V. Mulukha, “Priority queueing with finite buffer size and randomized push-out mechanism,” in *International Conference on Networking*, Menuires, Apr 11-16, 2010, pp. 316–320.
- [129] K. E. Avrachenkov, N. O. Vilchevsky, and G. L. Shevlyakov, “Priority queueing with finite buffer size and randomized push-out mechanism,” *Performance Evaluation*, vol. 61, pp. 1–16, Jun 2005.
- [130] G. Ahn, A. T. Campbell, A. Veres, and L. Sun, “Supporting service differentiation for real-time and best-effort traffic in stateless wireless ad hoc networks (SWAN),” *IEEE Transactions on Mobile Computing*, vol. 1, pp. 192–207, Sept 2002.
- [131] H. Narman and M. Atiquzzaman, “Energy aware scheduling and queue management for next generation Wi-Fi routers,” in *IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, New Orleans, LA, Mar 9-12, 2015.
- [132] S. Zeadally, S. U. Khan, and N. Chilamkurti, “Energy-efficient networking: past, present, and future,” *Journal of Supercomputing*, vol. 62, pp. 1093–1118, Dec 2012.
- [133] N. Mishra, K. Chebrolu, B. Raman, and A. Pathak, “Wake-on-WLAN,” in *15th International Conference on World Wide Web*, Edinburgh, Scotland, May 23-26, 2006, pp. 761–769.
- [134] W. Siriwongpairat, Z. Han, and K. Liu, “Power controlled channel allocation for multiuser multiband UWB systems,” *IEEE Transactions on Wireless Communications*, vol. 6, no. 2, pp. 583–592, Feb 2007.

- [135] H. Kim, C.-B. Chae, G. de Veciana, and R. Heath, “Energy-efficient adaptive MIMO systems leveraging dynamic spare capacity,” in *Information Sciences and Systems*, Princeton, NJ, Mar 19-21, 2008, pp. 68–73.
- [136] M. Tauber, S. Bhatti, and Y. Yu, “Towards energy-awareness in managing wireless LAN applications,” in *Network Operations and Management Symposium*, Maui, HI, Apr 16-20, 2012, pp. 453–461.
- [137] I. Stojmenovic, *Handbook of wireless networks and mobile computing*. John Wiley & Sons, 2003, vol. 27.

Appendix A

Author's List of Publications

A.1 Peer Reviewed Publications

- [1] Husnu Saner Narman, T. Korkmaz, and S. Limon, “Utilizing distance distribution in determining topological characteristics of multi-hop wireless networks,” *IEEE International Conference on Computing, Networking and Communications (ICNC)*, San Diego, CA, Jan 28-31, 2013, pp 149-154.
- [2] M. S. Hossain, Husnu Saner Narman, and M. Atiquzzaman, “A novel scheduling and queue management scheme for multi-band mobile routers,” *IEEE International Conference on Communications (ICC)*, Budapest, Hungary, June 9-13, 2013, pp. 3787-3791.
- [3] Husnu Saner Narman, M. S. Hossain, and M. Atiquzzaman, “Multi class traffic analysis of single and multi-band queuing system,” *IEEE Global Communications Conference (GLOBECOM)*, Atlanta, GA, Dec 9-13, 2013, pp. 1422-1427.
- [4] Husnu Saner Narman, M. S. Hossain, and M. Atiquzzaman, “h-DDSS: Heterogeneous Dynamic Dedicated servers scheduling in cloud computing,” *IEEE International Conference on Communications (ICC)*, Sydney, NSW, June 10-14, 2014, pp 3475-3480.
- [5] Husnu Saner Narman, M. S. Hossain, and M. Atiquzzaman, “DDSS: Dynamic dedicated servers scheduling for multi priority level classes in cloud computing,” *IEEE International Conference on Communications (ICC)*, Sydney, NSW, June 10-14, 2014, pp 3082-3087.
- [6] Husnu Saner Narman and M. Atiquzzaman, “Carrier components assignment method for LTE and LTE-A systems based on user profile and application,” *IEEE GLOBECOM*

Workshop on Broadband Wireless Access, Austin, TX, Dec 12, 2014, pp 1020-1025.

- [7] Husnu Saner Narman and M. Atiquzzaman, “Joint and partial carrier components assignment techniques based on user profile in LTE systems,” *IEEE Wireless Communications and Networking Conference (WCNC)*, New Orleans, LA, Mar 9-12, 2015, pp. 983-988.
- [8] Husnu Saner Narman and M. Atiquzzaman, “Energy aware scheduling and queue management for next generation Wi-Fi routers,” *IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, New Orleans, LA, Mar 9-12, 2015, pp 148-152.
- [9] Husnu Saner Narman, M. S. Hossain, and M. Atiquzzaman, “Management and analysis of multi class traffic in single and multi-band systems,” *Springer Wireless Personal Communications*, Feb 2015, pp 231-258.
- [10] Husnu Saner Narman and M. Atiquzzaman, “Analysis of joint and partial component carrier assignment techniques in LTE and LTE-A,” *IEEE Global Communications Conference (GLOBECOM)*, San Diego, CA, Dec 6-10, 2015.
- [11] Husnu Saner Narman and M. Atiquzzaman, “Selective periodic component carrier assignment technique in LTE and LTE-A systems,” *IEEE Global Communications Conference (GLOBECOM)*, San Diego, CA, Dec 6-10, 2015.

A.2 Papers under Review

- [1] Husnu Saner Narman, M. S. Hossain and M. Atiquzzaman, “Scheduling Internet of Things Applications in Cloud Computing,” *Submitted to Springer Annals of Telecommunications*.
- [2] Husnu Saner Narman and M. Atiquzzaman, “Joint and Selective Periodic Component Carrier Assignment for LTE-A,” *Submitted to Elsevier Computer Networks*.
- [3] Husnu Saner Narman and M. Atiquzzaman, “Carrier Components Assignment Method Based on User Profile in LTE and LTE-A,” *Submitted to Springer Wireless Networks*.

Appendix B

Acronyms

3GPP	The 3 rd Generation Partnership Project
ACK	Acknowledgment
ARQ	Automatic Repeat Request
BU	Binding Update
CA	Carrier Aggregation
CC	Component Carriers
CQI	Channel Quality Indicator
CQ	Channel Quality
DCI	Downlink Control Information
DQS	Disjoint Queue Scheduler
eNB	eNodeB - Evolved Node B
FDD	Frequency Division Duplex
FD	Frequency Division
FIFO	First In First Out
FSF	Fastest Server First
FSR	Frequency Selective Repeater
HARQ	Hybrid Automatic Repeat Request
IAC	Intra-Band Contiguous
IANC	Intra-Band Non-Contiguous

IENC	Inter-Band Non-Contiguous
IMT	International Mobile Telecommunications
IMT-A	International Mobile Telecommunications-Advanced
j-pCCA	Joint Periodic Component Carrier Assignment
JQS	Joint Queue Scheduler
LAA	Licensed Assisted Access
LA	Link Adaption
LBT	Listen-Before-Talk
LB	Load Balancing
LL	Least Load
LR	Least Load Rate
LTE-A	Long Term Evolution Advanced
LTE-U	Long Term Evolution Unlicensed
LTE	Long Term Evolution
LUF	Least Utilization First
MIMO	Multi-Input Multi-Output
MR	Mobile Router
NACK	Non-Acknowledgment
NRT	Non-Real-Time
OFDMA	Orthogonal Frequency Division Multiple Access
OFDM	Orthogonal Frequency Division Multiplexing
pCCA	Periodic Component Carrier Assignment
PCC	Primary Component Carriers
PDCCH	Physical Downlink Control Channel
PDSCH	Physical Downlink Shared Channel
PHY	Physical Layer
PUCCH	Physical Uplink Control Channel
PUSCH	Physical Uplink Shared Channel

QAM	Quadrature Amplitude Modulation
QCI	Quality Class of Identifier
QoE	Quality of Experience
QoS	Quality of Service
QPSK	Quadrature Phase-Shift Keying
RRC	Radio Resource Control
RRH	Remote Radio Head
RRM	Radio Resource Management
RT	Real-Time
R	Random
s-pCCA	Selective Periodic Component Carrier Assignment
SCC	Secondary Component Carriers
SR	Scheduling Request
SSF	Slowest Server First
TDD	Time Division Duplex
TD	Time Division
TTI	Transmission Time Interval
UCI	Uplink Control Information
UE	User Equipment