3-13-2015

Improving System Performance in Cellular and WBAN Networks via User-Specific QoS and MIMO *In Vivo* Technologies

Chao He
University of South Florida, he.chao@yahoo.com

Improving System Performance in Cellular and WBAN Networks

via User-Specific QoS and MIMO *In Vivo* Technologies

by

Chao He

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Electrical Engineering
College of Engineering
University of South Florida

Major Professor: Richard D. Gitlin, Sc.D.
Huseyin Arslan, Ph.D.
Nasir Ghani, Ph.D.
Srinivas Katkoori, Ph.D.
Zygmunt J. Haas, Ph.D.

Date of Approval:
March 13, 2015

Keywords: *In vivo* communications, MAC, MOS, rate adaptation, system capacity

DEDICATION

To my beloved wife Fei Yu, my darling son Ruiguo He, and my parents.

ACKNOWLEDGMENTS

I would like to express my greatest appreciation to my advisor Dr. Richard D. Gitlin, for his invaluable guidance, advice, and support during my Ph.D. studies. I am also grateful to Dr. Huseyin Arslan, Dr. Nasir Ghani, Dr. Srinivas Katkoori, and Dr. Zygmunt J. Haas for agreeing to serve in my doctoral committee, and for their constructive suggestions and feedback on my research. I would also like to thank Dr. Zygmunt J. Haas, Dr. Huseyin Arslan, Dr. Thomas P. Ketterl, and Dr. Gabriel E. Arrobo for many stimulating and enlightening discussions during the research.

I want to thank all my colleagues at the *i*nnovation in Wireless Information Networking Laboratory (*i*WINLAB). I also want to give my sincere thanks to all the staff at the Department of Electrical Engineering at USF for their kindness and valued assistance.

I also wish to give my deepest gratitude to my beloved wife Fei Yu, my darling son Ruiguo He, and my parents for their love, support and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

This dissertation is composed of two independent studies: Cellular research and WBAN (Wireless Body Area Network) research. Both investigations are directed towards improving the system performance in wireless communication systems in terms of Quality of Service (QoS) and system capacity.

For the Cellular research part, this dissertation will present novel user-specific QoS requirements as defined by their respective Mean Opinion Score (MOS) formulas, and associated schedulers for wireless applications and systems that optimize spectral allocation. User-specific QoS requirements are defined and several methods to make use of such requirements to maximum the spectral utilization are presented. Five User-Specific QoS Aware (USQA) schedulers are proposed that consider the user-specific QoS requirements in the allocation of spectral resources. Schedulers are introduced that dynamically adapt to the user-specific QoS requirements to improve quality as measured by the MOS, or the system capacity, or can improve both the quality and system capacity.

Due to the different cell deployment arrangements and inter-cell interference in heterogeneous networks in comparison to homogeneous networks, the USQA scheduling is also analyzed and the system performance is evaluated in such networks. Throughput improvements of File Transfer Protocol (FTP) applications benefiting from the rate adaptation and MAC (Media Access Control) scheduling algorithms for video applications that incorporate user-specific QoS requirements to improve system capacity are demonstrated.

Another novel approach recognizes that the user-specific frequency sensitivity can be used to improve capacity. There is considerable variation in the audible range of frequencies that can be perceived by individuals, especially at the high frequency end, which is primarily affected by a gradual decline with age. This can be utilized to improve the system performance by personalizing the VoIP

codecs and decreasing the user's source data rate for people from an older age group and thus increase the system capacity.

Given the potentially substantial system performance gain resulting from the USQA schedulers, it is critical to analyze their feasibility and complexity in practical LTE (4G cellular) and future wireless systems. From the LTE system perspective, LTE QoS end-to-end signaling procedures are addressed, and corresponding protocol adaptations are analyzed in order to support the USQA schedulers. In addition, the optimal scheduling period is analyzed that trades off between performance gain and implementation complexity.

In the WBAN research, MIMO (Multiple Input Multiple Output) *in vivo* antenna technologies are introduced and are motivated by the high data rate requirements of wirelessly transmitted low-delay High Definition (HD) video during Minimally Invasive Surgery (MIS). MIMO *in vivo* technologies are proposed to be used in the *in vivo* environments to enhance and determine the maximum data transmission rate while satisfying the Specific Absorption Rate (SAR) power limitations. Various factors are considered in the MIMO *in vivo* study including antenna separation, antenna angular positions, human body size, and system bandwidth to determinate the maximum data rate that can be supported.

CHAPTER 1. INTRODUCTION

Maximizing utilization of the limited spectral resources and satisfying the ever-increasing requirements for higher data rates are among the main research challenges in wireless communications. Numerous research studies are focused on various aspects of improving the spectral efficiency and maximizing the system capacity for these conditions. For example, in LTE-Advanced [1], [2], key wireless physical layer technologies such as small cells [3], [4], multi-antennas [5], CA (Carrier Aggregation) [6], and CoMP (Coordinated Multi-Point operation) [7] are introduced to improve the system performance to meet the 4G LTE-Advanced requirements [8], [9]. Spectral efficiency can also be improved in the upper layers via efficient MAC resource scheduling algorithms [10], [11].

The research presented in this dissertation is directed to improve the system performance in Cellular and WBAN by using novel user-specific QoS requirements in wireless resource scheduling for cellular systems and MIMO *in vivo* technologies in the WBAN physical layer respectively.

In this chapter, motivation for both the cellular research and WBAN research is presented, along with the research objectives for both type of networks.

1.1 Research Motivation

1.1.1 Motivation for Improving Cellular Utilization

In today's wireless 4G LTE networks, the spectral allocation of resources is either independent of the application's specific Quality of Service (QoS) requirements and of the users' specific perceived QoS, or at most relies on a set of pre-defined fixed priorities [12], [13]. Although in these standards, the MAC and the PHY layers have an increased role in optimizing the usage of the spectral resources and implementing link quality-aware techniques, nevertheless, optimization is still largely independent of the application context, the users' requirements, and the users' perception of performance degradation. In particular, the standards do not take into account the QoS required by different applications and their

users, beyond simply assigning fixed priorities to traffic classes. Indeed, from the user's perspective, the QoS required by different applications can be quite variable. Similarly, for a given application type, different users may require different levels of QoS.

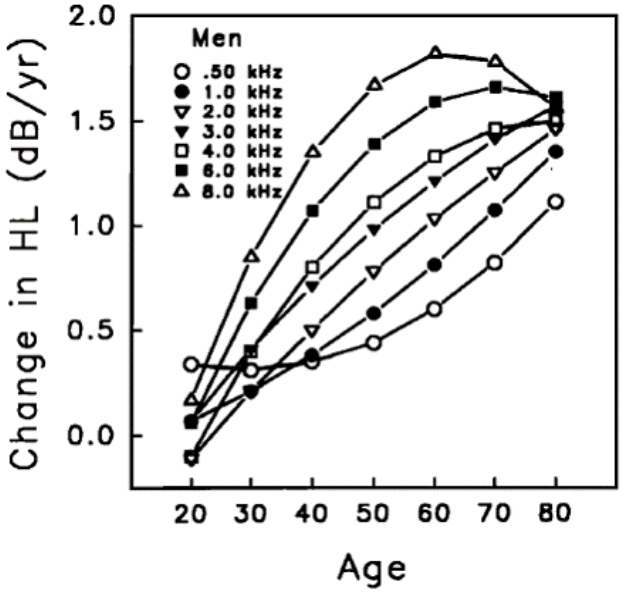

Figure 1.1 Hearing loss [HL] as a function of frequency and age [14]

For Voice over IP (VoIP) applications, as a motivating example, consider the fact that the perceived voice quality of different languages may differ substantially with the same data rate and BER (Bit Error Rate), because of the different spectral content of such languages and because of a particular user's auditory spectral response (with variations typically due to aging, as shown in Figure 1.1) [14], making the user more or less sensitive to a particular type of distortion or frequency band. Consequently, the same amount of degradation, as experienced by individual applications and their users, may have substantially different perceptual effects. Another example is the varying talking environments, where some users have a conversation under very noisy conditions, whereas some other users converse under very quiet conditions, thus making users more or less sensitive to packet losses. If the same spectral resource is allocated to users in very noisy and quiet backgrounds, then quite a different user experience will likely be incurred. As another example, consider that people from different age groups normally have

different sensitivity to high frequency sound content [14], which can be exploited to maximize the system capacity by reducing the bit rate for users with reduced frequency sensitivity.

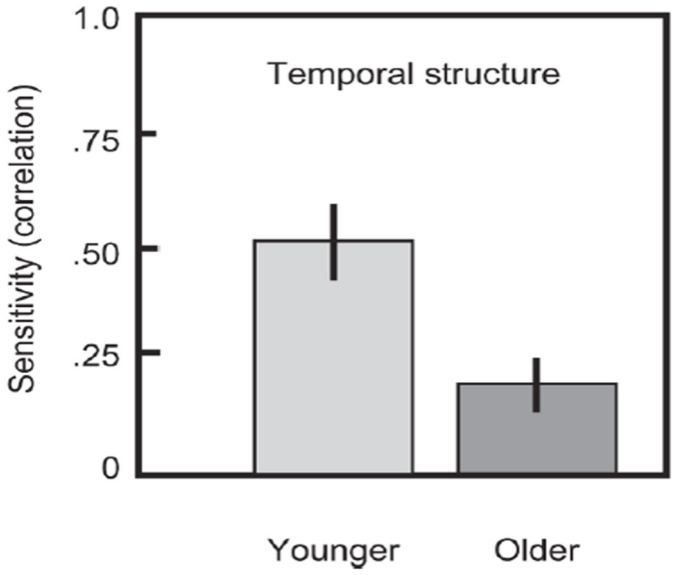

Figure 1.2 Video sensitivity as a function of different age group [15]

For video applications, older individuals, compared to young adults, are less sensitive to spatial forms defined by a temporal structure [15], as shown in Figure 1.2. So for many older people a lower video data rate provides the same user experience (i.e., QoS) as the full rate video does for younger people. Hence, the user-specific QoS requirements can be utilized by the scheduler to differentiate between the users and make better use of the wireless spectral resources.

According to the statistics for the population distribution by age done by the United States Census Bureau [16], about 25.8% of the population are older than 55 and 79.2% of the population are older than 15. So about 1/3 of the population among the people who can potentially use mobile phones (i.e., older than 15) are older than 55, and these people can be potentially identified as users with less sensitivity to distortion for VoIP applications and less sensitivity to frame rate for video applications. In fact, since the number of older people is continuously increasing, and more and more young people are subject to the hearing loss at high frequencies, this ratio can be higher [17], [18] as shown in Figure 1.3.

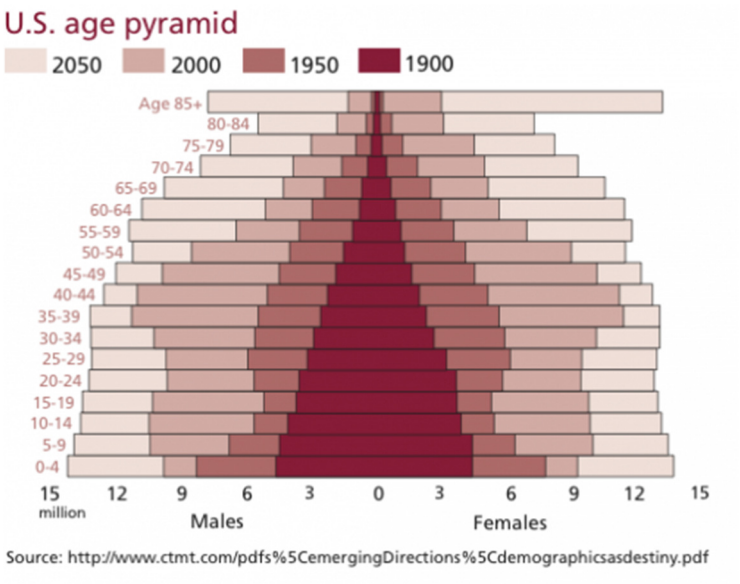

Figure 1.3 U.S. age pyramid

Furthermore, we observe that some previous studies (e.g., [19]–[22]), which use the QoS characteristics of an underlying application (typically expressed as a function of the MOS), allocate average spectral resources to applications, independently of the application's actual specific QoS requirement. In the literature, there are MAC schedulers that take into account instantaneous data rates and user's QoS [23]–[27]. In [28], content adaptation techniques and protocol adaptation techniques are used to optimize the interactions between applications/content and the underlying networks to accelerate the content delivery. However, in these schemes, no user-specific and age related QoS requirements have been considered in the MOS functions and the scheduling algorithms. Thus, in such schemes, especially for applications with widely varying QoS requirements (even for the same type of application), either the spectral resources are not efficiently utilized or the MOS is significantly degraded.

1.1.2 Motivation for Improving WBAN Performance

One appealing aspect of the emerging *Internet of Things* is to consider *in vivo* networking as a rich application domain for wireless technology in facilitating wirelessly enabled healthcare. Wireless technology has the potential to synergistically advance healthcare delivery solutions by creating new science and technology for *in vivo* wirelessly networked cyber-physical systems of embedded devices.

In vivo wireless networks have certain characteristics and requirements such as low complexity, limited transmission and processing power, low latency, high reliability, and transmission characteristics such as being a highly lossy [29] and dispersive radio frequency (RF) channel [30], with possible near-field operation [31]. It is the purpose of this dissertation to demonstrate that owing to the highly dispersive nature of the *in vivo* channel, achieving stringent performance requirements will be facilitated by the use of multiple-input multiple-output (MIMO) communications [32] to achieve enhanced data rates. For example, one potential application for MIMO *in vivo* communications is wirelessly transmitted low-delay High Definition video during Minimally Invasive Surgery (MIS) [33], as shown in Figure 1.4. With surgical and technological breakthroughs pioneered at the University of South Florida (USF) and Florida Hospital by our surgical colleagues, Alexander Rosemurgy and Sharona Ross, for certain procedures MIS has evolved into LESS (Laparoscopy-Endoscopy Single-Site surgery) [34]. Our research group has created a Miniature Anchored Remote Videoscope for Expedited Laparoscopy (*MARVEL*) robotic camera module, which is a wireless research platform for advancing MIS and requires high bit rates (~80–100 Mbps) for HD video transmission and low latency for proper operation during surgery.

It is widely recognized that MIMO technology, the use of multiple antennas both in transmitter and receiver, can significantly improve the capacity and performance of the communication system in comparison to the conventional system with single antenna. In modern communications systems, the combination of MIMO and OFDM technology [35] is extremely popular and takes advantage of multipath to materially improve the radio transmission performance. There has been some research that focuses on MIMO for WBANs, i.e., there are a few models for MIMO system that can be applied to WBANs. In [36], the authors place the antennas on human clothing and analyze the performance of the proposed wearable MIMO system, which has a significantly better performance than the previous system on a handheld platform. The wideband body-to-body radio channel with MIMO was investigated in [37] and the authors also present several characterizations of the channel such as path loss, body shadowing, and small-scale fading. However, to the best of our knowledge, MIMO system for *in vivo* environments has not been fully studied in the literature, and due to the tremendous opportunity to increase the reliability and the

communications efficiency for the *in vivo* channel and support novel applications (e.g., transmitting HD video from inside the human body) for the *in vivo* environment, in this dissertation, we present a detailed study of the performance of MIMO system in the *in vivo* environment.

1.2 Research Objectives

1.2.1 Cellular Objectives

- Either improve user satisfaction, as measured by the MOS, by adapting MAC scheduling algorithms to user-specific QoS requirements, and/or
- Improve system capacity by trading off the spectral resource allocations for the user-specific QoS requirements, while still maintaining acceptable levels of MOS.

1.2.2 WBAN Objectives

- Improve *in vivo* system performance by utilizing MIMO technologies; i.e., MIMO *in vivo*.
- MIMO *in vivo* performance optimization via antenna placement to achieve stringent requirements such as enhanced data rates and specified specific absorption rate (SAR) levels.

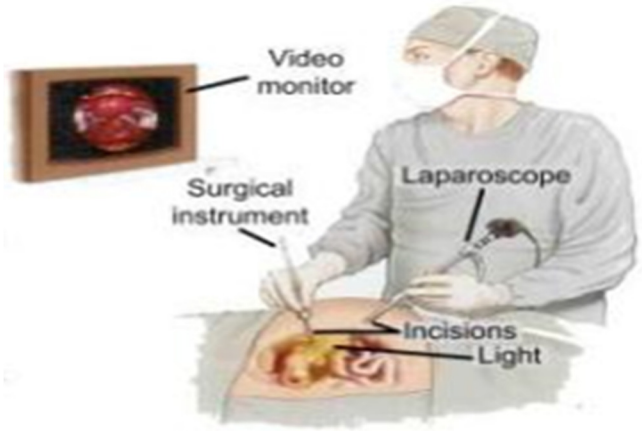

Figure 1.4 Minimally Invasive Surgery (MIS)

1.3 Dissertation Outline

The dissertation is organized as follows:

- CHAPTER 2 provides an overview of Cellular 4G LTE and WBAN networks that is a useful background to understand the subsequent chapters. In the LTE network overview, the

network architecture, QoS-related policy and charging control architecture, QoS parameters, and QoS parameters mapping functions are described. Another very important protocol, i.e., the Session Initiation Protocol (SIP), is also addressed, which is critical for the initial QoS negotiation during session initiation. In the WBAN network overview, the system characteristics, representative applications, design challenges, and application scenarios are described.

- CHAPTER 3 presents the MOS definitions in the existing literature for VoIP, video, and FTP applications. The user-specific MOS formulas are defined for these applications considering user-specific QoS requirements. To make better use of the user-specific QoS requirements and optimize the spectral allocation of wireless resources correspondingly, two alternatives to improve the system capacity are proposed, i.e., a user-specific MCS[1] selection method and a user-specific QoS aware scheduling method. The user-specific QoS aware scheduling method is the main focus of Chapter 4.
- CHAPTER 4 describes the user-specific QoS aware schedulers for VoIP, video, and FTP applications, considering the user-specific QoS requirements. These schedulers are proposed to either improve the MOS, the system capacity, or both the MOS and system capacity. Extensive system simulation has been done to verify the benefits of these schedulers in both the homogeneous and heterogeneous scenarios. Another very promising area of research described in this chapter is a user-specific frequency sensitivity QoS study, which maximize the system capacity by decreasing the user's source data rate through personalized VoIP codecs.
- CHAPTER 5 presents the implementation details of user-specific QoS aware schedulers in current LTE systems. These implementation details cover the protocol adaptation that must occur to support the user-specific QoS method including the scheduling period that relates to the complexity of user-specific QoS aware schedulers.

[1] MCS = Modulation and Coding Scheme.

- CHAPTER 6 provides the detailed analysis and performance evaluation of MIMO technologies used in human's *in vivo* environments based upon the ANSYS HFSS software that contains a complete human body model, and OFDM-based IEEE 802.11n. Different scenarios considering varying antenna separation distances, angular positions, human body size, and system bandwidth are analyzed and evaluated to determine the maximum supportable data transmission rate with specified SAR [Specific Absorption Rate] power limitations satisfied.
- CHAPTER 7 summarizes the research contributions in this dissertation, along with recommendations for future work.

CHAPTER 2. LTE AND WBAN NETWORK OVERVIEW

This chapter presents an overview of the LTE network with an emphasis on QoS and a WBAN network overview that provide the background materials for the subsequent chapters. The rest of the chapter is organized as follows. Section 2.1 gives an overview of the LTE network. The WBAN network overview is provided in section 2.2. Finally in section 2.3, we conclude this chapter.

2.1 LTE Network Overview

In the past decades, we have witnessed significant exciting progress in wireless communication systems from the first generation of the analog mobile systems, to the second generation of digital mobile systems, to the third generation (3G) of the first broadband data mobile system, to the current fourth generation (4G) LTE (Long Term Evolution) communication system. The evolution of 3G systems into LTE system is driven by the creation and development of new services for mobile devices, and is enabled by the advancement of technologies available for mobile systems [1], [10].

LTE is a standard for wireless communication based on and evolved from the GSM/EDGE 2G and UMTS/HSPA 3G network technologies, aimed to providing higher data rate, higher capacity and lower delay using an OFDM radio technology and improved all-IP core network technology. The LTE standard is developed by the 3GPP (3rd Generation Partnership Project) and its first release is specified in the Release 8 series, and continuously evolved to the current Release 12 series.

2.1.1 LTE Network Architecture

The LTE network architecture (non-roaming) including the network elements and the standardized interfaces [38] is presented in Figure 2.1. The LTE network is comprised of the EPC (Evolved Packet Core) and the E-UTRAN (Evolved Universal Terrestrial Radio Access Network). The EPC consists of many logical nodes, and the E-UTRAN is made up of eNodeBs (evolved NodeB) --- base stations. Each of these network elements is interconnected by means of standardized interfaces.

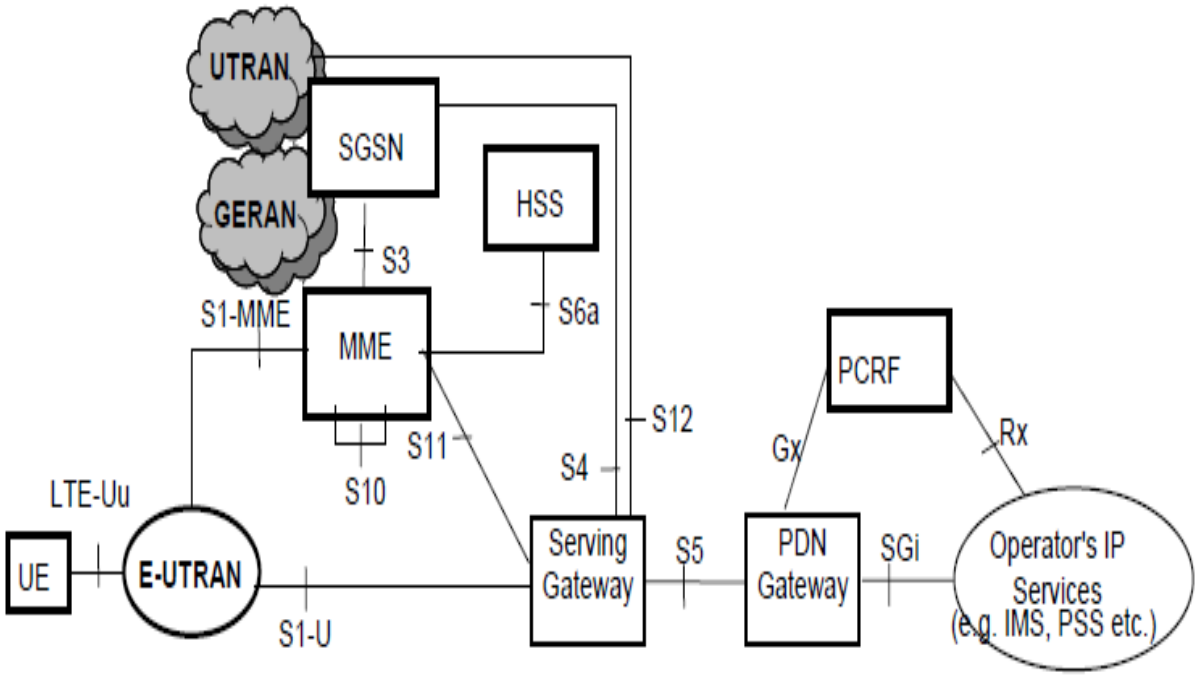

Figure 2.1 LTE network architecture (non-roaming)

The EPC is responsible for the overall control of the UE (User Equipment) and the establishment of the bearers. It is composed of the following major network elements: the S-GW (Serving Gateway), the PDN-GW (PDN [Packet Data Network] Gateway), the MME (Mobility Management Entity), the PCRF (Policy and Charging Rules Function), and the HSS (Home Subscriber Server). The EPC is connected to the external networks such as the IP Multimedia Subsystem (IMS) [39]. The S-GW is the point of interconnection between the eNodeB and the EPC. It serves the UE by routing and forwarding the IP packets from the eNodeB to the EPC. The PDN-GW is the point of interconnection between the EPC and the packet data network. The packet data network may be an operator external public or private packet data network or an intra-operator packet data network, e.g., for provision of IMS services. The PDN-GW routes packets to and from the external networks. The MME deals with the control plane. It handles the signaling related to mobility and security for E-UTRAN access. The HSS is a database that contains subscriber-related information. The PCRF is a logic entity in the EPC. It is mainly responsible for charging and policy control (e.g., QoS control) to the functionalities in the PCEF (Policy Control

Enforcement Function) that resides in the PDN-GW. The IMS is an architectural framework for delivering Internet Protocol (IP) multimedia services.

The following are the standardized interfaces in the LTE network:

- S1-MME: Reference point for the control plane protocol between the E-UTRAN and the MME.
- S1-U: Reference point for the data plane protocol between the E-UTRAN and the S-GW.
- S5: Provides user plane tunneling and tunnel management between the S-GW and the PDN-GW.
- Gx: Provides transfer of (QoS) policy and charging rules from the PCRF to the PCEF in the PDN-GW [40].
- S10: Reference point between MMEs for MME relocation and MME to MME information transfer.
- S11: Reference point between the MME and the S-GW.
- SGi**:** The reference point between the PDN-GW and the packet data network.
- Rx: The Rx reference point resides between the PCRF and the AF (Application Function) [41]. It is used to exchange application level session information between the PCRF and the AF.

2.1.2 Policy and Charging Control Architecture

The PCC functionality is comprised by the major functions of the Policy and Charging Enforcement Function (PCEF), the Policy and Charging Rules Function (PCRF), the Application Function (AF), and the Subscription Profile Repository (SPR) [13]. The PCC (Policy and Charging Control) architecture is shown in Figure 2.2.

The Sp reference point lies between the SPR and the PCRF. The Sp reference point allows the PCRF to request subscription information related to the transport level policies from the SPR based on a subscriber ID, a PDN identifier etc.

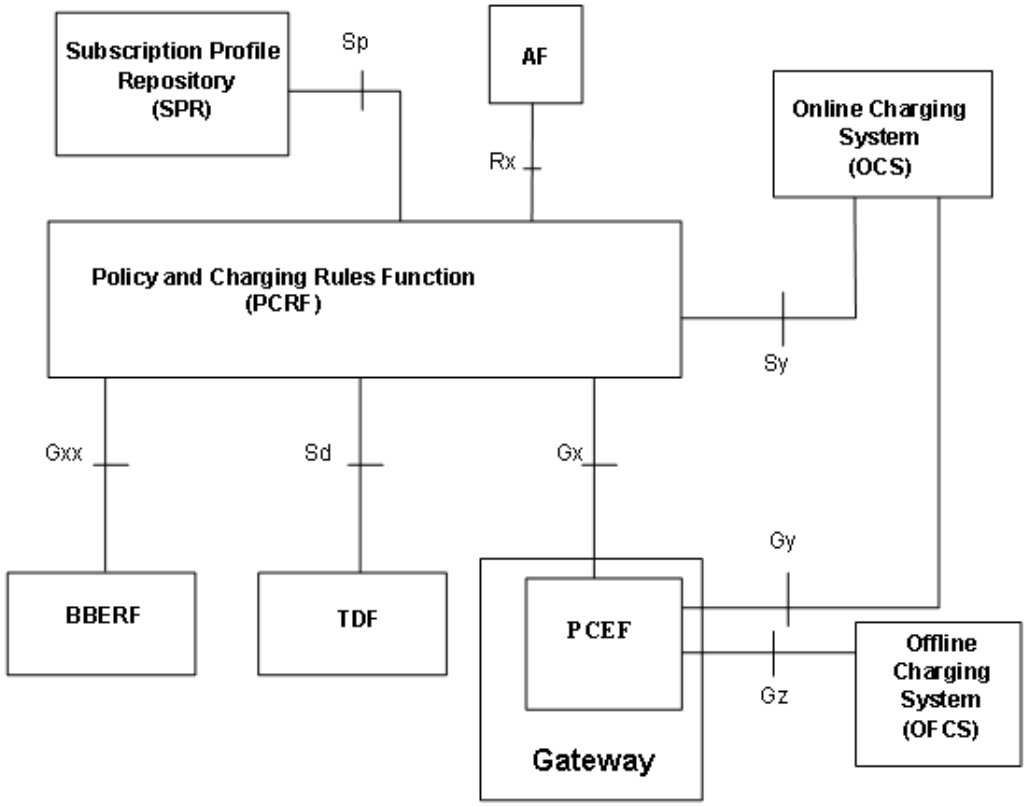

Figure 2.2 Policy and charging control architecture (non-roaming)

2.1.3 LTE QoS Parameters

The QoS parameters are QCI (QoS Class Identifier), ARP (Allocation and Retention Priority), GBR (Guaranteed Bit Rate), and MBR (Maximum Bit Rate) [13].

The term bearer refers to an edge-to-edge association between the UE and the Gateway, which uniquely identifies packet flows that receive a common QoS treatment between the UE and the Gateway [12], [42].

Each bearer is assigned a QCI by the network. Each QCI is characterized by a set of characteristics of resource type (GBR or Non-GBR), priority, packet delay budget and acceptable packet loss rate. The QCI is used in the eNodeB to determine the way it is handled (e.g., scheduling weights, admission thresholds, queue management thresholds, link layer protocol configuration, etc.) [1],[10]. QCIs have been standardized to ensure that applications / services mapped to that QCI receive the same

minimum level of QoS in multi-vendor network deployments. The set of standardized QCIs and their characteristics are provided in Table 2-1 [13].

The ARP specifies the control plane treatment that the bearers receive. More specifically, the ARP is used to decide whether a bearer establishment or modification request can be accepted or must be rejected due to resource limitations.

The MBR and GBR are defined only for GBR bearers. The MBR is defined as the bit rate that the bearer traffic may not exceed, and the GBR is the bit rate that the network guarantees (e.g., through the use of an admission control function).

2.1.4 QoS Parameters Mapping

Several QoS parameters mapping functions are needed during PCC interaction. These functions are located at the AF, PCRF, PCEF and UE. The main purpose of these mapping functions is the conversion of QoS parameters from one format to another.

The general QoS mapping framework is shown in Figure 2.3 and the network initiated QoS control procedures [43] is as follows:

- The AF can map from SDI (Session Description Language) within the AF session signaling to service information passed to the PCRF over the Rx interface.
- The PCRF shall map from the service information received over the Rx interface to the Authorized IP QoS parameters that shall be passed to the PCEF/BBERF (Bearer Binding and Event Reporting Function) via the Gx/Gxx (Gxx reference point is located between PCRF and the BBERF) interface. The mapping is performed for each IP flow. Upon a request from the PCEF/BBERF, the PCRF combines per direction the individual Authorized IP QoS parameters per flow.
- The PCEF/BBERF shall map from the Authorized IP QoS parameters received from the PCRF to the access specific QoS parameters, which are the QoS parameters that the MAC layer can access.

Table 2-1 Standardized QCI characteristics

QCI	Resource Type	Priority Level	Packet Delay Budget	Packet Error Loss Rate	Example Services
1	GBR	2	100 ms	10^{-2}	Conversational Voice
2		4	150 ms	10^{-3}	Conversational Video (Live Streaming)
3		3	50 ms	10^{-3}	Real Time Gaming
4		5	300 ms	10^{-6}	Non-Conversational Video (Buffered Streaming)
65		0.7	75ms	10^{-2}	Mission Critical user plane Push To Talk voice (e.g., MCPTT)
66		2	100 ms	10^{-2}	Non-Mission-Critical user plane Push To Talk voice
5	Non-GBR	1	100 ms	10^{-6}	IMS Signalling
6		6	300 ms	10^{-6}	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)
7		7	100 ms	10^{-3}	Voice, Video (Live Streaming) Interactive Gaming
8		8	300 ms	10^{-6}	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)
9		9			
69		0.5	60 ms	10^{-6}	Mission Critical delay sensitive signalling (e.g., MC-PTT signalling)
70		5.5	200 ms	10^{-6}	Mission Critical Data (e.g. example services are the same as QCI 6/8/9)

2.1.5 SIP Protocol

The Session Initiation Protocol (SIP) [44] is an application-layer control protocol for creating, modifying, and terminating sessions such as Internet multimedia conferences, Internet telephone calls, and multimedia distribution. The SIP messages used to create sessions carry session descriptions that allow participants to agree on a set of compatible media types. These session descriptions are commonly formatted using SDP (Session Description Protocol). When used with SIP, the offer/answer model provides a limited framework for negotiation using SDP.

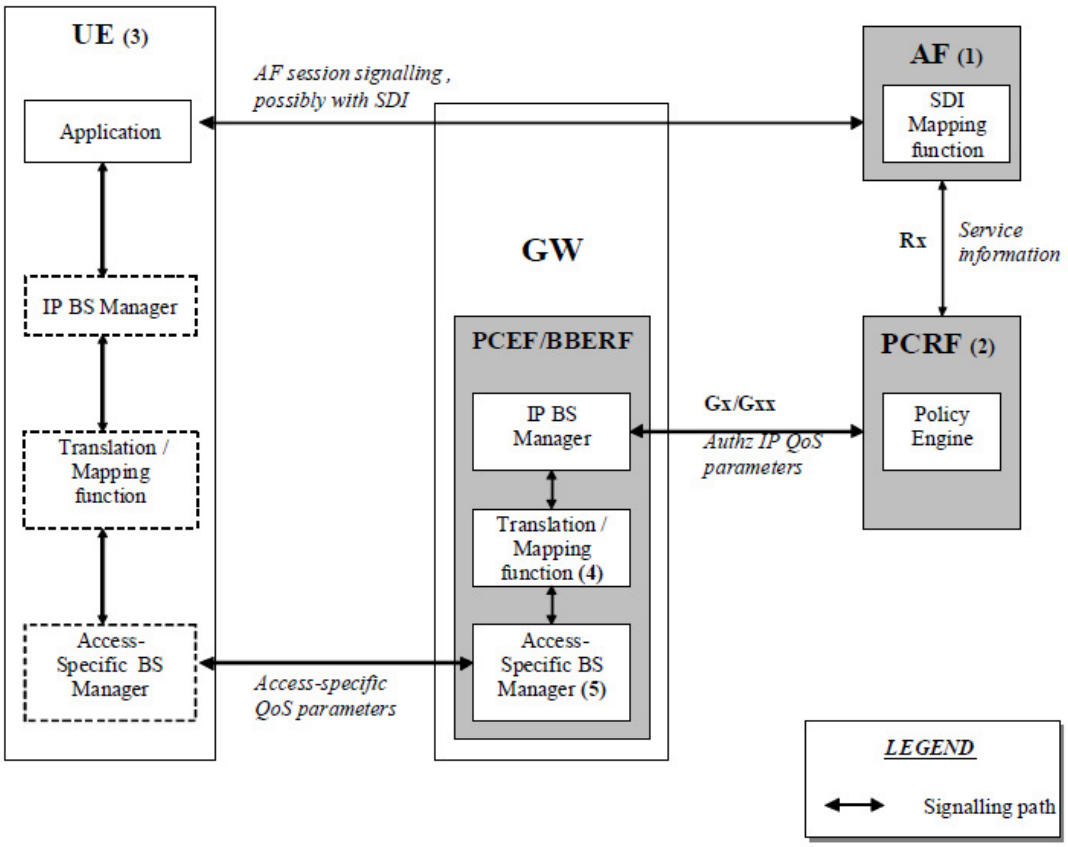

Figure 2.3 Framework for QoS mapping

SIP is not a vertically integrated communications system. SIP is rather a component that can be used with other IETF (Internet Engineering Task Force) protocols to build a complete multimedia architecture. Typically, these architectures will include protocols such as the Real-time Transport Protocol (RTP), described in the RFC 3550 [45], for transporting real-time data and providing QoS feedback, the Real-Time streaming protocol (RTSP) (RFC 2326 [46]) for controlling delivery of streaming media, and the Session Description Protocol (SDP) (RFC 4566 [47]) for describing multimedia sessions. Therefore, SIP should be used in conjunction with other protocols in order to provide complete services to the users. However, the basic functionality and operation of SIP does not depend on any of these protocols.

A SIP message is either a request from a client to a server, or a response from a server to a client. The details of the session, such as the type of media, codec, or sampling rate, are not described using SIP.

Rather, the body of a SIP message contains a description of the session, encoded in some other protocol format such as SDP (RFC 4566 [47]).

2.2 WBAN Network Overview

The increasingly aging population in many countries and the rising costs of healthcare have triggered the emergence and rapid development of a novel wireless network [48]–[50], referred to a WBAN (Wireless Body Area Network), to enhance current healthcare practices. A WBAN is a network typically formed by a collection of miniaturized, low-complexity, and low-power devices that are located on, in or around the human body that are used to continuously monitor human vital signs and advance human healthcare. Compared with traditional networks, WBANs have their own characteristics such as highly lossy communication medium, and very stringent application requirements in terms of high reliability, low energy efficiency, and low device complexity that constitute major challenges to the design and application of WBAN networks.

2.2.1 WBAN Applications

WBAN applications span a wide area such as medical care, military, and entertainment. Typical applications include [33], [49]:

- Wirelessly transmitted low-delay High Definition video during Minimally Invasive Surgery (MIS).
- Patient Monitoring including transmitting vital signals such as ECG/EMG/EEG, body temperature, respiration rate, blood pressure, body implant parameters, chest sounds, and glucose monitoring.
- Monitoring and program changes for pacemakers and implantable cardiac defibrillators, control of bladder function, and restoration of limb movement.
- Entertainment and gaming.
- Military applications such as assessing soldier fatigue and battle readiness.

2.2.2 WBAN PHY Layers

The IEEE Std 802.15.6 [51], [52] is a standard for short-range, wireless communications in the vicinity of, or inside, a human body (but not limited to humans). It uses the ISM (Industrial, Scientific and Medical) and MICS (Medical Implant Communication Service) bands, as well as frequency bands in compliance with applicable medical and communication regulatory authorities. It allows devices to operate on very low transmit power for safety to minimize the SAR into the body and increase the battery lifetime. It supports QoS, for example, to provide for emergency messaging. Since communications sessions can involve sensitive information, it also provides strong security. Three PHYs have been defined in IEEE Std 802.15.6 [52]:

- HBC (Human Body Communications) PHY:

  HBC PHY uses the human body itself as the communication medium. The center frequency is 21 MHz, with a bandwidth of 5.25 MHz.

- NB (Narrowband) PHY:

  NB PHY device support transmission and reception in at least one of the following optional frequency bands: 402-405 MHz, 420-450 MHz, 863-870 MHz, 902-928 MHz, 950-958 MHz, 2360-2400 MHz, and 2400-2483.5 MHz.

- UWB (Ultra-wide Band) PHY:

  UWB PHY is divided into a low (3.25-4.75 GHz) and a high (6.6-10.25 GHz) band, both having operating channels of 500 MHz bandwidth. UWB PHY is specifically used to provide performance for high quality, low complexity and ultralow power operations.

2.2.3 WBAN Design Challenges

The WBAN design challenges are as follows [48], [50], [53]:

- Antenna Design

  Antenna design for WBAN applications is a challenging issue due to restrictions on the antenna size, antenna material, shape of the antenna and hostile RF environment.

- Channel Model

Determining an accurate and statistical channel model plays a crucial role in the design of PHY technologies. Experimental channel modeling for implants and wearable devices is difficult due to the necessary involvement of human subjects.

- PHY Layer Protocol

  Physical layer protocol design requires minimizing power consumption and high reliability, without compromising the quality of communications.

- Low Power Consumption

  When one or more devices need to be implanted in or worn on a human body, it is of outmost importance to reduce the inconvenience and stress caused by the battery replacement or recharge, which in some cases may involve surgery.

- Security

  The communication of health related information between nodes in a WBAN and the communication from the WBAN to servers in other networks should be strictly private, confidential, and should be encrypted to protect the patient's privacy.

2.2.4 WBAN Scenarios

A list of WBAN scenarios are identified in IEEE 802.15.4 TG6 [54] as illustrated in Figure 2.4. The scenarios are determined based on the location of the communicating nodes (i.e., implant, body surface and external). The communication link can occur from an implanted node to an implanted node (i.e., CM1), an implanted node to a body surface node (i.e., CM2), a body surface node to a body surface node (i.e., CM3), or a body surface node to an external node (i.e., CM4).

2.3 Concluding Remarks

In this chapter, an overview of LTE and WBAN networks, which serves as a basis for the subsequent chapters was presented. The LTE network architecture, LTE QoS mechanisms and SIP

protocol were presented for the LTE networks. In WBAN networks, the WBAN applications, WBAN PHY layers, WBAN design challenges, and WBAN scenarios were described.

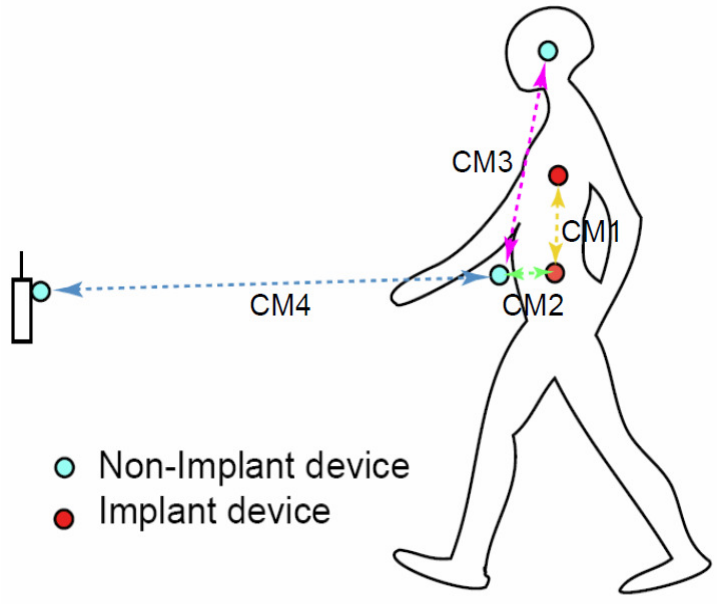

Figure 2.4 Communication links for WBANs

CHAPTER 3. USER-SPECIFIC QOS REQUIREMENTS

3.1 Introduction

As described in Chapter 1, from the user's perspective, the QoS required by different applications can be quite variable. Similarly, for a given application type, different users may be able to properly operate with different levels of QoS. In order to make better use of the user-specific QoS requirements to optimize the system performance, it is important to characterize the different user-specific QoS requirements, which is the aim of this chapter[2].

We begin by presenting MOS formulas for VoIP, video, and FTP applications from the literature. Based upon the existing MOS formulas, user-specific QoS formulas for VoIP, video, and FTP applications are proposed that consider different user QoS requirements. To make use of the user-specific QoS requirements to improve the system capacity, two such methods are proposed, where either the MCS index selection (user-specific MCS selection method) or the source data rate (user-specific QoS aware scheduling method) are optimized by considering user-specific QoS requirements.

The rest of the chapter is organized as follows. Section 3.2 presents the existing MOS formulas for VoIP, video, and FTP applications. Correspondingly, the user-specific QoS formulas for VoIP, video, and FTP applications are proposed in sections 3.3, 3.4, and 3.5 respectively. Section 3.6 gives two potential alternatives to improve system capacity by doing user-specific MCS selection or user-specific source data rate adaptation. Finally in section 3.7, we conclude this chapter.

3.2 MOS Formulas

3.2.1 VoIP Applications

The traditional method of assessing voice quality is to conduct subjective tests by groups of human listeners [19]. The results of these subjective tests are averaged to get a MOS value. However, it is

[2] Portions of this chapter were previously published in [70]. Permission is included in Appendix A.

not practical for such subjective tests to be done online to assess the voice quality. Because of this, the ITU-T has standardized a perceptual evaluation of speech quality (PESQ) model [55] that predicts with high correlation the MOS that would be given in a typical subjective test. The PESQ measures one-way voice quality, a signal is injected into the system under test and the degraded output is compared by PESQ with the input reference signal. The output of the PESQ algorithm is a numerical value of MOS. The mapping between MOS and user satisfaction is presented in Table 3-1. The PESQ algorithm is computationally too expensive to be used in real-time scenarios.

Table 3-1 Relation between the MOS and user satisfaction

MOS (Lower Limit)	User satisfaction
4.34	Very satisfied
4.03	Satisfied
3.60	Some users dissatisfied
3.10	Many users dissatisfied
2.58	Nearly all users dissatisfied

The Adaptive Multi-Rate (AMR) audio codec is an audio data compression scheme that is used in wireless communication systems (e.g., LTE, as well as 3G UMTS and 2G GSM) and is optimized for Voice Over IP (VoIP) speech coding. AMR consists of a multi-rate speech codec that encodes speech signals at variable bit rates ranging from 4.75 to 12.2 Kbps [56].

The E-Model algorithm [57] is a computational model for objective call quality assessment, as described in the G.107 recommendation by the ITU-T. The computation of the MOS is defined as follows:

$$R = R_0 - I_d - I_{eff} \tag{3-1}$$

where R_0 is the basic signal-to-noise ratio which has a default value of 93.2 [58], [59], I_d represents the impairments due to delay, which is the same for all the codec modes, and I_{eff} represents the effect of packet losses and depends on the codec (e.g., AMR, G.711, G.729) that is used. The parameter I_d is calculated as [59]:

$$I_d = 0.024d + 0.11(d - 177.3)U(d - 177.3) \tag{3-2}$$

where d is the end-to-end delay in milliseconds and U is the unit step function [59].

For AMR codecs [57],

$$I_{eff} = I_e + (95 - I_e)\left(\frac{100P_{pl}}{\frac{100P_{pl}}{BurstR} + B_{pl}}\right) \tag{3-3}$$

where P_{pl} represents packet loss ratio, $BurstR$ is the Average length of observed bursts in an arrival sequence to the Average length of bursts expected for the network under "random" loss ratio. In this dissertation we assume the packet loss is independent and hence we set $BurstR = 1$. The parameter B_{pl} is the robustness factor which is set to 10 for all AMR codec modes. The parameter I_e is defined for all AMR codec modes in Table 3-2 [60], where eight AMR-NB codec modes are defined for LTE as in [56].

Table 3-2 I_e value for all AMR codec modes

AMR codec mode	I_e
AMR12.2K	5
AMR10.2K	9
AMR7.95K	15
AMR7.4K	16
AMR6.7K	20
AMR5.9K	23
AMR 5.15K	27
AMR 4.75K	29

For G.711 codecs [59],

$$I_{eff} = 0 + 30\,ln(1 + 15P_{pl}) \tag{3-4}$$

For G.729 codecs [59],

$$I_{eff} = 11 + 40\,ln(1 + 10P_{pl}) \tag{3-5}$$

The parameter R is converted to MOS according to the following equation:

$$MOS = \begin{cases} 1, & \text{when } R < 0 \\ 1 + 0.035R + R(R - 60)(100 - R) & \\ 7 \cdot 10^{-6}, & \text{when } R \in [0, 100] \\ 4.5, & \text{when } R > 100 \end{cases} \tag{3-6}$$

With AMR10.2K VoIP users as an example, according to the ITU-T G.107 E-model, the MOS value depends upon both the packet loss ratio and delay. Figure 3.1 illustrates the relationship between the MOS value and the delay and packet loss ratio. The higher the delay, or the higher the packet loss ratio, the lower the MOS value. Thus, users' MOS can be improved either through the optimization of the time domain scheduler or frequency domain scheduler to reduce the delay or packet loss ratio, respectively.

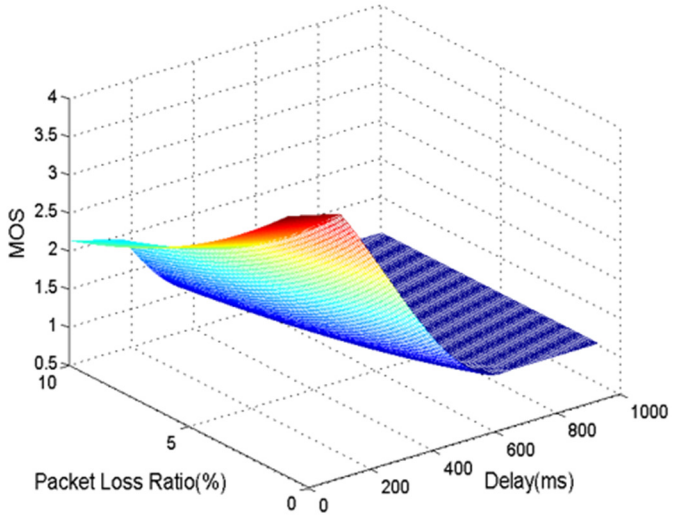

Figure 3.1 VoIP MOS as a function of packet loss and delay for AMR12.2K

For a given acceptable mouth-to-ear end-to-end delay of 150 ms [61], Figure 3.2 shows the MOS as a function of packet loss ratio for different VoIP codec types. From Figure 3.2 we can see that as the AMR codec rate decreases, the MOS value decreases, given the same packet loss ratio. If a lower AMR codec rate is used by the UEs, more users can be supported by the system if fewer allocated physical resources and a similar channel coding rate is assigned for each user, thus capacity improvement can be achieved, but the MOS will probably decrease.

3.2.2 Video Applications

In the current literature, there are several methods to objectively estimate the video quality. However, they are either dependent upon the reference video (i.e., original video) to compute the distortion [62]–[64] or too complex to be used in practical systems [65]–[67].

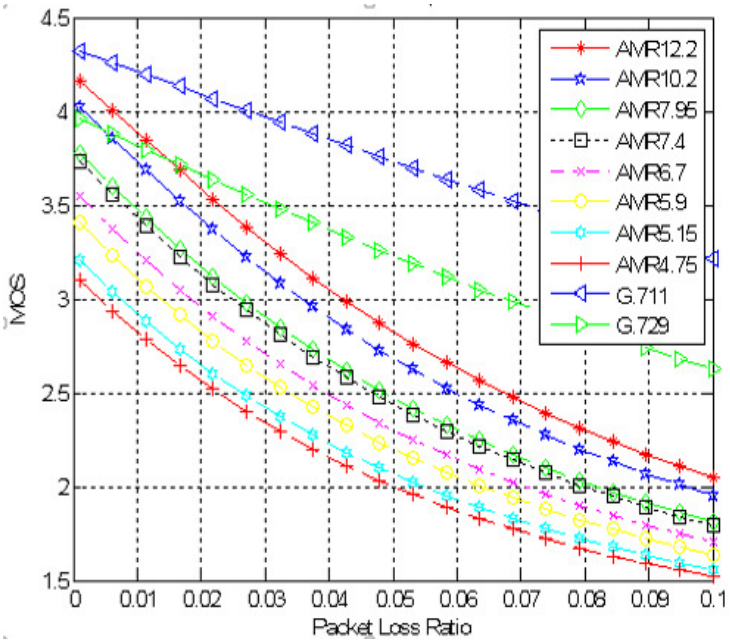

Figure 3.2 MOS as a function of packet loss ratio for different VoIP codec modes

In this dissertation, a simplified video MOS model [19], [68], [69] is used, where the distortion, as measured by the MSE (Mean Square Error), is assumed to be composed of two additive components, namely the source distortion D_S and the loss distortion D_L:

$$MSE = D_S + D_L = \eta \cdot R^{\xi} + \Gamma \cdot PEP \tag{3-7}$$

where η, ξ, and Γ are model parameters and PEP is the packet loss ratio. For different types of video sources, η, ξ, and Γ take different values. In this study, we assume $\eta = 1.76 \cdot 10^5$, $\xi = -0.658$, and $\Gamma = 1750$ as in [68]. The PSNR (Peak Signal-to-Noise Ratio) is a widely used objective measurement of video quality, and is related to the *MSE* by:

$$PSNR(\text{dB}) = 10 * \log_{10} \frac{255^2}{MSE} \tag{3-8}$$

A piecewise linear mapping [19] from the *PSNR* to *MOS* is:

$$MOS = \begin{cases} 1, & \text{when } PSNR < 20 \\ 1 + \frac{3.5}{20}(PSNR - 20), & \text{when } PSNR \epsilon\, [20, 40] \\ 4.5, & \text{when } PSNR > 40 \end{cases} \tag{3-9}$$

Figure 3.3 illustrates the relationship between the MOS value and the data rate and packet loss ratio. From Figure 3.3 and (3-7) to (3-9), the higher the data rate, or the lower the packet loss ratio, the higher the MOS value.

3.2.3 FTP Applications

As described in [19], the MOS of FTP users can be estimated based on the current rate R offered to the user by the system and packet error probability (PEP):

$$MOS = a * log_{10}(b * R * (1 - PEP)) \tag{3-10}$$

where a and b are determined from the maximum and minimum user perceived quality. The parameter R is the data rate a user is subscribed. Then, in case of no packet loss, user satisfaction on the MOS scale should be maximum, that is 4.5. On the other hand, a minimum transmission rate can be defined and assigned a MOS value of 1.0.

The PEP can be regarded approximately equivalent to one minus the ratio of the FTP file download time without packet loss to the actual FTP file download time. Therefore, (3-10) becomes:

$$MOS = a * log_{10}(b * R * \frac{\text{the FTP file download time without packet loss}}{\text{the actual FTP file download time}}) \tag{3-11}$$

Since $R *$ the FTP file download time without packet loss $\approx$ the FTP file siz, (3-11) becomes:

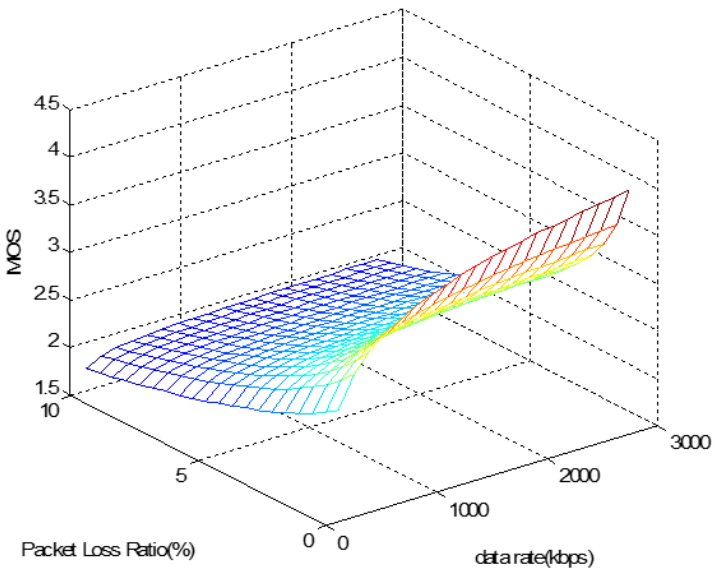

Figure 3.3 MOS as a function of data rate and packet loss ratio for video applications

$$MOS = a * log_{10}(b * \frac{\text{the FTP file size}}{\text{the actual FTP file download time}}) \quad (3\text{-}12)$$

Figure 3.4 shows the FTP MOS as a function of the actual FTP file download time assuming the FTP file size is 5 Mbytes. When a user's actual FTP file download time is 20 seconds, this user's MOS is set to 4.5. When a user's actual FTP file download time is 100 seconds, this user's MOS is set to 1.0. Based upon these two assumptions, the coefficients a and b in (3-9) can be determined as 5.0074 and $3.9595e^{-6}$, respectively.

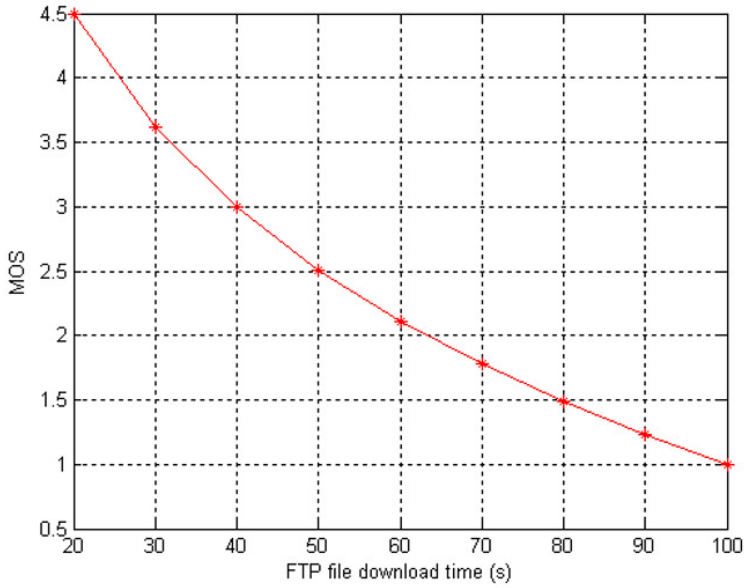

Figure 3.4 FTP MOS as a function of file download time given a file size of 5 Mbytes

3.3 User-Specific QoS Formulas for VoIP

It is assumed that different people have similar sensitivity to the end-to-end delay for VoIP applications, so that only user specific sensitivity to packet loss is studied. Therefore, User-specific QoS requirements are characterized by their different sensitivities to packet losses among other factors. To reflect this sensitivity, we introduce a user-specific packet loss sensitivity factor, α, to the ITU-T G.107 E-Model equation [70], [71]:

$$R = R_0 - I_d - \alpha \cdot I_{eff} \quad (3\text{-}13)$$

Without loss of generality and also for simplicity of illustration, the packet loss sensitivity factor α takes values from the following set $\{0.8, 1.0, 1.2\}$. Correspondingly users are classified into 3

categories: users with higher sensitivity factors (1.2), users with normal sensitivity factors (1.0), and users with lower sensitivity factors (0.8). The higher the value of the sensitivity factor α, the more the user is sensitive to packet loss.

Figure 3.5 shows the MOS as a function of different AMR codec data rates for different sensitivity factors α, given an end-to-end delay of 150 ms [14] and packet loss ratio of 0.05. For a comparison between AMR12.2K mode and $\alpha = 1.0$ with AMR10.2K/7.95K mode and $\alpha = 0.8$, we find the latter case may, under certain conditions, have a higher MOS than the former one. If the scheduler can know, or adaptively learn, each user's specific sensitivity factor, it can degrade the AMR mode for users with a lower sensitivity factor, while maintaining a comparable MOS as that of users with higher AMR mode but a normal sensitivity factor. With this approach, more users can be supported, thus achieving the target of improving system capacity.

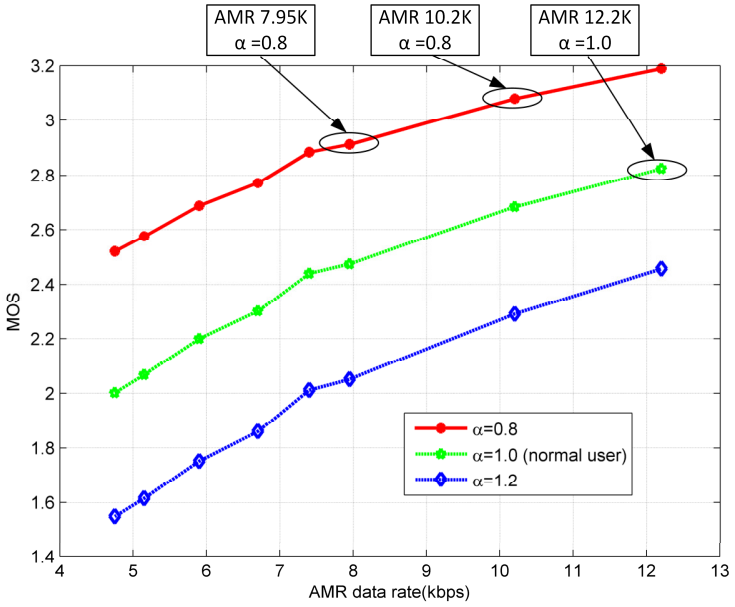

Figure 3.5 VoIP MOS as a function of AMR data rate given a packet loss ratio of 0.05 and end-to-end delay of 150 ms

3.4 User-Specific QoS Formulas for Video

To reflect users' different sensitivity to the data rate, a user-specific sensitivity factor γ to the data rate was introduced in (3-7), which becomes [71], [72]:

$$MSE = D_S + D_L = \gamma \cdot \eta \cdot R^{\xi} + \Gamma \cdot PEP \tag{3-14}$$

Without loss of generality and also for simplicity of illustration, the data rate sensitivity factor γ takes values from the following set $\{0.8, 1.0\}$. Correspondingly users are classified into 2 categories: users with normal sensitivity factors (1.0) and users with lower sensitivity factors (0.8). When γ takes the value 1.0, it is a normal user. When γ takes the value less than 1.0, it is less sensitive to the data rate compared with a normal user. The higher the value of the sensitivity factor γ, the user is more sensitive to the data rate.

For simplicity, in order to illustrate the main idea of the algorithm, 10 levels of data rate are used in this dissertation, which loosely correspond to the application requirements. For Level I, where I ∈ {1, 2,,,10}, the corresponding data rate is 135 * 128 pixels * (11-I) frames/s * 8bytes/pixel with data rates ranging from 138.2 Kbps to 1.382 Mbps.

Figure 3.6 shows the MOS as a function of data rate for different sensitivity factors γ under a given packet loss ratio of 0.001. From the figure, we observe that users' MOS differs even with the same data rate but with different sensitivity factors.

3.5 User-Specific QoS Formulas for FTP

To reflect user sensitivity to the FTP file download time, a UE specific sensitivity factor λ is added to (3-12), and (3-12) becomes:

$$MOS = a * log_{10}(b * \frac{\text{the FTP file size}}{(\text{the actual FTP file download time}) * \lambda}) \tag{3-15}$$

Without loss of generality and also for simplicity of illustration, the data rate sensitivity factor λ takes values from the following set $\{1.0, 1.2\}$. Correspondingly users are classified into 2 categories: users with normal sensitivity factors (1.0) and users with higher sensitivity factors (1.2). The higher the value of the sensitivity factor λ, the user is more sensitive to the FTP download time. When λ takes the value 1.0, it is a normal user. When λ takes the value greater than 1.0, the user is more sensitive to the FTP download time compared with a normal user. Figure 3.7 shows the MOS as a function of FTP file download time given a file size of 5 Mbytes for different sensitivity factors. From Figure 3.7 we can see that users with a

higher sensitivity factor may even have a lower MOS than users with a normal sensitivity factor even if they have the same file download time. So, this category of user with a higher sensitivity factor need to be treated differently in order to achieve a similar MOS as users with a normal sensitivity factor.

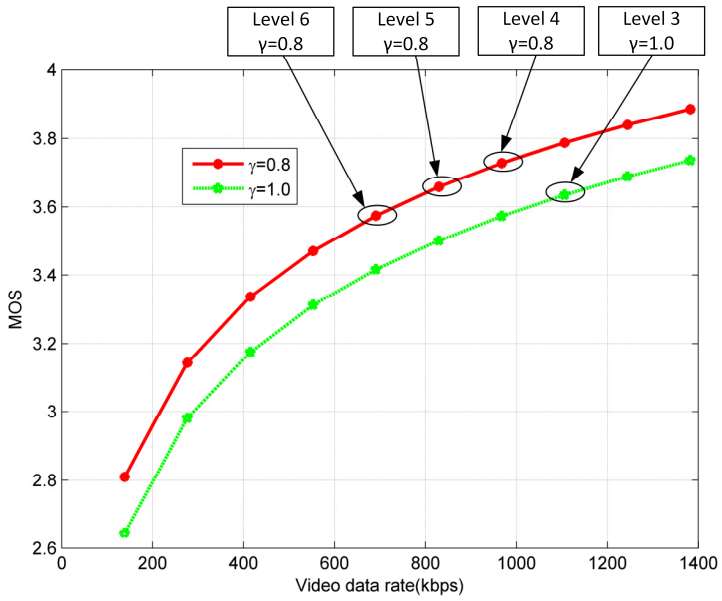

Figure 3.6 Video MOS as a function of data rate for different sensitivity factors, γ, given a packet loss ratio of 0.001

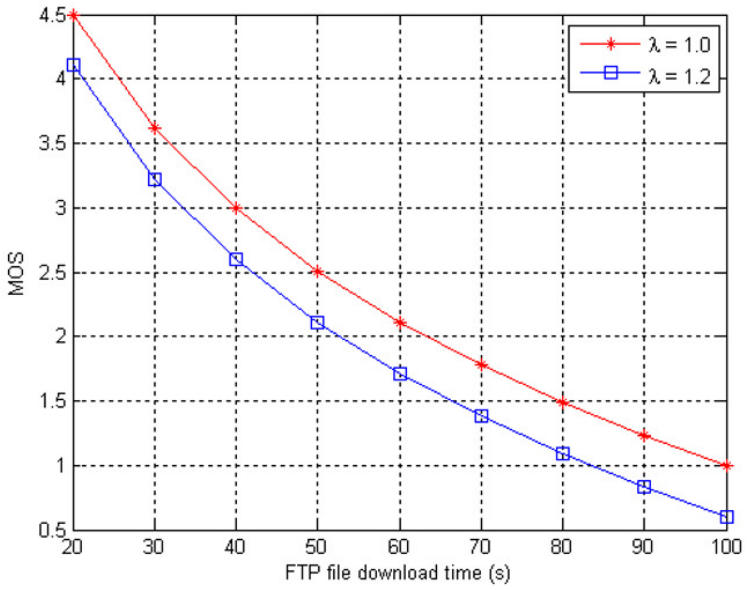

Figure 3.7 FTP MOS as a function of file download time given a file size of 5 Mbytes for different sensitivity factors λ

3.6 Alternatives for Improving System Capacity

There are two alternatives to improve system capacity considering user-specific QoS requirements. The first method is to optimize the spectral allocation of physical resources by utilizing the user-specific MCS selection method, where users with lower sensitivity factors are allocated less resources, and users with higher lower sensitivity factors are allocated more resources. The other method is to optimize the user's source data rate by dynamically adapting the data rate to the user-specific QoS requirements and network conditions. The first method is illustrated in this section, and the second method will be described in detail in the following chapters.

3.6.1 User-Specific MCS Selection Method

Before proceeding to the major part of the proposed user-specific QoS aware scheduler in the next section and the following chapters, in this section, a user-specific MCS selection method is introduced.

When considering the user-specific QoS requirements, it is straightforward and intuitive to considering allocating less wireless resources to users with lower sensitivity factors while maintaining the same source bit rate for these users, since these users can tolerate larger end-to-end delay and/or higher packet loss ratio. So in the user-specific MCS selection method, a higher MCS index, rather than a normally selected MCS index, can be designated for this class of users, so that fewer Physical Resource Blocks (PRBs)[3] may be assigned, and some wireless resources can be saved for use by other users. However, the user-specific MCS selection method has the following drawbacks:

- A higher MCS index doesn't necessarily mean fewer RBs (resource blocks) [73], [74], as seen in Figure 3.8. In this figure, the row stands for the required number of PRBs, while the column stands for the selected MCS index. Assuming the number of TB (Transport Block) bits to be transmitted in one TTI (Transmission Time Interval) is 515 bits, the normally selected MCS index and number of PRBs will be 9 and 4 respectively. If a higher MCS index (i.e., normally selected MCS index plus 1) is used, the selected MCS index and number of PRBs will be 10

[3] A PRB is the smallest resource entity assigned by the LTE resource scheduler. A PRB is defined as consisting of 12 consecutive subcarriers for one slot (0.5 msec) in duration.

and 4 respectively. That means there is not any savings on the number of PRBs, and thus the bit rate is not reduced.

- A higher MCS index than the normally selected MCS index, especially for high MCS indices, will result in a much higher BLER, thus many more frequent retransmissions, therefore even higher PRB consumptions, larger end-to-end delay and/or higher packet loss ratio, thus degraded voice quality (Figure 3.9 illustrates this case). Assuming a target BLER of 10% [73], [74] for the MCS index to be selected, when the MCS index 4 is used, the BLER is 10%. But when the MCS index 5 is used, the BLER is close to 50%.

A system simulation verified the above. The setup was 3 macro cells with 24 VoIP users in each macro cell, and a 10 MHz system bandwidth for both the normal and the user-specific MCS selection methods. The results showed that similar PDSCH loads (~4%) are incurred, which verified the inefficiency of this user-specific MCS selection method and motivated us to seek other user-specific QoS aware scheduling methods to improve system capacity as described in the following section and next chapters.

I_{TBS}	N_{PRB}									
	1	2	3	4	5	6	7	8	9	10
0	16	32	56	88	120	152	176	208	224	256
1	24	56	88	144	176	208	224	256	328	344
2	32	72	144	176	208	256	296	328	376	424
3	40	104	176	208	256	328	392	440	504	568
4	56	120	208	256	328	408	488	552	632	696
5	72	144	224	328	424	504	600	680	776	872
6	328	176	256	392	504	600	712	808	936	1032
7	104	224	328	472	584	712	840	968	1096	1224
8	120	256	392	536	680	808	968	1096	1256	1384
9	136	296	456	616	776	936	1096	1256	1416	1544
10	144	328	504	680	872	1032	1224	1384	1544	1736
11	176	376	584	776	1000	1192	1384	1608	1800	2024
12	208	440	680	904	1128	1352	1608	1800	2024	2280
13	224	488	744	1000	1256	1544	1800	2024	2280	2536
14	256	552	840	1128	1416	1736	1992	2280	2600	2856
15	280	600	904	1224	1544	1800	2152	2472	2728	3112

Figure 3.8 User-specific MCS selection method

3.6.2 User-Specific QoS Aware Scheduling Method

The user-specific QoS aware scheduling method optimizes a user's source data rate by dynamically adapting the data rate to the user-specific QoS requirements and network conditions. It consists of two components: a data rate adaptation algorithms and a MAC scheduling algorithm, which are described in the next chapter.

3.7 Concluding Remarks

In this chapter, quantitative metrics of the MOS for VoIP, video, and FTP applications were presented. User-specific QoS formulas for VoIP, video, and FTP applications were proposed that consider users specific QoS requirements. Two methods to improve system capacity were described, where either the MCS index selection (user-specific MCS selection method) or the source data rate (user-specific QoS aware scheduling method) are optimized by considering user-specific QoS requirements. It was shown that the user-specific MCS method is not very efficient in improving system capacity, which motivates us to investigate user-specific QoS aware scheduling methods that will be addressed in the following chapters.

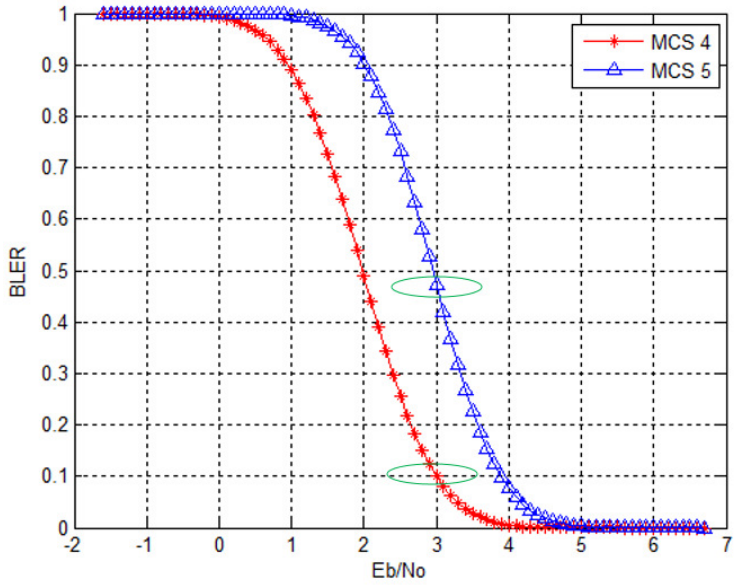

Figure 3.9 BLER comparison for different MCS values

CHAPTER 4. USER-SPECIFIC QOS AWARE SCHEDULER

4.1 Introduction

In this chapter[4], five user-specific QoS aware schedulers (USQA-M1, USQA-M2, USQA-MC, USQA-C, and USQA-F) that consider the user-specific QoS requirements in the allocation of spectral resources are proposed. Schedulers that dynamically adapt scheduling to the user-specific QoS requirements to improve quality as measured by the MOS (i.e., user experience), the system capacity (i.e., maximum number of supportable users), or both quality and capacity are introduced. Depending on whether the MOS, the system capacity, or both the MOS and system capacity are the optimization target, it is shown that appreciable MOS, system capacity improvement, or both appreciable MOS and system capacity improvement can be achieved when user-specific QoS requirements are utilized.

Compared with homogenous networks, the different cell deployment and different intra-cell and inter-cell interference distributions in heterogeneous networks may impact the system performance gain of the user-specific QoS aware schedulers. In this chapter, the system performance gain of the user-specific QoS aware scheduler (the USQA-C scheduler) is also analyzed and evaluated in heterogeneous networks. As the metric to measure the system capacity gain, the throughput improvement of FTP applications benefits from the rate adaptation and MAC scheduling algorithms for video applications that incorporate user-specific QoS requirements. These improvements are achieved while user satisfaction for video applications, as measured by the MOS, is maintained at an acceptable level.

Furthermore, user-specific QoS frequency QoS study is also addressed. It is found, when targeted to maximize spectrum utilization and combined with VoIP codecs matched to the auditory characteristics of users, higher system capacity as well as a comparable MOS level may be achieved.

[4] Portions of this chapter were previously published in [70]. Permission is included in Appendix A.

The rest of the chapter is organized as follows, section 4.2 gives the motivation for the MOS and data rate optimization. Several classical LTE MAC schedulers are provided in section 4.3. We present the currently existing rate adaptation mechanism of LTE systems in section 4.4. The overview of user-specific QoS-aware schedulers is given in section 4.5. The user-specific QoS-aware rate adaptation algorithms are described in section 4.6. The user-specific QoS-aware MAC scheduling algorithms are proposed in section 4.7. Section 4.8 presents the user-specific QoS frequency sensitivity QoS study. The USQA-C Scheduler in Heterogeneous Networks is presented in section 4.9. We present the simulation results in section 4.10. Finally in section 4.11, we present the concluding remarks of this chapter.

4.2 Motivation

4.2.1 Motivation for VoIP Optimization

Figure 4.1 shows the decreased or increased MOS percentage due to the different sensitivity factors α for different users. The MOS of VoIP users with α of 1.2 is decreased by ~15%, whereas the MOS of VoIP users with α of 0.8 is increased by ~15%, when a packet loss ratio of 5% and end-to-end delay of 150 ms are assumed. As the packet loss ratio increases, the MOS will decrease or increase even more depending on the user's sensitivity factor α.

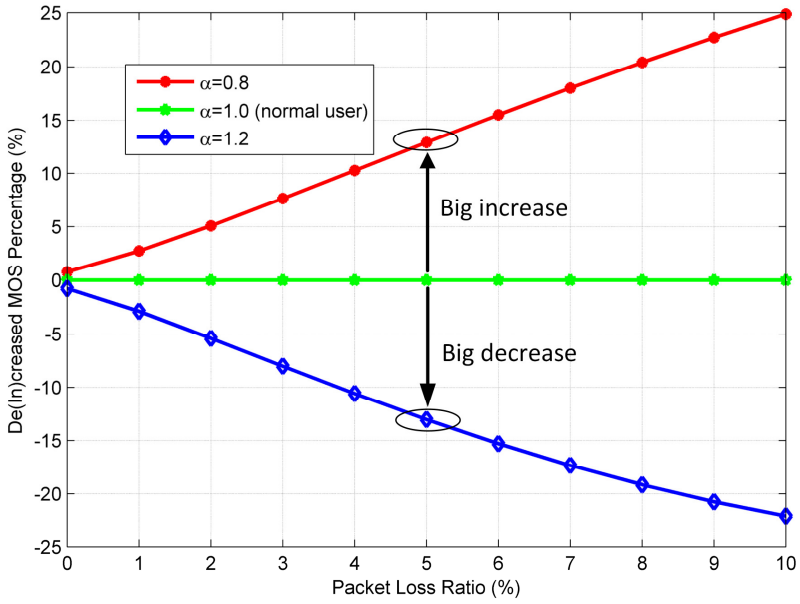

Figure 4.1 Decreased/Increased MOS as a function of packet loss ratio given an end-to-end delay of 150 ms and vocoder AMR12.2K

Therefore, the MOS of VoIP users with α of 1.2 needs to be improved to the corresponding MOS value with α of 1.0 and given a higher scheduling priority in the MAC Scheduler, whereas the MOS of VoIP users with α of 0.8 can be decreased to the MOS value of α of 1.0, as depicted in Figure 4.1. There are three approaches to deal with the MOS of VoIP users with α of 0.8:

- Approach I:

  This category of users can be deprioritized, that is, given a lower scheduling priority in the MAC scheduler. The MOS optimization-only scheme I in section 4.5 uses this method to optimize the MOS.

- Approach II:

  This category of users has the same scheduling as those normal users. Only the MOS of users with sensitivity factors α of 1.2 is optimized. The MOS optimization-only scheme II in section 4.5 uses this method to optimize the MOS.

- Approach III:

  The data rate (i.e., AMR mode) of this category of users can be degraded, so that a higher system capacity can be achieved. However, these users are scheduled as normal users in the MAC scheduler. The MOS optimization-plus-capacity improvement scheme and capacity improvement scheme in section 4.5 use this method to optimize the MOS.

4.2.2 Motivation for Video Data Rate Optimization

Similar to Chapter 3, 10 levels of data rate are used in this dissertation, which loosely correspond to the application requirements. For Level I, where $I \in \{1, 2, \ldots, 10\}$, the corresponding data rate is 135 * 128 pixels * (11-I) frames/s * 8bytes/pixel with data rates ranging from 138.2 Kbps to 1.382 Mbps.

An important observation that can be made from Figure 3.6 in Chapter 3 is that a user with a lower sensitivity factor of γ and a lower data rate may achieve a higher MOS value than that of users with a normal sensitivity factor and a higher data rate. If the user-specific QoS aware scheduler knows and

makes use of this user-specific sensitivity factor information to optimize the scheduling, it can decrease the data rate for users with a lower sensitivity factor to support more users with an acceptable MOS value.

4.3 LTE MAC Scheduler Overview

The MAC Scheduler is a key component of the eNodeB and is illustrated in Figure 4.2. The function of the scheduler is to facilitate the allocation of the available spectral resources (e.g., time and frequency resources), while striving to satisfy the QoS requirements for all the users. The MAC scheduler interfaces with the HARQ (Hybrid Automatic Retransmission Request) and Link Adaptation modules to acquire the related HARQ information and selected MCS value respectively. It also requires other inputs such as QoS related parameters and buffer status.

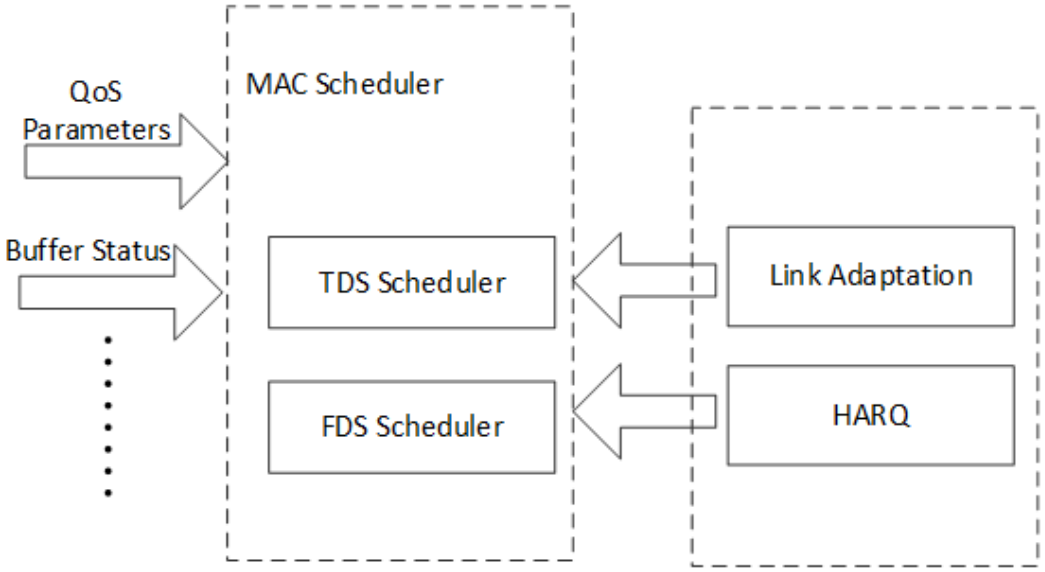

Figure 4.2 LTE MAC scheduler structure

The packet scheduling comprises two scheduling components. They are done sequentially in each scheduling time unit, known as Transmission Time Interval (TTI) in LTE (TTI = 1ms). The first component is the Time Domain Scheduler (TDS) and the second is the Frequency Domain Scheduler (FDS). Such a split is driven simply by the consideration of lower complexity and independent configurations for both domains. The objective of the Time Domain Scheduler is to choose a subset of all users requesting frequency resources, while the objective of Frequency Domain Scheduler is to allocate

physical resources for the candidate users provided by the time domain scheduler. Several classical MAC scheduling algorithms exist both in time and frequency domains [23]–[25]:

- Round-Robin scheduling algorithm

  Users are served in a Round-Robin way so that each user is served fairly, but at the expense of system throughput and spectral efficiency.

- Blind Equal Throughput (BET) scheduling algorithm

  The BET algorithm uses the reciprocal of the past average throughput achieved by each user as the metric. Throughput fairness can be achieved, but at the expense of system throughput and spectral efficiency.

- Maximum C/I scheduling algorithm

  Users with the maximum C/I (Carrier-to-Interference power ratio) are served first. This kind of scheduling aims to achieve maximum benefits in terms of system throughput and spectral efficiency but comes at the expense of fairness.

- Proportional-fair (PF) scheduling algorithm

  PF scheduling algorithm aims to trade off the system throughput for the users' fairness. The PF priority metric is calculated by dividing the predicted user's throughput, which is the instantaneous supportable data rate, by the estimation of the user's past average throughput.

4.4 Rate Adaptation for LTE

It is beneficial (from the OpEx [Operational Expenditure] and CapEx [Capital Expenditure] points of view) for operators to be able to control the data rate based on load criteria. At peak hour there could be a desire to trade off some quality for additional capacity [75], [76].

Adaptive mechanisms that act upon measured or signaled changes in the transport channel characteristics may be used in a conservative manner. Examples of measured changes in transport characteristics are variations in Packet Loss Rate (PLR) and delay jitter. An example of signaled changes in transport characteristics is ECN (Explicit Congestion Notification) Congestion Experienced (ECN-CE)

marking in IP packet headers. A conservative use of adaptation is characterized by a fast response to degrading conditions, and a slower, careful upwards adaptation intended to return the session media settings to the original default state of the session. The long-term goal of any adaptive mechanism is assumed to be a restoration of the session quality to the originally negotiated quality. The short-term goal is to maximize the session quality given the current transport characteristics, even if that means than the adapted state of the session will give a lower session quality compared to the session default state if transported on an undisturbed channel [76].

The E-UTRAN/UTRAN and the UE support ECN in the RFC 3168 [77]. The IP level ECN scheme enables the E-UTRAN/UTRAN to trigger a rate adaptation scheme at the application / service / transport layer. To make sufficient time available for end-to-end codec rate adaptation the E-UTRAN/UTRAN should attempt to not drop any packets on a bearer for a default grace period of at least 500 ms after it has indicated congestion with ECN on the bearer for packets within the packet delay budget. During this ECN grace period the E-UTRAN/UTRAN should also attempt to meet the QCI characteristics / QoS class associated with the bearer [76], [78].

The rate adaptation mechanism relies on existing end-to-end schemes for rate adaptation to dynamically adapt an individual real-time media component to changing conditions in the network. The receiving side can use RTCP (Real-time Transport Control Protocol) or RTP to explicitly control the data rate at the sending side, for example, CMR (Codec Mode Request) command for VoIP applications and TMMBR (Temporary Maximum Media Bit-rate Request) command for video applications [76], [79], [80].

4.5 User-Specific QoS Aware Scheduler

In the following text, the term USQA scheduler is shorthand for User-Specific QoS Aware Scheduler. There are five USQA schemes presented, with correspondingly five USQA schedulers. Depending upon whether improving the MOS or system capacity or both the MOS and system capacity is the goal, the following user-specific QoS aware schemes for the USQA schedulers are proposed. The novelty of the proposed user-specific QoS aware schedulers is the incorporation of user-specific QoS

requirements into the scheduling and personalized individual user's scheduling that uses this user-specific QoS information to improve system performance. The proposed scheduling algorithms are composed of two parts, the source (e.g., AMR, video) data rate adaptation algorithm and the MAC resource scheduling algorithms. The rationale of the proposed algorithms to improve the system capacity is by trading off the spectral resource allocations for the user-specific QoS requirements, while still maintaining acceptable levels of MOS. To be more specific, the system capacity can be improved through dynamically adapting the data rate of video users by considering user-specific QoS requirements, and allocating the saved spectral resource resulting from the degraded data rate of video users to other users as described in section 4.6. The rationale of the proposed algorithms to improve the MOS is through the MAC scheduling differentiation between users with different QoS requirements as described in section 4.7.

4.5.1 MOS Optimization only Scheme I

The MOS of users with $\alpha = 1.2$ will be increased, that is, given a higher scheduling priority, whereas users with $\alpha = 1.0$ have a normal scheduling priority, and the MOS of users with $\alpha = 0.8$ are decreased a little and given a lower scheduling priority, as described in section 4.7. This scheme is denoted as the USQA-M1 scheme. It is used in heavily loaded scenarios.

4.5.2 MOS Optimization only Scheme II

The MOS of users with $\alpha = 1.2$ will be increased, given a higher scheduling priority, whereas users with $\alpha = 1.0$ and $\alpha = 0.8$ have a normal scheduling priority, as described in section 4.7. This scheme is denoted as the USQA-M2 scheme. It is used in lightly loaded scenarios.

4.5.3 MOS Optimization plus Capacity Improvement Scheme

The MOS of users with $\alpha = 1.2$ will be increased, that is, given a higher scheduling priority, whereas users with $\alpha = 1.0$ or $\gamma = 1.0$ have a normal scheduling priority, as described in section 4.7, and users with $\alpha = 0.8$ or $\gamma = 0.8$ are used to improve the system capacity by degrading their AMR codec modes or video data rates as described in section 4.6.

This scheme is denoted as the USQA-MC scheme. It is used in heavily loaded scenarios to be capable of supporting more users in the system.

4.5.4 Capacity Improvement Scheme

This scheme is denoted as the USQA-C scheme, a variant of the USQA-MC scheme where all users are assumed to be characterized as $\alpha = 0.8$ and 1.0 or $\gamma = 0.8$ and 1.0. So the scheduling of these users with lower sensitivity factors is used to improve the capacity by degrading their AMR codec modes or video data rates as described in section 4.6.

The USQA-C scheme is used in heavily loaded scenarios to be capable of supporting more users in the system.

4.5.5 MOS Improvement Scheme for FTP Users

The MOS of users with $\lambda = 1.2$ will be increased, that is, given a higher scheduling priority, whereas users with $\lambda = 1.0$ have a normal scheduling priority. This scheme is denoted as the USQA-F scheme.

4.5.6 USQA Scheduler Summary

The USQA Schedulers are summarized as follows and in Table 4-1:

Table 4-1 USQA scheduler summary

Scheduler	Features	Descriptions
USQA-M1	VoIP MOS Optimization only	Improve the MOS of VoIP users with α of 1.2 Decrease the MOS of VoIP users with α of 0.8
USQA-M2	VoIP MOS Optimization only	Improve the MOS of VoIP users with α of 1.2 Maintain the MOS of VoIP users with α of 0.8
USQA-MC	VoIP MOS and capacity improvement	Improve the MOS of VoIP users with α of 1.2 Degrade the data rate of VoIP users with α of 0.8 Degrade the data rate of video users with γ of 0.8
USQA-C	Capacity improvement only	Degrade the data rate of VoIP users with α of 0.8 Degrade the data rate of video users with $\gamma$of 0.8
USQA-F	FTP MOS improvement only	Improve the MOS of FTP users with λ of 1.2

- USQA Schedulers: USQA-M1, USQA-M2, USQA-MC, USQA-C and USQA-F.
- USQA-M1, USQA-M2 optimizes VoIP MOS only.

- USQA-MC optimizes both VoIP MOS and Capacity.
- USQA-C optimizes capacity only.
- USQA-F optimizes MOS for FTP users only.
- The symbols used in the USQA schedulers denote: M for MOS, C for Capacity, MC for MOS and Capacity, and F for FTP.

4.6 User-Specific QoS Aware Rate Adaptation Algorithms

4.6.1 VoIP Rate Adaptation Algorithm

In order to illustrate the main idea of user-specific QoS capacity improvement scheduling for VoIP users, in this algorithm, we only consider three AMR modes (i.e., AMR 12.2K, 10.2K, and 7.95K), and the extension to other AMR modes (bit rates) is straightforward. The work flow of the AMR mode adaption is shown in Figure 4.3 [71]. The algorithm starts with the AMR12.2K mode. The thresholds to degrade the AMR mode can be configured to control the desired MOS levels. In this dissertation, they are set to 0.02, that is, the AMR mode will be degraded if the MOS is decreased by less than 0.02, compared with that of the MOS value for the non-degraded AMR mode with $\alpha = 1.0$. The input to the AMR mode adaption algorithm is the packet loss ratio in the receiver UE, while assuming an average end-to-end delay of 150 ms.

4.6.2 Video Rate Adaptation Algorithm

Similarly the work flow of the video data rate adaptation is shown in Figure 4.4 [71], [72]. For simplicity of illustration, four levels of data rate are assumed in the video data rate adaptation. The algorithm starts with the data rate of level 3. The video data rate level to be selected depends upon the respectively calculated MOS for each level of data rate. Similar to VoIP users, the thresholds to degrade the video data rate can be configured to control the desired MOS levels. In this dissertation, they are set to 0.01, which means that the data rate will be degraded when the MOS is not decreased by more than 0.01 compared with that of the MOS value for the non-degraded rate with $\gamma = 1.0$. The input to the video data rate adaption algorithm is the packet loss ratio in the receiver UE.

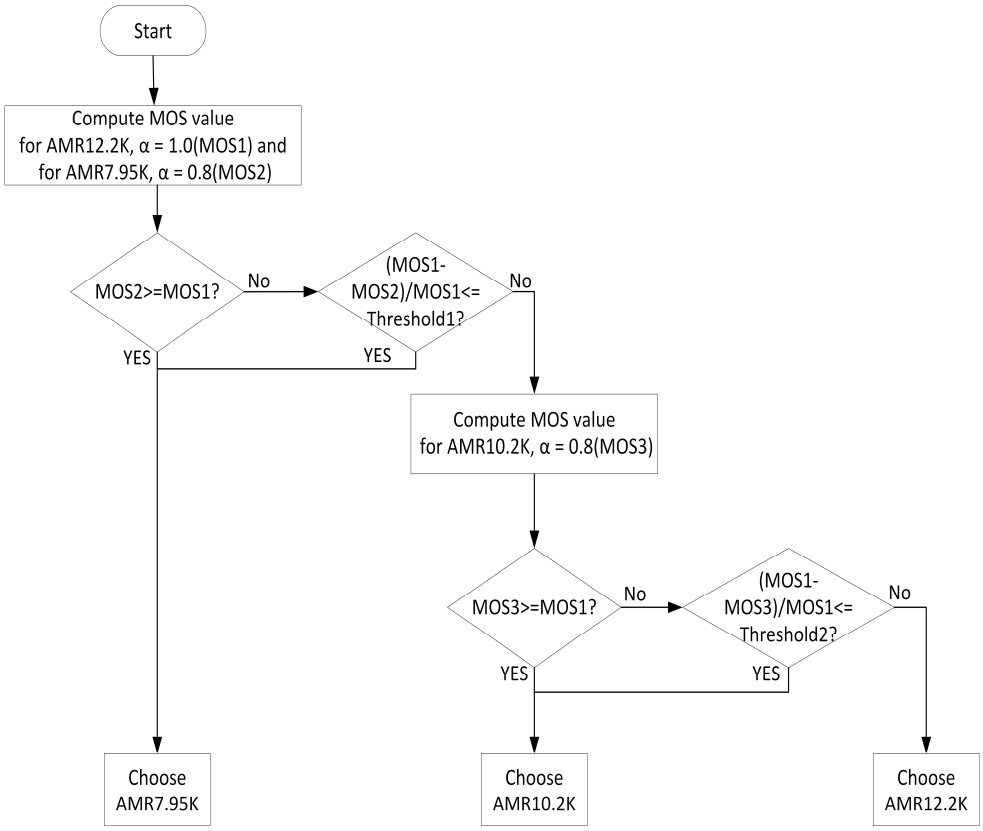

Figure 4.3 AMR mode adaptation work flow

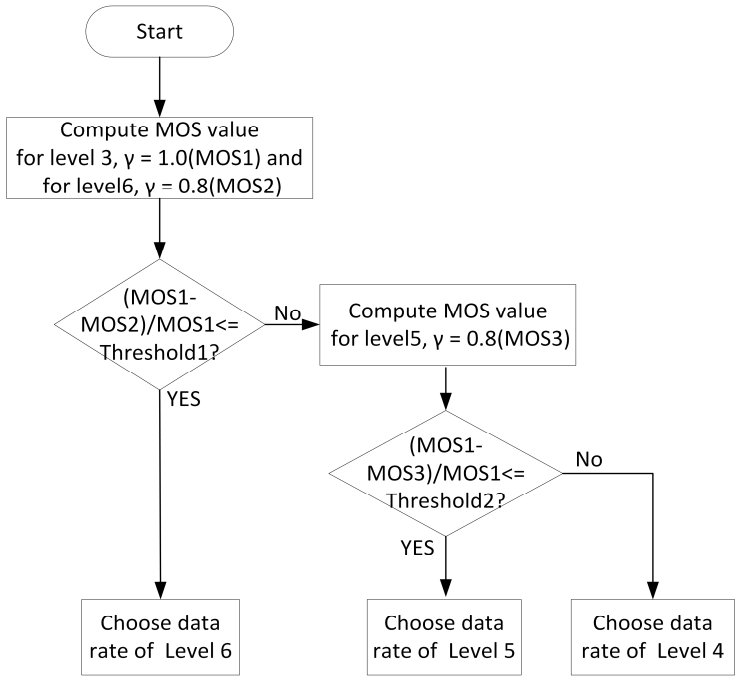

Figure 4.4 Video data rate adaptation work flow

4.7 User-Specific QoS Aware MAC Scheduling Algorithms

4.7.1 Time Domain Scheduler

The following is for GBR services.

GBR services refer to VoIP and video applications that are sensitive to the delay.Users with higher metrics can receive higher scheduling priority in the time domain. The metric for user *k* is defined as:

$$M_k = W_{pri} * TW_k * DOP \tag{4-1}$$

where W_{pri} is the priority factor for different GBR applications ($W_{pri} = 1.0$ for VoIP users, and $W_{pri} = 0.8$ for video users). The parameter TW_k is the user-specific time domain weight for user *k*. For the baseline scheduler, $TW_k = 0.8$ for all VoIP users, and $TW_k = 1.0$ for all video users, which means users are not differentiated by their user-specific QoS requirements. For the USQA-M1 scheduler, $TW_k = 1.0$ for VoIP users with $\alpha = 1.2$, $TW_k = 0.7$ for VoIP users with $\alpha = 0.8$, whereas for the USQA-M2 and USQA-MC schedulers, $TW_k = 1.0$ for VoIP users with $\alpha = 1.2$ and $TW_k = 0.8$ for other VoIP users. In (4-1) DOP is the packet delay in milliseconds in the MAC buffer. All VoIP users with M_k greater than 20 ms will be selected as the candidate users if the resource is available. The parameter TW_k is configurable based upon different application scenarios.

The following is for Non-GBR services.

Non-GBR services in this chapter refer to the FTP application that is a best effort application and not sensitivity to delay.

FTP Users are served in a Round-Robin way so that each user is served fairly. In this chapter, in each TTI, a maximum of 3 users are scheduled.

If a given user is more sensitive to the FTP download time, then this user will be scheduled in each TTI with a predefined probability P1 (e.g., P1 = 33.33% in this chapter) instead of normally waiting in queues to be served in a round-robin way.

FTP applications as Non-GBR services have lower scheduling priority than VoIP and video applications as GBR services in the time domain. That is, only if there are remaining resources after the GBR users are served, the Non-GBR services will be scheduled.

4.7.2 Frequency Domain Scheduler

Each user also has a frequency domain metric for each sub-band and this is sorted for each sub-band among all the scheduled users. Each sub-band is first allocated to the user that has the highest metric, then to the user with the second and third highest metric, and so on until all the resources of this given sub-band are allocated, as illustrated in Figure 4.5. The metric for user k in each sub-band n is defined by:

$$M_{n,k} = N_k * (MCS_{n,k} - MCS_{wb,k} + FW_{n,k}) \quad (4\text{-}2)$$

where N_k is the number of PRBs (Physical Radio Blocks) allocated to user k. The parameters $MCS_{n,k}$ and $MCS_{wb,k}$ are the MCS index of user k in sub-band n and wideband respectively. The parameter $FW_{n,k}$ is the user-specific frequency domain weight for user k in sub-band n. For the baseline scheduler, $FW_{n,k} = 0$ for all users, which means users are not differentiated by their specific QoS requirements. The parameters are introduced in order to allocate high priority users their preferred sub-bands, instead of other sub-bands. For the USQA-M1/M2 and USQA-MC schedulers, $FW_{n,k} = 1$ and $FW_{n,k} = -1$ for VoIP users with $\alpha = 1.2$ in their best sub-band and other sub-bands respectively, and $FW_{n,k} = 0$ for other users including video users. The parameter $FW_{n,k}$ is configurable. In scenarios with video users where video users typically require more PRBs, $FW_{n,k} = 25$ and $FW_{n,k} = -25$ for VoIP users with $\alpha = 1.2$ in their best sub-band and other sub-bands respectively. For the USQA-F scheduler, $FW_{n,k} = 10$ and $FW_{n,k} = -10$ for FTP users with $\lambda = 1.2$ in their best sub-band and other sub-bands respectively with a predefined probability P2 (e.g., P2 = 50 % in this chapter).

4.8 User-Specific Frequency Sensitivity QoS Study

Another very promising area of research is a user-specific frequency sensitivity QoS study. For humans, the audible range of frequencies is usually between 20 Hz and 20 kHz. However, there is

considerable variation between individuals - especially at the high frequency end, which is affected by a gradual decline with age. Elderly people normally are less sensitive to high frequencies, while younger people are more sensitive to higher frequencies.

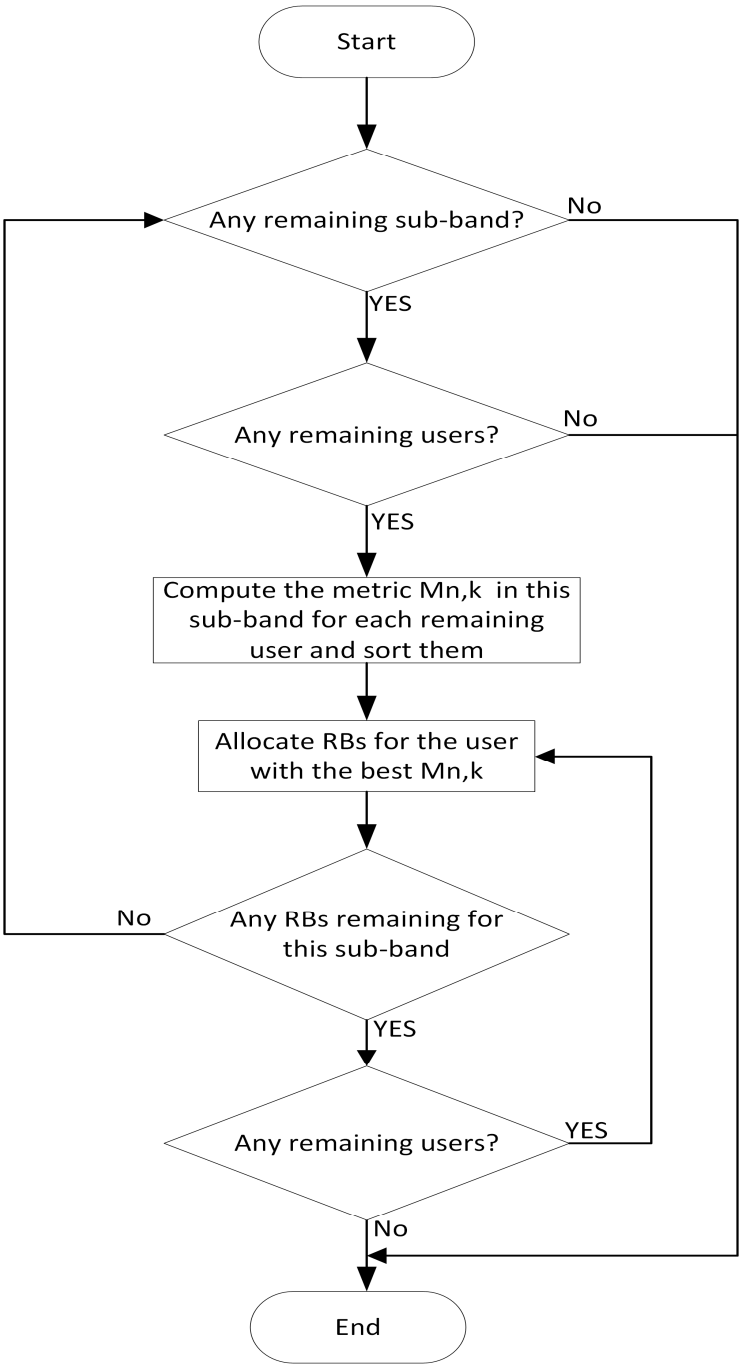

Figure 4.5 Frequency domain scheduler work flow

This difference in the sensitivity to higher frequencies can be utilized to further increase system capacity. A *frequency sensitivity factor* β is defined as the ratio of highest sensitive frequency of a given user to the standard sampling rate 8 KHz. If a user has a frequency sensitivity factor less than 1, the sampling rate can be reduced to 8β KHz, the data rate will be reduced, then the system capacity (i.e., maximum number of supportable users) will be increased correspondingly. An OPNET experiment in was setup to verify the capacity improvement. Details of the experiment are given in section 4.10.3.

4.9 USQA-C Scheduler in Heterogeneous Networks

Heterogeneous networks have been introduced in the LTE-Advanced standard. Driven by smart phones, tablets, and other data-hungry devices, data traffic demand in cellular networks is increasing exponentially. Since radio link performance is approaching theoretical limits with 3G/4G enhancements, the next performance leap in wireless networks will come from the network topology optimization as a means to provide higher network capacity and better coverage. A heterogeneous network uses a mixture of macrocells and small cells such as microcells, picocells, and femtocells, where small cells are overlaid in macro cells. These smallcells in heterogeneous networks can potentially improve system capacity and coverage by allowing future cellular systems to achieve higher data rates [81]–[85].

Compared with homogenous networks, the different cell deployment and different intra-cell and inter-cell interference distributions in heterogeneous networks may have a different impact on the system gain of the user-specific QoS aware scheduler. However, to date, no analysis and performance evaluation of user-specific QoS requirements and its applications in heterogeneous networks have been done.

Figure 4.6 describes the decrease in MOS percentage for different packet loss ratios calculated for a user with a lower sensitivity factor of 0.8 but a lower data rate of level 5, 6 and for a user with a normal sensitivity factor of 1.0 but a higher data rate of level 3, which corresponds to $\frac{MOS1-MOS3}{MOS1}$ and $\frac{MOS1-MOS2}{MOS1}$ respectively in the rate adaptation work flow in Figure 4.4 [72]. From Figure 4.6 it can be seen that, for a user with a lower sensitivity factor of 0.8 but a lower data rate of level 6 (red curve), the decreased MOS percentage decreases as the packet loss ratio increases. In order to maintain the target degraded MOS level

of 1%, the packet loss ratio needs to be greater than about 1%. That means, to maintain the desired degraded level of MOS, the higher packet loss ratio can result in a lower user data rate and the lower packet loss ratio can result in a higher user data rate. Therefore, in heterogeneous networks, if the SINR of video users can be improved, with a lower packet loss ratio, due to the deployment of LTE picocells and closer distances to the picocells, then the selected data rate of video users will be higher compared with homogeneous networks, as verified in the simulation results in section 4.10.4.

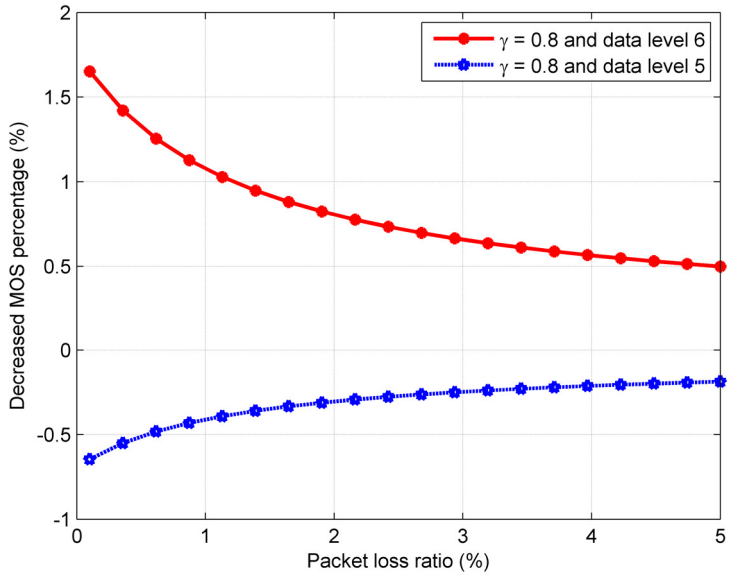

Figure 4.6 Decreased video MOS as a function of packet loss ratio for sensitivity factors γ of 0.8 and data level 5 and 6

4.10 Simulation Results

The system simulation was run using the OPNET 17.5 Modeler with the LTE modules [86]. The simulation was performed to evaluate the downlink scheduling, with an ideal uplink receiver.

The simulation environment is shown in Figure 4.7. The OPNET simulation project is composed of the following nodes:

- UE/eNodeB/EPC nodes

  The UE/eNodeB/EPC nodes represent the basic three network elements in LTE networks.

- LTE Attributes node

The LTE Attributes node is used to store PHY configurations and EPS bear definitions, which can be referenced by all LTE nodes in the network.

- Mobility Config node

  The Mobility Config node is used to define mobility profiles that individual nodes can reference to model mobility. It controls the movement of nodes based on the configured parameters.

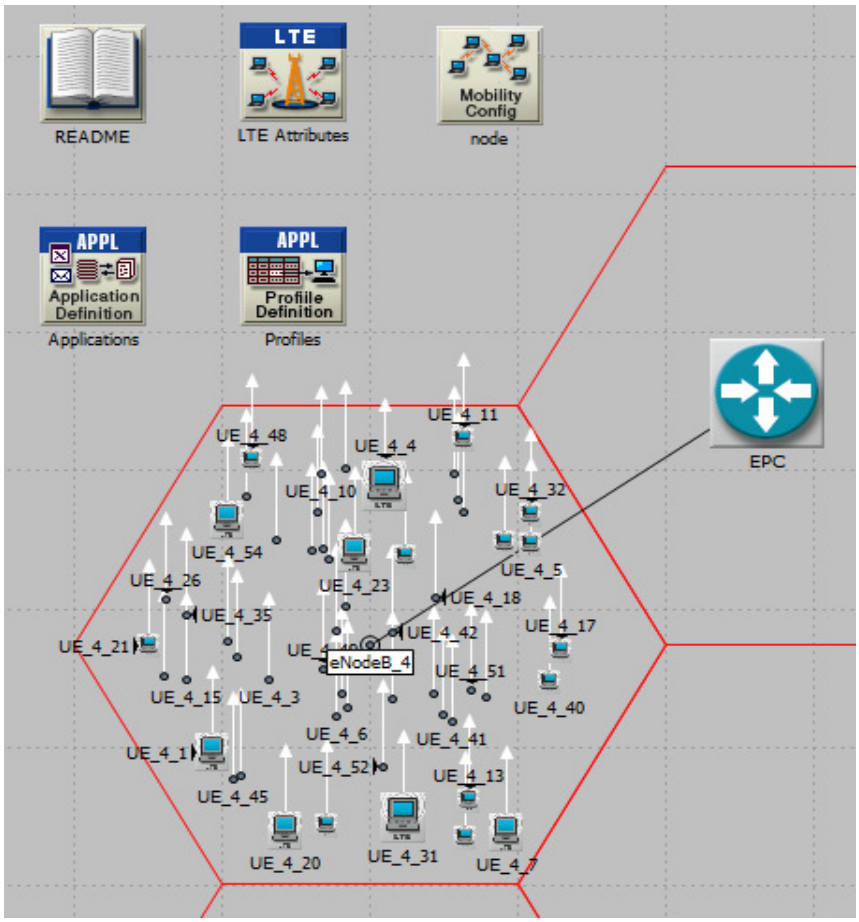

Figure 4.7 System environment setup

- Application Definition node

  The Application definition node is used to specify applications using available application types. We can specify a name and the corresponding description in the process of creating new applications. The specified application name will be used while creating user profiles on the Profile Definition node.

- Profile Definition node

  The Profile Definition node is used to create user profiles. These user profiles can then be specified on different nodes in the network to generate application layer traffic. The application defined in the Application Definition nodes is used by this object to configure profiles.

4.10.1 USQA-M1 & M2 & MC &C Schedulers

The system simulation configuration is partly based upon LTE macrocell system simulation baseline parameters [87] and Table 4-2.

Table 4-2 System simulation configuration for USQA schedulers

Parameter	Assumption
Cellular layout	1 macrocell
Cell radius	1kilometer
Path loss model	3GPP suburban macrocell
Mobility model	Random Way Point (30km/h)
Carrier frequency	Uplink:1920MHz Downlink:2110MHz
System bandwidth	10MHz
Channel model	ITU Vehicle A
Total BS TX power	40dBm
UE power	23dBm
VoIP codec modes	AMR12.2K, AMR10.2K, AMR7.95K
Video data rates	1105.920 Kbits/s (Level 3) 967.680 Kbits/s (Level 4) 829.440 Kbits/s (Level 5) 691.200 Kbits/s (Level 6)
Number of users	54 VoIP users w/o 12 video users
Scheduler	Dynamic scheduling USQA-M1/M2 scheduler, USQA-MC/C scheduler, and Baseline scheduler
Other assumptions	Ideal uplink receiver(no block error and packet loss)

1) Simulation Cases

Eight cases were simulated for user-specific QoS aware schedulers as described in Table 4-3. Cases 1-3 are used to evaluate the performance of the USQA-M1 and USQA-M2 schedulers, whereas cases 4-8 are used to evaluate the performance of the USQA-MC and USQA-C schedulers. In these cases, there are either 54 VoIP users having different sensitivity factors α, or both 54 VoIP users and 12 video users having different sensitivity factors α or γ.

Table 4-3 System simulation cases for USQA schedulers

Cases	Assumption	USQA scheduler
Case 1	54 VoIP users (18 users' α = 0.8, 18 users' α = 1.0, 18 users' α = 1.2)	USQA-M1 scheduler
Case 2	54 VoIP users plus 12 video users (18 VoIP users' α = 0.8, 18 VoIP users' α = 1.0, 18 VoIP users' α = 1.2, all video users' γ = 1.0)	USQA-M1 scheduler
Case 3	Same as Case 1	USQA-M2 scheduler
Case 4	Same as Case 1	USQA-MC scheduler
Case 5	Same as Case 2	USQA-MC scheduler
Case 6	54 VoIP users plus 12 video users (18 VoIP users' α = 0.8, 18 VoIP users' α = 1.0, 18 VoIP users' α = 1.2, 6 video users' γ = 1.0, 6 video users' γ = 0.8)	USQA-MC scheduler
Case 7	54 VoIP users with α = 0.8	USQA-C scheduler
Case 8	54 VoIP users with α = 0.8 plus 12 video users with γ = 0.8	USQA-C scheduler

To evaluate the impact of the ratio of users with lower sensitivity factors (i.e., 0.8) α or γ on the capacity improvement, 9 more cases are simulated as described in Table 4-4. Cases 1-3 are used to evaluate the VoIP performance for scenarios where all users are VoIP users, but differ in the ratio (i.e., 1/3, 2/3, 1) of users with lower sensitivity factors α of 0.8. Cases 4-6 are used to evaluate the VoIP performance in scenarios where normal video users exist, but VoIP users differ in the ratio (i.e., 1/3, 2/3, 1) of users with lower sensitivity factors α of 0.8. Cases 7-9 are used to

evaluate the video performance for scenarios where all VoIP users have sensitivity factors α of 0.8, but video users differ in the ratio (i.e., 1/2, 2/3, 1) of users with lower sensitivity factors γ of 0.8.

Table 4-4 System simulation cases for performance impact of the ratio of users with sensitivity factor of 0.8

Cases	Assumption	USQA scheduler
Case 1	54 VoIP users (18 users' α = 0.8, 18 users' α =1.0, 18 users' α = 1.2)	USQA-MC scheduler
Case 2	54 VoIP users (36 users' α = 0.8, 18 users' α = 1.2)	USQA-MC scheduler
Case 3	54 VoIP users (all users' α = 0.8)	USQA-C scheduler
Case 4	54 VoIP users plus 12 video users (18 users' α = 0.8, 18 users' α = 1.0, 18 users' α = 1.2, all video users' γ = 1.0)	USQA-MC scheduler
Case 5	54 VoIP users plus 12 video users (36 users' α = 0.8, 18 users' α = 1.2, all video users' γ = 1.0)	USQA-MC scheduler
Case 6	54 VoIP users plus 12 video users (all VoIP users' α = 0.8, all video users' γ = 1.0)	USQA-C scheduler
Case 7	54 VoIP users plus 12 video users (all VoIP users' α = 0.8, 6 video users' γ = 1.0, 6 video users' γ = 0.8)	USQA-MC scheduler
Case 8	54 VoIP users plus 12 video users (all VoIP users' α =0.8, 4 video users' γ = 1.0, 8 video users' γ = 0.8)	USQA-MC scheduler
Case 9	54 VoIP users plus 12 video users (all VoIP users' α =0.8, all video users' γ = 0.8)	USQA-C scheduler

2) Simulation Results

In this dissertation, the downlink MAC throughput is used to derive the approximate system capacity improvement. System capacity improvement is measured by the increase in the maximum supportable number of users by the system. A rough mapping from the downlink MAC throughput to the system capacity improvement can be done based upon (4-3):

$$\frac{1/(\text{MAC throughtput for USQA})}{1/(\text{MAC throughput for Baseline})} - 1 \qquad (4\text{-}3)$$

The simulation results for the MOS value and capacity improvement are shown in Tables 4-5, 4-6, and 4-7.

Table 4-5 shows the MOS value for the proposed schedulers of USQA-M1 and USQA-M2. Approximately 9%, 10%, and 6.5% MOS improvement can be achieved for users with a sensitivity factor α of 1.2 in cases 1, 2, and 3 respectively where the USQA-M1, USQA-M1, and USQA-M2 schedulers are used respectively. For cases 1-2, where the USQA-M1 scheduler is used, the MOS for users with a sensitivity factor α of 0.8 decreases by about 2.0% and 3.6% respectively. For case 3, where the USQA-M2 scheduler is used, the MOS for users with a sensitivity factor α of 0.8 decreases only by 0.7%. We can conclude that the USQA-M1 and USQA-M2 schedulers can greatly improve the MOS for users that have a sensitivity factor α of 1.2.

Table 4-5 Average MOS value for USQA-M1&M2

Cases	Scheduler	User category	MOS	MOS improvement
Case 1	USQA-M1	VoIP (1.2)	3.86	9%
		VoIP (1.0)	3.84	-1%
		VoIP (0.8)	3.87	-2%
	Basleine	VoIP (1.2)	3.54	N/A
		VoIP (1.0)	3.88	N/A
		VoIP (0.8)	3.95	N/A
Case 2	USQA-M1	VoIP (1.2)	2.91	10%
		VoIP (1.0)	3.07	4%
		VoIP (0.8)	2.68	-3.6%
		Video	3.03	-1.3%
	Baseline	VoIP (1.2)	2.65	N/A
		VoIP (1.0)	2.95	N/A
		VoIP (0.8)	2.78	N/A
		Video	3.07	N/A
Case 3	USQA-M2	VoIP (1.2)	3.77	6.5%
		VoIP (1.0)	3.88	0%
		VoIP (0.8)	3.93	-0.7%
	Baseline	VoIP (1.2)	3.54	N/A
		VoIP (1.0)	3.88	N/A
		VoIP (0.8)	3.95	N/A

Table 4-6 shows the capacity improvement for the proposed schedulers of USQA-MC and USQA-C. In this table, capacity improvement for VoIP and video users is listed separately. From the table, we find, approximately 4.5%, 8.2%, 7.6%, 15.1% and 25.1% capacity improvement can be achieved for VoIP users for case 4-8 respectively, whereas about 17.1% and 34.1% capacity improvement can be achieved for video users for cases 6 and 8 respectively. The capacity improvement gain increases as the ratio of users with sensitivity factors of 0.8 increases, which is also verified in the following simulation results of this section.

Table 4-7 shows the MOS for the proposed schedulers of USQA-MC and USQA-C. For cases 4-6, where the USQA-MC is used to optimize the MOS value for users with a sensitivity factor α of 1.2, we find approximately 6%, 5.5%, and 5.5% VoIP MOS improvement for users with a sensitivity factor α of 1.2 can be achieved respectively, whereas the MOS for users with a sensitivity factor α of 0.8 decreases by about 1.3%, 4.6%, and 4.5% respectively. For case 7, where the USQA-C scheduler is used, the MOS for users with a sensitivity factor α of 0.8 is decreased only by 2.8%.

Table 4-6 System capacity for USQA-MC&C

Cases	Scheduler	MAC throughput (Mbps)	Capacity improvement
Case 4	USQA-MC	1.000 (VoIP)	4.5%
	Basleine	1.045 (VoIP)	N/A
Case 5	USQA-MC	0.843 (VoIP)	8.2%
	Baseline	0.912(VoIP)	N/A
Case 6	USQA-MC	0.847 (VoIP)	7.6%
		11.456 (video)	17.1%
	Baseline	0.912 (VoIP)	N/A
		13.410 (video)	N/A
Case 7	USQA-C	0.908 (VoIP)	15.1%
	Baseline	1.045 (VoIP)	N/A
Case 8	USQA-C	0.728 (VoIP)	25.1%
		10.005(video)	34.0%
	Baseline	0.912 (VoIP)	N/A
		13.410 (video)	N/A

Table 4-7 Average MOS value for USQA-MC&C

Cases	Scheduler	User category	MOS	MOS improvement
Case 4	USQA-MC	VoIP (1.2)	3.75	6%
		VoIP (1.0)	3.83	-1.3%
		VoIP (0.8)	3.90	-1.3%
	Basleine	VoIP (1.2)	3.54	N/A
		VoIP (1.0)	3.88	N/A
		VoIP (0.8)	3.95	N/A
Case 5	USQA-MC	VoIP (1.2)	2.80	5.5%
		VoIP (1.0)	2.96	0.4%
		VoIP (0.8)	2.65	-4.6%
		Video	3.12	1.6%
	Baseline	VoIP (1.2)	2.65	N/A
		VoIP (1.0)	2.95	N/A
		VoIP (0.8)	2.78	N/A
		Video	3.07	N/A
Case 6	USQA-MC	VoIP (1.2)	2.80	5.5%
		VoIP (1.0)	2.98	0.9%
		VoIP (0.8)	2.66	-4.5%
		Video	3.33	7.1%
	Baseline	VoIP (1.2)	2.65	N/A
		VoIP (1.0)	2.95	N/A
		VoIP (0.8)	2.78	N/A
		Video	3.11	N/A
Case 7	USQA-C	VoIP (1.2)	N/A	N/A
		VoIP (1.0)	N/A	N/A
		VoIP (0.8)	3.79	-2.8%
	Baseline	VoIP (1.2)	N/A	N/A
		VoIP (1.0)	N/A	N/A
		VoIP (0.8)	3.90	N/A
Case 8	USQA-C	VoIP (1.2)	N/A	N/A
		VoIP (1.0)	N/A	N/A
		VoIP (0.8)	3.12	4.4%
		Video	3.40	8%
	Baseline	VoIP (1.2)	N/A	N/A
		VoIP (1.0)	N/A	N/A
		VoIP (0.8)	2.99	N/A
		Video	3.16	N/A

We can also observe from Table 4-7 that about 7.1% and 8.0% video MOS improvement can be achieved for cases 6 and 8 respectively. The MOS gain mostly comes from the big capacity improvement when the USQA-MC&C schedulers are used since the system is heavily

loaded. The same MOS improvement can also be observed from VoIP users for case 8, where the MOS is increased by 4.4% for users with a sensitivity factor α of 0.8.

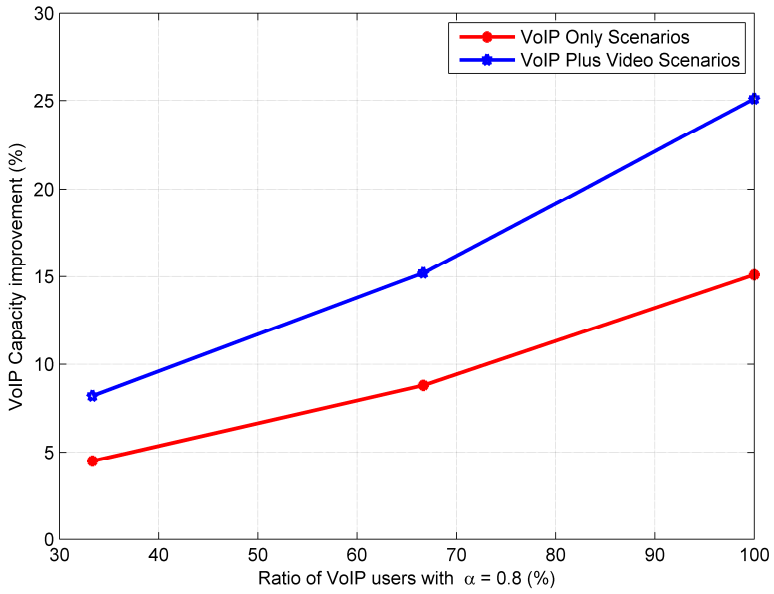

Figure 4.8 VoIP capacity improvement as a function of the ratio of VoIP users with $\alpha = 0.8$

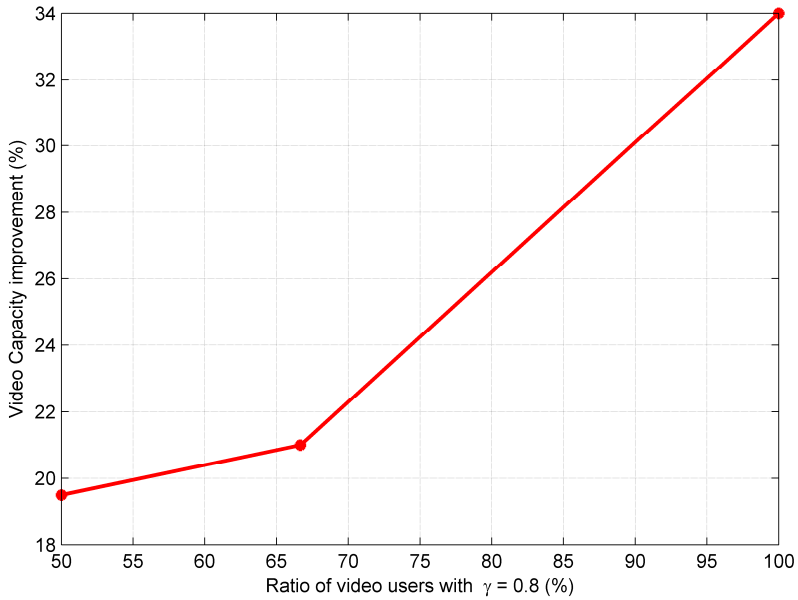

Figure 4.9 Video capacity improvement as a function of the ratio of video users with $\gamma = 0.8$

Figure 4.8 illustrates the VoIP capacity improvement as the ratio of VoIP users with lower sensitivity factors α of 0.8 in VoIP only scenarios corresponding to cases 1-3 in Table 4-4.

From this figure, we can see that the capacity improvement increases from 4.5% to 8.8% to 15.1% as the ratio increases from 1/3 to 2/3 to 1.

Figure 4.8 also illustrates the VoIP capacity improvement as a function of the ratio of VoIP users with lower sensitivity factors α of 0.8 in VoIP plus video scenarios corresponding to cases 4-6 in Table 4-4. From this figure, we can see that the capacity improvement increases from 8.2% to 15.2% to 25.1% as the ratio increases from 1/3 to 2/3 to 1.

Figure 4.9 illustrates the video capacity improvement as a function of the ratio of video users with lower sensitivity factors γ of 0.8 in VoIP plus video scenarios corresponding to cases 7-9 in Table 4-4. From this figure, we can that see the capacity improvement increases from 19.5% to 21.0% to 34.0% as the ratio increases from 1/2 to 2/3 to 1.

4.10.2 USQA-F Scheduler

Similarly, the system simulation was run using the OPNET 17.5 Modeler [86] with the LTE modules. The system simulation configuration is partly based upon LTE macrocell system simulation baseline parameters [87] as shown in Table 4-8. One macrocell with 10 FTP users was tested for a downlink USQA-F scheduler, with an ideal uplink receiver.

Table 4-8 System simulation configuration for USQA-F

Parameter	Assumption
Cellular Layout	1 Macrocell
Cell Radius	1Kilometer
Path loss model	3GPP suburban Macrocell
Mobility model	Random Way Point (RWP) with a speed of 0.1km/h
Carrier Frequency	Uplink: 1920 MHz Downlink: 2110 MHz
System Bandwidth	10MHz
Channel model	ITU Pedestrian A
Total BS TX power	40dBm
UE power class	23dBm
Number of Users	10 FTP users
FTP File Size	Constant 5MByte for each user
Scheduler	Dynamic scheduling: USQA-F vs. LTE baseline scheduler
Other assumptions	Ideal uplink receiver (no block error and packet loss)

1) Simulation Cases

Table 4-9 is the simulation cases used in this section. Case 1 is designed to verify the MOS improvement of the MOS targeted scheduler where user 2 have the sensitivity factors of 1.2 while other users have the sensitivity factors of 1. Case 2 is designed to verify the MOS improvement of the MOS targeted scheduler where users 1-2 have the sensitivity factors of 1.2 while other users have the sensitivity factors of 1. The coefficients a and b in (3-19) are determined as 5.0074 and 3.9595e-06 respectively.

2) Simulation Results

Table 4-9 System simulation cases for USQA-F

Cases	Assumptions
Case 1	User 2 has the sensitivity factor of 1.2 while the remaining users have the sensitivity factors of 1
Case 2	Users 1-2 have the sensitivity factor of 1.2 while the remaining users have the sensitivity factors of 1

From Figure 4.10, we find, with a predefined scheduling probability P1 (33.33 %) in time domain and with a predefined scheduling probability P2 (50 %) in frequency domain assigned to user 2, a 69 % MOS improvement can be achieved. The scheduler is flexible to control the target MOS of a specific user by changing or adapting its probability of the time domain scheduling and frequency domain scheduling priority. As we also see from Figure 4.10, some other users have somewhat decreased MOS. It is easily understandable since more resources are allocated to user 2 with a higher sensitive factor.

From Figure 4.11, we find, with a predefined scheduling probability P1 (33.33 %) in time domain and with a predefined scheduling probability P2 (50 %) in frequency domain assigned to users 1-2, a 16% and 68% MOS improvement can be achieved respectively for user 1 and 2. The scheduler is flexible to control the target MOS of a specific user by changing or adapting its probability of the time domain scheduling and frequency domain scheduling priority. As we also

see from Figure 4.11, some other users have somewhat decreased MOS. It is easily understandable since more resources are allocated to users 1-2 with a higher sensitive factor.

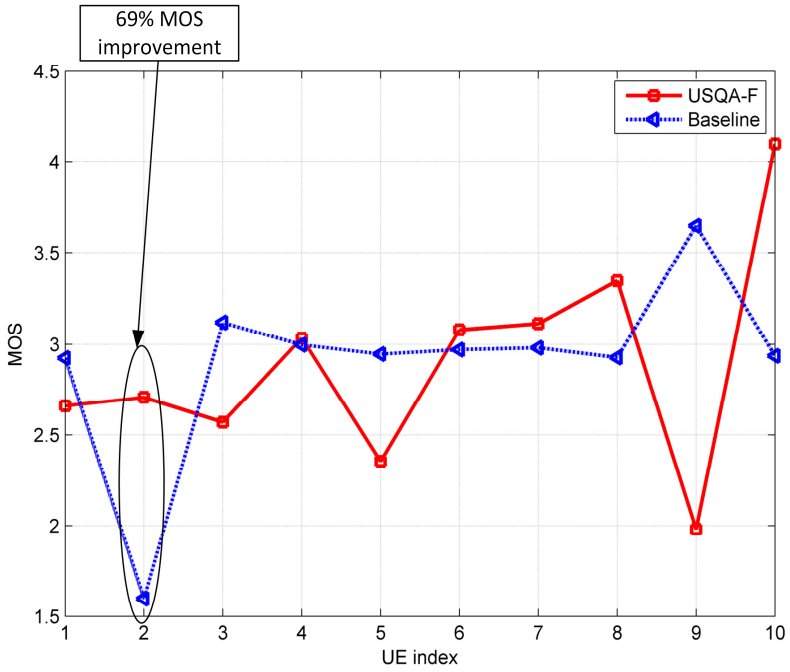

Figure 4.10 Average MOS as a function of UE index (case 1)

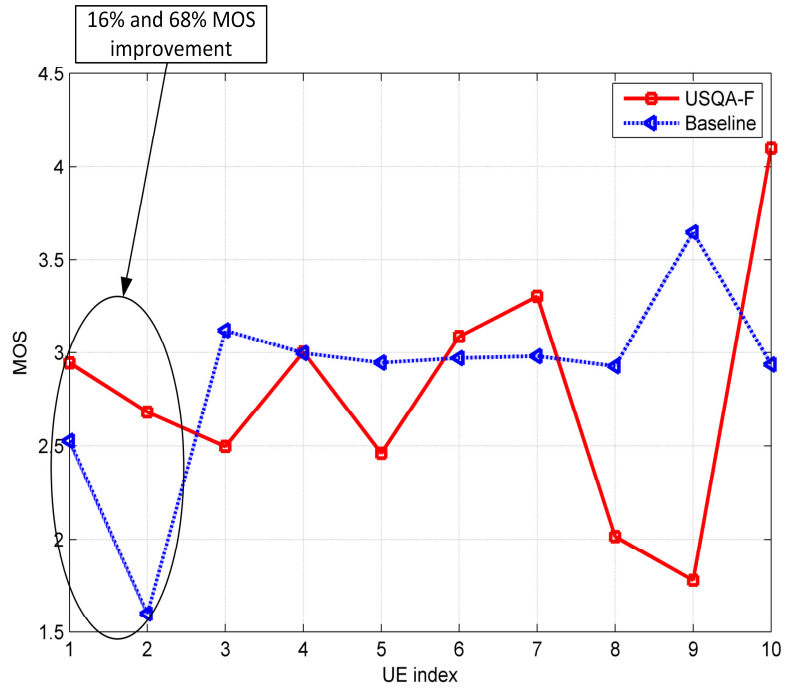

Figure 4.11 Average MOS as a function of UE index (case 2)

4.10.3 User-Specific Frequency Sensitivity QoS Study

Similarly, the system simulation was run using the OPNET 17.5 Modeler [86] with the LTE modules. The system simulation configuration is partly based upon LTE macrocell system simulation baseline parameters [87] as shown in Table 4-10.

1) Simulation Cases

Table 4-11 shows the test case for user-specific frequency sensitivity QoS study, where 24 G.711 VoIP users have different frequency sensitivity factors β of 1.0, 0.75, 0.5, and 0.25 respectively.

Table 4-10 System simulation configuration for user-specific frequency sensitivity QoS study

Parameter	Assumption
Cellular Layout	1 Macrocell
Cell Radius	1Kilometer
Path loss model	3GPP suburban Macrocell
Mobility model	Random Way Point (RWP) with a speed of 0.1km/h
Carrier Frequency	Uplink: 1920 MHz Downlink: 2110 MHz
System Bandwidth	10MHz
Channel model	ITU Pedestrian A
Total BS TX power	40dBm
UE power class	23dBm
Number of Users	24 G.711 VoIP users
Scheduler	Dynamic scheduling:
Other assumptions	Ideal uplink receiver (no block error and packet loss)

Table 4-11 System simulation cases for user-specific frequency sensitivity QoS study

Cases	Assumption
Case 1	24 G.711 VoIP users with a freqeuncy sensitivity factor $\beta = 1, 0.75, 0.5,$ and 025 respectively.

2) Simulation Results

Figure 4.12 plots the approximate capacity improvement (i.e., maximum supportable number of users) as a function of frequency sensitivity factor β. In the simulation, a rough

mapping from the PDSCH (Physical Downlink Shared Channel) load to the system capacity improvement can be done according to the following formula:

$$\text{Capacity impovement for factor } \beta(\%) = \frac{1/(\text{load for factor } \beta)}{1/(\text{load for factor } \beta = 1)} - 1 \quad (4\text{–}4)$$

From Figure 4.12, we can see that more than 100% capacity improvement can be achieved with a sensitivity factor β of 0.25, while an increase of around 30% can be achieved with a sensitivity factor β of 0.5 and 0.75.

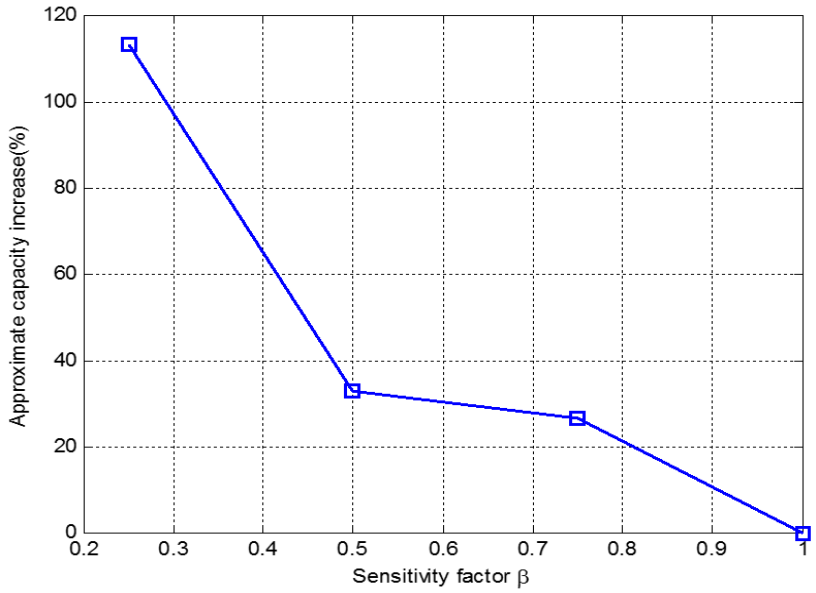

Figure 4.12 Approximate capacity improvement as a function of frequency sensitivity factor β

4.10.4 USQA-C Scheduler in Heterogeneous Networks

Similarly, the system simulation was run using the OPNET 17.5 Modeler with the LTE modules [86], as depicted in Table 4-12. In this section, we consider either one or three macrocell(s) and 2 overlaid picocells in each macrocell with 30/24/18 video users and 30/24/18 FTP users in each macro-plus-pico area. Video users and FTP users exist simultaneously in order to simulate practical scenarios with high traffic system loads. These configurations were evaluated for downlink scheduling, with an ideal uplink receiver, and Inter-cell Interference Coordination (ICIC) is not used.

1) Simulation Cases

Three cases corresponding to Configuration 4b of reference [88] were designed and simulated as described in Table 4-13. In case 1, 30 FTP users and 30 video users that have a data rate sensitivity factor of 0.8 are assumed in the simulation for each macro-plus-pico area. Among the 60 users, 10 FTP users and 10 video users are randomly and uniformly spaced within a 40m radius of each picocell. The remaining 10 FTP users and 10 video users are randomly and uniformly spaced within the remaining area of the macrocell.

Table 4-12 System simulation configuration for HetNet

Parameters	Assumption
Cellular layout	1/3 macrocell(s)
Cell radius	1kilometer
Pico cells	0 or 2 picocell(s)/macrocell (Configuration 4b)
Path loss model	3GPP suburban Macrocell
Mobility model	Random Way Point (RWP) with speed of 0.1km/h
Carrier frequency	Uplink:1920MHz Downlink:2110MHz
System bandwidth	10/20 MHz
Channel model	ITU Pedestrian A
Total BS TX power	macro: 46 dBm pico: 30 dBm
UE power	23 dBm
Video data rates	1105.920 Kbits/s (Level 3) 967.680 Kbits/s (Level 4) 829.440 Kbits/s (Level 5) 691.200 Kbits/s (Level 6)
FTP file size	File size: 15Mbytes
Number of users	30/24/18 video Users (each user takes a data rate sensitivity factor of 0.8), and 30/24/18 FTP users
Scheduler	Dynamic scheduling USQA-C scheduler and LTE baseline scheduler
ICIC	No ICIC used
Other assumptions	Ideal uplink receiver(no block error and packet loss)

In cases 2 and 3, 18 FTP users and 18 video users that have a data rate sensitivity factor of 0.8 are assumed in the simulation for each macro-plus-pico area. Among the 36 users, 6 FTP users and 6 video users are randomly and uniformly distributed within a 40m radius of each picocell. The remaining 6 FTP users and 6 video users are randomly and uniformly distributed within the remaining area of each macrocell. In case 3, three macrocells and two picocells in each macrocell exist.

For performance comparison, the scenarios where only macro cells are used, without picocells, for 30/18 FTP users and 30/18 video users in each macrocell are also simulated and the results appear below.

Table 4-13 System simulation cases for HetNet

Cases	Assumption
Case 1	10 MHz, 1 macrocell+2 picocells, 30 FTP users and 30 video users, each having a value for the data rate sensitivity factor γ of 0.8.
Case 2	10 MHz, 1 macrocell+2 picocells, 18 FTP users and 18 video users, each having a value for the data rate sensitivity factor γ of 0.8.
Case 3	20 MHz, 3 macrocells+6 picocells, 18 FTP users and 18 video users for each macrocell, each having a value for the data rate sensitivity factor γ of 0.8.

2) Simulation Results

In the following figures, green (left), blue (center) and red (right) bars represent, respectively, the performance for 1) 1/3 macrocell(s) without picocells and user-specific QoS consideration, 2) 1/3 macrocell(s) with 2/6 picocells but without user-specific QoS consideration, and 3) 1/3 macrocell(s) with 2/6 picocells and with user-specific QoS consideration.

a) FTP Throughput

In this dissertation, FTP throughput per cell is defined as the total received MAC traffic in Mbits/sec for all the FTP users divided by the number of macrocells.

Figures 4.13–4.15 show the FTP throughput comparison for cases 1, 2 and 3, respectively, with and without user-specific QoS considered in heterogeneous networks. For comparison, the performance for a reference scenario where no picocells exist is also presented in the figures (i.e., green bar).

From Figures 4.13–4.15 we can see that much better FTP throughput can be achieved when picocells are introduced in heterogeneous networks (i.e., between green and blue bars), and about 37%, 14% and 11% throughput improvement (i.e., between blue and red bars) can be further achieved in cases 1, 2 and 3 respectively when user-specific QoS is considered. The throughput improvement of FTP user results from the decreased throughput of video users when user-specific QoS is used in the proposed rate adaptation algorithm.

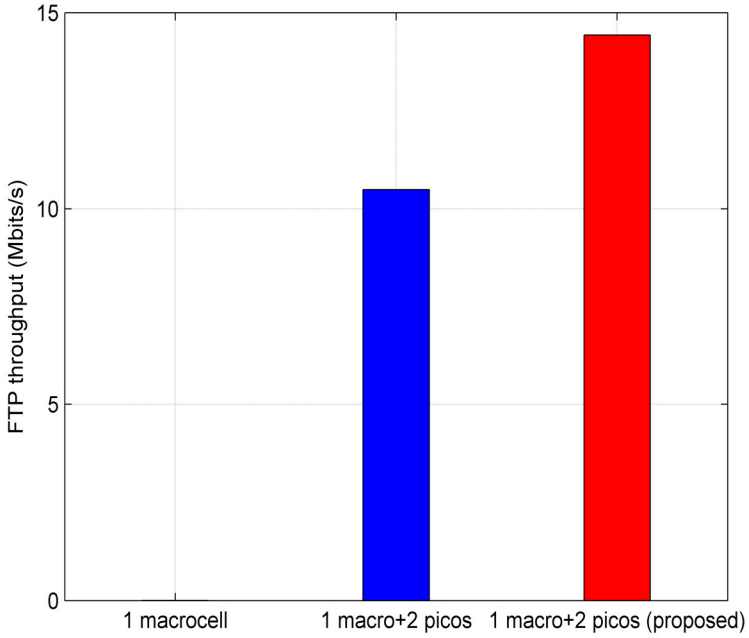

Figure 4.13 FTP throughput comparison for HetNet case 1

b) FTP Throughput Improvement vs. Number of Video Users

Figure 4.16 shows the FTP throughput improvement as a function of the number of video users. The FTP throughput improvement results from the lower data rate of video users when the user-specific QoS is considered. In this figure, 1 macrocell-plus-2 picocells are simulated with

varying number of video users and FTP users of 18, 24, and 30. It can be found as the number of video users increases, a higher FTP throughput improvement can be achieved. That is because as the number of video users increases, more resources can be saved from video users and used to convey FTP data.

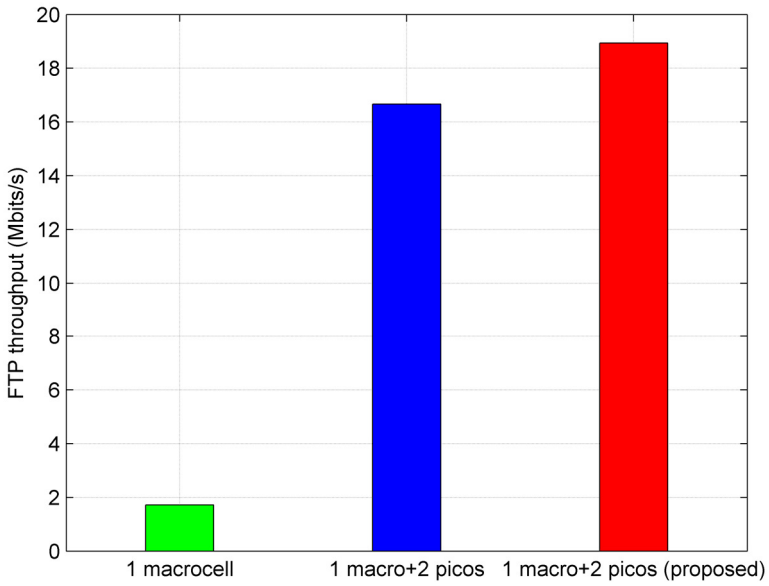

Figure 4.14 FTP throughput comparison for HetNet case 2

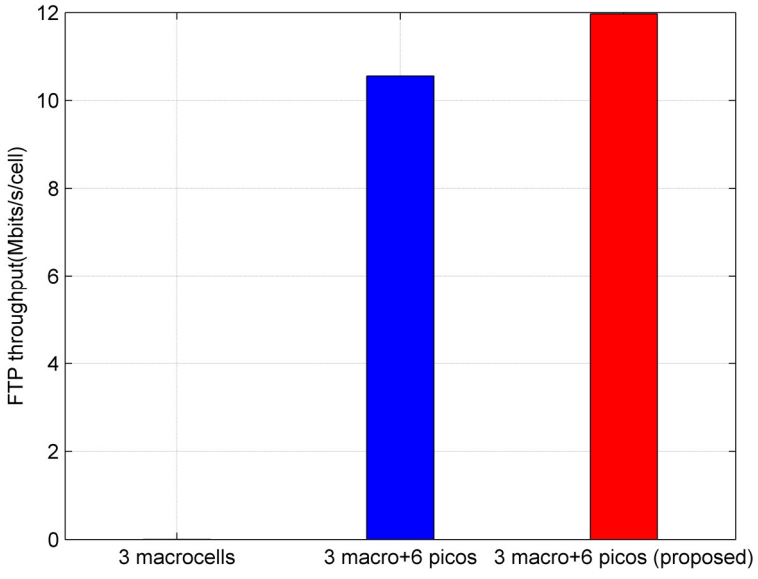

Figure 4.15 FTP throughput comparison for HetNet case 3

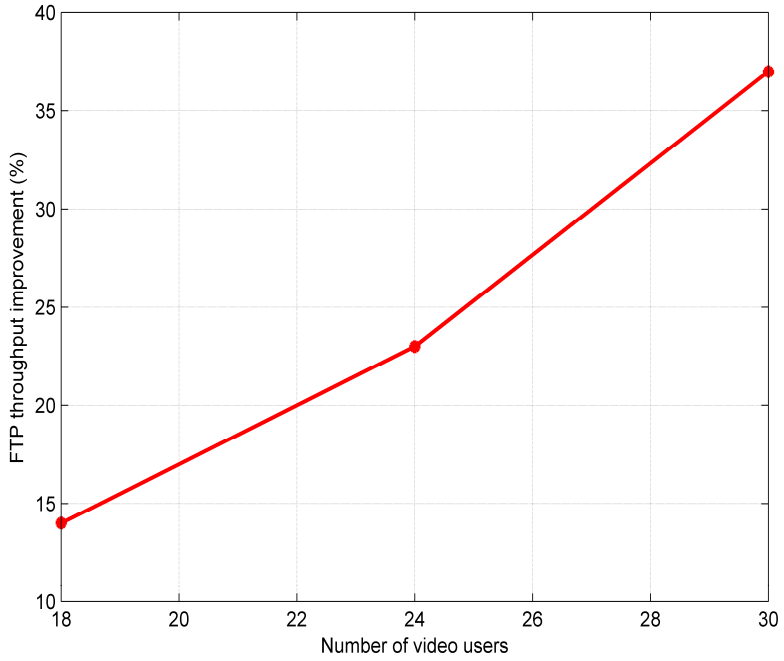

Figure 4.16 FTP throughput improvement as a function of the number of video users for 1 macrocell-plus-2 picocells

c) Video MOS

Figures 4.17-4.19 show the MOS comparison for cases 1, 2 and 3, respectively, with and without user-specific QoS considered in heterogeneous networks. For comparison, the performance for a reference scenario where no picocells exist is also presented in the figures (i.e., green bar).

From Figures 4.17-4.19, we can see a much better MOS value (56%, 12%, and 53% improvement respectively) can be achieved when picocells are introduced in heterogeneous networks (i.e., between green and blue bars), while similar MOS value (i.e., between blue and red bars) can be maintained when the user-specific QoS is considered. As noted above, FTP throughput can be improved significantly.

d) Homogeneous vs. Heterogeneous Networks

For case 2, 1 macrocell plus 2 picocells and purely 1 macrocell are compared, both taking into account user-specific QoS requirements. The average MAC throughput and SINR comparison for video users are listed in Table 4-14. From this table, we can see a ~5dB average SINR increase

in heterogeneous networks results in about only less than 1% MAC throughput increase (i.e., higher video data rate). Therefore, most of the system capacity gain is maintained in heterogeneous networks.

Similarly, for case 3, 3 macrocells plus 6 picocells and purely 3 macrocells are compared, both taking into account user-specific QoS requirements. The average MAC throughput and SINR comparison for video users are also listed in Table 4-14. From this table, we can see a ~13dB average SINR increase in heterogeneous networks results in ~12% MAC throughput increase (i.e., higher video data rate). Therefore, only part of the system capacity gain can be maintained in heterogeneous networks.

4.11 Concluding Remarks

In this chapter, we introduced the concept of user-specific QoS aware schedulers and demonstrated their utility in spectral allocation for MOS and/or system capacity gain. Five user-specific QoS aware schedulers are proposed aimed at improving the MOS or system capacity or both the MOS and system capacity in wireless systems. Using the OPNET Modeler for LTE, system simulations have been performed for a set of Voice over IP (VoIP) users, including cases with additional video users, where users have been assigned specific QoS target levels. Simulation results showed that appreciable MOS and system capacity improvements could be achieved when user-specific QoS requirements are utilized in the user-specific QoS aware schedulers.

OPNET system simulations were also performed for a set of FTP users and video users that were assigned specific QoS target levels in LTE heterogeneous networks Simulation results show that significant system capacity improvements and acceptable MOS levels can be achieved for FTP users and video users, respectively, if such user-specific QoS requirements are considered in heterogeneous networks in the user-specific QoS aware scheduler.

Moreover, the user-specific frequency sensitivity QoS results or guidelines, when combined with VoIP codecs matched to the different high frequency auditory characteristics of users, can achieve higher system capacity as well as a comparable MOS levels.

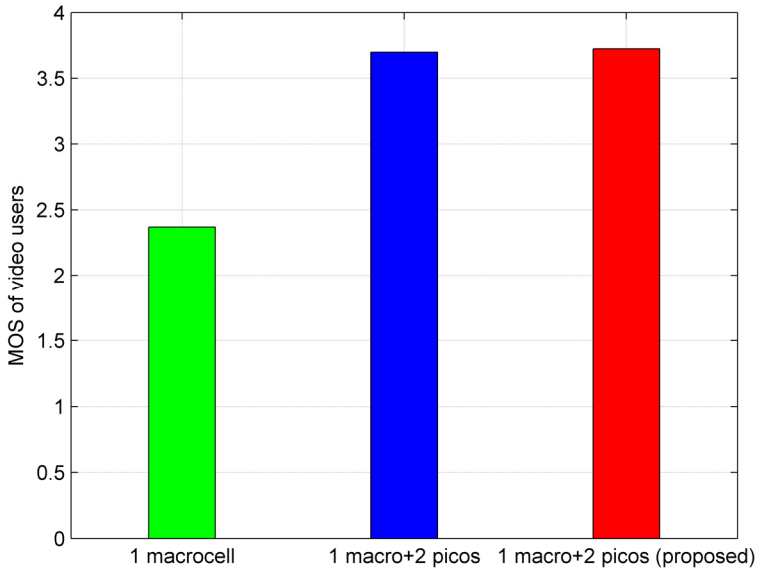

Figure 4.17 Video MOS comparison for HetNet case 1

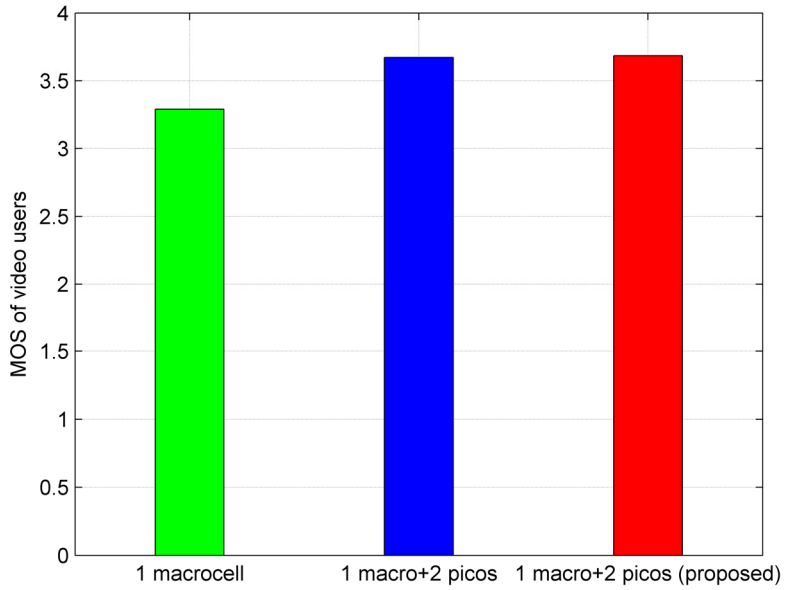

Figure 4.18 Video MOS comparison for HetNet case 2

Table 4-14 MAC throughput and SINR comparison

Cases	MAC throughput (Mbps)	SINR (dB)
1 macro	14.782	34.3
1 macro + 2 picos	14.905	39.5
3 macros	13.176	25.4
3 macros + 6 picos	14.762	38.7

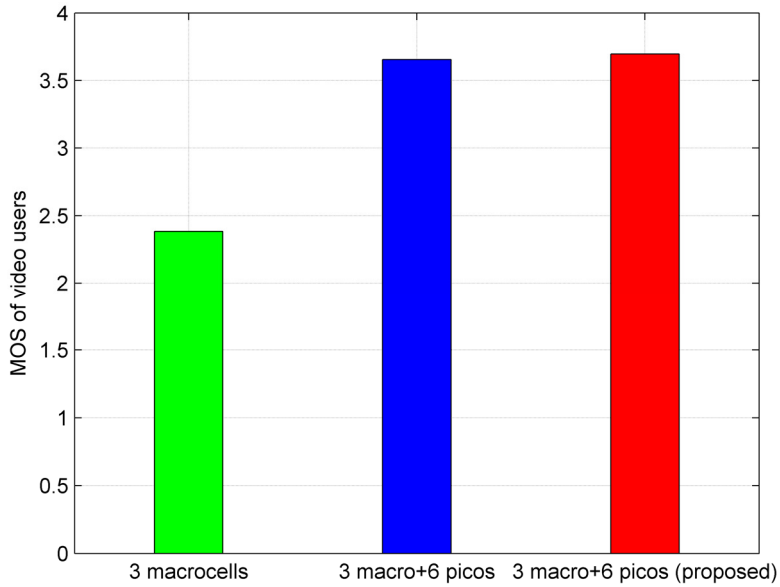

Figure 4.19 Video MOS comparison for HetNet case 3

CHAPTER 5. USER-SPECIFIC QOS AWARE SCHEDULER IMPLEMENTATION

5.1 Introduction

Since the system performance gain has been shown to be significant in the previous chapters when user-specific QoS requirements are considered in the scheduling algorithms, it is important to investigate the implementation details, such as the feasibility, complexity and protocol adaptation, so that the user-specific QoS aware scheduling algorithms can find application in LTE or future networks. In this chapter, these implementation details will be analyzed and the system performance considering the scheduling period will also be evaluated.

The rest of the chapter is organized as follows. LTE end-to-end procedures related to users' QoS are provided in section 5.2. Section 5.3 lists the LTE QoS related protocols that need to be adapted. Section 5.4 gives the general idea of how the user-specific QoS parameters can be obtained and used by the scheduling algorithms. Section 5.5 gives the system architecture of the user-specific QoS aware scheduling implementation. The user-specific QoS optimization process is provided in section 5.6. The detailed protocols adaptation related to the implementation of user-specific QoS aware schedulers is presented in section 5.7. The scheduler details of rate adaptation triggering, scheduling period, and integration with existing rate adaptation algorithms were described in section 5.8. We provided the system simulation in section 5.9. Finally in section 5.10, we conclude this chapter.

5.2 LTE End-to-End Procedures

In the following sections, for the simplicity of illustration, we consider only the network architecture for non-roaming scenarios. The consideration for roaming scenarios is a straightforward extension of the methodology that we present.

LTE end-to-end QoS-related procedures are shown in Figure 5.1 [89], [90] and discussed in the following sections. The LTE end-to-end procedures are composed of three major functions: 1) SIP

signaling, 2) AF session establishment/modification, and 3) EPS bearer establishment. These three major procedures will be illustrated respectively in detail in the following sections.

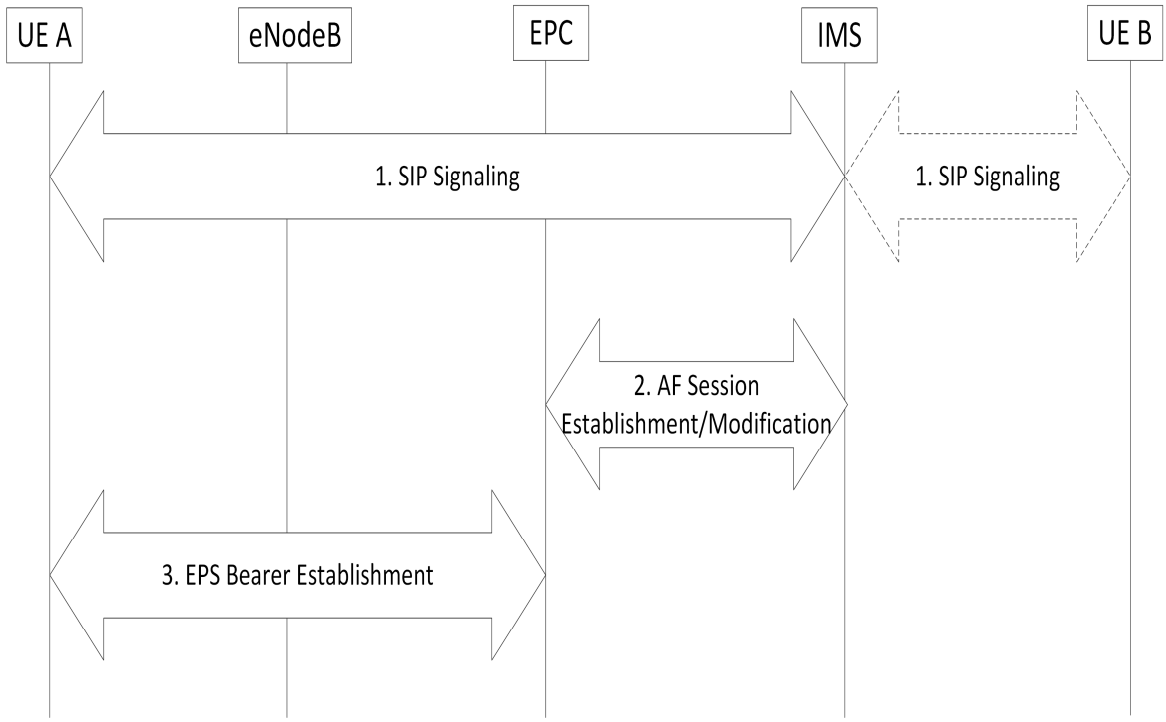

Figure 5.1 LTE end-to-end procedures

5.2.1 SIP Signaling

A typical SIP signaling to setup a voice call is shown in Figure 5.2 and illustrated as follows [44]. Here we ignore the registration procedures and assume the registration procedures have already been completed before the voice call setup procedure starts.

1) UE A sends the INVITE request to the IMS SIP server. The INVITE request contains a message body with the SDP media description of the type of session that UE A is willing to establish with UE B.
2) The IMS SIP server receives the INVITE request and forwards the INVITE request to UE B.
3) The IMS SIP server sends a 100 (Trying) response back to UE A. The 100 (Trying) response indicates that the INVITE request has been received and that the IMS is working to forward

the INVITE request to the destination. Typically responses in SIP use a three-digit code followed by a descriptive phrase.

4) When UE B receives the INVITE request and is alerted to the incoming call, it sends back a 180 (Ringing) response.
5) The IMS SIP server receives the 180 (Ringing) response and forwards it to UE A to indicate that UE B is alerted.
6) When UE B decides to answer the call, it sends a 200 (OK) response to indicate that the call has been answered. The 200 (OK) response contains a message body with the SDP media description of the type of session that UE B is willing to establish with UE A.
7) The IMS SIP sever receives the 200 (OK) response, and forwards it to UE A.
8) Finally, UE A sends an acknowledgement message, ACK, to UE B to confirm the reception of the final 200 (OK) response.

The media session now begins between UE A and UE B, and they can send media packets using the format to which they agreed in the exchange of SDP media description.

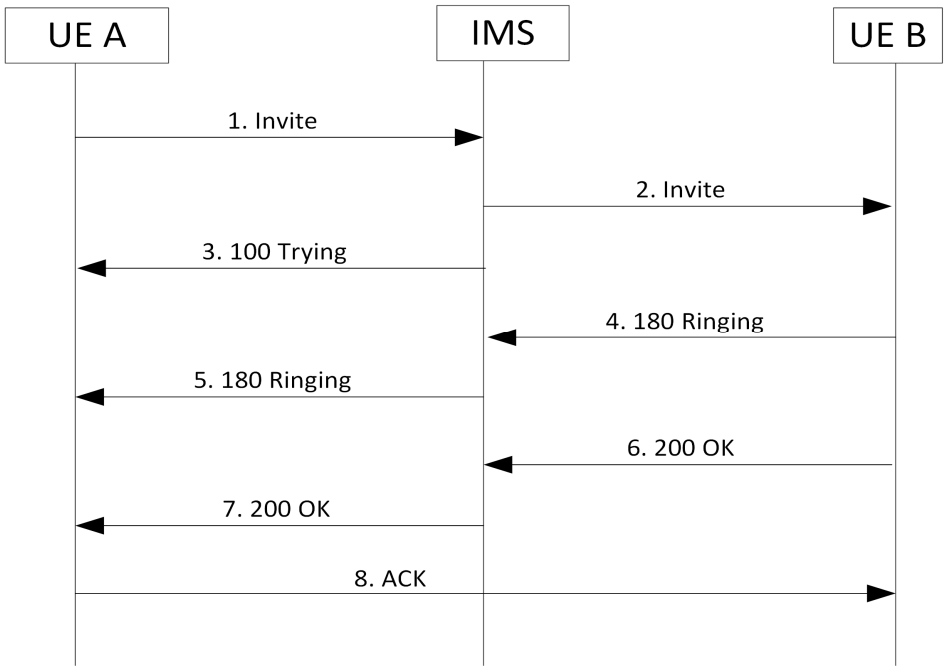

Figure 5.2 Sip signaling procedures

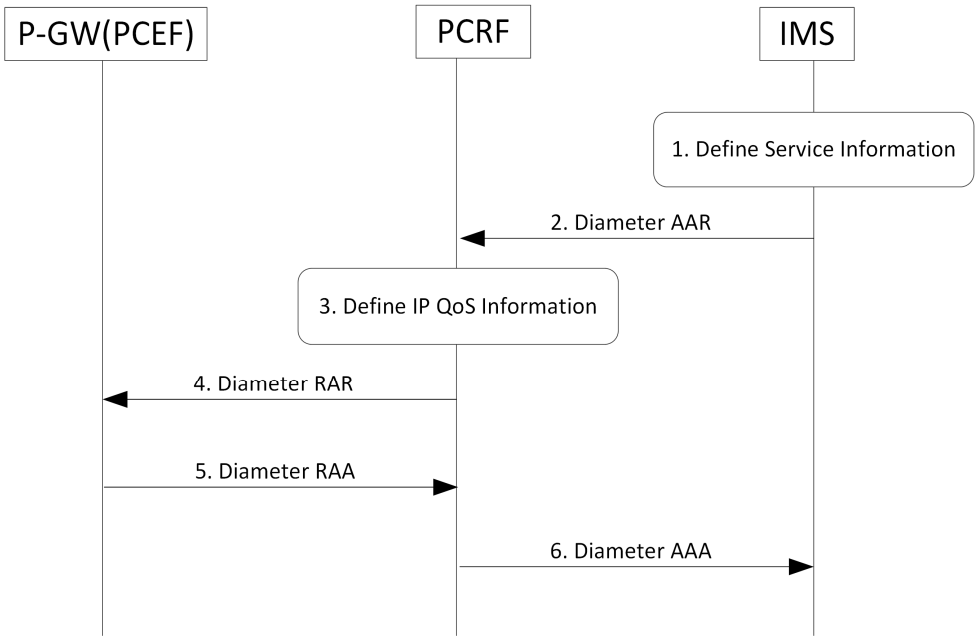

Figure 5.3 AF session establishment/modification procedures

5.2.2 AF Session Establishment/Modification

The AF session establishment/modification procedure is shown in Figure 5.3 and illustrated as follows [91]:

1) When a subscriber establish an IMS call using the SIP protocol, the IMS use the mapping rules to derive a Media-Component-Description AVP (Service Information) from the SDP parameters in the SIP message [43].
2) The IMS provides the Service Information to the PCRF by sending a Diameter AAR message to set up a new Rx Diameter session (Rx interface).
3) The PCRF derives the IP QoS information (e.g., QCI, GBR, MBR, ARP) [43] from the Service Information received in the Diameter AAR message.
4) The PCRF sends the Diameter RAR message to trigger the P-GW (PCEF) to request Charging Rules (Gx interface). The Diameter RAR message contains the IP QoS information derived in the previous step.

5) The P-GW (PCEF) sends the Diameter RAA message to acknowledge the Diameter RAR message.

5.2.3 EPS Bearer Establishment

The EPS bearer establishment procedure is shown in Figure 5.4 and illustrated as follows ([89], [92]):

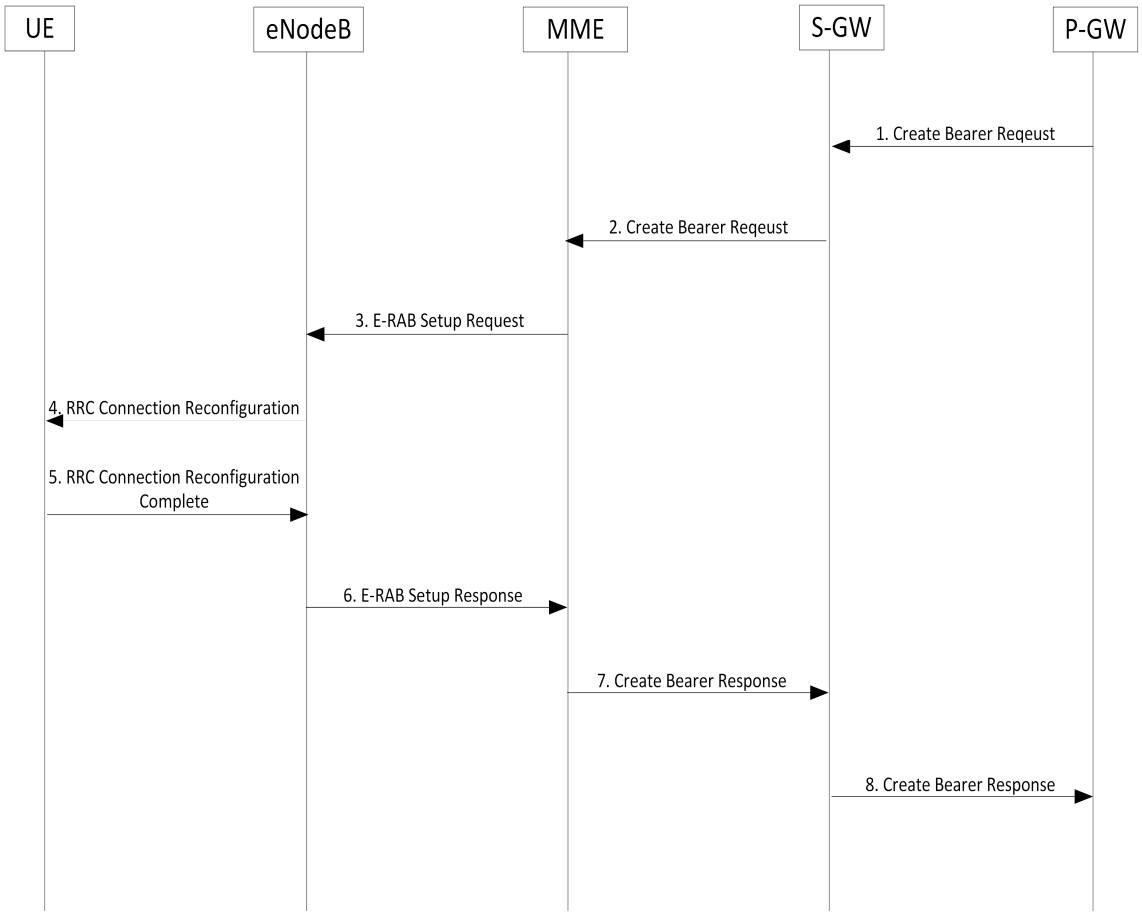

Figure 5.4 EPS bearer establishment procedures

1) The PDN-GW sends a Create Bearer Request message (containing IP QoS information received from the PCRF) towards the S-GW (S5 interface) to initiate the establishment process of the dedicated bearer.
2) The S-GW processes the Create Bearer Request message and forwards it to the MME (S11 interface) for further processing.

3) The MME now sends the E-RAB Setup Request message (containing IP QoS information received from the Create Bearer Request message) to the eNodeB (S1-MME interface) to allocate the bearer resource between the eNodeB and the S-GW.
4) The eNodeB allocates the radio resources for the radio bearers using a RRC Connection Reconfiguration Request message to the UE.
5) The UE establishes the radio bearers and sends back a RRC Connection Reconfiguration Complete message to the eNodeB.
6) The radio bearers are established between the eNodeB and the UE, so the eNodeB sends the E-RAB (Radio Access Bearer) Setup Response message to the MME.
7) The MME sends the Create Bearer Response message to the S-GW.
8) The S-GW forwards the Create Bearer Response message to the PDN-GW to complete the establishment process of the dedicated bearers.

5.3 LTE QoS Related Protocols

From the analysis of section 5.2, the LTE QoS related protocols are as follows:

- SIP protocol [44]

  The SIP protocol is used to create, modify, and terminate sessions such as Internet multimedia conferences, Internet telephone calls, and multimedia distribution.

- Diameter based protocol (Rx and Gx interfaces) [90]

  The Diameter base protocol provides an Authentication, Authorization and Accounting (AAA) framework for applications such as network access or IP mobility. Diameter is also intended to work in both local Authentication, Authorization & Accounting and roaming situations.

- GTP-C (Control) protocol (S5, S11 interfaces) [93]

  The control plane of the GPRS Tunneling Protocol (GTP) is responsible for creating, maintaining and deleting tunnels on Sx (e.g., S5, S11) interfaces.

- S1-AP protocol (S1-MME interface) [92]

  The S1-AP protocol provides the signaling service between the E-UTRAN and the EPC.

5.4 User-Specific QoS Parameter Acquisition

There are two methods to acquire the user-specific QoS parameters to be used by the user-specific QoS aware schedulers. The first one is to obtain the user-specific QoS parameters dynamically through the signaling messages (i.e., SIP, Diameter protocol etc.) that are delivered to the MAC layer. The other is to acquire the user-specific QoS parameters through the SPR database in the PCRF that are delivered to the MAC layer. The difference between these two methods is in how the PCRF obtains the user-specific QoS parameters. After the PCRF acquires the QoS parameters, the subsequent procedures will be the same so that the pertinent QoS parameters are conveyed to the MAC layer. These two methods will be described in detail in the following sections.

For the first method, as noted above, the user specific QoS parameters are obtained by the PCRF through signaling from the SIP and Rx interface protocols.

For the second method, no special SIP signaling is required before the PCRF sends the QoS parameters further to the PCEF through the Gx interface. In most commercial systems, the network operator can obtain the user-specific QoS requirements that are based primarily upon age. When users subscribe to a service from the network operator, they often provide their relevant information such as age, name, and nationality that can be used by the network to derive the user-specific QoS parameters. To be more specific, when a bearer is to be established or modified, the PCRF inquires of the SPR database about the relevant information of this user. If the relevant information shows that this user is older than a given age (e.g., 55) that can be configurable, this user will be considered as a user with a lower sensitivity factor, otherwise, it is regarded as a normal user.

5.5 System Architecture

The system architecture and interfaces based on the LTE system are illustrated in Figure 5.5, where only relevant modules are shown [94], [95]. In order to implement the user-specific QoS aware schedulers described in this dissertation, the AMR mode adaptation and video data rate adaptation

algorithms would be implemented in the Rate Adaptation module, whereas the user-specific MAC scheduling algorithm would be implemented in the eNodeB MAC layer. These are software only changes and can be readily accommodated in future versions of LTE [or in future 5G systems].

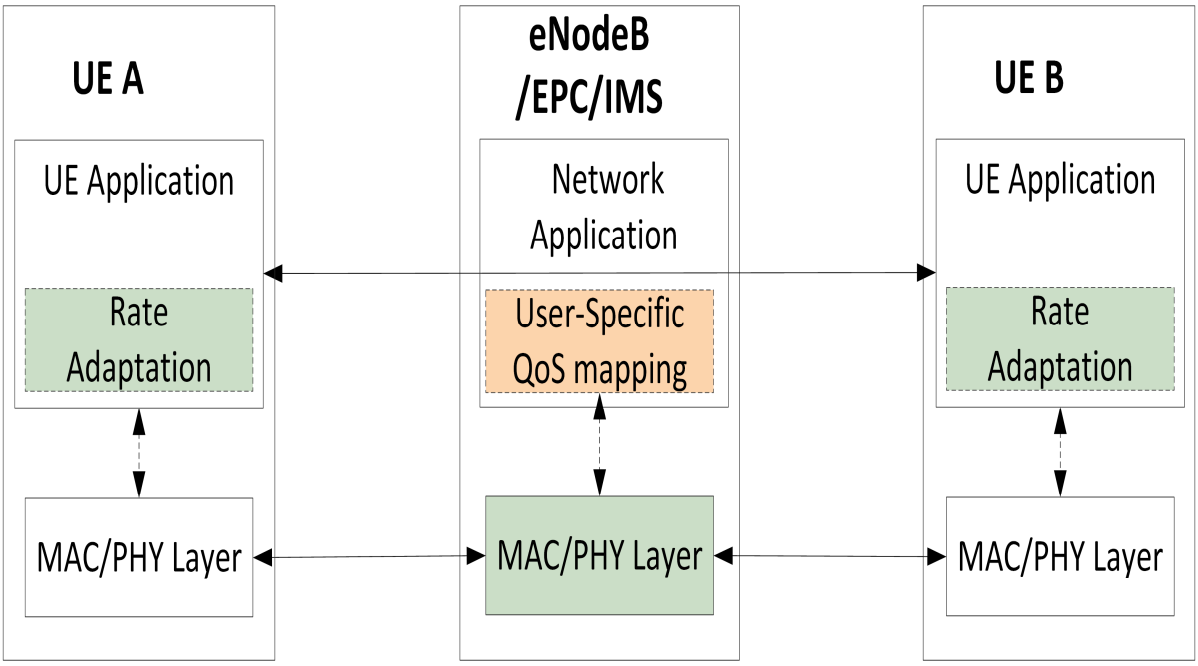

Figure 5.5 System architecture

5.6 Optimization Process

As an example of the optimization process, when a voice or video session is to be initialized through the SIP protocol [44], [96], the sender and receiver UE applications will negotiate with each other the application level QoS parameters such as supported AMR codec modes through the IMS (IP Multimedia Subsystem) [76]. User-specific QoS parameters could also be sent to the IMS by UE.

Next, user-specific QoS requirements will be mapped from the subscription database in the network (e.g., SPR [Subscriber Profile Repository]) [13], [39], [43] or user-specific QoS parameters obtained from UE during session initiation, as described in section 5.2. The user-specific QoS parameters shall be delivered to the MAC/PHY Layer in the eNodeB by the EPC/IMS and used to perform the user-specific QoS aware scheduling. Channel state information derived from physical layer measurements,

available spectral resources and buffer status should also be provided to the MAC scheduler to perform channel-dependent scheduling.

Finally, after the process of rate adaptation, the receiver UE will send the rate control command (e.g., CMR [Codec Mode Request] for VoIP) to the sender UE through the RTP or RTCP protocol [76], [79], [97] if the data rate is to be changed. Meanwhile the MAC layer in the eNodeB will perform the user-specific QoS aware scheduling.

It should be noted that, due to the changing channel environment for each user and varying network condition, the optimization process should be dynamic and periodic to achieve the maximum system performance gain.

5.7 Protocol Adaptation

Based upon the analysis in sections 5.3 and 5.4, the following protocols adaptation is proposed to support the user-specific QoS aware scheduling. As described in sections 5.2 - 5.4, the Gx interface, S5/S11 interface, and S1-MME interface protocols need to be adapted for the second user-specific QoS parameters derivation method. For the first acquisition method in addition, SIP and Rx interface protocol adaptation is needed [98]. Moreover, the RTCP protocol that is used to convey the control command is also analyzed to support the user-specific QoS aware scheduling.

5.7.1 SIP Protocol

When dynamic user-specific QoS information needs to be conveyed from the UE to the IMS, the SIP protocol [44], [96] needs to be adapted accordingly.

Both the SIP Request and Response messages use the basic format of RFC 2822 [99]. Both types of messages consist of a start-line, one or more header fields, an empty line indicating the end of the header fields, and an optional message-body, as shown in Figure 5.6.

One of the message-header (i.e., Content-Type) indicate the content-type of the message-body, One of the example is Content-Type: application/sdp, which means the message body is described by a SDF language [47].

The SDP protocol is intended for describing multimedia communication sessions for the purposes of session initiation, session termination, and parameter negotiation. When a session is initiated, the type of media, codec, and bandwidth are described using the SDP protocol.

```
generic-message  =  start-line
                    *message-header
                    CRLF
                    [ message-body ]
start-line       =  Request-Line / Status-Line
```

Figure 5.6 SIP message format

An SDP session description consists of a number of lines of text of the form as in Figure 5.7:

```
<type>=<value>
```

Figure 5.7 SDP format

An SDP session description consists of a session-level section followed by zero or more media-level sections. The session-level part starts with a "v=" line and continues to the first media-level section. Each media-level section starts with an "m=" line and continues to the next media-level section or end of the whole session description. The "v=" field gives the version of the Session Description Protocol. The "m=" line is defined as follows in Figure 5.8:

```
m=<media> <port> <proto> <fmt> ...
```

Figure 5.8 SDP media format

<media> is the media type. Currently defined media are "audio", "video", "text", "application", and "message".

<fmt> is a media format description. The fourth and any subsequent sub-fields describe the format of the media, e.g., codec modes.

The "b=" field denotes the proposed bandwidth to be used by the session or media as in Figure 5.9:

```
b=<bwtype>:<bandwidth>
```

Figure 5.9 SDP bandwidth format

So if user-specific QoS requirements need to be conveyed from the UE to the IMS, one method is to implicitly convey the user-specific QoS requirement through the media type field. Figure 5.10 shows the current defined media types for the media type field. A user-specific Audio or Video type can be added and defined, e.g., 100 indicates the user-specific Audio media type with a lower sensitivity factor, 200 indicates the user-specific Audio media type with a higher sensitivity factor, and 101 indicates the user-specific video media type with a lower sensitivity factor, as defined in Table 5-1. If new media types are defined this way, correspondingly, in the network and the peer UE, the media type field needs to be parsed differently.

- AUDIO (0)
- VIDEO (1)
- DATA (2)
- APPLICATION (3)
- CONTROL (4)
- TEXT (5)
- MESSAGE (6)
- OTHER (0xFFFFFFFF)

Figure 5.10 SDP media type defintion

5.7.2 RTCP Protocol

When an adapted data rate mode needs to be signaled, it uses the RTCP protocol. The current RTCP protocol supports the rate adaptation signaling. So it can be reused without any modifications on the RTCP protocol.

Specifically, for voice, the receiver UE can send RTP CMR message or RTCP-APP CMR message [76], [79], [97], to request the data rate adaptation from the sender UE through the RTP or RTCP protocol respectively.

For video, the Temporary Maximum Media Bit-rate Request (TMMBR) and Temporary Maximum Media Bit-rate Notification (TMMBN) messages of Codec-Control Messages (CCM) [76],

[100] are used in the session to indicate a desired video data rate to the sender UE and enable a dynamic video rate adaptation.

Table 5-1 User-specific SDP media type definition

User-specific QoS	Media type
Audio (low sensitivity factor)	0 + 100
Audio (normal sensitivity factor)	0
Audio (high sensitivity factor)	0 + 200
Video (low sensitivity factor)	1 + 100
Video (normal sensitivity factor)	1

5.7.3 Rx Interface

The Media-Component-Description AVP (Attribute Value Pair) is conveyed in the Diameter AAR message, and it contains Service Information for a single media component within an AF session [41]. The AAR message and Media-Component-Description AVP format are shown in Figures 5.11 and 5.12 respectively.

If the user-specific QoS parameters need to be conveyed from the UE, the Media-Component-Description AVP needs to be modified. The media-type field in the Media-Component-Description AVP can be used to convey the user-specific QoS requirements as section 5.7.1 describes. Similarly the PCRF needs to parse the media-type field differently as section 5.7.1 denotes.

5.7.4 Gx Interface

The PCRF may provide authorized QoS information to the PCEF after using the mapping rules to map from the Service Information to the authorized QoS information. The authorized QoS information shall be provisioned within a CCA or RAR Diameter message as QoS-Information AVP. The provisioning of the authorized QoS (which is composed of QCI, ARP and bitrates) is performed from the PCRF to the PCEF [40].

The RAR message and Media-Component-Description AVP format are shown in Figures 5.13 and 5.14 respectively.

```
<AA-Request> ::= < Diameter Header: 265, REQ, PXY >
                 < Session-Id >
                 { Auth-Application-Id }
                 { Origin-Host }
                 { Origin-Realm }
                 { Destination-Realm }
                 [ Destination-Host ]
                 [ IP-Domain-Id ]
                 [ AF-Application-Identifier ]
                *[ Media-Component-Description ]
                 [ Service-Info-Status ]
                 [ AF-Charging-Identifier ]
                 [ SIP-Forking-Indication ]
                *[ Specific-Action ]
                *[ Subscription-Id ]
                 [ OC-Supported-Features ]
                *[ Supported-Features ]
                 [ Reservation-Priority ]|
                 [ Framed-IP-Address ]
                 [ Framed-Ipv6-Prefix ]
                 [ Called-Station-Id ]
                 [ Service-URN ]
                 [ Sponsored-Connectivity-Data ]
                 [ MPS-Identifier ]
                 [ GCS-Identifier ]
                 [ Rx-Request-Type ]
                *[ Required-Access-Info ]
                 [ Origin-State-Id ]
                *[ Proxy-Info ]
                *[ Route-Record ]
                *[ AVP ]
```

Figure 5.11 Diameter AAR message format

```
Media-Component-Description ::= < AVP Header: 517 >
                                { Media-Component-Number } ; Ordinal number of the media comp.
                               *[ Media-Sub-Component ]    ; Set of flows for one flow identifier
                                [ AF-Application-Identifier ]
                                [ Media-Type ]
                                [ Max-Requested-Bandwidth-UL ]
                                [ Max-Requested-Bandwidth-DL ]
                                [ Min-Requested-Bandwidth-UL ]
                                [ Min-Requested-Bandwidth-DL ]
                                [ Flow-Status ]
                                [ Reservation-Priority ]
                                [ RS-Bandwidth ]
                                [ RR-Bandwidth ]
                               *[ Codec-Data ]
```

Figure 5.12 Media-Component-Description AVP format

```
<RA-Request> ::= < Diameter Header: 258, REQ, PXY >
                 < Session-Id >
                 { Auth-Application-Id }
                 { Origin-Host }
                 { Origin-Realm }
                 { Destination-Realm }
                 { Destination-Host }
                 { Re-Auth-Request-Type }
                 [ Session-Release-Cause ]
                 [ Origin-State-Id ]
                 [ OC-Supported-Features ]
                *[ Event-Trigger ]
                 [ Event-Report-Indication ]
                *[ Charging-Rule-Remove ]
                *[ Charging-Rule-Install ]
                 [ Default-EPS-Bearer-QoS ]
                *[ QoS-Information ]
                 [ Default-QoS-Information ]
                 [ Revalidation-Time ]
                *[ Usage-Monitoring-Information ]
                 [ PCSCF-Restoration-Indication ]
                *[ Proxy-Info ]
                *[ Route-Record ]
                *[ AVP ]
```

Figure 5.13 Diameter RAR message format

```
QoS-Information ::= < AVP Header: 1016 >
                    [ QoS-Class-Identifier ]
                    [ Max-Requested-Bandwidth-UL ]
                    [ Max-Requested-Bandwidth-DL ]
                    [ Guaranteed-Bitrate-UL ]
                    [ Guaranteed-Bitrate-DL ]
                    [ Bearer-Identifier ]
                    [ Allocation-Retention-Priority ]
                    [ APN-Aggregate-Max-Bitrate-UL ]
                    [ APN-Aggregate-Max-Bitrate-DL ]
                   *[ Conditional-APN-Aggregate-Max-Bitrate ]
                   *[ AVP ]
```

Figure 5.14 QoS information AVP format

In the PCRF, the QCI field needs to be derived based upon the SPR database or the Service Information obtained from the IMS through Rx interface as shown in section 5.7.3. If the user-specific QoS information is conveyed from the Rx interface, the PCRF can derive the QCI value according to the media type field in the Service Information. If the user-specific QoS information is not conveyed from the Rx interface, the PCRF can use the data from the SPR database to derive the user-specific QoS

parameters as shown in section 5.4. Specifically, since the QCI values 0, 10 – 64, 67-68, and 71 – 255 are reserved for future use as seen in Table 2-1, the basic QCI value (i.e., the QCI value derived when no user specific QoS requirements are considered) plus 100 can be used to denote the user-specific QoS with a lower sensitivity factor, while the basic QCI value plus 200 can be used to denote the user-specific QoS with a higher sensitivity factor. The mapping from the user-specific QoS information to the QCI value for conversational voice and video is shown in Table 5-2.

Table 5-2 Mapping from user-specific QoS to QCI

User-specific QoS	QCI
Voice (low sensitivity factor)	1+100
Voice (normal sensitivity factor)	1
Voice (high sensitivity factor)	1+200
Video (low sensitivity factor)	2+100
Video (normal sensitivity factor)	2

5.7.5 S5/S11 Interface

The Create Bearer Request message shall be sent on the S5 interface by the PDN-GW to the S-GW and on the S11 interface by the S-GW to the MME as part of the Dedicated Bearer Establishment procedure as seen in section 5.2.3 [93]. The Bearer Quality of Service (Bearer QoS) is transferred via GTP tunnels through the Create Bearer Request message. The Bearer QoS IE (Information Element) is shown in Figure 5.15. Because the QCI field has been redefined and added additional user-specific QoS values as described section 5.7.4, the IE doesn't need any further modification except different parsing in the respective protocols. The PDN-GW and S-GW only needs to forward the Bearer QoS information to the subsequent nodes of S-GW and MME respectively.

5.7.6 S1-MME Interface

The E-RAB Setup Request Message is sent by the MME to request the eNodeB to assign resources on Uu and S1 interfaces for one or several E-RABs [92]. The E-RAB Level QoS Parameters

are conveyed in the E-RAB Setup Request Message. Figures 5.16 and 5.17 show the E-RAB Setup Request Message and the E-RAB Level QoS Parameters.

Octets	Bits 8	7	6	5	4	3	2	1
1	Type = 80 (decimal)							
2-3	Length = n							
4	Spare				Instance			
5	Spare	PCI	PL				Spare	PVI
6	Label (QCI)							
7 to 11	Maximum bit rate for uplink							
12 to 16	Maximum bit rate for downlink							
17 to 21	Guaranteed bit rate for uplink							
22 to 26	Guaranteed bit rate for downlink							
27 to (n+4)	These octet(s) is/are present only if explicitly specified							

Figure 5.15 Bearer Quality of Service (Bearer QoS)

IE/Group Name	Presence	Range	IE type and reference	Semantics description	Criticality	Assigned Criticality
Message Type	M		9.2.1.1		YES	reject
MME UE S1AP ID	M		9.2.3.3		YES	reject
eNB UE S1AP ID	M		9.2.3.4		YES	reject
UE Aggregate Maximum Bit Rate	O		9.2.1.20		YES	reject
E-RAB to be Setup List		*1*			YES	reject
>E-RAB To Be Setup Item IEs		*1 to <maxnoof E-RABs>*			EACH	reject
>>E-RAB ID	M		9.2.1.2		-	
>>E-RAB Level QoS Parameters	M		9.2.1.15	Includes necessary QoS parameters	-	
>>Transport Layer Address	M		9.2.2.1		-	
>>GTP-TEID	M		9.2.2.2	EPC TEID	-	
>>NAS-PDU	M		9.2.3.5		-	
>>Correlation ID	O		9.2.2.80		YES	ignore

Figure 5.16 E-RAB Setup Request

Because the QCI has been redefined and added additional user-specific QoS information as described in section 5.7.4. The IE of this message doesn't need any further modification except different parsing in the respective protocols. But the MAC layer can make user of this user-specific QoS information to perform a more advanced resource scheduling, i.e., user-specific QoS aware scheduling.

IE/Group Name	Presence	Range	IE type and reference	Semantics description
E-RAB Level QoS Parameters				
>QCI	M		INTEGER (0..255)	QoS Class Identifier defined in TS 23.401 [11]. Coding specified in TS 23.203 [13].
>Allocation and Retention Priority	M		9.2.1.60	
>GBR QoS Information	O		9.2.1.18	This IE applies to GBR bearers only and shall be ignored otherwise.

Figure 5.17 E-RAB level QoS parameters

5.8 Other Scheduling Details

5.8.1 Rate Adaptation Triggering

The UE can trigger the rate adaptation based upon the ECN indication sent by the eNodeB when the network is congested as described in section 4.4, or the UE can perform the rate adaptation anyway regardless of the ECN indication from the network.

5.8.2 Scheduling Period

We assume the scheduling period of the rate adaptation algorithms is the frame rate of applications, that is, 20 ms for VoIP AMR applications and video frame rate for Video applications. It is necessary to explore what the system capacity gain will be if the scheduling period is increased to acquire the tradeoff for the reduced complexity. In section 5.9, the simulation results are shown when the scheduling period is increased from 20 to 2000 ms for VoIP users.

5.8.3 Integration with Existing Rate Adaptation Algorithms

The rate adaptation in the user-specific QoS aware scheduling can be easily integrated with existing rate adaptation algorithms. Figure 5.18 below illustrates the work flow of such an example for VoIP applications. In this figure, we assume AMR codec modes of M1, M2, M3,… in the decreasing order of AMR codec rates. The work flow starts from the initial AMR codec mode, e.g., M2, performs the USQA rate adaptation, and outputs the degraded AMR codec mode (e.g., M3) for VoIP users. During the VoIP session, if the transport conditions change and trigger the existing rate adaptation algorithm to

degrade or upgrade the AMR codec mode, the work flow will perform the USQA rate adaptation again, starting from the M3 or M1 AMR codec mode depending on whether the existing rate adaptation algorithm wants to degrade or upgrade the AMR codec mode. After that, this loop was completed and next loop will start from the next rate adaptation of existing rate adaptation algorithms.

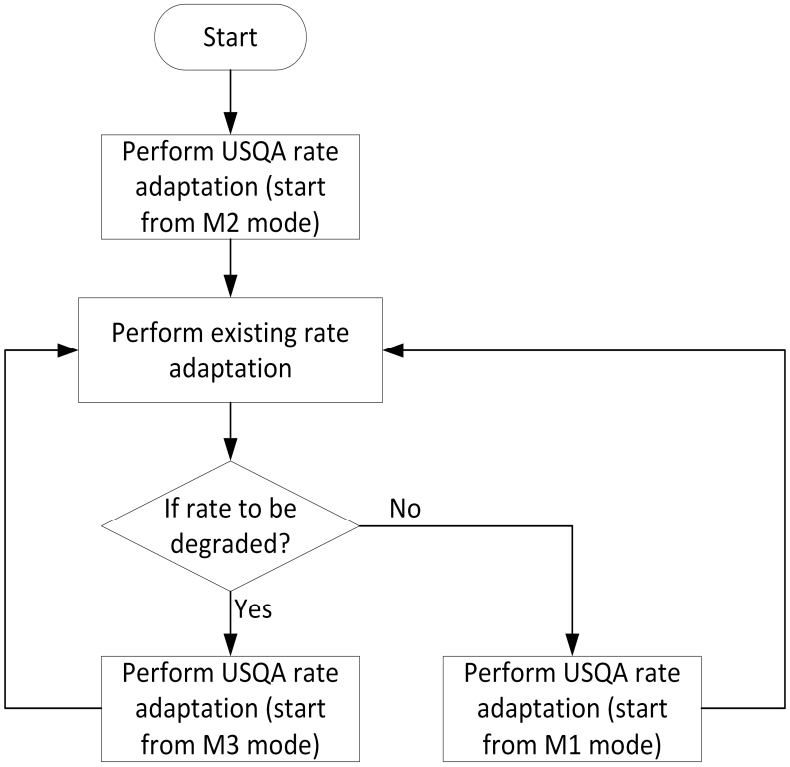

Figure 5.18 Integration with existing rate adaptation algorithms work flow

5.9 Simulation Results

The system simulation configuration is partly based upon LTE macrocell system simulation baseline parameters [87] and Table 5-3.

Fifty four VoIP users with a sensitivity factor of 0.8 are simulated for cases of vehicle speeds of 30 km/h and 60 km/h respectively.

Figure 5.19 shows the VoIP capacity as a function of scheduling period from 20 ms to 2000 ms. From Figure 5.19, we find that as the scheduling period increases, the performance gain will decrease correspondingly. As the scheduling period increases to 1000 ms, the capacity improvement will fall below 10% for the case of 60 km/h, whereas the capacity improvement is still good for the case of 30 km/h.

Table 5-3 System simulation configuration for scheduling period

Parameter	Assumption
Cellular layout	1 macrocell
Cell radius	1kilometer
Path loss model	3GPP suburban macrocell
Mobility model	Random Way Point (30/60 km/h)
Carrier frequency	Uplink:1920MHz Downlink:2110MHz
System bandwidth	10MHz
Channel model	ITU Vehicle A
Total BS TX power	40dBm
UE power	23dBm
VoIP codec modes	AMR12.2K, AMR10.2K, AMR7.95K
Number of users	54 VoIP users
Scheduler	USQA-C scheduler and Baseline scheduler
Other assumptions	Ideal uplink receiver(no block error and packet loss)

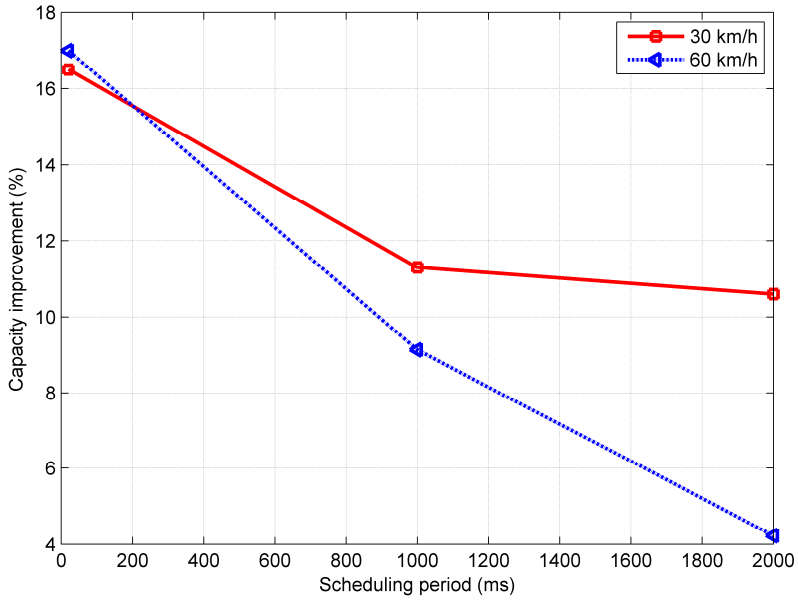

Figure 5.19 VoIP capacity improvement as a function of scheduling period

5.10 Concluding Remarks

In this chapter, we proposed and presented the user-specific QoS aware scheduler implementation details based upon current LTE systems. From the system's perspective, LTE QoS end-to-end signaling

procedures are addressed, and corresponding protocol adaptations are analyzed. It is found that very modest modification of the protocols is needed. To be more specific, when user-specific QoS information does not need to be conveyed from the UE to the network, user-specific QoS parameters can be obtained and derived from the SPR database in the PCRF, and a set of protocols (i.e., the Gx interface, S5/S11 interfaces, and S1-MME interface protocols) needs to be modified, where the QCI field needs to be redefined, added some more user-specific QoS values, and parsed differently in the corresponding protocols (i.e., the Gx interface, S5/S11 interfaces, and S1-MME interface protocols). However, when user-specific QoS information needs to be conveyed from the UE to the network, a set of additional protocols (i.e., the SIP and Rx interface protocols) need to be modified with new user-specific media types added.

Furthermore, the rate adaptation algorithms of the user-specific QoS aware schedulers can be triggered by the ECN indication or triggered regardless of the ECN indication, and they are able to be integrated very easily into currently existing rate adaptation algorithms. Also, it is shown that the scheduling period of rate adaptation algorithms of up to 1000 ms doesn't significantly impact the system performance.

CHAPTER 6. MIMO *IN VIVO* WBAN SYSTEMS

6.1 Introduction

In this chapter[5], we present the performance of MIMO wireless communications for *in vivo* environments by using the ANSYS HFSS human body model to determine the maximum data rates that can be achieved and the optimal placement of antennas using an OFDM-based IEEE 802.11n system architecture. In our performance optimization we consider various factors such as antenna separation distances, angular positions, human body size, and system bandwidth. We analyzed the performance of MIMO *in vivo* in terms of FER and system capacity statistically with one pair of antennas placed *in vivo* and the second pair placed inside and outside the body at various distances from the *in vivo* antennas and at various angular positions around the human body. The results were compared with a SISO *in vivo* system and we demonstrated that by using MIMO *in vivo*, significant performance gains may be achieved, with maximum SAR levels met, making it possible to achieve target data rates of 100 Mbps with a system bandwidth of 40 MHz when receiver antennas are located at the front or back surface of the body. Simulating the effect of different human body size by scaling the HFSS human body model, it was shown that varying human body size has a great impact on MIMO *in vivo* performance.

The rest of the chapter is organized as follows. We describe MIMO *in vivo* communications in section 6.2. In section 6.3, we present the MIMO *in vivo* capacity formulas based upon the IEEE 802.11n system. In sections 6.4 and 6.5 we present the simulation setup and results of the performance of MIMO *in vivo*, respectively. Finally, in section 6.6, we present our conclusions.

6.2 MIMO *In Vivo* Approach

Some *in vivo* sensor nodes may require high communication bandwidths such as the *MARVEL* platform that, with its camera module (CM), wirelessly transmits high definition (HD) video during

[5] Portions of this chapter were previously published in [101], [107]. Permission is included in Appendix A.

surgery [101], [102]. A MIMO *in vivo* system may be required to achieve high data rates using low transmission power to comply with the SAR requirements. SAR is the specific absorption rate of power absorption by human organs and is limited by the FCC (Federal Communications Commission), which in turn limits the transmission power [103].

The wireless *in vivo* channel is an exciting and challenging new environment that has not been well documented in the literature beyond limited analyses of signal attenuation and shadowing of human tissues limited to the Medical Implant Communication Service (MICS) frequency band (from 402 to 405 MHz) [54]. Moreover, the current studies use the Friis formula to calculate the path loss [104]. Note that the Friis formula is only valid for the far field and free space. The IEEE P802.15 TG6 WBANs channel model [54] provides guidance as to how the channel model should be developed for body area networks. For *in vivo* communications, however this document is limited and only provides the path loss exponent. In addition, three types of nodes and several communications links (scenarios) are defined as shown in Figure 6.1.

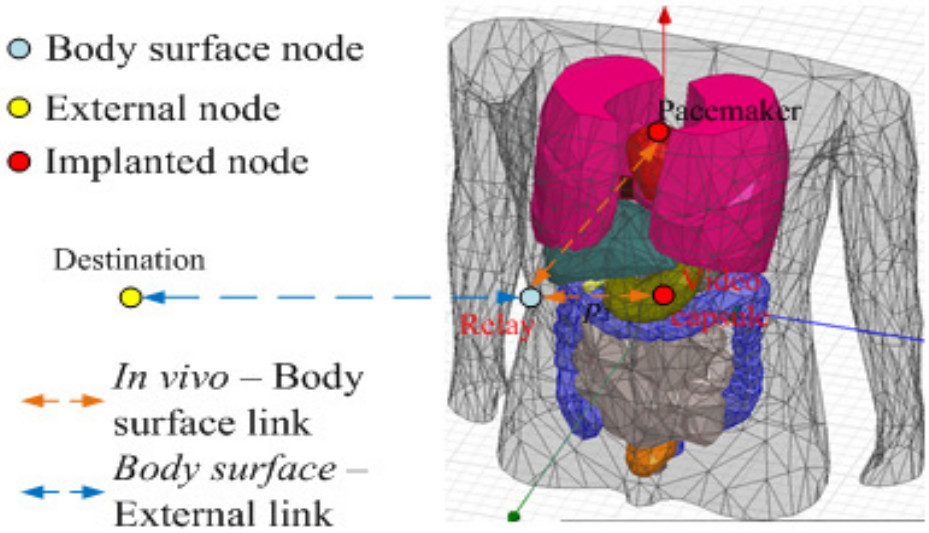

Figure 6.1 Communication links for a Wireless Body Area Network

In the far-field electromagnetic waves behave as plane waves and the total radiated power does not change with changing radial distance from the antenna since the distance between transmitter and receiver is large and free space is considered (relatively) lossless, as in cellular networks. However, since the *in vivo*

antennas are radiating into a complex lossy medium, the radiating near field will strongly couple to the lossy environment [105]. This means the radiated power is strongly dependent on the radial and angular position and the near field effects may have to be taken into account when operating in the *in vivo* environment. In the radiating near field, the electric and magnetic fields behave differently compared to the far field. Hence, the wireless channel inside the body requires different link equations [106]. Additionally, since the wavelength of the signal is much longer than the propagation environment in the near field, the delay spread concept and multi-path scattering of cellular network is not directly applicable to (near-field) channels inside the body. This will directly affect the correlation between the signals at the antennas, which is critically important for MIMO performance. Considering the near-field effect, the impact of various factors needs to be explored including antenna position and correlation, human body size, etc.

The achievable transmission rates in the *in vivo* environment are simulated using a model based on the IEEE 802.11n standard because this OFDM-based standard supports up to 4 spatial streams (4x4 MIMO). Because of the form factor constraint inside the human body, our current study is restricted to 2x2 MIMO. Moreover, the standard allows different Modulation and Coding Schemes (MCS) that are represented by a MCS index value and uses either a 20 MHz or 40 MHz bandwidth. Due to the target data rates for the *MARVEL* CM (~80–100 Mbps), the MCS index values of interest for MIS HD video applications are 13 and up for 20 MHz channels and 10 and up for 40 MHz channels.

The MIMO *in vivo* system capacity is the upper theoretical performance limit that can be achieved in practical systems, and can provide insight into how well the system can perform theoretically and give guidance on how to optimize the MIMO *in vivo* system. The system capacity for both MIMO and SISO *in vivo* can be calculated theoretically based upon the MIMO and SISO *in vivo* channels, as described in the next section.

6.3 MIMO *In Vivo* Capacity

6.3.1 MIMO *In Vivo* Capacity

Assuming two transmitter and receiver antennas are used in the MIMO *in vivo* system. The OFDM system can be modeled as [107]–[109]:

$$Y_k = H_k X_k + W_k, \quad k = 1,2,\ldots,N_{data} \tag{6-1}$$

where $Y_k, X_k, W_k \in C^2$ denote the received signal, transmitted signal, and white Gaussian noise with power density of N_0 respectively at subcarrier k. The symbol N_{data} is the total number of subcarriers configured in the system to carry data. The complex frequency channel response matrix at subcarrier k is denoted by $H_k \in C^{2*2}$.

The SVD (Singular Value Decomposition) of H_k is given as:

$$H_k = U_k \Lambda_k V_k \tag{6-2}$$

where $U_k, V_k \in C^{2*2}$ are unitary matrices, and Λ_k is the nonnegative diagonal matrix whose diagonal elements are singular values of $\sqrt{\lambda_{k1}}$, $\sqrt{\lambda_{k2}}$ respectively.

The system capacity for subcarrier k is:

$$C_k = E\left[\sum_{i=1}^{2} log_2\left(1 + \frac{\lambda_{ki} P}{2N_0 \cdot BW}\right)\right] \quad \text{bits/OFDM symbol} \tag{6-3}$$

where P is the total transmit signal power of the two transmitter antennas, BW is the configured system bandwidth in Hz, and E denotes expectation. In this dissertation, we consider only time invariant Gaussian channels, so we will ignore the expectation in the capacity calculation. The total system capacity is calculated as:

$$C = \frac{1}{T_{sym}} \sum_{k=1}^{N_{data}} C_k \quad \text{bits/s} \tag{6-4}$$

where T_{sym} is the duration of each OFDM symbol.

6.3.2 SISO *In Vivo* Capacity

For a performance comparison with MIMO *in vivo*, the SISO *in vivo* capacity is also calculated. The SISO system model is the same as defined in (6-1) except for the terms $Y_k, X_k, W_k \in C^1$.

The system capacity for SISO *in vivo* is:

$$C = \frac{1}{T_{sym}} E\left[\sum_{k=1}^{N_{data}} log_2\left(1 + \frac{H_k P}{N_0 \cdot BW}\right)\right] \quad \text{bits/s} \tag{6-5}$$

where $H_k \in C^1$, P, N_{data} and E mean the same as those for MIMO *in vivo*.

6.3.3 SNR and Bandwidth

For a 40 MHz system bandwidth, to maintain the similar SAR level, the power for each 20 MHz carrier should be half of that for a 20 MHz system bandwidth. The white noise power will also double due to the larger system bandwidth of 40 MHz. Hence the SNR (Signal to Noise Ratio) for a 20 MHz system bandwidth will be four times as high as that for a 40 MHz system bandwidth.

As indicated in (6-3) and (6-4), system capacity is basically a logarithm function, and depends upon the factors of both SNR (i.e., $\frac{\lambda_{ki}P}{2N_0 \cdot BW}$) and system bandwidth (i.e., BW). Since the logarithm function is a concave functions, it has the following two particular properties [108]:

$$log_2(1 + SNR) \approx SNR log_2 e \text{ when } SNR \approx 0 \quad (6\text{-}6)$$

$$log_2(1 + SNR) \approx log_2 SNR \text{ when } SNR \gg 1 \quad (6\text{-}7)$$

Based upon the properties in (6-6) and (6-7), from (6-3) and (6-4), when the SNR is low, the system capacity is proportional with SNR, so that the SNR is the dominant factor in determining the system capacity and the system capacity for a 20 MHz system bandwidth may be higher than that for a 40 MHz system bandwidth. When the SNR is high, the capacity is logarithmically proportional to the SNR, so that the system bandwidth is the dominant factor in determining the system capacity, and the system capacity for a 40 MHz system bandwidth will generally be higher, but not always, than that for a 20 MHz system bandwidth.

Therefore, as the system bandwidth doubles from 20 MHz to 40 MHz, depending upon different application scenarios, the resulting system capacity will not necessarily increase, as verified by the simulation result in section 6.5.

6.4 MIMO *In Vivo* Simulation Setup

6.4.1 Human Body Model

The simulations for the electromagnetic wave propagation were performed in ANSYS HFSS 15.0.3 using the ANSYS Human Body Model [110]. The model consists of a detailed adult male with over 300 muscles, organs, and bones with a geometrical accuracy of 1 mm and realistic frequency dependent

material parameters (conductivity and permittivity) from 20 Hz to 20 GHz. The antennas used in the simulations were monopoles designed to operate at the 2.4 GHz ISM band in their respective medium; free space for the *ex vivo* antennas and inside the body for *in vivo* antennas. We choose monopoles due to their smaller size, simplicity in design and omni-directionality [111]. For the *in vivo* case, the monopole's performance and radiation pattern will vary with position and orientation inside the body [30], [105], making the performance of MIMO *in vivo* strongly dependent on the antenna placement.

6.4.2 Simulation Cases

As shown in Figure 6.2, two *in vivo* antennas are placed inside the abdomen to simulate placement of transceivers in laparoscopic and intestinal medical applications. The two *ex vivo* antennas are placed at varying locations around the body at the same planar height as the *in vivo* antennas. The Transmitter (Tx) antennas are *in vivo* antennas, whereas the receiver (Rx) antennas can be either *in vivo* antennas or *ex vivo* antennas depending upon their positions and distances from the *in vivo* Tx antennas. The locations with respect to the *in vivo* antennas are given in Table 6-1 and Table 6-2. For the performance comparison with SISO *in vivo* systems, the locations of the Rx antennas for SISO cases are also listed in Table 6-1, whereas for Table 6-2, no SISO cases are listed for comparison because in these cases only MIMO performance are of concern.

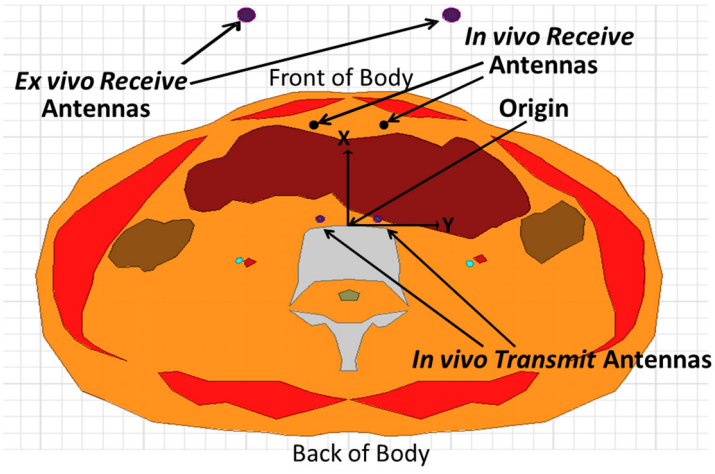

Figure 6.2 Antenna simulation setup showing locations of the MIMO antennas

It should be noted, since the permittivity of the body is much higher than that of free space, the wavelength is smaller inside the body and varies as it passes through various tissues and organs. On average, the wavelength is reduced by the square root of the dielectric constant and is approximately six times smaller *in vivo* than in free space. To guarantee the greatest MIMO spatial multiplexing and diversity gain, the *ex vivo* and *in vivo* antennas are separated by at least half of their respective *in vivo* and *ex vivo* wavelengths. The two Tx antennas are located 1.4 cm from either side of the origin along the X-axis or Y-axis, whereas in the SISO cases, the single Tx *in vivo* antenna is located at the origin. The two Rx antennas are located 5 cm from either side of the origin along the X-axis or Y-axis.

Simulation cases in Table 6-1 are used to evaluate the system performance for MIMO *in vivo* vs SISO *in vivo* in terms of both FER and system capacity under different Tx and Rx distances and angular positions. Cases 1-7 in Table 6-1 are cases with the Rx antennas placed in front of the body with the same angular positions, but with varying distances between Tx and Rx antennas. See section 6.5.1 for the simulation results for cases 1-7. Cases 7-10 in Table 6-1 are cases with the same distance between Tx and Rx antennas, but with different angular positions, which correspond to the front, right side, left side, and back of body, respectively. See section 6.5.2 for the simulation results for cases 7-10.

Table 6-1 Simulation cases for MIMO *in vivo* vs SISO *in vivo* with locations of antennas with respect to the origin (X=0, Y=0) shown in Figure 6.2

Cases	MIMO				SISO				Notes
	Receiver antennas		Transmitter antennas		Receiver antenna		Transmitter antenna		
	X (cm)	Y (cm)	X (cm)	Y (cm)	X (cm)	Y (cm)	X (cm)	Y (cm)	
1	7	±5	0	±1.4	7	0	0	0	Front of body
2	8.5	±5	0	±1.4	8.5	0	0	0	Front of body
3	10	±5	0	±1.4	10	0	0	0	Front of body
4	11	±5	0	±1.4	11	0	0	0	Front of body
5	13	±5	0	±1.4	13	0	0	0	Front of body
6	20	±5	0	±1.4	20	0	0	0	Front of body
7	30	±5	0	±1.4	30	0	0	0	Front of body
8	±5	30	±1.4	0	0	30	0	0	Right side of body
9	±5	-30	±1.4	0	0	-30	0	0	Left side of body
10	-30	±5	0	±1.4	-30	0	0	0	Back of body

Simulation cases in Table 6-2 are used to evaluate the system performance in terms of system capacity for on body MIMO *in vivo* performance and effect of body size on MIMO *in vivo* performances. To simulate the effect of different body size on MIMO *in vivo* performance, the HFSS human body model is uniformly scaled by a scaling size of 1.1 and 0.9 respectively. Cases 1-4 in Table 6-2 are cases where the Rx antennas are located on the surface of the body at four angular positions (i.e., front, right, left, and back). For a typical human body, the distance from the Tx antennas to the Rx antennas on the surface of front body is 11 cm, 9.6 cm to the surface of back body, and 15 cm to the surface of right and left body. See section 6.5.4 for the simulation results for cases 1-4. Cases 1, 5-6 in Table 6-2 are cases where the Rx antennas are located on the surface of the front body, but the distances of Tx and Rx antennas differ due to the different scaling factors of human body size. Cases 5-6 in Table 6-2 are cases where the scaling factor of human body size is 1.1 and 0.9 respectively so that the distance from the Tx antennas to the surface of front body is 1.1 x 11 cm and 0.9 x 11 cm respectively. See section 6.5.5 for the simulation results for cases 1, 5-6.

Table 6-2 Simulation cases for on body MIMO *in vivo* and effect of body size on MIMO *in vivo* with locations of antennas with respect to the origin (X=0, Y=0) shown in Figure 6.2

Cases	MIMO				Notes
	Receiver antennas		Transmitter antennas		
	X (cm)	Y (cm)	X (cm)	Y (cm)	
1	11	±5	0	±1.4	Front on body
2	±5	15	0	±1.4	Right side on body
3	±5	-15	0	±1.4	Left side on body
4	-9.6	±5	0	±1.4	Back on body
5	1.1x11	±5	0	±1.4	Front on body with a human body scaling factor of 1.1
6	0.9x11	±5	0	±1.4	Front on body with a human body scaling factor of 0.9

6.4.3 System Level Setup

The system capacity analysis and FER performance in the *in vivo* environment have been performed based on the IEEE 802.11n standard [112] transceiver. Agilent SystemVue [113] is used to simulate the FER performance and the system block diagram is shown in Figure 6.3. Because of the form factor constraint inside the human body, our initial study is restricted to 2x2 MIMO. System bandwidths

of 20 MHz and 40 MHz are used in the evaluation. The 802.11n standard supports different MCS represented by a MCS index. The transmission power is set to be 0.412 mW [11] for a 20 MHz system bandwidth, which gives the maximum local SAR level of 1.48 W/kg that will not exceed the maximum allowable SAR level of 1.6 W/kg [103]. The thermal noise power is set to -101 dBm for a 20 MHz system bandwidth and -98 dBm for a 40 MHz system bandwidth. Hence, in the system capacity analysis, the parameters in (6-3)-(6-5) are determined for a 20 MHz bandwidth as follows:

$$P = 0.412\,mW,\ N_0 = -174\,dBm,\ BW = 20\,MHz,\ N_{data} = 52,\ T_{sym} = 4\,us.$$

For a 40 MHz bandwidth, to maintain the similar SAR level to meet the maximum local SAR level of 1.6 W/kg, the power for each 20MHz carrier is one half of that for 20MHz bandwidth, that is, 0.206 mW. Correspondingly, the parameters in (6-3)-(6-5) are determined for a 40 MHz bandwidth as follows:

$$P = 0.206\,mW,\ N_0 = -174\,dBm,\ BW = 40\,MHz,\ N_{data} = 104,\ T_{sym} = 4\,us.$$

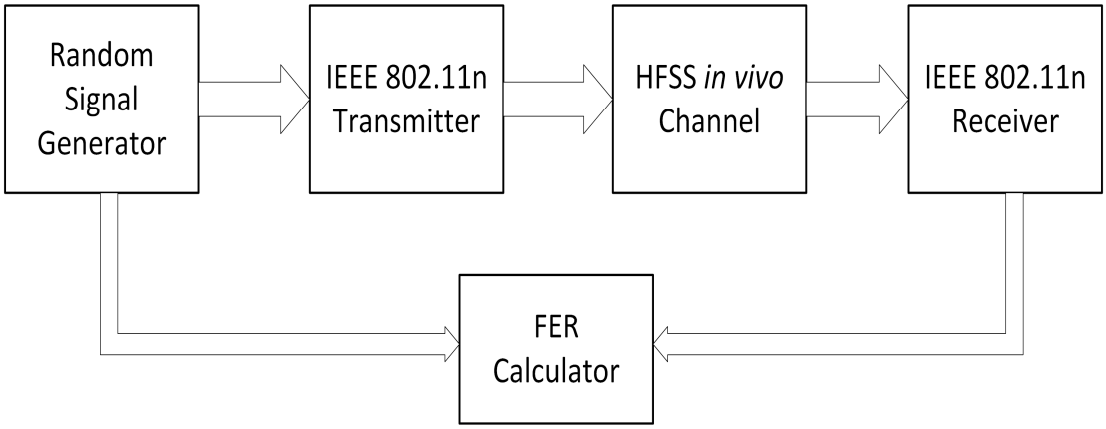

Figure 6.3 Block diagram of system level simulation with HFSS *in vivo* channel model

6.5 MIMO *In Vivo* Simulation Results

Simulations results have been obtained for a 2x2 MIMO and SISO setup with antennas operating at 2.4 GHz. From HFSS and the Human Body Model, the S-parameters between Tx and Rx antennas were extracted between 1 and 3 GHz [110]. Then, the FER for the IEEE 802.11n system was obtained by running 100 000 frames for each simulation for different MCS index values, for 20 MHz and 40 MHz, for

a 800 ns guard interval, and different frame lengths. The system capacity for both MIMO and SISO *in vivo* can be calculated based upon (6-2)-(6-5).

6.5.1 MIMO vs SISO *In Vivo*

Figure 6.4 shows the FER as a function of the MCS index value for both MIMO and SISO *in vivo* cases where Tx and Rx antennas are separated by 7 cm, 10 cm, and 13 cm respectively. Taking the case where the Rx antennas are located at 10 cm from the Tx antennas as the example, as observed in Figure 6.4, the data rate of 78 Mbps at MCS index 12 can be supported by MIMO *in vivo* for the target FER below 10^{-3}, while a data rate of only 19.5 Mbps at MCS index 2 can be supported by SISO *in vivo* with the same distance between antennas as MIMO *in vivo* case. We also observed that as the Tx and Rx antennas distance becomes smaller, the performance gain becomes even bigger.

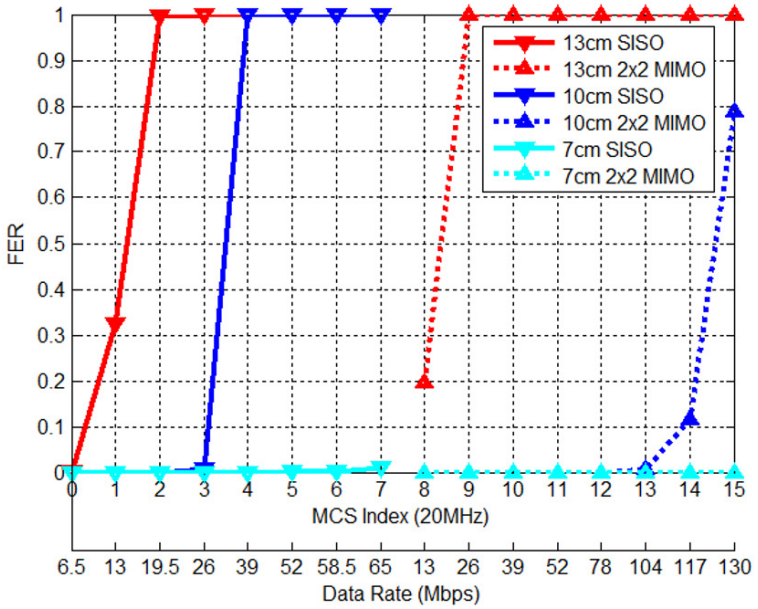

Figure 6.4 MIMO (2x2) and SISO *in vivo* FER performance comparison as a function of the MCS index value (20MHz channel)

Figure 6.5 shows the FER as a function of distance between the Tx and Rx antennas at different MCS index values equal to 12 and 13 corresponding to 78 and 104 Mbps respectively, which also correspond to our MIS data rate requirements of ~80-100 Mbps. When transmitting data at MCS equal to 13 (i.e., 104 Mbps), the Rx antennas need to be placed within a distance of less than 10 cm from the Tx antennas to achieve a minimum FER of 10^{-3} and to meet the data rate requirement of at least 100 Mbps.

Therefore, for our MIS application, it is possible to transmit high definition video with low latency only from deep inside the human body during MIS if MIMO *in vivo* is used.

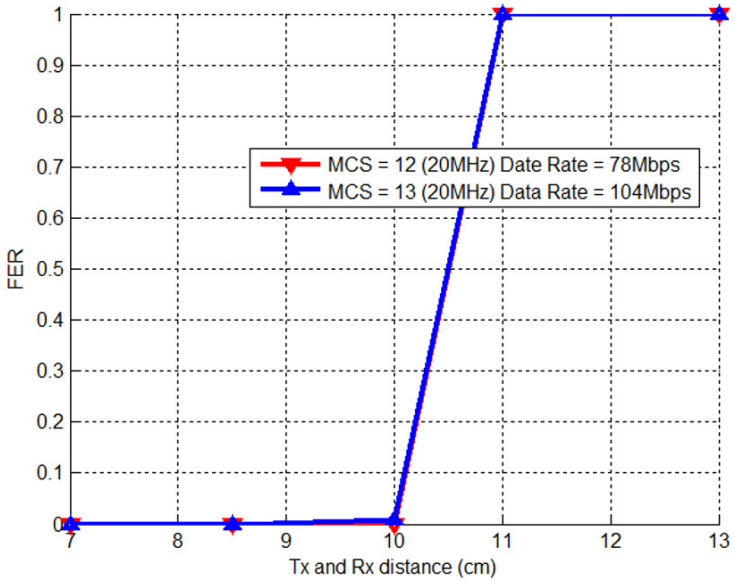

Figure 6.5 MIMO (2x2) *in vivo* FER performance comparison as a function of the distance for different MCS indexes (20MHz channel)

Figure 6.6 shows the system capacity for the cases of Rx antennas placed in front of the body with varying distances between Tx and Rx antennas. It can be seen that much less capacity will be achieved with increased distances. To support the required data rate of 100 Mbps, the distance cannot be greater than ~11cm. The system capacity decreases rapidly when the distance becomes greater, making necessary a larger system bandwidth or relay (or other forms of cooperative networked communications) and placing the receiver antennas as close to, or on, the surface of the body, in the WBAN network.

6.5.2 MIMO *In Vivo* Performance under Different Angular Positions

Figure 6.7 shows the system capacity for different angular positions around the human body with the same distance between Tx and Rx antennas of 30 cm. From Figure 6.7, we can observe the significant capacity gain compared with corresponding SISO cases, where the greatest capacity gain can be seen in the case of both MIMO antennas at the back of the body. We can also see from Figure 6.7 that the system capacity of MIMO *in vivo* for the cases of front and back body is much better than that of the other two cases of the side of the body. This is because much higher attenuation exists inside the body due to the

greater *in vivo* distance for the two cases of the side of the body. Furthermore, compared with the other three cases, MIMO *in vivo* for the back of the body case performs the best. From Figure 6.7, we see that with the greater distance of 30 cm between Tx and Rx antennas, the system capacity will fall below 25 Mbps for whatever angular positions the receiver antennas are located. Hence, for a 20 MHz system bandwidth, only a maximum data rate of less than 25 Mbps can be supported, which is a motivation to use a relay or other forms of cooperative networked communications and place the receiver antennas as close to, or on, the front or back of the body to support a data rate as high as 100 Mbps.

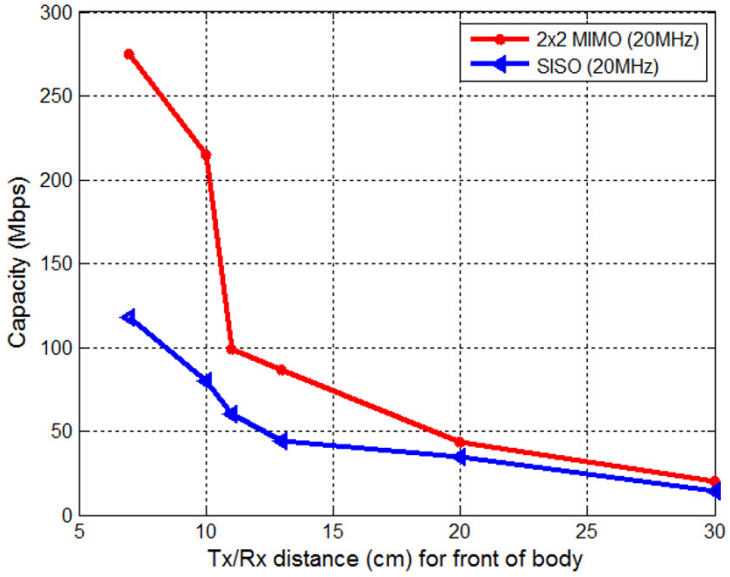

Figure 6.6 MIMO (2x2) and SISO *in vivo* capacity comparison as a function of the distance of the Tx and Rx antennas in front of the body (20MHz channel)

6.5.3 MIMO *In Vivo* Performance with a Larger Bandwidth

To support the required data rate as high as 100 Mbps, the Tx and Rx needs to be as low as 10-11 cm when Rx antennas are located at the front of the body. Since for a human body of average size, the distance from the Tx antennas to the front on body is at least 11 cm, even when relay of other forms of cooperative networked communications is used to improve the required minimum communication distance, the communication from *in vivo* to the surface of the body needs to be more reliable. In this section, we explore the possibility of improving MIMO *in vivo* capacity by considering a larger system bandwidth of 40 MHz [115].

Corresponding to Figures 6.4, 6.5, 6.6, and 6.7 for a 20 MHz bandwidth, Figures 6.8, 6.9, 6.10, and 6.11 show the simulation results respectively for a 40 MHz bandwidth. As discussed in section 6.3.3, the system capacity depends upon both SNR [which is influenced by the SAR limit] and bandwidth, and the system capacity for a 40 MHz system bandwidth will not necessarily be higher than that for a 20 MHz system bandwidth.

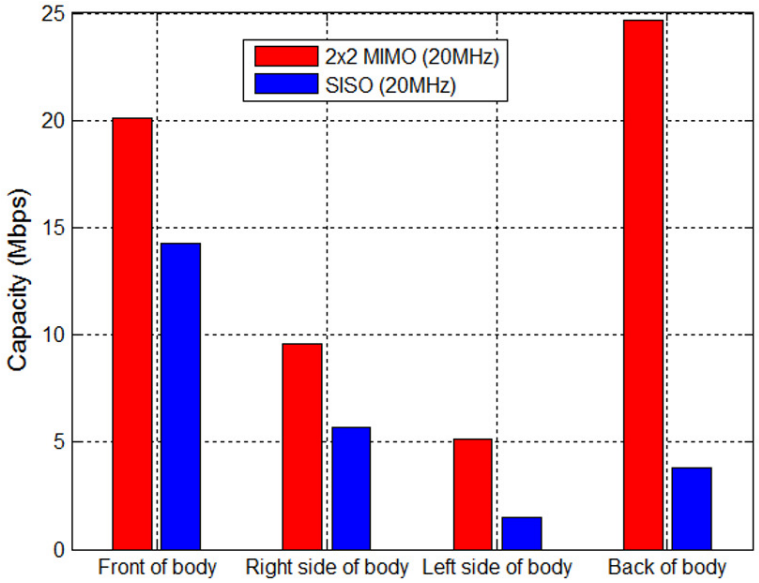

Figure 6.7 MIMO (2x2) and SISO *in vivo* system capacity comparison for front, right side, left side, and back of the body (20MHz channel)

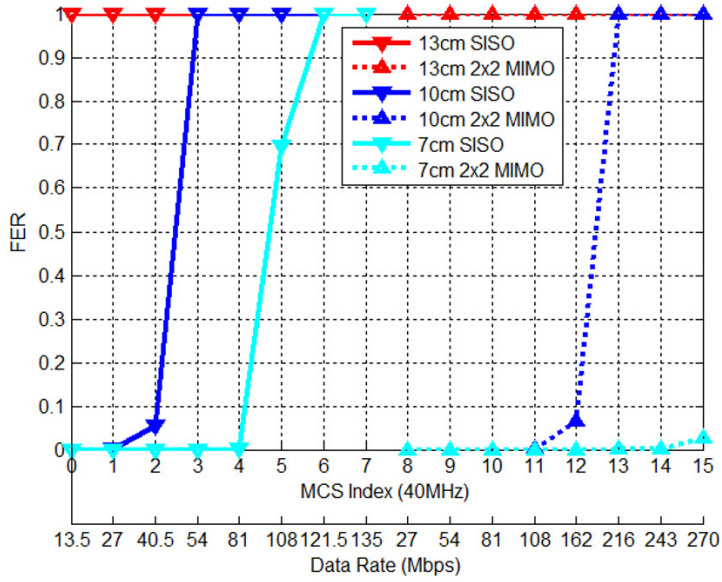

Figure 6.8 MIMO (2x2) and SISO *in vivo* FER performance comparison as a function of the MCS index value (40MHz channel)

Figure 6.8 shows the FER as a function of the MCS index value for both MIMO and SISO *in vivo* cases. To take the case where the Rx antennas are located at 10 cm from the Tx antennas as the example, as observed in Figure 6.8, the data rate of 108 Mbps at MCS index 11 can be supported by MIMO *in vivo* for the target FER below 10^{-3}, while only a data rate of 27 Mbps at MCS index 1 can be supported by SISO *in vivo* with the same distance between antennas as MIMO *in vivo* case. We also discovered that as the Tx and Rx antennas distance becomes smaller, the performance gain becomes even bigger.

Figure 6.9 shows the FER as a function of distance between the Tx and Rx antennas at different MCS index values equal to 10 and 11 corresponding to 82 and 108 Mbps respectively, which also correspond to our application data rate requirements of ~80-100 Mbps. In the case when transmitting data at MCS 11 (108 Mbps), the Rx antennas can be placed within a distance of 10 cm from the Tx antennas to achieve a minimum FER of 10^{-3} and meet the requirement of at least 100 Mbps, which is an improvement from the 20 MHz case.

Figure 6.10 shows the system capacity for the cases of Rx antennas placed in front of the body with varying distances between Tx and Rx antennas. To support the required data rate of 100 Mbps, the distance cannot be greater than ~13 cm, which is an improvement from ~11 cm for the 20 MHz case.

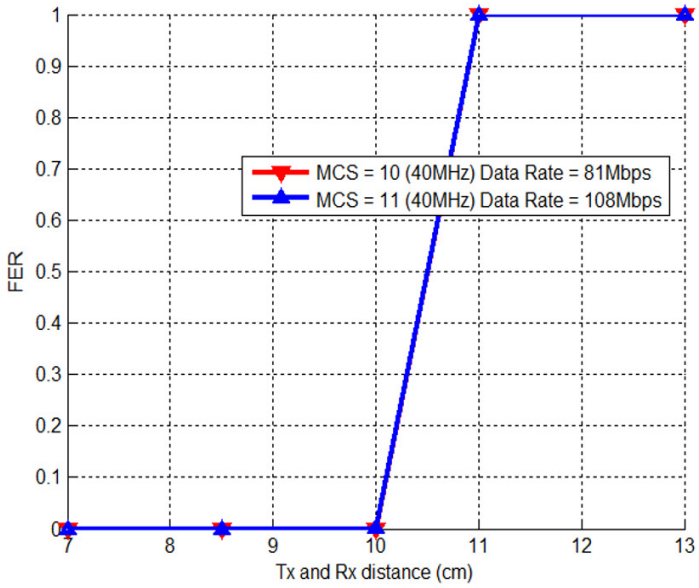

Figure 6.9 MIMO (2x2) *in vivo* FER performance comparison as a function of the distance for different MCS indexes (40MHz channel)

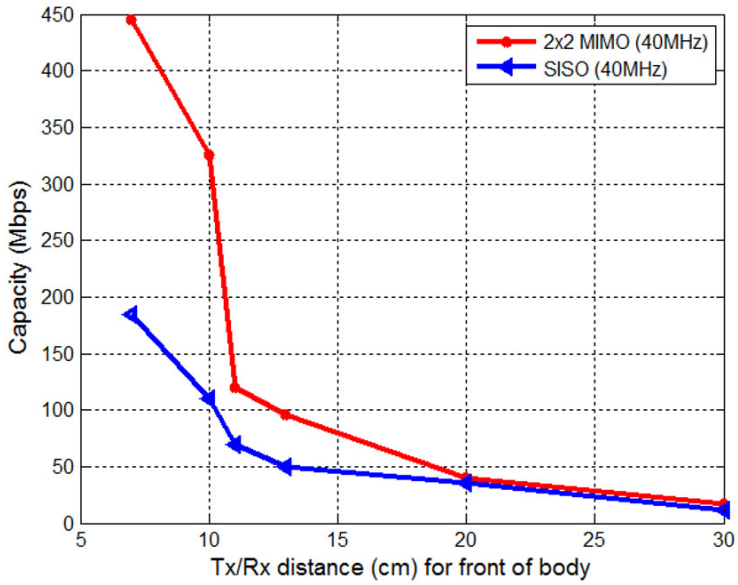

Figure 6.10 MIMO (2x2) and SISO *in vivo* capacity comparison as a function of the distance of the Tx and Rx antennas in front of the body (40MHz channel)

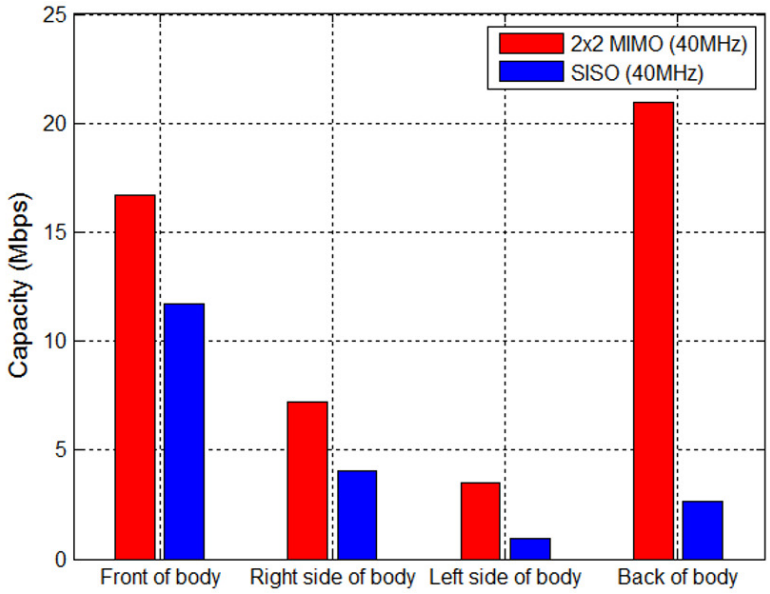

Figure 6.11 MIMO (2x2) and SISO *in vivo* system capacity comparison for front, right side, left side, and back of the body (40MHz channel).

Figure 6.11 shows the system capacity for different angular positions around the human body with the same distance of 30 cm. Compared with that of the 20 MHz case, contrarily the system capacity deceases, that is because transmitting power is reduced by half for each 20 MHz carrier and noise power doubles for a 40 MHz bandwidth, the SNR is very small (i.e., due to larger distance) and dominates the

system capacity more than the system bandwidth. For the cases of the front and the back of body the system performs much better due to the same reason as that of the 20 MHz case.

6.5.4 On Body MIMO *In Vivo* Performance

Compared with the MIMO performance in free space, the *in vivo* system performance constitutes more of a challenge to the system performance, which limits the maximum communication distance as described in sections 6.5.1-6.5.3. Therefore, it is necessary that the receiver antennas are placed as close as possible to the surface of the body or on the surface of the body, where we can use relay to communication from the surface of body to external nodes.

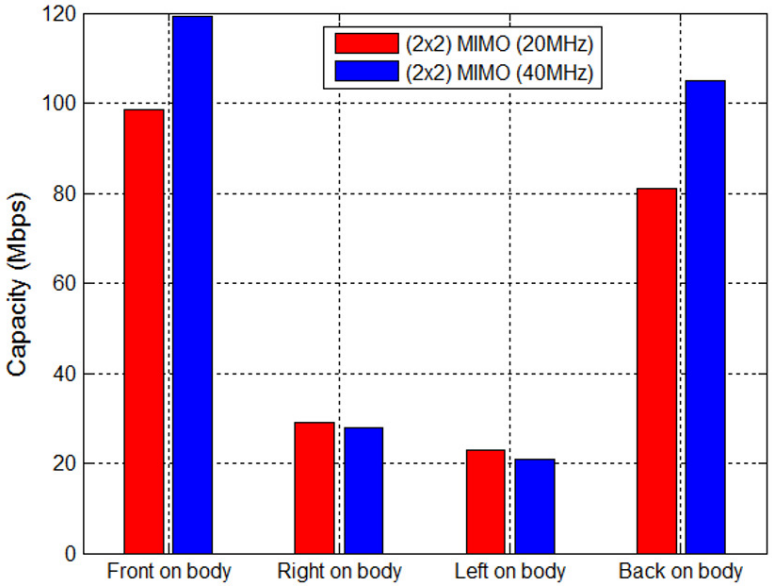

Figure 6.12 MIMO (2x2) *in vivo* system capacity comparison for front, right, left, and back on body

This section compares on body MIMO *in vivo* performance for different angular positions. Figure 6.12 shows the simulation result for on body MIMO *in vivo* performance for the corresponding cases 1-4 in Table 6-2 (i.e., front on body, right on body, left on body, and back on body) for both the 20 MHz and 40 MHz bandwidths. From Figure 6.12 we find the front and back body cases perform best for both the 20 MHz (i.e., >80 Mbps) and 40 MHz (i.e., >100 Mbps) bandwidth. This is because much higher attenuation exists inside the body due to the greater *in vivo* distance for the two cases of side on body.

Hence we can conclude that the placement of the receiver antennas on the front and back on body is optimal and sufficient to support the data rate of 100 Mbps under the larger system bandwidth of 40 MHz.

6.5.5 Effect of Human Body Size on MIMO *In Vivo* Performance

Figure 6.13 shows the performance comparison for different human body size. A default body size in the HFSS human body model is an adult male with a height of 183 cm. Scaling factors of 1.1 and 0.9 are added to the default adult male model to simulate the effect of body size on the MIMO *in vivo* performance. The Rx antennas are placed on the surface of the front body. From the figure, we find a significant increase can be achieved for the case of the scaling factor of 0.9 for both 20 MHz and 40 MHz bandwidths, while a substantial decrease is incurred for the case of the scaling factor of 1.1 for both bandwidths, where even for a 40 MHz bandwidth, the maximum system capacity can only reach as low as 50 Mbps, which is far lower than the application data rate requirement of 100 Mbps.

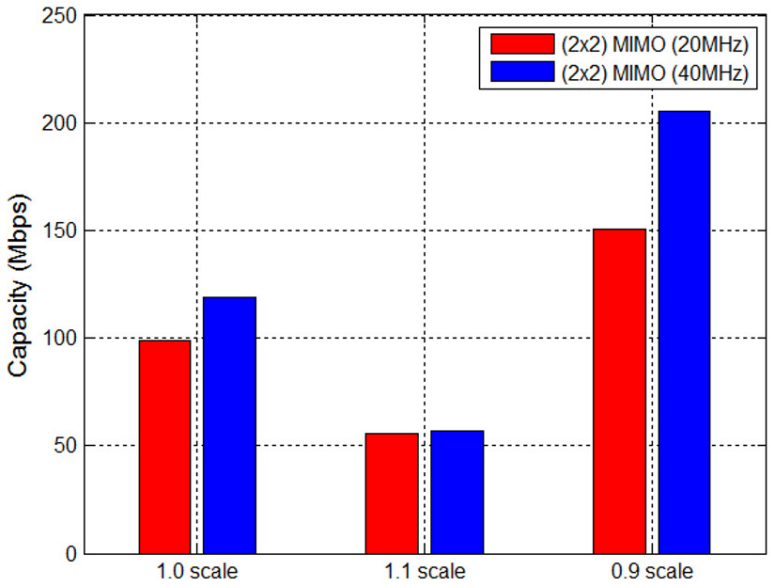

Figure 6.13 MIMO (2x2) *in vivo* system capacity comparison for different body size

6.6 Concluding Remarks

In this chapter, we present an extensive study of MIMO for *in vivo* environments to determine the maximum data rate that can be achieved in this challenging environment considering various factors. We simulated a MIMO OFDM-based system that complies with the IEEE 802.11n standard. The simulations

for the *in vivo* channel were obtained from the HFSS software that includes the Human Body Model. From the simulation result found in this study, the following conclusions can be drawn: 1) As expected, MIMO *in vivo* can achieve significant performance gains compared with SISO *in vivo*, and it is possible to achieve higher data rates when Rx antennas are moved closer to the surface of the body; 2) Significantly better system capacity can be observed when receiver antennas are paced at the front or the back of the body than when placed at the either side of the body; 3) An increased system bandwidth can help increase MIMO *in vivo* performance only when Tx and Rx distances are not great and SNR in the receiver is relatively good; and 4) Human body size can significantly impact MIMO *in vivo* performance.

CHAPTER 7. CONCLUSION AND FUTURE DIRECTIONS

In this dissertation we developed and analyzed the innovative concept of user-specific QoS requirements for cellular systems and described its realization via novel user-specific QoS-aware schedulers to improve the MOS or the system capacity, or both the MOS and system capacity. Part of the research on the user-specific QoS aware scheduling study in this dissertation has been published in [70]–[72], [98].

As an independent research, in this dissertation, we also presented the performance of MIMO *in vivo* technology for WBAN systems, using the ANSYS HFSS and their complete human body software model and an OFDM-based IEEE 802.11n system, to determine the maximum data rates (i.e., the capacity) that can be achieved and the optimal placement of antennas in a wide variety of circumstances. Part of the research on the MIMO *in vivo* study in this dissertation has been published [101], [107] or submitted in [115].

The main contributions and conclusions for each chapter are summarized and described in the following section.

7.1 Main Contributions and Conclusions

7.1.1 User-Specific QoS Requirements

User-Specific QoS requirements are characterized by their user-specific MOS formulas for VoIP, video, and FTP applications. These user-specific MOS formulas are based upon existing MOS formulas, and the introduction of novel user-specific sensitivity factors to the packet loss, data rate, or delay.

Two methods are proposed to improve the system capacity considering user-specific QoS requirements: a user-specific MCS selection method and user-specific QoS aware scheduling method.

7.1.2 User-Specific QoS Aware Scheduler

We proposed the user-specific QoS aware schedulers that, depending upon the system performance targets, can flexibly improve significantly the MOS (the USQA-M1, USQA-M2, and USQA-F schedulers) or system capacity (the USQA-C scheduler), or both the MOS and system capacity (the USQA-MC scheduler) when such user-specific QoS requirements are considered. The improvement of system capacity results from dynamic source data rate adaptation to the user-specific QoS requirements and network conditions. The MOS improvement results from the MAC scheduling differentiation between users with different QoS requirements.

The user-specific QoS aware schedulers used in heterogeneous scenarios were also first analyzed. The analysis and system performance evaluation show that the system gain can be mostly or partly maintained in heterogeneous networks depending upon different heterogeneous scenarios.

The novel user-specific frequency sensitivity QoS study where users from different age group are more or less sensitive to the high frequency contents was analyzed and evaluated. It was shown that when VoIP codecs are designed to be specifically matched to the different high frequency auditory characteristics of users, higher system capacity as well as a comparable MOS levels may be achieved.

7.1.3 User-Specific QoS Aware Scheduler Implementation

Based upon LTE systems operation, the user-specific QoS parameters can be derived from the SPR database in the PCRF or signaled dynamically through the SIP protocol and Rx interface protocol.

From the whole LTE system perspective, the user-specific QoS end-to-end signaling procedures were addressed, and corresponding protocol adaptations are analyzed. The potential protocols to be adapted include the SIP protocol, Rx interface, Gx interface, S5/S11 interfaces, and S1-MME interface protocols. Some new user-specific QCI values are needed to be defined and parsed differently by the Gx interface, S5/S11 interfaces, and S1-MME interface protocols. In addition, some new user-specific media type values need to be defined, signaled and parsed differently through the SIP protocol and Rx interface protocol if the user-specific QoS parameters are signaled dynamically.

The scheduling period and integration with existing rate adaptations algorithms in practical environments are also provided. It is shown that when the scheduling period is less than 1000ms, most of the capacity gain can be maintained by the user-specific QoS aware schedulers. It is also shown that the scheduling algorithms can be integrated very easily into existing rate adaptation algorithms.

The contributions of sections 7.1.1-7.1.3 constitute the whole user-specific QoS scheduling system.

7.1.4 MIMO *In Vivo* WBAN Systems

MIMO technologies are introduced for *in vivo* environments to improve the system performance in WBAN systems. An extensive MIMO *in vivo* study was done to determine the maximum data rate that can be achieved considering various factors such as antenna separation distances, antenna angular positions, human body size, and system bandwidth, using ANSYS HFSS and their complete software human body model and an OFDM-based IEEE 802.11n system, with the specified SAR satisfied.

System capacity was first introduced and evaluated in the MIMO *in vivo* environment to determine the maximum potentially achievable data transmission rate. It was found, that compared with SISO *in vivo*, MIMO *in vivo* can significantly improve system performance, with maximum SAR levels met, making it possible to achieve target data rates of 100 Mbps with a system bandwidth of 40 MHz when receiver antennas are located at the front or back surface of the body.

7.2 Future Directions

Beyond what has been presented throughout this dissertation, there are topics that can be further explored. For example:

- Analyze the impact of the movement of the human body, the dynamics of the human body, and the movement of *in vivo* antennas on MIMO *in vivo* performance.
- Research the potential application of other physical layer technologies such as UWB to improve *in vivo* system performance. UWB transmissions have received much attention because of its ultra-low power, low complexity, its small antenna size, higher available

system bandwidth, and low susceptibility to multipath fading, and the immunity to interference [48], [52], [116].

REFERENCES

[1] E. Dahlman, S. Parkvall, and J. Skold, *4G: LTE/LTE-Advanced for Mobile Broadband*, 1 edition. Academic Press, 2011.

[2] 3GPP, "Overview of 3GPP Release 12 V0.1.4." .

[3] 3GPP, "TS 36.872- V1.0.0 - 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Small cell enhancements for E-UTRA and E-UTRAN - Physical layer aspects (Release 12)." .

[4] 3GPP, "TS 36.932 - V12.1.0 - LTE; Scenarios and requirements for small cell enhancements for E-UTRA and E-UTRAN." .

[5] 3GPP, "RP-121416, Further Downlink MIMO Enhancement for LTE-Advanced." .

[6] 3GPP, "TS 36.808- V10.1.0 - 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Carrier Aggregation; Base Station (BS) radio transmission and reception (Release 10)." .

[7] 3GPP, "TS 36.874- V1.0.0 - 3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Coordinated multi-point operation for LTE with non-ideal backhaul (Release 12)." .

[8] 3GPP, "TS 36.913 - V12.0.0 - LTE; Requirements for further advancements for Evolved Universal Terrestrial Radio Access (E-UTRA) (LTE-Advanced)." .

[9] "Requirements related to technical performance for IMT-Advanced radio interface(s)," *ITU*. [Online]. Available: http://www.itu.int/pub/R-REP-M.2134/en. [Accessed: 20-Dec-2014].

[10] S. Sesia, I. Toufik, and M. Baker, *LTE - The UMTS Long Term Evolution: From Theory to Practice*, 2 edition. Chichester, West Sussex, U.K. ; Hoboken, N.J: Wiley, 2011.

[11] 3GPP, "TS 36.321 - V12.0.0 - Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) specification." .

[12] H. Ekstrom, "QoS control in the 3GPP evolved packet system," *Communications Magazine, IEEE*, vol. 47, no. 2, pp. 76–83, 2009.

[13] 3GPP, "TS 23.203 - V12.6.0 - Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Policy and charging control architecture." 3GPP.

[14] J. D. Pearson, C. H. Morrell, S. Gordon-Salant, L. J. Brant, E. J. Metter, L. L. Klein, and J. L. Fozard, "Gender differences in a longitudinal study of age-associated hearing loss," *The Journal of the Acoustical Society of America*, vol. 97, no. 2, pp. 1196–1205, Feb. 1995.

[15] R. Blake, M. Rizzo, and S. McEvoy, "Aging and Perception of Visual Form From Temporal Structure," *Psychol Aging*, vol. 23, no. 1, pp. 181–189, Mar. 2008.

[16] D. I. S. U. S. Census Bureau, "Age and Sex Composition in the United States: 2012." [Online]. Available: http://www.census.gov/population/age/data/2012comp.html. [Accessed: 18-Nov-2014].

[17] "HISTORY OF LIFE EXPECTANCY," *World Life Expectancy*. [Online]. Available: http://www.worldlifeexpectancy.com/history-of-life-expectancy. [Accessed: 21-Nov-2014].

[18] "FeatureAgingFigure1 » Sociological Images." .

[19] S. Khan, S. Duhovnikov, E. Steinbach, and W. Kellerer, "MOS-Based Multiuser Multiapplication Cross-Layer Optimization for Mobile Multimedia Communication," *Advances in Multimedia*, vol. 2007, pp. 1–11, 2007.

[20] A. Saul, Shoaib Khan, G. Auer, W. Kellerer, and E. Steinbach, "Cross-Layer Optimization With Model-Based Parameter Exchange," presented at the Communications, 2007. ICC '07. IEEE International Conference on, 2007, pp. 5701–5707.

[21] S. Thakolsri, S. Khan, E. Steinbach, and W. Kellerer, "QoE-Driven Cross-Layer Optimization for High Speed Downlink Packet Access," *Journal of Communications*, vol. 4, no. 9, Oct. 2009.

[22] A. Saul and G. Auer, "Multiuser Resource Allocation Maximizing the Perceived Quality," *EURASIP Journal on Wireless Communications and Networking*, vol. 2009, pp. 1–15, 2009.

[23] P. Kela, J. Puttonen, N. Kolehmainen, T. Ristaniemi, T. Henttonen, and M. Moisio, "Dynamic packet scheduling performance in UTRA Long Term Evolution downlink," presented at the 3rd International Symposium on Wireless Pervasive Computing, 2008. ISWPC 2008, 2008, pp. 308–313.

[24] K. C. Beh, S. Armour, and A. Doufexi, "Joint Time-Frequency Domain Proportional Fair Scheduler with HARQ for 3GPP LTE Systems," presented at the Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th, 2008, pp. 1–5.

[25] G. Monghal, K. I. Pedersen, I. Z. Kovacs, and P. E. Mogensen, "QoS Oriented Time and Frequency Domain Packet Schedulers for The UTRAN Long Term Evolution," in *IEEE Vehicular Technology Conference, 2008. VTC Spring 2008*, 2008, pp. 2532–2536.

[26] F. Capozzi, G. Piro, L. A. Grieco, G. Boggia, and P. Camarda, "Downlink Packet Scheduling in LTE Cellular Networks: Key Design Issues and a Survey," *IEEE Communications Surveys & Tutorials*, vol. 15, no. 2, pp. 678–700, 2013.

[27] Y. Zaki, T. Weerawardane, C. Gorg, and A. Timm-Giel, "Multi-QoS-Aware Fair Scheduling for LTE," presented at the Vehicular Technology Conference (VTC Spring), 2011 IEEE 73rd, 2011, pp. 1–5.

[28] T. Han, N. Ansari, M. Wu, and H. Yu, "On Accelerating Content Delivery in Mobile Networks," *IEEE Communications Surveys Tutorials*, vol. 15, no. 3, pp. 1314–1333, Third 2013.

[29] K. Sayrafian-Pour, W.-B. Yang, J. Hagedorn, J. Terrill, K. Yekeh Yazdandoost, and K. Hamaguchi, "Channel Models for Medical Implant Communication," *International Journal of Wireless Information Networks*, vol. 17, no. 3, pp. 105–112, 2010.

[30] T. P. Ketterl, G. E. Arrobo, A. Sahin, T. J. Tillman, H. Arslan, and R. D. Gitlin, "*In vivo* wireless communication channels," in *IEEE 13th Annual Wireless and Microwave Technology Conference*, 2012, pp. 1–3.

[31] Charles Capps, "Near field or far field?," *EDN*, pp. 95–102, 16-Aug-2001.

[32] Ezio Biglieri, Robert Calderbank, Anthony Constantinides, Andrea Goldsmith, Arogyaswami Paulraj, and H. Vincent Poor, *MIMO wireless communications*. Cambridge University Press, 2007.

[33] C. Castro, A. Alqassis, S. Smith, T. ketterl, Y. Sun, S. Ross, A. Rosemurgy, P. Savage, and R. Gitlin, "A Wireless Miniature Robot for Networked Expedited Laparoscopy," *IEEE Transactions on Biomedical Engineering*, pp. 930–936, Apr. 2013.

[34] S. E. Hodgett, J. M. Hernandez, C. A. Morton, S. B. Ross, M. Albrink, and A. S. Rosemurgy, "Laparoendoscopic single site (LESS) cholecystectomy," *J. Gastrointest. Surg.*, vol. 13, no. 2, pp. 188–192, Feb. 2009.

[35] G. L. Stuber, J. R. Barry, S. W. McLaughlin, Y. Li, M.-A. Ingram, and T. G. Pratt, "Broadband MIMO-OFDM wireless communications," *Proceedings of the IEEE*, vol. 92, no. 2, pp. 271–294, 2004.

[36] Y. Ouyang, D. J. Love, and W. J. Chappell, "Body-Worn Distributed MIMO System," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 1752–1765, 2009.

[37] Y. Wang, I. B. Bonev, J. O. Nielsen, I. Z. Kovacs, and G. F. Pedersen, "Characterization of the Indoor Multiantenna Body-to-Body Radio Channel," *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 4, pp. 972–979, 2009.

[38] 3GPP, "TS 23.401 - V10.10.0 - LTE; General Packet Radio Service (GPRS) enhancements for Evolved Universal Terrestrial Radio Access Network (E-UTRAN) access." 3GPP.

[39] 3GPP, "TS 23.228 - V12.6.0 - Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE;IP Multimedia Subsystem (IMS); Stage 2." 3GPP.

[40] 3GPP, "TS 29.212 - V12.6.0 -Universal Mobile Telecommunications System (UMTS); LTE; Policy and Charging Control (PCC); Reference points." 3GPP.

[41] 3GPP, "TS 29.214 - V12.5.0 -Universal Mobile Telecommunications System (UMTS); LTE; Policy and charging control over Rx reference point." 3GPP.

[42] R. Ludwig, H. Ekstrom, P. Willars, and N. Lundin, "An Evolved 3GPP QoS Concept," in *Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd*, 2006, vol. 1, pp. 388–392.

[43] 3GPP, "TS 29.213 - V11.8.0 - Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Policy and charging control signalling flows and Quality of Service (QoS) parameter mapping." 3GPP.

[44] G. Camarillo, M. Handley, J. Peterson, J. Rosenberg, A. Johnston, H. Schulzrinne, and R. Sparks, "SIP: Session Initiation Protocol." [Online]. Available: http://tools.ietf.org/html/rfc3261. [Accessed: 17-Nov-2014].

[45] V. Jacobson, R. Frederick, S. Casner, and H. Schulzrinne, "RTP: A Transport Protocol for Real-Time Applications." [Online]. Available: http://tools.ietf.org/html/rfc3550. [Accessed: 19-Nov-2014].

[46] H. S. <schulzrinne@cs.columbia.edu>, "Real Time Streaming Protocol (RTSP)." [Online]. Available: https://tools.ietf.org/html/rfc2326. [Accessed: 03-Dec-2014].

[47] M. Handley, C. Perkins, and V. Jacobson, "SDP: Session Description Protocol." [Online]. Available: https://tools.ietf.org/html/rfc4566. [Accessed: 17-Nov-2014].

[48] B. Latré, B. Braem, I. Moerman, C. Blondia, and P. Demeester, "A Survey on Wireless Body Area Networks," *Wirel. Netw.*, vol. 17, no. 1, pp. 1–18, Jan. 2011.

[49] S. Ullah, H. Higgins, B. Braem, B. Latre, C. Blondia, I. Moerman, S. Saleem, Z. Rahman, and K. S. Kwak, "A Comprehensive Survey of Wireless Body Area Networks," *J Med Syst*, vol. 36, no. 3, pp. 1065–1094, Jun. 2012.

[50] M. Patel and J. Wang, "Applications, challenges, and prospective in emerging body area networking technologies," *IEEE Wireless Communications*, vol. 17, no. 1, pp. 80–88, Feb. 2010.

[51] "IEEE 802.15 WPAN™ Task Group 6 (TG6) Body Area Networks," 15-Sep-2011. [Online]. Available: http://www.ieee802.org/15/pub/TG6.html. [Accessed: 15-Sep-2011].

[52] "IEEE Standard for Local and metropolitan area networks - Part 15.6: Wireless Body Area Networks," *IEEE Std 802.15.6-2012*, pp. 1–271, Feb. 2012.

[53] R. Cavallari, F. Martelli, R. Rosini, C. Buratti, and R. Verdone, "A Survey on Wireless Body Area Networks: Technologies and Design Challenges," *IEEE Communications Surveys Tutorials*, vol. 16, no. 3, pp. 1635–1657, Third 2014.

[54] IEEE P802.15 Working Group for Wireless Personal Area Networks (WPANs), "Channel Model for Body Area Network (BAN)," Sep. 2011.

[55] Telecommunication standardization sector of ITU, "P.862: Perceptual evaluation of speech quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs." International Telecommunication Union, Feb-2001.

[56] 3GPP, "TS 26.071 - V11.0.0 - Mandatory speech CODEC speech processing functions;AMR speech CODEC; General description." .

[57] ITU-T, "ITU-T G.107-The E-model: a computational model for use in transmission planning." Dec-2011.

[58] J. Fitzpatrick, "An E-Model based adaptation algorithm for AMR voice calls," in *Wireless Days (WD), 2011 IFIP*, 2011, pp. 1–6.

[59] R. G. Cole and J. H. Rosenbluth, "Voice over IP Performance Monitoring," *SIGCOMM Comput. Commun. Rev.*, vol. 31, no. 2, pp. 9–24, Apr. 2001.

[60] F. Mertz and P. Vary, “Efficient Voice Communication inWireless Packet Networks,” in *2008 ITG Conference on Voice Communication (SprachKommunikation)*, 2008, pp. 1–4.

[61] ITU-T, “ITU-T G.114-One-way transmission time.” May-2003.

[62] D. Wu, S. Ci, H. Luo, W. Zhang, and J. Zhang, “Cross-layer rate adaptation for video communications over LTE networks,” presented at the 2012 IEEE Global Communications Conference (GLOBECOM), 2012, pp. 4834–4839.

[63] A. Chan, K. Zeng, P. Mohapatra, S.-J. Lee, and S. Banerjee, “Metrics for Evaluating Video Streaming Quality in Lossy IEEE 802.11 Wireless Networks,” in *2010 Proceedings IEEE INFOCOM*, 2010, pp. 1–9.

[64] R. Zhang, S. L. Regunathan, and K. Rose, “Video coding with optimal inter/intra-mode switching for packet loss resilience,” *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 6, pp. 966–976, Jun. 2000.

[65] T. Brandao and M. P. Queluz, “No-Reference Quality Assessment of H.264/AVC Encoded Video,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 20, no. 11, pp. 1437–1447, Nov. 2010.

[66] K. Zhu, V. Asari, and D. Saupe, “No-reference quality assessment of H.264/AVC encoded video based on natural scene features,” 2013, vol. 8755, pp. 875505–875505–11.

[67] M. Naccari, M. Tagliasacchi, and S. Tubaro, “No-Reference Video Quality Monitoring for H.264/AVC Coded Video,” *IEEE Transactions on Multimedia*, vol. 11, no. 5, pp. 932–946, Aug. 2009.

[68] D. Jurca and P. Frossard, “Media Flow Rate Allocation in Multipath Networks,” *IEEE Transactions on Multimedia*, vol. 9, no. 6, pp. 1227–1240, Oct. 2007.

[69] L. U. Choi, M. T. Ivrlac, E. Steinbach, and J. A. Nossek, “Sequence-level models for distortion-rate behaviour of compressed video,” in *IEEE International Conference on Image Processing, 2005. ICIP 2005*, 2005, vol. 2, pp. II–486–9.

[70] C. He and R. D. Gitlin, “User Specific QoS and Its Application in Resources Scheduling for Wireless System,” in *Algorithms and Architectures for Parallel Processing*, X. Sun, W. Qu, I. Stojmenovic, W. Zhou, Z. Li, H. Guo, G. Min, T. Yang, Y. Wu, and L. Liu, Eds. Springer International Publishing, 2014, pp. 809–821.

[71] C. He and R. D. Gitlin, “Application-Specific and QoS-Aware Scheduling for Wireless Systems,” presented at the IEEE 25th International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 2014.

[72] C. He, G. E. Arrobo, and R. D. Gitlin, “Improving System Capacity Based upon User-Specific QoS for Heterogeneous Networks,” in 2015 IEEE Wireless Communications and Networking Conference (WCNC), 2015.

[73] 3GPP, “TS 36.211 - V12.0.0 - Evolved Universal Terrestrial Radio Access (E-UTRA);Physical channels and modulation.” .

[74] 3GPP, “TS 36.213 - V12.0.0 - Evolved Universal Terrestrial Radio Access (E-UTRA);Physical layer procedures.” .

[75] Alcatel-Lucent,AT&T, etc., 3GPP TSG-RAN Meeting 44, “RP-090660, Vocoder rate adaptation for LTE.” .

[76] 3GPP, “TS 26.114 - V10.3.0 - Universal Mobile Telecommunications System (UMTS); LTE; IP Multimedia Subsystem (IMS); Multimedia telephony; Media handling and interaction.” .

[77] S. Floyd, K. K. Ramakrishnan, and D. L. Black, “The Addition of Explicit Congestion Notification (ECN) to IP.” [Online]. Available: https://tools.ietf.org/html/rfc3168. [Accessed: 19-Nov-2014].

[78] 3GPP, “TR 23.860 - V10.0.0 - 3rd Generation Partnership Project; Technical Specification Group Services and System Aspects; Enabling Coder Selection and Rate Adaptation for UTRAN and E-UTRAN for Load daptive Applications; Stage 2.” .

[79] M. Westerlund, A. Lakaniemi, J. Sjoberg, and Q. Xie, “RTP Payload Format and File Storage Format for the Adaptive Multi-Rate (AMR) and Adaptive Multi-Rate Wideband (AMR-WB) Audio Codecs.” [Online]. Available: https://tools.ietf.org/html/rfc4867. [Accessed: 19-Nov-2014].

[80] U. Chandra, M. Westerlund, B. Burman, and S. Wenger, “Codec Control Messages in the RTP Audio-Visual Profile with Feedback (AVPF).” [Online]. Available: https://tools.ietf.org/html/rfc5104. [Accessed: 26-Nov-2014].

[81] A. Damnjanovic, J. Montojo, Y. Wei, T. Ji, T. Luo, M. Vajapeyam, T. Yoo, O. Song, and D. Malladi, “A survey on 3GPP heterogeneous networks,” *IEEE Wireless Communications*, vol. 18, no. 3, pp. 10–21, Jun. 2011.

[82] Y. S. Soh, T. Q. S. Quek, M. Kountouris, and H. Shin, “Energy Efficient Heterogeneous Cellular Networks,” *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 5, pp. 840–850, May 2013.

[83] M. Vajapeyam, A. Damnjanovic, J. Montojo, T. Ji, Y. Wei, and D. Malladi, “Downlink FTP Performance of Heterogeneous Networks for LTE-Advanced,” presented at the 2011 IEEE International Conference on Communications Workshops (ICC), 2011, pp. 1–5.

[84] M. Eguizabal and A. Hernandez, “Interference management and cell range expansion analysis for LTE picocell deployments,” presented at the 2013 IEEE 24th International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC), 2013, pp. 1592–1597.

[85] A. Daeinabi, K. Sandrasegaran, and X. Zhu, “Performance evaluation of cell selection techniques for picocells in LTE-advanced networks,” presented at the 2013 10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2013, pp. 1–6.

[86] “OPNET,” 12-Feb-2014. [Online]. Available: http://www.opnet.com/. [Accessed: 12-Feb-2014].

[87] 3GPP, “TS 25.814 - V7.1.0 - Physical layer aspects for evolved Universal Terrestrial Radio Access (UTRA).” .

[88] 3GPP, “TS 36.814 - V9.0.0 - Evolved Universal Terrestrial Radio Access (E-UTRA); Further advancements for E-UTRA physical layer aspects.” .

[89] 3GPP, “TS 24.301 - V12.6.0 - Universal Mobile Telecommunications System (UMTS); LTE; Non-Access-Stratum (NAS) protocol for Evolved Packet System (EPS); Stage 3.” .

[90] J. Arkko, E. Guttman, P. R. Calhoun, and J. Loughney, “Diameter Base Protocol.” [Online]. Available: https://tools.ietf.org/html/rfc3588. [Accessed: 01-Dec-2014].

[91] 3GPP, “TS 29.211 - V6.4.0 -Universal Mobile Telecommunications System (UMTS); Rx Interface and Rx/Gx signalling flows.” 3GPP.

[92] 3GPP, “TS 36.413 - V10.6.0 - LTE; Evolved Universal Terrestrial Radio Access Network (E-UTRAN); S1 Application Protocol (S1AP).” 3GPP.

[93] 3GPP, “TS 29.274 - V12.6.0 -Universal Mobile Telecommunications System (UMTS); LTE; 3GPP Evolved Packet System (EPS); Evolved General Packet Radio Service (GPRS) Tunnelling Protocol for Control plane (GTPv2-C); Stage 3.” 3GPP.

[94] 3GPP, “TS 36.300 - V12.2.0 - Evolved Universal Terrestrial Radio Access (E-UTRA);Overall description; Stage 2.” .

[95] 3GPP, “TS 36.304 - V11.6.0 - Evolved Universal Terrestrial Radio Access (E-UTRA);User Equipment (UE) procedures in idle mode.” .

[96] 3GPP, “TS 24.229 - V10.12.0 - Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE;IP multimedia call control protocol based on Session Initiation Protocol (SIP) and Session Description Protocol (SDP); Stage 3.” 3GPP.

[97] 3GPP, “TS 26.101 - V12.0.0 - Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Mandatory speech codec speech processing functions; Adaptive Multi-Rate (AMR) speech codec frame structure.” 3GPP.

[98] C. He and R. D. Gitlin, “User-Specific QoS Aware Scheduling and Implementation in Wireless Systems,” in the Wireless Telecommunications Symposium, 2015. WTS 2015, 2015, pp. 1–6.

[99] P. Resnick, “Internet Message Format.” [Online]. Available: https://www.ietf.org/rfc/rfc2822.txt. [Accessed: 17-Nov-2014].

[100] U. Chandra, M. Westerlund, B. Burman, and S. Wenger, “Codec Control Messages in the RTP Audio-Visual Profile with Feedback (AVPF).” [Online]. Available: http://tools.ietf.org/html/rfc5104. [Accessed: 19-Nov-2014].

[101] C. He, Y. Liu, T. P. Ketterl, G. E. Arrobo, and R. D. Gitlin, “MIMO *in vivo*,” in the Wireless and Microwave Technology Conference (WAMICON), 2014 IEEE 15th Annual, 2014, pp. 1–4.

[102] C. A. Castro, S. Smith, A. Alqassis, T. Ketterl, Y. Sun, S. Ross, A. Rosemurgy, P. P. Savage, and R. D. Gitlin, “MARVEL: A Wireless Miniature Anchored Robotic Videoscope for Expedited Laparoscopy,” presented at the IEEE International Conference on Robotics and Automation (ICRA), 2012, 2012, pp. 1–6.

[103] "IEEE Standard for Safety Levels With Respect to Human Exposure to Radio Frequency Electromagnetic Fields, 3 kHz to 300 GHz." IEEE, 2005.

[104] T. S. Rappaport, *Wireless Communications: Principles and Practice*, 2 edition. Upper Saddle River, N.J: Prentice Hall, 2002.

[105] A. K. Skrivervik, "Implantable antennas: The challenge of efficiency," in *2013 7th European Conference on Antennas and Propagation (EuCAP)*, 2013, pp. 3627–3631.

[106] H. G. Schantz, "Near field phase behavior," in *2005 IEEE Antennas and Propagation Society International Symposium*, 2005, vol. 3B, pp. 134–137 vol. 3B.

[107] C. He, Y. Liu, T. P. Ketterl, G. E. Arrobo, and R. D. Gitlin, "Performance Evaluation for MIMO *In Vivo* WBAN Systems," in the Microwave Workshop Series on RF and Wireless Technologies for Biomedical and Healthcare Applications (IMWS-BIO), 2014 IEEE MTT-S International, 2014, pp. 1–3.

[108] D. Tse, *Fundamentals of Wireless Communication*. Cambridge, UK ; New York: Cambridge University Press, 2005.

[109] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, 1 edition. Cambridge University Press, 2008.

[110] "ANSYS HFSS 15," 12-Jan-2013. [Online]. Available: http://www.ansys.com/Products/Simulation+Technology/Electromagnetics/High-Performance+Electronic+Design/ANSYS+HFSS. [Accessed: 12-Jan-2013].

[111] I. Khan, P. S. Hall, A. A. Serra, A. R. Guraliuc, and P. Nepa, "Diversity Performance Analysis for On-Body Communication Channels at 2.45 GHz," *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 4, pp. 956–963, Apr. 2009.

[112] "IEEE Standard for Information technology– Local and metropolitan area networks– Specific requirements– Part 11: Wireless LAN Medium Access Control (MAC)and Physical Layer (PHY) Specifications Amendment 5: Enhancements for Higher Throughput," *IEEE Std 802.11n-2009 (Amendment to IEEE Std 802.11-2007 as amended by IEEE Std 802.11k-2008, IEEE Std 802.11r-2008, IEEE Std 802.11y-2008, and IEEE Std 802.11w-2009)*, pp. 1–565, 2009.

[113] Agilent, "SystemVue Electronic System-Level (ESL) Design Software," 13-Jun-2012. [Online]. Available: http://www.home.agilent.com/agilent/product.jspx?cc=US&lc=eng&ckey=1297131&nid=-34264.0.00&id=1297131. [Accessed: 13-Jun-2012].

[114] T. P. Ketterl, G. E. Arrobo, and R. D. Gitlin, "SAR and BER Evaluation Using a Simulation Test Bench for *In Vivo* Communication at 2.4 GHz," in *IEEE 14th Annual Wireless and Microwave Technology Conference (WAMICON)*, 2013, pp. 1–4.

[115] C. He and R. D. Gitlin, "Improving WBAN System Performance Using MIMO *In Vivo*," *IEEE Transactions on Communications*, pp. 1–9, Dec. 2014.

[116] A. Ghildiyal, B. Godara, K. Amara, R. Dalmolin, and A. Amara, "UWB for low power, short range, in-body medical implants," in *2010 IEEE International Conference on Wireless Information Technology and Systems (ICWITS)*, 2010, pp. 1–4.

APPENDICES

Appendix A Copyright Permissions

Below is permission for the use of material in Chapter 3 and Chapter 4.

SPRINGER LICENSE
TERMS AND CONDITIONS

Feb 19, 2015

This is a License Agreement between Chao He ("You") and Springer ("Springer") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3572560424492
License date	Feb 19, 2015
Licensed content publisher	Springer
Licensed content publication	Springer eBook
Licensed content title	User Specific QoS and Its Application in Resources Scheduling for Wireless System
Licensed content author	Chao He
Licensed content date	Jan 1, 2014
Type of Use	Thesis/Dissertation
Portion	Full text
Number of copies	1
Author of this Springer article	Yes and you are a contributor of the new work
Order reference number	None
Title of your thesis / dissertation	Improving System Performance in Cellular and WBAN Networks via User-Specific QoS and MIMO In Vivo Technologies
Expected completion date	Apr 2015
Estimated size(pages)	140
Total	0.00 USD
Terms and Conditions	

Introduction
The publisher for this copyrighted material is Springer Science + Business Media. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightslink account and that are available at any time at http://myaccount.copyright.com).

Limited License
With reference to your request to reprint in your thesis material on which Springer Science and Business Media control the copyright, permission is granted, free of charge, for the use indicated in your enquiry.

Below is permission for the use of material in Chapter 6.

Thesis / Dissertation Reuse

The IEEE does not require individuals working on a thesis to obtain a formal reuse license, however, you may print out this statement to be used as a permission grant:

Requirements to be followed when using any portion (e.g., figure, graph, table, or textual material) of an IEEE copyrighted paper in a thesis:

1) In the case of textual material (e.g., using short quotes or referring to the work within these papers) users must give full credit to the original source (author, paper, publication) followed by the IEEE copyright line © 2011 IEEE.
2) In the case of illustrations or tabular material, we require that the copyright line © [Year of original publication] IEEE appear prominently with each reprinted figure and/or table.
3) If a substantial portion of the original paper is to be used, and if you are not the senior author, also obtain the senior author's approval.

Requirements to be followed when using an entire IEEE copyrighted paper in a thesis:

1) The following IEEE copyright/ credit notice should be placed prominently in the references: © [year of original publication] IEEE. Reprinted, with permission, from [author names, paper title, IEEE publication title, and month/year of publication]
2) Only the accepted version of an IEEE copyrighted paper can be used when posting the paper or your thesis on-line.
3) In placing the thesis on the author's university website, please display the following message in a prominent place on the website: In reference to IEEE copyrighted material which is used with permission in this thesis, the IEEE does not endorse any of [university/educational entity's name goes here]'s products or services. Internal or personal use of this material is permitted. If interested in reprinting/republishing IEEE copyrighted material for advertising or promotional purposes or for creating new collective works for resale or redistribution, please go to http://www.ieee.org/publications_standards/publications/rights/rights_link.html to learn how to obtain a License from RightsLink.

If applicable, University Microfilms and/or ProQuest Library, or the Archives of Canada may supply single copies of the dissertation.

BACK | CLOSE WINDOW

ABOUT THE AUTHOR

Chao He received his M.S. and B.S. degrees in Electrical Engineering from Shanghai Jiao Tong University, China, in 2000. He is currently working towards his Ph.D. degree in Electrical Engineering in University of South Florida, USA. Before he joined University of South Florida, he was with Huawei Technologies Co., Ltd., China, where he held various job positions ranging from project manager/leader to senior system engineer in the fields of LTE/LTE-A /WCDMA/TD-SCDMA wireless physical layer research, LTE/LTE-A baseband algorithms design and implementation, LTE/LTE-A baseband multi-core SoC design, and LTE smart phone product development. His research interests include advanced resource scheduling to optimize the performance of wireless networks and key technologies for Wireless Body Areas Networks (WBANs).