Comparative Performance Study of LTE Uplink Schedulers

by

Mohamed Salah

A thesis submitted to the
Department of Electrical and Computer Engineering
in conformity with the requirements for
the degree of Masters of Applied Science

Queen's University
Kingston, Ontario, Canada
April 2011

Abstract

Long Term Evolution (LTE) constitutes a significant milestone in the evolution of 3G systems towards fourth generation (4G) technologies. The performance targets promised by LTE makes it an ideal solution to accommodate the ever increasing demand for wireless broadband. LTE's promised performance targets were made possible due to improvements such as a simplified system access architecture and a fully IP-based platform. LTE has also great enhancements in its enabling radio technologies by introducing Orthogonal Frequency Division Multiplexing (OFDM) and advanced antenna technologies. In addition, LTE capabilities are further improved with enhanced Quality of Service (QoS) support for multiple data services, such as voice and other multimedia applications.

LTE packet scheduling plays an essential role as part of LTE's Radio Resource Management (RRM) to enhance the system's data rate and to support the diverse QoS requirements of mobile services. LTE packet scheduler should intelligently allocate radio resources to mobile User Equipments (UEs) such that the LTE network adheres to its performance requirements. In our work, we perform a performance evaluation of multiple LTE scheduling algorithms proposed for LTE uplink transmission. The evaluation takes place in single and mixed traffic scenarios to exploit the strengths and weaknesses of proposed algorithms. Simulation results illustrated the

importance of a scheduler's awareness of uplink channel conditions and QoS requirements in the presence of single and multiple traffic scenarios. Accordingly, we provide recommendations for future scheduling algorithm proposals, and ways to enhance the existing schedulers.

Acknowledgments

The work presented in this thesis is conducted as part of the Masters program at the Department of Electrical and Computer Engineering at Queen's University, under the supervision of Dr. Hossam Hassanein, along with my co-supervisor Dr. Najah Abu-Ali.

I would like to thank Dr. Hassanein from the Computer Science Department at Queen's University for his guidance, encouragement, and financial support throughout the duration of my thesis program. Also, I would like to thank Dr. Najah from the College of IT at UAE University for her professional support, valuable feedback, and also for ensuring my high morale throughout the course of this project.

In addition, I would like to provide my thanks to Dr. Taha from the Computer Science Department at Queen's University for his valuable guidance and instructions to assist me in completing the project.

Next, I would like to express my sincere gratitude to my friends and colleagues at the Telecommunication Research Lab at Queen's university for their friendship and support throughout my years at Queen's University.

And last, but not least, I would like to express my gratitude to my family and friends who have always been there for constantly providing me with their support and friendship as well.

List of Acronyms

2G	Second Generation Wireless
3G	Third Generation Wireless
3GPP	Third Generation Partnership Project
4G	Fourth Generation Wireless
ADSL	Asynchronous Digital Subscriber Line
ARP	Allocation Retention Priority
ARQ	Automatic Repeat reQuest
ATB	Adaptive Transmission Bandwidth
BFS	Breadth First Search
BMTP	Block Allocation for Minimum Total Power
BS	Base Station
BSR	Buffer Status Report
C-plane	Control plane
CAPEX	Capital Expenses
CP	Cyclic Prefix
CSI	Channel State Information
D-SR	Dedicated SR

DFDMA	Distributed FDMA
DFT	Discrete Fourier Transform
DFT-S-OFDMA	DFT-Spread-OFDMA
eNodeB	evolved NodeB
EDGE	Enhanced Data Rates for Global Evolution
EPC	Evolved Packet Core
E-UTRAN	Evolved-Universal Terrestrial Radio Access Network
FDD	Frequency Division Duplexing
FDPS	Frequency Domain Packet Scheduling
FME	First Maximum Expansion Algorithm
FTB	Fixed Transmission Bandwidth
GPRS	General Packet Radio Service
GSM	Global System for Mobile communication
GBR	Guaranteed Bit Rate
HSDPA	High Speed Downlink Packet Access
HARQ	Hybrid Automatic Repeat reQuest
HLGA	Heuristic Localized Gradient Algorithm
IP	Internet Protocol
IPTV	IP Television
ISI	Inter-Symbol-Interference
IRME	Improved RME Algorithm
ITRME	Improved Tree-based RME Algorithm
LCG	Logical Channel Group
LFDMA	Localized FDMA

LTE	Long Term Evolution
MAC	Medium Access Control
MAD	Minimum Area Difference Algorithm
MBR	Maximum Bit Rate
MME	Mobility Management Entity
MCS	Modulation and Coding Scheme
MC-SA	Multi-Carrier Scheduling Algorithm
MMSE	Minimum Mean Squared Error
MSC	Mobile Switching Center
OPEX	Operational Expenses
OFDM	Orthogonal Frequency Division Multiplexing
OFDMA	Orthogonal Frequency Division Multiple Access
PDCCH	Physical Downlink Control Channel
PDCP	Packet Data Convergence Protocol
PDN	Packet Data Network
PFGBR	Proportional Fairness with Guaranteed Bit Rate
PHY	Physical Layer
PRB	Physical Resource Block
PS	Packet Scheduling
PUCCH	Physical Uplink Control Channel
PUSCH	Physical Uplink Shared Channel
QoS	Quality of Service
QCI	QoS Class Identifier
RAC	Radio Admission Control

RA-SR	Random Access-based SR
RB	Radio Bearer
RBC	Radio Bearer Control
RC	Resource Chunk
RLC	Radio Link Control
RME	Recursive Maximum Expansion
RNC	Radio Network Controller
RR	Round Robin
RRC	Radio Resource Control
RRM	Radio Resource Management
SC-FDMA	Single Carrier Frequency Division Multiple Access
SC-SA	Single-Carrier Scheduling Algorithm
SDF	Service Data Flow
SFN	Sequence Frame Number
SIMO	Single Input Multiple Output
SINR	Signal to Interference and Noise Ratio
SR	Scheduling Request
SRS	Sounding Reference Signal
TB	Transport Block
TDD	Time Division Duplexing
TDPS	Time Domain Packet Scheduling
TTI	Transmission Time Interval
UE	User Equipment
WCDMA	Wideband Code Division Multiple Access

UMTS Universal Mobile Telecommunications System

U-plane User plane

Contents

List of Tables

List of Figures

Chapter 1

Introduction

The success story of mobile broadband has started along with the emergence of Third Generation (3G) mobile networks at the beginning the third millennium. The Third Generation Partnership Project (3GPP) has introduced 3G Universal Mobile Terrestrial System (UMTS) standard as 3GPP Release 5 in response to the continuous success of its predecessor Second Generation (2G) GSM/EDGE technology (Global System for Mobile Communications/Enhanced Data Rates for Global Evolution). UMTS enhances the data rate such that the experienced download speed can reach up to 2 Mbps, which is a significant improvement if compared to the download speed over EDGE networks.

UMTS has been greatly enhanced to arrive at 3GPP Release 6 of UMTS, alternatively known as High Speed Packet Access (HSPA). HSPA further increases the data rates over the wireless medium to download speeds up to 14 Mbps and upload speeds up to 5.8 Mbps. Internet applications and services could then be supported over 3G wireless interface with a reliable Quality of Service (QoS) support, which has attracted an increasing number of customers. As a result, the growth of 3G HSPA

has been accompanied by emerging smart phone devices with wireless data communication support and also emerging mobile services such as media streaming and video calling. As of today, HSPA is the dominant mobile data technology at the global scale, whose dominance is expected to last at least for another five to ten years [5]. According to [6], broadband subscription are expected to reach the 3.4 billion mark by 2014, 80% of which are going to be mobile broadband subscriptions.

Due to the exponential increase in mobile data usage, the need for more advanced broadband mobile technology presents itself as a necessity to support the forecasted traffic volumes over the long term of ten years or more. Henceforth, 3GPP has initiated the move towards standardizing the next generation of mobile broadband technology with the following motives in perspective [7]:

- Ensure the competitiveness of 3GPP's 3G systems in future markets.
- Meet the increasing demand for higher data rates and improved QoS support.
- Meet the demand for reducing Capital and Operational Expenses (CAPEX and OPEX) of mobile networks via
  - Moving into an optimized, fully packet switched network, and
  - Reducing the complexity in network protocols and network architecture.

In 2004, 3GPP commenced the work on their Release 8 of mobile broadband standard as a 3G-beyond solution, which was later termed as Long Term Evolution (LTE). Compared to its 3G predecessors, LTE is viewed as an important milestone in the path of mobile broadband evolution in terms of its enhanced features and enabling technologies, which makes LTE a strong competitor to wired broadband networks such as cable and ADSL (Asynchronous Digital Subscriber Line).

Release 8 of LTE was finalized in December 2008, which constitutes describing the system structure whether at the access level or at the core level. 3GPP then introduced more enhancements to LTE in 3GPP Release 9, such as the support of relay stations and femtocells, which was also finalized in December 2009 [7]. Many network operators around the globe has begun deployment trials of LTE networks as an intermediate step towards fully commercial deployment within the next few years [8]. Yet, LTE technology is still investigated by researchers, as many aspects of LTE performance still need further investigation to exploit LTE's full potential.

At the access level, an LTE network consists mainly of LTE base stations (BSs), termed as evolved NodeBs (eNodeBs). LTE standard has eliminated Mobile Switching Controller (MSC) and Radio Network Controller (RNC) units that are present in 2G and 3G access networks, respectively. MSC and RNC units perform higher level management of radio access network, such as mobility management and Radio Resource Management (RRM). eNodeB has inherited some of RNC's responsibilities, such as mobility management and RRM, while the other tasks have been moved up to the packet core network.

RRM plays a very crucial role in LTE networks by managing the limited radio resources such that the achievable data rates over LTE's radio interface becomes as high as possible. The focus of our work presented here is on a very important RRM component residing at the eNodeB, which is the packet scheduler. The packet scheduler in LTE is responsible for allocating shared radio resources among mobile User Equipments (UEs). The packet scheduler allocates radio resources to UEs both on the downlink (from the eNodeB down to the UE) and also on the uplink (from UE up to the eNodeB). Intelligent distribution of radio resources among active UEs

optimizes the system performance and also reduces the cost per bit transmitted over the radio interface.

The intelligence of eNodeB packet scheduler depends on its awareness of several factors that are important for high system performance. First of all, the packet scheduler at the eNodeB has to be fully aware of the types of traffic flows running over the LTE interface and their respective QoS requirements. For example, voice calls in LTE are supported over a packet switched, IP-based platform as Voice over IP (VoIP) services. VoIP services have stricter delay requirements than most other data services, and these requirements have to be met to preserve a good service quality as perceived by end users.

In addition, the LTE packet scheduler needs to be aware of the channel quality per UE to properly adapt the transmission rates on both the uplink and downlink directions. Rate adaptation based on channel quality has been enhanced in LTE due to introducing Orthogonal Frequency Division Multiplexing (OFDM) modulation technique. OFDM provides the scheduler the advantage to exploit channel conditions in both time and frequency, where UEs are assigned frequency subchannels over which these UEs experience good channel conditions.

Also, an LTE uplink scheduler has to consider extra factors not present in the downlink direction, such as UE power limitation and the contiguity constraint of resource allocation. The contiguity of resource allocation refers to avoiding segmentation of radio resource allocation to a single UE, where the contiguity restriction is imposed by the modulation technique used in the LTE uplink. The power limitation and contiguity constraint makes the design of an LTE uplink scheduler a more challenging task.

Numerous literature works on LTE scheduling was mainly concerned with LTE packet scheduling on the downlink, such as the ones proposed in [9, 10, 11, 12, 13, 14, 15]. Early interest in LTE uplink scheduling was shown in [16, 17]. Many more proposals have been presented since then to address the challenges in LTE uplink scheduling such as UE power limitation and contiguity of resource allocation. The performance evaluation setups that different researchers used to evaluate their proposed schedulers cannot be assumed to be the same, which makes it more difficult to just claim that an uplink packet scheduler outperforms another. Also, no study has been proposed so far that provides a common, comparative performance evaluation for uplink schedulers from multiple proposals to provide a basis for a fair, defendable comparison between different LTE uplink schedulers.

The purpose of our work is to conduct a comparative performance evaluation of LTE uplink packet schedulers proposed that have been proposed multiple literature works so far. The study encompasses a wide survey over the LTE uplink packet scheduling proposals in literature. The surveyed LTE uplink schedulers are then categorized based on certain performance characteristics, where representative schedulers are selected from each category for our performance evaluation. The findings from our performance evaluation presented to draw conclusions on the performance of the representative schedulers, and point out the strengths and weakness that are common to schedulers from each category.

1.1 Contributions

We focus our performance evaluation on packet scheduling algorithms that are proposed so far for LTE uplink. The performance study focuses on dynamic uplink

schedulers that is centralized at the the LTE eNodeB. The contribution presented in our work is summarized as follows:

- Conducted a wide survey on literature works for LTE uplink scheduling proposals, and identified the key features that uniquely distinguish different scheduling algorithms. Afterwards, the surveyed schedulers were categorized based common characteristics among schedulers of each group. Representative scheduling algorithms were selected from each group for the performance evaluation.
- Selected the evaluation metrics that can best differentiate the performance of the selected representative schedulers.
- Developed a system-level simulator in MATLAB to have a unified evaluation environment for the representative schedulers. The LTE uplink simulator was designed to adhere to the LTE standard, such as the physical layer frame structure, system access node structure, traffic models, and standardized QoS provisioning.
- Analyzed the simulation results obtained from the performance study to explore the strengths and weaknesses of the selected uplink schedulers, and also to provide recommendations on the LTE uplink scheduler design for future LTE uplink scheduler proposals.

1.2 Organization of Thesis

As for the remainder of the thesis, Chapter 2 provides background information on LTE standard that are related to the LTE network and protocol architecture at the system

access level. More details are provided on physical layer frame structure as well as the rule of the MAC layer in RRM and QoS support via packet scheduling procedure. The background information provided in Chapter 2 are necessary to understand the subsequent discussions on LTE uplink schedulers. Chapter 3 presents the definition of LTE uplink scheduling problem and the approach used in most literature works to solve the uplink scheduling problem. The chapter continues to provide a survey on the multiple literature proposals for LTE uplink scheduling, along with the proposed categorization of the scheduling schemes. Chapter 4 provides more details on the representative uplink scheduling algorithms whose performance are to be evaluated in our study. The detailed discussion is supported by pseudocode illustration of the schedulers' operations. Chapter 5 then describes the simulation environment within which the scheduling algorithms were evaluated. Performance results are graphically presented for the different experiments along with the commentary on the obtained results. Chapter 6 provides concluding remarks on the findings from the performance evaluation, and also gives recommendations for future LTE Uplink Schedulers based on the result analysis.

Chapter 2

Background

This chapter provides preliminary background information on various aspects of the LTE system. The background provided below on the LTE system presents the challenges imposed on LTE uplink scheduling and how they are handled in the proposed scheduling algorithms presented in subsequent chapters. First, the target performance requirements of LTE system are presented which 3GPP has agreed upon based on its initial studies. The discussion then entails more details about the LTE system architecture followed by a description of LTE's protocol layer architecture and the functionalities supported at each layer. The chapter then introduces RRM in LTE, with the focus on LTE uplink scheduling and control signaling involved in the scheduling process. The chapter then concludes with an overview of the QoS architecture in LTE and what on parameters that are associated with different QoS classes.

2.1 LTE Performance Targets

The initial studies conducted by different Work Groups (WGs) within 3GPP on the evolution of 3G were aimed to set the minimum target requirements that LTE has to meet. The requirements and goals that are set for LTE are defined in [18], which are as follows:

Peak Data rate The target peak data rate is set to 100 Mbps on the downlink and 50 Mbps on the uplink assuming a bandwidth of 20 MHz. The target peak data rates correspond to spectral efficiency levels of 5 bits/sec/Hz on the downlink and 2.5 bits/sec/Hz on the uplink. LTE target data rates just stated are almost ten times higher than what is achieved in Release 6 of HSPA. The high LTE data rates are one of the most critical requirements for handling forecasted traffic loads of future services.

Reduced Latency Latency requirements cover both control plane (C-plane) and user plane (U-plane) latencies [19]. C-plane latency refers to the time it takes a UE to transit from either the idle or dormant states to the active state so that the UE becomes ready for an upcoming radio transmission. Dormant state refers to the wait time starting from the end of a successful transmission session till either the UE starts another transmission or goes to the idle state. LTE targets a maximum of 100 ms for the transition from the idle state to the active state, and 50 ms from the dormant state to the active state.

U-plane latency refers to the time is takes a data packet to be transmitted from the UE's data buffer and the serving gateway of the core network and vice versa. U-plane latency is controlled by QoS parameters that are set by EPC, which is discussed

later in this chapter when talking about QoS in LTE.

Bandwidth Flexibility LTE is to support multiple bandwidth sizes, which can be 1.4, 5, 10, 15, or 20 MHz. The support for flexible bandwidth sizes is possible in LTE due to OFDM modulation, which is explained in more detail in subsequent sections. Bandwidth flexibility is an important feature of LTE due to the fragmentation of spectrum allocation, where an operator cannot guarantee a contiguous 20 MHz bandwidth in many of the frequency spectra in which it operates.

Support for Heterogeneous Network Deployment A key requirement of LTE is backward compatibility with previous 3GPP technologies, like GSM and HSPA. Initial deployments of LTE are going to be in areas that are already covered by pre-LTE systems. Also, LTE is expected to be deployed in areas that are covered by non-3GPP wireless technologies such as Wi-Fi and WiMAX.

The support of heterogeneous deployment with other 3GPP and non-3GPP networks provides UE mobility support across different wireless platforms, and leads to easier and more cost efficient LTE deployment.

Simplified Architecture Architecture simplification refers to reducing the number of network access nodes to have as few access nodes between the UE and the network core as possible. Simplifying the network architecture this way reduces the COPEX and OPEX, as well as the latency just mentioned above.

Also, simplifying the network architecture entails the introduction of a fully packet switched, IP-based platform. The IP-only-support provides a unified platform to support all services that are expected to run over LTE, instead of having separate

packet switched and circuit switched networks like in HSPA. Hence, the choice of full IP-support further simplifies the architecture of the LTE system.

Enhanced QoS Support The move to fully IP-based platform dictates supporting multiple services with a variety of QoS requirements, such as packet delay and packet loss rate. The LTE network should be able to provide enhanced QoS support in scenarios of any traffic mix.

2.2 LTE Access Network Architecture

Figure 2.1 provides a general overview of LTE access network architecture, which is alternatively referred to as Evolved-Universal Terrestrial Radio Access Network (E-UTRAN). As shown in the figure, E-UTRAN is composed of a single type of access level component, the eNodeB [2]. An eNodeB carries the functionalities of NodeB and RNC nodes that are present in HSPA, and constitutes the central unit that carries out all radio-related functions at the access level. It acts as the terminal point of all radio communications carried out by the UE, and relays data flows between the radio connection and the Evolved Packet Core (EPC) network. EPC is an IP-based, multi-access core network that makes it possible for operators to deploy one packet core network for multiple 3GPP radio access technologies (GSM, HSPA, and LTE) [20]. The functionality of EPC revolves mainly around mobility management, policy management, security, and acting as the Internet gateway for access level nodes. An eNodeB is connected to two units in EPC: mobility management entity (MME) and serving gateway (S-GW).

eNodeB also RRM tasks, where it manages the usage of radio resources by means

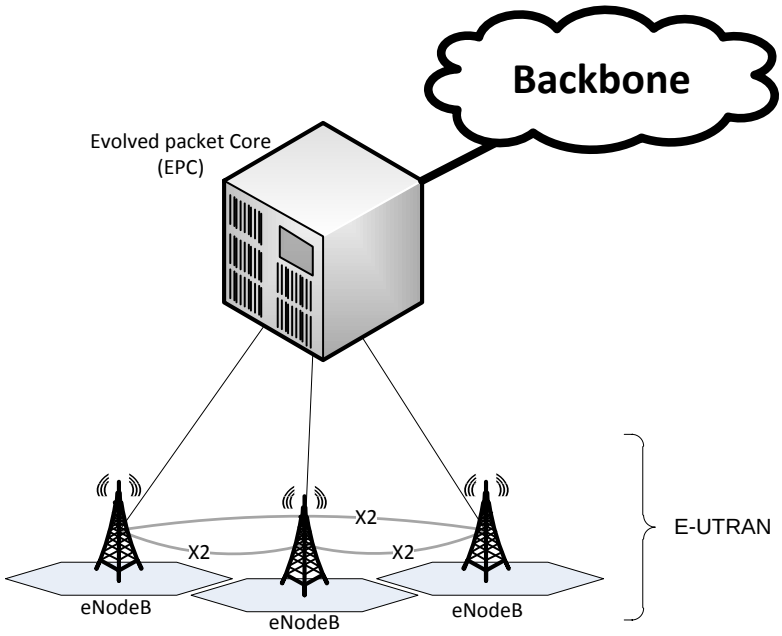

Figure 2.1: LTE access network (E-UTRAN) architecture, reproduced from [2]

of distributing radio resources among UEs with active transmissions, admitting or rejecting new connections arriving at the eNodeB's coverage, and prioritizing traffic flows according to their associated QoS attributes. eNodeB is responsible for mobility-related tasks as well such as handover mechanisms for UEs moving from one eNodeB coverage to another, and also UE tracking when in idle in cooperation with MME in EPC. The coordination among neighboring eNodeBs occurs through a logical interface that connects eNodeBs together that is known as the X2 interface.

2.3 LTE Radio Protocol Architecture

Figure 2.2 below provides an overall illustration of the protocol architecture at the E-UTRAN level [2]. The top three protocols are the sublayers of the TCP/IP Layer

2, while the PHY layer forms LTE's TCP/IP Layer 1.

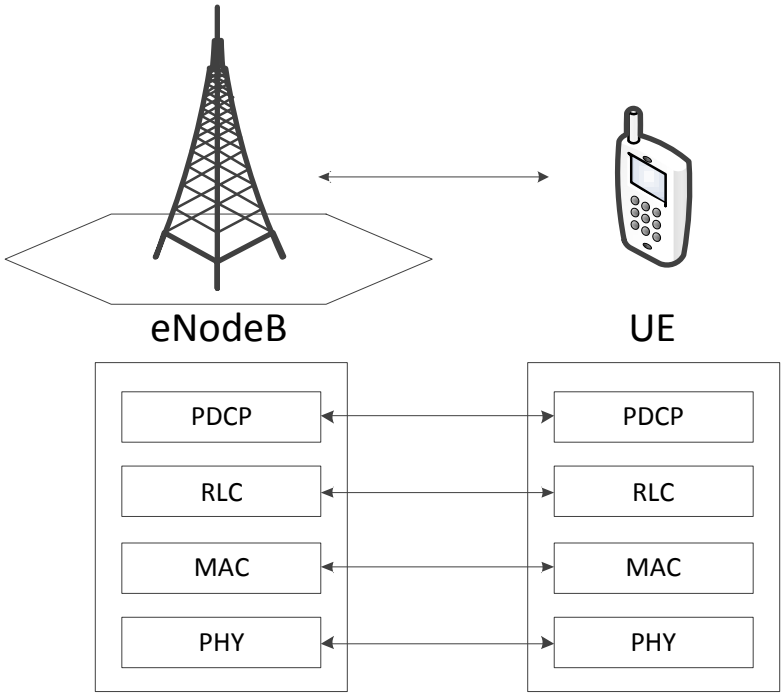

Figure 2.2: LTE E-UTRAN protocol architecture, reproduced from [2]

- Layer 2 (L2)
  - Packet Data Convergence Protocol (PDCP).
  - Radio Link Control (RLC).
  - Medium Access Protocol (MAC).
- Layer 1 (L1)
  - Physical Layer (PHY).

2.3.1 PDCP Layer

PDCP layer is mainly responsible for header compression/decompression of IP packets that are received from/transmitted to IP layer above, respectively. The header compression is critical in reducing the overhead of data communication over the wireless interface, which in turn increases the system's spectral efficiency.

PDCP also ensures in-sequence delivery of data packets either up to the IP layer or down to the RLC layer. The sequential delivery mechanism ensures that PDCP detects missing packets for which it can initiate retransmissions, or duplicate packets that are to be discarded.

PDCP also performs packet ciphering and packet encryption to ensure the secure delivery of data packets over the radio interface.

2.3.2 RLC Layer

RLC layer facilitates the transfer of data units between PDCP and MAC sublayers. In doing so, RLC performs tasks similar to PDCP, such as in-sequence packet delivery and duplicate packet detection. In addition to that, RLC performs error correction of data packets received from MAC layer below through window-based Automatic Repeat reQuest (ARQ) operation.

RLC also performs segmentation and re-assembly of data units that are passed down to/received from MAC sublayer below. To perform such a task, RLC layer contains the transmit/receive data buffers for different traffic flows, which are alternatively know as Radio Bearers (RBs). RB refers to a traffic flow, or a group of traffic flows, between a UE and the eNodeB over the radio interface that is characterized with certain QoS attributes. Keeping separate RB buffers ensures that RLC tailors

its services to RBs according to their QoS needs, such as providing ARQ services to RBs with best effort traffic low very tolerance to packet loss.

2.3.3 MAC Layer

The main task of MAC sublayer in LTE is to map between logical channels and PHY's transport channels. Logical channels are services provided by MAC to RLC sublayer above to accommodate different types of data exchange. Logical channels can be categorized into data traffic channels, for transferring data traffic, and control channels, for transferring control signals between UE and eNodeB. Each logical channel service is provided to a certain RB from RLC layer to ensure proper RB prioritization according to their QoS requirements.

MAC sublayer performs priority handling operation to map user and control data flows from different RBs to their appropriate physical channels on PHY layer via PHY's transport channels interface. Priority handling is performed either on RBs from different UEs, which is referred to as packet scheduling, or between RBs within the same UE.

In addition, MAC layer is responsible for, as part of the UE priority handling just described, detecting data transmission errors and correcting them via allocating time and frequency resources for data retransmissions. Data retransmissions are handled at the MAC sublayer through a process termed as Hybrid ARQ (HARQ), which is a combination of forward error-detection and correction via decoding process.

2.4 PHY Layer

HSPA employs Wideband Code Division Multiple Access (WCDMA) as the transmission method over HSPA's physical wireless channels within a 5 MHz spectrum. When it comes to LTE, WCDMA is a non-valid choice for fulfilling LTE requirements due to the difficulty of supporting bandwidth sizes larger than 5 MHz using single carrier radio interface like WCDMA. As LTE targets flexible spectrum allocations up to 20 MHz, the use of WCDMA would require a high Sgnal-to-Noise-Ratio (SINR) and more complex filter design and equalization schemes at the receiving antennas. Henceforth, WCDMA, as well as almost any single-carrier transmission scheme, becomes a poor choice for LTE. 3GPP directed its attention to the use of multiple-carrier transmissions for providing high data rates as efficiently as possible in terms of SINR [21].

Hence, OFDM modulation scheme was a more appropriate choice for LTE due to the multicarrier nature of OFDM that provides significant advantages in terms of high data rate support [22].

According to the 3GPP standard [3], Orthogonal Frequency Multiple Access (OFDMA) was chosen for the downlink direction, while Single Carrier Frequency Division Multiple Access (SC-FDMA) was chosen for the uplink.

2.4.1 OFDMA

The downlink in LTE uses OFDMA for its transmission scheme. The main principle behind OFDMA is breaking the radio spectrum into multiple orthogonal, narrowband subcarriers [23] to span the entire system bandwidth. With the use of narrowband subcarriers, data transmission is carried out over multiple streams with low data

rates, and hence larger duration of OFDM symbols carrying data traffic. Having larger symbol duration helps in neutralizing inter-symbol interference (ISI) effects to reduce error rates of data transmission. The increase in symbol period is further enhanced with adding guard bands between successive OFDM symbols in a process named Cyclic Prefix (CP). CP consists of a repetition of the last part of the preceding OFDM symbol so that the signal detection at the receiver side turns to a circular convolution for enhancing signal detection capability.

In addition to neutralizing ISI effects, using OFDM modulation for the radio interface allows for frequency domain equalization process to take place at the receiver end, which simplifies the hardware implementation of the receiver's equalizer.

2.4.2 SC-FDMA

When it comes to the LTE uplink, choosing OFDMA has proven to be less advantageous because of the high peak-to-average-power-ratio (PAPR) problem [23, 22]. The PAPR problem of OFDMA refers to having large variation of power transmission levels with the peak power of OFDMA transmission is high compared to the transmission power average. High PAPR would cause power inefficiency when transmitting data from the mobile terminal on the uplink channels, and hence imposes a challenge to the antenna design in mobile terminals. Therefore, SC-FDMA was found to be a better choice as LTE uplink transmission scheme.

SC-FDMA, which is alternatively known as DFT-Spread-OFDMA (DFT-S-OFDMA), is a variant of OFDMA. The main difference from OFDMA is the addition of Discrete Fourier Transform (DFT)-spread block prior to modulating the data symbols using OFDM modulation. SC-FDMA shares almost all the characteristics as OFDMA

transmission scheme, in addition to lower PAPR in comparison to OFDMA. A UE can benefit from SC-FDMA in the uplink transmission in terms of increased transmission power efficiency along with increased data rates, which eventually translates to improved battery life on the UE.

Despite the advantage SC-FDMA provides by lowering the PAPR effect, SC-FDMA requires that all subcarriers assigned to a single UE must be adjacent to each other on the frequency domain. Such restriction imposes a challenge on the allocation method of scheduling design, to be discussed later.

2.4.3 The SC-FDMA Frame Structure

Figure 2.3 illustrates the time-domain frame structure that is adopted for LTE uplink as well as downlink [3]. Although the discussion of the frame structure below focus on uplink frame structure, it is assumed to equally apply to LTE downlink as well.

The LTE uplink is divided time-wise into 10 ms long radio frames. Each frame is identified by a System Frame Number (SFN) to control different transmission cycles, such as paging and sleep mode cycles, that last more than one frame in duration.

A radio frame, in turn is divided into ten subframes. A subframe is 1 ms long, and is further divided into two 0.5 ms slots. The number of SC-FDM symbols in a time slot depends on which of two types of CP mode used, which are known as normal CP mode and extended mode. Normal CP is the default mode in which a time slot fits seven SC-FDM symbols. Extended CP mode is another mode that is defined for LTE uplink where the CP duration is extended when the system experiences higher ISI over the uplink. In this case, only six SC-FDMA symbols can fit a time slot under extended CP mode. The number of data bits carried by each SC-FDMA symbol is depends on

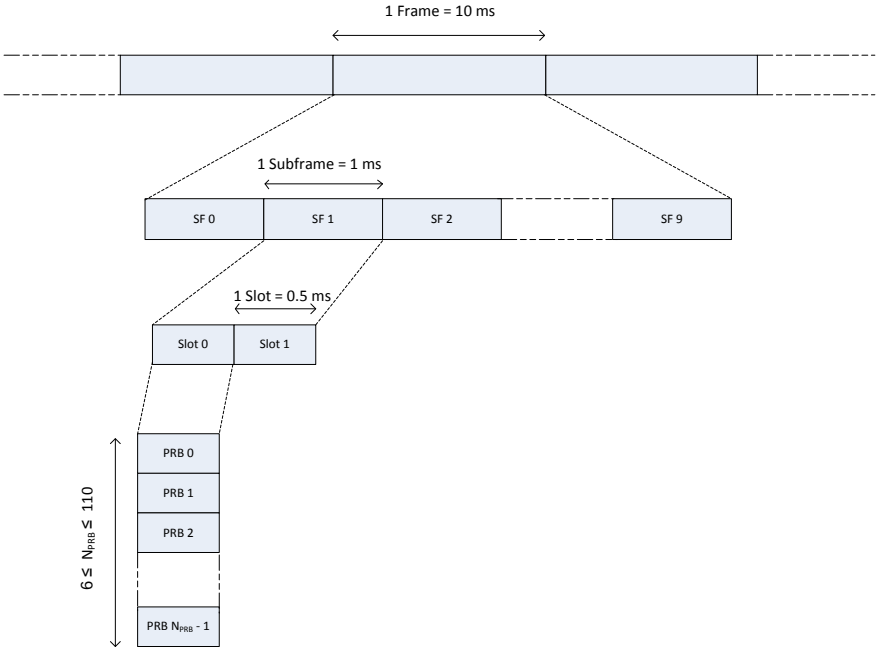

Figure 2.3: Frame structure of OFDMA and SC-FDMA radio interface, reproduced from [3]

the Modulation and Coding Scheme (MCS) used in the uplink transmission. In our study, we assume normal CP mode for the LTE uplink operation.

As illustrated in Figure 2.3, the frequency domain structure of a SC-FDMA time slot is divided into regions of 180 kHz that contain a contiguous set of twelve SC-FDMA subcarriers. The 0.5 ms $\times$ 180 kHz time-frequency block constitutes the basic radio resource unit termed as the Physical Resource Block (PRB). As shown in Figure 2.4, a PRB spans a group of twelve SC-FDM subcarriers, each carrying seven SC-FDM symbols, or resource elements as indicated in the figure, assuming normal CP mode.

The LTE standard defines two duplexing modes to allocate LTE frames among the uplink and downlink direcrtions: Frequency Division Duplexing (FDD) and Time

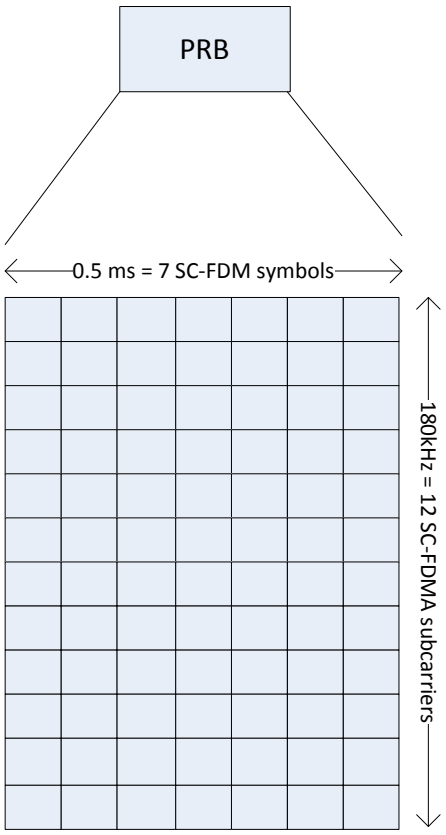

Figure 2.4: The Frequency Domain Structure of a single SC-FDMA time slot, reproduced from [3]

Division Duplexing (TDD) mode. In FDD mode, as the uplink and downlink transmissions are separated into different frequency bands, each subframe is treated as a whole unit for either uplink or downlink transmission.

In case of TDD, however, each frame contains both uplink and downlink transmissions, where distributions of the subframes between the two transmissions depend on the TDD configuration used, in addition to the presence of the special subframe that provides a guardband between the uplink and downlink transmissions.

2.5 Radio Resource Management in LTE

RRM in LTE is responsible for accepting/rejecting connection requests based on network policies and is responsible also for ensuring the efficient use of available radio resources [24]. In that sense, RRM provides LTE the necessary means to meet its target requirements that are discussed earlier in this chapter. In E-UTRAN, the role of RRM is focused on two major tasks: Radio Admission Control (RAC), and Packet Scheduling (PS).

2.5.1 Radio Admission Control

Radio Admission Control (RAC) is responsible for examining UEs' admission requests for new connections, whether on the downlink or uplink. RAC then accepts or rejects the admission requests based on the ability to accommodate these requests such that the network is still capable to accommodate the minimal QoS requirements of pre-existing and new RB connections as well.

In case of congestion, RAC can also accept new connections even with the scarcity of available radio resources within the system. In this situation, another RRM entity termed as Radio Bearer Control (RBC) examines the priorities of currently active sessions. The RBC entity may then choose to drop a connection if it has been active for a long period of time in order to admit a new connection with a higher priority. An example of such a newly admitted connection would be a UE with active session that goes through a handover process fas the UE moves from the neighboring cell to the target cell.

2.5.2 Packet Scheduling

Packet scheduling refers to allocating PRB resources to UEs with active, ongoing sessions on a periodic basis. Packet scheduling in LTE occurs as often as once every subframe, where a scheduling period of 1 subframe is alternatively known as Transmission Time Interval (TTI).

Packet scheduling in LTE involves selecting a set of UEs to be scheduled on an upcoming scheduling period. The packet scheduler then performs PRB allocation to decide which UE to utilize which group of PRBs on either the uplink or the downlink. A packet scheduler performs its allocation decision to maximize the satisfaction level system requirements. A scheduler measures system satisfaction based on a desirable performance metric, such as per UE's experienced data rate, fairness in resource allocation among UEs, average packet delay experienced by UEs, and so on. The choice of what performance metric to optimize influences how the scheduler resolves resource contention among UEs. In addition to resource allocation, the packet scheduler performs other tasks such as Link Adaptation (LA) to improve the chances of transmitted data packets to correctly arrive at their destination.

2.5.3 Control Signaling for Uplink Packet Scheduling and Link Adaptation

The LTE standard does not specify how the scheduler in LTE should be implemented. What the standard specifies is the exchange of packet scheduling control messages between the eNodeB and its associated UE as indicated in [24]. Our discussion in this section focuses on the signaling mechanism defined for uplink packet scheduling.

Uplink packet scheduling mechanism is triggered by the UE's sending the eNodeB

a scheduling request (SR) for uplink resources to transmit its data. The UE sends SR using one of two mechanisms defined in the LTE standard [24, 25]:

1. **Random Access-based SR (RA-SR).** The UE uses the random access (RA) procedure when it has no dedicated resources for it on the uplink control channels. The UE tries to time-align with the target eNodeB to obtain radio resources on the uplink, during which the UE sends a one-bit SR flag to the eNodeB. Once the access procedure is successful, the target eNodeB sends a scheduling grant to the UE on the Physical Downlink Control Channel (PDCCH) indicating on which PRBs the UE transmits its uplink data.

2. **Dedicated SR (D-SR).** The UE uses D-SR mechanism when it already has dedicated resources on the Physical Uplink Control Channel (PUCCH). The UE in this case sends a one-bit flag on the PUCCH to request resources for its uplink data. As in the case with RA-SR, the eNodeB responds by sending a scheduling grant to the UE over PDCCH.

In addition, the LTE standard provides other reporting mechanisms to provide the packet scheduler with valuable information about the cellular environment that can assist in increasing the scheduling operation in the uplink.

2.5.3.1 Channel State Information

Channel State Information (CSI) refers to the SINR measurement of the channel condition on the uplink direction between the UE and eNodeB. CSI reporting for the uplink channel is based on channel sounding techniques, where the UE sends a Sounding Reference Signal (SRS) of known magnitude that spans either part of or the

entire transmission bandwidth [26, 3]. The uplink packet scheduler uses SRS reports to determine the channel condition at each schedulable PRB, per UE. Several UEs can send their SRS signaling over the same bandwidth without interfering with each other due to the use of Zadoff-Chu sequence generation [3].

2.5.3.2 Buffer Status Reporting

Buffer Status Reporting (BSR) refers to the 3GPP standardized reporting mechanism a UE uses to communicate its buffer information to the eNodeB [24]. BSR mechanism plays an important role in the uplink scheduler's QoS provisioning where the scheduler gets a status report on how much data await transmission at the UE's uplink buffer. BSR is triggered in three scenarios as defined in the standard:

1. Uplink data packets belonging to a RB from a Logical Channel Group (LCG) become available for transmission. LCG is defined as a group of RBs that exhibit similar QoS characteristics. In the case that a UE supports multiple active RBs that belong to different LCGs, the UE provides long BSR report to provide the buffer status of all LCGs with active transmission. Otherwise, it sends a short BSR for a single LCG.

2. A UE is granted PRB resources for uplink transmission, and the size of padding bits is equal to or larger than the size of the BSR control element if inserted as part of the MAC packet.

3. Periodic BSR, where the period duration is set within the Radio Resource Control (RRC) layer at the eNodeB according to [27].

2.6 QoS Architecture

The 3GPP's concept on QoS is introduced in [28] and [29]. According to 3GPP, LTE organizes the different types of traffic flows into logical traffic pipes named *bearer services*. Each bearer service has certain QoS attributes associated with it, depending on the type of traffic it carries. Accordingly, traffic bearers are categorized into four QoS classes, based on the QoS constraints of the bearer's traffic:

- Conversational class.
- Streaming class.
- Interactive class.
- Background class.

Conversational Class Conversational class is the one with the most delay-sensitive profile. Example for traffic types that fall under this category are video conferencing and voice telephony. The delay requirements of such traffic comes from the human perception of delay in a conversation, which imposes high restrictions on telephony traffic.

Streaming Class The streaming class is also intended for real-time traffic, same as the conversational class. The main difference is that services from the streaming class have less stringent requirements than the conversational one on packet delay. Streaming class services usually assume a live entity (i.e. human) to be present at one end of the connection, which makes streaming services less sensitive to packet

delay than conversational services that assume live entities at both end terminals. An example of streaming services is on-line video streaming.

Interactive Class Services from the interactive class are Internet services that are based on an end client requesting data from a remote entity. Web-browsing and database information retrieval are examples of interactive services. Services belonging to the interactive class expect delivery of requested data within a certain time range. Hence, delay requirements are important for interactive traffic, yet these requirements are very loose when compared to those from the services belonging to either the conversational or streaming classes. Interactive services, on the other hand, place more emphasis on error-free delivery of data packets, and hence have stricter packet error rate requirements than conversational and streaming services.

Background Class Background traffic refers to data services running in the background. Background class has the least delay restrictions among the four QoS classes. Data delivery of background traffic tolerate larger delay in packet delivery to ensure the lowest error-rate possible. Examples of background services include Pear-to-Pear applications, File Transfer Protocol (FTP) services, and email delivery.

2.6.1 LTE and QoS

Each traffic flow with unique behaviors and certain QoS characteristics is termed as Service Data Flow (SDF). Being a fully packet-switched network, LTE supports different requirements of diverse SDFs to enhance the experienced QoS of as many end users as possible. QoS support in LTE is provided through EPS bearer, which is defined as the level of granularity for QoS control in the EPC/E-UTRAN at bearer

level. An EPS bearer may contain one or more SDFs, where they share the same forwarding policy set EPC network [2].

An EPS bearer is established when connecting a UE to the Packet Data Network (PDN) within EPS. PDN then assigns the UE an IP connection that stays on for the lifetime of the EPS bearer, which is referred to in this case as the default bearer. Other EPS bearers can get established within the default bearer for dedicated SDF types, which are thus termed as dedicated EPS bearers. An EPS bearer, whether default or dedicated, is categorized into either a GBR, or a non-GBR bearer. A GBR bearer is the one that carries SDF with Guaranteed Bit Rate (GBR) requirements, while a non-GBR bearer does not have any GBR requirements. Each dedicated bearer is characterized by certain QoS parameters that are set and controlled by Policy and Charging Control (PCC) architecture within the EPC network.

PCC architecture is responsible for policy and charging control mechanisms that are applied to each SDF bearer created between the UE and PDN. PCC regulates the SDF-filtering such that it either charges or rejects data packets that do not match any of the SDF filters. The SDF filtering is performed by setting the proper QoS parameters for each SDF, and is enforced by a PCC unit termed as Policy and Charging Enforcement Function (PCEF).

A EPS bearer is characterized by the following set of QoS attributes [2, 1]:

- **Allocation Retention Priority (ARP).** This refers to the allocation and retention priority mechanism of bearer resources at connection setup and handover. ARP aids the network in deciding which RBs are kept in situations such as congestion control. ARP's role is mainly confined to bearer establishment and RAC mechanisms, as it has no impact on post RAC tasks, such as packet

scheduling.

- **Maximum Bit Rate (MBR).** This parameter is used by the network to set a limit on the data rates for each radio bearer, so that no radio bearer connection would not exhaust the network's resources.

- **Guaranteed Bit Rate (GBR)**. GBR refers to sustaining a minimum bit rate based on the policy set by the network administrator.

- **QoS Class Identifier (QCI)**. QCI defines a set of characteristics that describe the packet forwarding treatment that an SDF receives between the UE and the PCEF in the EPC. The QCI parameters associated with each SDF are as follows:

  - Bearer Type. It is a parameter that reflects whether the associated bearer is a GBR or non-GBR one.
  - Packet Delay Budget (PDB). The maximum packet delay allowed between the UE (starting with the packet entering the transmission buffer) and PCEF.
  - Packet Loss Rate. The maximum number of erroneous bits that can be tolerated the traffic flow of a given QCI. This parameter is mainly used for HARQ operations at the MAC level.

  QCI parameters are presented in Table 2.1, which is reproduced from [1].

Table 2.1: QCI characteristics, reproduced from 3GPP standard[1].

QCI	**Priority**	**Resource Type**	**Packet Delay Budget**	**Packet Error Loss**	**Examples**
1	2	GBR	100 ms	10^{-2}	Conversational Voice
2	4	GBR	150 ms	10^{-3}	Conversational Video (Live Streaming)
3	3	GBR	50 ms	10^{-3}	Real Time Gaming
4	5	Non-GBR	300 ms	10^{-6}	Non-Conversational Video (Buffered Streaming)
5	1	Non-GBR	100 ms	10^{-6}	IMS Signalling
6	6	Non-GBR	300 ms	10^{-6}	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)
7	7	Non-GBR	100 ms	10^{-3}	Voice,Video (Live Streaming)Interactive Gaming
8	8	Non-GBR	300 ms	10^{-6}	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)
9	9	Non-GBR	300 ms	10^{-6}	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)

Chapter 3

LTE Packet Scheduling

This chapter focuses more on the uplink scheduling problem in LTE. Here, we define a problem statement for LTE uplink scheduling by first stating the considerations and challenges that must be taken into account when allocating resources to UEs. Then, we present a model for LTE uplink scheduler design which best describes the general operation of the scheduling algorithms proposed in literature so far. Thereafter, the chapter continues with a category-based survey that has been conducted on the proposed scheduling algorithms. The categorization scheme proposed here has been based on schedulers' performance objective, where a performance objective can be maximizing allocation fairness or optimizating a QoS requirement like experienced throughput relative to a UE's GBR or the UE's experienced average delay.

3.1 Defining the Scheduling Problem

The uplink packet scheduler is an RRM entity that exists at the eNodeB as part of the MAC layer, where it decides which PRB is to be assigned to which UE. SC-FDMA, being the standardized uplink radio interface, enables the scheduler to exploit per-UE channel variations in space, time, and frequency. The multi-dimensional channel exploitation can maximize the UE-diversity where UEs are mapped to PRBs over which they experience advantageous channel conditions. The channel-depending scheduling hence turns the presence of radio channel fading peaks on portions of the bandwidth from a significant limitation into a major advantage [30].

A good scheduler is one that can better predict the needs of each UE, such that a UE can satisfy the QoS requirements of its traffic while at the same time try to utilize the resources assigned to it as efficiently as possible. When allocating resources to multiple UEs over the shared radio interface, the uplink packet scheduler has to take multiple factors into consideration, such as:

- Payloads buffered at the transmission packet queue, as the scheduler has to ensure that packets do not stay within the transmission queue longer than required.
- H-ARQ retransmissions, as the scheduler needs to determine which PRBs are scheduled for retransmissions and which are scheduled for new transmissions.
- CSI reporting, which provides to the scheduler information about channel conditions per UE.
- QoS parameters of all active connections (delay, packet error rate, GBR, etc).

- UE transmission history, which is most often represented by the windowed, past average throughput experience by a UE.
- Maximum number of UEs allowed to be scheduled within the current TTI.
- Contiguity constraint, which refers to having all PRBs allocated to a single UE contiguous in the frequency domain.
- Limitation in uplink power transmission, due to the limited battery life of the UE.

Hence, the packet scheduler takes either some or all of the factors mentioned above to come up with an allocation pattern that maximizes a desired system objective. The objective can include either of the following:

- **Achievable Throughput**, where the packet scheduler ensures that the available bandwidth is utilized as efficiently as possible to maximize of amount of data successfully transmitted over the radio interface.
- **Fairness**, which is expressed either as fairness in resource allocation among UEs, intraclass fairness between connections of the same QoS class, or interclass proportional fairness between groups of connections belonging to different QoS classes.
- **QoS Satisfaction**, where the packet scheduler maximizes the positive QoS experience per UE, such as minimizing packet delay and packet loss.
- **Power Utilization**, where a PS scheduler attempts to minimize the per UE power utilization on the uplink.

3.2 Packet Scheduler Modeling

LTE uplink scheduling can be addressed as an optimization problem, where the desired solution is the mapping between the schedulable UEs to schedulable PRBs that maximizes the desired performance target. Solving the scheduling problem can be very complex given the number of factors to take into account, as well as the virtually unlimited number of scheduling patterns to examine. In addition, the packet scheduler faces the hard-time constraints where the scheduler has only one TTI length to come up with the optimal allocation scheme.

To solve this problem, designing a practical packet scheduler for LTE uplink can be broken into two stages:

1. **Defining a Utility Function.** A utility function is a mathematical model that measures the satisfaction level of the system performance towards meeting its target requirements. Target requirements refer to performance metrics such as data throughput (system or per-UE throughput), per-UE experienced packet delay, fairness in resource allocation among UEs, and so on. The utility function can measure the satisfaction of either one or more performance metrics, depending on the types of services expected to run within the system and depending on policies set by the network operator. The satisfaction level based on the utility function changes from one TTI to another as a consequence of how PRBs are distributed among active UEs. If we denote the utility function per UE u as U_u, then the system's utility function, U_{sys}, can be expressed as the summation of the utility functions of its individual UEs, $U_{sys} = \sum_i U_i$.

2. **Designing Search-based Allocation Scheme.** The packet scheduler executes a search algorithm to traverse through as many UE-to-PRB allocation choices as possible to end up with a UE-to-PRB mapping that best optimizes the scheduler's utility function. Depending on what utility function is used, the optimal resource mapping is either a one that minimizes a system's utility (e.g. packet delay) or maximizes it (e.g. throughput, fairness). Being limited to only one TTI, the search algorithm needs to be simplified to a heuristic method that makes UE-to-PRB mapping using simplified search methods.

Once the utility function choice and allocation algorithm design are finalized, the next phase is to realize them implementation-wise. Implementing the uplink packet scheduler can be modeled according to the model presented in [4]. A utility-based uplink scheduling operation can be decoupled into per-domain scheduling: a Time Domain Packet Scheduling (TDPS), and Frequency Domain Packet Scheduling (FDPS).

In TDPS, the scheduler performs prioritization of the currently active UEs to be scheduled for the upcoming TTI. The scheduler then chooses all UEs or a subgroup of UEs with the highest priorities for scheduling. FDPS then performs UE-to-PRB allocation that maximizes the LTE network's satisfaction level based on the scheduler's utility function. Figure 3.1 illustrates the overall design of an LTE uplink scheduler.

Most of the scheduling algorithms proposed for LTE uplink can fit the TDPS/FDPS scheduling model. The packet scheduler generates a utility-based metric(s) for each UE based on the utility function. The generated metric is either directly extracted or mathematically deduced from the utility function definition. A metric is a weight

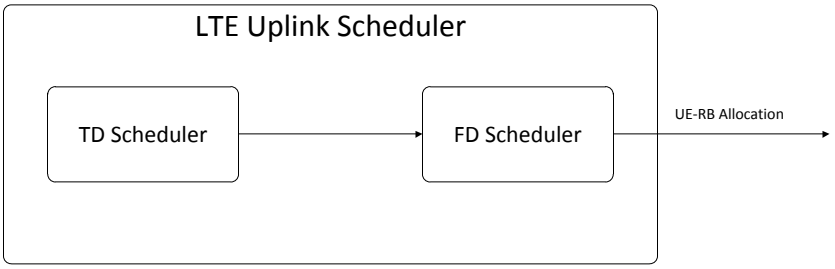

Figure 3.1: Block Diagram of Overall Design of LTE Uplink Scheduler, reproduced from [4]

that represents a UE's scheduling priority in a given TTI, the UE's gain when allocated either a certain PRB or a certain group of contiguous PRBs, or a combination of both.

TDPS scheduling is responsible for generating a per-UE weight that represents the UE's scheduling priority, where UEs with higher priorities are more likely to be allocated resources in an upcoming TTI. TDPS has not been fully addressed in many uplink scheduling proposals, where its main functionality has been limited to passing a list of all UEs with active transmissions to the FDPS.

The FDPS then calculates for each UE a utility-based metric at each PRB or for each subset of PRBs that can be allocated to the UE. Examples of Frequency Domain (FD) metrics include the achievable throughput and Proportional Fairness (PF). The calculated FD metrics for all UEs are fed into the allocation algorithm to determine the UE-to-PRB mapping that best optimizes the system performance gain as explained earlier.

3.3 Literature Review

This section provides a survey on the LTE scheduler proposals from literature. In conducting our study, we categorized the surveyed schedulers based on the scheduler's objective as follows:

1. Best Effort Schedulers
2. QoS-Based Schedulers
3. Power-Optimizing Schedulers

3.3.1 Best-Effort Schedulers

Most of the proposals made for LTE uplink scheduling focused on maximizing performance metrics such as data throughput and fairness, for example. Best-effort schedulers were designed such that their main target is to maximize the utilization of the radio resources and/or fairness of resource sharing among UEs.

The work presented in [17] is one of the earliest published works on this topic. The authors in [17] proposed two PF schedulers with two greedy allocation schemes, with one being proposed for Localized (L-)FDMA while the other one was for Distributed (D-)FDMA. For each UE, the scheduler calculates the PF metric for each PRB based on the estimated achievable throughput on the given PRB and also on the past average throughput of the UE. The work presented showed a significant gain in aggregated cell throughput in the case of the contiguous scheme compared to the interleaved allocation algorithm. Hence, the work demonstrated the advantage of using the contiguous allocation scheme in improving system performance.

Another LTE uplink scheduler was presented in [31], where the authors proposed a Heuristic Localized Gradient Algorithm (HLGA) to perform contiguous PRB allocation to schedulable UEs. HLGA performs dynamic resource allocation where the number of PRBs allocated to each UE dynamically changes in every TTI. The algorithm was proposed with H-ARQ awareness, during which the scheduler preserves some PRBs for retransmissions of previously unsuccessful transmissions. Afterwards, the scheduler allocates the remaining PRBs for UEs with new transmissions. The scheduler operates such that it keeps looking for the UE that has the largest metric at a given PRB. The scheduler assigns the PRB to the given UE as long as it does not break the allocation contiguity constraint. As a result, if the UE has been allocated resources previously, and the PRBs between the previously allocated PRBs and the PRB-to-be-assigned are assigned to no other UE, the scheduler assigns these PRBs to the same UE as well.

The authors from [31] extended their work to add a *pruning* process, where the scheduler modifies the PRB assignment to a UE based on its buffer status. If the number of contiguous PRBs assigned to a UE can hold more data than present in the UE's data buffer, the scheduler removes extra PRBs from the assigned PRB set, and allocates them to other UEs that are in more need for them. The results presented in [32] showed a performance improvement as a result of the buffer-awareness of the *pruning*-based HLGA scheduler. However, the buffer-awareness improvement of the scheduler assumes having detailed information of the buffer status at the UE end, which is not a realistic assumption according to the buffer reporting mechanism defined in [24] for LTE uplink.

Another set of dynamic LTE schedulers was proposed in [33], in which the authors

proposed three LTE uplink dynamic schedulers: First Maximum Expansion (FME), Recursive Maximum Expansion (RME), and Minimum Area Difference (MAD). All three schedulers were proposed with the same PF utility function, yet they mainly differ in the allocation scheme each follow. In FME, the algorithm starts with searching for the UE with the maximum PF metric at a given PRB. Once found, the algorithm expands the allocation for the UE till the algorithm finds no more PRBs whose maximum metrics belong to the same UE. At this point, the algorithm selects either side of the PRB set allocated to the first UE, and starts allocating PRBs in order to the UEs that hold the maximum PF metrics over the visited PRBs and at the same time does note break the contiguity constraint. Once all the PRBs on the one side is done, FME algorithm starts allocating PRBs on the other side in the same fashion.

In the case of RME, the allocation algorithm starts in the same fashion as FME, namely finds the UE-PRB pair with the maximum metric and expands the allocation for the given UE till there are no more PRBs whose maximum metric belongs to the same UE. RME scheduler then recursively performs the same procedure for the remaining UEs until all available PRBs get assigned.

The third scheduler proposed in this work, MAD, is search-tree based. The algorithm starts by isolating the maximum PF metric at every PRB and whose UE it belongs to. Afterwards, for each UE, the scheduler derives per-PRB MAD metric consisting of the difference between the UE's PF metric of a given PRB and the maximum metric of the PRB. Then, the PRBs are grouped into Resource Chunks (RCs), with each RC consists of a group of consecutive PRBs whose maximum PF metric belongs to the same UE. The next step is for the scheduler to construct all the possible UE-PRB allocation patterns in the form of a Breadth First Search (BFS)

tree, with each path consisting of one possible allocation pattern for the PRBs. The scheduler traverses through the tree to find the path that minimizes the sum of the MAD metrics of the different PRBs.

The results obtained in [33] demonstrated the superiority of RME and MAD over FME in terms of fairness and achievable throughput. However, there were no significant difference between the RME and MAD performance levels, despite MAD's higher computational complexity.

RME algorithm was further investigated in [34], where the authors derived two variants of RME algorithm, termed Improved RME (IRME) and Improved Tree-based RME (ITRME). The two RME variants were tested with Max-C/I metric instead of PF metric. Results shown in [34] suggested a performance improvement by 15% of proposed RME-variants over the RME algorithm in terms of spectral efficiency.

Another work on PF-based schedulers was proposed in [35]. The authors introduced the concept of Fixed Transmission Bandwidth (FTB), where the scheduler groups the PRBs into equally sized RCs. The size of RC depends on the scheduler's configuration. The authors first introduced a greedy algorithm where each UE is assigned a maximum of one RC based on the PF metric of each UE at each PRB. Then, the greedy algorithm introduced was further extended come up with a binary tree based greedy scheduler. The search-tree greedy scheduler examines each RC for the two UEs with first and second best metric. The algorithm eventually constructs a binary tree with the available with each path representing an allocation scheme from which the best one is selected.

3.3.2 QoS-Based Schedulers

Another group of literature work has investigated LTE uplink scheduling with QoS-awareness in mind. The QoS provisioning in the schedulers belonging to this group is implemented as part of the schedulers' utility functions.

One of the early contributions to QoS-based LTE uplink scheduling was proposed in [36], where the authors proposed a scheduling algorithm for supporting mixed traffic loads on LTE uplink. The proposed scheduler aimed at maximizing the proportional fairness of the active connections while maintaining QoS requirements using the GBR metric of the traffic flows. Hence, the scheduler was termed by the authors as Proportional Fairness with Guaranteed Bit Rate scheduling algorithm (PFGBR). PFGBR uses PF metric, by default, as a scheduling metric for UEs with non-GBR traffic flows. In the case of UEs with GBR-based traffic, the scheduler adds an extra weight to the PF metric that is an exponential function of GBR to provide better priority to UEs with stricter GBR requirements. The study showed an improved support for UEs with QoS requirements without leaving UEs with BE traffic as victims of starvation.

Another QoS-based scheduler was proposed in [37], where the authors proposed GBR/PF LTE uplink scheduler accompanies by a proposal for QoS-aware RAC algorithm. The packet scheduling algorithm proposed here explicitly decouples the scheduling process into TD and FD scheduling, as explained earlier in this chapter. QoS provisioning is achieved by introducing a term that is a function of the UE's average throughput normalized by its GBR. The introduced GBR-based term is used in TD as a prioritization method among UEs, and was also used as part of the FD metric as well.The study showed that the proposed scheduler provided better support

for QoS traffic streams, especially the ones with low GBR rate, such as VoIP services.

Other proposals for QoS-based packet schedulers incorporated packet delay provisioning in addition to GBR. A work on uplink scheduling with packet delay provisioning was presented in [38], where the authors proposed two delay-aware, PF-based scheduling algorithms. One of the algorithms uses a linear PF utility function, while the other one uses PF-marginal utility function. Both algorithms look for UEs whose experienced average delay exceeds their traffics' delay budgets, and schedule them first for transmission. Once all critical UEs are served, both schedulers distribute whatever resources available to remaining UEs using PF-based scheduling.

The authors in [38] have extended their work on QoS scheduling, where they proposed another two QoS-based schedulers. Instead of examining UEs in the order of their experienced packet delays, the two schedulers proposed in [39] use a utility-based metric with GBR and packet delay provisioning combined. Each of the two schedulers prioritizes UEs according to the QoS metric and assigns them PRB resources such that UEs with highest priority are served first. The difference between the two schedulers is that the first one, named Single-Carrier Scheduling Algorithm (SC-SA), assigns each UE a maximum of one PRB if the number of schedulable UEs is larger than the number of available PRBs. The second proposed scheduler, named Multi-Carrier Scheduling Algorithm (MC-SA), assigns a UE either one or more PRBs. The scheduler estimates the number of RBs to assign a UE based on the ratio between the UE's experienced throughput to its GBR.

3.3.3 Power-Optimizing Schedulers

The goal of schedulers in this category coincide with the purpose of choosing SC-FDMA as the radio interface for LTE uplink; the purpose is to reduce power consumption of mobile UEs on wireless transmissions. A scheduler in this category usually acquires some QoS aspects of the traffic flows transmitted on the LTE uplink, such as packet delay budget or GBR requirements. The schedulers in this case perform some algorithmic decisions such they reduce the transmission power of a UE down to a point that the UE can maintain its QoS requirements.

Power-optimizing schedulers were not investigated thoroughly in the literature work in LTE. Two scheduling algorithms were encountered in the survey that can be categorized as such. The first proposal was presented in [40], where the authors introduced power minimizing schedulers based on queue delay constraints. The concept of the scheduler here is that increasing the experienced packet delay by a small amount while respecting the UE's GBR requirements can lead to a significant savings in the UE's transmission power.

Another proposal for Power-based LTE uplink scheduling was proposed in [41]. The scheduler design was based on binary integer programming concept, where the scheduler starts with creating a matrix that represents all the allocation patterns possible of uplink PRBs such that it still respects the contiguity constraint. The scheduler then starts calculating the power allocating needed for each possible allocation pattern for each UE. Afterwards, the scheduler performs a greedy-based search algorithm to find the UE-PRB allocation pattern that minimizes the power expenditure on each UE while respecting their GBR requirements.

Chapter 4

Representative LTE Uplink Schedulers

This chapter provides more detailed discussion on the representative schedulers from the three categories discussed in Chapter 3. The chapter starts with defining necessary notations for describing how the schedulers operate. The mechanism of the representative schedulers are then illustrated in detail to reflect how the schedulers where implemented for the performance evaluation.

4.1 Utility Functions in Representative LTE Uplink Schedulers

All scheduling algorithms that have been selected for the performance evaluation are utility-based schedulers. As mentioned beforehand, it is the scheduler's utility function that defines the target performance metric a scheduler aims to maximize.

The scheduler uses its utility function to realize the TD and FD metrics used by the TDPS and FDPS, respectively. In other words,

$$U_u \longrightarrow M^{TD}(u),\ M^{FD}(u,r) \tag{4.1}$$

where $M^{TD}(u)$ is the TD metric of UE u at TTI t, $M^{FD}(u,r)$ is the FD metric of UE u at PRB r at TTI t. Again, for schedulers with no explicit definition of TDPS scheduling, M^{TD} is assumed to be unity, and all UEs with non-empty transmission buffer are passed to the FDPS for resource assignment. Table 4.1 defines terms that are used throughput this chapter for defining the mathematical expressions that describe the TD and FD metrics.

Table 4.1: Metrics-related terms.

Term	Definition
$\hat{R}(u,r)$	Estimated throughput for UE u on PRB r at TTI t
$\bar{R}(u)$	Moving, window-based average throughput for UE u at TTI t
$\bar{R}_{sch}(u)$	Average throughput for UE u at TTI t averaged only on TTI's within which UE u was allocated PRB resources
$\bar{D}(u)$	The average packet delay experience by UE u up to TTI t
$D_{MAX}(u)$	The packet delay of traffic carried out by UE u
$D^{Q}_{ave}(u)$	The average packet delay experienced by all UEs of QoS class q to which UE u belongs
$R_{GBR}(u)$	The GBR throughput limit assigned to UE u

4.1.1 Utility Metric Representations

Once the TD and FD metrics of the scheduler are calculated, the scheduler forms a metric-based matrix before FDPS's allocation algorithm takes place. For most schedulers examined in our work, the metric-based matrix provides a UE-to-PRB metric mapping, as shown in Table 4.2. The matrix entries represent FD metrics of UEs at all schedulable PRBs.

Table 4.2: A matrix that represents UE-to-PRB metric mapping.

RBs / **UEs**	RB_1	RB_2	$\cdots$	$RB_{N_{RB}}$
UE_1	$M_{1,1}$	$M_{1,2}$	$\cdots$	$M_{1,N_{RB}}$
UE_2	$M_{2,1}$	$M_{2,2}$	$\cdots$	$M_{2,N_{RB}}$
$\vdots$	$\cdots$			$\cdots$
UE_N	$M_{N,1}$	$M_{N,2}$	$\cdots$	$M_{N,N_{RB}}$

Some of the schedulers presented in this work, such as PFGBR and BMTP schedulers, contain metric values that map UEs to PRB-subsets rather than individual PRBs. The derivation of UE-to-PRB-subset mapping adopted in our study is based on the work presented in [41]. A PRB-subset is a group of contiguous PRBs that are assigned to a UE as a single unit. All PRB-subsets made of contiguous PRBs are represented in a matrix A of size $N_{PRB} \times N_C$, where N_{PRB} is the total number of PRBs and N_C is the number of all possible allocations. N_C is calculated as

$$N_C = 0.5N_{PRB}^2 + 0.5N_{PRB} + 1 \tag{4.2}$$

For example, in the case of an uplink bandwidth with 4 PRBs, matrix A is expressed as

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

Each column c in matric A represents one schedulable PRB-subset on the LTE uplink, with 1's entries within column c represent PRBs belonging to the PRB-subset.

The scheduler determines the FD metric for all PRB-subsets, after which the scheduler constructs a $N_{UE} \times C$ metric matrix in a similar fashion to UE-PRB metric matrix shown in Table 4.2.

4.2 Pseudocode Notations and Nomenclature

The following defines the symbolic notations used in our pseudocode descriptions:

- $|\cdot|$: the vertical line notation indicates the number of elements of a given set
- $\mathcal{U}$: set of all UEs that are available for scheduling at a given TTI.
- $\mathcal{I}$: set of total PRBs that are available for scheduling for all UEs in $\mathcal{U}$, where $\mathcal{I} = \{1, 2, \ldots, u, \ldots, N\}$.
- $\mathcal{I}_u$: the set of PRBs already assigned to UE u, $\mathcal{U} = \{1, 2, \ldots, r, \ldots, N_{PRB}\}$.
- $\mathcal{I}'_u$: set of unassigned PRBs that can still be scheduled UE u.
- $\mathcal{I}_{RC}$: set of RCs that are to be scheduled for UEs in $\mathcal{U}$. An RC, as described earlier, is a set of contiguous PRBs that are scheduled as a unit.

- $\mathcal{C}_u$: set of schedulable PRB-subsets u, where $\mathcal{C}_u = \{1, 2, \ldots, c, \ldots, N_C\}$. As such, $\mathcal{I}^C_{u,c}$ corresponds to the PRB-subset in the c^{th} column in matrix A. i.e. $\mathcal{I}^C_{u,c} = \{\text{PRB } r : A(r, c) = 1\}$.
- N: the total number of schedulable UEs.
- N_{PRB}: the total number of PRBs.
- N_{RC}: the total number of schedulable RCs.
- N_C: number of schedulable PRB-subsets in $\mathcal{C}_u$.
- M: a UE-to-PRB metric matrix of dimensions $N \times N_{PRB}$.
  - $M(u, r)$ denotes the scheduling metric for UE u at PRB r.
  - $M(u, :)$ denotes metrics for all PRBs for a given UE.
  - $M(:, r)$ denotes metrics of all UEs at a given PRB.

  In the case of PFGBR and BMTP, M represents a $N \times N_C$ UE-to-PRB allocation pattern metric-based matrix, where $M(u, c)$ represents the metric of the UE u for a PRB-subset $\mathcal{C}_u(c)$.
- Operations:
  - $a \leftarrow b$: Assign b to a
  - $a : b$: range between a and b, inclusive. e.g.: $A[a : b] \leftarrow i$ means to assign all RBs between a and b to UE i.
  - $a = b$ and $a \neq b$: used for logical equality testing.

- $S \setminus \{s\}$: remove element s from the set S. e.g., $R_B \leftarrow R_B \setminus \{b\}$ means to remove RB b from the set of RBs R_B, then replace the new set with the existing one.
- $\max_k(Arr)$: find the k_{th} max metric value in an array Arr.

4.3 Base Schedulers

4.3.1 Round Robin

RR scheduler is one of the classical schedulers that have been used in many legacy systems. RR scheduler is used here as a baseline for evaluating the schedulers implemented in our study.

Algorithm 1 Round Robin (RR)

1: Divide available RBs into groups of Resource Chunks according to $\lfloor \frac{|R_B|}{|U|} \rfloor$
2: For the remaining RBs from $\mod \left(\frac{|R_B|}{|U|} \right)$, distribute them among RBs
3: Distribute RCs among available UEs in an even fashion

4.4 Best-Effort Schedulers

4.4.1 Greedy Scheduler

The Greedy algorithm is considered an FTP allocation scheme, as discussed in Chapter 3, where PRBs are grouped into RCs, with each RC containing a set of contiguous PRBs as mentioned earlier. Each RC gets assigned to a UE whose metric at that RC is the highest. Afterwards, both the UE and its assigned RC get removed from the

schedulable UEs and schedulable RCs lists, respectively [35]. The algorithm performs very similarly to RR in the way that both distribute PRBs among schedulable UEs equally, except that Greedy uses channel dependent scheduling.

The Greedy algorithm uses a PF-based metric as a scheduling metric that is assigned to each PRB, and hence RC. The algorithm aims to maximize the fairness in resource allocation among UEs according to (4.3)

$$U = \sum_{u \in \mathcal{U}} \ln R(u) \tag{4.3}$$

Using the gradient rule, the utility function in (4.4) is used to derive the M^{FD} metric as follows

$$M^{FD}(u, r) = \frac{\hat{R}(u, r)}{\bar{R}(u)} \tag{4.4}$$

where $\hat{R}(u, r)$ and $\bar{R}(u)$ are explained according to Table 4.1.

Hence, the scheduler creates a $N \times N_{PRB}$ PF-metric matrix M, as described in Section 4.1.1, which is passed down to the FDPS allocation scheme in Algorithm 2. No TD metric was explicitly presented as part of Greedy's algorithm design.

Algorithm 2 Greedy Algorithm

1: Divide available PRBs into groups of Resource Chunks according to $\min\left(\lfloor \frac{N_{PRB}}{N} \rfloor, 1\right)$
2: Find the UE-RC pair with the highest metric.
3: Allocate RC to corresponding UE
4: Remove UE and RC from available UEs and RCs lists, respectively.
5: **Repeat** the above steps until all RCs are being allocated
6: Convert UE-RC allocation to UE-RB allocation

It is worth noting that the work in [35] presented binary search tree algorithm along with Greedy algorithm presented here. However, it was not possible to conduct

some of simulation experiments using search-tree based scheduler due to its very high execution time. As a result, the decision was to adopt Greedy algorithm as the algorithm of choice from [35] instead.

4.4.2 Recursive Maximum Expansion (RME)

RME is a PF-based scheduler that performs resource allocation using the *first maximum* search. That is, in each iteration the scheduler searches the for the UE with the maximum metric at a schedulable PRB. Once located, the scheduler expands the PRB allocation on both sides, in frequency domain, till the scheduler finds no more PRBs whose maximum metrics belongs to the given UE, at which point the scheduler removes the UE from the unscheduled UE list and places it in the already *served* list. The scheduler performs the same steps recursively for each UE till either all PRBs are scheduled or all UEs are served. In the case that all UEs are served but there still some PRBs unscheduled yet, the scheduler assigns PRBs to UEs that have neighboring PRBs already assigned to them.

As in the case of Greedy algorithm, RME uses the same PF metric to assign weights to PRBs for each UE according to (4.4).

Algorithm 3 Recursive maximum Expansion (RME) Algorithm [33]

Let $\mathcal{U}_{serv}$ be the list of users that have already been served by scheduler
$\mathcal{U}_{serv} \leftarrow \varnothing$
for all $u \in \mathcal{U}$ **do**
    Calculate FD metric of UE u, $M(u, r)$, $\forall r \in \mathcal{I}$ according to (4.4)
end for
while $\mathcal{I} \neq \varnothing$ and $\mathcal{U} \neq \varnothing$ **do**
    Find UE $u \in \mathcal{U}$ and PRB $r \in \mathcal{I}$ with maximum metric in M
    Assign PRB r to user u: $\mathcal{I}_u \leftarrow \mathcal{I}_u \cup \{r\}$
    $\mathcal{I} \leftarrow \mathcal{I} \setminus \{r\}$
    $M(u, r) \leftarrow 0$
    Let $r_l \leftarrow r - 1$, $r_r \leftarrow r + 1$
    while $M(u, r_l) = \max(M(:, r_l))$ **and** $r_1 \geq 1$ **do**
        $\mathcal{I}_u \leftarrow \mathcal{I}_u \cup \{r_l\}$; $\mathcal{I} \leftarrow \mathcal{I} \setminus \{r_l\}$; $r_l \leftarrow r_l - 1$
    end while
    while $M(u, r_r) = \max(M(:, r_r))$ **and** $r_r \leq N_{PRB}$ **do**
        $\mathcal{I}_u \leftarrow \mathcal{I}_u \cup \{r_r\}$; $\mathcal{I} \leftarrow \mathcal{I} \setminus \{r_r\}$; $r_r \leftarrow r_r + 1$
    end while
    $\mathcal{U} \leftarrow \mathcal{U} \setminus \{u\}$
    $\mathcal{U}_{serv} \leftarrow \mathcal{U}_{serv} \cup \{u\}$
end while
    ▷ Check if there are any resources left. If yes, assign them to adjacent UEs
while $\mathcal{I} \neq \varnothing$ **do**
    Assign PRBs to $u \in \mathcal{U}_{serv}$ such that the contiguity criteria is maintained
end while

4.5 QoS-Based Schedulers

4.5.1 Proportional Fairness with Guaranteed Bit Rate (PFGBR)

The utility function that PFGBR aims to maximize is the PF utility function defined in 4.3. However, as the name implies, PFGBR combines PF and GBR metrics to provide fair scheduling with QoS provisioning at the same time. Hence, the PF gradient-based utility function described in (4.4) is modified to account for the GBR requirements of the GBR-based traffic. Also, since the algorithm performs a UE-to-PRB-subset mapping rather than UE-to-PRB mapping, the FD metric used in FDPS scheduling is expressed as

$$M^{FD}(u,c) = \begin{cases} \exp\left(\alpha \cdot \left(R_{GBR}(u) - \bar{R}(u)\right)\right) \cdot \frac{\hat{R}(u,c)}{\bar{R}(u)} & \text{if } u \in \mathcal{U}_{GBR} \\ \frac{\hat{R}(u,c)}{\bar{R}(u)} & \text{if } u \in \mathcal{U}_{non-GBR} \end{cases} \tag{4.5}$$

where $\hat{R}(u,c)$ is the estimated achievable throughput over the PRB-subset $\mathcal{I}^{C}_{u,c}$.

The algorithm uses a PF metric for UEs that carry non-GBR traffic, $\mathcal{U}_{non-GBR}$. In the case of UEs with GBR traffic, $\mathcal{U}_{GBR}$, the PF metric is combined with an exponential weight that detects the difference between the UE's average throughput and its GBR requirement. Such GBR term provides higher weight to GBR UEs such that the scheduler ensure that their GBR requirements are satisfied.

As in the case with Greedy and RME schedulers, PFGBR scheduler was proposed in [36] with no explicit TDPS scheduling. Hence, the scheduler uses (4.5) to construct the $N \times N_C$ metric matrix M, with each column representing the metrics of all UEs for a given allocation pattern $c \in \mathcal{C}_u$. Consequently, the scheduler passes the matrix

M to its greedy-based allocation algorithm to perform UE-to-PRB assignment.

Algorithm 4 Proportional Fairness with Guaranteed Bit Rate Algorithm (PFGBR) [36]

Define PRB allocation matrix E according to (4.2)
Define M as a UE-to-PRB allocation pattern metric mapping matrix
for all $u \in \mathcal{U}$ **do**
    Calculate FD metric of UE u, $M(u,c)$, $\forall c \in \mathcal{C}_u$ according to (**??**)
    if $\exists c \in \mathcal{C}_u$ such that $M(u,c) > R_{MBR}(u)$ **then**
        $M(u,c) \leftarrow 0$
    end if
end for
while $\mathcal{I} \neq \varnothing$ **and** $\max(M) > 0$ **do**
    Find UE u and PRB allocation $u \in \mathcal{C}_u$ such that: $M(u,c) = \max(M)$
    $\mathcal{I}_u \leftarrow \mathcal{I}^C_{u,c}, \quad \mathcal{I} \leftarrow \mathcal{I} \setminus \mathcal{I}_u$
    for all $k \in \mathcal{C}_u$ such that $\mathcal{I}^C_{u,k} \cap \mathcal{I}^C_{u,c} \neq \varnothing$ **do**
        $M(:,k) \leftarrow 0$
        $\mathcal{I} \leftarrow \mathcal{I} \setminus \mathcal{I}^C_{u,k}$
    end for
end while
if $\mathcal{I} \neq \varnothing$ **then**
    Assign remaining PRBs randomly to UEs such that it maintains the contiguity constraint
end if

4.5.2 Guaranteed Bit Rate with Adaptive Transmission Bandwidth (GBR-ATB)

The authors [37] has proposed their scheduling algorithm accompanied by a QoS-aware RAC scheme, where the RAC algorithm admits new UEs as long as the network can satisfy their GBR. Afterwards, the scheduler allocates resources to newly and

already admitted UEs based on GBR and PF criteria. As described later on in Chapter 5, no RAC is assumed for the LTE uplink simulator, as the number of UEs remain constant for the duration of the simulator. Hence, only the scheduling algorithm was implemented here.

Unlike schedulers presented so far, GBR-ATB scheduler was proposed with explicit definition of TDPS and FDPS scheduling. In TDPS, UEs are prioritized according to TD metric as defined in (4.6).

$$M^{TD} = \frac{R_{GBR}(u)}{\bar{R}(u)} \tag{4.6}$$

no UE subset size limitation was mentioned in [37], and hence all UEs with active uplink transmission are passed down to the FDPS scheduler. FDPS scheduler creates a $N \times N_{PRB}$ matrix M based on FD metric which is defined as follows

$$M(u,r) = \frac{R_{GBR}(u)}{\bar{R}(u)} \cdot \frac{\hat{R}(u,r)}{\bar{R}_{sch}(u)} \tag{4.7}$$

The variable $\bar{R}_{sch}(u)$ is calculated as the average throughput experienced by UE u for TTIs within which the UE is assigned PRBs. Afterwards, the matrix M is passed down to the resource allocation scheme described in Algorithm 5 for resource assignment.

Algorithm 5 Guaranteed Bit Rate with Adaptive Transmission Bandwidth (GBR-ATB) [37]

Let U_{serv} be the list of users that have already been served by scheduler
$U_{serv} \leftarrow \varnothing$
Let $P_{tx}(u, \mathcal{I}_u)$ be the power budget of UE u with allocation $\mathcal{I}_u$
for all $u \in \mathcal{U}$ **do**
    Calculate FD metric of UE u, $M(u, r)$, $\forall r \in \mathcal{I}$ according to (4.7)
end for
while $\mathcal{I} \neq \varnothing$ **and** $\mathcal{U} \neq \varnothing$ **do**
    Find UE $u^* \in \mathcal{U}$ and PRB $r^* \in \mathcal{I}$ such that $M(u^*, r^*) = \max(M(u, r))$, $\forall u \in \mathcal{U}, \forall r \in \mathcal{I}$
    $\mathcal{I}_{u^*} \leftarrow \mathcal{I}_{u^*} \cup \{r^*\}$, $\mathcal{I} \leftarrow \mathcal{I} \setminus \{r^*\}$
    Let $r_l \leftarrow r^* - 1$, $r_r \leftarrow r^* + 1$, finish $\leftarrow 0$
    while finish $= 0$ **and** ($(M(u^*, r_l) = \max(M(u, r_l))$ **and** $r_1 \geq 1$) **or**
      $(M(u^*, r_r) = \max(M(u, r_r))$ **and** $r_r \leq N_{RB}$)), $\forall u \in \mathcal{U}$ **do**
        Choose PRB r_{opt} such that: $M(u^*, r_{opt}) = \max(M(u^*, r_l), M(u^*, r_r))$
        if $P_{tx}(u^*, \mathcal{I}_u \cup \{r_{opt}\}) > P_{tx,MAX}$ **then** finish $\leftarrow 1$
        else
            $\mathcal{I}_{u^*} \leftarrow \mathcal{I}_{u^*} \cup \{r_{opt}\}$; $\mathcal{I} \leftarrow \mathcal{I} \setminus \{r_{opt}\}$
            if $r_{opt} \leftarrow r_l$ **then** $r_l \leftarrow r_l + 1$;
            else $r_r \leftarrow r_r + 1$;
            end if
        end if
    end while
    $\mathcal{U} \leftarrow \mathcal{U} \setminus \{u^*\}$, $\mathcal{U}_{serv} \leftarrow \mathcal{U}_{serv} \cup \{u^*\}$
end while
while $\mathcal{I} \neq \varnothing$ **do**
    Assign PRBs to UEs such that the contiguity criteria and power constraints are maintained
end while

4.5.3 Multi-Carrier Scheduling Algorithm (MC-SA)

The Multi-Carrier Scheduling Algorithm is another QoS-aware scheduler that was proposed in [39]. A main distinguishing feature from PFGBR and GBR-ATB is that it relies on two QoS parameters rather than just one, GBR and packet delay budget. The algorithm proposed in [39] aims to maximize a QoS-based PF utility function which is shown in

$$U = \sum_{u \in \mathcal{U}} \sum_{r \in \mathcal{I}} \alpha(u, r) \cdot F(u) \tag{4.8}$$

where $\alpha(u, r)$ equals 1 if PRB r is assigned to UE u, and 0 otherwise.

$$F(u) = \frac{D_{MAX}(u)}{D^{Q}_{ave}(u)} \cdot \frac{\bar{R}(u)}{R_{GBR}(u)} \tag{4.9}$$

As in the case of GBR-ATB, the algorithm here was proposed with explicit definition of the TDPS and FDPS operations, with no proposed limitation on the number of UEs that can be passed down to the FDPS scheduler. However, the authors in [39] combine the TDPS and FDPS scheduling operation rather than execute both of them separately. The TDPS and FDPS scheduling is more apparent when the number of UEs become more than the number of PRBs in the system. In this case, the scheduler starts by determining the UEs priorities according to (4.9). Afterwards, rather than passing down the UEs list to the FDPS scheduling, the scheduler traverses through the prioritized UEs list, one UE at a time. In each iteration, the scheduler assigns a UE one or more PRBs, then moves down the UE priority list to schedule for the next UE. The number of PRBs assigned to a UE is based on its experienced average throughput relative to the UE's GBR requirement. The scheduler stops traversing

through the prioritized UE list once all PRBs get allocated.

Algorithm 6 Multi-Carrier Scheduling Algorithm (MC-SA) [39]

for all UE $u \in \mathcal{U}$ **do**
  Calculate $F(u)$, according to (4.9)
  if $\bar{D}(u) < D_{MAX}(u)$ **then** $F(u) \leftarrow 0$
  end if
end for
if $N < N_{PRB}$ **then**
  MCSA-SUB1 (M)
else
  MCSA-SUB2 (M)
end if

Algorithm 7 MC-SA subroutine 1

procedure MCSA-SUB1(M)
  Determine RC size: $m = \left\lfloor \frac{N_{PRB}}{N} \right\rfloor$
  Determine RC size with residual blocks: $m' = m + \mod\left(\frac{N_{PRB}}{N}\right)$
  Find UE $u^* \in \mathcal{U}$ with highest priority such that: $F(u^*) = \min(F)$
  Find PRB $r^* \in \mathcal{I}'_u$ such that: $\hat{R}(u^*, r^*) \leftarrow \max\left(\hat{R}(u^*, r)\right), \forall r \in \mathcal{I}$
  Create $N \times N_C$ matrix $\hat{R}_C$ from $\hat{R}$ such that:
    RC j^* has m' PRBs and: $\hat{R}_C(u^*, j^*) = \max(R_C)$
  Let $\mathcal{I}_j^{RC}$ represent all PRBs belonging to RC j
  while $\mathcal{U} \neq \varnothing$ **do**
    Find the $u \in \mathcal{U}$ and RC j such that: $M_C(u, j) = \max(M_C)$.
    $\forall r \in \mathcal{I}_j^{RC}$: $\mathcal{I}_u \leftarrow \mathcal{I}_u \cup \{r\}$, $\mathcal{I} \leftarrow \mathcal{I} \setminus \{r\}$
    $F(u^*) \leftarrow \infty$, $\mathcal{U} \leftarrow \mathcal{U} \setminus \{u^*\}$
  end while
end procedure

Algorithm 8 MC-SA subroutine 2

procedure MCSA-SUB2
  Let $\hat{R}_{PRB}(u)$ be the average estimated throughput of $u \in \mathcal{U}$ per PRB
  Calculate $\hat{R}_{PRB}(u)$, $\forall u \in \mathcal{U}$
  while $\min(F) < \infty$ **and** $\mathcal{I} \neq \varnothing$ **do**
    Find UE $u^* \in \mathcal{U}$ such that: $F(u^*) = \min(F)$
    if $\bar{R}(u^*, r^*) < R_{GBR}(u^*)$ **then**
      Determine number of PRBs to assign u^*: $n_{PRB} = \left\lceil \frac{R_{GBR}(u^*)}{\bar{R}_{PRB}(u^*)} \right\rceil$
    else
      $n_{PRB} \leftarrow 1$
    end if
    Let $x \leftarrow 1$, $\mathcal{I}'_{u^*} \leftarrow \mathcal{I}$
    while $r_x \leq n_{PRB}$ **and** $\mathcal{I}'_{u^*} \neq \varnothing$ **do**
      Find PRB $r^* \in \mathcal{I}'_{u^*}$ such that: $\hat{R}(u^*, r^*) \leftarrow \max\left(\hat{R}(u^*, r)\right)$, $\forall r \in \mathcal{I}'_{u^*}$
      $\mathcal{I}_{u^*} \leftarrow \mathcal{I}_{u^*} \cup \{r^*\}$, $\mathcal{I} \leftarrow \mathcal{I} \setminus \{r^*\}$
      $\mathcal{I}'_{u^*} \leftarrow \{\min(\mathcal{I}_{u^*}) - 1, \max(\mathcal{I}_{u^*}) + 1\} \cap \mathcal{I}$
      $x \leftarrow x + 1$
    end while
    $\mathcal{U} \leftarrow \mathcal{U} \setminus \{u\}$
  end while
end procedure

4.6 Power-Optimizing Schedulers

4.6.1 Block Allocation for Minimum Total Power (BMTP)

The work presented in [41] presents one of the few scheduling schemes that have been presented to account for transmission power conservation over uplink SC-FDMA radio interface. The idea behind the algorithm is that it tries to determine all the allocation patterns that are possible for a UE. Then, the scheduler tries to estimate the amount

of transmission power over each PRB assignment pattern such that it minimizes the transmission rate to a minimal level that satisfies the GBR requirement of each UE. That is, if $p_{tx}(u)$ represents the amount of power assigned to UE u for uplink transmission, the utility function of BMTP scheduler can be expressed as

$$U = \sum_{u \in \mathcal{U}} p_{tx}(n), \quad \text{satisfying } R(u) \geq R_{GBR}(u), \forall u \in \mathcal{U} \tag{4.10}$$

Algorithm 9 BMTP: Part 1: Initialization

for all UE $u \in \mathcal{U}$ **do**
  $p_{tx}(u) \leftarrow \infty$, $\mathcal{I}_u \leftarrow \varnothing$, $\mathcal{I}'_u \leftarrow \mathcal{I}$
  for all $c \in \mathcal{C}_u$ **do**
    Calculate Tx power for u to transmit on $\mathcal{I}^C_{u,c}$: $M(u,c)$
    Set $M(u,c) \leftarrow \infty, \forall M(u,c) > P_{tx,MAX}$
  end for
end for
finish $\leftarrow 0$
if there exists $u \in \mathcal{U}$ with $c \in \mathcal{C}_u$ such that: $M(u,c) < \infty$ **then**
  for all UE $u \in \mathcal{U}$ **do**
    Find PRB-allocation $c \in \mathcal{C}_u$ with minimum number of PRBs, $B_{min}(u)$, such that:
    $B_{min}(u) \leftarrow \min\left\{|\mathcal{I}^C_{u,c}| : M(u,c) < \infty\right\}$
    if $\sum_{u\in\mathcal{U}} B_{min}(u) > N_{PRB}$ **then**
      finish $\leftarrow 1$
    else
      Find the maximum number of PRB allocations per UE u
      $B_{max}(u) \leftarrow N_{PRB} - \sum_{n\neq u\in\mathcal{U}} B_{min}(n)$
      for all $c' \in \mathcal{C}_u$ such that $|\mathcal{I}^C_{u,c'}| > B_{max}(u)$ **do**
        $M(u,c') \leftarrow \infty$
      end for
    end if
  end for
else
  finish $\leftarrow 1$
end if

Algorithm 10 BMTP: Part 2: Scheduling

Call initialization routine in Algorithm 9
while finish $\leftarrow 0$ **do**
    for all UE $u \in \mathcal{U}$ **do**
        Determine the 1^{st} and 2^{nd} smallest power allocations for all $c \in \mathcal{C}_u$: $e_{u,1}$ and $e_{u,2}$
    end for
    Select UE $u^* \in \mathcal{U}$ that satisfies $\min(e_{u,2} - e_{u,1})$
    if $e_{u^*,1} < \infty$ **then**
        Find $c^* \in \mathcal{C}_{u^*}$ such that: $M(u^*, c^*) = e_{u^*,1}$
        $\mathcal{I}_{u^*} \leftarrow \mathcal{I}^C_{u^*,c^*}$
        $\mathcal{I} \leftarrow \mathcal{I} \setminus \mathcal{I}_{u^*}$
        $\mathcal{I}'_{u^*} \leftarrow \{\min(\mathcal{I}_{u^*}) - 1, \max(\mathcal{I}_{u^*}) + 1\} \cap \mathcal{I}$
        $p_{tx}(u^*) \leftarrow e_{u^*,1}$
        for all UE $u \in \mathcal{U}$ **do**
            for all $c \in \mathcal{C}_u$ such that $\mathcal{C}_u \cap \mathcal{I}_{u^*} \neq \varnothing$ **do**
                $M(u, c) \leftarrow \infty$
                $\mathcal{I}'_u \leftarrow \mathcal{I}'_u \cap \mathcal{I}$
            end for
        end for
        $\mathcal{U} \leftarrow \mathcal{U} \setminus \{u^*\}$
    else
        finish $\leftarrow 1$
    end if
end while
if $\mathcal{I} \neq \varnothing$ **then**
    Execute **MPD** procedure
end if

Algorithm 11 MPD

procedure MPD(Inherits all variables from Algorithm 10)
  Let $\Delta p_{tx}(u, r)$ be the change in power allocation if a PRB r is assigned to UE u
  finish $\leftarrow 0$
  for all UE $u \in \mathcal{U}$ **do**
    for all PRB $r \in \mathcal{I}'_u$ **do**
      Find $c \in \mathcal{C}_u$ such that: $\mathcal{I}^C_{u,c} = \mathcal{I}_u \cup \{r\}$
      $\Delta p_{tx}(u, r) \leftarrow p_{tx}(u) - M(u, c)$
    end for
  end for
  while finish $\leftarrow 0$ **do**
    Find UE u^* and PRB r^* such that: $\Delta p_{tx}(u^*, r^*) = \max(\Delta p_{tx})$
    if $\Delta p_{tx}(u^*, r^*) > 0$ **then**
      Find $c^* \in \mathcal{C}_{u^*}$ such that: $\mathcal{I}^C_{u^*,c^*} = \mathcal{I}_{u^*} \cup \{r^*\}$
      $p_{tx}(u^*) \leftarrow M(u^*, c^*)$
      $\mathcal{I} \leftarrow \mathcal{I} \setminus \{r^*\}$
      $\mathcal{I}'_{u^*} \leftarrow \{\min(\mathcal{I}_{u^*}) - 1, \max(\mathcal{I}_{u^*}) + 1\} \cap \mathcal{I}$
      $\Delta p_{tx}(u, r^*) \leftarrow 0$, for all UE $u \in \mathcal{U}$
      $\Delta p_{tx}(u^*, r) \leftarrow 0$, for all PRB $r \in \mathcal{I}$
      for all $u \in \mathcal{U}$ **do**,
        $\forall r \in \mathcal{I}'_u \cap \mathcal{I}^C_{u,c}$, $\Delta p_{tx}(u^*, r) \leftarrow p_{tx}(u^*) - M(u^*, c)$
      end for
    else
      finish $\leftarrow 1$
    end if
  end while
end procedure

Chapter 5

Performance Analysis

This chapter describes the experimental work done to evaluate the schedulers' performance, and discusses the results obtained from experiments conducted using the LTE uplink simulator developed. The design of the LTE uplink simulator is discussed in detail, covering different aspects such as network topology, channel model, link adaptation model, and traffic types simulated. Simulation experiments were setup based on the developed simulator such that they measure different aspects of LTE uplink performance under different scenarios. The metrics used for performance evaluation include frame utilization, fairness between different traffic classes, fairness between UEs within the same traffic class, frame utilization, and QoS metrics such as packet delay and packet loss.

5.1 Simulation Setup

An LTE uplink simulator was designed and developed in MATLAB to have a unified, standard-compliant simulation settings which provides a valid comparison among

different proposals of LTE uplink scheduling algorithms. The simulator design focused on having the necessary modules that fully simulate the events occurring in LTE uplink while having it easily extendible for future work.

Before going into explaining the LTE simulator, we first discuss the system setup assumed when designing the simulator and the different models used as well.

5.1.1 System Topology

Figure 5.1 provides a block diagram illustration of the LTE uplink system topology developed in MATLAB. The system represents the LTE uplink transmission within a single-cell environment with single-sector with 1km $\times$ 1km grid area. The cell's topology consists of one eNodeB at the center of the cell, with UEs located around the eNodeB in a spatially uniform pattern.

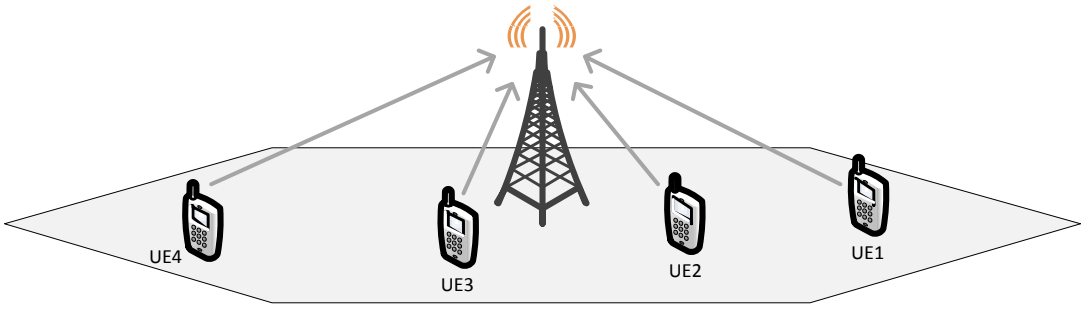

Figure 5.1: System topology of LTE simulated LTE system

Every UE is equipped with a single transmit antenna, while the eNodeB is equipped with two received antennas. Hence, the uplink transmission is assumed to have a 1×2 Single-Input-Multiple-Output (SIMO) antenna configuration. By having a single transmit antenna at the UE end, the uplink transmission is assumed to be a

single stream with transmit/receive diversity. Transmit/receive diversity refers to using both antennas at the eNodeB to receive the same uplink stream from a UE, which improves the received SINR and hence increases the power efficiency of uplink transmission.

No RAC is implemented at the eNodeB, since the main focus of the performance evaluation is on uplink scheduling. All UEs are created at the beginning of each simulation run, and remain active for the entire simulation duration. UEs are presumed to be stationary for the whole simulation time as well to observe the maximum performance level possible under the operation of each scheduler.

5.1.2 Wireless Radio Channel Model

The operating uplink bandwidth is set to 5 MHz with FDD configuration. The uplink bandwidth is hence divided into twenty-five PRBs as specified in the standard [3]. Twenty four PRBs constitute the physical uplink shared channel (PUSCH), which is used as a shared medium among UEs for data transmission. The remaining one PRB is assumed to be reversed for uplink control channels, and is chosen to be at the end of the spectrum to ensure contiguity of available PUSCH resources.

The modeling of the wireless propagation channel is broken down to a macroscopic model and a microscopic model. The macroscopic channel model describes large-scale channel variations that depend on the UE's spatial position relative to the eNodeB, such as path loss, shadowing, and penetration loss. The microscopic channel model mainly focuses on describing fast variations of the channel gain, which mainly entails multipath fading. Multi-path effect can either be slow fading or fast fading, depending on the UE mobility (i.e. the Doppler Effect).

5.1.2.1 Macroscopic Channel Model

The macroscopic propagation model we chose here is to represent propagation path loss , PL, in the typical urban setup as found in [42], and it is expressed as

$$PL[\text{dB}] = I + \alpha \cdot 10 \log_{10}(d[\text{km}]) + \chi_0 + P_{loss} \tag{5.1}$$

where I captures the free space propagation loss, which is equal to 128.1 dB at 2 GHz frequency. α is the path loss exponent, which is experimentally determined to be 3.76 in typical urban environment. d is the distance between UE and eNodeB in km.χ_0 is a random variable that represents the shadowing effect in the typical urban, which follows a lognormal distribution with mean of 0 dB and standard deviation of 8 dB. P_{loss} is the penetration loss caused by signal penetration through obstacles, which is assumed to be a constant of 20 dB.

5.1.2.2 Microscopic Channel Model

We used a Tapped Delay Line (TDL) to model the microscopic effects of the channel. The microscopic channel model follows the 12-tap Typical Urban (TU) Power Delay Profile (PDP) for pedestrian UEs as described in [43].

TDL taps are generated once every TTI, after which the time-domain delay lines are converted to frequency domain using the Fast Fourier Transform (FFT), resulting in frequency response with gain taps separated by 15 kHz (subcarrier spacing). The resulting taps are normalized such that the expected value of the summation of the normalized PDP taps becomes 1.

5.1.3 Link Adaptation Model

Based on both the macroscopic and microscopic parts of the channel model, the CSI per SC-FDMA subcarrier, for each UE, is used to represent the path gain experienced by UE u at subcarrier k, $CSI_{u,k}$, and is determined as follows:

$$CSI_{u,k} = \frac{G_{UE} \cdot G_{eNB} \cdot (|H_{u,k,1}|^2 + |H_{u,k,2}|^2)}{PL_u \cdot \sigma_n^2 \mathcal{N} \Delta f} \tag{5.2}$$

where G_{UE} is the UE's antenna gain, G_{eNB} is the eNodeB's antenna gain, PL_u is the power loss experience by user u (equation (5.1)), σ_n^2 is the noise density per Hz, $\mathcal{N}$ is the noise figure of the receiver at eNodeB, and Δf is the subcarrier spacing. The terms $|H_{u,k,1}|^2$ and $|H_{u,k,2}|^2$ refer to the normalized multipath gains at eNodeB's receive antenna 1 and antenna 2, respectively. The summation of both gains is based on the transmit/receive diversity assumption stated earlier.

The link adaptation model is then used in our simulator to predict the appropriate MCS to use when transmitting the data on assigned resource blocks. Once the packet scheduler assigns UEs their corresponding PRBs, the effective SINR, γ_u, is determined for the assigned PRBs assuming a Minimum Mean Squared Error (MMSE) receiver [44],

$$\gamma_u = \left(\frac{1}{\frac{1}{N_u} \sum_{k=1}^{N_u} \frac{\gamma_{u,k}}{\gamma_{u,k}+1}} - 1 \right)^{-1} \tag{5.3}$$

where $\gamma_{u,k}$ is the SINR for UE u at subcarrier k, and N_u is the number of contiguous subcarriers assigned to UE u. $\gamma_{u,k}$ is determined per subcarrier k as follows

$$\gamma_{u,k} = \frac{P_u}{N_u} \cdot CSI_{u,k} \tag{5.4}$$

where P_u is the total transmit power assigned to UE u by the eNodeB.

Determining the UE's total uplink transmit power is based on the Open Loop Power Control (OLPC) mechanism where the LA process aims to compensate for macroscopic effects experienced in the uplink channel between a UE and the eNodeB. Hence, the UE transmit power can be calculated according to [45]

$$P = \min\left(P_{t,MAX}, P_0 + 10 \cdot \log_{10} N_{u,PRB} + \alpha \cdot PL_u\right) \quad [dBm] \tag{5.5}$$

Where $P_{t,MAX}$ is the maximum UE transmission power set to 24 dBm as shown in Table 5.1, $N_{u,PRB}$ is the number of PRBs allocated to UE u, PL_u is the path loss expressed in (5.1), and P_0 and α are both cell-specific parameters, where P_0 is the base transmit power per PRB, and α is the fraction of path loss to be compensated for. The value of $P_{t,MAX}$ was set based on UE profile settings defined in Table 4.8 in [42]. For simplicity, UE's total uplink transmission power is assumed to be allocated for data transmission only.

Afterwards, the effective SINR calculated in (5.3) is mapped to pre-determined SINR ranges to determine the MCS to be used for data transmission. Based on the MCS selected and the number of PRBs assigned to the UE, the Transport Block (TB) size is calculated in bits. A TB is composed of the MAC packet data unit (includes the MAC header, RLC header, and the data payload) and a 24-bit Cyclic Redundancy Check (CRC) for error detection.

Table 5.1 summarizes the simulation parameters chosen for our LTE uplink packet-level simulator.

Table 5.1: System Simulation Parameters

Cellular Layout	Single-Cell with Omnidirectional Antenna
System Bandwidth	5 MHz
Carrier Frequency	2 GHz
Number of Resource Blocks	25
TTI Duration	1 ms
Path Loss Model	$128.1 + 37.6\log_{10}(d[\text{km}])$
Penetration Loss	20 dB
Shadowing	Lognormal: $\mu = 0$, $\sigma = 8$dB
Minimum Distance Between UE and Cell	90 m
Power Delay Profile	TU12 Profile, 12 taps
Channel Estimation	Ideal
MCS Settings	QPSK [1/6 1/4 1/3 1/2 2/3 3/4] 16QAM [1/2 2/3 3/4]
HARQ Process	OFF
eNodeB Antenna Gain	15 dBi
UE Antenna Gain	0 dBi
eNodeB Noise Figure	5 dB
Max. UE Transmit Power	24 dBm
Power Compensation	$P_0 = -58$ dBm, $\alpha = 0.6$
UE Speed	0 km/h
Frequency Re-use Factor	1
Simulation Time	10000 TTIs

5.1.4 Traffic Model

The traffic models developed for the LTE uplink simulator are adopted from [46], and are to be explained in more detail shortly. The QoS-based packet parameters for each traffic type are based on the QCI parameters illustrated in Table 2.1. The PDB value defined in 2.1 represents the maximum packet delay allowed between the UE and PCEF in the network core. According to NOTE 1 in Table 6.1.7 in [1], the offset in packet delay between eNodeB and PCEF ranges between 10 ms up to 50 ms if the PCEF is far away from eNodeB. Accordingly, we chose to set the PDB offset to 50

ms as to drive the access network to perform in tight delay scenarios where PCEF entity is furthest away from the access network.

In addition, to better evaluate the performance of the scheduler and its distinction between the different traffic types, each UE is assumed to carry a single traffic stream.

5.1.4.1 VoIP Traffic

The VoIP traffic of choice in our work was chosen to be 12.2 kbps AMR VoIP service with silence suppression [46]. In the active state, the VoIP source is to generate a 40 bytes VoIP packet once every 20 ms subframe (or once every two LTE frames), with each packet consisting of 244 bits of payload data and the remaining space consisting of the packet's overhead. Due to the limitation of the simulation time of our experiments to 10 seconds only, we have decided to fixate all VoIP streams at the active state to regulate the offered VoIP traffic load.

5.1.4.2 Video Streaming Traffic

We have chosen the video streaming traffic modeling presented in [46], which represents a low quality video stream running at a minimum guaranteed bit rate of 64 kbps. When system capacity permits, the uplink video streaming traffic load is allowed to increase per UE up to a maximum of 1024 kbps.

5.1.4.3 FTP Traffic

In the FTP traffic case, we have chosen to implement the traffic generator with a CBR-based packet stream instead of the FTP traffic model described in the standard [46]. CBR-based FTP traffic model aids our study in creating a fixed offered traffic

load per UE to better map between the offered FTP traffic load per UE against its measured throughput. A constant packet size of 256 bytes with a traffic rate of 128 kbps per FTP connection has been chosen in our case. MBR of the offered FTP traffic load per UE was set to 1024 kbps. As a Best Effort traffic type, FTP has no GBR nor delay requirements associated with it. However, as some of the schedulers implemented here require a GBR value for their operation, such as PFGBR and MC-SA, a GBR of 10 kbps in this case is assumed for each FTP connection.

Table 5.2 summarizes the simulation parameters of the traffic models used.

Table 5.2: Traffic Models Used in Simulation Experiments.

		QCI Parameters					
Traffic	Traffic Class	QCI#	Priority	Type	PDB (ms)	GBR (kbps)	MBR (kbps)
VoIP	Conversational	1	2	GBR	50	12.2	64
Video Streaming	Streaming	2	4	GBR	100	64	512
FTP	Background	6	6	non-GBR	—	10	1024

5.1.5 Performance Metrics

The following metrics are measured in order to evaluate several aspects of the system performance under the operation of different uplink schedulers.

- **The cell's aggregate throughput**, which is measured as

$$\bar{T}_{cell} = \frac{B}{t_{sim}} \tag{5.6}$$

Where B is the total number of bits successfully transmitted over the air interface from the UE up to the eNodeB, t_{sim} is the total simulation time.

- **Intra-Class Fairness**, which represents the fairness among UEs of the same class. The Intra-class fairness is calculated using the min-max fairness index. Assuming UE i is the one with the maximum throughput, while UE j is the user with the minimum throughput, the inter-class fairness index can be calculated as:

$$F_{min-max} = \frac{\bar{T}_i}{\bar{T}_j} \tag{5.7}$$

where $\bar{T}_i$ is the throughput of UE i, and $\bar{T}_j$ is the throughput of UE j.

- **Inter-Class Fairness**, which is a measure of the fairness among UEs with different traffic classes. The measurement of the Inter-class fairness is performed using the well-known Jain's Index, which is calculate as follows:

$$F_{Jain} = \frac{\left|\sum_{i=1}^{N_{UE}} x_i\right|^2}{N_{UE} \sum_{i=1}^{N_{UE}} x_i^2} \tag{5.8}$$

where x_i represents the normalized average throughput of user i. To achieve the interclass fairness between different QoS classes, the UE's average throughput is normalized with respect to the UE's maximum bit rate (MBR) that is defined for each traffic type.

- **Packet Loss**, which is a measure of the percentage of packets of a certain traffic class dropped of the transmission packet queue due to exceeding their packet

delay budget. The transmission buffer length in this experiment is assumed to be infinite to exclude packet drops due to packet buffer congestion.

- **Packet Delay**. The delay per packet is measured from the time the packet enters the RLC queue till the time it successfully arrives at the eNodeB. Packet delay is measured by collecting the delay stamps for all packets being sent within the entire simulation time, then determining the experienced average delay of all UEs within a given traffic class. The packet delay measurements, along with the measurements from the packet drops, can give us an indication on the ability of a scheduling algorithm to respect the QoS requirements of each traffic class with an active transmission within our simulator.

- **TB Utilization**. TB utilization refers to the averaged percentage of TB size used for transmitting the data payload. TB in LTE refers to PHY payload to be transmitted over the radio interface, which comprises the MAC packet plus a 24-bit CRC overhead. The TB utilization statistics are collected per UEs carrying traffic of the same QoS class, and averaged over the number of UEs belonging that traffic class. TB utilization provides a measure on how well a scheduler predicts the needs of each UE in the system, and its ability to minimize the resources wasted as a result from resource allocation mismanagement.

5.1.6 LTE Uplink Simulator

The LTE uplink system was designed as an event-driven, packet-level simulator using the Monte-Carlo method. The simulator runs in discrete time with time steps as small as one TTI. Event flow within the system is mainly controlled via the system

clock and event manager modules to ensure proper sequencing in executing system events.

UE and eNodeB modules were designed and developed to constitute the active network components within the LTE system. Figure 5.2 illustrates the hierarchy and the interaction between different simulation blocks. Each UE is linked to the eNodeB via a wireless channel pipe module which simulates the SC-FDMA-based wireless medium through which control and data communication from a UE to the eNodeB takes place. Each of the UE, eNodeB, and channel modules are structured as follows:

- eNodeB Module
  - Parameters
    - Spacial Position
    - List of UEs connected to the eNodeB
    - Antenna parameters, such as antenna gain and antenna pattern
  - Internal Modules
    - Packet scheduler
    - Link adaptation module
    - Two omnidirectional, receive antennas
    - Receive packet buffer for each connected UE
    - Sink for collecting statistics about received packets
    - Control unit that receives control messages and feedback from UEs, and passes them on to other modules within the eNodeB
- UE Module

  - Parameters
    - UE ID
    - Spacial Position
    - QoS profile of carried traffic
  - Internal Modules
    - Packet source
    - RLC transmission buffer
    - One omnidirectional, transmit antenna
    - Control unit

- Wireless Channel Module
  - Parameters
    - System bandwidth
    - Carrier Frequency
    - UE-eNodeB distance
    - Noise figure of wireless medium
    - Multipath power delay profile
    - Environment-based parameters: penetration loss, shadowing
  - Internal Modules
    - Path Loss Module
    - per-PRB CSI generator

Each of the UE and eNodeB modules contain a control module, shown in Figure 5.2, that emulates the transmission and reception of control messages between the

two modules, such as scheduling requests, scheduling grants, packet acknowledgments, and CSI reports. The control module communicates the control feedback received to the other internal modules within the internal blocks (e.g. packet schedulers, packet buffers, etc) when needed.

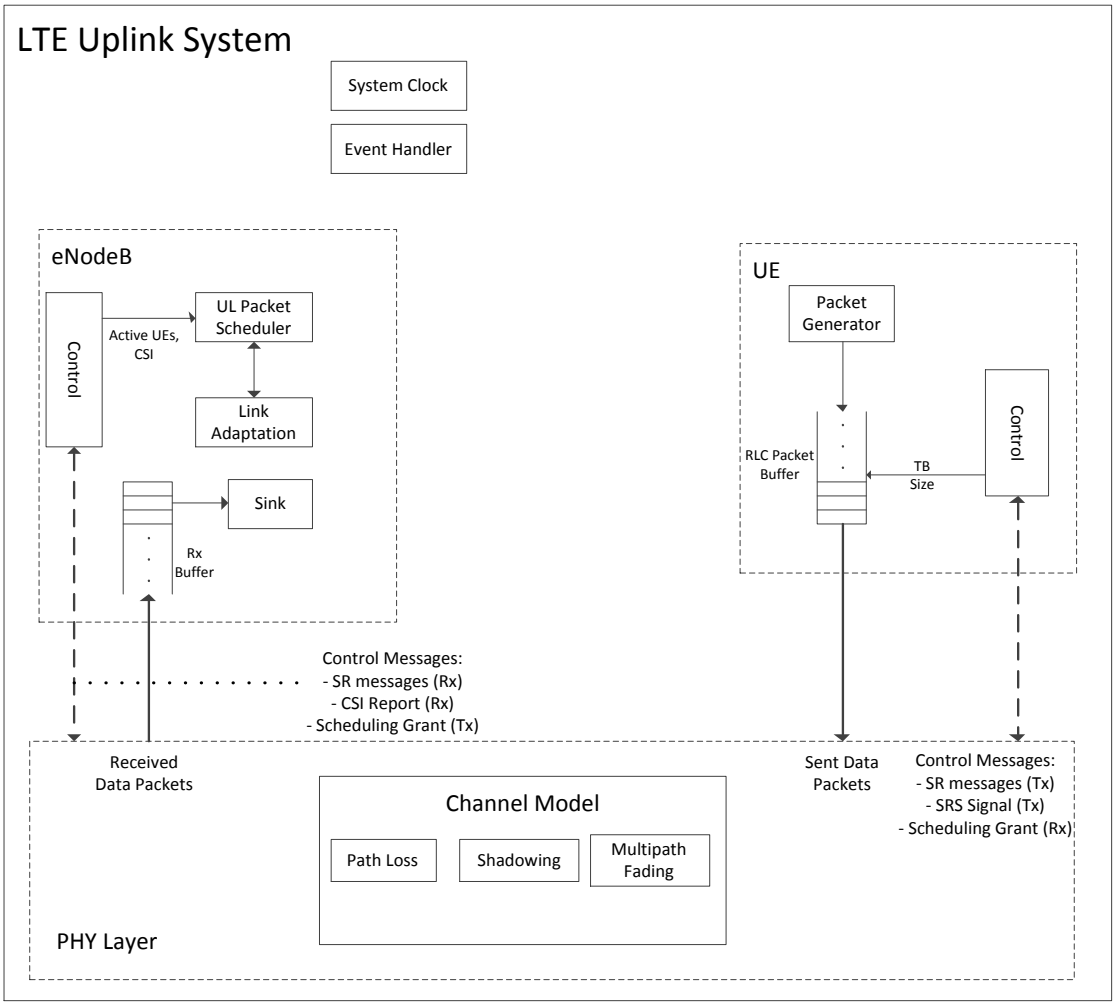

Figure 5.2: Modeling of LTE system components

5.2 Simulation Results

In this section, we present the results obtained from the experiments that were conducted on the representative schedulers using the simulator discussed in Section 5.1. The experiments conducted on the LTE uplink schedulers were designed to look at the following aspects of the system performance under different scenarios:

- The performance of LTE system under conditions of full traffic load of a single traffic profile.
- The effect of different traffic mixes on throughput, fairness, packet delays, and packet drops experienced within the system.
- The effect of varying the number of UEs on the frame utilization, as well as packet loss and packet delay
- The uplink power usage under the supervision of different uplink packet schedulers.

The results obtained from our experiments are analyzed to understand the schedulers' behaviors, after which we provide an analysis on the complexity of the schedulers' allocation methods.

5.2.1 Experiment 1: Effect of Varying the Total Number of UEs Under Heavy Traffic Load

The experiment conducted here examine the FD scheduling in a scenario where there is only a single traffic profile in the system. Each UE generates 1 Mbps FTP stream

assuming no packet drops are taking place at the UE's transmission buffer. The use of a single traffic profile assists us to determine the maximum attainable data throughput of the system under the packet scheduler's supervision. The results collected from this experiment aids us to better understand the schedulers' performance with the presence of different traffic mixes in the experiments that are discussed later on in this chapter.

As mentioned earlier in Chapter 3, the resource allocation algorithm that is implemented as part of the FDPS scheme aims to maximize the efficiency of bandwidth utilization while respecting the contiguity constraint of each PRBs subset allocated per single UE. Hence, the experiment conducted here mainly tests the resource allocation algorithm design of the schedulers being implemented. The ability of the FDPS resource allocation is to be measured in terms of system's aggregated throughput, as well as fairness. The simulation parameters of the experiment conducted are illustrated in

Table 5.3: Simulation Parameters of Experiment 1A

Parameter	**Description**
Number of UEs	10, 20, 30, 40, 50
Traffic Profile	1 Mbps FTP stream per UE

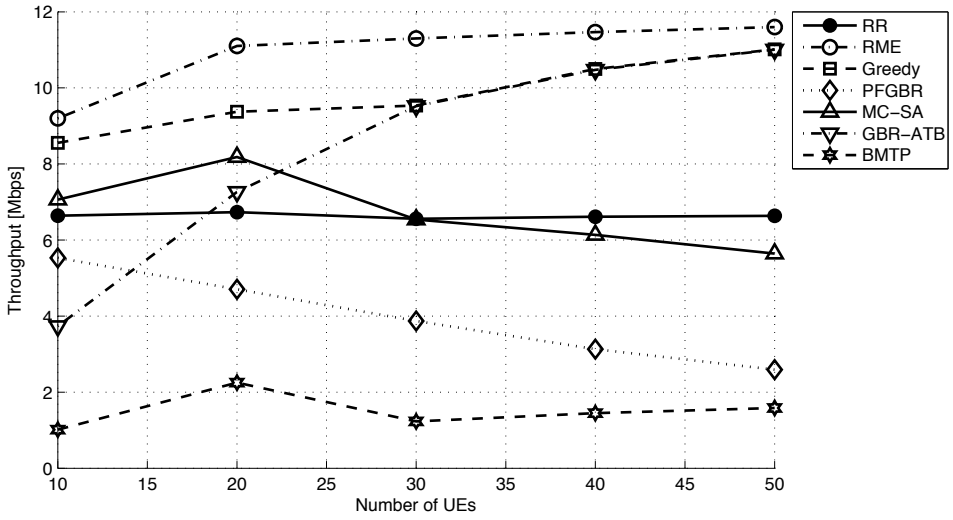

Figure 5.3: Experiment 1: Aggregated Throughput

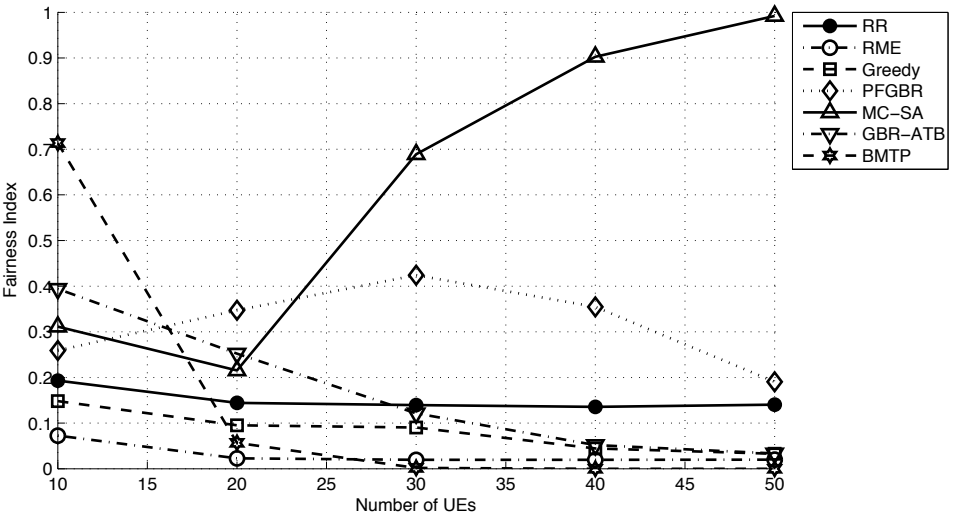

Figure 5.4: Experiment 1: Aggregated Fairness

Results obtained on the aggregate throughput and interclass fairness were presented in Figures 5.3 and 5.4, respectively. Best-effort schedulers demonstrated higher throughput levels than the schedulers from the other two categories. The use of PF

utility function in RME and Greedy algorithms increases of spectrum utilization while showing some degree of fairness in allocating PRBs among UEs. As the number increasing the number of UEs in the cell, more UEs get positioned close to the eNodeB. Such UEs are able to pack large amount of data over their assigned PRBs due to using higher MCS most of the time. The spatially uniform distribution of UEs around the eNodeB also means that increasing the number of UEs in the cell leads to having more UEs closer to the cell edge that are competing over the same amount of limited radio resources, while transmitting at lower throughput level for most of the simulation duration. This explains the decrease in fairness level as the number of UEs increases, as shown in Figure 5.4.

RME achieves higher throughput levels due to its dynamic allocation scheme, where the PRB allocations per UE dynamically changes such that the total PF utility for all UEs is maximized. However, the RME's dynamic allocation method degrades the algorithm's fairness. RME's dynamic allocation scheme has no upper bounds on how many PRBs a UE can be allocated, as long as continuing to allocate PRBs to a given UE maximizes the proportional fairness utility function of the system. Most of the UE-to-PRB assignments can lead to scheduling far less UEs than the number of UEs that are candidates for scheduling in a given TTI. Greedy algorithm solves this problem by performing the FTB scheduling explained earlier, though at the expense of a throughput degradation.

Greedy algorithm, shown in Algorithm 2, groups available PRBs into RCs, where the number of PRBs per RC is based on the number of UEs in the system. The scheduler then assigns at most one RC per UE to provide the fairest opportunities to UEs in terms of resource allocation. FTB-based greedy scheme has achieved a better

Min-Max fairness level than RME for high number of UEs, yet it was still below 0.2. Giving equal chances to UEs to transmit does not necessarily lead to equal per-UE average throughput. The different channel quality experienced by each UE dictates that UEs are to transmit at different data rates if given equal chances.

QoS-based schedulers were simulated under a single non-GBR FTP traffic load which neutralizes the QoS metrics. The absence of QoS distinction in this case makes QoS-based schedulers having PF-like schedulers, and thus their observed performance levels become mainly due to the allocation algorithm of the scheduler. As for the GBR-ATB scheduler, the FDPS allocation scheme is very similar to RME, except that the scheduler stops allocating PRBs to a UE at the extra condition of exceeding the power constraint. In addition, the scheduler's utility function incorporates the average throughput term based on the moving average throughput defined in Table 4.1, as well as the term representing the window-based average throughput calculated only over scheduled TTIs. Hence, the PF-based terms and the power-limitation in resource allocation increases the fairness level of GBR-ATB at the expense of lowering the aggregated throughput, as shown both figures. However, as the number of UEs increases, the scheduler starts to show similar behavior to Greedy and RME algorithms, which is the increase in data throughput and degradation of inter-class fairness level. As noted in Figure 5.3, the similarity in performance results from GBR-ATB and Greedy as the number of UEs increases indicates GBR-ATB behaving like Greedy algorithm in the sense that it maximizes the fairness in terms of having almost equal transmission chances per UE.

The other QoS schedulers, PFGBR and MC-SA, do demonstrate better fairness levels than the schedulers discussed above. PFGBR algorithm demonstrated improved

fairness level than all of RME, Greedy, and GBR-ATB as well. However, PFGBR's fairness still falls short of 0.4. PFGBR's fairness level achieved does not justify the degradation in throughput levels when compared to the schedulers just discussed. When looking at the PFGBR's allocation scheme in Algorithm 4, we see that the scheduler computes the utility-based metric per PRB-subset rather than per PRB, as explained in Chapter 4. Hence, when performing greedy-based allocation on the per-PRB-subset metrics, each scheduling iteration results in assigning a subset of contiguous PRBs to UE, as long as the PRB-subset has the maximum PF metric. As seen in Figures 5.3 and 5.4, the algorithm's performance declines in terms of fairness and throughput as a result of having more UEs with relatively moderate or poor channel quality as result of being spatially located further away from the eNodeB.

Next, the performance of MC-SA scheduler has shown to be the one closest to achieving fair usage of the network with large number of UEs present within the cell. Looking at the algorithm of MC-SA, we see that it executes one of two different sub-routines depending on whether the number of UEs are less or larger than the number of available PRBs. When the number of UEs get smaller than the number of PRBs, MC-SA scheduler executes an allocation scheme similar to that of Greedy algorithm. The difference is that MC-SA assigns more PRBs to the UE with the highest priority. This explains the similarity in performance trends between MC-SA and Greedy when the number of UEs are below the number of PRBs. For thirty UEs and more, MC-SA executes a different algorithm where the scheduler traverses through a sorted UE priority list, starting with the UE with the highest priority. The scheduler then assigns at least one PRB to each traversed UE as long as there are PRBs still available. The scheduler can assign more PRBs to a UE such that

the GBR requirement of that UE is satisfied. The traversal through the sorted list of UEs causes the aggregate throughput of the system to gradually decrease as the number of UEs increases further. Higher priority UEs do not necessarily have as efficient utilization of the bandwidth as lower priority UEs, since per-UE traversal gives higher priority to satisfying the QoS requirements per UE rather than increasing the system's throughput. This is reflected in the fairness level of MC-SA in Figure 5.4 as the Min-Max fairness index gets closer to unity as the number of UEs decrease, which contradicts the fairness behavior of all other algorithms examined here.

BMTP algorithm shows the lowest overall performance level. There are two concepts which BMTP relies on to lower per-UE transmission power. First, to maintain a certain rate, allocating more PRBs to a UE means that the MCS needed to achieve a given minimum bit rate becomes lower. Consequently, lowering the MCS level means that the threshold SINR to be met will be lower as well, which in turn lowers the uplink transmission power. In other words, no matter how many PRBs get allocated to a UE, the scheduler will lower the transmission power such that a UE is to maintain a certain data rate. Second, the BMTP's FDPS allocation scheme follows a UE-to-PRB-subset allocation scheme that causes only few UEs to get resource allocation every TTI, with each UE transmitting at a fixed target bit rate. This method of power optimization explains the relatively degraded throughput performance of the scheduler as shown in Figure 5.3.

Another point that needs to be clarified here is that the power-saving approach chosen by BMTP has the disadvantage of being very sensitive to UE's experienced channel condition, which is essentially a function of the UE's distance from the eNodeB. Such disadvantage can have a detrimental effect on the eNodeB performance

with widespread coverage area. Therefore, BMTP shares the same disadvantage as max-SINR schedulers introduced in earlier wireless systems where the scheduler favors UEs with better SINR, leaving UEs further away victims for starvation.

5.2.2 Experiment 2: The Effect of Different Traffic Mixes on System Performance

The purpose of this experiment is to examine the schedulers' performance under different traffic mixes. The LTE uplink schedulers at the eNodeB do not distinguish between the different traffic types at a single UE when scheduling resources. Hence, we chose to provide each UE with one SDF carrying a single traffic stream, to see the impact of the scheduler on the QoS experienced in traffic mix scenarios. The total traffic load for the entire experiment is set to 8145 kbps, of which 465 kbps is given to VoIP traffic, 3840 kbps to video streaming, and 3840 kbps to FTP traffic. The remaining parameters are listed in Table 5.4. The UE ratios presented in the table reflect the ratio of the number of UEs with a given traffic class to the number of UEs from the other traffic classes.

Table 5.4: The Effect of Traffic Mix Ratios - Parameters

Parameters	**Value**
Number of UEs	25
UE Ratios (VoIP, Video, FTP)	1:1:3, 1:2:2, 1:3:1, 2:2:1, 3:1:1

When looking at the results obtained for cell's aggregate throughput in Figures 5.5 through 5.7, a common behavior among all schedulers is that the total system throughput of a given traffic class improves as the concentration of UEs belonging to that class increase within the system.

The performance of RME has degraded significantly compared to its performance from Experiment 1. RME's performance degradation is due to the significant variations between offered traffic loads from UEs of different traffic classes. In general, UEs with VoIP traffic have low data rate requirements compared to UEs from the video and FTP traffic classes. Hence, the use of PF metric causes UEs with VoIP traffic to be either scheduled more frequently or allocated more PRBs than other UEs. In addition, RME adopts a dynamic resource allocation scheme where the size of allocated PRB set dynamically changes for each UE, with no limitations on how many PRBs can be allocated to a single UE within a given TTI. Therefore, UEs with VoIP traffic can often get allocated many PRBs which can be more than their needs. This effect becomes more dominant as the portion of UEs with VoIP traffic becomes higher.

The results obtained in Figures 5.5 through 5.15 show that Greedy's and RR's performances in mixed traffic scenarios was competitive to that of QoS-based schedulers, despite the absence of QoS provisioning in both schemes. This competitiveness is related to the FTB-based allocation scheme of both algorithms, which maximizes the fairness of transmission chances among UEs by setting an upper bound on how many PRBs are allocated per UE. This gives both Greedy and RR a great advantage over RME in terms of their ability to support the traffic mix profiles that were simulated in the system, despite RME's higher performance in Experiment 1.

When looking at the experienced average delay results of both VoIP and video traffic classes in Figures 5.8 and 5.9, respectively, and also the percentage of packet drops of the same two traffic classes in Figures 5.10 and 5.11, the results show the

ability of Greedy and other QoS-based schedulers to accommodate the QoS requirements of VoIP traffic. We note that GBR-ATB scheduler, which uses a similar FDPS allocation scheme to that of RME, can better accommodate for traffic mixes than RME can. This is a strong indication that the modification from classic PF utility function to a QoS-based utility function can significantly improve the performance of adaptive allocation schemes in accommodating traffic mixes.

Also, PFGBR has shown better adaptivity to mixed uplink traffic if compared to RR, RME, and Greedy. A key concept in PFGBR performance is that it limits maximum number of available PRB-subsets per UE based on the UE's defined maximum bit rate. Such constraints have proven critical in the sense that they prevent allocating a UE more PRBs than it needs. In Addition, PFGBR scheduler shows good compromise when supporting both GBR and non-GBR traffic streams, as demonstrated by examining the FTP traffic performance under PFGBR, as shown in the FTP results shown in Figures 5.7 and 5.14.

FTP traffic is the most traffic type that suffers from resource starvation with the presence of higher concentration of other GBR traffic types of higher priority and lower experienced traffic rates. However, the performance of FTP traffic was degraded the least with PFGBR scheduler, since the scheduler uses two separate metrics for GBR and non-GBR traffic types. The introduction of such metric system has shown to allow PFGBR to better prevent lower priority, non-GBR traffic such as FTP from starvation in the presence of other GBR traffic with higher priority.

MC-SA has shown good support for handling UEs of different QoS classes relative to other schedulers, even the schedulers from its own class. In this experiment, MC-SA scheduler performs the allocation subroutine where the scheduler prioritizes

UEs according to their past average throughput and experienced QoS, then traverses through the prioritized list one UE at a time to estimate the resources it needs to satisfy its GBR as well as its delay requirements. This scheme gets differentiated from other schedulers when examining the experienced performance of video streaming traffic. We find that MC-SA is the only scheduler that is able to accommodate the video's offered traffic load with the concentration of UEs with video streaming services is the highest, considering that aggregate throughput of video traffic in this case is very close to the offered traffic load with very minimal packet drops experienced here.

The performance of MC-SA is superseded by PFGBR, GBR-ATB, Greedy, and also RR in the case of FTP traffic. MC-SA performance with FTP traffic is the outcome of its UE prioritization that leaves UEs with the non-GBR FTP traffic always at the bottom of the prioritized list most of the time. As a result, FTP would have few chances of getting any resources allocated especially if they are present with low concentrations along with other UEs with higher priority traffic.

On the other hand, BMTP scheduler has shown poor performance with all three traffic classes simulated here. One main reason for such a poor performance level is the low efficiency of bandwidth usage for lowering the uplink transmission power, as mentioned in earlier discussions of Experiment 1. Allocating a VoIP UE three PRBs or more for transmitting a single VoIP packet, for example, causes other UEs of traffic classes to have higher starvation levels, eventually causing overall lower satisfaction levels for active upstream traffic.

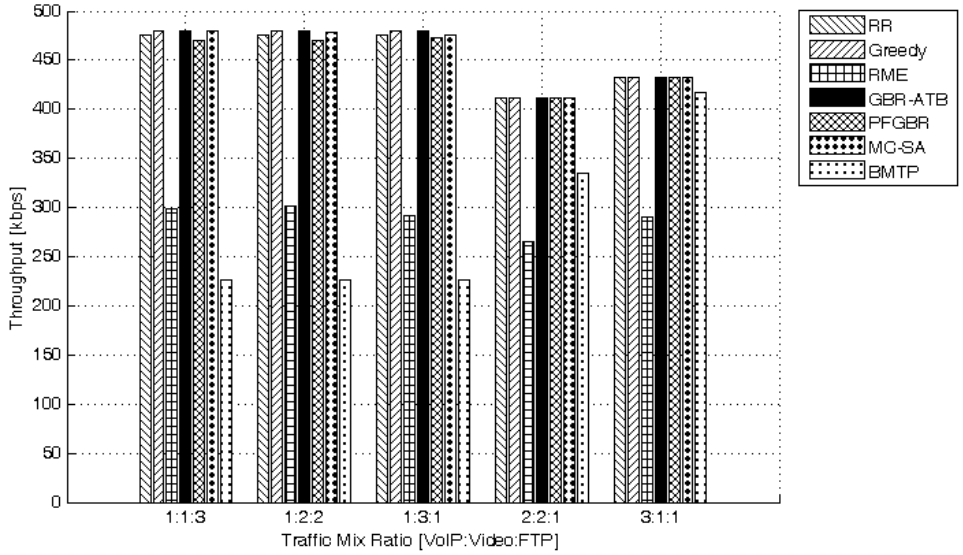

Figure 5.5: Experiment 2: VoIP Aggregated Throughput

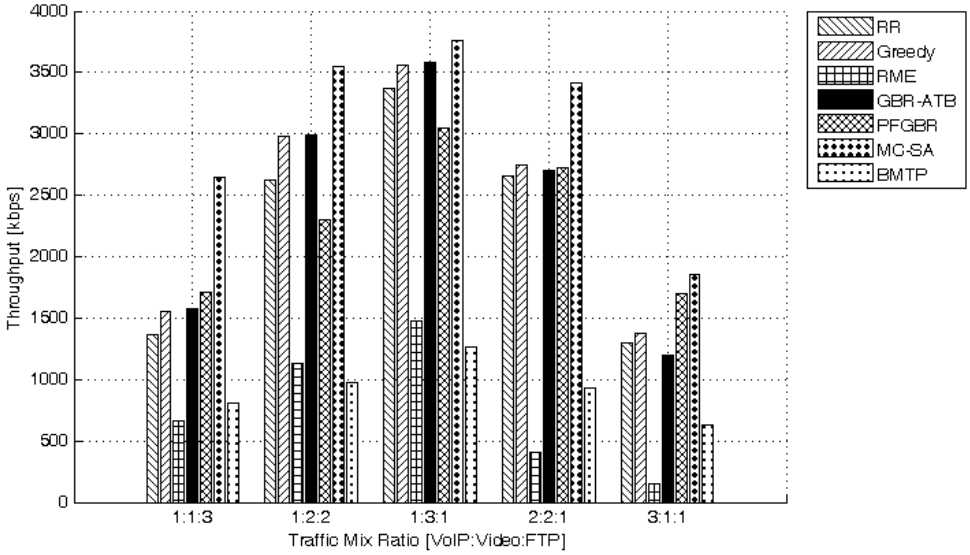

Figure 5.6: Experiment 2: Video Aggregated Throughput

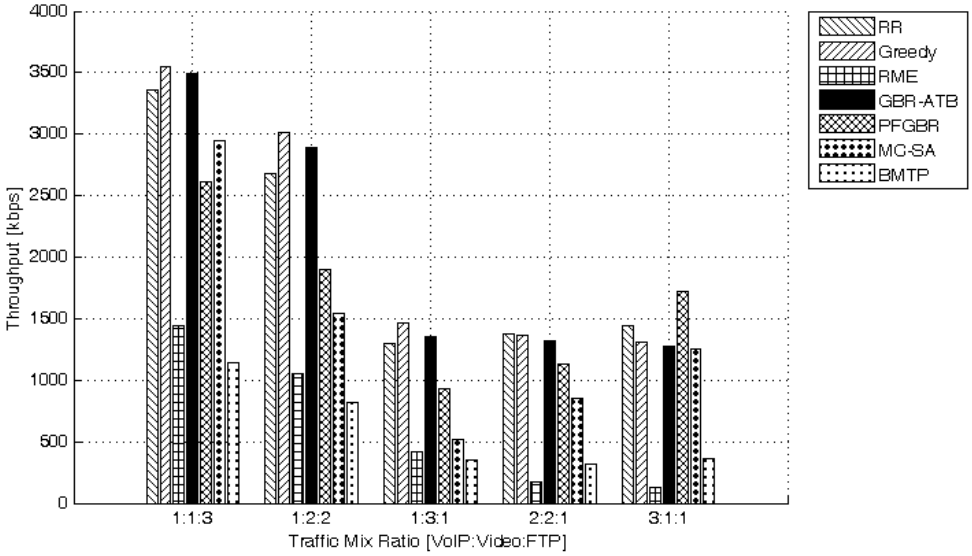

Figure 5.7: Experiment 2: FTP Aggregated Throughput

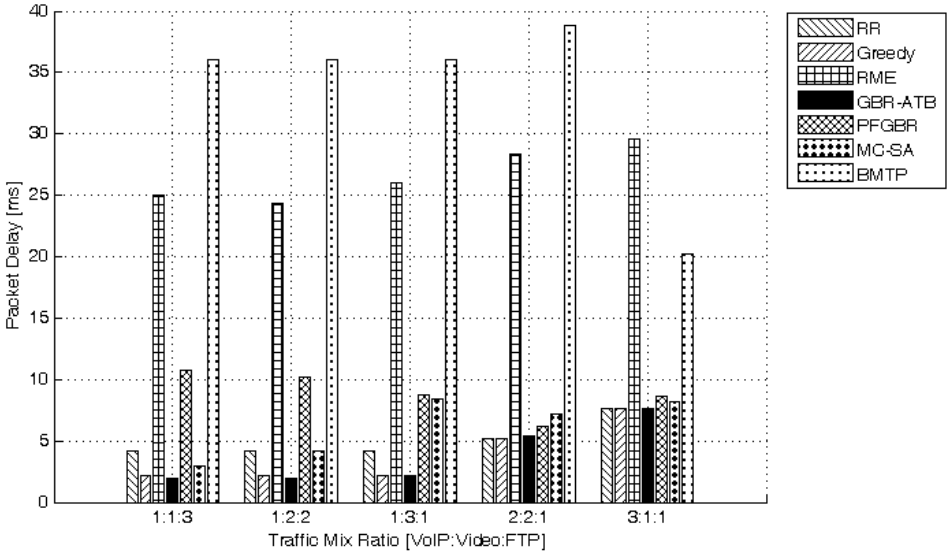

Figure 5.8: Experiment 2: VoIP Average Delay

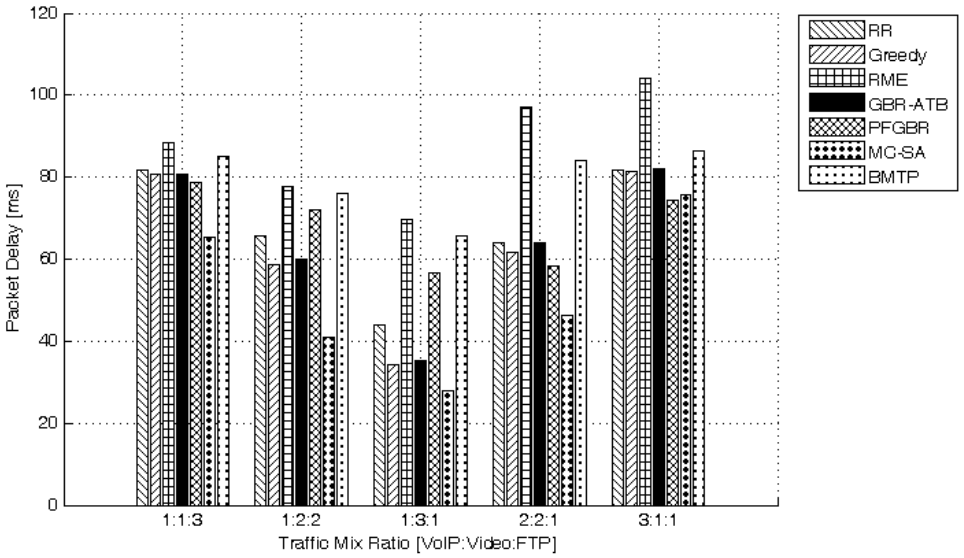

Figure 5.9: Experiment 2: Video Average Delay

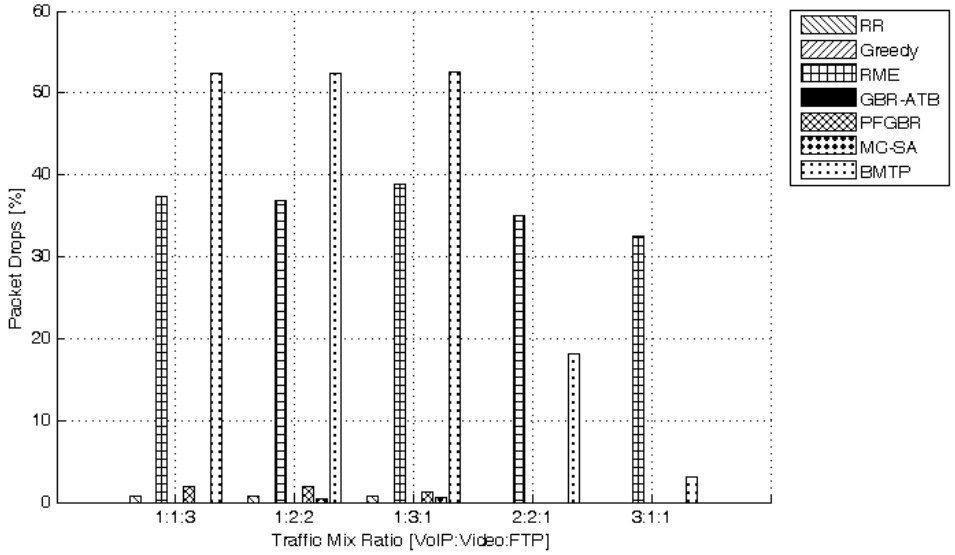

Figure 5.10: Experiment 2: VoIP Packet Drops

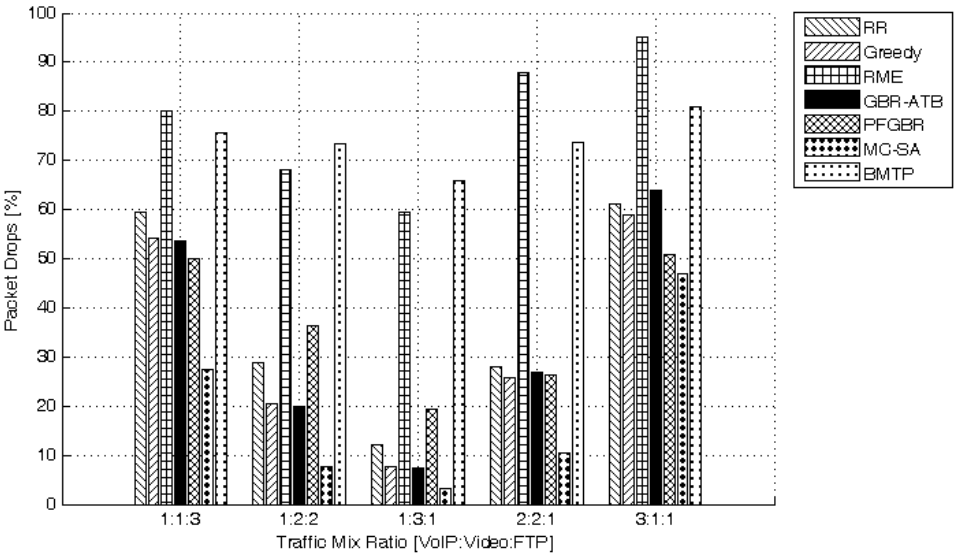

Figure 5.11: Experiment 2: Video Packet Drops

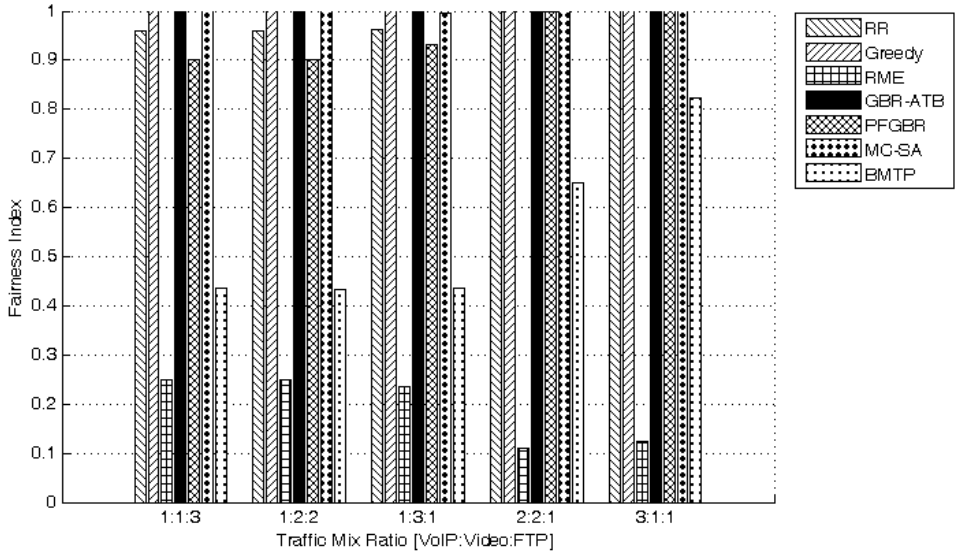

Figure 5.12: Experiment 2: VoIP Min-Max Fairness

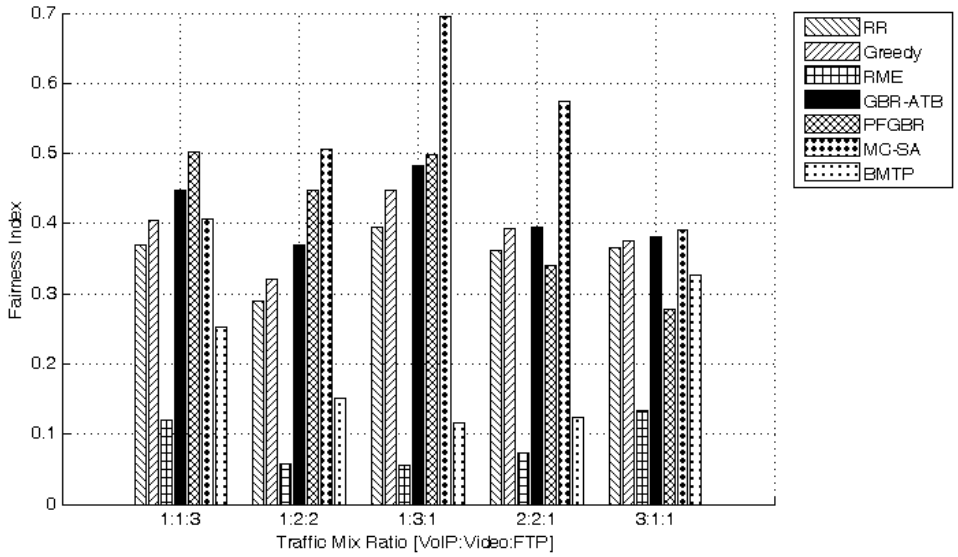

Figure 5.13: Experiment 2: Video Min-Max Fairness

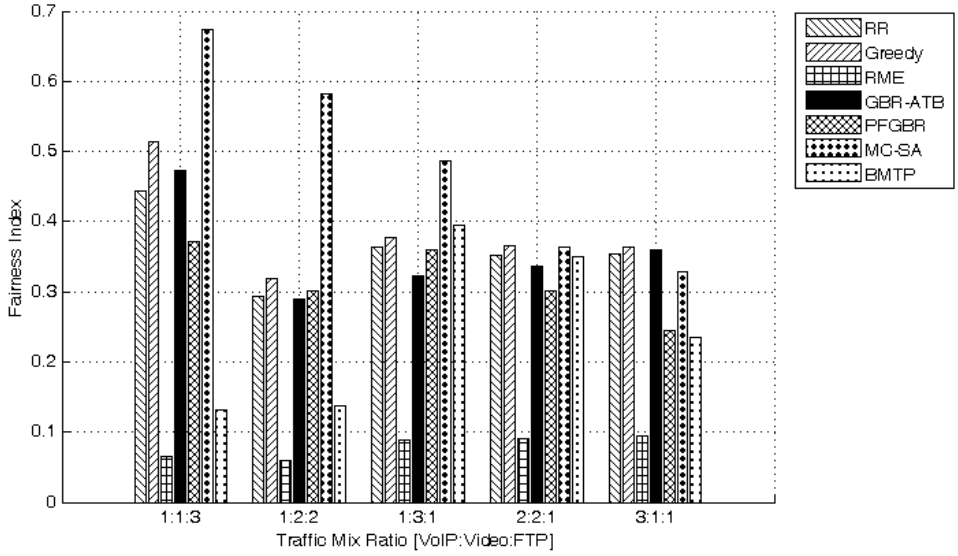

Figure 5.14: Experiment 2: FTP Min-Max Fairness

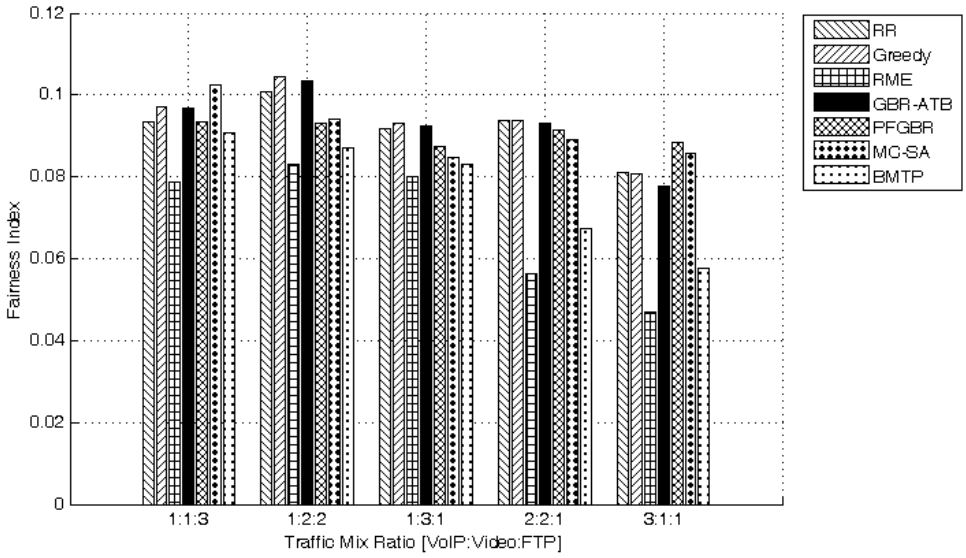

Figure 5.15: Experiment 2: Inter-Class Fairness

5.2.3 Experiment 3: Effect of Varying Number of UEs Under Mixed Traffic Scenarios

The number of UEs within the cell and the amount of offered traffic load per UE affects the contention over the available, limited radio resources. Increasing the number of UEs leads to increasing the competition among UEs over the limited number of PRBs, leading to decreasing the chances of transmission for each UE. Also, increasing the traffic load per UE leads to increasing the contention over the limited PRBs assigned to the UE. On the other hand, having a small number of UEs with light load can lead to insufficient utilization of the assigned resources, resulting in a waste of transmission power and hence a waste of radio resources that could have been allocated elsewhere.

5.2.3.1 Varying UEs under Per-UE Fixed Uplink Traffic Load

The purpose of this experiment is to measure the utilization efficiency in terms of how well a UE of a certain QoS class can utilize its assigned TB. The Simulation parameters for this experiment are shown in Table 5.5. The traffic load per UE was fixed to how the contention increase per assigned PRBs per UE affects the performance of each scheduler. The traffic rates of UEs with VoIP, video streaming, and FTP traffic streams were set to 14.4 kbps, 64 kbps, and 128 kbps, respectively.

When constructing the TB for each transmission, a 6-byte overhead is assumed; 3 bytes are generated for the MAC and RLC headers, and 3 bytes are occupied by the CRC checksum.

Table 5.5: Effect of Number of UEs on System Performance - Parameters.

Parameters	**Value**
Number of UEs	10, 20, 30, 40, 50
UEs Ratio (VoIP, Video, FTP)	2:2:1

The results of this experiment were collected and are shown in Figures 5.16 through 5.28. A general trend observed here is that as long as there are low contentions on available radio resources, both the aggregate traffic rate and the utilization of assigned TB per UE increases as well. As shown in Figures 5.19, 5.20, and 5.21, the TB utilization starts to saturate as the number of UEs in the system surpass the total number of available PRBs. The aggregate throughput of MC-SA increases when the number of UEs changes from twenty to thirty UEs, which triggers the switch from scheduling method change subroutine 1 shown in Algorithm 7, executed when number of UEs is smaller than number of PRBs, to subroutine 2 in Algorithm 8.

UEs that experience relatively poor channel conditions, which are close to the

cell edge, will suffer from poor spectral efficiency due to using low MCS and hence smaller TB sizes. The effect of the overhead from MAC header and CRC checksum collectively further reduces the amount of space available for the data payload. On the other hand, UEs closer to the eNodeB with better channel conditions transmit at higher MCS, which reduces TB overhead impact on the TB utilization. The variation of TB overhead just explained reduces the inter-class fairness among UEs of the same traffic class, which is shown in Figures 5.26, 5.27, and 5.28 for VoIP, video, and FTP traffic classes, respectively.

UEs with VoIP traffic have shown poor average TB utilization of assigned resources, which never exceeded 55%, due to VoIP UE's low traffic loads relative to video and FTP UEs, as shown in Figure 5.19. The low TB utilization is mainly contributed by UEs close to the eNodeB, where they get assigned MCS high enough that only small portion of the TB is used for data payloads, while the rest of the TB space is filled with padding. This is even observed with RR and Greedy schedulers where a UE is assigned only one PRB for thirty UEs and above. The results observed for VoIP traffic is a strong indication that supporting voice services over the SC-FDMA radio interface can be easily achieved since a VoIP packet can easily fit into a single PRB if the UE is close enough to the eNodeB. Therefore, UEs close to eNodeB should experience very minimal delays and packet drops. On the contrary, UEs closer to the cell edge tend to transmit with low MCS. As a result, TB sizes tend to get smaller to the point that the overhead size becomes more significant compared to the small space available for data payloads. This clearly indicates that packet drops within the system, shown in Figures 5.22 and 5.24, respectively, are mostly contributed from cell-edge UEs. VoIP packet drops and packet delays do not seem to be affected by

the number of VoIP UEs in the system. When increasing the number of VoIP UEs in the cell, the probability of placing UEs in close proximity from the eNodeB, midway between eNodeB and the cell edge, and close to the cell edge is the same. Hence, with the total offered VoIP increasing with the number of UEs, the percentage of VoIP packets that get successfully transmitted stays the same still.

Figure 5.16 shows that UEs under RME scheduling show the lowest average TB utilization of assigned resources due to the low traffic loads, and hence the experienced data throughput, of VoIP UEs. For RME being blind to the total offered load by each UE, the low throughput of VoIP UEs influences RME's allocation decision by allocating them more PRBs than necessary. Such allocation behavior is affecting the resource allocation of video and FTP traffic as shown in Figure 5.17 and Figure 5.18. The impact of VoIP's low data rate on video and FTP traffic flows becomes more obvious as each of video and FTP traffic saturates beyond a certain number of UEs within the cell. Under RME scheduling, FTP aggregate throughput saturates as the number of UEs becomes thirty and above. As for the video traffic, the aggregate throughput starts to saturate as the number of UEs start to cross forty UEs under RME scheduling as well. The leveling of aggregate throughput of the video and FTP traffic classes under RME scheduling is for for increasing the number of VoIP UEs that compete with UEs from the other two classes on the ever-constant number of PRBs. The difference in throughput saturation between video and FTP is mainly due to having twice as many video UEs as FTP UEs, while the total offered traffic load of both traffic flows is the same for any given number of UEs in the system. As a result, the starvation for radio resources caused by VoIP traffic ,as explained above, has a greater negative effect on FTP UEs than on video UEs.

As for GBR-ATB, the TB utilization of all the traffic classes seem to be consistent regardless of the number of UEs in the system. Compared to the TB utilization of the other schedulers where the number of PRBs allocated per UE is inversely proportional to the number of UEs in the system, GBR-ATB shows better distribution of the PRBs among the UEs in the system to maintain a certain level of TB utilization. The behavior of GBR-ATB shows to be similar with video and FTP traffic.

BMTP has shown similar TB utilization behavior with the three traffic classes, where the the TB utilization is highest (55%-65%) with twenty UEs or less, then drops to about 45%-50% when the number of UEs further increase. The abrupt decrease of TB utilization as the number of UEs increase from twenty to thirty UEs comes from having UEs added to the system that are assigned TB sizes small enough for the overhead to make up at least 50% or more of the total TB size. Also, with thirty UEs and above, there would be a higher probability that some UEs get TB assignments that are not big enough to hold any data, and thus end up wasting some radio resources. In addition, increasing the number of UEs lead to having more cell-edge UEs that do not have any resource assignments due to not being able to satisfy their GBR requirements without violating the maximum power constraints. This is deduced from the zero min-max fairness level experienced under BMTP scheduling for thirty UEs and above in Figures 5.26, 5.27, and 5.28. The types of UEs just mentioned here can greatly reduce the average TB utilization of the entire UEs present in the system down to as low as 45%.

The general trend of packet delays and packet drops of both VoIP traffic (Figures 5.22 and 5.24) and video traffic (Figures 5.23 and 5.25) under the different schedulers

shows little variations as the number of UEs increase. Also, since the offered traffic load increases with the number of UEs, the overall delay and packet drops stay relatively the same for any number of UEs within the cell coverage.

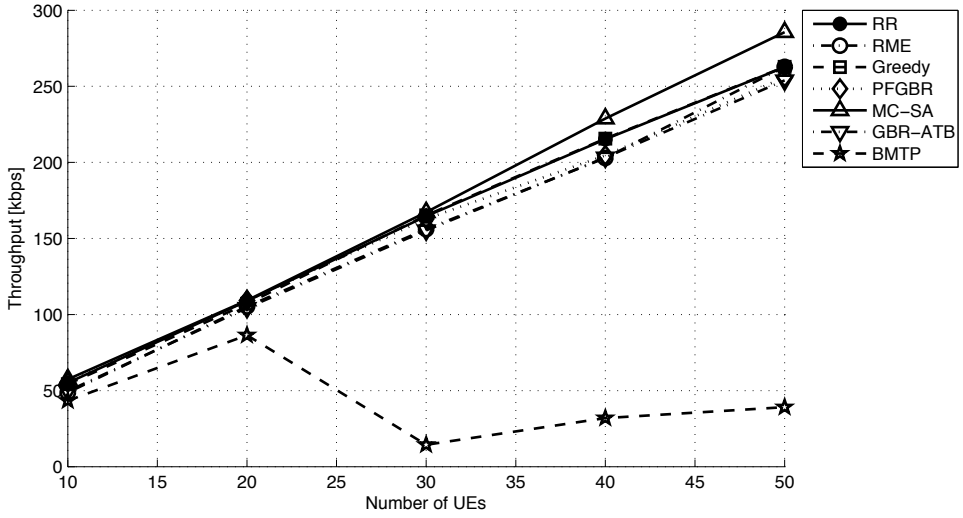

Figure 5.16: Experiment 3-1: VoIP Aggregated Throughput

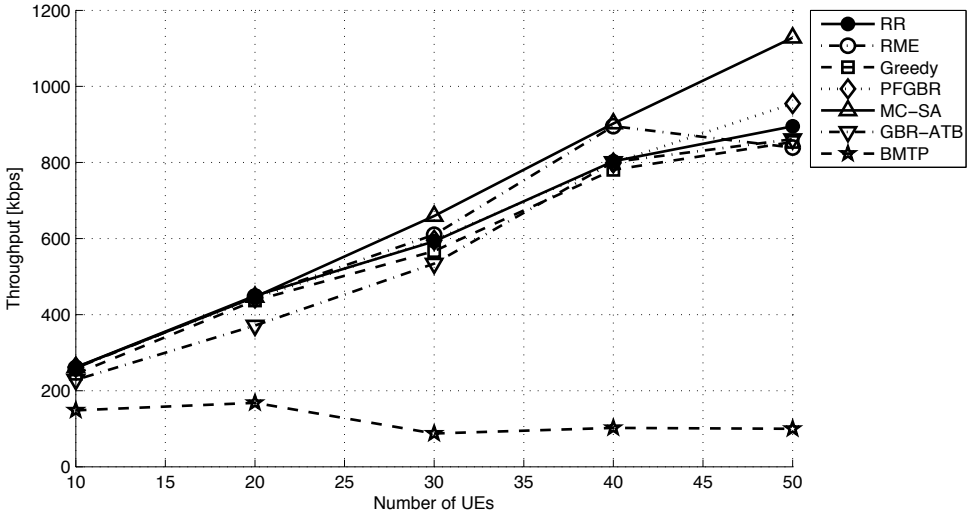

Figure 5.17: Experiment 3-1: Video Aggregated Throughput

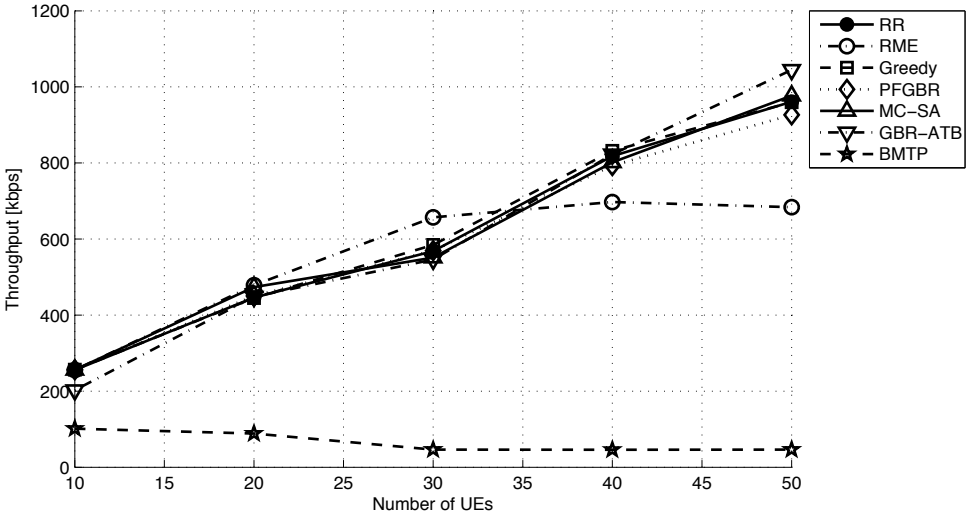

Figure 5.18: Experiment 3-1: FTP Aggregated Throughput

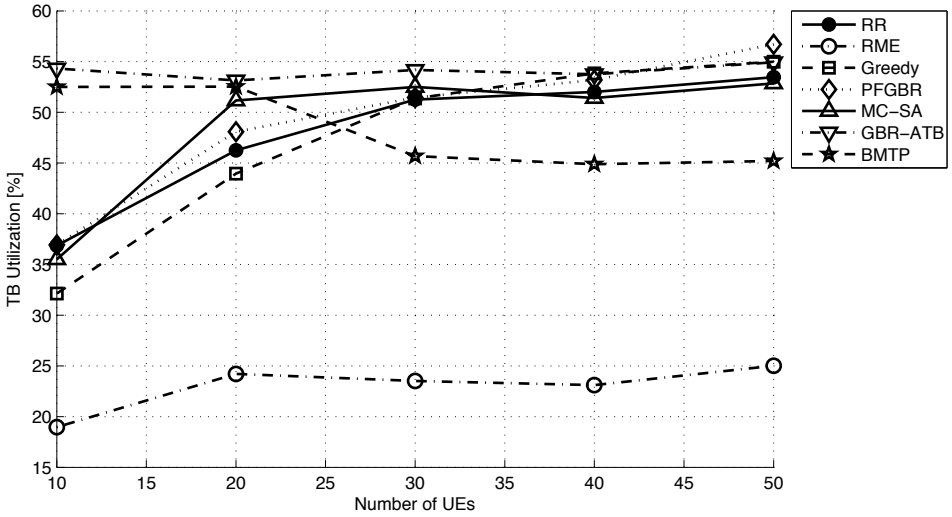

Figure 5.19: Experiment 3-1: VoIP TB Utilization

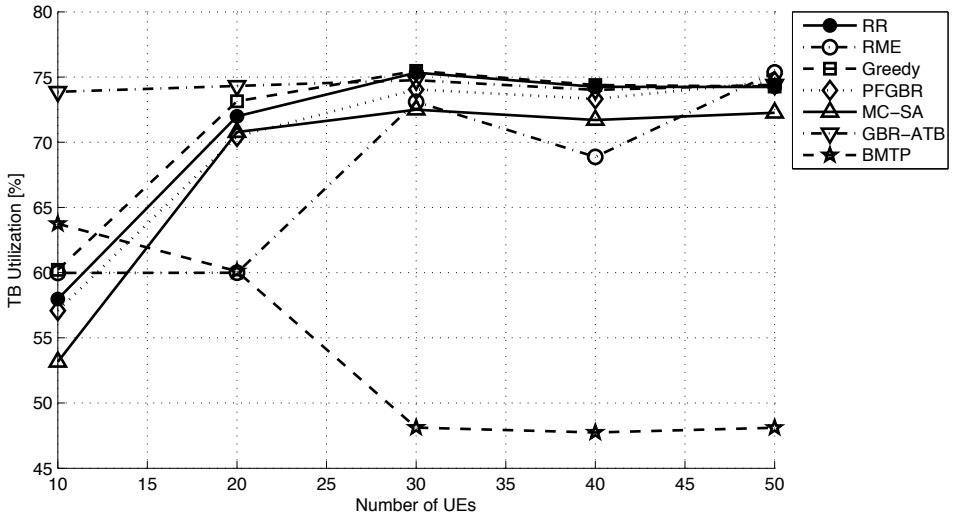

Figure 5.20: Experiment 3-1: Video TB Utilization

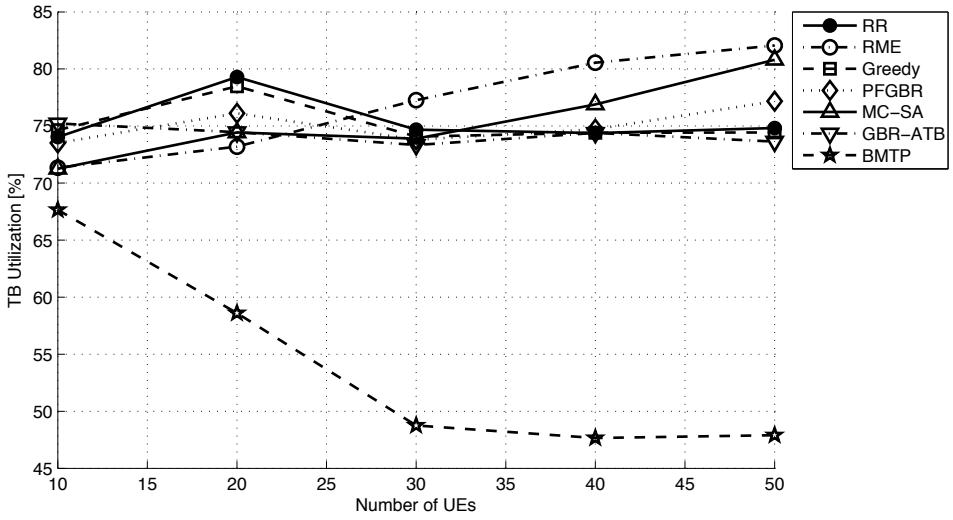

Figure 5.21: Experiment 3-1: FTP TB Utilization

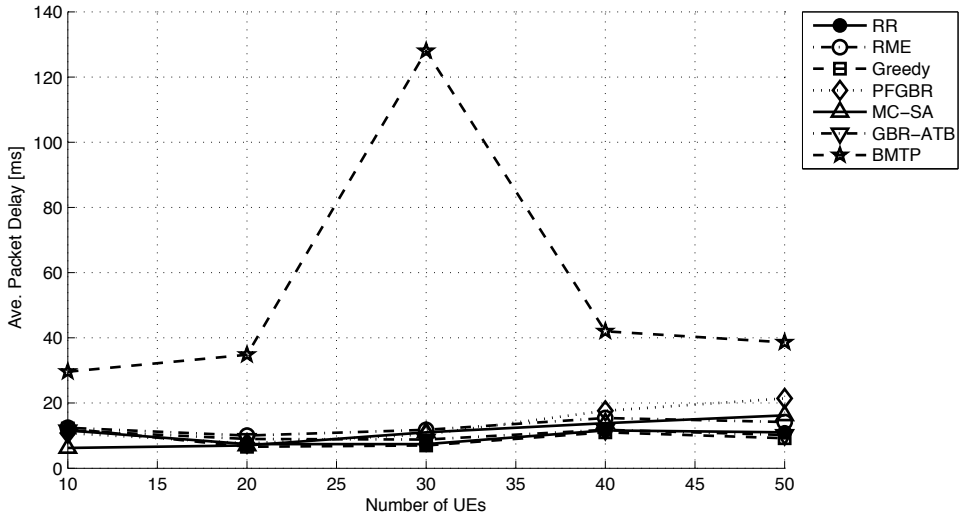

Figure 5.22: Experiment 3-1: VoIP Average Delay

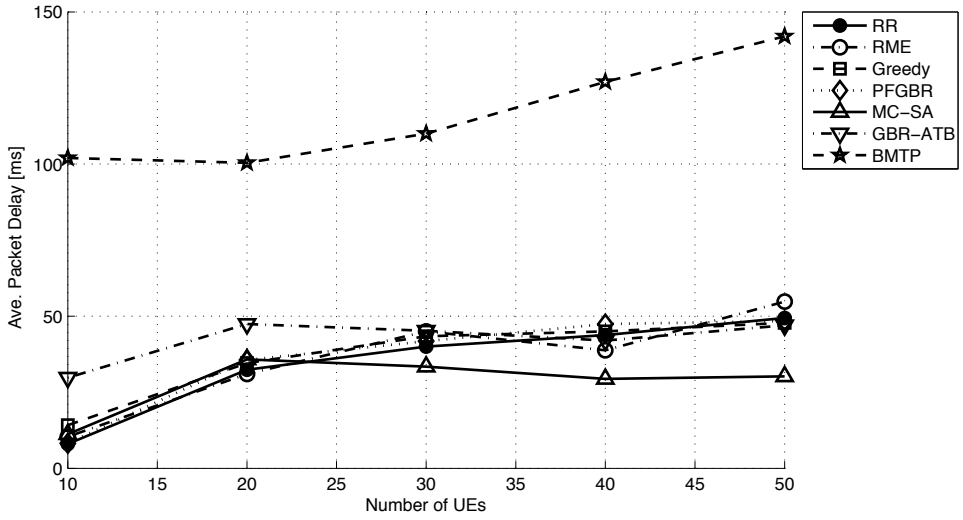

Figure 5.23: Experiment 3-1: Video Average Delay

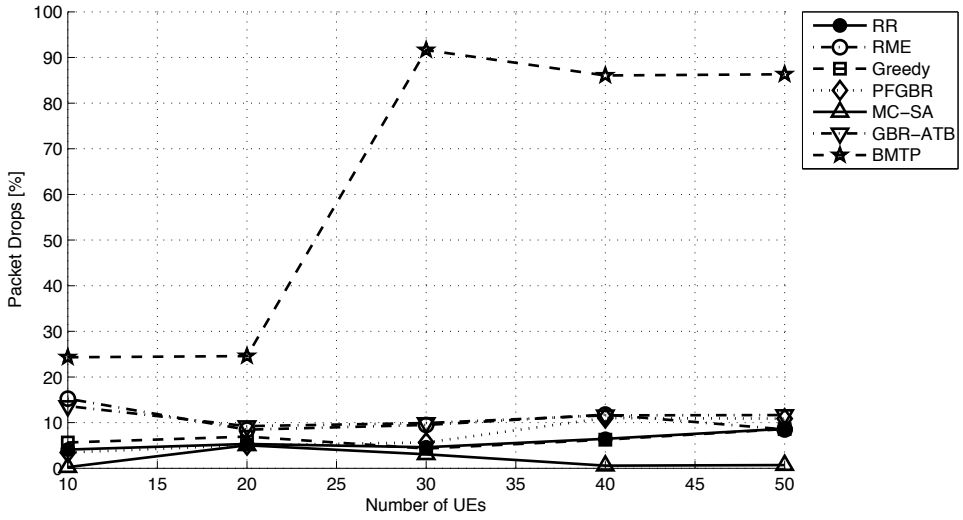

Figure 5.24: Experiment 3-1: VoIP Packet Drops

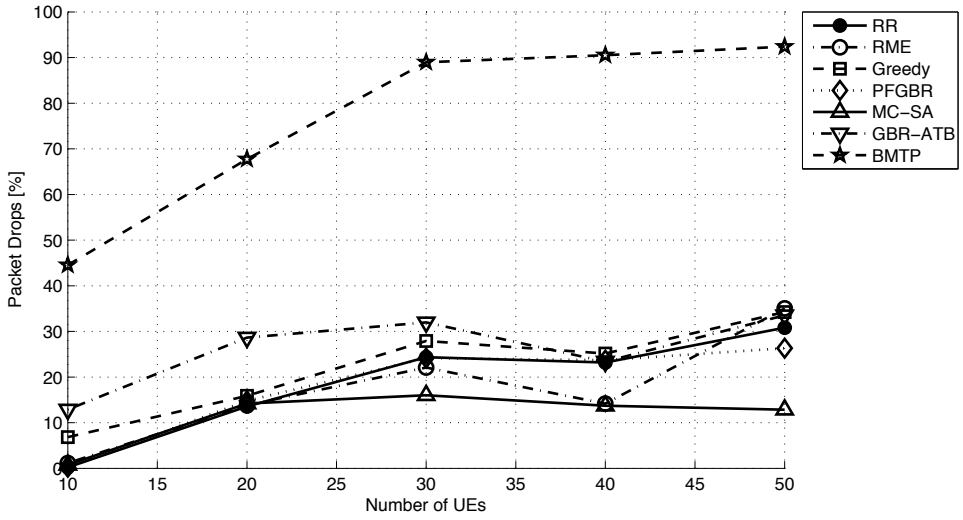

Figure 5.25: Experiment 3-1: Video Packet Drops

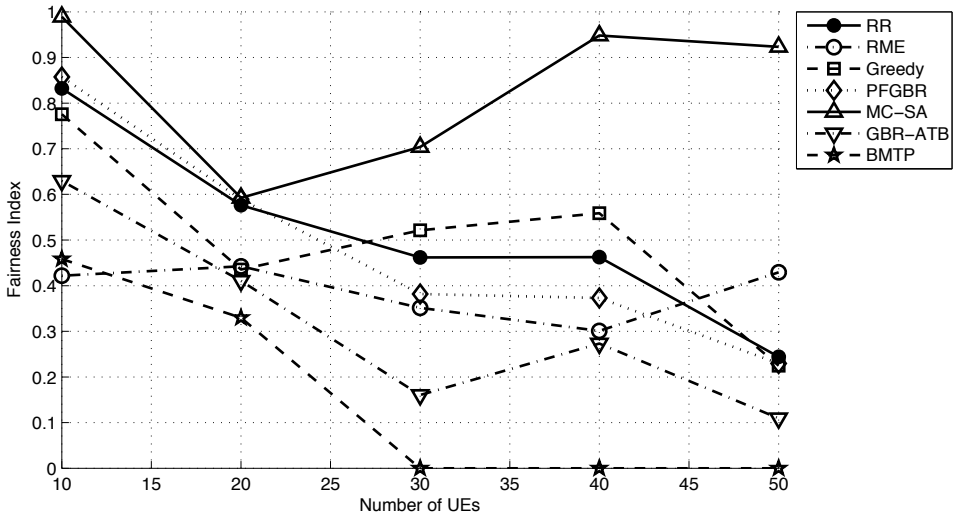

Figure 5.26: Experiment 3-1: VoIP Min-Max Fairness

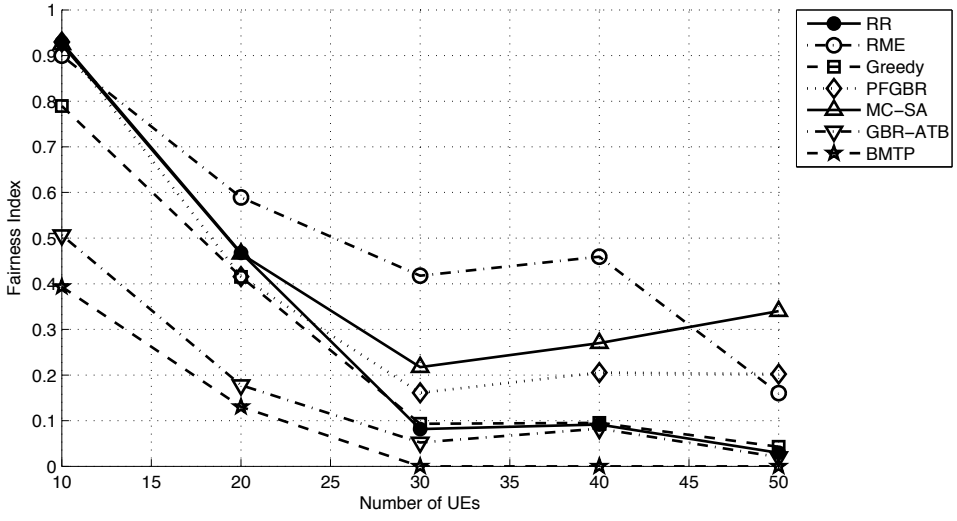

Figure 5.27: Experiment 3-1: Video Min-Max Fairness

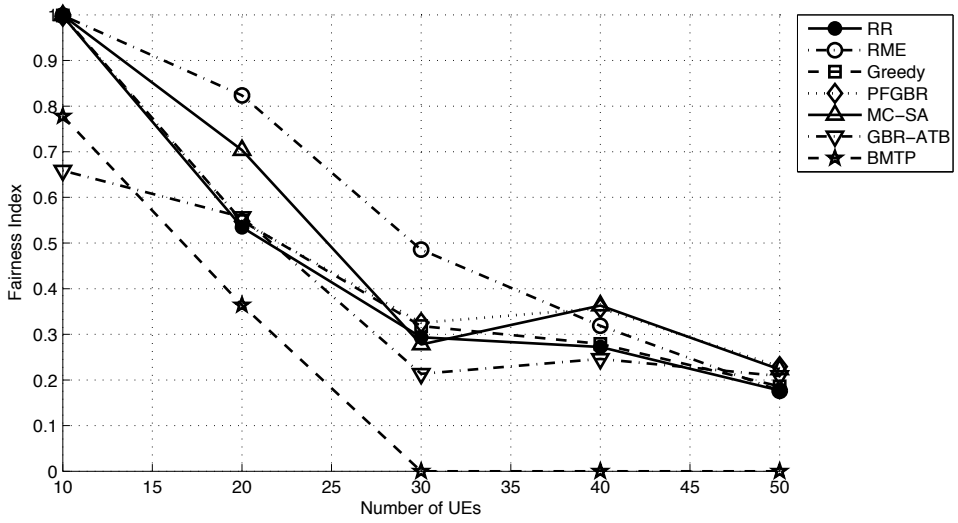

Figure 5.28: Experiment 3-1: FTP Min-Max Fairness

5.2.3.2 Varying Number of UEs Under Fixed Total Uplink Traffic Load

In this experiment, we want to see the effect of varying the number of UEs while fixing the total traffic load on system performance. Varying the number of UEs under fixed traffic illustrates how the total throughput of the system can vary as the traffic contention over the resources allocated per UE. The findings in this experiment are to be compared to the performance results obtained from Experiment 1 on the system settings that best suit the schedulers under study. The simulation parameters are listed in Table 5.6.

Table 5.6: Effect of number of UEs on TB utilization with fixed load - Parameters.

Parameters	**Value**
Number of UEs	10, 20, 30, 40, 50
UEs Ratio (VoIP, Video, FTP)	4:3:3

The total network traffic load is to be fixed at 6400 kbps, where 256 kbps being

reserved for VoIP, 3072 kbps is to be reserved for video streaming, while 3072 kbps is to be allocated to FTP traffic.

The simulation results of this experiment are demonstrated the Figure 5.29 through Figure 5.38. An interesting note on the aggregate throughput from the different traffic classes is that the system's total throughput from UEs of all the three traffic classes constitutes about 30-50 % of the entire offered traffic load, which amounts for 6400 kbps. Most of the loss in throughput comes from FTP traffic, followed by video. UEs of the VoIP traffic class suffered the least performance degradation as the average packet loss experienced did not exceed 20%, except in the case of PFGBR and BMTP. As previously pointed out, the offered load from UEs of the VoIP traffic class is very low compared to that of video and FTP. In addition, the utility functions of all the schedulers here, except BMTP and MC-SA, use PF term as part of the utility function metric based on which these schedulers assign weights to each schedulable PRB for each each UE. Consequently, the PF term in the utility functions in the presence of UEs with significant gap between the offered traffic loads of UEs from different classes disrupts the PRB allocation decisions of such UEs. Even in the case of QoS schedulers like GBR-ATB and PFGBR, the PF term in their utility functions even outweigh the QoS terms, as the performance of QoS-blind schedulers, such as RR, RME, and Greedy, have comparable performance levels to the schedulers from the QoS-based category.

PFGBR, in contrary to other schedulers, has shown degraded performance of VoIP traffic as the number of UEs increases beyond thirty. The degradation noted with PFGBR can be related to PFGBR's scheduling scheme described in Algorithm 4.

That is, as part of the scheduler's initiation steps, the scheduler computes the estimated throughput for each UE at all possible PRB-subsets, after which the scheduler eliminates all PRB-subsets that provide data rates exceeding the UE's maximum bit rates. As the maximum bit rates of UEs from video and FTP traffic classes far exceed that of UEs from the VoIP classes, UEs from video and FTP classes are allowed to have larger choices of PRB allocation sets than VoIP UEs. Therefore, as the number of UEs increase, the competition on radio resources increase among UEs as well, were the possibility to allocate PRB subsets with larger sizes to video and FTP can easily deplete VoIP UEs from their possible PRB allocations.

MC-SA, on the other hand, has shown better handling of traffic mixes than other schedulers, which is mainly due to UE-prioritization shown in equation (4.9). Rather than incorporating PF term that is based on ratio of achievable throughput on a given PRB to the UE's average throughput, the metric rather relies on two QoS-based ratios. One ratio tracks the experienced average delay relative to the PDB of the UE of a certain traffic class, and also the other ratio tracks the UE's average throughput relative to its GBR. Therefore, the scheduler can better track the QoS experienced by UEs from VoIP and video classes, and ensure that their QoS needs are satisfied as best as possible.

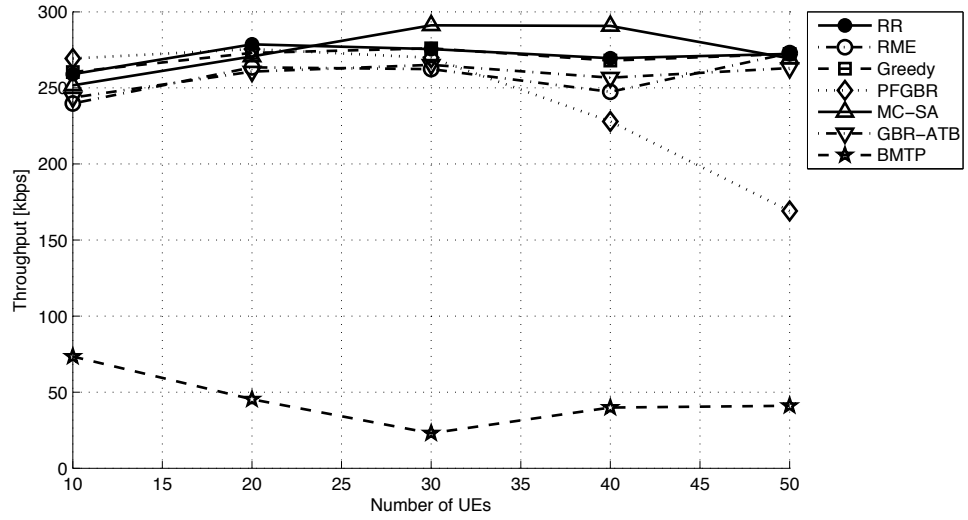

Figure 5.29: Experiment 3-2: VoIP Aggregated Throughput

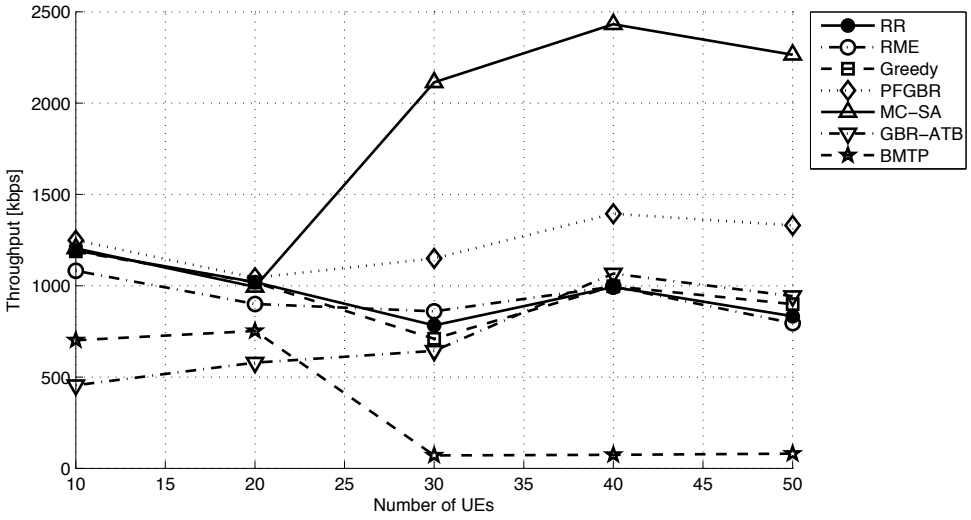

Figure 5.30: Experiment 3-2: Video Aggregated Throughput

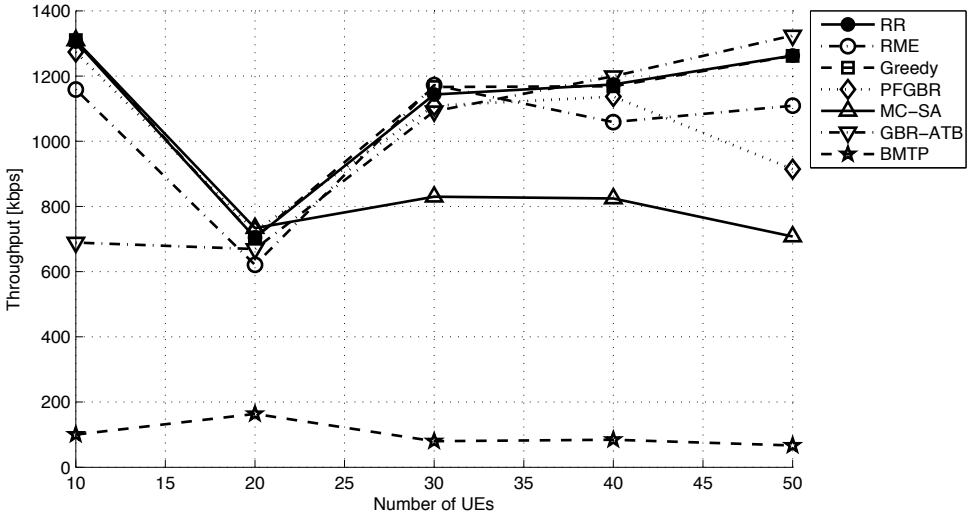

Figure 5.31: Experiment 3-2: FTP Aggregated Throughput

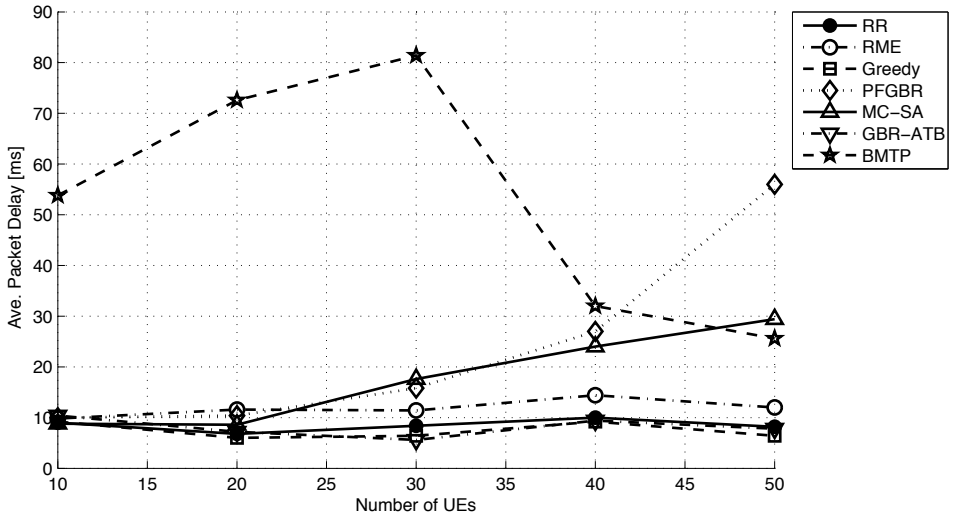

Figure 5.32: Experiment 3-2: VoIP Average Delay

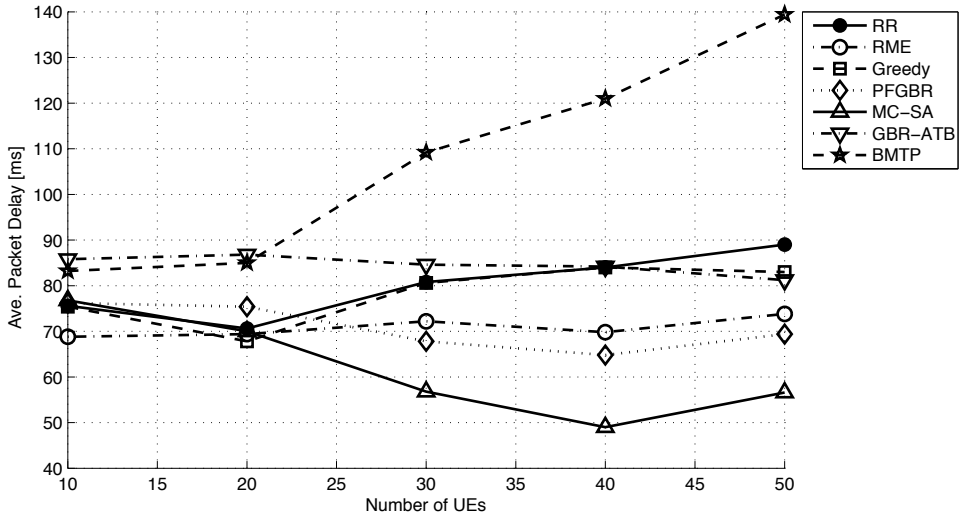

Figure 5.33: Experiment 3-2: Video Average Delay

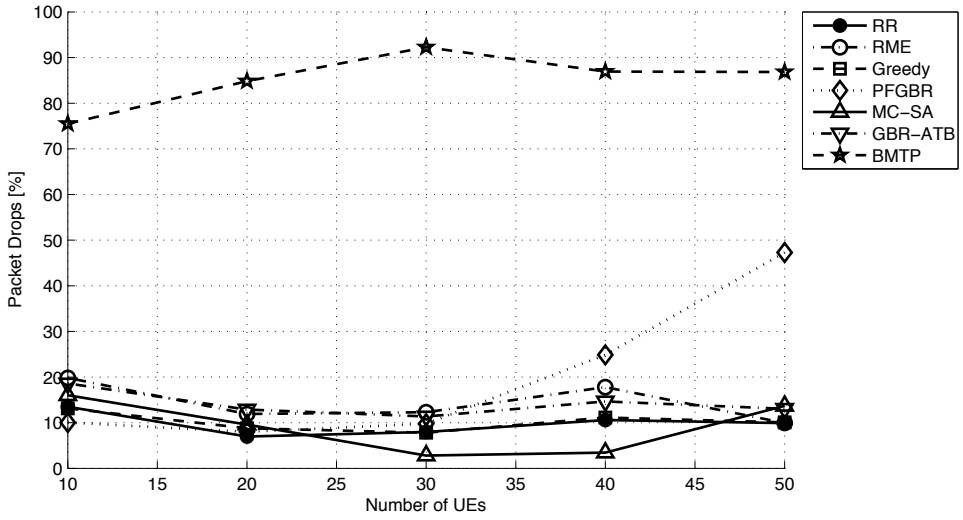

Figure 5.34: Experiment 3-2: VoIP Packet Drops

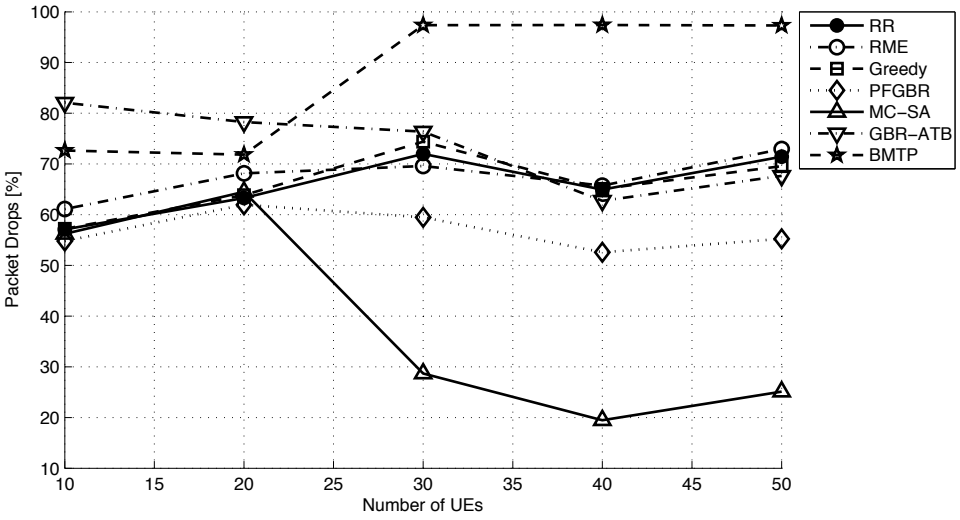

Figure 5.35: Experiment 3-2: Video Packet Drops

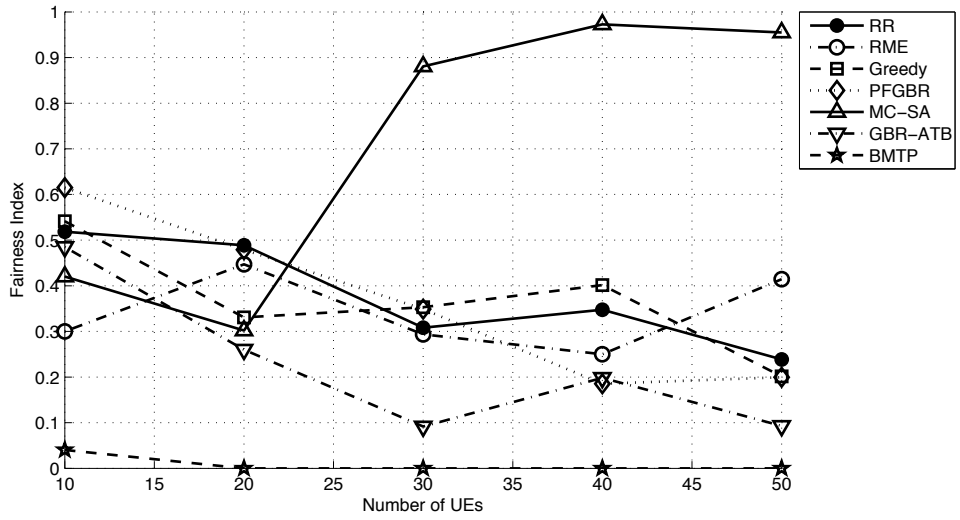

Figure 5.36: Experiment 3-2: VoIP Min-Max Fairness

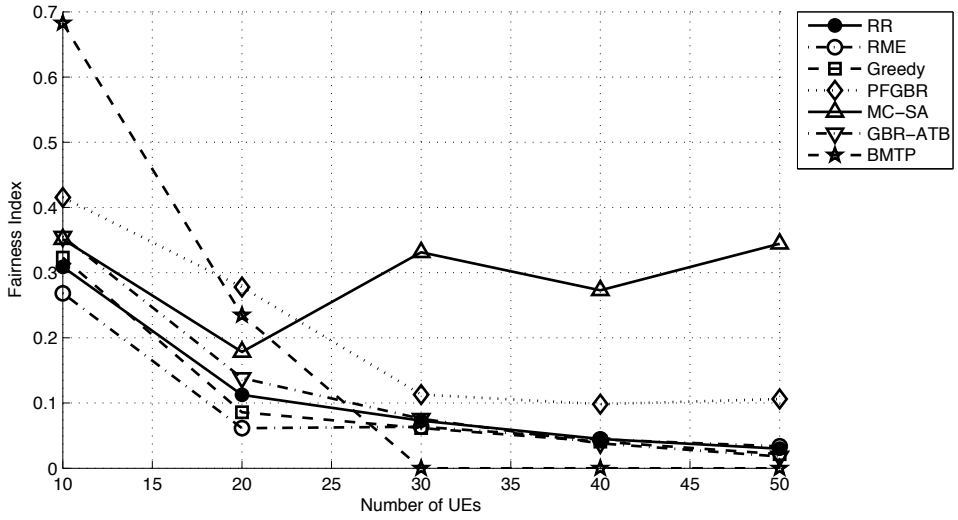

Figure 5.37: Experiment 3-2: Video Min-Max Fairness

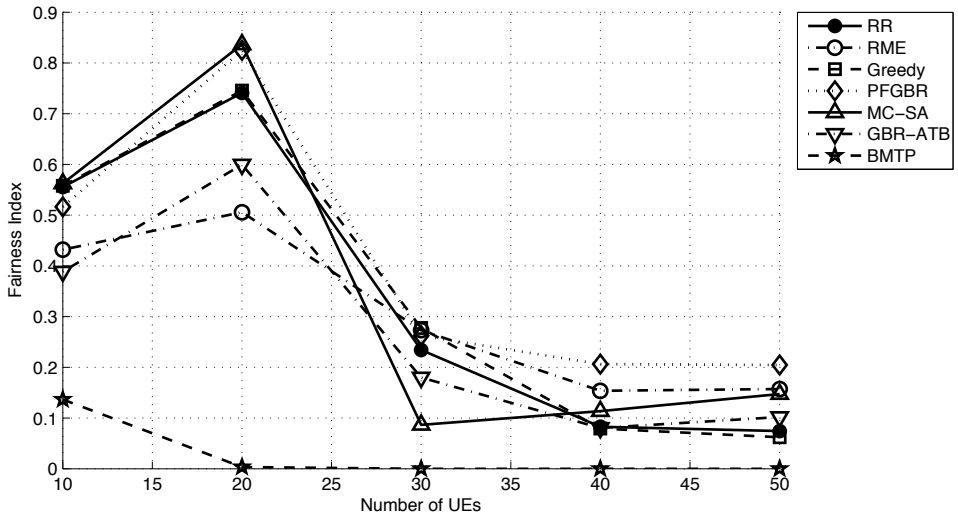

Figure 5.38: Experiment 3-2: FTP Min-Max Fairness

When comparing the throughput levels of UEs obtained in this experiment to those obtained from Experiment 1, we realize the large gaps between the total aggregate throughput of each scheduler from the two experiments. For optimal performance

of the schedulers under traffic mix conditions, it is recommended that QoS profiles of most of the UEs from different traffic classes to be similar, or at least the offered traffic loads from different UEs within the network to be of comparable magnitudes. The presence of large number of UEs with low traffic loads influences the schedulers to degrade the network's uplink performance, as seen before. VoIP traffic would be the most common low-rate service that is heavily present within LTE networks along with other higher data rate services. Solutions for supporting VoIP separately from other services has been addressed in LTE standard and literature work as well. However, these solutions are beyond the scope of our work.

5.2.4 Experiment 4: Effect of Variable Per-TTI Schedulable UE Subset Sizes on System Performance

As mentioned before, one of the main tasks of TDPS unit the the uplink packet scheduler is to perform UE prioritization and UE subset selection. UE subset selection involves selecting only a subset out of the all the UEs actively transmitting over the uplink. In the next TTI, the scheduler re-performs the UE prioritization in which UEs scheduled in the previous TTI get deprioritized further down the priority list, allowing for other UEs to be selected for scheduling in the next TTI. This addition of UE subset selecting was not proposed in any of the literature work from which schedulers under examination were selected. However, UE subset selection can be easily added to all the schedulers here as part of TD scheduling illustrated in Figure 3.1.

One of the authors in [35] has addressed the UE subset selection as part of the TD scheduling in his extended work in [47]. The simplest TD UE prioritization proposed

in [47] is a RR TD scheduling based on the UE's elapsed time since the last time it was allocated any PRB resources,

$$M^{TD} = tti_{curr} - tti_{sched} \tag{5.9}$$

where tti_{curr} refers to the current TTI in ms, and tti_{sched} represents the last TTI, measured in ms, a UE was allocated any resources.

In this experiment, we examine the effect of varying the UE subset size on the system performance in terms of throughput and fairness. The experiment was conducted with a total of forty UEs present in the system. Four different subset sizes were selected for evaluation, as indicated in Table 5.7. The traffic profile chosen in this experiment is the same as the one chosen in Experiment 1 to better evaluate the impact on the maximum attainable throughput of the system.

UE prioritization presented in (5.9) above is assumed for all schedulers, except for GBR-ATB and MC-SA. Both schedulers were proposed with explicit indication of UE prioritization based on prioritization metrics proposed in both literature work. As for the remaining schedulers, there were no explicit indication of creating a UE prioritized list when scheduling, and hence the RR-based prioritization was assumed.

Table 5.7: Simulation Parameters of Experiment 4

Parameter	**Description**
Number of UEs	40
Schedulable UE Set Sizes	10, 20, 30, 40
Traffic Profile	1 Mbps FTP stream per UE

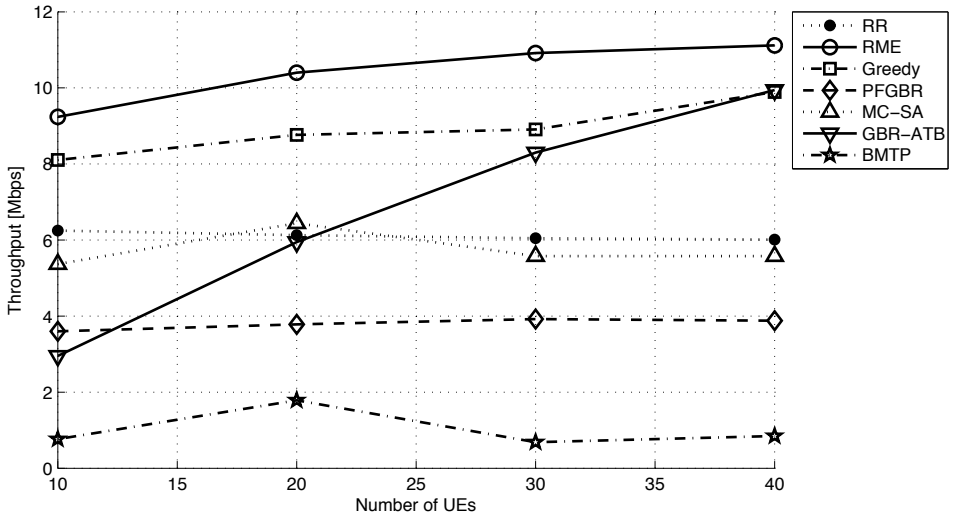

Figure 5.39: Experiment 4: Aggregated Throughput

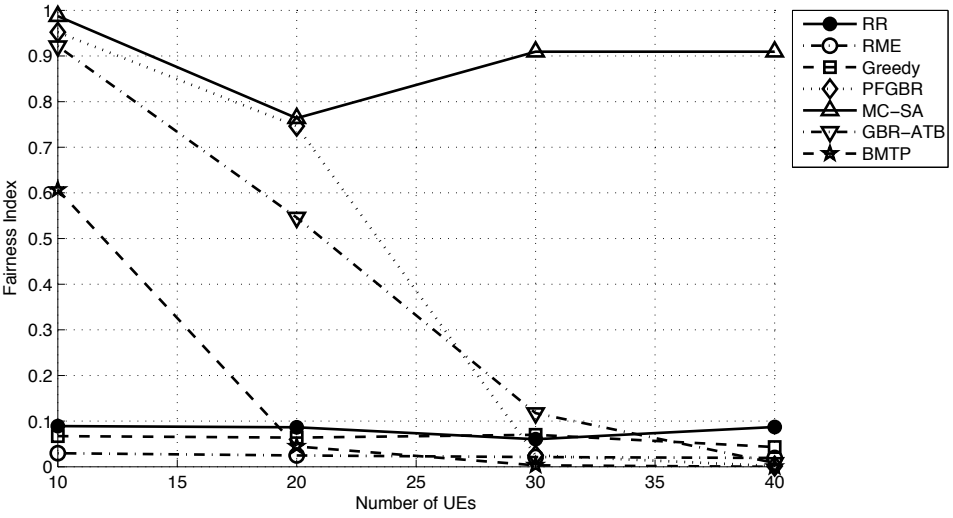

Figure 5.40: Experiment 4: Mix-Max Fairness Level

The results of the aggregate throughput and min-max fairness levels shown in Figures 5.39 and 5.40, respectively, illustrate that changing the size of schedulable UE subset per TTI has mainly affected the performance of GBR-ATB scheduler, while

RME and Greedy were less affected by the schedulable UE subset sizes. The remaining schedulers have shown steady performance levels regardless of the of the maximum schedulable UE subset size at any given TTI. Having RR-based TD scheduling has shown no performance change with respect to changing the schedulable UE subset size. In the case schedulers based on fixed sized PRB allocation such as Greedy and RR, reducing the maximum number of UEs that can be scheduled per TTI is counteracted by increasing the size of RC to be allocated to each UE. That is, with smaller UE subsets, UEs get less frequent, yet larger, scheduling grants.

As for RME and PFGBR schedulers, the number of PRBs allocated to each UE within a TTI greatly varies. However, the small impact of UE subset sizes on the performance is a good indication that the number of the allocation schemes result in not many UEs get PRBs assigned since the schedulers impose no limit on the number of PRBs that can be assigned to a single UE. For example, when comparing the throughput and fairness levels of both RME and PFGBR, we come to the conclusion that the when allowing up to forty UEs to be scheduled at any given TTI, there is a high probability that the number of UEs end up with resource allocation is ten or less.

GBR-ATB was the only scheduler that showed a considerable performance change due to changing the size of schedulable UE set per TTI. The performance of scheduler in this case is very similar to that of Experiment 1 where the schedulable set is the same as the total number of UEs within the system. First of all, this is an indication of the effectiveness of the TD metric used for UE prioritization shown in (4.6) which is a ratio of the GBR defined for the UE and its experienced average throughput. In this experiment where there is only one type of traffic, only the average throughput is

what plays a key role in prioritizing UEs. This way, the algorithm ensures that UEs with the worst average throughput experienced are always pushed towards the top of the list. As a result, with smaller size schedulable UE sets, the algorithm can provide better fairness level though at the expense of degrading the experienced aggregate throughput of the system.

5.2.5 UE Uplink Power Utilization Under Different Uplink Packet Schedulers

We recall from Chapter 2 that one of the major reasons for choosing SC-FDMA rather than OFDMA for LTE uplink is to reduce PAPR level. Except for BMTP and GBR-ATB, schedulers presented in our our study as well as most schedulers proposed in literature overlook monitoring uplink transmission power consumption for the sake of optimizing objectives like fairness and QoS.

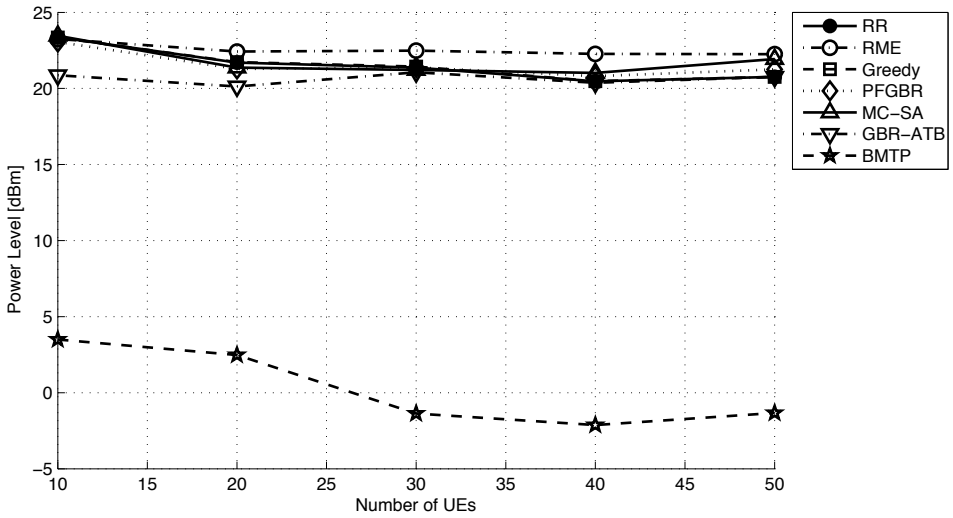

Figure 5.41: Per UE average uplink power utilization obtained from Experiment 3

Figure 5.41 demonstrates the uplink power utilization behavior from Experiment 3. Except for the BMTP, all schedulers show power usage around 22 dBm, which is close to the maximum power usage listed in Table 5.1. When the number of UEs is small, each UE's tends to be assigned PRBs more frequently. The power level assigned per UE based on (5.5) is directly proportional to the number of PRBs it is assigned. Hence, UEs in this case will tend to transmit at a larger power level for error-free data transmission over the assigned PRBs.

The distance between the UE and eNodeB is another important factor that affects the uplink power usage, where both the UE-to-eNodeB distance and the uplink power usage are directly proportional as shown in (5.5) as well. Due to the path loss and shadowing experienced by UEs, the further away the UEs are from the eNodeB, the more power UEs would need to maintain a certain error rate given the same MCS level the same number of allocated PRBs. With the exception of BMTP and GBR-ATB, schedulers do not provide any power usage provision when allocating PRBs to UEs that are far away from the eNodeB. Hence, when there are higher number of UEs in the system, more UEs use full power to transmit their data, where such UEs are usually located either midrange or close to the cell edge.

BMTP's greedy allocation method allows UEs to be scheduled if they are close enough to the eNodeB. Hence, as explained earlier, BMTP allocates close UEs either most or all the PRBs to reduce the average uplink power usage. UEs that are further away rarely get any PRBs allocation because they would need higher power level than their maximum power budget to maintain a certain minimum bit rate. Lowering the minimum bit rate means less power is needed for uplink transmission. Hence, cell edge UEs with low traffic loads, such as UEs with VoIP traffic, has better chances of

transmitting data than UEs with higher data rate demands.

Hence, BMTP is not recommended as an LTE uplink scheduler at the eNodeB within an E-UTRAN system. BMTP scheduler can be, however, implemented for cases where limiting the coverage range of a cell becomes desirable, as in the case of femtocells. The femtocells are by default low power eNodeBs which were introduced to the LTE standard in 3GPP Release 9 as Home eNodeB [48]. LTE femtocells were designed to operate with limited coverage, low power, and with small number of UEs [48]. Hence, BMTP would be more suitable in femtocell environments rather than in micro or macro cell setup.

5.2.6 Results Summary

The performance evaluation of the representative scheduling algorithms examined here comprise of four different experiments. The first experiment was conducted under full non-GBR traffic load to examine both the ability of the schedulers to maximize the utilization of the available bandwidth, as well as the fairness level they can achieve. The second experiment was to simulate the schedulers under different scenarios with several traffic mixes of VoIP, video streaming, and FTP. The third experiment was to look at the system performance achievable under the provisioning of the selected schedulers with variable and fixed traffic load.

RME algorithm demonstrated high throughput levels in single traffic case. RME's adaptive allocation scheme provides more freedom to assign PRBs to whichever UE that maximizes the PF utility function of the system, even if all PRBs are allocated to as few as one UE. However, RME fairness level degrades rapidly with increasing UEs in the cell. This is due to the failure of PF metric to provide the necessary fairness

level in cases of micro and macro cell coverage. Also, the performance of RME tends to suffer in the presence of mixed traffic scenarios. Again, this is an indication that UEs under RME management need to have the similar offered load to perform proper PF-based resource allocation.

Greedy algorithm is another best-effort scheduler, like RME, that relies on PF utility metric when assigning PRBs. However, Greedy quantizes schedulable PRBs to contiguous PRB groups, which are also known as RCs. Throughput level under fixed-size PRB allocation scheme such as Greedy scheme might be degraded with few UEs in the system, but the throughput becomes comparable to RME's with higher number of UEs. Imposing an upper limit on the number of schedulable PRBs per a single UE prevented Greedy's performance in the case of multiple traffic classes as in teh case of RME. Hence, Greedy algorithm's results indicated that a PF scheduler with an upper limitation per UE PRB grants can provide an improved support for UEs with mixed QoS requirements.

GBR-ATB algorithm can be viewed as adaptation of RME with the introduction of a utility function based on both GBR and PF. GBR-ATB also differs from RME in having an additional restriction stopping resource allocation to a UE if the UE is to exceed its power limitation. In the case of UEs with single traffic profile scenario, GBR-ATB's throughput suffers from the power-based restriction on the number of UEs to be assigned per UE. The throughput level improves with higher number of UEs, though the scheduler's fairness level goes down with more UEs in the cell, as in the case with RME and Greedy. However, with the presence of large number of UEs, GBR-ATB fairness level can be improved when reducing the size of UE subset that can be scheduled per TTI. Also, GBR-ATB has performed as expected by providing fair

support for the multiple UEs with mixed traffic requirements due to the scheduler's GBR-awareness.

PFGBR scheduler has demonstrated a good QoS support in the case of multiple traffic mixes in Experiment 2. The good performance level was mainly due to the different MBR values of the simulated traffic classes within the network. Having UEs with varied MBRs provides different cut-offs for the maximum number of PRBs that are allowed to be scheduled per UE, which gives more chances for transmissions from different QoS classes. However, PFGBR's performance seems to degrade in the case of single traffic profile as the number of schedulable UEs increase. With the presence of single MBR limit for all UEs, the effect of resource allocation scheme takes place, which is based on per-PRB-set allocation rather than per-PRB allocation. Once a UE is selected for scheduling, the higher metric usually tends to refer to the largest PRB-set. Hence, the scheduler ends up scheduling few UEs in every iteration, which depraves PRBs from other UEs that can better optimize the overall system performance.

MC-SA provides a better fairness level among all the selected schedulers in either the single-traffic scenarios and for higher priority traffic in the case of having traffic mix. This is due to the UE-prioritization scheme that ensures all UEs to get as fair a chance for transmission as possible. MC-SA's fairness level for traffic classes of lower priority tends to have performance levels close to the other schedulers. MC-SA reduces the availability of PRB resources to lower priority traffic classes, which explains the degradation their fairness levels, especially when the number of UEs increase in the cell.

BMTP has the lowest power utilization out of all the schedulers, due to its concept

of allocating more PRBs to a UE while maintaining the same throughput level to reduce the transmitted power. BMTP scheduling scheme, although has proven to be effective in conserving power, yet has a degraded performance level due to inefficient usage of the available spectrum.

5.3 Complexity Analysis

The complexity analysis of the examined schedulers is based on the number of iterations it takes a scheduler's algorithm in searching for the final UE-to-PRB mapping. The search algorithm is the one that consumes most of the scheduler's total operation time in comparison to the computation of the utility-based metrics.

Except for RR, the complexity of the scheduler algorithms examined here are mainly a function of the number of UEs, N and number of PRBs N_{PRB} in the system. The baseline RR algorithm is known to have a constant complexity of $O(1)$ irrespective of the number of UEs in the system. Greedy, RME, and GBR-ATB are similar in the sense that they linear matrix-based search, where they construct the UE-PRB metric matrix M and perform a greedy-based, linear search where a PRB is never to be revisited once it is allocated to a UE. The worst case scenario of Greedy algorithm is when N is larger than N_{PRB}, in which case the size of an RC corresponds to one PRB. In that case, a total of $\sum_{u=1}^{N} u \cdot (N_{PRB} - (N - u))$. Hence, the overall complexity of the scheduler is $O(N \cdot N_{PRB})$.

As for RME, the number of search operations depend on experienced metric values [33]. A worst case scenario occurs when the algorithm performs recursive iterations for all UEs, in which no UE can expand on the both sides of their first assigned PRB with their maximum metric. In that case, the number of iterations is the same as in

the case of Greedy. For the remaining $N_{PRB} - N$ PRBs, the algorithm performs a search over the entire available UE-PRB metrics. Hence, at any given iteration with b PRBs still available, there is a total of $N \cdot b$ search operations to find which PRB is to assigned to which UE. Hence, the total number of search iterations becomes $\sum_{u=1}^{N} u \cdot (N_{PRB} - (N - u)) + \sum b = 1^{N_{PRB}-N} N \cdot b$, resulting also in a complexity of $O(N \cdot N_{PRB})$. GBR-ATB search algorithm performs metric search very similar to RME, and hence has the same complexity level.

As for PFGBR, the worse case scenario occurs when the maximum bit rates of all UEs are set such that a UE can be allocated the entire bandwidth. Then, the first iteration of the scheduler would search through the entire $N \times N_C$ possible allocations to find the one with the maximum metric. When this happens, the allocated UE is removed from the schedulable UE list. Also, the algorithm searches through the entire set of possible PRB-subset for each UE, and removes any PRB-subset that share any PRB that has already been allocated. Hence, the total number of search operations can be expressed as $\sum_{u=1}^{N} (N - u) \cdot (N_C - x_u)$, where x_u is the number of PRB-allocation patterns that have been removed from list of possible PRB allocation sets up till iteration u. Hence, the number of possible allocation scenarios get reduced at an accelerating rate as scheduling progresses, having the algorithm terminating in far less iterations than the number of allocation possibilities the algorithm started with. As a result, The PFGBR's complexity can be expressed as $O(\log(N) \cdot \log(N_{PRB}))$

BMTP phase 1 algorithm's complexity can be derived in a similar fashion to that of PFGBR. However, if there are n_{PRB} PRBs remain unassigned from the first procedure, the algorithm moves to BMTP phase 2, which referred in [41] as MPD algorithm. The worst case scenario of MPD is when no PRBs get assigned in the

BMTP phase 1, then allocates all PRBs. In this case, the algorithm performs a greedy search for the PRB that minimizes the power decrease of a UE if assigned the designated PRB, resulting in $\sum_{u=1}^{N} N \cdot (N_{PRB} - u) = N \cdot N_{PRB}(N_{PRB} - 1)/2$ iterations. MPD's complexity is expressed as $O(N \cdot N_{PRB}^2)$, thus dominating the complexity of the first phase in BMTP.

As for MC-SA, the complexity of the scheduler differs depending on which subroutine is executed. In the case of the first subroutine, the scheduling procedure here is very similar to the one in Greedy algorithm, and hence the complexity in this case is the same as that of Greedy. In the second subroutine where the number of UEs is greater than the number of available PRBs, the scheduler iterates through UEs in the priority list, and searches for PRB with the best channel quality. In the worst case, the scheduler iterates through the first N_{PRB} UEs and assign each one PRB. Therefore, the complexity in this case is expressed as $O(N_{PRB}^2)$.

Table 5.8 provides the complexity notations of all the examined schedulers.

Table 5.8: Summary of Schedulers' Complexity Levels.

Scheduler	**Complexity**
RR	$O(1)$
RME	$O(N \cdot N_{PRB})$
Greedy	$O(N \cdot N_{PRB})$
GBR-ATB	$O(N \cdot N_{PRB})$
PFGBR	$O(\log(N) \cdot \log(N_{PRB}))$
MC-SA	$O(N_{PRB}^2)$
BMTP	$O(N \cdot N_{PRB}^2)$

Chapter 6

Summary and Conclusions

The success of the HSPA deployments caused a revolution in the development of more advanced mobile hardware and also the emergence of Internet services that run over mobile networks such as mobile browsing, video streaming, and interactive gaming. The rapid increase in the mobile traffic load will lead to the inability of HSPA at one point to accommodate forecasted volumes.

LTE is considered a major milestone towards the evolution of currently deployed 3G technologies. 3GPP has proposed LTE to ensure a competitive future mobile market where mobile broadband provides means to deploy mobile networks with higher data rate networks at considerable costs. LTE promises performance levels that supersede its predecessor HSPA in terms of higher throughput, lower latency, more simplified architecture and easier deployment. The LTE standard defines Layer 1 (PHY) and Layer 2 (MAC/RLC/PDCP) specifications. The PHY layer in LTE is based on OFDM modulation which supports higher data rates than WCDMA while reducing the ISI interference of successive transmitted OFDM symbols. The LTE uplink radio interface is SC-FDMA, an OFDMA-variant that has the main feature of

reducing the PAPR to reduce the power usage on uplink transmissions. LTE defines the Physical Resource Block (PRB) as the smallest time/frequency unit that spans 180 kHz in frequency and 0.5 ms in duration.

LTE's move towards a flat, fully IP-based platform dictates an enhanced support of QoS requirements to address the needs of a variety of mobile services with multiple QoS requirements. QoS enforcement is supported at the system access level through radio resource management entities like LTE packet scheduling.

Packet scheduling in LTE refers to deciding which UE is to transmit over which SC-FDMA radio resources. The scheduling decision is repeated periodically once every 1 ms, which is alternatively known as a TTI. An LTE packet scheduler should optimally predict the needs of UEs with different requirements, such that a UE can satisfy its QoS requirements while at the same time try to utilize the resources assigned to it as efficiently as possible.

The focus of the performance evaluation was to provide a comparative study on LTE uplink scheduler proposals from multiple proposals in literature. Many LTE uplink schedulers were surveyed and categorized based on commonly shared characteristics among each group of schedulers. Representative schedulers were selected from each category and examined in this study were Best-Effort schedulers, QoS-based packet schedulers, and power-optimizing packet schedulers.

The performance analysis on the LTE uplink schedulers was conducted in an in-house packet-level simulator that was developed in adherence to the LTE standards. The performance metrics used to measure the performance levels of the schedulers were: aggregate throughput, inter-class fairness, intra-class fairness, average packet delay, packet drops, transport block utilization, and uplink power utilization.

Best-effort schedulers such as RME and Greedy are best suited for supporting single traffic scenarios or for traffic mixes of similar QoS attributes, such as FTP and web browsing from non-GBR traffic. The RME algorithm is more sensitive to variations in traffic loads among UEs, as its reliance on the PF metric alone and its dynamic scheduling scheme leads to resource starvation of UEs with high data rates services in the presence of UEs with low data rate services. This leads to assigning UEs with low traffic rates more PRBs than necessary, while other UEs with heavier traffic loads can suffer from resource starvation. Greedy accounts for such scenarios by equally dividing PRBs equally among the UEs to provide as equal chances for uplink transmission as possible. Also, the fixed resource allocation method used in Greedy makes it also perform better than RME in scenarios with mixed traffic loads. In conclusion, Best-Effort schedulers with no QoS supervision can perform best if combined with other QoS-aware schedulers where each scheduler is to support different traffic classes.

QoS-based schedulers perform well in scenarios of mixed traffic profiles. GBR-ATB showed comparable performance level to best-effort schedulers like greedy in the case of single traffic scenarios. However, the scheduler demonstrated better accommodation for multiple traffic scenarios. GBR-ATB, shows that combining dynamic scheduling similar to RME with a QoS-based utility function can greatly improve the QoS performance of the scheduler. PFGBR also provides good scheduling support for mixed traffic scenarios. However, when the network experiences heavy traffic load from a single traffic type, PFGBR falls short in performance as its throughput, as well as its intra-class fairness level, decline with the increase in the number of UEs. This is due to UE-to-PRB-subset allocation scheme in which the scheduler assigns chunks

of PRBs in each iteration, ending with only a few UEs with allocations. MC-SA's two allocation schemes, greedy-subroutine and UE-prioritization-based subroutines showed to provide high fairness levels in allocating PRBs to UEs, while ensuring UEs with higher quality get high priority in QoS guarantees. All the three QoS-based schedulers exhibit similar behavior in the sense that they are able to prioritize FTP traffic such that they would not suffer starvation even with the presence of high concentration of other traffic types.

BMTP has shown considerable power savings compared to other schedulers. However, BMTP performs allocation that causes very poor utilization of the radio resources, hence the scheduler is unable to provide the necessary capacity to support large numbers of UEs. Also, since the scheduler performs a greedy selection based on minimum power usage, the scheduler only selects UEs that are closer to the eNodeB, while UEs that are closer to the cell edge are never scheduled. Accordingly, the power saving techniques employed in BMTP might prove useful in small scale broadband coverage, such as femtocells. Because femtocells are designed for small coverage and a small number of UEs, BMTP can show more promising performance levels if it is implemented at a femtocell base station.

Bibliography

[1] 3GPP, "Policy and Charging Control Architecture," TR 23.203, 3rd Generation Partnership Project (3GPP), Mar. 2010.

[2] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2," TS 36.300, 3rd Generation Partnership Project (3GPP), Dec. 2010.

[3] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channels and Modulation," TR 36.211, 3rd Generation Partnership Project (3GPP), Mar. 2010.

[4] A. Pokhariyal, K. Pedersen, G. Monghal, I. Kovacs, C. Rosa, T. Kolding, and P. Mogensen, "HARQ Aware Frequency Domain Packet Scheduler with Different Degrees of Fairness for the UTRAN Long Term Evolution," in *Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th*, pp. 2761 –2765, Apr. 2007.

[5] G. Americas, "HSPA to LTE-Advanced: 3GPP Broadband Evolution to IMT-Advanced," ts, 4G Americas.org, Sept. 2009.

[6] Ericsson, "Long Term Evolution (LTE) - An Introduction," ts, Ericsson, June 2009.

[7] "3gpp," 2011. Available from http://www.3gpp.org/; accessed 15 February 2011.

[8] "4g americas," 2011. Available from http://www.4gamericas.org/; accessed 11 March 2011.

[9] A. Pokhariyal, T. Kolding, and P. Mogensen, "Performance of downlink frequency domain packet scheduling for the utran long term evolution," in *Personal, Indoor and Mobile Radio Communications, 2006 IEEE 17th International Symposium on*, pp. 1 –5, Sept. 2006.

[10] S. Nagata, Y. Ofuji, K. Higuchi, and M. Sawahashi, "Optimum resource block bandwidth for frequency domain channel-dependent scheduling in evolved utra downlink ofdm radio access," in *Vehicular Technology Conference, 2006. VTC 2006-Spring. IEEE 63rd*, vol. 1, pp. 206 –210, May 2006.

[11] N. Miki, Y. Kishiyama, K. Higuchi, and M. Sawahashi, "Optimum Adaptive Modulation and Channel Coding Scheme for Frequency Domain Channel-Dependent Scheduling in OFDM Based Evolved UTRA Downlink," in *Wireless Communications and Networking Conference, 2007.WCNC 2007. IEEE*, pp. 1783 –1788, Mar. 2007.

[12] A. Pokhariyal, G. Monghal, K. Pedersen, P. Mogensen, I. Kovacs, C. Rosa, and T. Kolding, "Frequency domain packet scheduling under fractional load for the utran lte downlink," in *Vehicular Technology Conference, 2007. VTC2007-Spring. IEEE 65th*, pp. 699 –703, Apr. 2007.

[13] P. Kela, J. Puttonen, N. Kolehmainen, T. Ristaniemi, T. Henttonen, and M. Moisio, "Dynamic packet scheduling performance in utra long term evolution downlink," in *Wireless Pervasive Computing, 2008. ISWPC 2008. 3rd International Symposium on*, pp. 308 –313, May 2008.

[14] R. Kwan, C. Leung, and J. Zhang, "Proportional fair multiuser scheduling in lte," *Signal Processing Letters, IEEE*, vol. 16, pp. 461 –464, June 2009.

[15] B. Sadiq, R. Madan, and A. Sampath, "Downlink scheduling for multiclass traffic in lte," *EURASIP J. Wirel. Commun. Netw.*, vol. 2009, pp. 14:9–14:9, mar 2009.

[16] H. G. Myung, J. Lim, and D. J. Goodman, "Single carrier fdma for uplink wireless transmission," *Vehicular Technology Magazine, IEEE*, vol. 1, pp. 30 –38, Sept. 2006.

[17] J. Lim, H. Myung, K. Oh, and D. Goodman, "Proportional Fair Scheduling of Uplink Single-Carrier FDMA Systems," pp. 1 –6, sep 2006.

[18] 3GPP, "Technical Specification Group Radio Access Network; Requirements for Evolved UTRA (E-UTRA) and Evolved UTRAN (E-UTRAN)," TR 25.913, 3rd Generation Partnership Project (3GPP), Dec. 2009.

[19] 3GPP, "Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Architecture description," TR 36.401, 3rd Generation Partnership Project (3GPP), June 2010.

[20] K. Bogineni, R. Ludwig, P. Mogensen, V. Nandlall, V. Vucetic, B. Yi, and Z. Zvonar, "LTE Part I: Core network," *Communications Magazine, IEEE*, vol. 47, pp. 40 –43, Feb. 2009.

[21] E. Dahlman, S. Parkvall, J. Skold, and P. Beming, *3G Evolution, Second Edition: HSPA and LTE for Mobile Broadband*. Academic Press, 2008.

[22] C. Ciochina and H. Sari, "A review of ofdma and single-carrier fdma," in *Wireless Conference (EW), 2010 European*, pp. 706 –710, Apr. 2010.

[23] G. Berardinelli, L. Ruiz de Temino, S. Frattasi, M. Rahman, and P. Mogensen, "Ofdma vs. sc-fdma: performance comparison in local area imt-a scenarios," *Wireless Communications, IEEE*, vol. 15, pp. 64 –72, Oct. 2008.

[24] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Medium Access Control (MAC) Protocol Specification," TS 36.321, 3rd Generation Partnership Project (3GPP), Jan. 2010.

[25] A. Larmo, M. Lindstrom, M. Meyer, G. Pelletier, J. Torsner, and H. Wiemann, "The LTE Link-Layer Design," *Communications Magazine, IEEE*, vol. 47, pp. 52 –59, Apr. 2009.

[26] 3GPP, "Physical Layer Aspect for Evolved Universal Terrestrial Radio Access (UTRA)," TR 25.814, 3rd Generation Partnership Project (3GPP), Oct. 2006.

[27] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Resource Control (RRC); Protocol Specification," TS 36.331, 3rd Generation Partnership Project (3GPP), June 2010.

[28] 3GPP, "Services and service capabilities (Release 9)," TR 22.105, 3rd Generation Partnership Project (3GPP), Dec. 2008.

[29] 3GPP, "Quality of Service (QoS) concept and architecture (Release 9)," TR 23.107, 3rd Generation Partnership Project (3GPP), June 2010.

[30] 3GPP, "Channel-Dependent Scheduling Method for Single-Carrier FDMA Radio Access in Evolved UTRA Uplink," TS R1-050701, 3rd Generatin Partnership Project (3GPP), Sept. 2005.

[31] M. Al-Rawi, R. Jantti, J. Torsner, and M. Sagfors, "Opportunistic Uplilnk Scheduling for 3G LTE Systems," in *Innovations in Information Technology, 2007. IIT '07. 4th International Conference on*, pp. 705 –709, 18-20 2007.

[32] M. Al-Rawi, R. Jantti, J. Torsner, and M. Sagfors, "On the Performance of Heuristic Opportunistic Scheduling in the Uplink of 3G LTE Networks," in *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*, pp. 1 –6, 15-18 2008.

[33] L. Ruiz de Temino, G. Berardinelli, S. Frattasi, and P. Mogensen, "Channel-Aware Scheduling Algorithms for SC-FDMA in LTE Uplink," in *Personal, Indoor and Mobile Radio Communications, 2008. PIMRC 2008. IEEE 19th International Symposium on*, pp. 1 –6, 15-18 2008.

[34] F. Liu, X. She, L. Chen, and H. Otsuka, "Improved Recursive Maximum Expansion Scheduling Algorithms for Uplink Single Carrier FDMA System," pp. 1 –5, may. 2010.

[35] F. Calabrese, P. Michaelsen, C. Rosa, M. Anas, C. Castellanos, D. Villa, K. Pedersen, and P. Mogensen, "Search-Tree Based Uplink Channel Aware Packet Scheduling for UTRAN LTE," in *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE*, pp. 1949 –1953, 11-14 2008.

[36] P. Wen, M. You, S. Wu, and J. Liu, "Scheduling for streaming application over wideband cellular network in mixed service scenarios," in *Military Communications Conference, 2007. MILCOM 2007. IEEE*, pp. 1 –7, Oct. 2007.

[37] M. Anas, C. Rosa, F. Calabrese, K. Pedersen, and P. Mogensen, "Combined Admission Control and Scheduling for QoS Differentiation in LTE Uplink," in *Vehicular Technology Conference, 2008. VTC 2008-Fall. IEEE 68th*, pp. 1 –5, Sept. 2008.

[38] O. Delgado and B. Jaumard, "Scheduling and Resource Allocation in LTE Uplink with a Delay Requirement," in *Communication Networks and Services Research Conference (CNSR), 2010 Eighth Annual*, pp. 268 –275, May 2010.

[39] O. Delgado and B. Jaumard, "Scheduling and Resource Allocation for Multiclass Services in LTE Uplink Systems," in *Wireless and Mobile Computing, Networking and Communications (WiMob), 2010 IEEE 6th International Conference on*, pp. 355 –360, Oct. 2010.

[40] Z. Li, C. Yin, and G. Yue, "Delay-bounded power-efficient packet scheduling for uplink systems of lte," in *Wireless Communications, Networking and Mobile Computing, 2009. WiCom '09. 5th International Conference on*, pp. 1 –4, Sept. 2009.

[41] F. Sokmen and T. Girici, "Uplink Resource Allocation Algorithms for Single-Carrier FDMA Systems," in *Wireless Conference (EW), 2010 European*, pp. 339 –345, Apr. 2010.

[42] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Radio Frequency (RF) System Scenarios," TR 36.942, 3rd Generation Partnership Project (3GPP), Apr. 2010.

[43] 3GPP, "Radio transmission and reception," TR 45.005, 3rd Generation Partnership Project (3GPP), June 2010.

[44] F. Calabrese, *Adaptive Radio Resource Management for Uplink Wireless Networks*. Phd thesis, Polytechnic University, jun 2006.

[45] 3GPP, "Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Layer Procedures," TR 36.213, 3rd Generation Partnership Project (3GPP), Mar. 2010.

[46] Orange, China Mobile, KPN, NTT DoCoMo, Sprint, T-Mobile, Vodafone, Telecom Italia, "LTE Physical Layer Framework for Performance Verification," TS R1-070674, 3rd Generatin Partnership Project (3GPP), Feb. 2007.

[47] F. Calabrese, *Scheduling and Link Adaptation for Uplink SC-FDMA Systems: A LTE Case Study*. Phd thesis, Aalborg University, apr 2009.

[48] A. Golaup, M. Mustapha, and L. Patanapongpibul, "Femtocell Access Control Strategy in UMTS and LTE," *Communications Magazine, IEEE*, vol. 47, pp. 117 –123, Sept. 2009.