

# Dynamic Resource Allocation Algorithms for Long Term Evolution (LTE) Wireless Broadband Networks

by

Peng Sun

Supervisor: Dr. Hassan Naser

A Thesis Presented to Lakehead University

in Partial Fulfillment of the Requirement for the Degree of

Master of Science in Electrical and Computer Engineering

Faculty of Engineering

Lakehead University

Thunder Bay, Ontario, Canada 2010

Copyright © Peng Sun, 2010

# Abstract

Following the successful standardization of High-Speed Packet Access (HSPA), the 3rd Generation Partnership Project (3GPP) recently specified the Long Term Evolution (LTE) as a next generation radio network technology to meet the increasing performance requirements of mobile broadband. The results include a flexible and spectrally efficient radio link protocol design with low overhead. The first release of LTE provides peak rates of 300 Mbps in downlink and 75 Mbps in uplink. It is a significant increase in spectrum efficiency compared to the previous cellular systems. Single-Carrier Frequency Division Multiple Access (SC-FDMA) has been selected as the uplink access scheme in the LTE. With SC-FDMA, the frequency spectrum resource is divided into time-frequency grids, referred to as resource blocks (RBs). Multiple-access is achieved by distributing resource blocks to users. The function of resource block allocation algorithms is to distribute resource blocks among users in a fair and efficient manner. The Modulation and coding scheme is determined adaptively according to the time-varying channel conditions. Sounding Reference Signals (SRS) are transmitted in the uplink direction to allow for the base station to estimate the uplink channel quality at different frequencies. The LTE system supports wideband SRS and narrowband SRS.

In this thesis, we propose two resource-block allocation algorithms in or-

der to provide quality of service (QoS) for the uplink transmission in the LTE system. Various classes of traffic are transmitted based on their QoS requirements, transmission history, and their channel conditions. The first algorithm is called: Class-of-service Aware Scheduling Algorithm (CASA), which employs the wideband SRS technique together with a credit pooling technique to allocate resource blocks among customers' classes of service in a fair manner. In CASA, a weighted round robin arbitration mechanism is employed to determine the priority of users in each Class-of-service (CoS) by their long term transmission records. The second algorithm is called: Independent Channel Aware Scheduling Algorithm (ICAS), which employs the narrowband SRS technique. ICAS algorithm divides resource blocks into a number of groups, and employs a priority based scheduling to distribute these groups among customers. Both of these algorithms support Class-of-service (CoS) capabilities. In particular, we test both algorithms with three CoSs which are based on the well-known Differentiated Services IP (DiffServ). These are called: Expedited Forwarding (EF), Assured Forwarding (AF) and Best Effort (BE). We also consider the channel condition in our algorithms based on the Adaptive Modulation and Coding (AMC) scheme.

We have developed an in-house simulation program in C++ to evaluate and compare the performance of CASA and ICAS algorithms in terms of packet loss ratio, delay, and throughput. Simulation results show that the proposed algorithms are able to satisfy the QoS requirements. Both of proposed algorithms support multiple CoSs simultaneously without impeding the first class (Expedited Forwarding) transmission. Also, both of the proposed algorithms achieve high throughput in a large range cell.

# Acknowledgment

I would like to thank my supervisor Dr. Hassan Naser for his guidance and patience throughout while completing this thesis. It is my firm belief that the best professors only need to provide a guiding hand to their graduate students.

I would also like to extend my thanks to my family and friends. Without all your love and support, this would not be possible.

# Contents

<b>1</b>	<b>Introduction</b>	<b>12</b>
1.1	Review of previous works . . . . .	14
1.2	Thesis Contributions . . . . .	17
1.3	Thesis Outline . . . . .	18
<b>2</b>	<b>Physical Layer for LTE Uplink</b>	<b>19</b>
2.1	Introduction . . . . .	19
2.2	Single Carrier Frequency Division Multiple Access Principles . . . . .	20
2.3	Uplink Physical Resource . . . . .	21
2.4	Uplink Reference Signals . . . . .	25
2.4.1	Sounding Reference Signal . . . . .	26
2.5	Summary . . . . .	30
<b>3</b>	<b>System Architecture</b>	<b>31</b>
3.1	Basic Model of LTE Wireless System . . . . .	31
3.2	Transmission Cycle . . . . .	32
3.3	Architecture of the scheduling algorithm . . . . .	34
3.4	Uplink Channel Model . . . . .	38
3.5	Parameters of AMC mode . . . . .	39
3.6	Buffer management in UE . . . . .	40

<i>CONTENTS</i>	5
3.7 Summary . . . . .	43
<b>4 Class-of-service Aware Scheduling Algorithm</b>	<b>44</b>
4.1 System Structure . . . . .	44
4.2 Grant Processing . . . . .	49
<b>5 Independent Channel Aware Scheduling Algorithm</b>	<b>53</b>
5.1 System Structure . . . . .	53
5.2 Grant Processing . . . . .	56
<b>6 Simulation Analysis</b>	<b>65</b>
6.1 The traffic models for CoS . . . . .	65
6.2 Simulation setup . . . . .	69
6.3 Simulation results and analysis . . . . .	70
6.3.1 Simulation results with 0.5 ms TTI . . . . .	71
6.3.2 Simulation results with 1.0 ms TTI . . . . .	76
6.3.3 Simulation results comparison with 0.5 ms TTI and 1.0 ms TTI . . . . .	82
6.3.3.1 The comparison of ICAS performance with 0.5 ms TTI and 1.0 ms TTI . . . . .	82
6.3.3.2 The comparison of CASA performance with 0.5 ms TTI and 1.0 ms TTI . . . . .	86
<b>7 Conclusion and Future Works</b>	<b>91</b>

# List of Figures

2.2.1 SC-FDMA transmitter and Receiver . . . . .	21
2.3.1 General LTE uplink time-domain structure . . . . .	22
2.3.2 Details of the LTE uplink structure . . . . .	23
2.3.3 Structure of Resource Block in frequency domain . . . . .	24
2.3.4 Resource block structure in both frequency and time domain . . . . .	25
2.3.5 Illustrations for both FDD and TDD . . . . .	26
2.4.1 Illustration of the SRS position with 2 <i>ms</i> interval . . . . .	27
2.4.2 Illustration of the SRS parallel transmission . . . . .	28
2.4.3 Operations of Wideband SRS (a) and Narrowband SRS (b) . . . . .	29
3.2.1 Illustration of transmission cycle . . . . .	33
3.3.1 The general architecture of the scheduling algorithm . . . . .	34
3.3.2 The resource-allocation matrix for 0.5 <i>ms</i> resource blocks . . . . .	36
3.3.3 The resource-allocation matrix for 1 <i>ms</i> resource block pairs . . . . .	37
3.6.1 Illustration of the WRED processing . . . . .	42
4.1.1 Structure of the CASA system . . . . .	45
4.1.2 Illustration of the wideband SRS transmission . . . . .	47
4.2.1 Illustration of the grant processing . . . . .	52

5.1.1 Structure of the ICAS system . . . . .	54
5.1.2 The transmission of the narrowband SRS in ICAS . . . . .	55
5.2.1 The processing for choosing resource blocks group for UE $i$ . . . . .	57
5.2.2 Illustration of remove of the resource blocks group of UE $i$ . . . . .	59
5.2.3 Illustration of the resource blocks assignment when $Q_i < N_i$ . . . . .	62
5.2.4 Illustration of the grant processing . . . . .	63
5.2.5 Illustration of ICAS grant processing . . . . .	64
6.1.1 Pareto probability density function (PDF) for with $x_m = 1$ . . . . .	67
6.1.2 Pareto cumulative distribution function (CDF) for with $x_m = 1$ . . . . .	68
6.3.1 CASA and ICAS average delay with 0.5 ms TTI . . . . .	71
6.3.2 CASA and ICAS Average Packet Loss Ratio with 0.5 ms TTI . . . . .	73
6.3.3 CASA and ICAS network throughput with 0.5 ms TTI . . . . .	74
6.3.4 CASA and ICAS CoS throughput with 0.5 ms TTI . . . . .	75
6.3.5 CASA and ICAS resource block allocated ratio with 0.5 ms TTI . . . . .	76
6.3.6 CASA and ICAS average packet delay with 1.0 ms TTI . . . . .	77
6.3.7 CASA and ICAS Average Packet Loss Ratio with 1.0 ms TTI . . . . .	78
6.3.8 CASA and ICAS network throughput with 1.0 ms TTI . . . . .	79
6.3.9 CASA and ICAS CoS throughput with 1.0 ms TTI . . . . .	80
6.3.10 CASA and ICAS resource allocated ratio with 1.0 ms TTI . . . . .	81
6.3.11 ICAS Average delay with 0.5 ms and 1.0 ms TTI . . . . .	82
6.3.12 ICAS Average Packet Loss Ratio with 0.5 ms and 1.0 ms TTI . . . . .	83
6.3.13 ICAS network throughput with 0.5 ms and 1.0 ms TTI . . . . .	84
6.3.14 ICAS CoS throughput with 0.5 ms and 1.0 ms TTI . . . . .	85
6.3.15 Example of the resource allocation in case of 0.5 ms TTI and 1.0 ms TTI . . . . .	86
6.3.16 ICAS resource blocks allocated ratio with 0.5 ms and 1.0 ms TTI . . . . .	87

<i>LIST OF FIGURES</i>	8
6.3.17 Comparison of average delay performance of CASA with 0.5 ms and 1.0 ms TTI . . . . .	88
6.3.18 Comparison of average Packet Loss Ratio of CASA with 0.5 ms and 1.0 ms TTI . . . . .	89
6.3.19 Comparison of average throughput of CASA with 0.5 ms and 1.0 ms TTI . . . . .	89
6.3.20 Comparison of average CoS throughput of CASA with 0.5 ms and 1.0 ms TTI . . . . .	90
6.3.21 Comparison of resource blocks allocated ratio of CASA with 0.5 ms and 1.0 ms TTI . . . . .	90

# List of Tables

3.2	Summary of the test LTE system . . . . .	32
3.4	Summary of the Adaptive Modulation and Coding scheme mode	40

# Nomenclature

$\bar{\beta}^m$	mean of received SINR
$T_{cp}$	Length of Cyclic Prefix
$N$	Number of active UEs in the system
$\Gamma(m)$	Gamma function
$m$	Nakagami fading parameter
$R_{RB}$	Information transmission rate of the resource block
$R_{bpsymbol}$	Bit rate of each symbol
$v_j$	CoS $j$ 's weight
$V_j$	CoS $j$ 's credit pool size
$u_i$	UE's weight ( $u_i$ )
$N_i$	UE's credit pool size
$\mu_{ij}$	Weight of UE $i$ in CoS $j$
$N_{total}$	The total number of available resource blocks within a given transmission cycle
$\sum_a u_a$	The sum of all active UEs' weight
$\sum_b v_b$	The sum of CoSs' weight
$Req_{ij}$	The number of Bytes of data requested from UE $i$ for CoS $j$
$n_{reqij}$	The number of resource blocks requested from UE $i$ for CoS $j$
$N_{remainder}$	All remaining credits in the CoS credit pools.
$S_{gi}$	Group size for UE $i$

$M_i$	Amount of data in byte that can be supported within each resource group in the 0.5 ms timeslot for UE $i$
$Q_i$	Total request size for UE $i$

# Chapter 1

## Introduction

With the social and economic progresses, more and more people have started using mobile communications, leading to a tremendous increase in wireless data transmission. According to a research report from ERICSSON [34], in 2008 the increase in data transmission was up by a factor of five, and large operators' network are now carrying data on the order of 5~10 Tbytes/day. This causes an extremely heavy load for the existing networks. Customers also want better services from their network operators. So the operators have to find a new technology to provide new and more advanced services at a reasonable cost. The new technology should also provide existing services in a better and more cost-efficient way. In order to solve these challenges, the Long Term Evolution (LTE) was designed as a project of the Third Generation Partnership Project (3GPP).

From a technical point of view, the LTE provides a downlink peak data rate of at least 100 Mbps, an uplink data rate of at least 50 Mbps, and round-trip times of less than 10 ms. LTE supports a wide range of bandwidth, from 1.4 MHz up to 20 MHz. It also supports both Frequency Division Duplex (FDD)

and Time Division Duplex (TDD). In order to replace the General Packet Radio Service (GPRS) Core Network and provide compatibility with some legacy or non-3GPP systems, the LTE system was designed as a flat IP-based network architecture. The main advantages with LTE are high throughput, low latency, plug and play, FDD and TDD in the same platform, an improved end-user experience and a simple architecture resulting in low operating costs. LTE will also support seamless passing to cell towers with older network technologies such as GSM, WCDMA, and CDMA2000[36]. The next step for LTE evolution is LTE Advanced which is currently being standardized in 3GPP Release 10.

The multiple-access scheme in LTE downlink transmission is Orthogonal Frequency Division Multiple Access (OFDMA), which is based on Orthogonal Frequency Division Multiplexing (OFDM). Based on OFDMA, the entire bandwidth will be divided into a set of narrowband sub-channels (subcarriers). Based on this property, LTE network can support a more flexible frequency selective scheme. During the transmission, the LTE will insert a Cyclic Prefix (CP) for each data symbol to eliminate the Inter-symbol Interference (ISI). Based on this technique, the User Equipment (UE) can easily detect the data by a low-complexity single tap equalizer. Cyclic Prefix is also important for the Multiple-input Multiple-output (MIMO) technique implementation.

Because of the power limitation in UE and the relatively larger peak to average power ratio (PAPR) in OFDMA, the multiple-access for LTE uplink transmission is chosen as Single-carrier Frequency Division Multiple Access (SC-FDMA). Unlike OFDMA, in SC-FDMA the system will choose the appropriate modulation scheme for the entire subcarrier without the single symbol in the subcarrier. This can improve the Radio Frequency (RF) transmission power efficiency. Like OFDMA, the uplink transmission also employs the Cyclic Prefix, which will help base station to detect the received signal easily.

In order to estimate the channel condition, the LTE system employs two types of Sounding Reference Signals (SRS): one is wideband Sounding Reference Signal; the other one is narrowband Sounding Reference Signal. By employing a sufficiently wideband SRS, the entire transmission bandwidth can be covered by a single wideband SRS. Because the narrowband SRS has much more narrow bandwidth than wideband SRS, a sequence of narrowband SRS sent from the UE will jointly cover the entire bandwidth. Based on the strength of the received SRS signal in base station, the Adaptive Modulation and Coding (AMC) scheme can be employed in the LTE to maximize the system throughput by chosen appropriate Modulation and Coding scheme, like BPSK, QPSK and QAM. When the system employs the 64-QAM, the peak data rate for the downlink can achieve 300 Mbps; the peak data rate for uplink transmission is approximately 75 Mbps. This is a sufficient transmission capacity in wireless networks, which allows LTE to support multiple services, such as real-time voice, video, and non-real time applications (email, website, etc).

In order to support multiple services, the LTE system has to face different Quality-of-service (QoS) requirements, such as delay, throughput and packet loss ratio. So, it is very important to design an efficient resource allocation algorithm for the LTE data transmission. We also need to consider long term fair share of resource between the UEs. In this thesis, we propose two new resource allocation algorithms for the LTE uplink transmission. One is called Class-of-service Aware Scheduling Algorithm (CASA); and the other is called Independent Channel Aware Scheduling Algorithm (ICAS). Both of these algorithms consider the Class-of-service and the channel condition.

## 1.1 Review of previous works

Most of the previous works about channel aware scheduling algorithms fo-

cused on the OFDMA, like [9]. But the methods they proposed can give us some insights. Recently, some researchers have conducted research on LTE up-link transmission with SC-FDMA. In [12], the authors presented the Channel-Dependent Scheduling (CDS) methods based on the different subcarrier mapping scheme; Localized FDMA and Interleaved FDMA. Unfortunately, these algorithms did not consider that the allocated resource blocks have to be kept continuously in the frequency domain.

In [28], a channel aware scheduling algorithm was proposed to exploit the bandwidth flexibility. This algorithm employs a resource block-user matrix to distribute the resource blocks. The system started the distributing processing from the resource block-user pair with the maximum value in the matrix. The system expanded bandwidth for this user until another user has a higher metric value on the adjacent resource block or reached physical constraints (the spectrum edge or the resource block had been assigned to another user) on both sides. This algorithm is simple and flexible. However, it did not consider the fair share between different UEs. It also did not consider the QoS.

In [10], the authors proposed three channel-aware scheduling algorithms: First Maximum Expansion (FME), the Recursive Maximum Expansion (RME) and the minimum area-difference to the envelope ( $MAD^E$ ). Similar to [28], all of these three algorithms are implemented based on a two dimensional user-resource block ( $user-RB$ ) matrix. The value of the matrix element is dependent on the channel condition. In FME, the algorithm found combination  $user-RB$  with the highest metric value (i.e.  $user_0-RB_0$ ) and assign  $RB_0$  to  $user_0$ . Then the system expanded user's bandwidth on one side until reached physical constraints (the bandwidth edge or the resource block had been assigned to another user). The RME is similar to FME, but it employs a recursive search for the largest value. Unlike FME, The expansion for each user in RME is on both

sides. The expansion will not stop until reached physical constraints or another user has a higher metric value on the adjacent resource block. The  $MAD^E$  is an improvement for RME. This algorithm employs another envelope-metric, which consists of all the maximum values in all RBs' column. After the normal RME processing finished, the algorithm will reassemble the user-RB combination based on the value in envelope-metric to achieve the minimum difference between user-RB matrix and envelope-metric. All these three algorithms are simple and flexible. In  $MAD^E$ , authors also considered the fair share between different users to some extent. But they still have some limitations: If a user always has a better channel condition than other users, it will receive much more resource blocks in the whole simulation time. The long term fair share between different users can not be achieved. Additionally, all these three did not consider the Class-of-service.

In [24], the authors proposed a method that allows the TDD-LTE uplink to support the VoIP service. This paper introduced the operation of the TDD-LTE and shown the simulation method for the VoIP and analyzed some restrictive factors' effect on the VoIP capacity. Based on this paper and G.711 standard, we employed VoIP transmission in our simulation to represent the first class real-time voice data. We simulated our designed system with the FDD operation.

In [26] and [27], new kinds of the channel estimation methods were introduced. They focus on the different area of the LTE network. This is an important issue that we considered in the resource allocation algorithm designing processing.

In [30], Weighted Random Early Detection (WRED) buffer management mechanism is introduced and explained. The WRED measures the average queue size and drops packets based on probability. If most of the buffer space is empty, all incoming packets will be accepted. However, when the average queue

size exceeds a preconfigured minimum threshold, the WRED will drop or accept the arriving packet based on a certain probability, where the probability is a function of the average queue size. As the queue size grows, the probability for dropping an incoming packet grows. When the buffer is full, the probability has reached 1 and all incoming packets are dropped. WRED will keep the average queue size low while allowing occasional bursts of packets in the queue. We adopted WRED to manage the Class-of-service (CoS) queues in our proposed system.

## 1.2 Thesis Contributions

In order to provide quality of service (QoS) for the uplink transmission in the LTE system, we designed two resource-block allocation algorithms. Various classes of traffic are transmitted based on their QoS requirements, transmission history, and their channel conditions. The first algorithm is called: Class-of-service Aware Scheduling Algorithm (CASA), which employs the wideband SRS technique together with a credit pooling technique to allocate resource blocks among customers' classes of service in a fair manner. The second algorithm is called: Independent Channel Aware Scheduling Algorithm (ICAS), which employs the narrowband SRS technique. ICAS algorithm divides resource blocks into a number of groups, and employs a priority based scheduling to distribute these groups among customers. Both of these algorithms support Class-of-service (CoS) capabilities. The weighted round robin arbitration mechanism is employed in both algorithms to determine the priority of users by their long term transmission records. In particular, we test both algorithms with three CoSs which are based on the well-known Differentiated Services IP (DiffServ). These are called: Expedited Forwarding (EF), Assured Forwarding (AF) and Best Effort (BE). We also consider the channel condition in our algorithms based

on the Adaptive Modulation and Coding (AMC) scheme.

### 1.3 Thesis Outline

This thesis will be organized as follows. In Chapter 2, we will introduce the background knowledge about the LTE system and SC-FDMA scheme. Chapter 3 shows the simulated system architecture, which includes: system model, transmission cycle, architecture of the scheduling algorithm and the uplink channel model. Chapter 4 and Chapter 5 show the Class-of-service Aware Scheduling Algorithm and Independent Channel Aware Scheduling Algorithm, respectively. In Chapter 6, we will introduce the simulation setup and analyze the simulation results. In Chapter 7, we will provide concluding and the possible future work.

## Chapter 2

# Physical Layer for LTE

## Uplink

### 2.1 Introduction

In this section, we summarized some of the desirable attributes of the LTE uplink. These attributes are shown below in no particular order:

1. In order to minimize intra-cell interference and maximize capacity, the uplink transmission for each User Equipment (UE) should be kept orthogonal.
2. Adaptive Modulation and Coding scheme will be employed to choose a wide range of data rates based on the received Signal-to-Interference plus Noise Ratio (SINR) of sounding reference signal.
3. In order to achieve higher power efficiency, transmission waveform should have low Peak-to-Average Power Ratio (PAPR).

4. It should fully exploit the system bandwidth (up to 20 MHz), even when the data is being transmitted at low data rates.
5. Frequency-selective scheduling will be employed.
6. In order to exploit spatial diversity and enhance uplink capacity, advanced multiple-antenna techniques should be employed in the system.

Single-Carrier Frequency Division Multiple Access (SC-FDMA) has been chosen as the multiple-access scheme for the LTE uplink transmission to achieve the principles shown above.

## 2.2 Single Carrier Frequency Division Multiple Access Principles

Based on the SC-FDMA scheme, the entire bandwidth is divided into multiple parallel subcarriers, which is similar to OFDMA. By using a Cyclic Prefix (CP) or guard time, different frequency-selective channels will be kept orthogonal. Using a Cyclic Prefix can also eliminate the impact of the Inter-Symbol Interference (ISI) between SC-FDMA information blocks. It transforms the linear convolution of the multipath-channel into a circular convolution. The receiver can simply use a complex gain factor to scale each subcarrier in order to equalize the channel.

However, unlike OFDMA, where the data symbols directly modulate each subcarrier independently (such that the amplitude of each subcarrier at a given time interval is set by the constellation points of the digital modulation scheme), in SC-FDMA the signal modulated onto a given subcarrier is a linear combi-

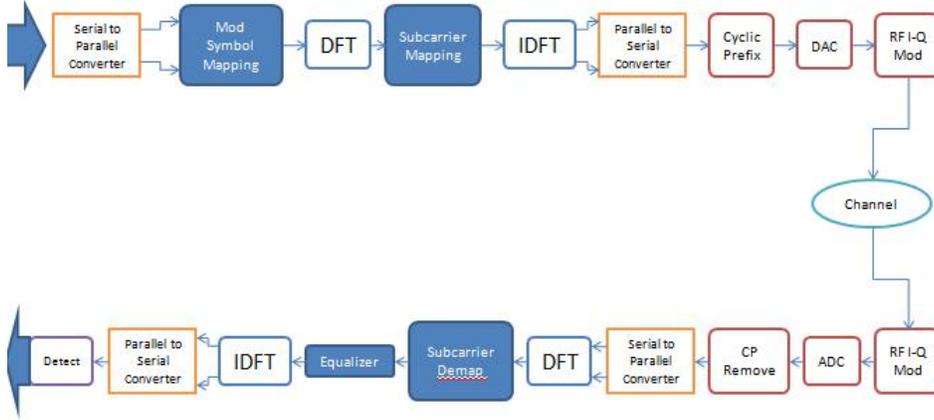


Figure 2.2.1: SC-FDMA transmitter and Receiver

nation of all the data symbols transmitted at the same time interval. This is achieved by employing the Discrete Fourier Transform (DFT) element in the transmitter [1]. Thus in each symbol period, all the transmitted subcarriers of an SC-FDMA signal carry a component of each modulated data symbol. This gives SC-FDMA its crucial single-carrier property, which results in the PAPR being significantly lower than pure multicarrier transmission schemes such as OFDMA. The receiver and transmitter structure for SC-FDMA are shown in Figure 2.2.1.

## 2.3 Uplink Physical Resource

In the LTE system, the uplink transmission scheme is based on the SC-FDMA. The wireless resource is specified in both time-domain and frequency-domain. According to the proposed LTE standard [5], in the time-domain, the LTE uplink consists of 10 ms radio frame, divided into 10 equal size sub-frames with 1.0 ms length. Figure 2.3.1 shows the general structure of the LTE uplink.

To be more specific, each 1 ms sub-frame consists of two equal size 0.5 ms

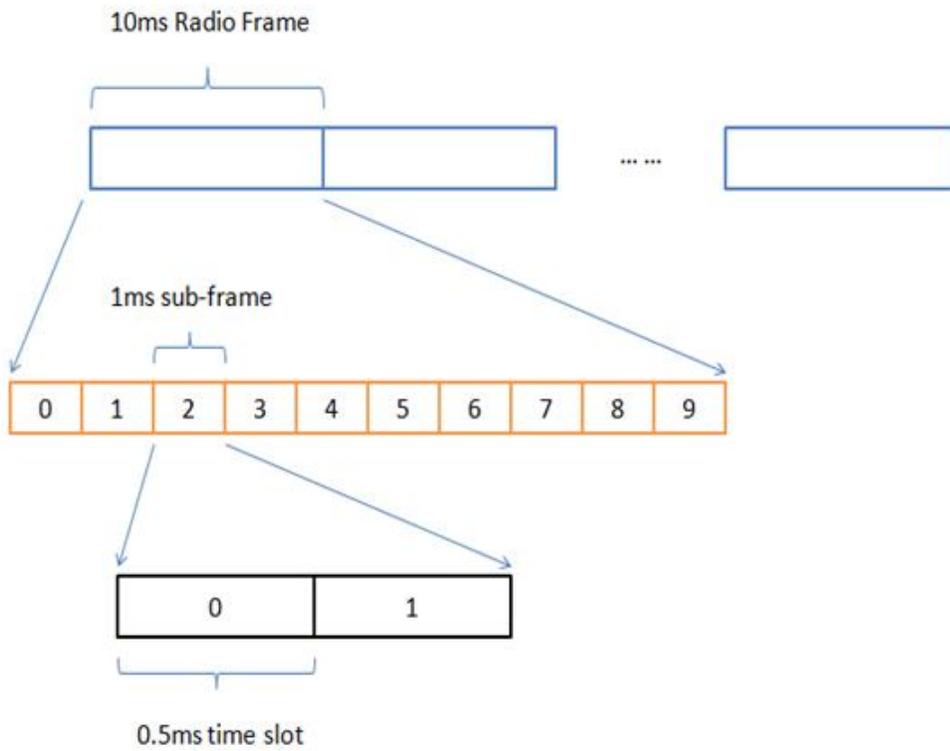


Figure 2.3.1: General LTE uplink time-domain structure

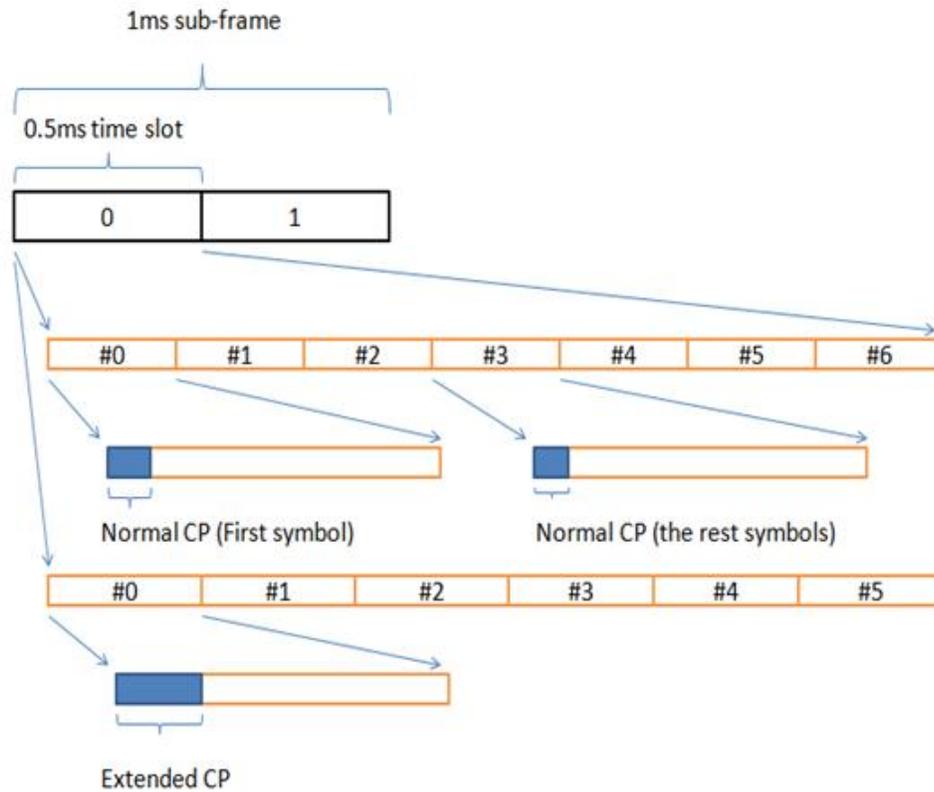


Figure 2.3.2: Details of the LTE uplink structure

time slots, which is shown in Figure 2.3.2. In order to eliminate the Inter-Symbol Interference (ISI), the LTE also employs the Cyclic Prefix (CP). There are two kinds of CP. One is called normal CP, whose length is  $T_{cp} \approx 5.1\mu s$  in the first symbol and  $T_{cp} \approx 4.7\mu s$  in all other symbols. The other is called extended CP, whose length is  $T_{cp} \approx 16.7\mu s$ .

In frequency domain, the LTE system considers the radio resource as Resource Blocks (RBs). The LTE uplink consists of a number of resource blocks, ranging from 6 to 110 resource blocks. This corresponds to the entire uplink bandwidth from 1 MHz to 20 MHz. As shown in Figure 2.3.3, each resource block consists of 12 equal size subcarriers in frequency domain. Each subcarrier

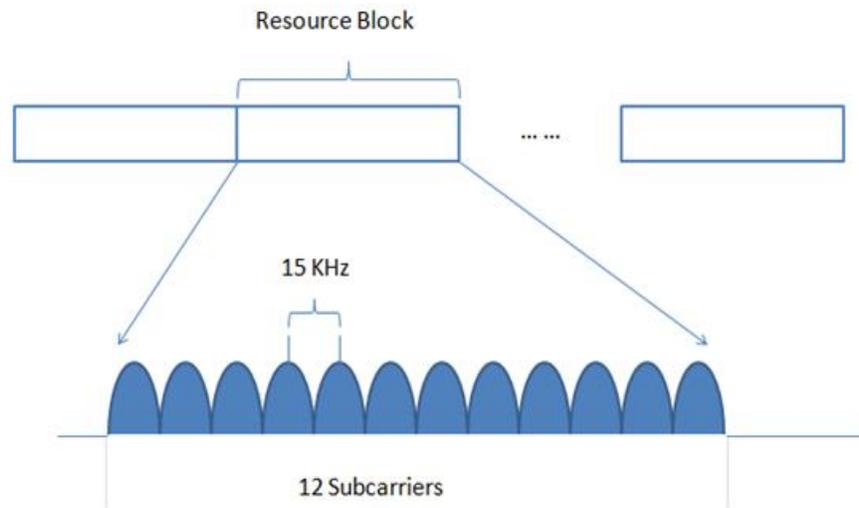


Figure 2.3.3: Structure of Resource Block in frequency domain

has 15 KHz space.

Figure 2.3.4 shows the structure of a resource block in both time and frequency domain. Each resource block can be considered as a two dimensional radio resource unit. It consists of 12 subcarriers during 0.5 ms slot. Each subcarrier contains seven or six resource elements, according to normal CP and extended CP. So each resource block square is consisted of 84 or 72 resource elements.

LTE uplink transmission supports both Frequency Division Duplex (FDD) and Time Division Duplex (TDD) as shown in Figure 2.3.5. In FDD technique, the LTE employs two separated transmission band, one for uplink transmission and the other for downlink transmission. Both uplink and downlink transmissions can be implemented simultaneously. TDD is a kind of Time Division Multiplex (TDM) technique. Both uplink and downlink transmissions share a transmission band. In order to eliminate the transmission collision happened during the uplink and downlink switching, the system employs a special frame.

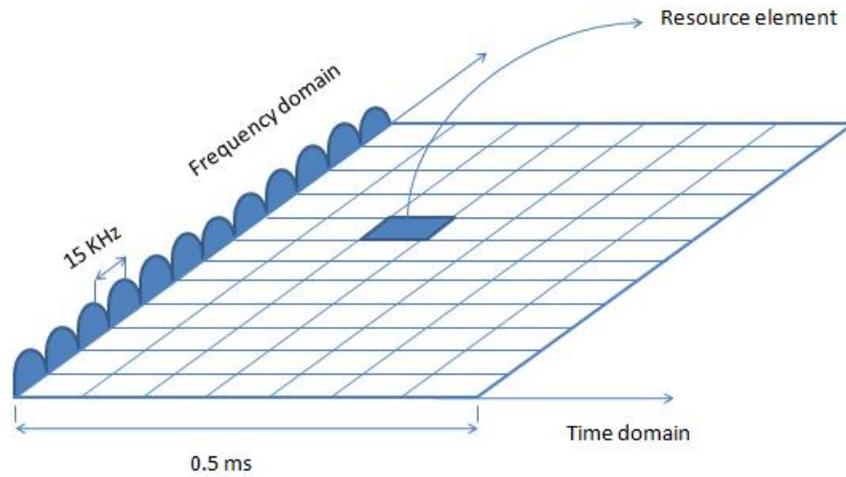


Figure 2.3.4: Resource block structure in both frequency and time domain

As shown in Figure 2.3.5.(b), the special frame contains the Uplink Part, Guard Period, and the Downlink Period. Both Uplink Part and Downlink Part can support the normal transmission data.

## 2.4 Uplink Reference Signals

In order to help the system to perform demodulation and channel condition estimation, the LTE uplink transmission employs two kinds of reference signals. One is Demodulation Reference Signal (DRS) and the other is Sounding Reference Signal (SRS). Demodulation Reference Signal is used to help the system to determine the demodulation scheme, so it will be transmitted with the uplink data on Physical Uplink Shared Channel (PUSCH). If there is no data transmitted on the uplink, the DRS will be transmitted on the Physical Uplink Control Channel (PUCCH). Unlike DRS, the Sounding Reference Signal will be used to help the base station to estimate the channel condition and to enable the frequency selective scheme. The Sounding Reference Signal will cover the whole

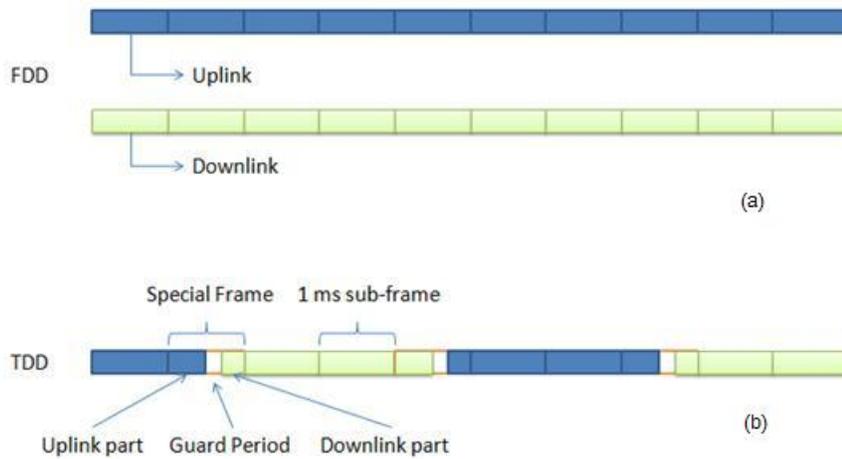


Figure 2.3.5: Illustrations for both FDD and TDD

bandwidth, so it will not be associated with uplink data. Both the Demodulation Reference Signal and Sounding Reference Signal are time multiplexed with the uplink data. In our paper, we focus on the scheduling algorithms, so the Sounding Reference Signal is more important than the Demodulation Reference Signal.

### 2.4.1 Sounding Reference Signal

As we mentioned above, the Sounding Reference Signal will be used to estimate the transmission channel condition. The channel condition is an important factor that is considered in our algorithm. For this reason, the detailed explanation of Sounding Reference Signal is given below.

The Sounding Reference Signal will be transmitted from UEs periodically. The transmission interval could be chosen as 2, 5, 10, 20, 40, 80, 160 and 320 ms. When the Sounding Reference Signal is being transmitted, it will always

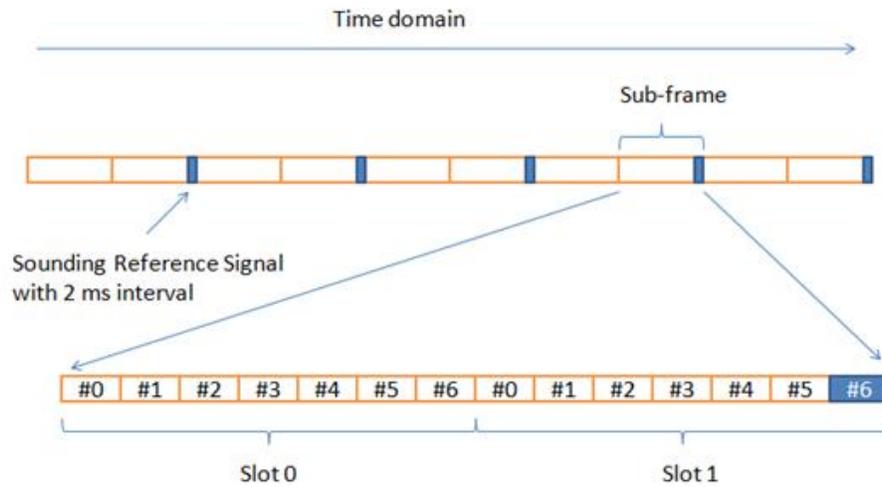


Figure 2.4.1: Illustration of the SRS position with 2 *ms* interval

occupy the last symbol of the sub-frame, which is shown in Figure 2.4.1.

Since the base station has to estimate the channel condition for all the UEs, the Sounding Reference Signal should be overlapped very well. In order to avoid collision between Sounding Reference Signal and the PUSCH transmission from other UEs, all UEs will be aware of which sub-frame will be used to transmit the Sounding Reference Signal. On the other hand, if the Sounding Reference Signal from different UEs will cover the same bandwidth, the system can transmit them simultaneously based on the Phase Rotation technique. Different Sounding Reference Signals can be orthogonal with each other by assigning different phase rotations. When the Sounding Reference signal is mapped to subcarriers, the spectrum can be considered as a ‘comb’. By assigning different offsets to different Sounding Reference Signal, the phase rotation can be achieved. Multiple Sounding Reference Signals can be transmitted in parallel, which is shown in Figure 2.4.2.

In order to cover the interest bandwidth in frequency domain, the LTE

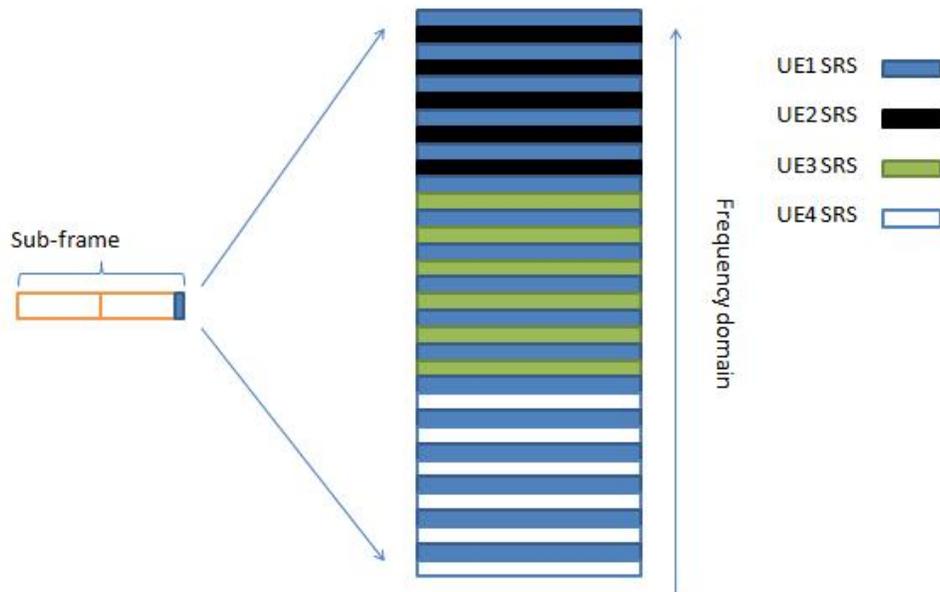


Figure 2.4.2: Illustration of the SRS parallel transmission

system employs two types of Sounding Reference Signals: one is wideband SRS; the other one is narrowband SRS. Each type of Sounding Reference Signal works as follows [5]:

1. By employing a sufficiently wideband SRS, a single wideband SRS can cover the entire transmission spectrum.
2. The narrowband SRS has much more narrow bandwidth than wideband SRS. In order to cover the transmission bandwidth, the LTE system supports frequency hopping of narrowband SRS. As shown in Figure 2.4.3.(b), in order to cover the entire spectrum, a sequence of narrowband SRS signals will be sent periodically. Each narrowband SRS signal will cover a part of spectrum. The base station will use the received SINR of each narrowband SRS to estimate the channel condition for its covered spectrum.

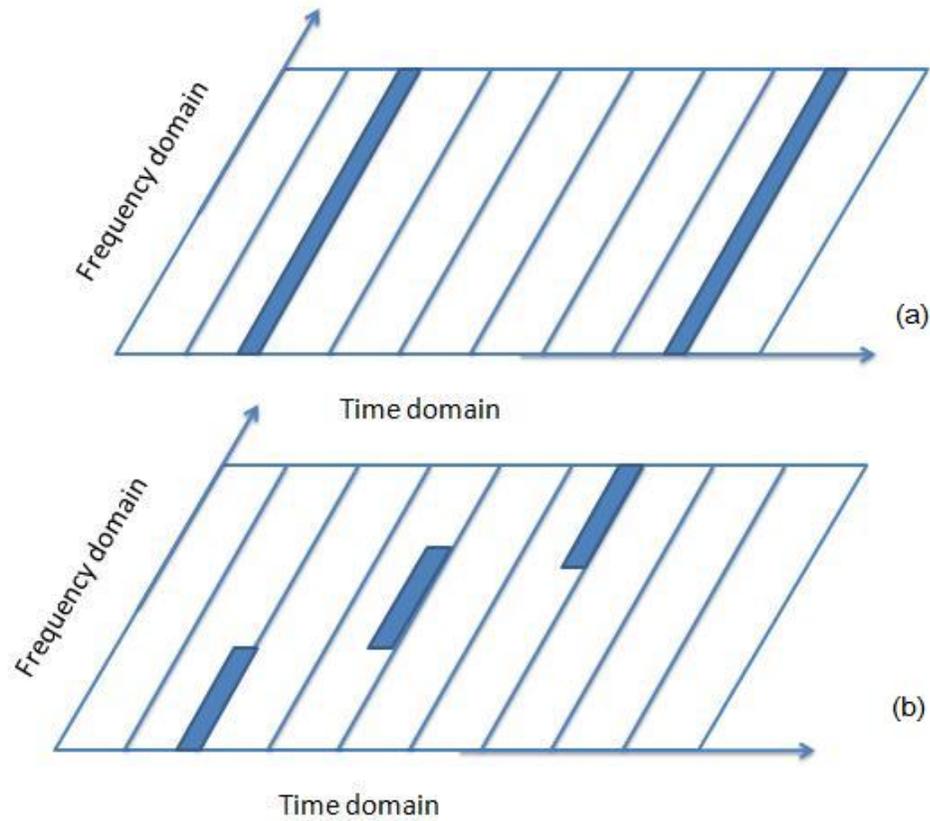


Figure 2.4.3: Operations of Wideband SRS (a) and Narrowband SRS (b)

As shown in Figure 2.4.3, we can see that the main benefit of the wideband SRS is that the entire bandwidth can be covered by a single wideband SRS. As the Sounding Reference Signal is time division multiplexed with the transmitted data, the wideband SRS has a higher bandwidth efficiency than the narrowband SRS. However, if the system is running in a bad channel condition (i.e. high uplink path loss), the received wideband SRS may have a relatively lower power strength. And this may reduce the quality of the channel condition estimation, even though there are some high quality sub-channel is existing. So in this case, the narrowband SRS will be more efficient in the channel condition

estimation processing. The narrowband SRS will give the base station more specific information about the channel condition.

## 2.5 Summary

In this chapter, we introduced some important properties of the SC-FDMA transmission scheme used for the LTE uplink. We introduce physical resource in LTE uplink, with which we will use to transmit data. We also introduced details about the sounding reference signal technique, which is one of the most important techniques in our design. Based on the back ground knowledge mentioned above, we will introduce the specific information about the LTE system used in our simulation in Chapter 3.

## Chapter 3

# System Architecture

### 3.1 Basic Model of LTE Wireless System

In this thesis, we consider the uplink of a single-cell operating in Frequency Division Duplex (FDD). The cell contains one base station communicating simultaneously with  $N$  mobile user equipments (UEs). For convenience, we only focus on uplink transmission. We assume this LTE communication system has a 20 MHz bandwidth. In the frequency domain, the channel is grouped into 100 available RBs. Each RB consists of 12 subcarriers each with a 0.5 ms slot. Each subcarrier has a 15 KHz space. The multiple-access is achieved by distributing RBs to UEs. In the frequency domain, the RBs allocated to each UE should be continuous.

The modulation and coding mode for any UE is specified adaptively according to the UE's channel condition. The representative modulation schemes, which include BPSK, QPSK, 16QAM and 64QAM, are adopted in our discussion according to the instantaneous Signal-to-Interference plus Noise Ratio (SINR) of the entire channel or the narrowband sub-channel, based on the wideband

Parameter	Value
System	LTE
Uplink Channel Bandwidth	20 MHz
Subcarrier spacing	15 KHz
RB size	12 subcarriers
Total number of RBs in frequency domain	100
Frame period	10 ms
Sub-frame period	1 ms
Timeslot period	0.5 ms
Transmission Time Interval (TTI)	0.5 ms / 1.0 ms

Table 3.2: Summary of the test LTE system

SRS or narrowband SRS technique, respectively. According to proposed LTE standard, the resource is considered as resource block (RB) when Transmission Time Interval (TTI) is set to 0.5 ms, and a pair of adjacent resource blocks when TTI is set to 1 ms. The parameters of this LTE communication system, which we consider in this thesis, are summarized in Table 3.1.

## 3.2 Transmission Cycle

A typical transmission cycle begins with the execution of the proposed algorithm in the base station. After the scheduling process completed, the base station sends the grant message to the users with the downlink data. After the users received the grant message, users will use the assigned resource to transmit data at the assigned time interval in parallel during next 2 ms time interval. In our simulation, the transmission cycle consists of scheduling time, grant transmission time and the data transmission time, as shown in the figure

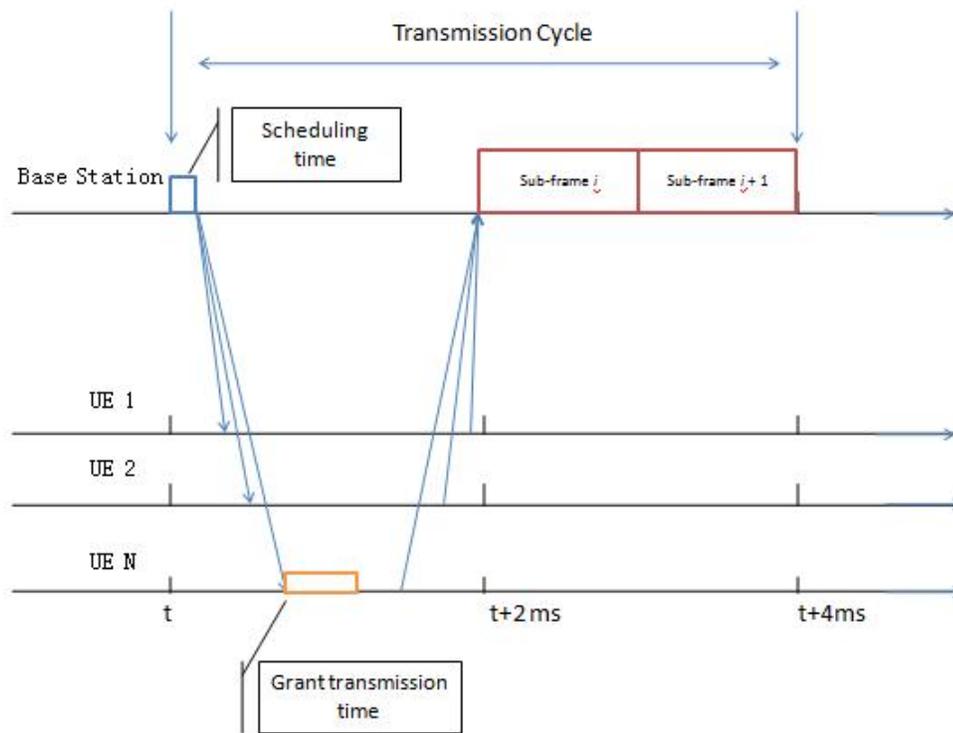


Figure 3.2.1: Illustration of transmission cycle

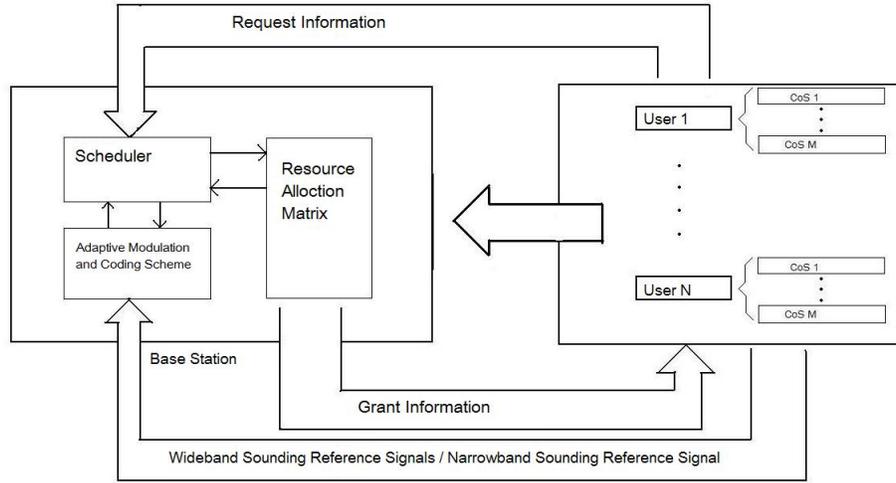


Figure 3.3.1: The general architecture of the scheduling algorithm

3.2.1.

### 3.3 Architecture of the scheduling algorithm

In this thesis, we focus on the uplink transmission for simplification. Figure 3.3.1 shows the general architecture of the scheduling algorithm for the uplink data transmission. We assume that the user can generate multiple classes of data, such as real time video, real time voice and non-real time data. UE will employ multiple CoS queue buffer for different CoS data. The user's generated data will be classified according to their different classes and stored in different CoS queues. Based on the CoS queue size, the user will send the request information to the base station. The request information contains the user I.D, CoS I.D and the waiting data size in *byte*. Also, during the transmission cycle, the user will send the Sounding Reference Signal to the base station.

After the transmission cycle is finished, the scheduler will start the scheduling process. Based on the channel estimates obtained at the base station, the

Adaptive Modulation and Coding (AMC) selector determines the modulation-coding pair, whose index will be sent back to user with the grant information. The scheduler will use this value to calculate the transmission rates for users, which will be used in the scheduling algorithm. This processing will be shown in Chapter 3.5.

The base station also employs a resource-allocation matrix, as shown in both Figure 3.3.2 and Figure 3.3.3. The Transmission Time Interval (TTI) in LTE has two length, one is 0.5 ms; the other is 1.0 ms. If the system employs the 0.5 ms TTI, the minimum unit of the available resource in the system is RB. The two dimensional resource-allocation matrix will consist of four 0.5 ms column in time domain and 100 RB in frequency domain, which is shown in Figure 3.3.2. The total amount of resource in this resource-allocation matrix is 400 RBs. If the system employs the 1.0 ms TTI, the minimum unit of the available resource in the system is RB-pair. The two dimensional resource-allocation matrix will consist of two 1.0 ms column in time domain and 100 RB-pairs in frequency domain, which is shown in Figure 3.3.3. The total amount of resource in this resource-allocation matrix is 200 RB-pairs.

In our proposed Class-of-service Aware Scheduling Algorithm (CASA), the resource-allocation matrix will work as a re-arrange matrix. After the scheduler allocated the resource to users based on the Class-of-service requirement, the matrix will add the resource block I.D information to the grant information from the scheduler, which only has the grant size information.

In the Independent Channel Aware Scheduling (ICAS) algorithm, the resource-allocation matrix works with the scheduler together. During the algorithm working time, the matrix will share the resource information with the scheduler. And after the scheduling processing finished, the matrix will also stop.

After the distribution processing is finished, the base station will send the

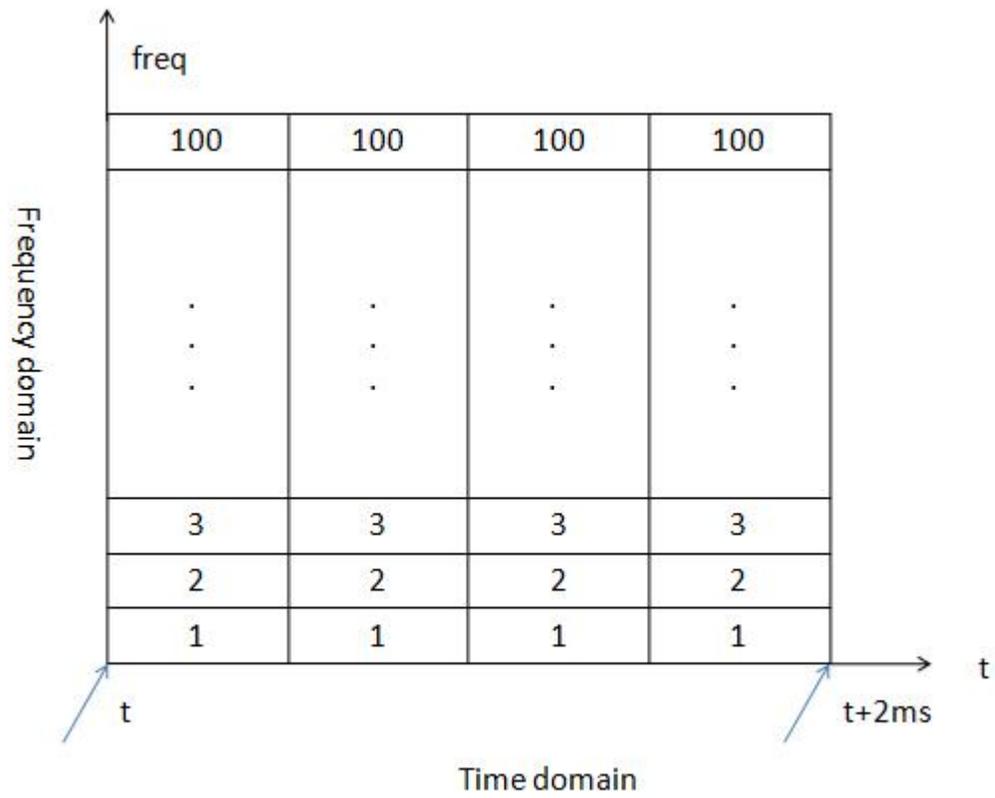
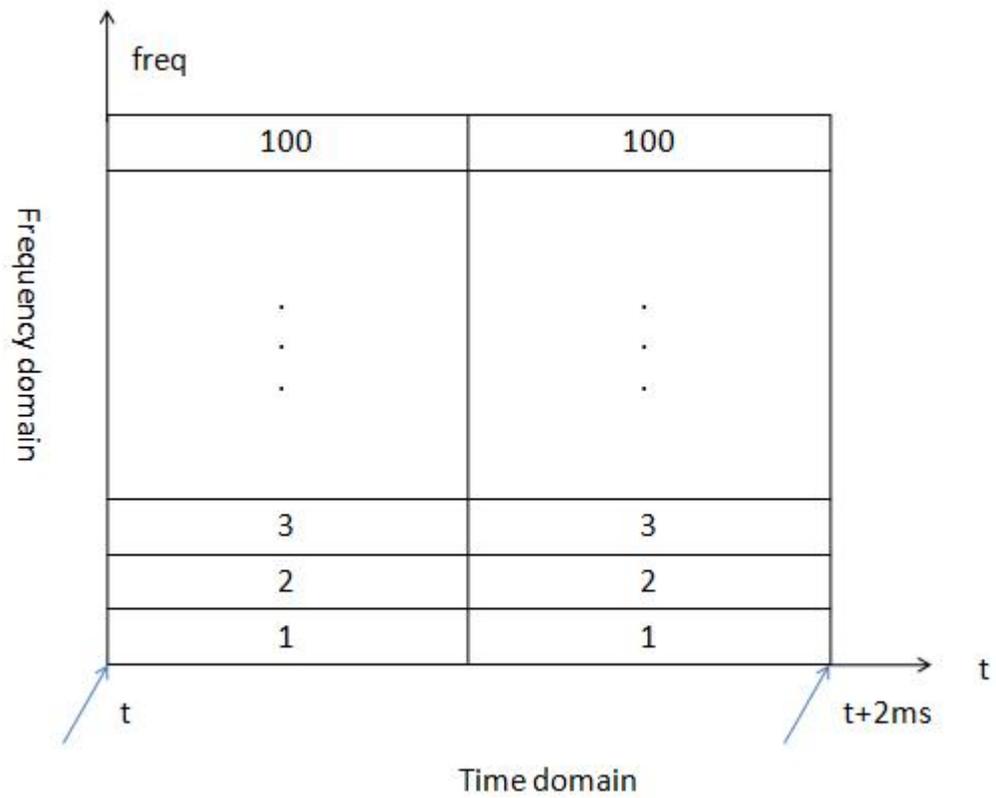


Figure 3.3.2: The resource-allocation matrix for 0.5 ms resource blocks

Figure 3.3.3: The resource-allocation matrix for 1 *ms* resource block pairs

grant information to the users. The grant message includes the information about the number of resource blocks for each CoS, the resource block I.D. and the user I.D.

### 3.4 Uplink Channel Model

According to [9], we assume that the uplink channel is a general *Nakagami- $m$  model* in our simulation. The channel state stays essentially invariant during the information transmission time. In order to estimate the channel condition, the received SINR  $\beta$  is thus a random variable with a Gamma probability density function [32], which is given by:

$$p(\beta) = \frac{m^m \beta^{m-1}}{\bar{\beta}^m \Gamma(m)} \exp\left(-\frac{m\beta}{\bar{\beta}}\right), \quad \beta \geq 0 \quad (3.4.1)$$

$$\Gamma(m) = \int_0^{\infty} t^{m-1} e^{-t} dt, \quad m \geq \frac{1}{2} \quad (3.4.2)$$

where  $\bar{\beta}^m$  is the mean of received SINR,  $\Gamma(m)$  is the Gamma function, and  $m$  is the Nakagami fading parameter. According to [35], the channel model is suitable for flat-fading channels as well as frequency-selective fading channels cooperated with orthogonal frequency-division multiplexing (OFDM) systems, when  $m = 1$ .

When  $m = 1$ , the SINR of the received SRS signal  $\beta$ , which is a random variable with a probability density function given by:

$$p(\beta) = \frac{1}{\bar{\beta}} \exp\left(-\frac{\beta}{\bar{\beta}}\right) = \frac{1}{\bar{\beta}} \exp\left(-\frac{\beta}{\bar{\beta}}\right), \quad \beta \geq 0 \quad (3.4.3)$$

where  $\bar{\beta}$  is the mean of SINR and is assumed to be a function of distance in our

simulations, as shown below:

$$\bar{\beta} = \begin{cases} 19 & 0 \text{ Km} < D \leq 2 \text{ Km} \\ 15 & 2 \text{ Km} < D \leq 10 \text{ Km} \\ 7 & 10 \text{ Km} < D \leq 20 \text{ Km} \\ 4 & \textit{else} \end{cases} \quad (3.4.4)$$

### 3.5 Parameters of AMC mode

Adaptive Modulation and Coding (AMC) scheme is used to adaptively adjust the modulation-coding pair for each user or each resource block according to the wideband SRS or narrowband SRS technique respectively. With AMC scheme, the power of the transmitted signal is kept constant during the transmission cycle. The modulation and coding format is adjusted according to the strength of the received SRS signals, which is represented by Signal-to-Interference plus Noise Ratio (SINR). Based on the AMC, the base station determines the transmission rate for users, which will help the system to achieve a good balance between transmission capacity and the transmission accuracy. Consequently, the transmission rate of each resource block changes periodically according to its AMC mode, which is shown in Table 3.2.

For example, if the received SINR of the SRS is in the range of 7 ~ 11 dB, then the base station will choose mode 3 for this user. In this case, we know that one symbol represents 2 bits with QPSK modulation scheme. With the coding rate 3/4, the information bit rate on each symbol is thus:

AMC	SINR( <i>dB</i> )	Modulation	Coding	Bit/symbol	$R_{RB}$ (Kbps)
1	1.5~4	BPSK	1/2	0.5	84
2	4~7	QPSK	1/2	1.0	164
3	7~11	QPSK	3/4	1.5	252
4	11.5~13.5	16QAM	9/16	2.25	378
5	13.5~18.5	16QAM	3/4	3.0	504
6	>18.5	64QAM	3/4	4.5	756

Table 3.4: Summary of the Adaptive Modulation and Coding scheme mode

$$R_{bpsymbol} = 2 \text{ bit/symbol} \times \frac{3}{4} = 1.5 \text{ bit/symbol} \quad (3.5.1)$$

Based on the structure of the LTE uplink resource block we mentioned in section 3.4, the information transmission rate of the resource block is given by:

$$\begin{aligned} R_{RB} &= R_{bpsymbol} \times 7 \text{ symbol/subc} \times 12 \text{ subc/RB} \times \frac{1 \text{ RB}}{0.5 \text{ ms}} \\ &= 1.5 \text{ bit/symbol} \times 7 \text{ symbol/subc} \times 12 \text{ subc/RB} \times \frac{1 \text{ RB}}{0.5 \text{ ms}} \quad (3.5.2) \\ &= 252 \text{ Kbps} \end{aligned}$$

Similarly, we can obtain the information transmission rate for the resource block with other AMC modes.

### 3.6 Buffer management in UE

In our design, every UE will generate multiple classes of data based on the Class-of-Service (CoS). Before the data can be transmitted, it has to be stored in the UE's buffer. So we assume that each UE will separate the buffer into a number of queues, each queue serving a specific CoS. The CoSs may be used

to deliver a type of data, such as real-time voice, video and non-real-time data traffic. They can also be mapped to standardized classes defined in DiffServ. DiffServ is a computer networking architecture that specifies a simple, scalable and coarse-grained mechanism for classifying, managing network traffic and providing Quality of Service (QoS) guarantees on modern IP networks. In our thesis, CoS1 will be mapped to Expedited Forwarding (EF), which provides for low loss and delay-sensitive services; CoS2 will be mapped to Assured Forwarding (AF), which provides for low loss services; and CoS3 will be mapped to Best Effort (BE), which does not require any commitment from the network. All of the queues in each UE share a common buffer space. The well-known *Weighted Random Early Detection* (WRED) buffer management mechanism [30] is implemented in every UE in order to monitor these queues and penalize them if they take up more than their fair share of the buffer space.

WRED operates on the average queue size  $S_\alpha$  for every CoS  $n$  (or queue  $n$ ), where  $\alpha$  is from 1 to  $n$ . On the arrival of a CoS  $n$  packet at the UE buffer, the WRED computes a new value for  $S_\alpha$  using an exponential weighted moving average:

$$S_\alpha = \eta \times S_i + (1 - \eta) \times S'_\alpha \quad (3.6.1)$$

, where  $\eta$  is a weight parameter, and  $S_i$  is the current queue size for all CoSs 1 to  $n$ . The WRED will then decide to drop the newly arrived packet based on the outcome of the following tests:

1. If  $S_\alpha$  is less than a predetermined minimum queue threshold ( $S_{min}$ ), the packet will not be dropped;
2. Else, if  $S_\alpha$  exceeds a predetermined maximum queue threshold ( $S_{max}$ ),

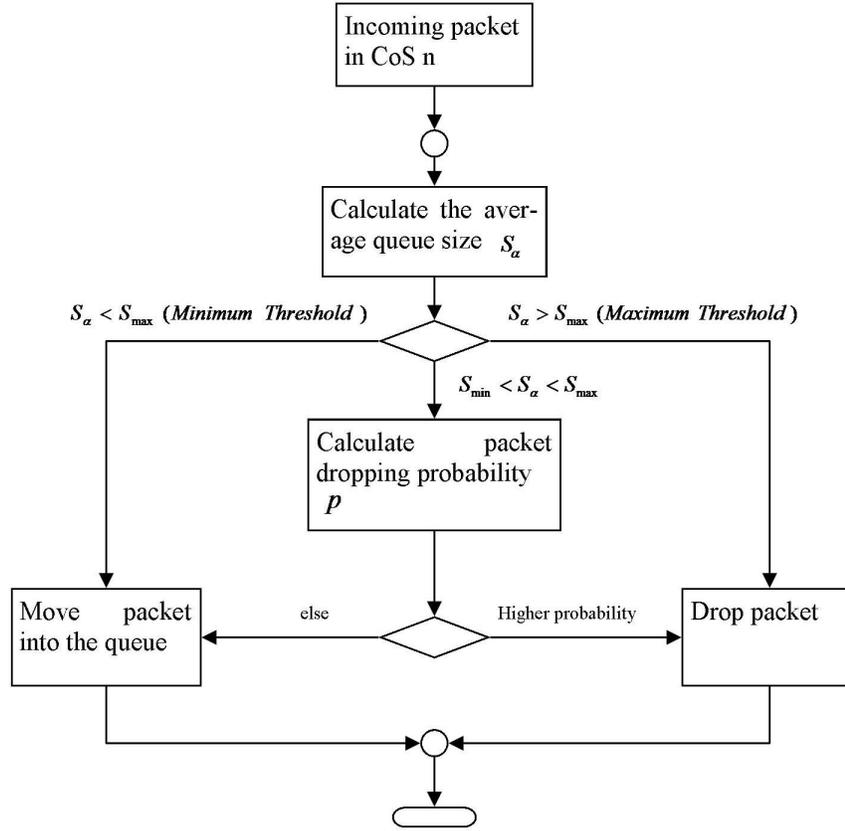


Figure 3.6.1: Illustration of the WRED processing

the packet will be dropped;

3. Else, the packet will be dropped with the probability ( $p$ )

$$p = P_d \times (S_\alpha - S_{min}) / (S_{max} - S_{min}) \quad (3.6.2)$$

, where  $P_d$  is a maximum drop probability [30]. This processing is illustrated by Figure 3.6.1

### 3.7 Summary

In this chapter, we introduced some details of our design, such as system model, transmission cycle, architecture of the scheduling algorithm, uplink channel model and so on. In section 3.5, we introduced AMC scheme. Using the AMC scheme, the base station can choose the transmission rates for users based on their channel condition, which is an important factor considered in our proposed algorithm. In section 3.6, WRED buffer management buffer management mechanism is introduced. WRED keep the average queue size low while allowing occasional bursts of packets in the queue. It helps our proposed algorithms provide a more fair service for all the incoming packets. According the information we mentioned above, we will introduce CASA and ICAS algorithms in coming chapters.

## Chapter 4

# Class-of-service Aware Scheduling Algorithm

In this chapter, we will introduce the Class-of-service Aware Scheduling Algorithm (CASA). This algorithm employs the *wideband SRS* technique together with a *credit pooling* technique to distribute resource among customers' classes of service in a fair manner.

### 4.1 System Structure

Figure 4.1.1 shows the structure of the CASA implemented in the base station. The base station maintains credit pools for both CoSs and UEs. The credit pool for a given CoS is used to enforce a long-term average rate of CoS traffic transmitted from all users, while still permitting short-term bursts above the allowed amount of resources. The credit pool for CoS  $j$  has a set of two programmable parameters: weight ( $v_j$ ) and the number of resource blocks ( $V_j$ ), which will fill into the credit pool before scheduler processing. The base station

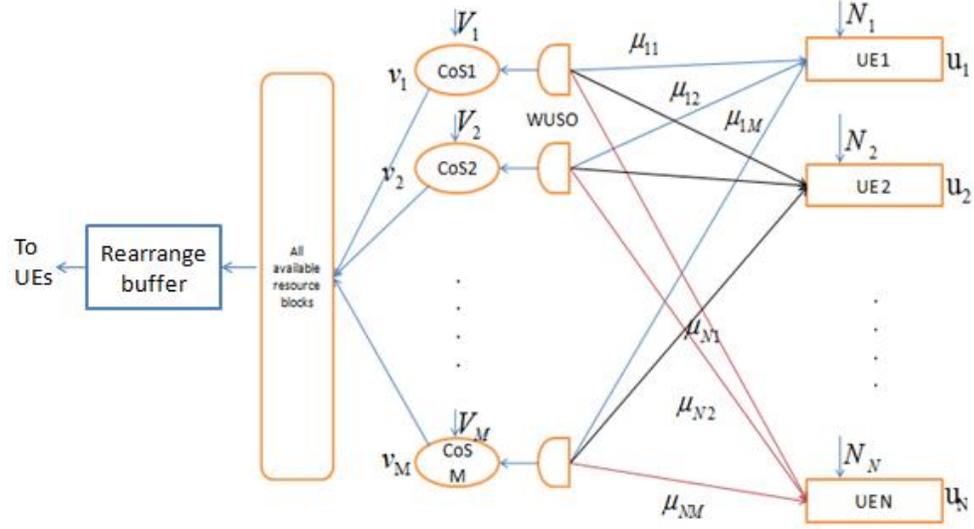


Figure 4.1.1: Structure of the CASA system

also maintains  $N$  credit pools for UEs, which are used to control the usage of the resource blocks by UEs.  $N$  is the number of active UEs in the system. The credit pool for UE  $i$  has two parameters: weight ( $u_i$ ) and the number of resource blocks ( $N_i$ ), which will fill into the credit pool before scheduler processing. In our design, the value of UE  $i$ 's weight  $u_i$  is configured by the user's service level agreement (SLA) [37]. In our thesis, the scheduling interval of CASA is set to be 2 ms.

At the beginning of every cycle, all the CoSs' and UEs' credit pools are initialized with their pre-configured  $V_j$  and  $N_i$  for the next cycle, respectively. A request from an active UE  $i$  for a certain CoS  $j$  is granted if there are enough resource blocks in both the UE and CoS credit pool. A fair distribution of CoS  $j$  credits among UEs is established through the assignment of the weight  $\mu_{ij}$ . In this algorithm, we use Weighted UE scheduling Order (WUSO) arbitration mechanism to enforce these weights [6].

There are  $M$  WUSO arbiters in the system, one arbiter before each CoS. Each WUSO arbiter  $j$  ensures that the contending UEs for CoS  $j$  will receive a fair share of the CoS credit pool, based on their weights. The arbiter uses the number of CoS  $j$  bytes scheduled so far for every UE  $i$  ( $a_{ij}$ ) and the total number of CoS  $j$  bytes scheduled so far ( $A_j$ ) in order to obtain a relative priority  $p_{ij}$  with which new requests from UEs will be processed. Ideally, the ratio of  $a_{ij}$  to  $A_j$  should be  $\mu_{ij}$ . The goal of the WUSO arbiter  $j$  is to achieve this optimum ratio for every UE. For example, both UE  $x$  and UE  $y$  are requesting resources from CoS  $j$ 's credit pool:

$$p_{xj} = \frac{\frac{a_{xj}}{A_j} - \mu_{xj}}{\mu_{xj}} < p_{yj} = \frac{\frac{a_{yj}}{A_j} - \mu_{yj}}{\mu_{yj}} \quad (4.1.1)$$

If expression 4.1.1 is true then UE  $x$  will be considered for CoS  $j$  credits before UE  $y$ . Otherwise, UE  $y$  will be considered before UE  $x$ . The advantage of the WUSO arbitration mechanism is that it can be executed offline (i.e., during the previous cycle) in order to produce the UE orders for the current cycle.

The wideband SRS technique is employed with this algorithm. During the transmission cycle, each UE will send a single wideband SRS signal to the base station. As we mentioned in Chapter 3.2, the transmission cycle in our simulation is 4 ms. As shown in Figure 4.1.2, the wideband SRS will be sent by UE periodically in every 2 ms. The single wideband SRS will cover the whole uplink bandwidth. In our thesis, bandwidth is 20 MHz space in the frequency domain. The whole bandwidth consists of 100 resource blocks in the frequency domain. After the base station received this wideband SRS, the Adaptive Modulation and Coding (AMC) selector will choose the appropriate transmission rate for the UE, based on the SINR of the received reference signal.

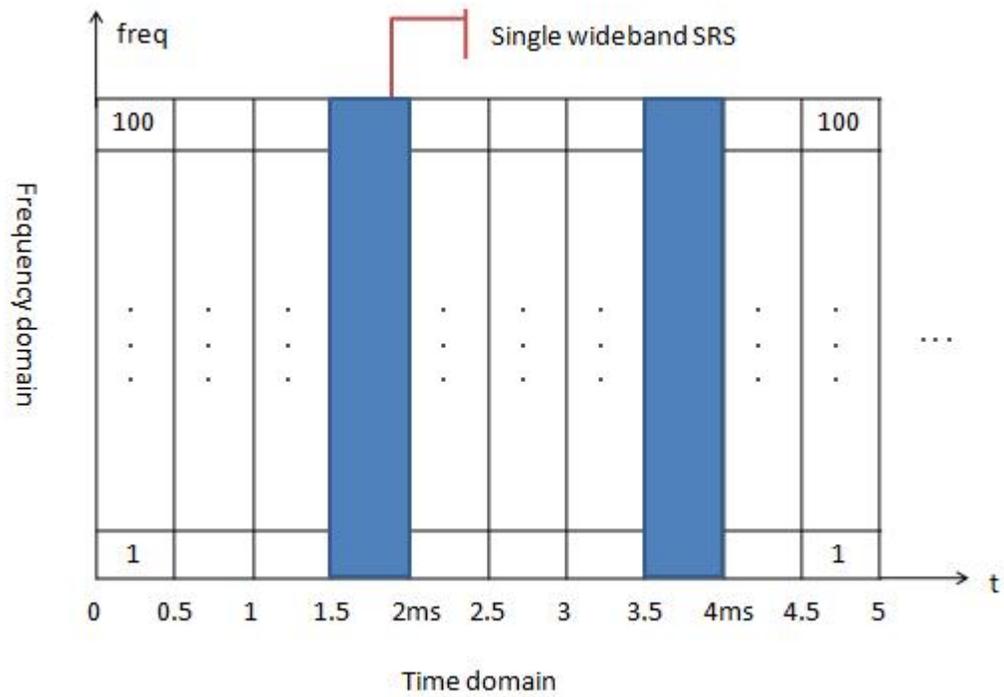


Figure 4.1.2: Illustration of the wideband SRS transmission

For example, if the received SINR of this wideband SRS is in the range of  $7 \sim 11$  dB, then the base station will choose mode 3 for this UE based on Table 3.2. In this case, one symbol represents 2 bits with QPSK modulation scheme. With the coding rate  $3/4$ , the information bit rate on each symbol is given by:

$$R_{bpsymbol} = 2 \text{ bit/symbol} \times \frac{3}{4} = 1.5 \text{ bit/symbol} \quad (4.1.2)$$

Based on the structure of the LTE uplink resource block we mentioned in section 3.4, the information transmission rate of the resource block is given by:

$$\begin{aligned} R_{RB} &= R_{bpsymbol} \times 7 \text{ symbol/subc} \times 12 \text{ subc/RB} \times \frac{1 \text{ RB}}{0.5 \text{ ms}} \\ &= 1.5 \text{ bit/symbol} \times 7 \text{ symbol/subc} \times 12 \text{ subc/RB} \times \frac{1 \text{ RB}}{0.5 \text{ ms}} \quad (4.1.3) \\ &= 252 \text{ Kbps} \end{aligned}$$

Because of the single wideband SRS signal from UE will cover the whole bandwidth, only one SINR value can be obtained. So, for each UE, all the resource blocks in the system will have the same transmission rate during a given transmission cycle.

A re-arrange buffer is employed with CASA. The re-arrange buffer stores the credit calculated for all UEs during the scheduler processing. Each entry of the re-arrange buffer contains three information elements: UE I.D, CoS I.D and the number of resource blocks given to that specific UE in that CoS. At the end of the scheduling time, the total number of resource blocks stored in the rearrange buffer will not be more than the maximum number of resource blocks in the transmission cycle. After the scheduling is completed, the base station will rearrange the resource blocks so that all grants for a certain UE will

be grouped together. Then the rearrange buffer generates grant messages and sends them to UEs. The grant message contains: UE I.D, CoS I.D, RB I.D, and the amount of allocated RBs. The data transmission will start from the farthest UE to the nearest UE.

## 4.2 Grant Processing

The base station will use the following algorithm (CASA) to calculate grants for each UE. The CASA algorithm includes two rounds of execution. The first round is mandatory that gives rise to the distribution of the most of the available resource blocks between UEs. At the end of this round, some of the resource blocks may still remain in the credit pools. The second round is designed to redistribute the remaining resource blocks between UEs. And the second round of the algorithm will lead to a work-conserving scheduling service.

### Round 1

**Step 1:** The base station will choose an appropriate transmission rate  $R_i$  for each UE based on the SINR of the received reference signal as we mentioned before.

**Step 2:** The base station will fill up both CoS and UE credit pools, which are given by:

$$N_i = \left\lfloor \frac{u_i}{\sum_a u_a} \times N_{total} \right\rfloor \quad (4.2.1)$$

$$V_j = \left\lfloor \frac{v_j}{\sum_b v_b} \times N_{total} \right\rfloor \quad (4.2.2)$$

,where  $N_{total}$  is the total number of available resource blocks within a given transmission cycle.  $\sum_a u_a$  is the sum of all active UEs' weight.  $\sum_b v_b$  is the

sum of CoSs' weight.

**Step 3:** For each UE, the base station will convert the number of CoS Bytes of data requested from each UE to the number of resource blocks:

$$n_{reqij} = \left\lceil \frac{Req_{ij}}{R_i \times 0.5ms} \right\rceil \quad (4.2.3)$$

,where  $Req_{ij}$  is the number of Bytes of data requested from UE  $i$  for CoS  $j$ .

**Step 4:** The base station begins processing grants for CoS 1 first. After CoS 1 grants for all UEs are calculated, the base station will process CoS 2 grants. In general, the base station will begin processing CoS  $j$  grants only after all CoS  $(j - 1)$  grants have been calculated.

The calculation of CoS  $j$  grant for each UE is as follows. The base station takes the first UE, for example UE  $i$  stored in the priority database  $p_{ij}$ , and compares the requested number of resource blocks  $n_{reqij}$  with both CoS  $j$  and UE  $i$  credit pool ( $N_i$  and  $V_j$ ), and will grant this request if there is an equivalent amount of credits exist in both CoS  $j$  and UE  $i$  credit pools. If either of these credit pools does not have enough credits, UE  $i$  will get all the existing credits from the smaller credit pool. The base station will put UE  $i$  into the partial grant set. Whenever the base station issues a full or partial grant, it subtracts the number of resource blocks from both the credit pools. The base station then calculates a grant for the next UE in the priority database using the above algorithm.

## Round 2

The base station first examines the total number of credits in all CoS credit pools. If the total number of credits is zero, the grant processing will be finished. If the total number of credits is not zero, it means that there are some credits still remaining in the CoS credit pools altogether, which is represented

by  $N_{remainder}$ . These unused credits can be recollected and distributed between UEs that have received partial or no grant in Round 1. The base station distributes these unused credits between the CoSs in proportion ( $v_j$ ):

$$V'_j = \begin{cases} \left\lfloor \frac{v_j}{\sum_b v_b} \times N_{remainder} \right\rfloor & j=1 \\ \left\lfloor \frac{v_j}{\sum_b v_b} \times N_{remainder} \right\rfloor + \alpha_{j-1} & j>1 \end{cases} \quad (4.2.4)$$

The first term in the right side of (4.2.1) represents the fair share of the remaining credits (from Round 1) that will be given to each CoS in the second round. While CoS 1 gets its fair share only, a lower priority CoS  $j$  ( $j > 1$ ) gets its fair share plus extra  $\alpha_{j-1}$  credits. The  $\alpha_{j-1}$  is the total credits that will remain unused in CoS  $j-1$  credit pool when the processing of grants for that CoS is completed in the second round.

The base station executes the same algorithm as in Round 1, but this time it only considers the UEs stay in the partial grant set. Like in Round 1, the CoS 1 grants are processed first, followed by other CoSs in sequence. Once the processing of grants for CoS 1 is completed, the base station knows exactly how many credits are still remaining in the CoS 1 credit pool, i.e.  $\alpha_1$ . The base station will then configure CoS 1 credit pool with its fair share of the remaining bandwidth plus  $\alpha_1$ . In general, the exact value for  $\alpha_{j-1}$  is known once the processing of grants for CoS ( $j - 1$ ) is completed.

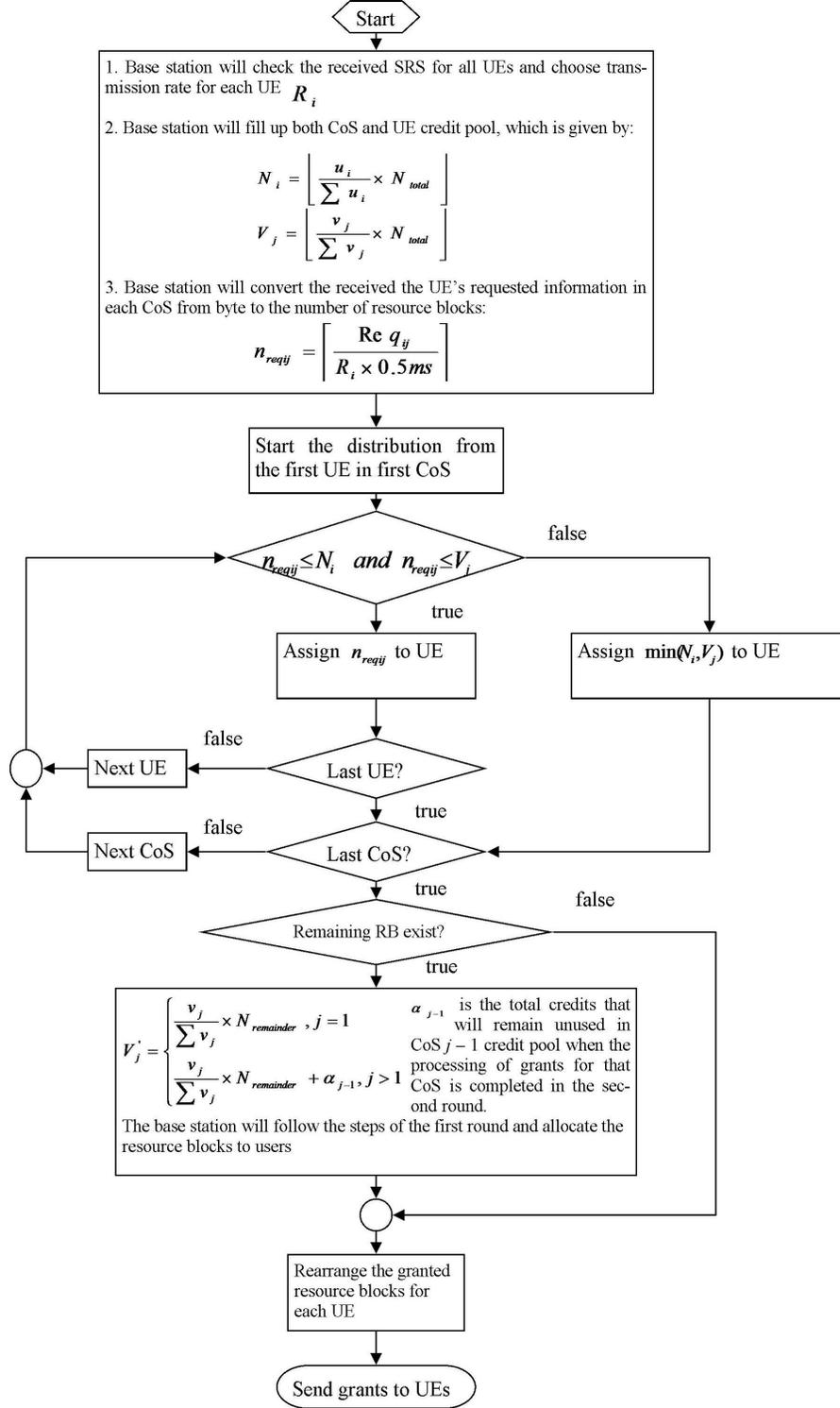


Figure 4.2.1: Illustration of the grant processing

## Chapter 5

# Independent Channel Aware Scheduling Algorithm

In this chapter, we will introduce the Independent Channel Aware Scheduling Algorithm (ICAS). This algorithm employs the *narrowband SRS* technique together with a *credit pooling* technique to allocate resource blocks among customers' classes of service in a fair manner.

### 5.1 System Structure

Figure 5.1.1 shows, the structure of the ICAS implemented in the base station. The base station maintains  $N$  credit pools for UEs, which are used to control the usage of the resource blocks by UEs. The credit pool for UE  $i$  has two parameters: weight ( $u_i$ ) and the size ( $N_i$ ).  $u_i$  is a pre-configured value, which is based on the Service Level Agreement (SLA) [37].  $N_i$  will be initialized in grant processing. In our thesis, the scheduling interval of ICAS is set to be 2 ms.

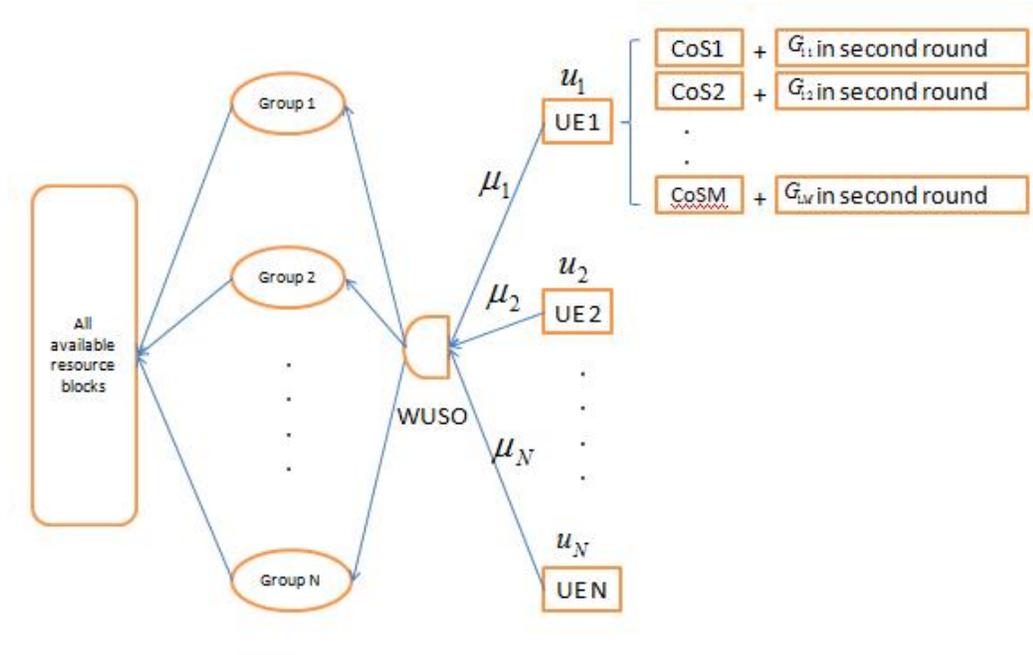


Figure 5.1.1: Structure of the ICAS system

Similar to CASA, the Weighted UE scheduling Order (WUSO) arbitration mechanism is also employed with the ICAS. But in ICAS, the base station only employs one WUSO arbiter to ensure that the contending UE  $i$  will receive a fair share of the resource, based on its weight ( $\mu_i$ ) in WUSO. In our thesis, we set  $\mu_i = u_i$ . In order to achieve expected Class-of-service, UE will distribute assigned resource blocks to different CoSs based on their weight.

During the transmission cycle, each UE will send four groups of 25 narrowband SRS signals to the base station. Each group of narrowband SRS signals will cover one fourth of the entire 100 resource blocks in the frequency domain, which is shown in Figure 5.1.2. Since different resource block will be covered by different narrowband SRS, the base station will consider the resource blocks in the frequency domain to be different. So in this algorithm, the base station will choose the transmission rate for each resource block in the frequency domain

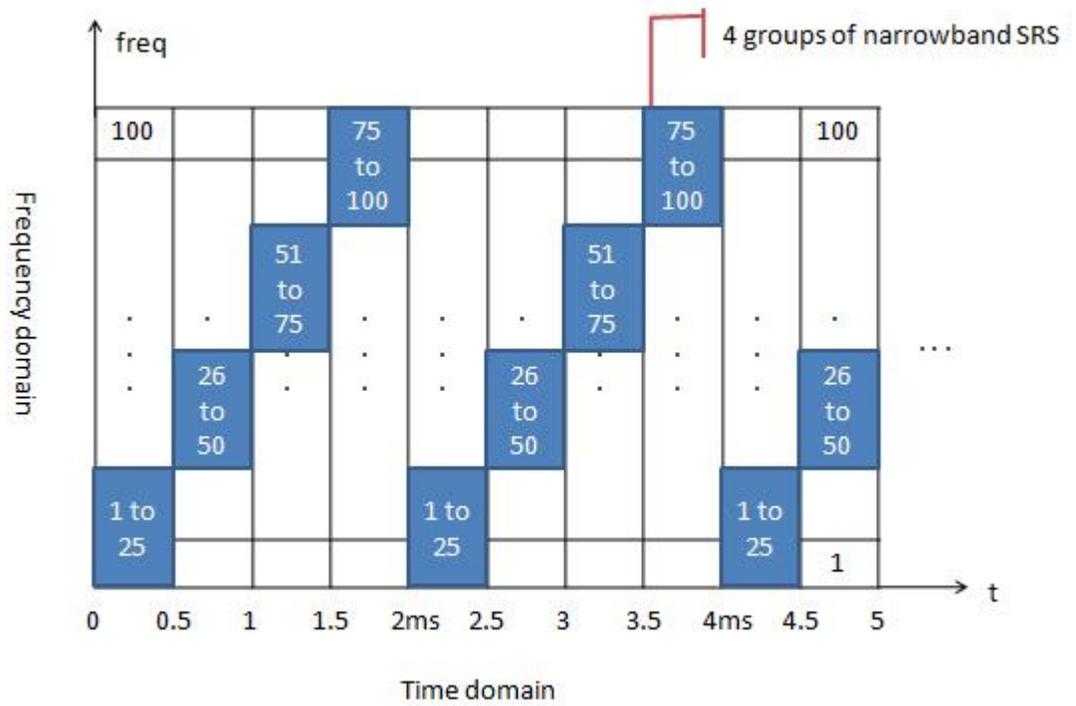


Figure 5.1.2: The transmission of the narrowband SRS in ICAS

based on the received SINR of each narrowband SRS signal.

## 5.2 Grant Processing

The base station will use the following algorithm (ICAS) to calculate grants for each UE. The ICAS algorithm includes two rounds of execution. The first round is mandatory that gives rise to the assignment of most of the available resource blocks between UEs. At the end of this round, some of resource blocks may still remain in the system. The second round is designed to reassign the remaining resource blocks between UEs, which received partial grant in the first round. The second round of the algorithm will lead to a work-conserving scheduling service.

### Round 1

**Step 1:** The base station will choose an appropriate transmission rate  $R_{im}$  for UE  $i$  on the  $m$ th ( $m = 1, \dots, 100$ ) resource block in the frequency domain based on the  $m$ th SINR of the received narrowband SRS signal, this processing had been mentioned in Chapter 3.5.

**Step 2:** The base station takes the first UE, for example UE  $i$ , stored in the priority database WUSO based on its priority value  $p_i$ :

$$p_i = \frac{\frac{a_i}{A} - \mu_i}{\mu_i} \quad (5.2.1)$$

,where  $a_i$  is the total amount of data (Bytes) sent by UE  $i$  in the previous transmission, and  $A$  is the total amount of data (Bytes) transmitted in the system.

**Step 3:** The base station will assign a group of resource blocks to UE  $i$ , which consists of four 0.5 ms time slots. The group size is defined by the

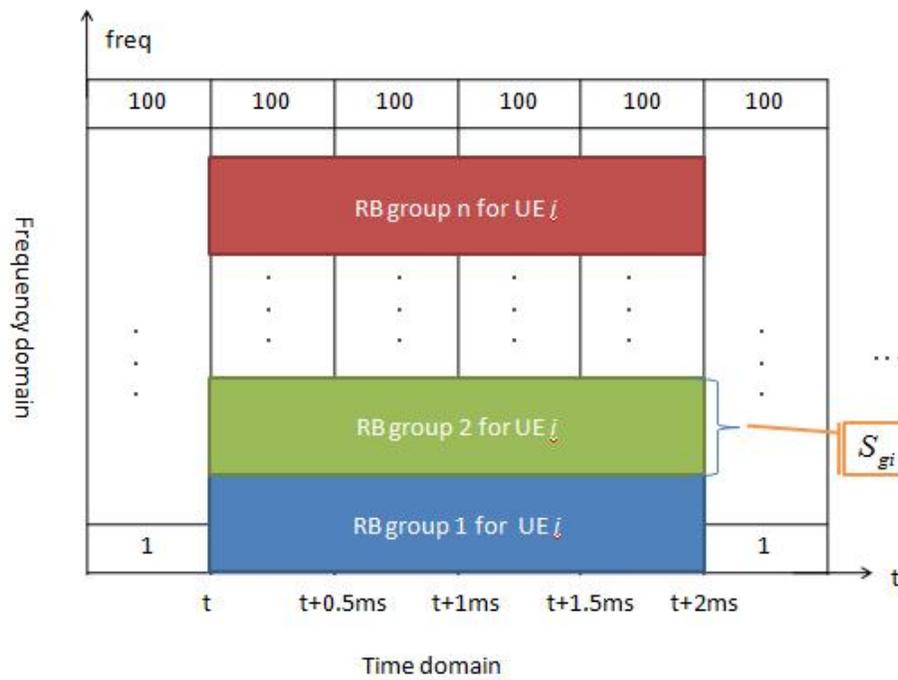


Figure 5.2.1: The processing for choosing resource blocks group for UE  $i$

number of resource blocks in a 0.5 ms timeslot. So the group size for UE  $i$  is given by:

$$S_{gi} = \left\lfloor \frac{\mu_i}{\sum_s \mu_s} \times 100 RB \right\rfloor \quad (5.2.2)$$

where the  $\sum_s \mu_s$  is the sum of all active UEs' weights. Since the resource blocks within each group ( $S_{gi}$ ) have to be adjacent, there are at most  $n_i$  different locations within the 0.5 ms timeslot where the group can be located, where  $n_i$  is given by:

$$n_i = \left\lfloor \frac{100}{S_{gi}} \right\rfloor \quad (5.2.3)$$

Then the base station will decide the position of the group for UE  $i$  based on the UE channel condition. As shown in Figure 5.2.1, the base station will calculate the amount of data in byte ( $M_{ia}$  and  $a$  is from 1 to  $n_i$ ) that can be supported within each resource group in the 0.5 ms timeslot for UE  $i$ :

$$\begin{aligned} M_{i1} &= \sum_{m=1}^{S_{gi}} (R_{im} \times 0.5 \text{ ms}) \\ M_{i2} &= \sum_{m=S_{gi}+1}^{2 \times S_{gi}} (R_{im} \times 0.5 \text{ ms}) \\ &\dots \\ M_{in_i} &= \sum_{m=(n_i-1) \times S_{gi}+1}^{n_i \times S_{gi}} (R_{im} \times 0.5 \text{ ms}) \end{aligned}$$

Then the base station will assign the group with maximum value ( $Max_i = MAX(M_{i1}, \dots, M_{in_i})$ ) to UE  $i$  and initialize the UE credit pool size to be

$$N_i = 4 \times Max_i \quad (5.2.4)$$

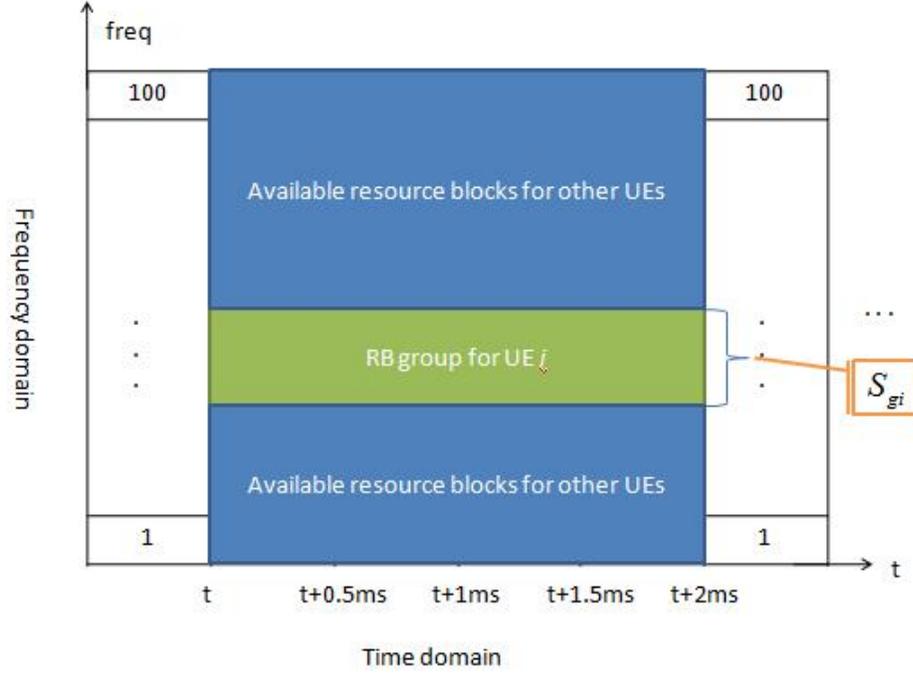


Figure 5.2.2: Illustration of remove of the resource blocks group of UE  $i$

Because of the scheduling interval is set to be 2 ms, there are four 0.5 ms time slots in the group in equation (5.2.4). After the group is assigned to UE  $i$ , the base station will remove the allocated resource group to UE  $i$  from the available spectrum, which is shown in Figure 5.2.2.  $S_{gi}$  will be subtracted from the total 100 available resource blocks in 0.5 ms timeslot.

**Step 4:**

1. The base station will calculate the total request size  $Q_i$  for UE  $i$ :

$$Q_i = \sum_{j=1}^l Req_{ij} \quad (5.2.5)$$

where  $l$  is the number of CoSs.  $Req_{ij}$  is the number of Bytes of data requested from UE  $i$  for CoS  $j$ .

2. The base station compares the request size  $Q_i$  with the UE  $i$  credit pool size  $N_i$ . If  $Q_i > N_i$ , the base station will assign  $N_i$  to UE  $i$ , which means the whole group of resource blocks will be assigned to UE  $i$ . The base station will begin the allocation from CoS 1. And the grant for CoS  $j$  will be calculated only after CoS  $j - 1$  grants have been calculated. Each CoS' grant is calculated based on its predefined proportion. UE  $i$  will be moved to partial grant set. If  $Q_i = N_i$ , the base station will assign  $N_i$  to UE  $i$ . Base station will consider UE  $i$  received a full grant. If  $Q_i < N_i$ , then the base station will check whether  $Q_i$  can be accommodated by a multiple integral time slots in the group. The number of integral time slots ( $c$ ) is given by

$$c = \left\lfloor \frac{Q_i}{Max_i} \right\rfloor$$

where  $c$  could be 1, 2 or 3. If  $Q_i = c \times Max_i$ , the base station will assign  $c$  time slots of resource blocks in the group to UE  $i$ . If  $Q_i > c \times Max_i$ , the base station will assign the integral time slots ( $c$ ) of resource blocks to UE  $i$  firstly and subtract this part of data from  $Q_i$ , since the resource blocks in one time slot must keep continuously. Since the base station knows the transmission rates of all resource blocks in the  $(c+1)$ th time slot, the base station can pick up a minimum set of continuous resource blocks for UE  $i$  to transmit the rest requested data ( $Q_i - c \times Max_i$ ). UE  $i$  will receive a full grant. Figure 5.2.3 shows an example of the processing we mentioned above, in this example we assume  $Q_i = 70$  and  $Max_i = 20$ . There are eight resource blocks in one time slot.

3. After the grant calculation for UE  $i$ , the base station then calculates a grant for the next UE in the priority database using the above algorithm.

## Round 2

Firstly, the base station examines the partial grant set. If the partial grant set is empty then finish the processing and send the grant information to UEs. If the partial grant set is not zero, it means that there are some UEs that received a partial grant in Round 1. Like Round 1, the base station will pick up the UE in the partial grant set who has a lower priority value. Since the resource blocks in the frequency domain must be continuous, the base station will make sure that it assigns the resource blocks to the UE continuously in the frequency domain. So the base station will check if there are any unused resource blocks existing in both the upper side and lower side of the UE group. The base station will check whether these unused resource blocks connect with the UE group. If it is true, the base station will assign these extra resource blocks to the UE. Otherwise, these unused resource blocks will not be assigned to UE. The processing is shown in the back slash part of Figure 5.2.4. In Figure 5.2.4, the UE assigned in group  $i$  received a partial grant in round 1. In round 2, the base station will check if there are unused resource blocks in both group  $i - 1$  and group  $i + 1$ . If it is true, the base station will check whether these unused resource blocks connect with group  $i$ . In Figure 5.2.4, we can see that in group  $i - 1$  there are some resource blocks that had been assigned in time slot 3. The base station can only allocate the resource blocks above those assigned resource blocks to the UE in group  $i$ , which is shown by the black slash part in the Figure. In group  $i + 1$ , the processing is the same. If all resource blocks in

$Q_i = 70$
$c = \left\lfloor \frac{65}{20} \right\rfloor = 3$

	Slot 1	Slot 2	Slot 3	Slot 4
RB1	1	1	1	1
RB2	2	2	2	2
RB3	4	4	4	4
RB4	3	3	3	3
RB5	3	3	3	3
RB6	2	2	2	2
RB7	2	2	2	2
RB8	3	3	3	3

The base station will assign 3 slots of resource blocks to UE firstly and subtract this part of data from $Q_i$
$Q_i = 70 - 60 = 10$

The base station will find all combinations of resource blocks in slot 4, which can transmit the rest of the requested size.  $Q_i = 10$

All available combinations are shown below: {RB1,RB2,RB3,R4}, {RB2,RB3,RB4,RB5}, {RB3,RB4,RB5}, {RB4,RB5,RB6,RB7}, {RB5,RB6,RB7,RB8}. We can see the third combination has the minimum number of resource blocks, so the base station will assign this set of resource blocks to UE.

	Slot 1	Slot 2	Slot 3	Slot 4
RB1	1	1	1	1
RB2	2	2	2	2
RB3	4	4	4	4
RB4	3	3	3	3
RB5	3	3	3	3
RB6	2	2	2	2
RB7	2	2	2	2
RB8	3	3	3	3

RB represents resource block;  
Slot represents the time slot in the resource group.

Figure 5.2.3: Illustration of the resource blocks assignment when  $Q_i < N_i$

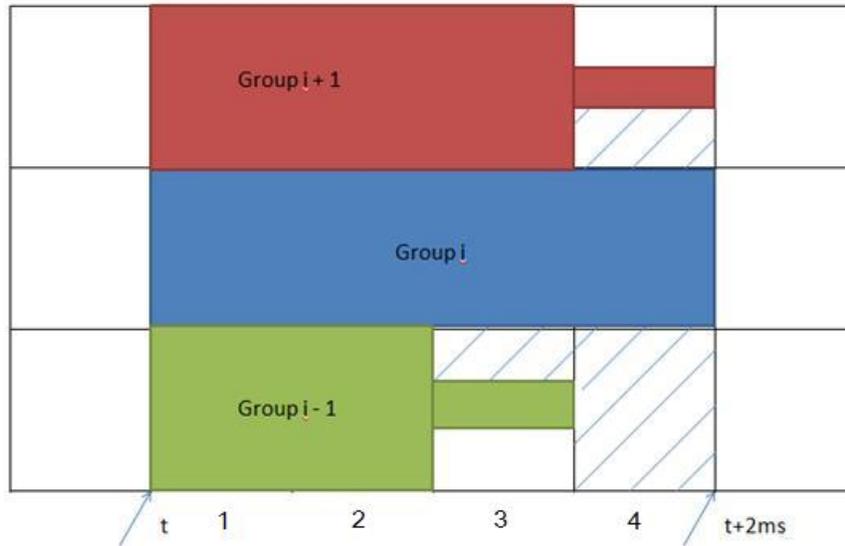


Figure 5.2.4: Illustration of the grant processing

a time slot are unused, all resource blocks will be assigned to the UE in group  $i$ , which is shown in the time slot 4 in group  $i - 1$  in the Figure 5.2.4.

Like in Round 1, for each UE the CoS 1 grants are processed first, followed by other CoSs in sequence.

After the grant processing is finished, the base station will send the grant message to all UEs. The grant message contains: UE I.D, CoS I.D, and allocated RB I.D. The grant processing is shown in Figure 5.2.5.

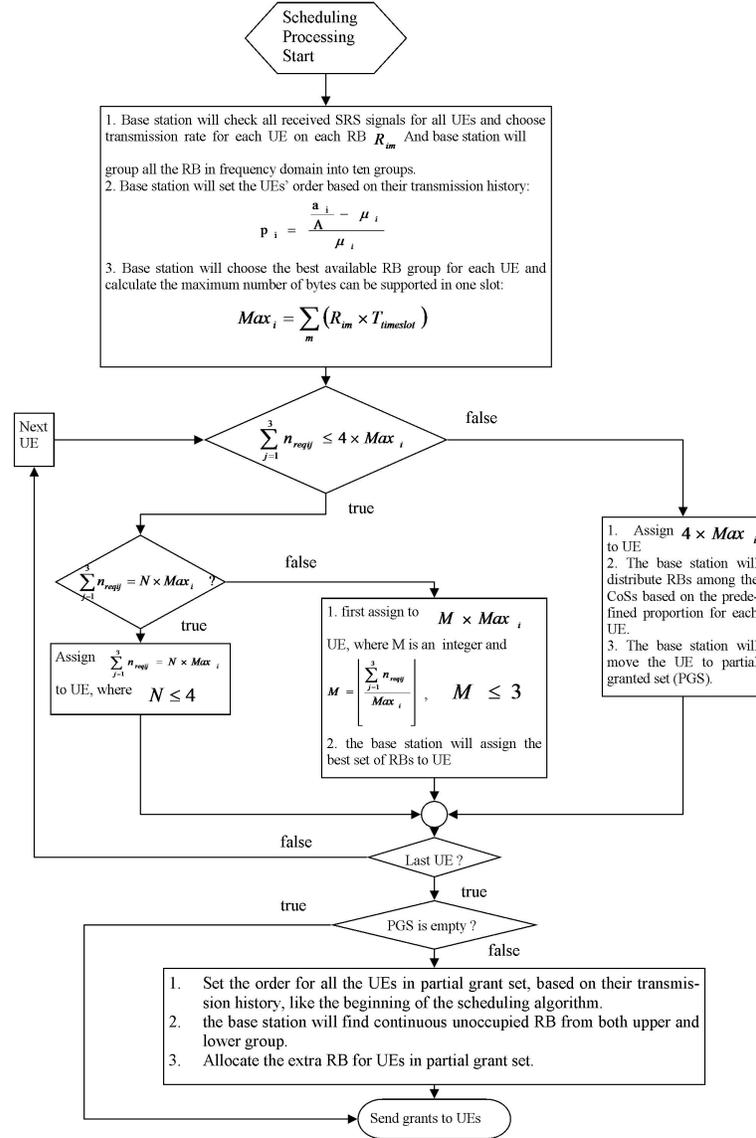


Figure 5.2.5: Illustration of ICAS grant processing

## Chapter 6

# Simulation Analysis

In this chapter, we will present our simulation framework and introduce the traffic models for different classes of data. We have developed an in-house simulation program in C++ to evaluate and compare the performances of CASA and ICAS algorithms in terms of packet loss ratio, packet delay and throughput.

### 6.1 The traffic models for CoS

We adopt three classes of traffic in our simulation: CoS 1, CoS 2 and CoS 3. These CoSs share a common buffer space of  $F = 5$  Mbytes at each UE. CoS 1 represents low-loss, delay-sensitive Expedited Forwarding (EF) service, which typically provides for CBR voice. CoS 2 represents Assured Forwarding (AF) service, which provides for non-delay sensitive, bandwidth guaranteed VBR (Variable Bit Rate) video/data. CoS 3 represents Best Effort (BE) service, which requires no bandwidth commitment from the network.

In each UE, the CoS 1 is represented by a CBR VoIP traffic. Based on G.711 standard, the data rate for the VoIP stream is chosen as  $r_{CBR} = 66 \text{ packets/s}$

and packet length  $L_{CBR} = 120$  Bytes. The amount of CBR traffic is kept constant for all simulations. The offered load at each UE is varied by changing the rate with which the traffic is generated for CoS 2 and CoS 3.

We used an ON-OFF source to generate traffic for each of these two CoSs at each UE. The two sources have identical parameters. For each source, the ON and OFF (silent) intervals are drawn according to a *Pareto distribution* [33].

If  $X$  is a random variable with a Pareto distribution, then the probability that  $X$  is greater than some number  $x$  is given by:

$$Pr(X > x) = \begin{cases} \left(\frac{x_m}{x}\right)^k & \text{for } x \geq x_m \\ 1 & \text{for } x < x_m \end{cases} \quad (6.1.1)$$

, where  $x_m$  is called the location parameter, and is the minimum possible value of random value  $X$ .  $k$  is called a shape parameter. Both  $x_m$  and  $k$  are positive in value. The probability density function (PDF) of a Pareto distribution is given by:

$$f_x(x) = \begin{cases} k \frac{x_m^k}{x^{k+1}} & \text{for } x \geq x_m \\ 0 & \text{for } x < x_m \end{cases} \quad (6.1.2)$$

The cumulative distribution function (CDF) is given by:

$$F_X(x) = \begin{cases} 1 - \left(\frac{x_m}{x}\right)^k & \text{for } x \geq x_m \\ 0 & \text{for } x < x_m \end{cases} \quad (6.1.3)$$

The expected mean value and variance of the Pareto distribution are given by:

$$E(X) = \frac{kx_m}{k-1} \quad (6.1.4)$$

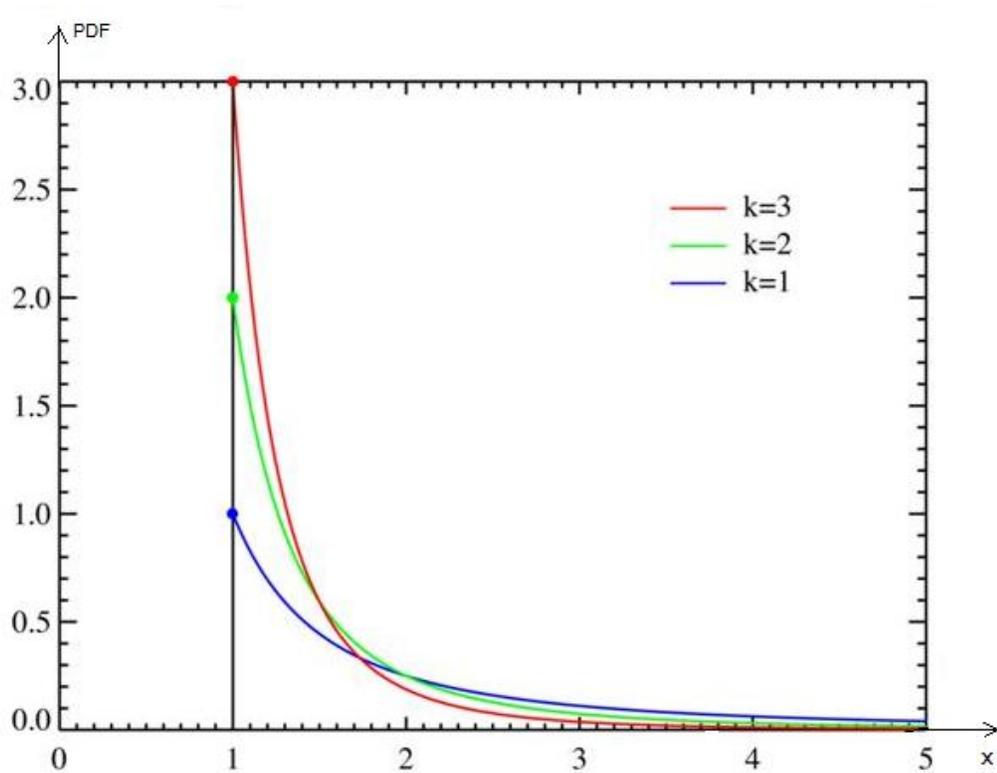


Figure 6.1.1: Pareto probability density function (PDF) for with  $x_m = 1$

The PDF and CDF of a Pareto distribution with  $x_m = 1$  are illustrated in Figure 6.1.1 and Figure 6.1.2, respectively.

The Pareto distribution is a heavy-tailed distribution characterized by the shape parameter  $k$  and the location parameter  $x_m$ . This distribution has a finite mean and infinite variance when the shape parameter is chosen in the interval  $[1, 2]$  [6]. In our simulations, the shape parameter for the ON and OFF intervals is set to 1.4 and 1.2, respectively. The location parameter is a function of the shape parameter and the mean of the distribution, shown in Formula 6.1.5,

$$x_m = E(X) \times (k - 1.0) / k \quad (6.1.5)$$

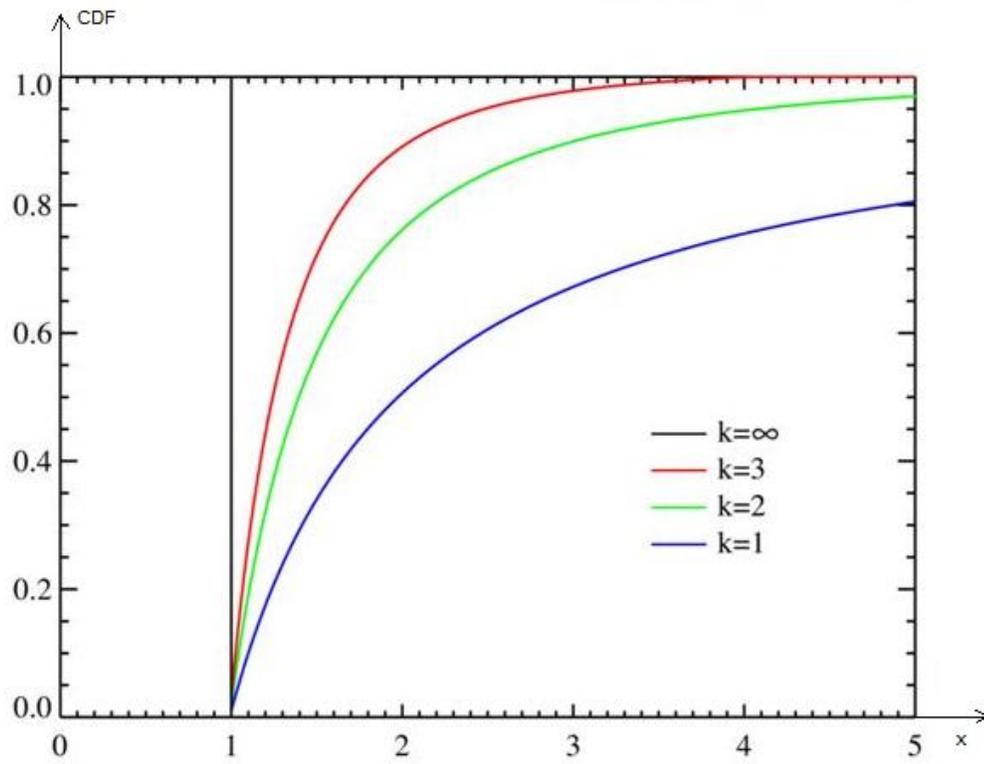


Figure 6.1.2: Pareto cumulative distribution function (CDF) for with  $x_m = 1$

We assume that the amount of data in CoS 3 is twice than CoS 2. To achieve the desired offered load in the system, for each UE the average arrival rates of each ON-OFF source for CoS 2 and CoS 3 varied in the interval  $[0.10, 2.48]$  and  $[0.20, 4.96]$  Mbps, respectively. The length of packets generated during an ON state follows the tri-modal distribution used in [6]. These three modes correspond to most frequent packet sizes 64, 594 and 1518 Bytes. In our simulations, each of these packets is generated with a frequency of 62%, 10%, and 28%, respectively.

## 6.2 Simulation setup

In this thesis, we intend to evaluate the performance of the CASA and ICAS algorithms to support multiple users in a large cell. In our simulation, we will set the distance between a given UE and the base station randomly in the interval  $[0\text{Km}, 30\text{Km}]$ . In order to test the two algorithms fairly and make sure that the UEs will be well-distributed in the cell, we assume that the UEs will stay stationary during the simulation.

Based on the transmission cycle of the sounding reference signal, the scheduling interval for both CASA and ICAS is chosen to be 2 ms. All the UEs have identical traffic parameters and offered loads. The offered load for system is defined as a ratio of total incoming data rate to the maximum transmission capacity of the LTE uplink. Based on the specified LTE standard, we know the peak rate of the LTE uplink transmission is 75 Mbps. So we will use this value to represent the maximum uplink transmission capacity. In our simulation we assume that all UEs will share the entire wireless resource. We set the weight of every UE in WUSO to be the same value  $\mu_i = \frac{1}{N}$  in the case of ICAS and  $\mu_{ij} = \frac{1}{N \times M}$  in the case of CASA, where  $N$  is the number of the UEs and  $M$  is the number of the CoSs. In our simulation we assume  $N = 10$  and  $M = 3$ . So

the offered load for each UE will be given by:

$$CoS1 = 0.064 \text{ Mbps} \quad (6.2.1)$$

$$CoS2 = (75 \times \rho - (10 \times 0.064)) \times \frac{1}{3}/N \text{ Mbps} \quad (6.2.2)$$

$$CoS3 = (75 \times \rho - (10 \times 0.064)) \times \frac{2}{3}/N \text{ Mbps} \quad (6.2.3)$$

, where  $\rho$  is the offered load for the system, which ranges from 0.1 to 1.0.

### 6.3 Simulation results and analysis

In our simulation, we use the delay, packet loss ratio, throughput and resource blocks allocated ratio to evaluate the performance of the proposed algorithms. In our thesis, the packet delay is measured from the packet generated in UE to it received by the base station. It consists of queuing delay, transmission delay and propagation delay. The packet loss ratio is given by

$$\text{packet loss ratio} = \frac{\text{number of dropped packets in CoS queue}}{\text{Total packets generated in CoS}}$$

The throughput is given by

$$\text{throughput} = \frac{\text{total amount of data received in the base station}}{\text{total simulation time}}$$

In our thesis, the resource block allocated ratio is used to represent the normal bandwidth utilization. The resource block allocated ratio is given by

$$\text{resource block allocated ratio} = \frac{\text{total allocated resource blocks}}{\text{total available resource blocks in system}}$$

Since the scheduling interval is 2.0 ms, the total available resource blocks in the system are 400. In [2], the authors' simulation results show that in the

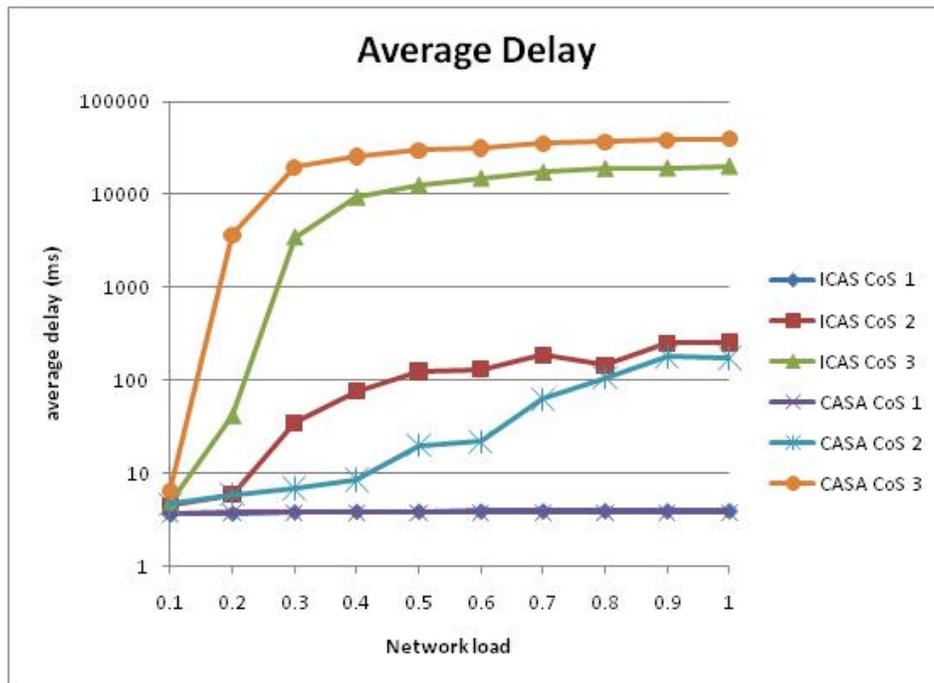


Figure 6.3.1: CASA and ICAS average delay with 0.5 ms TTI

case of a 0.5 ms Transmission Time Interval (TTI), the system can achieve higher capacity. In the case of 1.0 ms TTI, the system will achieve higher power efficiency. Based on this property, we will run our simulation with both 0.5 ms and 1.0 ms TTI.

### 6.3.1 Simulation results with 0.5 ms TTI

In case of 0.5 ms TTI, the minimum resource unit can be allocated to UEs is resource block. The simulation results of both CASA and ICAS with 0.5 ms TTI will be shown below in the Figure 6.3.1 through Figure 6.3.5. The simulation results of CoS  $i$  ( $i = 1, 2, 3$ ) is labeled as CASA CoS  $i$  and ICAS CoS  $i$  in the figures, respectively.

Figure 6.3.1 shows the average packet delay of the each CoS traffic achieved

by CASA and ICAS when the system is running with 0.5 ms TTI. We measure the average delay as a function of the network offered load.

From Figure 6.3.1, we can see that CoS 1 data in both algorithms got almost the same performance; the average delay of CoS 1 increases slightly while the network loads increase. In our simulation, the average delay stays in the interval [3.6 ms, 4.0 ms].

The CoS 2 average delay in both CASA and ICAS increases as the network load increases. The CoS 2 delay of CASA is significantly less than ICAS when the network load is in [0.3, 0.7] interval. When the network load is higher than 0.7, the performance of CASA approaches ICAS. As we mentioned in Chapter 5, the ICAS scheduler will assign the resource group to UE. In some transmission cycle, the UE with high WUSO priority value may get a group in bad channel, since the good channel will be used to transmit the data from the UE has lower priority value. In this case, the UE can only send a small amount of data which is stored in the UE's CoS queue with low transmission rate. The rest of the data have to wait for another grant when UE will use a good group to send data. This will increase the queuing delay of these stored data. It will cause the total average packet delay of ICAS to increase without affect the average throughput.

For CoS 3, ICAS has a better performance than CASA. Based on the results shown in Figure 6.3.4, we can see that ICAS CoS 3 has a much higher throughput than CASA CoS 3. This means that there are more packets can be transmitted from ICAS CoS3 queue. For each packet in ICAS CoS queue the packet queuing delay will be reduced. So the total queuing delay in the ICAS CoS 3 is lower than CASA CoS 3.

Figure 6.3.2 shows the average packet loss ratio for both CASA and ICAS with 0.5 ms TTI. For CoS 1, the packet loss ratio of both CASA and ICAS

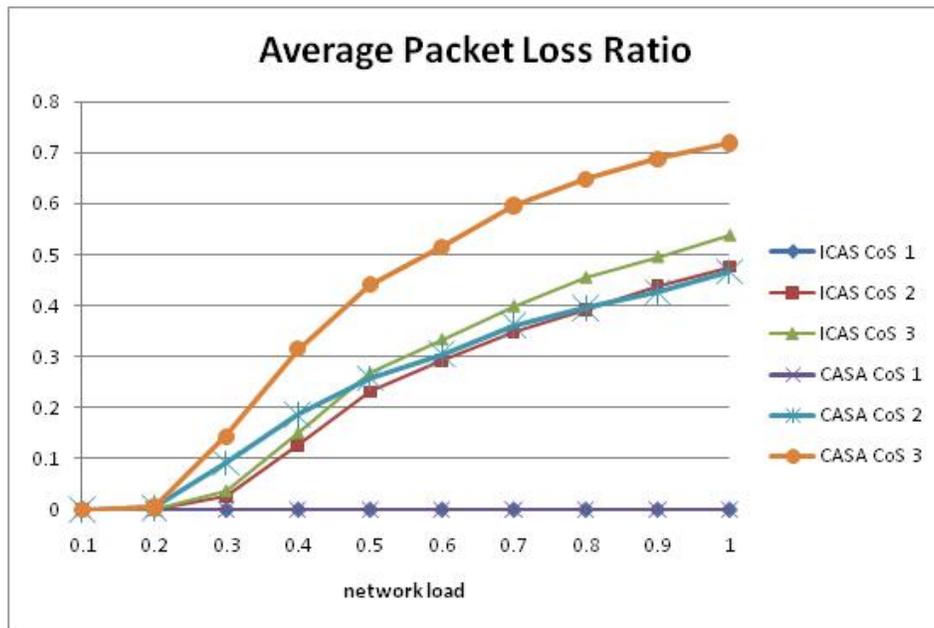


Figure 6.3.2: CASA and ICAS Average Packet Loss Ratio with 0.5 ms TTI

stays zero, which means the VoIP service will not be degraded by other CoS transmission. For both CASA and ICAS, the CoS 2 packet loss ratio increases as the network load increases. The CoS 2 packet loss ratio of ICAS is less than ICAS when the network load is in the  $[0.3, 0.7]$  interval. When the network load is higher than 0.7, the performance of CASA approaches ICAS. CASA CoS 3 has higher packet loss ratio than ICAS CoS 3.

Figure 6.3.3 shows both CASA and ICAS performance in throughput with 0.5 ms TTI. From this figure we can see that ICAS has higher throughput than CASA. When the offered load increases, the maximum throughput of ICAS increases and eventually reaches to a maximum value of 35 Mbps. The throughput of CASA however reaches to a saturation level of 23 Mbps, when the offered load exceeds 0.3. As we mentioned in Chapter 5, ICAS employs the narrowband SRS technique. The ICAS scheduler can obtain more specific information on the

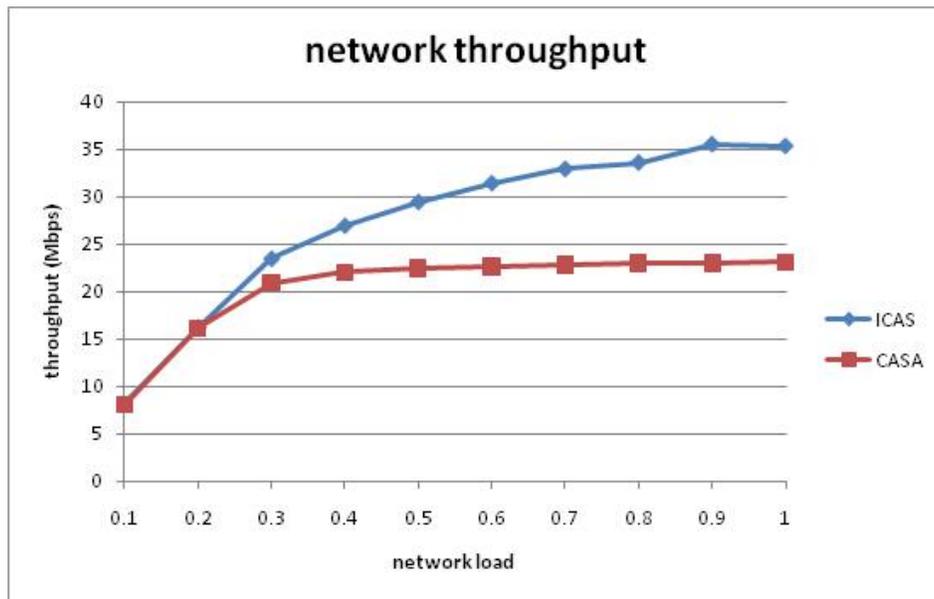


Figure 6.3.3: CASA and ICAS network throughput with 0.5 ms TTI

entire transmission channel. Even though the UE will transmit data through a high path-loss channel, the UE can send data with the resource blocks has a relatively higher transmission rate. But in CASA, wideband SRS is employed. If UE is in a heavy path-loss channel, the SINR of received wideband SRS may have a relatively lower value. This may reduce the quality of the channel condition estimation, even though there are some high quality sub-channels existing. UE will use a relatively lower data rate to transmit the data. So CASA has a lower throughput than ICAS.

Figure 6.3.4 shows the average throughput performance of each CoS for both CASA and ICAS when TTI is set to be 0.5 ms. Both CASA CoS 1 and ICAS CoS 1 throughputs show exactly what is expected of the CBR traffic. There was no change in CoS 1 traffic. Both CASA CoS 1 and ICAS CoS 1 throughputs remain constant as the offered load increases. For CoS 2, both CASA and ICAS have similar performance. As the offered load increases, both

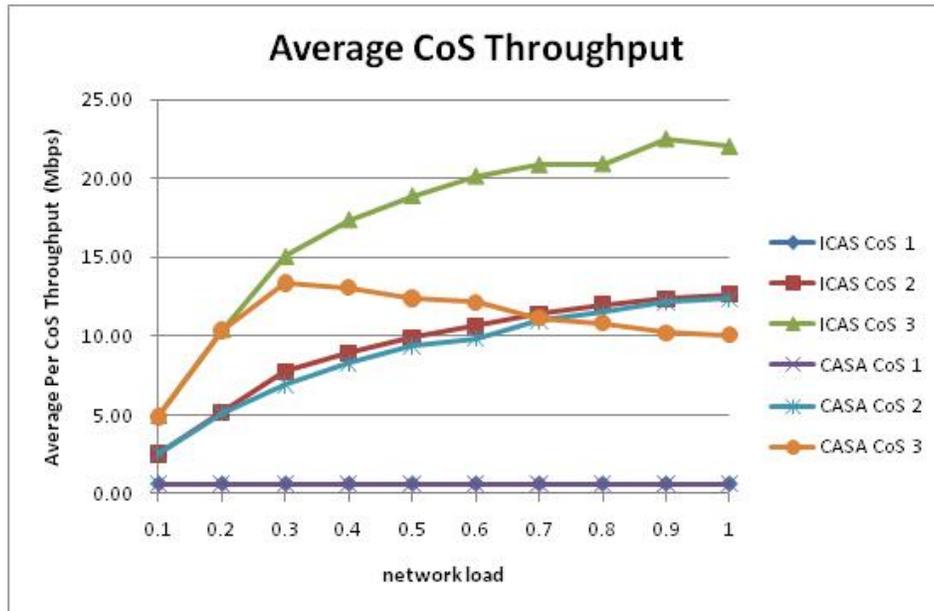


Figure 6.3.4: CASA and ICAS CoS throughput with 0.5 ms TTI

CASA CoS 2 and ICAS CoS 2 throughputs increase. For CoS 3, as the offered load increases ICAS CoS 3 throughput increases. As shown in the figure, the performance of CASA CoS 3 starts deteriorating as the offered load exceeds 0.3. As shown in 6.3.3, after the offered load exceeds 0.3, the throughput of CASA approaches saturation. In order to support the data transmission in CASA CoS 2, the base station will assign more resource blocks to CASA CoS 2. This will reduce the amount of available resource blocks for CASA CoS 3. This causes the deterioration in CASA CoS 3 throughput.

Figure 6.3.5 shows the resource block allocated ratio performance of both CASA and ICAS. As the offered load increases, the resource block allocated ratio of ICAS keeps increasing and reaches saturation after the offered load exceeds 0.7. Under different offered load, CASA has almost the same performance. In some cases, the wideband SRS technique in CASA may reduce the quality of channel condition estimation. This will cause the CASA to use more resource

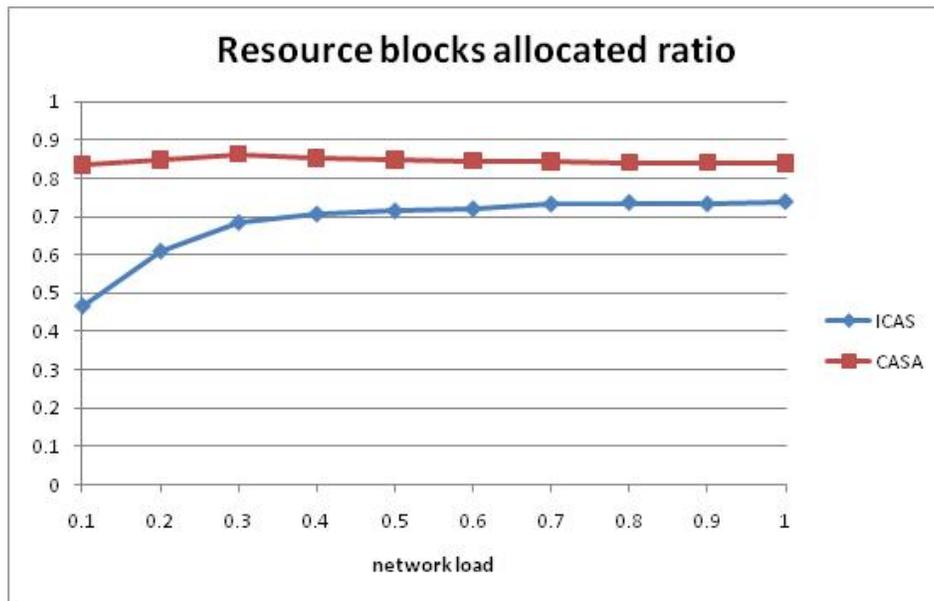


Figure 6.3.5: CASA and ICAS resource block allocated ratio with 0.5 ms TTI

blocks than ICAS to transmit same amount of data.

### 6.3.2 Simulation results with 1.0 ms TTI

In case of 1.0 ms TTI, the minimum resource unit can be allocated to the UEs is a pair of resource blocks. The simulation results of both CASA and ICAS with 1.0 ms TTI will be shown in the Figure 6.3.6 through Figure 6.3.10 below. The simulation results of CoS  $i$  ( $i = 1, 2, 3$ ) is labeled as CASA CoS  $i$  1.0 ms and ICAS CoS  $i$  1.0 ms in figures, respectively.

Figure 6.3.6 shows the average packet delay of the each CoS traffic achieved by CASA and ICAS algorithms when the system running with 1.0 ms TTI. We measure the average delay as a function of the network offered load.

From Figure 6.3.6, we can see that CoS 1 data in both algorithms have almost the same performance; the average delay of CoS 1 increases slightly while the network loads increase. In our simulation, the average delay stays in the interval

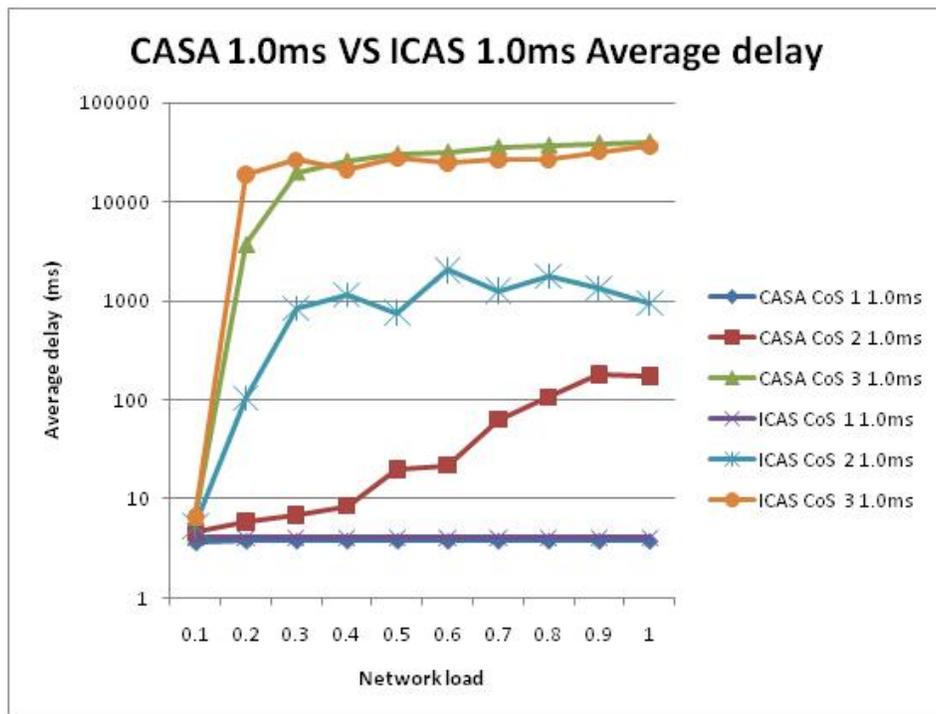


Figure 6.3.6: CASA and ICAS average packet delay with 1.0 ms TTI

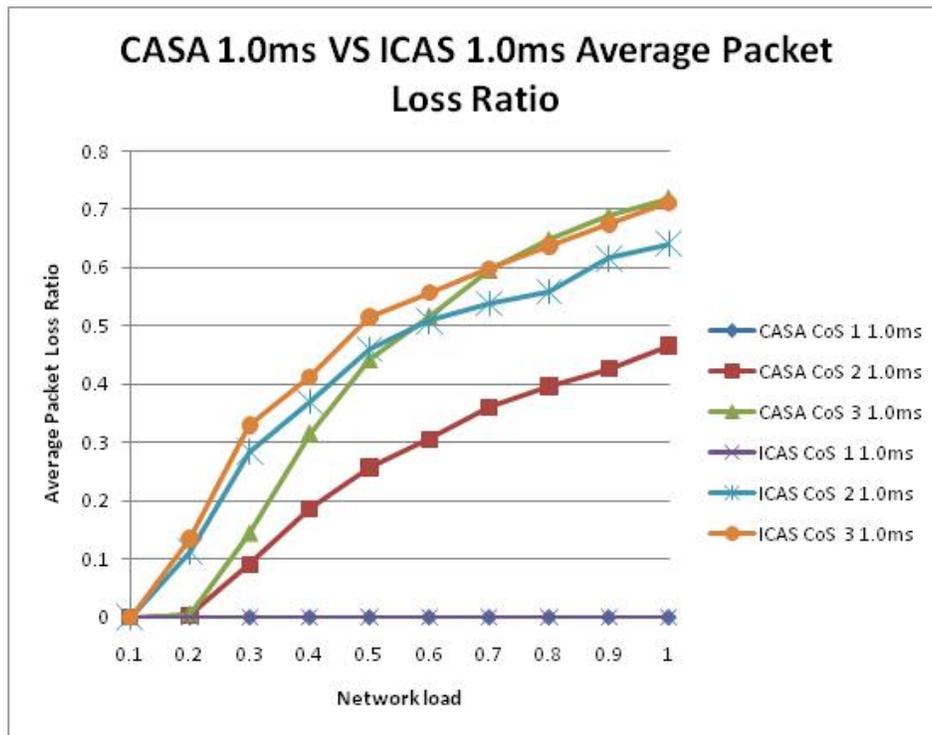


Figure 6.3.7: CASA and ICAS Average Packet Loss Ratio with 1.0 ms TTI

[3.6 ms, 4.0 ms]. The CoS 2 average delay in both CASA and ICAS increases as the network load increases. For CoS 2, the delay performance of CASA is significantly less than ICAS. For CoS 3, ICAS and CASA have almost the same performance.

Figure 6.3.7 shows the average packet loss ratio for both CASA and ICAS with 1.0 ms TTI. For CoS 1, the packet loss ratio of both CASA and ICAS stay zero, which means the VoIP service also will not be degraded by any other CoS transmission with 1.0 ms TTI. For CoS 2, the packet loss ratio of both CASA and ICAS increases as the network load increases. ICAS has higher packet loss ratio than CASA after the offered load exceeds 0.1. CASA CoS 3 1.0 ms has less packet loss ratio than ICAS CoS 3 1.0 ms when the offered load is below 0.7.

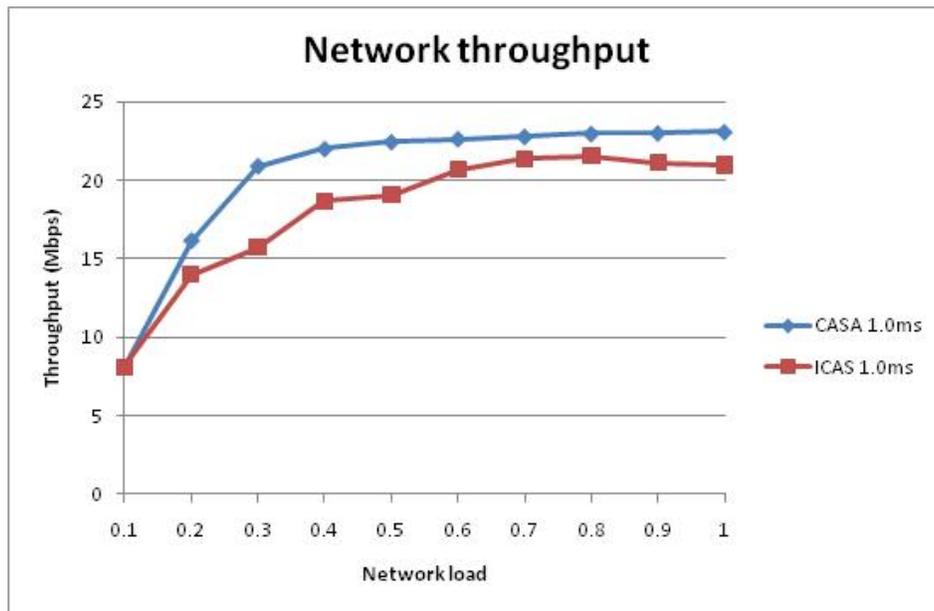


Figure 6.3.8: CASA and ICAS network throughput with 1.0 ms TTI

When the network load is higher than 0.7, the performance of CASA approaches ICAS.

Figure 6.3.8 shows both CASA and ICAS performance in throughput with 1.0 ms TTI. From this figure we can see that CASA has higher throughput than ICAS. As the offered load increases, the maximum throughput of ICAS increases and eventually reaches a maximum value of 21 Mbps. The throughput of CASA however reaches a saturation level of 23 Mbps, when the offered load exceeds 0.4.

Figure 6.3.9 shows the average throughput performance of each CoS for both CASA and ICAS when TTI is set to be 1.0 ms. Both CASA CoS 1 and ICAS CoS 1 throughputs remain constant as the offered load increases. For CoS 2 throughputs, both CASA and ICAS increase as the offered load increases. CASA has higher throughput than ICAS in CoS 2. For CoS 3, as the offered load increases ICAS CoS 3 throughput keep increasing until it reached saturation

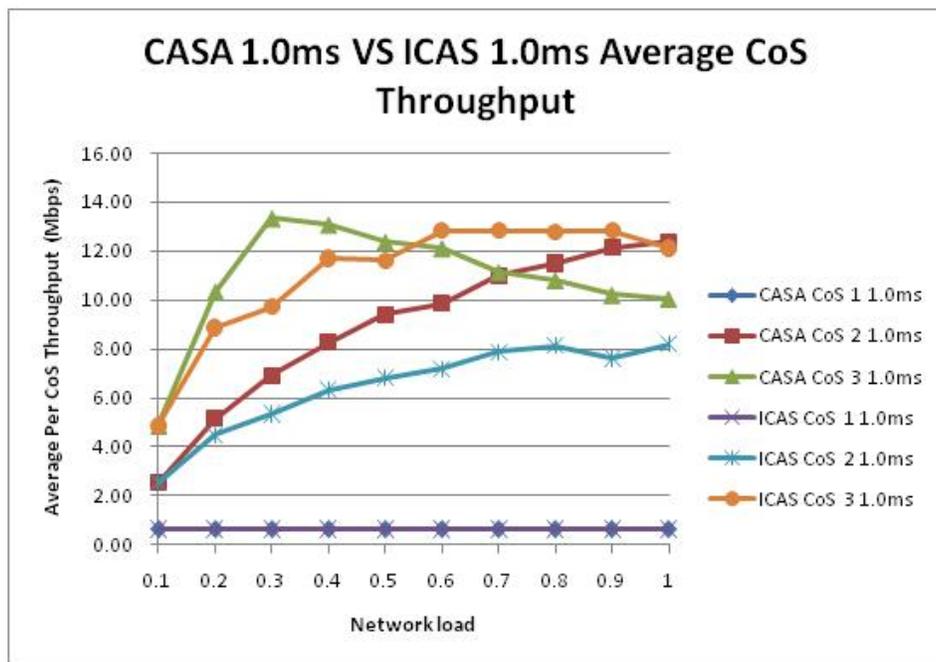


Figure 6.3.9: CASA and ICAS CoS throughput with 1.0 ms TTI

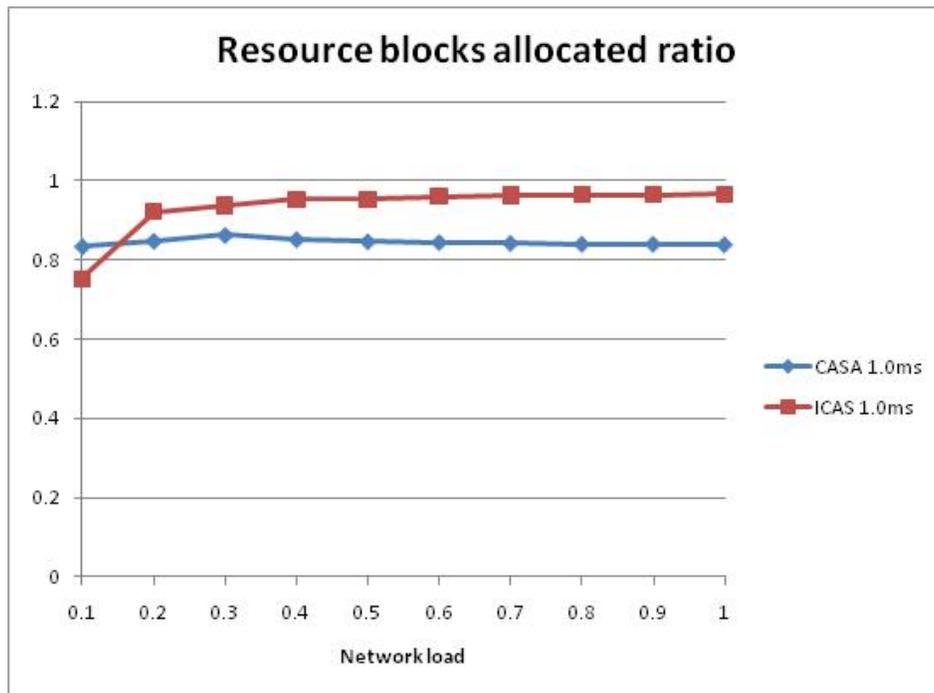


Figure 6.3.10: CASA and ICAS resource allocated ratio with 1.0 ms TTI

when the offered load exceeded 0.6. The performance of CASA in CoS 3 starts deteriorating at an offered load higher than 0.3. As shown in 6.3.8, after the offered load exceeds 0.3, the throughput of CASA reaches saturation. In order to support the data transmission in CASA CoS 2, the base station will assign more resource blocks to CASA CoS 2. This will deduct the amount of available resource blocks for CASA CoS 3. This causes the deterioration in CASA CoS 3 throughput.

Figure 6.3.10 shows the resource block allocated ratio performance of both CASA and ICAS. As the offered load increases, the resource block allocated ratio of ICAS keeps increasing and reaches saturation after the offered load exceeded 0.4. Under different offered load, CASA has almost the same performance.

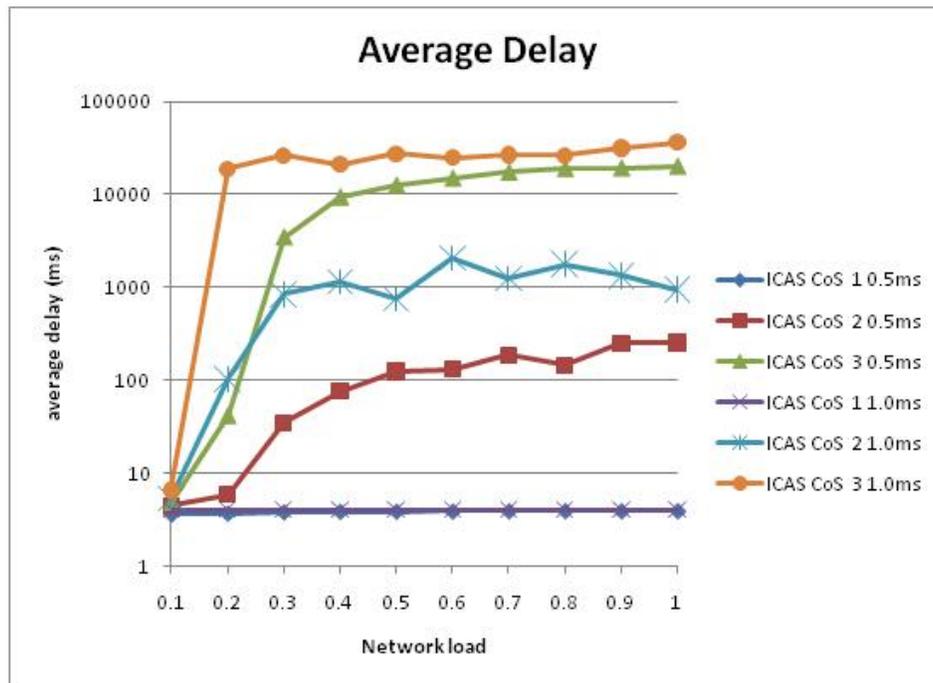


Figure 6.3.11: ICAS Average delay with 0.5 ms and 1.0 ms TTI

### 6.3.3 Simulation results comparison with 0.5 ms TTI and 1.0 ms TTI

In this section we will compare the performance of ICAS and CASA with 0.5 ms TTI and 1.0 ms TTI respectively.

#### 6.3.3.1 The comparison of ICAS performance with 0.5 ms TTI and 1.0 ms TTI

The simulation results of CoS  $i$  ( $i = 1, 2, 3$ ) is labeled as ICAS CoS  $i$  0.5 ms and ICAS CoS  $i$  1.0 ms in figures, respectively.

Figure 6.3.11 shows the average packet delay of the each CoS traffic achieved by the ICAS algorithm with 0.5 ms TTI and 1.0 ms TTI, respectively.

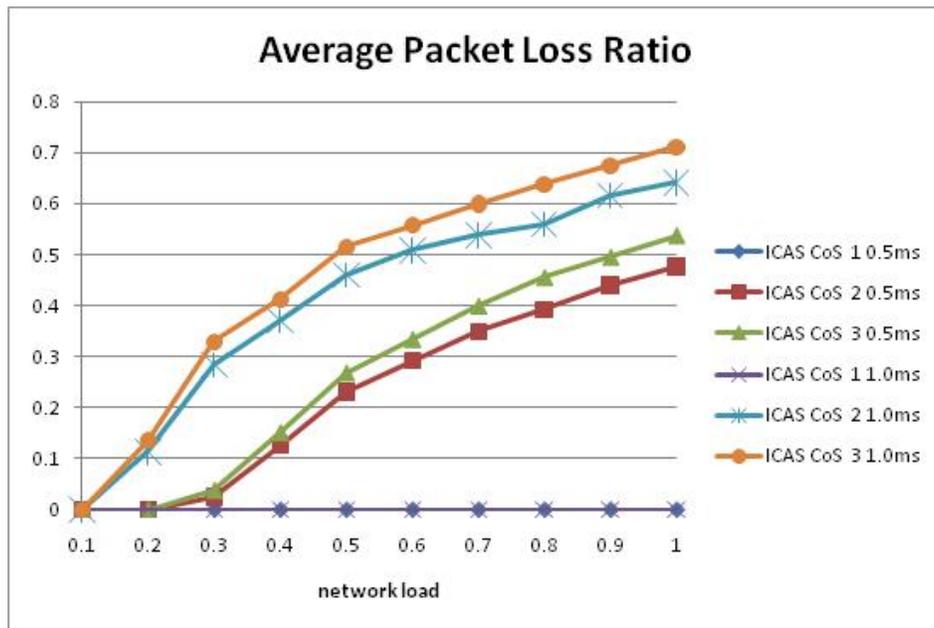


Figure 6.3.12: ICAS Average Packet Loss Ratio with 0.5 ms and 1.0 ms TTI

Looking at Figure 6.3.11, we can see that CoS 1 data of ICAS with both 0.5 ms TTI and 1.0 ms TTI got almost the same performance; the average delay of CoS 1 increases slightly while the network loads increase. In our simulation, the average delay stays in the interval [3.6 ms, 4.0 ms]. The CoS 2 average delay of ICAS in both 0.5 ms TTI and 1.0 ms TTI increases as the network load increases. The CoS 2 delay of ICAS CoS 2 0.5 ms is significantly less than ICAS CoS 2 1.0 ms when the network is loaded. For CoS 3, ICAS CoS 3 0.5 ms has lower packet delay than ICAS CoS 2 1.0 ms. Based on the observation, we can see that the ICAS with 0.5 ms TTI has a much better packet delay performance than with 1.0 ms TTI.

Figure 6.3.12 shows the average packet loss ratio for ICAS with both 0.5 ms TTI and 1.0 ms TTI. For CoS 1, the packet loss ratio of ICAS with both 0.5 ms TTI and 1.0 ms TTI kept at zero, which means the VoIP service of ICAS

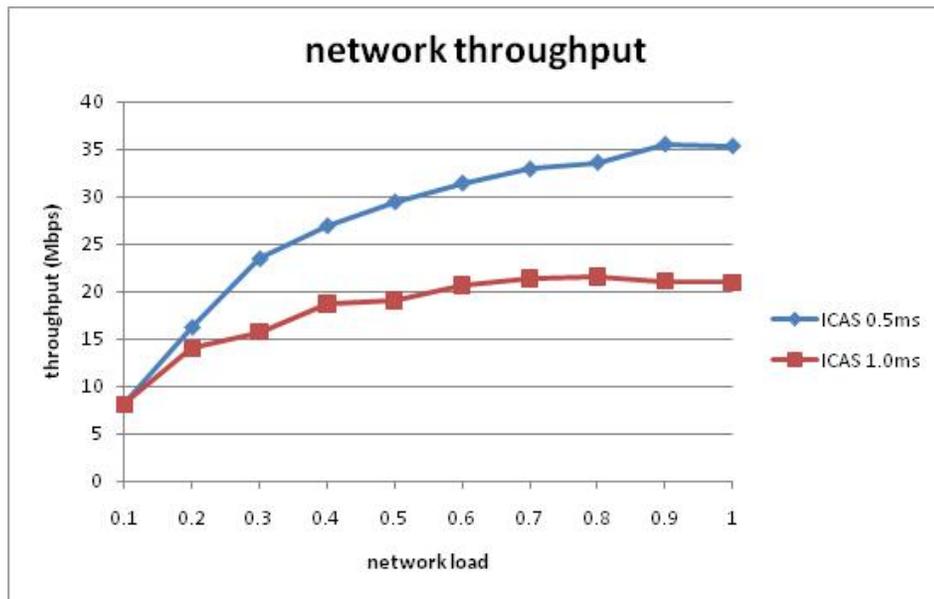


Figure 6.3.13: ICAS network throughput with 0.5 ms and 1.0 ms TTI

can be guaranteed with both 0.5 ms TTI and 1.0 ms TTI. Both CoS 2 and CoS 3 packet loss ratio in ICAS with both 0.5 ms TTI and 1.0 ms TTI increases as the network load increases. For both CoS 2 and CoS 3, ICAS with 0.5 ms TTI has a much better packet loss ratio performance than it with 1.0 ms TTI.

Figure 6.3.13 shows ICAS performance in throughput with both 0.5 ms and 1.0 ms TTI, respectively. From this figure we can see that ICAS with 0.5 ms TTI has higher throughput than with 1.0 ms TTI. When the offered load increases, the throughput of ICAS with 0.5 ms TTI increases and eventually reaches a maximum value of 35 Mbps. When the offered load increases, the throughput of ICAS with 1.0 ms TTI increases and reaches to a saturation value of 21 Mbps.

Figure 6.3.14 shows the average throughput performance of each CoS for ICAS with both 0.5 ms and 1.0 ms TTI. For CoS 1, throughputs of ICAS with both 0.5 ms and 1.0 ms TTI remain constant as the offered load increases. For both CoS 2 and CoS 3, throughputs of ICAS with both 0.5 ms and 1.0 ms

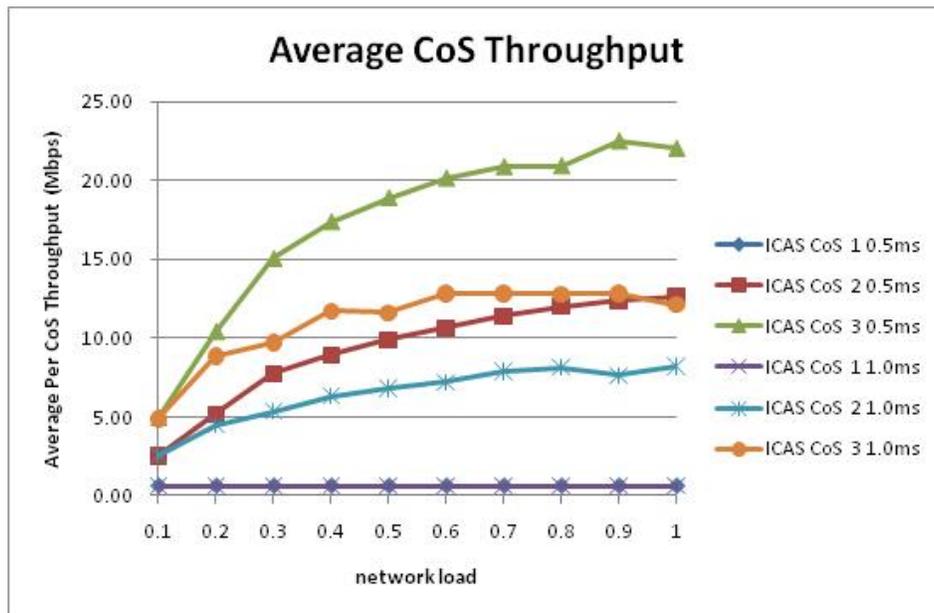


Figure 6.3.14: ICAS CoS throughput with 0.5 ms and 1.0 ms TTI

TTI increase as the offered load increases. ICAS with 0.5 ms TTI has higher throughput than it with 1.0 ms TTI. For CoS 3, as the offered load increases, throughput of ICAS with 1.0 ms TTI also increases and reaches saturation when the offered load exceeded 0.6.

Based on the observations shown above, we can see that the ICAS with 0.5 ms TTI has much better performance in packet delay, packet loss ratio and throughput than it with 1.0 ms. As shown in Figure 6.3.15, if there is a small size of data request to transmit, in the case of 0.5 ms TTI, the ICAS scheduler can only assign a resource block to a UE. But in case of 1.0 ms TTI, the ICAS scheduler has to assign the whole pair of adjacent resource blocks (1.0 ms) to a UE, which will waste more than half of the resource. It reduces the efficiency of ICAS scheduling. This causes the ICAS to have poor performance in the case of 1.0 ms TTI. Also, it makes ICAS allocated more resource to UEs, which cause ICAS with 1.0 ms TTI has higher resource blocks allocated ratio than it with

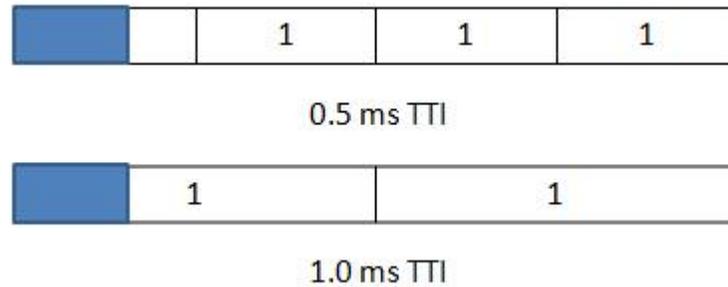


Figure 6.3.15: Example of the resource allocation in case of 0.5 ms TTI and 1.0 ms TTI

0.5 ms TTI. It is shown in Figure 6.3.16.

Figure 6.3.16 shows the resource block allocated ratio performance of ICAS with both 0.5 ms and 1.0 ms TTI. As the offered load increases, the resource block allocated ratio of ICAS with both 0.5 ms and 1.0 ms TTI increases and reaches saturation after the offered load exceeds 0.4. The ICAS with 1.0 ms TTI has higher resource block allocated ratio than it with 0.5 ms TTI.

### 6.3.3.2 The comparison of CASA performance with 0.5 ms TTI and 1.0 ms TTI

The simulation results of CoS  $i$  ( $i = 1, 2, 3$ ) are labeled as CASA CoS  $i$  0.5 ms and CASA CoS  $i$  1.0 ms in the figures, respectively. The performance of CASA with 0.5 ms TTI and 1.0 ms TTI in packet delay, packet loss ratio, throughput and resource blocks allocated ratio is shown in Figure 6.3.17 through to Figure 6.3.21, respectively.

As the Figures shown below, we can see that CASA with 0.5 ms TTI has almost the same performance in packet delay, packet loss ratio, throughput, and resource blocks allocated ratio as 1.0 ms TTI. As we mentioned in Chapter 4, a rearrange buffer is employed in CASA. In the case of 1.0 ms TTI, the minimum

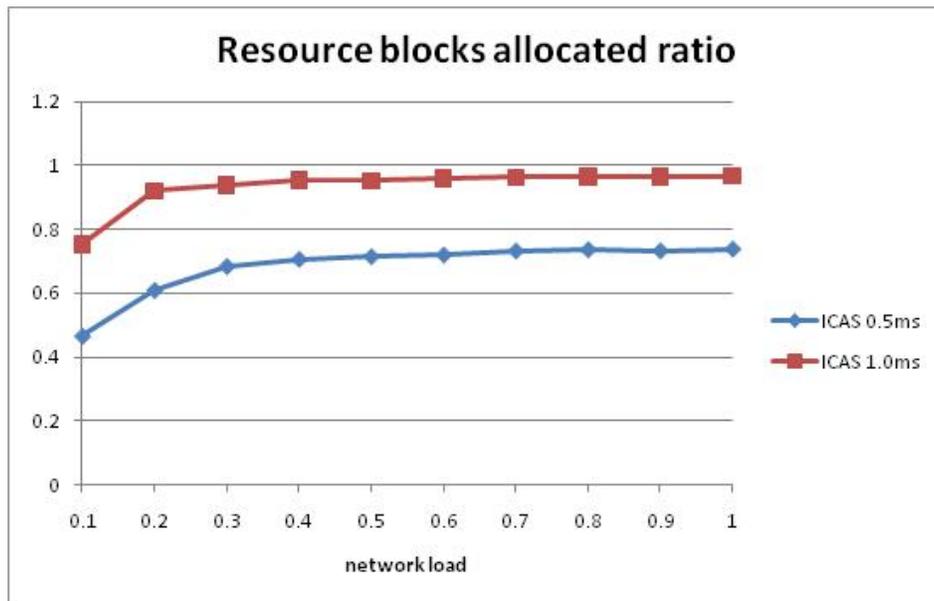


Figure 6.3.16: ICAS resource blocks allocated ratio with 0.5 ms and 1.0 ms TTI

unit of the resource is a pair of resource blocks. Every time after the first round of CASA scheduling processing is finished, the rearrange buffer will help the system rearrange the assigned resources for UEs. For each UE, the rearrange buffer will move the grant into continuous pairs of adjacent resource blocks. It will fill up the pair of resource blocks one by one. It helps CASA make up for the reduced resource usage efficiency caused by the pair of resource blocks, which is shown in Figure 6.3.15. This helps CASA to achieve the same performance with both 0.5 ms TTI and 1.0 ms TTI.

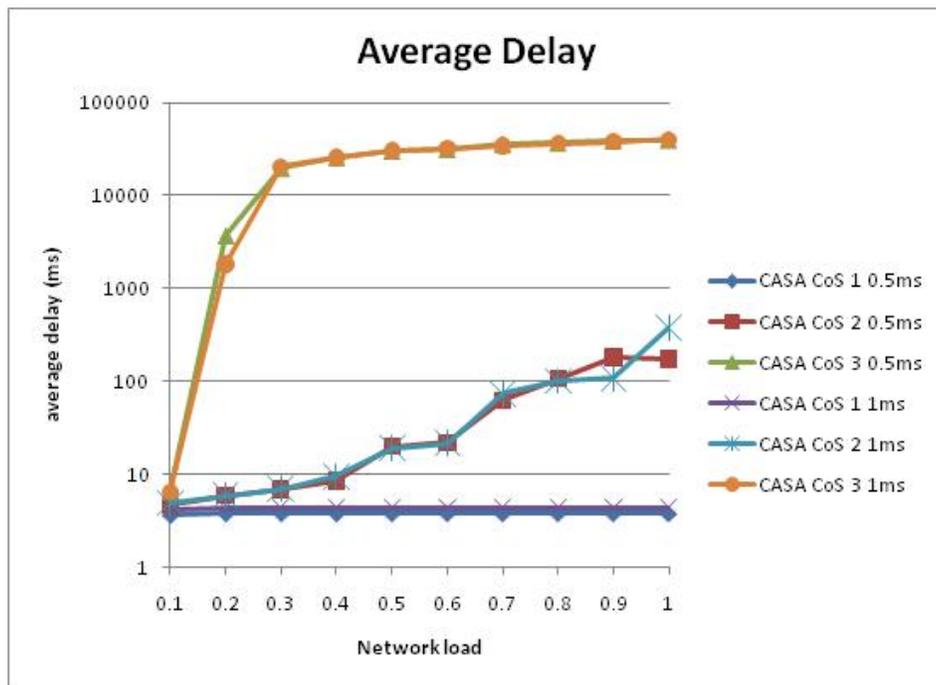


Figure 6.3.17: Comparison of average delay performance of CASA with 0.5 ms and 1.0 ms TTI

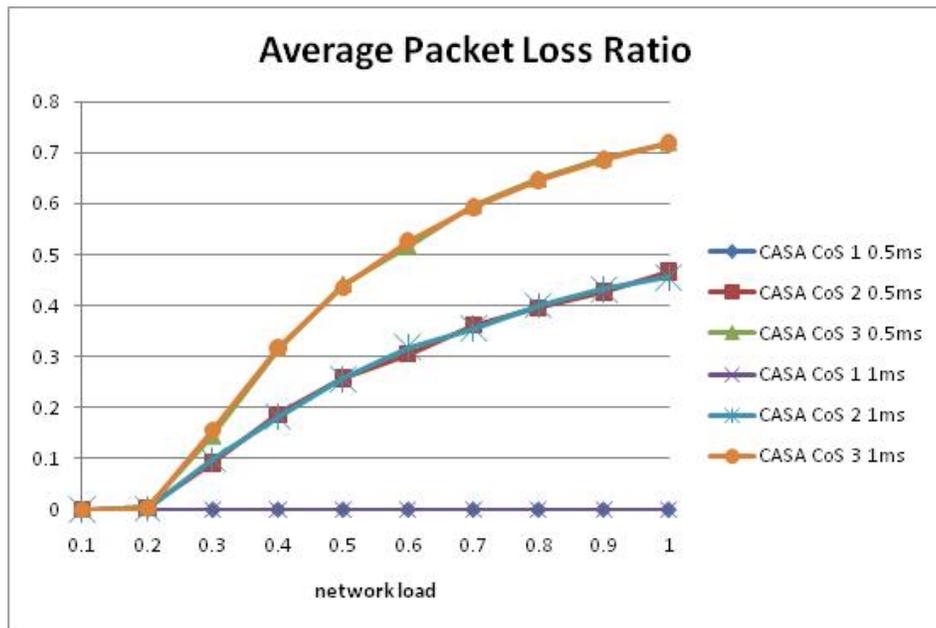


Figure 6.3.18: Comparison of average Packet Loss Ratio of CASA with 0.5 ms and 1.0 ms TTI

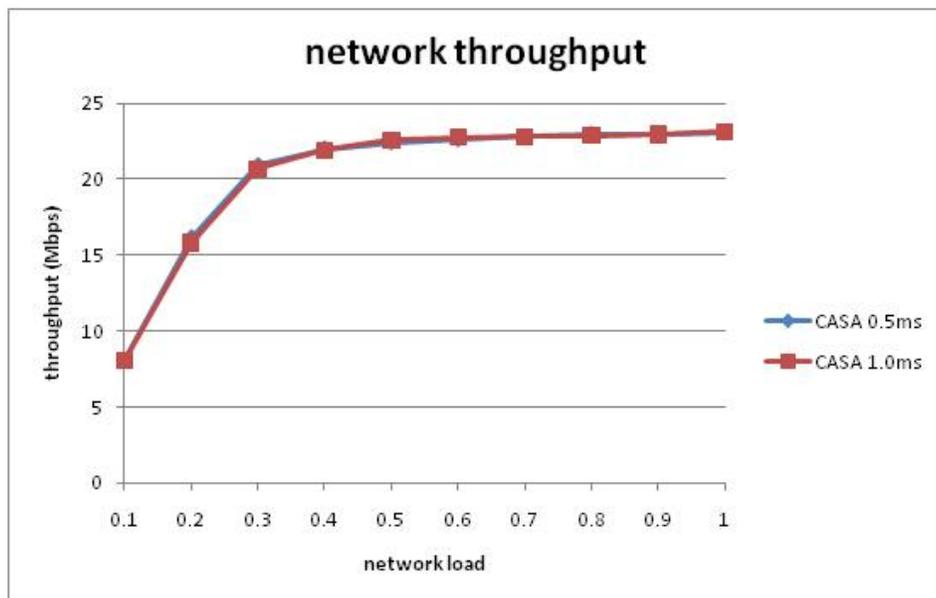


Figure 6.3.19: Comparison of average throughput of CASA with 0.5 ms and 1.0 ms TTI

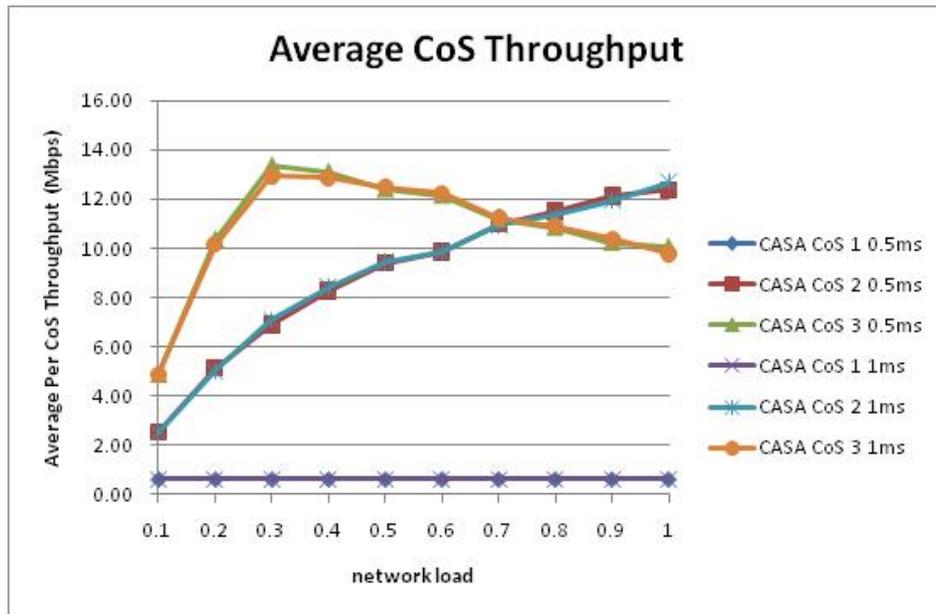


Figure 6.3.20: Comparison of average CoS throughput of CASA with 0.5 ms and 1.0 ms TTI

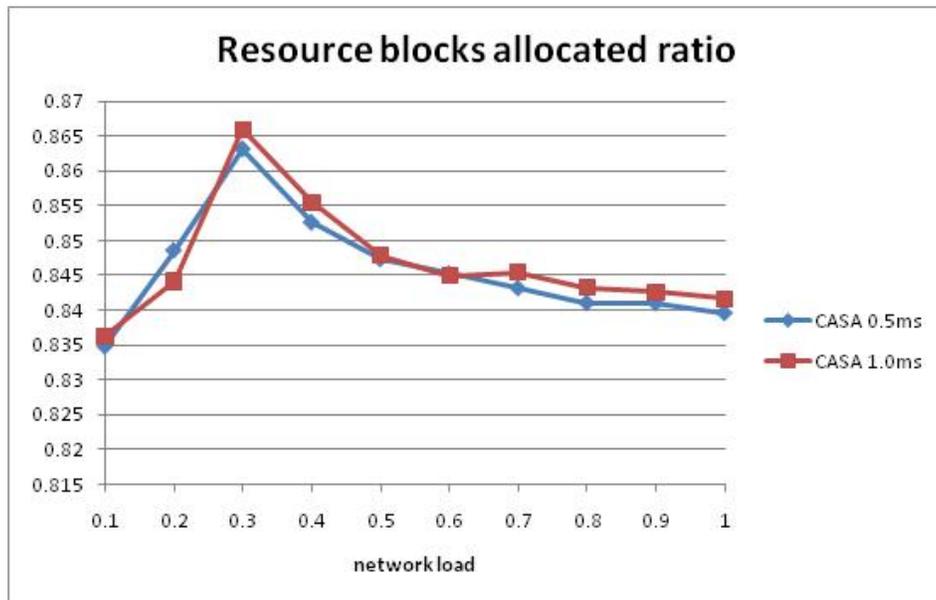


Figure 6.3.21: Comparison of resource blocks allocated ratio of CASA with 0.5 ms and 1.0 ms TTI

## Chapter 7

# Conclusion and Future Works

In this thesis, we proposed two novel resource allocation algorithms for LTE uplink transmission. One is Class-of-service Aware Scheduling Algorithm (CASA) and the other one is Independent Channel Aware Scheduling Algorithm (ICAS). Both of algorithms support Class-of-service (CoS). We tested them with three classes: Expedited Forwarding (EF), Assured Forwarding (AF) and Best Effort (BE). CASA employs the *wideband SRS* technique together with a *credit pooling* technique to distribute resource among customers' classes of service in a fair manner. In CASA, Weighted UE scheduling Order (WUSO) arbitration mechanism is employed to determine the priority of UE in each CoS by its long term transmission record. In ICAS, the *narrowband SRS* technique is employed. Similar to CASA, ICAS also employs credit pooling technique and WUSO arbitration mechanism. Unlike CASA, in ICAS the WUSO arbiter is only used to ensure that the contending UE  $i$  will receive a fair share of the resource. We also account the channel condition in our algorithms based on the Adaptive Modulation and Coding (AMC) scheme.

We have developed an in-house simulation program in C++ to evaluate and

compare the performance of these two algorithms in terms of packet loss ratio, delay, and throughput. We compared the performance of these two algorithms with both 0.5 ms Transmission Time Interval (TTI) and 1.0 ms TTI. In our thesis, the resource is considered as resource block (RB) when TTI is set to 0.5 ms, and as a pair of resource blocks when TTI is set to 1 ms. Under the different simulation conditions (different offered loads), the obtained simulation results showed that the proposed two algorithms can match the Class-of-service requirements of different CoSs. The system supports all three classes traffics simultaneously without impeding the transmission of data in the first class (Expedited Forwarding). Also, both of CASA and ICAS achieved relatively high throughputs in the simulation. Based on the simulation results, we also see that if the power efficiency is an important factor, CASA algorithm can be chosen with 1.0 ms TTI. If the transmission capacity will be considered preferentially, ICAS with 0.5 ms TTI can be employed in the system.

In our works, we chose *Nakagami-m model* and did not consider the efficiency of Sounding Reference Signal. We did not compare our algorithms with other algorithms. In the future, we will run simulation with other kinds of channel models for the LTE uplink transmission. We will also design an efficient channel condition estimation algorithm for the Sounding Reference Signals in the LTE uplink transmission. We will simulate the channel estimation algorithm with the CASA and ICAS together. We will compare the performance of the proposed algorithms with other algorithms to find possible future improvements. We will also consider integrating the LTE system with other access networks, such as Passive Optical Networks together.

# Bibliography

- [1] E. Dahlman et al., *3G Evolution: HSPA and LTE for Mobile Broadband*, 2nd ed., Academic Press, 2008.
- [2] Stefania Sesia, Issam Toufik and Matthew Baker, *LTE-The UMTS Long Term Evolution: From Theory to Practice*, John Wiley & Sons, Ltd., 2009
- [3] Averill M. Law, W. David Kelton, *Simulation Modeling and Analysis*, Third Edition, The McGraw-Hill Companies, 2000.
- [4] 3GPP TS36.300, “Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN): Overall Description.”
- [5] 3GPP TS 36.211, “3rd Generation Partnership Project; Technical Specification Group Radio Access Network; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical Channel and Modulation (Release 8),” v8.9.0, Dec. 2009.
- [6] Hassan Naser, Hussein T. Mouftah, “A Joint-ONU Interval-Based Dynamic Scheduling Algorithm for Ethernet Passive Optical Networks”, *IEEE/ACM Transactions on Networking*, Vol.14, No. 4, August 2006.
- [7] A. Larmo et al., “The LTE link Layer Design,” *IEEE Communication Magazine*, Apr. 2009.

- [8] D. Astély, et al., "LTE-The Evolution of Mobile Broadband," IEEE Communication Magazine, Apr. 2009.
- [9] Bai Zhang, Zhiwei Mao, "Packet Scheduling for AMC-Based OFDMA Wireless Communication Systems", A thesis submitted to the Faculty of Graduate Studies, Lakehead University, July 2008.
- [10] Ruiz de Temiño, L. et al., "Channel-Aware Scheduling Algorithms for SC-FDMA in LTE uplink," in PIMRC 2008, IEEE 19th International Symposium on, pp 1-6.
- [11] Al-Rawi, M. et al., "On the Performance of Heuristic Opportunistic Scheduling in the Uplink of 3G LTE Networks", in IEEE PIMRC 2008, pp 1-6.
- [12] Junsung Lim, H.G. Myung, Kyungjin Oh and D.J. Goodman, "Channel-Dependent Scheduling of Uplink Single Carrier FDMA systems", IEEE Vehicular Technology Conference, Sept. 2006, pp 25-28.
- [13] Hyung G. Myung, Junsung Lim and David J. Goodman, "Single Carrier FDMA for Uplink Wireless Transmission", IEEE Vehicular Technology Magazine, Sept. 2006, pp 30-38.
- [14] T. Kawamura, Y. Kishiyama, A. Morimoto and M. Sawahashi, "Experiments on IP and physical layer packet throughput performance in evolved UTRA uplink using SC-FDMA radio access", IEEE International Symposium on Wireless Communication Systems, Oct. 2008, pp 103-107.
- [15] I.C. Wong, O. Oteri and W. McCoy, "Optimal resource allocation in uplink SC-FDMA systems", IEEE Transactions on Wireless communications, May 2009, Volume 8, Issue 5, pp 2161-2165.

- [16] Yan Meng, Mingli You, Jin Liu and Hanwen Luo, "A Novel Space-Frequency Block Coding Scheme for SC-FDMA," in IEEE 70th Vehicular Technology Conference, Anchorage, AK, 20-23 Sept. 2009, pp 1-5.
- [17] Zihuai Lin, Pei Xiao, B. Vucetic and M. Sellathurai, "Analysis of receiver algorithms for lte LTE SC-FDMA based uplink MIMO systems", IEEE Transactions on Wireless Communications, January 2010, pp 60-65.
- [18] Hyoung-Jin Lim, Tae-Won Yune and Gi-Hong Im, "QoS-Constrained Opportunistic Scheduling for SC-FDMA with Iterative Multiuser Detection", IEEE Communications Letters, January 2009, Volume 13, Issue 1, pp 4-6.
- [19] Xiaolin Hou, Zhan Zhang and H. Kayama, "DMRS Design and Channel Estimation for LTE-Advanced MIMO Uplink", IEEE 70th Vehicular Technology Conference, Anchorage, AK, 20-23 Sept. 2009, pp 1-5.
- [20] F.D. Calabrese, C. Rosa, K.I. Pedersen and P.E. Mogensen, "Performance of proportional fair frequency and time domain scheduling in LTE uplink", European Wireless Conference, Aalborg, 17-20 May 2009, pp 271-275.
- [21] Suk-Bok Lee, I. Pefkianakis, A. Meyerson, Shugong Xu and Songwu Lu, "Proportional Fair Frequency-Domain Packet Scheduling for 3GPP LTE Uplink", IEEE INFOCOM 2009, Rio de Janeiro, 19-25 April 2009, pp 2611-2615.
- [22] Zhengwei Li, Changchuan Yin and Guangxin Yue, "Delay-Bounded Power-Efficient Packet Scheduling for Uplink Systems of LTE", 5th International Conference on Wireless Communications, Networking and Mobile Computing, Beijing, 24-26 Sept. 2009, pp 1-4.
- [23] F.D. Calabrese, M. Anas, C. Rosa, P.E. Mogensen and K.I. Pedersen, "Performance of a Radio Resource Allocation Algorithm for UTRAN LTE Up-

- link”, IEEE 65th Vehicular Technology Conference, Dublin, 22-25 April 2007, pp 2895-2899.
- [24] Zhijie Wang, Yafeng Wang, Dajie Jiang, Chunchang Tian and Dacheng Yang, “Scheduling and Link Adaptations for VoIP in TDD-LTE Uplink”, 5th International Conference on Wireless Communications, Networking and Mobile Computing, Beijing, 24-26 Sept. 2009, pp 1-5.
- [25] Hongkun Yang, Fengyuan Ren, Chuang Lin and Jiao Zhang, “Frequency-Domain Packet Scheduling for 3GPP LTE Uplink”, 2010 Proceedings IEEE INFOCOM, San Diego, CA, USA, 14-19 March 2010, pp 1-9.
- [26] Yafeng Wang, Aoyang Zheng, Junfeng Zhang and Dacheng Yang, “A novel channel estimation algorithm for sounding reference signal in LTE uplink transmission”, IEEE International Conference on Communications Technology and Applications, Beijing, 16-18 Oct. 2009, pp 412-415.
- [27] Meili Zhou, Bin Jiang, Wen Zhong and Xiqi Gao, “Efficient channel estimation for LTE uplink”, International Conference on Wireless Communications & Signal Processing, Nanjing, 13-15 Nov. 2009, pp 1-5.
- [28] F.D. Calabrese, C. Rosa, M. Anas, P.H. Michaelsen, K.I. Pedersen and P.E. Mogensen, “Adaptive Transmission Bandwidth Based Packet Scheduling for LTE Uplink”, IEEE 68th Vehicular Technology Conference, Calgary, 21-24 Sept. 2008, pp 1-5.
- [29] P. Kysti et al., “Winner II Channel Models”, IST-4-027756 Winner II Deliverable 1.1.2 V 1.0, Sept. 2007.
- [30] S. Floyd and V. Jacobson, “Random early detection gateways for congestion avoidance,” IEEE/ACM Trans. Netw., vol. 1, no. 4, pp. 397–413, Aug. 1993.

- [31] F.D. Calabrese et al., “Search-Tree Based Uplink Channel Aware Packet Scheduling for UTRAN LTE,” IEEE Proc. of the 67th Veh. Tech. Conf., May 2008.
- [32] Q. Liu, X. Wang and G.B. Giannakis, “A Cross-Layer Scheduling Algorithm With QoS Support in Wireless Networks,” IEEE Trans. Veh. Technol., vol. 55, no. 3, pp. 1-48, 2001
- [33] Pareto, Vilfredo, Cours d’Économie Politique: Nouvelle édition par G.-H. Bousquet et G. Busino, Librairie Droz, Geneva, 1964, pages 299–345.
- [34] Ericsson to make World-first demonstration of end-to-end LTE call on handheld devices at Mobile World Congress, Barcelona, 2008
- [35] Z. Kang, K. Yao, and F. Lorenzelli, “Nakagami-m fading modeling in the frequency domain for OFDM system analysis,” IEEE Commun. Lett., vol. 7, no. 10, pp. 484–486, Oct. 2003.
- [36] LTE – an introduction, Ericsson Inc Whitepapers, [Online]. Available: [http://www.ericsson.com/res/docs/whitepapers/lte\\_overview.pdf](http://www.ericsson.com/res/docs/whitepapers/lte_overview.pdf), 2009.
- [37] Service Level Agreements, IT-Tude.com, [Online]. Available: <http://www.it-tude.com/sla-article.html>