Analytical Modeling of HSUPA-Enabled UMTS Networks for Capacity Planning

A dissertation
submitted to the School of Information Technologies
at the University of Sydney
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

Tuo Liu
October 2008

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

(David Everitt) Principal Adviser

Approved for the University Committee on Graduate Studies.

Abstract

In recent years, mobile communication networks have experienced significant evolution. The 3rd Generation (3G) mobile communication system, Universal Mobile Telecommunication System (UMTS), employs Wideband Code Division Multiple Access (WCDMA) as the air interface standard, which leads to quite different mobile network planning and dimensioning processes compared with 2nd Generation (2G) systems. The UMTS system capacity is limited by the received interference at NodeBs due to the unique features of WCDMA, which is denoted as 'soft capacity'. Consequently, the key challenge in UMTS radio network planning has been shifted from channel allocation in the channelized 2G systems to blocking and outage probabilities computation under the 'cell breathing' effects which are due to the relationship between network coverage and capacity. The interference characterization, especially for the other-cell interference, is one of the most important components in 3G mobile networks planning.

This monograph firstly investigates the system behavior in the operation of UMTS uplink, and develops the analytic techniques to model interference and system load as fully-characterized random variables, which can be directly applicable to the performance modeling of such networks. When the analysis progresses from single-cell scenario to multi-cell scenario, as the target Signal-to-Interference-Ratio (SIR) oriented power control mechanism is employed for maximum capacity, more sophisticated system operation, 'feedback behavior', has emerged, as the interference levels at different cells depend on each other. Such behaviors are also captured into the constructed interference model by iterative and approximation approaches.

The models are then extended to cater for the features of the newly introduced

High Speed Uplink Packet Access (HSUPA), which provides enhanced dedicated channels for the packet switched data services such that much higher bandwidth can be achieved for best-effort elastic traffic, which allows network operators to cope with the coexistence of both circuit-switched and packet-switched traffic and guarantee the QoS requirements. During the derivation, we consider various propagation models, traffic models, resource allocation schemes for many possible scenarios, each of which may lead to different analytical models. All the suggested models are validated with either Monte-Carlo simulations or discrete event simulations, where excellent matches between results are always achieved.

Furthermore, this monograph studies the optimization-based resource allocation strategies in the UMTS uplink with integrated QoS/best-effort traffic. Optimization techniques, both linear-programming based and non-linear-programming based, are used to determine how much resource should be assigned to each enhanced uplink user in the multi-cell environment where each NodeB possesses full knowledge of the whole network. The system performance under such resource allocation schemes are analyzed and compared via Monte-Carlo simulations, which verifies that the proposed framework may serve as a good estimation and optimal reference to study how systems perform for network operators.

Acknowledgements

My PhD study is financially supported by the University of Sydney as well as the Smart Internet Technology CRC, to which I would show my greatest appreciation. These generous scholarships not only help me concentrate on my research, but also provide me valuable opportunities for wonderful research visiting experience.

There are many people I would like to thank for their help during the last a few years. Without their contribution and support this thesis would have never been possible.

First of all I am deeply grateful to my supervisor, Prof. David Everitt for his guidance and wise advice during my whole PhD work. David introduced me into the world of mobile networks, handed me a great research topic and kept providing insightful ideas for the research direction. I would especially thank him for his constant encouragement of my efforts.

I am also sincerely thankful to Dr. Dirk Staehle and Andreas Mäder, who I worked with during my research visit at the University of Würzburg, Germany. I definitely consider the joint research in Germany as the highlight of my PhD study, from both professional and personal points of view. I would also thank Prof. Phuoc Tran-Gia for offering me such a great visiting opportunity. I would also like to express my appreciation to all the group members in i3 for having so much fun together, and for their efforts in usage of English whenever I was present.

Many thanks to Dr. Björn Landfeldt for being my associate supervisor and to Prof. Albert Zomaya for providing such an excellent research environment in the advanced network research group. I would also like to express my gratitude to all the other group members, in particular Young Choon Lee, Penghao Wang, Weisheng Si, Riky Subrata, Tanveer Zia, Mohsin Iftikhar and Wibowo Hardjawana for constant

discussions on everything except research.

Finally, I would thank my parents for their unconditional love and care even though they are not in Australia. And foremost, I would like to thank my girlfriend, who shared all the happiness and depression with me, and kept supporting me at all times. Moreover, according to what she always told me, she decided to do her PhD study in order to have more 'common topics' with me. Thanks for her continuous love. This monograph is dedicated to her.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Problem Statement and Motivation

In recent years, mobile communication networks have experienced significant evolution. Unlike 2nd Generation (2G) networks which support only voice telephony and short message services, 3rd Generation (3G) mobile communication systems make it possible to bring the available services in the Internet, such as WWW browsing or file sharing, into mobile networks. The Universal Mobile Telecommunication System (UMTS) is the European standard of 3G mobile communication system with Wideband Code Division Multiple Access (WCDMA) employed as its air interface. It offers much higher data rates compared with 2G communication systems, such as Global System for Mobile Communication (GSM) and Interim Standard 95 (IS-95), and this is the foundation for the mobile users to enjoy various multimedia applications besides the traditional telephony service. In particular, the newly introduced High Speed Downlink Packet Access (HSDPA) and High Speed Uplink Packet Access (HSUPA) (also known as Enhanced Uplink) provide dedicated channels in order to cater for the growing demands of packet-switched data transmission, such that much higher data rate and superior performance can be achieved for best-effort elastic traffic. This enables the mobile network operators to cope with the coexistence of both circuit-switched and packet-switched traffic and guarantee the Quality of Service (QoS) requirements.

Network planning and dimensioning are essential steps in the deployment of mobile

networks, where proper planning strategies should lead to superior capacity which translates to more revenue, while maintaining high service availability from both spatial and temporal points of view at minimum infrastructure costs. From the mobile users' side, they are expected to enjoy various multimedia services 'anywhere and anytime', which implies the demands in both coverage and capacity. In more technical words, the network operators aim to search for the best configuration with minimum number of base stations over the planning area, to fulfill the capacity requirements in terms of blocking probability and outage probability, coverage probability as well as QoS of mobile users. The most crucial issue in the 2G GSM radio network planning is frequency reuse as the user channels are divided by frequency bands and time slots. In the 3G UMTS mobile network, larger capacity is gained with WCDMA as its air interface, which also implicates new paradigms in the radio network planning process. The frequency reuse issue can be put aside in WCDMA systems as the same frequency band can be used throughout the network, but the 'soft capacity' feature in the interference-limited WCDMA systems shifts the key challenge in UMTS radio network planning to blocking and outage probability computation under the 'cell breathing' effects which are caused by the dependence among the coverage area, network capacity and the user QoS.

To find the optimal configuration of a certain network, the most straightforward way is based on the 'trial and error' method, which consists of an iterative process of adjusting the network dimension and system performance evaluation on the given network for the feasibility evaluation. The simplest and most flexible approach for the performance evaluation is generally through Monte Carlo simulations, however the major drawback is that it is too time-consuming.

In this thesis, we investigate some analytic methods to analyze the performance of given network configurations to a similar degree of accuracy as the simulations, but on a much more efficient basis. In UMTS networks, the interference characterization, especially for the interference contributed by the users located in the neighboring cells (known as other-cell interference), is the key component for such analysis. Once the interference factor is well modeled analytically, each time with an input network configuration composed of the number of NodeBs and their layout as well as the traffic distribution, the feasibility of the input network, specified as whether capacity and

coverage can meet the demands or not, can be determined instantly. The operators can benefit much from it in terms of efficiency, since the computation task during the capacity planning process is significantly reduced whenever a new WCDMA network is to be deployed.

Besides the benefits in the mobile network deployment process, the analytical models can also contribute to the network maintenance for the operators. For instance, for the evaluation prior to the introduction of new data services by the network provider, or, for any change which may impact the traffic demand patterns (e.g. World Youth Day 2008 where more than 400,000 people came to Sydney staying for a week to attend), the network performance under new conditions can be analyzed more efficiently by the proposed analytical models.

1.2 Summary of Contributions

The first major contribution of this monograph is to derive the analytical models of some key system performance parameters in the UMTS uplinks with QoS traffic, such as received interference power at NodeBs, transmit power from mobile users, user throughput, system blocking and outage probability, etc. Among these factors, the other-cell interference is the most important and the most difficult one to characterize, and thus much attention has been directed towards analysis of it, and the other factors can be derived straightforwardly once the other-cell interference is modeled. An iterative method is presented to solve the constructed fixed-point equations with respect to the other-cell interference, followed by a log-normal approximation technique in order to simplify the computational complexity, such that the approximation model can be easily employed by the network planning software products. The accuracy of the proposed analytical models are validated through comparison with Monte Carlo and discrete event simulations, where excellent matches in the results are achieved. Therefore, such analytical approximation modeling techniques for the interference variables in UMTS networks can be directly applied into capacity planning, where the operators can benefit much since the computation task during the planning and dimensioning process could be considerably reduced each time a new UMTS network is to be deployed. This contribution is to be discussed in Chapter 3.

The second major contribution is to extend the above models to be applicable to the newly introduced Enhanced Uplink in UMTS networks, which supports both QoS traffic and best-effort packet data traffic by allocating different dedicated channels. Besides the user categories, the extended models consider many other scenarios, such as heterogenous services classes, distance loss based or log-normal shadowing based propagation models, time based or volume based traffic models, power oriented or Signal-to-Interference-Ratio (SIR) oriented power control schemes, pole-capacity based or target load based admission control policies, etc. The analytical approximation models for these different scenarios are presented and validated by simulations, and again the results indicate the suggested models can be directly applied into the radio network capacity planning. This contribution is to be discussed in Chapter 4.

The third major contribution of this monograph is to present interference and load models for the HSUPA-enabled UMTS uplink under optimized resource allocation strategies with both QoS traffic and best-effort traffic. Such optimization-based algorithms may achieve optimal resource assignment in the multi-cell environment if each NodeB has full knowledge of network loads, but at the price of high computational intensity, therefore the proposed framework can assist the network operators to analyze the system performance as an optimal reference. The network performance is evaluated by Monte-Carlo simulations and demonstrated by some key system parameters, as well as by the feasible load regions. The impact of 'down grants', which is a new feature in HSUPA, is also studied and illustrated. This part of work is jointly contributed by Dirk Staehle and Andreas Maeder from University of Würzburg, Germany, where the feasible load regions comparison in Section 5.4 and the results (from Fig. 5.10 to Fig. 5.13) are mainly contributed by Andreas. This contribution is to be discussed in Chapter 5.

1.3 Thesis Outline

We begin in Chapter 2 with an overview of popular mobile communication systems in chronological order, followed by some detailed background knowledge in UMTS including network architecture, available services, Radio Resource Management (RRM) schemes and principles of its air interface - WCDMA.

In Chapter 3, we develop efficient analytic techniques for characterization of other-cell interference in the uplink of WCDMA cellular networks, and later the system outage probability. We first describe the propagation model including both distance loss and log-normal shadowing effects, power control mechanism employed in the capacity model, and the time-based traffic model with heterogeneous service classes of different data rates, target bit error rates and voice activity factors. Then the Monte-Carlo simulation approach is presented to understand the system behavior, followed by the analytic methods based on an iterative process of solving fixed-point equations to characterize the distribution function of the other-cell interference, and also with a corresponding log-normal approximation to simplify the computation. Then, one important application using the derived distribution of other-cell interference, the calculation of outage probability, is demonstrated. Finally, the suggested analytical model is validated through numerical results.

Chapter 4 extends the analytical model obtained in Chapter 3 for the other-cell interference in the multi-cell UMTS with the recent addition to the standards enabled, which is the enhanced uplink for high speed packet data access of best-effort traffic in the uplink direction. The feature of variable data rate for best-effort services is considered in this model, thus the resultant 'greedy' radio resource allocation scheme is first introduced in this chapter. The volume-based traffic model for the best-effort traffic is described next. By a modified iterative approach, the moments in the log-normal approximations are determined, and these are later justified by comparison with the results from discrete event simulations. The accuracy of the approximation can be further enhanced and applied to outage probability calculation.

In Chapter 5, we present an optimization-based centralized resource allocation framework in the multi-cell enhanced uplink. The principle of centralized resource management is given first, followed by a description of the potential unnecessary outage problem caused by centralized schemes. To address this problem, a few optimization-based strategies are introduced, followed by a derivation of the corresponding interference and load models to characterize the system behavior and evaluate the system performance. The 'down grants' feature of HSUPA, which is to suppress the interference from soft handover is also included in the analysis. The feasible load regions and other performance metrics under each scheme are demonstrated

and compared, demonstrating that the proposed framework may serve as a good estimation and reference for the network operators to study how systems perform under such resource allocation mechanisms.

Chapter 6 concludes this monograph by summarizing the completed work, and providing an outlook into the future research possibilities.

1.4 List of Publications

The following consists of a list of publications related to this monograph. They are classified based on the contributions, which form corresponding chapters in this monograph.

Chapter 3: Modeling of UMTS Networks with QoS Traffic

- Tuo Liu and David Everitt, "Interference and Outage Probability Evaluation in UMTS Network Capacity Planning," in *European Transactions on Telecommunications*, **JOURNAL PAPER**, Vol.18, No.6, 2007.

- Tuo Liu and David Everitt, "Analytical Approximation of Other-cell Interference in the Uplink of CDMA Cellular Systems", in *Proc. 63rd IEEE Vehicular Technology Conference* **VTC2006-S**, Melbourne, Australia, May, 2006.

- Tuo Liu and David Everitt, "Other-cell Interference Characterization in the UMTS Systems with Shadowing Effect", in *Proc. 12th European Wireless Conference* **EW2006**, Athens, Greece, Apr., 2006.

Chapter 4: Modeling of UMTS Enhanced Uplink with Best-Effort Traffic

- Tuo Liu, Andreas Mäder, Dirk Staehle, Phuoc Tran-Gia and David Everitt, "Analytic Modeling of the UMTS Enhanced Uplink with Best-Effort Traffic," accepted by *Journal of Networks*, **JOURNAL PAPER**, in press.

- Tuo Liu, Andreas Mäder, Dirk Staehle and David Everitt, "Analytic Modeling of the UMTS Enhanced Uplink in Multi-cell Environments with Volume-Based Best-Effort Traffic," in *Proc. 7th IEEE International Symposium on Communications and Information Technologies* **ISCIT2007**, Sydney, Australia, Oct., 2007.

- Tuo Liu, Andreas Mäder and Dirk Staehle, "Analytical Other-cell Interference Characterization over HSUPA-Enabled Multi-cell UMTS Networks," in *Proc. 66th IEEE Vehicular Technology Conference* **VTC2007-F**, Baltimore, USA, Oct., 2007.

Chapter 5: Optimization-Based Resource Allocation Strategies

- Tuo Liu, Andreas Mäder and Dirk Staehle, "A Novel Linear-Programming Based Approach for Near-Optimal Rate Allocation Computation in the HSUPA-Enabled UMTS," in *Proc. 3rd IEEE International Conference on Wireless and Mobile Computing, Networking and Communications* **WiMob2007**, New York, USA, Oct., 2007.

- Andreas Mäder, Dirk Staehle, Tuo Liu and Hans Barth, "Feasible Load Regions for Different RRM Strategies for the Enhanced Uplink in UMTS Networks," in *Springer Lecture Notes in Computer Science, Wireless Systems and Mobility in Next Generation Internet*, **LNCS** Volume 4369, Jan., 2007.

Chapter 2

The Principles of the Universal Mobile Telecommunication System

This chapter gives an introduction to a few well-known mobile communication systems and later focusses on the UMTS systems. We first briefly describe the mobile network evolution history over the last few decades, and then present the services and applications in UMTS and its network architecture. The operational principles and features of WCDMA which is the air interface of UMTS are discussed, followed by a brief introduction on all the functional modules for radio resource management in UMTS.

2.1 Mobile Network Evolution

Mobile communication is not a recent technology, but it keeps evolving rapidly, especially in the past few decades since the cellular concept and frequency reuse techniques were introduced. The first US cellular telephone system named the Advanced Mobile Phone Service (AMPS) [72] was developed by AT&T Bell Laboratories in the late 1970s. It is classified as the first generation communication system where analog technology is implemented. Frequency Modulation (FM), Frequency Division Duplex (FDD) and Frequency Division Multiple Access (FDMA) are employed for voice modulation, duplexing and multiple access in the AMPS air interface, respectively. Thus the whole allocated spectrum is divided into multiple frequency bands,

which can serve as voice channels or control channels. A seven-cell frequency reuse pattern is used in the system with provisions of sectoring and cell splitting to increase capacity if required. To maintain the AMPS subscribers' service quality, especially in a heavily populated area, is a particularly hard task for the network operators due to tremendous system complexity and lack of control. Other first generation mobile networks include European Total Access Communication System (ETACS) deployed in Europe which is virtually identical to the AMPS except for some minor differences in frequency bandwidth, as well as the Nordic Mobile Telephone (NMT) system, which operates in a unified way in whole Scandinavia.

In the early 1990s, as the techniques in the first generation analog system were not able to satisfy the growing demand for capacity, cellular systems using digital modulation techniques emerged. These digital systems offer large improvement in capacity and system performance [53], have replaced first generation networks, reached massive market gradually and remain the most ubiquitous cellular networks in today's world. Unlike the first generation networks that rely exclusively on FDMA/FDD and analog FM, the 2G systems conform to the standards which use digital modulation formats and TDMA/FDD and CDMA/FDD multiple access techniques.

The GSM [51] is the 2G mobile network widely deployed in Europe, which later gained worldwide acceptance and became the world's most popular mobile technology. FDD and a combination of Time Division Multiple Access (TDMA) and Frequency Hopped Multiple Access (FHMA) schemes are employed for multiple access to mobile users, while Gaussian Minimum Shift Keying (GMSK) is used for digital modulation. Each frequency channel is further divided into eight time slots so that they can be shared by several subscribers in order to achieve higher capacity.

IS-95 (also known as cdmaOne), one of the 2G cellular systems in the U.S., is based on Code Division Multiple Access (CDMA) developed by Qualcomm. With CDMA air interface, the system can accommodate variable number of users in one frequency band with wider bandwidth using direct sequence spread spectrum. Although the CDMA systems may experience much higher co-channel interference level than FDMA or TDMA systems due to multiple users occupying a same frequency channel, their inherent interference resistance ability allows the systems to operate properly with very low SIR. A promised capacity increase in such system is verified in [26] with

power control employed.

With more advanced technology, the 2G mobile networks serve as a high capacity replacement for the older first generation cellular mobile communication systems. They are able to support conventional telephony services as well as the short messaging service. However, as the standards for 2G mobile networks were mostly specified before the prevailing use of the Internet, these systems are usually designed for circuit-switched data only and have very limited throughput for packet-switched data, which cannot satisfy the increasing demands for more and more popular services like Internet browsing, email, etc. In order to support these modern Internet applications, some new standards have been developed on the basis of 2G technologies, such that the existing 2G equipment can be upgraded with higher data rate transmissions, while still operating on the same carriers. They are generally categorized as 2.5G mobile communication systems, among which General Packet Radio Service (GPRS) and Enhanced Data Rates for GSM Evolution (EDGE) are some notable technologies.

The 2.5G technologies only serve as an interim data solution for the exploding Internet services. The eventual 3G systems provide much higher data rates, much larger system capacity, as well as much more services. The 3G technologies emerge with the International Mobile Telephone 2000 (IMT-2000) plan suggested by the International Telecommunications Union (ITU), which aims to implement a global ubiquitous mobile communication standard based on wideband-CDMA throughout the world. However, as different 2G systems were already deployed in Europe and U.S., the 3G evolution path diverges as well in order to be backward compatible, and accordingly forms two major camps. The GSM systems in Europe lead to UMTS which is standardized by the 3rd Generation Partnership Project (3GPP) community, while in U.S., with the existing 2G CDMA systems, the standardization activities of cdma2000 are organized in the 3rd Generation Partnership Project 2 (3GPP2) association. In this thesis, as we are focused on the UMTS networks, all the releases and protocols mentioned in the later refer to those in the 3GPP only.

Voice conversation is the most popular service at the start of the 3G era, but later on the share of data traffic increases considerably, and has the trend to be the dominant source of traffic volume since eventually the conversational services can also

be delivered as packet data if sufficient QoS is being supported. Therefore, increasing the packet data throughput is the next challenge for mobile network evolution. By the introduction of HSDPA [3] in Release 5, the packet throughput is boosted tremendously for the increasing bandwidth demands in the downlink direction. And the recent proposal of Enhanced Uplink (also known as HSUPA) [4] in Release 6 aims to meet the growing traffic demands in the uplink direction. Both HSDPA and HSUPA enable the efficient transport of packet-switched Internet traffic, thus they are sometimes respectively referred to as 3.5G and 3.75G mobile systems.

The work in this monograph is mainly focused on UMTS networks, hence in the remainder of this chapter, UMTS architecture, services, features and radio resource management strategies will be elaborated.

2.2 The UMTS Services and Network Architecture

Compared to 2G mobile networks, UMTS can provide much higher user bit rates: 384kbps for circuit-switched traffic and up to 2Mbps for packet-switched traffic. With such a high data rate, it is naturally possible for the system to support a wide range of service applications with different QoS requirements. Four QoS traffic classes have been identified in UMTS by 3GPP [2] based on the delay-sensitivity of the applications, which are list as follows,

- conversational: the applications of this class include most services of real-time conversation, such as traditional speech over circuit-switched bearers, Voice over IP (VoIP), and videotelephony, etc. Due to the nature of conversation, this type of service requires most stringent QoS on delay and jitter.

- streaming: the popular applications of this class are Web broadcast and Video on Demand (VoD). Since multimedia streaming can store the received data into a buffer, they can generally tolerate more delay and jitter in transmission than conversational class. However, as a steady bit rate stream is required, they still have higher requirement on delay than the interactive and background services.

One common point between conversational and streaming classes is that both need to preserve the synchronization (real-time) when transmitting data.

- interactive: Web browsing is the most well-known application in this service class, while others include server telnet access, etc. As interactions are involved for the traffic of this class, it is generally characterized by the request response pattern where a response message is always expected within a certain period from the other party, thus the round-trip delay is the main concern for this traffic class.

- background: this type of services consists of those application which can be delivered in background without immediate interaction, such as E-mail, SMS, file transfer, etc. This is the most delay-insensitive class where the delay may be in the order of minutes. The data integrity of both interactive and background services should be preserved, where the data must be transmitted error free.

Based on different QoS requirements on delay and jitter of these four classes, the first two are usually transmitted over a circuit-switched network, while the last two are typically transmitted with packet scheduling.

The UMTS system is composed of various logical network elements, each of which has its own functionality, based on which these network elements can be grouped into two parts: UMTS Terrestrial Radio Access Network (UTRAN) and Core Network (CN). The former manipulates all the radio related functionality, and the latter is responsible for switching and routing voice and data connections to the external networks. The mobile station in UMTS network is referred to as User Equipment (UE). The interface that connects UTRAN and CN is called Iu-Interface, and the WCDMA radio interface between UE and UTRAN is denoted as Uu-Interface, where most research in this monograph is located. Fig. 2.1 illustrates the network architecture of a general UMTS system, and Table. 2.1 lists some important air interface parameters of UMTS networks.

The UTRAN consists of two network equipments: NodeB and Radio Network Controller (RNC), which play similar roles as Base Station and Base Station Controller in the GSM system. The NodeB is responsible for the radio transmission with

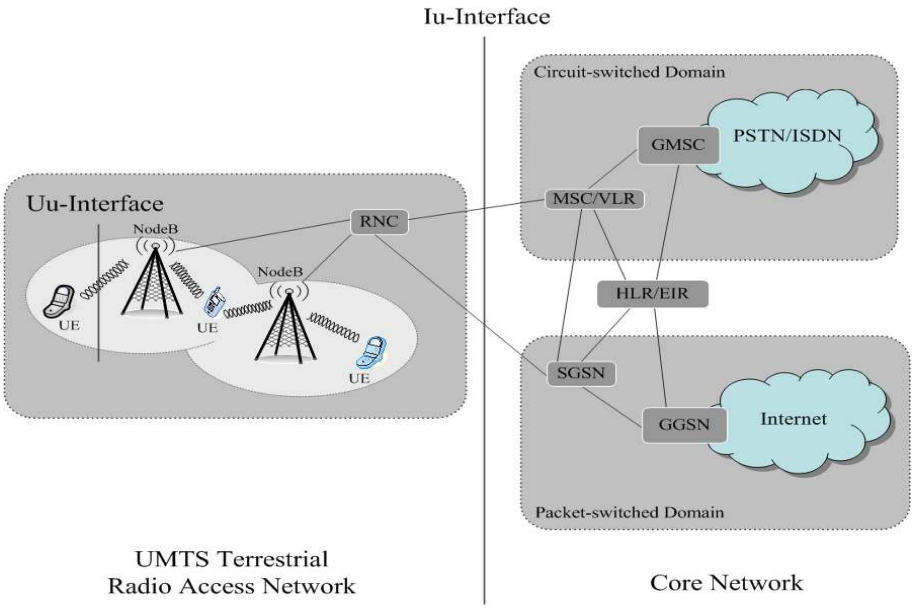


Figure 2.1: UMTS network architecture

UEs in the physical layer, which includes the duties of spreading, modulation, encoding, etc. It is also involved in some radio resource management activities, such as fast power control, fast load control, and packet scheduling in the HSUPA/HSDPA enabled systems. The RNC performs most tasks related to the radio resource management for several NodeBs connected to it, including outer-loop power control, admission control, handover control, load control and packet scheduling. The details regarding radio resource management functionalities will be discussed in Section 2.4.

Although the air interface in UMTS has been completely evolved to WCDMA, the design of equipments and protocols in the CN side is inherited from existing GSM systems, hence both UTRAN and GSM/EDGE Radio Access Network (GERAN) can connect to the same CN, which is quite economic during upgrade without losing performance since the radio access side is the main bottleneck which limits the network capacity. The core network has two domains depending on different traffic types, which are circuit switched domain and packet switched domain. In the circuit switched domain which generally caters for the real-time traditional voice or video streaming traffic, there are Mobile Switching Centre (MSC) and Visitor Location Register (VLR) which serve as the switch and database for circuit switched

Table 2.1: Air interface parameters for WCDMA in UMTS

Carrier spacing	5 MHz
Chip rate	3.84 Mcps
Power control frequency	1500 Hz
Multiple access	CDMA
Duplex	FDD/TDD
Modulation	BPSK (uplink) / QPSK (downlink)
Spreading factor	4-256 (uplink) / 4-512 (downlink)
Allocated frequency bands	1920-1980 MHz (uplink) / 2110-2170 MHz (downlink)

services, and Gateway Mobile Switching Centre (GMSC) connecting to the external circuit switched networks (e.g. Public Switched Telephone Network (PSTN) and Integrated Services Digital Network (ISDN)). The packet switched domain consists of Serving GPRS Support Node (SGSN) which functions similarly to the MSC/VLR but for packet switched services, and Gateway GPRS Support Node (GGSN) analog to GMSC while connecting to the external packet switched networks (e.g. Internet). Additionally, several registers, such as Home Location Register (HLR), Equipment Identity Register (EIR), etc., are employed as system database to store UE-related information.

In the recent 3GPP Release 4/5/6, the network architectures have been improved in order to offer significantly higher data rate than the legacy UMTS network. The main difference between the Release 99 architecture and Release 4 architecture is that the CN circuit-switched domain becomes a distributed network, where the traditional circuit-switched MSC is divided into an MSC server and a Media Gateway (MGW), and also the GMSC is divided into a GMSC server and a MGW. The next step in the UMTS evolution is the introduction of an all-IP multimedia network architecture. In Release 5, it contains first phase of IP Multimedia Subsystem (IMS), which enables the standard approach for IP based service provision via packet-switched domain. The functions of the IMS are further enhanced in Release 6, where the services similar to circuit-switched domain are allowed to be provided via the packet-switched domain. In this architecture, both voice and data traffic are largely handled in the same manner all the way from UE to the ultimate destination, which can be considered as

the convergence of voice and data.

2.3 The Principles and Features of WCDMA

In mobile communications, since there are multiple users transmitting data simultaneously, the sharing of radio spectrum is always required. It is performed by allocating the available bandwidth to active users with certain multiple access scheme, where FDMA, TDMA and CDMA are three major options. As mentioned above, F/TDMA are standard air interfaces in most first and second generation mobile networks, while CDMA was first employed in the IS-95 system, and later its enhanced version, wideband CDMA, is selected as the multiple access technique for the 3G communication systems. Therefore we give a brief introduction to the principles and features of CDMA-based systems in this section.

In CDMA-based systems, the desired narrowband signal is spread to a pseudo-noise code sequence which has much wider frequency band than the message signal. Every mobile subscriber in the system continuously uses the entire spectrum, since each of which has its unique pseudo-random codeword to separate each other due to its orthogonality. On the received side, by again multiplying the incoming signal with its own codeword, the original message can be detected and recovered, whereas all the other signals including the data signal transmitted by other subscribers and background noise appear as interference. A more detailed description of CDMA systems is available in [39,64]. It can be seen that the operation principles in CDMA-based systems are quite different from those in F/TDMA-based networks, and in the following we summarize a few distinct features which have great impact in the later capacity analysis.

Power Control: If the transmit powers of mobile subscribers within a cell are not controlled, they are generally received at NodeB at different power levels, most likely a mobile closer to the NodeB will be received at greater power than a mobile located on the cell fringe. This may incur the problem of near-far effect, which gives unfair advantages to the users with stronger received signal levels, and decreases the demodulating probability for weaker signals since the stronger ones raise the noise

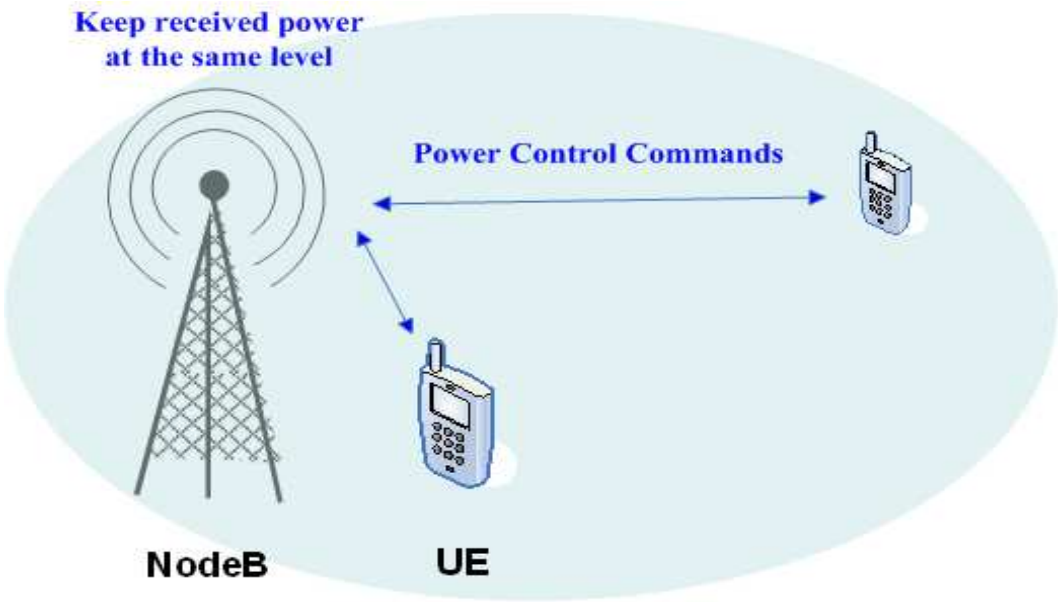


Figure 2.2: Power control in CDMA networks

floor at NodeB. In some cases, the mobile users at the cell boundary can not communicate with the NodeBs at all. To eliminate this effect, power control is introduced in most CDMA-based systems, which aims to keep all the mobile subscribers within one cell received at the same power level, no matter how far from the NodeB it is located. Power control is implemented by constantly sampling the radio signal strength indicator levels of each mobile, and sending the corresponding commands to adjust the transmit power of mobile subscribers. The power control problem and solution is shown in Fig. 2.2.

Soft Capacity: In the orthogonally channelized F/TDMA systems, bandwidth is divided into 'channels' based on frequency or time. There is a hard limit on the maximum number of available channels, which implies no additional user can be accepted when the user number reaches this bound, and thus they are named as bandwidth-limited systems. CDMA systems, however, employ the entire spectrum for all the traffic and modulate the data sequence with a very large number of pseudo-orthogonal codes which are always sufficient. Increasing the number of users in the system results in a certain noise floor rise in a linear manner. One additional user can be admitted into the system at the price of slight degradation of speech quality

for all users. Hence in CDMA networks, rather than an absolute limit applied for the number of users it can accommodate, the system performance drops gradually with increasing number of users. Alternatively, if the service quality is maintained to be constant, then the range of the geographical area covered by a NodeB depends on the instantaneous amount of traffic. When a cell becomes heavily loaded, the coverage area would shrink and the mobile users at cell boundary may be directed to a neighboring more lightly loaded NodeB. This effect is named as 'cell breathing', which is used to balance the load between neighboring cells. Hence, the system capacity, coverage area and user service quality in CDMA networks relate to each other and become a trade-off. Since the capacity depends on the total interference experienced in the system, it is known as interference-limited system.

Soft Handover: When a user moves across the boundary of cells in CDMA cellular networks, since the same frequency bands are used in all cells, CDMA does not need to switch between channels as the case in F/TDMA networks. Instead, the macroscopic spatial diversity allows one user to connect to more than one NodeB simultaneously and gradually handover to the new cell simply by choosing the strongest signal, which makes the handover procedure smoother and more reliable. This is called 'soft handover', which may alleviate the shadowing effect as well [22]. The user may also simultaneously connect to NodeBs located in multiple cell sectors but within the same cell, in which scenario it is referred to as 'softer handover'. The soft handover feature is depicted in Fig. 2.3.

Universal Frequency Reuse: The frequency band allocation used to be a critical task in frequency-reuse based systems, because the same frequency bands need to be allocated apart from each other with at least frequency reuse distance, otherwise the communication quality would be severely affected by the co-channel interference. This is the most complex part in the planning procedure for such networks, where various channel assignment mechanisms (e.g. fixed and dynamic channel assignment) have been intensively researched in order to achieve maximum system capacity gain. However, in CDMA-based networks, since all the mobile users share the entire spectrum in each cell concurrently, there is no need at all to assign the frequency channels for individual cells. Although the universal frequency reuse eradicates much complexity from the network planning, the soft capacity makes estimation of blocking

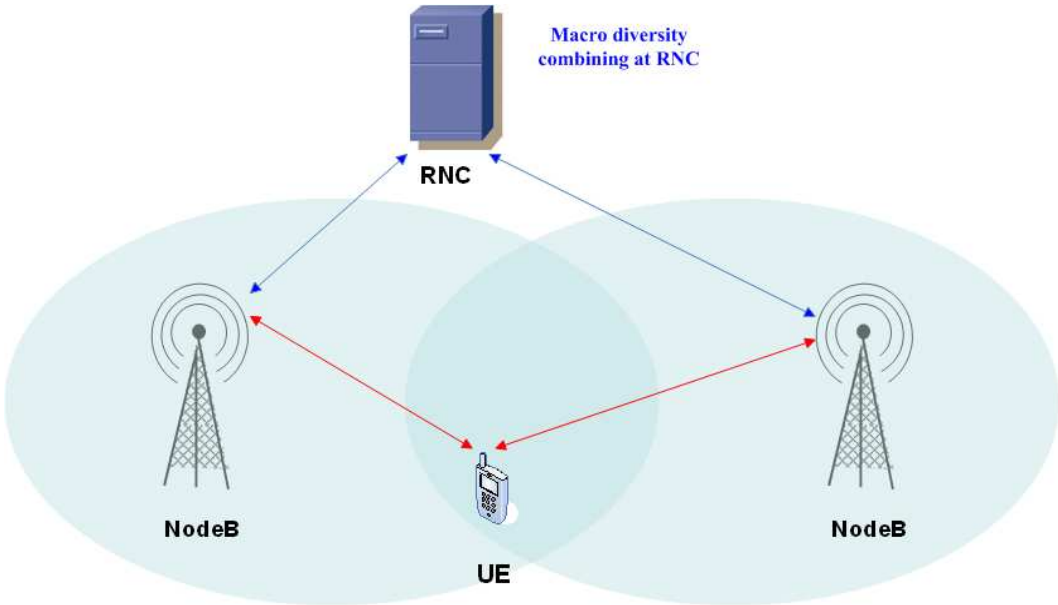


Figure 2.3: Soft handover in CDMA networks

probability and outage probability as well as the coverage area in CDMA-based networks much harder than that in the traditional frequency-reuse based networks, hence the main task in the CDMA network planning has moved into this part.

Voice activity: During a typical two way conversation, each party only spends a portion of time speaking, with the rest of the time listening to the other party. This property is known as 'voice activity' of a conversation, which indicates that the voice data generated by each user only occupy less than 50% of the conversation duration. In F/TDMA networks, it is difficult to take advantage of such property, since the channel is normally allocated to individual user continuously throughout the whole conversation, which is hard to make other subscribers share this channel during the silence period. But in CDMA networks, it is simply an inherent feature to exploit this property as it is an interference-limited system. Each time when the user does not generate voice traffic (i.e. silence period), it does not contribute to interference seen by other users as well, which then translates to higher capacity. This can be equivalent to statistical multiplexing of many ON-OFF sources, which leads to capacity gains, but also results in some uncertain factors for the network planning.

2.4 Radio Resource Management in UMTS

In the 3G mobile communication systems, the mobile users demand higher transmission rates, more reliability and various services support. RRM plays a crucial role in the operation of such systems, which is responsible for efficient utilization of limited radio resources and provision of guaranteed QoS for the mobile users. The functionalities of RRM in UMTS are divided into five components in [30], which are

- Power Control: to allocate the proper transmission power for mobiles so as to minimize the total system interference level;
- Handover Control: to deal with the mobility of users across adjacent cells;
- Admission Control: to decide whether the incoming traffic can be admitted into the system or not in order to guarantee the QoS of the current admitted users;
- Load Control: to handle the situation when the system encounters overload (i.e. congestion);
- Packet Scheduling: to allocate available resources for the packet data traffic and schedule the transmission.

We briefly describe each functional component in the following, with more concentration on the power control and admission control since they are the key factors when determining the capacity model later in this monograph. Note that some of the components (e.g. power control) are the same as what are listed in the previous section, however, in this section, we focus more on the principles and impacts on the aspects of UMTS radio resource management mechanisms by these components, whereas Section 2.3 describes more on the general properties of WCDMA. More implementation details regarding radio resource management in UMTS can be found in the specification [5].

2.4.1 Power Control

The resources for the F/TDMA systems are frequency bands and time slots which can be allocated to individual users, however in the CDMA-based systems, as mobile

users constantly share the same frequency spectrum, the transmission signals from one user are regarded as interference by the others, thus the essential resource in such systems is power. A mobile user with high received power may gain an unnecessarily high SIR, but generate more interference to other users and in turn decrease the system capacity. It is pointed out by [26] that the theoretical maximum capacity of CDMA network is achieved only when the system is under proper power control such that the near-far problem can be effectively overcome.

In a single-cell system, as long as the received power from all users in this cell can be equal to each other, the system is optimized from both capacity and received SIR points of view. Now considering a more general system with multiple cells, the received SIR would depend on not only the power from own-cell users, but also interference from other-cell users, while at the same time, the transmit power would impact the SIR of other-cell users as well. In this sense, in order to achieve the overall maximum system capacity, power control aims to adjust the transmit power of each mobile to just satisfy the required QoS exactly, which in turn lowers the interference contributed to other users. In addition to capacity benefits, the battery lifetime of mobiles are extended as they only transmit at the lowest possible power.

The power control mechanism implemented in the UMTS uplink comprises both open loop power control and closed loop power control. As the path loss can be hardly estimated accurately from the downlink beacon signals, the open loop power control only provides a rough initial power setting for newly connected mobiles. It is the closed loop power control in UMTS, which is composed of inner loop and outer loop, that actually maintains the received SIR to reach the proper target SIR. The outer loop power control is responsible for choosing a suitable target SIR value so as to meet the Bit Error Rate (BER) requirement. The inner loop power control measures the received SIR, compares the measured SIR with the set target SIR, and sends power-up or power-down commands frequently according to the comparison result so as to adapt the transmit power to achieve the target SIR, which forms a measure-command-react cycle. The inner loop power control is performed between the mobile user and the NodeB, whereas the outer loop is between the NodeB and the RNC. Such power control process is demonstrated in Fig. 2.4. In case of soft handover where one mobile user is connecting to several NodeBs (referred to as the

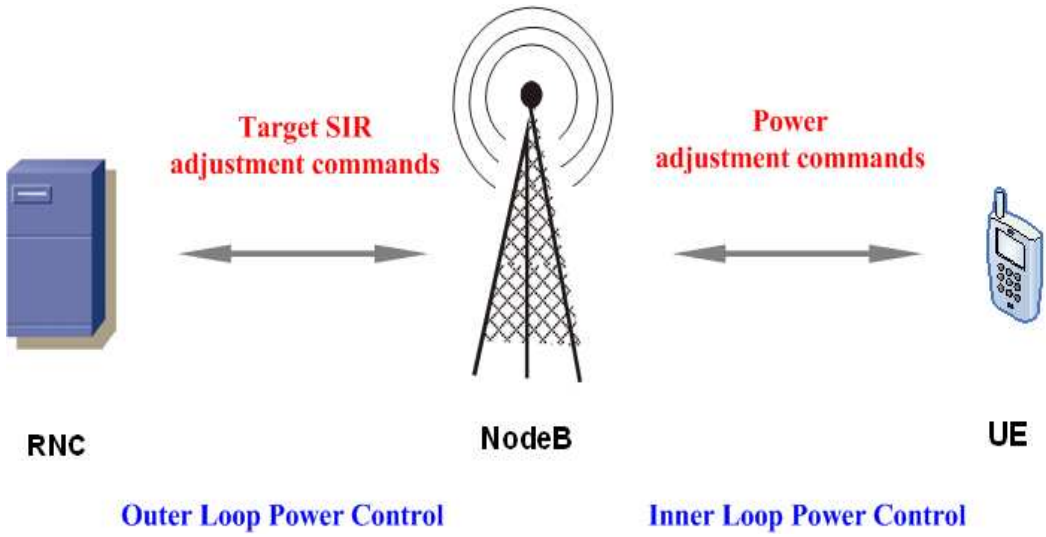


Figure 2.4: Inner and outer loop power control

active set), the power control process would be executed between the mobile and all the NodeBs in the active set.

2.4.2 Handover Control

The soft handover feature of CDMA-based system allows the mobile users to connect to more than one NodeBs concurrently, thus the connection to the target cell can be established before the connection to the original cell is terminated during handover period. This makes the handover process more smooth, and consequently reduces the dropout probability which is quite an important QoS metric. The mobile stations keep measuring the power of the pilot signal broadcast from all NodeBs in terms of chip-energy-to-interference ratio, and then report to the RNC, in which all the signal strengths are compared with certain parameters. Based on the comparison results, the corresponding NodeBs with power stronger than the specified thresholds are categorized into either active set or neighbor set, where active set consists of NodeBs having connections to the mobile, and neighbor set is a list of NodeBs whose pilot signals are not strong enough for the active set, but the mobile station would continuously monitor.

The mobile station connects to multiple NodeBs not only during the handover

period, but also anytime when the pilot signal power satisfies the threshold. In the uplink, the transmit data from mobile station are received at all the NodeBs in the active set, and then each of which relay the data to the RNC where the best candidate frame is selected. This effect, known as macro diversity, is another advantage of soft handover, which decreases the system error rate to a large extent.

Besides the soft handover, softer handover is also implemented in UMTS networks, where the mobile user connects to multiple sectors within one cell simultaneously. Furthermore, inter-frequency handover within WCDMA and inter-system handover between WCDMA and GSM are also supported in UMTS, where the former is to balance the load among different WCDMA carriers, and to extend the coverage area if the original frequency band is limited, and the latter is needed not only for the coverage and load balancing reasons, but also to direct services to the systems which provide better service quality (e.g. GSM).

2.4.3 Admission Control

If there is no limitation applied on the number of serving users, the coverage area will keep shrinking until the territory cannot be fully covered, or the user service quality will drop below the acceptable threshold, either of which may lead to a poorly functioning system. To avoid such cases, Call Admission Control (CAC) is employed to make decisions of acceptance or rejection for the newly incoming requests by estimating the load increase that the new establishment may cause. The principle of admission control in CDMA-based systems is analogous to dynamic channel assignment in F/TDMA networks because both schemes manage the resources in the own cell and neighboring cells. Generally speaking, the aim of CAC module is to admit as much traffic as possible for high utilization without breaching the QoS requirements.

The following part of this section gives a brief review on some typical CAC schemes in CDMA-based systems from the literature, since the 3GPP standard does not specify a uniform admission control technique.

The paper [43] proposes a SIR-based CAC algorithm where it defines the term residual capacity as additional number of initial calls that can be accepted into the system subject to guaranteed outage probability constraints. Then two algorithms are suggested to calculate the residual capacity, where the first one is a localized

algorithm by comparing the difference between SIR measured at local cell base station and the SIR threshold (design parameter), and the second one makes CAC decisions by utilizing all the SIR measurements of adjacent cells and a coupling coefficient which is used to estimate the interference between neighboring cells. The performance of the algorithms is compared with that of fixed CAC scheme through simulation results in terms of blocking probability and outage probability under various conditions.

The authors of [57] point out that the algorithms in [43] are inconsistent with reality, since SIR is kept constant by power control in practical systems, hence it proposes a CAC scheme that is based on total measured interference received at the base station instead of SIR. A channel can be assigned to the new connection if the corresponding interference margin is less than the allowed interference, and distinct interference thresholds can give priority to the handoff connections. In [15], the concept of directed retry is applied to strengthen the CAC scheme suggested by [57], where the new incoming connections from overlapped regions between neighboring cells may be directed to the one with light load, and thus traffic load can be balanced. The paper [13] evaluates the performance of this interference-based CAC scheme and one number-based CAC method [45] in CDMA systems through simulation. The paper [14] also compares these two kinds of CAC mechanisms with the same model as in [13], but for Time Division CDMA (TD-CDMA) systems where time-slot assignment issue for admitted connections should be considered as well.

The work in [17] suggests a CAC algorithm aiming at minimizing the blocking probability without altering the handoff dropping probability so as to maximize the utilization of the CDMA system. Guard channel scheme is employed to give priority to handoff connections, but the guard channel threshold can vary so that new connections can be admitted as much as possible provided that there is no expected arrival of handoff connections in the near future called look-ahead time. Prediction of user's movements is accomplished via base stations by monitoring all the mobiles' power.

The papers [48,68,71] present approximate analytical formulations for CAC problem, mainly focused on handling multiclass services in cellular networks. Firstly the capacity is represented in terms of the number of channels, and thus these are number-based CAC schemes. Then three resource sharing policies among multiple classes are analyzed, which are Complete Partitioning (CP), Complete Sharing (CS)

and Virtual Partitioning (VP), all with guard channels to give priority to handoff connections. CP and CS are two extreme mechanisms where resources are totally isolated or unrestrictedly shared among classes respectively, while VP manages to strike a balance between CP and CS. VP divides traffic classes into various groups, while the resources of under-utilized groups can be used by the excess traffic from overloaded groups, subject to preemption. Hence it behaves like CS when the overall traffic is light and like CP when the overall traffic is heavy. The call-level QoS constraints (i.e. blocking probability, dropping probability) are achieved through these resource sharing policies. Furthermore, the voice activity is considered so that more connections can be admitted into the system without violating the packet-level QoS constraints (packet loss rate). Through joint call-level and packet-level optimization, the system utilization enhancement is attained with satisfying the QoS requirements in both levels, which is the main aim of these papers. The analytical models are derived using Markov chains, and the numerical and simulation results for QoS metrics show great agreement.

In the analysis later in this work, the CAC we employ is based on the idea similar to that in [57], as this admission control is much practical, as more straightforward compared with others. As the admission control is not the main focus in this thesis, the simpler one is more suitable.

2.4.4 Load Control

If the other functionality of RRM such as admission control and packet scheduling work well, there would be only occasional situations where the load experienced at NodeB exceed the limit. If such cases occur, load control manages to recover the system from overload status and maintain it back to the target load. As mentioned above, outer loop power control can reduce the target load, and soft handover can contribute to the load balancing by assigning the mobile users to the other underload frequencies or systems. All these actions can be utilized for load control, and some other approaches include lowering the packet data throughput via packet scheduling, as well as dropping some low priority mobile stations in some extreme overload cases.

The possible load control approaches are summarized in [30] as:

- Downlink fast load control: to deny the power-up commands received from the

UEs;

- Uplink fast load control: to reduce the target E_b/I_0 for the outer loop power control;
- Reduce the throughput of packet data traffic;
- Inter-frequency handover, i.e. to another WCDMA carrier;
- Inter-system handover, i.e. to another system such as GSM;
- Decrease the bit rates of real-time traffic, e.g. employ another speech codec;
- Drop low priority ongoing calls.

The first two approaches can react in a very fast manner since they are carried out within NodeB. The best-effort packet data traffic can be reduced by certain packet scheduling schemes, which is to be introduced in the next part.

2.4.5 Packet Scheduling

As UMTS systems support not only real-time QoS traffic, but also elastic traffic based on packet access, it is the responsibility of RRM to allocate appropriate data rate for these packet switched services. The functions of the packet scheduler are summarized as:

- Allocate the available capacity resources (apart from the resources occupied by the QoS traffic) to each best-effort traffic user;
- Determine the transport channels for the packet data transmission for each user;
- Monitor and maintain the system load.

The packet scheduler is typically operating in RNC, which possesses the load information of multiple adjacent cells, such that the scheduling can be much efficient. The air interface load information required by the packet scheduler is measured at the NodeB side. The packet scheduler keeps monitoring the system load, attempting to adjust the data rates for best-effort users so that the total received load can be as

close as possible but not beyond the target load. Later in Section 5.2, this principle is elaborated and illustrated in more details. As mentioned in the previous part, when system overload status is encountered, packet scheduler can decrease the network load contributed by the best-effort packet data traffic for the load control.

In WCDMA systems, three types of transport channels can be used to transmit the packet data, which are common channels, dedicated channels and shared channels. The packet scheduler chooses the most suitable channels for the packet data transmission.

When allocating the available load and channels to individual users, it complies with certain rules specified by the network operators such as fairness, maximum system throughput, etc. There are a number of scheduling algorithms that can be employed when dividing the available resources for the simultaneous users, such as *fair throughput scheduling*, *fair time scheduling*, *C/I scheduling* and *prioritized scheduling*. Later in the analysis in Chapter 4 and 5, we choose the fair throughput scheduling scheme, whose goal is to give all best-effort users the same packet data throughput, no matter where they are.

2.4.6 Related Work in RRM

There are also some other studies on the radio resource management in CDMA-based systems. The work in [74] suggests a comprehensive RRM scheme for the cellular CDMA-based systems with support of heterogeneous services, which comprises of power distribution, rate allocation, service scheduling and call admission control. The service model is first constructed by dividing the traffic into four types, each of which corresponds to the standard QoS classes defined by 3GPP and is characterized by a few service requirements such as rate, delay, jitter, BER, etc. Then the algorithms and functionalities for each module are stated. Rate allocation calculates the transmission rate for each connection to guarantee the minimum transmission rate for non-real-time traffic and the delay/jitter constraints for real-time traffic through a fixed-sized buffer as well as the packet loss rate due to buffer overflow. Power distribution determines the target receive power at the base station so as to control the packet loss caused by interference for all the mobiles. CAC decides to accept or reject the incoming traffic based on the rate allocation and power distribution results so

that the QoS requirements of all the admitted connections can be maintained. Moreover, CAC makes use of user mobility information to reserve resources for potential handover requests in order to achieve acceptable Grade of Service (GoS) performance as well in terms of new connection blocking probability, handover connection dropping probability and resource utilization at the network layer. Packet scheduling at the link layer assigns the appropriate rate and power for admitted connections to efficiently utilize the system resources while achieving QoS guarantees.

Chapter 3

Modeling of UMTS Networks with QoS Traffic

In this chapter, we aim to present an analytical capacity model of the UMTS uplink, which is a multi-cell WCDMA-based system. As the system is interference-based, the interference power is the key factor to build the capacity model, in which the interference contributed by mobile users in other cells is quite difficult to characterize due to the interactions among multiple NodeBs. Therefore, the main task in the construction of a capacity model is to propose efficient analytic techniques for derivation of the distribution function of such other-cell interference random variable. The employed practical power control mechanism aims to maintain the received SIR, rather than received power, of each mobile user at the same level so as to maximize the system capacity. The system behavior with multiple service classes and log-normal shadowing effects are studied. After obtaining the fully analytical characterization of the other-cell interference, our computation of the interference is directly applied to the analysis of outage probability, which is a significant parameter during the network capacity planning.

In Section 3.1, a brief literature review in the area of capacity evaluation for a few types of mobile networks is presented. Section 3.2 describes the general system assumptions, propagation model with both distance loss and shadowing effects included, traffic model with spatial user distribution complied with homogeneous Poisson process, and the employed power control mechanism in capacity evaluation. The Monte

Carlo simulation approach is first demonstrated in Section 3.3 to understand system behaviors and produce reference results for later analytical model verification, followed by proposing a complete analytical model for the other-cell interference characterization in Section 3.4. An iterative approach is employed in this section to solve fixed-point equations for the distribution function and the corresponding log-normal approximation parameters in Section 3.5 in order to simplify the computational complexity. With fully characterized other-cell interference, the outage probability in WCDMA networks, which is the most significant capacity metric, is accordingly formulated in Section 3.6. Section 3.7 demonstrates some numerical results through which the suggested model is successfully validated. Heterogeneous service classes with different data rates, BERs and activity factors are supported in the proposed model, but in this chapter all the users are treated as QoS traffic. Each incoming user is assigned to a certain pre-defined class, and its data rate is fixed once admitted into the system. The analysis with integration of QoS and best-effort users will be elaborated in the next chapter.

3.1 Introduction

The capacity model is quite a useful tool during the radio network planning process, especially in the initial dimensioning phase. The goal in the initial phase is to give a rough estimation on the number of NodeBs required and their basic spatial configuration in a certain area with pre-identified user side QoS (e.g. target SIR) and network side GoS (e.g. outage probability). There are two types of approaches for this purpose, by 'trial and error' simulations and by analytical models. The former method is based on picking up some design patterns randomly and comparing the system performance of each scenario, which generally may achieve more accurate results with given configuration as input and more flexibility for system implementation. However, since there are too many variables involved at this stage, the computational task would be tremendous, which consequently impacts the efficiency of searching the optimal pattern. On the contrary, with analytical models, we can quickly decide the network feasibility with much less computational effort, although the resultant accuracy depends on the modeling techniques. Apparently this latter approach is more

suitable for the initial planning process due to its efficiency, since the accuracy of the method can be later compensated in the following stages. Here and in subsequent chapters, we focus on the construction of capacity models which can be utilized for network planning under various conditions. The capacity models are characterized in terms of a few network parameters such as blocking probability, outage probability, etc., and thus the maximum offered load which maintains such probabilities below certain thresholds can be derived straightforwardly.

3.1.1 Capacity Models in the Reuse-Based Networks

We first briefly give an overview on the capacity models in F/TDMA cellular networks in second-generation mobile communications systems. In both types of networks, there are fixed number of either radio channels or time slots to be allocated to the users subject to certain constraints. The traffic capacity under a stochastic load depends largely on the channel assignment scheme adopted, which is fundamental to the operation.

The simplest strategy for channel allocation is Fixed Channel Assignment (FCA), where channels are assigned at planning stage and the assignment remains fixed during operation. The mobiles connected to a particular base station may only occupy the channels allocated to this cell even the channels in neighboring cells are free. The utilization of resources is poor because traffic patterns may change from time to time. The capacity model is rather simple under such scenario, where an $M/G/N/N$ queue is used to characterize the user number in each cell with some standard assumptions, which can be easily evaluated with Erlang-B formula [37].

Dynamic Channel Assignment (DCA), on the contrary, can fully exploit the capability of the system due to its flexibility. The user channels are assigned to cells when required as long as the reuse constraints can be satisfied. Better utilization leads to the performance gain over FCA in capacity for typical systems between 10% and 20% [24]. The capacity models for the reuse-based systems operating with DCA are suggested in [22–24], where the joint distribution of the connections number in each cell, either exactly or as a good bound, takes a product form in a truncated state space. This product form model also applies to several variants of DCA implementations such as hybrid channel assignment (i.e. a balance between FCA and

DCA), DCA with directed retry, DCA with directed handoff, etc. [22], and even CDMA-based cellular networks in certain scenarios according to [21].

3.1.2 Capacity Models in the CDMA-Based Systems

In CDMA-based systems, quite different traffic behaviors are observed from those of above reuse-based networks due to distinct operational features, which lead to diverse capacity analyzing method in the planning process [22]. One prominent distinction is the relationship between coverage planning and capacity planning. In F/TDMA networks, these are completely two separate stages as the capacity depends only on the number of user channels, while coverage area depends on the transmission power. However in CDMA-based systems, the number of users that can be accommodated in the system has tight relation with the target SIR requirement and received power. More mobile users connecting to a particular NodeB results in higher interference experienced, which consequently leads to higher required received power if constant target SIR assumed, and thus smaller coverage area. This trade-off between capacity and coverage causes both issues to be considered together in the network planning process, which indicates more complexity involved.

Single Cell Model

We start with investigating the analytical capacity model of single-cell CDMA system, where the fundamental expression is developed in [26]. With the assumptions of perfect power control with respect to received power level, the capacity in terms of number of users supported is

$$N = 1 + \frac{W/R}{E_b/I_0} - \frac{N_0}{S} \tag{3.1}$$

where W is the total spread bandwidth, R is the data rate, E_b/I_0 (bit energy to interference density ratio) is the requirement in order to satisfy an acceptable BER, N_0 is the background noise and S is the received signal strength at the NodeB.

If sectorization and voice activity detection are considered, the performance can

be augmented. Now the average number of users in one cell can be expressed as

$$N = \beta \cdot \left[1 + \frac{1}{\alpha} \cdot \left(\frac{W/R}{E_b/I_0} - \frac{N_0}{S}\right)\right] \tag{3.2}$$

where α represents the 'voice activity factor' which is the average ratio when the speaker is active and generally take the value 35% to 40% [11]. As discussed in the previous chapter, the transmit power can be suppressed during the period when no voice is present. In other words, no interference would be contributed during this interval by this particular user, which leads to an approximately multiplicative capacity gain. Another parameter β is the sectorization gain, for example, with three directional antennas per based station, each of which occupies 120° beamwidths, the gain is just three.

In the above analysis, the term 'capacity' refers to the maximum number of users that can be accommodated in the reverse-link with the resulting outage probability staying below a given threshold. This can be regarded as the 'static' capacity of the system since no traffic model is included in this expression. In order to measure the capacity more closed to the real system, the performance of the system under stochastically varying traffic loads becomes a crucial factor since it demonstrates the real-time operation of CDMA system. This concept is introduced in [20, 67], in which the capacity is defined as 'Erlang capacity', or 'traffic capacity'. Rather than the maximum number of users, it studies the maximum offered traffic load per cell that can be supported while maintaining a certain outage probability (or blocking probability), and thus it can be regarded as the 'dynamic' capacity of the system in contrast to the previous definition of 'static' capacity, since the stochastic nature of call arrivals and departures is encapsulated in this model. In [67], three random variables are developed to characterize the source traffic, which are a Poisson random variable to model the number of users in each cell rather than a constant number, a Bernoulli random variable for the voice activity again and a log-normal random variable for the target E_b/I_0 of each user since it may vary with propagation conditions due to imperfect power control. A new random variable is introduced as

$$Z = \sum_{i=1}^{k} \upsilon_i \varepsilon_i \tag{3.3}$$

where k, υ and ε are the Poisson, Bernoulli and log-normal random variables mentioned above, respectively. Then the outage probability is represented as the ratio when the total interference level exceeds a predetermined value

$$P_{outage} = \Pr\left\{\sum_{i=1}^{k} \upsilon_i \varepsilon_i > \frac{W}{R} \cdot \left(1 - \frac{N_0}{I_0}\right)\right\} \tag{3.4}$$

where N_0 and I_0 are the background noise and maximum total acceptable interference density. Two techniques are employed to characterize Z, by using a modified Chernoff bound [12] to obtain an upper bound, or by central limit theorem to gain an approximation. After setting this probability equal to a pre-defined value, the Erlang capacity can be determined with the given distribution of user number k. A number of papers further the research on the traffic capacity in this framework, such as in [56], one additional constraint on the transmit power applies into the calculation, and in [55], the capacity model is extended to support integrated voice and data traffic.

Multiple Cell Model

If perfect power control is assumed, the capacity of a single-cell CDMA-based system is independent of the user location since they are controlled to be received at the same level at NodeB. When it comes to a multi-cell environment, the capacity evaluation becomes much more complicated, which mainly comes from an additional component - the total interference generated by the users from other cells. This factor depends not only on the number of active users in the other cells, but also on the location of each of them since these users are power controlled by different NodeBs. Therefore, the key factor in the multi-cell capacity model is the characterization of other-cell interference. Once the distribution of this variable is determined, the multi-cell capacity can be analyzed by incorporating the other-cell interference to the above single-cell model. We again start with the analysis of 'static' capacity where the number of users located in the surrounding cells are assumed to be constants. The cases with variation in the number of users will be treated in the later part when traffic issue is included.

As power control being applied in the CDMA-based systems as shown in Fig. 3.1, the received interference at NodeB x from user k in other cell depends not only

on the propagation conditions along the path to the target NodeB x, but also on that of the path from the user k to its own serving NodeB y, since this NodeB may increase or decrease the transmit power of the user through power control. The work in [26] is one of the first capacity analyses of CDMA-based networks, in which the other-cell interference is modeled as a Gaussian random variable and its moments are represented as functions of the own-cell interference, or the mean user number with the assumption that required received power are equal at all the NodeBs. The employed propagation model presumes the path loss between the transmitter and the receiver is proportional to the product of the minus fourth power of distance and a log-normal random variable with standard deviation 8dB (i.e. $10^{(\zeta/10)}r^{-4}$). Let r_m be the distance from the interfering user k to its home NodeB y, and r_0 be the distance to the target NodeB x, the interference produced by this particular user compared with the power controlled received unit power S can be represented as

$$\frac{I_k\left(r_0, r_m\right)}{S} = \left(\frac{r_m}{r_0}\right)^4 \cdot 10^{(\zeta_0-\zeta_m)/10} \tag{3.5}$$

where the first term is due to the attenuation by distance factor assuming the path loss component fixed at -4 and the second term is by the shadowing effects. Then by integrating the above ratio multiplied by the user density over the entire network area, the moments of the ratio of other-cell interference to the required received power I_{oc}/S are obtained. Then with the own-cell interference modeled as a sum of Bernoulli random variables, and the other-cell interference modeled as a Gaussian random variable, the outage probability in multi-cell system is just in a similar form to that of the single-cell scenario except the additional other-cell interference component. It is given as

$$P_{outage} = \Pr\left\{\sum_{i=1}^{k} \upsilon_i + \frac{I_{oc}}{S} > \frac{W/R}{E_b/I_0} - \frac{N_0}{S}\right\}, \tag{3.6}$$

which can be computed straightforwardly once the moments of other-cell interference random variable is known.

Another issue considered in this work is the cell membership of the users. In practical systems, each mobile user chooses the NodeB with least attenuation, which is compared among the reference pilot power the mobile receives, as its serving NodeB.

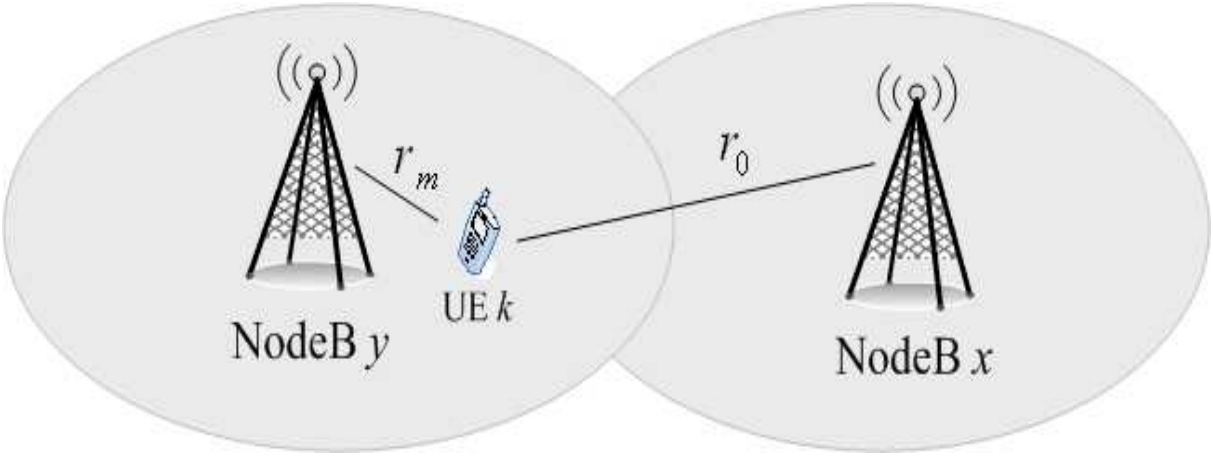


Figure 3.1: Mobile k in the multi-cell environment

As a result, based on the employed propagation model, both distance and shadowing effects factors determine which NodeB the user joins. In [26], the smallest distance rather than the least attenuation is assumed as the decision factor during the cell membership selection, which simplifies the analysis a great deal. More on this issue is discussed in [66] and is elaborated in the following.

Another analytical expression of the other-cell interference can be found in [36] with a simpler propagation model, which merely considers the distance factor. With perfect power control, the interference made by one user located in the surrounding cells becomes $\tilde{I}_k(r_0, r_m) = \left(\frac{r_m}{r_0}\right)^4 S$ instead. This assumption also implicitly diminishes the complexity of cell membership issue, since the NodeB with the smallest distance is always the one with the least attenuation in this model. Furthermore, it is assumed that there are equal and fixed number of users uniformly distributed in each cell, and all the hexagonal cells are approximated by circular cells in the calculation. The interference generated from one of surrounding cells $\tilde{I}_m$ can be derived by a similar integration of $\tilde{I}(r_0, r_m)$ multiplied by user density over that particular circular cell region, and the resultant $\tilde{I}_m$ is a function of distance from NodeB y to the target NodeB x, the user density and the controlled received power at NodeB. With all the above assumptions and a constant Path Loss Exponent (PLE), an analytic expression of other-cell interference is given in this work.

The paper [66] is an extension work of [26] in terms of modeling the other-cell

interference in cellular power controlled CDMA networks, especially on the cell membership issue. Rather than choosing the closest NodeB, the mobile user is assumed to be served by the one with the least attenuation among the set of N_c nearest NodeBs. Similar to [26], the other-cell interference is again modeled as a factor f, which is known as relative other-cell interference factor, of the own-cell interference. With the factor f, the other-cell interference can be understood as the 'virtual' effective interference contributed by each own-cell user. Therefore, $(1+f)N$ can be recognized as the total interference received (intra-cell and inter-cell) at the target NodeB in the power controlled CDMA-based system. During the calculation, the integration region has to be split into two areas depending on whether the N_c nearest cells include the target NodeB or not, and the sum of these two components determines the expected interference generated from all the uniformly distributed users in other cells. Due to tremendous complexity incurred by such assumption, only the first moment of the other-cell interference is derived in this work, although it is still not an easy thing. The impact of N_c is also discussed in this paper through the numerical results, in which it is shown that a much lower f is yielded from $N_c = 1$ to $N_c = 2$ in normal propagation environment, while just limited further reduction of other-cell interference for the cases $N_c > 2$.

By incorporating the factor f, the average number of users per cell in the multi-cell environment can be simply obtained in a similar way to the single-cell capacity analysis, which is expressed as in [21, 22]

$$N = \frac{1}{1+f} \cdot \left[1 + \frac{W/R}{\alpha} \cdot \left(\frac{1}{E_b/I_0} - \frac{N_0}{S}\right)\right]. \tag{3.7}$$

The capacity analysis for wideband-CDMA system with multi-class traffic is presented in [70] by treating the cases where different traffic classes have different spreading gains. The outage probabilities for each class are formulated in terms of the number of users in each class, the voice activity factors and spreading gains, the intra-cell received power, and the inter-cell interference. Through that the capacity for a system with K-classes can be obtained in a K-dimensional space. The paper [69] extends this work by analyzing the wideband-CDMA systems with variable bit rate multi-class services where each user has multiple spreading codes.

The multi-cell 'static' capacity models discussed above are analyzed based on the assumption of maximum number of users in each cell. The drawback of such evaluation is that it only reflects the worst case in the system since maximum number of users implies maximum interference from other cells, and thus the results would be quite conservative. In order to take the multiplexing gain into account, again the traffic capacity under varying load is worth investigation, where different source traffic models may lead to various system performance.

The simplest way to model the traffic behaviors in the CDMA-based cellular networks is to employ $M/G/N/N$ queues. Each cell in the network is assumed to be able to accommodate N channels, thus the user number in each cell can be modeled as an independent $M/G/N/N$ queue. This model can be analogous to the fixed channel assignment scheme in orthogonally channelized system such as F/TDMA because no user can borrow the available capacity from adjacent cells. The major advantage of this model is its simplicity when calculating the blocking probability since the incoming user is blocked each time when the user number reaches N, in which case Erlang-B formula can be used. However, the apparent shortcoming is that soft capacity, one important inherent feature of interference-based CDMA against F/TDMA that one cell can automatically accommodate more than N users without breaching the BER requirements in case there are light traffic loads in surrounding cells, is not exploited in this model yet, and thus the system capacity is underestimated.

In [20, 32, 67], an $M/G/\infty$ queue model is adopted for each cell, where ∞ implies that any incoming call can be admitted into the system due to soft capacity. In this model, the fixed user number in each cell is replaced by an independent Poisson random variable. Since there is no blocking occurred in such system model, the outage probability, which is defined as the probability that the BER and bit rate requirements cannot be satisfied, is studied instead.

The paper [67] is also the extension work of [26] in terms of Erlang capacity. Based on the assumption that each cell is equally loaded with same user arrival and departure rates, the relative other-cell interference factor f is again used under $M/G/\infty$ traffic model, and the problem reduces to the single cell case with k replaced by $k(1+f)$. The outage probability by Gaussian approximation is presented in this work and after inverting the resultant equation, the formula for the Erlang capacity under a fixed

outage probability is given.

The paper [20] also studies the outage probability based on $M/G/\infty$ traffic model, but with different approach to characterize the other-cell interference. With common assumptions like simple propagation model, perfect power control to unit power and approximation of the hexagonal cells by circles, the interference caused at the target NodeB located at (a, π) in polar coordinates by a user located at (r, ϕ) and served by the NodeB at origin can be represented as

$$I(r, \phi) = \left(\frac{r}{\sqrt{r^2 + a^2 + 2ar \cdot \cos\phi}} \right)^{\gamma} \tag{3.8}$$

where γ is the PLE. Rather than deriving the distribution function of the random variable I_{oc} which is the total other-cell interference directly in previous work, it focuses on the inter-cell interference random variable $I(r, \phi)$ generated by an individual user in the other cell, which can be modeled from the distribution functions of r and ϕ. The total other-cell interference is consequently represented as a compound Poisson sum of $I(r, \phi)$, and the problem then reduces to calculation of the Poisson sum of a series of identical and independent distributed random variables, which can be computed through central limit theorem approximation or Chernoff bound. The effect of voice activity can be included as a multiplicative factor, and the log-normal shadowing effect is also included in the later part where the incurred cell membership problem is solved by choosing between the closest NodeB and the target NodeB. Although only the results under uniform user distribution are presented in this paper, it can be extended for arbitrary user distributions in the similar way, nevertheless it may become much more complicated generally. Another advantage of this approach is that when characterizing the total received interference, the convolutions of two random variables, which are intra-cell and inter-cell interferences respectively, can be avoided.

The paper [32] develops the techniques to evaluate the Erlang capacity of the CDMA-based systems more accurately. It suggests the numerical integration method over the two dimensional hexagonal regions to study the moments of the individual interference random variable. The suggested work is more efficient in case the entire distribution function is not required, but in fact, it shares the same principle with [20].

Again the central limit theorem is applied for the compound Poisson sum and it extends [20] by introducing the Edgeworth asymptotic correction to remove the error in central limit theorem approximation. The numerical integration approach is also efficient and accurate in the calculation of Chernoff bound, extension to non-uniform user distribution and irregular cell layouts.

There are also traffic models to describe the system with admission control mechanisms in order to guarantee the QoS of ongoing calls. A product form model on the truncated state space is suggested in [19] where the effective bandwidth concept [18,35,49] for variable bit rate sources in broadband ISDN networks using Asynchronous Transfer Mode (ATM) is applied in CDMA-based cellular mobile networks to characterize each individual mobile dependent on its class and location. The call admission policy is then defined by the resultant admissible region, which has a quite similar form to that of circuit-switched networks operating with fixed routing. In this sense, the general theory of multi-service loss networks [34] can be directly applied to the mobile environment. In particular, the stationary user number state distribution has a product form on the truncated state space. The major sources of variability included in this model are random user locations within a cell and bursty SIR requirement for each mobile.

The paper [40] develops a queueing model with the SIR-based CAC [43] applied in the CDMA systems. The ongoing calls are also subject to be dropped due to poor SIR. A quasi-birth-and-death process is used to model the number of users in each cell, and then an iterative algorithm is presented to determine the stationary distribution of the system. The performance measures such as blocking, dropping and outage probability as well as carried traffic are given in this work.

3.2 System Model Description

3.2.1 General System Assumptions

Throughout this monograph the following general assumptions for system description, which are common assumptions in much of the literature, are made. The standard uniform hexagonal cell layout with a NodeB at the center of each cell as illustrated in Fig. 3.2 is assumed. The uplink and downlink are assumed to utilize disjoint frequency

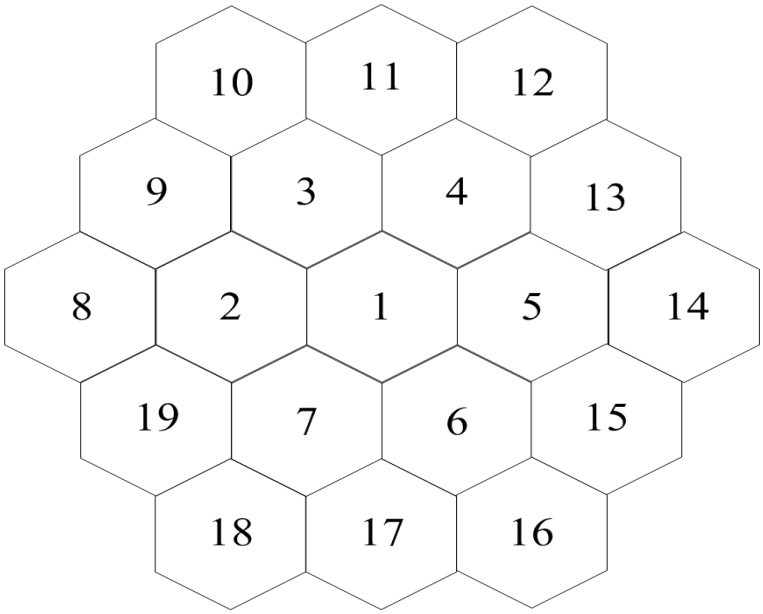


Figure 3.2: Standard uniform hexagonal cellular network layout

bands. Furthermore, it is assumed that there are always sufficient available codes so that the system is interference limited only. Finally, without loss of generality, all the distance values are normalized by the distance between any two adjacent NodeBs. All mentioned quantities in the sequel (e.g. path loss, received power, etc.) refer to the uplink only which is the main focus of this monograph.

3.2.2 Propagation Model

The simplest propagation model for a communication channel in the mobile radio environment is the log-distance path loss model, where the attenuation incurred is inversely proportional to the distance d between the transmitter and the receiver raised to the PLE γ [28]. If the distance between the transmitter and the receiver is represented by variable d, the received power P_R is given as

$$P_R = P_T C_0 d^{-\gamma} \tag{3.9}$$

where P_T denotes the transmit power. C_0 is a function of carrier frequency, antenna gains, etc. which is independent of distance and thus assumed to be a constant in

this model. The PLE γ depends on the antenna heights, and is typically in the range between two and six.

However, in order to employ such a simple model, many restrictions, such as minimum and maximum distances, terrain profile variation, should apply. In practice, due to variations in terrain contour and shadowing from buildings along the propagation path, measurements have shown that the path loss at a particular location is random and distributed log-normally about the above mean distance-dependent value [38, 54]. Incorporation of this phenomena, which is generally referred to as log-normal shadowing, leads to the following equation

$$P_R = P_T C_0 d^{-\gamma} 10^{\zeta/10} \tag{3.10}$$

where ζ is a zero mean Gaussian random variable with standard deviation σ typically in the range six to twelve. The received power P_R has the log-normal density function given in [20] as

$$f_P(z) = \frac{1}{\sigma' \sqrt{2\pi}} e^{-(\ln z - \mu)^2 / 2\sigma'^2} \tag{3.11}$$

where $\mu = \ln P_T C_0 - \gamma \ln d$ and $\sigma' = \sigma \ln 10/10$.

3.2.3 Power Control Mechanism

The transmit powers of each UE are the main resources for CDMA-based systems, thus power control plays an important role in such systems. In many publications regarding CDMA-based network modeling listed above, power control is assumed to achieve a fixed received power globally for simplicity reasons. However, under such an assumption, the optimal system performance cannot be achieved in the multi-cell environment. This is because each NodeB experiences a different level of interference, if the received power are same for all the UEs, the received bit-energy-to-interference-density ratio would accordingly differ from each other. This implies some UEs would achieve unnecessarily higher QoS than the target, but at the price of causing more interference to the system, which in turn lowers the overall system capacity.

In order to utilizes the resources more efficiently, in real systems, the practical

power control algorithms manage to maintain the received bit-energy-to-interference-density ratio of UEs, rather than the received power, to their target levels. This ensures that the mutual interference is minimized so that significant capacity gain can be achieved. Under such power control mechanism, the inner and outer loop power control aim to minimize each UE's transmit power whereas still satisfying the QoS, which translates to minimum interference and thus maximum network capacity.

In mathematical terms, each UE k operating at bit-rate R_k is power controlled by its serving NodeB x to maintain the target E_b/I_0 requirement ε_k^*. For the cell membership issue, the criterion for the choice of serving NodeB is the one with least attenuation to UE k. The voice activity of one mobile user is modeled by a Bernoulli random variable υ_k with ON status at a probability ψ and OFF status at a probability $1-\psi$. The power control equation which needs to be satisfied by all the UEs k connecting to their serving NodeB x is then given by

$$\varepsilon_k^* = \frac{W}{R_k} \cdot \frac{S_{k,x}^R}{WN_0 + \sum_{i \neq k} S_{i,x}^R \upsilon_k} \tag{3.12}$$

where $S_{k,x}^R$ denotes the received power of UE k at NodeB x, W the system bandwidth and N_0 the background thermal noise spectral density. The ratio W/R_k is also referred to as the processing gain of UE k.

If the above equation can be fulfilled by all the UEs in the system simultaneously, and updated instantly at each moment whenever system status varies (i.e. arrival of new users or departure of existing users), the theoretical global minimum transmit power and in turn minimum interference can be achieved, we thus define such optimal scenario as *perfect power control*. However, in the real systems, due to complex propagation conditions, the power control equations do not always hold for all UEs. The *imperfect power control* effect can be approximated by modeling the target E_b/I_0 as a log-normally distributed random variable from empirical results in the literatures [63, 65].

3.2.4 Traffic Model

We assume there are T available streaming service classes in the system, each of which has its particular QoS requirements such as data rate, target E_b/I_0 requirement, etc.

The stochastic user population and distribution of each class on a two-dimensional surface are generated according to a spatial homogeneous Poisson process [16] with λ_t denoting the spatial traffic intensity (i.e. mean number of mobile users per unit area size) of class t. Therefore, the probability distribution of the number of active users $\bar{n}_x = (n_{1,x}, \ldots, n_{T,x})$ connecting to NodeB x spanning on a surface with area A_x for any arbitrary observation instant falls into the truncated product form as in [34]:

$$P(\bar{n}_x) = \begin{cases} P_0 \prod_{t=1}^{T} \frac{(\lambda_t A_x \upsilon_t)^{n_{t,x}}}{n_{t,x}!}, & \bar{n}_x \in \mathbf{S} \\ 0, & \text{otherwise} \end{cases} \tag{3.13}$$

and the normalization constant

$$P_0 = \frac{1}{\sum_{\bar{n}'_x \in \mathbf{S}} \prod_{t=1}^{T} \frac{(\lambda_t A_x \upsilon_t)^{n'_{t,x}}}{n'_{t,x}!}} \tag{3.14}$$

where υ_t is the activity factor of class t UEs. The admissible region $\mathbf{S}$ is defined by the pole capacity [63] as $\bar{n}_x \in \mathbf{S}$ if

$$\sum_{t=1}^{T} \frac{n_{t,x} \varepsilon_t^* R_t}{W + \varepsilon_t^* R_t \upsilon_t} < 1. \tag{3.15}$$

3.3 Monte Carlo Simulation Approach

To investigate the system behavior, Monte Carlo simulation is employed, which generates a large set of user patterns (i.e. user population and locations), and under each scenario,the system parameters (e.g. interference power) are evaluated repetitively so that the distribution of interested random variable can be approximated.

In the uplink of a UMTS network with T service classes, each user pattern is formed based on a spatial homogeneous Poisson process according to the above traffic model, where all the mobile users in the same class have the bit rate R_t, target E_b/I_0 requirement ε_t^* and voice activity factor υ_t. Then the power control equation (3.12)

which need to be fulfilled by all the UEs can be rewritten as the following equation:

$$\varepsilon_t^* = \frac{W}{R_t} \cdot \frac{S_{k,x}^R}{I_x^{own} + I_x^{oc} + WN_0 - S_{k,x}^R \upsilon_t} \tag{3.16}$$

where the interference component I_x^{own} refers to the total power received from all the UEs that connect to the same NodeB x, and I_x^{oc} corresponds to the sum of power received at NodeB x from the UEs which are served by all the NodeBs other than x.

Following the assumption that the power control mechanism aims to maintain the target E_b/I_0 for any UE k of service t as a constant ε_t^*, since all the UEs served by one NodeB experience the same level interference, the received powers for each UE with same class connecting to NodeB x are in turn controlled equal to each other (i.e. $S_{t,x}^R = S_{k,x:[k\in t]}^R, \forall k$). Thus, solving (3.16) yields

$$S_{t,x}^R = \frac{\varepsilon_t^* R_t}{W + \varepsilon_t^* R_t \upsilon_t} \left(I_x^{own} + I_x^{oc} + WN_0\right). \tag{3.17}$$

It can be seen from this equation that the differences of received power for different users within the same cell only come from the first part of this product, which can be expressed by a class-dependent term ω_t for simple representation as

$$\omega_t = \frac{\varepsilon_t^* R_t}{W + \varepsilon_t^* R_t}. \tag{3.18}$$

The activity factor υ_t is omitted in the above expression as a mobile produces the interference only in its active status. In the later chapter after the load concept is introduced, there will be more discussions on the underlying meaning issues of the variable ω.

Alternatively, the own-cell interference at NodeB x can also be expressed by a sum of received powers $S_{k,x}^R$ from all the UEs connecting to it, which is given by

$$I_x^{own} = \sum_{k\in x} S_{k,x}^R \upsilon_k = \sum_{t=1}^{T} S_{t,x}^R n_{t,x} \upsilon_t \tag{3.19}$$

where the notation $k \in x$ indicates that the UE k is being served by the NodeB x.

Then substitution of (3.17) for $S_{t,x}^{R}$ in the above equation leads to

$$I_x^{own} = (I_x^{own} + I_x^{oc} + WN_0) \cdot \sum_{t=1}^{T} \omega_t n_{t,x} \upsilon_t. \tag{3.20}$$

Solving this for I_x^{own} yields

$$I_x^{own} = \frac{\sum_{t=1}^{T} \omega_t n_{t,x} \upsilon_t}{1 - \sum_{t=1}^{T} \omega_t n_{t,x} \upsilon_t} (I_x^{oc} + WN_0) \tag{3.21}$$

which can be substituted back into (3.17), and the intra-cell received power $S_{t,x}^{R}$ accordingly becomes

$$S_{t,x}^{R} = \frac{\omega_t}{1 - \sum_{t=1}^{T} \omega_t n_{t,x} \upsilon_t} \cdot (I_x^{oc} + WN_0). \tag{3.22}$$

In the multi-cell environment, the received interference at NodeB x comes also from the UEs located in the surrounding cells, which is defined as other-cell interference and denoted as I_x^{oc} in the above equations. As the received power from own-cell users I_x^{own} has been given as a function of I_x^{oc} in (3.21), once the distribution function of I_x^{oc} is determined, the total received interference can be consequently fully characterized, and so is the system outage probability. Therefore, to derive the distribution function of I_x^{oc} is a key task in network planning.

To model the other-cell interference I_x^{oc}, we begin with investigating the inter-cell received power $S_{k,x}^{inter}$, which is the power caused by one UE with class t that does not belong to the target NodeB x (i.e. $NodeB\,(k) \neq x$). This variable depends not only on the intra-cell received power at its serving NodeB (assumed to be y), but also on the propagation attenuation from UE k to the target NodeB x as well as to the serving NodeB y. If we initially suppose that each UE connects to the closest NodeB, then by the suggested log-normal shadowing propagation model, $S_{k,x}^{inter}$ can be represented as

$$S_{k,x}^{inter} = S_{k,y}^{R} \left(\frac{d_{k,y}}{d_{k,x}}\right)^{\gamma} 10^{(\zeta_{k,x} - \zeta_{k,y})/10} = S_{t,y}^{R} \Delta_{y \to x}^{k} 10^{\zeta_k/10} \tag{3.23}$$

where $d_{k,x}$ and $d_{k,y}$ refer to the distance from UE k to NodeB x and y, respectively.

ζ_k is the difference of two independent Gaussian r.v. with zero mean and standard deviation σ, and thus is a zero mean Gaussian r.v. with standard deviation $\sqrt{2}\sigma$. For simple expression, we use $\Delta^k_{y\to x}$ to denote the attenuation ratio $(d_{k,y}/d_{k,x})^\gamma$, which depends on the location of UE k only.

With the inclusion of shadowing effects, not only does the received power become more varied, but even the cell membership becomes a more complicated issue than before, since now the NodeB with the least distance is not always to be the one with the least propagation attenuation. In theory, any UE connects to the NodeB with the strongest received power, however to be practically feasible, it only selects the one with the least attenuation from a limited number of closest candidate NodeBs. This scenario has been studied in [66], which leads to much complicated computations and thus only mean values are obtained. To reduce the complexity, we follow the same approach as in [20, 26], where the choice is made merely between the closest NodeB and the target NodeB, thus the inter-cell received power would take the same value as either in (3.23) if controlled by the target NodeB, or simply the own-cell received power if controlled by the closest NodeB. In order to include such effects into our model, the received power at each cell are assumed to be equal to each other as in [20, 26] at this stage so that the min operator can be applied. The inter-cell received power under shadowing effects in (3.23) can be rewritten as

$$S^{inter}_{k,x} = S^R_{t,y} \cdot \min\left[1, \Delta^k_{y\to x} 10^{\zeta_k/10}\right]. \tag{3.24}$$

Then the other-cell interference I^{oc}_x is the sum of inter-cell received power $S^{inter}_{k,x}$ from the UEs connecting to all the NodeBs in the network other than x:

$$I^{oc}_x = \sum_{y\neq x}\sum_{k\in y} \upsilon_k S^{inter}_{k,x}. \tag{3.25}$$

Hence, if we substitute (3.22) and (3.24) into (3.25), a set of N equations, where N is the total number of NodeBs in the system, with respect to the variables I^{oc} can be constructed. For easy representation, two new variables, $I^{out}_{y\to x}$ and $F_{y\to x}$, are introduced, where $I^{out}_{y\to x}$ denotes the part of inter-cell interference experienced at NodeB x caused by all the UEs associated with NodeB y, and $F_{y\to x}$ is defined as a

coefficient as follows:

$$F_{y\to x} = \frac{1}{1-\sum_{m=1}^{T}\omega_m n_{m,y}\upsilon_m} \cdot \sum_{t=1}^{T}\omega_t\upsilon_t \sum_{k\in y}\min\left[1, \Delta_{y\to x}^{k} 10^{\zeta_k/10}\right]. \tag{3.26}$$

Then the above equations regarding to other-cell interference can be rewritten as

$$I_x^{oc} = \sum_{y\neq x} I_{y\to x}^{out} \tag{3.27}$$

and

$$I_{y\to x}^{out} = \left(I_y^{oc} + WN_0\right) \cdot F_{y\to x}. \tag{3.28}$$

Let $\bar{I}$ be a row vector where $\bar{I} = [I_1^{oc}, \ldots, I_N^{oc}]$, $\bar{N}$ be a constant row vector with all entries equaled to WN_0, and $\tilde{F}$ be a matrix where $\tilde{F}[x,y] = F_{x\to y}$ when $x \neq y$ and all the diagonal entries $\tilde{F}[x,x] = 0$, then the above linear equations (3.27) and (3.28) can be formulated as a fixed-point matrix equation

$$\bar{I} = (\bar{I} + \bar{N}) \cdot \tilde{F} \tag{3.29}$$

and solved for the other-cell interference vector $\bar{I}$ by matrix inversion

$$\bar{I} = \bar{N}\tilde{F}(\tilde{E} - \tilde{F})^{-1} \tag{3.30}$$

where $\tilde{E}$ is the identity matrix.

As this set of linear equations are constructed based on the cells in the mobile network, where each row in the matrix $(\tilde{E} - \tilde{F})$ represents the interference coefficient for one cell of the system, it can be safely assumed that they are linear-independent of each other, and thus such matrix is non-singular, which implies we can have one and only one solution for the interference vector from the equations above.

From above equations, it can be seen that the variable $F_{y\to x}$ in (3.26), which consists of user population and location information, is independent of I_x^{oc} and $I_{y\to x}^{out}$. In each scenario generated in the Monte Carlo simulation, all these information are deterministic as $n_{t,x}$, $\Delta_{x\to y}^{k}$ and ζ_k are all given, the variable $F_{y\to x}$ becomes a constant accordingly. Then in order to evaluate the other-cell interference, we may either

construct a system of linear equations with respect to I_x^{oc} with rank equal to the number of NodeBs based on (3.27) and (3.28), or solve it directly from the matrix equation as in (3.29) and (3.30). If sufficient samples can be obtained, the distribution of I_x^{oc} can be approximately characterized. Once the other-cell interference is well modeled, the other system parameters, such as own-cell interference, received power at NodeB, and transmit power of the UEs, can be consequently derived.

The Monte Carlo simulation is easy to implement, and thus is quite a flexible approach since we can easily change the employed propagation model, traffic model, cell layout and even resource management schemes. However, as each time a large number of samples are required in order to determine the distribution of interested variables, it is not efficient for the network planning tools. Thus we mainly use this technique to verify the proposed analytical models in the following chapters.

3.4 An Iterative-Based Analytical Approach

In the above section, we have already discussed the importance of the other-cell interference component, which is also the most difficult part to determine in the network planning process. The arrival and departure events of each mobile user in any location in the coverage area of UMTS network, or even a short movement of a single user, might cause the fluctuation of this value, since the other-cell interference of one cell depends on the transmit power and location of every mobile user in the network. In the other way around, a fluctuation in the other-cell interference level of a certain cell results in corresponding variation in transmit power of all the mobiles in this cell in order to satisfy the power control equation, which in turn influences the other-cell interference received at all neighboring NodeBs and then the interference from surrounding cells changes again. This is due to the 'feedback behavior' defined in [59], which makes the modeling task not straightforward any more. In this section, we propose an iterative approach in order to capture such behavior and suggest a pure analytical model for the other-cell interference in UMTS uplink under stochastic user patterns.

3.4.1 Problem Formulation and Iterative Approach

In the UMTS uplink, the interference received at each NodeB comprises the own-cell interference and the other-cell interference according to different membership of each contributing UE. From (3.22), we can see that the intra-cell received power $S^R_{t,x}$ at NodeB x depends on the other-cell interference I^{oc}_x experienced at the same NodeB by mathematical transformation of the power control equation. The required received power in turn determines the transmit power of all the UEs power controlled by this NodeB, which received at the neighboring NodeBs (e.g. NodeB y) as other-cell interference I^{oc}_y. Then the variation of I^{oc}_y may affect the transmit power of all the UEs connecting to NodeB y, which again lead to an updated I^{oc}_x. Thus we can see the other-cell interference at different NodeB mutually depends on each other, where the relation is illustrated in Fig. 3.3 and characterized by the stochastic fixed-point equations representation of (3.27) and (3.28) as

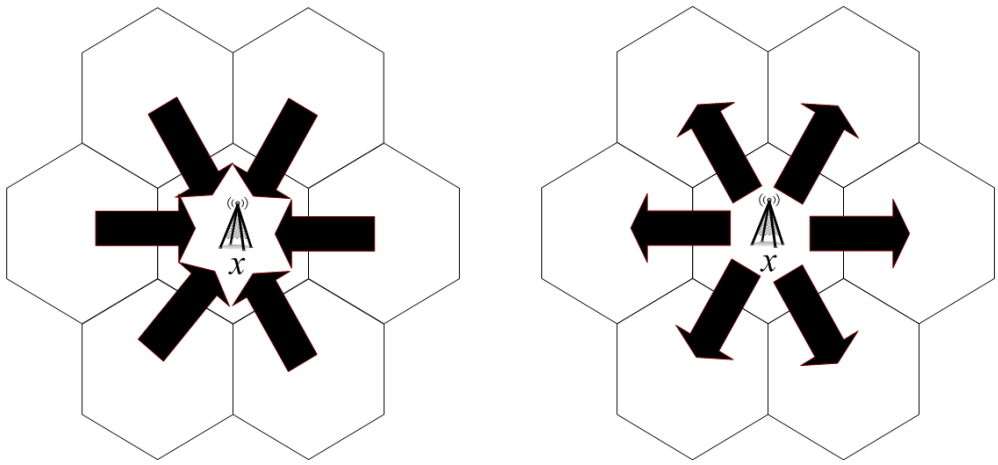
$$\mathcal{I}^{oc}_x = \sum_{y \neq x} \mathcal{I}^{out}_{y \to x} \tag{3.31}$$

and

$$\mathcal{I}^{out}_{y \to x} = \left(\mathcal{I}^{oc}_y + W N_0\right) \mathcal{F}_{y \to x}. \tag{3.32}$$

where $\mathcal{I}^{oc}_x$, $\mathcal{I}^{out}_{y \to x}$ and $\mathcal{F}_{y \to x}$ are the corresponding random variables.

In theory, the distribution function of other-cell interference $\mathcal{I}^{oc}$ can be derived from the above fixed-point stochastic equations based on an iterative approach if the distribution of $\mathcal{F}_{y \to x}$ is known. In the first iteration, we can begin with setting the random variable $\mathcal{I}^{oc}_x$ to zero, and thus calculating the distribution of $\mathcal{I}^{out}_{y \to x}$ for all possible pairs of x and y from (3.32) without considering $\mathcal{I}^{oc}_y$ in it, followed by updating the distribution for $\mathcal{I}^{oc}_x$ from the sum of $\mathcal{I}^{out}_{y \to x}$ over all NodeBs other than x according to (3.31). Then in the next and following iterations, we can substitute the updated distribution of the other-cell interference $\mathcal{I}^{oc}_y$ from previous iteration into (3.32) and determine a new distribution for $\mathcal{I}^{out}_{y \to x}$, where the rest of computation is same as the prior iteration (i.e. update distribution for $\mathcal{I}^{oc}_x$). The distribution function of $\mathcal{I}^{oc}$ would finally converge if the relative error falls below a certain threshold. To employ this approach, the prerequisite is a fully characterization of the random variable $\mathcal{F}_{y \to x}$,



(a) Interference received at cell x (b) Interference generated from cell x

Figure 3.3: Interference relationship among neighboring NodeBs

which is dealt with in the following part.

3.4.2 Derivation of $\mathcal{F}_{y\to x}$

With stochastic user patterns, the user population and distribution become random variables with distribution functions specified by the traffic model, and all these random factors are included in the random variable $\mathcal{F}_{y\to x}$, which is the stochastic representation of the variable $F_{y\to x}$ in (3.26) as

$$\mathcal{F}_{y\to x} = \frac{1}{1-\sum_{m=1}^{T} n_{m,y}\omega_m \upsilon_m} \cdot \sum_{t=1}^{T} n_{t,y}\omega_t \upsilon_t \cdot \min\left[1, \Delta_{y\to x} 10^{\zeta/10}\right] \tag{3.33}$$

with all the expressions inside such as $n_{t,y}$, $\Delta_{y\to x}$ become corresponding random variables now.

To characterize this random variable, we can first focus on the conditional distribution by fixing the number of users in each class, and then uncondition it as a sum over all the possible user number combinations in cell y according to the total probability theorem

$$P\left(\mathcal{F}_{y\to x} \le z\right) = \sum_{\bar{n}_y \in \mathbf{S}} P(\bar{n}_y) \cdot P\left(\mathcal{F}_{y\to x} \le z | \bar{n}_y\right) \tag{3.34}$$

where $P(\bar{n}_y)$ is the probability that $(n_{1,y}, \ldots, n_{T,y})$ active UEs connecting to NodeB y, which is given as a product form in (3.13)-(3.15).

For the conditional distribution $P(\mathcal{F}_{y\to x} \leq z|\bar{n}_y)$, since now the number of users in each class are given, the only part that remains undetermined is $\mathcal{D}_{y\to x} = \min\left[1, \Delta_{y\to x}10^{\zeta/10}\right]$, which depends on the individual user location within one cell and shadowing factor. To characterize this random variable, we can employ the approach outlined in [20] where a similar random variable is modeled based on integration of the conditional distribution function. By this mean we can obtain the distribution of $\mathcal{D}_{y\to x}$, and calculate the first and second moments for later approximation use. For clarity, we first focus on the distribution function of $\Delta_{y\to x}$, then extend it to that of $\mathcal{D}_{y\to x}$.

With the assumption that users are uniformly distributed within one cell, which is equivalent to the assumed traffic model under spatial homogeneous Poisson process, this problem can be simplified by approximating the hexagonal cells by circles with radius b and consequently the coordinate system is converted from Cartesian to polar. The reason for such conversion is to simplify the computation complexity which is caused by diverse orientations of a hexagonal cell as well as the dependence of coordinates. Each mobile user is assigned a pair of values (r, ϕ) to represent its current location with distance and corresponding phase to the serving NodeB. Then if the UE at location (r, ϕ) is served by NodeB y at the origin, and the target NodeB x is at location (a, π) (i.e. the distance between these two NodeBs is a), the interested random variable $\Delta^a_{y\to x}$ becomes

$$\Delta^a_{y\to x}(r, \phi) = \left(\frac{r}{\sqrt{r^2 + a^2 + 2ar\cos\phi}}\right)^{\gamma} \tag{3.35}$$

where the random variables r and ϕ in the above formula have independent distribution functions as

$$P(r \leq z) = \frac{z^2}{b^2} \tag{3.36}$$

and

$$P(\phi \leq z) = \frac{z}{\pi}. \tag{3.37}$$

Both r and ϕ are defined in the range $[0, b]$ and $[0, \pi]$ respectively since we only need

consider a semi-circle due to symmetry.

As the distribution functions of both random variables in (3.35) have been given, the modeling of $\Delta^a_{y\to x}$ is not a hard task. By firstly fixing ϕ and calculating the conditional distribution function $P\left(\Delta^a_{y\to x}\leq z|\phi\right)$ and then unconditioning it with a definite integral on ϕ from 0 to π, the distribution function of $\Delta^a_{y\to x}$ is given same as in [20]

$$P\left(\Delta^a_{y\to x}\leq z\right)=\begin{cases}0, & z<0\\ g_1\left(z\right), & 0\leq z<\left(\frac{a}{b}+1\right)^{-\gamma}\\ g_2\left(z\right), & \left(\frac{a}{b}+1\right)^{-\gamma}\leq z<\left(\frac{a}{b}-1\right)^{-\gamma}, z\neq 1\\ g_3\left(z\right), & 1<\left(\frac{a}{b}-1\right)^{-\gamma},\ z=1\\ 1, & \left(\frac{a}{b}-1\right)^{-\gamma}\leq z\end{cases} \tag{3.38}$$

where

$$\begin{aligned}g_1\left(z\right)&=\frac{a^2z^{2/\gamma}}{b^2\left(z^{2/\gamma}-1\right)^2}\\ g_2\left(z\right)&=\frac{1}{\pi}\arccos\left(h_1\left(z\right)\right)+\frac{1}{\pi}g_1\left(z\right)\Big[\pi-2z^{2/\gamma}h_1\left(z\right)h_2\left(z\right)\\ &\quad-\frac{a}{b}z^{2/\gamma}h_2\left(z\right)-\arccos\left(h_1\left(z\right)\right)-\arcsin\left(z^{1/\gamma}h_2\left(z\right)\right)\Big]\\ g_3\left(z\right)&=\frac{1}{\pi}\arccos\left(-\frac{a}{2b}\right)+\frac{a}{4\pi b^2}\sqrt{4b^2-a^2}\end{aligned} \tag{3.39}$$

and

$$\begin{aligned}h_1\left(z\right)&=\frac{-a^2-b^2+b^2z^{-2/\gamma}}{2ab}\\ h_2\left(z\right)&=\sqrt{1-h_1^2\left(z\right)}.\end{aligned} \tag{3.40}$$

The last thing to determine is the value of approximating circle radius b with the distance between two adjacent NodeBs normalized to 1. A few possible options for b have been proposed in [20], where $b=1/\sqrt{3}$ such that the circle contains the hexagon, $b=1/2$ such that the circle is contained in the hexagon, and $b=\sqrt[4]{3}/\sqrt{2\pi}\approx 0.53$ such that the area of circle is equal to that of the hexagon. Following the same way in [20], the approximation results under all three b values and same NodeB distance $a=1$ are

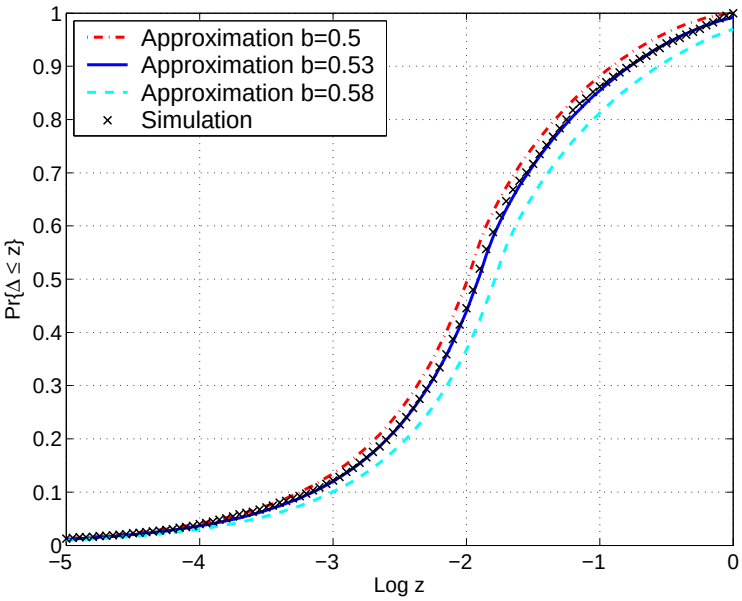


Figure 3.4: Comparison of b values

compared with the Monte Carlo simulation results and presented in Fig. 3.4, with the conclusion that the option $b = 0.53$ gives the best approximation accuracy. Fig. 3.5 verifies the chosen b value in approximation with various NodeB distances $a = 1$, $a = \sqrt{3}$ and $a = 2$ (i.e. all the possible distances between any NodeB and the central one in a two-tier layout) and excellent matches are achieved, which implies the distribution function in (3.38) can well model the random variable $\Delta_{y\to x}$. The PLE value in the experiments is assumed to be a constant as $\gamma = 4$.

Next in order to compute the complete distribution function of $\mathcal{D}_{y\to x}$, we begin with deriving the distribution of $\hat{\mathcal{D}}_{y\to x} = \Delta_{y\to x} 10^{\zeta/10}$, followed by applying the min operator to get the final form for $\mathcal{D}_{y\to x}$. Since the randomness resulted from different locations has been characterized by the above result, we can first give the distribution function of $\hat{\mathcal{D}}_{y\to x}$ conditioned on known $\Delta_{y\to x}$ as

$$P\left(\hat{\mathcal{D}}_{y\to x} \leq z | \Delta_{y\to x}\right) = Q\left(\frac{\ln z - \ln \Delta_{y\to x}}{\sqrt{2}\beta\sigma}\right) \tag{3.41}$$

where

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} \mathbf{d}t, \tag{3.42}$$

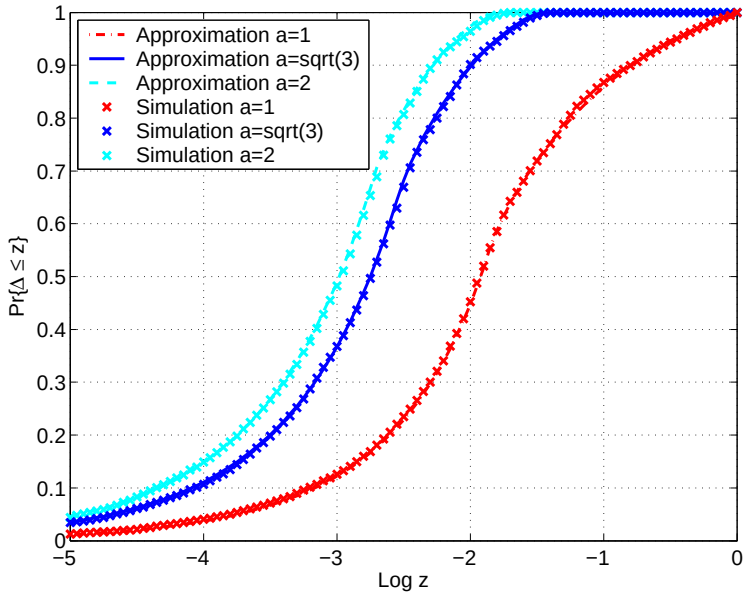


Figure 3.5: Approximation vs. simulation results with different NodeB distances

$\beta = \ln 10/10$ and $\sqrt{2}\sigma$ is the standard deviation of ζ which is the difference of two independent zero mean Gaussian random variables defined as before.

Then (3.41) can be unconditioned with known distribution function of Δ in (3.38)

$$\begin{aligned} P\left(\hat{\mathcal{D}}_{y\to x}\leq z\right) = &\int_0^{\left(\frac{a}{b}+1\right)^{-\gamma}} Q\left(\frac{\ln z - \ln \Delta_{y\to x}}{\sqrt{2}\beta\sigma}\right) g_1'(\Delta)\,\mathbf{d}\Delta \\ &+ \int_{\left(\frac{a}{b}+1\right)^{-\gamma}}^{\left(\frac{a}{b}-1\right)^{-\gamma}} Q\left(\frac{\ln z - \ln \Delta_{y\to x}}{\sqrt{2}\beta\sigma}\right) g_2'(\Delta)\,\mathbf{d}\Delta. \end{aligned} \tag{3.43}$$

And eventually with inclusion of the min operator the complete distribution function of $\mathcal{D}_{y\to x}$ can be expressed as

$$P\left(\mathcal{D}_{y\to x}\leq z\right) = \begin{cases} 0, & z < 0 \\ P\left(\hat{\mathcal{D}}_{y\to x}\leq z\right), & 0 \leq z < 1 \\ 1, & 1 \leq z \end{cases} \tag{3.44}$$

The distribution of $\mathcal{F}_{y\to x}$ can be fully derived in theory since all the random variables in (3.34) have been characterized. Then the iterative approach introduced

early in this section can be used to analytically model the other-cell interference $\mathcal{I}^{oc}$. However, it is quite a hard task since the computation would involve many convolutions and is numerically intractable. Hence in the next section some approximation techniques would be employed to reduce the computational complexity significantly.

3.5 Log-Normal Approximation

As pointed out in the last section, the direct approach for the characterization of other-cell interference is quite time-consuming, therefore, we consider applying approximation techniques to solve this problem. Through some trial experiments, the normal probability plot in Fig. 3.6 demonstrates comparison between the simulation results of the logarithm of $\mathcal{F}_{y\to x}$ and the corresponding normal distribution. In such normal probability figure, the closer the dashed line to the solid line, the more similar the distribution of the simulation results to the normal distribution. From the figure, it is shown that the logarithm of majority samples of $\mathcal{F}_{y\to x}$ match the normal distribution very well, which implies $\mathcal{F}_{y\to x}$ itself can be well approximated by the log-normal distribution. This reduces the problem down to determining the mean and variance of $\mathcal{F}_{y\to x}$ since only these two parameters are enough to characterize a log-normal random variable. To calculate both moments of $\mathcal{F}_{y\to x}$, we would firstly fix the number of users served by NodeB y in each class as $n_{t,y}$ and develop the conditional moments $\mathcal{F}_{y\to x}(\bar{n}_y)$ from (3.33) as

$$E\left[\mathcal{F}_{y\to x}(\bar{n}_y)\right] = \sum_{t=1}^{T} n_{t,y} \cdot \frac{\omega_t \upsilon_t \cdot E\left[\mathcal{D}_{y\to x}\right]}{1-\sum_{m=1}^{T} n_{m,y}\omega_m \upsilon_m}, \tag{3.45}$$

$$E\left[\mathcal{F}_{y\to x}(\bar{n}_y)^2\right] = VAR\left[\mathcal{F}_{y\to x}(\bar{n}_y)\right] + E\left[\mathcal{F}_{y\to x}(\bar{n}_y)\right]^2 \tag{3.46}$$

where the conditional variance is given by

$$VAR\left[\mathcal{F}_{y\to x}(\bar{n}_y)\right] = \sum_{t=1}^{T} n_{t,y} \cdot \frac{(\omega_t \upsilon_t)^2 \cdot VAR\left[\mathcal{D}_{y\to x}\right]}{\left(1-\sum_{m=1}^{T} n_{m,y}\omega_m \upsilon_m\right)^2}. \tag{3.47}$$

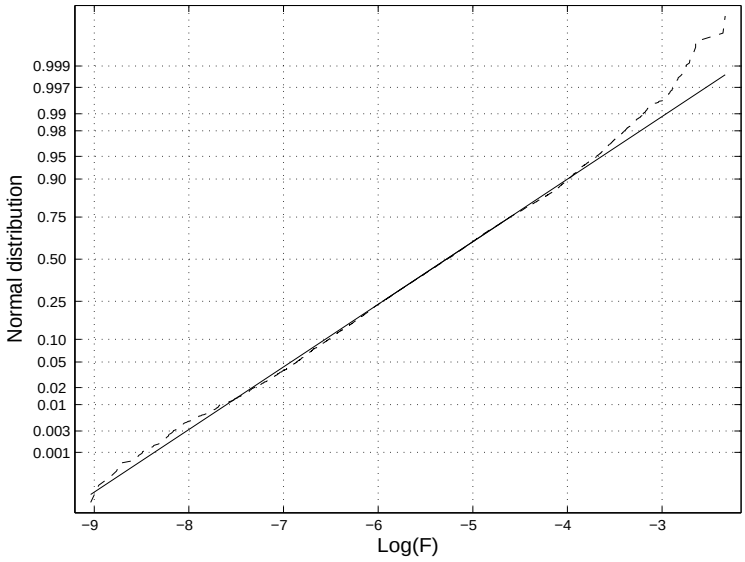


Figure 3.6: Normal probability plot of the logarithm of $\mathcal{F}$

Then again applying the theorem of total probabilities to characterize the stochastic number of users, the first and second moments of $\mathcal{F}_{y\to x}$ are given as

$$E\left[\mathcal{F}_{y\to x}\right] = \sum_{\bar{n}_y \in \mathbf{S}} P(\bar{n}_y) E\left[\mathcal{F}_{y\to x}(\bar{n}_y)\right], \tag{3.48}$$

$$E\left[\mathcal{F}^2_{y\to x}\right] = \sum_{\bar{n}_y \in \mathbf{S}} P(\bar{n}_y) E\left[\mathcal{F}_{y\to x}(\bar{n}_y)^2\right] \tag{3.49}$$

where $P(\bar{n}_y)$ and $\mathbf{S}$ are the corresponding user distribution and admissible region in (3.13) and (3.15), respectively. And finally the variance of $\mathcal{F}_{y\to x}$ can be simply derived from these two moments as

$$VAR\left[\mathcal{F}_{y\to x}\right] = E\left[\mathcal{F}^2_{y\to x}\right] - E\left[\mathcal{F}_{y\to x}\right]^2. \tag{3.50}$$

The mean and variance of random variable $\mathcal{D}_{y\to x}$ appeared in the above equations can be computed via its Cumulative Distribution Function (CDF) derived before in (3.38) - (3.44). One thing which is worth noticing is that during the calculation of moments of $\mathcal{D}_{y\to x} \in [0, 1]$, it is hard to compute the derivative of (3.43) with respect

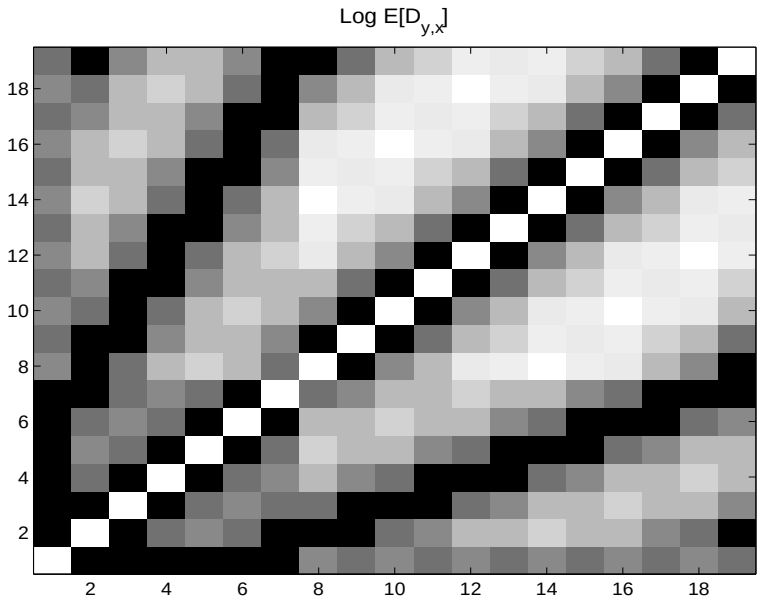


Figure 3.7: Mean attenuation ratios in dB

to z directly, therefore we use the following alternative equations instead

$$\begin{aligned} E\left[\mathcal{D}_{y\to x}\right] &= \int_0^1 zF'_{\mathcal{D}}\left(z\right)\mathbf{d}z \\ &= zF_{\mathcal{D}}\left(z\right)\big|_0^1 - \int_0^1 F_{\mathcal{D}}\left(z\right)\mathbf{d}z, \end{aligned} \tag{3.51}$$

$$\begin{aligned} E\left[\mathcal{D}^2_{y\to x}\right] &= \int_0^1 z^2F'_{\mathcal{D}}\left(z\right)\mathbf{d}z \\ &= z^2F_{\mathcal{D}}\left(z\right)\big|_0^1 - 2\int_0^1 zF_{\mathcal{D}}\left(z\right)\mathbf{d}z \end{aligned} \tag{3.52}$$

where $F_{\mathcal{D}}\left(z\right)$ refers to the distribution function. The mean value results of $\mathcal{D}_{y\to x}$ for each possible NodeB pair under a two-tier cell layout are illustrated by the color intensity in Fig. 3.7. The darker squares correspond to larger mean attenuation ratios $E[\mathcal{D}_{y\to x}]$ between two NodeBs specified by both subscripts where the cell numbers follow Fig. 3.2, and vice versa.

Now that both moments of $\mathcal{F}$ are obtained, we can progress to the approximation of $\mathcal{I}^{oc}$. The excellent log-normal approximation of $\mathcal{F}$ suggests that $\mathcal{I}^{oc}$ might also

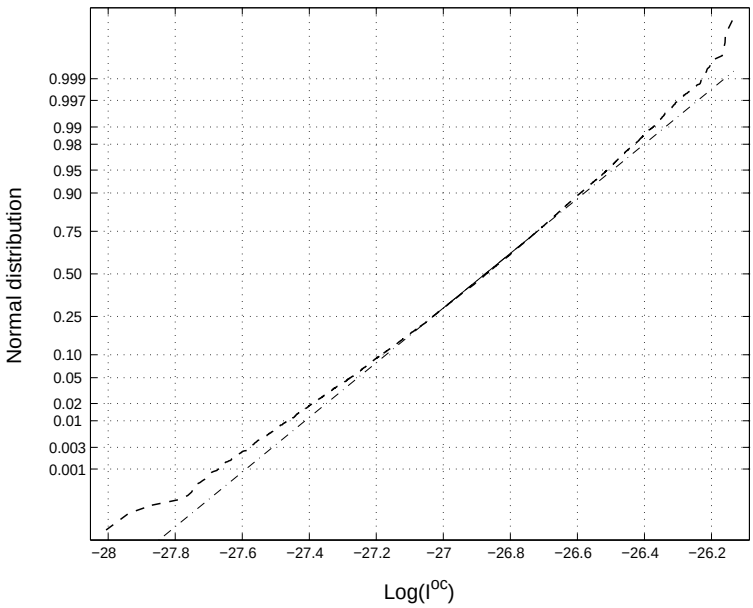


Figure 3.8: Normal probability plot of the logarithm of $\mathcal{I}^{oc}$

follow a log-normal distribution. This is verified through Monte Carlo simulations and presented in Fig. 3.8, besides similar conclusion has been drawn in [58] as well. This implies again the mean and variance suffice for characterization of $\mathcal{I}^{oc}$. To calculate both moments, we can follow a similar iterative approach described in the previous section.

In the first step, the random variable $\mathcal{I}_x^{oc}$ is set to zero, thus $\mathcal{I}_{y\to x}^{out}$ is simply the product of WN_0 and $\mathcal{F}_{y\to x}$, which in turn follows log-normal distribution. The mean and variance can be easily determined from (3.32) as

$$E\left[\mathcal{I}_{y\to x}^{out}\right] = WN_0 \cdot E\left[\mathcal{F}_{y\to x}\right], \tag{3.53}$$

$$VAR\left[\mathcal{I}_{y\to x}^{out}\right] = (WN_0)^2 \cdot VAR\left[\mathcal{F}_{y\to x}\right]. \tag{3.54}$$

Then following the same assumption in [59] that $\mathcal{I}_{y\to x}^{out}$ are independent of each other for all pairs $x \neq y$, the other-cell interference $\mathcal{I}_x^{oc}$, which comprises the sum of a series of log-normally distributed random variables based on (3.31), can also be approximated

by a log-normal random variable with both parameters as

$$E\left[\mathcal{I}_x^{oc}\right] = \sum_{y \neq x} E\left[\mathcal{I}_{y\to x}^{out}\right], \tag{3.55}$$

$$VAR\left[\mathcal{I}_x^{oc}\right] = \sum_{y \neq x} VAR\left[\mathcal{I}_{y\to x}^{out}\right]. \tag{3.56}$$

With the updated moments of $\mathcal{I}_x^{oc}$, $\mathcal{I}_{y\to x}^{out}$ become the product of two log-normal random variables, $(WN_0 + \mathcal{I}_y^{oc})$ and $\mathcal{F}_{y\to x}$, which again can be regarded as a log-normal random variable in all the following iterations. The multiplication is performed by summing up the corresponding parameters as

$$\mu_{\mathcal{I}_{y\to x}^{out}} = \mu_{\left(WN_0 + \mathcal{I}_y^{oc}\right)} + \mu_{\mathcal{F}_{y\to x}}, \tag{3.57}$$

$$\sigma^2_{\mathcal{I}_{y\to x}^{out}} = \sigma^2_{\left(WN_0 + \mathcal{I}_y^{oc}\right)} + \sigma^2_{\mathcal{F}_{y\to x}} \tag{3.58}$$

where μ_X and σ_X^2 are the median and variance of the random variable X's logarithm, which can be calculated from the mean and variance of X as

$$\sigma_X^2 = \ln\left(\frac{VAR[X]}{E[X]^2} + 1\right), \tag{3.59}$$

$$\mu_X = \ln\left(E[X]\right) - \frac{\sigma_X^2}{2}. \tag{3.60}$$

Since the mean and variance of $\mathcal{F}_{y\to x}$ are known, we can substitute the moments of $\mathcal{I}_y^{oc}$ from the previous step into calculation, and then compute the μ and σ^2 parameters of $\mathcal{I}_{y\to x}^{out}$, which consequently lead to $E[\mathcal{I}_{y\to x}^{out}]$ and $VAR[\mathcal{I}_{y\to x}^{out}]$ from an inversion of (3.59) and (3.60) as

$$E[X] = e^{\mu_X + \frac{\sigma_X^2}{2}}, \tag{3.61}$$

$$VAR[X] = E[X]^2\left(e^{\sigma_X^2} - 1\right). \tag{3.62}$$

Again the mean and variance of $\mathcal{I}_x^{oc}$ can be determined by (3.55) and (3.56). In the following, each iteration can be summarized as

1. Compute $\mu_{\mathcal{I}_x^{oc}}$ and $\sigma^2_{\mathcal{I}_x^{oc}}$ according to (3.59) and (3.60) for all NodeB x from updated $E[\mathcal{I}_x^{oc}]$ and $VAR[\mathcal{I}_x^{oc}]$ in the previous step;

2. Compute $\mu_{\mathcal{I}^{out}_{y\to x}}$ and $\sigma^2_{\mathcal{I}^{out}_{y\to x}}$ according to (3.57) and (3.58) for all possible NodeB pairs x and y;

3. Compute $E[\mathcal{I}^{out}_{y\to x}]$ and $VAR[\mathcal{I}^{out}_{y\to x}]$ according to (3.61) and (3.62) for all possible NodeB pairs x and y;

4. Compute $E[\mathcal{I}^{oc}_x]$ and $VAR[\mathcal{I}^{oc}_x]$ according to (3.55) and (3.56) for all NodeB x.

These steps would be repeated until the relative change of interested parameters fall below certain thresholds. The mean and variance of other-cell interference are then obtained after convergence.

3.6 Outage Probability Analysis

In mobile networks, the system capacity is always defined as the maximum number of mobile users that the system can support with the probability of an outage event occurring kept below a given threshold. Therefore, once the outage probability is known, the derivation for system capacity is straightforward. In this section, we demonstrate how to compute the CDF of system outage probability, which is an important application of the other-cell interference characterization in the network planning process.

In [20], the outage probability is introduced as a QoS indicator for capacity analysis in radio network planning. It is assumed in this work that received power of UEs at all their serving NodeBs are controlled at the same level, thus the outage probability is defined as the probability that a UE receives an insufficient SIR, which can be easily translated into a constraint on the total interference at one NodeB. Then the calculation of outage probability reduces to the evaluation of the probability that a compound Poisson random variable exceeds a certain threshold.

However, in our current model, since it is presumed that power control aims to maintain E_b/I_0 for each UE to be constant, thus the required received power increases as the number of UEs grows. A UE can always satisfy the requirement for acceptable QoS unless the requested transmission power, which is proportional to the received power at NodeB, goes beyond its capability. In UMTS networks where multiple classes

are supported, the outage probability is approximately defined by the received power as

$$P_{out} = 1 - \Pr\left\{S_{1,x}^R < S^*, S_{2,x}^R < S^*, \ldots, S_{T,x}^R < S^*\right\} \tag{3.63}$$

where S^* represents a certain received power threshold and $S_{t,x}^R$ is given in (3.22).

From (3.22), it can be seen that within a given cell, the received power $S_{t,x}^R$ of each class only differ from each other in ω_t. Hence, if the received power of the user class with maximum ω_t is less than the threshold, the other classes must also satisfy the requirements. Then the outage probability can be rewritten as

$$\begin{aligned} P_{out} &= 1 - \Pr\left\{\max_{t\in T}\left|S_{t,x}^R\right| < S^*\right\} \\ &= 1 - \Pr\left\{S_{\tilde{t},x}^R < S^*, \tilde{t} : \max_{t\in T}|\omega_t|\right\}. \end{aligned} \tag{3.64}$$

Based on the expression for $S_{\tilde{t},x}^R$ in (3.22), the above probability can be firstly conditioned on the given user distribution, followed by applying total probability theorem

$$\Pr\left\{S_{\tilde{t},x}^R < S^*\right\} = \sum_{\bar{n}_x\in\mathbf{S}} P\left(\bar{n}_x\right)\Pr\left\{\mathcal{I}_x^{oc} < S^* \cdot \frac{\left(1 - \sum\limits_{m=1}^{T} n_{m,x}\omega_m\upsilon_m\right)}{\omega_{\tilde{t}}} - WN_0\right\} \tag{3.65}$$

with the notation $P(\bar{n}_x)$ and $\mathbf{S}$ referred to (3.13) and (3.15).

Since the right hand of the above inequality is a constant number, such probability can be easily determined with acquired log-normal distribution of $\mathcal{I}^{oc}$. Under some user patterns, the right hand of the above inequality might be less than zero, in which cases the probability of this summand component simply takes zero since the other-cell interference is always positive. In most cases when the right hand expression is larger than zero, natural logarithm can be applied on both sides. Because $\ln\left[\mathcal{I}^{oc}\right]$ follows standard Gaussian distribution with known parameters $\mu_{\mathcal{I}^{oc}}$ and $\sigma_{\mathcal{I}^{oc}}$, the probability

inside the summand component is actually the tail of Gaussian distribution

$$\Pr\left\{\ln\left[\mathcal{I}^{oc}\right] < \ln\left[S^{*}\cdot\frac{\left(1-\sum\limits_{m=1}^{T} n_{m,x}\omega_m \upsilon_m\right)}{\omega_{\tilde{t}}} - WN_0\right]\right\}$$
$$= Q\left(\frac{\ln\left[S^{*}\cdot\frac{\left(1-\sum\limits_{m=1}^{T} n_{m,x}\omega_m \upsilon_m\right)}{\omega_{\tilde{t}}} - WN_0\right] - \mu_{\mathcal{I}^{oc}}}{\sigma_{\mathcal{I}^{oc}}}\right) \tag{3.66}$$

where $Q(x)$ is the function given in (3.42).

Hence, together with the parameters $\mu_{\mathcal{I}^{oc}}$ and $\sigma_{\mathcal{I}^{oc}}$ of other-cell interference derived in the previous section, the outage probability can be finally characterized, which translates to system capacity by definition.

3.7 Models Validation

For numerical tractability, we evaluate the system performance in a two-tier hexagonal cell ring area which consists of 19 NodeBs. Firstly the parameters of interested random variables can be calculated numerically via the suggested analytical models, then the validation of such analytical models can be performed by Monte Carlo simulations described above, where the user patterns are generated randomly according to spatial homogeneous Poisson processes, followed by solving a set of linear equations to compute the other-cell interference for each NodeB. To avoid the border effect due to less neighbors for the cells located at the boundary, only the sample values at the central cell are counted and compared. The system parameters are assumed as follows: system chip rate $W = 3.84$MHz, background thermal noise density $N_0 = -108$dBm, and PLE $\gamma = 4$.

We start with investigation on the system performance with single-class service only, with operation parameters as in Table. 3.1 below. Fig. 3.9 and Fig. 3.10 illustrate below the mean and standard deviation numerical results comparisons from analytical

Table 3.1: Operation parameters of the single-class service model

Bit Rate $[R]$	Target E_b/I_0 $[\varepsilon^*]$	Activity $[\upsilon]$
12.2kbps	5.5dB	1

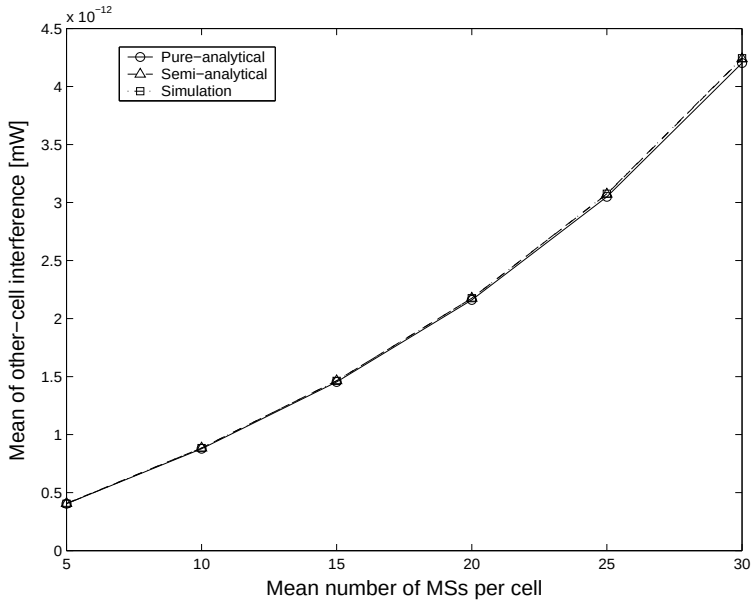


Figure 3.9: Comparison of mean other-cell interference with single service

model, semi-analytical model and simulation, respectively. Here the semi-analytical model refers to the one in which moments of random variable $\mathcal{D}$ are acquired empirically, while in pure-analytical model they are derived analytically.

From the above figures, we can declare the analytical results and simulation results achieve excellent matches with single service class in both first and second moments. The maximum relative error is no more than 1% for the mean and 3% for the standard deviation. The obtained parameters for other-cell interference can be applied into the outage probability analysis, where the analytical and simulation results are presented below. In Fig. 3.11, both outage probability results are calculated with maximum received power $S^* = -125$dBm, while in Fig. 3.12, how the outage probabilities vary with different S^* are presented. It can be seen in all these cases the outage probabilities are accurately characterized by the suggested analytical model.

Next we present the results with multiple service classes. Three service classes,

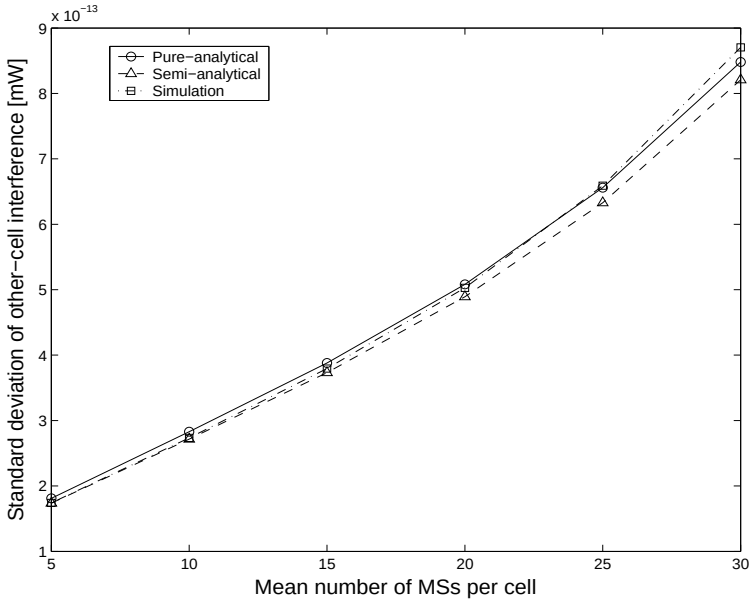


Figure 3.10: Comparison of standard deviation of other-cell interference with single service

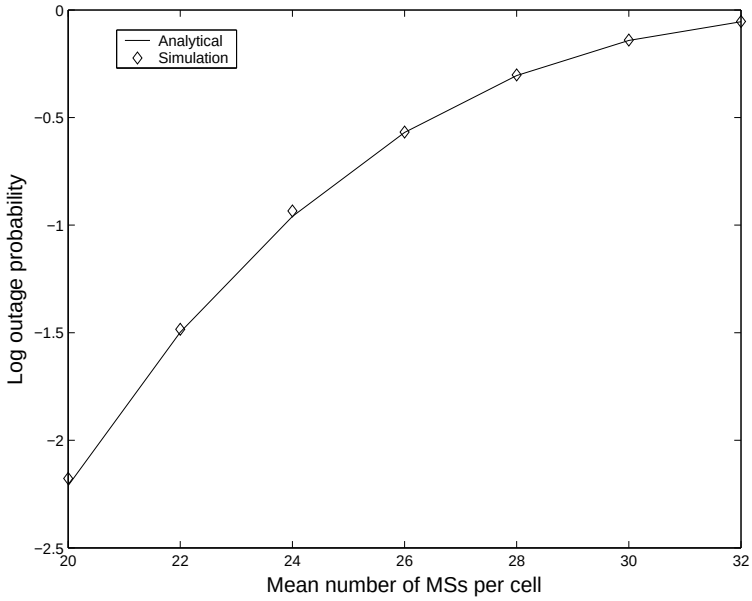


Figure 3.11: Logarithm outage probability with single service

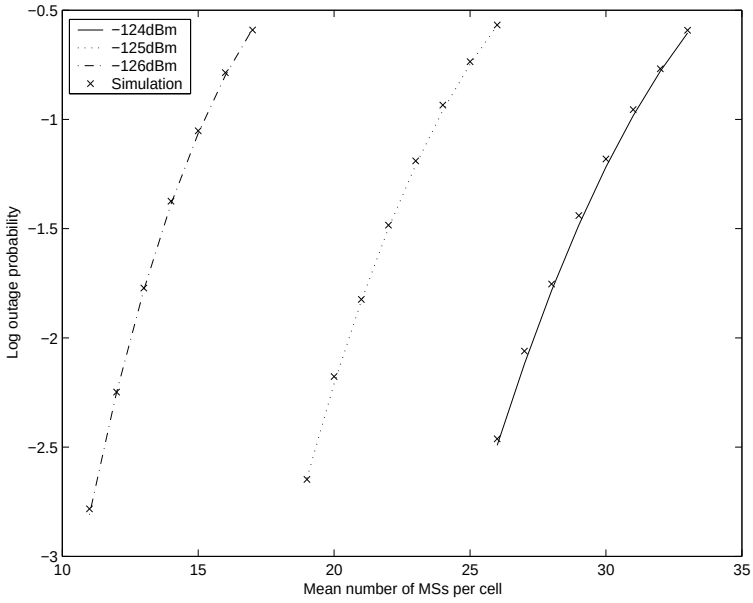


Figure 3.12: Variation of outage probability with maximum received power S^*

which are voice users, low-rate data users and high-rate data users, are assumed to be supported in the system, whose operation parameters are given in the following table.

Table 3.2: Operation parameters of the multi-class service model

Service	Bit Rate $[R]$	Target E_b/I_0 $[\varepsilon^*]$	Activity $[\upsilon]$	Ratio
voice	12.2kbps	5.5dB	1	75%
low-rate data	28.8kbps	4.0dB	1	20%
high-rate data	64kbps	3.5dB	1	5%

In Fig. 3.13 and Fig. 3.14, the corresponding comparisons of mean and standard deviation of other-cell interference with various load under multi-class traffic model are displayed. Again the mean values from both models match quite well, while the standard deviations show slightly greater discrepancy, especially in higher load region. The reason for this is due to the mutual independence assumption of $\mathcal{I}^{out}_{y\to x}$ made during the iterative calculation. It impacts on the variance computation and thus underestimates the standard deviation of $\mathcal{I}^{oc}$ with high loads. Applying the acquired moments of other-cell interference into outage probability analysis, the analytical

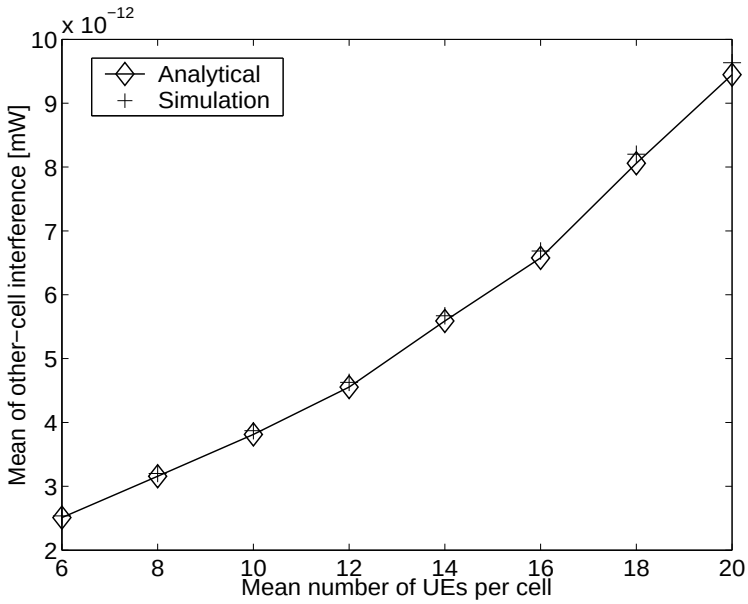


Figure 3.13: Comparison of mean other-cell interference with multiple services

and simulation results of P_{outage} are shown in log scale in Fig. 3.15 with maximum received power $S^* = -119$dBm assumed. Sufficient accuracy of such approximation on the system capacity has been validated through these numerical results, which can be justified to be an efficient approach for the the system capacity model in UMTS systems.

3.8 Discussion

In this chapter, we have presented a purely analytical model for the characterization of other-cell interference in UMTS networks with log-normal shadowing effects, which is crucial for efficient capacity evaluation during network planning. In theory, the distribution function of other-cell interference can be computed based on solving fixed-point equations iteratively where many convolutions may be involved. A log-normal approximation model is then suggested and verified so that the calculation can be simplified significantly. Finally, one important metric for network capacity planning, outage probability, is introduced and its derivation is demonstrated with fully analytically characterized other-cell interference.

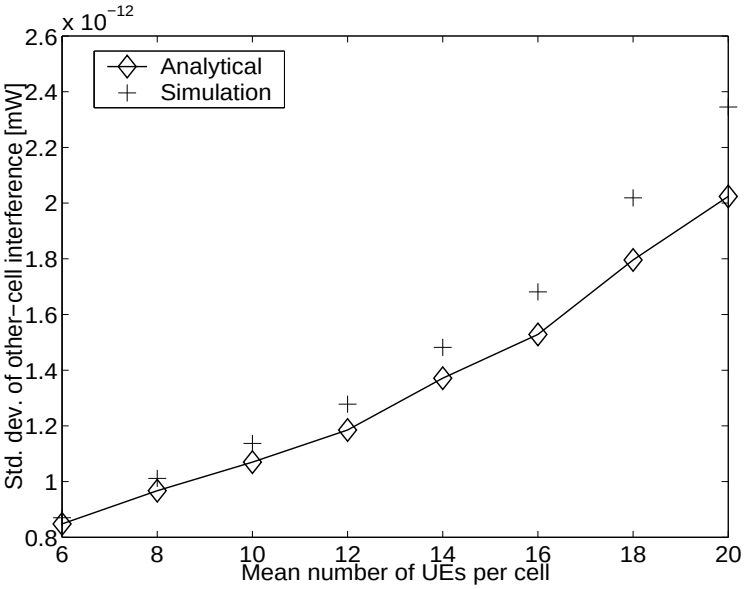


Figure 3.14: Comparison of standard deviation of other-cell interference with multiple services

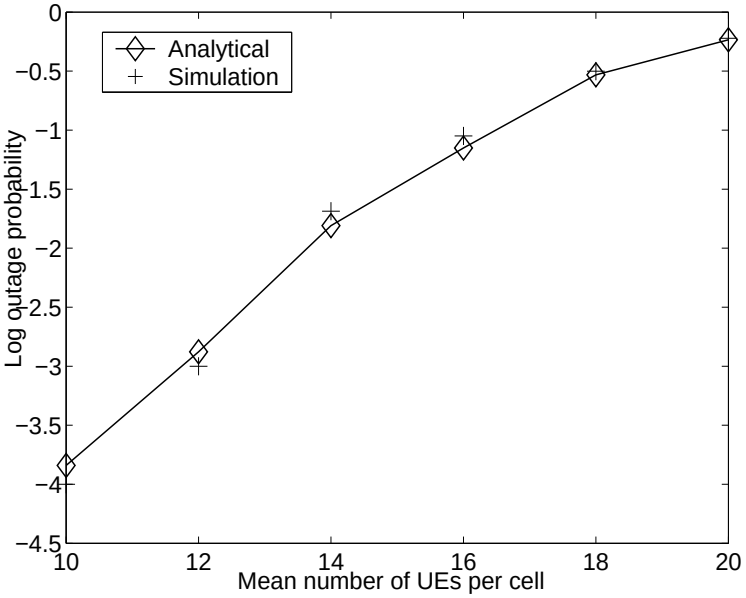


Figure 3.15: Analytical and simulation results for logarithm outage probability with multiple services

As mentioned before, the main motivation for the study of the analytical interference models in the UMTS networks is to increase the efficiency for the capacity evaluation, thus the computation time is another concern. In the models validation process above, the analytical approach generally returns the results instantly (no more than 5 seconds). With simulations, depending on the number of times repeated in the Monte-Carlo simulation, the calculation time varies in order of minutes. In my experiments, the simulation generates 100,000 snapshots of different network scenarios, and the final results are obtained in several minutes (generally no more than 10 minutes). It can be seen that the computation time is reduced significantly by the analytical approach.

Although the analysis in this chapter involves multi-class traffic model, the best-effort services are treated as the traditional voice services, which negotiate the proper fixed data rates at the time of being admitted into the networks. To shape the constructed model more closely to the practical systems, the feature on variable data rates for packet-switched best-effort users should be included, which are discussed in the following chapters.

Chapter 4

Modeling of UMTS Enhanced Uplink with Best-Effort Traffic

The 3G mobile communication systems allow network operators to provide a large variety of services, which can be categorized into four main classes, that is, conversational, streaming, interactive and background services. The conversational and streaming services require minimum bandwidth and maximum delay, thus they are always described as QoS traffic. The interactive and background services consume the remaining resources available in the systems, and they are generally denoted as best-effort traffic. In the recent evolution of UMTS, the enhanced uplink enables the efficient transport of packet-switched best-effort traffic. With bit rates comparable to DSL links, it is expected that interactive and background services like Internet browsing and file sharing will become increasingly popular in UMTS networks.

As the UMTS networks are rapidly deployed worldwide and the UMTS subscribers base continues to grow, the network operators want to support next generation packet data services that require very high data rates, both in uplink and downlink. To compete with other technologies such as 1xEV-DO, the UMTS standards have defined HSUPA for the uplink and HSDPA for the downlink. In the previous chapter, we have constructed analytical models for the UMTS systems supporting users with mere fixed-rate QoS traffic, while in this chapter, we extend the analysis on the basis of the previous chapter to present an analytical interference model in the multi-cell UMTS HSUPA with users composed of both QoS and best-effort traffic. Two types of traffic

models are employed respectively to characterize the incoming best-effort traffic, as these services can be regarded as either time-based which means the sojourn time of a connection depends on how long the user stays in the network (e.g. Internet browsing, chatting), or volume-based which means the sojourn time depends on the transmitted data volume (e.g. file transferring, e-mail sending). In this work, we propose analytic models for the interference and capacity in the UMTS enhanced uplink, which considers the impact of best-effort traffic on system capacity, other-cell interference and co-existing QoS users. Furthermore, to avoid the numerous convolutions which are involved in the direct distribution derivation, log-normal approximations are suggested with further enhancement of the accuracy of approximation discussed. The analytic approximation models are validated with both Monte Carlo and flow-level discrete event simulations.

4.1 Introduction

In recent years, the performance of packet access in UMTS keeps improving, firstly by the introduction of HSDPA [3] in Release 5 for increasing bandwidth demands in the downlink direction, and then recently with the proposal of Enhanced Uplink (HSUPA) [4] in Release 6 to meet the growing traffic requests in the uplink direction. With such enhancements, packet switched data are transmitted over the newly introduced Enhanced Dedicated Channel (E-DCH) so that they can be distinguished from the streaming traffic on the Dedicated Channel (DCH), and in turn achieving superior performance.

As illustrated in Fig. 4.1, three new major features are employed in the UMTS HSUPA in order to fulfill the higher bandwidth requirements, which are *fast hybrid Automatic Repeat reQuest (ARQ)*, *fast scheduling* implemented in NodeB and *short Transmission Time Interval (TTI)* of 2ms [52]. These features lead to shorter signaling delays and consequently enable fast reactions in the resource allocation processes. Among these features, the relocation of the scheduler from RNC to NodeB allows much more rapidity and flexibility for the implementation of RRM. In particular in the enhanced uplink, *fast rate control*, which enables the efficient transport of elastic traffic, is implemented. Elastic traffic means that an application does not have

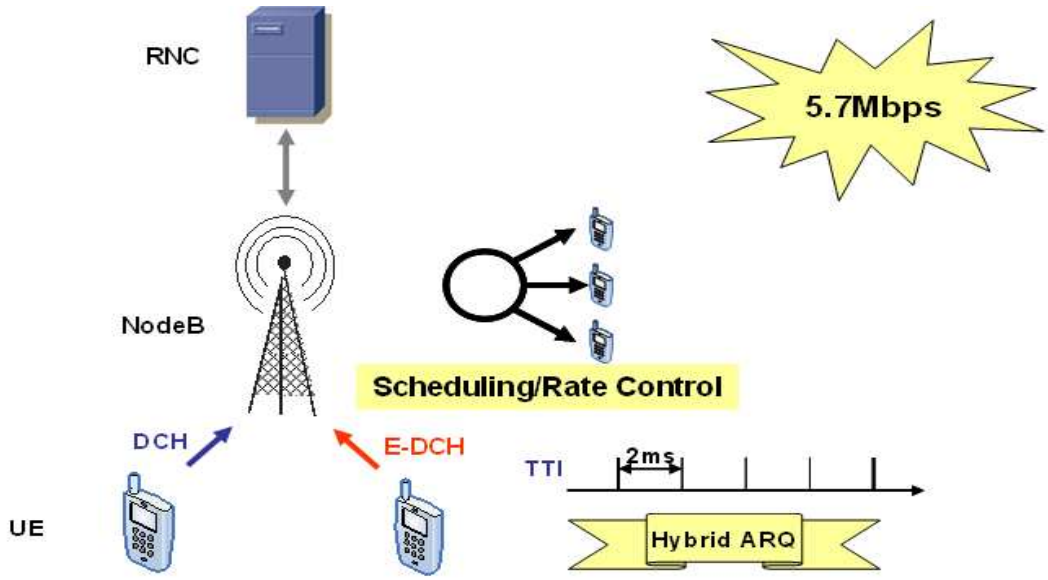


Figure 4.1: New features in the UMTS enhanced uplink

stringent throughput requirements (such as voice traffic) but will use the offered link capacity as best as possible. Elastic traffic is therefore often used as a synonym for best-effort traffic as generated by applications like File Transfer Protocol (FTP) or peer-to-peer file sharing applications. Such applications have in common that a certain data volume has to be transmitted, for example a large document, and the connection or session ends if the file transport is completed. Therefore, in contrast to QoS services like voice telephony, which are generally time-based, some of these applications follow a volume-based traffic model.

The analytical models for the interference in the conventional non-rate-controlled systems with only QoS traffic supported operating with either target SIR oriented RRM [42,46,59] or target received power oriented RRM [20,67] have been intensively studied. An analytical approximation for the other-cell interference under the fixed SIR power control mechanism is presented in [59], in which an iterative method is proposed to solve the fixed-point equations in order to determine the distribution. Furthermore, the other-cell interference is proved to follow a log-normal distribution approximately in this scenario, and thus the complexity of computation reduces dramatically since only the mean and variance are sufficient to characterize this random

variable. However, in doing so, it requires some empirical results, which can be regarded as semi-analytical approximation. Moreover, much effort has been spent on the interference model construction and performance analysis for the QoS/best-effort data integrated multi-service CDMA systems [8, 41] and recent 3G networks [29].

For the newly evolved HSUPA-enabled 3G systems with packet data traffic scheduled by the rate controller, analytical interference and load models of the UMTS enhanced uplink have been constructed in [47], where the blocking probability, cell load and bit rate are derived. However, the other-cell interference in this analysis is just modeled as an independent log-normally distributed random variable, thus it only accounts for the single-cell scenarios. In the multi-cell environments, the interference level at each NodeB depends on each other due to the feedback behavior explained in the previous chapter. Hence the correlations and interactions of the interferences among all the NodeBs are not included in this model.

Some other related work can be found e.g. in a few publications [7, 25, 29], where the uplink capacity of a general rate-controlled WCDMA system is investigated. The paper [7] considers time-based traffic only with a queueing analysis for the CDMA uplink with best-effort services, while in [25], which is an extension work of [7], the dynamic adjustment of the slow down rates for elastic traffic is presented and analyzed with a Markov model in single-cell scenario. The paper [29] also considers mixed traffic scenarios with volume-based traffic. However, in the case of multi-cell scenarios the other-cell interference is approximated with an f-factor approach, which assumes a linear relation between own-cell and other-cell interference.

In this work, we extend the analysis on the basis of the previous chapter to present an analytical model for the other-cell interference in multi-cell UMTS enhanced uplink with best-effort traffic supported. The rest of this chapter is organized as follows. In Section 4.2, the target own-cell load oriented radio resource allocation mechanism for the UMTS enhanced uplink is presented. Section 4.3 derives the user distribution based on the time-based and volume-based traffic models which are employed for the incoming best-effort users. In Section 4.4, we present the uplink interference model with deterministic user patterns, and introduce the principles of Monte Carlo as well as discrete event simulation techniques to investigate the system behavior under both

traffic models respectively. Section 4.5 suggests the stochastic model in UMTS enhanced uplink with random user patterns, followed by a log-normal approximation approach to reduce computation complexity. An accuracy enhancement technique to capture the effects of interference feedback and a capacity evaluation approach based on the interference model are explained in the latter part of this section. Simulation results from Monte Carlo and discrete event simulations are compared with the numerical results from analytical models in Section 4.6 to verify the proposed analytical approximation approach. In Section 4.7, the models are summarized, and the scope for future work is highlighted.

4.2 The Radio Resource Allocation for the UMTS Enhanced Uplink

The resource allocation scheme plays an important role in the system behaviors, and in turn the performance. Therefore, before going into the interference model construction, we specify in this section the employed rate allocation algorithms as well as blocking criteria, etc., since the later proposed analytical models depend much on such schemes.

One major advance of the enhanced uplink compared with the conventional uplink is the fast scheduling, which allows the resource allocation tasks to be performed at each individual NodeB, rather than at the RNC side in previous UMTS releases. This leads to much more rapid rate assignment, so that the target of enhanced uplink, such as reduced delays, increased data rates and capacity, can be well achieved. A new Medium Access Control (MAC) entity has been introduced in each NodeB for this purpose, thus the resource allocation in the enhanced uplink become decentralized, which differs most from that of traditional UMTS uplink. In conventional UMTS uplink, the RNC is in full possession of the information of received interference at each NodeB, and thus can make centralized resource allocation. While in the enhanced uplink, due to the decentralized resource allocation, one NodeB cannot adjust its received load contributed by the users from other cells directly since they are controlled by other NodeBs, such that the interference model differs from previous one accordingly.

Under target E_b/I_0 based power control mechanism, the essence of resource allocation is actually to determine the value of ω_t for each class t from a mathematical point of view. This is because in CDMA-based systems, power is the major capacity resource. In (3.17), it can be seen that the received power is expressed as a product of the ω component and the interference component. As the interference component only depends on the total received interference at each NodeB, they are equal to each other for all the users connected to the same NodeB, thus the resources assigned to each user only vary in the ω_t components. For the QoS users (also dubbed as DCH users in the following as they occupy the dedicated channels in HSUPA-enabled UMTS), ω_t are simply class-dependent pre-defined constants during system operation as the variables ε_t^* and R_t in (3.18) are pre-determined. However, for best-effort users (also dubbed as E-DCH users in the following as they occupy the enhanced dedicated channels in HSUPA-enabled UMTS), since R_t vary with the system status, ω_t are no longer pre-determined constants but variables changing with time where the values taken depend on the employed resource allocation schemes. In general, the higher the data rate of one user, the higher the interference and load this user would contribute, and in turn the more resources consumption. The key point in the resource allocation over UMTS enhanced uplink thus translates to the instant computation of R_t and ω_t. Note that certain minimum rates to guarantee continuous elastic data transmission for each class (denoted as $R_{t,min}^E$) must be satisfied by all the admitted E-DCH users in the system.

In [30], a new term 'cell load', which can be uniquely determined by the received interference, is introduced and the load η_x in cell x is defined as

$$\eta_x = \frac{I_x}{I_x + WN_0} \tag{4.1}$$

where I_x is the received interference at NodeB x from all the DCH and E-DCH users over the system. The received interference consists of power from DCH users and E-DCH users in own cell and other cells respectively as

$$I_x = I_{x,own}^D + I_{x,own}^E + I_{x,oc}^D + I_{x,oc}^E. \tag{4.2}$$

Hence the cell load can be consequently decomposed into four corresponding parts

according to different sources as

$$\begin{aligned}\eta_x &= \frac{I^D_{x,own}}{I_x + WN_0} + \frac{I^E_{x,own}}{I_x + WN_0} + \frac{I^D_{x,oc}}{I_x + WN_0} + \frac{I^E_{x,oc}}{I_x + WN_0} \\ &= \eta^D_{x,own} + \eta^E_{x,own} + \eta^D_{x,oc} + \eta^E_{x,oc}\end{aligned} \tag{4.3}$$

subject to

$$\eta_x < 1. \tag{4.4}$$

The aim of resource allocation now becomes a trade-off. On one hand, the received cell load should be kept below the maximum allowable load limit (denoted as η_x^{max}), so that the outage cases can be avoided as much as possible. The *Outage* event is defined as the cell load η_x exceeding η_x^{max}, which causes the QoS of all the connected DCH and E-DCH users to drop below the acceptable thresholds. On the other hand, the shared resources should be utilized as efficiently as possible in order to reach the maximum capacity. This trade-off translates into the goal that maintaining the received load η_x as close to η_x^{max} as possible but not exceeding it.

In order to satisfy such a goal, since the own-cell DCH load $\eta^D_{x,own}$ and the other-cell load $\eta_{x,oc}$ cannot be easily adapted, the adjustable own-cell E-DCH load component $\eta^E_{x,own}$ is the best way to deploy fast scheduling, such that the received load can be 'waterfilled' up to a desired target by the best-effort users. The decentralized resource allocation is then performed with mere local load information (i.e. $\eta^D_{x,own}$ and $\eta^E_{x,own}$), which aims to keep the own-cell received load as own-cell target load η^*_{own}

$$\eta_{x,own} = \eta^D_{x,own} + \eta^E_{x,own} = \eta^*_{own} \tag{4.5}$$

Since the own-cell load is kept as a pre-defined target load η^*_{own} under this mechanism, once the other-cell load in such scenarios can be analytically characterized as a function of η^*_{own}, the outage probability and in turn network capacity can be analyzed accordingly. Thus we can simply use the parameter η^*_{own} to evaluate the system performance during network planning process with such a model. The technique to derive the CDF for $\eta_{x,oc}$ under such 'waterfilling' resource allocation is the focus in this chapter.

Note that even with such own-cell target load oriented resource allocation which

considers the local load information, interference levels at different NodeBs still depend on each other during operation due to the feedback behavior, which is the key point in this analysis. Another point is there is one exception that equation (4.5) does not hold, which is no E-DCH user in the current cell. In such cases, $\eta_{x,own}$ is simply equal to the own-cell DCH load $\eta^D_{x,own}$ which cannot be adjusted.

The own-cell target load η^*_{own} also serves as a reference for admission control in the resource allocation. If the received load from own-cell DCH users together with the minimum load contributed by own-cell E-DCH users (i.e. each E-DCH user is only allocated R^E_{min}) exceed the threshold η^*_{own}, a *Blocking* event occurs, during which no incoming users can be admitted. Thus η^*_{own} is also an important parameter for the blocking probability.

4.3 Traffic Model in the UMTS Enhanced Uplink

In this section, we introduce two types of source traffic models: *time-based* and *volume-based*, respectively. We will investigate the impacts on system behavior under both models later in this chapter.

4.3.1 Time-Based Traffic Model

The time-based source traffic model is similar to the one employed in the previous chapter, which complies with a spatial homogeneous Poisson process [16]. It assumes Poisson arrival of both DCH and E-DCH users and negative exponentially distributed sojourn time, thus the user distributions can be easily calculated from a $T+1$ dimensional Markov chain, where T is the total number of DCH user classes, as product form solutions. A state is defined by the number of DCH and E-DCH users as $(\bar{n}^D_x, \bar{n}^E_x)$ where $\bar{n}^D_x$ stands for the vector consisting of the user number in each class $\left[n^D_{1,x}, \ldots, n^D_{T,x}\right]$ in cell x. The state space Ω is restricted by the admissible region $\mathbf{S}$, which is defined by the own-cell target load oriented admission control policy, where the sum of own-cell DCH load and minimum E-DCH load cannot exceed η^*_{own}

$$\bar{n}^D_x, \bar{n}^E_x \in \mathbf{S} \quad \text{if} \quad \bar{n}^E_x \omega^E_{min} + \sum_{t=1}^{T} \bar{n}^D_{t,x} \omega^D_{t,x} < \eta^*_{own}. \tag{4.6}$$

The user distribution $P(\bar{n}_x^D, \bar{n}_x^E)$ is given as

$$P\left(\bar{n}_x^D, \bar{n}_x^E\right) = \begin{cases} P_0 \cdot \frac{\left(N_x^E\right)^{\bar{n}_x^E}}{\bar{n}_x^E!} \prod_{t=1}^{T} \frac{\left(N_{t,x}^D\right)^{n_{t,x}^D}}{n_{t,x}^D!}, & \bar{n}_x^D, \bar{n}_x^E \in \mathbf{S} \\ 0, & \text{otherwise} \end{cases} \tag{4.7}$$

and

$$P_0 = \frac{1}{\sum_{\bar{n}_x^D, \bar{n}_x^E \in \mathbf{S}} \frac{(N_x^E)^{\bar{n}_x^E}}{\bar{n}_x^E!} \prod_{t=1}^{T} \frac{\left(N_{t,x}^D\right)^{\bar{n}_{t,x}^D}}{\bar{n}_{t,x}^D!}} \tag{4.8}$$

where $N_{t,x}^D$ denotes the mean number of DCH users of class t connecting to NodeB x and N_x^E refers to that of E-DCH users, both of which can be determined straightforwardly from the mean arrival and departure rates.

4.3.2 Volume-Based Traffic Model

The traffic generated by best-effort users are commonly elastic, which implies these users would not leave the system until all the data have been transmitted. Based on this features, there can be another way to model the incoming traffic, which is to assume the volume size to be transmitted by each E-DCH user, rather than the sojourn time, is negative exponentially distributed. We denote such model as volume-based traffic model. Since the data rates for E-DCH users vary with the remaining cell load capacity, the sojourn time of E-DCH users also depend on the system state (i.e. the number of concurrently active DCH and E-DCH flows). From such model we can derive the state-dependent statistical distribution of sojourn time and apply it into interference modeling.

With such assumptions, the derivation of user state distribution is not as straightforward as the above method, since the application of product form solution is no more valid. We then need to adapt the $T+1$ dimensional Markov chain to characterize this scenario. If considered in isolation with only E-DCH users in the volume-based traffic model, such a system can be interpreted as a classical generalized processor sharing queue, for which closed-form solutions for the steady-state probabilities exist. In an integrated system supporting both DCH and E-DCH users as here, since the joint Markov process is not time-reversible which can be instantly verified with

the Kolomogorov's reversibility criterion, the closed-form solution does not exist. We need to construct a matrix equation in order to solve the steady-state probabilities.

Similar to the above approach, we firstly identify the state space Ω using (4.6), then the steady-state probabilities can be calculated with the generator matrix Q, where the entries q are defined with help of an index function $\phi(\bar{n}_x^D, \bar{n}_x^E) : \Omega \rightarrow \mathbb{N}$ as follows:

$$\begin{aligned}
q[\phi(\bar{n}_x^D, \bar{n}_x^E), \phi(\bar{n}_x^D + 1, \bar{n}_x^E)] &= \lambda_t^D \\
q[\phi(\bar{n}_x^D, \bar{n}_x^E), \phi(\bar{n}_x^D, \bar{n}_x^E + 1)] &= \lambda^E \\
q[\phi(\bar{n}_x^D, \bar{n}_x^E), \phi(\bar{n}_x^D - 1, \bar{n}_x^E)] &= n_{t,x}^D \cdot \mu_t^D \\
q[\phi(\bar{n}_x^D, \bar{n}_x^E), \phi(\bar{n}_x^D, \bar{n}_x^E - 1)] &= \bar{n}_x^E \cdot \mu_{\bar{n}_x^D, \bar{n}_x^E}^E
\end{aligned} \tag{4.9}$$

where λ_t^D, μ_t^D represent the mean arrival and departure rate for class t DCH users and λ^E is the mean arrival rate for the E-DCH users. All these three rates are simply known constant values, while the state-dependent departure rate for the enhanced uplink flows $\mu_{\bar{n}_x^D, \bar{n}_x^E}^E$ follows as the reciprocal of the conditional mean sojourn time which is the duration to finish data transfer with given data rate $R_{\bar{n}_x^D, \bar{n}_x^E}^E$ and mean data volume size $E[V^E]$:

$$\mu_{\bar{n}_x^D, \bar{n}_x^E}^E = \frac{R_{\bar{n}_x^D, \bar{n}_x^E}^E}{E[V^E]} \tag{4.10}$$

where $R_{\bar{n}_x^D, \bar{n}_x^E}^E$ is calculated from (3.18) as

$$R_{\bar{n}_x^D, \bar{n}_x^E}^E = \frac{W}{\varepsilon_E^*} \cdot \frac{\omega_{\bar{n}_x^D, \bar{n}_x^E}^E}{1 - \omega_{\bar{n}_x^D, \bar{n}_x^E}^E}. \tag{4.11}$$

All the other entries in the generator matrix Q that are not listed in (4.9) are set to 0 with exception of the diagonal entries, which are calculated such that the row sum of the matrix is 0. The steady-state distribution can be calculated by solving the classical equation $\bar{\pi} \cdot Q = 0$ subject to $\sum \pi = 1$, where $\pi[\phi(\bar{n}_x^D, \bar{n}_x^E)] = P(\bar{n}_x^D, \bar{n}_x^E)$ which refers to the stationary probability of state $(\bar{n}_x^D, \bar{n}_x^E)$. A single local balance

equation for the state $(\bar{n}_x^D, \bar{n}_x^E)$ is expressed as an example as follows

$$\begin{aligned}
&\left(\sum_{t=1}^{T} \lambda_t^D + \lambda^E + \sum_{t=1}^{T} n_{t,x}^D \mu_t^D + \bar{n}_x^E \mu_{\bar{n}_x^D, \bar{n}_x^E}^E\right) \cdot P(\bar{n}_x^D, \bar{n}_x^E) \\
=&\sum_{t=1}^{T} \lambda_t^D P(n_{1,x}^D, \ldots, n_{t,x}^D - 1, \ldots, n_{T,x}^D, \bar{n}_x^E) \\
&+ \sum_{t=1}^{T} (n_{t,x}^D + 1)\mu_t^D P(n_{1,x}^D, \ldots, n_{t,x}^D + 1, \ldots, n_{T,x}^D, \bar{n}_x^E) \\
&+ \lambda^E P(\bar{n}_x^D, \bar{n}_x^E - 1) + (\bar{n}_x^E + 1)\mu_{\bar{n}_x^D, \bar{n}_x^E + 1}^E P(\bar{n}_x^D, \bar{n}_x^E + 1).
\end{aligned} \tag{4.12}$$

Note that if any state in the right hand of the above equation falls outside the state space Ω, the corresponding summand becomes zero. With the constructed set of linear equations with respect to $P(\bar{n}_x^D, \bar{n}_x^E)$, the analytical stationary user distribution can be solved and employed for the later computation of the other-cell interference.

The state diagrams of the modified Markov chain are shown in the following figures. Fig. 4.2 demonstrates an overview of a Markov chain representation of the volume-based source traffic model, while Fig. 4.3 presents the transition rates related to a general user state $(\bar{n}_x^D, \bar{n}_x^E)$. For the clarity of illustration, only one class of DCH users is assumed in the system shown in this figure, thus the Markov chain can be plotted in a two-dimensional state space.

To verify the above analytical approach, the derived result of user distribution are compared with that from simulations in the following figures. Again one class of DCH user is assumed for display reason, such that the distribution can be plotted as a mesh surface. Quite similar shapes of the mesh surfaces observed in Fig. 4.4 and 4.5 indicate good match between the analytical and simulation results.

4.4 Simulations in the UMTS Enhanced Uplink

Again we firstly use simulations to get a brief understanding of the system behavior, which also serve as the benchmark for the later proposed analytical model. Since two types of traffic models are considered, we correspondingly need to employ both Monte Carlo and discrete event simulations, which will be introduced respectively in

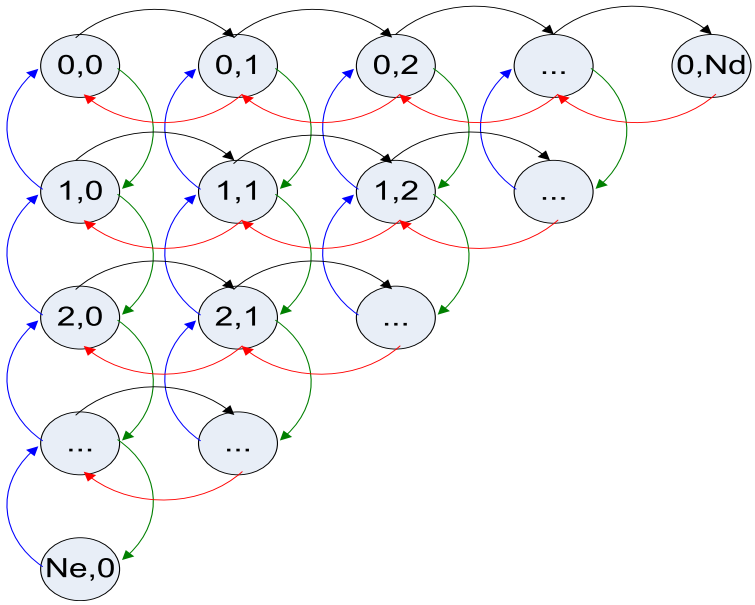


Figure 4.2: Markov chain for volume-based source traffic model

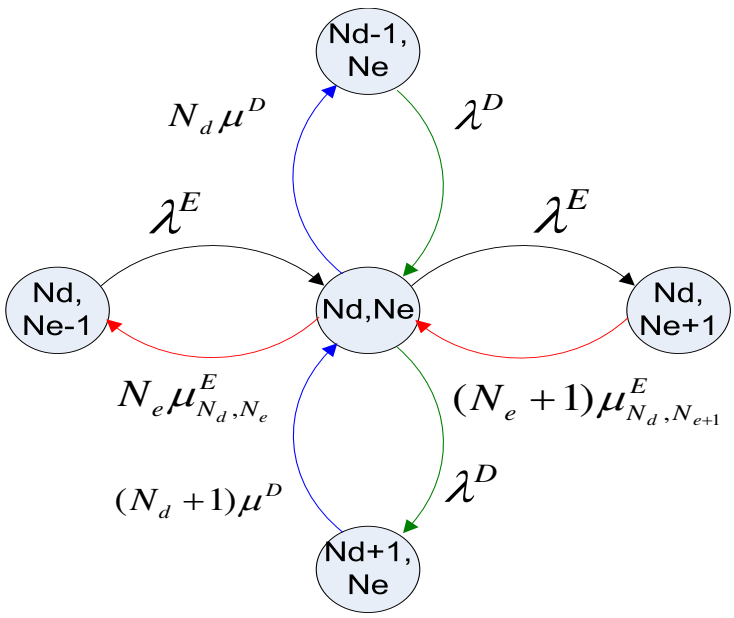


Figure 4.3: The transition rates of a general user state

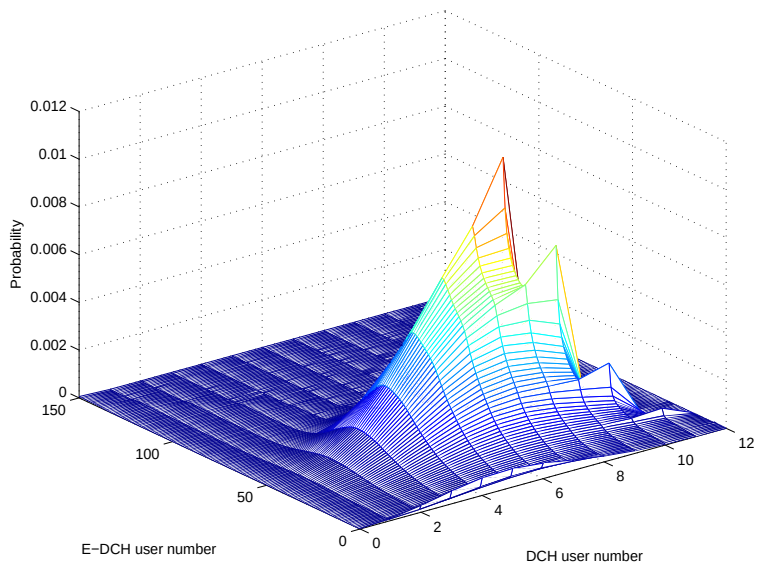


Figure 4.4: Analytical results of the volume-based traffic model

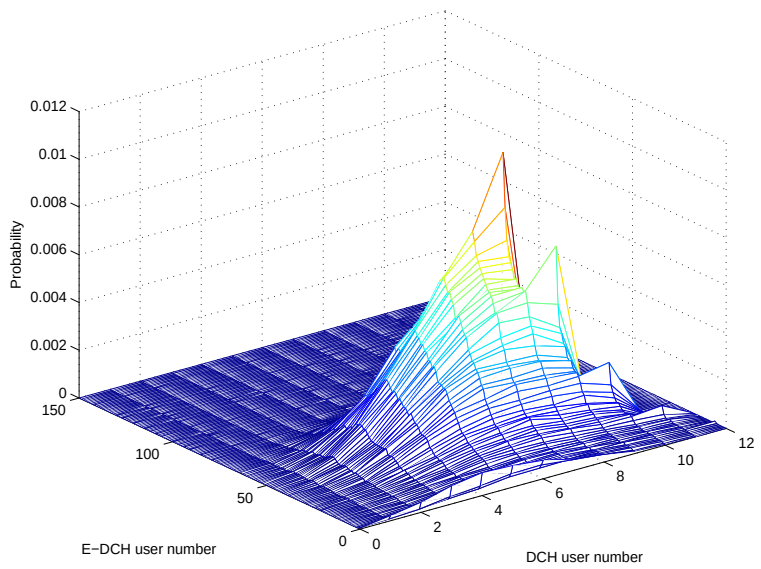


Figure 4.5: Simulation results of the volume-based traffic model

this section.

The UMTS enhanced uplink follows the similar power control equation:

$$\varepsilon_k^* = \frac{W}{R_k} \cdot \frac{S_{k,x}^R}{I_x^{own} + I_x^{oc} + WN_0 - S_{k,x}^R} \tag{4.13}$$

with same notations as in (3.16), from which the received power $S_{k,x}^R$ is solved as

$$S_{k,x}^R = \omega_k \left(I_x^{own} + I_x^{oc} + WN_0\right) \tag{4.14}$$

where ω_k has the same definition as:

$$\omega_k = \frac{\varepsilon_k^* R_k}{W + \varepsilon_k^* R_k}. \tag{4.15}$$

Then together with the cell load definition in (4.1) and decomposition in (4.3), the own-cell load from DCH and E-DCH users can be rewritten as

$$\eta_{x,own}^D = \frac{I_{x,own}^D}{I_x^{own} + I_x^{oc} + WN_0} = \frac{\sum\limits_{k \in \mathcal{D}_x} S_{k,x}^R}{I_x^{own} + I_x^{oc} + WN_0} = \sum_{k \in \mathcal{D}_x} \omega_k^D \tag{4.16}$$

and

$$\eta_{x,own}^E = \frac{I_{x,own}^E}{I_x^{own} + I_x^{oc} + WN_0} = \frac{\sum\limits_{j \in \mathcal{E}_x} S_{j,x}^R}{I_x^{own} + I_x^{oc} + WN_0} = \sum_{j \in \mathcal{E}_x} \omega_j^E \tag{4.17}$$

where $\mathcal{D}_x$ and $\mathcal{E}_x$ refer to all the DCH and E-DCH users controlled by the NodeB x. Following the same techniques in the previous chapter, we can cancel the I_x^{own} component and rewrite (4.14) as

$$S_{k,x}^R = \frac{\omega_k}{1 - \left(\eta_{x,own}^D + \eta_{x,own}^E\right)} \left(I_x^{oc} + WN_0\right). \tag{4.18}$$

From the above equations, it can be seen that ω is in essence the effective load contributed by each individual user to its serving NodeB, thus referred to as *Service Load Factor (SLF)* similar to [62]. According to the expression in (4.15), if the system bandwidth W is always assumed as a constant, the variable ω_k depends on bit rate and

target E_b/I_0 values only, which implies that for DCH users, once they are admitted into the network, ω_k^D can be regarded as a constant, while for E-DCH users, ω_j^E is a function of the allocated bit rate R_j, hence it can always be adapted during the service period to meet the resource allocation algorithm in (4.5). The rate allocation for each E-DCH user within one cell depends on the employed scheduling discipline. If assuming parallel equal-rate scheduling scheme, then every E-DCH user within one cell shares the same SLF ω_j^E, and the instant ω_j^E is determined by evenly dividing the available load resources for E-DCH users as

$$\omega_j^E = \frac{\eta_{own}^* - \eta_{x,own}^D}{\bar{n}_x^E}, \tag{4.19}$$

while the bit rate R_j can be accordingly assigned for each E-DCH user.

Then we proceed to model the other-cell interference I_x^{oc}, which is the sum of the inter-cell received power $S_{y\to x}^k$ at NodeB x contributed by the UEs served by all the NodeBs other than x, and in turn the load components η_x^{oc} received from other-cell users. This time the log-distance path loss propagation model is assumed for simplicity reasons, which of course can be extended to log-normal shadowing propagation model by similar approaches introduced before. Again we have a similar form of linear equations as in the previous chapter

$$I_x^{oc} = \sum_{y\neq x} I_{y\to x}^{out} \tag{4.20}$$

$$I_{y\to x}^{out} = \left(I_y^{oc} + WN_0\right) F_{y\to x}, \tag{4.21}$$

but now they differ in the parameter $F_{y\to x}$, where an additional component for E-DCH traffic is included, which becomes

$$F_{y\to x} = \frac{1}{1-\left(\eta_{y,own}^D + \eta_{y,own}^E\right)} \left[\sum_{k\in\mathcal{D}_y} \omega_k^D \Delta_{y\to x}^k + \sum_{j\in\mathcal{E}_y} \omega_j^E \Delta_{y\to x}^j\right] \tag{4.22}$$

Since the variable $F_{y\to x}$ in (4.22) contains all the user population and location information, distinct traffic models would only affect the computation of $F_{y\to x}$, while the latter matrix equation solving part keeps the same. With the time-based traffic

model, we can generate each user pattern based on the spatial homogeneous Poisson process with known mean user numbers, thus the Monte Carlo simulation can be directly applied. From each generated user pattern, the moments of I_x^{oc} are computed and average values of both moments can be accordingly obtained from a large number of samples.

However, with the volume-based model, since the E-DCH user departure rate depends on the system state, the mean user number cannot be simply calculated as the quotient of arrival rate by departure rate. Thus the Monte Carlo simulation technique does not suit such scenarios, instead, we need to apply a discrete event simulation in order to capture the instant E-DCH data rate information. The key principles regarding the employed simulation are summarized below.

1. This discrete-event simulation consists of a chronological sequence of user arrival and departure events. It starts at time 0, and ends when the total number of other-cell interference samples reaches a certain value.

2. Each time when there is an incoming user into the target cell, no matter DCH or E-DCH, the other-cell interference at this instant is calculated and sampled. This is according to PASTA (Poisson Arrivals See Time Average) as both arrival processes are Poisson processes.

3. For incoming users into the other cells, the other-cell interference is not sampled. Instead, only the location information are recorded for later interference calculations.

4. In each user arrival and departure event (both DCH and E-DCH), the data rates of all the E-DCH users in the same cell as the incoming user need to be updated. Accordingly the departure time of all these E-DCH users would be re-calculated and re-scheduled.

5. After the simulation finishes, we can estimate the possible distribution and derive the moments from sampled values.

The simulation results are good references to investigate the UMTS enhanced uplink, however, it is not an efficient approach to predict the system behavior in the

network planning process because the required number of samples would be quite large and thus it might become very time-consuming, especially for the discrete event simulations. We would propose the analytical models and more efficient approximation techniques in the next section for this reason, which is the key contribution in this chapter.

4.5 Stochastic Interference Model in the UMTS Enhanced Uplink

The interference model with deterministic user pattern has been investigated by simulations in the previous section, then in order to derive the analytical model of the other-cell interference, the random factors within user pattern such as traffic intensity and spatial distribution should be considered in this part.

4.5.1 Analytical Model Formulation and Direct Approach

The stochastic fixed-point equations representation of (4.20) and (4.21) are formulated as

$$\mathcal{I}_x^{oc} = \sum_{y \neq x} \mathcal{I}_{y \to x}^{out} \quad \text{and} \quad \mathcal{I}_{y \to x}^{out} = \left(\mathcal{I}_y^{oc} + W N_0\right) \mathcal{F}_{y \to x}. \tag{4.23}$$

where $\mathcal{I}_x^{oc}$, $\mathcal{I}_{y \to x}^{out}$ and $\mathcal{F}_{y \to x}$ are the corresponding random variables. We can see that this set of fixed-point equations have exactly the same structure as those in our previous work with the only exception existing in the parameter $\mathcal{F}_{y \to x}$ due to the inclusion of best-effort traffic. It now becomes

$$\mathcal{F}_{y \to x} = \frac{\sum_{t=1}^{T} \omega_{t,y}^{D} \sum_{k=1}^{n_{t,y}^{D}} \Delta_{y \to x} + \omega_y^{E} \sum_{j=1}^{\bar{n}_y^{E}} \Delta_{y \to x}}{1 - \left(\eta_{y,own}^{D} + \eta_{y,own}^{E}\right)}. \tag{4.24}$$

Therefore, the task now reduces to the characterization of the distribution of $\mathcal{F}_{y \to x}$ in our current model, followed by a similar approach to determine the distribution of other-cell interference.

If the number of DCH and E-DCH users in the above expression are fixed as $\hat{n}_y^D$

and $\hat{n}_y^E$ where $\hat{n}_y^D$ stands for the vector $\left(\hat{n}_{1,y}^D, \ldots, \hat{n}_{T,y}^D\right)$, the values of $\eta_{y,own}^D$, $\eta_{y,own}^E$, $\omega_{t,y}^D$ and ω_y^E can be easily calculated from (4.15)-(4.19). Then together with the CDF of $\Delta_{y\to x}$ approximated in closed form in (3.38) - (3.40), the conditional CDF $P\left(\mathcal{F}_{y\to x} \leq z | \hat{n}_y^D, \hat{n}_y^E\right)$ can be derived in theory. Finally for the complete CDF of $\mathcal{F}_{y\to x}$, we apply the total probability theorem to uncondition $P\left(\mathcal{F}_{y\to x} \leq z | \hat{n}_y^D, \hat{n}_y^E\right)$ as

$$P\left(\mathcal{F}_{y\to x} \leq z\right) = \sum_{\hat{n}_y^D, \hat{n}_y^E \in \mathbf{S}} P(\hat{n}_y^D, \hat{n}_y^E) \cdot P\left(\mathcal{F}_{y\to x} \leq z | \hat{n}_y^D, \hat{n}_y^E\right) \tag{4.25}$$

where the user distribution $P(\hat{n}_y^D, \hat{n}_y^E)$ is given according to either time-based or volume-based traffic models.

Theoretically, with all the acquired CDFs of the random variables above, the CDF of $\mathcal{F}_{y\to x}$ can be computed in a similar iterative way as in Section 3.4.1, however, it is again quite a hard task due to the involvement of numerous convolutions which becomes in fact numerical intractable. Thus we will investigate some approximation techniques for $\mathcal{F}_{y\to x}$ and in turn $\mathcal{I}_x^{oc}$ in the next step to reduce the computational complexity.

4.5.2 Log-normal Approximation Approach

In light of the excellent log-normal approximation of the other-cell interference in the UMTS network with only QoS users demonstrated in the previous chapter, we perform similar verification experiments but with additional best-effort users under both time-based and volume-based traffic models. We present the normal probability plot which illustrates the likelihood between the normal distribution and the logarithm of $\mathcal{I}_x^{oc}$ under volume-based traffic model in Fig. 4.6 as an example. From such figure, it is great to see that again this random variables is still shown to be log-normally distributed as expected. This again reduces the problem to determining only the first and second moments of both random variables. Other experiments demonstrate this property exists for both $\mathcal{F}_{y\to x}$ and $\mathcal{I}_x^{oc}$ in both scenarios.

We first concentrate on the random variable $\mathcal{F}_{y\to x}$ which is the major different factor compared with the analytical interference model proposed before due to inclusion of best effort traffic. Applying the total probability theorem, the first moment

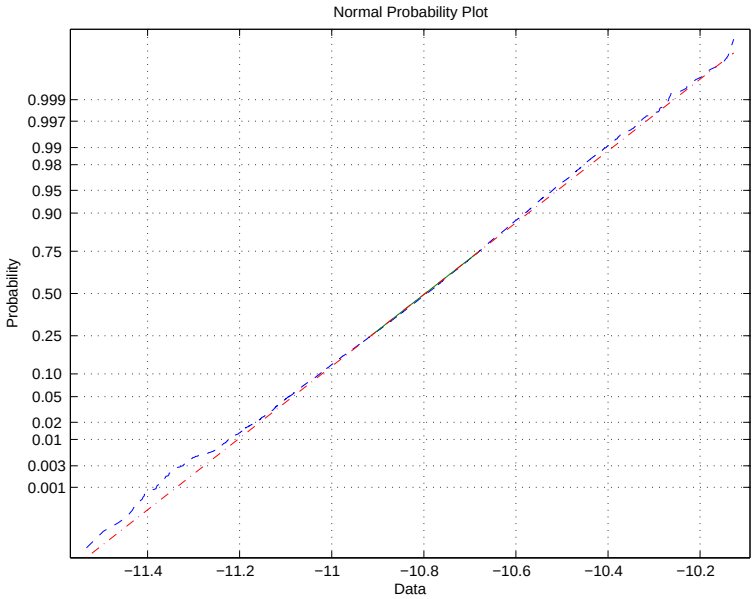


Figure 4.6: Normal probability plot of the logarithm of other-cell interference under volume-based traffic model

of $\mathcal{F}_{y\to x}$ is given as

$$E\left[\mathcal{F}_{y\to x}\right] = \sum_{\hat{n}_y^D,\hat{n}_y^E\in\mathbf{S}} P(\hat{n}_y^D,\hat{n}_y^E)\cdot E\left[\mathcal{F}_{y\to x}\left(\hat{n}_y^D,\hat{n}_y^E\right)\right] \tag{4.26}$$

where $\mathcal{F}_{y\to x}\left(\hat{n}_y^D,\hat{n}_y^E\right)$ denotes the random variable $\mathcal{F}_{y\to x}$ conditioned on known user combination $(\hat{n}_y^D,\hat{n}_y^E)$. Depending on whether there is E-DCH user in cell y or not, the received load at NodeB y is either equal to the DCH load $\eta_{y,own}^D$ or waterfilled up to the target load η_{own}^*. Thus the conditional mean is calculated as

$$E\left[\mathcal{F}_{y\to x}\left(\hat{n}_y^D,\hat{n}_y^E\right)\right] = \begin{cases} E\left[\Delta_{y\to x}\right]\cdot\frac{\eta_{y,own}^D}{1-\eta_{y,own}^D}, & \hat{n}_y^E = 0 \\ E\left[\Delta_{y\to x}\right]\cdot\frac{\eta_{own}^*}{1-\eta_{own}^*}, & \hat{n}_y^E \neq 0. \end{cases} \tag{4.27}$$

And the second moment is

$$\begin{aligned} E\left[\mathcal{F}_{y\to x}^2\right] &= \sum_{\hat{n}_y^D,\hat{n}_y^E\in\mathbf{S}} P(\hat{n}_y^D,\hat{n}_y^E)\cdot E\left[\mathcal{F}_{y\to x}\left(\hat{n}_y^D,\hat{n}_y^E\right)^2\right] \\ &= \sum_{\hat{n}_y^D,\hat{n}_y^E\in\mathbf{S}} P(\hat{n}_y^D,\hat{n}_y^E)\cdot\left(VAR\left[\mathcal{F}_{y\to x}\left(\hat{n}_y^D,\hat{n}_y^E\right)\right]+E\left[\mathcal{F}_{y\to x}\left(\hat{n}_y^D,\hat{n}_y^E\right)\right]^2\right) \end{aligned} \tag{4.28}$$

in which the conditional variance is calculated as

$$VAR\left[\mathcal{F}_{y\to x}\left(\hat{n}_y^D,\hat{n}_y^E\right)\right] = \begin{cases} VAR\left[\Delta_{y\to x}\right]\cdot\frac{\sum_{t=1}^{T}\hat{n}_{t,y}^D\left(\omega_{t,y}^D\right)^2}{\left(1-\eta_{y,own}^D\right)^2}, & \hat{n}_y^E=0 \\ VAR\left[\Delta_{y\to x}\right]\cdot\frac{\sum_{t=1}^{T}\hat{n}_{t,y}^D\left(\omega_{t,y}^D\right)^2+\frac{1}{\hat{n}_y^E}\left(\eta_{own}^*-\eta_{y,own}^D\right)^2}{\left(1-\eta_{own}^*\right)^2}, & \hat{n}_y^E\neq 0. \end{cases} \tag{4.29}$$

The components $E[\Delta_{y\to x}]$ and $VAR[\Delta_{y\to x}]$ in the above equations take the same values as those computed in the previous chapter with derived distribution function of $\Delta_{y\to x}$. And finally the variance of $\mathcal{F}_{y\to x}$ is given as

$$VAR\left[\mathcal{F}_{y\to x}\right] = E\left[\mathcal{F}_{y\to x}^2\right] - E\left[\mathcal{F}_{y\to x}\right]^2. \tag{4.30}$$

Once $\mathcal{F}_{y\to x}$ is characterized, we can use the similar iterative method elaborated before to derive the first and second moments of the other-cell interference. Or instead, we can also apply a direct matrix transform approach to obtain the moments of $\mathcal{I}^{oc}$, which will be described as follows.

With the assumptions of mutual independence between $\mathcal{F}_{y\to x}$ and $\mathcal{I}_y^{oc}$, the first moments of $\mathcal{I}_x^{oc}$ and $\mathcal{I}_{y\to x}^{out}$ can be derived from (4.23) as

$$E\left[\mathcal{I}_x^{oc}\right] = \sum_{y\neq x} E\left[\mathcal{I}_{y\to x}^{out}\right], \tag{4.31}$$

$$E\left[\mathcal{I}_{y\to x}^{out}\right] = E\left[\mathcal{F}_{y\to x}\right]\left(WN_0 + E\left[\mathcal{I}_y^{oc}\right]\right). \tag{4.32}$$

With further independence among $\mathcal{I}_{y\to x}^{out}$ assumed, the variance and second moment

of $\mathcal{I}_x^{oc}$ are represented as

$$VAR\left[\mathcal{I}_x^{oc}\right] = \sum_{y \neq x} VAR\left[\mathcal{I}_{y\to x}^{out}\right], \tag{4.33}$$

$$\begin{aligned} VAR\left[\mathcal{I}_{y\to x}^{out}\right] = & \left(\left[WN_0\right]^2 + 2WN_0 \cdot E\left[\mathcal{I}_y^{oc}\right]\right) \cdot VAR\left[\mathcal{F}_{y\to x}\right] \\ & - E\left[\mathcal{I}_y^{oc}\right]^2 E\left[\mathcal{F}_{y\to x}\right]^2 + E\left[\left(\mathcal{I}_y^{oc}\right)^2\right] E\left[\mathcal{F}_{y\to x}^2\right], \end{aligned} \tag{4.34}$$

and

$$E\left[\left(\mathcal{I}_y^{oc}\right)^2\right] = E\left[\mathcal{I}_y^{oc}\right]^2 + VAR\left[\mathcal{I}_y^{oc}\right] = \mathcal{H}_y + \sum_{z \neq y} E\left[(\mathcal{I}_z^{oc})^2\right] E\left[\mathcal{F}_{z\to y}^2\right] \tag{4.35}$$

where the auxiliary variable

$$\mathcal{H}_y = E\left[\mathcal{I}_y^{oc}\right]^2 + \sum_{z \neq y} \left(\left[WN_0\right]^2 + 2WN_0 \cdot E\left[\mathcal{I}_z^{oc}\right]\right) VAR\left[\mathcal{F}_{z\to y}\right] - \sum_{z \neq y} E\left[\mathcal{I}_z^{oc}\right]^2 E\left[\mathcal{F}_{z\to y}\right]^2. \tag{4.36}$$

To compute the moments of other-cell interference in one general cell, we formulate the above equations as matrix equations, such that the results can be derived from stochastic fixed-point equations. In order to do so, we define the row vectors

$$E\left[\left(\bar{\mathcal{I}}^{oc}\right)^k\right][x] = E\left[\left(\mathcal{I}_x^{oc}\right)^k\right], \tag{4.37}$$

$$\bar{\mathcal{N}}[x] = WN_0, \tag{4.38}$$

and the matrix

$$E\left[\tilde{\mathcal{F}}^k\right][x,y] = \begin{cases} E\left[\left(\mathcal{F}_{x\to y}\right)^k\right], & \text{if } x \neq y \\ 0, & \text{if } x = y \end{cases}, \tag{4.39}$$

then the system of linear equations in (4.31) and (4.32) can be written as

$$E\left[\bar{\mathcal{I}}^{oc}\right] = \left(E\left[\bar{\mathcal{I}}^{oc}\right] + \bar{\mathcal{N}}\right) \cdot E\left[\tilde{\mathcal{F}}\right]. \tag{4.40}$$

Let $\tilde{I}$ be the identity matrix, the mean interference vector $E\left[\bar{\mathcal{I}}^{oc}\right]$ can be solved from

(4.40) through matrix inversion

$$E\left[\bar{\mathcal{I}}^{oc}\right] = \bar{\mathcal{N}} \cdot E\left[\tilde{\mathcal{F}}\right] \cdot \left(\tilde{I} - E\left[\tilde{\mathcal{F}}\right]\right)^{-1}. \tag{4.41}$$

Similarly, the linear equations in (4.35) is formulated as

$$E\left[\left(\bar{\mathcal{I}}^{oc}\right)^2\right] = \bar{\mathcal{H}} + E\left[\left(\bar{\mathcal{I}}^{oc}\right)^2\right] E\left[\tilde{\mathcal{F}}^2\right] \tag{4.42}$$

with $\bar{\mathcal{H}}$ being the corresponding row vector of $\mathcal{H}_x$, and the vector of second moments of other-cell interference is accordingly derived as

$$E\left[\left(\bar{\mathcal{I}}^{oc}\right)^2\right] = \bar{\mathcal{H}} \cdot \left(\tilde{I} - E\left[\tilde{\mathcal{F}}^2\right]\right)^{-1}. \tag{4.43}$$

Then the corresponding variance values can be computed straightforwardly with both moments solved.

4.5.3 Accuracy Enhancement of the Approximation

In the derivation above, we made the assumption of complete independence between $\mathcal{F}_{y\to x}$ and $\mathcal{I}_y^{oc}$. However, due to the feedback behavior caused by mutual influences among transmission power of all the UEs, the correlation between these two random variables would introduce certain error in the suggested analytical model. In this section, we try to investigate how to alleviate such errors.

If the dependence between $\mathcal{F}_{y\to x}$ and $\mathcal{I}_y^{oc}$ is considered, the first moment of $\mathcal{I}_x^{oc}$ becomes from the equations (4.31) - (4.32) to

$$E\left[\mathcal{I}_x^{oc}\right] = \sum_{y\neq x} \left[W N_0 \cdot E\left[\mathcal{F}_{y\to x}\right] + E\left[\mathcal{I}_y^{oc}\right] E\left[\mathcal{F}_{y\to x}\right] \cdot \left(1 + \mathcal{G}_{\mathcal{I}_y^{oc},\mathcal{F}_{y\to x}}\right)\right] \tag{4.44}$$

with a new function $\mathcal{G}$ introduced to represent the correlation between $\mathcal{F}_{y\to x}$ and $\mathcal{I}_y^{oc}$. It is defined as

$$\mathcal{G}_{x,y} = c_x \cdot c_y \cdot \rho_{x,y} \tag{4.45}$$

where c_x, c_y are the coefficients of variation of x and y, and $\rho_{x,y}$ is the correlation coefficient between both random variables. If we determine the value of $\mathcal{G}_{\mathcal{I}_y^{oc},\mathcal{F}_{y\to x}}$

from the Monte-Carlo simulations, the modified mean of other-cell interference can be accordingly calculated from matrix inversion.

Similarly, for the second moment, it can be represented as

$$E\left[(\mathcal{I}_x^{oc})^2\right] = \hat{\mathcal{H}}_x + \sum_{y\neq x} E\left[\left(\mathcal{I}_y^{oc}\right)^2\right] E\left[\mathcal{F}_{y\to x}^2\right] \cdot \left(1 + \mathcal{G}_{\left(\mathcal{I}_y^{oc}\right)^2, \mathcal{F}_{y\to x}^2}\right) \tag{4.46}$$

where

$$\begin{aligned}\hat{\mathcal{H}}_x =& E\left[\mathcal{I}_x^{oc}\right]^2 + \sum_{y\neq x}\left[(WN_0)^2 \cdot E\left[\mathcal{F}_{y\to x}^2\right] + 2WN_0 \cdot E\left[\mathcal{I}_y^{oc}\right] \cdot E\left[\mathcal{F}_{y\to x}^2\right]\left(1 + \mathcal{G}_{\mathcal{I}_y^{oc}, \mathcal{F}_{y\to x}^2}\right)\right] \\ &- \sum_{y\neq x}\left[WN_0 \cdot E\left[\mathcal{F}_{y\to x}\right] + E\left[\mathcal{I}_y^{oc}\right] E\left[\mathcal{F}_{y\to x}\right]\left(1 + \mathcal{G}_{\mathcal{I}_y^{oc}, \mathcal{F}_{y\to x}}\right)\right]^2,\end{aligned} \tag{4.47}$$

which can be solved following the same approach before, with the values of $\mathcal{G}_{\mathcal{I}_y^{oc}, \mathcal{F}_{y\to x}}$, $\mathcal{G}_{\mathcal{I}_y^{oc}, \mathcal{F}_{y\to x}^2}$ and $\mathcal{G}_{\left(\mathcal{I}_y^{oc}\right)^2, \mathcal{F}_{y\to x}^2}$ given from simulations.

It can be seen that if we can determine function $\mathcal{G}$ analytically, the accuracy enhancement can be performed in a totally analytic way. Taking the first moment derivation as an example, $c_{\mathcal{F}_{y\to x}}$ is given in previous section, while $c_{\mathcal{I}_y^{oc}}$ can be iterative computed since (4.44) can be regarded as a fixed-point equation, thus the only part undetermined is the correlation coefficient $\rho_{\mathcal{I}_y^{oc}, \mathcal{F}_{y\to x}}$, on which more investigation should be devoted in the future.

4.5.4 Outage Probability Analysis

Unlike the outage events defined before by the limitation of received power in the UMTS networks with QoS traffic, while in the UMTS enhanced uplink, the outage event is now defined by the received load at NodeB described in the resource allocation section. An outage event occurs if

$$\eta_x > \eta_x^{max}, \tag{4.48}$$

which can be easily mapped into the interference representation as

$$I_x > \frac{\eta_x^{max}}{1-\eta_x^{max}} W N_0 = I_x^{max}. \tag{4.49}$$

The outage probability is given as the tail probability of the log-normal distributed random variable $\mathcal{I}_x^{oc}$, depending on whether the E-DCH traffic can feed the own-cell received interference to I_{own}^* or not

$$P_x^{outage} = \sum_{\hat{n}_x^D \in \mathbf{S}, \hat{n}_x^E = 0} P(\hat{n}_x^D, 0) P_{out}^D\left(\hat{n}_x^D, 0\right) + \sum_{\hat{n}_x^D, \hat{n}_x^E \in \mathbf{S} \setminus (\hat{n}_x^E = 0)} P(\hat{n}_x^D, \hat{n}_x^E) P_{out}^*. \tag{4.50}$$

The conditional outage probabilities $P_{out}^D\left(\hat{n}_x^D, 0\right)$ and P_{out}^* are given by

$$\begin{aligned} P_{out}^D\left(\hat{n}_x^D, 0\right) &= \Pr\left\{\mathcal{I}_x^{oc} > I_x^{max} - I_x^D\right\} \\ &= 1 - LN_{\mu_{\mathcal{I}_x^{oc}}, \sigma_{\mathcal{I}_x^{oc}}}\left[\left(1-\eta_{x,own}^D\right) \cdot W N_0 \left(\frac{\eta_x^{max}}{1-\eta_x^{max}} - \frac{\eta_{x,own}^D}{1-\eta_{x,own}^D}\right)\right], \end{aligned} \tag{4.51}$$

and

$$\begin{aligned} P_{out}^* &= \Pr\left\{\mathcal{I}_x^{oc} > I_x^{max} - I_{own}^*\right\} \\ &= 1 - LN_{\mu_{\mathcal{I}_x^{oc}}, \sigma_{\mathcal{I}_x^{oc}}}\left[\left(1-\eta_{own}^*\right) \cdot W N_0 \left(\frac{\eta_x^{max}}{1-\eta_x^{max}} - \frac{\eta_{own}^*}{1-\eta_{own}^*}\right)\right] \end{aligned} \tag{4.52}$$

where the variables $\mu_{\mathcal{I}_x^{oc}}$ and $\sigma_{\mathcal{I}_x^{oc}}$ are the corresponding parameters of Gaussian random variable $\ln\left(\mathcal{I}_x^{oc}\right)$, which can be derived straightforwardly from the moments of $\mathcal{I}_x^{oc}$ as in (3.59) - (3.60), and LN is the log-normal distribution function as

$$LN_{\mu,\sigma}\left(x\right) = \frac{1}{2} + \frac{1}{2} erf\left[\frac{\ln\left(x\right) - \mu}{\sigma\sqrt{2}}\right] \tag{4.53}$$

where the error function is given as

$$erf\left(x\right) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \mathbf{d}t. \tag{4.54}$$

4.6 Models Validation

Again we consider a two-tier hexagonal cell ring area which consists of 19 NodeBs under both time-based and volume-based source traffic models. The numerical results are calculated from the proposed analytical models according to the assumed values of system and traffic parameters, while the simulation results are obtained with same input parameters in order to validate the models. The following system parameters are assumed throughout all the results in this section, which are system bandwidth $W = 3.84$MHz, background thermal noise density $N_0 = -174$dBm, and PLE $\gamma = 4$. Again only the samples from central cell are counted in order to avoid border effect.

4.6.1 Results with Time-Based Traffic Model

Two service categories, DCH and E-DCH traffic, are supported in the system, where the traffic parameters under time-based model are assumed in Table. 4.1. Since the user distribution under such traffic model complies with a truncated Poisson distribution as discussed in the previous section, the mean user number can be represented straightforwardly as a product form solution. Thus we can simply assume the mean user number directly rather than arrival and departure rates for each class and apply Monte Carlo simulation.

Table 4.1: Traffic parameters of time-based traffic model

Service Categories	Bit Rate R	Target E_b/I_0 ε^*	Activity υ	Mean User Number
DCH	12.2kbps	5.5dB	1	66.7%
E-DCH	$\geq$5kbps	5.5dB	1	33.3%

Based on the values of the assumed traffic parameters in the above table, Fig. 4.7 illustrates the comparisons of mean other-cell interference obtained from Monte Carlo simulations, from proposed analytical model, and from analytical model with accuracy enhancement, while Fig. 4.8 demonstrates the corresponding standard deviation comparisons. The unit of Y axis in Fig. 4.7 - 4.10 is mW. The mean number of total users in one cell N (i.e. the mean in the spatial homogeneous Poisson process) is assumed to be 9, 15 and 21 respectively, with the ratio between DCH and E-DCH

users fixed at 2:1 as shown in the table. Since the mean numbers of users assumed in these three scenarios are relatively large, the probabilities of no E-DCH user in one cell are therefore quite small, thus the means of other-cell interference would be similar to each other, and then only one set of results with $N = 15$ is shown. To better illustrate the system behavior under lower target load, the load range from 0.2 to 0.5 is selected. From the figures, we can see the mean values from all the models achieve excellent match, while the standard deviations show slightly greater discrepancy in the case of higher load, which is due to the mutual independence assumptions made among $\mathcal{I}_{y\to x}^{out}$ as well as between $\mathcal{I}_x^{oc}$ and $\mathcal{F}_{y\to x}$. As the latter independence assumption has been released in the enhanced analytical model (solid line), better accuracy has been achieved in most cases than the plain analytical approximation model (dashed line).

Fig. 4.9 and 4.10 demonstrate the same comparisons, but in the cases where the mean DCH user number is fixed at 10 and the mean E-DCH user numbers are quite small. The reason of such validation is to test the models with cases when there is no E-DCH user in the system, which is also common in practice. Again excellent matches have been achieved, which implies our models are also valid in these scenarios. The good matches under all these user patterns not only verify the validity the log-normal approximation model, but also suggest the possibilities of applying such a model into practical network planning.

In Fig. 4.11, the logarithm outage probability versus own-cell target load is shown under the assumption that maximum load $\eta^{max} = 0.85$. Again, three scenarios are considered and all the analytical and simulation results match well. It can also be seen that the larger the number of users, the lower the outage probability due to lower standard deviation of the other-cell interference as shown in Fig. 4.8. In this figure, the load range is picked up from 0.42 to 0.54. The reason for this is that the outage events occur much rarely in the lower load end, thus the results do not make much sense if the load range is set lower than 0.42; in the higher load end, the outage probabilities reach as high as 10% when target load is 0.54, which is actually already not acceptable in practice. Similar reason applies when choosing the load range in Fig. 4.14.

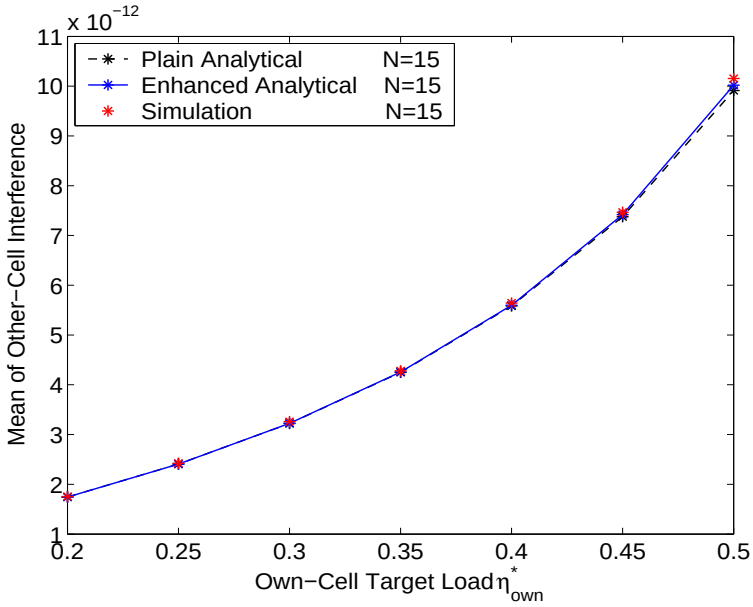


Figure 4.7: Mean of other-cell interference under time-based traffic model

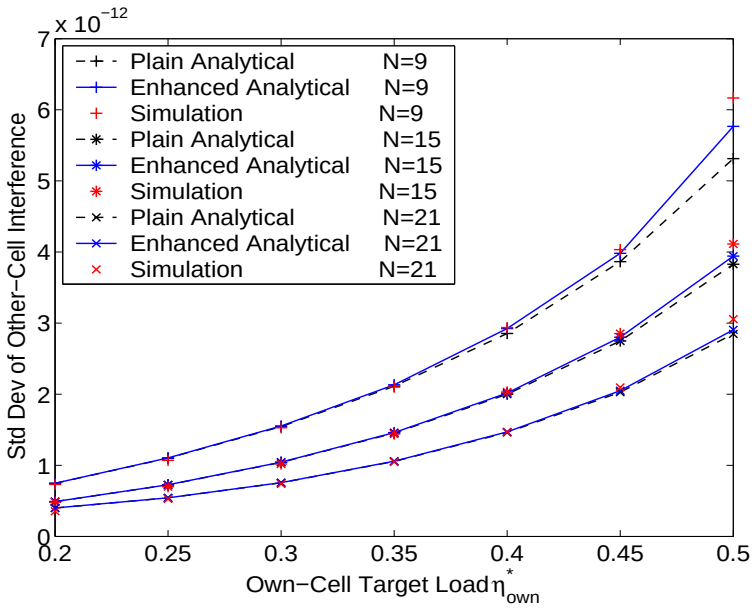


Figure 4.8: Standard deviation of other-cell interference under time-based traffic model

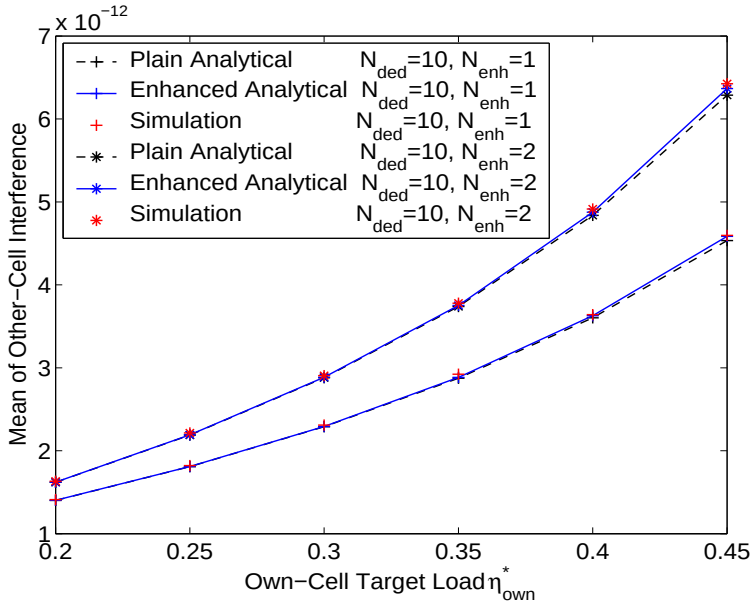


Figure 4.9: Mean of other-cell interference when E-DCH user number is relatively small

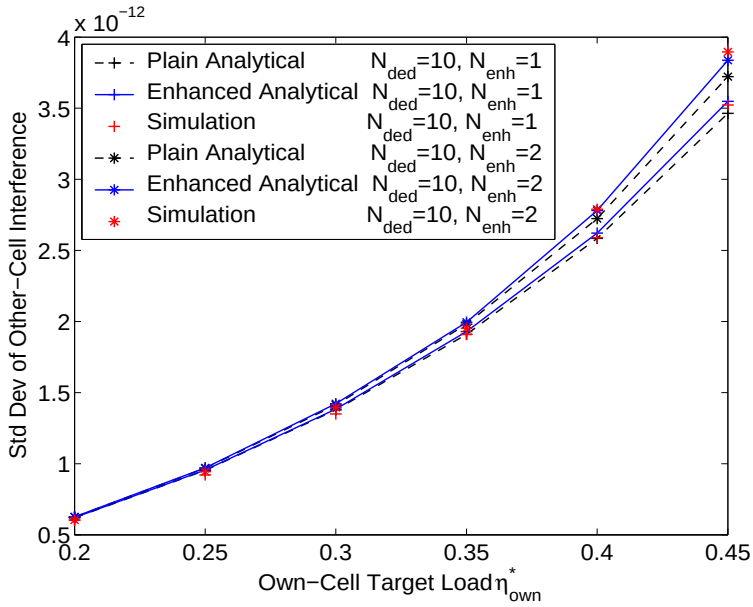


Figure 4.10: Standard deviation of other-cell interference when E-DCH user number is relatively small

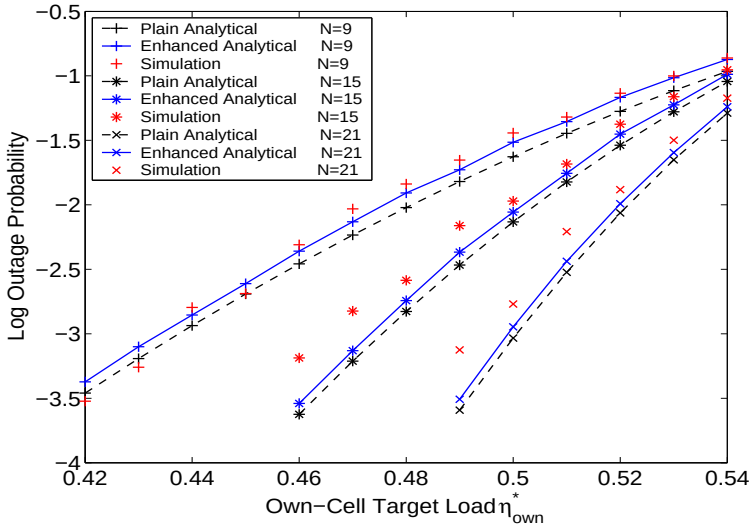


Figure 4.11: Log outage probability under time-based traffic model

4.6.2 Results with Volume-Based Traffic Model

The traffic parameters for the DCH and E-DCH service categories under volume-based model are presented in Table. 4.2. As the E-DCH user departure rate depends on the system state in this traffic model, we do not know the mean service time for E-DCH users, thus cannot suggest the mean user number straightly. Instead, we would assume the arrival rates for both DCH and E-DCH services, departure rates for the DCH service, and data volume size for the E-DCH service. We choose some different traffic parameter values from the above experiment to verify the analytical models more thoroughly.

Table 4.2: Traffic parameters of volume-based traffic model

Service Categories	Bit Rate R	Target E_b/I_0 ε^*	Activity υ	Mean Interarrival Time T_{arr}	Mean Service Time T_{svc}
DCH	64kbps	3dB	1	10s	100s
E-DCH	≥9kbps	3dB	1	40s	V^E/R^E

The Fig. 4.12 and 4.13 illustrate the mean and variance comparisons of other-cell interference from analytical model and simulations under various target load η^*, but this time with volume-based traffic generator. The mean volume size of E-DCH users

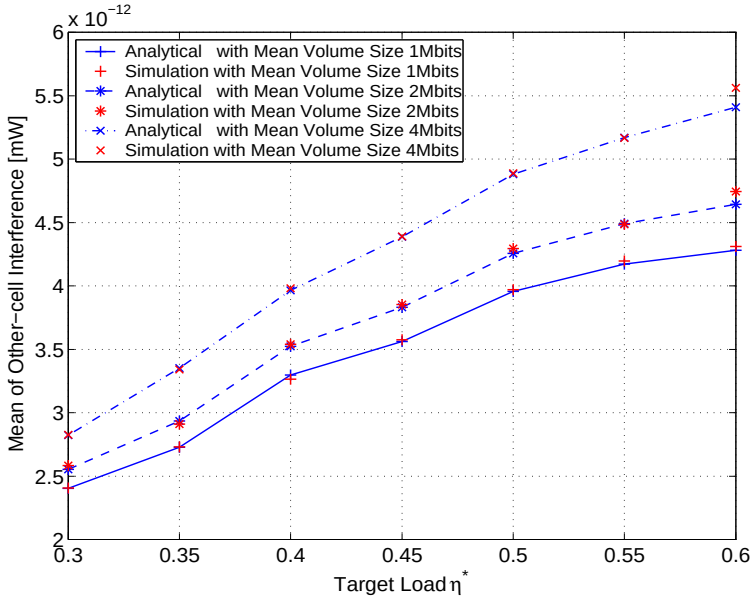


Figure 4.12: Mean of other-cell interference under volume-based traffic model

V_{svc}^{E} are set to 1Mbits, 2Mbits and 4Mbits, corresponding to three curves in each figure respectively. Again from such figures we can prove the good match, where the errors caused by independence assumption can be alleviated by the accuracy enhancement techniques. In Fig. 4.14, the logarithm outage probability versus target load η^* is shown as an application of other-cell interference modeling for the network planning.

4.7 Discussion

In this chapter, we built analytical interference models for the UMTS enhanced uplink, which can be regarded as an extension work of the models constructed in the previous chapter. The incoming elastic traffic are characterized by both time-based and volume-based source traffic models, where the user distribution are obtained by either applying product form or solving multi-dimensional state-dependent rate Markov chain, respectively. We again present the analytical interference models firstly from the iterative approach where many convolutions involved, followed by an approximation model such that the computation complexity has been greatly reduced and thus numerical tractable in light of the approximation techniques before. Then, the

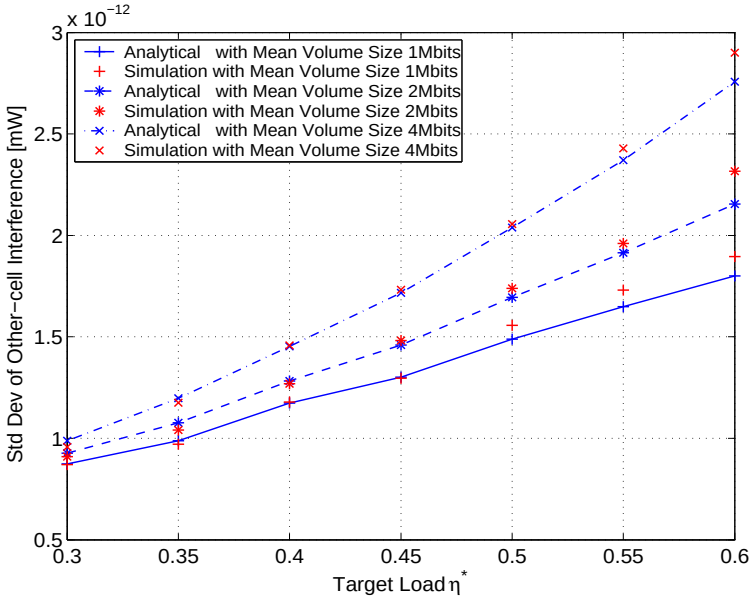


Figure 4.13: Standard deviation of other-cell interference under volume-based traffic model

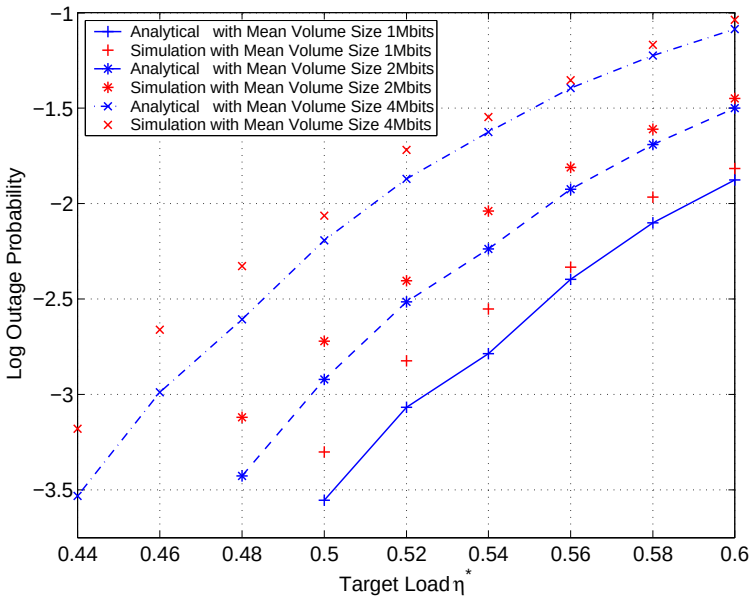


Figure 4.14: Log outage probability under volume-based traffic model

approaches to enhance the accuracy of the suggested approximation models as well as to apply other-cell interference characterization into outage probability estimation are introduced afterwards.

Both the numerical results from analytical models and the simulation results are illustrated to verify the suggested approximation model. The first and second moments of other-cell interference obtained from analytical models and simulations show excellent match, especially in the lower load region. When the target load increases, the errors over standard deviations slightly go up due to weak independence assumptions made during calculations, which is later compensated by introducing the accuracy enhancement techniques.

In our current analysis, the volume-based source traffic model consists of a $T+1$ dimensional Markov chain and thus large number of states involved during the calculation. In the future, we can aim to further reduce this complexity by employing the Kaufman-Roberts recursion similar to [60], such that the T dimensional state space for DCH users can be combined into one dimension. Furthermore, investigation on the interference model under different resource allocation schemes in the UMTS enhanced uplink is also worth more effort.

The work in this chapter focuses on the system behavior with best-effort traffic under time-based and volume-based traffic models. When comparing these two models, we study the system performance with various input values of the parameters which may impact the results significantly, for example, the number of total users, the volume of enhanced uplink users, etc. For those parameters which generally have similar impacts on both traffic models, they are used as constants during the work. In the future, it would be useful to study the system behavior with different values of these parameters, for example, voice activity factor, bit rate, target E_b/I_0, etc.

Chapter 5

Optimization-Based Resource Allocation Strategies in the UMTS Enhanced Uplink

In this chapter, in contrast to study on the distributed radio resource management, we investigate the system behavior under centralized resource assignment. Such 'greedy' scheme manages to maintain the total received load up to the target load. This is achieved by a few suggested optimization-based resource allocation strategies, both non-linear global optimization and linear-programming approaches. The motivation of employing linear-programming techniques over non-linear global optimization is due to much less computational complexity, and from the comparison of theoretical optimal results between these techniques demonstrates that linear-programming approach can also achieve near-optimal performance. The major work of this chapter is to provide a framework based on optimization to calculate the resource assigned to each E-DCH user in the multi-cell environment. The system performance under various resource allocation schemes are analyzed and compared by Monte-Carlo simulations, which shows the proposed framework can serve as a good estimation and reference to study how systems perform for the network operators.

5.1 Introduction

The analysis in chapter 4 is concentrated on the distributed resource management scheme, which is implemented in the UMTS enhanced uplink by moving the RRM entity from RNC side to each NodeB. This feature allows much rapidity and flexibility, as well as leads to shorter signaling delays and consequently enables fast reactions in the resource allocation processes. However, the concomitant drawback of such scheme is that resource allocation is performed with local knowledge of own-cell load only. This makes it more difficult to avoid the load overshoot, thus it is expected more outage cases may occur. In this chapter, we focus on the centralized resource management schemes, which is based on the assumption that the NodeBs are aware of the load information of neighboring cells. In such scheme, the interference contributed by users in other cells should be considered when making decisions for resource allocation. A few optimization-based approaches are proposed for this purpose, where the system performance and feasible load regions under these strategies are studied and compared.

For the 3G systems with packet data traffic, one analytical model of the UMTS enhanced uplink has been constructed in [47], where the 'greedy' resource allocation strategy, which aims to maximize the resource utilization, is employed. The other-cell interference in this analysis is modeled as an independent log-normally distributed random variable, but in fact in the multi-cell environment, it depends on the interference level at each NodeB due to mutual influence over the whole system. If such correlations and interactions are considered, the 'greedy' resource allocation may result in unnecessary outage cases.

The work in [73] is one of the very first to investigate the target SIR power control schemes for the co-channel interference management in the general cellular radio systems. The outage probability, which is defined as the probability that interference exceeds certain threshold, is selected as the performance measure. The results show the proposed centralized RRM performs much better than the target received power RRM schemes.

Several works in the existing literature formulate the resource allocation as optimization problems as well, with different constraints and utility functions. In [6], the goal is to maximize the network throughput of CDMA system, subject to the upper

bound of blocking probability and the lower bound of bit energy to interference ratio. The recent work in [44] studies both centralized and decentralized RRM strategies in the UMTS enhanced uplink, with and without one central node coordinating the resource assignment according to the network interference. The solution also comes out from the optimization problem, with again the network throughput utility function, and several constraints that guarantee the system feasibility. In [10], some important results on the feasibility region of the CDMA uplink power assignment problem have been found. It is shown that the solution set of the problem is log-convex if the QoS requirements for the link are convex in the log domain itself, which makes such problem solvable within reasonable time with standard algorithms. The work in [27] discusses a distributed mechanism to jointly optimize the SIR and transmit power in a cellular mobile network. A new characterization of feasible SIR region in terms of load factor as well as potential interference from mobile users called spillage are introduced, based on which the optimal SIR assignment can be calculated without centralized computation. In [31,61], the rate allocation schemes are designed not only to maximize the overall system throughput, but also to concern the fairness of various criteria. The non-linear global optimization problems in both papers are converted into convex problems with linear constraints in order to reduce the complexity.

The first main contribution of this chapter is to raise a so-called 'soft outage' problem which elucidates the reason why the 'greedy' centralized resource allocation mechanisms employed in the single-cell model [47] cannot be directly applied into multi-cell UMTS environments. We then suggest some optimization based approaches for resource allocation to account for the above problem via considering various constraints. An interference model that describes the interactions among the NodeBs is constructed on the basis of such RRM in order to investigate the system performance in terms of outage probability, average traffic load, resource utilization, etc. Additionally, the newly introduced 'down grants' mechanism in the enhanced uplink, which aims to lower the impact of other-cell interference caused by the mobiles in soft handover area, can also be easily included through some specific constraints. Therefore, the suggested optimization-based resource allocation framework can be employed by the network operators for efficient estimations of system performance.

The rest of this chapter is organized as follows. Section 2 describes the basic

centralized resource management principles in the UMTS enhanced uplink, and the 'soft outage' problem caused by 'greedy' resource allocation. In Section 3, we first construct the interference model under centralized RRM, and then present a few optimization based resource allocation strategies to analyze the system behavior, with newly introduced enhanced uplink feature - 'down grants' possibly included. Feasible load regions under the proposed resource allocation mechanisms are investigated and compared in Section 4. Section 5 illustrates numerical results obtained from Monte-Carlo simulations, and the work in this chapter is summarized in Section 6.

5.2 Centralized Radio Resource Management Principles in the Enhanced Uplink

As discussed and illustrated in the interference and load models for the UMTS networks before, the shared resources in the uplink are the interference power or the load received at NodeB, while the remaining capacity is defined by the margin to the target load η^*. A higher target load means more resources and higher bit rates for the enhanced uplink users, but also implies an increase of the probability of load overshoots which may lead to outage events in the worst cases. On the other side, a lower target load leads to a more stable system, but may also result in insufficient resources for the best-effort users and in turn low resource utilization. How resources are assigned to the admitted users is in the hand of the network operators. In general, the resources consumption depends on the data rate of each user. Since the data rates for DCH users are fixed once they are admitted into the system, in order to maximize the system capacity, the rate scheduler is responsible for instantly adjusting the data rates, and in turn the transmission power of best effort users in the enhanced uplink, such that the received uplink load is kept as close as possible to but below the target load.

There are various resource management ways to maintain the received load as the target load. In the previous chapter, we have demonstrated the distributed resource assignment scheme where the allocation decisions are made by individual NodeBs with mere local knowledge. In such mechanism, each NodeB attempts to maintain the load received from own-cell users to the pre-defined system parameter: target own-cell load

η^*_{own}. The received load from other-cell users can be accordingly approximated as a log-normal random variable η_{oc} with distribution represented as a function of η^*_{own}. Then, by adjusting the single parameter η^*_{own}, we can adapt the total received load at one NodeB, which is simply the sum of both components $\eta = \eta^*_{own} + \eta_{oc}$, to the target load η^* mentioned above.

Another approach for the resource management is based on centralized allocation, which is to be discussed in this chapter. Suppose all the NodeBs in the network share local load information with each other, they can make decisions considering global received loads. Following the same notations on the cell load as before, the goal of resource allocation is simply represented as $\eta = \eta^*$ for each NodeB. Such 'greedy' resource allocation strategy, which aims to 'waterfill' the total cell load up to the target load, is illustrated in Fig. 5.1. This 'greedy' RRM can be interpreted as the uplink equivalent of resource allocation strategies for the best-effort services in downlink bearers (e.g. HSDPA or 1xEV-DO), since it assigns as much as possible resources to the elastic traffic. If the target loads are reached while not exceeded for all the NodeBs, the system is operating in an optimal status, since the resources are fully utilized. The received loads and interference in such ideal scenarios can be determined straightforwardly by solving the corresponding linear equations matrix. However, in some cases, the above requirements cannot be satisfied globally, where the scenarios and underlying reasons are discussed in the following.

As stated above, one goal of the resource assignment mechanism is to ensure that the cell load does not go beyond the target load. If it overshoots, the QoS of all the connected DCH and E-DCH users may drop below the acceptable threshold, which we define as an *Outage* event similar as before. With distributed resource management described in the previous chapter, an outage event occurs when the loads received from other-cell users exceed allowed load margin, while with centralized resource management, there are generally three kinds of scenarios which may cause outage occurrences.

1. At one NodeB in the system, the received load contributed by only DCH users (including own-cell and other-cell) exceeds the target load η^*, which is, the maximum load is overshoot by mere DCH users.

2. Each E-DCH user in the system must have a certain minimum rate which is

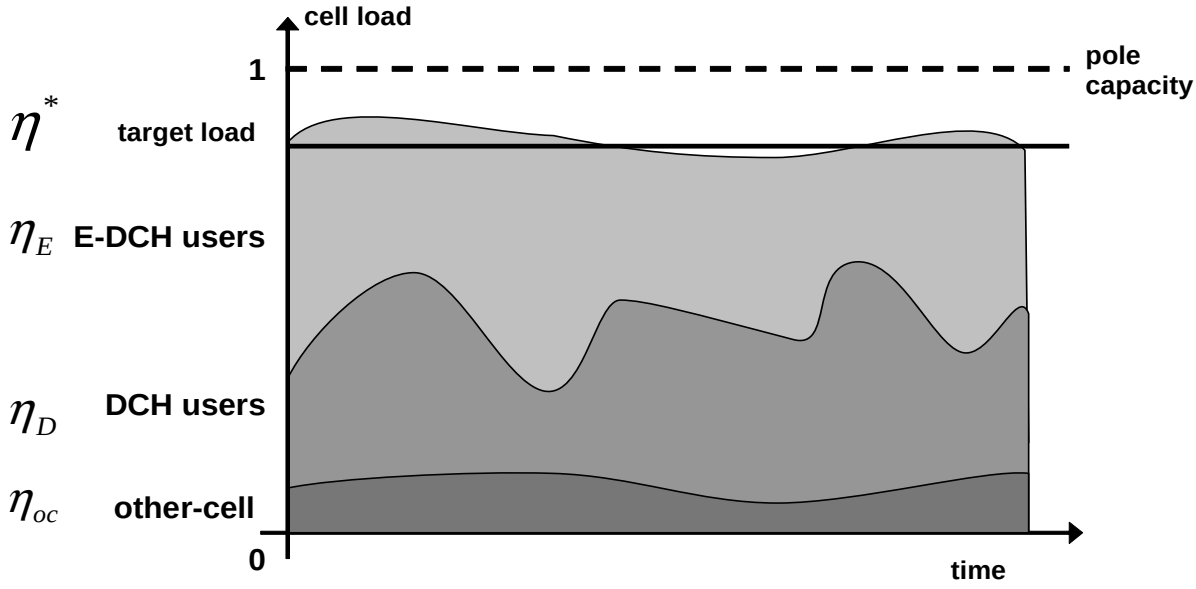


Figure 5.1: Cell load under greedy resource allocation strategy

larger than zero for elastic data transmission, which is denoted as R^E_{min}. This corresponds to the minimum E-DCH load experienced by the NodeB. Therefore, the second scenario can be described as the received load from DCH users is less than η^*, but with the minimum load generated by E-DCH users (i.e. each E-DCH user is only allocated R^E_{min}), the cell load goes beyond its limitation η^*. Now the outage occurs due to DCH users and minimum rate E-DCH users.

3. Suppose the system can accommodate all the DCH users as well as the E-DCH users with minimum rates R^E_{min}, however, under the greedy resource allocation scheme, if each NodeB manages to waterfill up to the target load, the received cell load at some NodeBs will overshoot η^* due to the interference from E-DCH users in the other cells. The following simple scenario of a two-cell system demonstrates why such outage occurs. Assuming there are one DCH user $D1$ and one E-DCH user $E1$, both of which are served by NodeB 1, and also one E-DCH user $E2$ which is connecting to NodeB 2, with location of each user as shown in Fig. 5.2. Since both $D1$ and $E1$ are closer to NodeB 1 than to NodeB 2, the received interference due to these two users at NodeB 1 must be higher

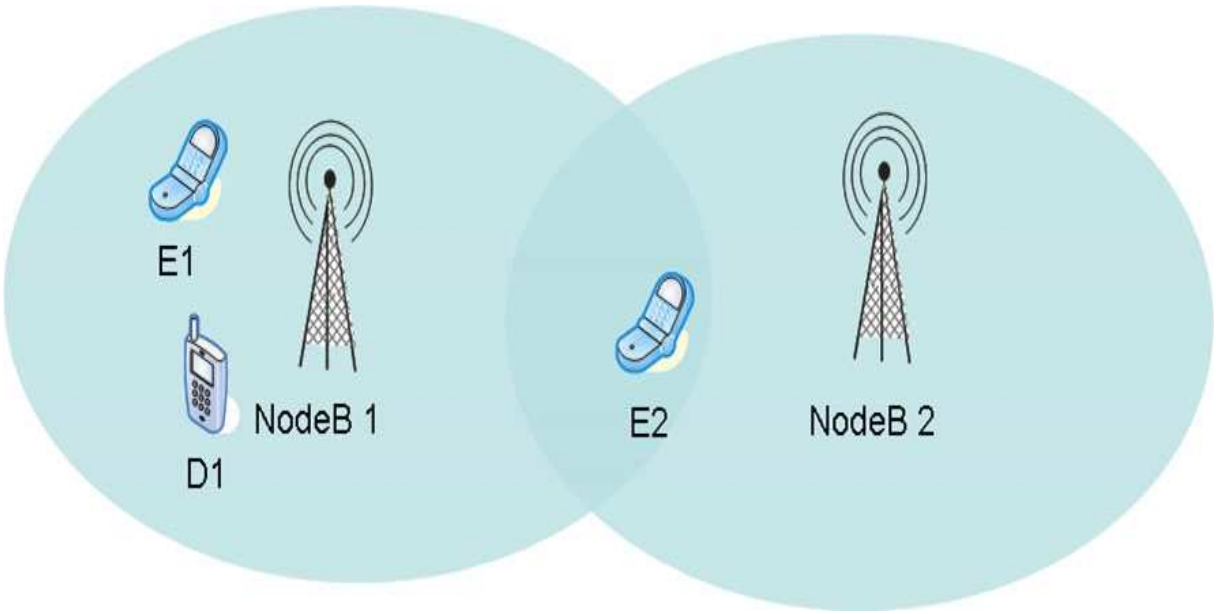


Figure 5.2: Simple scenario

than that of NodeB 2, which implies some difference between the remaining load capacity at these two NodeBs. The user $E2$ is located in the edge of cell 2, such that the interferences contributed by this user to both NodeBs are similar. Then if the difference of the remaining load is large enough, no matter what rate $E2$ is choosing, the target load at both NodeBs cannot be satisfied simultaneously. From such a simple example, it can be seen that there are certain scenarios in which the target load cannot be achieved globally across the whole network. If the rate scheduler in this example manages to achieve the target load in NodeB 2, there would be an outage event occurred at NodeB 1. This type of outage is due to improper resource allocation scheme, since it can be actually avoided.

We categorize scenario 1 and 2 as *Hard Outage*, since the rates of DCH users and minimum rates of E-DCH users cannot be adjusted, and thus this is due to the system capacity limitation, which can be mostly avoided by introducing certain admission control mechanism. The analysis of such blocking probability is similar to the process demonstrated in Chapter 3. In scenario 3, such an allocation of rates to E-DCH users will lead to the unnecessary degradation of the QoS of DCH users, and thus we define

this as *Soft Outage* since it is not inevitable. For instance, in the above example, if the schedulers decrease the transmit power of UEs such that NodeB 1 satisfies the target load while NodeB 2 receives lower load, the outage will not occur any more although the resource utilization drops. One principle of resource allocation is that the best-effort elastic traffic from E-DCH users cannot affect the QoS of streaming traffic from DCH users, therefore in the following, we present a few optimization-based resource allocation strategies to account for the soft outage problem. Such strategies lower the load in certain cells in order to prevent the soft outage, while they make distinct decisions to choose in which cells and how much to decrease the received load due to different implementations. In theory, the soft outage cases can always be avoided (worst case is that all E-DCH users take minimum rate), thus the key point for this problem is how to keep the resource utilization as high as possible in the feasible solutions.

5.3 Optimization-Based Resource Allocation Strategies

In this section, we present the optimization-based resource assignment mechanisms to address 'soft outage' problem, while at the same time aiming at highest resource utilization. For better understanding of the proposed mechanisms, we would introduce the constructed interference model for the enhanced uplink first, followed by a few optimization implementations based on various utility functions and constraints. We define some common constraints as follows.

1. The target load or interference should not be exceeded. The purpose is to guarantee a stable system, since if the cell loads get too high, the required transmit powers for the mobiles tend to approach infinity, which makes it impossible for them to reach their required target E_b/I_0.

2. All E-DCH users have a certain minimum bandwidth guarantee, which corresponds to a minimum bit rate and thus to a minimum SLF ω^E_{min}. This condition avoids quasi-outage of users. Moreover, there is a maximum SLF ω^E_{max} limitation applying to E-DCH users as well.

3. Each mobile has its maximum transmit power. The instant transmit power can be calculated as the required receive power at serving NodeB divided by the propagation loss between the mobile and the serving NodeB. As this type of constraints varies with the mobile terminal technology, it is not included into the calculation in this chapter.

4. The 3GPP standards for the enhanced uplink states that DOWN grants are sent to mobiles in adjacent cells if the ratio between the other-cell interference from E-DCH users in the soft handover area and own-cell E-DCH interference exceeds a certain operator-defined threshold. This alleviates the flooding of cells due to high bit-rate mobiles near the cell borders.

The goal of the resource assignment is to make all these conditions fulfilled. Otherwise it may lead to the outage cases described as above. We name the load region defined by the above constraints as feasible load region. Depending on the resource allocation implementations and the degree of knowledge on the global load situation that the decision entity has, the feasible load region significantly differ from each other, which is to be discussed and compared after these optimization schemes being introduced.

5.3.1 Interference Model of the Enhanced Uplink

Assuming there are totally $\mathcal{N}$ cells in the system, then $I_1, ..., I_{\mathcal{N}}$ and $\omega_1^E, ..., \omega_{\mathcal{N}}^E$ are the received interference and allocated SLF from cell 1 to $\mathcal{N}$, respectively. The own-cell parameters for a NodeB x, such as $I_{x,own}^D$, $I_{x,own}^E$, $\eta_{x,own}^D$ and $\eta_{x,own}^E$ under distributed resource allocation have been derived from the power control equation in the previous chapter. They are still applicable here under centralized resource management in this chapter as the equations for these variables are independent of the RRM schemes. For the other-cell parameters, the interference can be also divided into two components, $I_{x,oc}^D$ and $I_{x,oc}^E$, depending on the user categories, which are the corresponding sum of inter-cell received power $S_{k,x}^{inter}$ over all the DCH and E-DCH users connecting to the NodeBs in the system other than x

$$I_{x,oc}^D = \sum_{y \neq x} \sum_{k \in \mathcal{D}_y} S_{k,x}^{inter} = \sum_{y \neq x} \sum_{k \in \mathcal{D}_y} S_{k,y}^R \Delta_{y \to x}^k, \tag{5.1}$$

$$I^E_{x,oc} = \sum_{y \neq x} \sum_{j \in \mathcal{E}_y} S^{inter}_{j,x} = \sum_{y \neq x} \sum_{j \in \mathcal{E}_y} S^R_{j,y} \Delta^j_{y \to x}. \tag{5.2}$$

Then the sum of received power over all the sources yields the total interference I_x at NodeB x as

$$\begin{aligned} I_x &= I^D_{x,own} + I^E_{x,own} + I^D_{x,oc} + I^E_{x,oc} \\ &= \sum_{k \in \mathcal{D}_x} \omega^D_k \left(I_x + WN_0\right) + \sum_{j \in \mathcal{E}_x} \omega^E_j \left(I_x + WN_0\right) \\ &\quad + \sum_{y \neq x} \sum_{k \in \mathcal{D}_y} \omega^D_k \Delta^k_{y \to x} \left(I_y + WN_0\right) + \sum_{y \neq x} \sum_{j \in \mathcal{E}_y} \omega^E_j \Delta^j_{y \to x} \left(I_y + WN_0\right) \end{aligned} \tag{5.3}$$

which can be formulated as the following matrix equation form

$$\bar{I} = (\bar{I} + \bar{N}_0)\tilde{G}^D_{own} + (\bar{I} + \bar{N}_0)\tilde{G}^E_{own} + (\bar{I} + \bar{N}_0)\tilde{G}^D_{oc} + (\bar{I} + \bar{N}_0)\tilde{G}^E_{oc} \tag{5.4}$$

with $\bar{I}$ representing the received interference row vector $\bar{I} = [I_1, \ldots, I_N]$, $\bar{N}_0$ denoting a constant row vector $\bar{N}_0 = [WN_0, \ldots, WN_0]$. The own-cell coefficients matrices $\tilde{G}^D_{own}$ and $\tilde{G}^E_{own}$ are diagonal matrices as

$$\tilde{G}^D_{own}[x, y] = \begin{cases} \sum\limits_{k \in \mathcal{D}_x} \omega^D_k, & x = y \\ 0, & x \neq y \end{cases} \tag{5.5}$$

$$\tilde{G}^E_{own}[x, y] = \begin{cases} \sum\limits_{j \in \mathcal{E}_x} \omega^E_j, & x = y \\ 0, & x \neq y \end{cases} \tag{5.6}$$

and the other-cell coefficients matrices $\tilde{G}^D_{oc}$ and $\tilde{G}^E_{oc}$ consist of zeros in the diagonal and the sum of SLFs multiplied with the link gain ratios on the remaining entries, which are expressed as

$$\tilde{G}^D_{oc}[x, y] = \begin{cases} 0, & x = y \\ \sum\limits_{k \in \mathcal{D}_x} \omega^D_k \Delta^k_{x \to y}, & x \neq y \end{cases} \tag{5.7}$$

$$\tilde{G}_{oc}^{E}[x,y] = \begin{cases} 0, & x = y \\ \sum\limits_{j \in \mathcal{E}_x} \omega_j^E \Delta_{x \to y}^j, & x \neq y \end{cases} \tag{5.8}$$

There are two sets of variables undetermined in this model under centralized resource allocation, which are the interference vector $\bar{I}$ and the E-DCH users SLF vector $\bar{\omega}^E$. This is the major point which differentiates the interference models in the previous chapters, where only the interference vector is left unknown and thus it can be derived by mathematical transformation. There are several approaches to allocate resources to the E-DCH users in the enhanced uplink with centralized RRM, each of which leads to dissimilar system performance. In the following, we present a few of such strategies and compare them through performance evaluation.

5.3.2 Direct Assignment Approach

In most cases, the target interference is reached at all the NodeBs in the system, that is,

$$I_x = I^*, \quad \forall x \tag{5.9}$$

where the target interference uniquely corresponds to the equivalent target load as

$$I^* = \frac{\eta^*}{1 - \eta^*} \cdot W N_0. \tag{5.10}$$

Since the received interference equal to each other, we can accordingly divide the $I_x + W N_0$ expression at both sides of the above enhanced uplink interference model, which reduces it into a simple load model as

$$\eta^* = \sum_{k \in \mathcal{D}_x} \omega_k^D + \sum_{j \in \mathcal{E}_x} \omega_j^E + \sum_{y \neq x} \sum_{k \in \mathcal{D}_y} \omega_k^D \Delta_{y \to x}^k + \sum_{y \neq x} \sum_{j \in \mathcal{E}_y} \omega_j^E \Delta_{y \to x}^j. \tag{5.11}$$

As there is only one set of variable in such model, the way to calculate SLFs for the E-DCH users is straightforward, which is to solve the above load equation system for SLF vector $\bar{\omega}^E$. This consequently yields the rate assignment for each individual E-DCH user.

This direct assignment approach would lead to negative results in the soft outage scenarios, because it does not make any effort to avoid such outages. In these cases,

we assume the users with allocated SLF less than ω^E_{min} are just set to be equal to ω^E_{min}, which would result in load overshoot. The resultant E-DCH load as well as user utilization from this scheme would be very high since it does not remove the soft outage, and thus can be regarded as theoretical supremum reference.

5.3.3 Global Optimization Resource Allocation Scheme

During the normal system operation, the total received interference $\bar{I}$ in above equation at each NodeB should be equal to the target interference I^*, such that the static assignment policy applies. However, in the soft outage cases, as demonstrated in the example, the target interference cannot be satisfied at all NodeBs simultaneously, then the received interferences distinguish from each other and thus $\bar{I}$ becomes a vector of variables. To achieve high utilization, the aim of resource allocation is to maintain the received interference at each NodeB as close to the target interference I^* as possible. Since there exist the products of two sets of variables I and ω^E in the interference model, the resource allocation is actually a general non-linear global optimization problem.

The first step is to figure out the utility function in the optimization. The most straightforward one is to sum over all individual SLF of E-DCH users. However, this approach results in unfair assignment pattern in the sense that the users closer to the serving NodeB would get higher rate allocated than the users which are further away. A frequently mentioned generic fairness criterion is that of α-fairness, where the optimization converges to different fairness goals according to the setting of parameter α [50]. With such criterion, the utility function is expressed as

$$U = \sum_{x=1}^{\mathcal{N}} \sum_{j \in \mathcal{E}_x} \frac{\left(\omega^E_j\right)^{1-\alpha}}{1-\alpha} \tag{5.12}$$

With the above utility function, proportional fairness [33] can be achieved with $\alpha \rightarrow 1$ and max-min fairness [9] can be fulfilled with $\alpha \rightarrow \infty$. The optimization aims to maximize the utility function while subject to the following constraints

$$\omega^E_{min} \leq \omega^E_j \leq \omega^E_{max} \tag{5.13}$$

$$I_x \leq I^* \tag{5.14}$$

with x from 1 to $\mathcal{N}$ and I_x is formulated in (5.3).

Such problems are generally difficult to solve, however, in [10] it has been shown that the feasible set is convex if the optimization function is log-convex, that is, concave in the log-domain. Since the SLFs form a convex set, we can use standard non-linear programming solver like the gradient search algorithm to find the global maximum.

5.3.4 Linear-Programming Based Resource Allocation Schemes

The global non-linear optimization can be employed to find the optimal pattern, but it is quite a time-consuming task. Although for convex optimization problems the best solution can be found in a reasonable time for a small to medium-scale number of variables, the optimization of a whole UMTS network with possibly thousands of sectors and mobiles may become difficult. To reduce the computational complexity, we try to solve it based on some linear-programming techniques. Since there are two sets of variables existing in the problem which is impossible for linear-programming to optimize at the same time, the general idea is to fix one set as constants while optimizing only the other one set, followed by switching them step by step.

Unlike the above global optimization where each E-DCH user may be allocated its own rate, we assume parallel equal-rate scheduling discipline is employed with linear-programming rate assignment, such that the E-DCH users within one cell share the same rate, as well as the same SLF ω_j^E.

Non-Iterative Allocation

First we propose a non-iterative resource allocation scheme. When calculating the values for $I_1, ..., I_{\mathcal{N}}$ according to (5.3), we replace all the variables I at the right hand side with constant I^*, after which the problem can be solved with linear-programming techniques with respect to $\omega_1^E, ..., \omega_{\mathcal{N}}^E$.

For the utility function to be maximized, as it must be linear, we can not use the

same one as in (5.12), but the total E-DCH load in the system, which is

$$U = \sum_{x=1}^{\mathcal{N}} \eta_x^E = \sum_{x=1}^{\mathcal{N}} n_x^E \omega_x^E \tag{5.15}$$

where n_x^E is the number of E-DCH users connecting to NodeB x. And the constraints to be satisfied are

$$\omega_{min}^E \leq \omega_x^E \leq \omega_{max}^E \tag{5.16}$$

$$(I^* + WN_0)\left(\sum_{k\in\mathcal{D}_x} \omega_k^D + \sum_{j\in\mathcal{E}_x} \omega_x^E + \sum_{y\neq x}\sum_{k\in\mathcal{D}_y} \omega_k^D \Delta_{y\to x}^k + \sum_{y\neq x}\sum_{j\in\mathcal{E}_y} \omega_y^E \Delta_{y\to x}^j\right) \leq I^* \tag{5.17}$$

with x from 1 to $\mathcal{N}$.

However, the obtained solution from this approach may not be the optimal one due to its conservativeness. During the linear-programming optimization, the I_x for each NodeB is calculated based on the assumption that the received interferences at all the other NodeBs reach the target interference I^*, which results in lower ω_x^E accordingly. But in fact, the received interferences might be lower than I^* as well. To utilize this feature, the iterative approach is proposed in the following.

Iterative Allocation

The interference model in (5.3) can be interpreted as a set of fixed-point equations since the received interferences in different cells have the same distribution, therefore, some iteration-based methods may be applied to reach a more exact solution rather than the simple calculation with I^* approximation in the above approach.

Starting with the same linear-programming problem defined in (5.15)-(5.17) (i.e. target interference is assumed to be satisfied at each NodeB in the initial step), a set of 'semi-optimal' SLF ω_x^E for each cell x can be computed in the same way. Then if we substitute the obtained ω_x^E back into (5.3), together with the current values of each I_x (all are equal to I^* in the first step), the new received interference I_x at each NodeB can be simply calculated from the right hand expression of (5.3), and now they begin to distinguish from each other due to different user patterns in each cell. Again the linear-programming function should be called in the remaining steps, with

the constraints of (5.17) replaced by

$$(I_x + WN_0)\left(\sum_{k\in\mathcal{D}_x}\omega_k^D + \sum_{j\in\mathcal{E}_x}\omega_x^E\right) + \sum_{y\neq x}(I_y + WN_0)\left(\sum_{k\in\mathcal{D}_y}\omega_k^D\Delta_{y\to x}^k + \sum_{j\in\mathcal{E}_y}\omega_y^E\Delta_{y\to x}^j\right) \leq I^* \tag{5.18}$$

where the variables I_x and I_y take the updated values.

Therefore, each iteration in the following comprises the linear-programming optimization according to (5.15) (5.16) and (5.18) with updated received interference from the previous step and the updating calculation of received interference in each cell according to (5.3). The iteration converges if the relative changes of all I_x fall below a certain threshold. All the I_x will then be certainly lower than the target interference I^* after convergence due to the linear-programming constraints, thus the obtained ω_x^E can be used for rate allocation for each cell.

In some scenarios, the final convergence cannot be reached under certain user patterns. In such cases, one modification should be applied to the above iterative algorithm each time when updating received interference. Since iteratively calculating I_x from the right hand expression of (5.3) does not converge, we could construct (5.3) as a linear equations system with size $\mathcal{N}$ instead and solve for I_x. The verification of whether all I_x are lower than the target interference I^* is now required for the solution feasibility. Then with updated received interference, the linear-programming optimization remains same as before for ω_x^E. After certain repetitions, the best rate assignment can be selected from the candidate solutions.

5.3.5 Impacts of Down Grants Commands

As described before, the scheduling in the enhanced uplink is implemented based on requests (from UE to NodeB) and grants (from NodeB to UE) framework. Two types of grants, absolute grants and relative grants, are used, where absolute grants set the absolute value of the transmission power upper limit of an E-DCH user, and relative grants update the power limitation based on the previous value to go 'UP', 'DOWN' or simply 'HOLD'.

Soft handover is also supported in the enhanced uplink which provides macro

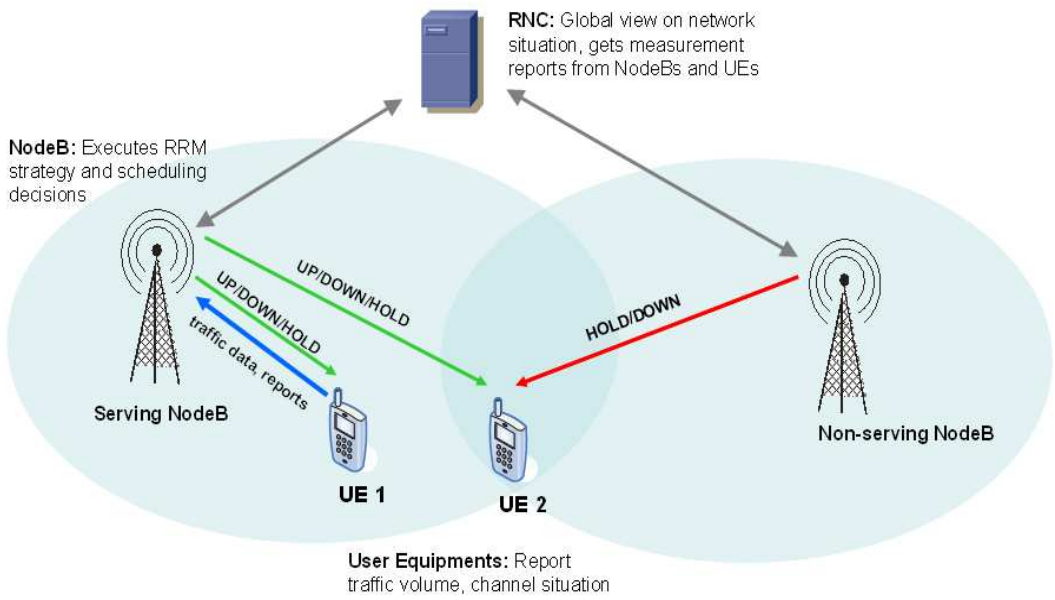


Figure 5.3: Scheduling in UMTS enhanced uplink

diversity gains. Hence grants can be received from both serving NodeB and non-serving NodeBs. The absolute grants can only be received from serving NodeB, while relative grants may be received from both serving and non-serving NodeBs. However, the non-serving NodeBs can only send the relative grants of 'DOWN' or 'HOLD' to decrease the other-cell interference that they are experiencing. The scheduling scheme in UMTS enhanced uplink is illustrated in Fig. 5.3.

The relative grants 'DOWN' command sent by non-serving NodeBs is a new feature introduced in the enhanced uplink. It serves as an 'overload indicator' which is mainly responsible to suppress the other-cell interference from soft handover area. Two criteria for non-serving NodeBs to determine when to send 'down grants' are listed in [4], where the first one restrains the total received interference and the second one limits the E-DCH load contributed by the UEs in soft handover area of each NodeB.

It would not be a hard task to include the impact on the system behavior caused by such 'down grants' into the interference model since this only implies that more constraints should be augmented in the optimization. In particular, the first criterion regarding total received interference have already been included in the model in (5.14)

in global optimization, or (5.17) and (5.18) in linear-programming optimization. The remaining criterion concerning E-DCH load from users in soft handover area can be written as

$$\frac{\sum_{y \neq x} \sum_{j \in \mathcal{E}_{y,x}^{SHO}} \omega_y^E \Delta_{y \to x}^j}{\eta_{x,own}^E} \leq \alpha \tag{5.19}$$

for all x from 1 to $\mathcal{N}$, where α denotes the certain threshold for this ratio and $\mathcal{E}_{y,x}^{SHO}$ refers to the E-DCH users in soft handover area with NodeB y as their serving NodeB and NodeB x as their non-serving NodeB. These constraints can be applied in both non-linear and linear optimization algorithms.

It can be seen in (5.19) that the SLF ω^E are still assumed to be equal to each other for all the E-DCH users within one cell, however, to achieve better performance, it would be better to distinguish users who are in the central area from those who are in the edge (i.e. allocate different ω^E according to user locations). The system performance in such scenarios would be studied in the future.

5.4 Feasible Load Regions Comparison

In this section, we demonstrate a simple example to give a brief comparison among various resource allocation schemes described above. Two E-DCH users and another two DCH users located in cell A and cell B respectively are assumed in this example, where each NodeB serve a pair of DCH and E-DCH users. The parameter values in the example scenario are summarized in the Table. 5.1 below. The parameter Δ

Table 5.1: Parameters in the example scenario

	E-DCH #1	E-DCH #2	DCH #1	DCH #2
Serving NodeB	A	B	A	B
ω			0.1	0.05
Δ	0.9	$6 \cdot 10^{-4}$	$3 \cdot 10^{-4}$	$1 \cdot 10^{-4}$

correspond to the path loss ratio between the serving and non-serving NodeBs. The first E-DCH user is close to the cell edge, which leads to a high Δ of 0.9. The second E-DCH user in cell B is close to its serving NodeB. The DCH users have moderate distances to their serving NodeBs. As for fairness criterion in the global resource

allocation schemes, we chose max-min-fairness since it is closest to the behavior of the direct assignment approach.

The resulting feasible SLF regions for the two E-DCH users are shown in Fig. 5.4, where each curve corresponds to one set of constraints. The points positioned below a particular curve compose the feasible SLF region defined by this set of constraints. The curve 'load only' refers to the load region under global optimization with only load and SLF constraints, while the curve 'row-sum' represents the linear-programming based optimization since the constraints under this scheme have close relationship with the row sum of the system load matrix $\tilde{G}$. The 'power' and 'down-grants' curves denote the transmit power and down grants constraints, respectively. The max-min-optimal points for the global RRM differ significantly from the direct approach (pointed as 'local RRM'), where the latter scheme yields a very unbalanced result between the two E-DCH users but still lies within the feasible region. The transmit power constraints in this scenario lead to a SLF configuration which favors the first E-DCH user, while for the load, down-grant and linear constraints the SLF values are balanced. The direct approach for the local RRM corresponds to the linear constrained RRM with sum-optimal utility function. The feasible region does not reach the maximum possible SLF ω_{max} due to the load from the DCH users. The optimal solution for the down-grant constraints correspond in this case to the solution with load constraints only, however this would change if the maximum allowed ratio between soft handover area and own-cell E-DCH load is set to a lower value.

The corresponding cell load η_A at NodeB A is illustrated in Fig. 5.5. The loads for the non-linear and linear cases begin to diverge on the solution point of the direct approach. The effect of the linear constraint on the load is that the target load is not reached for a large range of the feasible SLF region. Furthermore, the max-min-optimal point in this case is significantly lower than the one in the non-linear case. The direct approach naturally reaches the target load at both NodeBs, but at the expense of a very low SLF for the first E-DCH user. It should be mentioned that this scenario is quite extreme, which is the reason for the different results of the approaches. As we will see in the next section, with more users the results get more close to each other.

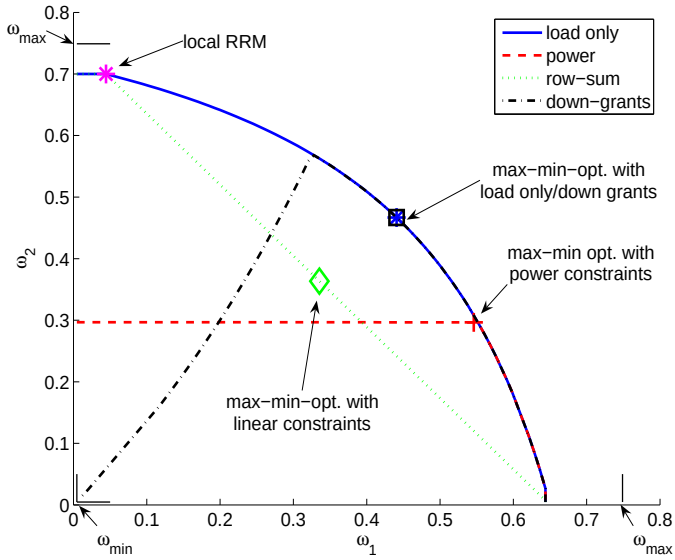


Figure 5.4: Feasible SLF regions for the two-cell scenario

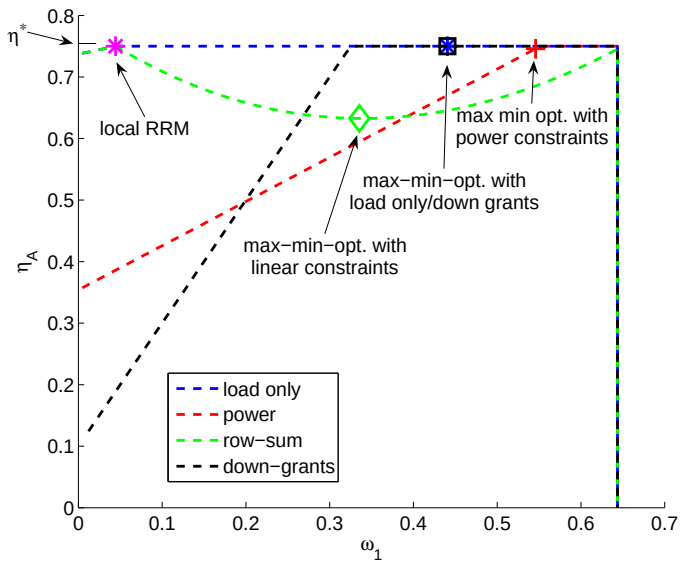


Figure 5.5: Feasible cell loads at NodeB A

5.5 Numerical Results

In this section, we present the results from Monte Carlo simulations to evaluate the system performance under different resource allocation schemes described above. The first set of simulations run in a two-tier hexagonal cell ring layout. The system bandwidth $W = 3.84$MHz, background thermal noise density $N_0 = -174$dBm, and PLE $\gamma = 3.76$. The target system load η^* is assumed to be 0.85. Only the results in the central cell are sampled. Two service classes, DCH and E-DCH, are assumed in the system, where the traffic parameters are listed in Table. 5.2.

Table 5.2: Traffic parameters for optimization-based resource allocation simulations

Service Categories	Bit Rate R	Target E_b/I_0 ε^*	Activity υ
DCH	12.2kbps	5.5dB	1
E-DCH	$\geq$5kbps	5.5dB	1

The results are sampled from Monte-Carlo simulations, where the user patterns are generated randomly according to a spatial homogeneous Poisson process, followed by applying certain resource allocation to determine the rates for each user and finally working out the desired results such as outage probability and received load from each source.

Fig. 5.6 illustrates the soft outage probabilities with different DCH and E-DCH user density. Note that this probability only concerns the soft outage which occurs as in scenario 3 described in Section 5.2. It can be seen that the more DCH users, the higher the outage probability due to less load capacity for E-DCH users. However, the more E-DCH users, the lower the outage probability. The reason is that with more E-DCH users, there would be more possibilities for the distribution of available resource, and thus averages out some outage cases. Another thing that is worth noticing is the soft outage probability could go beyond 10% in the cases with only a few E-DCH users, which implies the effort to avoid such outages is not unnecessary.

In Fig. 5.7, the total received loads at central NodeB from both DCH users and E-DCH users under various resource allocation algorithms are presented. The mean number of E-DCH users is assumed to be 6 and all the results are only sampled from the soft outage cases because in the normal cases the optimal results can be

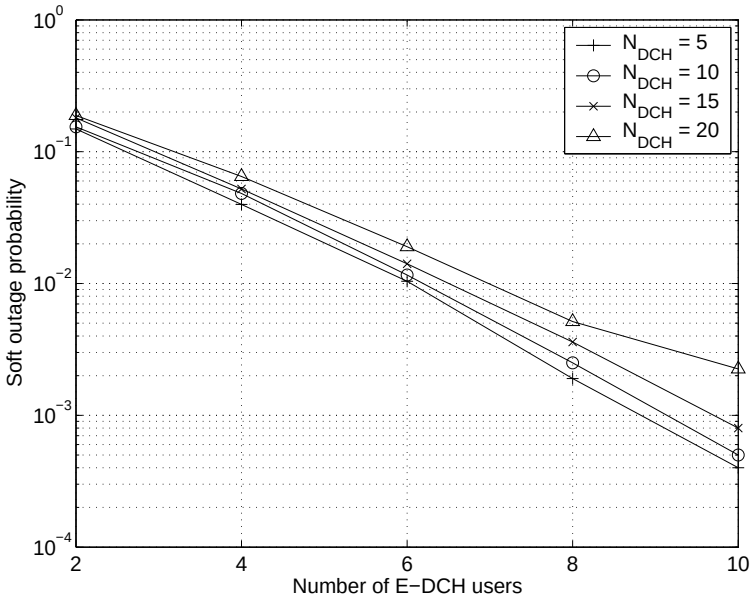


Figure 5.6: Soft outage probabilities under different user combinations

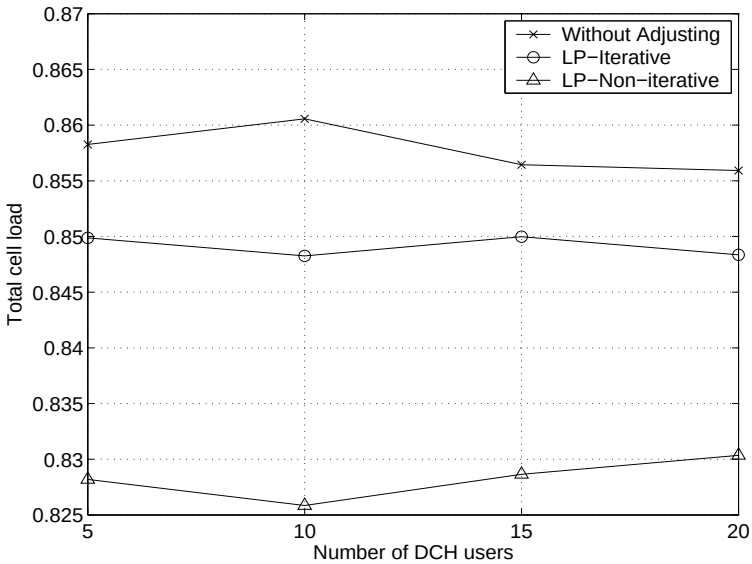


Figure 5.7: Mean total load under different RRM strategies with target load of 0.85

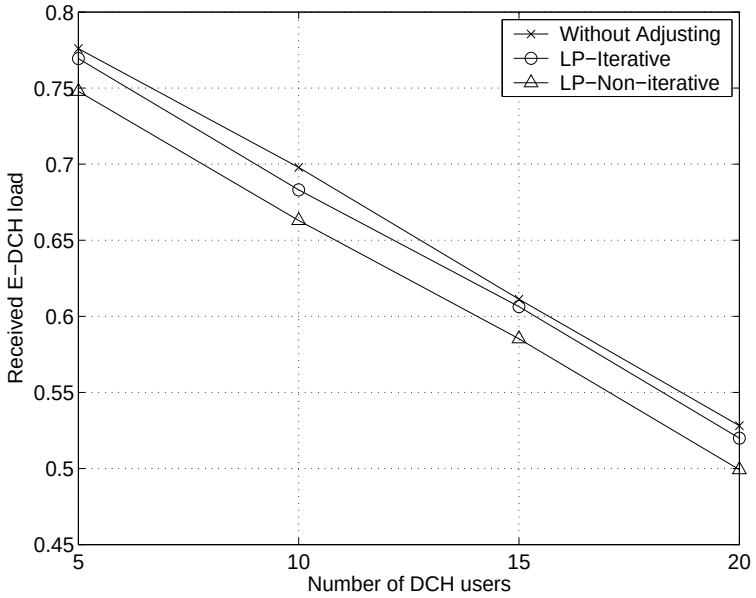


Figure 5.8: Mean E-DCH load under different RRM strategies

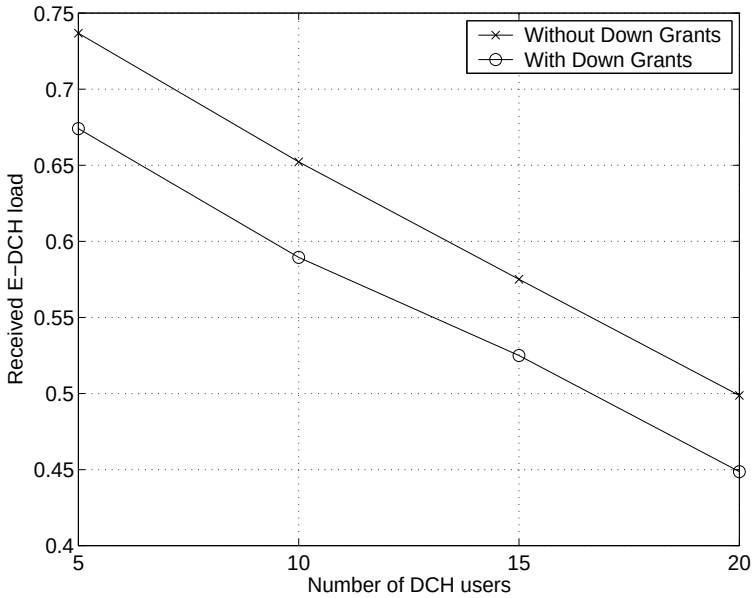


Figure 5.9: Impact of down grants on mean E-DCH load

simply obtained from solving the linear equations system, where any RRM would yield the same assignment pattern. The approach 'without adjusting' listed in the figure refers to the direct assignment. From the figure, we can see the resultant E-DCH loads under this scheme achieve highest utilization, but also exceed the target system load η^*, which implies soft outages occur. The loads under 'Linear-Programming non-iterative' scheme are lower than η^* due to its conservativeness. The results from 'Linear-Programming iterative' approach stay quite close to but below η^*, under which high resource utilization is achieved and soft outage case is avoided.

Fig. 5.8 shows the corresponding mean loads from only E-DCH users. A similar utilization order is expected and verified as the received DCH loads are independent of the rate allocation scheme, which can be regarded as constants. The received loads from 'LP iterative' are quite close to those from 'without adjusting' that are the possible maximum, which indicates the former is an efficient algorithm to find the near-optimal solution.

The impact of inclusion of 'down grants' in the linear-programming algorithm is illustrated in Fig. 5.9. Both are based on the non-iterative resource allocation, while the latter includes 'down grants' constraints in the RRM algorithm. The threshold α takes the value of 0.7. It can be seen with more constraints considered, the SLF ω^E is lower and in turn the mean E-DCH load. As stated in previous, if the E-DCH users belonging to the same cell can be allocated different ω^E according to their locations, the E-DCH load is expected to be higher.

We have also run another set of Monte Carlo simulations with different parameters in a one-tier cell ring layout to evaluate system performance under the resource allocation schemes described above. The path loss is calculated from the COST-231 small urban Hata model [1], and the target load is set to $\eta^* = 0.75$, which corresponds to a target interference of -103dBm. To see the influence of the number of E-DCH and DCH users, the total number of the users in one cell is fixed at 10, while the fraction of E-DCH user grows from 2 to 8 users. The rest are the DCH users with bit rate of 64kbps. Since it is not practical for $\alpha \to \infty$ in (5.12) in the numerical calculation,

the utility function for the global optimization is defined as $U(\omega_m) = \sum_m \omega_m^{-1}$,

$$U = \sum_{x=1}^{\mathcal{N}} \sum_{j \in \mathcal{E}_x} \left(\omega_j^E\right)^{-1} \tag{5.20}$$

which corresponds approximately to the max-min-fair resource assignment in [9].

The total loads under a few resource allocation mechanisms are shown in Fig. 5.10. With an increasing number of E-DCH users, the mean total load decreases from 0.752 to the target load of 0.75. Correspondingly, the the loads for linear-programming based RRM increase until the target load is reached. The total loads for the global RRM with different constraints stay all below the target load but are close to each other. Only in the case of two E-DCH users it can be observed that the down-grants and the power constraints lead to a lower received load than with load and SLF constraints only.

In Fig. 5.11, the other-cell load for the direct approach is significantly higher than the loads for the global approaches, and the distance between the curves is more or less constant for all considered E-DCH to DCH ratios. As expected, the linear-programming based RRM receives the highest other-cell loads for the optimization-based RRM implementations, and the down-grant constrained RRM has the lowest, although the difference to the other non-linear approaches is not very high. This leads also to the highest own-cell E-DCH load shown in Fig. 5.12, although the difference is even smaller among the different optimization approaches. The direct approach yields, corresponding to the highest other-cell load, the lowest E-DCH loads with a nearly constant difference to the global RRM strategies of 5%. Note that the own cell E-DCH load grows almost linear with the number of E-DCH users. Corresponding to the E-DCH own cell load the mean assigned SLFs are shown in Fig. 5.13. The highest SLF is 0.1 for the global optimization RRM, which corresponds to bit rate of approximately 220kbps. With two E-DCH users, the SLF for the direct approach is around 0.07, which corresponds to bit rate of 150kbps.

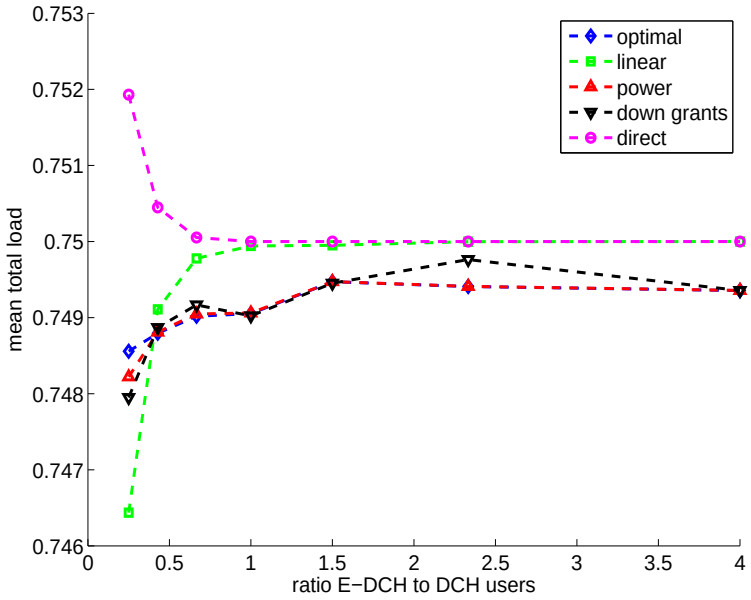


Figure 5.10: Mean total load η

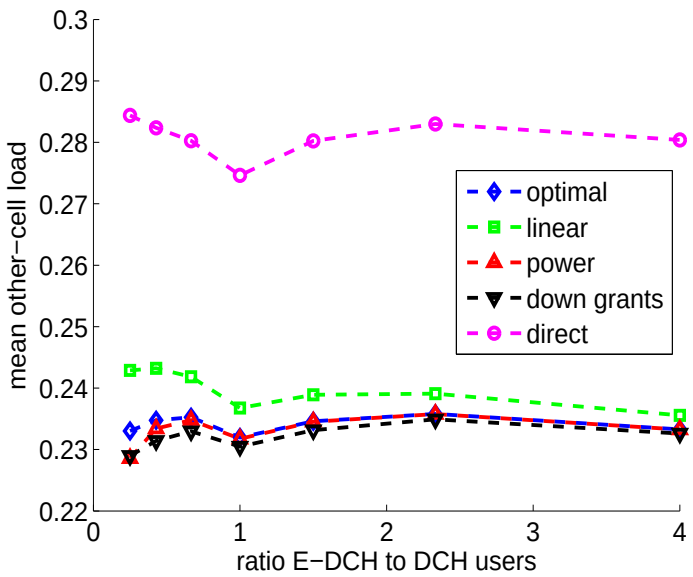


Figure 5.11: Mean other-cell load η^{oc}

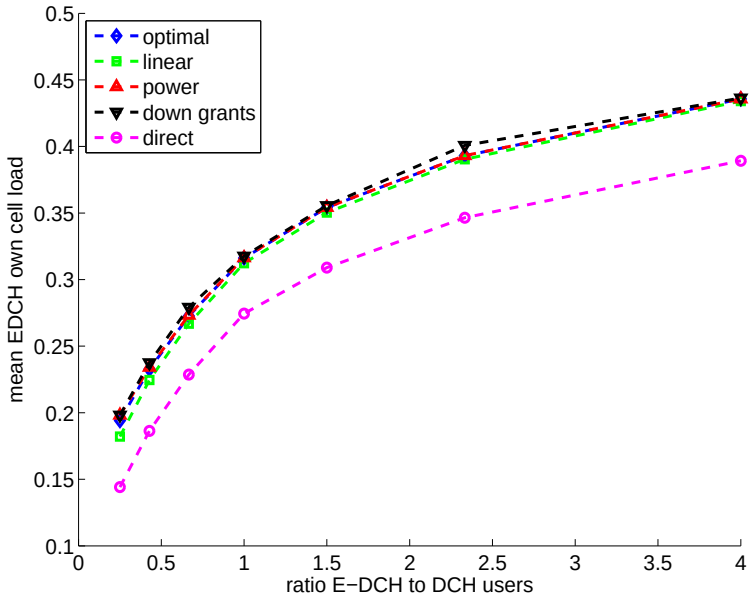


Figure 5.12: Mean E-DCH own cell load

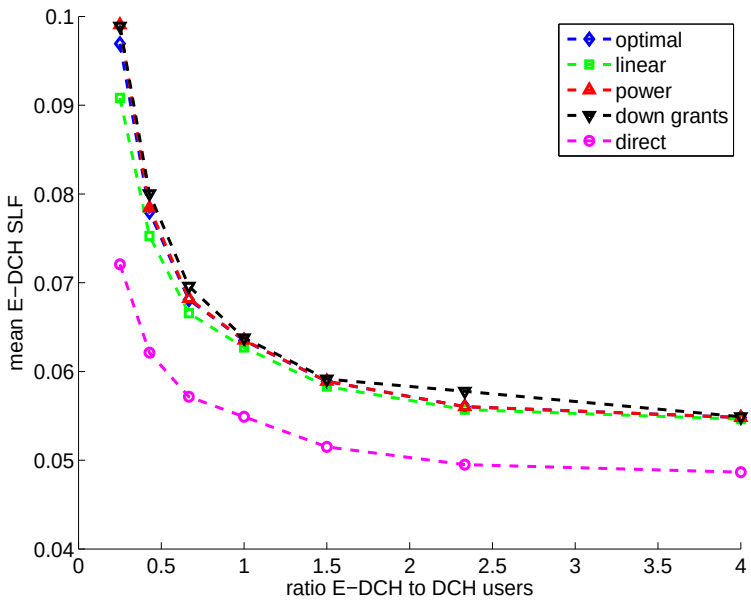


Figure 5.13: Mean E-DCH SLFs

5.6 Discussion

In this chapter, we have presented a framework to assign resources to E-DCH users in UMTS enhanced uplink under optimization based resource allocation schemes, which include direct, global optimization, non-iterative and iterative linear programming approaches. The suggested resource allocation strategies not only maximize the utilization through some specific utility functions, but also avoid the occurrence of unnecessary outage cases where the elastic traffic degrade the performance of streaming traffic by proper constraints. The simulation results illustrate the performance under various rate allocation mechanisms. In particular linear-programming based resource allocation achieves similar performance as the global non-linear optimization, while the computation complexity is much less than the latter. Therefore, they are quite efficient to study the UMTS system behaviors.

As mentioned in the previous section, our future work includes performance evaluation for the RRM which can distinguish users by allocating various rates according to user locations. And the final goal is to build analytic models to approximately characterize the performance under these RRM schemes in the UMTS enhanced uplink, such that the time and complexity for the decision of resource allocation can be further reduced.

Chapter 6

Conclusion

This monograph develops the analytical models of UMTS networks in various scenarios (e.g. different radio resource management schemes, source traffic models, propagation models, etc.), justifies the accuracy of proposed models through numerical results from Monte Carlo and discrete event simulations, and evaluates the corresponding system performance with these models.

We start with investigation on the system behaviors in the operation of UMTS uplink, and development of the analytic techniques to model interference and system load as fully-characterized random variables, which can be directly applicable to the performance evaluation of such networks in the existing radio network planning tools. The models are constructed for the systems with mere circuit-switched fixed-rate QoS traffic first, and then extended to those with additional packet-switched best-effort traffic with varying data rates in the enhanced uplink. In order to achieve maximum capacity, the target SIR oriented power control mechanism is employed, which leads to more sophisticated system manner, denoted as 'feedback behavior', in the multi-cell environments as the interference levels at different cells depend on each other. To capture such behavior when deriving the full distribution functions of the interference variables, we first follow iterative method to solve constructed fixed-point equations, and then introduce a log-normal approximation technique to reduce the computational complexity significantly. During the derivation, we consider various propagation models, traffic models, resource allocation schemes for many possible scenarios, each of which may lead to different analytical models. However, it has

been verified in all these scenarios that the other-cell interference complies with the log-normal distribution, which makes the approximation techniques always applicable. All the suggested models are validated with either Monte-Carlo simulations or discrete event simulations, where excellent match results are always achieved.

The analytical interference model constructed for the HSUPA-enabled UMTS networks can be employed by the network operators to understand the newly evolved system behavior, as well as applied into outage probability analysis which is crucial for capacity evaluation. This model can also be embedded into the corresponding network planning software, such that the efficiency can be largely increased compared with the conventional trial-and-error planning process.

Another innovation in this work is the optimization-based framework for the performance evaluation under centralized resource allocation strategies in the UMTS enhanced uplink. The resources available for the enhanced uplink users depend on several factors like the spatial configuration of the mobiles in the cells, the number of QoS users and the employed resource management strategy. This framework is based on the centralized resource allocation mechanism, which aims to search the best allocation pattern with highest utilization under the assumption that the instant load information from all the NodeBs is known at the rate controller. The basic principles of centralized resource management are described and the corresponding interference and load models are formulated for the understanding of the system behavior. A few rate assignment strategies for the best effort packet data traffic in the enhanced uplink are stated afterwards, which consists of direct, linear-programming based and global optimization based rate allocation schemes. The suggested framework also considers the 'down grants' for restraining the other-cell interference. Then with the comparison of results from feasible load regions, as well as of the performance metric such as total received load, SLF of E-DCH users, etc., the suggested optimization-based resource management schemes are analyzed and evaluated. The Monte Carlo simulation results can be used for the UMTS network operators to gain better estimation of how the UMTS enhanced uplink performs in real scenarios.

6.1 Outlook

In each summary section of previous chapters, we have stated a few points for possible future research which mainly address the topic in that chapter. In this section, we list a few potential research directions from an overall point of view. There is one assumption always made during the analysis in this monograph, which is the hexagonal cell layout. It is quite a general assumption in the mobile network research, however, this restricts the input network configuration to some extent. Actually at this stage, it is more like a dimensioning tool which is able to estimate the quantity of NodeBs, but cannot handle the questions of how to place these NodeBs. As the final goal of the long-term research is to justify the feasibility of any input network configuration with derived analytical models, the proposed models should be verified under random NodeB layouts. Another assumption of uniform spatial traffic distribution applies in most cases, however there are also scenarios where 'hot-spots' with high traffic density exist on the traffic map, thus heterogenous traffic distribution may be another improvement for the current models. Generally speaking, the more assumptions released, the more general the proposed models are, and in turn can be applied into more scenarios.

The work in this monograph only covers the UMTS network in the uplink direction, this is because the uplink is always considered as the performance bottleneck, especially in the conventional voice-only mobile networks. With the multimedia streaming such as VoD and file download services become more and more popular in 3G networks, the performance in the downlink direction would attract much more attention, especially after HSDPA technology being introduced. Constructing the analytical models for the system downlink parameters with integrated QoS and best-effort traffic may be another issue in the future research.

The 'down grants' is a useful approach introduced in the enhanced uplink to suppress the other-cell interference when needed. In this work, only the simulation results are presented for the system behavior under such commands. Inclusion of such factor into more comprehensive analytical models is also an interesting issue for future research. Moreover, various resource allocation and admission control schemes could be studied and reflected by future analytical models.

Appendix A

List of Acronyms

2G	2nd Generation
3G	3rd Generation
3GPP	3rd Generation Partnership Project
3GPP2	3rd Generation Partnership Project 2
AMPS	Advanced Mobile Phone Service
ATM	Asynchronous Transfer Mode
ARQ	Automatic Repeat reQuest
BER	Bit Error Rate
CAC	Call Admission Control
CDF	Cumulative Distribution Function
CDMA	Code Division Multiple Access
CN	Core Network
CP	Complete Partitioning
CS	Complete Sharing

DCA Dynamic Channel Assignment

DCH Dedicated Channel

E-DCH Enhanced Dedicated Channel

EDGE Enhanced Data Rates for GSM Evolution

EIR Equipment Identity Register

ETACS European Total Access Communication System

FCA Fixed Channel Assignment

FDD Frequency Division Duplex

FDMA Frequency Division Multiple Access

FHMA Frequency Hopped Multiple Access

FM Frequency Modulation

FTP File Transfer Protocol

GERAN GSM/EDGE Radio Access Network

GGSN Gateway GPRS Support Node

GMSC Gateway Mobile Switching Centre

GMSK Gaussian Minimum Shift Keying

GoS Grade of Service

GPRS General Packet Radio Service

GSM Global System for Mobile Communication

HLR Home Location Register

HSDPA High Speed Downlink Packet Access

HSUPA	High Speed Uplink Packet Access
IMS	IP Multimedia Subsystem
IMT-2000	International Mobile Telephone 2000
IS-95	Interim Standard 95
ISDN	Integrated Services Digital Network
ITU	International Telecommunications Union
MAC	Medium Access Control
MGW	Media Gateway
MSC	Mobile Switching Centre
NMT	Nordic Mobile Telephone
PLE	Path Loss Exponent
PSTN	Public Switched Telephone Network
QoS	Quality of Service
RNC	Radio Network Controller
RRM	Radio Resource Management
SGSN	Serving GPRS Support Node
SIR	Signal-to-Interference-Ratio
SLF	Service Load Factor
TD-CDMA	Time Division CDMA
TDMA	Time Division Multiple Access
TTI	Transmission Time Interval

UE User Equipment

UMTS Universal Mobile Telecommunication System

UTRAN UMTS Terrestrial Radio Access Network

VLR Visitor Location Register

VP Virtual Partitioning

VoD Video on Demand

VoIP Voice over IP

WCDMA Wideband Code Division Multiple Access

Bibliography

[1] COST Action 231. Digital mobile radio towards future generation systems, final report. Technical Report EUR 18957, European Communities, 1999.

[2] 3GPP TS 23.107. QoS concept and architecture.

[3] 3GPP TS 25.308. High speed downlink packet access (HSDPA); overall description.

[4] 3GPP TS 25.309. FDD enhanced uplink; overall description.

[5] 3GPP TS 25.331. Radio resource control.

[6] Robert Akl, Mort Naraghi-Pour, and Manju Hegde. Throughput optimization in multi-cell CDMA networks. In *Proc. IEEE Wireless Communications and Network Conference*, volume 3, pages 1292–1297, New Orleans, USA, Mar. 2005.

[7] Eitan Altman. Capacity of multi-service cellular networks with transmission-rate control: a queueing analysis. In *Proc. ACM MobiCom*, pages 205–214, Atlanta, USA, Sep. 2002.

[8] S. Anand and A. Chockalingam. Performance analysis of voice/data cellular CDMA with SIR based admission control. *IEEE Journal on Selected Areas in Communications*, 21(10):1674–1684, 2003.

[9] Holger Boche and Martin Schubert. Resource allocation in multi-antenna systems - achieving max-min fairness by optimizing a sum of inverse SIR. *IEEE Transactions on Signal Processing*, 54(6), 2006.

[10] Holger Boche and Slawomir Stanczak. Optimal QoS tradeoff and power control in CDMA systems. In *Proc. of IEEE INFOCOM*, 2004.

[11] Paul T. Brady. A statistical analysis of on-off patterns in 16 conversations. *Bell System Technical Journal*, 47:73–91, 1968.

[12] James A. Bucklew. *Large deviation techniques in decision, simulation, and estimation.* John Wiley & Sons, New York, 1990.

[13] Antonio Capone and Simone Redana. Call admission control techniques for UMTS. In *Proc. IEEE 54th Vehicular Technology Conference*, volume 2, pages 925–929, Oct. 2001.

[14] Antonio Capone and Simone Redana. Performance of call admission control techniques for TD-CDMA systems. In *Proc. IEEE 58th Vehicular Technology Conference*, volume 3, pages 1673–1677, Oct. 2003.

[15] Ruay-Shiung Chang and Chih-Tsung Hsieh. Uplink channel assignment with balanced load for code division multiple access cellular systems. *Wireless Communications and Mobile Computing*, 3:63–72, 2003.

[16] Noel A. Cressie. *Statistics for spatial data.* John Wiley & Sons, New York, 1991.

[17] Mostafa M. El-Said, Anup Kumar, E. Elliott Bennett, Adam N. Foley, and Travis J. Hillenbrand. Call admission control using dynamic cell virtual traffic in CDMA systems. In *Proc. IEEE 58th Vehicular Technology Conference*, volume 5, pages 3508–3511, Oct. 2003.

[18] Anwar I. Elwalid and Debasis Mitra. Effective bandwidth of general Markovian traffic sources and admission control of high speede networks. *IEEE/ACM Transactions on Networking*, 1:329–343, 1993.

[19] Jamie S. Evans and David Everitt. Effective bandwidth-based admission control for multiservice CDMA cellular networks. *IEEE Transactions on Vehicular Technology*, 48:36–46, 1999.

[20] Jamie S. Evans and David Everitt. On the teletraffic capacity of CDMA cellular networks. *IEEE Transactions on Vehicular Technology*, 48(1):153–165, 1999.

[21] David Everitt. Analytic traffic models of CDMA cellular networks. In *Proc. 14th International Teletraffic Congress*, pages 349–356, Jun. 1994.

[22] David Everitt. Traffic engineering of the radio interface for cellular mobile networks. *Proc. IEEE*, 82(9):1371–1382, 1994.

[23] David Everitt and N. W. Macfadyen. Analysis of multicellular mobile radiotelephone systems with loss. *British Telecom Technology Journal*, 1:37–45, 1983.

[24] David Everitt and David Manfield. Performance analysis of cellular mobile communication systems with dynamic channel assignment. *IEEE Journal on Selected Areas in Communications*, 7:1172–1180, 1989.

[25] Gábor Fodor and Miklós Telek. Performance analysis of the uplink of a CDMA cell supporting elastic services. In *Proc. NETWORKING*, pages 205–216, Waterloo, Canada, May 2005.

[26] Klein S. Gilhousen, Irwin M. Jacobs, Roberto Padovani, Andrew J. Viterbi, Lindsay A. Weaver, and Charles E. Wheatley. On the capacity of a cellular CDMA system. *IEEE Transactions on Vehicular Technology*, 40(2):303–312, 1991.

[27] Prashanth Hande, Sundeep Rangan, and Mung Chiang. Distributed uplink power control for optimal SIR assignment in cellular data networks. In *Proc. IEEE INFOCOM*, Barcelona, Spain, Apr. 2006.

[28] Masaharu Hata. Empirical formula for propagation loss in land mobile radio services. *IEEE Transactions on Vehicular Technology*, 29(3):317–325, 1980.

[29] Nidhi Hegde and Eitan Altman. Capacity of multiservice WCDMA networks with variable GoS. *Wireless Networks*, 12(2):241–253, 2006.

[30] Harri Holma and Antti Toskala. *WCDMA for UMTS - radio access for third generation mobile communications.* John Wiley, Chichester, 2000.

[31] Tara Javidi. Decentralized rate assignments in a multi-sector CDMA network. In *Proc. of IEEE GLOBECOM*, volume 1, pages 65–69, San Francisco, USA, Dec. 2003.

[32] Gulshan Karmani and Kumar N. Sivarajan. Capacity evaluation for CDMA cellular systems. In *Proc. IEEE INFOCOM*, volume 1, pages 601–610, 2001.

[33] F. P. Kelly, A. K. Maulloo, and D. K. H. Tan. Rate control in communication networks: shadow prices, proportional fairness and stability. *Journal of the Operational Research Society*, 49(3):237–252, 1998.

[34] Frank P. Kelly. Loss networks. *Annals of Applied Probability*, 1(3):319–378, 1991.

[35] George Kesidis, Jean Walrand, and Cheng-Shang Chang. Effective bandwidths for multiclass Markov fluids and other ATM sources. *IEEE/ACM Transactions on Networking*, 1:424–428, 1993.

[36] Kyoung Il Kim. CDMA cellular engineering issues. *IEEE Transactions on Vehicular Technology*, 42(3):345–350, 1993.

[37] Leonard Kleinrock. *Queueing systems volume I: theory.* John Wiley, New York, 1975.

[38] William C. Y. Lee. *Mobile cellular telecommunications systems.* McGraw-Hill, Singapore, 1990.

[39] William C. Y. Lee. Overview of cellular CDMA. *IEEE Transactions on Vehicular Technology*, 40:291–302, 1991.

[40] Bin Liu and Attahiru Sule Alfa. A queueing model with time-varying QoS and call dropping for evaluating the performance of CDMA cellular systems. *Wireless Communications and Mobile Computing*, 4(4):439–447, 2004.

[41] Te Kai Liu and John A. Silvester. Joint admission/congestion control for wireless CDMA systems supporting integrated services. *IEEE Journal on Selected Areas in Communications*, 16(6):845–857, 1998.

[42] Tuo Liu and David Everitt. Analytical approximation of other-cell interference in the uplink of CDMA cellular systems. In *Proc. IEEE 63rd Vehicular Technology Conference*, volume 2, pages 693–697, Melbourne, Australia, May 2006.

[43] Zhao Liu and Magda El Zarki. SIR-based call admission control for DS-CDMA cellular systems. *IEEE Journal on Selected Areas in Communications*, 12:638–644, 1994.

[44] Erik Geijer Lundin, Fredrik Gunnarsson, and Fredrik Gustafsson. Robust uplink resource allocation in CDMA cellular radio systems. In *Proc. of Joint IEEE Conference on Decision and Control and European Control Conference*, pages 1848–1853, Seville, Spain, Dec. 2005.

[45] Yue Ma, James J. Han, and Kishor S. Trivedi. Call admission control for reducing dropped calls in code division multiple access (CDMA) cellular systems. In *Proc. IEEE INFOCOM*, volume 3, pages 1481–1490, Mar. 2000.

[46] Andreas Mäder and Dirk Staehle. Uplink blocking probabilities in heterogeneous WCDMA networks considering other-cell interference. In *Proc. Southern African Telecommunication Networks & Applications Conference*, South Western Cape, South Africa, Sep. 2004.

[47] Andreas Mäder and Dirk Staehle. An analytical model for best-effort traffic over the UMTS enhanced uplink. In *Proc. IEEE Vehicular Technology Conference*, Montréal, Canada, Sep. 2006.

[48] J. W. Mark, T. C. Wong, M. Jin, B. Bensaou, and K. C. Chua. Resource allocation for multiclass services in cellular systems. In *Proc. IEEE International Conference on Communication Systems*, Nov. 2000.

[49] Debasis Mitra and John A. Morrison. Erlang capacity and uniform approximations for shared unbuffered resources. *IEEE/ACM Transactions on Networking*, 2(6):558–570, 1994.

[50] Jeonghoon Mo and Jean Walrand. Fair end-to-end window-based congestion control. *IEEE/ACM Transactions on Networking*, 8(5):556–567, 2000.

[51] Michel Mouly and Marie-Bernadette Pautet. *The GSM system for mobile communications*. Cell & Sys, Palaiseau, France, 1992.

[52] Stefan Parkvall, Janne Peisa, Johan Torsner, Mats Sågfors, and Peter Malm. WCDMA enhanced uplink - principles and basic operation. In *Proc. IEEE Vehicular Technology Conference*, volume 3, pages 1411–1415, Stockholm, Sweden, May 2005.

[53] Krister Raith and Jan Uddenfeldt. Capacity of digital cellular TDMA systems. *IEEE Transactions on Vehicular Technology*, 40(2):323–331, 1991.

[54] Theodore S. Rappaport. *Wireless communications: principles and practice.* Prentice Hall, New Jersey, 2002.

[55] Ashwin Sampath, Narayan B.Mandayam, and Jack M. Holtzman. Erlang capacity of a power controlled integrated voice and data CDMA system. In *Proc. IEEE 47th Vehicular Technology Conference*, volume 3, pages 1557–1561, Phoenix, USA, May 1997.

[56] Ashwin Sampath, P. Sarath Kumar, and Jack M. Holtzman. Power control and resource management for a multimedia CDMA wireless system. In *Proc. IEEE PIMRC*, volume 1, pages 21–25, Sep. 1995.

[57] Sungmoon M. Shin, Cheol-Hye Cho, and Dan Keun Sung. Interference-based channel assignment for DS-CDMA cellular systems. *IEEE Transactions on Vehicular Technology*, 48(1):233–239, 1999.

[58] Dirk Staehle, Kenji Leibnitz, Klaus Heck, Bernd Schröder, Phuoc Tran-Gia, and Albert Weller. An approximation of othercell interference distributions for UMTS systems using fixed-point equations. Research Report 292, Institute of Computer Science, University of Wurzburg, Germany, Jan. 2002.

[59] Dirk Staehle, Kenji Leibnitz, Klaus Heck, Bernd Schröder, Albert Weller, and Phuoc Tran-Gia. Approximating the othercell interference distribution in inhomogeneous UMTS networks. In *Proc. IEEE 55th Vehicular Technology Conference*, volume 4, pages 1640–1644, Birmingham, USA, May 2002.

[60] Dirk Staehle and Andreas Mäder. An analytic approximation of the uplink capacity in a UMTS network with heterogeneous traffic. In *Proc. 18th International Teletraffic Congress*, Berlin, Germany, Sep. 2003.

[61] Corinne Touati, Eitan Altman, and Jérôme Galtier. Fair power and transmission rate control in wireless networks. In *Proc. of IEEE GLOBECOM*, volume 2, pages 1229–1233, Taipei, Nov. 2002.

[62] Venugopal Veeravalli and Andrew Sendonaris. The coverage-capacity trade-off in cellular CDMA systems. *IEEE Transactions on Vehicular Technology*, 48(5):1443–1450, 1999.

[63] Venugopal V. Veeravalli, Andrew Sendonaris, and Nikhil Jain. CDMA coverage, capacity and pole capacity. In *Proc. IEEE 47th Vehicular Technology Conference*, volume 3, pages 1450–1454, Phoenix, USA, May 1997.

[64] Andrew J. Viterbi. *CDMA: principles of spread spectrum communication.* Addison-Wesley, 1997.

[65] Andrew J. Viterbi, Audrey M. Viterbi, and Ephraim Zehavi. Performance of power controlled wideband terrestrial digital communications. *IEEE Transactions on Communications*, 1993.

[66] Andrew J. Viterbi, Audrey M. Viterbi, and Ephraim Zehavi. Other-cell interference in cellular power-controlled CDMA. *IEEE Transactions on Communications*, 42(2/3/4):1501–1504, 1994.

[67] Audrey M. Viterbi and Andrew J. Viterbi. Erlang capacity of a power controlled CDMA system. *IEEE Journal on Selected Areas in Communications*, 11(6):892–900, 1993.

[68] T. C. Wong, J. W. Mark, and K. C. Chua. Resource allocation in mobile cellular networks with QoS constraints. In *Proc. IEEE Wireless Communications and Networking Conference*, volume 2, pages 717–722, Mar. 2002.

[69] T. C. Wong, J. W. Mark, K. C. Chua, and B. Kannan. Performance analysis of variable bit rate multiclass services in the uplink of wideband CDMA. In *Proc. IEEE International Conference on Communications*, volume 1, pages 363–367, May 2003.

[70] T. C. Wong, J. W. Mark, K. C. Chua, J. Yao, and Y. H. Chew. Performance analysis of multiclass services in the uplink of wideband CDMA. In *Proc. 8th International Conference on Communication Systems*, volume 2, pages 692–696, Nov. 2002.

[71] Jianxin Yao, Jon W. Mark, Tung Chong Wong, Yong Huat Chew, Kin Mun Lye, and Kee-Chaing Chua. Virtual partitioning resource allocation for multiclass traffic in cellular systems with QoS constraints. *IEEE Transactions on Vehicular Technology*, 53:847–864, 2004.

[72] W. Rae Young. Advanced mpbile phone service: introduction, background, and objectives. *Bell Systems Technical Journal*, 58(1):1–14, Jan. 1979.

[73] Jens Zander. Performance of optimum transmitter power control in cellular radio systems. *IEEE Transactions on Vehicular Technology*, 41(1):57–62, 1992.

[74] Dongmei Zhao, Xuemin Shen, and Jon W. Mark. Radio resource management for cellular CDMA systems supporting heterogeneous services. *IEEE Transactions on Mobile Computing*, 2(2):147–160, 2003.