RESOURCE ALLOCATION IN CELLULAR CDMA SYSTEMS WITH CROSS-LAYER OPTIMIZATION

YAO JIANXIN

NATIONAL UNIVERSITY OF SINGAPORE

2005

RESOURCE ALLOCATION IN CELLULAR CDMA SYSTEMS WITH CROSS-LAYER OPTIMIZATION

YAO JIANXIN

(B.Eng., M.Eng.)

A THESIS SUBMITTED

FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN

ELECTRICAL ENGINEERING

DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

NATIONAL UNIVERSITY OF SINGAPORE

2005

Acknowledgements

I would like to extend my heartfelt appreciation and deepest gratitude to the followings:

Dr. Yong Huat CHEW and Dr. Tung Chong WONG for their guidance, help and patience in the accomplishment of this thesis. Their deep insights and advices beyond academic and research were and will always be well appreciated. It has been a great honor to have them as my supervisors, and to work together with them on research and papers;

Prof. Jon W. MARK and Dr. Chin Choy CHAI for their precious instructions and kind assistance in the research work;

NUS and I^{2}R for giving me an opportunity and providing a wonderful technical environment for the completion of my research;

Agency for Science, Technology and Research (A*STAR) and STMicroelectronics for supporting the research work as part of the Singapore-Ontario Joint Research Programme;

My friends and my laboratory mates, Hu Xiaoyu, Zhou Kainan, Mo Ronghong, Wang Jia, Zhang Songhua, Xiao Lei, Nie Chun, etc., for their support and kind assistance in my study and life and making my stay in the laboratory a very enjoyable and memorable one.

Last but not least, I am always grateful to my dearest family and my girlfriend Wang Shugui, for their substantial support with their endless love, caring, understanding and encouragement in my life.

Contents

Summary

This thesis studies resource allocation in cellular CDMA systems with cross-layer optimization across the physical layer, the link layer and the network layer.

The analyses for the network layer and the link layer are firstly presented. The physical layer features are included into the link layer analysis.

In the network layer, two connection admission control (CAC) schemes, complete sharing (CS) and virtual partitioning (VP), are investigated. The analytical models are based on a K-dimensional Markov chain and solved using a number of preemption rules. The formulae of the grade of service (GoS) metrics at the connection-level and the quality of service (QoS) metrics at the packet-level for different CAC schemes are derived. The GoS metrics include the new-call-blocking probability, the handoff-call-dropping probability and the system utilization. The QoS metric includes the packet-loss probability. A method to maximize system utilization through joint optimization of connection/packet levels parameters is proposed. Numerical results indicate that significant gain in system utilization is achieved using the joint optimization compared to the case without the joint optimization.

In the link layer, the interference models are carefully built with soft handoff, diversity and statistical multiplexing in both the reverse and forward links. The analytical models are based on the largest received power base station (BS) selection criterion. In the forward link, an approximation selection method combining the advantages of previously used approximations and adapting to the QoS specification for different services is proposed. Furthermore, different power control schemes for mobile users in soft handoff are investigated and compared. The signal-to-interference

ratios (SIR) and the outage probabilities for multi-class services at the BSs (for the reverse link) and in the mobile users (for the forward link) are formulated. By constraining the outage probability to be within its requirement value, admission regions are obtained.

The motivation to employ cross-layer optimization in wireless networks comes from the recognition and understanding of the time-varying parameters, such as channel gains, in the wireless links. The time-varying characteristics in the wireless link cause statistical behavior among layers and consequently lead to the need of statistical QoS guarantees in the higher layers. A function block, the cross-layer decision-maker (DM), through which the cross-layer optimization will be applied without disturbing the integrity of the conventional protocol structure is proposed. The parameters intertwined among layers, including the QoS and GoS metrics, are considered together to achieve cross-layer optimization in the DM. Besides the intertwining parameters, there are connection parameters between layers and the system configuration parameters which are the outputs of the optimization problem for each layer. The general optimization problem is constructed to maximize the system utilization subject to the QoS requirements.

Based on the general cross-layer model, the capacity unbalance problem is solved with an adaptive soft handoff probability (SHP) scheme. The physical parameter, SHP, is controlled adaptively along with the changing traffic volumes in the reverse or forward link. The influences of the QoS requirements from diverse services are also presented. The QoS requirements from different services affect the efficiency of resource allocation in the lower layers, such as the adjustment of the SHP in the physical layer and the determination of the admission region in the link layer.

List of Figures

List of Tables

Glossary of Symbols

α	activity factor
A	area of a cell
$\mathbf{A}$	admission region for the link layer considering both the reverse and forward links
$\mathbf{A}_F$	admission region in the forward link
$\mathbf{A}_R$	admission region in the reverse link
b_{nk}	new class k call blocking probability considering only class k traffic
b_{hk}	handoff class k call dropping probability considering only class k traffic
B+1	total number of BS in the system
B_{nk}	new-call-blocking probability for class k traffic
B_{hk}	handoff-call-blocking probability for class k traffic
B_n	total new-call-blocking probability
B_h	total handoff-call-dropping probability
C	physical capacity
C_i , $i = 1, 2$	nominal capacity for each group, $C_1 + C_2 = N$
C_G	number of guard channels in a cell
δ_k	outage probability requirement value of class k traffic
Δ_{dB}	hysteresis margin
ξ	Gaussian random variables with zero mean and σ_ξ^2 variance

η	power spectral density of background noise
$d_{b,i}$	distance between a mobile *i* and the reference BS *b*
$(E_b/I_0)_{k,i}$	SIR for user *i* in class *k*
f	orthogonality factor
$\psi_{k,j}^{(h)} \in \{0,1\}$	connection activity random variable of HBR spreading code for the *j*th mobile user of class *k* traffic
$\psi_{k,j}^{(l)}$	connection activity random variable of the number of active LBR spreading codes for the *j*th mobile user of class *k* traffic
κ	slack factor
K	total number of traffic classes
G_k	spreading gain of class *k* traffic
$\gamma_{F,k}^{*}$	SIR requirement of class *k* traffic in the forward link
$\gamma_{R,k}^{*}$	SIR requirement of class *k* traffic in the reverse link
I_k	inter-cell interference from class *k* traffic
L_k	packet-loss probability for class *k* traffic
$L_{b,i}$	channel gain of a mobile *i* at a distance $d_{b,i}$ from the reference BS *b*
M	maximum number of active LBR spreading codes
μ_{ck}^{-1}	mean class *k* call holding time
μ_{hk}^{-1}	mean class *k* call dwell time
μ_{OFF}	transition rate from the *OFF* state to the *ON* state
n_k	instantaneous channel occupancy for class *k* traffic
$\vec{n}$	state of the system with the number of class *k* users, n_k, in each of

	the K classes
N	nominal capacity
N_{uk}	system utilization for class k traffic
N_u	total system utilization
λ_{nk}	arrival rate of new class k calls
λ_{hk}	arrival rate of class k handoff calls
λ_{ON}	transition rate from the *ON* state to the *OFF* state
$p_{c,j,b}$	transmission power to mobile user j of class c traffic from BS b
$P(\vec{n})$	steady state probability
$P_{\text{outage},k,i}(\vec{n})$	outage probability for user i of class k traffic
$P^{R}_{\text{outage},k,i}(\vec{n})$	outage probability for user i of class k traffic in the reverse link
$P^{F}_{\text{outage},k,i}(\vec{n})$	outage probability for user i of class k traffic in the forward link
P_{nj}	preemption probability for class j call when a new call arrives
P_{hj}	preemption probability for class j call when a handoff call arrives
$\text{Penalty}_{\text{call blocking}}$	penalty of call blocking
r	soft handoff probability
$r_k^{(h)}$	number of basic channels for users using HBR spreading codes
$r_k^{(l)}$	number of basic channels for users using LBR spreading codes
r_k	number of basic channels required by each class k call
$\vec{r}$	number of basic channels, r_k, required for each class k call with K classes
R_k	transmission rate of class k user

revenue_k	revenue for class k traffic
ρ_k	average number of users of class k per unit area
s	cell length of a square cell
S	state space of the system
S_k	received power of class k mobile user
σ_ξ^2	variance of the Gaussian random variable, ξ
θ_{nk}	probability that the next arrival is new call from class k
θ_{hk}	probability that the next arrival is handoff call from class k
θ_k	probability that the next arrival is from class k
τ	path loss exponent
v_k	speed for a class k mobile
W	total spread spectrum bandwidth

Abbreviations

1G	First Generation
2G	Second Generation
3G	Third Generation
3GPP	Third Generation Partnership Project
ATM	Asynchronous Transfer Mode
BER	Bit Error Ratio
BS	Base Station
CAC	Connection Admission Control
CDMA	Code Division Multiple Access
CP	Complete Partitioning
CPICH	Common Pilot Channel
CS	Complete Sharing
DM	Decision-Maker
DS-CDMA	Direct-Sequence Code Division Multiple Access
FDD	Frequency Division Duplex
FDMA	Frequency Division Multiple Access
FEC	Forward Error Correction
FTP	File Transfer Protocol
GoS	Grade of Service
HBR	High-Bit-Rate
ISO	International Standards Organization
ISR	Interference-to-Signal Ratio

LBR	Low-Bit-Rate
LU	Lower triangular/Upper triangular
MAC	Medium Access Control
MAI	Multi-Access Interference
MRC	Maximal Ratio Combining
NRT	Non-Real-Time
OSI	Open Systems Interconnection
p.d.f.	probability density function
PN	Pseudorandom Noise
QoS	Quality of Service
RT	Real-Time
RV	Random Variable
SIR	Signal-to-Interference Ratio
SHP	Soft Handoff Probability
SSDT	Site Selection Diversity Transmission
TCP	Transmission Control Protocol
TDD	Time Division Duplex
TDMA	Time Division Multiple Access
UMTS	Universal Mobile Telecommunications Services
VP	Virtual Partitioning
WCDMA	Wideband Code Division Multiple Access

Chapter 1

Introduction

During the last decade of the twentieth century, one’s life style has been changed by the prevailing personal communication systems including Internet (wireline) and mobile phones (wireless). The purpose of wireless cellular communications is to deliver information to people anytime and anywhere so that the interpersonal relationship is enhanced. Rapid growth of mobile phone users and the demand for broadband data service spur the successive development of wireless cellular communication techniques evolving from the first generation (1G), through the second generation (2G), to the newly deployed third generation (3G) cellular systems.

Cellular communication network provides flexible information transport platform for mobile users, so that they can roam without suffering intolerable performance degradation [1]. However, there are three major problems in wireless communications for the support of information transport between mobile users. The main problems come from (*a*) the hostile wireless propagation medium which leads to a time-varying channel condition, (*b*) the user mobility which leads to handoff and location management during communications, and (*c*) the scarce radio resource which leads to frequency reuse and hence resulting in inter-cell or intra-cell interference from other co-transmitters. These three problems cause the providers of wireless communication networks to face more challenges in providing reliable services than those using wireline communication networks.

The providers of the cellular communication networks must overcome these difficulties and provide the required service quality and system performance for mobile users. The system performance can be measured by mapping quality of service (QoS) criteria in the physical layer, the link layer and the network layer (the packet-level), and grade of service (GoS) criteria in the network layer (the connection-level). The providers must fulfill their guarantees in all layers so that satisfactory information transportation between mobile users can be realized.

With conventional protocol structure, wireless networks can be divided into various layers, such as the physical layer, the link layer, the network layer, the application layer, etc. In the literature, the signal-to-interference ratio (SIR) is a commonly used link layer QoS measure, including SIR in the base station (BS) and SIR in the mobile user [2]. The SIR values are primarily determined by channel impairments which mostly come from multi-path, fading, intra-cell interference, inter-cell interference, ambient noise, etc. SIR should be maintained above a required threshold so that mobile users can communicate with each other reliably. By properly applying resource allocation schemes in the link layer, such as power control and rate allocation, SIRs in both the BS and the mobile user can be controlled to meet the SIR requirements. However, due to the irregularly changing environment, e.g., explosively increased load in neighboring cells, instantaneous channel impairments, etc., SIR may not be perfectly controlled to meet the requirements all the time and data loss occurs during the degradation of SIR. In a practical system, the short-term degradation of SIR is tolerable as the data loss can be compensated by retransmission in the transmitters or recovered through forward error correction (FEC) in the receivers. The outage probability is thus used to measure the degradation probability that SIR is below its requirement. For services without retransmission and FEC, the outage probability in the link layer is the instantaneous packet-loss probability in the network layer.

In the network layer, user mobility produces handoff processes which are non-existent in wireline networks. Thus, besides the new-call-blocking probability, the handoff-call-dropping probability is another important criterion of system performance in the wireless networks. Meanwhile, to minimize the call blocking probability is equivalent to maximizing the system utilization and achieving optimal system performance. Thus, the connection-level GoS, such as the new-call-blocking probability, the handoff-call-dropping probability and the system utilization, and the packet-level QoS, such as the packet-loss probability, are actually related to each other and selected as performance metrics in the network layer [3-5].

The growing demand for wireless access with GoS and QoS satisfaction necessitates the efficient use and reuse of the scarce radio resources. Thus, to achieve effective and efficient resource allocation in different layers is extremely important. In recent years, resource allocation for GoS and QoS provisioning in wireless communication networks has received much attention and is a hot spot in wireless communication research [5-8]. In the thesis, the investigation on resource allocation in cellular code division multiple access (CDMA) systems with cross-layer optimization is conducted. The motivation to employ cross-layer optimization in wireless networks comes from the recognition and understanding of the time-varying characteristics in the wireless links. The system performance in the higher layers will vary along with the changing conditions of the lower layers in the wireless networks and thus should be provided statistical guarantees (‘soft’ guarantees) instead of hard guarantees [9-18]. By jointly optimizing the various performance criteria in different layers, e.g., to maximize GoS, such as the system utilization, subject to the minimum QoS requirements, system resource can be efficiently utilized in layers and the network can therefore accommodate more mobile users. This translates to more revenue for network providers or equivalently the lower charges for individual mobile user. In the thesis, the GoS and QoS metrics in the cross-layer model will be derived analytically and the cross-layer optimization will be achieved by providing these performance metrics with statistical guarantees.

1.1 Motivation

Generally speaking, 1G and 2G are narrowband cellular communication systems, while 3G is wideband. As shown in Fig. 1.1, the multiple access technology used in 1G is frequency division multiple access (FDMA). The multiple access technologies used in 2G are FDMA, time division multiple access (TDMA) and CDMA. The third generation (3G) wireless communication systems are based on CDMA.

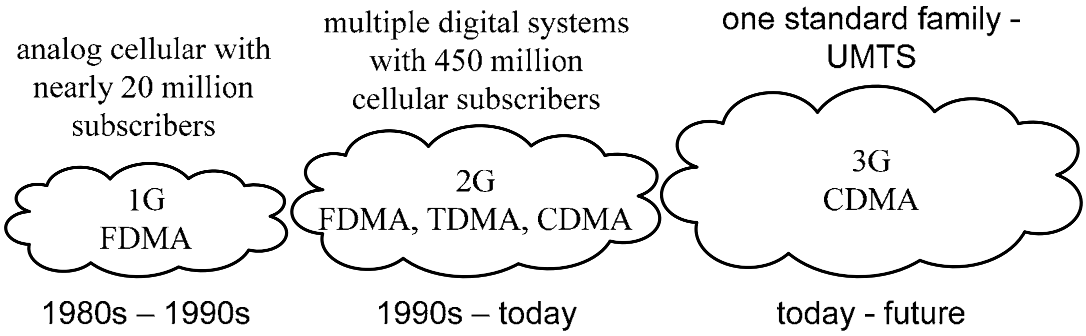

Fig. 1.1. The Evolution of the Cellular System

The differences among the mechanisms of FDMA, TDMA and CDMA are presented in Fig. 1.2. For FDMA, the communication between two transmitters occupies a frequency band during the whole conversation time. For TDMA, the communication between two transmitters uses the whole bandwidth but only in a given time slot. For CDMA, the frequency and time assignments are different from those in FDMA and TDMA. The communication between two transmitters is identified by an orthogonal code spreading across the whole bandwidth during the whole conversation time.

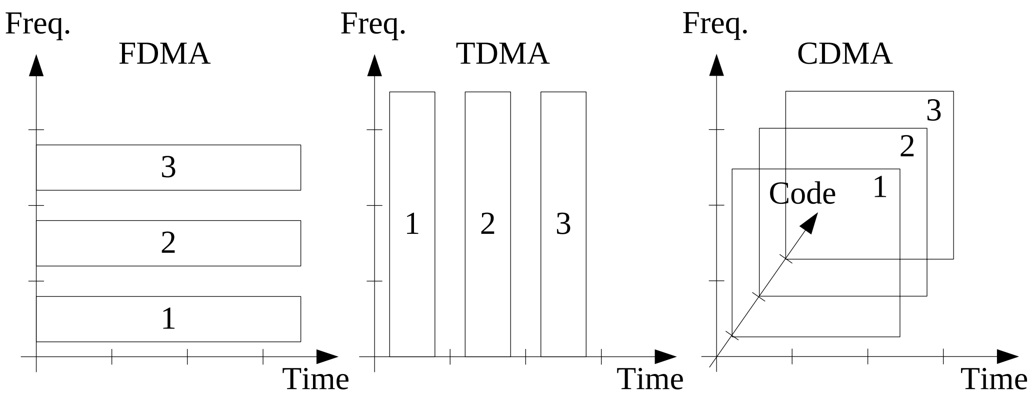

Fig. 1.2. The Mechanisms of FDMA, TDMA and CDMA

In [19], Viterbi applied “the three lessons” from Shannon’s Information Theory for the comparison among CDMA, TDMA and FDMA techniques in personal communications. The author evaluated Shannon's Information Theory as the foundations for the design of efficient wireless communication systems, including those seeking multiple access to a common medium, and concluded that only CDMA and spread spectrum are possible on rendering the interference benign and excel over FDMA and TMDA to approach the ultimate performance limit. The development of cellular systems [1, 2] in recent decade validates the predictions of Dr. Andrew Viterbi.

CDMA is a spread spectrum multiple access method. The bandwidth of the information codes from different users is spread by codes with a bandwidth much larger than that of the information codes. The spreading codes are referred to as pseudorandom noise (PN) sequences. Ideally, PN sequences used for different users are orthogonal to each other. Fig. 1.3 shows the structure of a fraction of a cellular system with hexagonal cells. There are 19 cells with BSs located at their middle points and 3 mobile users, *a*, *b* and *c*, in cells 0 and 4, respectively.

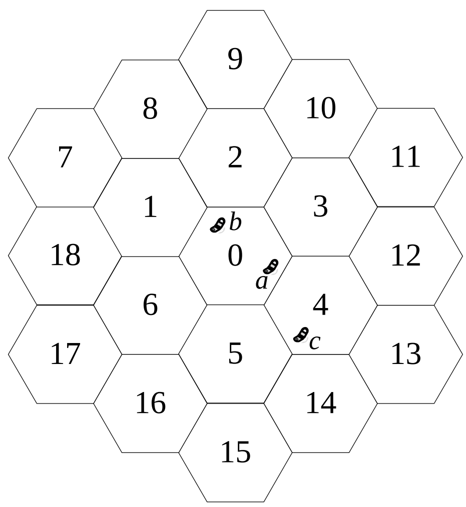

Fig. 1.3. The Structure of a Cellular System

The five salient characteristics of CDMA are summarized using the system structure as follows [1, 20]:

1. Universal frequency reuse: In FDMA, to enlarge the distance between users using same frequency band and reduce the interference from them, the total frequency bandwidth is collectively used by B cells. For example, with $B = 4$, cells 8, 9, 1, and 2 will use the total bandwidth and cells 6, 0, 16, 5 will reuse the bandwidth. However, in CDMA, as the interference is mainly determined by the correlation between PN sequences, the total frequency bandwidth allocated to the system can be reused from cell to cell, i.e., $B = 1$;
2. Soft handoff: Because of the universal frequency reuse, a mobile user is able to simultaneously communicate with several nearby BSs. Thus, during handoff, the mobile user terminates the old BS after the new connection is steadily established;
3. High transmission accuracy: Due to the wide bandwidth used in the spread spectrum, Rake receivers are possible to mitigate the fading dispersive channel impairments and, therefore, improve transmission accuracy [21];
4. Soft capacity: In FDMA or TDMA, the capacity is constrained by the constant number of allocated frequency pairs or time slots. In CDMA, the constraint comes from the intra- and inter-cell interference [19, 22], i.e., the less the interference, the more the number of users supported in the cell, and vice versa. For example, with mobile users *a* and *b* communicating with BS 0 and mobile user *c* communicating with BS 4 in Fig. 1.3, the signals from mobile user *a* will

be received by BS 0 with the mixture of signals from mobile users *b* and *c* as interference. More specifically, the interference from users connected to the same BS is called the intra-cell interference, e.g., the interference from user *b*, and the interference from users connected to other BSs is called the inter-cell interference, e.g., the interference from user *c*. Besides the interference, the QoS requirements from different services are other factors to create soft capacity, i.e., the higher the QoS requirements of the service, the fewer the users of the service coexisting in the system, and vice versa;

5. Flexibility: In CDMA, when a user does not transmit, it will introduce no interference to others and therefore consume no resource. Thus, statistical multiplexing for *ON/OFF* voice traffic and bursty data traffic is easier to implement in CDMA than in TDMA.

The discussions on the five salient characteristics of CDMA demonstrate the necessity of using cross-layer approach in studying cellular CDMA systems. It is because, in CDMA, the instantaneous variation in the physical layer will extensively affect and determine the performance in the higher layers.

For example, statistical multiplexing plays an important role to efficiently utilize system resources. In TDMA or FDMA, statistical multiplexing is normally ignored in the physical layer analysis as a mobile user will exclusively occupy the time slot or frequency pair during communication. However, in CDMA, statistical multiplexing is automatically taken into consideration as a mobile user will not consume any system resource when no signal is transmitting. As a result, by considering statistical multiplexing in the physical layer, the system utilization in the higher layers will inherently be increased by saving the resource wasted during the *OFF* periods [23-25].

For another example, soft handoff plays an important role to determine the intra- and inter-cell interference and can further influence both the reverse and forward links capacities in CDMA systems. Soft handoff decreases the interference and increases the system capacity in the reverse link [26]. On the other hand, in the forward link, soft handoff increases the interference [27] and excessive number of soft handoff users causes the loss from multiple BSs in assigning resources to the same mobile user and diminishes the diversity gain [28-31]. As a result, soft handoff setting in the physical layer will inevitably determine the system capacity and further influence the connection admission decisions in the network layer.

Besides the information from the physical layer, the information from the higher layers should also be shared. The QoS requirements from different services in the higher layers will make a difference between the resource allocations in the lower layers which will be shown in the latter part of the thesis. Thus, the analytical model must involve the lower and higher layers and the parameters must be exchanged among layers to facilitate such a cross-layer design approach.

Besides the necessity of using cross-layer consideration, the cross-layer model also has advantages for optimizing the system performance. With conventional protocol structure, the optimization of resource allocation is processed separately in each layer without information from other layers. Although the system in that particular layer is optimized, the whole system may not be optimal as only the worst-case performance is assumed in other segregated layers. The assumption brings hard guarantees for the QoS and GoS requirements and these requirements are guaranteed all the time even at the worst-case situation in other layers. In the wireless links, the time-varying characteristics will more likely produce instantaneous performance degradation, i.e., even worse worst-case situation. Therefore, the hard guarantees will exceedingly fulfill the QoS and GoS requirements and lead to the wastage of system resource as system performance is typically limited by the average (rather than the worst-case) conditions in CDMA systems [32]. The cross-layer approach is looking at the integrated studies of exploiting the statistical behavior between the performance metrics in various layers to obtain optimal system performance. With the cross-layer optimization, the QoS and GoS are provided with statistical guarantees, e.g., allowing 1% performance degradation, instead of hard guarantees. Thus, the cross-layer optimization will certainly produce a system utilization gain as compared to that in the segregated-layer design.

In the thesis, resource allocation in cellular CDMA systems with cross-layer optimization across the physical layer, the link layer and the network layer is investigated. The cross-layer model is shown in Fig. 1.4 with involving the physical layer, the link layer and the network layer. In the physical layer, system characteristics, including universal frequency reuse, soft handoff, using Rake receiver, soft capacity and statistical multiplexing, are considered. In the link layer, the QoS metrics, including SIR and the outage probability, are considered. In the network layer, the QoS metrics, including the packet-loss probability, and the GoS metrics, including the new-call-blocking probability, the handoff-call-dropping probability and the system utilization, are considered. Formulae for these QoS and GoS metrics are derived analytically. The cross-layer optimization is then proposed based on these metrics and the statistical behaviors between these metrics are coherently investigated. The system utilization and the blocking probability are chosen as the criteria to measure the system performance and the optimal performance is equivalent to the maximization of system utilization and the minimization of blocking probability. Thus, the optimization problem can be described as to maximize the system utilization or to minimize the blocking probability subject to the QoS constraints.

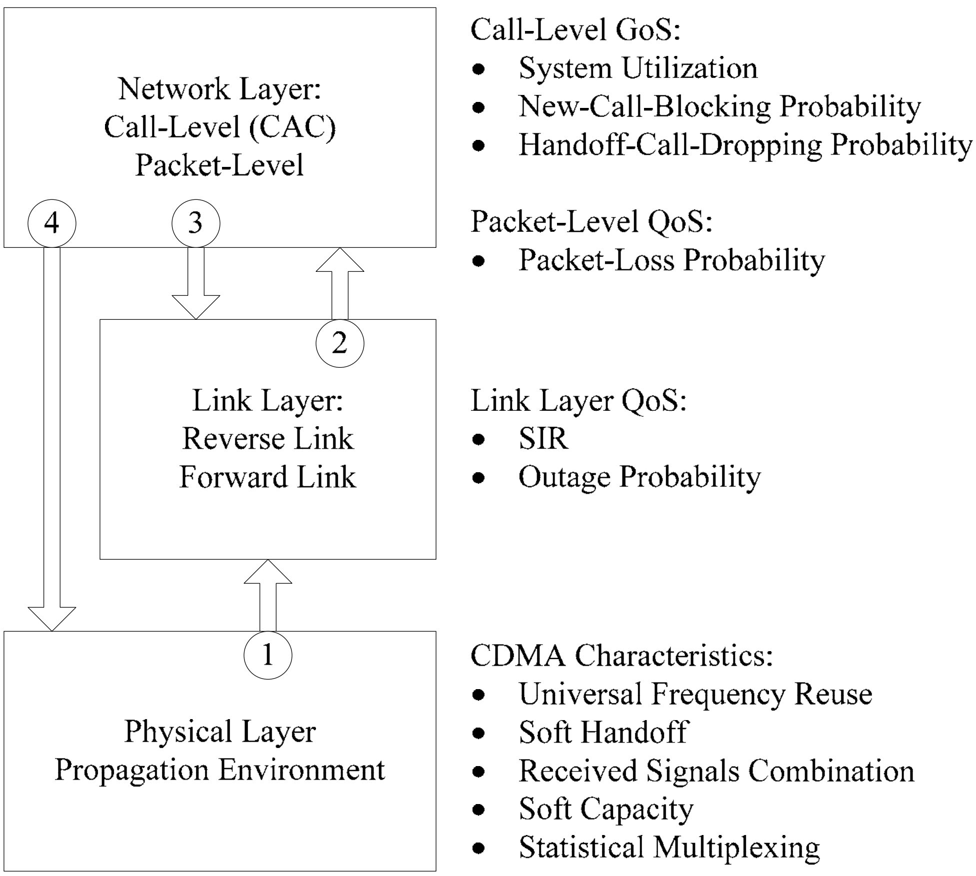

Fig. 1.4. The Cross-Layer Model

In Fig. 1.4, the arrowheads between the blocks represent the exchange of the shared parameters among layers and the influences produced by the shared parameters between layers. Arrowhead 1 represents the influence from the physical layer characteristics, e.g., soft handoff, signals combination, etc., to the link layer QoS in both the reverse and forward links. Arrowhead 2 represents the sharing of the outage in the link layer which is equivalent to the instantaneous packet loss in the network layer.

Arrowheads 3 and 4 represent the QoS requirements from different services affecting the system resource allocated in the link layer and the physical layer, respectively. In the model, dynamic connection admission control (CAC) schemes in the network layer, such as virtual partitioning (VP), will categorize services with different QoS requirements by assigning ranked priorities for the services. The bias in priority in the network layer causes the link and physical layers to allocate unequal resource for different QoS requirement services.

In the thesis, these cross-influence phenomena are exploited and the benefits from the cross-layer design compared to the segregated-layer design are investigated. The numerical results in the latter parts of the thesis will display the benefits. The system parameters settings in the numerical results are following those in the standardized wideband CDMA (WCDMA) technology which is the main 3G air interface in the world [2].

1.2 Outline of the Thesis

Chapter 2 reviews the related work in three major areas, resource allocation in the network layer, issues in the reverse and forward links and the cross-layer optimization. The literature surveys on CAC schemes, including complete sharing (CS), complete partitioning (CP) and VP, joint packet and connection levels QoS optimization, soft capacity, soft handoff, the reverse and forward links performances, and capacity unbalance problem between the reverse and forward links, etc.

The studies on modeling the cross-layer optimization in cellular CDMA system are arranged into four chapters, Chapter 3 to Chapter 6. Chapter 3 aims at providing a platform for the analyses in the following chapters. The system structure, system parameters and assumptions are introduced here.

In Chapter 4, the CAC schemes, including CS, VP with preemption for all classes, and best effort and guarantee access with preemption for best effort traffic are investigated. The general method to implement VP for multi-class traffic with guard channel in cellular systems is proposed. The joint connection and packet levels optimization method is employed to maximize the system utilization.

The link layer analyses for both the reverse and forward links are presented in Chapter 5. The focus of the chapter is to derive the formulae for SIRs and the outage probabilities in the reverse and forward links, respectively. The admission regions in the link layer are then obtained using the derived formulae.

The GoS and QoS metrics are formulated in Chapter 4 and Chapter 5. In Chapter 6, the cross-layer model in Fig. 1.4 is built based on these GoS and QoS metrics of the segregated layers. The general cross-layer model is proposed to maximize system utilization subject to the QoS constraints. Based on the general model, a practical problem, the capacity unbalance problem, is solved with an adaptive soft handoff probability (SHP) scheme. The comparisons between the cross-layer design and the segregated-layer design are made. Finally, conclusions and future work are discussed in Chapter 7.

1.3 Contributions of the Thesis

This section summarizes the main contributions of the thesis.

In Chapter 3, a platform for cross-layer analysis as shown in Fig. 1.4 is proposed. The system structure which covers the salient features of cellular CDMA systems, including soft handoff, PN sequences, path loss model, power control schemes, interference model, etc., is illustrated. The relationship between the SHP and the hysteresis margin is derived. The traffic models for multimedia services, such as voice, video, web-browsing and data services, etc., are presented. The CAC schemes, including CS and VP, are introduced to demonstrate the techniques of sharing a common resource for users. The overview of cross-layer optimization is also presented as the guide for the rest of the thesis.

In Chapter 4, the investigation on the network layer issues is processed. The aim of the work is to give a general solution method that admits extension to multi-class traffic for different CAC schemes. However, using the existing technique [47-50, 52, 53], to extend VP scheme to multi-class traffic is very difficult. The contribution in the chapter is to propose such a general method with which other CAC schemes can be formulated to multi-class traffic in the same way. Three CAC schemes for handling multi-class service with guard channels, including CS, VP with preemption for all classes (VP-Case 1) and best effort and guarantee access with preemption for best effort traffic (VP-Case 2), are analytically formulated. The analytical models, derived using a *K*-dimensional Markov chain, are solved using preemption rules for VP schemes. The formulae for the GoS metrics, including the new-call-blocking probability, the handoff-call-dropping probability and the system utilization at the connection-level, and the QoS metrics, including the packet-loss probability at the packet-level, for different CAC schemes, are derived. A method to maximize system utilization through joint optimization of connection and packet levels parameters is proposed. Numerical results indicate that over 30% gain in system utilization can be achieved using the joint levels optimization.

In Chapter 5, the admission regions that represent the largest set of QoS points delivered under any CAC schemes are derived in the forward and reverse links and in the link layer [33]. The analytical models are based on the largest received power BS selection criterion in both the reverse and forward links instead of the nearest BS selection criterion which is applied in literature. Interference models are carefully developed with soft handoff, diversity and statistical multiplexing, and impractical assumptions are avoided. In the forward link, an approximation selection method combining the advantages of the previously used Gaussian and lognormal approximations and adapting to the QoS specification for different services is proposed. Another contribution beyond the previous literature [63] is the removal of the need for simulation to obtain the intermediate parameters. Furthermore, different power control schemes for mobile users in soft handoff are investigated and compared. The SIRs and the outage probabilities for multi-class services in BSs (the reverse link) and in mobile users (the forward link) are formulated as the output of the link layer analysis. The admission region represents the region with hard outage guarantee.

In Chapter 6, a cross-layer optimization across the physical layer, the link layer and the network layer to model CDMA cellular systems is designed and solved analytically. This work differing from those in the literature [9, 13, 15] where the three layers are jointly considered is that the salient features of CDMA systems, including universal frequency reuse, soft handoff, soft capacity, micro-diversity, and statistical multiplexing, are all modeled analytically into the cross-layer optimization. A function block, the cross-layer decision-maker (DM) through which the cross-layer optimization will be applied without disturbing the integrity of the conventional protocol structure is proposed. The general cross-layer optimization method is proposed to maximize system utilization subject to the QoS constraints. With the general model, a practical problem, the capacity unbalance problem is solved by an adaptive SHP scheme which deals with the problem, for the first time, using the cross-layer model. In the numerical results, over 60% gain in system utilization can be achieved with the cross-layer optimization and the adaptive SHP scheme over the conventional segregated-layer design.

The research projects presented in the thesis are either published, or have been submitted for publication [34-40].

Chapter 2

Literature Review

This chapter reviews three major areas related to the research described herein. These are resource allocation in the network layer, issues in the reverse and forward links and the cross-layer optimization.

2.1 Resource Allocation in the Network Layer

The network layer is normally divided into two levels, the connection-level and the packet-level. The resource allocation in the connection-level is known as the connection admission control (CAC). There are many existing CAC schemes to manage the connection admission decisions. In the thesis, three schemes, namely, complete sharing (CS), complete partitioning (CP) and virtual partitioning (VP), are focused upon. In the packet-level, the packet loss is investigated. The loss is caused by the excess of the required resource from the co-transmitting users over the total system resource. By considering both the connection-level and the packet-level, the joint connection and packet levels optimization will be studied.

2.1.1 Connection Admission Control Schemes

Two commonly used GoS measures in CAC schemes are the new-call-blocking probability and the handoff-call-dropping probability. Based on the fact that maintaining an ongoing call is more important than admitting a new call, the admission of new and handoff calls has to be treated differently in resource allocation [3]. Many resource allocation schemes have been proposed to meet both GoS constraints and the need to maintain service continuity [41-45]. In [41], CAC with guard channels which assign higher priority to handoff calls over new calls is proposed. Certain amounts of capacity are reserved as guard channels for accepting handoff requests. The guard channel method in [41] is a fixed reservation strategy. In [42], dynamic reservation methods, including new call bounding, cutoff priority and new call thinning schemes, are investigated. In [43-45], other dynamic resource allocation schemes are proposed and investigated. The dynamic resource allocation schemes change the admission rules according to the variation in the system parameter values. Thus, these schemes are more complex but give better system performance by lowering the blocking and dropping probabilities.

Besides the priorities for handoff calls over new calls, it is also necessary to assign different priorities for multiservice. In future wireless systems, the provision of multimedia services with QoS guarantees is a crucial requirement. The multimedia services include audio, video, web-browsing, data services, etc. Each service has distinctive characteristics differentiated from others and the characteristics of different services are sensitive factors to the system performance [46]. In general, applications and services can be divided into different classes according to their required bit rates and QoS guarantees. For example, considering the delay sensitivity for different services, there are real-time (RT) services, e.g., the video service, and non-real-time (NRT) services, e.g., the data service. Considering the loss sensitivity for different services, there are nearly error free requirement services, e.g., the data service, and acceptable high bit error rate services, e.g., the audio service. Considering traffic symmetry between the forward and reverse links, there are symmetrical services, e.g., the VoIP service, and asymmetrical services, e.g., the web-browsing service. Services with different QoS requirements possess different priorities. Thus, CAC schemes have to play the role of distinguishing different services from each other by applying prioritized admission rules.

Two classical CAC schemes in multi-class traffic networks are CS, which allows all classes to share the resource with equal priorities indiscriminately, and CP, which statically divides the resource among the classes and exclusively allows each class to use its allocated capacity [47]. Under CS, system resource could be fully utilized when the system load is light. However, one class may overwhelm all the others when the particular class users are overloaded. Whilst under CP, the flood from any class can be prohibited. However, the resource may be underutilized when the total system load is light. To combine the advantages and to diminish the drawbacks of these two schemes, there is another CAC scheme, VP.

VP strikes a balance between unrestricted sharing in CS and unrestricted isolation in CP [48]. VP was originally proposed by Wu and Mark in [49]. The concept of VP is that each individual traffic class is allocated a nominal amount of resources with different priorities and with the provision that under-utilized resources can be used by the excess traffic of an overloaded class, subject to preemption. The under-utilized resources come from the traffic classes whose arrival rates are below thresholds which are set based on past statistics. In this situation, the nominal allocation for under-loaded classes can be utilized by other traffic classes. However, VP performs preemption for the under-loaded classes when their arrival rates revert to their thresholds. For traffic whose arrival rates are higher than the thresholds, if the overall traffic is light, the overloaded classes can use the nominal allocation of other traffic classes just like in CS. If the overall traffic becomes heavy, the overloaded classes are preempted by other traffic classes and can only use the nominal allocation for themselves just like in CP. VP behaves like unrestricted sharing when the overall traffic is light and complete isolation when the overall traffic is heavy [50]. Thus, VP combines the best characteristics of CS and CP under different loadings.

Research on VP has received much attention in recent years. In [47], Mitra and Ziedins applied VP in cellular system and considered only the single class case. In [50], Borst and Mitra extend their work to 2 classes. Comparisons between CS, CP and VP with guard channels are made in [41]. Results for 2 classes system are presented. In [51], Wong et al. consider CS and CP with guard channels for K classes. The transmission bandwidth of each class can be an integer multiple of those for other classes. In [52, 53], Wong et al. consider VP with preemption for groups 1 and 2 with guard channels pertaining to no more than 2 classes. To extend VP scheme to multi-class traffic using the existing technique is very difficult. Thus, to propose a more general way that admits extension to multi-class traffic will be valuable.

2.1.2 Joint Connection and Packet Levels Optimization

Resource allocation for connection admission is basically a connection-level problem, but satisfying GoS constraints at the connection-level alone may limit the traffic load admitted. When traffic flows are admitted into the network proper, QoS is measured in terms of packet loss rate, packet delay and packet delay variation at the packet-level. Scheduling and statistical multiplexing gains play a crucial role in determining the amount of traffic that can be admitted into the network proper while still satisfying the packet-level QoS. With the gain from scheduling and statistical multiplexing, the packet-level can sustain a larger load than that constrained by satisfying the GoS at the connection-level. Thus, there could still be excessive allocation of resources between these two levels. It is believed that making use of both the connection-level and packet-level properties can enhance system utilization with GoS and QoS constraints at both levels.

Beshai et al. [54] are the first to suggest using the interaction between the connection-level and the packet-level GoS and QoS in ATM networks to improve system performance. However, the work in the paper does not support user mobility. In the case of wireless networks, the ability to support user roaming is the key feature, and user mobility affects the attainable system throughput and the satisfaction of GoS and QoS requirements. Cheung and Mark [23] have proposed a resource allocation strategy subject to joint connection and packet levels GoS and QoS constraints. They found that there is a significant improvement in system utilization when the deployment of system resources is subjected to simultaneous satisfaction of both connection-level and packet-level GoS and QoS constraints.

2.2 Issues in the Reverse and Forward Links

The reverse and forward links performances are investigated herein. The five salient physical layer characteristics in CDMA, including universal frequency reuse, soft handoff, using Rake receiver, soft capacity and statistical multiplexing, will determine the link layer performance. The universal frequency reuse is the fundamental characteristic of CDMA to differ from FDMA and TDMA. The other characteristics in CDMA are also due to the universal frequency reuse. For examples, using Rake receiver improves transmission accuracy with mitigating the fading dispersive channel impairments [21]. Statistical multiplexing is the inherent feature of CDMA which is automatically applied in the network layer [23], the link layer [24] and the physical layer [25]. The three characteristics, universal frequency reuse, using Rake receiver and statistical multiplexing, will be modeled into the analytical model in the following chapters. In this section, literature reviews on soft capacity and soft handoff are presented. Furthermore, the capacity unbalance problem, i.e., inefficient use of radio resource induced by asymmetrical traffic in the reverse and forward links, is studied and past research on this problem is investigated.

2.2.1 Soft Capacity and Soft Handoff in the Reverse Link

The capacity bound in CDMA systems is interference-limited and this capacity is referred to as soft capacity. The soft capacity in CDMA systems includes the reverse link capacity and the forward link capacity. Some work has been done to model the soft capacity in the reverse and forward links, respectively. For the reverse link, in [55], the soft capacity was modeled by an upper limit and a lower limit of the capacity bounds. The capacity bounds are changed randomly in the model. However, the change of the capacity bounds should occur according to the variation of the interference level. In [56], based on the reverse link capacity model in [2], the soft capacity for 3G (WCDMA) system was investigated through simulation. The four services carried by 3G systems are discussed. In [34], the soft capacity model in one BS was determined analytically based on the assumption of uniformly distributed mobile users for a multi-service CDMA system. However, in these work, the effect from soft handoff on the reverse link capacity is not considered.

In CDMA, soft handoff is a technique that extends cell coverage and increases the reverse link capacity [26]. The technique allows the mobile users in soft handoff to be power-controlled by the smallest attenuation BS while still connected to other nearby BSs involved. In [57], soft handoff in multi-code CDMA system was modeled with a traffic management scheme. A Markov process-based approach was used to study the system performance for a multi-service system in the network layer. However, the interference model was built on the worst-case performance assumption where a hard bound for SIR requirement was applied and the inter-cell interference was not statistically expressed. In [58], the interference model for a DS-CDMA cellular system with soft handoff was built to evaluate the network performance in terms of call blocking probability. However, the analytical model presented is difficult to be extended to multi-class traffic. In [59], a traffic model using the interference-based CAC with soft handoff was developed. The time-varying status of multi-class traffic in the neighboring cells was considered. A geographical structure with three regions was assumed to identify the mobile users in soft handoff or non-soft handoff. However, this assumption is not practical and will over-estimate the system performance as the unanticipated users, e.g., a user in soft handoff but is located within the non-soft handoff region, will produce a lot of interference to the whole system.

2.2.2 Soft Capacity and Soft Handoff in the Forward Link

In future wireless networks, the provision of multimedia services is crucial. Some services have asymmetric transmission traffic streams between the reverse and forward links [60, 61], such as video, web-browsing, data services, etc. Traffic streams are normally heavier in the forward link especially for services such as web browsing. This would make the forward link the bottleneck in the future systems. As a result, providing accurate capacity estimation in the forward link is useful and necessary.

Practically, in the forward link, a mobile user in non-soft handoff is served by the BS from which it received the largest power [22] and a mobile user in soft handoff is served by the first two or three largest received power BSs. This largest received power BS criterion is referred to as the practical BS selection criterion. However, most of the research work evaluating the forward link capacity makes the assumption that a mobile user is simply connected to the closest BS/BSs due to the difficulty in formulating an expression for the interference power at the mobile user. In [62], Kim and Stuber provided a forward link model to study the effects of soft handoff and assumed that only mobile users in a specified region are allowed to perform soft handoff. However, the assumption is not reasonable in practice as soft handoff may happen anywhere. To build a more realistic model, Zhuge and Li derived the modified path loss ratio using the practical BS selection criterion in [63]. However, Monte Carlo simulation is needed to obtain the interference characteristics before computing the forward link capacity analytically. To provide a more accurate and convenient forward link capacity estimation method, an analytical framework with the practical BS selection criterion and without the need of simulation to obtain the intermediate parameters is necessary.

In the forward link, excessive number of soft handoff users will cause the loss in assigning resources from multiple BSs to the same mobile user and thus diminishes the gain of diversity [29-31]. In [29], the extra load on the system due to the increase in the number of channels occupied and reserved for macro-diversity was investigated. The information related to the required number of channels, the offered load and the blocking probability under various ratios of soft handoff users to the total users was provided. In [30], the losses of soft and softer handoff on the forward link capacity for unsectorized and sectorized cells due to the use of at least two BSs in the handoff process are evaluated analytically in various propagation environments. In [31], the forward link capacity is found to decrease if the ratio of number of soft handoff users to the total number of users becomes too large. These pieces of work provide valuable guide to explore the effect from the soft handoff on the forward link capacity.

2.2.3 Capacity Balancing between the Reverse and Forward Links

The performance evaluations on the reverse and forward links were reviewed in the two sections above. A brief conclusion can be drawn that the forward link would more likely be the bottleneck of the system capacity if there were increasingly more traffic streams in the forward link than those in the reverse link. Thus, the incoming service may not be admitted due to the finite amount of radio resource in the forward link even if there is still sufficient radio resource in the reverse link to accommodate more services. This inefficient use of radio resource induced by asymmetric traffic is referred to as the capacity unbalance problem. To solve the capacity unbalance problem, many schemes have been proposed [60, 61, 64-66].

In [60], a concept of CDMA/Time Division Duplex (TDD) systems with different number of time slots between the two links was proposed. System utilization was maximized by allocating an asymmetrical number of time slots to the two links in proportion to their traffic volumes and this is easily realizable through the use of TDD instead of Frequency Division Duplex (FDD) mode. In [61], the problem of "which link is the limiting link of system capacity and what are the factors determining this limiting link" was investigated in the link layer. The influencing factors considered in their work include spatial distribution of mobile users, inter- and intra-cell interference, and traffic asymmetry, etc. In [64, 65], another influencing factor, handoff, was considered. In [64], the number of mobile users in handoff between cells was adjusted to achieve traffic volumes balancing in separate cells through the control of a hysteresis margin in the physical layer when performing power control. In [65], asymmetric bandwidth allocation was applied with guard bandwidth for handoff calls. System utilization is formulated and optimized with a Markovian analysis in the network layer. In [66], Maeda et al. proposed broadband packet wireless access for next generation networks in FDD mode. The asymmetric bandwidths for the forward and reverse links were respectively allocated with 101.5 MHz and 40 MHz. However, in the literature, resource allocation with capacity balancing is optimized separately either in the network layer, or in the link layer, or in the physical layer only. Although the system in that particular layer can be optimized, the whole system may not be optimal as only the worst-case performance is studied in the segregated layer.

From the literature review in sections 2.2.1 and 2.2.2, it can be seen that soft handoff decreases the interference and increases the system capacity in the reverse link. However, excessive number of soft handoff users causes the loss from multiple BSs in assigning resources to the same mobile user and diminishes the gain of diversity in the forward link. Thus, the capacity unbalance problem can be solved with an adaptive soft handoff probability (SHP) scheme. The principles of the adaptive SHP scheme are, when the traffic asymmetry is biased toward the forward link, the forward link capacity can be increased by decreasing SHP, and when the traffic is symmetrical, the reverse link capacity can be enlarged by increasing SHP. The SHP is controlled adaptively along with the changing traffic volumes in the reverse or forward link. The hysteresis margin can be used to control the number of mobile users in soft handoff [64, 67] in the physical layer. For example, in WCDMA, the use of the common pilot channel (CPICH) reception level in the mobile user for handoff measurements has the consequence that by adjusting the CPICH power level the number of users involved in the handoff can be controlled [2].

2.3 Cross-Layer Optimization

The ISO (International Standards Organization)/OSI (Open Systems Interconnection) 7-layer model for networking is the most popular networking protocol applied in the wireline networks. The communication tasks are divided into 7 layers and implemented independently. The correlative dependence of one layer on other layers is weak. However, in the development of wireless networks, the cross-layer optimization has become a research hot spot, which conflicts to the OSI 7-layer model requirements since sharing of information among different layers becomes necessary in order to achieve performance optimization. The realization of the need for cross-layer optimization comes from the recognition and understanding of the time-varying characteristics in the wireless link so that system resources in the higher layers could be allocated efficiently based on instantaneous information instead of the worst-case assumption from the lower layers. Providing QoS in the higher layers then becomes ‘soft’ as QoS parameters are adjusted along with changing channel conditions and can be guaranteed statistically [14].

With cross-layer optimization, relevant parameters from segregated layers are extracted for use in cross layers. The extracted parameters and the application approaches with these parameters have been prevalently reported in the literature. In [12], information of the wireless medium from the physical layer and the medium access control (MAC) layer is shared with higher layers to provide efficient resource allocation applications over the Internet. With the improved transmission control protocol (TCP), the source will be informed of the losses caused by congestion and the losses over the wireless channel so that the TCP mechanism can reduce the redundant retransmissions and the throughput loss. In [11, 17], the updates of measurements in the physical layer are used to optimize the MAC schemes. System efficiency improvement is achieved and the cross-layer design leads to a gain in performance. In [10, 16], the cross-layer optimization is processed between the physical layer and the link layer to achieve optimal power control. In [18], the influence of the physical layer constraints, such as the multi-access interference (MAI), on the link and network layers performance, such as BER, the throughput, and the network connectivity, etc., in a wireless CDMA sensor network is studied. In [9, 13, and 15], cellular CDMA systems that jointly consider characteristics of the physical, link and network layers were proposed to achieve effective QoS guarantee and efficient resource allocation.

Although the ideas of the cross-layer design to share parameters among layers in the literature are similar, the work in different papers is to fulfill different optimization objectives and apply different schemes with different shared parameters among layers. The use of different parameters in different schemes introduces difficulties to integrate various cross-layer design schemes in one system. This is because as the number of cross-layer adaptation schemes in one system increases, the chance that the optimal parameter values required by different schemes may conflict with each other increases [68]. Thus, a protocol architecture that considers cross-layer interactions is required [69]. To set up protocols for cross-layer design, research on cross-layer optimization should be further carried out. Various cross-layer models can be built to measure their effectiveness to optimize systems and different shared parameters among layers should be investigated to evaluate their effectiveness in representing their own layers.

Chapter 3

Analytical Platform of Cellular CDMA System

The system models for the ensuing analysis are first introduced. These models include the physical layer model, the link layer model and the network layer model. These models of the segregated layer are harmoniously integrated into the cross-layer model in the rest part of the thesis. In this chapter, the system structure which covers the salient features of cellular CDMA systems, including soft handoff, PN sequences, path loss model, power control schemes, interference model, etc., are firstly illustrated. Then, the traffic models for multimedia services, such as audio, video, web-browsing and data services, etc., are introduced. Besides, the CAC schemes in the network layer, including CS and VP, are presented to show their mechanisms of sharing a common resource for multi-class services. Guard capacity to reserve resource for handoff calls is also considered. Finally, an overview on cross-layer optimization is given and an indication of where the system utilization gain comes from is presented. Some assumptions are made in the models. The thesis will show that these assumptions are normally acceptable in the literature and adequately preserve the precision for the cross-layer design.

3.1 Overview of System Structure

Fig. 3.1 shows a simplified CDMA system where two BSs (*n* and *h*) are drawn to represent the home BS/BSs while other inter-cell interference BSs are all represented by BS *b*. The cell shape is not specified in this thesis, i.e., the analytical model is applicable for any kind of cell shape. There are two major cell shapes used in the literature, hexagonal cell and square cell. The hexagonal cell structure has been illustrated in Fig. 1.3. The square cell structure is also known as the Manhattan model [70-72] and will be shown in Fig. 4.1. In this thesis, to simplify the calculations in the analysis and simulation, the square cell structure is used in Chapters 4, 5 and 6.

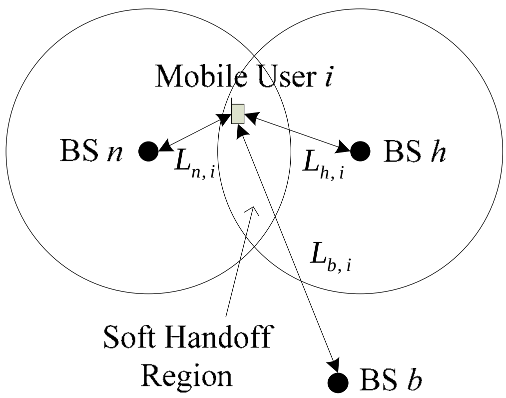

Fig. 3.1. System Structure

The channels of all cells are assumed to be statistically identical and mobile users are uniformly distributed. Although in the modeling, soft handoff can take place anywhere in the cell, most of them will still take place only in the regions between neighboring cells. A two-leg soft handoff model is used in this thesis. This means that the mobile user in soft handoff connects to its two smallest attenuation BSs (active BSs), in both the reverse and forward links communications.

3.1.1 Power Control in the Reverse and Forward Links

Power control techniques are applied in cellular networks. In the reverse link, mobile users in a cell transmit their power to the same BS. If there is no power control technique applied, the closest transmitter is often able to capture the BS due to its smallest attenuation to the BS and it introduces more interference than those from the other transmitters. This is called the near-far problem. Power control is applied to avoid the near-far problem in the reverse link. On the other hand, in the forward link, due to the one-to-many scenario, there is no near-far problem since all the signals within one cell originate from one BS to all mobile users. However, power control is desirable to provide a marginal amount of additional power to mobile users at the cell edge as they suffer from increased inter-cell interference. The closed-loop power control is designed as a rigorous solution for power setting in future cellular networks [2]. When performing power control, SIR is a prevalently used signal quality measure. In this thesis, the perfect SIR-based power control schemes in both the reverse and forward links are assumed. That is, the BS has equal received powers from all the same service mobile users belonging to that cell in the reverse link and all the same service mobile users in the cell have the same SIR target in the forward link [4].

In the forward link, long scrambling codes are used to distinguish the cells.

Due to the asynchronous transmission from BSs, the total transmission powers from nearby BSs must be included into the inter-cell interference since the long scrambling code does not have an impulsive autocorrelation. In each cell, orthogonal codes are used by mobile users and this greatly reduces the interference among them. However, perfect orthogonality cannot be achieved in a multi-path environment, resulting in intra-cell interference. An orthogonality factor f is introduced to model this loss, where f is defined as the ratio of the intra-cell interference to the total transmission power from the home BS.

In the reverse link, the transmissions from different sources are separated by the long scrambling codes. The transmission powers from transmitters with long scrambling codes should be totally treated as interference. Thus, the transmission powers from any other mobile users except itself are included into the interference. As the reverse link performance is mainly determined by the cross correlation between codes, the multi-path (fast fading) effect is ignored in the reverse link analysis [21].

During soft handoff, mobile users will connect to two BSs simultaneously. Thus, in the reverse link, two active BSs will receive the signal from the mobile user, and in the forward link, the mobile user will receive two signals from the two active BSs. Diversity can be employed to improve transmission performance by combining the independent signals. There are at least four combining techniques, namely, selection diversity, maximal ratio combining (MRC), equal gain combining and feedback diversity [73]. In the reverse link, selection diversity that only the primary BS (the largest channel gain BS) power-controls the mobile user and relays the signals to the destination is applied. Another active BS will simply receive the signal. In the forward link, MRC is applied in the mobile user [2]. Thus, the received signals from the active BSs are summed up to achieve the best performance.

Power control schemes for mobile users in soft handoff in the forward link are more complex than those in the reverse link. The power control schemes in the forward link will determine the levels of transmission powers from active BSs so that the received SIR in the mobile user can fulfill the desired threshold. The power control schemes have a great influence on the forward link performance from many aspects, such as the interference, the total allocated power and the link layer outage, etc., and even on the reverse link performance due to the increased overheads used in the reverse link control signals. The investigation and comparison for four forward link power control schemes for mobile users in soft handoff will be presented in Chapter 5. They are the equal signal-to-noise ratio per bit (E_b/I_0) power control scheme [74, 75], the balancing power control scheme [2, 76, 77], the unbalancing power control scheme [78-80] and the site selection diversity transmission power control (SSDT) scheme [81].

3.1.2 Path Loss Model in the Propagation Environment

The propagation environment is modeled using path loss with the large-scale (or long-term) shadowing component, and, like those in the literature [22, 45, 63], the small-scale (or short-term) fading caused by multi-path propagation is ignored. The generally adopted lognormal transmission model is used. The channel gain of a mobile user i at a distance $d_{b,i}$ from the referenced BS b is given by

$$L_{b,i} = -10\tau \log d_{b,i} + \xi , \qquad (3.1)$$

where ξ is a Gaussian random variable (RV) with standard deviation, σ_ξ, and zero mean, and τ is the path loss exponent. The path loss exponent, which is normally determined from measurements, is in the range of 2-5. As perfect power control is assumed, the transmission powers at the mobile users (or BSs) will have to compensate for such variations in the reverse (or forward) link and hence they are also lognormally distributed. Thus, in the following chapters, when analyzing interference, the lognormal distribution is assumed.

3.1.3 Soft Handoff Decisions in the Reverse and Forward Links

During communication, soft handoff does not occur until the pilot signal that a mobile user receives from the neighboring BS is higher than that from the local BS by a hysteresis margin, Δ_{dB}. In the reverse link, the mobile user will always be power-controlled by the largest channel gain BS. Thus, for a total of $(B+1)$ BSs, if a mobile user i in non-soft handoff is power-controlled by BS n, the conditions of the channel gains from BSs to mobile user i are

$$L_{b,i} < L_{n,i} - \Delta_{\mathrm{dB}}, \text{ for } b \neq n \text{ and } b = 0, \ldots, B. \qquad (3.2)$$

Let a mobile user i in soft handoff be connected to BSs n and h, and power-controlled by BS n. Thus, the conditions of the channel gains from BSs to mobile user i are

$$L_{b,i} < L_{h,i} < L_{n,i} \text{ and } L_{h,i} \geq L_{n,i} - \Delta_{\mathrm{dB}}, \text{ for } b \neq n, \ b \neq h, \ b = 0, \ldots, B, \qquad (3.3)$$

which guarantees that the channel gains from BSs n and h to the mobile user are within the hysteresis margin.

In the thesis, the duplex mode is not specified, i.e., the analytical platform is built for both the TDD and FDD modes. In the TDD mode, the channel gains in the reverse link and in the forward link between mobile user and BS are same to each other as the frequency bands used in the reverse and forward links are same. Thus, the BS/BSs connected to the mobile user in the reverse link are the same as the one/those in the forward link. In the FDD mode, the channel gains in the reverse link and in the forward link between mobile user and BS may be different to each other as the frequency bands used in the reverse and forward links are different. However, the possible difference on BS assignments in the reverse and forward links in the FDD mode will not produce an effect on the performance evaluation in the reference BS if the channels of all cells are assumed to be statistically identical and independent.

In this thesis, the BS/BSs connected to the mobile user in the reverse link are the same as the one/those in the forward link in both the TDD and FDD modes. This enables soft handoff in the reverse and forward links to be considered together. In the forward link, if a mobile user in non-soft handoff is connected to BS n, the conditions of the channel gains are the same as (3.2) in the reverse link. If mobile user i in soft handoff is connect to BSs n and h, the conditions of the channel gains from mobile user i to the BSs are

$$L_{b,i} < \min\left(L_{n,i}, L_{h,i}\right) \text{ and } L_{h,i} - \Delta_{\mathrm{dB}} \le L_{n,i} \le L_{h,i} + \Delta_{\mathrm{dB}},$$

$$\text{for } b \ne n,\; b \ne h,\; b = 0, \ldots, B, \tag{3.4}$$

which guarantee that either BS n or h has the largest channel gain to the mobile user and the channel gains from BSs n and h to the mobile user are within the hysteresis margin.

During soft handoff, the handoff process is initialized by the mobile users through measurement of the received pilot signal [2]. Thus, the soft handoff decision is made in the forward link by the mobile users. SHP, denoted by r, is determined by Δ_{dB} with

$$r = \iint_A \sum_{n=0}^{B} \sum_{h=0,h\neq n}^{B} \Pr\begin{pmatrix} \max\{L_{b,i} \mid b = 0,\dots B, b \neq n, b \neq h\} \leq \min(L_{h,i}, L_{n,i}), \\ L_{h,i} - \Delta_{\mathrm{dB}} \leq L_{n,i} \leq L_{h,i} + \Delta_{\mathrm{dB}} \end{pmatrix} dA, \tag{3.5}$$

where

$$\begin{aligned} &\Pr\left(\max\{L_{b,i} \mid b = 0,\dots B, b \neq n, b \neq h\} \leq \min(L_{h,i}, L_{n,i}), L_{h,i} - \Delta_{\mathrm{dB}} \leq L_{n,i} \leq L_{h,i} + \Delta_{\mathrm{dB}}\right) \\ &= \Pr\left(\max\{L_{b,i} \mid b = 0,\dots B, b \neq n, b \neq h\} \leq L_{h,i}, L_{h,i} \leq L_{n,i} \leq L_{h,i} + \Delta_{\mathrm{dB}}\right) \\ &+ \Pr\left(\max\{L_{b,i} \mid b = 0,\dots B, b \neq n, b \neq h\} \leq L_{n,i}, L_{n,i} \leq L_{h,i} \leq L_{n,i} + \Delta_{\mathrm{dB}}\right) \\ &= \int_{-\infty}^{x_h} \cdots \int_{-\infty}^{+\infty} \cdots \int_{x_h}^{x_h+\Delta_{\mathrm{dB}}} \cdots \int_{-\infty}^{x_h} [f_1(x_1) \cdots f_h(x_h) \cdots f_n(x_n) \cdots f_B(x_B)] dx_B \cdots dx_n \cdots dx_h \cdots dx_1 \\ &+ \int_{-\infty}^{x_n} \cdots \int_{x_n}^{x_n+\Delta_{\mathrm{dB}}} \cdots \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{x_n} [f_1(x_1) \cdots f_h(x_h) \cdots f_n(x_n) \cdots f_B(x_B)] dx_B \cdots dx_n \cdots dx_h \cdots dx_1 \end{aligned},$$

A represents the whole system area and $f_b(x_b) = \exp\left(-\left(x_b + 10\tau \log d_{b,i}\right)^2 / \left(2\sigma_\xi^2\right)\right) / \left(\sqrt{2\pi}\sigma_\xi\right)$ is Gaussian distributed. Note that the probability density function of jointly Gaussian distribution is the product of the probability density functions (p.d.f.) of Gaussian distribution if the RVs are independent.

The relationship between SHP and the hysteresis margin with a setting of $\sigma_\xi = 6$ dB and $\tau = 4$ is presented in Fig. 3.2. The result is consistent to the simulation results presented in [82]. SHP increases almost linearly with the increase of the hysteresis margin.

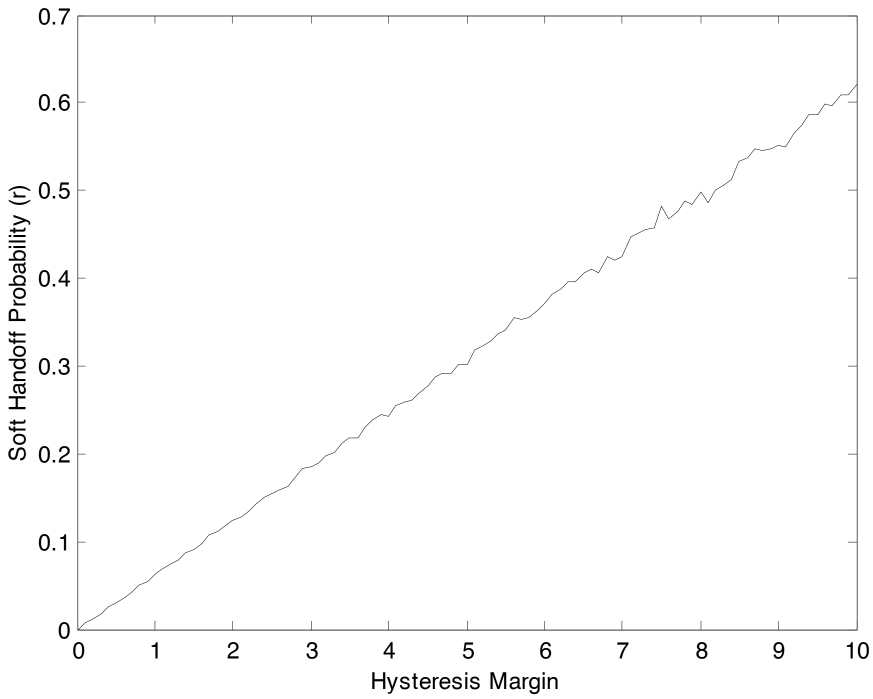

Fig. 3.2. Soft Handoff Probability vs. Hysteresis Margin

SHP increases linearly with the increase of hysteresis margin. This is expected. With a large value of hysteresis margin, the difference among the channel gains between BSs and mobile users will be more likely involved in the margin, which means more mobile users are in soft handoff. Thus, the number of mobile users in soft handoff can be controlled through the control of the hysteresis margin.

3.2 Traffic Model for Multimedia Services

In future cellular networks, to provide multi-class services is an essential requirement. For example, in 3G networks, four traffic classes, including conversational, streaming, interactive, and background classes, are identified. One distinguishing factor between these classes is the delay-sensitivity of the traffic [2]. The conversational and streaming classes can be grossly classified as real-time (RT) connections, while the interactive and background classes are non-real-time (NRT) packet data. In this thesis, these four traffic classes are considered in the system model and the traffic models, and parameter values in the numerical results will follow those settings in the WCDMA technology.

The objective of the thesis is to provide an analytical model to capture the practical cellular CDMA networks without resorting to any simulation tools. As modeling the NRT services with buffering and retransmission is normally based on simulations [83-85], in the thesis, only the worst-case performance for the NRT services without buffering and retransmission is studied [86]. In practical systems, NRT services should have buffering and retransmission. Some work has been done in [111]. In the thesis, the packet delay for the NRT services is therefore excluded from the system model and the packet loss for the NRT services given in the thesis will be consequently high as no retransmission is applied.

As RT services are assigned higher priorities than those of NRT services, whether there is buffering or not and whether there is retransmission or not for NRT services will not affect the RT services. Besides, as only the worst-case performance for the NRT services is studied, better performance in practical system with buffering and retransmission will be obtained compared to the results in the later numerical result sections. In other words, more NRT service users can be admitted into the system simultaneously.

Two traffic models for the services are introduced in this thesis. The voice service and the NRT services, such as web-browsing and data services, are modeled as *ON-OFF* sources with exponential *ON*/exponential *OFF* [87], as shown in Fig. 3.3. In Fig. 3.3, λ_{ON} is the transition rate from the *ON* state to the *OFF* state and μ_{OFF} is the transition rate from the *OFF* state to the *ON* state. Thus, assuming no buffering and no retransmission, the activity factor, which is the probability that the process stays in the *ON* state, is

$$\alpha = \lambda_{ON} / (\lambda_{ON} + \mu_{OFF}). \tag{3.6}$$

For the voice service, the *ON* period represents 'talking' and the *OFF* period represents the pause between speeches. For the NRT services, the traffic models developed nowadays are with heavy-tailed distribution [88-91]. As deriving the closed-form with these models is difficult without aid from computer simulations, the NRT services traffic sources are usually characterized by an activity factor that is defined as a duty cycle or the data bursty state [9, 74].

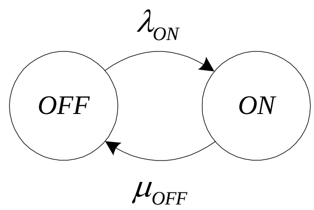

Fig. 3.3. Traffic Model with 2-State Markov Chain for an *ON-OFF* Source

Comparatively, the video service is usually treated as the variable-bit-rate (VBR) service and modeled as a classical two-dimensional discrete-state, continuous-time Markov chain [92], as shown in Fig. 3.4, which is known as Sen's model. The source changes its rate among different fixed-rate levels. The possible data rate levels are built up from integer combinations of low-bit-rate (LBR) and high-bit-rate (HBR) processes. LBRs account for scenes with slow motion, while HBRs account for scenes with fast motion. Each mobile user uses a combination of HBR and LBR spreading codes. The superscript (l) is used to represent mobile users using LBR spreading codes, while the superscript (h) is used to represent mobile users using HBR spreading codes throughout the thesis.

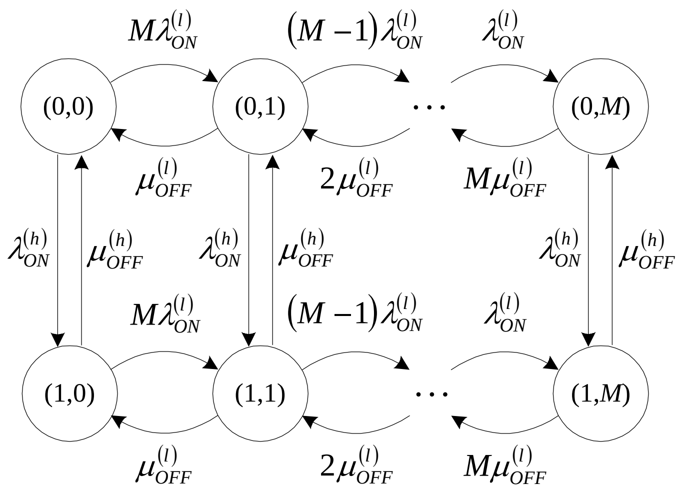

Fig. 3.4. Traffic Model with 2-Dimensional Markov Chain for a Video Source

In Fig. 3.4, each state (x, y) represents the combined discrete level of LBR and HBR that are generated by a single source. The combined data rate of each source is $xR_k^{(h)} + yR_k^{(l)}$, where $R_k^{(h)}$ is the bit rate for class k users using one HBR spreading code and $R_k^{(l)}$ is the bit rate for class k users using one LBR spreading code. Each LBR level is modeled by a two-state mini-source with an increase rate of $\lambda_{ON}^{(l)}$ and a decrease rate of $\mu_{OFF}^{(l)}$. Each HBR fluctuation is modeled by a two-state Markov chain with an increase rate of $\lambda_{ON}^{(h)}$ and a decrease rate of $\mu_{OFF}^{(h)}$. A single source can generate one active HBR mini-source and a maximum of M LBR mini-sources. The activity factors for each of the LBR and HBR mini-sources are respectively

$$\alpha^{(l)} = \lambda_{ON}^{(l)} \Big/ \left(\lambda_{ON}^{(l)} + \mu_{OFF}^{(l)}\right), \tag{3.7}$$

and

$$\alpha^{(h)} = \lambda_{ON}^{(h)} \Big/ \left(\lambda_{ON}^{(h)} + \mu_{OFF}^{(h)}\right). \tag{3.8}$$

In Chapter 5, in the analysis of the SIR and the outage probability, a multi-class traffic model containing K classes traffic using Sen's model is considered. This is a more general model compared to the 4-class traffic models and the *ON-OFF* source can be treated as the special case of Sen's model without the active HBR mini-source states and with only one two-state LBR mini-source. Furthermore, the same traffic models are assumed in both the forward and reverse links.

3.3 Mechanisms of Connection Admission Control Schemes

With the development of the cellular networks, the single service 'voice' in the previous 1G system is extended to the multiservice in the modern 3G system. Thus, the use of 'call', such as call blocking, call admission, etc., should be evolved to the use of 'connection' so that other services except for voice service can be represented. In this thesis, the concept of 'call' is equivalent to the concept of 'connection'. The multi-class traffic is divided into two groups. Within the group, users have similar characteristics, such as the delay constraint, the loss sensitivity, the traffic symmetry, the revenue, etc. Thus, group 1 and group 2 will differ from each other in terms of priorities, and within each group, the users share the same priority.

3.3.1 Complete Sharing and Virtual Partitioning

The cell site [BS] supports K classes of users that can be admitted into at most N channels. Group 1 includes traffic for classes 1 to p and group 2 includes traffic for classes p+1 to K. In FDMA or TDMA system, the concept of channel represents the frequency pair or the time slot. In CDMA systems, although there is no such practical channel, a logical channel could be presented in both time and frequency as in Fig. 1.2. To simplify the network layer analysis, the system capacity is normalized as the resource consumed by different users using some basic units. The basic units are referred to as channels in the thesis. Furthermore, the number of basic channels required by class k call, $r_k^{(h)}$ for users using HBR spreading codes and $r_k^{(l)}$ for users using LBR spreading codes, may not be an integer. Thus, the average number of basic channels required by class k call is

$$r_k = \alpha_k^{(h)} r_k^{(h)} + M_k \alpha_k^{(l)} r_k^{(l)}, \tag{3.9}$$

where M_k is the maximum number of active LBR spreading codes supported by class k traffic, $\alpha^{(l)}$ and $\alpha^{(h)}$ are the activity factors defined in (3.7) and (3.8), respectively.

N is an important system-design parameter which presents the upper bound for admitting call at the connection-level. The choice of this parameter will determine the statistical multiplexing gains and deeply influence the performance in the packet and connection levels. With a small value of N, statistical multiplexing gains are not fully utilized and system admits fewer calls below its capacity. The call blocking probability is high. With a large value of *N*, statistical multiplexing gains are over utilized and system suffers higher load beyond its capacity. The packet-loss probability grows high. The decrease of call blocking probability and the increase of packet-loss probability with increasing nominal admission bound, *N*, are shown in Fig. 3.5. The system optimization is thus presented to maximize the system utilization, N_u, (at the optimal *N*) subject to the call blocking and the packet loss constraints (B^* and L^* in the figure).

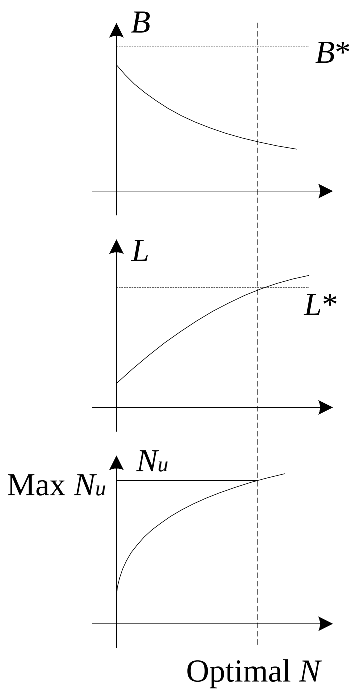

Fig. 3.5. Selection of Optimal Nominal Admission Bound

The optimal *N* is the point at which the packet loss probability equals to the packet loss probability constraint. The selection of optimal *N* is then an educated search considering the packet loss probability rather than the blocking probability with an iterative procedure. This educated search looks for the point before the packet loss probability requirement is violated. Note that the values of N are discrete. The educated search is applied in the following chapters to solve the constrained optimization problems.

Two connection admission control schemes, CS and VP, are investigated in this thesis. With CS, groups 1 and 2 share the capacity without priority assignment. Calls in either group 1 or group 2 are admitted if there are available resources, and rejected otherwise without preemption of users in another group.

VP is a type of dynamical resource allocation scheme. Two cases for VP are considered in this thesis, including VP with preemption of groups 1 and 2 (case 1) and VP with preemption of group 2 (case 2). Let n_k denote the instantaneous number of users for class k traffic and C_i, $i = 1, 2$ denote the nominal capacity for each group with $C_1 + C_2 = N$. The choice of C_i values will be related to the revenues of groups which include the economic considerations from many aspects, such as call arrival rates, traffic volumes, real-time or non-real-time property, etc.

With case 1, the total number of channels are to be shared by groups 1 and 2 using virtual partitioning. Virtual partitioning is defined as that, if $\sum_{k=1}^{p} n_k r_k \le C_1$, group 1 is in the underload state; if $\sum_{k=p+1}^{K} n_k r_k \le C_2$, group 2 is in the underload state; if $\sum_{k=1}^{p} n_k r_k > C_1$, group 1 is in the overload state, if $\sum_{k=p+1}^{K} n_k r_k > C_2$, group 2 is in the overload state. Thus, the way of implementing VP with the preemption over groups 1 and 2 is as follows. When group 1 (2) is in the underload state while group 2 (1) is in the overload state, if the arriving group 1 (2) services cannot be admitted, group 1 (2) can preempt group 2 (1) users up to the channel occupancy tending to the connection-level capacity bound *N*. Within groups, CS is applied.

With case 2, group 1 is offered guaranteed access while group 2 is offered best effort service. With the nominal capacity partitioned into C_1 and C_2, the allocation is $\hat{C}_1 = \sum_{k=1}^{p} n_k r_k = C_1$ (guaranteed access) and $\hat{C}_2 = \sum_{k=p+1}^{K} \eta_k n_k r_k \le C_2$ (best effort), where η_k is the ratio of the number of class *k* calls admitted into the cell without being preempted during their lifetimes to the total number of admitted class *k* calls. Group 1 is given hard capacity to guarantee access. This group is given higher QoS protection and can preempt group 2 calls. Unused capacity during any epoch is available to group 2.

3.3.2 Guard Capacity and Admission Rules

The dynamics of a radio cell is driven by new call requests, call terminations, and handoffs induced by user mobility. Maintaining an ongoing call is more important than admitting a new call. Hence, handoff calls should be given a higher access priority, or a lower dropping probability than the probability of blocking new calls. One way to facilitate this is to reserve capacity for admitting handoff calls, which is not accessible by new requests. The reserved capacity is referred to as guard capacity. In this thesis, the simplest reservation strategy, the fixed reservation strategy with the guard capacity set to a fixed constant value, is applied in the system.

The number of channels available for admitting new and handoff calls is $N - \sum_{k=1}^{K} n_k r_k$. Let C_G denote the number of guard channels in a cell. The admission rules are defined as:

1) Admit both new and handoff calls if $N - \sum_{k=1}^{K} n_k r_k > C_G$;

2) Reject new calls and admit only handoff calls if $0 < N - \sum_{k=1}^{K} n_k r_k \le C_G$.

The admission rules will be applied in the network analysis along with the proposed CAC schemes.

3.4 Overview of Cross-Layer Optimization

The cross-layer optimization proposed in the thesis produces gains in system utilization which will benefit both the system providers and subscribers. The gains come from exploiting the statistical behavior between the performance metrics in various layers.

In the link layer, the SIR is traditionally used as the performance metric to determine the admission region which is the largest set of QoS points delivered under any CAC schemes. Thus,

$$SIR(\vec{n}) = \frac{G \times S}{I_{\text{intra-cell}} + I_{\text{inter-cell}} + \eta} \ge SIR^*, \tag{3.10}$$

where $\vec{n}$ represents the state in the admission region, G is the spreading gain, S is the received power, $I_{\text{intra-cell}}$ and $I_{\text{inter-cell}}$ are respectively the intra-cell interference and the inter-cell interference which are determined by the number of users, the traffic models and the physical layer characteristics, such as the hysteresis margin and the path loss model, etc., η is the background noise and SIR^* is the SIR requirement. The SIRs should at all time meet the requirements in both the reverse and forward links in this design. This is a very stringent requirement especially when all the interference and noise terms are stochastic processes and the worst case conditions of all parameters need to be used in this approach.

However, the short-term degradation of SIR is tolerable as the communication can tolerate a small percent of packet loss since small loss can be compensated by retransmission in the transmitters or recovered through FEC in the receivers. The outage probability is the degradation probability that SIR below its requirement. Thus, the outage probability is selected as another performance metric to determine the admission region. In the reverse link,

$$\Pr_{\text{outage}}^{R}(\vec{n}) = \Pr\left(SIR_R(\vec{n}) \le SIR_R^*\right), \tag{3.11}$$

and in the forward link,

$$\Pr_{\text{outage}}^{F}(\vec{n}) = \Pr\left(SIR_F(\vec{n}) \le SIR_F^*\right). \tag{3.12}$$

The outage probabilities must meet their requirements in both the reverse and forward links. Thus,

$$\Pr_{\text{outage}}(\vec{n}) = \max\left(\Pr_{\text{outage}}^{R}(\vec{n}), \Pr_{\text{outage}}^{F}(\vec{n})\right) \le \delta^*, \tag{3.13}$$

where δ^* is the outage probability requirement.

However, the existence of short-term degradation in the outage probability sometimes is still tolerable. This is because, even if there are states for users suffering outage degradation larger than the requirement, the steady state probabilities, $P(\vec{n})$, of users in those states are comparatively very small. Thus, the statistical behavior of users in the network layer should also be put into consideration. In this thesis, with the consideration of the statistical behaviors across the network and link layers, the packet-loss probability is proposed as the performance metric to determine the bound of admission. Thus,

$$L = \mathcal{H}\left(\Pr_{\text{outage}}(\vec{n}),\, P(\vec{n})\right) \le L^{*}, \qquad (3.14)$$

where $\mathcal{H}(.)$ is the notation used to denote the analytical procedure to obtain the packet-loss probability taking the user steady state probabilities into consideration as presented in Chapter 6, and L^{*} is the packet-loss requirement. As $P(\vec{n})$ is correlated to the choice of CAC schemes, the bound for admitting connection at the connection-level will combine the information from the link layer and the network layer.

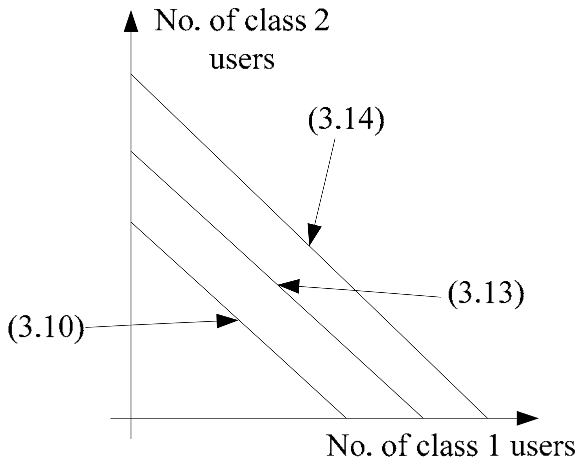

Fig. 3.6. The Performance Gains from the Cross-Layer Optimization

The performance gains from the cross-layer optimization can be illustrated using Fig. 3.6 with a line constraining a two-dimensional area presenting the bound of admission. The definition in (3.10) limits the admission region with the hard guarantee on the SIR. The definition in (3.13) improves the admission region by taking the average on the SIR. However, (3.13) still underestimates the admission region with the hard guarantee on the outage probability. With the cross-layer approach, (3.14) takes the average on the outage probability in the link layer and maximally expands the bound of admission in the network layer.

3.5 Summary

The analytical platform for cellular CDMA systems has been described in this chapter. With the consideration of the cross-layer design in the following chapters, system structure in the physical layer, power control and traffic models in the link layer, and the CAC schemes in the network layer are described and investigated. The overview on cross-layer optimization is also presented.

Based on the system models in this chapter, the analytical procedures will be presented for the network layer, the link layer and the cross-layer optimization in Chapters 4, 5 and 6, respectively. In the next chapter, based on the proposed CAC schemes, the formulae for the GoS metrics, including the new-call-blocking probability, the handoff-call-dropping probability and the system utilization at the connection-level, and the QoS metrics, including the packet-loss probability at the packet-level, will be derived. With these formulae, joint packet and connection levels optimization is investigated.

Chapter 4

Resource Allocation in the Network Layer

The goal for the network service providers is to maximize the revenue by improving system utilization, which is usually associated with minimizing the new-call-blocking probability while keeping the handoff-call-dropping probability below a certain threshold [5]. With the packet and connection levels jointly considered, the packet-loss probability should also be constrained below certain thresholds. In this chapter, the analytical model for the network layer including the connection-level and the packet-level is built. The connection admission control (CAC) schemes considered include complete sharing (CS), virtual partitioning (VP) with preemption for all classes (VP Case 1) and best effort and guarantee access with preemption for best effort traffic (VP Case 2). The contribution of the work is to propose a more general way that admits extension to multi-class traffic for VP schemes. The formulae for the GoS metrics at the connection-level and the QoS metrics at the packet-level for different CAC schemes are derived. The GoS metrics include the new-call-blocking probability, the handoff-call-dropping probability and the system utilization. The QoS metric includes the packet-loss probability. The packet delay is ignored in this thesis under the assumption that there is no buffering and retransmission. With these formulae, joint packet and connection levels optimization is presented.

In this chapter, issues in the physical and link layers are all represented by a defined physical capacity of C channels per cell. C is the constraint of the number of instantaneously transmitted packets of admitted calls. The characteristics of wireless system in the physical layer and in the link layer, such as interference, soft handoff, power control, SIR, and the outage probability, etc., are all represented by the physical capacity, C. Thus, the network layer analysis can be simplified to build on this physical capacity and the cross-layer model will update the physical capacity with ensuing results in the link layer analysis in Chapter 5.

As stated before, connection-level performance measures such as blocking and dropping probabilities are known as GoS. For convenience of representation, QoS is used to represent connection-level as well as packet-level performance measures in the remaining part of the thesis.

4.1 Problem Statement

4.1.1 Connection-Level and Packet-Level Parameters

Consider a class k user. Its QoS is specified by the new-call-blocking probability, B_{nk}, the handoff-call-dropping probability, B_{hk}, and the system utilization, N_{uk}, at the connection-level and the packet-loss probability, L_k, at the packet-level. At the BS, the estimation of new call arrival rate can be done using a jumping window or a moving window method. With generally accepted approach, the arrivals of calls are assumed as Poisson distributed and the call holding time and the call dwell time are exponentially distributed. Although these assumptions may not be true in practical mobile networks, they have been widely used [41, 42, 44, 93, and 94] to provide approximate solutions for cellular systems.

The connection-level and the packet-level parameters used throughout the paper are listed below. The connection-level parameters are listed as following:

B_{nk} $= \dfrac{\text{Number of new class } k \text{ calls blocked}}{\text{Total number of call arrivals}}$: new-call-blocking probability for class k;

B_n $= \sum_{k=1}^{K} B_{nk}$: total new-call-blocking probability;

B_{hk} $= \dfrac{\text{Number of handoff class } k \text{ calls dropped}}{\text{Total number of call arrivals}}$: handoff-call-dropping probability for class k;

B_h $= \sum_{k=1}^{K} B_{hk}$: total handoff-call-dropping probability;

N_{uk} $= \dfrac{\int_T n_k r_k dt}{T}$: system utilization for class k calls;

N_u $= \sum_{k=1}^{K} N_{uk}$: total system utilization (average channel occupation, $N_u < C$);

λ_{nk} arrival rate of new class k calls;

λ_{hk} arrival rate of class k handoff calls;

λ_k $= \lambda_{nk} + \lambda_{hk}$: total arrival rate of class k calls;

μ_{ck}^{-1} mean class k call holding time (lifetime);

μ_{hk}^{-1} mean class k call dwell time (interhandoff time).

The packet-level parameters are list as following:

L_k packet-loss probability for class k call;

L $=\sum_{k=1}^{K} L_k$: total packet-loss probability.

4.1.2 Formulae on Arrival Rates

Let θ_{nk} denote the probability that the next arrival is a new call from class k:

$$\theta_{nk} = \frac{\lambda_{nk}}{\sum_{k=1}^{K} \lambda_k}. \tag{4.1}$$

Let θ_{hk} denote the probability that the next arrival is a handoff call from class k:

$$\theta_{hk} = \frac{\lambda_{hk}}{\sum_{k=1}^{K} \lambda_k}. \tag{4.2}$$

Let θ_k denote the probability that the next arrival is from class k:

$$\theta_k = \frac{\lambda_k}{\sum_{k=1}^{K} \lambda_k}. \tag{4.3}$$

All cells are assumed to be statistically identical. Thus, the rate of handoff departing from a cell equals the rate at which handoff calls entering the cell. From the complete partitioning policy, equating the class k handoff call arrival rate to the product of the average handoff rate, μ_{hk}, and the average number of class k calls, the class k handoff call arrival rate is obtained as follows [51, 95]:

$$\lambda_{hk} = \mu_{hk} \times \frac{N_{uk}}{r_k} = \mu_{hk} \times \frac{(1-b_{nk})\lambda_{nk} + (1-b_{hk})\lambda_{hk}}{\mu_{ck} + \mu_{hk}}, \tag{4.4a}$$

Solving for λ_{hk} in (4.4a),

$$\lambda_{hk} = \frac{\mu_{hk}(1-b_{nk})\lambda_{nk}}{\mu_{ck} + b_{hk}\mu_{ck}} = \frac{\mu_{hk}(1-B_{nk}/\theta_{nk})\lambda_{nk}}{\mu_{ck} + (B_{hk}/\theta_{hk})\mu_{ck}}, \tag{4.4b}$$

where $b_{nk} = B_{nk} / \theta_{nk}$ is the new class k call blocking probability considering only class k traffic, $b_{hk} = B_{hk} / \theta_{hk}$ is the handoff class k call dropping probability considering only class k traffic, $\mu_{hk} = v_k / s$, v_k is the speed of the class k mobile user and s is the size of a square cell. More details on the demonstration of (4.4) could be looked up from [95] for a single traffic system and [51] for a multi-traffic system.

Note that both the new-call-blocking and handoff-call-dropping probabilities are functions of the new call arrival rate and the handoff call arrival rate, respectively. The handoff-call-dropping probability can only be calculated after the handoff call arrival rate has been determined. On the other hand, the handoff call arrival rate can only be calculated using (4.4b) when B_{nk} and B_{hk} are known. To resolve this paradox, (4.4b) can be written as a set of recurrence equations. At the initial time instant $l = 1$, the handoff call arrival rate, $\lambda_{hk}(1)$ is set equal to $\mu_{hk}\lambda_{nk} / \mu_{ck}$ on the assumption that $b_{hk}(1)$ and $b_{nk}(1)$ are negligibly small. Then $\lambda_{hk}(1)$ is used to compute $b_{hk}(2)$ and $b_{nk}(2)$ using the analytical results presented in sections 4.2, 4.3 and 4.4. For example, for CS, (4.8) is used to solve for the equilibrium system state probability $P(\vec{n})$. $P(\vec{n})$ is then used in (4.11)-(4.12) to evaluate the new-call-blocking and handoff-call-dropping probabilities which in turn are substituted into (4.4b). Iterating in this manner, the recurrence equations can be written as

$$\lambda_{hk}(l) = \frac{\mu_{hk}(1 - b_{nk}(l))\lambda_{nk}}{\mu_{ck} + b_{hk}(l)\mu_{ck}}, \tag{4.5a}$$

$$b_{hk}(l+1) = \mathscr{F}(\lambda_{hk}(l)), \tag{4.5b}$$

$$b_{nk}(l+1) = \mathscr{G}(\lambda_{hk}(l)), \tag{4.5c}$$

where $\mathscr{F}(.)$ and $\mathscr{G}(.)$ are notations used to denote the analytical procedure as described above.

4.2 CS - Complete Sharing

CS is the simplest CAC scheme which has almost no rules for admitting calls. However, the admission policy must be considered as the guard capacity is proposed to protect the handoff calls.

4.2.1 CS – Connection-Level Analysis

Assuming Poisson distributed arrivals for new and handoff calls and exponentially distributed call holding time, the channel occupancy can be modeled as a *K*-dimensional Markov chain and solve it using the techniques in [55, 96, and 97]. The key of the techniques is to formulate the global balance equations based on a *K*-dimensional Markov chain.

Let $\vec{n} = \left(n_1, n_2, \ldots, n_K\right)$ denote the state of the system with the number of class *k* users, n_k, in the *k*th class and $\vec{r} = \left(r_1, r_2, \ldots, r_K\right)$ denote the vector of basic channels, where r_k is the number of basic channels required for each class *k* call and $r_k = \alpha_k^{(h)} r_k^{(h)} + M_k \alpha_k^{(l)} r_k^{(l)}$. Let $\lambda_k\left(\vec{n}\right)$ denote the arrival rate and $\mu_k\left(\vec{n}\right)$ the departure rate for class *k* calls in the system. Thus the state space of the system, denoted by *S*, is given by $S := \left\{ \vec{n} \mid \vec{r} \cdot \vec{n} \le N \right\}$.

The admission policy for CS will be defined here. When the system is in state $\vec{n}$ and a class *k* call (new or handoff) arrives, an admission policy determines whether or not the call is admitted into the system. The admission policy is specified by mapping $\vec{f}(\vec{n}) = (f_1(\vec{n}), f_2(\vec{n}), \ldots, f_K(\vec{n}))$ for both the new and handoff calls excluding the guard channels, and $\vec{f}_G(\vec{n}) = (f_{G1}(\vec{n}), f_{G2}(\vec{n}), \ldots, f_{GK}(\vec{n}))$ for only the handoff calls within the guard channels, where the union of $S(\vec{f}(\vec{n}))$ and $S(\vec{f}_G(\vec{n}))$ gives the system state space *S*, and $f_k(\vec{n})$ and $f_{Gk}(\vec{n})$ each takes on the value 0 or 1 if a class *k* call is rejected or admitted, respectively, when the system state is $\vec{n}$. They are defined by the following equations:

$$f_k(\vec{n}) = \begin{cases} 1, & \vec{r} \cdot \vec{n} + r_k \le N - C_G \\ 0, & \text{otherwise} \end{cases}, \tag{4.6}$$

and

$$f_{Gk}(\vec{n}) = \begin{cases} 1, & N - C_G < \vec{r} \cdot \vec{n} + r_k \le N \\ 0, & \text{otherwise} \end{cases}. \tag{4.7}$$

The global balance equations for the Markov process under the admission policy for CS are given by:

$$\begin{aligned} \sum_{k=1}^{K} \{\lambda_k(\vec{n})[f_k(\vec{n}) + f_{Gk}(\vec{n})] + \mu_k(\vec{n})\} P(\vec{n}) &= \sum_{k=1}^{K} P(\vec{n} + \vec{e}_k)\mu_k(\vec{n} + \vec{e}_k) \\ &+ \sum_{k=1}^{K} P(\vec{n} - \vec{e}_k)\lambda_k(\vec{n} - \vec{e}_k)[f_k(\vec{n} - \vec{e}_k) + f_{Gk}(\vec{n} - \vec{e}_k)] \end{aligned}, \tag{4.8}$$

where $\vec{e}_k$ is a *K*-dimensional vector of all zeros except for a one in the *k*th place, i.e., $\vec{n} + \vec{e}_k$ means to admit a class *k* call and $\vec{n} - \vec{e}_k$ means to terminate a class *k* call who is preempted by another call which will happen in VP or has just completed its communication, and

$$\lambda_k(\vec{n}) = \begin{cases} \lambda_{nk} + \lambda_{hk}, & \text{if} \quad f_k(\vec{n}) = 1 \\ \lambda_{hk}, & \text{if} \quad f_{Gk}(\vec{n}) = 1 \end{cases}, \tag{4.9}$$

$$\mu_k(\vec{n}) = n_k(\mu_{ck} + \mu_{hk}), \; 0 < \vec{r} \cdot \vec{n} \le N. \tag{4.10}$$

Equation (4.8) can be solved using lower triangular/upper triangular (LU) decomposition together with the condition that the sum of all the state probabilities is 1 to obtain $P(\vec{n})$. LU decomposition is a common numerical technique to solve linear algebraic equations [98].

A new class *k* call is blocked from entering the system (and is assumed lost) if upon arrival it finds that it cannot be accommodated because the number of available channels (excluding the guard channels) is less than r_k. Therefore the blocking probability for a new class *k* call considering all classes is given by

$$B_{nk} = \sum_{n_1=0}^{\lfloor N/r_1 \rfloor} \sum_{n_2=0}^{\lfloor (N-n_1 r_1)/r_2 \rfloor} \cdots \sum_{n_K=0}^{\left\lfloor \left(N - \sum_{i=1}^{K-1} n_i r_i \right) \big/ r_k \right\rfloor} P(\vec{n}) \theta_{nk} \,, \tag{4.11}$$

for all $\vec{n}$ satisfying $(N - C_G - \vec{r} \cdot \vec{n}) < r_k$. $\lfloor x \rfloor$ is the floor function which gives the largest integer less than or equal to *x*.

A class *k* handoff call is blocked from entering the system (and is assumed lost) if upon arrival it finds that it cannot be accommodated because the number of available channels (including the guard channels) is less than r_k. Therefore the dropping probability for a class *k* handoff call considering all classes is given by

$$B_{hk} = \sum_{n_1=0}^{\lfloor N/r_1 \rfloor} \sum_{n_2=0}^{\lfloor (N-n_1 r_1)/r_2 \rfloor} \cdots \sum_{n_K=0}^{\left\lfloor \left(N - \sum_{i=1}^{K-1} n_i r_i \right) \big/ r_k \right\rfloor} P(\vec{n}) \theta_{hk} \,, \tag{4.12}$$

for all $\vec{n}$ satisfying $(N - \vec{r} \cdot \vec{n}) < r_k$.

The utilization for class *k* is given by:

$$N_{uk} = \sum_{n_1=0}^{\lfloor N/r_1 \rfloor} \sum_{n_2=0}^{\lfloor (N-n_1 r_1)/r_2 \rfloor} \cdots \sum_{n_K=0}^{\left\lfloor \left(N - \sum_{i=1}^{K-1} n_i r_i \right) \big/ r_K \right\rfloor} n_k P(\vec{n}) r_k \,. \tag{4.13}$$

The total utilization is thus,

$$N_u = \sum_{k=1}^{K} N_{uk} \ . \tag{4.14}$$

In the thesis, the utilization represents the average number of channels. The utilization is not normalized. Thus, the value can be greater than 1.

4.2.2 CS – Packet-Level Analysis

The statistical multiplexing is modeled at the packet-level. The exponentially distributed *ON*/*OFF* mini-sources in Sen's model for each call have been assumed. For mobile users using HBR spreading codes, let h_k denote the number of in-progress mini-sources. The probability that h_k in-progress mini-sources are in the *ON* state given that there are n_k calls in progress is given by

$$P_h\left(h_k \mid n_k\right) = \binom{n_k}{h_k}\left(\alpha_k^{(h)}\right)^{h_k}\left(1-\alpha_k^{(h)}\right)^{n_k-h_k} . \tag{4.15}$$

For mobile users using LBR spreading codes, let l_k denote the number of in-progress mini-sources. The probability that l_k in-progress mini-sources are in the *ON* state given that there are n_k calls in progress is given by

$$P_l\left(l_k \mid n_k\right) = \binom{M_k n_k}{l_k}\left(\alpha_k^{(l)}\right)^{l_k}\left(1-\alpha_k^{(l)}\right)^{M_k n_k-l_k} , \tag{4.16}$$

where M_k is the maximum number of active LBR spreading codes supported by class *k* traffic.

Because the design parameter *N* must not be less than the physical capacity *C*, when the total number of channels from calls in the *ON* state exceeds the physical capacity, they will suffer packet losses. Assuming a priority structure with class 1 having the highest priority and no packet buffer, the equivalent class *k* packet-loss probability is given by [51]:

$$L_k = \frac{\displaystyle\sum_{n_1=0}^{\lfloor N/r_1\rfloor}\sum_{n_2=0}^{\lfloor (N-n_1r_1)/r_2\rfloor}\cdots\sum_{n_K=0}^{\left\lfloor\left(N-\sum_{i=1}^{K-1}n_ir_i\right)/r_K\right\rfloor} P(\vec{n})\sum_{h_1=0}^{n_1}\sum_{h_2=0}^{n_2}\cdots\sum_{h_K=0}^{n_K}P_h(h_1\mid n_1)P_h(h_2\mid n_2)\cdots P_h(h_K\mid n_K) \times\sum_{l_1=0}^{M_1n_1}\sum_{l_2=0}^{M_2n_2}\cdots\sum_{l_K=0}^{M_Kn_K}P_l(l_1\mid n_1)P_l(l_2\mid n_2)\cdots P_l(l_K\mid n_K) \times\begin{cases}\left[\sum_{i=1}^{k}\left(h_ir_i^{(h)}+l_ir_i^{(l)}\right)-C\right]^+, & \text{if } \sum_{i=1}^{k-1}\left(h_ir_i^{(h)}+l_ir_i^{(l)}\right)\le C\\ h_kr_k^{(h)}+l_kr_k^{(l)}, & \text{if } \sum_{i=1}^{k-1}\left(h_ir_i^{(h)}+l_ir_i^{(l)}\right)> C\end{cases}}{L_{sum}},$$

(4.17)

where

$$L_{sum} = \sum_{n_1=0}^{\lfloor N/r_1\rfloor}\sum_{n_2=0}^{\lfloor (N-n_1r_1)/r_2\rfloor}\cdots\sum_{n_K=0}^{\left\lfloor\left(N-\sum_{i=1}^{K-1}n_ir_i\right)/r_K\right\rfloor} P(\vec{n})\sum_{h_1=0}^{n_1}\sum_{h_2=0}^{n_2}\cdots\sum_{h_K=0}^{n_K}P_h(h_1\mid n_1)P_h(h_2\mid n_2)\cdots P_h(h_K\mid n_K) \times\sum_{l_1=0}^{M_1n_1}\sum_{l_2=0}^{M_2n_2}\cdots\sum_{l_K=0}^{M_Kn_K}P_l(l_1\mid n_1)P_l(l_2\mid n_2)\cdots P_l(l_K\mid n_K)\times\sum_{i=1}^{K}\left(h_ir_i^{(h)}+l_ir_i^{(l)}\right),$$

(4.18)

$[x]^+ = \max[x,0]$, $r_k^{(h)}$ denotes the channel number for class k users using HBR spreading codes and $r_k^{(l)}$ denotes the channel number for class k users using LBR spreading codes.

The total packet-loss probability is then given by

$$L = \sum_{k=1}^{K} L_k \ . \tag{4.19}$$

4.2.3 CS - Performance Evaluation

The performance for CS by means of numerical analysis and simulation is evaluated. The Manhattan model used in the simulation is shown in Fig. 4.1. The block length is 200m with 100 BSs, the filled circle, placed at every street corner. The dashed line represents the curbside and the solid line represents the cell boundary.

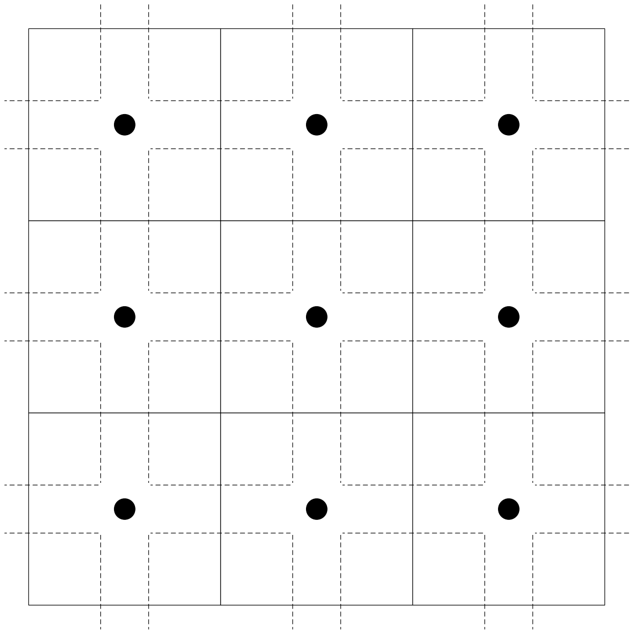

Fig. 4.1. The Manhattan Model in the Simulation

4.2.3.1 Simulation Model

The simulation is based on an "event-oriented" approach: a sequence of events corresponding to the actions that modify the state of the system is processed. The events include a new call arrival, a handoff call arrival, a call departure and a new system utilization updating. The resource allocation schemes are negotiated in each BS when a new call or a handoff call arrives. The assumptions of Poisson distributed arrivals for new calls and exponentially distributed call holding time are used. The handoff call arrival rate is not determined by (4.4) or (4.5), but based on the mobility of the calls. Performance statistics are obtained from the middle cell in the system. By collecting the number of arriving calls, new blocked calls, and handoff dropped calls and the system utilization per time unit, the blocking probability and the system utilization can be computed as described in section 4.1.1.

Furthermore, when a new call is generated, it can move in any one of the four directions: north, south, east or west. The movement of the mobile users is restricted along the streets, and the birth positions of the new calls are randomly selected. Handoff occurs only when the residual call holding time is longer than the time duration that a mobile user will travel in a cell. So, the residual dwell time is determined by the birth position of the new calls, the moving directions and the speed of the mobile users. In the simulation, the mobile users travel in one of the directions at a constant speed. Thus, excluding the cases for the dwell times in the birth cell and the ending cell of a call, the dwell times of a call in the cells between the birth cell and the ending cell are fixed constant times. That is, these dwell times are equal to the block length divided by the speed of the mobile user shown in (4.4). Thus, the handoff call arrival rate is determined by the speed of the mobile users and further influences the blocking probability and system utilization. The performances of fast and slow mobile users, representing users in vehicles and pedestrian users, respectively, are compared.

The performance metrics include the new-call-blocking probability, the handoff-call-dropping probability and the system utilization for different traffic services with fast and slow mobile users. The purpose of the simulation is to justify the reasonableness for the formulae in the analysis.

Each simulation result is generated up to 100 million simulation minutes. A warm-up period of 10 million simulation minutes has also been used to minimize the effects of initial simulation transients. Four traffic classes (K=4) are considered. The first, third and fourth classes are modeled as *ON-OFF* sources that are Sen's model without the active HBR mini-source states and with only one two-state LBR mini-source each representing the voice, web-browsing and data services, while the second class is modeled as Sen's model with the maximum number of active LBR spreading codes, $M_2 = 8$, representing the video service.

The settings of the traffic models are following the definition in 3GPP. According to [2, 99], the parameters used in the numerical examples are tabulated in Table 4.1.

Table 4.1. The Parameters Values in the Network Layer Analysis

Class	1	2	3	4
Parameter Type	Voice	Video	Web-browsing	Data
Activity Factor (α_k)	0.4	0.3867 (LBR) 0.5 (HBR)	0.1176	0.1028
Channel Number (r_k)	2	1 (LBR) 2 (HBR)	8	16
Holding Time ($1/\mu_{ck}$)	1 min	1 min	3 mins	3 mins
Slow User Speed (v_k)	4.0 km/hr	4.0 km/hr	4.0 km/hr	4.0 km/hr
Fast User Speed (v_k)	40 km/hr	40 km/hr	40 km/hr	40 km/hr
Nominal Capacity (N)		16		
Physical Capacity (C)		12		
Guard Capacity (C_G)		4		
Nominal Capacity for Group 1 (C_1)		12		
Nominal Capacity for Group 2 (C_2)		4		
Cell Size (s)		200m		

4.2.3.2 Connection-Level Blocking Probability and System Utilizations for CS

The connection-level analytical and simulation results of the new-call-blocking and handoff-call-dropping probabilities, B_n and B_h, for classes 1 and 2, and 3 and 4 with fast and slow mobile users are shown in Figs. 4.2, 4.3, 4.4 and 4.5, respectively. The new call arrival rates are all the same for the 4 traffic classes.

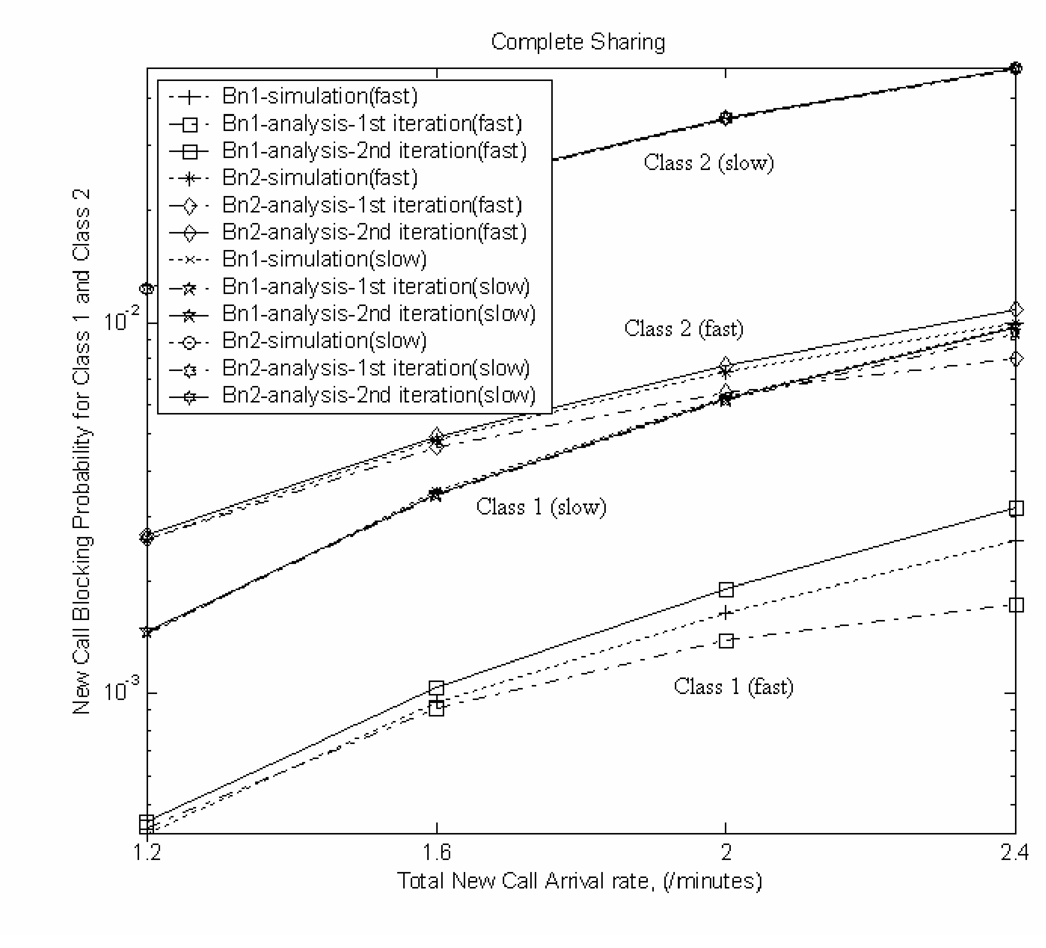

Fig. 4.2. New-call-blocking probabilities for classes 1 and 2 (CS)

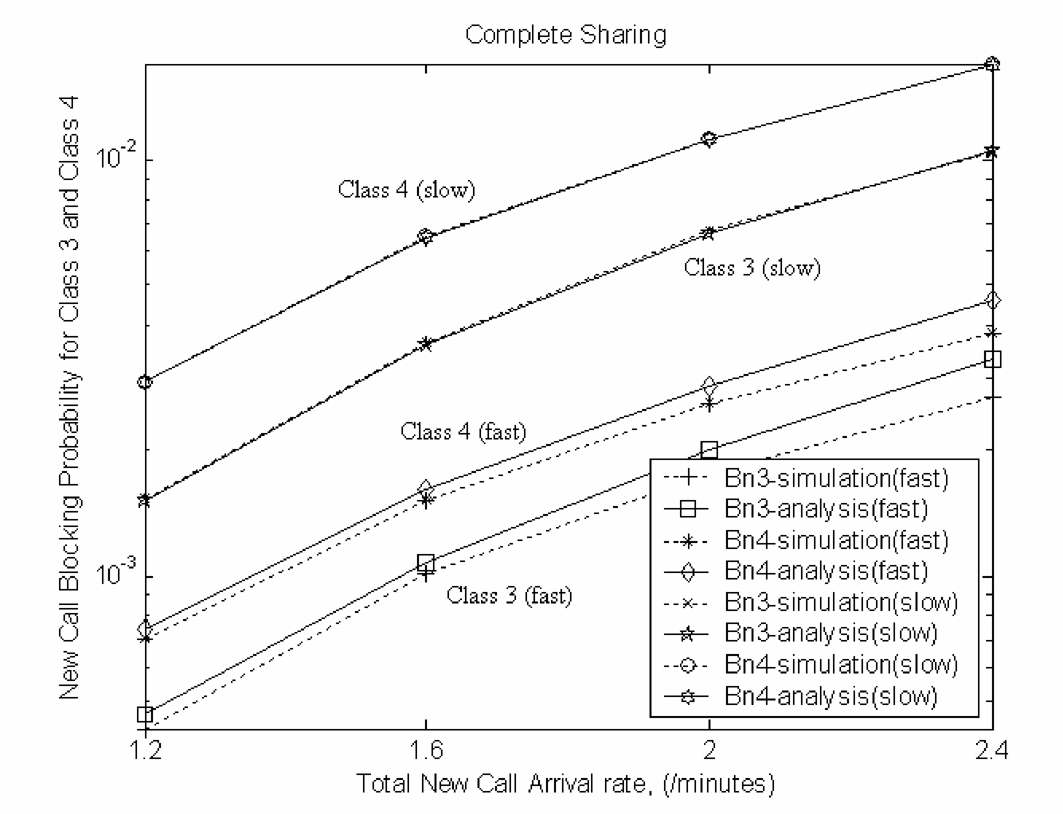

Fig. 4.3. New-call-blocking probabilities for classes 3 and 4 (CS)

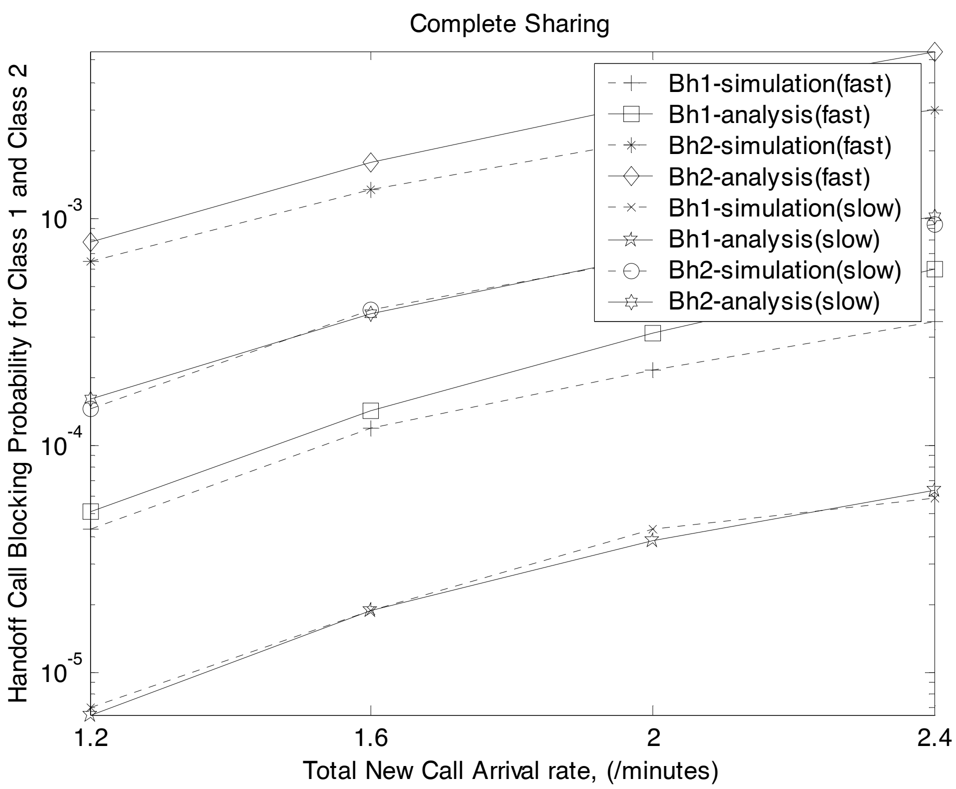

Fig. 4.4. Handoff-call-dropping probabilities for classes 1 and 2 (CS)

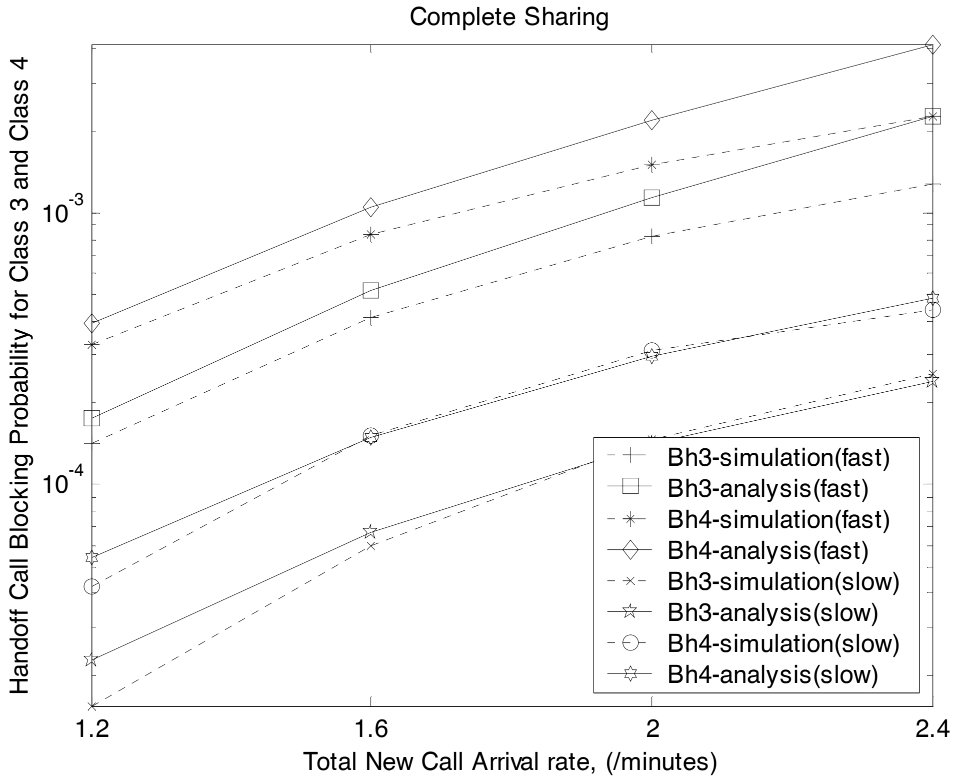

Fig. 4.5. Handoff-call-dropping probabilities for classes 3 and 4 (CS)

The agreement between the analytical results and simulation results are excellent. Although this is expected as the simulations use the same model and assumptions, it does verify the reasonableness for the formulae and showed that two iterations of (4.5) are sufficient to obtain quite accurate results. The effect of the iterations is shown in Fig. 4.2. Under the same total traffic load input (the total new call arrival rate) in both the fast and slow users' cases, the fast users' case will lead to more handoff processes than that for the slow users' case as the dwell time for the fast user is shorter than that for the slow user. Therefore, the chance of the handoff call dropping is higher in the fast users' case than that in the slow users' case. That is, the handoff-call-dropping probabilities in the fast users' case are larger than those in the slow users' case (as shown in Figs. 4.4 and 4.5). On the other hand, the larger handoff-call-dropping probabilities in the fast users' case bring more available capacity for the new arrival calls compared to that in the slow users' case. As a result, the new-call-blocking probabilities in the fast users' case are lower than those in the slow users' case (as shown in Figs. 4.2 and 4.3). The numerical results have been obtained using only two iterations in the calculation of the handoff arrival rate. The analytical and simulation results agree better in the slow users' case than those in the fast users' case. This is because the handoff call arrival rate in the slow users' case is comparatively smaller than that in the fast users' case. Thus, the estimation error of the handoff call arrival rate in (4.4) due to limited iterative times will have little effect on the performance evaluation in the slow users' case.

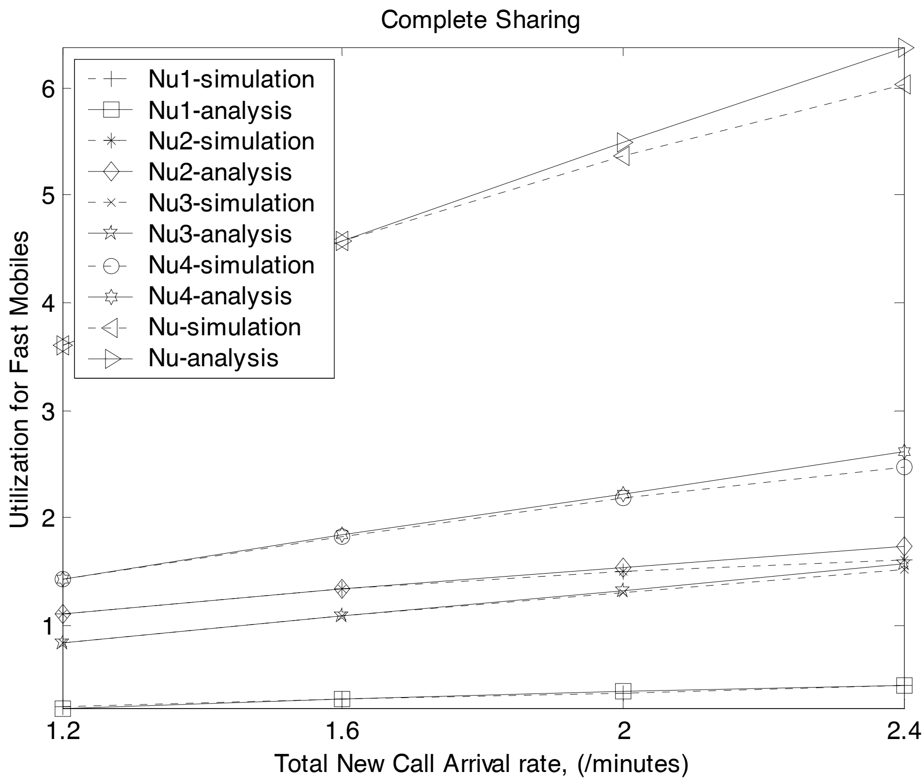

Fig. 4.6. System utilization for fast mobile users (CS)

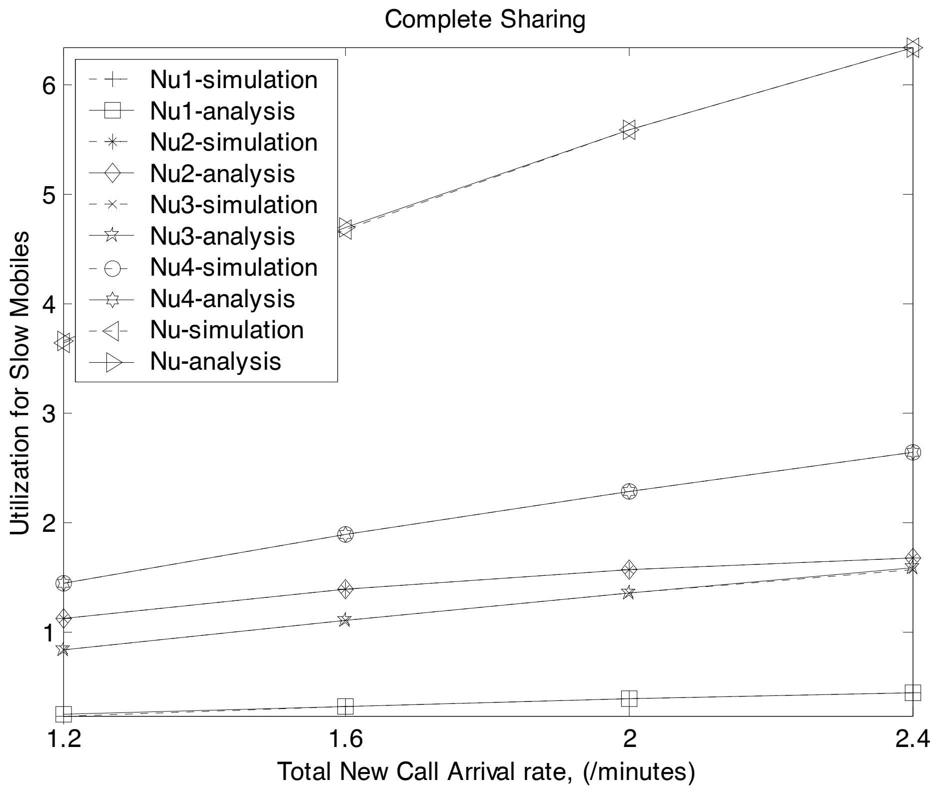

Fig. 4.7. System utilization for slow mobile users (CS)

Figs. 4.6 and 4.7 show the connection-level analytical and simulation results of system utilization, N_u, for fast and slow mobile users, respectively. Similarly, the analytical and simulation results agree better in the slow users' case than those in the fast users' case.

4.3 VP Case 1 - VP with Preemption for Groups 1 and 2

The difference between CS and VP lies on the preemptions in VP. The global balance equations for CS have been obtained. Focusing on the characteristic of VP, five rules to present the preemptions are defined. The five preemption rules will describe all characteristics of preemptions that are possible in VP. They will be used as the criteria to determine whether or not the preemption can happen in a particular state of the system. Thus, there are the admission policies and the preemption rules to determine the admission, the rejection and the preemption decisions for the new and handoff calls.

4.3.1 Case 1 – Connection-Level Analysis

The multi-class users are divided into two groups. Group 1 ($G_1, \vec{n}_1$) contains classes from 1 to p with the nominal capacity of C_1; group 2 ($G_2, \vec{n}_2$) contains classes from p+1 to K with the nominal capacity of C_2.

In mathematical terms, the admission policies for case 1 are defined as

$$f_k(\vec{n}) = \begin{cases} 1, & \vec{r} \cdot \vec{n} + r_k \le N - C_G \\ 0, & \text{otherwise} \end{cases}, \tag{4.20}$$

and

$$f_{Gk}(\vec{n}) = \begin{cases} 1, & N - C_G < \vec{r} \cdot \vec{n} + r_k \le N \\ 0, & \text{otherwise} \end{cases}, \tag{4.21}$$

which are the same as those for CS.

Preemption in case 1 is governed by the following preemption rules:

1. Preemption could happen only under the condition that $\sum_{k=p+1}^{K} n_k r_k > C_2$, or, $\sum_{k=1}^{p} n_k r_k > C_1$, i.e., group 2 users occupying capacity nominally allocated to group 1; group 1 users occupying capacity nominally allocated to group 2.
2. Preemption happens when a new class j call arrives under the conditions that $C_G + \vec{r} \cdot \vec{n} = N$, or {$C_G + \vec{r} \cdot \vec{n} < N$ and $C_G + \vec{r} \cdot \vec{n} > N - r_j$}. Note that no matter which type of preemption happens, new calls can only be admitted into the system when the guard channels are not occupied.

3. Preemption happens when a handoff class j call arrives under the condition that $\vec{r} \cdot \vec{n} > N - r_j$.

4. When there is preemption for class $j \in G_1$ over group 2 (class j adds one user and one/some user(s) is/are deleted from group 2), the number of terminated users from each class within group 2 must be calculated individually. Here the problem is simplified to assume that the termination happens to only one class, e.g., class k. So, the number of class k users terminated from the system will be less than or equal to $\left\lceil \frac{r_j}{r_k} \right\rceil$, where $\lceil x \rceil$ is the ceiling function which gives the smallest integer larger than or equal to x. Likewise, when there is preemption for class $k \in G_2$ over class j of group 1, the number of class j users terminated from the system will be less than or equal to $\left\lceil \frac{r_k}{r_j} \right\rceil$.

5. If preemption happens over class k, it means that class k is the most overloaded class in its own group and the class k will be determined by the following criteria: $k = \arg\min\left\{\left(C_1 - n_1 \times r_1\right), \ldots, \left(C_1 - n_p \times r_p\right)\right\}$ for $k \in G_1$, and $k = \arg\min\left\{\left(C_2 - n_{p+1} \times r_{p+1}\right), \ldots, \left(C_2 - n_K \times r_K\right)\right\}$ for $k \in G_2$. Another criterion could be to preempt the lowest priority class in that group and this is a subject for future investigation.

Due to the fact that the only difference between CS and VP is the preemption in VP, to formulate the global balance equations for VP, the global balance equations for CS will be formulated first, and then each state will be filtered by the five preemption rules. If preemption happens in that state, the corresponding state transitions will be added to complete the global balance equations. Let $P(\vec{n})$ denote the equilibrium probability that the system is in state $\vec{n}$. The global balance equations for the Markov process under the preemption rules and the admission policy for case 1 are given by:

$$
\begin{aligned}
&\sum_{k=1}^{K}\{\lambda_k(\vec{n})[f_k(\vec{n})+f_{Gk}(\vec{n})]+\lambda_{Pk}(\vec{n})+\mu_k(\vec{n})\}P(\vec{n}) \\
&=\sum_{k=1}^{K}P(\vec{n}-\vec{e}_k)\lambda_k(\vec{n}-\vec{e}_k)[f_k(\vec{n}-\vec{e}_k)+f_{Gk}(\vec{n}-\vec{e}_k)]+\sum_{k=1}^{K}P(\vec{n}+\vec{e}_k)\mu_k(\vec{n}+\vec{e}_k)\ , \qquad (4.22)\\
&+\sum_{k=1}^{K}\sum_{j=1,\,j\neq k}^{K}\left[\sum_{i=1}^{\lceil r_j/r_k\rceil}P(\vec{n}+i\vec{e}_k-\vec{e}_j)\lambda_{Pk}(\vec{n}+i\vec{e}_k-\vec{e}_j)\right]
\end{aligned}
$$

where $\lambda_k(\vec{n})$ is defined in (4.9), $\mu_k(\vec{n})$ is defined in (4.10), and $\lambda_{Pk}(\vec{n})$ is defined as follows:

$$
\lambda_{Pk}(\vec{n})=\begin{cases}\lambda_{nk}, & \text{by rule 2}\\ \lambda_{hk}, & \text{by rule 3}\end{cases}. \qquad (4.23)
$$

Without the $\lambda_{Pk}(\vec{n})$ terms, (4.22) is equivalent to the global balance equations for CS.

A new class *k* call is blocked from entering the system (and is assumed lost) if upon arrival it finds that it cannot be accommodated because the number of available channels (excluding the guard channels) is less than r_k. Therefore the blocking probability for a new class *k* call considering all classes is given by

$$
B_{nk}=\sum_{n_1=0}^{\lfloor N/r_1\rfloor}\sum_{n_2=0}^{\lfloor (N-n_1r_1)/r_2\rfloor}\cdots\sum_{n_K=0}^{\left\lfloor\left(N-\sum_{i=1}^{K-1}n_ir_i\right)\Big/r_k\right\rfloor}P(\vec{n})\theta_{nk}-\sum_{j\text{ in the other group}}P_{nj}\ , \qquad (4.24)
$$

for all $\vec{n}$ satisfying $(N-C_G-\vec{r}\cdot\vec{n})<r_k$, where P_{nj} is the preemption probability for class *j* call when a new class *k* call arrives, as defined in (4.26) below.

A class *k* handoff call is blocked from entering the system (and is assumed lost) if upon arrival it finds that it cannot be accommodated because the number of available channels (including the guard channels) is less than r_k. Therefore the dropping probability for a class *k* handoff call considering all classes is given by

$$B_{hk} = \sum_{n_1=0}^{\lfloor N/r_1 \rfloor} \sum_{n_2=0}^{\lfloor (N-n_1 r_1)/r_2 \rfloor} \cdots \sum_{n_K=0}^{\left\lfloor \left(N - \sum_{i=1}^{K-1} n_i r_i \right) / r_k \right\rfloor} P(\vec{n})\theta_{hk} - \sum_{j \text{ in the other group}} P_{hj} \quad , \tag{4.25}$$

for all $\vec{n}$ satisfying $\left(N - \vec{r} \cdot \vec{n}\right) < r_k$, where P_{hj} is the preemption probability for class *j* call when a handoff class *k* call arrives, as defined in (4.27) below.

The preemption probability for a class *k* call, when a class *j* new call arrives, is given by:

$$P_{nk} = \sum_{n_1=0}^{\lfloor N/r_1 \rfloor} \sum_{n_2=0}^{\lfloor (N-n_1 r_1)/r_2 \rfloor} \cdots \sum_{n_K=0}^{\left\lfloor \left(N - \sum_{i=1}^{K-1} n_i r_i \right) / r_k \right\rfloor} P(\vec{n})\theta_{nj} \ , \tag{4.26}$$

for all $\vec{n}$ satisfying rules 2 and 4.

The preemption probability for a class *k* call, when a class *j* handoff call arrives, is given by:

$$P_{hk} = \sum_{n_1=0}^{\lfloor N/r_1 \rfloor} \sum_{n_2=0}^{\lfloor (N-n_1 r_1)/r_2 \rfloor} \cdots \sum_{n_K=0}^{\left\lfloor \left(N - \sum_{i=1}^{K-1} n_i r_i \right) / r_k \right\rfloor} P(\vec{n})\theta_{hj} \ , \tag{4.27}$$

for all $\vec{n}$ satisfying rules 3 and 4.

The utilization for class *k* is given by:

$$N_{uk} = \sum_{n_1=0}^{\lfloor N/r_1 \rfloor} \sum_{n_2=0}^{\lfloor (N-n_1 r_1)/r_2 \rfloor} \cdots \sum_{n_K=0}^{\left\lfloor \left(N - \sum_{i=1}^{K-1} n_i r_i \right) / r_K \right\rfloor} n_k P(\vec{n}) r_k \ . \tag{4.28}$$

The total utilization is thus,

$$N_u = \sum_{k=1}^{K} N_{uk} \ . \tag{4.29}$$

4.3.2 Case 1 - Packet-Level Analysis

The analytical formulations of the packet-loss probability for case 1 are the same as those for CS in section 4.2.2.

4.3.3 Case 1 - Performance Evaluation

The performance for case 1 by means of numerical analysis and simulation is evaluated. The simulation model and the parameters used in the numerical examples are all the same as those in CS. The new call arrival rates are all the same for the 4 traffic classes. The connection-level analytical and simulation results of new-call-blocking and handoff-call-dropping probabilities, B_n and B_h, for classes 1 and 2, and 3 and 4 for fast and slow users are shown in Figs. 4.8, 4.9, 4.10 and 4.11, respectively.

The agreement between the analytical results and simulation results are excellent. This justifies the reasonableness of the defined preemption rules. VP case 1 assigns different priorities for group 1 and 2 users and group 1 users have higher priorities than group 2 users. Thus, the new-call-blocking and handoff-call-dropping probabilities of group 1 users (classes 1 and 2) are lower than those in CS where group 1 and 2 users have the same priority.

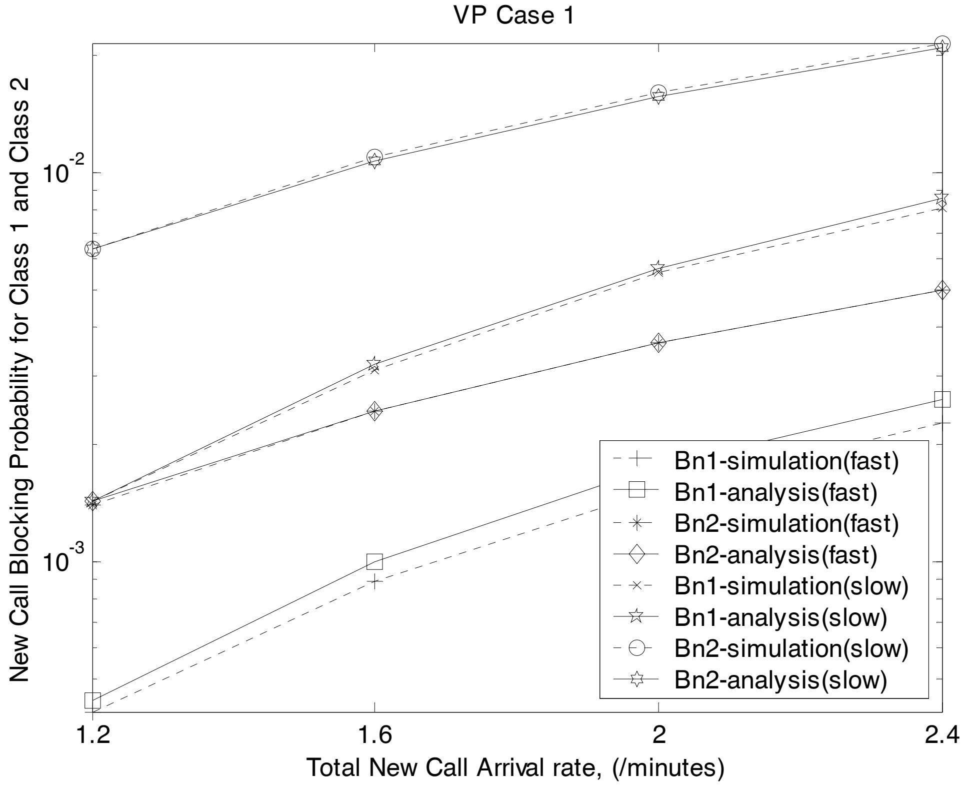

Fig. 4.8. New-call-blocking probabilities for classes 1 and 2 (VP Case 1)

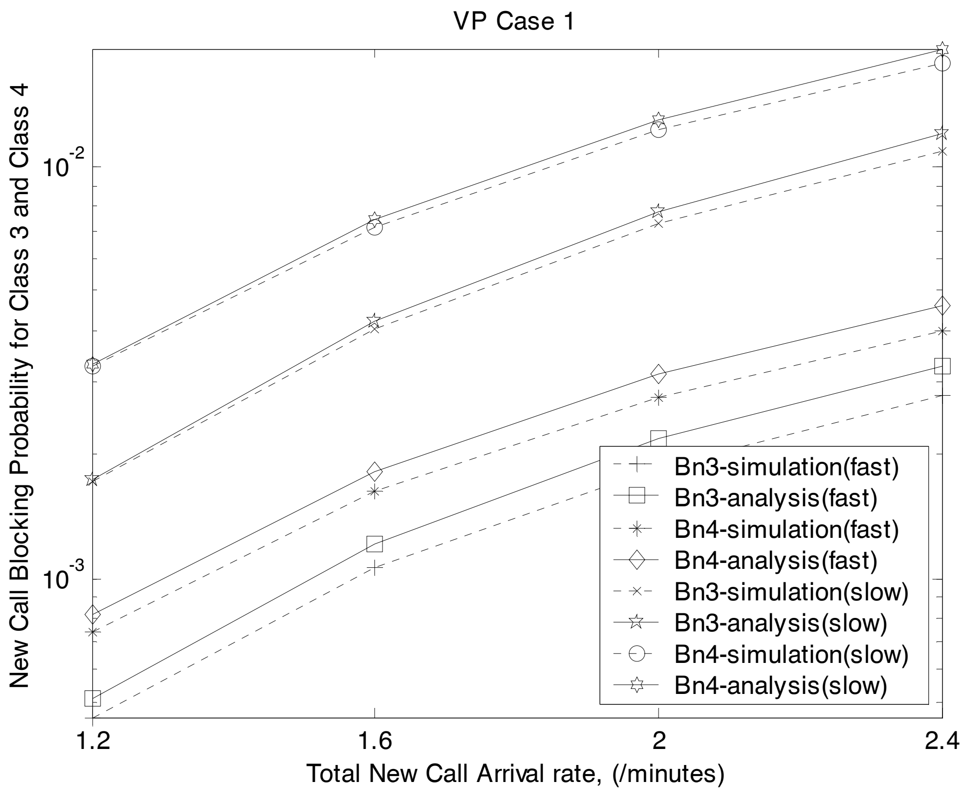

Fig. 4.9. New-call-blocking probabilities for classes 3 and 4 (VP Case 1)

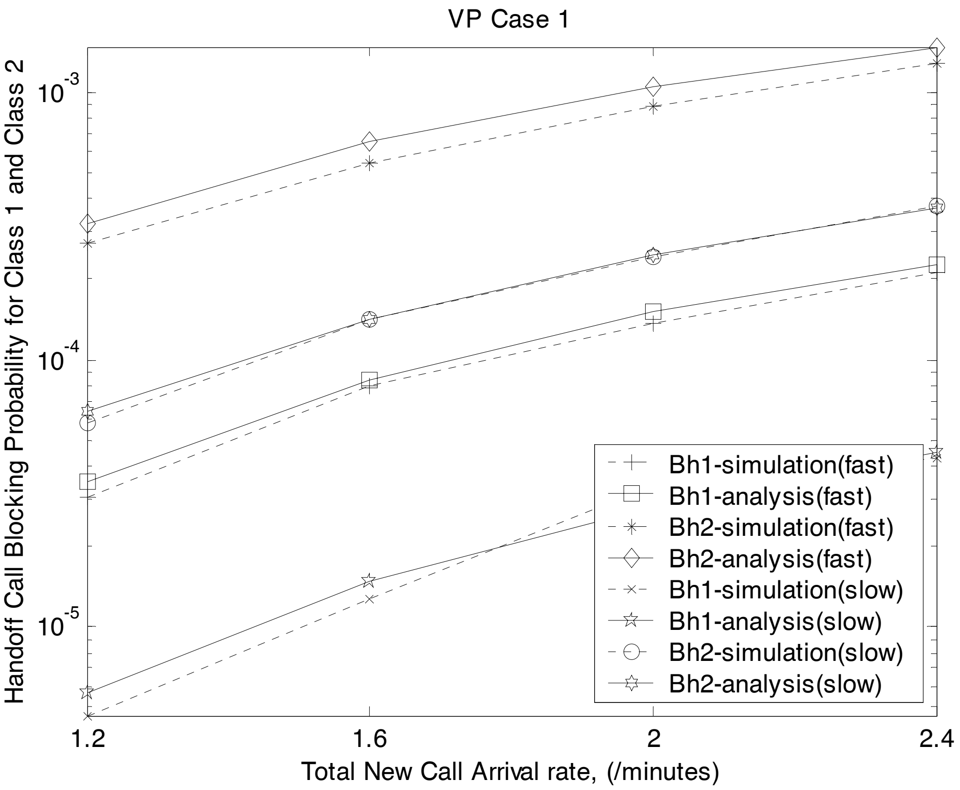

Fig. 4.10. Handoff-call-dropping probabilities for classes 1 and 2 (VP Case 1)

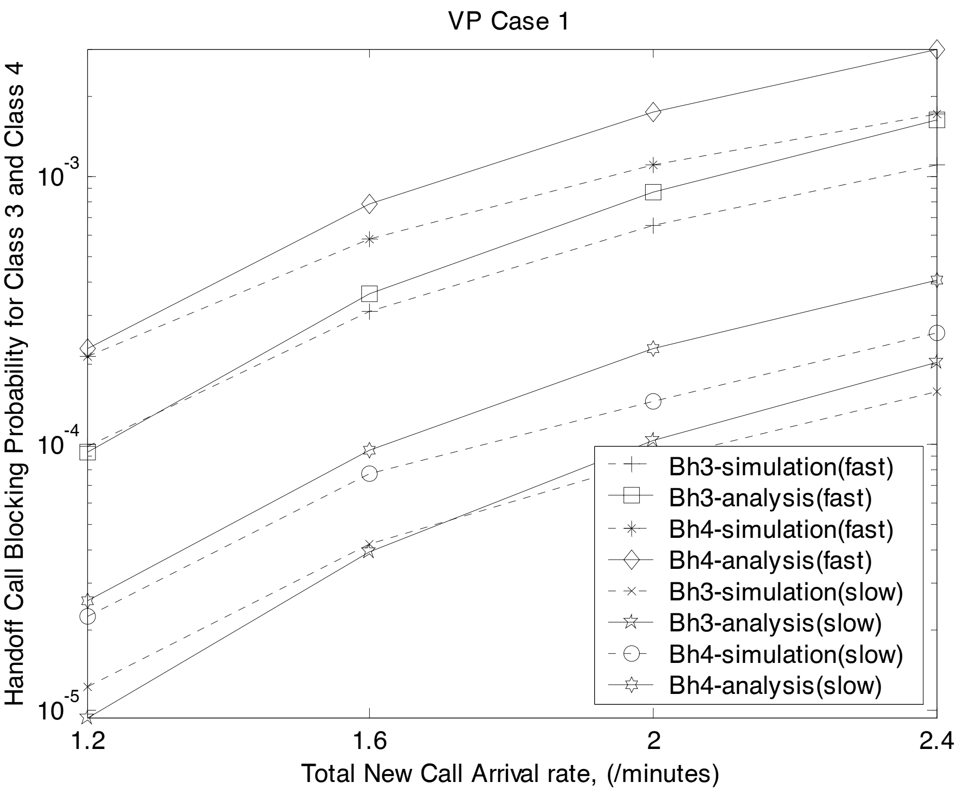

Fig. 4.11. Handoff-call-dropping probabilities for classes 3 and 4 (VP Case 1)

Figs. 4.12 and 4.13 show the connection-level analytical and simulation results of system utilization, N_u, for fast and slow users, respectively. Similar to CS, the analytical and simulation results agree better for the slow users' case than those for the fast users' case. Besides, due to the different priorities for group 1 and 2 users in VP case 1, the system utilizations for group 1 users in VP case 1 become higher than those in CS. For example, N_{u2} in Figs. 4.12 and 4.13 is higher than that in Figs. 4.6 and 4.7 and N_{u3} in Figs. 4.12 and 4.13 is lower than that in Figs. 4.6 and 4.7.

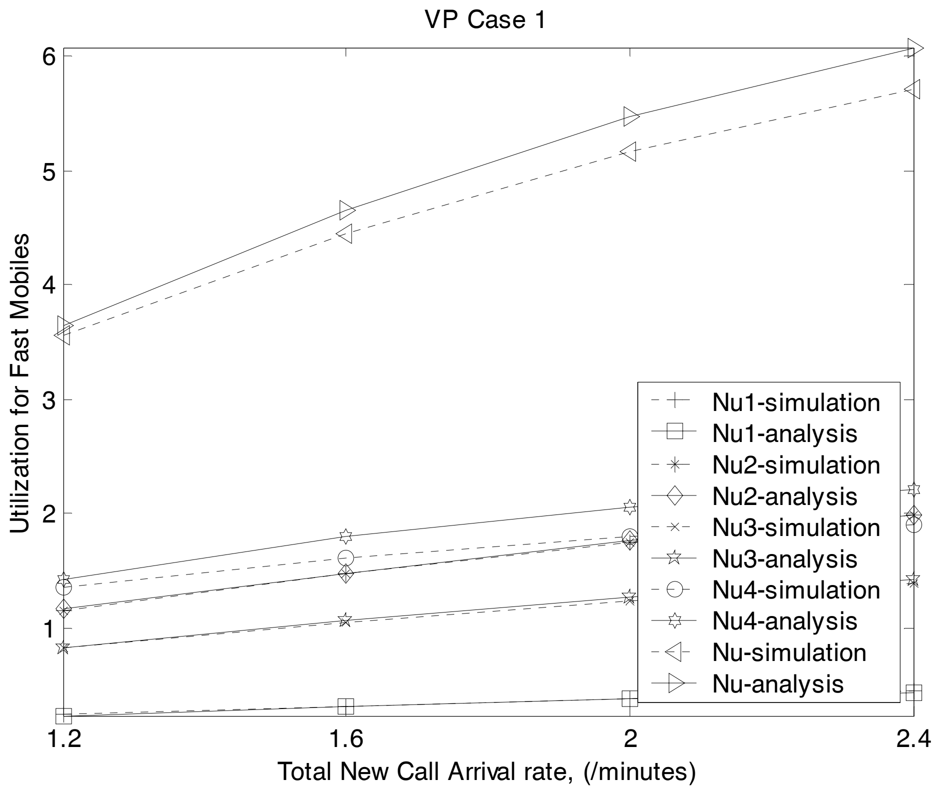

Fig. 4.12. System utilization for fast mobile users (VP Case 1)

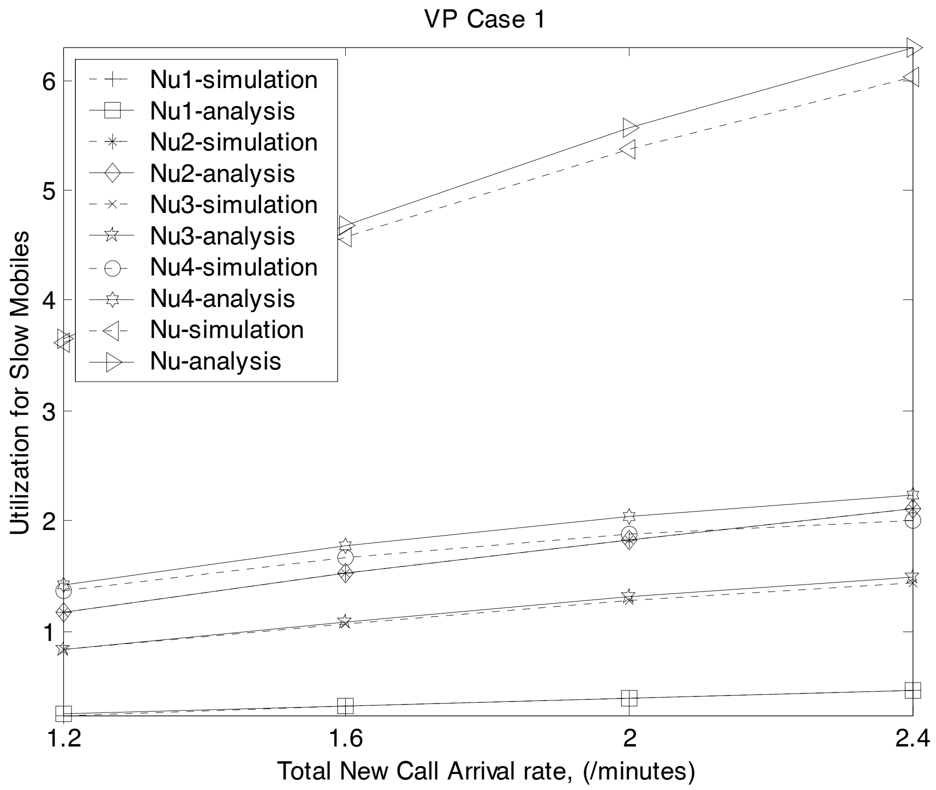

Fig. 4.13. System utilization for slow mobile users (VP Case 1)

4.4 VP Case 2 - VP with Preemption for Group 2

The difference between case 1 and case 2 lies on that, in case 1, group 2 can preempt group 1. In case 2, there is no preemption for group 1 and group 1 is offered guaranteed access while group 2 is offered best-effort service. As a result, the admission policies and the preemption rules in case 2 will differ from those in case 1.

4.4.1 Case 2 – Connection-Level Analysis

The admission policies for case 2 are defined as:

$$f_k(\vec{n}) = \begin{cases} 1, & \text{condition 1} \\ 0, & \text{otherwise} \end{cases}, \tag{4.30}$$

where condition 1 refers to: for $k \in G_2$, $\sum_{i=p+1}^{K} r_i n_i + r_k \leq N - C_G$; or for $k \in G_1$,

$\sum_{i=1}^{p} r_i n_i + r_k \leq C_1 - C_G$, and

$$f_{Gk}(\vec{n}) = \begin{cases} 1, & \text{condition 2} \\ 0, & \text{otherwise} \end{cases}, \tag{4.31}$$

where condition 2 refers to: for $k \in G_2$, $N - C_G < \sum_{i=p+1}^{K} r_i n_i + r_k \leq N$; or for $k \in G_1$, $C_1 - C_G < \sum_{i=1}^{p} r_i n_i + r_k \leq C_1$.

Preemption in case 2 is governed by the following preemption rules:

1. Preemption could happen only under the condition that $\sum_{k=p+1}^{K} n_k r_k > C_2$, i.e., group 2 users occupying capacity nominally allocated to group 1.

2. Preemption happens when new class $j \in G_1$ call arrives under the conditions that $C_G + \vec{r} \cdot \vec{n} = N$, or { $C_G + \vec{r} \cdot \vec{n} < N$ and $C_G + \vec{r} \cdot \vec{n} > N - r_j$ }.

3. Preemption happens when handoff class $j \in G_1$ call arrives under the condition that $\vec{r} \cdot \vec{n} > N - r_j$.

4. When there is preemption for class $j \in G_1$ over class *k* in group 2, the number of class *k* users terminated from the system will be less than or equal to $\left\lceil \frac{r_j}{r_k} \right\rceil$.

5. If preemption happens over class *k*, it means that class *k* is the most overload class in group 2 and the class *k* will be determined by the criterion $k = \arg\min\left\{\left(C_2 - n_{p+1} \times r_{p+1}\right), \ldots, \left(C_2 - n_K \times r_K\right)\right\}$.

Let $P(\vec{n})$ denote the equilibrium probability that the system is in state $\vec{n}$. The global balance equations for the Markov process under the preemption rules and the admission policy for case 2 are given by:

$$
\begin{aligned}
&\sum_{k=1}^{K}\{\lambda_k(\vec{n})[f_k(\vec{n})+f_{Gk}(\vec{n})]+\lambda_{Pk}(\vec{n})+\mu_k(\vec{n})\}P(\vec{n}) \\
&=\sum_{k=1}^{K}P(\vec{n}-\vec{e}_k)\lambda_k(\vec{n}-\vec{e}_k)[f_k(\vec{n}-\vec{e}_k)+f_{Gk}(\vec{n}-\vec{e}_k)]+\sum_{k=1}^{K}P(\vec{n}+\vec{e}_k)\mu_k(\vec{n}+\vec{e}_k), \quad (4.32) \\
&+\sum_{k=p+1}^{K}\sum_{j=1}^{p}\left[\sum_{i=1}^{\lceil r_j/r_k\rceil}P(\vec{n}+i\vec{e}_k-\vec{e}_j)\lambda_{Pk}(\vec{n}+i\vec{e}_k-\vec{e}_j)\right]
\end{aligned}
$$

where $\lambda_k(\vec{n})$ is defined in (4.9), $\mu_k(\vec{n})$ is defined in (4.10), and $\lambda_{Pk}(\vec{n})$ is defined in (4.23).

A new class $k \in G_1$ call is blocked from entering the system (and is assumed lost) if upon arrival it finds that it cannot be accommodated because the number of available channels (excluding the guard channels) is less than r_k. Therefore the blocking probability for a new class $k \in G_1$ call, considering all classes, is given by

$$B_{nk} = \sum_{n_1=0}^{\lfloor C_1/r_1\rfloor}\cdots\sum_{n_p=0}^{\left\lfloor\left(C_1-\sum_{i=1}^{p-1}n_ir_i\right)/r_p\right\rfloor}\;\sum_{n_{p+1}=0}^{\left\lfloor\left(N-\sum_{i=1}^{p}n_ir_i\right)/r_{p+1}\right\rfloor}\cdots\sum_{n_K=0}^{\left\lfloor\left(N-\sum_{i=1}^{K-1}n_ir_i\right)/r_K\right\rfloor}P(\vec{n})\theta_{nk}-\sum_{i=p+1}^{K}P_{ni}\,, \quad (4.33)$$

for all $\vec{n}$ satisfying $(N-C_G-\vec{r}\cdot\vec{n})<r_k$, where P_{ni} is the preemption probability for class *i* call when a new class *k* call arrives, which is defined in (4.37).

Similarly, the blocking probability for a class $k \in G_2$ new call, considering all classes, is given by:

$$B_{nk} = \sum_{n_1=0}^{\lfloor C_1/r_1\rfloor}\cdots\sum_{n_p=0}^{\left\lfloor\left(C_1-\sum_{i=1}^{p-1}n_ir_i\right)/r_p\right\rfloor}\;\sum_{n_{p+1}=0}^{\left\lfloor\left(N-\sum_{i=1}^{p}n_ir_i\right)/r_{p+1}\right\rfloor}\cdots\sum_{n_K=0}^{\left\lfloor\left(N-\sum_{i=1}^{K-1}n_ir_i\right)/r_K\right\rfloor}P(\vec{n})\theta_{nk}\,, \quad (4.34)$$

for all $\vec{n}$ satisfying $(N-C_G-\vec{r}\cdot\vec{n})<r_k$.

A class $k \in G_1$ handoff call is blocked from entering the system (and is assumed lost) if upon arrival it finds that it cannot be accommodated because the number of available channels (including the guard channels) is less than r_k. Therefore, the dropping probability for a class $k \in G_1$ handoff call, considering all classes, is given by:

$$B_{hk} = \sum_{n_1=0}^{\lfloor C_1/r_1 \rfloor} \cdots \sum_{n_p=0}^{\left\lfloor \left(C_1 - \sum_{i=1}^{p-1} n_i r_i \right) \middle/ r_p \right\rfloor} \sum_{n_{p+1}=0}^{\left\lfloor \left(N - \sum_{i=1}^{p} n_i r_i \right) \middle/ r_{p+1} \right\rfloor} \cdots \sum_{n_K=0}^{\left\lfloor \left(N - \sum_{i=1}^{K-1} n_i r_i \right) \middle/ r_K \right\rfloor} P(\vec{n})\theta_{hk} , \qquad (4.35)$$

for all $\vec{n}$ satisfying $\left(C_1 - \sum_{i=1}^{p} n_i r_i\right) < r_k$.

Similarly, the dropping probability for a class $k \in G_2$ handoff call considering all classes is given by:

$$B_{hk} = \sum_{n_1=0}^{\lfloor C_1/r_1 \rfloor} \cdots \sum_{n_p=0}^{\left\lfloor \left(C_1 - \sum_{i=1}^{p-1} n_i r_i \right) \middle/ r_p \right\rfloor} \sum_{n_{p+1}=0}^{\left\lfloor \left(N - \sum_{i=1}^{p} n_i r_i \right) \middle/ r_{p+1} \right\rfloor} \cdots \sum_{n_K=0}^{\left\lfloor \left(N - \sum_{i=1}^{K-1} n_i r_i \right) \middle/ r_K \right\rfloor} P(\vec{n})\theta_{hk} , \qquad (4.36)$$

for all $\vec{n}$ satisfying $\left(N - \vec{r} \cdot \vec{n}\right) < r_k$.

The preemption probability for a class $k \in G_2$ call, when a class $j \in G_1$ new call arrives, is given by:

$$P_{nk} = \sum_{n_1=0}^{\lfloor C_1/r_1 \rfloor} \cdots \sum_{n_p=0}^{\left\lfloor \left(C_1 - \sum_{i=1}^{p-1} n_i r_i \right) \middle/ r_p \right\rfloor} \sum_{n_{p+1}=0}^{\left\lfloor \left(N - \sum_{i=1}^{p} n_i r_i \right) \middle/ r_{p+1} \right\rfloor} \cdots \sum_{n_K=0}^{\left\lfloor \left(N - \sum_{i=1}^{K-1} n_i r_i \right) \middle/ r_K \right\rfloor} P(\vec{n})\theta_{nj} , \qquad (4.37)$$

for all $\vec{n}$ satisfying rules 2 and 5.

The preemption probability for a class $k \in G_2$ call, when a class $j \in G_1$ handoff call arrives, is given by:

$$P_{hk} = \sum_{n_1=0}^{\lfloor C_1/r_1 \rfloor} \cdots \sum_{n_p=0}^{\left\lfloor \left(C_1 - \sum_{i=1}^{p-1} n_i r_i \right) \middle/ r_p \right\rfloor} \sum_{n_{p+1}=0}^{\left\lfloor \left(N - \sum_{i=1}^{p} n_i r_i \right) \middle/ r_{p+1} \right\rfloor} \cdots \sum_{n_K=0}^{\left\lfloor \left(N - \sum_{i=1}^{K-1} n_i r_i \right) \middle/ r_K \right\rfloor} P(\vec{n})\theta_{hj} , \qquad (4.38)$$

for all $\vec{n}$ satisfying rules 3 and 5.

The utilization for class *k* is given by:

$$N_{uk} = \sum_{n_1=0}^{\lfloor C_1/r_1 \rfloor} \cdots \sum_{n_p=0}^{\left\lfloor \left(C_1 - \sum_{i=1}^{p-1} n_i r_i \right) / r_p \right\rfloor} \sum_{n_{p+1}=0}^{\left\lfloor \left(N - \sum_{i=1}^{p} n_i r_i \right) / r_{p+1} \right\rfloor} \cdots \sum_{n_K=0}^{\left\lfloor \left(N - \sum_{i=1}^{K-1} n_i r_i \right) / r_K \right\rfloor} n_k P(\vec{n}) r_k \,. \qquad (4.39)$$

The total utilization is thus,

$$N_u = \sum_{k=1}^{K} N_{uk} \,. \qquad (4.40)$$

4.4.2 Case 2 – Packet-Level Analysis

The analytical formulation of the packet-loss probability for case 2 is the same as that for CS in section 4.2.2.

4.4.3 Case 2 - Performance Evaluation

The performance for case 2 by means of numerical analysis and simulation is evaluated. The simulation model and the parameters used in the numerical examples are all the same as those in CS. The new call arrival rates are all the same for the 4 traffic classes. The connection-level analytical and simulation results of new-call-blocking and handoff-call-dropping probabilities, B_n and B_h, for classes 1 and 2, and 3 and 4 for fast and slow mobile users are shown in Figs. 4.14, 4.15, 4.16 and 4.17, respectively. The preemption rules for VP case 2 are proven to be reasonable as the agreement of the analytical results and simulation results are excellent. The preemption happens when the system is in the over-loaded situation, i.e., VP behaves like complete partitioning when the overall traffic is heavy. The preemption protects both the new call and handoff call arrivals. Thus, the call blocking probabilities are lower than those in CS especially in the high load situation (where the total new call arrival rate is 2.4/minute).

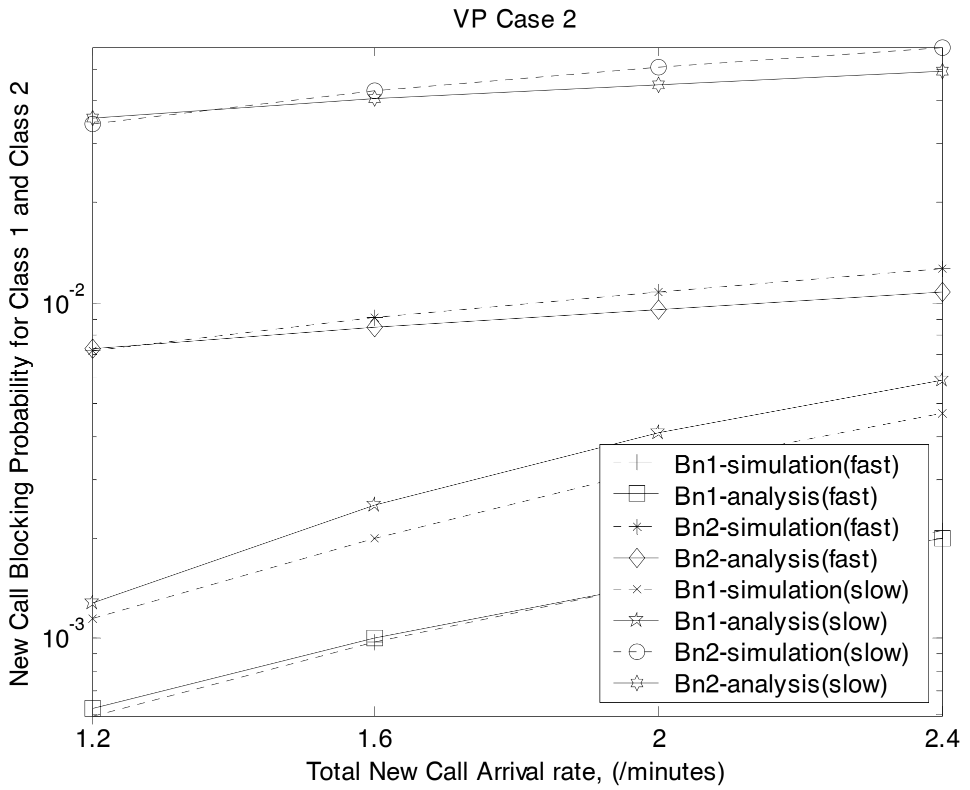

Fig. 4.14. New-call-blocking probabilities for classes 1 and 2 (VP Case 2)

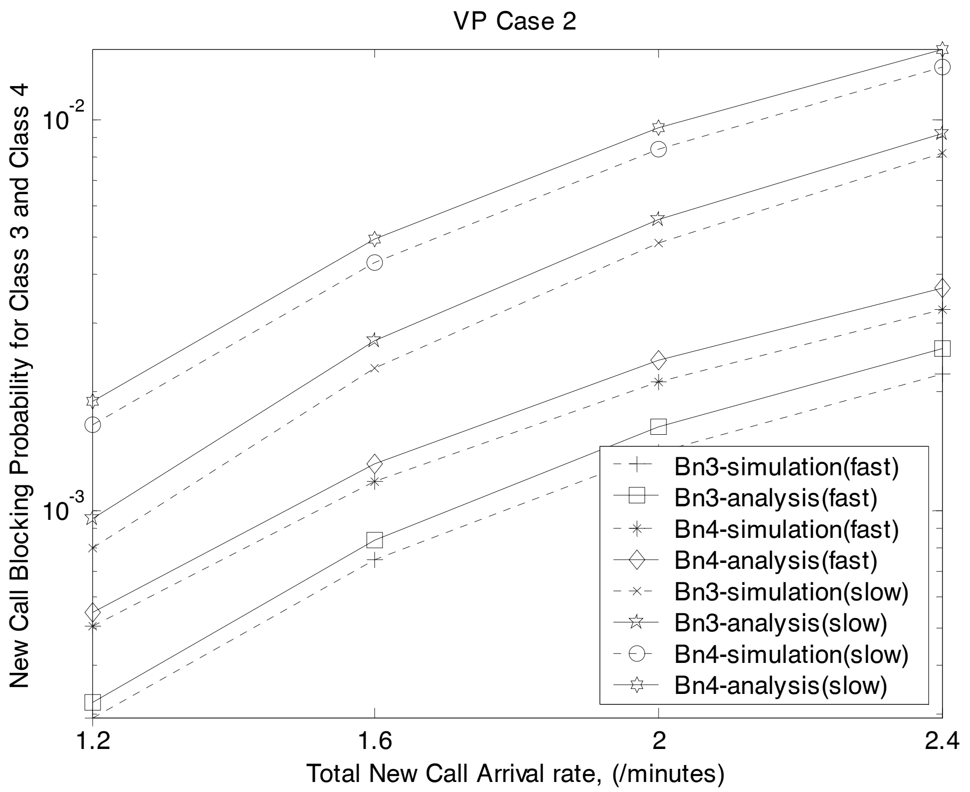

Fig. 4.15. New-call-blocking probabilities for classes 3 and 4 (VP Case 2)

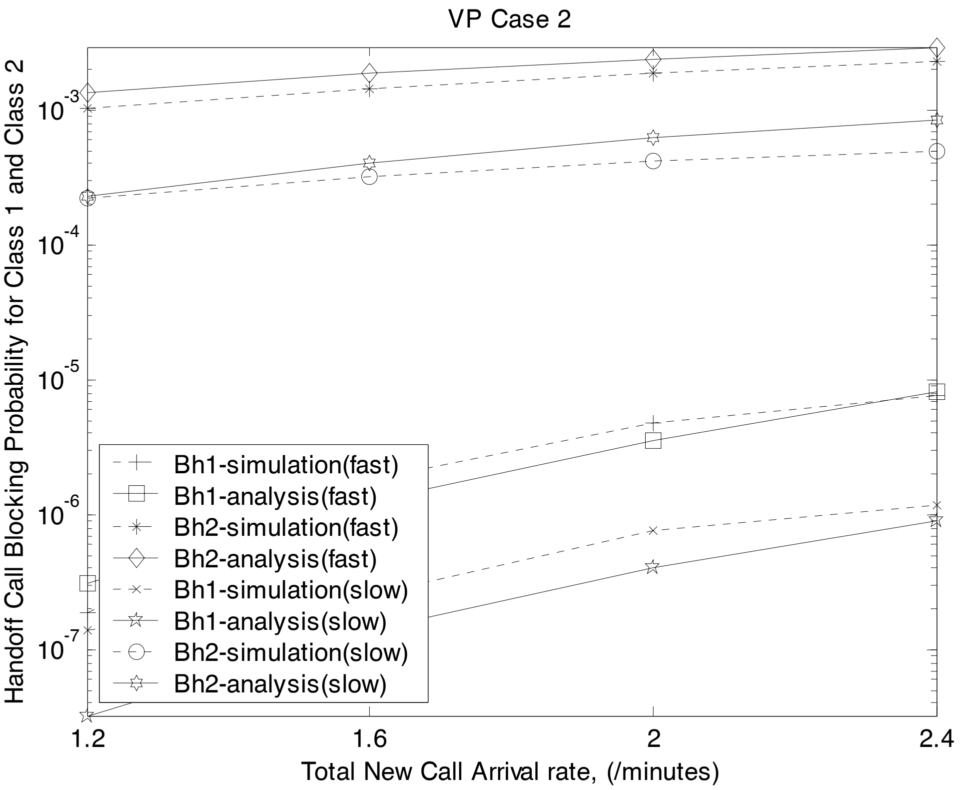

Fig. 4.16. Handoff-call-dropping probabilities for classes 1 and 2 (VP Case 2)

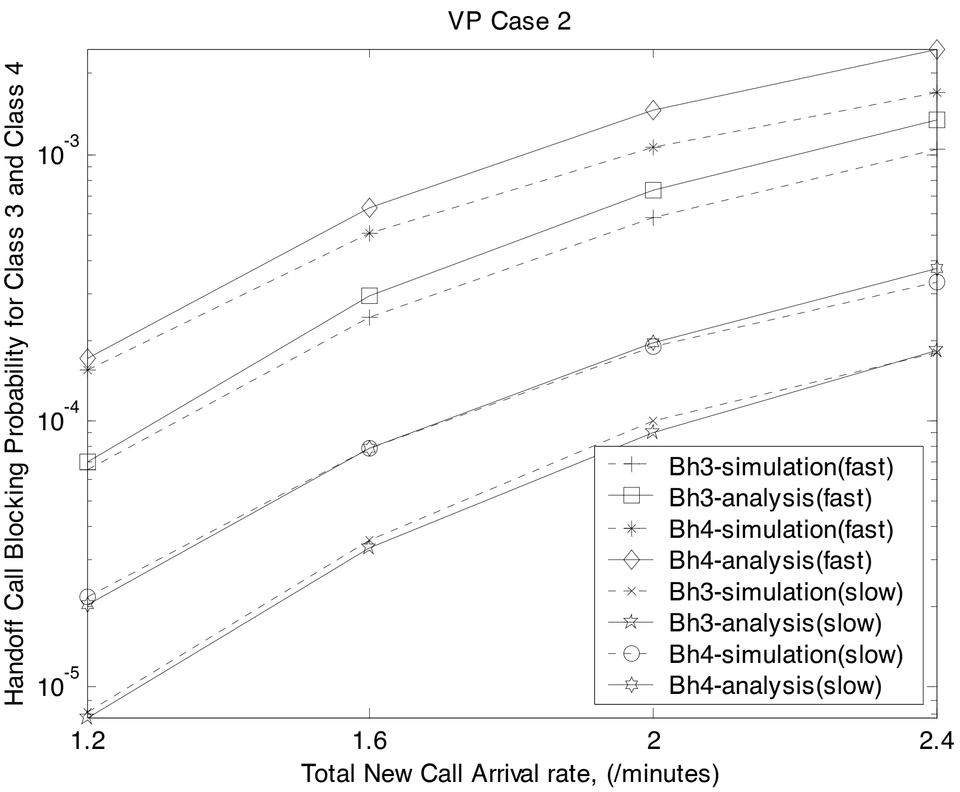

Fig. 4.17. Handoff-call-dropping probabilities for classes 3 and 4 (VP Case 2)

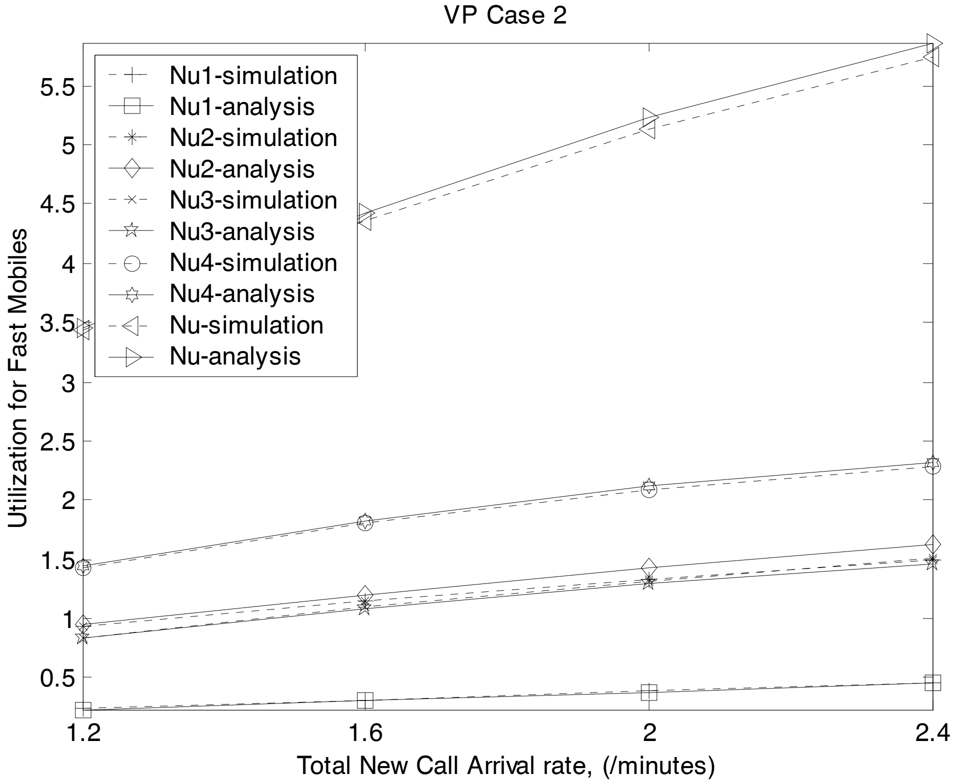

Fig. 4.18. System utilization for fast mobile users (VP Case 2)

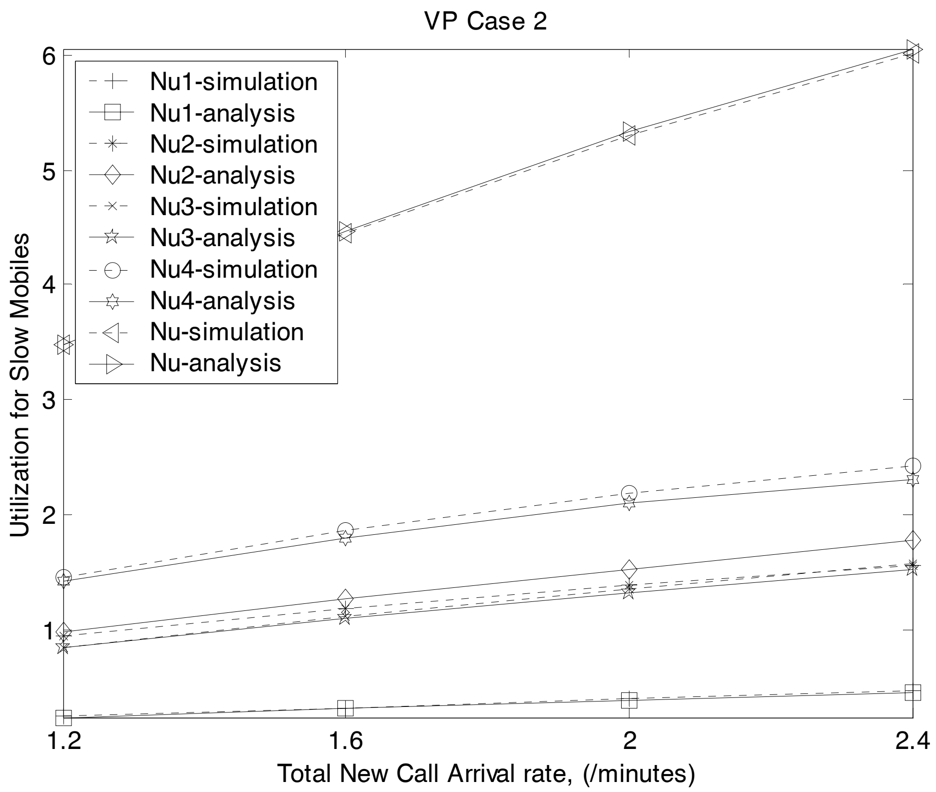

Fig. 4.19. System utilization for slow mobile users (VP Case 2)

Figs. 4.18 and 4.19 show the connection-level analytical and simulation results of system utilization, N_u, for fast and slow mobile users, respectively. Similar to CS, the analytical and simulation results agree better for the slow users' case than those for the fast users' case.

4.5 Joint Connection and Packet Levels Optimization

4.5.1 The Joint Levels Optimization Analysis

The formulae for the QoS metrics, including the new-call-blocking probability, the handoff-call-dropping probability and the system utilization at the connection-level, and the packet-loss probability at the packet-level have been defined for CS and VP. Maximizing system utilization can translate to more revenue for the network providers who in turn can lower the charges for the mobile users. So, the goal of joint connection and packet levels optimization is to maximize system utilization subject to the QoS constraints in both levels. It can be achieved by choosing the optimal parameter *N*. Every value of *N* is tested as a system parameter in a range of values, and the optimum *N* that gives the maximum system utilization and satisfies the QoS constraints at the same time is selected. The optimization problem is then presented as:

$$\max\{N_u\}, \tag{4.41}$$

subject to

$$B_n \le B_n^*,$$

$$B_h \le B_h^*,$$

$$L_k \le L_k^*,$$

where the superscript * denotes the QoS requirements of the corresponding parameters. The selection of optimal N uses the educated search with an iterative procedure described in section 3.3.1.

4.5.2 Performance Evaluation on the Joint Levels Optimization

The parameters used in the numerical examples are all the same as those in the CS performance evaluation, except that $B_{nk}^{*} = 0.1$, $B_{hk}^{*} = 0.1$ and $L_{k}^{*} = 1 \times 10^{-3}$. The new call arrival rates are all the same for the 4 traffic classes.

4.5.2.1 Joint Levels Optimization for CS

Figs. 4.20 and 4.21 show the gain in system utilization through joint connection and packet levels QoS optimization for fast and slow mobile users, respectively.

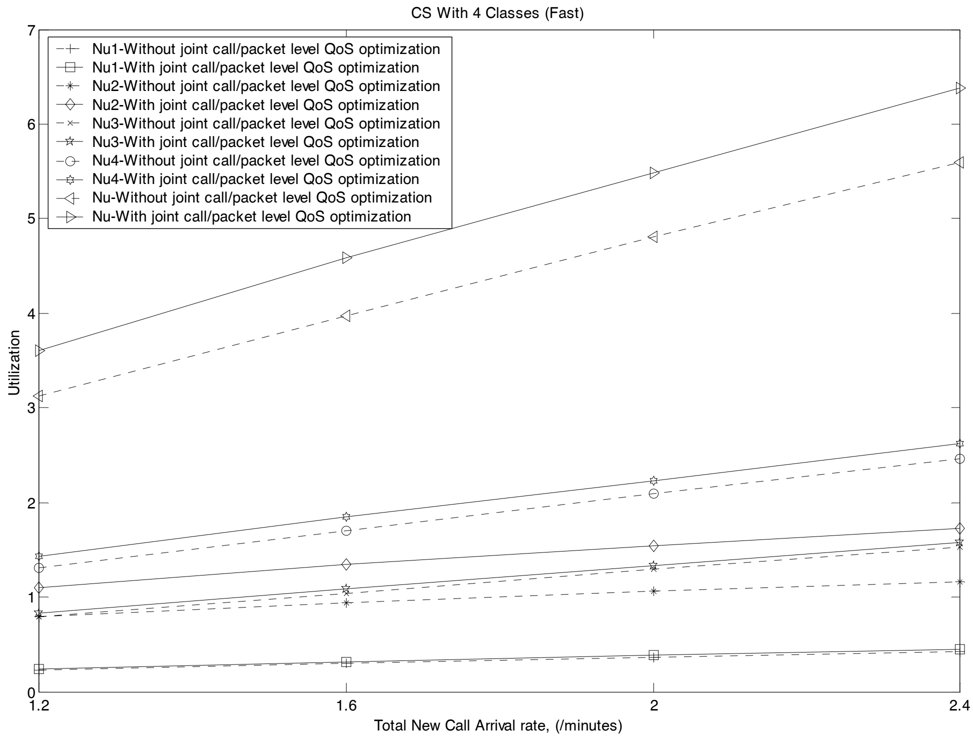

Fig. 4.20. System utilization with and without joint connection and packet levels QoS optimization for fast mobile users (CS)

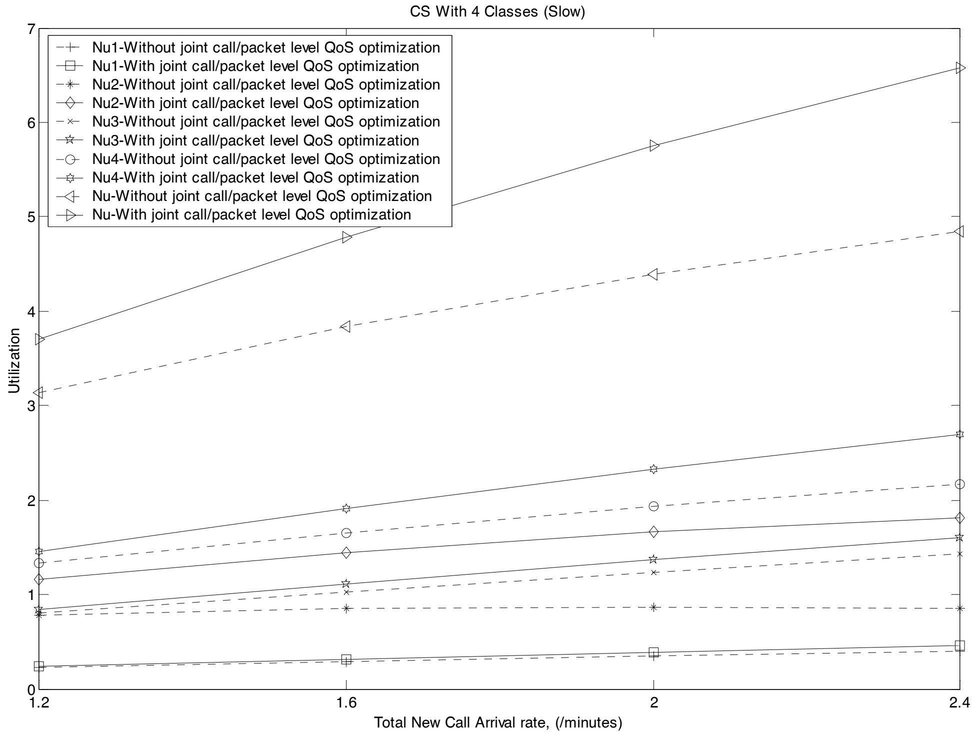

Fig. 4.21. System utilization with and without joint connection and packet levels QoS optimization for slow mobile users (CS)

Without joint connection and packet levels optimization, *N* is set to equal to the physical capacity. Thus, the packet loss probability will be zero for the case without optimization. With joint connection and packet levels optimization, the parameter *N* is maximized. In this numerical example, the optimal *N* is 17 for both fast and slow users' scenarios. The system can admit more new and handoff calls though the packet loss probabilities are also increased but still subject to the QoS constraint. In the figures, the larger the load (total mean new call arrival rate), the larger the gain in system utilization, e.g., in Fig. 4.21, at $\sum_{k=1}^{K} \lambda_k = 1.2$, the total system utilization gain is about 0.6, while at $\sum_{k=1}^{K} \lambda_k = 2.4$, the gain becomes over 1.5.

4.5.2.2 Joint Levels Optimization for Case 1

Figs. 4.22 and 4.23 show the gain in system utilization through joint connection and packet levels QoS optimization for fast and slow mobile users, respectively.

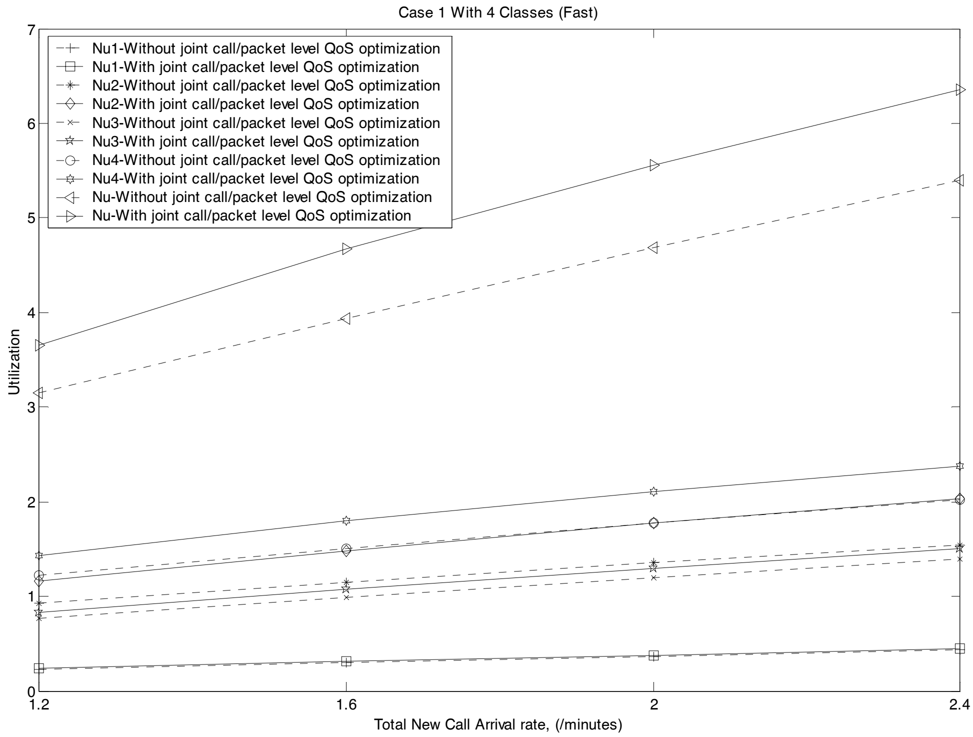

Fig. 4.22. System utilization with and without joint connection and packet levels QoS optimization for fast mobile users (VP Case 1)

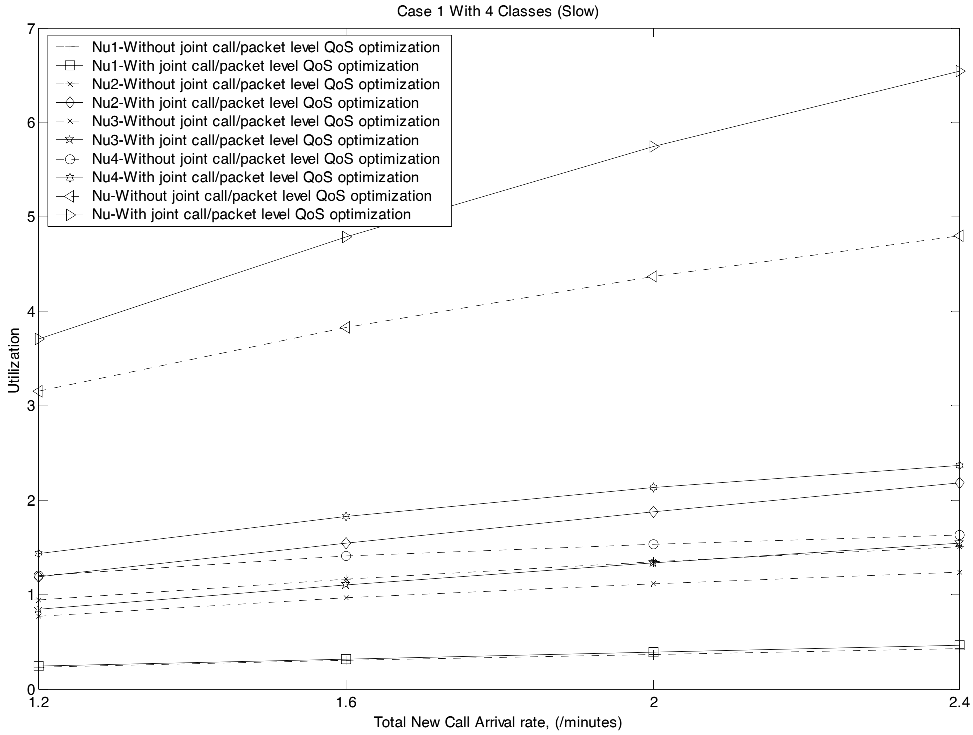

Fig. 4.23. System utilization with and without joint connection and packet levels QoS optimization for slow mobile users (VP Case 1)

The joint levels optimization exploits the statistical behaviors between the packet level and the connection level to implement the statistical multiplexing in the model. The optimization is independent from CAC schemes applied in the network layer. Thus, the joint levels optimization also brings about system utilization gain for the VP scheme. In this numerical example, the optimal *N* is also 17 for both fast and slow users' scenarios.

4.5.2.3 Joint Levels Optimization for Case 2

Figs. 4.24 and 4.25 show the gain in system utilization through joint connection and packet levels QoS optimization for fast and slow mobile users, respectively.

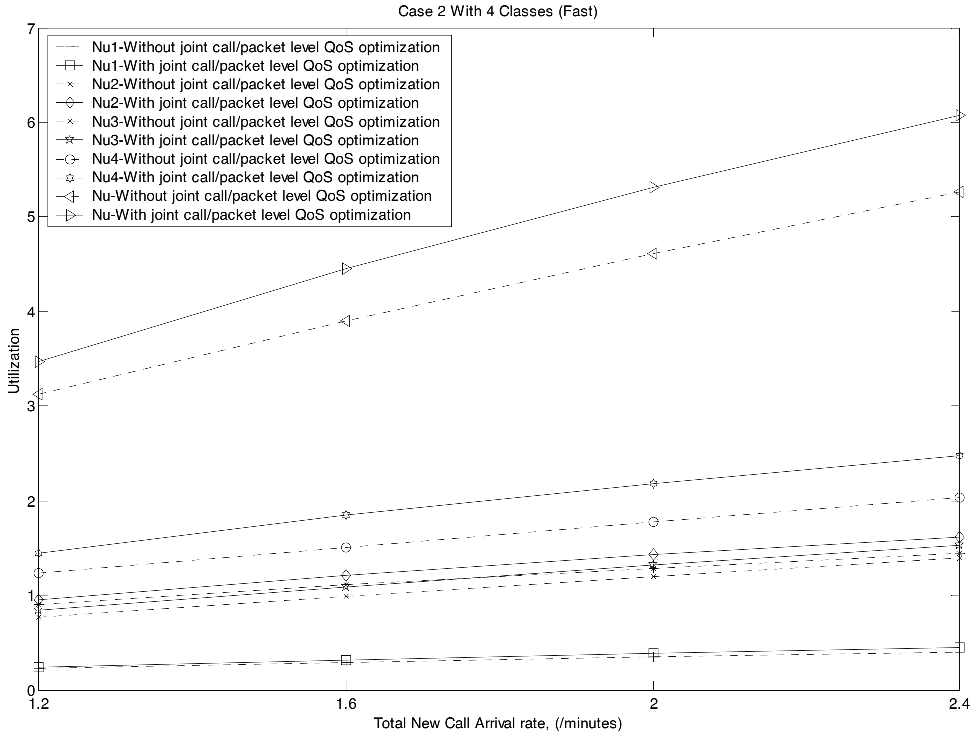

Fig. 4.24. System utilization with and without joint connection and packet levels QoS optimization for fast mobile users (VP Case 2)

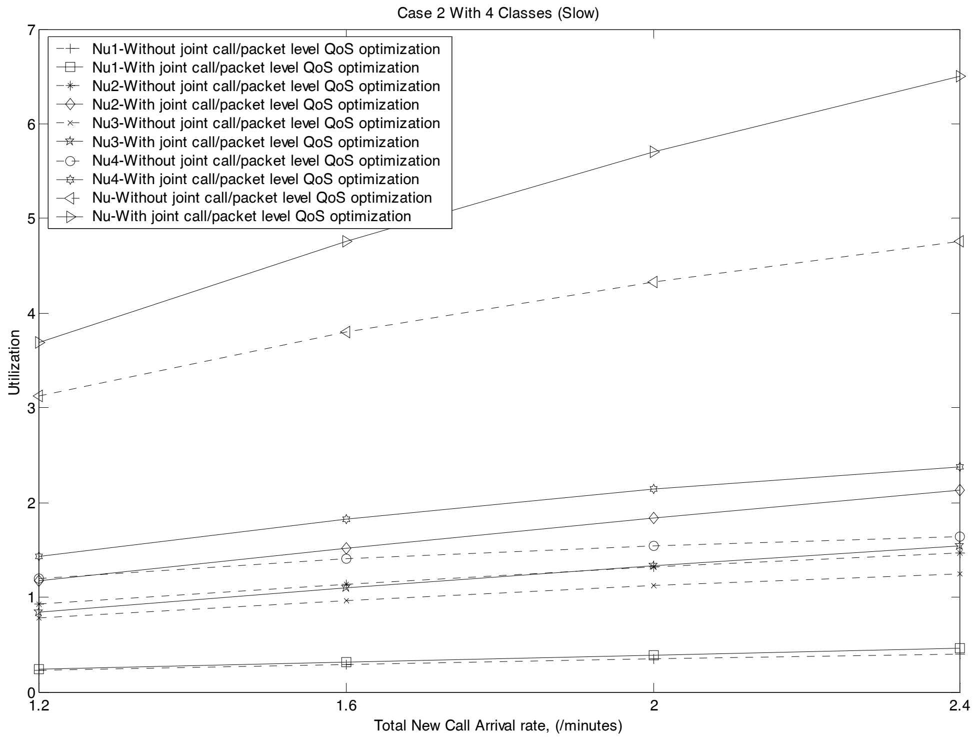

Fig. 4.25. System utilization with and without joint connection and packet levels QoS optimization for slow mobile users (VP Case 2)

In this numerical example, the optimal N is also 17 for both fast and slow users' scenarios. Thus, for group 1 (classes 1 and 2), case 2 has better gain in system utilization than case 1. For group 2 (classes 3 and 4), case 1 has better gain in system utilization than case 2. These are expected because group 1 users (guaranteed access) are assigned higher priorities than group 2 users (best effort) in case 2 and groups 1 and 2 users are assigned same priorities in case 1. Thus, group 1 users in case 2 will be allocated more system resources than group 1 users in case 1 and group 2 users in case 2 will be allocated less system resources than group 2 users in case 1. In Figs. 4.22, 4.23, 4.24 and 4.25, system utilization gains of group 1 users in case 2 exceed those in case 1 and system utilization gains of group 2 users in case 1 exceed those in case 2.

4.6 Summary

An approximate analytical formulation of complete sharing and virtual partitioning resource allocation schemes for handling multi-class traffic with guard channels in a cellular system have been presented in this chapter. The technique that derives the analytical formulation of VP schemes for multi-class traffic is proposed for the first time. The analytical models are based on a K-dimensional Markov chain and solved using the preemption rules for the schemes. The QoS performance metrics are call blocking probability, system utilization and packet-loss probability. Joint optimization through the connection and packet levels QoS is investigated to achieve higher gain in system utilization. This will translate to more revenue for network providers who, in turn, can lower the charges for subscribers.

To simplify the analysis procedure, a physical capacity, *C*, is defined to represent the characteristics in the physical and link layers. However, to represent the physical capacity of a CDMA cellular system with a constant is far from satisfying as the interference-limited CDMA cellular system has the soft capacity. Thus, to build the interference model is necessary for system designers to look into the influence of the characteristics in the physical and link layers on the physical capacity. The physical and link layers issues are investigated in Chapter 5, including the interference model, soft handoff, power control, SIR, and the outage probability, etc., in both the reverse and forward links. The outage probabilities are formulated for the further usage on the determination of the admission region which is the largest set of QoS points delivered under any CAC schemes.

Chapter 5

Admission Regions in the Reverse and Forward Links

The salient features of a CDMA system differing from FDMA and TDMA systems include universal frequency reuse, soft handoff, using Rake receiver, soft capacity, statistical multiplexing, etc. These features will deeply influence the whole system performance penetrating from the lower layers to the higher layers. Investigation on these features is important as they are the foundations for the whole system. In this chapter, the link layer QoS metrics, including SIRs and the outage probabilities, for multi-class services in BSs (the reverse link) and in mobile users (the forward link) are derived with the salient features in the CDMA system described above. The contribution of the work is to originally derive the mathematical expressions for the interference model, SIRs and the outage probability without poor simplification and assumptions. The admission regions in the reverse link, in the forward link and for the link layer considering both the reverse and forward links are formulated.

During the evaluation on the SIRs in the reverse and forward links, a single-connection system model of one mobile user supporting only a single-connection service at any time is assumed. Despite the single-connection assumption made in the thesis, the analytical model is also suitable for the multi-connection system that one mobile user supports more than one connection service at one time. The service is modeled as Sen's model containing the HBR and LBR spreading codes and the same traffic models are assumed in both the forward and reverse links. Formulating the SIRs for services in the reverse and forward links, the final SIR results are the same for mobile users using HBR spreading codes and using LBR spreading codes [100]. Thus, the final SIR results analysis can be just represented for mobile users using HBR or LBR spreading codes even though the results are functions of signals or interference powers from both HBR and LBR sources.

5.1 Evaluation on the Reverse Link

The signal-to-interference ratio (SIR) is an important QoS measurement and should be maintained above a required threshold so that mobile users can communicate to each other smoothly. However, the short-term degradation of SIR is tolerable as the data loss can be compensated by retransmission in the transmitters or recovered through forward error correction in the receivers. The outage probability is then used as the degradation probability that SIR is below its requirement. To obtain the expression for SIR and the outage probability, the interference model should be carefully built. In the reverse link, soft handoff will decrease the inter-cell interference and improve the SIR performance.

5.1.1 Interference Models in the Reverse Link

The interested cell is treated as the given zeroth cell and the BS in it is treated as BS 0. The reverse link geometry is shown in Fig. 5.1. The cell structure is presented without the streets in the Manhattan model shown in Fig. 4.1. The arrowhead with solid line represents the power from the mobile user to the BS that power-controls it. The arrowhead with dashed line represents the power from the mobile user to the BS that connects but does not power-control the mobile user. The arrowhead with dotted line represents the interference to the BS. The mobile users in the system are divided into four types. Type I includes the mobile users power-controlled by BS 0, e.g., mobile users *a* and *b*. Type II includes the mobile users that are in soft handoff and connected to but not power-controlled by BS 0, e.g., mobile user *c*. Type III includes the mobile users that are in soft handoff and not connected to BS 0, e.g., mobile user *d*. Type IV includes the mobile users that are in non-soft handoff and not connected to BS 0, e.g., mobile user *e*. Thus, mobile users of types I and II will introduce intra-cell interference to BS 0 and mobile users of types III and IV will introduce inter-cell interference to BS 0.

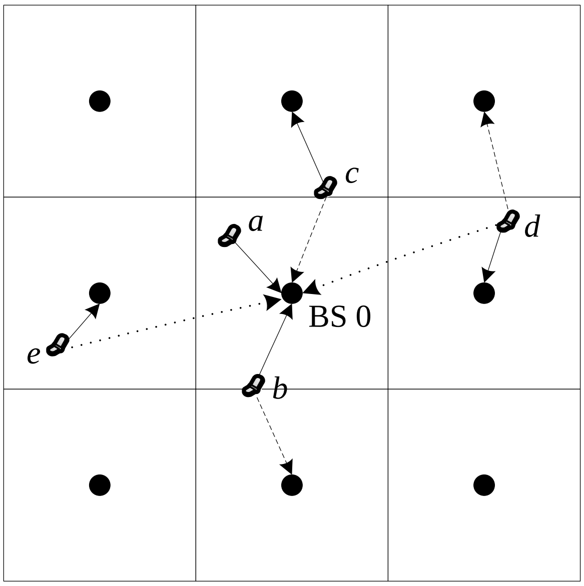

Fig. 5.1. The Reverse Link Geometry

5.1.1.1 Interference Model for Type I Mobile Users

The interference-to-signal ratio (ISR) is considered. Due to the assumption of perfect SIR power control, the intra-cell ISR can be assumed as a mean value related to the number of mobile users in the given zeroth cell. Thus, for class *k* traffic, the mean of ISR from type I mobile users using HBR spreading code, denoted by $I_{I,k}^{(h)} \big/ S_k^{(h)}$, is $m_{I,k}^{(h)}$, where *k* is the index of class *k* traffic, $I_{I,k}^{(h)}$ is the interference power from class *k* traffic using HBR spreading code of type I, $S_k^{(h)}$ is the received power of class *k* traffic using HBR spreading code. $m_{I,k}^{(h)}$ can be shown that

$$m_{I,k}^{(h)} = \alpha_k^{(h)} n_k, \tag{5.1}$$

where $\alpha_k^{(h)}$ is the connection activity factor for the HBR mini-source defined in (3.8), and n_k is the instantaneous number of users for class *k* traffic.

Similarly, the mean of ISR from type I mobile users using LBR spreading code is

$$m_{I,k}^{(l)} = M_k \alpha_k^{(l)} n_k, \tag{5.2}$$

where $\alpha_k^{(l)}$ is the connection activity factor for the LBR mini-source in (3.7) and M_k is the maximum number of active LBR spreading codes supported by class *k* traffic.

5.1.1.2 Interference Model for Type II Mobile Users

The mean of ISR from type II mobile users using HBR spreading code, denoted by $I_{II,k}^{(h)} / S_k^{(h)}$, is $m_{II,k}^{(h)}$, where $I_{II,k}^{(h)}$ is the interference power from class *k* traffic using HBR spreading code of type II.

$$m_{II,k}^{(h)} = E\left(\frac{I_{II,k}^{(h)}}{S_k^{(h)}}\right) = \iint_{II} E\left(\frac{\psi_{k,j}^{(h)} L_{0,i}}{L_{b,i}}\right)\rho_k dA = \alpha_k^{(h)} \iint_{II} E\left(\frac{L_{0,i}}{L_{b,i}}\right)\rho_k dA, \tag{5.3}$$

where $\psi_{k,j}^{(h)} \in \{0, 1\}$ is the connection activity random variable (RV) of HBR spreading code for the *j*th mobile user of class *k* traffic, $\alpha_k^{(h)} = \Pr\left(\psi_{k,j}^{(h)} = 1\right)$, A represents the area of a cell, $\rho_k = n_k / A$ is the average number of users of class *k* per unit area, $L_{b,i}$ is the path loss to mobile user *i* from BS *b* defined in (3.1), *b* is the BS index by which mobile user is power-controlled, *i* is herein the index for any mobile user of type II. The first two moments of L_0 / L_b are derived using the condition in (3.3) as following where BS *b* is the BS that power control the mobile user *i*.

Let $X = \ln L_{0,i}$, $Y = \ln L_{b,i}$ and $Z = X - Y$, subject to $Y \geq X \geq Y - \beta\Delta_{\mathrm{dB}}$ which is defined by (3.3), where $\beta = \ln 10/10$. To determine the mean and variance of e^Z, the joint statistics of X and Y should be obtained first. The joint cumulative distribution function of X and Y is given by

$$F_{XY}(x, y) = P\left(X \leq x, Y \leq y \middle| Y \geq X \geq Y - \beta\Delta_{\mathrm{dB}}\right) = \frac{P\left(X \leq \min\left(x, Y - \beta\Delta_{\mathrm{dB}}\right), Y \leq y\right)}{P\left(Y \geq X \geq Y - \beta\Delta_{\mathrm{dB}}\right)}. \tag{5.4}$$

Y is conditioned and the expectation is taken over Y.

$$
\begin{aligned}
&F_{XY}(x, y)P(Y \geq X \geq Y - \beta\Delta_{\mathrm{dB}}) \\
&= E\left[P\left(X \leq \min(x, s - \beta\Delta_{\mathrm{dB}}), s \leq y \middle| Y = s\right)\right] \\
&= \int_{-\infty}^{+\infty} P(s \leq y)P\left(X \leq \min(x, s - \beta\Delta_{\mathrm{dB}})\right) f_Y(s)ds \\
&= \int_{-\infty}^{+\infty} (1 - u(s - y))P\left(X \leq \min(x, s - \beta\Delta_{\mathrm{dB}})\right) f_Y(s)ds.
\end{aligned}
\tag{5.5}
$$

Thus, the joint p.d.f. of X and Y is

$$
\begin{aligned}
f_{XY}(x, y) &= \frac{\partial^2 F_{XY}(x, y)}{\partial x \partial y} = \frac{1}{P(Y \geq X \geq Y - \beta\Delta_{\mathrm{dB}})} \int_{-\infty}^{+\infty} \delta(s - y) f_X(x) f_Y(s)ds \\
&= \frac{1}{P(Y \geq X \geq Y - \beta\Delta_{\mathrm{dB}})} f_X(x) f_Y(y)
\end{aligned}
, \tag{5.6}
$$

where $y \geq x \geq y - \beta\Delta_{\mathrm{dB}}$ and, as X and Y are Gaussian RVs with mean, $-\mu \ln d_{0,i}$ and $-\mu \ln d_{b,i}$, respectively, and standard deviation, $\beta\sigma$,

$$
P(Y \geq X \geq Y - \beta\Delta_{\mathrm{dB}}) = Q\left(\frac{-\mu \ln\left(d_{0,i}/d_{b,i}\right)}{\sqrt{2}\beta\sigma}\right) - Q\left(\frac{-\mu \ln\left(d_{0,i}/d_{b,i}\right) + \beta\Delta_{\mathrm{dB}}}{\sqrt{2}\beta\sigma}\right). \tag{5.7}
$$

where $Q(y) = \int_y^{\infty} \exp\left(-x^2/2\right)/\sqrt{2\pi}\, dx$.

The p.d.f. of $Z = X - Y$ is then given by

$$
f_Z(z) = \int_{-\infty}^{+\infty} f_{XY}(z + y, y)dy\,. \tag{5.8}
$$

Note that $0 \geq Z \geq -\beta\Delta_{\mathrm{dB}}$, then the mean and variance of e^Z are respectively given by

$$
m_{e^Z} = \int_{-\beta\Delta_{\mathrm{dB}}}^{0} e^z f_Z(z)dz\,, \tag{5.9}
$$

and

$$
\sigma_{e^Z}^2 = \int_{-\beta\Delta_{\mathrm{dB}}}^{0} e^{2z} f_Z(z)dz - \mu_{e^Z}^2\,. \tag{5.10}
$$

Similarly, the mean of ISR from type II mobile users using LBR spreading code is

$$
m_{II,k}^{(l)} = E\left(\frac{I_{II,k}^{(l)}}{S_k^{(l)}}\right) = \iint_{II} E\left(\frac{\psi_{k,j}^{(l)} L_{0,i}}{L_{b,i}}\right)\rho_k dA = M_k \alpha_k^{(l)} \iint_{II} E\left(\frac{L_{0,i}}{L_{b,i}}\right)\rho_k dA\,, \tag{5.11}
$$

where $\psi_{k,j}^{(l)} \in \{0, 1, \ldots, M_k\}$ is the connection activity RV of the number of active LBR spreading codes for the *j*th mobile user of class *k* traffic.

5.1.1.3 Interference Models for Type III and IV Mobile Users

Based on the central limit theorem, the inter-cell ISR is assumed as Gaussian distributed with mean, $m_k^{(h)} = m_{III,k}^{(h)} + m_{IV,k}^{(h)}$, and variance, $\left(\sigma_k^{(h)}\right)^2 = \left(\sigma_{III,k}^{(h)}\right)^2 + \left(\sigma_{IV,k}^{(h)}\right)^2$, for class *k* traffic using HBR spreading code, and mean, $m_k^{(l)} = m_{III,k}^{(l)} + m_{IV,k}^{(l)}$, and variance, $\left(\sigma_k^{(l)}\right)^2 = \left(\sigma_{III,k}^{(l)}\right)^2 + \left(\sigma_{IV,k}^{(l)}\right)^2$, for class *k* traffic using LBR spreading codes. The first two moments, $m_{III,k}^{(h)}$ and $\left(\sigma_{III,k}^{(h)}\right)^2$, are derived as following using the condition in (3.3) which is the case whereby the mobile user is power-controlled by the smallest attenuation BS.

$$m_{III,k}^{(h)} = \alpha_k^{(h)} \iint_{III} E\left(\frac{L_{0,i}}{\max\left(L_{b,i}, L_{n,i}\right)}\right)\rho_k dA, \tag{5.12}$$

$$\left(\sigma_{III,k}^{(h)}\right)^2 = \iint_{III} Var\left(\frac{\psi_{k,j}^{(h)} L_{0,i}}{\max\left(L_{b,i}, L_{n,i}\right)}\right)\rho_k dA, \tag{5.13}$$

where *b* and *n* are the BS indexes to which the mobile user is connected, *i* is herein the index for any mobile user of type III. The mobile user is power-controlled by BS *b* or *n* depending on the intensities of their attenuations. To derive the first two moments of $L_{0,i} / \max\left(L_{b,i}, L_{n,i}\right)$, the same technique as that in section 5.1.1.2 is used, except when dealing with $Y = \ln\left\{\max\left(L_{b,i}, L_{n,i}\right)\right\} = \max\left(Y_1, Y_2\right)$ in (3.3), where $Y_1 = \ln L_{b,i}$, $Y_2 = \ln L_{n,i}$, and $L_{b,i} - \Delta_{\mathrm{dB}} \le L_{n,i} \le L_{b,i} + \Delta_{\mathrm{dB}}$. Thus, the p.d.f. of Y can be derived as

$$\begin{aligned} F_Y(y) &= P\left(\max\left(Y_1, Y_2\right) \le y \middle| Y_2 + \beta\Delta_{\mathrm{dB}} \ge Y_1 \ge Y_2 - \beta\Delta_{\mathrm{dB}}\right) \\ &= \frac{P\left(\max\left(Y_1, Y_2\right) \le y, Y_2 + \beta\Delta_{\mathrm{dB}} \ge Y_1 \ge Y_2 - \beta\Delta_{\mathrm{dB}}\right)}{P\left(Y_2 + \beta\Delta_{\mathrm{dB}} \ge Y_1 \ge Y_2 - \beta\Delta_{\mathrm{dB}}\right)}, \end{aligned}$$

where

$$
\begin{aligned}
&P\left(\max(Y_1, Y_2) \le y, Y_2 + \beta\Delta_{\mathrm{dB}} \ge Y_1 \ge Y_2 - \beta\Delta_{\mathrm{dB}}\right) \\
&= P\left(Y_1 \le y, Y_1 \ge Y_2 \ge Y_1 - \beta\Delta_{\mathrm{dB}}\right) + P\left(Y_2 \le y, Y_1 \le Y_2 \le Y_1 + \beta\Delta_{\mathrm{dB}}\right) \quad . \\
&= \int_{-\infty}^{y}\int_{y_1-\beta\Delta_{\mathrm{dB}}}^{y_1} f_{Y_2}(y_2) f_{Y_1}(y_1) dy_2 dy_1 + \int_{-\infty}^{y}\int_{y_2-\beta\Delta_{\mathrm{dB}}}^{y_2} f_{Y_1}(y_1) f_{Y_2}(y_2) dy_1 dy_2
\end{aligned}
$$

Thus,

$$
\begin{aligned}
f_Y(y) &= \frac{\partial F_Y(y)}{\partial y} = \frac{1}{P\left(Y_2 + \beta\Delta_{\mathrm{dB}} \ge Y_1 \ge Y_2 - \beta\Delta_{\mathrm{dB}}\right)} \\
&\times \frac{\int_{-\infty}^{y}\int_{y_1-\beta\Delta_{\mathrm{dB}}}^{y_1} f_{Y_2}(y_2) f_{Y_1}(y_1) dy_2 dy_1 + \int_{-\infty}^{y}\int_{y_2-\beta\Delta_{\mathrm{dB}}}^{y_2} f_{Y_1}(y_1) f_{Y_2}(y_2) dy_1 dy_2}{\partial y} \\
&= \frac{\int_{y_1-\beta\Delta_{\mathrm{dB}}}^{y_1} f_{Y_2}(y_2) f_{Y_1}(y_1) dy_2 \Big|_{y_1=y} + \int_{y_2-\beta\Delta_{\mathrm{dB}}}^{y_2} f_{Y_1}(y_1) f_{Y_2}(y_2) dy_1 \Big|_{y_2=y}}{P\left(Y_2 + \beta\Delta_{\mathrm{dB}} \ge Y_1 \ge Y_2 - \beta\Delta_{\mathrm{dB}}\right)} \\
&= \frac{\left[F_{Y_2}(y_1) - F_{Y_2}(y_1 - \beta\Delta_{\mathrm{dB}})\right] f_{Y_1}(y_1)\Big|_{y_1=y} + \left[F_{Y_1}(y_2) - F_{Y_1}(y_2 - \beta\Delta_{\mathrm{dB}})\right] f_{Y_2}(y_2)\Big|_{y_2=y}}{P\left(Y_2 + \beta\Delta_{\mathrm{dB}} \ge Y_1 \ge Y_2 - \beta\Delta_{\mathrm{dB}}\right)} \\
&= \frac{\left\{F_{Y_2}(y) - F_{Y_2}(y - \beta\Delta_{\mathrm{dB}})\right\} f_{Y_1}(y) + \left\{F_{Y_1}(y) - F_{Y_1}(y - \beta\Delta_{\mathrm{dB}})\right\} f_{Y_2}(y)}{P\left(Y_1 - \beta\Delta_{\mathrm{dB}} \le Y_2 \le Y_1 + \beta\Delta_{\mathrm{dB}}\right)}
\end{aligned}
\quad . \quad (5.14)
$$

where

$$
\begin{aligned}
&P\left(Y_1 - \beta\Delta_{\mathrm{dB}} \le Y_2 \le Y_1 + \beta\Delta_{\mathrm{dB}}\right) \\
&= Q\left\{\frac{-\mu\ln\left(d_{b,i}/d_{n,i}\right) - \beta\Delta_{\mathrm{dB}}}{\sqrt{2}\beta\sigma}\right\} - Q\left\{\frac{-\mu\ln\left(d_{b,i}/d_{n,i}\right) + \beta\Delta_{\mathrm{dB}}}{\sqrt{2}\beta\sigma}\right\} .
\end{aligned}
\qquad (5.15)
$$

Furthermore, because the interference is smaller than the smaller channel gain BS of BSs *b* and *n* in (3.3), during the derivation, the constraint condition should be $X \le \min(Y_1, Y_2)$. Thus,

$$
f_{XY}(x, y) = \frac{1}{P\left(X \le \min(Y_1, Y_2)\right)} f_X(x) f_Y(y), \qquad (5.16)
$$

where

$$
\begin{aligned}
&P\left(X \le \min(Y_1, Y_2)\right) = P(X \le Y_1) P(Y_1 \le Y_2) + P(X \le Y_2) P(Y_1 \ge Y_2) \\
&= Q\left\{\frac{-\mu\ln\left(d_{0,i}/d_{b,i}\right)}{\sqrt{2}\beta\sigma}\right\} Q\left\{\frac{-\mu\ln\left(d_{b,i}/d_{n,i}\right)}{\sqrt{2}\beta\sigma}\right\} + Q\left\{\frac{-\mu\ln\left(d_{0,i}/d_{n,i}\right)}{\sqrt{2}\beta\sigma}\right\} Q\left\{\frac{-\mu\ln\left(d_{n,i}/d_{b,i}\right)}{\sqrt{2}\beta\sigma}\right\} .
\end{aligned}
$$

(5.17)

Equations (5.16) and (5.17) perform the correspondingly similar functions as

(5.6) and (5.7), respectively.

Similarly, the mean and the variance of ISR from type III mobile users using LBR spreading code are

$$m_{III,k}^{(l)} = M_k \alpha_k^{(l)} \iint_{III} E\left(\frac{L_{0,i}}{\max\left(L_{b,i},\, L_{n,i}\right)} \right) \rho_k \, dA, \tag{5.18}$$

$$\left(\sigma_{III,k}^{(l)}\right)^2 = \iint_{III} Var\left(\frac{\psi_{k,j}^{(l)} L_{0,i}}{\max\left(L_{b,i},\, L_{n,i}\right)} \right) \rho_k \, dA, \tag{5.19}$$

The analysis of the interference model for type IV mobile users is similar to those for type II mobile users except the condition (3.2) instead of (3.3) is used. Thus, the constraint for X and Y is $X \le Y - \beta\Delta_{\mathrm{dB}}$. The first two moments, $m_{IV,k}^{(h)}$ and $\left(\sigma_{IV,k}^{(h)}\right)^2$, $m_{IV,k}^{(l)}$ and $\left(\sigma_{IV,k}^{(l)}\right)^2$, are derived as follows.

$$m_{IV,k}^{(h)} = \alpha_k^{(h)} \iint_{IV} E\left(L_{0,i} / L_{b,i}\right) \rho_k \, dA, \tag{5.20}$$

$$\left(\sigma_{IV,k}^{(h)}\right)^2 = \iint_{IV} Var\left(\psi_{k,j}^{(h)} \, L_{0,i} / L_{b,i}\right) \rho_k \, dA, \tag{5.21}$$

and

$$m_{IV,k}^{(l)} = M_k \alpha_k^{(l)} \iint_{IV} E\left(L_{0,i} / L_{b,i}\right) \rho_k \, dA, \tag{5.22}$$

$$\left(\sigma_{IV,k}^{(l)}\right)^2 = \iint_{IV} Var\left(\psi_{k,j}^{(l)} \, L_{0,i} / L_{b,i}\right) \rho_k \, dA, \tag{5.23}$$

where b is the BS index by which the mobile user is power-controlled, i is herein the index for any mobile user of type IV.

5.1.2 SIR Analysis in the Reverse Link

The received E_b / I_0 in BS 0 for mobile users in class k using HBR spreading codes is

$$\left(\frac{E_b}{I_0}\right)_k^{(h)} = \frac{G_k^{(h)} S_k^{(h)}}{\left[\begin{array}{l}\sum_{j=1}^{n_k-1}\left(\psi_{k,j}^{(h)} S_k^{(h)} + \psi_{k,j}^{(l)} S_k^{(l)}\right) + \sum_{c=1,\, c\neq k}^{K} \sum_{j=1}^{n_c}\left(\psi_{c,j}^{(h)} S_c^{(h)} + \psi_{c,j}^{(l)} S_c^{(l)}\right) \\ + \sum_{c=1}^{K}\left(I_{II,c}^{(h)} + I_{II,c}^{(l)}\right) + \sum_{c=1}^{K}\left(I_c^{(h)} + I_c^{(l)}\right) + \eta\end{array}\right]}, \qquad (5.24)$$

where $G_k^{(h)} = W/R_k^{(h)}$ is the spreading gain of class k traffic using HBR spreading code, W is the total spread spectrum bandwidth chip rate, $R_k^{(h)}$ is the transmission rate of class k traffic using HBR spreading code, $S_k^{(h)}$ is the received power of class k traffic using HBR spreading code which is a constant for ideal power control, $S_k^{(l)}$ is the received power of class k traffic using LBR spreading code, $I_{II,k}^{(h)}$ is the intra-cell interference from class k traffic of type II using HBR spreading code, $I_{II,k}^{(l)}$ is the intra-cell interference from class k traffic of type II using LBR spreading code, $I_k^{(h)}$ is the inter-cell interference from class k traffic of types III and IV using HBR spreading code, $I_k^{(l)}$ is the inter-cell interference from class k traffic of types III and IV using LBR spreading code, and η is the background noise. In the denominator, the first term is due to the intra-cell interference of type I from other mobile users in class k. The second term is due to the intra-cell interference of type I from mobile users in other classes. The third term is due to the intra-cell interference of type II. The fourth term is due to the inter-cell interference of types III and IV.

In the home BS, n_k is the number of users for class k traffic whose transmission signals are processed by the BS. As this BS power-controls the mobile user and processes its signals, n_k also represents the number of mobile users power-controlled by the home BS and it is identical in all cells. Then, each BS will connect to $(1+r)n_k$ mobile users that are power-controlled or not power-controlled but connected by the BS, where r is the soft handoff probability (SHP). As the SHP can determine the number of mobile users in each type and further affect the intra- and inter-cell interference, the SHP will influence the SIR value.

The power ratios, $S_k^{(l)}/S_k^{(h)}$, $S_c^{(h)}/S_k^{(h)}$ and $S_c^{(l)}/S_k^{(h)}$ in (5.24), can be obtained using the methods in [100].

$$\frac{S_k^{(l)}}{S_k^{(h)}}=\frac{G_k^{(h)}/\gamma_{R,k}^{*(h)}}{G_k^{(l)}/\gamma_{R,k}^{*(l)}}, \tag{5.25}$$

$$\frac{S_c^{(h)}}{S_k^{(h)}}=\frac{\left(G_k^{(h)}/\gamma_{R,k}^{*(h)}\right)\times[1+M_k\alpha_k^{(l)}/\left(G_k^{(l)}/\gamma_{R,k}^{*(l)}\right)]+\alpha_k^{(h)}}{\left(G_c^{(h)}/\gamma_{R,c}^{*(h)}\right)\times[1+M_c\alpha_c^{(l)}/\left(G_c^{(l)}/\gamma_{R,c}^{*(l)}\right)]+\alpha_c^{(h)}}, \tag{5.26}$$

$$\frac{S_c^{(l)}}{S_k^{(h)}}=\frac{\left(G_k^{(h)}/\gamma_{R,k}^{*(h)}\right)\times[1+M_k\alpha_k^{(l)}/\left(G_k^{(l)}/\gamma_{R,k}^{*(l)}\right)]+\alpha_k^{(h)}}{\left(G_c^{(l)}/\gamma_{R,c}^{*(l)}\right)\times[1+\alpha_c^{(h)}/\left(G_c^{(h)}/\gamma_{R,c}^{*(h)}\right)]+M_c\alpha_c^{(l)}}, \tag{5.27}$$

where $\gamma_{R,k}^{*}$ denotes the SIR requirement of class *k* traffic in the reverse link.

Using the statistics of ISR, $I_{II,k}^{(h)}/S_k^{(h)}$, $I_{II,k}^{(l)}/S_k^{(l)}$, $I_k^{(h)}/S_k^{(h)}$ and $I_k^{(l)}/S_k^{(l)}$ above, the formula for the outage probability can be mathematically derived. Let l_c represent the number of LBR mini-sources for class *c* traffic and h_c represent the number of HBR mini-sources for class *c* traffic. The outage probability is solved by conditioning on the activity factors for LBR and HBR mini-source, e.g., on the state $(l_1,\cdots l_K,h_1,\cdots h_K)$. The probability that the system is operating on this state is $P(l_1,\cdots l_K,h_1,\cdots h_K)$. The probability that the instantaneous SIR is smaller than γ_R^* on this state is $P_{SIR\le\gamma_R^*}(l_1,\cdots l_K,h_1,\cdots h_K)$. Thus, the outage probability for class *k* traffic in the reverse link is

$$\begin{aligned}P_{\text{outage},k}^{R}(\vec{n})&=\Pr\left(SIR_{R,k}\le\gamma_{R,k}^{*}\right)=\Pr\left(SIR_{R,k}^{(h)}\le\gamma_{R,k}^{*(h)}\right)\\&=\sum_{l_1=0}^{M_1n_1}\cdots\sum_{l_k=0}^{M_k(n_k-1)}\cdots\sum_{l_K=0}^{M_Kn_K}\sum_{h_1=0}^{n_1}\cdots\sum_{h_k=0}^{n_k-1}\cdots\sum_{h_K=0}^{n_K}P(l_1,\cdots l_K,h_1,\cdots h_K)\times P_{SIR\le\gamma_R^{*(h)}}(l_1,\cdots l_K,h_1,\cdots h_K),\end{aligned} \tag{5.28}$$

where

$$P(l_1,\cdots l_K,h_1,\cdots h_K)=\binom{n_k-1}{h_k}\left(\alpha_k^{(h)}\right)^{h_k}\left(1-\alpha_k^{(h)}\right)^{n_k-1-h_k}\binom{M_k(n_k-1)}{l_k}\left(\alpha_k^{(l)}\right)^{l_k}\left(1-\alpha_k^{(l)}\right)^{M_k(n_k-1)-l_k}$$
$$\times\prod_{c=1,\,c\neq k}^{K}\left[\binom{M_c n_c}{l_c}\left(\alpha_c^{(l)}\right)^{l_c}\left(1-\alpha_c^{(l)}\right)^{M_c n_c-l_c}\binom{n_c}{h_c}\left(\alpha_c^{(h)}\right)^{h_c}\left(1-\alpha_c^{(h)}\right)^{n_c-h_c}\right],$$

$$P_{SIR\le\gamma_R^{*(h)}}(l_1,\cdots l_K,h_1,\cdots h_K)=Q\left(\frac{z_{R,k}^{(h)}-m_{R,k}^{(h)}}{\sigma_{R,k}^{(h)}}\right),$$

$$m_{R,k}^{(h)}=\sum_{c=1}^{K}\left(h_c\frac{S_c^{(h)}}{S_k^{(h)}}+l_c\frac{S_c^{(l)}}{S_k^{(h)}}\right)+\sum_{c=1}^{K}\left(m_{II,c}^{(h)}\frac{S_c^{(h)}}{S_k^{(h)}}+m_{II,c}^{(l)}\frac{S_c^{(l)}}{S_k^{(h)}}\right)+\sum_{c=1}^{K}\left(m_c^{(h)}\frac{S_c^{(h)}}{S_k^{(h)}}+m_c^{(l)}\frac{S_c^{(l)}}{S_k^{(h)}}\right),$$

$$\sigma_{R,k}^{(h)}=\sqrt{\sum_{c=1}^{K}\left[\left(S_c^{(h)}/S_k^{(h)}\right)^2\left(\sigma_c^{(h)}\right)^2+\left(S_c^{(l)}/S_k^{(h)}\right)^2\left(\sigma_c^{(l)}\right)^2\right]},\text{ and }\; z_{R,k}^{(h)}=G_k^{(h)}/\gamma_{R,k}^{*(h)}-\eta/S_k^{(h)}.$$

5.1.3 Reverse Link - Performance Evaluation

The performance for the reverse link by means of numerical analysis and simulation is evaluated. Four traffic classes are considered. The first, third and fourth classes are modeled as *ON-OFF* sources that are Sen’s model without the active HBR mini-source states and with only one two-state LBR mini-source each representing the voice, web-browsing and data services, while the second class is modeled as Sen’s model with the maximum number of active LBR spreading codes, $M_2=8$, representing the video service. In 3G, WCDMA is characterized by a wide bandwidth of 5 MHz and a constant high chip rate of 3.84 Mcps [2]. The parameters used are the same as those in Chapter 4. The parameters used in the numerical examples are tabulated in Table 5.1.

Table 5.1. The Parameters Values in the Link Layer Analysis

Parameter Type	Voice	Video	Web-browsing	Data
Activity Factor (α_k)	0.4	0.3867 (LBR) 0.5 (HBR)	0.1176	0.1028
Transmission Rate (R_k)	60 kbps	30 kbps (LBR) 60 kbps (HBR)	120 kbps	240 kbps
SIR Requirement ($\gamma^*_{R,k}$)	2 dB	2 dB (LBR) 2 dB (HBR)	3 dB	3 dB
Total Bandwidth (W)		5 MHz		
Shadowing Deviation (σ_ξ)		6 dB		
Path Loss Factor (τ)		4		
BS Number (B)		24*		
S_1/η		-1 dB		

* the 25 BSs (including BS 0) form a 5x5 square cells structure.

In the simulation, the Monte Carlo method is used to obtain the instantaneous interference power and the instantaneous SIR. The outage happens when the achieved SIR is smaller than the SIR requirement.

The analytical and the simulation results of the outage probabilities for classes 1, 2, 3, and 4 are shown in Figs. 5.2, 5.3, 5.4, and 5.5, respectively. In these figures, the number of users for one selected service is varied from 1 to a maximum, while the numbers of users for the other services are fixed. The fixed numbers of users for various services are defined as $n_1 = 10$, $n_2 = 2$, $n_3 = 10$, and $n_4 = 5$. The analytical results are close to the simulation results which verify the correctness of the formulae.

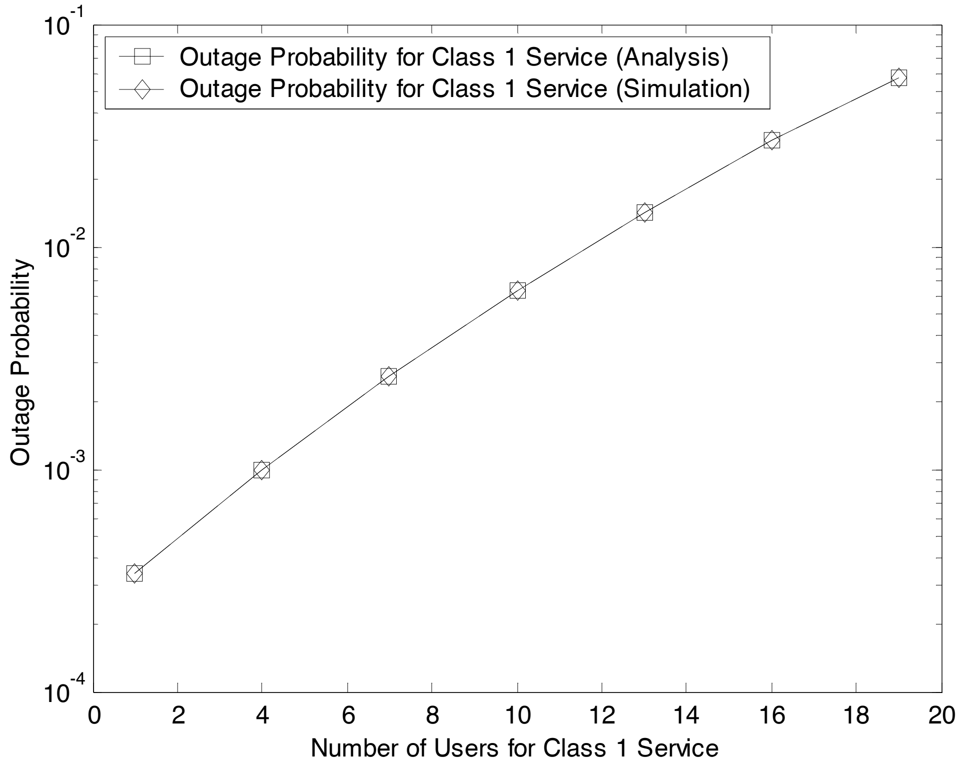

Fig. 5.2. Outage Probability for Class 1 Service

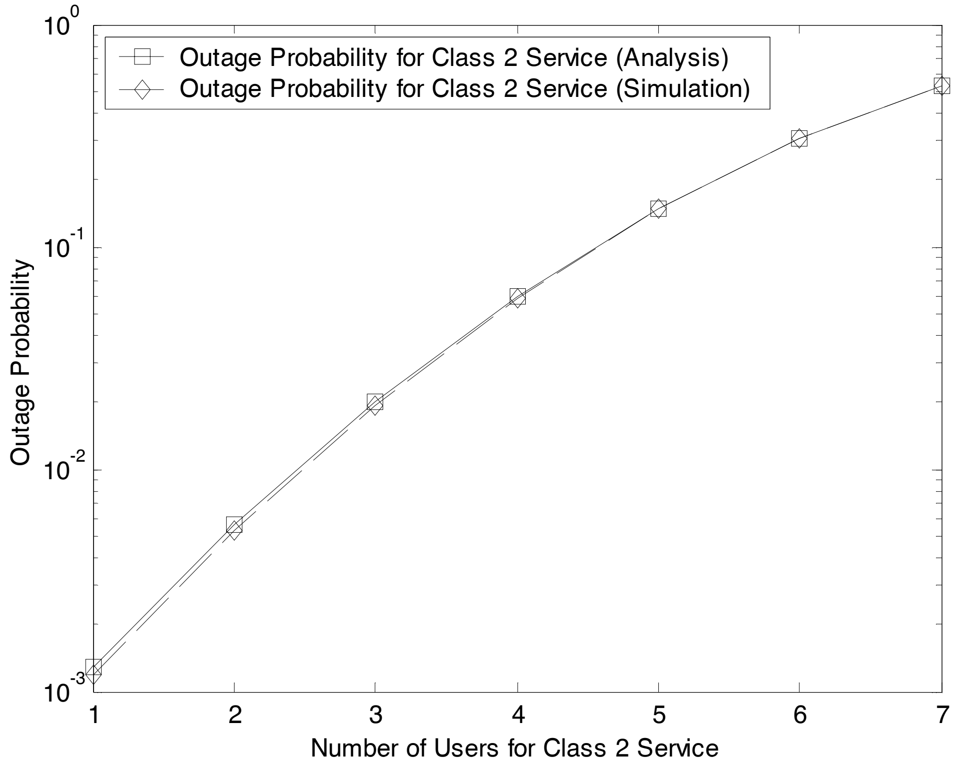

Fig. 5.3. Outage Probability for Class 2 Service

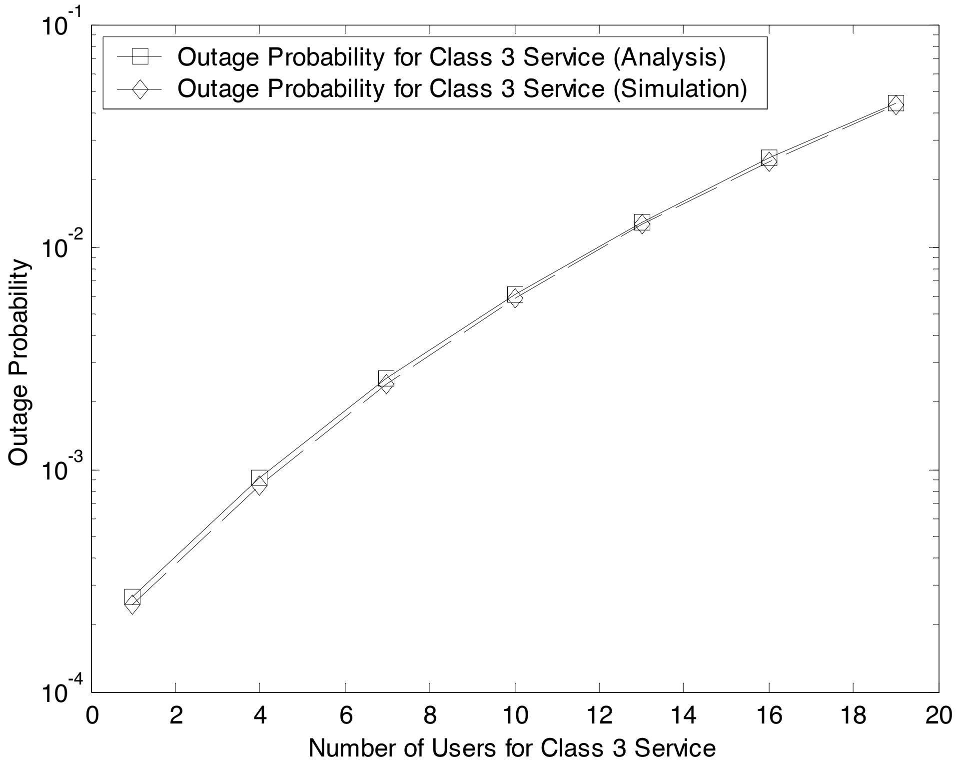

Fig. 5.4. Outage Probability for Class 3 Service

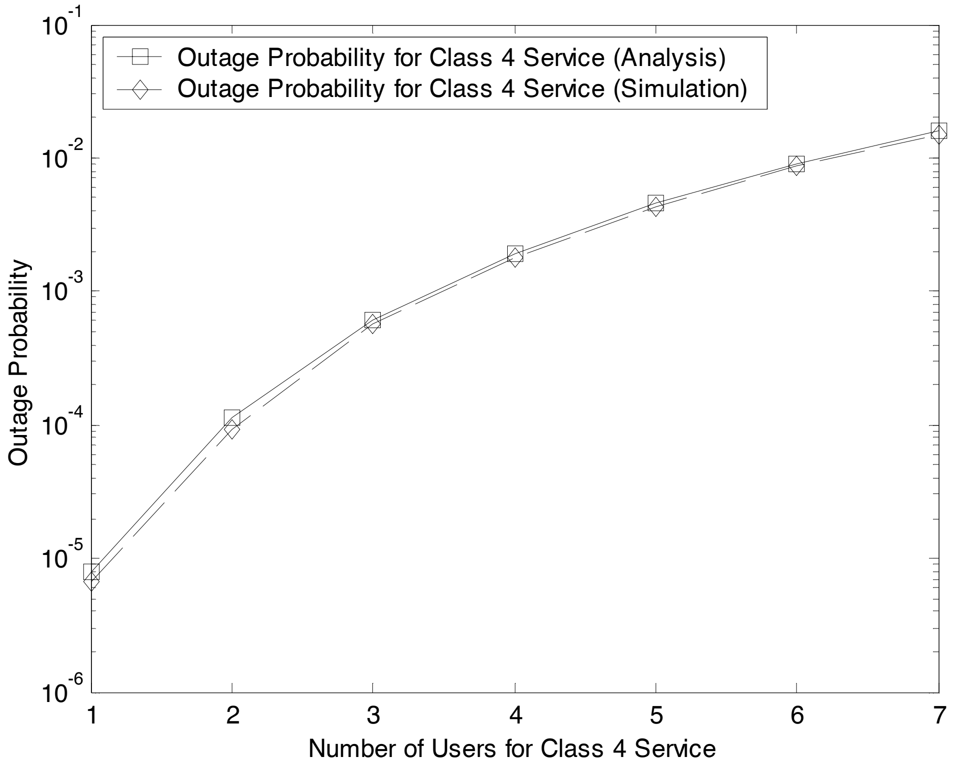

Fig. 5.5. Outage Probability for Class 4 Service

5.2 Evaluation on the Forward Link

The analytical procedure for the forward link is almost the same as that for the reverse link. In the forward link, the signal quality of users in soft handoff is improved by combining the received signals from active BSs that are connected to the mobile users. Such a “diversity gain” can help to reduce transmission power with the signal quality remaining nearly constant. However, this power reduction does not directly imply a capacity improvement because multiple BSs are assigned resources to the same mobile user. Thus, resources assignment schemes for mobile users in soft handoff possess strategical importance to determine the forward link and even the reverse link performance and will lead to various results of the forward and reverse link capacity.

5.2.1 Interference Models in the Forward Link

The forward link geometry is shown in Fig. 5.6. The arrowhead with solid line represents the signal power from the BS to its connecting mobile user. The arrowhead with dotted line represents the interference to the mobile user. Mobile user *a* is in non-soft handoff and connected to BS 0 and mobile user *b* is in soft handoff and connected to BSs 0 and 1.

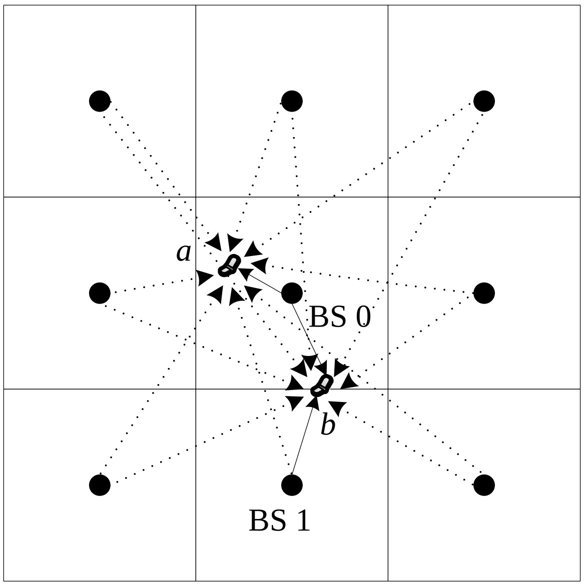

Fig. 5.6. The Forward Link Geometry

The received E_b/I_0 from the given zeroth cell is considered. For mobile user i of class k traffic using HBR spreading code in non-soft handoff, the received E_b/I_0 is

$$\left(\frac{E_b}{I_0}\right)_{k,i,0}^{(h)} = \frac{G_k^{(h)} p_{k,i,0}^{I(h)} L_{0,i}}{\left(p_{\text{total},0} - \alpha_k^{(h)} p_{k,i,0}^{I(h)}\right) L_{0,i} f + \sum_{b=1}^{B} L_{b,i} p_{\text{total},b} + \eta}, \tag{5.29}$$

where $p_{\text{total},b} = \sum_{c=1}^{K}\left[\sum_{j=1}^{\lfloor n_c(1-r)\rfloor}\left(\psi_{c,j}^{(h)} p_{c,j,b}^{I(h)} + \psi_{c,j}^{(l)} p_{c,j,b}^{I(l)}\right) + \sum_{j=1}^{\lfloor 2rn_c\rfloor}\left(\psi_{c,j}^{(h)} p_{c,j,b}^{II(h)} + \psi_{c,j}^{(l)} p_{c,j,b}^{II(l)}\right)\right]$ is the total transmission power from BS b, $p_{c,j,b}^{I}$ is the transmission power to mobile user j of class c traffic in non-soft handoff from BS b, $p_{c,j,b}^{II}$ is the transmission power to mobile user j of class c traffic in soft handoff from BS b, and $\lfloor x \rfloor$ is the floor function which gives the largest integer less than or equal to x. In the denominator, the first term is due to the intra-cell interference and the second term is due to the inter-cell interference.

Assuming that $\left(E_b/I_0\right)_{k,i,0}^{(h)}$ achieves the SIR requirement of class k traffic, denoted by $\gamma_{F,k}^{*(h)}$,

$$G_k^{(h)} p_{k,i,0}^{I(h)} \big/ \gamma_{F,k}^{*(h)} = -\alpha_k^{(h)} p_{k,i,0}^{I(h)} f + \left(f + \sum_{b=1}^{B} \frac{L_{b,i}}{L_{0,i}} \right) p_{\text{total},b} + \eta \,. \tag{5.30}$$

The background noise is negligible compared to the total signal power. Assuming that the number of users is sufficiently large, the term, $\alpha_k^{(h)} p_{k,i,0}^{I(h)} f$, can also be neglected. Thus, the statistical characteristic of ISR is determined by p.d.f. of $\sum_{b=1}^{B} L_{b,i} \big/ L_{0,i}$.

In the reverse link, the inter-cell ISR from large numbers of mobile users is assumed as Gaussian distributed by the central limit theorem. However, the theorem may not be suitable in the forward link as the number of BSs is limited. There has been a general agreement that a sum of independent lognormal RVs can be well approximated by another lognormal RV [101, 102]. And the inter-cell ISR in the forward link has been wildly modeled as lognormal RV in the literature [103-105]. However, with the conditions on ISR, such as (3.4), $L_{b,i} \big/ L_{0,i}$ is actually not a lognormal RV [106]. Thus, the inter-cell ISR subject to the conditions may not be well approximated by a lognormal RV.

In the following, Gaussian and lognormal approximate methods are used to calculate the p.d.f. of $\sum_{b=1}^{B} L_{b,i} \big/ L_{0,i}$ and are further compared with simulation results. Based on the comparison, an approximation selection method is proposed to approximate the inter-cell ISR. To obtain the p.d.f. of $\sum_{b=1}^{B} L_{b,i} \big/ L_{0,i}$, the parameters used are as following: $B = 24$, $\Delta_{\text{dB}} = 6$ dB and $r = 0.36$. User *i* in non-soft handoff is connected to BS 0 and the 25 BSs (including BS 0) form a 5x5 square cells structure. The mobile user is located at a given distance, $d_{0,i}$, from BS 0. The mean and variance of the Gaussian approximation are obtained by matching the first two moments of $\sum_{b=1}^{B} L_{b,i}/L_{0,i}$, by using the first two moments of $L_{b,i}/L_{0,i}$ which can be derived using the same technique as that used in the reverse link analysis. The mean and variance of the lognormal approximation are obtained by using the Fenton-Wilkinson method [101]. Figs. 5.7 and 5.8 show the p.d.f. of $\sum_{b=1}^{B} L_{b,i}/L_{0,i}$ by simulation, and with lognormal and Gaussian approximations, under different shadowing standard deviation, σ, and path loss exponent, μ, of (3.1).

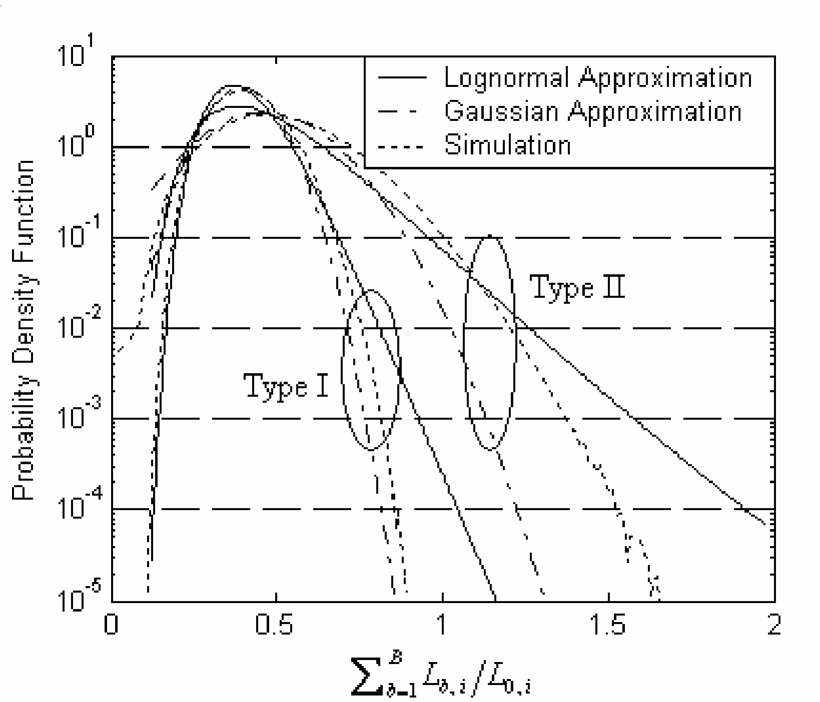

Fig. 5.7. p.d.f. Comparison for Lognormal Approximation, Gaussian Approximation and Simulation with $\mu = 4$, Type I: $\sigma = 2$ dB and Type II: $\sigma = 6$ dB

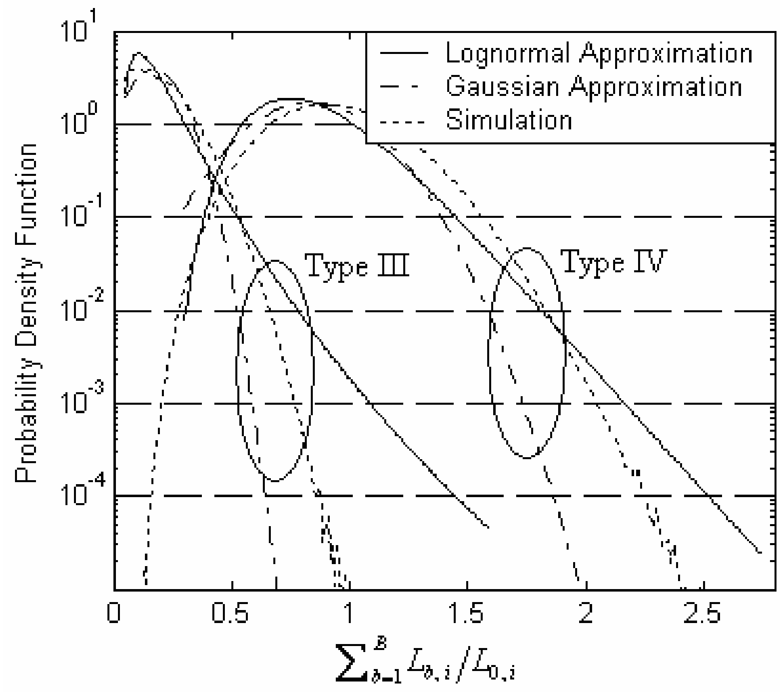

Fig. 5.8. p.d.f. Comparison for Lognormal Approximation, Gaussian Approximation and Simulation with $\sigma = 6$ dB, Type III: $\mu = 6$ and Type IV: $\mu = 3$

When calculating the outage probability, it is the tail part of the p.d.f. of $\sum_{b=1}^{B} L_{b,i}/L_{0,i}$ that is usually of concern. From Figs. 5.7 and 5.8, it can be seen that, in the range of 1×10^{-1} to 1×10^{-3}, the lognormal approximation is closer to the simulation than the Gaussian approximation, while in the range of 1×10^{-3} to 1×10^{-5}, the Gaussian approximation is closer to the simulation than the lognormal approximation, for typical shadowing and path loss parameter values that may be used.

As the concerned range of the tail part of the p.d.f. of $\sum_{b=1}^{B} L_{b,i}/L_{0,i}$ depends on the outage requirement from different services, the selection of lognormal or Gaussian approximation is determined by the QoS specification of these services. Thus, services with a large outage requirement, such as the voice service, will apply the lognormal approximation, and services with a small outage requirement, such as the email and web browsing services, will apply the Gaussian approximation. Thus, different services with different outage requirements will select different approximate methods. The proposed approximation selection method fully utilizes and handily combines the advantages of both lognormal and Gaussian approximate methods.

From (5.29), assuming that $\left(E_b/I_0\right)_{k,i,0}^{(h)}$ achieves the SIR requirement of class k traffic using HBR spreading code, denoted by $\gamma_{F,k}^{*(h)}$, the transmission power ratio of mobile user j of class c traffic in non-soft handoff using HBR spreading code to mobile user i can be represented as $p_{c,j,0}^{I(h)} \Big/ p_{k,i,0}^{I(h)} = \varphi_{1,j}\left(R_c^{(h)}\gamma_{F,c}^{*(h)}\right) \Big/ \left(R_k^{(h)}\gamma_{F,k}^{*(h)}\right)$, where $\varphi_{1,j} = \left(f + \sum_{b=1}^{B} L_{b,j}/L_{0,j}\right) \Big/ \left(f + \sum_{b=1}^{B} L_{b,i}/L_{0,i}\right)$. Similarly, the transmission power ratio of mobile user j of class c traffic in non-soft handoff using LBR spreading code to mobile user i can be represented as $p_{c,j,0}^{I(l)} \Big/ p_{k,i,0}^{I(h)} = \varphi_{1,j}\left(R_c^{(l)}\gamma_{F,c}^{*(l)}\right) \Big/ \left(R_k^{(h)}\gamma_{F,k}^{*(h)}\right)$.

5.2.2 Power Control Schemes in the Forward Link

Next, mobile users in soft handoff are considered. With MRC, the combined E_b/I_0 at the output is the linear sum of the received E_b/I_0. Thus, the combined E_b/I_0 of mobile user i of class k traffic in soft handoff using HBR spreading code is

$$\begin{aligned}\left(E_b/I_0\right)_{k,i}^{(h)} &= \left(E_b/I_0\right)_{k,i,0}^{II(h)} + \left(E_b/I_0\right)_{k,i,1}^{II(h)} \\ &= \frac{G_k^{(h)} p_{k,i,0}^{II(h)}}{\left(f + \sum_{b=1}^{B} \frac{L_{b,i}}{L_{0,i}}\right) p_{\mathrm{total},b}} + \frac{G_k^{(h)} p_{k,i,1}^{II(h)}}{\left(f + \sum_{b=0,\, b\neq 1}^{B} \frac{L_{b,i}}{L_{1,i}}\right) p_{\mathrm{total},b}}.\end{aligned} \tag{5.31}$$

In (5.31), the combined E_b/I_0 could achieve its SIR requirement with multiple combinations of transmission power allocations from the two active BSs. The power control schemes thus will greatly influence the SIR in the forward link as the transmission power from the two active BSs behaves like interference to other mobile users. There are at least four power control schemes for mobile users in soft handoff. They are the equal E_b/I_0 power control, the balancing power control, the unbalancing power control and the site selection diversity transmission power control (SSDT). The comparisons of these four schemes are presented below.

For the equal E_b/I_0 power control, the E_b/I_0s contributed from the two active BSs are the same. Thus, $(E_b/I_0)_{k,i,0}^{II(h)} = (E_b/I_0)_{k,i,1}^{II(h)}$ and $(E_b/I_0)_{k,i}^{(h)} = (E_b/I_0)_{k,i,0}^{II(h)} + (E_b/I_0)_{k,i,1}^{II(h)} = 2(E_b/I_0)_{k,i,1}^{II(h)}$.

For the balancing power control, the powers transmitted from the two active BSs are the same. Thus, $p_{k,i,0}^{II(h)} = p_{k,i,1}^{II(h)}$.

For the unbalancing power control, various levels of power are transmitted from the two active BSs. The differentiated powers are adaptively weighted to the channel gains from the active BSs to the mobile user. $p_{k,i,0}^{II(h)} / p_{k,i,1}^{II(h)} = L_{0,i}/L_{1,i}$ is taken as an example of the unbalancing power control.

For SSDT, the differentiated powers are allocated in an extreme way that the active BS with smaller channel gain will not transmit any signal to the mobile user while another active BS will provide full SIR requirement for the mobile user, that is, if $L_{0,i} < L_{1,i}$, $p_{k,i,0}^{II(h)} = 0$ and $p_{k,i,1}^{II(h)} = P$. Thus, $(E_b/I_0)_{k,i}^{(h)} = (E_b/I_0)_{k,i,1}^{II(h)}$.

These schemes have been presented in [74-81]. However, the performance comparison between them is not yet clear. To compare the performance of different schemes, the sum of transmission powers from two active BSs to a certain mobile user in soft handoff is chosen as the criterion. The scheme with smaller transmission power will introduce less interference to other mobile users and therefore provide better performance to the whole networks. The ratio of the sum of transmission powers from two active BSs to the mobile user j in soft handoff, denoted by $p_{k,j}^{II(h)} = p_{k,j,0}^{II(h)} + p_{k,j,1}^{II(h)}$, to the transmission power to another mobile user i in non-soft handoff, denoted by $p_{k,i,0}^{I(h)}$ can be obtained as follows.

For the equal E_b/I_0 power control, $\left(E_b/I_0\right)_{k,i,0}^{II(h)} = \left(E_b/I_0\right)_{k,i,1}^{II(h)} = \left(E_b/I_0\right)_{k,i}^{(h)}\Big/2$. Thus,

$$G_k^{(h)} p_{k,j,0}^{II(h)} \Big/ \gamma_{F,k}^{*(h)} = \left(f + \sum_{b=1}^{B} \frac{L_{b,j}}{L_{0,j}} \right) p_{\text{total},b} \Bigg/ 2, \tag{5.32}$$

and

$$G_k^{(h)} p_{k,j,1}^{II(h)} \Big/ \gamma_{F,k}^{*(h)} = \left(f + \sum_{b=0,b\neq 1}^{B} \frac{L_{b,j}}{L_{1,j}} \right) p_{\text{total},b} \Bigg/ 2. \tag{5.33}$$

With (5.29), (5.32) and (5.33),

$$\frac{p_{k,j}^{II(h)}}{p_{k,i}^{I(h)}} = \frac{f + \left(\sum_{b=1}^{B} \frac{L_{b,j}}{L_{0,j}} + \sum_{b=0,\, b\neq 1}^{B} \frac{L_{b,j}}{L_{1,j}} \right) \Bigg/ 2}{f + \sum_{b=1}^{B} \frac{L_{b,i}}{L_{0,i}}}. \tag{5.34}$$

For the balancing power control, $p_{k,i,0}^{II(h)} = p_{k,i,1}^{II(h)}$. Thus,

$$G_k^{(h)} p_{k,i}^{II(h)} \Big/ \gamma_{F,k}^{*(h)} = 2 \Bigg/ \left(\frac{1}{\left(f + \sum_{b=1}^{B} \frac{L_{b,i}}{L_{0,i}} \right) p_{\text{total},b}} + \frac{1}{\left(f + \sum_{b=0,b\neq 1}^{B} \frac{L_{b,i}}{L_{1,i}} \right) p_{\text{total},b}} \right). \tag{5.35}$$

With (5.29) and (5.35),

$$\frac{p_{k,j}^{II(h)}}{p_{k,i}^{I(h)}} = 2 \Bigg/ \left(\frac{\left(f + \sum_{b=1}^{B} \frac{L_{b,i}}{L_{0,i}} \right)}{\left(f + \sum_{b=1}^{B} \frac{L_{b,j}}{L_{0,j}} \right)} + \frac{\left(f + \sum_{b=1}^{B} \frac{L_{b,i}}{L_{0,i}} \right)}{\left(f + \sum_{b=0,b\neq 1}^{B} \frac{L_{b,j}}{L_{1,j}} \right)} \right). \tag{5.36}$$

For the unbalancing power control, $p_{k,i,0}^{II(h)} \big/ p_{k,i,1}^{II(h)} = L_{0,i} \big/ L_{1,i}$. Thus,

$$G_k^{(h)} p_{k,i}^{II(h)} \big/ \gamma_{F,k}^{*(h)} = 2 \Bigg/ \left(\frac{1 \Big/ \left(1 + \frac{L_{1,i}}{L_{0,i}}\right)}{\left(f + \sum_{b=1}^{B} \frac{L_{b,i}}{L_{0,i}}\right) p_{\mathrm{total},b}} + \frac{1 \Big/ \left(1 + \frac{L_{0,i}}{L_{1,i}}\right)}{\left(f + \sum_{b=0,\, b\neq 1}^{B} \frac{L_{b,i}}{L_{1,i}}\right) p_{\mathrm{total},b}} \right). \quad (5.37)$$

With (5.29) and (5.37),

$$\frac{p_{k,j}^{II(h)}}{p_{k,i}^{I(h)}} = \frac{L_{0,j} + L_{1,j}}{\dfrac{f + \sum_{b=1}^{B} \frac{L_{b,i}}{L_{0,i}}}{f + \sum_{b=1}^{B} \frac{L_{b,j}}{L_{0,j}}} L_{0,j} + \dfrac{f + \sum_{b=1}^{B} \frac{L_{b,i}}{L_{0,i}}}{f + \sum_{b=0,\, b\neq 1}^{B} \frac{L_{b,j}}{L_{1,j}}} L_{1,j}}. \quad (5.38)$$

The comparison of the transmission power among these three schemes is calculated at follows. Let $x = \left(f + \sum_{b=1}^{B} \frac{L_{b,i}}{L_{0,i}}\right) \Big/ \left(f + \sum_{b=1}^{B} \frac{L_{b,j}}{L_{0,j}}\right)$ and $y = \left(f + \sum_{b=1}^{B} \frac{L_{b,i}}{L_{0,i}}\right) \Big/ \left(f + \sum_{b=0,\, b\neq 1}^{B} \frac{L_{b,j}}{L_{1,j}}\right)$. Then, For the equal E_b / I_0 power control, $\frac{p_{k,j}^{II(h)}}{p_{k,i}^{I(h)}} = \frac{1}{2x} + \frac{1}{2y}$. For the balancing power control, $\frac{p_{k,j}^{II(h)}}{p_{k,i}^{I(h)}} = \frac{2}{x+y}$. For the unbalancing power control, $\frac{p_{k,j}^{II(h)}}{p_{k,i}^{I(h)}} = \frac{L_{0,j} + L_{1,j}}{xL_{0,j} + yL_{1,j}}$.

Because $\left(\frac{1}{2x} + \frac{1}{2y}\right) \Big/ \left(\frac{2}{x+y}\right) = \frac{(x+y)^2}{4xy} = \frac{(x-y)^2 + 4xy}{4xy} \geq 1$, the sum of transmission power from two active BSs in (5.36) is smaller than that in (5.34).

If $L_{0,j} \geq L_{1,j}$, then $x \geq y$, and vice versa. As a result, $\frac{2}{x+y} - \frac{L_{0,j} + L_{1,j}}{xL_{0,j} + yL_{1,j}} = \frac{(x-y)(L_{0,j} - L_{1,j})}{(xL_{0,j} + yL_{1,j})(x+y)} \geq 0$. Thus, the sum of transmission power from two active BSs in (5.38) is smaller than that in (5.36).

The unbalancing power control is superior to the balancing power control and the balancing power control is superior to the equal E_b/I_0 power control.

For SSDT, because the only BS with the largest channel gain transmits signal to the mobile user regardless of whether the mobile user is in soft handoff or not, the analysis on the transmission power for user in soft handoff is the same as that for mobile user in non-soft handoff. Thus, with SSDT, the analysis in the forward link will be the same as that in the reverse link in section 5.1, i.e., soft handoff in the forward link will improve system performance by decreasing the inter-cell interference.

Because SSDT maximally reduces the total transmission power to the mobile users in soft handoff through maintaining only one connection between the active BSs and the mobile user at all times, SSDT is superior to other three schemes in the forward link.

In 3GPP, the balancing power control and SSDT are proposed as the recommended power control schemes for mobile users in soft handoff [76]. The differences between the balancing power control and SSDT are

1. In the forward link, the total transmission power to the mobile user in soft handoff is minimized in SSDT so that the forward link capacity is maximized. However, because the capability for tracking the changing primary BS degrades with the increase in the Doppler frequency, the gain derived from SSDT over the balancing power allocation decreases as the Doppler frequencies increase;
2. In the reverse link, due to the need in transmitting the site selection messages in SSDT, the overhead signaling in the reverse link are increased. It has been

reported that the degradation estimated by the simulations is 0.1-0.2 dB in the reverse link capacity [107].

The unbalancing power control provides a compromise between the balancing power control and SSDT. However, due to the complexity in the application of the unbalancing power control, more work should be done before standardizing this scheme for usage in practical systems.

In this thesis, the balancing power control scheme is employed. Thus, the transmission power ratio of mobile user j of class c traffic in soft handoff using HBR spreading code to mobile user i is $p_{c,\,j,\,0}^{II(h)} \big/ p_{k,\,i,\,0}^{I(h)} = \varphi_{2,\,j}\left(R_c^{(h)}\gamma_{F,\,c}^{*(h)}\right) \big/ \left(R_k^{(h)}\gamma_{F,\,k}^{*(h)}\right)$, where $\dfrac{1}{\varphi_{2,\,j}} = \left(f + \sum_{b=1}^{B}\dfrac{L_{b,\,i}}{L_{0,\,i}}\right) \Bigg/ \left(f + \sum_{b=1}^{B}\dfrac{L_{b,\,j}}{L_{0,\,j}}\right) + \left(f + \sum_{b=1}^{B}\dfrac{L_{b,\,i}}{L_{0,\,i}}\right) \Bigg/ \left(f + \sum_{b=0,\,b\neq 1}^{B}\dfrac{L_{b,\,j}}{L_{1,\,j}}\right)$. Similarly, the transmission power ratio of mobile user j of class c traffic in soft handoff using LBR spreading code to mobile user i is $p_{c,\,j,\,0}^{II(l)} \big/ p_{k,\,i,\,0}^{I(h)} = \varphi_{2,\,j}\left(R_c^{(l)}\gamma_{F,\,c}^{*(l)}\right) \big/ \left(R_k^{(h)}\gamma_{F,\,k}^{*(h)}\right)$.

To derive the first two moments of $\sum_{b=1}^{B} L_{b,\,i} \big/ L_{0,\,i}$ with the condition (3.4), the same technique is applied as that in the reverse link analysis, except when dealing with $b = 1$, so that BSs 0 and 1 are the two active BSs. Thus, $X = \ln L_{1,i}$, $Y = \ln L_{0,i}$ and

$$f_{XY}(x,\,y) = \frac{1}{P(Y - \beta\Delta_{\mathrm{dB}} \le X \le Y + \beta\Delta_{\mathrm{dB}})} f_X(x) f_Y(y), \qquad (5.39)$$

where $y - \beta\Delta_{\mathrm{dB}} \le x \le y + \beta\Delta_{\mathrm{dB}}$ and

$$\begin{aligned} &P(Y - \beta\Delta_{\mathrm{dB}} \le X \le Y + \beta\Delta_{\mathrm{dB}}) \\ &= Q\left\{\frac{-\mu\ln\left(d_{1,i}/d_{0,i}\right) - \beta\Delta_{\mathrm{dB}}}{\sqrt{2}\beta\sigma}\right\} - Q\left\{\frac{-\mu\ln\left(d_{1,i}/d_{0,i}\right) + \beta\Delta_{\mathrm{dB}}}{\sqrt{2}\beta\sigma}\right\}. \end{aligned} \qquad (5.40)$$

The mean and variance of e^Z are respectively given by

$$\mu_{e^Z} = \int_{-\beta\Delta_{\mathrm{dB}}}^{+\beta\Delta_{\mathrm{dB}}} e^z f_Z(z)\,dz\,, \qquad (5.41)$$

and

$$\sigma_{e^z}^2 = \int_{-\beta\Delta_{\text{dB}}}^{+\beta\Delta_{\text{dB}}} e^{2z} f_Z(z)dz - \mu_{e^z}^2 . \tag{5.42}$$

Equations (5.39), (5.40), (5.41) and (5.42) perform the correspondingly similar functions as (5.6), (5.7), (5.9) and (5.10), respectively, in the reverse link.

5.2.3 SIR Analysis in the Forward Link

The outage probability, that is, the probability that the SIR is smaller than γ_F^*, is presented. The perfect SIR-based power control scheme assumption makes the SIR for the same service to remain the same for all mobile users, regardless of whether it is in soft handoff or in non-soft handoff. The outage probability for mobile user *i* of class *k* traffic in non-soft handoff using HBR spreading code in the forward link is

$$\begin{aligned} P_{\text{outage},k,i}^F(\vec{n}) &= \Pr\left(SIR_{F,k} \le \gamma_{F,k}^*\right)_i = \Pr\left(SIR_{F,k}^{(h)} \le \gamma_{F,k}^{*(h)}\right)_i \\ &= \sum_{l_1=0}^{\lfloor M_1 n_1(1+r)\rfloor} \cdots \sum_{l_K=0}^{\lfloor M_K n_K(1+r)\rfloor} \sum_{h_1=0}^{\lfloor n_1(1+r)\rfloor} \cdots \sum_{h_K=0}^{\lfloor n_K(1+r)\rfloor} P(l_1,\cdots l_K, h_1,\cdots h_K) \times P_{SIR\le\gamma_F^{*(h)}}(l_1,\cdots l_K, h_1,\cdots h_K), \end{aligned} \tag{5.43}$$

where

$$\begin{aligned} P_{SIR\le\gamma_F^{*(h)}}(l_1,\cdots l_K, h_1,\cdots h_K) &= \Pr\left[\frac{G_k^{(h)}}{\gamma_{F,k}^{*(h)}} \le \left(f + \sum_{b=1}^{B}\frac{L_{b,i}}{L_{0,i}}\right)\frac{p_{\text{total},b}}{p_{k,i,0}^{I(h)}}\right] \\ &= \Pr\left\{\frac{G_k^{(h)}}{\gamma_{F,k}^{*(h)}} \le \left(f + \sum_{b=1}^{B}\frac{L_{b,i}}{L_{0,i}}\right)\sum_{c=1}^{K}\left[h_c\left(\frac{1-r}{1+r}\varphi_1 + \frac{2r}{1+r}\varphi_2\right)\frac{R_c^{(h)}\gamma_{F,c}^{*(h)}}{R_k^{(h)}\gamma_{F,k}^{*(h)}}\right.\right. \\ &\quad \left.\left. + l_c\left(\frac{1-r}{1+r}\varphi_1 + \frac{2r}{1+r}\varphi_2\right)\frac{R_c^{(l)}\gamma_{F,c}^{*(l)}}{R_k^{(h)}\gamma_{F,k}^{*(h)}}\right]\right\}, \\ &= \Pr\left\{W\Big/\left[\left(\frac{1-r}{1+r}\varphi_1 + \frac{2r}{1+r}\varphi_2\right)\sum_{c=1}^{K}\left(h_c R_c^{(h)}\gamma_{F,c}^{*(h)} + l_c R_c^{(l)}\gamma_{F,c}^{*(l)}\right)\right] - f \le \sum_{b=1}^{B}\frac{L_{b,i}}{L_{0,i}}\right\} \\ &= Q\left(\frac{z_{F,i} - m_{F,i}}{\sigma_{F,i}}\right) \end{aligned}$$

$$P(l_1,\cdots l_K,h_1,\cdots h_K)=$$
$$\prod_{c=1}^{K}\left[\binom{\lfloor M_c n_c(1+r)\rfloor}{l_c}\left(\alpha_c^{(l)}\right)^{l_c}\left(1-\alpha_c^{(l)}\right)^{\lfloor M_c n_c(1+r)\rfloor-l_c}\binom{\lfloor n_c(1+r)\rfloor}{h_c}\left(\alpha_c^{(h)}\right)^{h_c}\left(1-\alpha_c^{(h)}\right)^{\lfloor n_c(1+r)\rfloor-h_c}\right],$$

$$z_{F,i}=\ln\left\{W\Big/\left[\left(\frac{1-r}{1+r}\varphi_1+\frac{2r}{1+r}\varphi_2\right)\sum_{c=1}^{K}\left(h_c R_c^{(h)}\gamma_{F,c}^{*(h)}+l_c R_c^{(l)}\gamma_{F,c}^{*(l)}\right)\right]-f\right\}, \quad (5.44)$$

for the lognormal approximation, and

$$z_{F,i}=W\Big/\left[\left(\frac{1-r}{1+r}\varphi_1+\frac{2r}{1+r}\varphi_2\right)\sum_{c=1}^{K}\left(h_c R_c^{(h)}\gamma_{F,c}^{*(h)}+l_c R_c^{(l)}\gamma_{F,c}^{*(l)}\right)\right]-f, \quad (5.45)$$

for the Gaussian approximation. φ_1 is the mean of $\varphi_{1,j}$, φ_2 is the mean of $\varphi_{2,j}$, and $m_{F,i}$ and $\sigma_{F,i}^2$ are respectively the mean and variance of $\sum_{b=1}^{B}L_{b,i}/L_{0,i}$. In evaluating the values of φ_1 and φ_2, a simplification is made. The smallest distance rather than the smallest attenuation is used to determine the distributions of users that are in soft handoff or non-soft handoff. The calculation is much simplified and the estimations of the outage probability are only slightly higher [22].

5.3 Admission Region

5.3.1 The Formulae for Admission Regions

In Chapter 5, the term ‘admission region’ will be used instead of ‘physical capacity’ defined in Chapter 4. The difference between these two terms lies in two aspects. 1) The physical capacity used in the network layer analysis represent the capacity not only in CDMA system but also in other cellular systems, such as FDMA and TDMA systems, etc., and the network layer model in Chapter 4 can be generally applied in many cellular systems. However, the admission region defined here is only for CDMA system. 2) In the previous network layer analysis, the physical capacity presents the constraint of the number of instantaneously transmitted packets of admitted calls, which neglects statistical multiplexing in the lower layers. However, in CDMA system, the statistical multiplexing is automatically applied in all layers. As a result, with consideration to these two aspects, the ‘admission region’ in the link layer and the ‘physical capacity’ in the network layer are used to differentiate the difference.

The admission region in the link layer is defined as that, within the admission region, the outage probability is smaller than its requirement value, δ_k. That is, the admission region in the reverse link, denoted by $\mathbf{A}_R$, is given by

$$\mathbf{A}_R = \left\{ \vec{n} \left| P^R_{\text{outage},k}(\vec{n}) \le \delta_{R,k},\ k = 1, \ldots, K \right. \right\}. \tag{5.46}$$

The admission region in the forward link, denoted by $\mathbf{A}_F$, is given by

$$\mathbf{A}_F = \left\{ \vec{n} \left| P^F_{\text{outage},k}(\vec{n}) \le \delta_{F,k},\ k = 1, \ldots, K \right. \right\}. \tag{5.47}$$

The admission region for the link layer considering both the reverse and forward links, denoted by $\mathbf{A}$, is given by

$$\mathbf{A} = \left\{ \vec{n} \left| P_{\text{outage},k}(\vec{n}) \le \delta_k,\ k = 1, \ldots, K \right. \right\}, \tag{5.48}$$

where $P_{\text{outage},k}(\vec{n}) = \max\left(P^R_{\text{outage},k}(\vec{n}),\ P^F_{\text{outage},k}(\vec{n})\right)$.

5.3.2 Admission Region in the Reverse Link

To obtain the admission regions in the reverse link, the traffic models for 4 classes services used in Chapter 4 is assumed. The parameters used are the same as those defined in Table 5.1. Besides, $\delta_{R,k} = 1.0 \times 10^{-3}$. The admission regions in the reverse link with the variation of SHP from 0 to 0.7 for classes 1 and 2, and classes 3 and 4 are shown in Figs. 5.9 and 5.10, respectively.

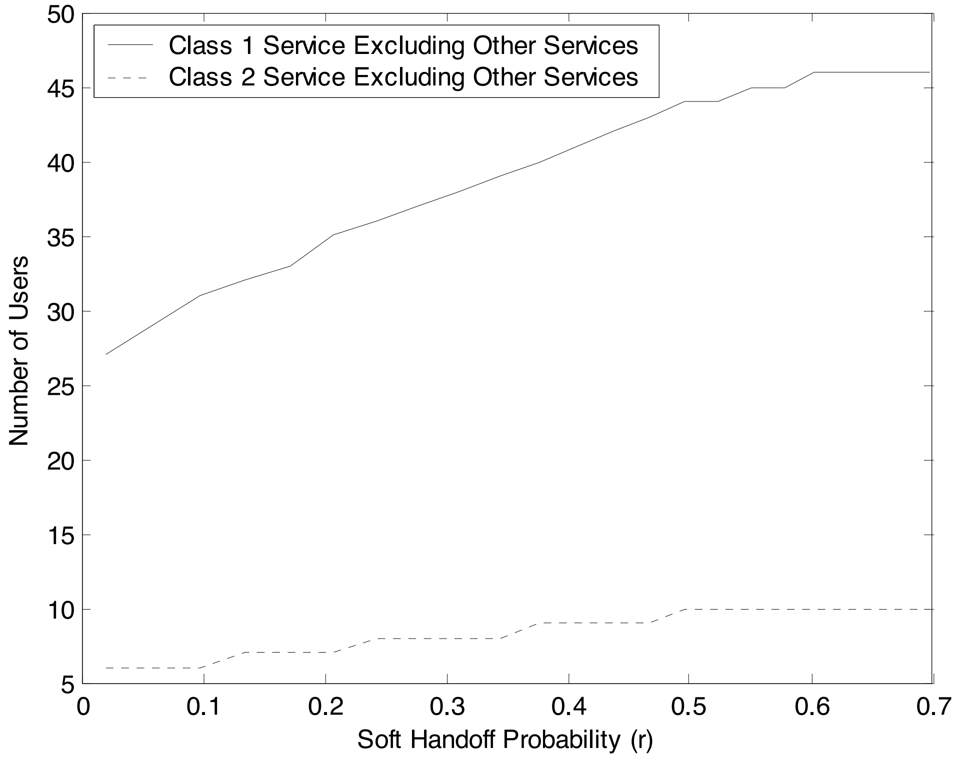

Fig. 5.9. Admission Region in the Reverse Link for Classes 1 and 2

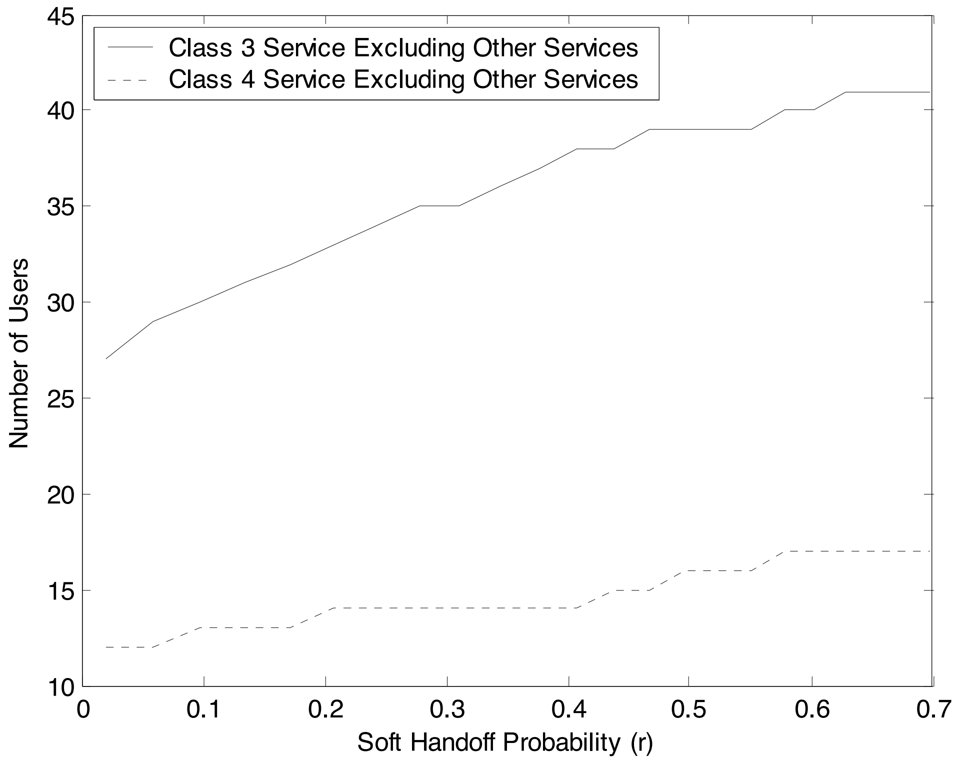

Fig. 5.10. Admission Region in the Reverse Link for Classes 3 and 4

When SHP increases from its low value, the admission regions for classes 1, 2, 3, and 4 increase too. However, with the further increase of SHP beyond a certain point (the knee of the curve), the increase of the admission regions become non-obvious. The observation can be explained as follows. With the increase of SHP, mobile users in soft handoff increase and these mobile users will be power-controlled by the larger channel gain BS. Therefore, mobile users in soft handoff lead to less transmission power and the interference is reduced. SIR behaves better and the admission regions increase accordingly. However, when SHP becomes extremely large, the influence from the soft handoff on the interference decreases as mobile users that lean to one BS and are far away from other BSs would more likely be power-controlled by the most nearby BS. This phenomenon is equivalent to that happening in the non-soft handoff case. Thus, to keep increasing SHP becomes meaningless as a large SHP cannot bring gains for the system any more. The knee of the curve is at around $r = 0.5 \sim 0.6$.

5.3.3 Admission Region in the Forward Link

The same traffic models are assumed in both the forward and reverse links. Besides the parameters defined in the reverse link in Table 5.1, the parameters values in the forward link are chosen as: $\gamma^*_{F,1} = \gamma^{*(l)}_{F,2} = \gamma^{*(h)}_{F,2} = 2$ dB , $\gamma^*_{F,3} = \gamma^*_{F,4} = 3$ dB , $\delta_{F,k} = 1.0 \times 10^{-3}$, and $f = 0.2$.

To investigate the proposed approximation selection method in section 5.2.1, the admission region in the forward link at $r = 0.36$ for class 1 and class 4 services excluding class 2 and class 3 services is presented. Fig. 5.11 shows the four admission regions, including the results obtained with the Gaussian approximation, the proposed approximation selection method and the lognormal approximation under the practical BS selection criterion and the result obtained under the nearest BS selection criterion.

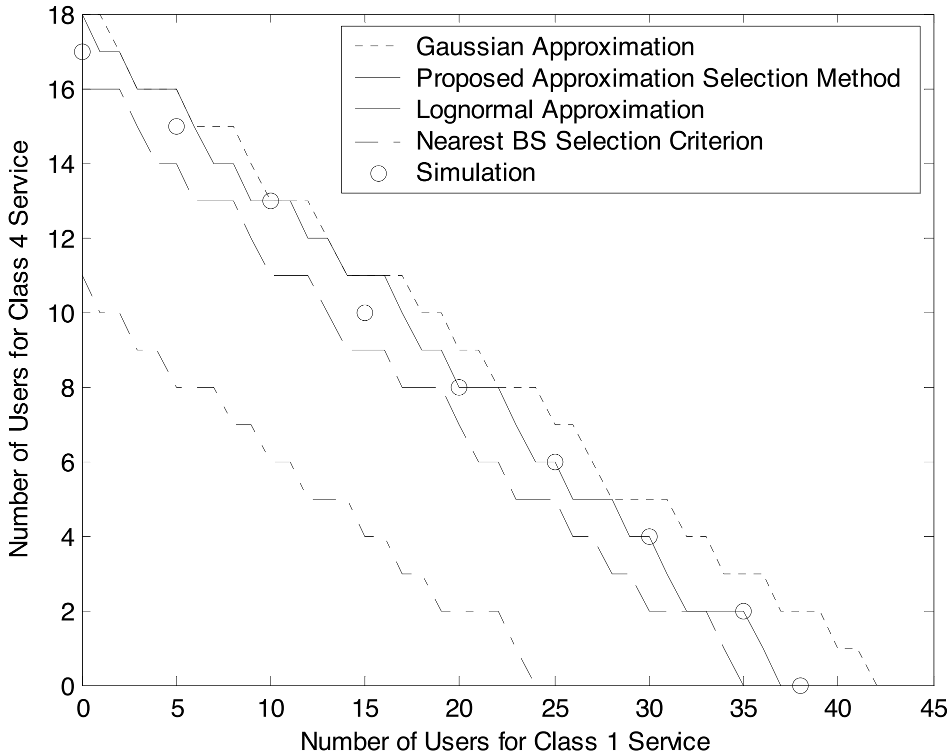

Fig. 5.11. Comparison among the Approximation Methods

The admission region obtained with the Gaussian approximation is larger than that obtained with the proposed approximation selection method and the admission region obtained with the proposed method is larger than that obtained with the lognormal approximation. The observation can be explained as follows. From Figs. 5.7 and 5.8, it can be seen that the lognormal approximation gives a higher estimation whilst the Gaussian approximation gives a lower estimation in the tail part of the p.d.f. of $\sum_{m=1}^{M} L_{m,i} / L_{0,i}$. Thus, the use of lognormal approximation gives conservative outage estimation, which leads to a smaller admission region, and the use of Gaussian approximation gives optimistic outage estimation, which leads to a larger admission region. However, compared to the Gaussian and lognormal approximations, the proposed approximation selection method combines the advantages of these two approximations and gives more appropriate estimation on the forward link capacity. In Fig. 5.11, the estimation by the proposed approximation selection method also matches the simulation results very well. Furthermore, there appears a significant difference between the result with the nearest BS selection criterion and the results with the practical BS selection criterion, which justifies the need to develop more accurate interference models.

The admission regions in the forward link with the change of SHP from 0 to 0.7 for classes 1 and 2, and classes 3 and 4 are shown in Figs. 5.12 and 5.13, respectively.

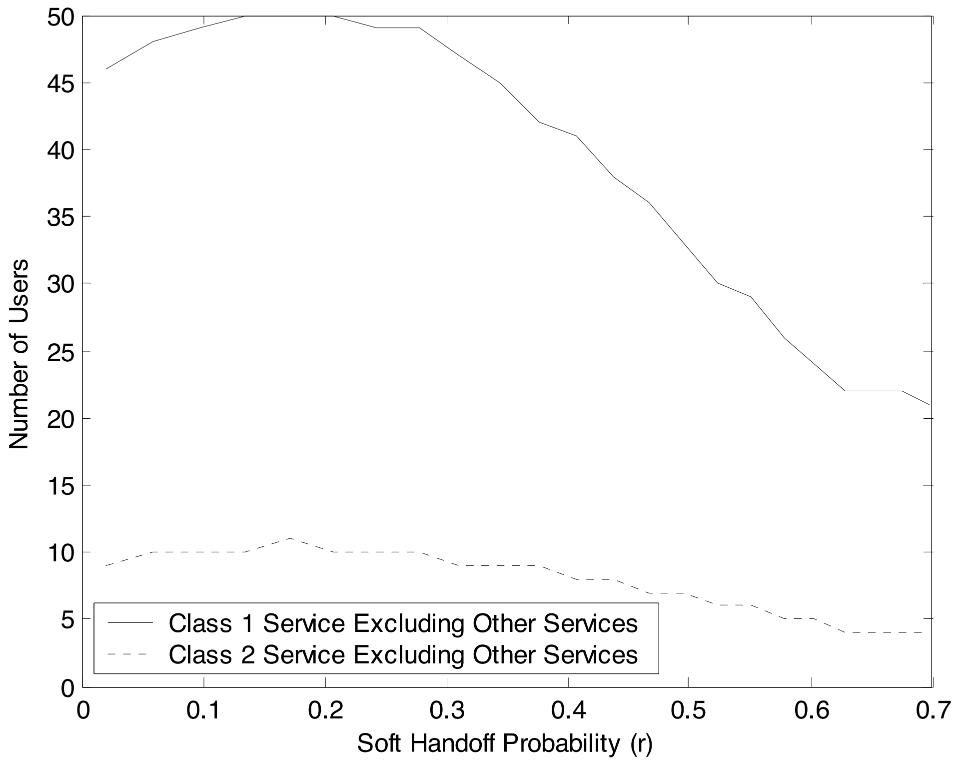

Fig. 5.12. Admission Region in the Forward Link for Classes 1 and 2

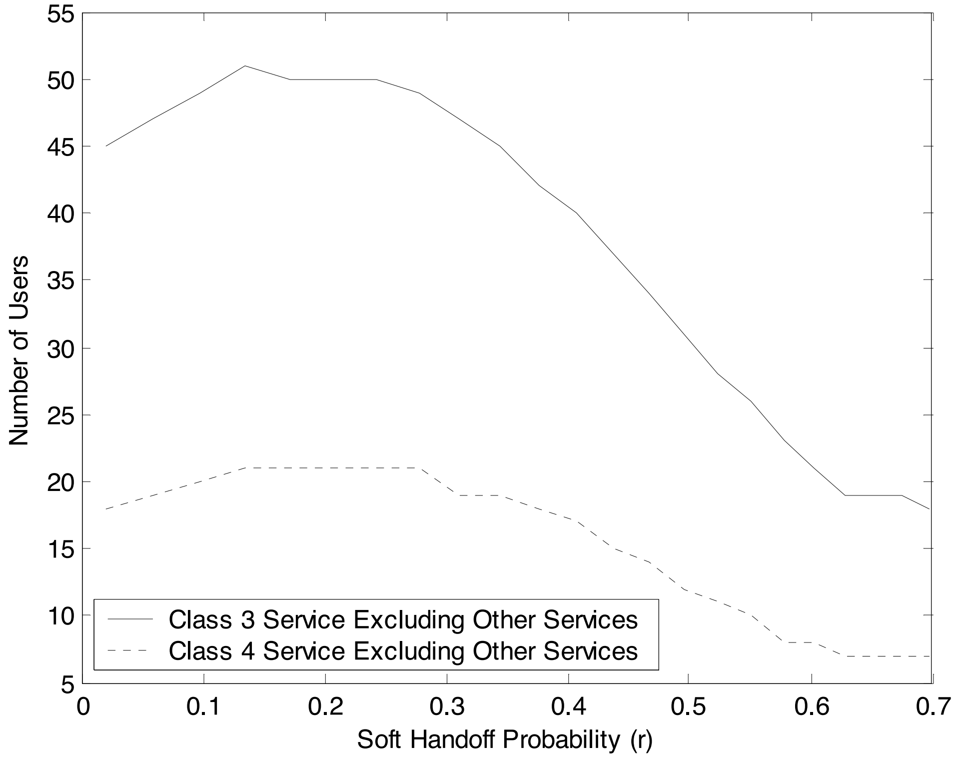

Fig. 5.13. Admission Region in the Forward Link for Classes 3 and 4

In the figures, the soft handoff makes gain for admission regions in the forward link with small SHP values within around $r = 0 \sim 0.15$. However, the SHPs set in practical systems are normally beyond these values, which makes the gain to appear insignificant. When SHP increases to a large value, the forward link admission regions begin to decrease. Although the signal qualities of users in soft handoff is improved through combining the received signals from BSs connected to the mobile users, this diversity gain cannot make up the loss from multiple BSs assigning resources to the same user especially when the number of users in soft handoff becomes extremely large.

As the forward link capacity is normally not the bottleneck of total capacity in a voice-oriented cellular system, the effect of soft handoff on the forward link capacity does not have a significant impact on the system performance in 1G and 2G systems. However, in the future cellular systems, the increasing demand in video and data services with heavier forward link traffic streams may cause the bottleneck capacity to be at the forward link. Thus, SHP should be carefully pre-determined to balance the benefits of soft handoff in the reverse link and the capacity loss induced in the forward link.

5.3.4 Admission Region for the Link Layer

The admission regions for the link layer considering both the reverse and forward links with $\delta_k = 1.0\times10^{-3}$ for classes 1 and 2, and classes 3 and 4 are shown in Figs. 5.14 and 5.15, respectively.

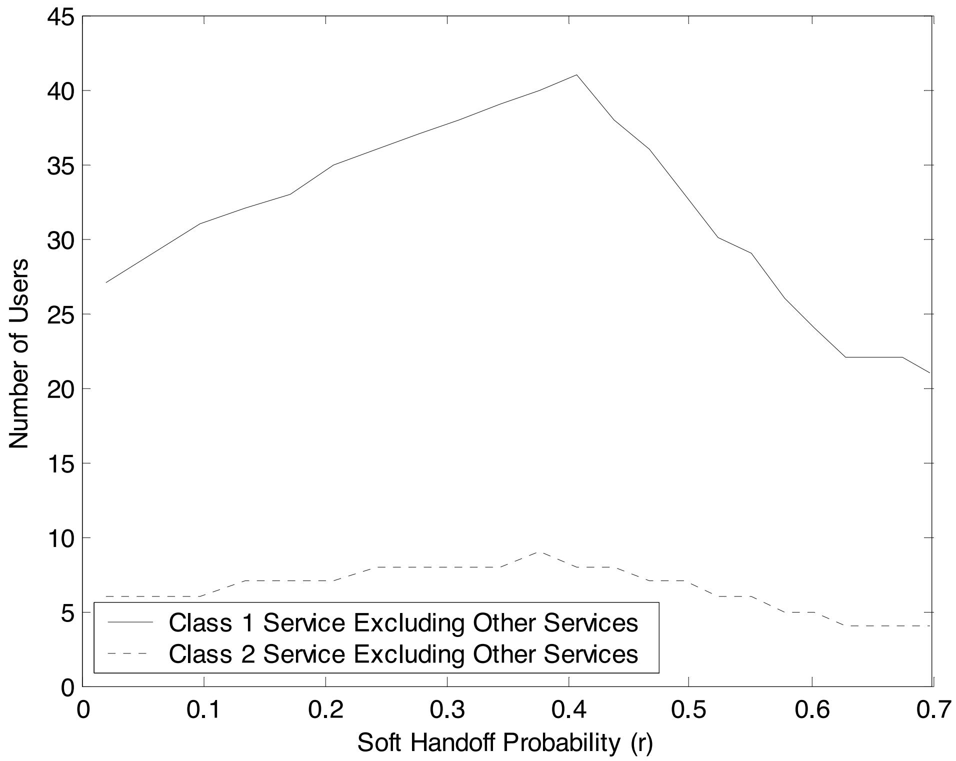

Fig. 5.14. Admission Region for the Link Layer for Classes 1 and 2

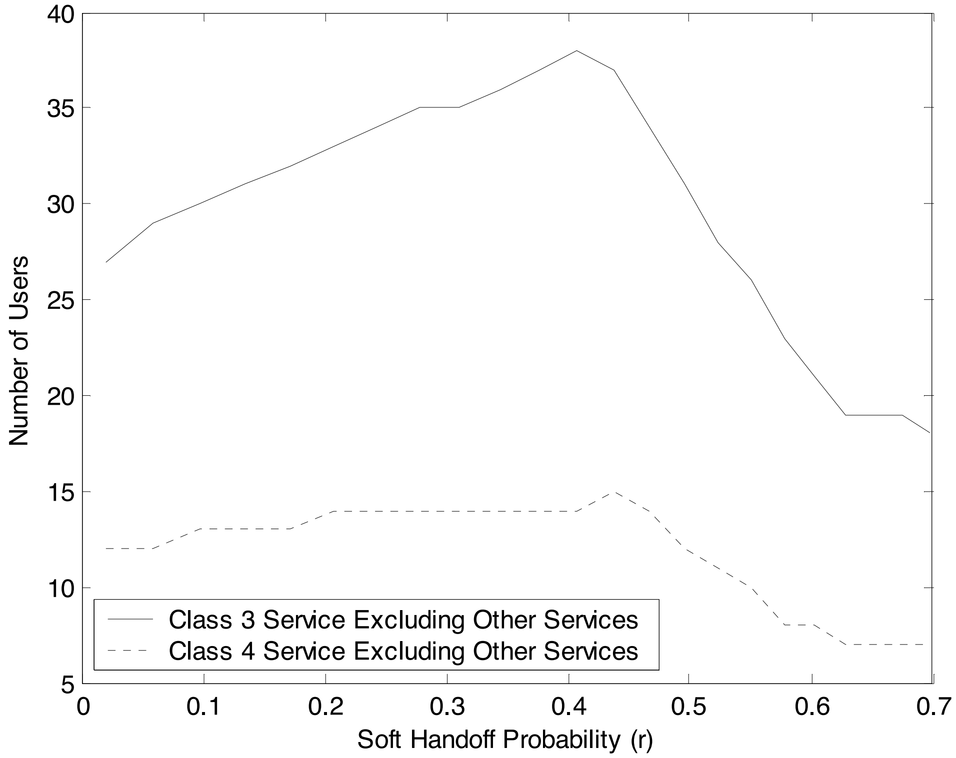

Fig. 5.15. Admission Region for the Link Layer for Classes 3 and 4

When SHP increases from a low value, the admission regions increase. However, with further increase of SHP beyond a certain point, the admission regions decrease. The largest admission region is obtained at around $r = 0.4 \sim 0.5$. The observation can be explained as follows. At the lower end of SHP, the reverse link is the bottleneck for the joint capacity and the gain of soft handoff in the reverse link dominates the admission region. Therefore, the admission regions increase. However, excessive soft handoff decreases the forward link capacity due to the loss from multiple BSs assigning resources to the same user with further increase of SHP. The forward link becomes the bottleneck of the joint capacity and the loss dominates the admission region. Thus, the admission regions decrease.

Note that the traffic volumes in the reverse and forward links in the numerical results are assumed to be the same. In Chapter 6, the capacity unbalance problem will be solved with a proposed adaptive SHP scheme. With this scheme, SHP will be adaptively adjusted along with the variation of the traffic volumes in the reverse or forward link.

There are two knees of the curves for SHP in the reverse and forward link admission regions, $r = 0.15$ and $r = 0.55$. Within $r = 0 \sim 0.15$, the admission regions in the reverse and forward links keep growing. Beyond $r = 0.55$, the admission region in the reverse link changes very little and the admission region in the forward link keeps shrinking. Thus, within $r = 0 \sim 0.15$ and beyond $r = 0.55$, to adjust SHP becomes meaningless as the two admission regions keep changing in the same direction. The adaptive SHP scheme will then operate within the restricted region, $r = 0.15 \sim 0.55$.

5.4 Summary

The SIRs and outage probabilities for multi-class services in BSs (the reverse link) and in mobile users (the forward link) have been formulated with consideration of the salient features in cellular CDMA systems. The analytical models are based on the largest received power BS selection criterion. Interference models are carefully built with soft handoff, diversity and statistical multiplexing in the reverse and forward links. In the forward link, the approximation selection method combining the advantages of previous approximation methods is proposed. Furthermore, the power control schemes for mobile users in soft handoff are investigated and compared. The balancing power control scheme in 3GPP is chosen as the power control scheme for mobile users in soft handoff. The objective for formulating the SIRs and outage probabilities is to obtain the admission region, i.e., the physical capacity defined in Chapter 4. This will make connection between the network layer analysis in Chapter 4 and the link layer analysis in Chapter 5.

Up to now, the segregated layers models for the network layer and the link layer are built. The physical layer features are included into the link layer analysis. The linkage between the network layer analysis and the link layer analysis is the physical capacity, *C*, i.e., the admission region, $\mathbf{A}$, which is used in the network layer analysis and defined in the link layer analysis. The difference in these two concepts will be harmonized in Chapter 6. In Chapter 6, the cross-layer model is presented and the two layers analyses are integrated. With the cross-layer model, the capacity unbalance problem, i.e., the inefficient use of radio resource induced by asymmetric traffic between the reverse and forward links, is solved by controlling various parameters in different layers, including SHP in the physical layer and the priorities of services in the network layer, adapting to the traffic volumes.

Chapter 6

Analysis of Cross-Layer Optimization

The conventional protocol structure divides the wireless networks into various protocol layers, such as the physical layer, the link layer, the network layer and the application layer, etc. This protocol structure allows system designers to build the system easier as the peer-to-peer layers in every communication nodes are transparent to each other. However, this structure also brings inflexibility because the design for each isolated layer lacks the information from other layers, which may lead to the wastage of resource allocation. As a result, although the performance in a particular layer is optimized, the whole system may not be optimal as only the worst-case performance is studied in respective segregated layer. In the wireless network design, the drawbacks of the segregated layer protocol structure turn to be more obvious as the hostile wireless propagation medium will degrade the performance in segregated layers and worsen the worst-case performance used in the design. The cross-layer optimization then becomes more necessary and meaningful in the wireless network design than that in the wireline network design.

The cross-layer approach is looking at the integrated studies of exploiting the statistical behavior among the intertwining parameters in various layers to obtain optimal system performance. The intertwining parameters are generally the performance metrics defined in all layers. In this thesis, the QoS metrics formulated include the new-call-blocking probability, the handoff-call-dropping probability, the system utilization and the packet-loss probability in the network layer, SIR and the outage probability in the link layer, and the soft handoff probability (SHP) and the hysteresis margin in the physical layer. To fulfill the cross-layer approach without disturbing the integrality of the conventional protocol structure which then can be reserved for further development and update, a function block, the cross-layer decision-maker (DM), is proposed herein. The cross-layer optimization with the intertwining parameters is applied in the cross-layer DM.

The intertwining parameters are transferred through an interface from the segregated layers into the cross-layer DM. The system configuration parameters which are the outputs of the optimization problem for each layer are also transferred back through this interface. Besides the intertwining parameters and system configuration parameters, there are the connection parameters to exchange information between layers.

Furthermore, to balance the resources required by multi-class traffic with different QoS requirements, the revenues from multi-class traffic are exploited in the system model. The revenue includes the economic concerns from many aspects, such as call arrival rates, traffic volumes, blocking rate requirements, loss rate requirements, etc. The revenues will be used to determine the nominal capacity of different groups, C_i, $i = 1, 2$, defined in section 3.3.1, i.e., the higher revenue of users in the group, the bigger nominal capacity allocated for that group, and vice versa. Thus, the revenues will affect priorities of the multi-class traffic, e.g., higher revenue users are elevated to higher priorities.

A practical problem, the capacity unbalance problem, is solved with an adaptive SHP scheme in this chapter as an example of applying the proposed cross-layer optimization.

6.1 Introduction to the Cross-Layer Decision-Maker

The cross-layer DM, the conventional layered system and the interface between them are shown in Fig. 6.1.

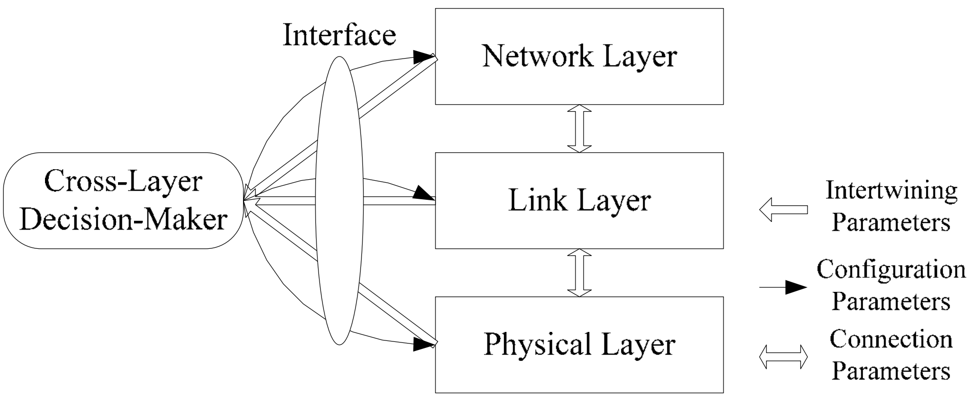

Fig. 6.1. The Cross-Layer Decision-Maker

The segregated layer analyses for the network layer and the link layer have been presented in Chapters 4 and 5, respectively. The physical layer features are included into the link layer analysis. The analyses in previous chapters actually have exhibited some simple models of the cross-layer design, e.g., the joint packet and connection levels optimization in the network layer analysis, and the investigation of the influence from the physical layer consideration, the soft handoff, on the SIRs in both the reverse and forward links in the link layer analysis, etc.

A general cross-layer optimization model will be proposed in section 6.3. The general model will be applied in the cross-layer DM. Three types of system parameter are defined in the DM. They are the intertwining parameters, the connection parameters and the system configuration parameters.

The intertwining parameters, the connection parameters and the system configuration parameters used in the thesis are listed below.

The intertwining parameters from each layer are listed as follows:

B_n	total new-call-blocking probability in the network layer;
B_h	total handoff-call-dropping probability in the network layer;
N_u	total system utilization in the network layer;
L_k	packet-loss probability for class *k* call in the network layer;
$(E_b/I_0)_{k,i}$	SIR for user *i* in class *k* in the link layer;
$P_{\text{outage},k}(\vec{n})$	outage probability for class *k* traffic in the link layer.

The connection parameters between layers are listed as follows:

C	physical capacity connecting the network and link layers;
r_k	number of basic channels required for each class *k* call connecting the network and link layers;
r	soft handoff probability connecting the link and physical layers.

The system configuration parameters fed back to each layer are listed as follows:

N upper bound for admitting calls in the network layer;

Δ_{dB} hysteresis margin in the physical layer.

6.2 Parameters in the Cross-Layer Decision-Maker

The parameters in the cross-layer DM have been listed in section 6.1. Many of these parameters have been defined and formulated in previous study, such as blocking probability, system utilization, SIR, outage probability, etc. However, there are still some parameters, such as the connection parameters, SHP and the channel number, and the intertwining parameter, the packet-loss probability, which are formulated in the segregated layer and should be developed to adapt their definitions in the cross-layer model. Besides, with the introduction of the revenues for multi-class users, another intertwining parameter, the penalty of call blocking that represents the revenue loss from different users because of call blocking, is brought in and should be derived as a QoS metric representing the efficiency of resource utilization.

6.2.1 Soft Handoff Probability

In the link layer analysis, the SHP, i.e., the ratio of the number of soft handoff users permitted in the network to the total number of users, is defined. It is also shown that the SHP will affect the intra- and inter-cell interference in both the reverse and forward links by determining the number of mobile users in soft handoff and non-soft handoff.

Thus, SHP can determine SIR and the outage probability in the reverse and forward links.

The proposed adaptive SHP scheme has been described in Chapter 2. When the traffic asymmetry is biased toward the forward link, SHP is decreased to increase the forward link capacity, and when the traffic is symmetrical, SHP is increased to enlarge the reverse link capacity. Although this may result in a slight degradation in the handoff reliability, it is an efficient way to improve the overall system utilization which is shown in the latter part of this chapter. With the adaptive SHP scheme, SHP adapts to the change in the reverse and forward links transmission traffic volumes. Under the optimal SHP, the reverse and forward links capacities are balanced and the system achieves its performance optimization.

6.2.2 Number of Basic Channels

This section explains how results in Chapter 5 can be extended to be used in the cross-layer design.

The number of basic channels of multi-class traffic has been used in the network layer analysis. The number of channels actually represents the system resource consumed by the individual traffic, i.e., the larger the resource consumed by a single user of a given traffic, the larger the number of channels assigned to the traffic. The system resources allocated to multi-class traffic, such as bandwidth, power, rate, etc., are actually the link layer issues. Thus, the number of channels used in the network layer analysis should be based on the link layer analysis. This can be done in the following way.

With a given traffic, the resource consumed by a single user of the traffic is inversely proportional to the maximum number of users supported by the system, i.e., the larger the maximum number of users, the lesser the resource required by a single user. Thus, the number of channels can be approximated with the aid from the admission region defined in section 5.3 where the maximum number of users for the traffic can be obtained without including users from any other traffic. Assuming the channel for class j traffic users as the basic channel (units), the number of basic channel for class k traffic users is given by

$$r_k = \frac{n_j^{\max}}{n_k^{\max}} r_j , \tag{6.1}$$

where $n_k^{\max}$ denotes the maximum number of users for class k traffic supported in admission region $\mathbf{A}$ and $n_j^{\max}$ denotes the maximum number of users for class j traffic supported in admission region $\mathbf{A}$.

Combining the physical capacity, C, and the number of channels, r_k, the admission region in the link layer can be remodeled by the network layer parameters as

$$\mathbf{A}' = \left\{ \vec{n} \left| \vec{n} \cdot \vec{r} \le C \right. \right\}, \tag{6.2}$$

where $\vec{r} = (r_1, r_2, \ldots, r_K)$.

Because the admission region obtained by the numerical programming actually has a linear surface, e.g., Fig. 5.11 shows the linear surfaces of the admission regions in the forward link, $\mathbf{A}'$ can provide an accurate estimation on the admission region. Thus, by using the connection parameters, C and r_k, the network layer analysis can make use of the admission region defined in the link layer analysis.

6.2.3 Packet-Loss Probability

This section explains how results in Chapter 4 can be extended to be used in the cross-layer design.

The packet-loss probability has been discussed in the network layer analysis. In the network layer analysis, although the statistical multiplexing is considered at the packet-level and formulated in the packet-loss probability, the physical capacity, C, ignores statistical multiplexing in the lower layers. However, in CDMA, due to the burst of traffic, statistical multiplexing happens inherently in all layers. Thus, the packet-loss probability should be reformulated and matched to the outage probability in the cross-layer design.

The outage probability in the link layer represents the ratio of the lost packets to the total packets for the scenario without buffering and retransmission. By combining the outage probability with a weighting factor, the steady state probability, the equivalent packet-loss probability can capture the activities of the mobile users [108]. In other words, the packet-loss probability in the network layer is equivalent to the outage probability in the link layer as a result of the activities of the mobile users in the network layer for the scenario without buffering and retransmission.

Thus, the equivalent class k packet-loss probability normalized over all classes defined below can be use to replace (4.17) in the network layer analysis. The class k equivalent packet loss probability is given by

$$L_k = \frac{\sum_{n_1=0}^{\lfloor N/r_1 \rfloor} \cdots \sum_{n_K=0}^{\left\lfloor \left(N - \sum_{l=1}^{K-1} n_l r_l \right) \big/ r_K \right\rfloor} P(\vec{n}) \times n_k \times P_{\mathrm{outage},k,i}(\vec{n}) \times r_k}{L_{sum}}, \tag{6.3}$$

where

$$L_{sum} = \sum_{n_1=0}^{\lfloor N/r_1 \rfloor} \cdots \sum_{n_K=0}^{\left\lfloor \left(N - \sum_{l=1}^{K-1} n_l r_l \right) \big/ r_K \right\rfloor} P(\vec{n}) \times \sum_{k=1}^{K} n_k r_k , \tag{6.4}$$

and $P_{\mathrm{outage},k,i}(\vec{n}) = \max\left(P^{R}_{\mathrm{outage},k,i}(\vec{n}), P^{F}_{\mathrm{outage},k,i}(\vec{n})\right)$ is the outage probability for user i of class k traffic where $P^{R}_{\mathrm{outage},k,i}(\vec{n})$ is the outage probability for user i of class k traffic in the reverse link defined in (5.28) and $P^{F}_{\mathrm{outage},k,i}(\vec{n})$ is the outage probability for user i of class k traffic in the forward link defined in (5.43).

6.2.4 Penalty of Call Blocking

The proposed adaptive SHP scheme improves the system utilization through balancing the capacities in the reverse and forward links. However, the improvements of different services are unbalanced as symmetrical traffic services prefer a fixed SHP while asymmetrical traffic services prefer a dynamic SHP. Applying adaptive SHP scheme will then be at the price of the performance degradation for symmetrical traffic services. To balance system performance for symmetrical and asymmetrical traffic services in the adaptive SHP scheme, the revenues for different services, besides their call blocking probabilities, are taken into consideration. The penalty of call blocking is defined with combining the new-call-blocking probability, the handoff-call-dropping probability and the revenues of different services together. As the optimization objective is to maximize either the total utility summed over all users, or total revenue generated by the users

[109], the optimal SHP is selected by looking for the lowest penalty of call blocking. Thus, at the optimal SHP, the maximum revenue is achieved. The penalty of call blocking is defined as

$$\text{Penalty}_{\text{call blocking}} = \sum_{k=1}^{K} \left(B_{nk} + \kappa \times B_{hk}\right) \times \text{Revenue}_k \ , \tag{6.5}$$

where $\kappa \geq 1$ is a slack factor which assigns a higher priority to the handoff call connections than the new call connections.

With CS, because each service shares the capacity without priority assignment, the revenue of different service is assumed to be the same with $\text{revenue}_k = 1$, for $k = 1, \ldots, K$. With VP, the revenues of all services in group 1 are assumed as *x* and the revenues of all services in group 2 are *y*. Thus, VP scheme can be specified as VP (*x*-*y*) with $\text{revenue}_k = x$, for $k = 1, \ldots, p$, and $\text{revenue}_k = y$, for $k = p+1, \ldots, K$. If $x > y$, group 1 services have higher revenue and priority than group 2 services. If $x < y$, group 2 services have higher revenue and priority than group 1 services.

6.3 Cross-Layer Optimization in Decision-Maker

The formulae of the intertwining parameters have been defined. Maximizing system utilization can translate to more revenue for the network providers who in turn can lower the charges for the mobile users. Thus, the goal of system optimization is to maximize system utilization subject to the QoS constraints.

The joint connection and packet levels optimization has been presented in Chapter 4. Combining the link layer analysis, the optimization problem will be extended for cross layers. The optimization problem processed in the cross-layer DM is then proposed as

$$\max N_u , \tag{6.6}$$

subject to

$$B_n \le B_n^* ,$$

$$B_h \le B_h^* ,$$

$$L_k \le L_k^* ,$$

where the superscript * denotes the QoS requirements of the corresponding parameters. Equation (6.6) corresponds to (3.14) in Chapter 3 where function $\mathscr{H}(.)$ is (6.3).

System optimization is achieved by choosing the optimal nominal capacity N which is a design parameter representing the upper bound for admitting calls. The selection of optimal N uses the educated search with an iterative procedure described in section 3.3.1. Every value of N in a range of values is tested and the optimum N that gives the maximum system utilization and satisfies the QoS constraints at the same time is obtained. Note that the equivalent packet-loss probability contains information from the network layer, the link layer and the physical layer. Equation (6.6) thus optimizes the system utilization by cross considering the issues in the network layer, the link layer and the physical layer.

After an off-line study using the analytical model, a high-dimensional table mapping the system environment (e.g., interference, QoS requirements) to the system configuration will be stored in the decision-maker. During the execution phase of the system, if the environment is changed, the system will inform the decision-maker about the changes through the intertwining parameters. These changes could come from the change of the distribution of users, traffic burst, propagation environment, etc. The decision-maker searches the mapping table and feeds back the optimal configuration parameters to the original system. The connection parameters between layers represent the information exchanges during the analysis.

In this thesis, due to the assumptions of the statistically identical cells and the low data burst from the services, the environment could be assumed as stable. Thus, there is no need for iteration and the initial optimal status will last for the whole duration of the system operation time.

Equation (6.6) represents the general formulae for the cross-layer optimization where the system utilization is maximized subject to the requirements of QoS metrics in all layers. The system configuration parameter fed back to each layer is the nominal capacity, N, in the network layer. To solve other practical problems with the cross-layer model, the optimization includes other specified system configuration parameters for a certain problem and the general cross-layer optimization method would be slightly revised to adapt to the problems.

In this thesis, a practical problem, the capacity unbalance problem, is solved with the adaptive SHP scheme. Thus, the hysteresis margin in the physical layer which determines the SHP value is another feedback system configuration parameter. The revenues for different services and the penalty of call blocking are considered into the system model. Note that to minimize the penalty of call blocking probability has the same meaning to maximize the system utilization as more high revenue calls can be admitted into the system instantaneously. Thus, the optimization problem for applying the adaptive SHP scheme is

$$\min \text{Penalty}_{\text{call blocking}}, \tag{6.7}$$

subject to

$$B_n \leq B_n^*,$$

$$B_h \leq B_h^*,$$

$$L_k \leq L_k^*,$$

$$r = 0.15 \sim 0.55.$$

where the range of SHP is specified according to the statements in section 5.3.4. Every value of N and SHP in their ranges of values is tested and the optimum N and SHP that give the minimum penalty of call blocking, i.e., the maximum revenue, and satisfy the QoS constraints at the same time are obtained.

To differentiate the adaptive SHP scheme from the general cross-layer optimization, the general case in (6.6) is specified as the joint optimization in the later part of this chapter.

The analytical model will select the optimal SHP value giving the ratio of the traffics volume of the forward link to the reverse link and the priorities for different traffics. After an off-line study for different traffic volumes combinations and traffic priorities assignments, a high-dimensional result can be printed out or stored in memory. The input parameters will be different traffic volumes combinations, traffic priorities assignments, interference model settings and QoS requirements, etc., and the output parameters are the optimal nominal admission bound and the optimal SHP. Tracing the surface of the output parameter values, the adaptive SHP scheme can be implemented in practice.

The traffic ratio of the forward link to the reverse link can be measured in the BSs. Jumping window or moving window can be applied for the traffic volume measurement in the forward and reverse links. The details of the measurement methods should be included into system model in the future. The search algorithm to accelerate the searching of the high-dimensional table during operation should also be designed in the future to bring the cross-layer model into practice. In the thesis, the exhaustive search method of the high-dimensional table is applied.

For the purpose of comparing the performance of the segregated-layer design and the cross-layer design, a method assessing the performance of the segregated-layer design is presented here. Corresponding to (3.13), the outage probability is constrained in the segregated link layer without being included in the packet-loss probability. The nominal capacity, *N*, is then equivalent to the admission region given in (5.48). Thus, the approach to obtain the system utilization through the admission control in the segregated layers is given by

$$\max N_u , \tag{6.8}$$

subject to

$$B_n \le B_n^* ,$$

$$B_h \le B_h^* ,$$

$$P_{\text{outage},k,i}(\vec{n}) \le P^*_{\text{outage},k,i}(\vec{n}) .$$

Equation (6.8) is used to obtain the results by the conventional segregated-layer design.

6.4 Cross-Layer Optimization - Performance Evaluation

Following the services definition in performance evaluations in the network and link layers analyses, 4 classes that represent voice, video, web-browsing and data services, respectively, are considered. The voice, web-browsing and data services are modeled as *ON-OFF* sources which are Sen's model without the active HBR mini-source states and with only one two-state LBR mini-source each, while the video source is modeled as Sen's model with the maximum number of active LBR spreading codes, $M_2 = 8$. The users from group 1 services (voice and video) are assumed symmetric, while the users from group 2 services (web-browsing and data) are asymmetric. Thus, while increasing the amount of transmission traffic in the forward link, the traffic for group 2 is increased and the traffic for group 1 is kept unchanged.

The CAC schemes in the network layer include CS and VP with preemption for all classes (VP Case 1). However, the best effort and guarantee access with preemption for best effort traffic (VP Case 2) is excluded as the technique to implement VP Case 2 in the cross-layer optimization is almost the same as that for VP Case 1.

The parameters used in this chapter inherit those from previous study as in Table 4.1 and Table 5.1. The parameters values are shown in Table 6.1.

Table 6.1. The Parameters Values in the Cross-Layer Optimization

Parameter Type	Voice	Video	Web-browsing	Data
Activity Factor (α_k)	0.4	0.3867 (LBR) 0.5 (HBR)	0.1176	0.1028
Transmission Rate (R_k)	60 kbps	30 kbps (LBR) 60 kbps (HBR)	120 kbps	240 kbps
SIR Requirement (γ_k^*)	2 dB	2 dB (LBR) 2 dB (HBR)	3 dB	3 dB
Holding Time ($1/\mu_{ck}$)	1 min	1 min	2 mins	2 mins
User Speed (v_k)	4.0 km/hr	4.0 km/hr	4.0 km/hr	4.0 km/hr
Total Bandwidth (W)		5 MHz		
Guard Capacity (C_G)		8		
Shadowing Deviation (σ_ξ)		6 dB		
Path Loss Factor (τ)		4		
Outage Requirement (δ_k)		1.0×10^{-3}		
Orthogonality Factor (*f*)		0.2		
BS Number (B)		24*		
S_1/η		-1 dB		

* the 25 BSs (including BS 0) form a 5x5 square cells structure.

6.4.1 Selection of the Optimal SHP for CS and VP

The penalties of call blocking under various SHP for CS, VP (5-1) and VP (1-5) defined in section 6.2.4 are shown in Figs. 6.2, 6.3 and 6.4, respectively. The new call arrival rates are all the same for the 4 traffic classes with the total call arrival rate of 3.6/minute. The SHP increases from 0 with the step size 0.01. When SHP increases from a low value, the penalty of call blocking decreases. However, with further increase of SHP beyond a certain point, the penalty of call blocking increases. The above observation can be explained as follows. At the lower end of the SHP, the reverse link is the bottleneck for the joint capacity. The call blocking decreases due to the gain in soft handoff in the reverse link. However, excessive soft handoff decreases the forward link capacity. After a certain point (the optimal SHP), the forward link becomes the bottleneck of the joint capacity. The call blocking increases due to the loss from multiple BSs assigning resources to the same mobile user in the forward link.

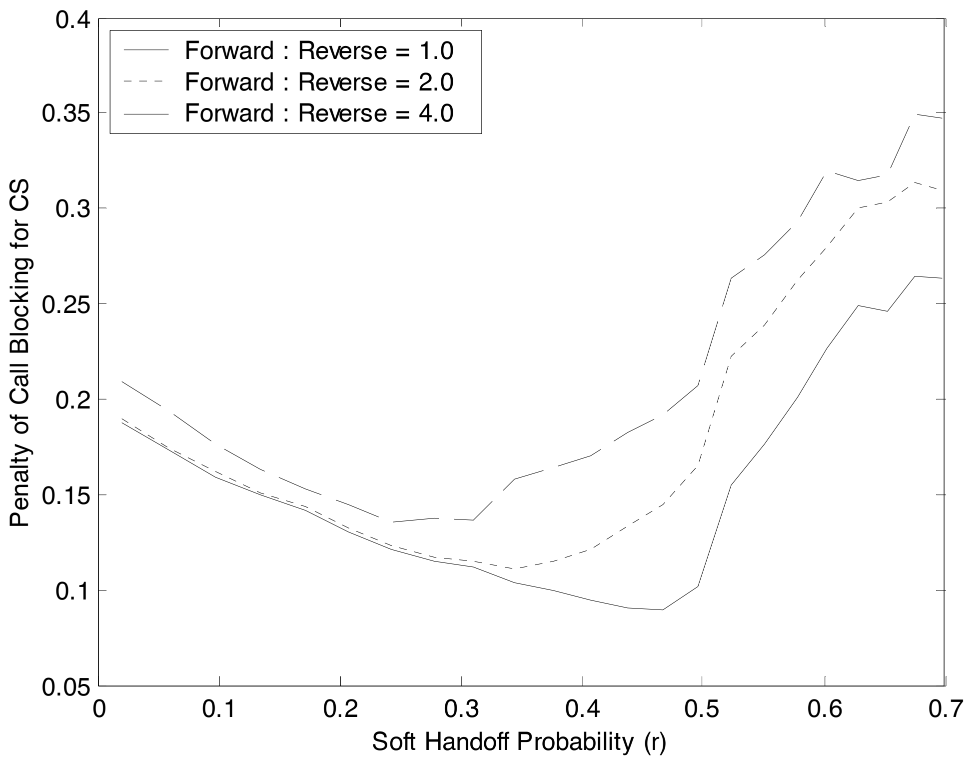

Fig. 6.2. Penalty of Call Blocking vs. SHP for CS

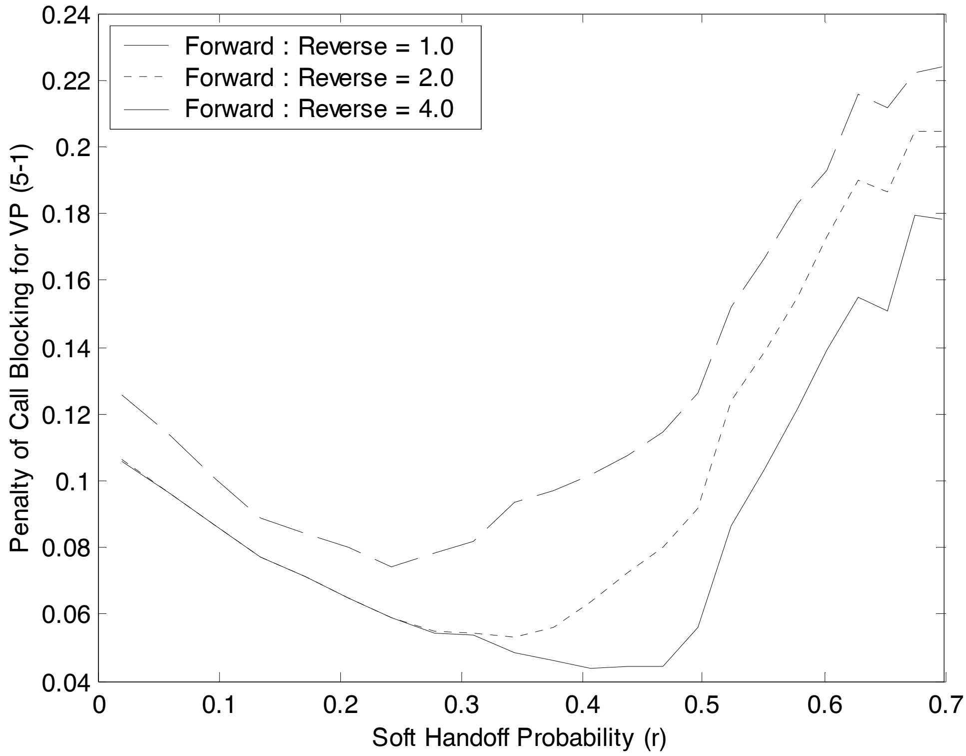

Fig. 6.3. Penalty of Call Blocking vs. SHP for VP (5-1)

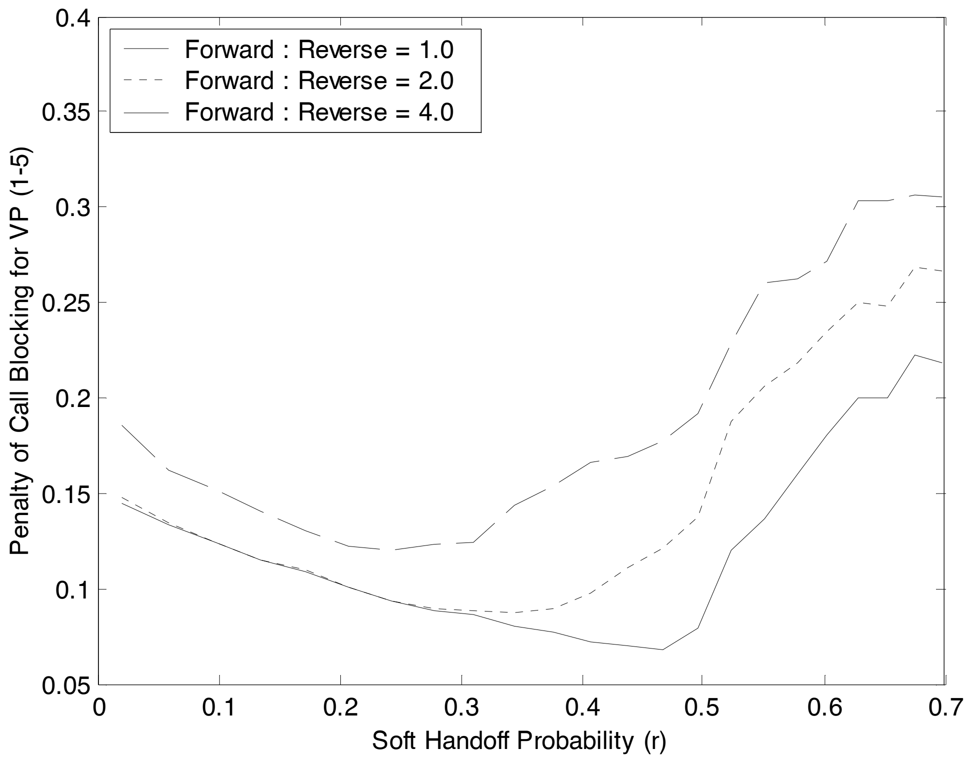

Fig. 6.4. Penalty of Call Blocking vs. SHP for VP (1-5)

In Figs. 6.2, 6.3 and 6.4, the ratios of the forward link to the reverse link represent the traffic ratios for group 2. This is done by scaling the connection activity factors, α_3 and α_4, in the forward link. The increase of the transmission traffic from group 2 shrinks the admission region and intensifies the call blocking. The optimal SHP shifts to a lower value. Moreover, the higher the increase of transmission traffic from group 2, the lower the value of the optimal SHP. The optimal SHP values under various forward link transmission traffic volumes for group 2 with CS, VP (5-1) and VP (1-5) are presented in Table 6.2. This mapping table is a specific example of the cross-layer optimization model and stored in the decision-maker.

Table 6.2. The Optimal SHP Values with CS, VP (5-1) and VP (1-5)

SHP	$F / R = 1.0$	$F / R = 2.0$	$F / R = 4.0$
CS	0.45	0.33	0.25
VP (5-1)	0.45	0.33	0.28
VP (1-5)	0.45	0.31	0.19

In Table 6.2, F represents the forward link transmission traffic volumes for group 2 and R represents the reverse link transmission traffic volumes for group 2.

The optimal SHP values for VP (5-1) are larger than that for CS, while the optimal SHP values for VP (1-5) are smaller than that for CS, especially for a large traffic ratio of the forward link to the reverse link. The observation can be explained as follows. For VP (5-1) scheme, group 1 services have higher revenues, which means group 1 services have higher priority than group 2 services and the selection of the optimal SHP would always prefer group 1 services. Because of the symmetric traffic for group 1, the optimal SHP for group 1 tends to be constant. Thus, when there is asymmetric traffic for group 2, the shift of optimal SHP for whole system would be small and the optimal SHP values for VP (5-1) are larger than that for CS. Due to the same reasons, for VP (1-5), group 2 services have higher priority than group 1 services and group 2 services are more likely to have smaller values of the optimal SHP that result in larger forward link capacities for group 2. Thus, the optimal SHP values for VP (1-5) are smaller than that for CS.

6.4.2 System Utilization Gain for CS and VP

With the adaptive SHP scheme, the system is adjusted to operate on the smallest penalty call blocking point (the optimal SHP). At the optimal SHP, the capacities between the reverse and forward links are balanced and the system performance is optimized.

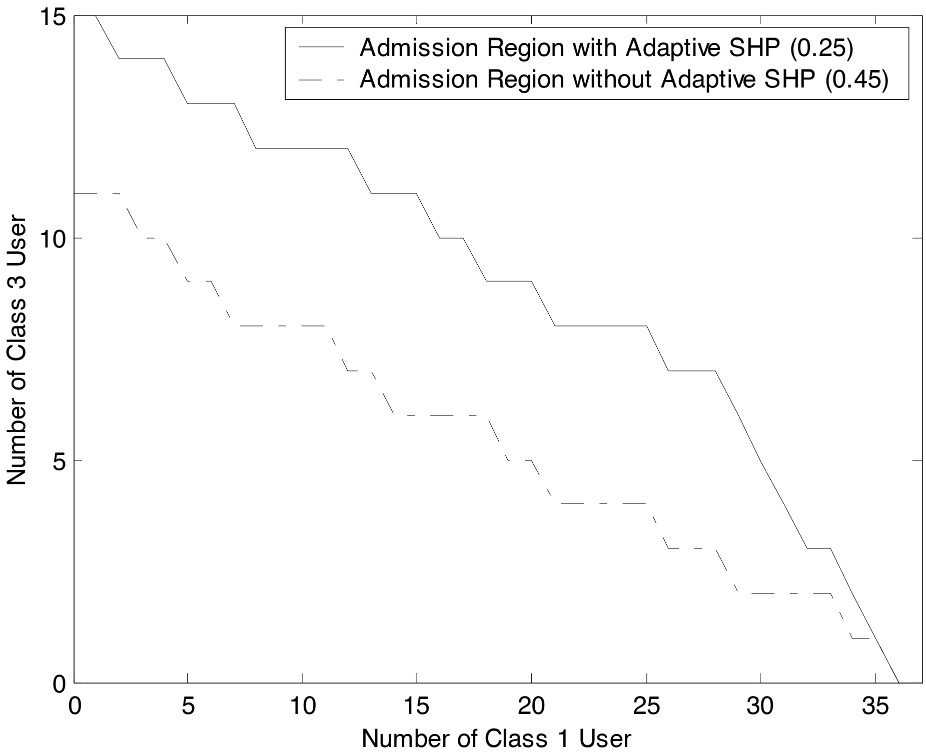

Fig. 6.5. Capacity Gain with Adaptive SHP Scheme

Fig. 6.5 shows the capacity gain with the adaptive SHP scheme compared to that without SHP. The ratio of the transmission traffic for group 2 in the forward link to that in the reverse link is 4.0. In Table 6.2, the optimal SHP for CS is 0.25 with adaptive SHP scheme, while the SHP for CS is 0.45 without the adaptive SHP scheme. In order to see the admission region of the whole system as a two-dimensional graph, the admission region for classes 1 and 3 services is considered without the existence of any class 2 or 4 service. The capacity gain is obvious for the adaptive SHP scheme. However, as the advantages of the adaptive SHP scheme would not be shown if there are no asymmetric users, the capacity gain for class 1 service vanishes quickly when the number of class 3 users decreases to zero.

The system utilizations for VP (5-1) with adaptive SHP in (6.7), VP (1-5) with adaptive SHP in (6.7), CS with adaptive SHP in (6.7), CS with joint optimization but without adaptive SHP in (6.6) and CS with segregated-layer design in (6.8) for classes 1, 2, 3, and 4, and the total system utilization are shown in Figs. 6.6, 6.7, 6.8, 6.9 and 6.10, respectively. The ratio of the transmission traffic for group 2 in the forward link to that in the reverse link is 4.0.

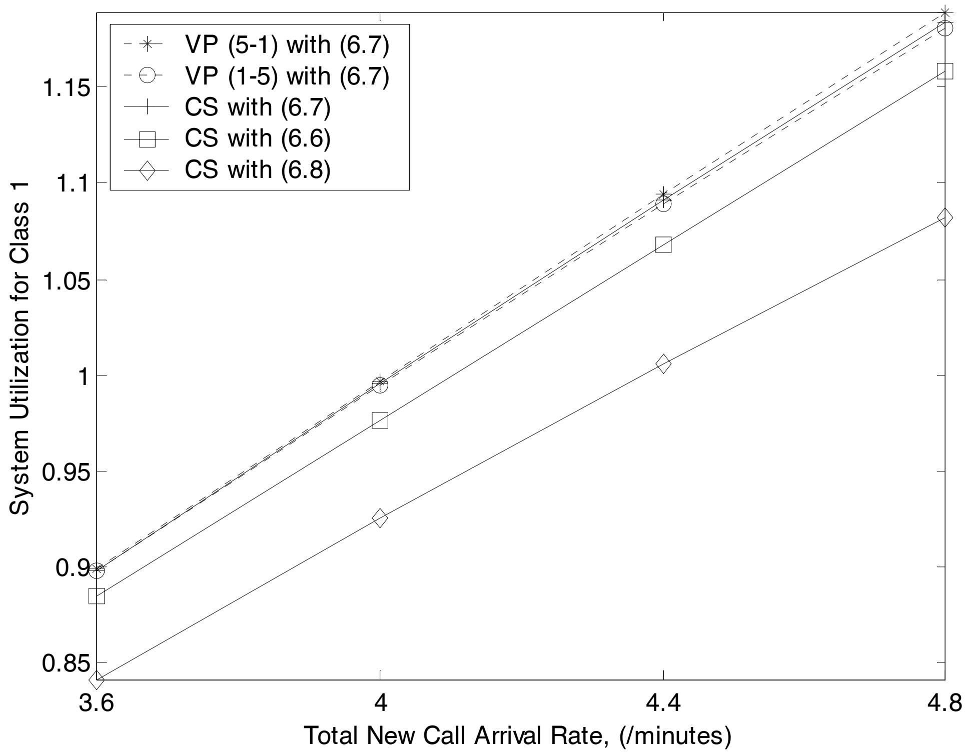

Fig. 6.6. System Utilization for Class 1 of Various Schemes

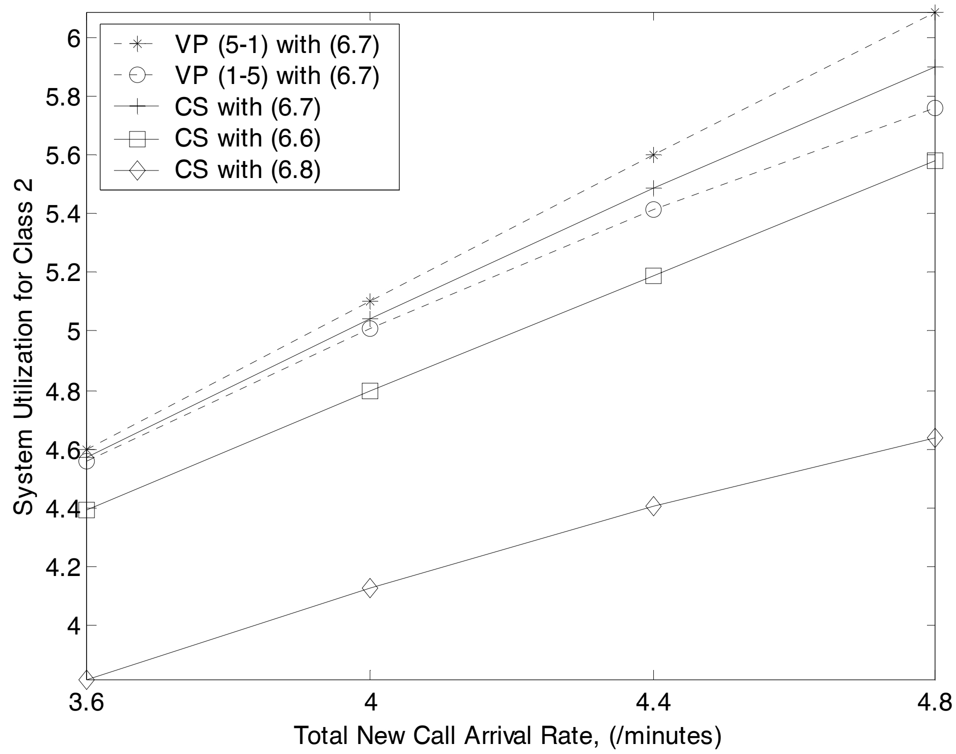

Fig. 6.7. System Utilization for Class 2 of Various Schemes

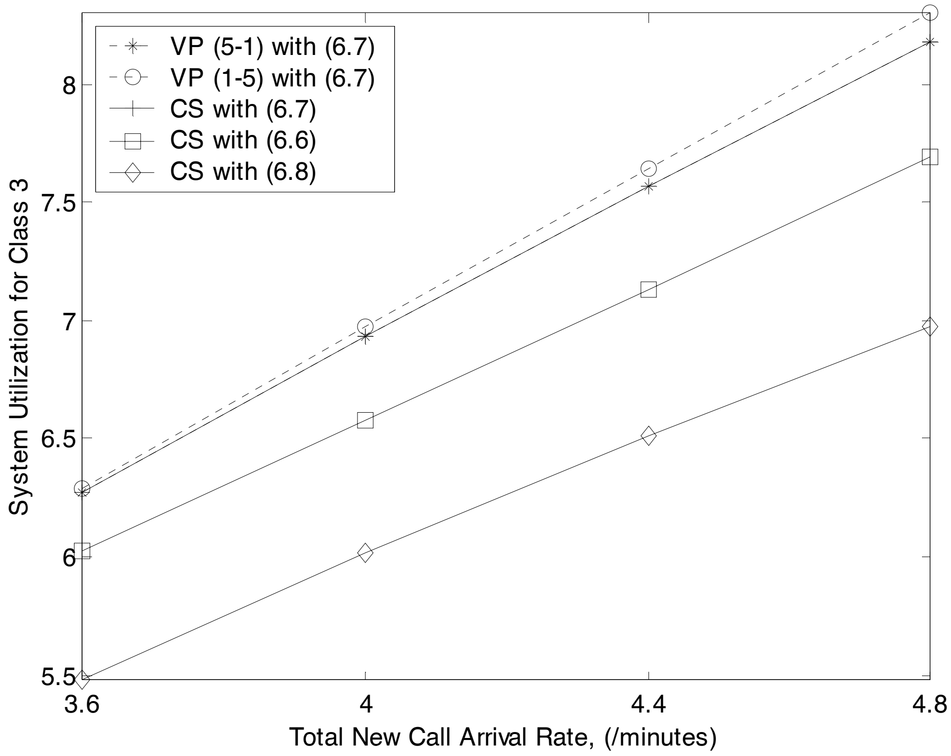

Fig. 6.8. System Utilization for Class 3 of Various Schemes

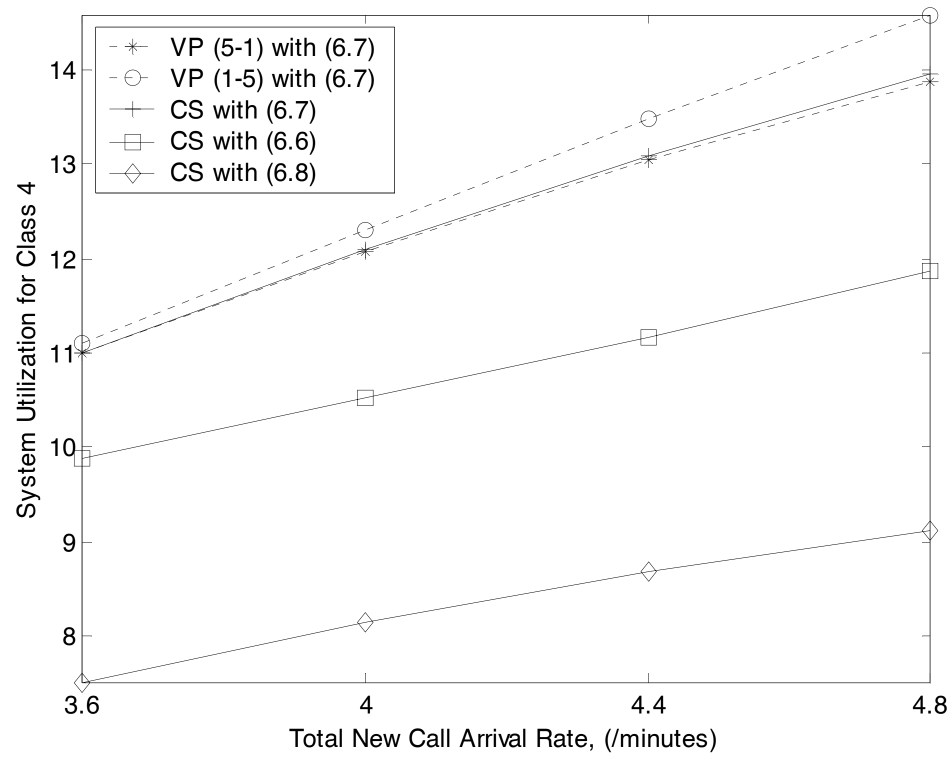

Fig. 6.9. System Utilization for Class 4 of Various Schemes

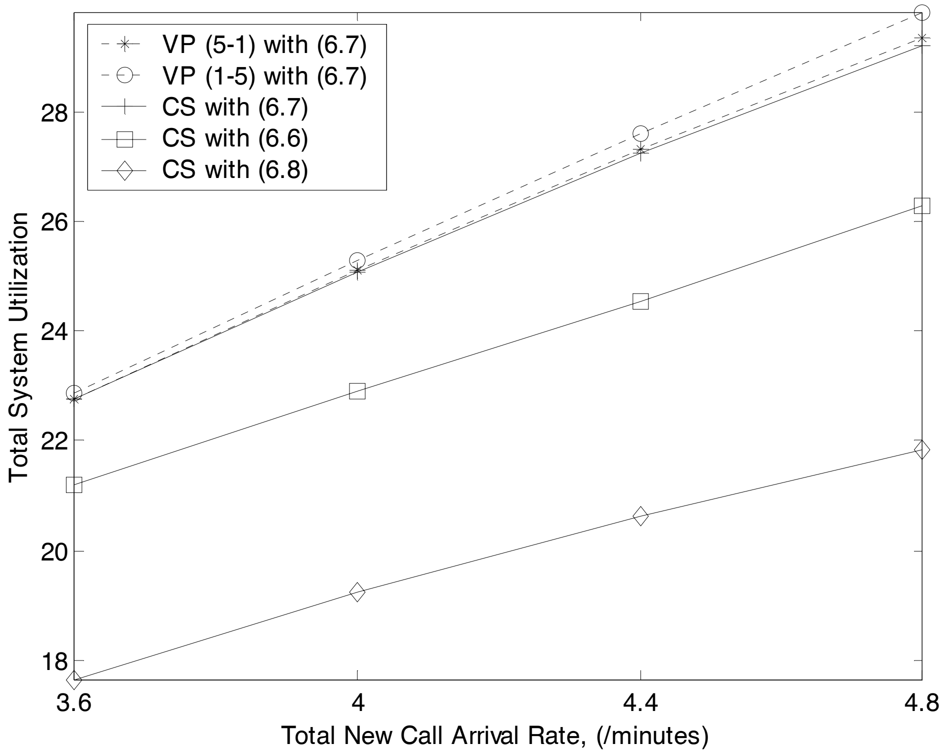

Fig. 6.10. Total System Utilization of Various Schemes

CS with joint optimization but without adaptive SHP in (6.6) is superior to CS with segregated-layer design in (6.8), achieving the maximum system utilization gain of about 30%. CS with adaptive SHP in (6.7) is superior to CS with joint optimization but without adaptive SHP in (6.6), achieving the maximum system utilization gain of about 20%. The maximum system utilization gain for CS with adaptive SHP in (6.7) over CS with segregated-layer design in (6.8) achieves about 50%. The system utilization gains show the superiority of the proposed cross-layer model over the conventional segregated-layer model. The gains in the cross-layer model come from exploiting the statistical behavior among layers and providing the 'soft' guarantee for QoS metrics in all the layers. However, in segregated-layer model, the QoS metrics are provided the 'hard' guarantee. Besides, the proposed adaptive SHP scheme balances the reverse and forward links capacities and offers high system utilization gain compared to the fixed SHP scheme.

Furthermore, from the figures, VP (5-1) which assigns a higher priority for classes 1 and 2 is superior to CS for classes 1 and 2 with the price of system utilization losses for classes 3 and 4 having low revenues. Similarly, VP (1-5) which assigns a higher priority for classes 3 and 4 is superior to CS for classes 3 and 4 with the price of system utilization losses for classes 1 and 2. Thus, with the adaptive SHP scheme, VP protects the high revenue services and brings high system utilization gain to them as compared to CS. Combining the benefits from the adaptive SHP scheme and VP schemes, the maximum achievable gain in system utilization is about 60% as compared to CS with segregated-layer design in (6.8).

VP provides different priorities for multi-class traffic and leads to changes in the system configuration in lower layers, such as the selection of the optimal SHP for VP in section 6.4.1. Thus, the higher layer parameters, such as the revenues of services, affect the resource allocated for different services in the lower layers so that services with higher revenues can occupy more resource in the system. With the cross-layer optimization, not only the characteristics in the lower layers determine the performance in the higher layers, the QoS requirements from different services in the higher layers also affect the resource allocation in the lower layers.

6.5 Summary

The cross-layer model involving the physical layer, the link layer and the network layer has been designed. A cross-layer DM through which the cross-layer optimization is applied without disturbing the integrity of the conventional protocol structure is introduced. The parameters intertwined among layers are considered together to achieve cross-layer optimization in the DM. Some parameters, such as SHP, channel number, packet-loss probability and penalty of call blocking, are further formulated to supersede their previous definitions. The general cross-layer optimization method is proposed to maximize system utilization subject to the QoS constraints. Based on the general model, a practical problem, the capacity unbalance problem, is solved with the adaptive SHP scheme. The SHP is controlled adaptively along with the changing traffic volumes in the reverse or forward link. System utilization gains are achieved with the cross-layer optimization over the conventional segregated-layer design.

The cross-influence phenomena among layers are investigated. The cross-influence phenomena are caused by two factors, the time-varying parameters in the wireless link and the QoS requirements from different services, which are the same factors as to create the soft capacity in the CDMA system. In the lower layers, the characteristics of the wireless link cause the statistical behavior among layers and consequently lead to the need of the statistical QoS guarantees in the higher layers. In the higher layers, the QoS requirements from different services give rise to the changes of priorities for the services and affect the efficiency of resource allocation in the lower layers, such as the adjustment of the SHP in the physical layer and the determination of the admission region in the link layer.

Chapter 7

Conclusions and Future Work

7.1 Conclusions

Resource allocation in cellular CDMA systems with cross-layer optimization across the physical layer, the link layer and the network layer has been analyzed theoretically in the thesis. The motivation to employ the cross-layer optimization in wireless networks comes from the recognition and understanding of the time-varying characteristics in the wireless links. The system performance in the higher layers is adjusted along with the changing conditions in the lower layers. The statistical behavior among layers leads to the necessity for the cross-layer optimization in the wireless networks. With the cross-layer optimization, QoS requirements are provided with statistical guarantees ('soft' guarantees) which allow some performance degradation within a short time but still fulfilling the QoS requirements on average, instead of hard guarantees which assume the worst-case performance in other segregated layers.

The analyses of the segregated layers are firstly processed for the network and link layers. The physical layer features are included into the link layer analysis. In the network layer analysis, two CAC schemes, CS and VP are investigated. The analytical models are based on a *K*-dimensional Markov chain and solved using the preemption rules for the schemes. The formulae for the GoS metrics at the connection-level and the QoS metrics at the packet-level for different CAC schemes are derived. The GoS metrics include the new-call-blocking probability, the handoff-call-dropping probability and the system utilization. The QoS metric includes the packet-loss probability. A method to maximize system utilization through joint optimization of connection and packet levels parameters is proposed. Numerical results indicate that significant gain in system utilization is achieved using the joint optimization.

In the link layer, the interference models are carefully built with soft handoff, diversity and statistical multiplexing in both the reverse and forward links. The analytical models are based on the largest received power BS selection criterion. In the forward link, an approximation selection method combining the advantages of previously used approximations and adapting to the QoS specification for different services is proposed. Furthermore, different power control schemes for mobile users in soft handoff are investigated and compared. The SIRs and the outage probabilities for multi-class services in the BSs (for the reverse link) and in the mobile users (for the forward link) are formulated. By constraining the outage probability to its requirement value, the admission regions are obtained.

The QoS metrics formulated in the segregated layers analyses include the new-call-blocking probability, the handoff-call-dropping probability, the system utilization and the packet-loss probability in the network layer, SIR and the outage probability in the link layer, and the hysteresis margin and the soft handoff probability (SHP) in the physical layer. These QoS metrics are selected as the intertwining parameters among layers which are transferred through an interface from the conventional segregated-layer system into a defined function block, the cross-layer decision-maker (DM). In the cross-layer DM, the cross-layer optimization is employed and the system configuration parameters are fed back to each layer as the outputs of the cross-layer DM. The designed cross-layer DM preserves the integrality of the conventional protocol structure which then can be further developed and updated.

The aim of the optimization is to maximize the system utilization subject to the QoS constraints and feed back the optimal system configurations to each layer. Based on the general cross-layer model, a practical problem, the capacity unbalance problem, is solved with an adaptive SHP scheme. The SHP is controlled adaptively along with the changing traffic volumes in the reverse or forward link. Thus, the hysteresis margin in the physical layer which determines the SHP becomes another feedback system configuration parameter. Furthermore, the QoS requirements from different services give rise to the changes of priorities for services and affect the efficiency of resource allocation in lower layers. This cross-influence phenomenon in the cross-layer model has been investigated as an interesting topic in this thesis.

Performance evaluations are processed in the network layer analysis, the link layer analysis and the cross-layer analysis. The numerical results verify the reasonableness of the modeling in the network layer and in the physical and link layers.

The proposed cross-layer model optimized the resource allocation in the segregated layers and improved the system utilization. The adaptive SHP scheme works well to minimize the call blocking and improve the efficiency of resource allocation among multiple services compared to the fixed SHP case. Combining the benefits from the cross-layer optimization and the adaptive SHP scheme, the maximum achievable gain in system utilization is about 60% as compared to the case with segregated-layer design in the numerical results.

The resource which is not optimally exploited in the segregated-layer design is released in the cross-layer model. With the cross-layer model, the parameters with statistic behavior can share information with each other and the QoS metrics then can be provided the ‘soft’ guarantees. Therefore, the cross-layer model is superior to the segregated-layer model as no assumption on the worst-case performance in any layer is needed.

7.2 Future Work

The work in the thesis proposes new methods to realize the cross-layer design with different resource allocation approaches to achieve high system utilization. To convert the current segregated-layer system to the proposed cross-layer system, the cross-layer DM is required to be applied in a practical way and the interface between the conventional system and the DM should be built. However, the practical implementations are yet to be considered in the study although they are beyond the main focus of the thesis. The implementation methods should be investigated in future work, which would realize the proposed model and prove its applications in practical systems.

Furthermore, the assumption of the worst-case performance on the NRT services should be improved by considering buffering and retransmission. For the NRT services with buffering and retransmission, the power and rate can be adaptively allocated so that the transmission for the NRT services could be slowed down and even be ceased when there are overloads from the RT services. The NRT services with buffering and retransmission should be considered into the system model analytically in the future. The tradeoff between energy efficiency in the physical and link layers and packet delay in the network layer is another topic in cross-layer wireless resource allocation [110].

Last but not least, in the link layer, the power control schemes for mobile users in soft handoff should be investigated to achieve the minimum interference in both the reverse and forward links. The promising SSDT power control scheme which outperforms other power control schemes in the forward link should be formulated analytically in both the reverse and forward links to estimate its system performance in the cross-layer model.

Bibliography

[1] J.W. Mark and W. Zhuang, *Wireless Communications and Networking*, Pearson Education, Inc., 2003.

[2] H. Holma and A. Toskala, *WCDMA for UMTS – Radio Access for Third Generation Mobile Communications*, New York: Wiley, 2001.

[3] D. Hong and S.S. Rapport, “Traffic Modeling and Performance Analysis for Cellular Mobile Radio Telephone Systems with Prioritized and Nonprioritized Handoff Procedures,” CEAS Technical Report No. 773, College of Engineering and Applied Sciences, State University of New York, Stony Brook, NY 11794, USA, Jun. 1999. (Version 2 of the 1986 *IEEE Trans. Veh. Technol.* paper)

[4] L. Jorguseski, E. Fledderus, J. Farserotu and R. Prasad, “Radio Resource Allocation in Third-Generation Mobile Communication Systems,” *IEEE Commun. Mag.*, vol. 39, no. 2, pp, 117-123, Feb. 2001.

[5] L. Huang, S. Kumar and C.-C.J. Kuo, “Adaptive Resource Allocation for Multimedia QoS Management in Wireless Networks,” *IEEE Trans. Veh. Technol.*, vol. 53, no. 2, pp. 547-558, Mar. 2004.

[6] S.A. Grandhi, R.D. Yates and D.J. Goodman, “Resource Allocation for Cellular Radio System,” *IEEE Trans. Veh. Technol.*, vol. 46, no. 3, pp. 581-587, Aug. 1997.

[7] S. Lal and E.S. Sousa, "Distributed Resource Allocation for DS-CDMA-Based Multimedia ad hoc Wireless LAN's," *IEEE J. Select. Areas Commun.*, vol. 17, no. 5, pp. 947-967, May 1999.

[8] S.-J. Oh, D. Zhang and K.M. Wasserman, "Optimal Resource Allocation in Multiservice CDMA Networks," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 811-821, Jul. 2003.

[9] D. Zhao, X. Shen and J.W. Mark, "Radio Resource Management for Cellular CDMA Systems Supporting Heterogeneous Services," *IEEE Trans. Mobile Comput.*, vol. 2, no. 2, Apr.-Jun. 2003.

[10] C. Comaniciu and H.V. Poor, "Jointly Optimal Power and Admission Control for Delay Sensitive Traffic in CDMA Networks with LMMSE Receivers," *IEEE Trans. Signal Processing*, vol. 51, no. 8, pp. 2031-2042, Aug. 2003.

[11] A. Maharshi, L. Tong and A. Swami, "Cross-Layer Designs of Multichannel Reservation MAC under Rayleigh Fading," *IEEE Trans. Signal Processing*, vol. 51, no. 8, pp. 2054-2067, Aug. 2003.

[12] S. Shakkottai, T.S. Rappaport and P.C. Karlsson, "Cross-Layer Design for Wireless Networks," *IEEE Commun. Mag.*, vol. 41, no. 10, pp, 74-80, Oct. 2003.

[13] T.C. Wong, J.W. Mark and K.C. Chua, "Joint Connection Level, Packet Level and Link Layer Resource Allocation for Variable Bit Rate Multiclass Services in Cellular DS-CDMA Networks with QoS Constraints," *IEEE J. Select. Areas Commun.*, vol. 21, no. 10, pp. 1536-1545, Dec. 2003.

[14] W. Kumwilaisak, Y.T. Hou, Q. Zhang, W. Zhu, C.-C.J. Kuo and Y-Q. Zhang, “A Cross-layer Quality-of-Service Mapping Architecture for Video Delivery in Wireless Networks,” *IEEE J. Select. Areas Commun.*, vol. 21, no. 10, pp. 1685-1698, Dec. 2003.

[15] L. Xu, X. Shen and J.W. Mark, “Dynamic Fair Scheduling with QoS Constraints in Multimedia Wideband CDMA Cellular Networks,” *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 60-73, Jan. 2004.

[16] T. Elbatt and A. Ephremides, “Joint Scheduling and Power Control for Wireless *Ad Hoc* Networks,” *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 74-85, Jan. 2004.

[17] L. Alonso and R. Agusti, “Automatic Rate Adaptation and Energy-Saving Mechanisms Based on Cross-Layer Information for Packet-Switched Data Networks,” *IEEE Commun. Mag.*, vol. 42, no. 3, pp. S15-S20, Mar. 2004.

[18] S. De, C. Qiao, D.A. Pados, M. Chatterjee and S.J. Philip, “An Integrated Cross-Layer Study of Wireless CDMA Sensor Networks,” *IEEE J. Select. Areas Commun.*, vol. 22, no. 7, pp. 1271-1285, Sept. 2004.

[19] A.J. Viterbi, “Wireless Digital Communication: a View Based on Three Lessons Learned,” *IEEE Commun. Mag.*, vol. 29, no. 9, pp, 33-36, Sep. 1991.

[20] J. Zhang, J. Huai, R. Xiao and B. Li, “Resource Management in the Next-Generation DS-CDMA Cellular Networks,” *IEEE Wireless Commun.*, vol. 11, no. 4, pp. 52-58, Aug. 2004.

[21] T.S. Rappaport, *Wireless Communications: Principles and Practice (Second*

Edition), Prentice Hall, PTR, 2002.

[22] K.S. Gilhousen, I.M. Jacobs, R. Padovani, A.J. Viterbi, L.A. Weaver and C.E. Wheatley, "On the Capacity of a Cellular CDMA System," *IEEE Trans. Veh. Technol.*, vol. 40, no. 2, pp. 303-312, May 1991.

[23] M. Cheung and J.W. Mark, "Resource Allocation in Wireless Networks Based on Joint Packet/Call Levels QoS Constraints," in *Proc. IEEE GLOBECOM*, pp. 271-275, Nov. 2000.

[24] V. Huang and W. Zhuang, "Optimal Resource Management in Packet-Switching TDD CDMA Systems," *IEEE Personal Commun.*, vol. 7, no. 6, pp. 26-31, Dec. 2000.

[25] A.C. Kam, T. Minn and K-Y. Siu, "Supporting Rate Guarantee and Fair Access for Bursty Data Traffic in W-CDMA," *IEEE J. Select. Areas Commun.*, vol. 19, no. 11, pp. 2121-2130, Nov. 2001.

[26] A.J. Viterbi, A.M. Viterbi, K.S. Gilhousen and E. Zehave, "Soft Handoff Extends CDMA Cell Coverage and Increases Reverse Link Capacity," *IEEE J. Select. Areas Commun.*, vol. 12, no. 8, pp. 1281-1288, Oct. 1994.

[27] D. Wong and T.J. Lim, "Soft Handoff in CDMA Mobile Systems," *IEEE Personal Commun.*, vol. 4, no. 6, pp. 6-17, Dec. 1997.

[28] A.J. Viterbi, *CDMA: Principles of Spread Spectrum Communication*, Addison-Wesley Publishing Company, 1995.

[29] D-W. Tcha, S-Y. Kang and G-W. Jin, "Load Analysis of the Soft Handoff Scheme in a CDMA Cellular System," *IEEE J. Select. Areas Commun.*, vol.

19, no. 6, pp. 1147-1152, Jun. 2001.

[30] C.E.U-. Rios and D.L-. Rodriguez, “Forward Link Capacity Losses for Soft and Softer Handoff in Cellular Systems,” in *Proc. IEEE PIMRC*, pp. 48-53, Sept. 2001.

[31] Y.H. Kwon, D.C. Lee and W. Park, “Capacity analysis of forward link with deterministic power model in CDMA systems with adaptive antenna array and soft handoff,” in *Proc IEEE VTC-Spring*, pp. 335-339, May 2002.

[32] R. Kohno, R. Meidan and L.B. Milstein, “Spread Spectrum Access Methods for Wireless Communications,” *IEEE Commun. Mag.*, vol. 33, no. 1, pp. 58-67, Jan. 1995.

[33] J.M. Capone and I. Stavrakakis, “Determining the Call Admission Region for Real-Time Heterogeneous Applications in Wireless TDMA Networks,” *IEEE Network*, vol. 12, no. 2, pp. 38-47, Mar.-Apr. 1998.

[34] 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. IEEE ICCS*, vol. 2, pp. 692-696, 2002.

[35] J. Yao, T.C. Wong, Y.H. Chew and C.C. Chai, “Resource Allocation for Multiclass Traffic in Cellular CDMA Networks with Soft Capacity and Soft Handoff,” in *Proc. CIC*, Conference CD-ROM, 28-31 Oct. 2003.

[36] J. Yao, T.C. Wong and Y.H. Chew, “Cross-Layer Design on the Reverse and Forward Links Capacities Balancing in Cellular CDMA Systems,” in *Proc. IEEE WCNC*, pp. 21-25, Mar. 2004.

[37] J. Yao, J.W. Mark, T.C. Wong, Y.H. Chew, K.M. Lye and K.C. Chua, “Virtual Partitioning Resource Allocation for Multiclass Traffic in Cellular Systems with QoS Constraints,” *IEEE Trans. Veh. Technol.*, vol. 53, no. 3, pp. 847-864, May 2004.

[38] J. Yao, Y.H. Chew and T.C. Wong, “Forward Link Capacity in Multi-service Cellular CDMA Systems with Soft Handoff in the Presence of Path Loss and Lognormal Shadowing”, in *Proc. IEEE ICCS*, 2004.

[39] J. Yao, T.C. Wong and Y.H. Chew, “Capacity Balancing between the Reverse and Forward Links in Multiservice CDMA Cellular Networks with Cross-Layer Design,” to appear in *IEEE Trans. Veh. Technol.*.

[40] J. Yao, T.C. Wong and Y.H. Chew, “A Cross-Layer Design Structure in Cellular CDMA Systems,” in preparation.

[41] 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 ICCS*, Nov. 2000.

[42] Y. Fang and Y. Zhang, “Call Admission Control Schemes and Performance Analysis in Wireless Mobile Networks,” *IEEE Trans. Veh. Technol.*, vol. 51, no. 2, pp. 371-382, Mar. 2002.

[43] A. Malla, M. El-Kadi, S. Olariu and P. Todorova, “A Fair Resource Allocation Protocol for Multimedia Wireless Networks,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 14, no. 1, pp. 63-71, Jan. 2003.

[44] Y. Haung and J. Ho, “Distributed Call Admission Control for a Heterogeneous

PCS Network," *IEEE Trans. Comput.*, vol. 51, no. 12, pp. 1400-1409, Dec. 2002.

[45] S. Shen, C-J. Chang, C. Huang and Q. Bi, "Intelligent Call Admission Control for Wideband CDMA Cellular Systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 5, pp. 1810-1821, Sept. 2004.

[46] N. Dimitriou, R. Tafazolli and G. Sfikas, "Quality of Service for Multimedia CDMA," *IEEE Commun. Mag.*, vol. 38, no. 7, pp, 88-94, Jul. 2000.

[47] D. Mitra and I. Ziedins, "Virtual Partitioning by Dynamic Priorities: Fair and Efficient Resource Sharing by Several Services," in *Broadband communications*, B. Plattner, Ed: Springer-Verlag, pp. 173-185, 1996.

[48] D. Mitra, M.I. Reiman and J. Wang, "Robust Dynamic Admission Control for Unified Cell and Call QoS in Statistical Multiplexers," *IEEE J. Select. Areas Commun.*, vol. 16, no. 5, pp. 692-707, Jun. 1998.

[49] G.L. Wu and J.W. Mark, "A Buffer Allocation Scheme for ATM Networks: Complete Sharing Based on Virtual Partition," *IEEE/ACM Trans. Networking*, vol. 3, no. 6, pp. 660-670, Dec. 1995.

[50] S.C. Borst and D. Mitra, "Virtual Partitioning for Robust Resource Sharing: Computational Techniques for Heterogeneous Traffic," *IEEE J. Select. Areas Commun.*, vol. 16, no. 5, pp. 668-678, Jun. 1998.

[51] T.C. Wong, J.W. Mark and K.C. Chua, "Resource Allocation in Mobile Cellular Networks with QoS Constraints," in *Proc. IEEE WCNC*, pp. 717-722, Mar. 2002.

[52] T.C. Wong, J.W. Mark, M. Jin, B. Bensaou and K.C. Chua, “Resource Allocation in Mobile Multimedia Networks – Analytical and Simulation Results for Case 1 to 4 with Fixed Guard Channels,” Technical Report RA-001-07-2000, Centre for Wireless Communications, National University of Singapore, Jul. 2000.

[53] T.C. Wong, J.W. Mark and K.C. Chua, “Resource Allocation in Mobile Multimedia Networks with User Mobility – Case 4: Virtual Partitioning with Preemption for Groups 1 and 2,” Technical Report RA-002-10-2001, Centre for Wireless Communications, National University of Singapore, Oct. 2001.

[54] M. Beshai, R. Kositpaiboon and J. Yan, “Interaction of Call Blocking and Cell Loss in an ATM Network,” *IEEE J. Select. Areas Commun.*, vol. 12, no. 6, pp. 1051-1058, Aug. 1994.

[55] C.-N. Wu, Y.-R. Tsai and J.-F. Chang, “A Quality-Based Birth-and-Death Queueing Model for Evaluating the Performance of an Integrated Voice/Data CDMA Cellular System,” *IEEE Trans. Veh. Technol.*, vol. 48, no. 1, pp. 83-89, Jan. 1999.

[56] T. Dahlberg, K.R. Subramanian and B. Cao, “Soft Capacity Modeling for Third Generation Radio Resource Management,” in *Proc. MobiWac*, pp. 33-39, Oct. 2002.

[57] D.K. Kim and D.K. Sung, “Traffic management in a multicode CDMA system supporting soft handoffs,” *IEEE Trans. Veh. Technol.*, vol. 51, no. 1, pp. 52-62, Jan. 2002.

[58] R.P. Narrainen and F. Takawira, “Performance Analysis of Soft Handoff in CDMA Cellular Networks,” *IEEE Trans. Veh. Technol.*, vol. 50, no. 6, pp. 1507-1517, Jan. 2001.

[59] F. Ashtiani, J.A. Salehi and M.R. Aref, “A Flexible Dynamic Traffic Model for Reverse Link CDMA Cellular Networks,” *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 159-164, Jan. 2004.

[60] D.G. Jeong and W.S. Jeon, “CDMA/TDD system for wireless multimedia services with traffic unbalance between uplink and downlink,” *IEEE J. Select. Areas Commun.*, vol. 17, no. 5, pp. 939-946, May 1999.

[61] D. Kim and D.G. Jeong, “Capacity Unbalance between Uplink and Downlink in Spectrally Overlaid Narrow-Band and Wide-Band CDMA Mobile Systems,” *IEEE Trans. Veh. Technol.*, vol. 49, no. 4, pp. 1086-1093, Jul. 2000.

[62] J.Y. Kim and G.L. Stuber, “CDMA Soft Handoff Analysis in the Presence of Power Control Error and Shadowing Correlation,” *IEEE Trans. Wireless Commun.*, vol. 1, no. 2, pp. 245-255, Apr. 2002.

[63] L. Zhuge and V.O.K. Li, “Forward Link Capacity in Multi-Service DS-CDMA Systems,” in *Proc. IEEE GLOBECOM*, vol. 1, pp. 609-613, Nov. 2001.

[64] F.J. J-Romero and D. M-Rodriguez, “Generalized Bayesian Hypothesis Testing for Cell Coverage Determination,” *IEEE Trans. Veh. Technol.*, vol. 49, no. 4, pp. 1102-1109, Jul. 2000.

[65] W.S. Jeon and D.G. Jeong, “Call Admission Control for Mobile Multimedia Communications with Traffic Asymmetry between Uplink and Downlink,”

IEEE Trans. Veh. Technol., vol. 50, no. 1, pp. 59-66, Jan. 2001.

[66] N. Maeda, Y. Kishiyama, K. Higuchi, H. Atarachi and M. Sawahashi, “Experimental Evaluation of Throughput Performance in Broadband Packet Wireless Access Based on VSF-OFCDM and VSF-CDMA,” in *Proc. IEEE PIMRC*, vol. 1, pp. 6-11, Sept. 2003.

[67] J. Wu, S. Affes and P. Mermelstein, “Forward-link soft-handoff in CDMA with multiple-antenna selection and fast joint power control,” *IEEE Trans. Wireless Commun.*, vol. 2, no. 3, pp. 459-471, May 2003.

[68] X. Gao, G. Wu and T. Miki, “End-to-End QoS Provisioning in Mobile Heterogeneous Networks,” *IEEE Wireless Commun.*, vol. 11, no. 3, pp. 24-34, Jun. 2004.

[69] G. Carneiro, J. Ruela and M. Ricardo, “Cross-Layer Design in 4G Wireless Terminals,” *IEEE Wireless Commun.*, vol. 11, no. 2, pp. 7-13, Apr. 2004.

[70] M.D. Kulavaratharasah and A.H. Aghvami, “Teletraffic Performance Evaluation of Microcellular Personal Communication Networks (PCN’s) with Prioritized Handoff Procedures,” *IEEE Trans. Veh. Technol.*, vol. 48, no. 1, pp. 137-152, Jan. 1999.

[71] M. Bozinovski, P. Popovski and L. Gavrilovska, “QoS-based Policy for Call Admission Control in Mobile Cellular Network,” in *Proc. IEEE WCNC*, pp. 502-506, 2000.

[72] D. Calin and M. Areny, “Impact of radio resource allocation policies on the TD-CDMA system performance: evaluation of major critical parameters,”

IEEE J. Select. Areas Commun., vol. 19, no. 10, pp. 1847-1859, Oct. 2001.

[73] W.C. Jakes, *Microwave Mobile Communications*, New York: IEEE Press, 1974.

[74] W. Choi and J.Y. Kim, “Forward-link capacity of a DS/CDMA system with mixed multirate sources,” *IEEE Trans. Veh. Tech.*, vol. 50, no. 3, pp. 737-749, May 2001.

[75] W.-U. Pistelli and R. Verdone, “Power allocation strategies for the downlink in a W-CDMA system with soft and softer handover: the impact on capacity,” in *proc. IEEE PIMRC*, pp. 5-10, 2002.

[76] 3GPP TR 25.922 v5.0.0, “Radio resource management strategies,” 2003.

[77] Y. Chen and L.G. Cuthbert, “Downlink performance of different soft handover algorithms in 3G multi-service environments,” in *proc. IEEE MWCN*, pp. 406-410, Sept. 2002.

[78] A. Morimoto, K. Higuchi and M. Sawahashi, “Site independent diversity transmit power control for inter-cell site diversity in W-CDMA forward link,” in *proc. IEEE VTC-fall*, pp. 645-649, Oct. 2001.

[79] D. Staehle, K. Leibnitz and K. Heck, “Effects of soft handover on the UMTS downlink performance,” in *proc. IEEE VTC-fall*, pp. 960-964, Sept. 2002.

[80] F. Blaise, L. Elicegui, F. Goeusse and G. Vivier, “Power control algorithms for soft handoff users in UMTS,” in *proc. IEEE VTC-fall*, pp. 1110-1114, Sept. 2002.

[81] H. Furukawa, K. Hamabe and A. Ushirokawa, “SSDT – site selection

diversity transmission power control for CDMA forward link," *IEEE J. Select. Areas Commun.*, vol. 18, no. 8, pp. 1546-1554, Aug. 2000.

[82] K. Takeo and S. Sato, "Adaptive power allocation for MSs under soft handoff considering traffic deviation in CDMA systems," in *Proc. IEEE ICCT*, pp. 21-25, Aug. 2000.

[83] L. Tong and P. Ramanathan, "Adaptive Power and Rate Allocation for Service Curve Assurance in DS-CDMA Network," *IEEE Trans. Wireless Commun.*, vol. 3, no. 2, pp. 555-564, Mar. 2004.

[84] F. Berggren and S.-J. Kim, "Energy-Efficient Control of Rate and Power in DS-CDMA System," *IEEE Trans. Wireless Commun.*, vol. 3, no. 3, pp. 725-733, May 2004.

[85] D.G. Jeong and W.S. Jeon, "Congestion Control Schemes for Reverse Link Data Transmission in Multimedia CDMA Systems," *IEEE Trans. Veh. Technol.*, vol. 52, no. 6, pp. 1489-1496, Nov. 2003.

[86] L.O.- Guerrero and A.H. Aghvami, "A Distributed Dynamic Resource Allocation for a Hybrid TDMA/CDMA System," *IEEE Trans. Veh. Technol.*, vol. 47, no. 4, pp. 1162-1178, Nov. 1998.

[87] Selection procedures for the choice of radio transmission technologies of the UMTS, ETSI TR 101 112, 1998.

[88] M. Cheng and L.F. Chang, "Wireless Dynamic Channel Assignment Performance under Packet Data Traffic," *IEEE J. Select. Areas Commun.*, vol. 17, no. 7, pp. 1257-1269, Jul. 1999.

[89] R.G. Garroppo, S. Giordano, M. Isopi and M. Pagano, “On the Implications of the OFF Periods Distribution in Two-State Traffic Models,” *IEEE Commun. Letters*, vol. 3, no. 7, pp. 220-222, Jul. 1999.

[90] S. Thilakawardana and R. Tafazolli, “Effect of Service Modeling on Medium Access Control Performance,” in *Proc. PIMRC*, pp. 60-63, Sept. 2001.

[91] M. Zukerman, T.D. Neame and R.G. Addie, “Internet Traffic Modeling and Future Technology Implications,” in *Proc. INFOCOM*, pp. 587-596, Mar. 2003.

[92] P. Sen, B. Maglaris, N.E. Rikli and D. Anastassiou, “Model for packet switching of variable-bit-rate video sources,” *IEEE J. Select. Areas Commun.*, vol. 7, no. 5, pp. 865-869, Jun. 1989.

[93] P. Lin and Y. Lin, “Channel allocation for GPRS,” *IEEE Trans. Veh. Technol.*, vol. 50, no. 2, pp. 375-387, Mar. 2001.

[94] W. Shan, P. Fan and Y. Pan, “Performance evaluation of a hierarchical cellular system with mobile velocity-based bi-directional call-overflow scheme,” *IEEE Trans. Parallel Distrib. Syst.*, vol. 14, no. 1, pp. 72-83, Jan. 2003.

[95] G. Haring, R. Marie, R. Puigjaner and K. Trivedi, “Loss formulas and their application to optimization for cellular networks,” *IEEE Trans. Veh. Technol.*, vol. 50, no. 3, pp. 664-673, May 2001.

[96] K.W. Ross, *Multiservice Loss Models for Broadband Telecommunication Networks*, Springer-Verlag, 1995.

[97] R.P. Narrainen and F. Takawira, “A traffic model for CDMA cellular systems

with soft capacity taken into account," *IEEE Sixth International Symposium on Spread Spectrum Techniques and Applications*, vol. 1, pp. 325-329, 2000.

[98] W.H. Press, S.A. Teukolsky, W.T. Vettering and B.P. Flannery, *Numerical recipes in C – The art of scientific computing (Second Edition)*, Cambridge Univ. Press, 2002.

[99] Quality of service (QoS) concept and architecture, 3GPP, TS 23.107, 2002.

[100] T.C. Wong, J.W. Mark and K.C. Chua, "Performance evaluation of video services in a multirate DS-CDMA system," in *Proc. IEEE PIMRC,* pp. 1490-1495, Sept. 2003.

[101] L.F. Fenton, "The Sum of Log-Normal Probability Distributions in Scatter Transmission Systems", *IEEE Trans. Syst.*, vol. CS-8, pp. 57-67, Mar. 1960.

[102] S.C. Schwartz and Y.S. Yeh, "On the Distribution Function and Moments of Power Sums with Log-Normal Components," *Bell Syst. Tech. J.*, vol. 61, pp. 1441-1462, Sept. 1982.

[103] A. Safak, "Statistical Analysis of the Power Sum of Multiple Correlated Log-Normal Components," *IEEE Trans. Veh. Technol.*, vol. 42, no. 1, pp. 58-61, Feb. 1993.

[104] A.A.A- Dayya and N.C. Beaulieu, "Outage Probabilities in the Presence of Correlated Lognormal Interferers," *IEEE Trans. Veh. Technol.*, vol. 43, no. 1, pp. 164-173, Feb. 1994.

[105] N.C. Beaulieu, A.A.A- Dayya and P.J. McLane, "Estimating the Distribution of a Sum of Independent Lognormal Random Variables," *IEEE Trans.*

Commun., vol. 43, no. 12, pp. 2869-2873, Dec. 1995.

[106] M. Zorzi, "On the Analytical Computation of the Interference Statistics with Applications to the Performance Evaluation of Mobile Radio Systems," *IEEE Trans. Commun.*, vol. 45, no. 1, pp. 103-109, Jan. 1997.

[107] CSELT, "Impact of the Number of Pilot Bits on the Uplink Performance," Rep. 3GPP TSG RAN WG1, TSGR 1#3(99)170, Mar. 1999.

[108] J. Zhang, J.W. Mark and X. Shen, "A Novel Resource Reservation Scheme for Handoff in CDMA Wireless Cellular Networks," in *Proc. IEEE WCNC*, pp. 2069-2074, Mar. 2003.

[109] P. Liu, P. Zhang, S. Jordan and M.L. Honig, "Single-Cell Forward Link Power Allocation Using Pricing in Wireless Networks," *IEEE Trans. Wireless Commun.*, vol. 3, no. 2, pp. 533-543, Mar. 2004.

[110] R.A. Berry and E.M. Yeb, "Cross-Layer Wireless Resource Allocation," *IEEE Signal Proc. Mag.*, vol. 21, no. 5, pp. 59-68, Sept. 2004.

[111] J. Yao, etc., "Resource Allocation for End-to-End QoS Provisioning in a Hybrid Wireless Wideband CDMA and Wireline IP-based Network," the final report of the Singapore – Ontario Joint Research Programme, Apr. 2005.