MOBILITY MANAGEMENT IN WIRELESS ATM NETWORK
(PENGURUSAN MOBILITI DALAM RANGKAIAN ATM WIRELES)

SHARIFAH HAFIZAH SYED ARIFFIN

RESEARCH VOT NO:
71464

Faculty of Electrical Engineering
Universiti Teknologi Malaysia

2008

MOBILITY MANAGEMENT IN WIRELESS ATM NETWORK

(Keywords: Wireless ATM, handoff probability, blocking probability)

Traffic control and resource management are complex issues in wireless asynchronous transfer mode (WATM) due to broad range quality-of-services (QoS), limited bandwidth spectrum and susceptibility to error. In WATM, traffic control is essential in order to protect the network from congestion and to achieve realistic network efficiency in compliance with the QoS. Micro and pico cell sites are usually used in the WATM environment. Exchanging access point to the backbone network is thus common and this may contribute to connection rejections when channels are busy. The effect to QoS at call level is the high probability of handoff dropping if the targeted call site has no available channels. The objective of this project is to design a WATM system using virtual connection tree (VCT) architecture model that can give high QoS performance in terms of probability of call lost and handoff call drop and can provide connections of a bandwidth rate that will not jeopardize QoS requirements in terms of a bandwidth given to a connection which is not less than the set threshold bandwidth. All of these will be done in mathematical analysis. The design was initially started with a scheme called With Reserved Channel (WRC), which has reserved channels for handoff calls. A low probability of handoff dropping was obtained but a set back was found which involves the probability of new call block. Hence, an improvement had been made and it is called With Queuing and Reserved Channels (WQRC), which allows queuing of new incoming calls and reserved channels for handoff calls. Using WQRC, low probability of handoff dropping was maintained and new call blocking probability was eliminated

Key researchers:
Sharifah Hafizah Syed Ariffin (head)
Norsheila Fisal
Rozeha Rashid
Fatimah Mohamed

Email : sharifah@fke.utm.my
Tel. No: 07 5535348
Vote No.: 71464

PENGURUSAN MOBILITI DALAM RANGKAIAN ATM WIRELES

(Katakunci: wireless, ATM kebarangkaian hilang, kebarangkalian tukar panggilan)

Pengawalan trafik dan pengurusan sumber adalah isu kompleks dalam mod penghantaran tidak serentak tanpa wayar (WATM) disebabkan oleh julat kualiti perkhidmatan (QoS) yang besar, had spektrum lebarjalur dan mudah memberi ralat. Dalam WATM, pengawalan trafik adalah penting untuk mengelakkan rangkaian dari kesesakan dan mencapai tahap rangkaian yang cekap. Walau bagaimanapun pengawalan trafik harus patuh kepada kehendak parameter QoS iaitu kehilangan panggilan baru, kehilangan panggilan tukar, kelewatan panggilan dan pemberian kadar lebarjalur yang sesuai. Dalam WATM sel tapak yang digunakan adalah jenis mikro dan piko. Maka apabila pengguna bergerak dari satu sel tapak ke sel tapak yang lain penukaran poin masuk ke rangkaian tulang belakang akan meningkat. Ini menyebabkan penolakan panggilan oleh rangkaian yang sibuk selalu terjadi. Kesannya adalah pada QoS, iaitu kebarangkalian yang tinggi terhadap kehilangan panggilan tukar akan terjadi jika saluran di sel tapak sasaran telah penuh. Objektif projek ini adalah untuk mereka satu sistem WATM menggunakan model pokok sambungan maya (VCT) yang boleh memberikan QoS yang baik serta boleh memberikan kadar lebarjalur pada sambungan yang tidak mengugat kehendak QoS. Analisis dalam tesis ini berdasarkan matematik. Rekaan sistem telah dimulakan dengan menggunakan sistem WRC yang mempunyai saluran simpanan untuk panggilan tukar. Kebarangkalian yang rendah untuk kehilangan panggilan tukar diperolehi tetapi kebarangkalian terhadap kehilangan panggilan baru masih tinggi. Pembaikan terhadap sistem ini telah dilakukan dan sistem baru telah dibangunkan dan dikenali sebagai WQRC. Kebarangkalian yang rendah terhadap kehilangan panggilan tukar dapat dikekalkan dan kebarangkalian panggilan baru dapat dikurangkan. Disebabkan oleh pengumpulan stesen tapak dalam VCT kebarangkalian beban lebih dalam rangkaian dapat dikurangkan.

Penyelidik-penyelidik utama:
Sharifah Hafizah Syed Ariffin (ketua)
Norsheila Fisal
Rozeha Rashid
Fatimah Mohamed

Email : sharifah@fke.utm.my
Tel. No: 07 5535348
Vote No.: 71464

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

GLOSSARY OF SYMBOLS AND ABBREVIATIONS

λ_{rt}	new incoming call arrival rate for real-time connection
λ_{ert}	effective new incoming call arrival rate for real-time connection
λ_{nrt}	new incoming call arrival rate for non real-time connection
λ_{rtI}	new incoming call arrival rate for real-time connection in WORC
λ_{rtII}	new incoming call arrival rate for real-time connection in WRC
λ_{II}	new incoming call arrival rate for real-time connection in WQRC
μ_{rt}	departure call rate for real-time connection
μ_{ert}	effective departure call rate for real-time connection
μ_{nrt}	departure call rate for non real-time connection
μ_{II}	departure call rate for real-time connection for WQRC system
μ_{rtI}	departure call rate for real-time connection for WORC
μ_{rtII}	departure call rate for real-time connection for WRC
γ_{rt}	departure of handoff call rate for real-time connection
γ_{a}	arrival of handoff call rate for real-time connection
m	number of connection a base station can handle
m_{II}	number of connection a base station can handle in WQRC system
B	number of base station in a cell cluster
B_{II}	number of base station in a cell cluster in WQRC system
P_{HD}	probability of handoff call drop
P_{AB}	probability of new call block
$P_{rt}(i)$	probability of i real-time connection in a system
$P_{nrt}(s)$	probability of s non real-time connection in a system
P_c	probability of call block
$p(j)$	probability of connection j in a system
$p_A(j)$	probability of connection j in a truncated A system
P_o	overload probability
Nrt	pre determined threshold of number of call accepted
Nnrt	maximum number of non real-time connection to a network
N_{II}	pre determined threshold of number of call accepted for WQRC system
B_r	connection burstinest
B_p	peak bit rate
B_m	mean bit rate
B_E	effective bandwidth
$Ploss$	probability of cell lost
C	capacity of a base station
$C1$	amount of capacity assigned for real-time connection
C-$C1$	amount of capacity assigned for non real-time connection
BW1	bandwidth rate assigned for real-time connection
UB	unit bandwidth

ABR	available bit rate
ATM	asynchronous transfer mode
BISDN	broadband integrated digital network
BER	bit error rate
CPE	customer premises equipment
CBR	constant bit rate
CAC	call admission control
FPLMTS	Future Public Land Mobile Telecom system
HLR	home location register
MT	mobile terminal
MAC	medium access control
MSC	mobile switch connection
nrt-VBR	non real-time variable bit rate
NCP	network call processor
PCN	personal communication network
QoS	quality-of-service
RSS	radio signal strength
RS	root switch
rt-VBR	real-time variable bit rate
TE	terminal equipment
UPC	usage parameter control
UBR	unspecified bit rate
UMTS	Universal Mobile Telecommunication system
VC	virtual connection
VLR	visitor location register
VCT	virtual connection tree
WRC	with reserved channels
WORC	without reserved channels
WQRC	with queuing of new incoming calls
WATM	wireless asynchronous transfer mode
g	percentage of the received channels

LIST OF APPENDIX

CHAPTER I

INTRODUCTION

1.1 Standard ATM

Broadband telecommunication networks, according to the I-300 series of the International Telecommunication Union - Telecommunication Sector (ITU-T) recommendations, are based on a packet switching technique, established in 1990/91, the so-called asynchronous transfer mode (ATM) [1]. ATM is a data transport technology that supports a single high-speed infrastructure for integrated broadband communications involving voice, video and data. It achieves bandwidth efficiency through statistical multiplexing of transmission bandwidth. ATM networks are characterized by virtual channel connections (VCCs) that carry small, fixed size packets (53 bytes) called cells within the network irrespective of the applications being supported [2].

Instead of assigning a fixed capacity during call establishment on the transmission paths of the network for connections between communicating terminals as done in the integrated service digital network (N-ISDN), the broadband ISDN (B-ISDN) uses VCCs to route ATM cells through the network. The data are usually transmitted over a coaxial or optic fibre cable of high transfer rate (155, 600, 1200 Mb/s), low bit error ratio (e.g., 10^{-9}) and cell loss ratio at about 10^{-6} [3]. At the end of the equipment, an ATM adaptation layer (AAL) maps the service offered by the ATM network to the services required by the application. This enables ATM to handle a wide range of information bit rate together with various types of real-time and non real-time service classes with different traffic attributes and quality-of-service (QoS) guarantees at cell and call levels. Unlike local area network (LAN) technologies such as Ethernet, ATM is distance-independent and can be deployed as local or wide area networks.

1.2 Wireless ATM (WATM)

A significant development emerged in the 1990's, which is the area of wireless personal communications. Wireless communication offer borderless freedom to the customer who is using mobile devices such as phones, television and wireless computers. According to some forecast the number of mobile phone subscribers may reach one billion by the year 2010 and surpass fixed phone lines [4]. As market demands for wireless communication continue to grow rapidly, it is apparent that a new generation of wireless networks will be needed for two main reasons:

- To enable different wireless technologies to interworking seamlessly with existing wired networks;
- To meet the diverse traffic demands (e.g. voice, video, data) required by current and future customers.

The convergence of ATM in the wired and wireless domains serves as an effective platform in achieving these objectives. This trend can be found in several projects within research program of the European Community (EC). There, a new generation of mobile communications technology had been introduced as the Universal Mobile Telecommunication System (UMTS) for the RACE project, which focus on the design of the Mobile Broadband System (MBS), a wireless cellular network fully integrated into B-ISDN. This mobile application can reach an approximate range of 2 to 155 Mb/s and RACE mobile is now studying a Mobile Broadband System operating in 60 GHz bands [5].

1.3 Challenges in WATM

The extension of ATM to WATM however, creates a new set of challenging issues. The ATM standards are developed based on high data rates and reliable transmission links. This is contrast to wireless channels; hence, WATM raises a number of challenges.

- Integrated multimedia services require high user data rates (from 2Mb/s up), which are nowadays common in high-speed wired networks but still a challenge over the radio link.
- The limited radio spectrum where the capacity available for wireless access service is generally limited by regulation. Thus, unlike wired communications wherein an increasing user population can easily be served by deploying additional wired or fiber facilities to connect those users to the network, the available radio spectrum cannot arbitrarily be expanded [6].
- Mobility management handles the mobility of terminals in a wireless ATM networks. It consists of three basic issues, which are location management, connection management and handoff management. In handoff management connectionless network are having problems with handing off calls where it have to tear down and set up segments or establish extra segments for handoff. Hence, the handoff procedure gives a significant impact to the wireless network system. The probability of handoff dropping increases as cell sites sizes diminishes and mobile subscriber growing rapidly. The need to concentrate on handoff calls is because, to a MT, the premature termination of a call is less desirable than a new attempt call. Thus, a scheme is needed to decrease the probability of handoff dropping to give better service to the customer.
- The integration of mobility within B-ISDN implies the dynamic reestablishment of the ATM virtual circuits (VCs) within the short time span of the MT handoff from one cell site to another. In addition, an important goal of the VC reestablishment procedure is to ensure in sequence and loss-free delivery of the ATM cells containing user data in order to guarantee the QoS requirements on the connection.

1.4 Mobility Management in WATM

The three basic issues are [7]:

1) Location management
2) Connection management
3) Handoff management

1.4.1 Location Management

Location management has two functions: tracking the position of a mobile (with registration) and handling queries regarding the location of a mobile (for example, prior to call delivery to a mobile) which in Public Land and Mobile Network (PLMN). The location information is usually maintained and used by the network to locate the user for call routing purposes [8]. This may be a valid model to adopt for wide area WATM networks. In location management, each base station keeps broadcasting on a periodic basis, the cell sites identities on the “broadcast control channels” of the cell site under its coverage. The mobile terminal within each cell site keeps monitoring such information. As changes in location are detected (from the last information recorded by them), they each report the new location to the base station, which route it to the Visitor Location Register (VLR) of the Mobile Switch Control (MSC) to which it is connected. The VLR in turn, send the location information to the users’ HLR where it is also recorded. In the meantime, the Home Location Register (HLR) directs the old VLR to delete the old visiting location of the mobile from its database and also sends a copy of the users service profile to the new VLR.

1.4.2 Connection Management

Connection management deals with connection routing and QoS maintenance. ATM is a connection-oriented technology with a connection establishment phase prior to data exchange and a connection release phase after data exchange. In wireless networks, as end points move, segments of connection need to be torn down and reestablished. Meanwhile, maintaining call sequence and connection QoS while performing handoff are important requirement in wireless ATM network.

In chapter 2 section 2.7, different types of connection rerouting methods for handoff will be discuss. Some schemes require segment tear down and set up during handoff. Some other schemes minimize handoff time and packet loss during handoff by using pre-establishing virtual connection tree approach.

The latter schemes avoid segment tear down and setup during handoff by setting up extra segments of a connection tree (from a root switch to a set of base stations to which the mobile terminal could potentially move) during the connection setup phase. In WATM network, it completely avoids this problem of having to tear down and setup segment to reduce delay in data transmission. Hence, handoff-rerouting methods such as pre-establishing and increment extensions (which are discussed in detail in chapter 2) are widely proposed for WATM networks.

1.4.3 Handoff Management

Handoff management; quickly and timely handoff has crucial effect on how users perceive QoS [9]. However, handoff strategies should not be complicated. Wireless ATM provides mobility and freedom of borderless area to its users. Thus as a user moves from one area to another, the connection of that particular MT needs to handoff to a new base station in the current cell site or another adjacent cell site and the process continues as the user moves again. WATM environment uses small cell sites (pico cell -

~ 1km and micro cell - ~ 0.5km), compared to the amount of handoff experienced by the MT in a macro cell site (>1.5 km). to accommodate its growing population of mobile users [10]. This phenomenon affects the performance of the system especially during handoff events. In no-protection call handling schemes, handoff requests are treated the same manner as the new incoming call, so the probability of handoff dropping equals the probability of blocking of the new incoming calls. However from the MT point of view, the dropping of handoff calls is less desirable than the new call attempt. Hence some measurement must be taken to avoid too many terminations of these handoff calls. In this case handoff management of the network have to determine a way to reduce unnecessary termination.

There are various technical approaches that have been used to improve the probability of handoff dropping in WATM mobile system. The most common used technique is the prioritizing the handoff calls. There are two types of prioritizing schemes; a) Queuing of the handoff request and b) Reserving channels for handoff. Both prioritizing schemes had been investigated by many researches.

1.5 QoS Level in ATM Traffic

A three level hierarchical model with different time scales can describe ATM traffic. In this work the QoS to be upgrade will be at call level as well as cell level. The call level has a typical time of seconds (for example a short telephone call) up to hours (for example, a long lasting video conference). Their burst level is in **millisecond range** up to seconds and the cell level is in microseconds range. These levels have different impacts on an ATM network [11].

***Call level*:** The level that controls by the connection admission control (CAC) algorithms, which decide whether a connection can be accepted or not. If too many connections are rejected, the utilization of the network will degrade. Additional network resources should be deployed. The number of end users who want to transmit information concurrently will have to determine the time scale at this level.

***Burst level*:** This level can be controlled by fast resource reservation mechanism and adaptive flow control protocols like ABR. It determines the large buffers needed for non real-time connection. Furthermore, specific CAC algorithm may allow connection to be statistically multiplied at burst level. Burst periods are generated by end application, which transmit more or less information during the connection lifetime.

***Cell level*:** This level can be controlled by mechanism such as policing, priority controller or traffic shaping. At multiplexing points, the traffic patterns at cell level determines the buffer size required for real-time connections. Cell arrival variations are caused by the principles of asynchronous multiplexing of cells.

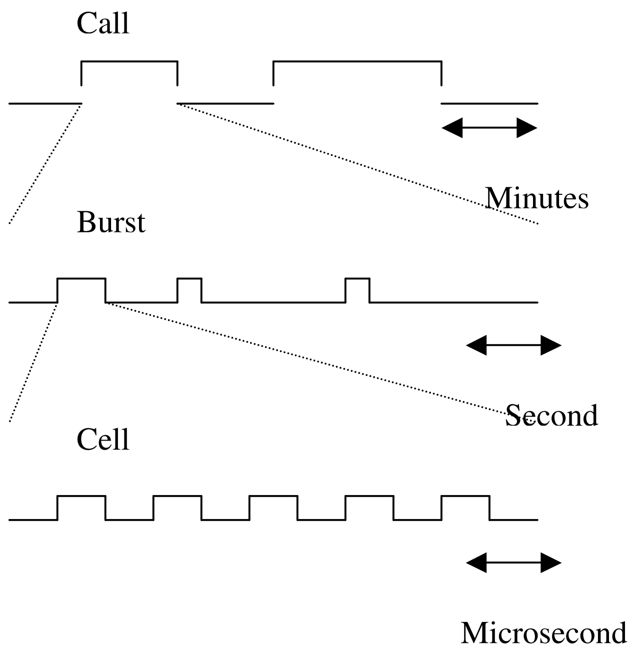


Figure 1.1: The three level quantification of traffic

Based on the traffic model shown in figure 1.1 mathematical models have been defined with different levels of abstraction [12]: a) Burtiness (peak bit rate/ average bit rate) b) Geometrically distributed burst lengths c) Switched Poisson process d) Markov Modulated Poisson Process e) Generally Modulated Deterministic Process. In this thesis the QoS at call are refers to traffic dependent performance metrics such as message loss and message delay, unless it is stated other wise.

1.6 Objective of the Study

The objective of this research is to design a scheme that can improve QoS performance of VCT network. The QoS performance will be in terms of probability of new call lost, the probability of handoff call dropping for real-time connections and the probability of overload for non real-time connections. The scheme should also maintain each probability without affecting the other.

1.7 Scope of the Study

The study uses virtual connection tree (VCT) as the architecture model in WATM environment. VCT grouped adjacent cell sites into a cell cluster and each cell sites has at least one base station. The tree contains wired network and wireless network but this study concentrate on the wireless part from MT to the base station. The types of call generated into the network will be the real-time connection and the non real-time connections. The call arrivals are generating as Poisson process and the call departure rate is generated exponential distribution.

The study concentrate on the performance of WATM at call level where probability of new call block, probability of handoff call drop, probability of call delay and probability of overload are the quality-of-service metrics. Considering network that support variety of traffic services, the value of the effective bandwidth given to a non real-time connection will also affect the performance of WATM in VCT. The study improves the VCT by allowing reserved channels for handoff calls during busy hours and allowing queuing for new incoming calls. However, this should not affect the calls that are already connected to the network.

CHAPTER II

LITERATURE REVIEW

2.1 ATM traffic Characteristics

ATM network is designed to support many types of traffic simultaneously, including real-time application such as voice, video and non real-time application such as data. Even though both applications are transmitted in one network, each of the different traffic is handle depending on the characteristics and the requirement of the application [13]. For example, real-time video traffic must be delivered within minimum variation in delay. Hence, the ATM forum has defined the following characteristics:

- *Real-time Service*: Constant bit rate (CBR), Real-time variable bit rate (rt-VBR)
- *Non real-time service*: Non real-time variable bit rate (nrt-VBR), Available bit rate (ABR), Unspecified bit rate (UBR)

2.2 Real-Time Service

The most distinction among applications concerns the amount of delay and the variability of delay, referred to as jitter that the application can tolerate. Real-time application typically involves a flow of information to a user that is intended to reproduce that flow at a source. For example, a user expects a flow of audio or video information to be presented in a continuous, smooth fashion. A lack of continuity or excessive loss results in significant loss of quality. Applications that involve interaction between people have tight constraints on delay. Typically, any delay above a few hundred milliseconds becomes noticeable and annoying. Accordingly, the demands in the ATM network for switching and delivery of real-time are high. The types of real-time and non real-time connections are defined by [13].

2.2.1 Constant Bit Rate (CBR)

The CBR service is perhaps the simplest service to define. It is used by applications that require a fixed data rate that is continuously available during the connection lifetime and a relatively tight upper bound on transfer delay. CBR is commonly used for uncompressed audio and video information. Example of CBR applications include videoconferencing, interactive audio (e.g. telephony), audio/video distribution (e.g. television, distance learning, pay-per-view), audio/video retrieval (e.g. video-on-demand, audio library)

2.2.2 Real-Time Variable Bit Rate (rt-VBR)

The rt-VBR category is intended for time-sensitive applications; that is, those requiring tightly constrained delay and delay variation. The principle difference between applications appropriate for rt-VBR and those appropriate for CBR is that rt-VBR application transmits at a rate that varies with time. Equivalently, an rt-VBR source can be characterized as somewhat bursty. For example, the standard approach to video compression results in a sequence of image frames of varying sizes. Because real-time video requires a uniform frame transmission rate, the actual data rate varies. The rt-VBR service allows the network more flexibility than CBR. The network is able to statistically multiplex a number of connections over the same dedicated capacity and still provide the required service to each connection.

2.3 Non Real Time Services

Non real-time services are intended for applications that have bursty traffic characteristics and do not have tight constraints on delay and delay variation. Accordingly, the network has greater flexibility in handling such traffic flows and can make greater use of statistical multiplexing to increase network efficiency. The non real-time services are further categorized as nrt-VBR (nrt-variable bit rate), UBR (unspecified bit rate) and ABR (available bit rate).

2.3.1 Non Real-Time Variable Bit Rate (nrt-VBR)

For some non real-time applications, it is possible to characterize the expected traffic flow so that the network can provide substantially improved quality of service (QoS) in the areas of loss and delay at cell level. Such applications can use the nrt-VBR service. With this service, the end system specifies a peak cell rate, a sustainable or average cell rate, and a measure of how bursty or clumped the cells. With this information, the network can allocate resources to provide relatively low delay and minimal cell loss. The nrt-VBR service can be used for data transfers that have critical response-time requirements. Examples include airline reservations, banking transactions and process monitoring.

2.3.2 Unspecified Bit Rate (UBR)

At any given time, a certain amount of the capacity of an ATM network is consumed in carrying CBR and the two types of VBR traffic. Additional capacity is available for one or both of the following reasons: a) not all of the total resources have been committed to CBR and VBR traffic, and b) the bursty nature of VBR traffic means that at some times less than the committed capacity is being used. All of this unused capacity could be made available for the UBR service. This service is suitable for applications that can tolerate variable delays and some cell loses, which is typically true of TCP-based traffic with UBR, cells are forwarded on a first-in-first out (FIFO) basis using the capacity not consumed by other services; both delays and variable losses are possible. No initial commitment is made to a UBR source and no feedback concerning congestion is provided; this is referred to as a best-effort service. Example of UBR applications include, text/data/image transfer, messaging, distribution, retrieval, remote terminal (e.g. telecommuting)

2.3.3 Available Bit Rate (ABR)

Bursty applications that use a reliable end-to-end protocol such as TCP can detect congestion in a network by means of increased round-trip delays and packet discarding. However TCP has no mechanism for causing the resources within the network to be shared fairly among many TCP connections. Further, TCP does not minimize congestion as efficiently as is possible using explicit information from congested nodes within the network.

To improve the service provided to bursty sources that would otherwise use UBR, the ABR service has been defined. An application using ABR specifies a peak cell rate (PCR) that it will use and a minimum cell rate (MCR) that it requires. The network allocates resources so that all ABR applications receive at least their MCR capacity. Any unused capacity is then shared in a fair and controlled fashion among all ABR sources. The ABR mechanism uses explicit feedback to sources to assure that capacity is fairly allocated. Ant capacity not used by ABR sources remains available for UBR traffic. An example of an application using ABR is LAN interconnection. In this case, the end systems attached to the ATM networks are routers. Figure 2.1 suggests on how a network allocates resources during a steady-state period of time (no additions or deletions of virtual networks.

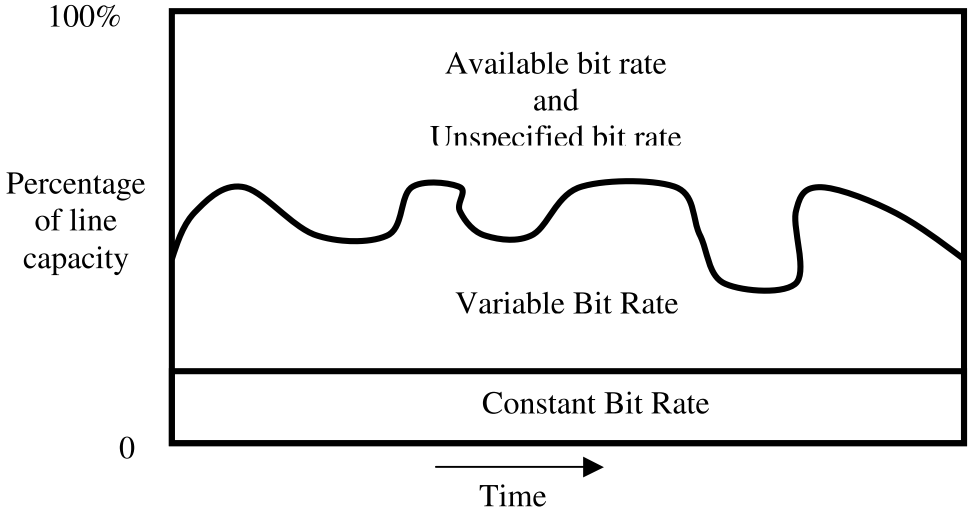


Figure 2.1: ATM Bit Rate Services

2.4 WATM Traffic Control

The purpose of traffic control is to optimize network resources, provide QoS at call (i.e. size of bandwidth for a non real time connection) and cell levels (i.e. cell delay and cell loss) for existing connections and avoid or limit congestion. To meet these objectives, ITU-T and the ATM forum have defined a collection of traffic control functions that operates across a spectrum of timing intervals. Table 2.1 lists these functions with respect to the response times within which they operate.

Table 2.1: WATM Traffic Control Functions

Response Time	Traffic Control Functions	Congestion Control Functions
Long Term	• **Resource management using virtual paths**	
Connection Duration	• **Connection admission control (CAC)**	
Round-trip propagation Time	• **Fast resource management**	• **Explicit forward congestion indication (EFCI)** • **ABR flow control**
Cell Insertion Time	• **Usage parameter control (UPC)** • **Priority control** • **Traffic shaping**	• **Selective cell discard**

- Cell insertion time: Functions at this level react immediately to cells as they are transmitted.
- Round trip propagation time: At this level, the network responds within the lifetime of a cell in the network and may provide feedback indications to the source.
- Connection duration: At this level, the network determines whether a new connection at a given QoS (at cell level) can be accommodated and what performance levels will be agreed to
- Long Term: These are controls that affect more than one ATM connection and are established for long-term use.

The essence of the traffic control strategy is based on determining whether a given new ATM connection can be accommodated and agreeing with the subscriber on the performance parameters that will be supported. In effect, the subscriber and the network enter into a traffic contract where the network agrees to support traffic at a certain level of performance on this connection, and the subscriber agrees not to exceed traffic parameters limits. Traffic control functions are concerned with establishing these traffic parameters and enforcing them. Thus, they are concerned with congestion avoidance. If traffic control fails in certain instances, then congestion may occur. At this point, congestion control functions are invoked to respond to and recover from the congestion.

2.4.1 Connection Admission Control (CAC)

Connection Admission Control performed by the network to determine whether a newly arrival call should be admitted to the network or not. One of the main considerations here is that the QoS at cell level for the existing connections are no under circumstances jeopardized. A connection request for a given call is accepted only when sufficient resources are available to carry the new connection through the whole network as it's requested QoS at cell level while maintaining the agreed QoS at cell level of already established connections in the network.

During connection establishment procedure (i.e. at call set-up phase) the following information, embedded in a traffic contract specification, has to be negotiated and agreed between "user" and " network" to enable CAC to make reliable connection acceptance/ denial decisions [14]:

- Specific limits on the traffic volume the network are expected to carry in terms of well-chosen traffic descriptors.
- A requested QoS class at cell level expressed in terms of cell transfer delay, delay jitter and cell loss ratio
- A tolerance to accommodate cell delay variation introduced by for example the Terminal Equipment (TE) or the Customer Premises Equipment (CPE), which may alter the negotiated limits of the expected traffic volume.

This information may be negotiated during the lifetime of the connection at the request of the user. The network itself may limit the frequency of these **renegotiations.** CAC schemes are currently not standardized and are at the discretion of the network operators. For WATM an implementation of the CAC function might incorporate Radio Signal Strength (RSS) or Bit Error Rate (BER) measurements and this is illustrates in figure 2.2.

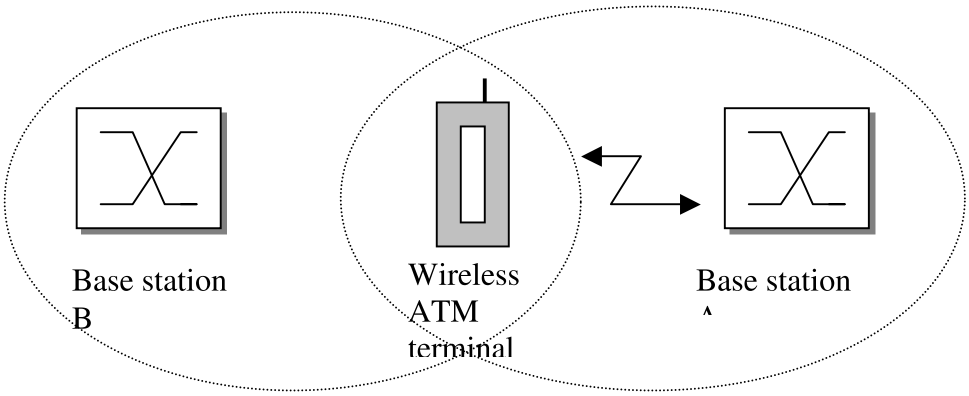


Figure 2.2: RSS measurement location

2.4.2 Usage Parameter Control

Once a connection has been accepted by the connection admission control function, the usage parameter control (UPC) function of the network monitors the connection to determine whether the traffic conforms to the traffic contracts. This associates at cell level. The main purpose of UPC is to protect network resources from an overload on one connection that would adversely affect the QoS (at cell level in terms of cell rate and cell burst size) on the other connections by detecting violations of assigned parameters and taking appropriate actions.

In the fixed ATM network, the UPC functions are located at the entry point of a connection into a network, which is in the most cases the port of an ATM switch
(See figure 2.3).

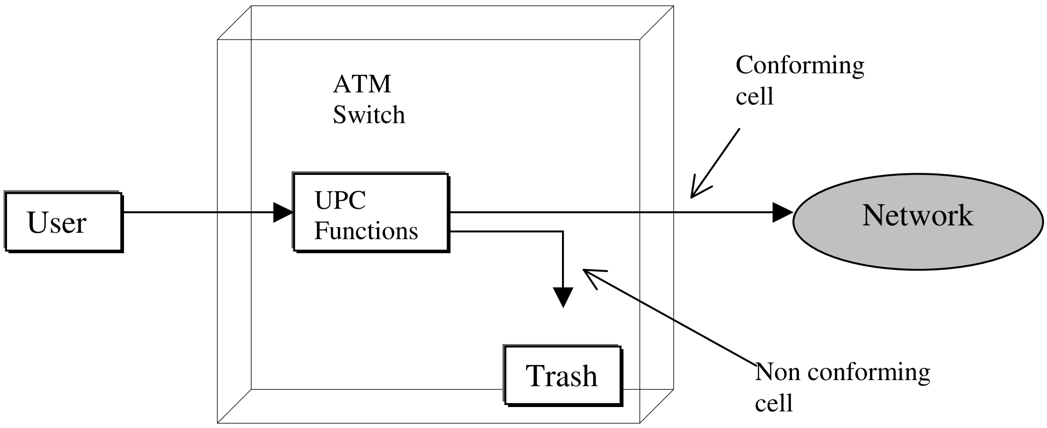


Figure 2.3: UPC location on fixed network

In wireless ATM the UPC functions have to be integrated with the radio resource management [15]. User has to be prevented from using more radio bandwidth than specified in the traffic contract and the WATM medium access control (MAC) layer provides this functionality. For a centralized radio resource management in the base station, the UPC functions can be integrated with MAC layer scheduling of the bandwidth requested by the wireless ATM terminals. If the users request to transmit more cells than allowed, the exceeding requests will not be granted by the WATM base station controller. Another solution is to reject the request in the wireless terminal. In this case the location of the UPC functions has to be change to the wireless terminal.

2.5 Call Level Congestion Control

Congestion at call level occurs when all of the channels in the network are busy to accept a connection request. When a real time application call enters a congested area, the call will be blocked resulting in the termination of the call. The QoS metrics in this thesis are define to be the new incoming call blocking probability if it is the case of a new attempt call to the network. However, if it is the case of an on-going call, it is define as the handoff dropping probability. The non real-time application is for data oriented connection and are less sensitive to delay. When a non real-time application call enters a congested area the call will queue. This situation is referred as an overload state where the total wireless capacity available is smaller than the total instantaneous capacity required by mobile terminal in that area. At this time packets will suffer loss or delay but when the overload is terminated the packet will flow normally. The QoS metrics in this case will be the overload probability and blocking of the new call when all channels are occupied. Table 2.2 summarizes this application with its QoS metric based on [16].

Table 2.2: Classes of service and their QoS metrics

Classes	Real Time	Non Real Time
Characteristics	Highly delay sensitive	Delay less sensitive
Application	Voice	Transmission Control Protocol (TCP)
	Video	Remote log-in
QoS Metric at call level	Blocking Probability	Blocking Probability
	Dropping Probability	
QoS Metric at cell level	Cell Loss	Overload Probability
	Cell Delay	Cell Delay

Regarding wireless resources, the purpose of CAC is to limit the number of in progress wireless calls such that, once a wireless call is admitted, the probability of its encountering a congested radio state is acceptably low as to provide the required QoS. Blocking new wireless call setup requests when the number of existing call has reached this limit does this.

2.6 The Traffic Contract and Optimization of Bandwidth Allocation

When an ATM end user request a connection by signaling or subscription, along with this connection request, comes source traffic descriptor and QoS parameters at cell level. This will characterize quantity and quality of the traffic to be transported over the connection and police the traffic as it enters the network. Hence, this agreement between the ATM end user and the network is called the traffic contract [17]. The ATM network however, will support the connection and associated QoS at cell as well as call level as long as the ATM end user stays within the agreed upon parameters. For example, if a non real time call needs a 64kb/s connection, it will request for 64kb/s bandwidth from the network. If the amount of bandwidths in the network is sufficient enough to provide 64kb/s for the connection, the network will accept the call. Otherwise, the network will negotiate for a lower amount of bandwidth, for example 40kb/s for VBR connection. However, with less bandwidth provided, it will jeopardize the QoS in terms of cell lost, hence, call lost. For an example, the probability of cell lost for VBR connection is about 10^{-6} but with less bandwidth given, the probability of cell lost will be about 10^{-4}.

Amount of bandwidth allocated for any call is performed at each transmission call set-up in all the multiplexer located on the selected virtual path. The algorithm in charge of admitting or rejecting the user requests to establish a transmission path through ATM network is called the Call Admission Control (CAC) [18]. This allocation is carried out after each transaction of call admittance or rejection, which include bandwidth renegotiation. The purpose of admission control based on allocating bandwidth is to establish a fair blocking against any new coming call and to ensure that the resources or bandwidth available to each admitted connection remains sufficient to meet the performance objectives. Indeed, any new connection can only be admitted to the network if sufficient resources are available. The bandwidth to be allocated to a connection in each multiplexer will be expressed in [bits per seconds, bps] and defined as a fraction of the transmission bit rate which empties the multiplexing queue with a exponential service time.

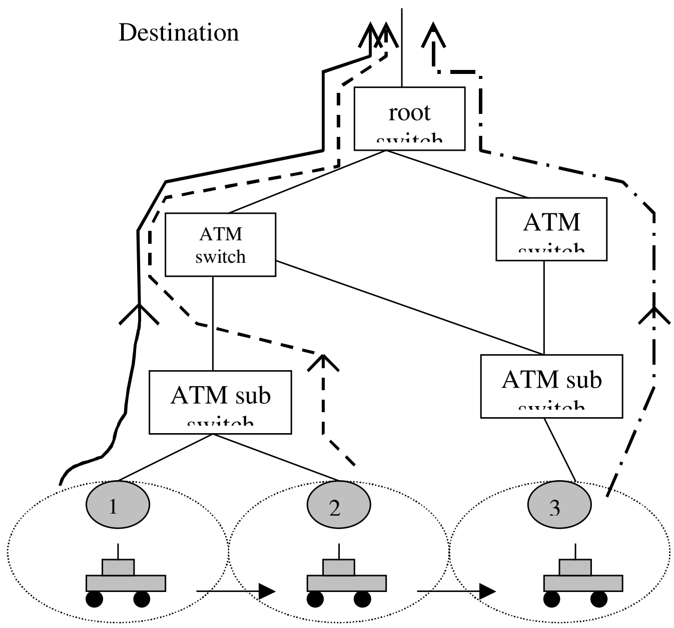
2.7 Handoff Mechanism in WATM

Handoff is a procedure by which a mobile user's radio link (or connection) is transferred from one base station to another through the network without any interruption of the user's connection. It requires a new access-point to the fixed network and the fixed connections are rerouted to transmit calls to the destination. A network needs to decide on how it will reestablish the route of the new incoming calls and handoff calls starting from their given base station to the destination which will either be in the fixed network or in the wireless network.

2.7.1 Connection Rerouting Methods for Handoff

Different approaches had been taken to handle the handoff situation, which have completely different characteristics, performance and impact on the ATM standards [19,20,21,22,23]. They are the full establishment approach, the connection extension case, incremental reestablishment technique and multicast establishment.

i) The full establishment approach requires the setup of a completely new connection between the source and the destination terminals [19]. This is one of the earliest proposals that have minor impact on the fixed network but it takes longer time to establish the connection. Thus it will cause termination of the handoff calls. Figure 2.4 illustrates the full establishment extension.



Original VC - - - **path extension1** — · **path extension 2**

Figure 2.4: MT roaming through three cell sites in the full establishment case.

ii) In the connection extension case, each time a MT handoffs to a new base station the connection from the new base station will reconnect with the old connection at the first base station to the destination terminal. This technique is used in [20,21] to simplify the rerouting connection in the fixed and wireless network is shown in figure 2.5. In [20] it is called Homing Algorithm, which includes simple control and preservation of FIFO cell sequence within VCs. The locations of the local exchange or ATM sub switch need to be updated frequently and this prolongs the time of connection to the destination and remarkably a waste of resource.

iii) Incremental reestablishment technique connects the first connection to the original connection and the same with the second connection. In this way the original path will be reused extensively throughout the connection's lifetime. This technique requires only the establishment on a new partial path, which connects to a portion of the original connection path, allowing VCs to be partly reused. This technique is fast and efficient [19] but will overlap the use of the same path for the same connection. The network example is shown in figure 2.6

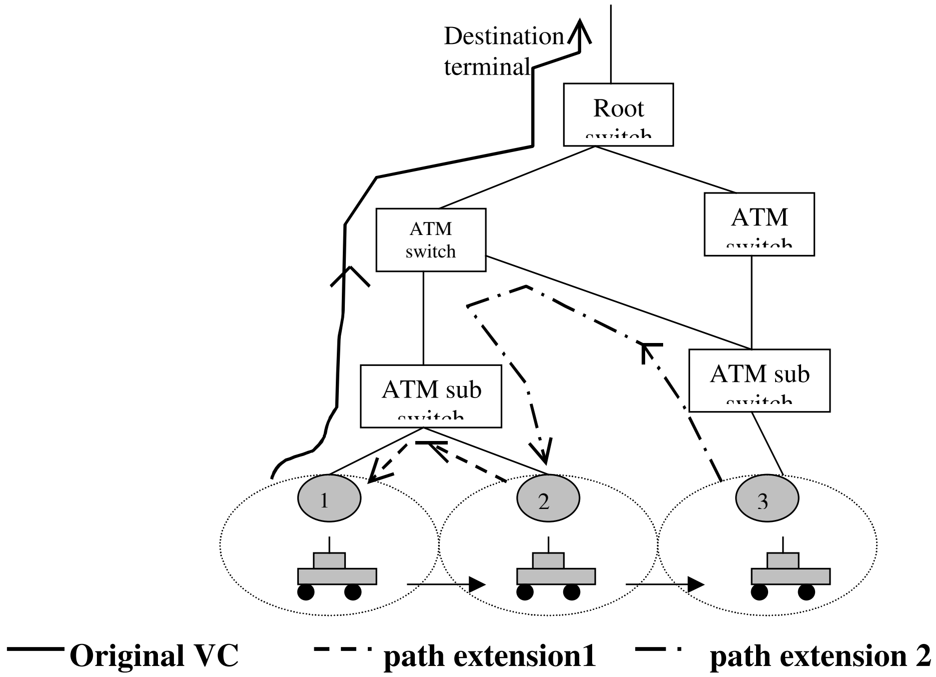


Figure 2.5: VC extent from one base station to another in the connection extension case

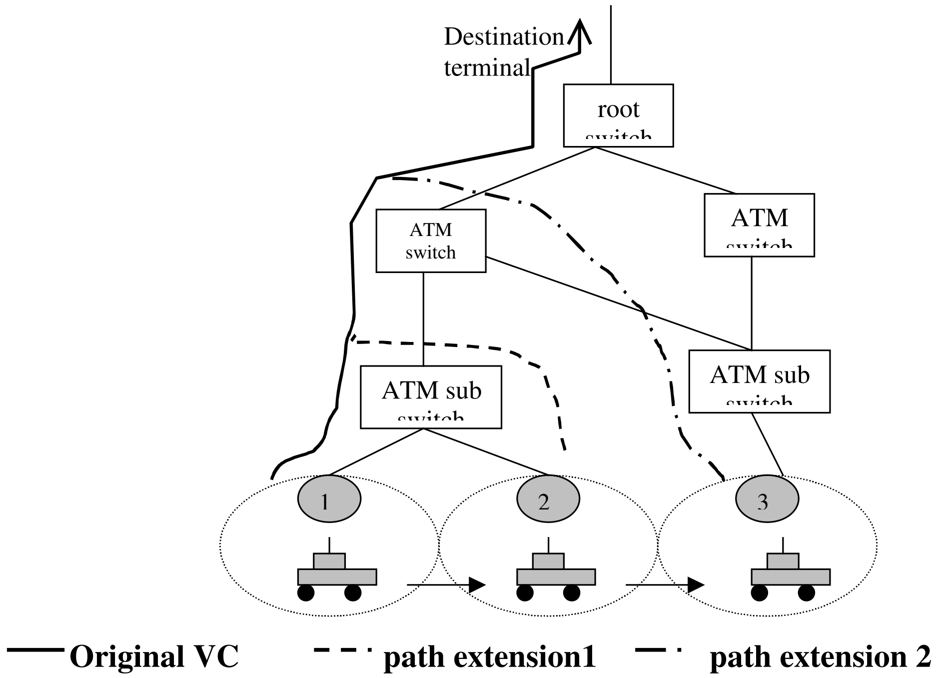


Figure 2.6: MT roaming through three cell sites in increment reestablishment case.

iv) Multicast establishment reallocates resources in the network portion surrounding the cell sites where the mobile user is located. When a new MT is connected to a base station, a set of virtual connections, named virtual connection tree (VCT) is created reaching all base stations in the adjacent cell sites. Thus, the MT can freely roam in the area covered by the tree without invoking the network call processor (NCP) during local and global handoff events. The virtual connection tree can either be static [21] or dynamic [23]. Compare to other methods, this approach is fast and guarantees the QoS contract (at cell level) because the negotiation is executed only once, at connection establishment, allocating resources in the entire area where the MT is expected to roam.

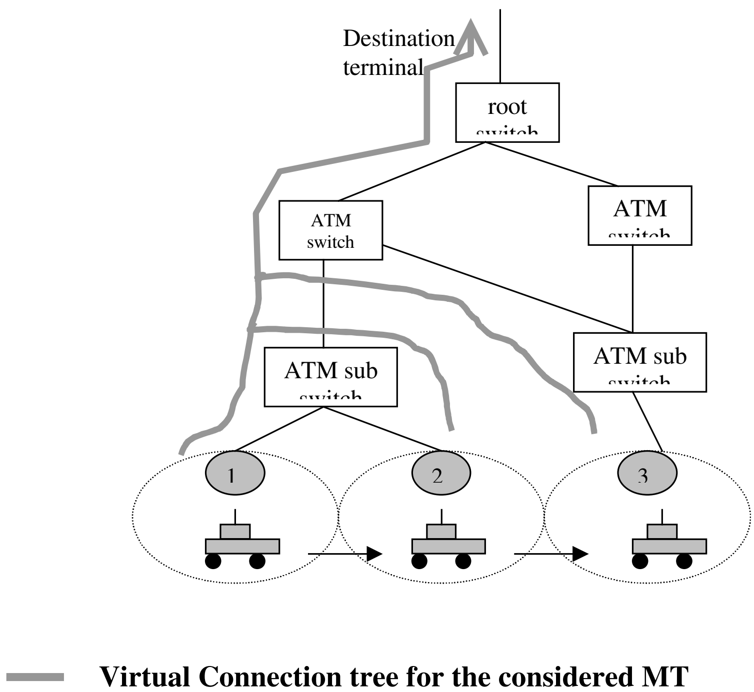


Figure 2.7: MT roaming through three cell sites in multicast establishment

2.8 The Virtual Connection Tree (VCT)

A wireless ATM network system consists of the wireless and the fixed part. A virtual connection tree (VCT) will have the root, the branch and the leaves. The root is the root switch (RS) of the network, the branch is the connection in the fixed network and the leaves are the base stations. To support the frequent amount of handoffs in a small cell site, the network control processor (NCP) have to be invoked every time handoff happens. The NCP is a processor that decides on the route and address number for each route, it also decides on the acceptance or rejection of a call coming into the network. The multicast concept is a method of rerouting handoff call where mentioned in section 2.8.1 iv). It reduces the need to invoke the NCP for handoff by grouping a number of radio cell sites to a cell cluster. When a mobile terminal (MT) enters a connection network, the NCP will allocate a multicast route in the fixed and wireless environment. This is to ensure the mobile terminal is free to handoff to any base stations in the connection network. Multimedia devices, mobile phones, wireless notebooks are some of the examples of wireless terminals.

The fixed portion includes the base station, sub ATM switch (sub SW), ATM switch (SW) and the root switch (RS). The adjacent group cell sites is included in a cell cluster, the root switch and the NCP will control each cell cluster (See figure 2.8). This architecture reduces call set up and routing load on the NCP. As a mobile enters a VCT, it is free to handoff to any base station in the VCT without involving the NCP, thus it will decrease the process of need to set up and tear down the connection during handoff event. The fixed portion of the connection will be maintained as the user moves within the virtual connection network. Thus QoS at cell level as well as call level are guaranteed to the connection in the network throughout the connection lifetime. A virtual connection tree has a control mechanism to monitor the network performance [21].

2.8.1 Implementation of VCT

For new incoming call, when a MT is accepted to a VCT, a collection of virtual circuit numbers (VCN) is assigned to the call. Each of these VCN defines a path between the root of a connection tree and a distinct base station within the neighboring mobile access region. A given MT selects its base station from among those in its connection tree by transmitting ATM cells with the appropriate VCN. Hence, an ATM cells associated with a given tetherless connection will eventually flow through the root of the connection tree and from there, to its destination by means of a fixed wired path if the connection is to a hard wired network port or to a root of another mobile user.

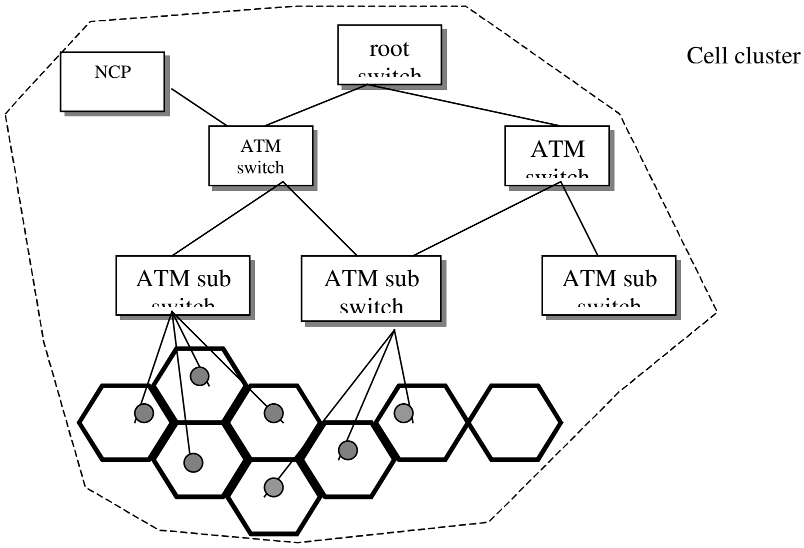


Figure 2.8: The hierarchical network based with cell cluster

For handoff call, when a mobile user already admitted to a virtual connection tree wishes to handoff- to another base station in the same virtual connection tree, it simply begins to transmit ATM cells with the connection number assigned for use between itself and the new base station. Using the pre-established path between the new base stations at the root of the tree, that mobile's ATM cells will flow to the root, across the fixed portion of the network, and to their ultimate destination. In this way, the call processor is not involved in the handover.

In ATM different type of service classes will be identify by the cell header where it carries sufficient information for the network to route and distinguish each cell to the destination port and end users.

2.8.2 Advantage of VCT to WATM

One of the issues mentioned in section 1.3 involves the use of geographically small cell sites to increase overall system capacity, as needed to support a large number of users. As cell site size diminishes, the rate of cell handoff escalates. Hence, a strategy is needed to quickly reroute each handed-off virtual connection to its new cell site and maintain QoS guarantees without burdening the admission controller. The VCT described in section 2.3 meets these criteria [16]. The VCT totally decentralizes handoff operations, with each mobile responsible for managing its own handoff events without requiring intervention of the NCP. The admission controller is involved only in establishing new virtual connections and in handling connections to adjacent trees, but not in handing off among cell sites within the same tree.

At the time a mobile connection is admitted to a tree, a set of virtual connection networks (VCNs) are assigned to that connection. Each VCN defines a path between a specified base station and the root of the tree. A given mobile selects its base station from among those in its connection tree merely by transmitting ATM cells with the appropriate VCN. In this way, an ATM cell associated with a given virtual connection will eventually flow through the root of the connection tree and from there, to it destination by means of fixed wired path if the connection is to a hard wired network port, or to the root of another connection tree, the VCN used to route the cell along the appropriate wired path. Thus, creation of VCT and assignment of a set of VCNs for each connection wastes neither channel capacity nor wire link capacity [6]. Also, the call processor is involved only at connection setup, but is totally uninvolved in processing handoffs within a connection tree, which are handled entirely by handed-off mobile itself in a totally distributed fashion. The call processor becomes involved only in handoffs to different VCT; since the geographical area served by a given connection tree, it may include many base stations and becomes quite large, the frequency of call processor involvement remains low and any potential handoff problems that might have resulted from the use of small cell sites is avoided

2.9 Queuing in WATM network

In ATM wireless network, when very small cell sites are used, the handoff procedure affects the performance of the system. In no-protection call handling schemes, handoff requests are treated the same manner as the new incoming calls, so the probability of handoff dropping equals the probability of blocking of the new incoming calls. However from the MT point of view, the dropping of handoff calls is less desirable than the new call attempt. In order to decrease the probability of blocking of the on going calls, handoff protection schemes had been proposed. Two commonly used handoff prioritization schemes are 1) guard channels that reserved a percentage of channels exclusively for handoff calls [16,23,24] and 2) queuing handoff request with priority [25] or without priority [26]. Handoff protection scheme, in general decrease the probability of handoff dropping at the cost of increasing of new incoming call and reducing the total admitted traffic. The use of guard channels is a means of improving the probability of successful handoffs by simply reserving a fixed or dynamically adjustable number of channels exclusively for handoff requests. The determination of the optimum number of guard channel requires the knowledge of the traffic pattern and estimation of channel occupancy times.

For wireless ATM network that support real-time and non real-time connection, both traffic have to be taken in account to ensure the smooth flow of integrated transmission. The previous queuing system researches concentrated more on real-time connection only. Tekinay and Jabbari [25] used non-preemptive priority schemes where the queuing discipline clearly depends on the power measurements on the radio channels. The handoff area is the area where the ratio of the received power levels from the current and target base station's is between the handoff and the receiver threshold. The MT, whose power level is closest to the receiver threshold, has the higher priority to handoff first. However, if the entire channel in the target cell is occupied after the timeout expired, the call will then be terminated. Instead of using priority queuing, Gaasvik et al used FIFO queuing for handoff call with direct retry [26], which proved to upgrade the probability of handoff dropping. This scheme might be proper for data transmission but for CBR application, delay will cause termination of the on going calls. Since only non real-time connection can tolerate delay, we conclude that queuing for handoff calls is possible ATM wireless network. However for real-time connection queuing scheme can be applied to the new incoming call with considerate amount of queue. Guerin proposed a prioritization scheme for handoff for cellular system [24]. A certain number of channels are reserved for handoff and the new incoming calls are allowed to queue with infinite queue. This thesis will investigate how efficient the prioritizing scheme with guard channels and queue for new in coming calls using integrated WATM wireless with VCT.

2.10 Rerouting Handoff Signaling

When setting up a point-to-point call the ITU-T Recommendations Q.2931 and Q.2961 shall apply. When a VCN value is assigned, the allocation is in both directions even when the connection is unidirectional (has zero backward bandwidth). The VPN value cannot be used for another connection until the call is cleared [27]. Selected extensions to ATM signaling syntax (e.g. Q.2931) together with a new wireless control (meta-signaling) sublayer is needed to support mobile ATM users at the connection control level [28]. Specific mobile network functions requiring signaling/ control support include address registration for mobile users, wireless network QoS parameter specification/renegotiation and handoff.

A basic issue to be addressed is that of designing wireless control and ATM signaling to support relatively transparent handover operations. In general, this will require re-routing of the ATM connection fro the base station to another, while also moving wireless network state information to smoothly resume communication with a minimum of cell loss. An experimentally validated example of possible signaling and meta-signaling exchanges during an ATM handoff is shown in Figure 2.9. In this example, a new signal (handoff request) is used to initiates transfer of group of VC's from one base station to another.

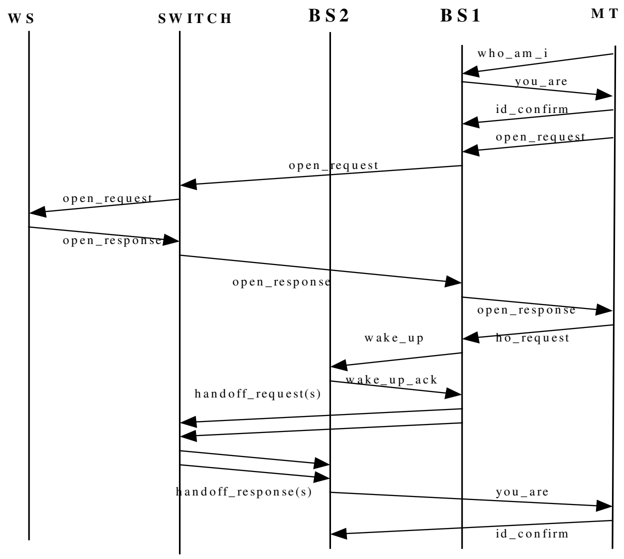


Figure 2.9: Example of handoff control procedure in wireless ATM

Whilst in VCT architecture as a mobile terminal (MT) admitted to a connection tree, the call setup procedure is executed in two steps:

The fixed portion of the virtual connection is established between the root of the tree and the appropriate fixed point of the wired network (i.e.: the fixed user terminal or the root of a destination tree).

Within the connection tree, two sets of connection numbers are assigned to that mobile connection (one in each direction) with one member of each set used to define a path from the root to one of the leaves, and the routing table of the switches within the connection tree are appropriately updated to include the new connection numbers.

Illustrated in figure 2.10, base station X is used by MT to make connection to the fixed VC and VC_X had been assigned for it as the connection number. When the MT wishes to handoff to base station Y (which is in the same connection tree), it simply begins to transmit the packet with the connection numbers assigned for use between itself and the new base station.

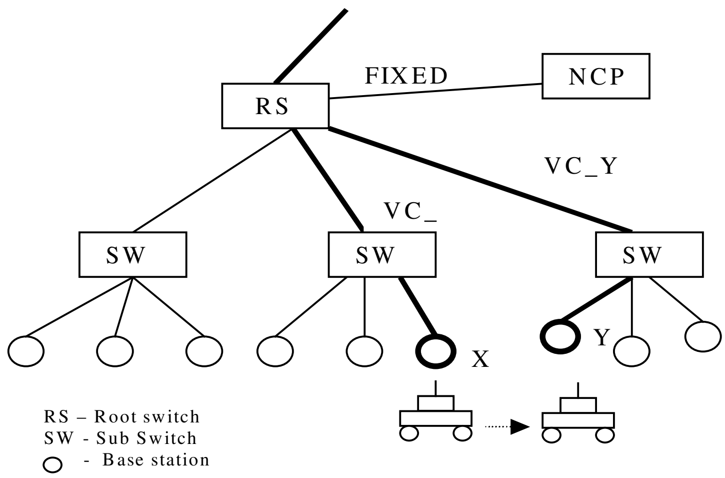


Figure 2.10: Call set up for handoff event.

Using the pre-established path between the new base stations. The mobile packets will flow to the root and across the fixed portion to the destination. In this case the NCP will not be involved. Routing is essential to accommodate the new arrival and handoff calls. The function of NCP is to setup the route and ensure the new established route maintain acceptable QoS at cell level as well as call level to both the wireless connection and pre-existing connection sharing links of the new route. To execute handoff (to another tree), the NCP must first ensure that the new wireless connection does not overload the base station and then it will create a link between the MT and the new base station. Figure 2.11 shows the signaling from base station to the NCP when handoff calls are execute in figure 2.10. It shows that after setting up has been acknowledge the NCP will not be invoke until the call needs to handoff to another VCT.

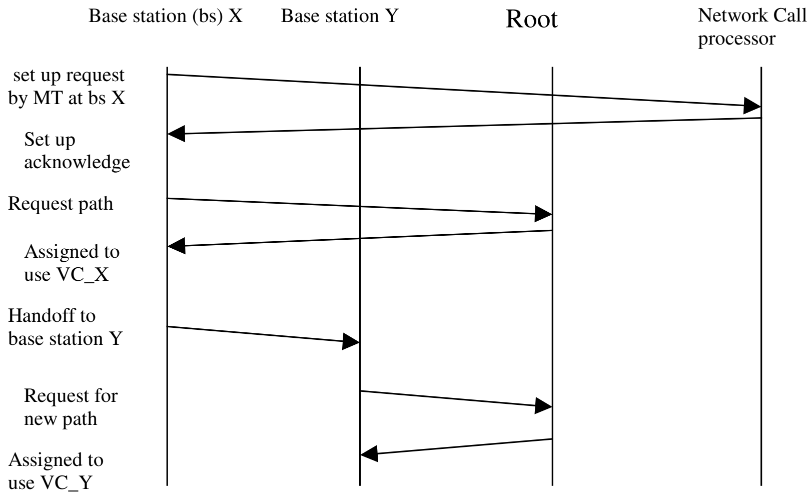


Figure 2.11: An example of a connection tree operation set up

2.11 Summary

In the survey, virtual connection tree (VCT) using multicast concept is chosen as the architecture model wireless ATM. It is chosen because of its advantages stated in section 2.9.1. The advantages are in terms of cost and time it consumes whenever handoff calls are conducted. Hence, VCT architecture model will be use to determine a scheme that can upgrade the WATM performances in terms of probability of call blocking and probability of handoff call dropping.

CHAPTER III

CALL ADMISSION CONTROL

3.1 Introduction

Call admission control (CAC) is the first line of defense for a network in protecting itself from excessive loads. In wireless ATM (WATM), the CAC function is located in the base station. Usually, the CAC function has to be called before a handoff procedure can be completed. In order to guarantee a certain degree of mobility for the wireless ATM terminals, the CAC function needs to reserve some of the available radio bandwidth for handoff calls. A handoff call drop results an interruption of on going call and occurs if the bandwidth required by a connection cannot be supplied in the new cell site location.

In virtual connection tree (VCT), the CAC function is a mechanism that limits the incoming calls into the network in order to protect the already connected calls from being interrupted. An accepted connection request by a mobile terminal (MT) is by the network controller will be given pre-established virtual connections to all the base stations in the tree so that the MT can freely move within the tree. The VCT network also monitors the existing connections in the network in order not to exceed the pre determined threshold limit. Moreover, it uses a scheme that gives priority to the handoff calls where a percentage of channels are reserved for handoff calls whenever the channels in the base station are busy.

3.2 CAC for Real-Time Connection Without Reserve Channels

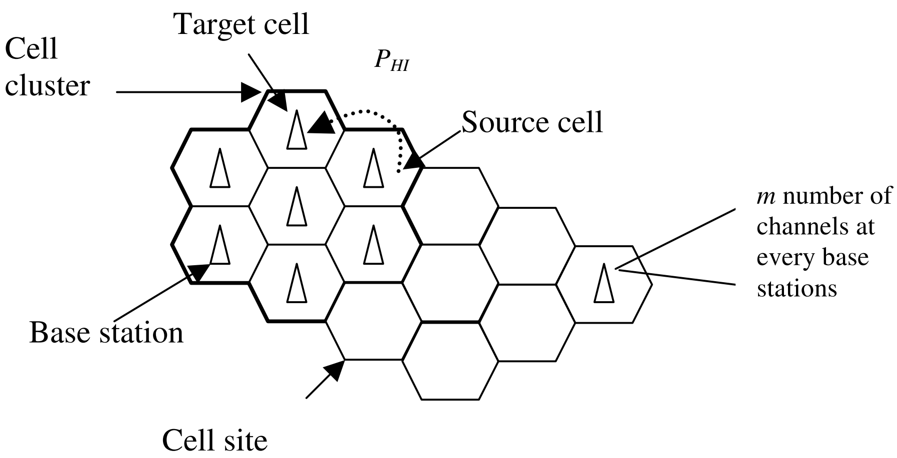


Figure 3.1: The model of ATM wireless network

In VCT, adjacent cell sites are grouped into a cell cluster and *B* is the number of base stations in a cell cluster. This model is illustrated in figure 3.1. It is assume that all cell sites have a similar size and shape, which is considered hexagonal. The probability of a MT from a source cell site to handoff to another (target) cell site within a cluster is:

$$P_{HI} = 1/(B - 1) \qquad (3.1)$$

Here it is assumed that a target cell site is the cell site the MT intended to move to and a source cell site is the cell site the MT is currently located. Hence, the probability that user will handoff to the entire cell site is equal to 1:

$$\sum_{1}^{B-1} P_{HI} = 1 \qquad (3.2)$$

To ensure high efficiency of the limited spectrum for a large number of mobile users, smaller cell sites such as pico and micro cell sites are used. They are in the range of 0 - 500m and less than 1 km radius respectively. As the user moves from a source cell site to other cell sites in the cell cluster, it will visits a number of cell sites before coming back to the source cell site. The probability that the mobile user visit k cell sites is given by:

$$P_v(k) = P_{HI}(1 - P_{HI})^{k-1} \qquad k = 1, 2, \ldots \qquad (3.3)$$

The real-time connection for CAC without reserve channels is said to using the M/ M/ 1 queuing system. The birth-death process of M/M/1 queuing system is illustrated as state-transition-rate diagram of Markov Chain in figure 3.2.

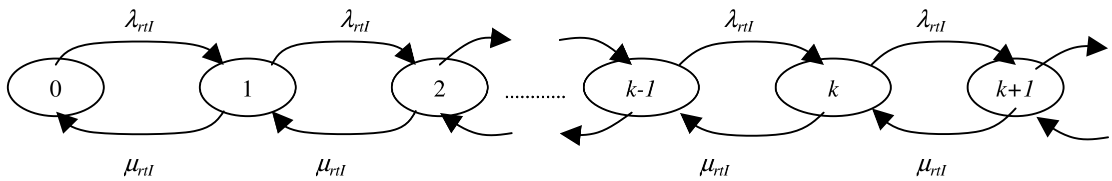


Figure 3.2: State transition rate diagram for M/M/1 queuing system.

From the state-transition-rate diagram in figure 3.2, the equilibrium different equation for state k can be obtain:

$$\lambda_{rtI\,k-1} p_{k-1} + \mu_{rtI_{k+1}} p_{k+1} = (\lambda_{rtI_k} + \mu_{rtI\,k}) p_k \qquad (3.4)$$

where λ_{rtI} is the call arrival rate, μ_{rtI} is the call departure rate and p_k is the probability that the system has k members.

Hence, the probability of finding k mobile terminals in a system given

$$p_k = p_o \prod_{i=0}^{k-1} \frac{\lambda_{rtI}}{\mu_{rtI_{i+1}}} \qquad k = 1,2\ldots \qquad (3.5)$$

$$\text{where } p_o = \frac{1}{1 + \sum_{k+1}^{\infty} \prod_{i=0}^{k-1} \frac{\lambda_{rtI}}{\mu_{rtI_{i+1}}}} \qquad k = 1, 2, \ldots.$$

Applying equation (3.5) to M/M/1 queuing system, we have:

$$p_k = p_o \prod_{i=0}^{k-1} \frac{\lambda_{rtI}}{(i+1)\mu_{rtI}} \qquad k \leq m \qquad (3.6)$$

$$\text{where } p_o = \left[\sum_{k=0}^{m} \left(\frac{\lambda_{rtI}}{\mu_{rtI}} \right)^k \frac{1}{k!} \right]^{-1}$$

Hence the probability of new call blocking and the handoff call dropping for CAC WORC is given by:

$$P_{ABI} = \frac{(\lambda_{rtI}/\mu_{rtI})^m}{m!} \Bigg/ \sum_{k=0}^{m} \frac{(\lambda_{rtI}/\mu_{rtI})^k}{k!} \qquad (3.7)$$

In mechanism without reserve channel (WORC), where all the channels in the network are for new incoming call and handoff calls, probability of blocking for new call and handoff calls will depend on the call arrival rate into the network. This is because both types of calls are treated the same manner, as illustrated in figure 3.3.

One of the QoS parameters required in a network is a low probability of handoff drop in order to provide better service to the network customers. Hence CAC with reserved channels for handoff calls or WRC is presented.

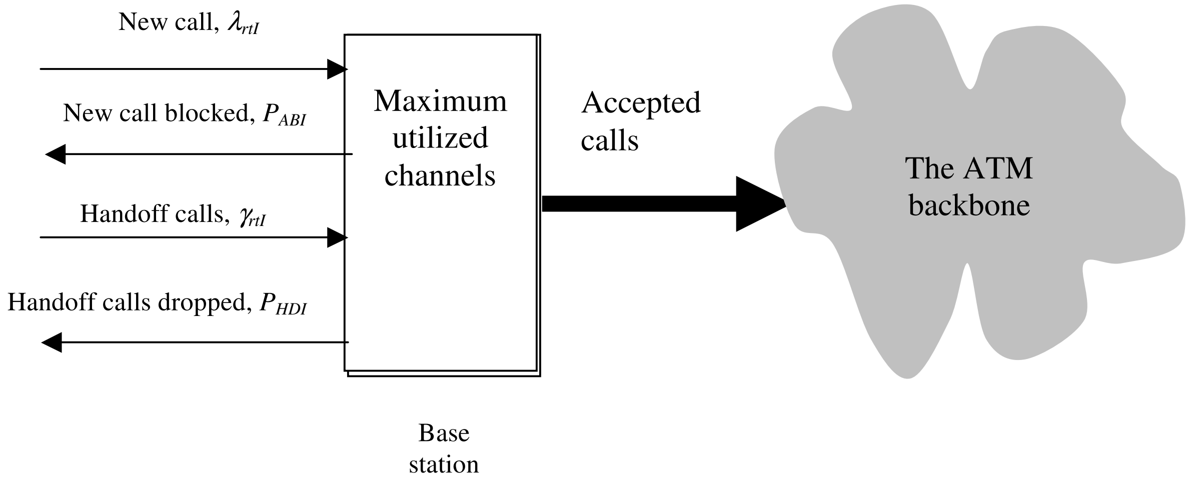


Figure 3.3: CAC of WORC system from the base stations to the backbone network

3.3 CAC for Real-Time Connection With Reserve Channels

To solve the problem of rapid decreasing of handoff call droppings when all channels are busy in CAC WORC, reserve channel for handoff calls are introduced. When a new call enters a network, it can be blocked if the number of connections in the network exceeded the pre-determined threshold, *Nrt* and cannot support any additional connections. Here, *Nrt* is given by:

$$Nrt = mgB \qquad (3.8)$$

where B is the number of base station in a cell cluster, m is the number of call a base station can handle and g is the percentage of reserved channels.

The figure of g can be in between 0 to 100% but for a reasonable value in order not to block new incoming call too much, the percentage of reserved channels must not be less than 20% of the total channel capacity. When a communicating MT moves out of the source cell site, the handoff attempt is generated. If there are not enough channels in the target cell site to accommodate the handoff call, the call will be terminated. This is called handoff dropping of the on going call, which is the interruption of successfully connected call. Since a termination due to failure of handoff attempt is more obtrusive than the new call blocking, in this system the handoff call may be given priority access to the channels in the following way.

Each base station reserves a certain percentage of the total channel allocated to it for handoff. A new incoming call that originates can use any idle channels, provided that they are less than $Nrt\,/\,B$ channels, which are used in the base station at that particular time. If $Nrt\,/\,B$ or more channels are already being used, the new call will be blocked. A handoff call however, will be able to use any idle channels of the *m* channels in the base station. In this way the handoff calls have the priority of accessing channels in any base station in the network.

Applying equation (3.7), the probability of having new call blocked for CAC WRC is given by

$$P_{rt}(f) = \frac{(\lambda_{rtII}/\mu_{rtII})^f}{f!} \Bigg/ \sum_{k=0}^{N_{rtI}} \frac{(\lambda_{rtII}/\mu_{rtII})^k}{k!} \qquad (3.9)$$

where λ_{rtII} is the new call arrival rate and μ_{rtII} is the new call departure rate for WRC.

However, considering the real phenomena, call departures from a cell site consist of calls that desire to be terminated and calls that desire to handoff to another base station. Thus, the actual or effective call departure rate, μ_{ert}, is higher that the natural call departure rate, μ_{rtII}. If observation is made in a cell site, there are changes of call coming in and going out of the cell site as follows (see figure 3.4):

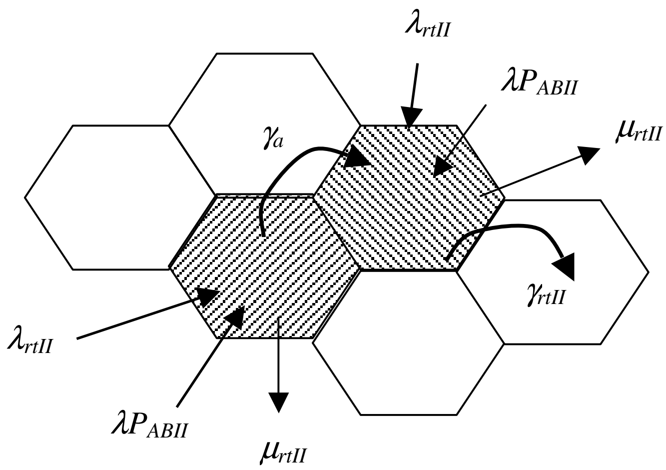


Figure 3.4: The changes of calls going in and coming out of one cell site to another.

λ_{rtII} - The new incoming calls in the source cell site

μ_{rtII} - The completion of calls in the source cell site

γ_a - The arrival of handoff calls in the source cell site

γ_{rtII} - The departure of handoff calls from the source cell to the target cell cite

where

$$\gamma_a = \gamma_{rtII} (1 - P_{HDII}) \quad (3.10)$$

The effective departure rate of wireless calls from a source cell is given by:

$$\mu_{ert} = \mu_{rtII} + \gamma_{rtII} P_{HDII} \quad (3.11)$$

The effective arrival rate is given by:

$$\lambda_{ert} = \lambda_{rtII} (1 - P_{ABII}) \quad (3.12)$$

where P_{HDII} and P_{ABII} are the probability of handoff dropping and the probability of new call block, respectively for CAC WRC

The reserved channels for handoff calls vary with the capacity of population of an area where *Nrt* is the maximum number of connections allowed to a network. In other words *Nrt* is the threshold for real-time calls into a VCT network. Since the network grouped the adjacent base stations into a cell cluster, the number of base stations, *B*, will limit the size of a cell cluster with *Nrt* being the threshold of the allowed number of calls into the VCT. Hence, the probability of the new incoming calls being block by the CAC function in WRC is given by:

$$P_{ABII} = B\frac{(\lambda_{rtII}/\mu_{ert})^{Nrt}}{Nrt!} \Bigg/ \sum_{k=0}^{Nrt}\frac{(\lambda_{rtII}/\mu_{ert})^{k}}{k!} \qquad (3.13)$$

Assumed that the mobile movement patterns of handoff are independent and similar to the new incoming calls to any cell site. The model of aggregate calls arrival to any cell sites in the cell cluster is assumed as Poisson process. When a call departs from a cell site while still communicating, not all of these calls will be successfully handed off to the target cell sites. Assumed that the handoff call going out of the network is statistically identical to the process of handoffs going into a network. Again, by using figure 3.2 and the equilibrium equation can be found. Note that to find the probability of handoff calls that will probably be dropped, actual load that had been accepted to a base station will only be considered. Hence, with *m* number of channels at a base station the probability of handoff dropping is given by:

$$P_{HDII} = \frac{(\lambda_{ert}/\mu_{ert})^{m}}{m!} \Bigg/ \sum_{k=0}^{m}\frac{(\lambda_{ert}/\mu_{ert})^{k}}{k!} \qquad (3.15)$$

However, during heavy traffic, λ_{rtII} tends to infinity, the cell cluster will always have *Nrt* number of calls. Hence, the Erlang load at each base station in the network is Nrt / B, the probability of handoff in heavy traffic:

$$P_{HDII} = \lim_{\lambda \to \infty} P_{HDII} = \frac{(\lambda_{ert,\infty}/\mu_{ert})^m}{m!} \bigg/ \sum_{k=0}^{m} \frac{(\lambda_{ert,\infty}/\mu_{ert})^k}{k!}$$
$$= \frac{(Nrt/B)^m}{m!} \bigg/ \sum_{k=0}^{m} \frac{(Nrt/B)^k}{k!} \qquad (3.16)$$

The block diagram in figure 3.5 illustrated the calls accepted and rejected in a base station using CAC WRC.

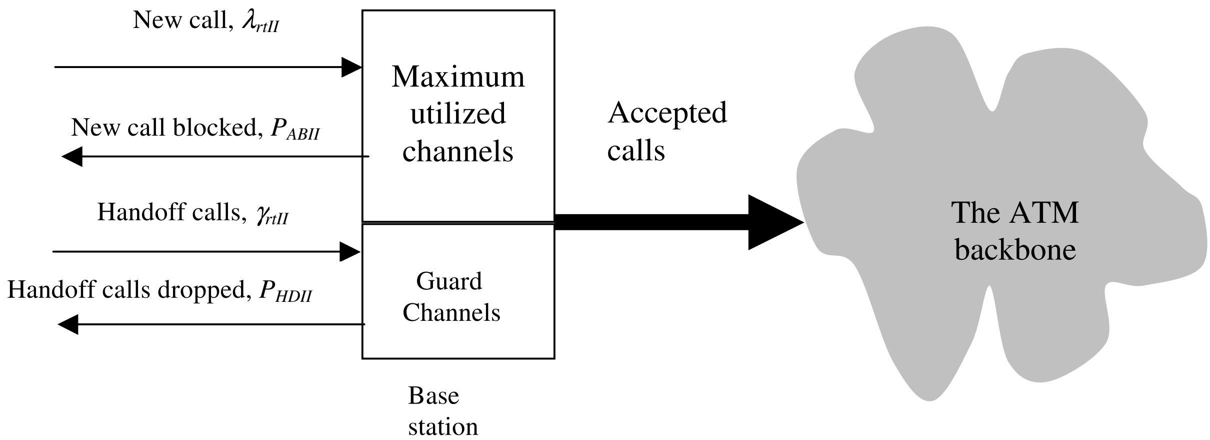


Figure 3.5: CAC of WRC system from the base stations to the backbone network

3.3.1 CAC For Non Real-Time Connection

The ATM network supports multiclass applications. Since it accepts different classes of service, each of them is characterized by a specific transmission quality-of-service. For non-real time connection, data are transmitted despite of real time communication. Hence, they can queue at any base station whenever there are no available channels. Generally a queuing network consists in principle of finite number of nodes or base stations, where the queuing capacities are supposed to be finite or infinite.

In an open queuing system, the arrival process entering the network are assumed to be independent of each other and independent of the queuing service processes. All entering calls are supposed to leave the network eventually. This is opposite to the closed system where a fixed number of calls flow in the system from node to node without either entering or departing from the network. A VCT network that support non real-time connection and do not limit incoming calls will have problems regarding overload in the network. To solve this problem a system using VCT architecture model with control mechanism of incoming calls is introduced. Assume that the system is homogeneous for non real-time connection in this project to be an open queuing system of M / M/ ∞ with the probability of j connections in a system is [32]:

$$p(j) = \left(\lambda_{nrt} / \mu_{nrt} \right)^{j} e^{-(\lambda_{nrt}/\mu_{nrt})} \Big/ j! \qquad (3.17)$$

In VCT network, cell sites are grouped into a cell cluster. In order to limit the number of calls coming into the network a threshold is set to control the maximum number of connections, *Nnrt*, in a network with *B* number of base stations in a cell cluster. Hence, the VCT system is said to be the truncation of the open queuing system (see figure 3.6).

Here, $Z \subset \Im$ having the equilibrium distribution of [33]:

$$p_A(j) = \frac{p(j)}{\sum_{k \in A} p(k)} \qquad j \in Z \qquad (3.18)$$

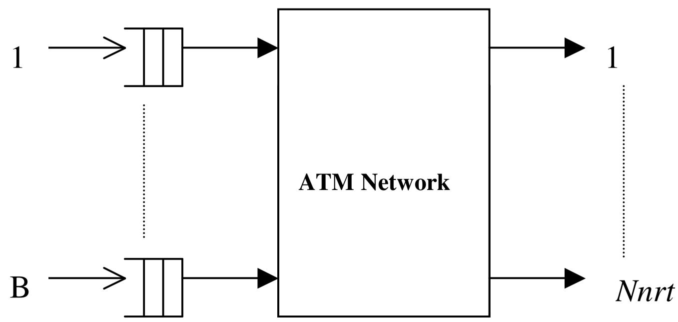


Figure 3.6: The truncated open queuing system where B is the number of base station in a cluster.

$\Im$ is a state space in reversible Markov process. The call arrivals are in Poisson process with arrival rate, λ_{nrt}, per base station and call duration is exponentially distributed with mean $1 / \mu_{nrt}$. In this homogeneous system, *Nnrt* is the maximum number of non real-time connection allowed to a network and *s* is the number of connections to a base station. The process is illustrated as state-transition-rate diagram of Markov Chain in figure 3.7. The equilibrium distribution of the truncated process is a conditional probability distribution that the original process is in state *j* given that it is somewhere in *Z*.

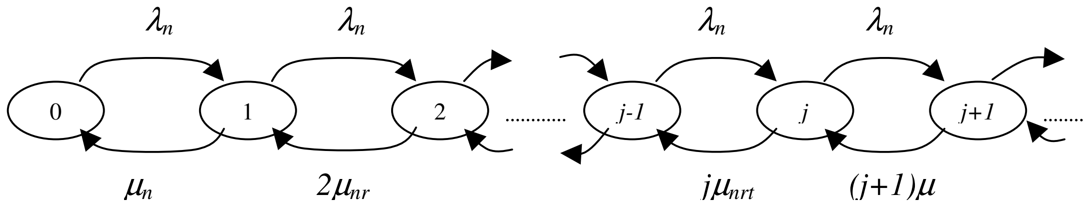


Figure 3.7: State transition diagram of an open system, M /M /∞.

State space Z containing all $\{j_1, j_2, \ldots, j_B\}$ with:

$$\sum_{i=1}^{B} j_i \leq Nnrt \qquad (3.19)$$

where $i = 1,2,..B$.

Assumed that j_i is the number of connection at base station i. Also, assumed that Z_k is the subspace of Z with:

$$\sum_{i=1}^{B} j_i = k \qquad (3.20)$$

where $i = 1, 2, \ldots, B$.

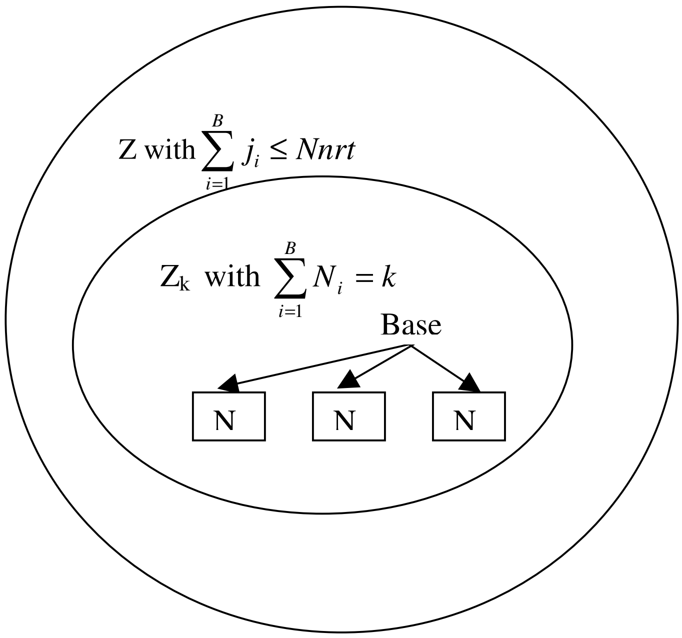


Figure 3.8: Illustration of the truncated open queuing system

The truncated open queuing system is shown in figure 3.8. Expand equation-(3.18) using equation (3.17) with:

$$e^{-\rho} = 1 \Big/ 1 + \sum_{k=1}^{Nnrt} \rho^k \big/ k! \qquad (3.21)$$

where $\rho = \lambda_{nrt} / \mu_{nrt}$,

Hence, the probability of s connection at a base station is:

$$P_{nrt}(s) = \frac{\left[(\lambda_{nrt}/\mu_{nrt})^s \Big/ s! \right] \sum_{k=0}^{Nnrt-s} (\lambda_{nrt}/\mu_{nrt})^k \Big/ k!}{\sum_{k \in A} \frac{\left(\lambda_{nrt} \big/ \mu_{nrt} \right)^k}{k!} \sum_{k=0}^{Nnrt} \left(\lambda_{nrt} \big/ \mu_{nrt} \right)^k \Big/ k!} \qquad s = 0,1,.. \qquad (3.22)$$

Since *(j_1, j_2,…j_B)* for all base stations in the cell cluster is equal to less than *Nnrt*, equation (3.22) can be simplified to:

$$P_{nrt}(s) = \frac{\frac{(\lambda_{nrt}/\mu_{nrt})^s}{s!} \sum_{k=0}^{Nnrt-s} \frac{[(B-1)(\lambda_{nrt}/\mu_{nrt})]^k}{k!}}{\sum_{k=0}^{Nnrt} \frac{(B\,\lambda_{nrt}/\mu_{nrt})^k}{k!}} \qquad s = 0,1,.. \qquad (3.23)$$

However, in heavy traffic condition the equilibrium distribution of number of connection in any base station approaches a Binomial distribution:

$$\lim_{\rho \to \infty} P_s = \lim_{\rho \to \infty} \rho^s \Big/ s! \, \frac{\sum_{k=0}^{Nnrt-s} (B-1)\rho^k \Big/ k!}{\sum_{k=0}^{Nnrt} B\rho^k \Big/ k!} \qquad s = 0,1,.. \qquad (3.24)$$

$$= \binom{Nnrt}{s} \cdot \left(\frac{1}{B-1} \right)^s \cdot \left(1 - \frac{1}{B} \right)^{Nnrt}$$

The QoS performance for non real-time connection is the overload probability at a base station. This is because a call that enters a congested area will queue until there are available channels for it to continue transmitting and the situation will continue until the base station is overloaded.

The overload probability is the summation of the probability that there are more than m connections in a network with limitation of Nnrt and is given by:

$$P_{ov} = \sum_{s=m+1}^{Nnrt} P_{nrt}(s) \qquad \textbf{(3.25)}$$

3.4 The VCT Model

The study will concentrate on the architecture of wireless ATM environment using virtual connection tree (VCT). The VCT architecture and its advantages has already mentioned in section 2.9. This architecture grouped adjacent base stations into a cell cluster. In reality, the cell site shape is affected by the geographical structure of an area but here it is assumed that all cell site shape is hexagonal. Every cell sites has at least one base station and each cell site is given a set of channel frequencies. Adjacent cell site have different set of frequencies, however the same set of frequency can only be use after a certain distance. This is call frequency reuse [11].

In the VCT network, there are two types of calls coming in and going out, the new incoming calls, which are the newly initiated call and the handoff calls, which are the already connected calls. It is assumed that the call arrivals are in Poisson process and call duration is exponentially distributed. Also, it is assumed that the handoff rate of each MT, from any cell site to another, experiences the same rate of arrival of handoff calls. Furthermore, all wireless real time connections are homogenous (e.g. 64kbps-voice connection) and any base station can support up to *m* number of calls.

A predetermined threshold, *Nrt* is set to limit the number of connections within the VCT network. This is the CAC function in VCT where priority of access channels is given to the handoff calls in the network and the new incoming call will be blocked whenever all of the channels in the base station are busy. Each base station in a cell site reserved a certain percentage of the total channels allocated to it for handoff calls. These reserved channels vary with the capacity of population of an area where *Nrt* is the allowed number of connection to a network. The percentage amount of the reserved channels however is fixed for a particular area and will be upgraded or otherwise, yearly depending on request. The speed mobility of a MT will range in between 20 to 40 kmph.

3.5 Simulation of VCT network model

COMNET III a performance analysis tool for computer and communication networks. Base on description of a network, its control algorithms and workload, COMNET III simulates the operation of the network and provides measures of the network performance. No programming is required. Also, using COMNET III, network descriptions are created graphically through a highly intuitive interface that speeds up model formulation and experimentation. COMNET III is integrated into a single windowed package, which performs all functions of model design, model execution and presentation of results. A model is built and executed in straightforward steps. COMNET III can be used to model Wide Area Networks (WANs) and Local Area Networks (LANs).

A simulation of the VCT model for real-time connection was analyzed using discrete simulation COMNET III release 2.1. The system is assumed to have 32 numbers of channels at a base station where each channel is 64kb/s and the number of base station is 3. This analysis was based on the probability that call attempt will be blocked when all of the channels in a base station are busy. The parameters inputs for MT are i) *the interarrival time* per seconds ii) *the call duration* and the Erlang load is equivalent to the mean holding time (call duration) divide by the mean interarrival time. The parameters for the base station and the link are i) *the number of circ*uits and ii) *bandwidth per circuit*. The simulation was briefly done to make a comparison between theoretical and simulation to verify that mathematical is comparable to the simulation results.

The simulation was done at call level using COMNET III circuit switch package and the model playback list, point-to point link, call source, call duration and destination type are as attached in appendix E1-E5. The model playback consists in detail the parameters inputs at each block simulated and not simulated by the COMNET. However in the appendix attached, the simulated block and their input parameters are highlighted. The result of the simulation will then compare with the result of WRC with the same parameters. The simulation model is shown in figure 3.9 for three base stations.

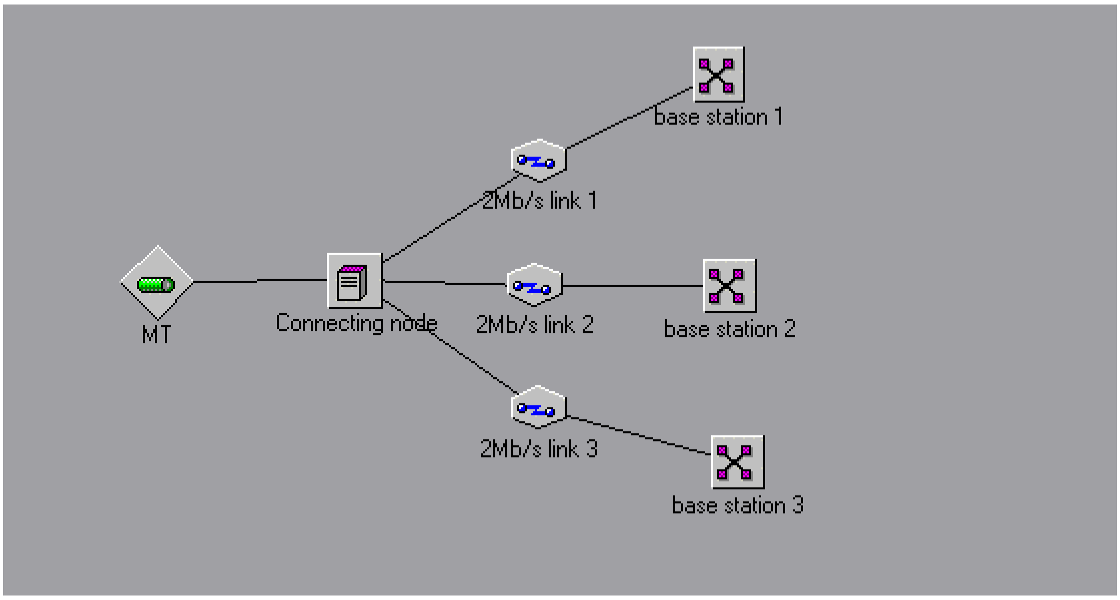


Figure 3.9: The simulation model of a VCT with four base stations in a cell cluster.

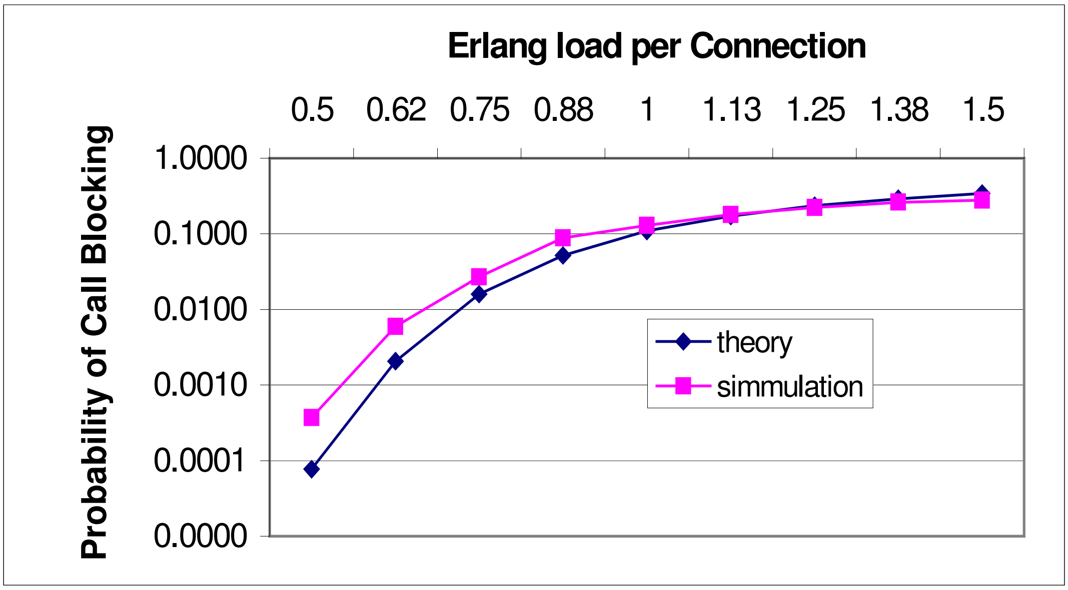


Figure 3.10: Comparison of probability of call block using simulation and theory. (Inputs at MT - interarrival time =0.025 - 0.008 seconds, call duration = 0.007 min) (Inputs at 2Mbit link and base station - bandwidth =62kb/s per circuit, Circuits = 32)(B = 1, μ_{rt1} = 2.5, m = 32 [64kb x 32 = 2Mb/s])

The graph showing the comparison between simulation and theoretical method is plotted in figure 3.10 and the data is shown in appendix A1. The call arrival rate was varied from 40-120 calls per seconds. The call departure rate is fixed to 2.5 calls per seconds. The simulation shows an increment of probability of call block when traffic increases. The result of the probability of call block is comparable to the theory (the difference in the last data is due to random process in discrete event simulation).

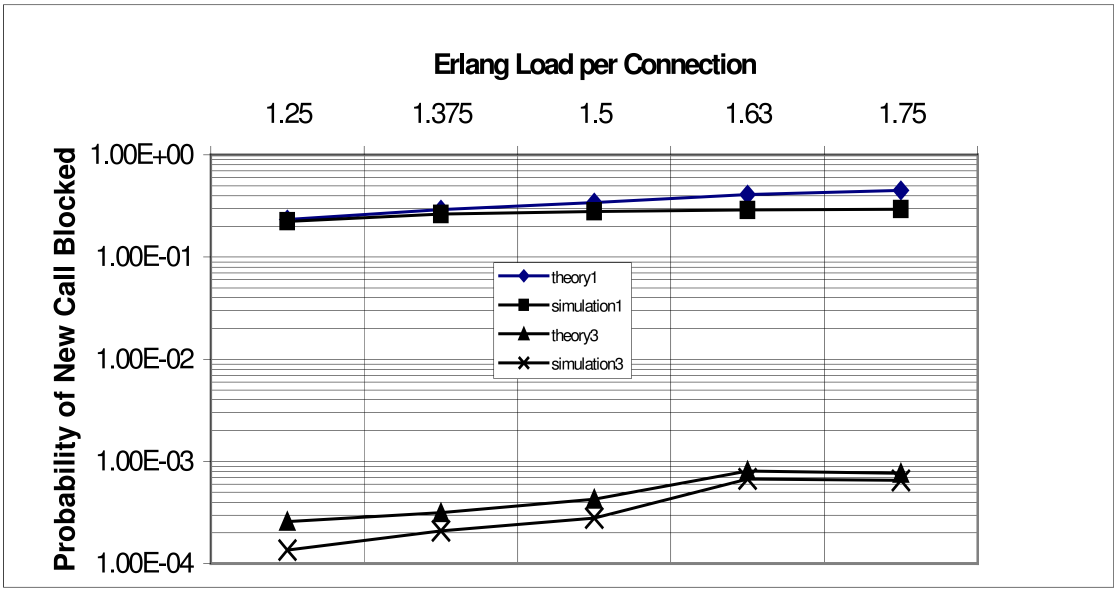


Figure 3.11: The probability of call block against Erlang load for simulation. (Inputs at MT - interarrival time =0.001 - 0.0072 seconds, call duration = 0.007 min) (Inputs at 2Mbit link and base station - bandwidth =62kb/s for every 32 circuits (*B* = 3, μ_{rt} = 2.5, *m* = 32, theory1- for 1 base station, theory3 - for 3 base stations, simulation1- for 1 base station, simulation3 - for 3 base stations)

Figure 3.11 is a graph of probability of call block against Erlang load per cell site and the data of the graph is in appendix A2. The call arrival rate varies from 100 to 140 calls per seconds. The call departure rate is 2.5 calls per seconds and the number of base station are 1 and 3. For call arrival 40 -90 calls per seconds, using 3 base stations the probability of call block is extremely low and in the simulation, they are assumed as no blocking probability. Assumed that every base station can support up to 32 calls. The result shows that the probability of call block increases with Erlang load for both simulation and theory. However, as the number of base stations increases, the probability of calls blocking decreases (the difference between theory and simulation calculated is approximately 17 to 47 percent and these difference are due to the random process in the discrete event simulation). This because there are larger capacities provided by the network. From the graph it shows that even though the base station in the cell luster increases the probability of call block is still comparable with theory. Hence, quality-of-service can be analyzed using mathematical analysis, which is done in the next sections.

3.6 Mathematical Analysis of VCT Network Model

The mathematical analysis used Mathcad as the analysis tool. Mathcad is a combination of a powerful technical computing environment centered on real math notation, flexible and full featured technical word processor. With Mathcad, the tasks of performing computations and documenting them are integrated into one seamless process, resulting in substantial increases in productivity. Unlike other software, Mathcad performs mathematics the same way it is written in paper and pencil. Mathcad's on screen interface is a blank worksheet on which equations, graph data or functions are entered with annotate with text, anywhere on the page. The mathematical expression in Mathcad look the way it would seem in text or a reference. The only difference is that Mathcad's equation and graphs are "live". Change any data, variable or equation and Mathcad recalculates the worksheet instantly. Mathcad handles complex numbers as seamlessly as its handles units, and the standard engineering and math functions are also built into the environment. The analysis for this chapter was done using Mathcad simulation tool where the parameters or inputs to the equation used are the assumptions for that particular situation. The Mathcad programmes are as attached in appendix D1 - D4.

3.6.1 CAC Without Reserved Channels -WORC For Real Time Connection

Assuming that the network has a group of 7 base stations with 20 connections per base station, the new incoming calls arrival rate varies from 20 to 45 calls per seconds. The handoff call rate is 2 calls per seconds with call departure at 1 call per seconds.

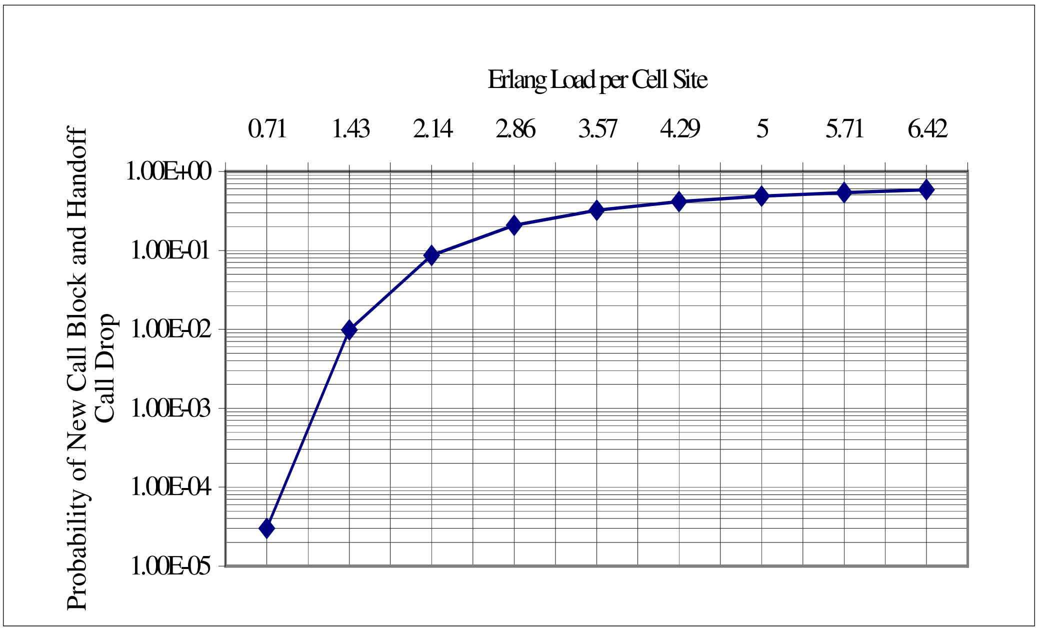


Figure 3.12: The probability of new call block and handoff call drop against Erlang load per cell site for CAC WORC (B = 7, γ_{rt} = 10, μ_{rt} = 2)

Using the assumptions in equation (3.7), figure 3.12 shows the probability of new incoming call and handoff calls increase with Erlang load for WORC. The data of the graph can be obtained in A3. However when channels are busy both types of calls will be block. It means that handoff calls will be block when it enters congested are in WORC. This is because the total channels are for handoff calls and new incoming calls.

3.6.2 CAC With Reserved Channel - WRC For Real Time Connection

A comparison is done between WRC and WORC. Assuming that the network has a group of 7 base stations with 20 connections per base station and q equals to 0.85. The new incoming calls arrival rate varies from 20 to 45 calls per seconds.

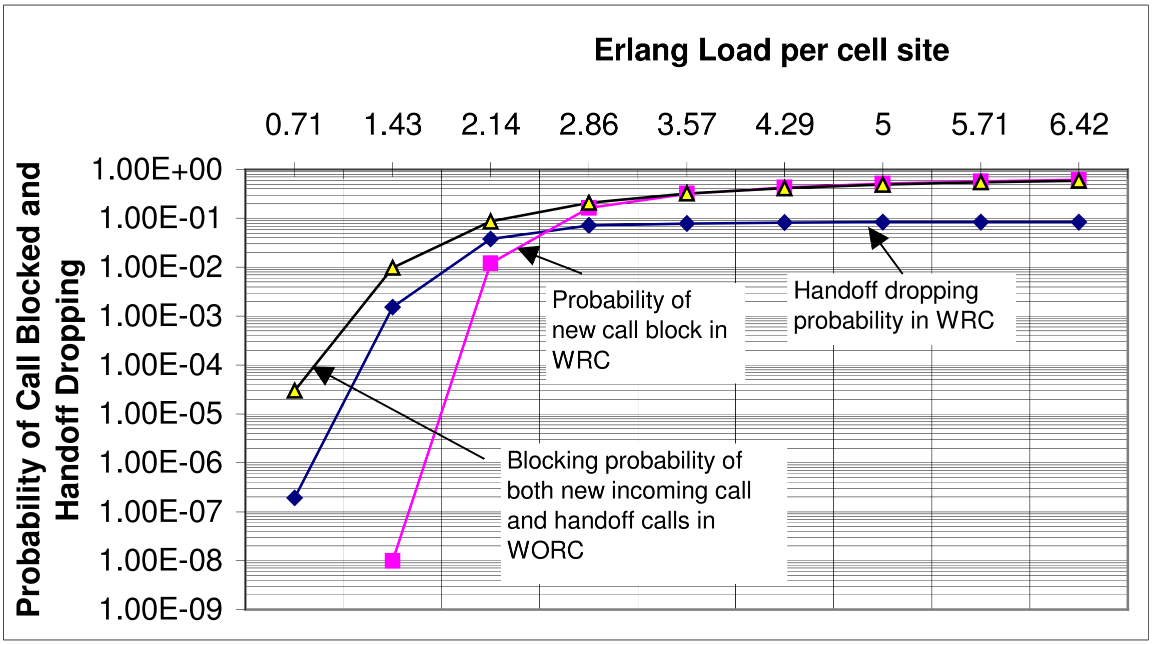


Figure 3.13: Probability of call blocked and handoff dropping for WRC and WORC against Erlang load per cell sites

(B = 7, γ_{rt} = 10, μ_{rt} = 2, g = 0.85)

Using the above assumptions in equation (3.7), (3.13) and (3.15), figure 3.13 shows the performance of WRC and WORC where the data can be obtain from appendix A4. At normal load, which is less than 3.57 Erlang load per cell sites, both mechanisms gives acceptable rate probability of new call blocking and handoff drop but as the load increases higher than 3.57 Erlang load per cell site, WORC cannot seem to maintain the handoff calls. This is shown when the probability to lose call and drop of handoff calls in WORC is as same as the probability of new call block in WRC. This is because in WORC, when all of the channels in a base station are occupied, the handoff calls will be blocked as well as the new arrival calls. However, in WRC priority are given to handoff calls during heavy load and this is shown in figure 3.13 where the probability of handoff call drop are maintain at a constant value at heavy load.

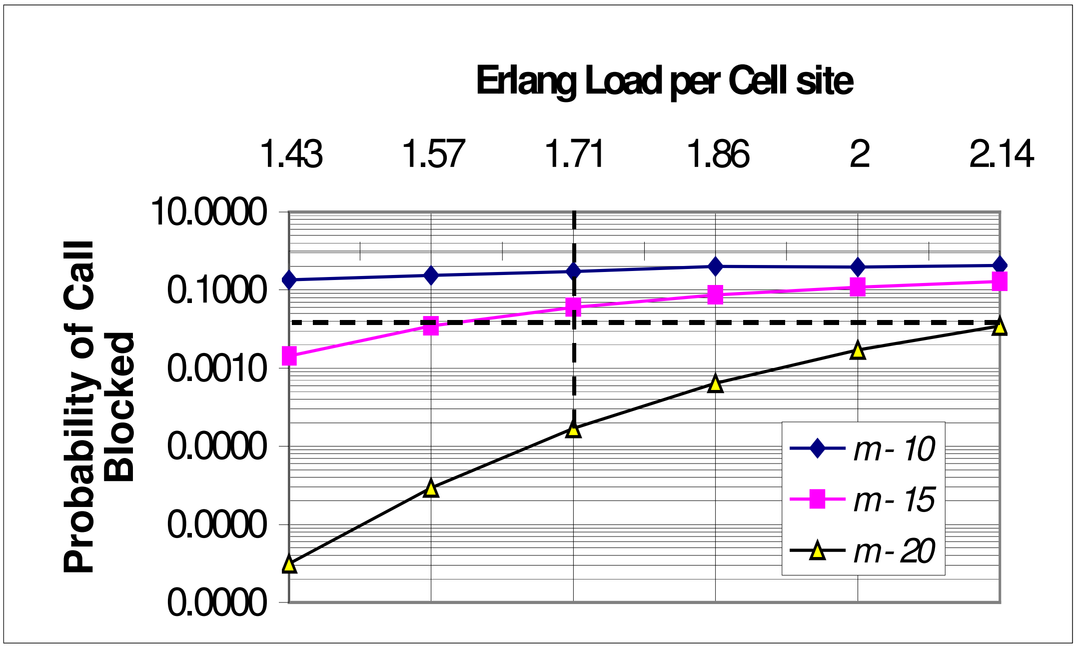


Figure 3.14: The probability of call block against the Erlang load per cell site for different size of capacity a base station can handle.
(B = 7, γ_{rtII} = 10, μ_{rtII} = 5, g = 0.85)

The performance of real-time connections was analyzed in WRC system. Assuming that call arrival rates vary from 50 to 75 calls per seconds. The handoff rate is fixed at 10 calls per seconds. The call departure rate is fixed at 5 calls per second. The network has 7 base stations and q equals to 0.85. The probability of handoff block is assumed 0.01. Using the assumptions in equation (3.13), figure 3.14 shows a graph of probability of call block against Erlang load per cell site at base stations that can supported up to 10, 15 and 20 connections (m = 10, 15, 20). The data of this graph is in appendix A5. A reduction of new call blocking probability was obtained when the number of connections supported by a base station, m, is increased. For instance at 1.71 Erlang load per cell site, the new call blocking probability for a base station with m = 10 connections is 0.1 but with m = 20 connections per base station the new call blocking probability is 2.85e-5. This is because when a base station can support more number of connections, more new incoming call will be accepted by the CAC.

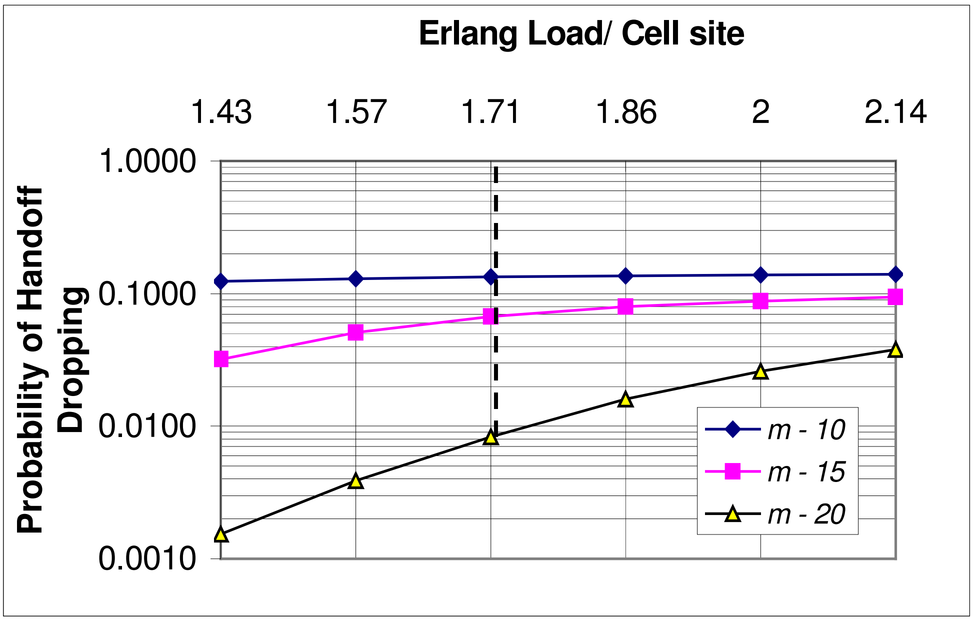


Figure 3.15: The probability of handoff dropping against the Erlang load per cell site.
(B = 7, μ_{rtII} = 5, γ_{rtII} = 10, g = 0.85)

Figure 3.15 shows a graph of probability of handoff dropping against Erlang load per cell site at a base station that can support, m, up to 10, 15 and 20 connections. The data of this graph can be obtained in appendix A6. Using the assumptions in equation (3.15), in figure 3.15 shows that a reduction of handoff dropping probability was also obtain when more number of calls can be support by a base station. This is because when a base station can support more number of connections, more reserved channels will be provided to handoff calls.

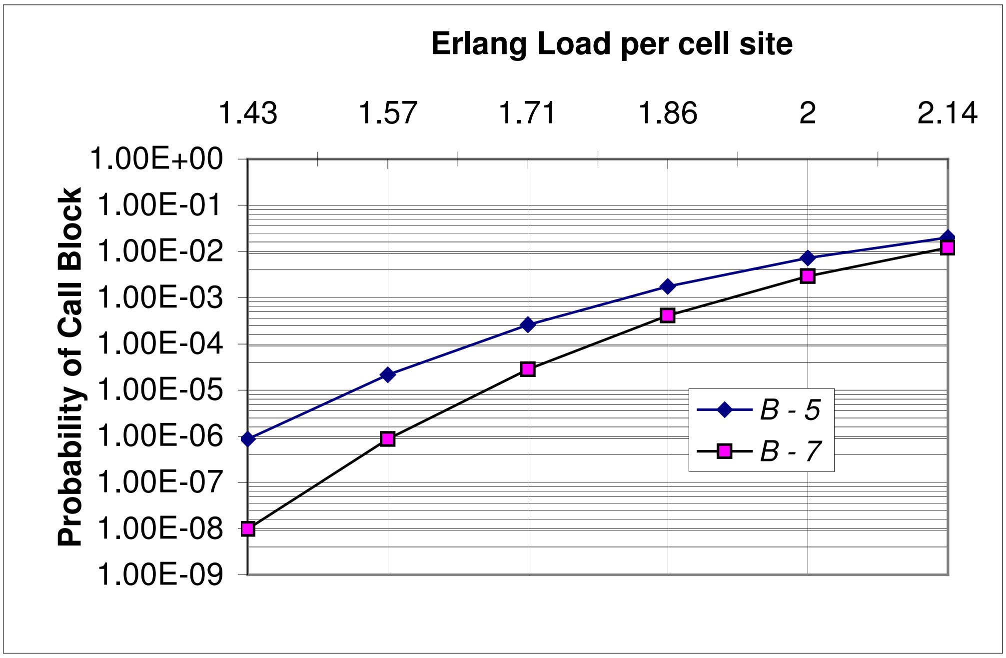


Figure 3.16: The probability of call block against the Erlang load per cell site for different size of cell cluster.
(μ_{rtII} = 5 , γ_{rtII} = 10 , m = 20 , g = 0.85)

An analysis is also done to observe the affect of cell cluster size to the QoS of a VCT network. Assuming that call arrival rates vary from 50 to 75 calls per seconds. The handoff rate is fixed at 10 calls per seconds. The call departure rate is fixed at 5 calls per second. The network has 7 base stations and m equals to 20 with q equals to 0.85. In figure 3.16 a graph of probability of new call block against Erlang load per cell site is plotted and the data is as attached in appendix A7. Using the above assumptions in equation (3.9), it is found that larger cell site reduces the probability that a new call will be block. This is because when more adjacent cell sites are grouped, there will be more capacity in a network, hence, this allows for more connection to be serve by the network.

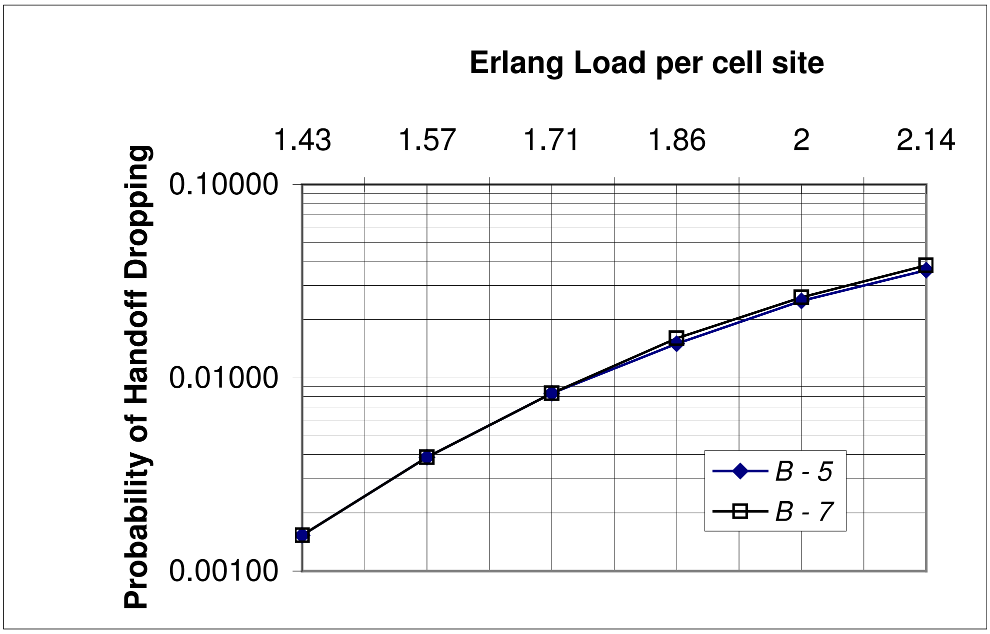


Figure 3.17: Probability of handoff dropping against the Erlang load per cell site
(μ_{rtII} = 5 , γ_{rtII} = 10 , m = 20 , q = 0.85)

In figure 3.17 a graph of probability of handoff call block against Erlang load per cell site is plotted using the assumptions used in the previous graph, in equation (3.15). The data of this graph is in appendix A8. It is found that the size of cell cluster does not affect the probability of handoff dropping. This is due to the constant value of handoff call rate through both size of cell clusters and the smaller difference of the reserved channels in both cell clusters.

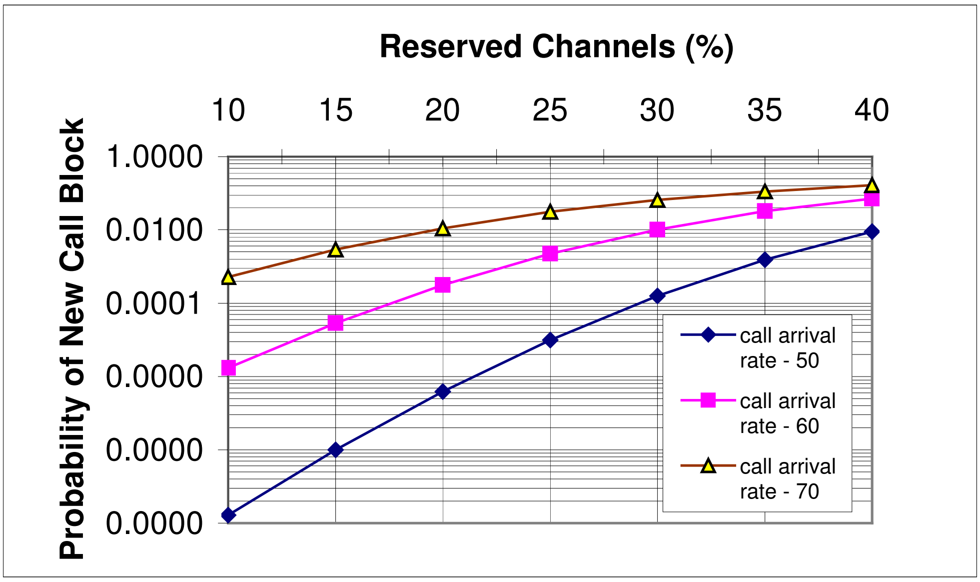


Figure 3.18: Probability of new call block against reserved channels for different value of call arrival.
(B = 7, γ_{rtII} = 10, μ_{rtII} = 5)

An analysis was done to observe the affect of call arrival rate to the QoS performance of the network at certain percentage of reserved channels. In figure 3.18, a graph of probability of new call block against reserved channels was plotted and the data can be found in appendix A9. Assuming that call arrival rates vary from 50, 60 and 70 calls per seconds. The handoff rate is fixed at 10 calls per seconds. The call departure rate is fixed at 5 calls per second. The network has 7 base stations and m equals to 20 with q equals to 0.85.

Using the assumptions in equation (3.13), it is found that when the amount of reserved channels is large, the probability of call block is higher than when the amount of reserved channel is small. This is because when more bandwidth are reserved for handoff calls, the remaining bandwidth capacity are too small to cater the arrival rate per second with only 5 calls depart per second from the network. Figure 1.8 also shows that the more calls arrive at a network, the higher the probability of new call block. For example, when the reserved channel is 20% out of the total base station capacity, with call arrival rate is 50 calls per second, the probability of blocking is 1.65e-10 but when the call arrival rate increases to 70 calls per second, the probability of call block is 0,01. In this case this is because 80% of the capacity have to cater the increment of call arrival and probability that a call will not be accepted is higher when traffic is more.

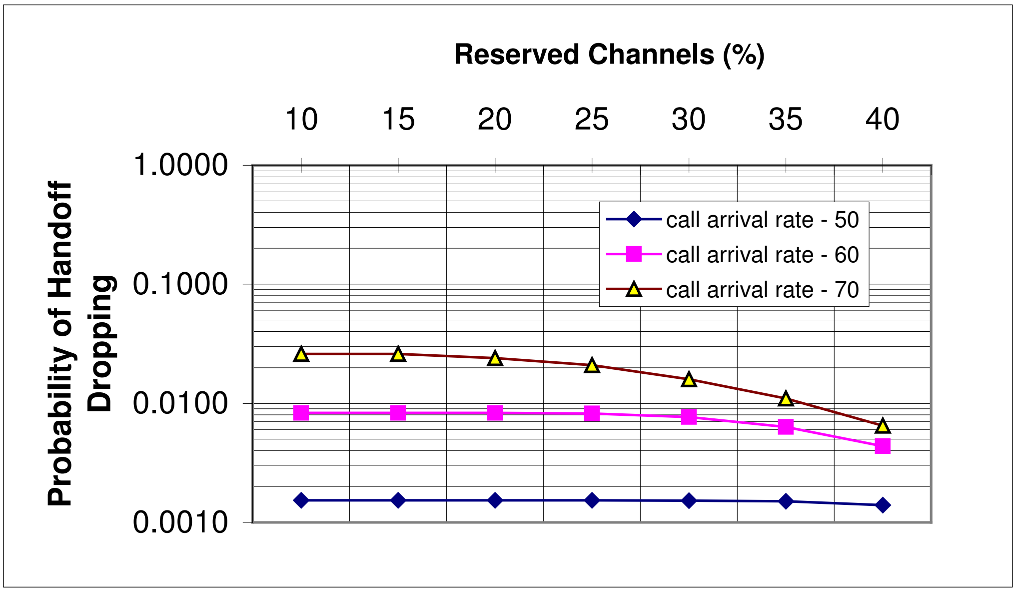


Figure 3.19: The probability of handoff dropping against reserved channels. (B = 7, γ_{rtII} = 10 , μ_{rtII} = 5)

In figure 3.19 a graph of probability of handoff dropping against reserved channels was plotted. The data of the graph can be found in appendix A10. Using assumptions below figure 3.19 in equation (3.15), it is found that as the percentage of reserved channels increases the probability of handoff dropping decreased. This is due to the more number of handoff connections that can be maintain by the reserved channels. However, from the observation, it is noticed that when the reserved channels is 10, 15 and 20% the probability of handoff drop is almost the same. This may be because the value of handoff rate is fixed at 10 calls per second and the increment of reserved channels affects much to the probability of new call block rather than the probability of handoff drop.

3.6.3 CAC for Non Real-Time Connection

For non real-time application, the overload state is considered as a serious issue whenever many calls arrive simultaneously at a base station. Certain amounts of bandwidth are assigned to a base station and depending on these amounts the probability of overload state can be found. First, a comparison was done on two different systems which using VCT architecture as the model. One has a control mechanism of maximum number of connection that can be allowed and another without the control mechanism. The performance of non real-time connections in both systems is then compared. The base station in a cell cluster is assumed to be 7. Average bandwidth for each non real-time connection will be given 1 UB and the call departure is 1. Each base station had been assigned 30 UB and the allowed number of connection in the controlled system is 80. An interesting result was obtain in figure 3.20 where the system controlled mechanism maintain the QoS at below 10^{-7} even in heavy load. The data can be found in appendix A11.

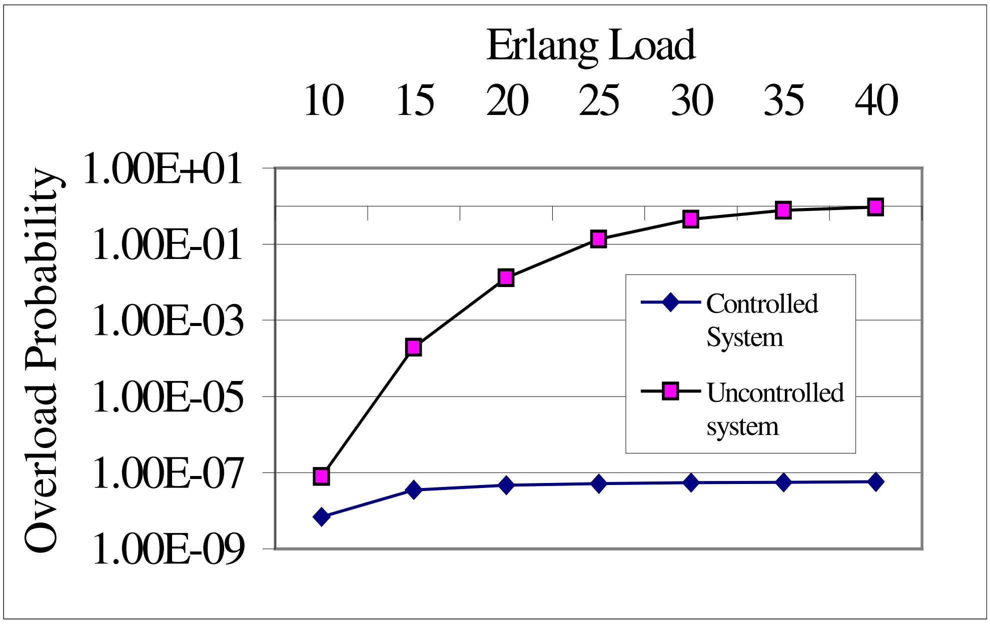


Figure 3.20: Overload probability against Erlang load for two different systems.
(B=7, C = 30UB, μ_{nrt} = 1, $Nnrt$ = 80)

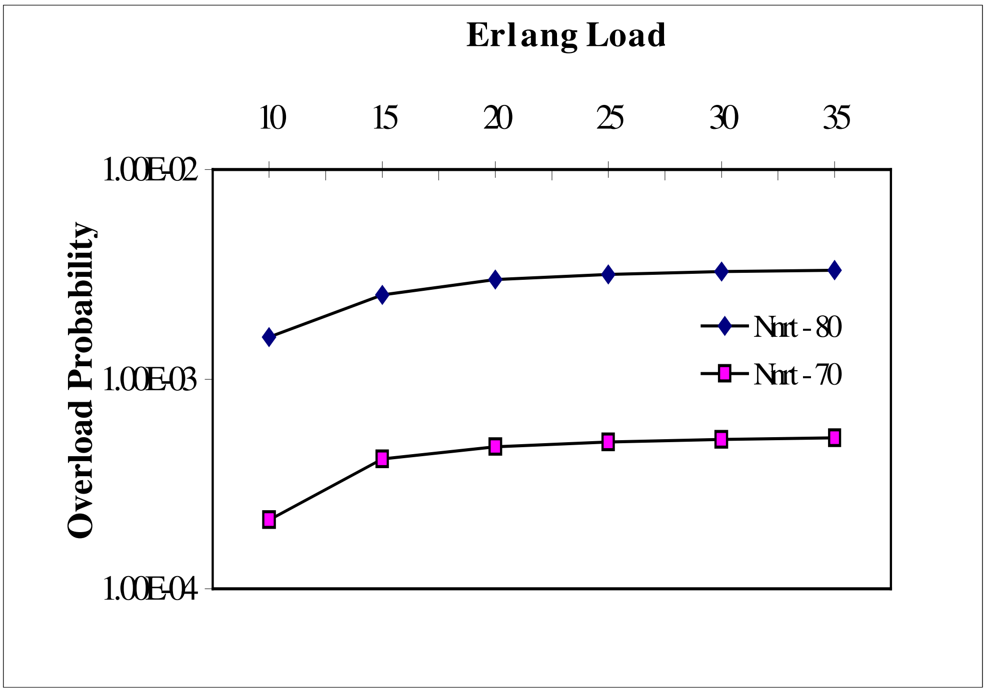


Figure 3.21:Overload probability against Erlang load for total capacity of 20UB (*B*=7, *C* = 20UB, μ_{nrt} = 1)

From figure 3.20, using system without control mechanism, the overload probability grows very rapid as the call arrival increased. This is because in system with control mechanism a threshold is set to limit the incoming calls and this is done to avoid overflow. Now, using system with control mechanism assuming 7 base stations are connected to a VCT, the average call departure is 1 call per seconds. Consider an ideal case where 1 call will have 1units of bandwidth (UB), which approximately equals to 64kb/s. Here, two situations are analyzed where 20 UB and 30 UB are assigned to two base stations. Using the assumptions in equation (3.25), figure 3.21 was obtained and as the results, when the number of connections allowed to a base station is 80, higher probability of overload is obtained and this is because more connection can be connected to the network. The data can be obtained in appendix A12.

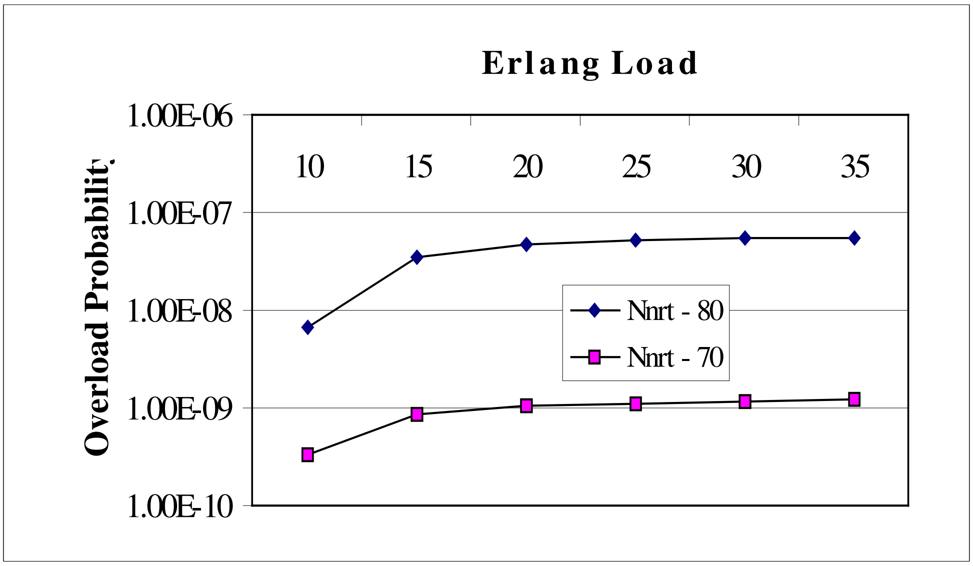


Figure 3.21: Overload probability against Erlang load total capacity of 30UB

(B=7, C = 30UB, μ_{nrt} = 1)

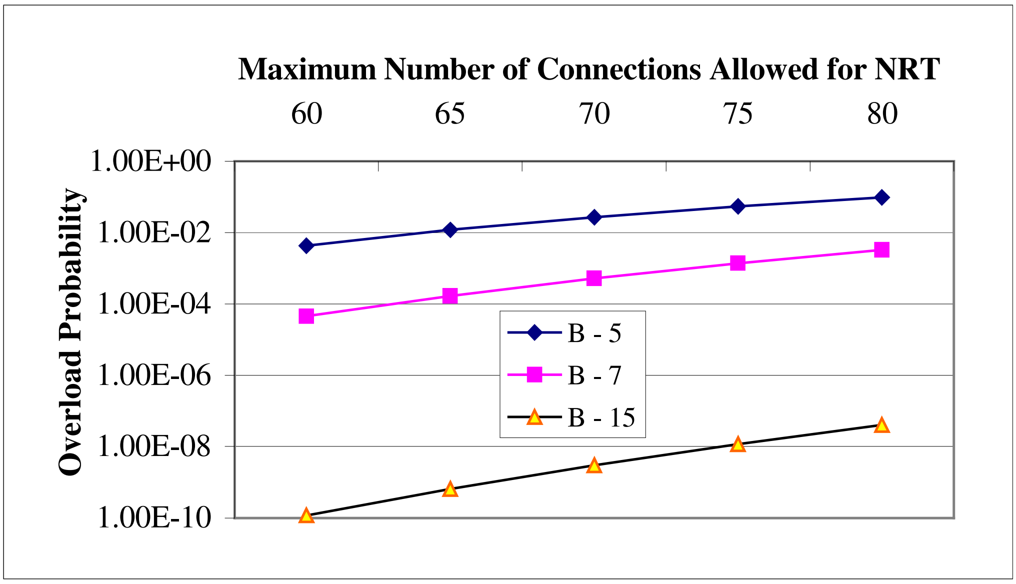


Figure 3.22: Overload probability against maximum number of connection allowed for NRT

. (B=7, C = 20UB, μ_{nrt} = 1, λ_{nrt} = 30)

As the size of base station is increased a much lower probability of overload at a base station is obtained for both *Nnrt*-80 and *Nnrt*-70. This is shown in figure 3.22 and the data of the graph is in appendix A13. However, this did not consider the case where statistical multiplexing is applied. To reduce congestion, the size of the cell cluster is expanded. The arrival rate is 30 per seconds with 20UB of base station capacity. The average bandwidth for each connection is assumed to be 1 UB and the call departure rate is 1 call per seconds. Using the assumptions in equation (3.25), figure 3.23 shows that less overload probability is obtained with more base station in a cell cluster. The data can be obtained in appendix A14. This is because the larger capacity allows more calls and reduces congestions.

3.7 Summary

In this chapter, the performance of real-time and non real-time connection had been analyzed theoretically. First, the real-time connection was analyzed using call admission control without reserve channels, WORC. It was found that using WORC, the probability of handoff calls is treated the same manner as the new incoming calls. Hence when all of the channels in the base station are busy, the new incoming calls as well as the handoff calls will be block. Rejection of newly initiated call is not as annoying as the rejection of the already connected calls. These reserve channels for handoff calls are introduce in mechanism with reserve channels, WRC.

Using threshold *Nrt*, the calls coming into a network are limited. The effect of Erlang load and reserve channels to the probability of new incoming call block and handoff calls is analyzed for WRC. As Erlang load increased, the probability of new incoming call block and handoff call drop increases because the more traffic coming into the network, the more the chances of calls being rejected due to limited channels. However, by expanding the capacity of base station, the probability of handoff call drop is lowered. When the capacity of a base station is expanded, the available bandwidth will be more and chances of call rejection will be less. Expanding the cell cluster will also result the same as expanding the capacity of base station for WRC.

Analysis was also done on the size of the reserved channels. It is found that in WRC, the probability of call block is directly proportional to the size of the reserved channel but the probability of handoff drop is indirectly proportional to the size of the reserved channels. This is because when the percentage of reserved channels is large, q channels is less and less new incoming calls will be accepted when q is bigger, the reserved channels will be less and more calls will be rejected. In comparison of simulation and theoretical results that shows mathematical analysis are valid to be analyzed.

In VCT network that supports homogenous non real-time connections, the QoS performance depends very much on the overload probability. This was shown in the results obtained where the overload probability increased with the increment of Erlang load. However, this can be control through the adjustment of the maximum number of non real-time connection allowed into a network and the capacity assigned for a base station. The efficiency of VCT network model in supporting non real-time connection is highlighted when larger size of cell cluster is used.

Thus from this analysis and comparison of WORC and WRC, it is found that WRC upgrade the QoS performance in terms of probability of handoff call dropping using reserved channels. Whenever a handoff call enters a congested area, it will not be block as it has been given a priority to access and use the reserved channels.

CHAPTER IV

CALL ADMISSION CONTROL WITH QUEUEING

4.1 Introduction

In chapter III, the reserved channels for handoff calls were introduced. Unfortunately this WRC scheme increases the probability of blocking of the new incoming calls by blocking all new incoming calls when the channels are busy. In this chapter queuing for the new incoming call is presented at the same time the probability of handoff dropping will be maintain. The call admission control (CAC) with queue allows queuing of the new incoming calls when there are no available channels at the targeted base station. It also has reserved channels for handoff calls to decrease the probability of handoff drop when channels are busy. Hence, it is called the CAC with queue and reserve channels, WQRC.

In the public switching telephone network (PSTN), the signaling needed for dialing is done on the communication channel. Queuing new incoming calls would therefore result in multiple dialing that would unnecessarily occupy some communication channel. In the cellular system, the dialing (the setup of a call) is done on a separate control channel, which can provide the system with a way of queuing new incoming call without affecting the transmission channels.

In this chapter a method of queuing incoming calls will be analyze using the same virtual connection tree (VCT) as the architecture model. In this scheme, new incoming call requests for connection to a congested network will not be block, but instead they will be queued, waiting for the connection to be accepted to the network. WQRC can be applied to any network despite of their sensitiveness to delay. The queue only involves the request for connection service and not during the call already connected. Such application is propose for wireless telephony and teleconferences systems.

4.2 The Queuing Model

Prioritizing scheme is one of many ways to upgrade quality-of-service (QoS) of network to customers. Though not all real-time connection call can tolerate delay, some (e.g. rt-VBR) may consider a redial scheme to ensure the request will be accepted despite of the waiting time. Assumed that all wireless connections are homogenous real-time connections. Figure 4.1 describes the arrival of calls to a network. There are two types of calls; the handoff calls with handoff call rate of γ_{II}, and the new incoming calls with call rate of λ_{II}.

Both types of calls have access to all channels. When all of the channels are busy, the new incoming calls will be blocked but the handoff calls have priority access to the network. In other words, the new incoming calls will have free access to the system as long as N_{II} channels are available. N_{II} is the pre determined threshold for WQRC, where:

$$N_{II} = g(B_{II} \cdot m_{II}) \qquad (4.1)$$

where B_{II} is the number of base station in WQRC and m_{II} is the number of connection a base station can handle. As mentioned the figure of g can be in between 0 to 100% but for a reasonable value in order not to block new incoming call too much, the percentage of reserved channels must not be less than 20% of the total channel capacity

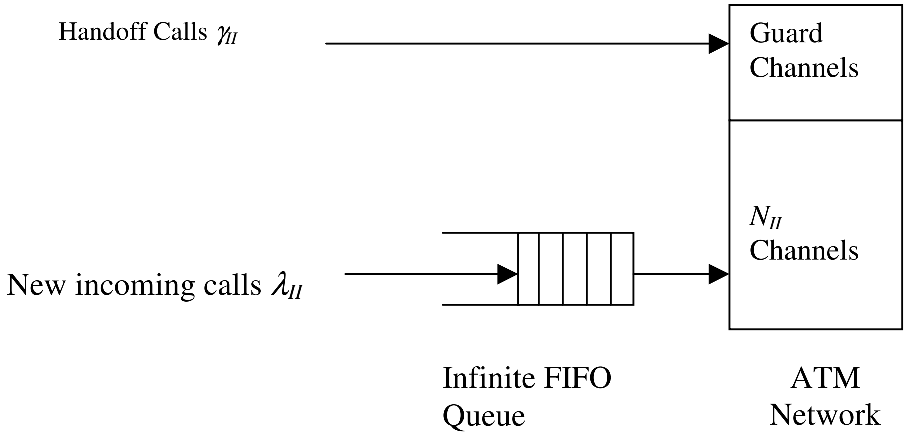


Figure 4.1: The CAC with queue of the new incoming calls during heavy loads.

However, if all N_{II} channels are fully occupied the new incoming calls will be put in the queue and will receive service on first in first out (FIFO) basis as soon as less than N_{II} channels are idle. In a wireless ATM network, setting up of a call is done on a separate control channel. This function however will require a slight modification of the existing connection setup protocol. A MT that wants to initiate a call sends a call request to the ATM network on the control channel, and must wait for an answer from the system. The connection setup protocol will have an expire-time procedure that makes the MT automatically resend the call request after the expire-time period. This will then notify the MT that it has been put in a queue waiting for service. This is illustrated in figure 4.2.

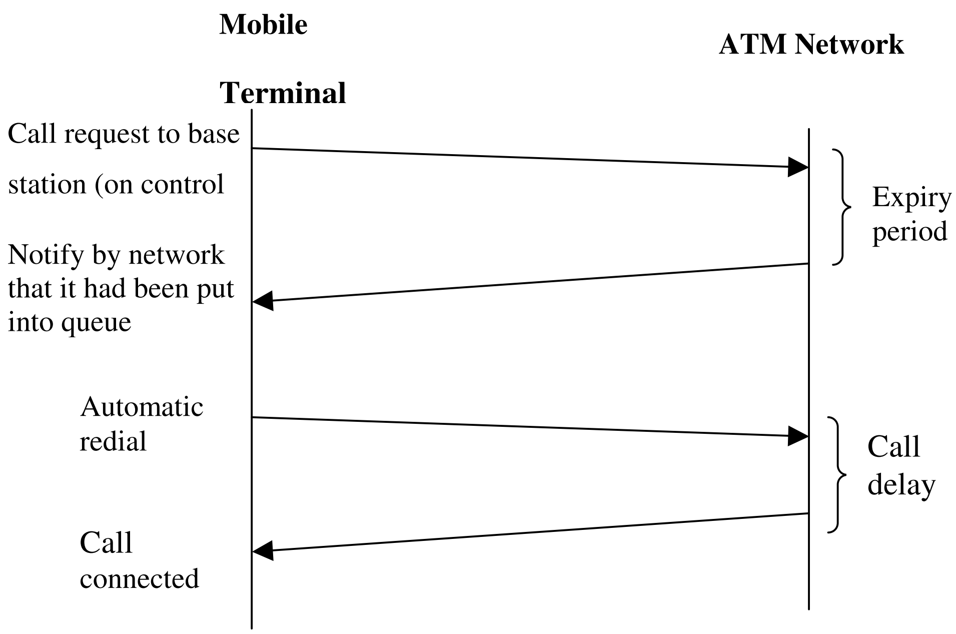


Figure 4.2: The call set-up procedure prior to connection acceptance.

4.3 The Representation of Queuing of Incoming Call

A two-dimensional state diagram of figure 4.3 describes the flow of two different types of calls in a cell cluster. The state diagram is divided into two parts. Part 1 consists of less than N_{II} channels are busy and it is the lower part of the column in the state diagram. This is as same as the pure blocking system. Part 2 is when more than N_{II} server is busy and no more new incoming calls will be allow to be connected. However, priority is given to the handoff calls. Assuming the pair (i_1, i_2) denotes the following notation: i_1 will be the number of customers in the queue, and i_2 will be the number of customers in the service.

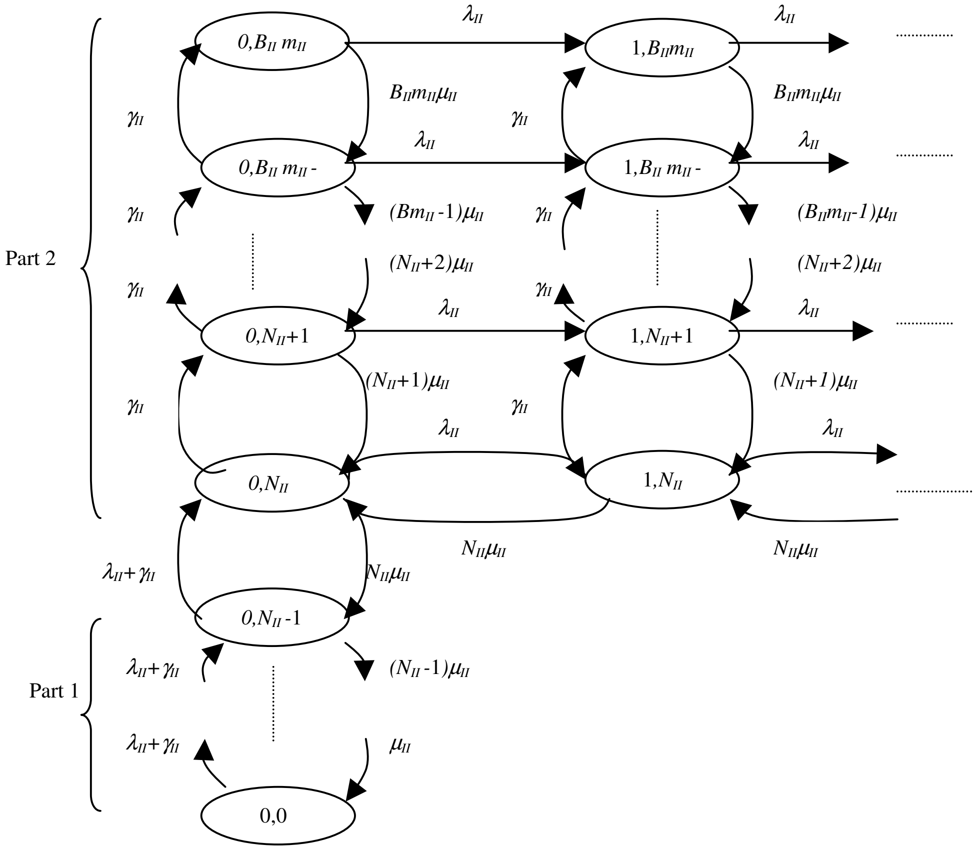


Figure 4.3: Two-dimensional state diagram of WQRC

In figure 4.3 $B_{II}m_{II,}$ is the maximum capacity the network can handle and $\mu_{II\,,}$ is the call departure rate of WQRC. The state diagram can be derived as the balance equations of E_1, E_2 and E_3:

E_1: $(i_1= 0,\ 0 \le i_2 \le N_{II}\text{-}1) \rightarrow P\ (0,\ N_{II}\text{-}1)$:

$$(i_2+1)\ P(0,i_2+1)+\ (b+c)\ P(0,i_2\text{-}1) = (b+c+i_2)\ P(0,i_2) \qquad (4.2)$$

E_2: $(i_1=0,\ i_2 = N_{II}) \rightarrow P\ (0,\ N_{II})$:

$$(i_2+1)\ P(0,\ i_2+1) + i_2\, P(1,i_2)+(b+c)\ P(0,i_2\text{-}1) = (b+c+i_2)\ P(0,i_2) \qquad (4.3)$$

$(i_1=0,\ N_{II}<i_2<B_{II}m_{II}) \rightarrow P(0,N_{II}+1)$:

$$(N_{II}+2)\ P(0,N_{II}+2)+c\ P(0,N_{II}) = [c+b=(N_{II}+2)]\ P(0,N_{II}+1) \qquad (4.4)$$

$(i_1=0,\ i_2= B_{II}m_{II}) \rightarrow P(0,\ B_{II}m_{II})$:

$$c\ P(0,i_2 -1) = [B_{II}m_{II} +b]\ P(0,i_2) \qquad (4.5)$$

E_3: $(i_1 \geq 1,\ i_2 = N_{II}) \rightarrow P\ (1,\ N)$:

$$(i_2+1)\ P\ (1,i_2+1) + i_2\ P\ (2,\ i_2)+ b\ P\ (0,\ i_2)= (b+c+i_2)\ P\ (1,\ i_2) \qquad (4.6)$$

$(i_1 \geq 1,\ N_{II}< i_2< B_{II}m_{II}) \rightarrow P\ (1,N_{II}+1)$:

$$(N_{II}+2)\ P\ (1,N_{II}+2)+b\ P\ (0,N_{II}+1)+c\ P\ (1,N_{II}) = [c+b+(N_{II}+1)]\ P\ (1,N_{II}+1) \quad (4.7)$$

$(i_1 \geq 1,\ i_2 = B_{II}m_{II}) \rightarrow P\ (1,\ B_{II}m_{II})$:

$$b\ P\ (0,i_2)+c\ P\ (1,i_2-1) =(b+i_2)\ P\ (1,i_2) \qquad (4.8)$$

where $b = \lambda_{II} / \mu_{II}$ and $c = \gamma_{II} / \mu_{II}$

E_1 is as same as the pure blocking system in [21] such that:

$$P(0,i_2) = \frac{a^{i_2}}{i_2!} P(0,0), \qquad (4.9)$$

where $0 \leq i_2 \leq N_{II}$ and $a = \dfrac{\lambda_{II} + \gamma_{II}}{\mu_{II}}$

To find the general solution for the state diagram, we need to solve the balance equation. Referring to figure 4.2, for the first column starting from the top:

$$P(0, B_{II}m_{II} - k) = q_k \qquad 0 \leq k \leq B_{II}m_{II} - N_{II}$$

$$P(0, N_{II}) = q_g = \frac{a^{N_{II}}}{N_{II}!} P(0,0) \qquad (4.10)$$

To express $P\ (0,i_2)$, $(0 \leq i_2 \leq B_{II}m_{II})$ in terms of P (0,0), from the balance equation, we have the following relations for q_k's:

$$q_o = q_o \Rightarrow P(0, B_{II} m_{II}) = P(0,0) \tag{4.11}$$

$$q_1 = r + B_{II} m_{II} c^{-1} P(0,0) = (B_{II} m_{II} c^{-1} + r) q_o \tag{4.12}$$

$$q_2 = \left[(B_{II} m_{II} + 1)c^{-1} + (r+1) - 2c^{-1}\right]_{q_{k-1}} - \left[(B_{II} m_{II} + 2)c^{-1} - 2c^{-1}\right]_{q_{k-2}} \tag{4.13}$$

$$q_k = \left(\overline{\alpha} - k\overline{\delta}\right)_{q_{k-1}} - \left(\overline{\beta} - k\overline{\delta}\right)_{q_{k-2}} \tag{4.14}$$

where

$$r = \frac{\lambda_{II}}{\gamma_{II}} = \frac{b}{c}, \overline{\beta} = (B_{II} m_{II} + 2)c^{-1}$$

$$\overline{\alpha} = (B_{II} m_{II} + 1)c^{-1} + (r+1), \overline{\delta} = c^{-1}$$

Equation (4.13) however is only valid for $2 \le k \le B_{II}m_{II} - N_{II}$ but for all $k \ge 2$ we need to define the sequence into a differential equation and dependent transform (refer to appendix C1). So we get $k = 1 \ldots B_{II}m_{II}$ [18]:

$$q_{B_{II} m_{II} - k} = q_o \sum_{j=k}^{B_{II} m_{II}} c^j \binom{j}{j-k} (-1)^{j-k} \cdot \left[\sum_{i=j}^{B_{II} m_{II}} (b + j + 1)_{i-j} \binom{B_{II} m_{II}}{i} c^{-i} \right] \tag{4.15}$$

Where $(x)_k = \dfrac{\Gamma(x+k)}{\Gamma(x)}$

(Refer Appendix C2 for function Γ)

Since $q_o = P\,(0,\, B_{II}m_{II})$ and $q_g = P(0, N_{II}) = \dfrac{a^{N_{II}}}{N_{II}!} \cdot P(0,0)$ as in equation (4.10), equation (4.15) can be written as:

$$P(0, N_{II}) = P(0, B_{II}m_{II}) \sum_{j=N_{II}}^{B_{II}m_{II}} c^j \binom{j}{j - N_{II}} (-1)^{j-N_{II}} \cdot \left[\sum_{i=j}^{B_{II}m_{II}} (b + j + 1)_{i-j} \binom{B_{II}m_{II}}{i} c^{-i} \right] \quad (4.16)$$

Hence, $P\,(0, B_{II}m_{II})$ can be expressed in terms of $P\,(0,0)$ and $P\,(0,k)$ is used in terms of $P\,(0,0)$ in equation (4.15), giving:

$$P(0,i_2) = P(0, B_{II}m_{II}) \sum_{j=N_{II}}^{B_{II}m_{II}} c^j \binom{j}{j - i_2} (-1)^{j-i_2} \cdot \left[\sum_{i=j}^{B_{II}m_{II}} (b + j + 1)_{i-j} \binom{B_{II}m_{II}}{i} c^{-i} \right] \quad (4.17)$$

To simplify the equation, substitute the indexes i_2 for k and ℓ for j - i_2 in equation (4.17), giving:

$$P(0,i_2) = P(0, B_{II}m_{II}) \sum_{\ell=0}^{B_{II}m_{II}} \frac{c^\ell + c^{i_2}}{i_2!} \left[\frac{(\ell + i_2)!}{\ell!} \right] (-1)^\ell \cdot \left[\sum_{i=\ell+i_2}^{B_{II}m_{II}} (b + j + 1)_{i+\ell+i_2} \binom{B_{II}m_{II}}{i} c^{-i} \right] \quad (4.18)$$

Let,

$$\lambda_{II}(B_{II}m_{II}, i_2) = \sum_{\ell=0}^{B_{II}m_{II}-i_2} c^\ell \left[\frac{(\ell + i_2)!}{\ell!} \right] (-1)^\ell \cdot \left[\sum_{i=\ell+i_2}^{B_{II}m_{II}} (b + j + 1)_{i+\ell+i_2} \binom{B_{II}m_{II}}{i} c^{-i} \right] \quad (4.19)$$

Hence,

$$P(0,i_2) = \frac{c^{i_2}}{i_2!} \lambda_{II}(B_{II}m_{II}, i_2) P(0, B_{II}m_{II}) \quad (4.20)$$

This implies that $P\,(0,\, i_2)$ can be written in terms of $\lambda_{II}\,(Bm_{II},\, i_2)$ and $\lambda_{II}\,(Bm_{II},\, N_{II})$ as

$$P(0,i_2)=\frac{c^{i_2}}{i_2!}\lambda_{II}(Bm_{II},i_2)\times\frac{N_{II}!}{c^{N_{II}}}\frac{P(0,N_{II})}{\lambda_{II}(Bm_{II},N_{II})}$$

$$=\frac{N_{II}!}{c^{N_{II}}}\frac{c^{i_2}}{i_2!}\frac{\lambda_{II}(Bm_{II},i_2)}{\lambda_{II}(Bm_{II},N_{II})}P(0,N_{II}) \qquad (4.21)$$

The state probabilities of the first column are obtained as functions of P (0,0) and are given by ($0 \le Bm_{II} - N_{II} \le Bm_{II}$):

$$P(0,i_2)=\begin{cases} \dfrac{a^{i_2}}{i_2!}P(0,0) & \text{for} \quad 0 \le i_2 \le N_{II} \\ \left(\dfrac{a}{c}\right)^{N_{II}} \dfrac{c^{i_2}}{i_2!}\cdot\dfrac{\lambda_{II}(Bm_{II},i_2)}{\lambda_{II}(Bm_{II},N_{II})}P(0,0) & \text{for} \quad N_{II} \le i_2 \le Bm_{II} \end{cases} \qquad (4.22)$$

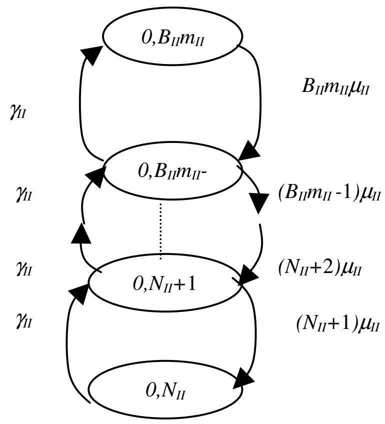


Figure 4.4: The state diagram of busy servers

To find the blocking probability of handoff calls, we have to solve the cases of $i_2 \ge N_{II}$ (i_2 servers are busy).

$$Pi_2=\sum_{i_1=0}^{\infty}P(i_1,i_2) \qquad (4.23)$$

A state diagram involving these larger states (for $N_{II} \leq i_2 \leq B_{II}m_{II}$) is shown in figure 4.4, which is a compressed version of figure 4.3. The diagram is a typical classic purely blocking system and the state probabilities are easily expressed in terms of $P_{N_{II}}$.

$$Pi_2 = c^{i_2 - N_{II}} \frac{N_{II}!}{i_2!} P_{N_{II}} \qquad (4.24)$$

which gives

$$\sum_{i_1=0}^{\infty} P(i_1, i_2) = c^{i_2 - N_{II}} \frac{N_{II}!}{i_2!} \sum_{i_1=0}^{\infty} P(i_1, N_{II}) \qquad (4.25)$$

where $N_{II} \leq i_2 \leq B_{II}m_{II}$. From the column of the system (in the two-dimensional state diagram of figure 4.3) it is found that the outgoing flux (to the right) is balance by the incoming flux from state (i_1+1, N_{II}). Hence, it can be written as:

$$P(i_1 + 1, N_{II}) = \frac{b}{N_{II}} \sum_{i_2=N_{II}}^{B_{II}m_{II}} P(i_1, i_2) \qquad (4.26)$$

where $i_1 \geq 0$

To solve Pi_2, the sum of equation for $i_2 \geq 0$ using equation (4.25) and (4.26) has to be computed. For equation (4.26), for all values of $i_1 \geq 0$:

$$\sum_{i_1=0}^{\infty} P(i_1 + 1, N_{II}) = \frac{b}{N_{II}} \sum_{i_1=0}^{\infty} \left(\sum_{i_2=N_{II}}^{B_{II}m_{II}} P(i_1, i_2) \right) \qquad (4.26a)$$

$$\therefore \sum_{i_1=1}^{\infty} P(i_1, N_{II}) = \frac{b}{N_{II}} \sum_{i_2=N_{II}}^{B_{II}m_{II}} \left(\sum_{i_1=0}^{\infty} P(i_1, i_2) \right) \qquad (4.26b)$$

$$\therefore \sum_{i_1=0}^{\infty} P(i_1, N_{II}) = P(0, N_{II}) + \frac{b}{N_{II}} \sum_{\ell=N_{II}}^{B_{II}m_{II}} \left(\sum_{i_1=0}^{\infty} P(i_1, \ell) \right) \qquad (4.26c)$$

In equations (4.26a-c) i_2 is substituted by λ. Substituting equation (4.26c) in equation (4.25) gives:

$$\sum_{i_1=0}^{\infty} P(i_1,i_2) = c^{i_2-N_{II}} \frac{N_{II}!}{i_2!} \cdot \left[P(0,N_{II}) + \frac{b}{N_{II}} \sum_{\ell=N_{II}}^{B_{II}m_{II}} \left(\sum_{i_1=0}^{\infty} P(i_1,\ell) \right) \right] \qquad (4.26d)$$

The double summation in equation (4.25d) represents the sum of all state probabilities, except for cases where less than N_{II} servers are busy. These missing probabilities are known as functions of P (0,0) (see equation (4.21)).

For stability, all states probabilities must add up to 1, so:

$$\sum_{\ell=N_{II}}^{B_{II}m_{II}} \left(\sum_{i_1=0}^{\infty} P(i_1,\ell) \right) = 1 - \sum_{\ell=0}^{N_{II}-1} P(0,\ell) \qquad (4.27)$$

Thus,

$$Pi_2 = \sum_{i_1=0}^{\infty} P(i_1,i_2)$$

$$= c^{i_2-N_{II}} \frac{N_{II}!}{i_2!} \cdot \left[P(0,N_{II}) + \frac{b}{N_{II}} \left(1 - \sum_{\ell=0}^{N_{II}-1} P(0,\ell) \right) \right]$$

$$= P(0{,}0) \left[\left(\frac{a}{c} \right)^{N_{II}} \frac{c^{i_2}}{i_2!} - \frac{b}{N_{II}} \frac{N_{II}!}{c^{N_{II}}} \frac{c^{i_2}}{i_2!} \left(\sum_{\ell=0}^{N_{II}-1} \frac{a^{\ell}}{\ell!} \right) \right] + \frac{b}{N_{II}} \frac{N_{II}!}{c^{N_{II}}} \frac{c^{i}}{i_2!} \qquad (4.28)$$

Also stable system holds that:

$$\sum_{i_2=0}^{N_{II}-1} P(0,i_2) + \sum_{i_2=N_{II}}^{B_{II}m_{II}} Pi_2 = 1 \qquad (4.29)$$

From (4.9),

$$\sum_{i_2=0}^{N_{II}-1} P(0,i_2) = P(0,0) \sum_{i_2=0}^{N_{II}-1} \frac{a^{i_2}}{i_2!}$$

Hence,

$$1 = P(0,0) \sum_{i_2=0}^{N_{II}-1} \frac{a^{i_2}}{i_2!} + P(0,0)\left(\frac{a}{c}\right)^{N_{II}} \sum_{i_2=N_{II}}^{B_{II}m_{II}} \frac{c^{i_2}}{i_2!} - P(0,0)\frac{b}{N_{II}}\frac{N_{II}!}{c^{N_{II}}}\left(\sum_{i_2=N_{II}}^{B_{II}m_{II}} \frac{c^{i_2}}{i_2!}\right)\left(\sum_{\ell=0}^{N_{II}-1} \frac{a^{\ell}}{\ell!}\right)$$
$$+ \frac{b}{N_{II}}\frac{N_{II}!}{c^{N_{II}}} \cdot \left(\sum_{i_2=N_{II}}^{B_{II}m} \frac{c^{i_2}}{i_2!}\right)$$

Let

$$A = \sum_{i_2=0}^{N_{II}-1} \frac{a^{i_2}}{i_2!} + \left(\frac{a}{c}\right)^{N_{II}} \sum_{i_2=N_{II}}^{B_{II}m_{II}} \frac{c^{i_2}}{i_2!} - \frac{b}{N_{II}}\frac{N_{II}!}{c^{N_{II}}} \cdot \left(\sum_{i_2=N_{II}}^{B_{II}m_{II}} \frac{c^{i_2}}{i_2!}\right) \cdot \left(\sum_{\ell=0}^{N_{II}-1} \frac{a^{\ell}}{\ell!}\right) \qquad (4.30)$$

or

$$1 = \left[P(0,0) \cdot A\right] + \frac{b}{N_{II}}\frac{N_{II}!}{c^{N_{II}}}\left(\sum_{i_2=N_{II}}^{B_{II}m_{II}} \frac{c^{i_2}}{i_2!}\right) \qquad (4.31)$$

From this, *P* (0,0) can be found as:

$$P(0,0) = 1 - \frac{b}{N_{II}}\frac{N_{II}!}{c^{N_{II}}}\left(\sum_{i_2=N_{II}}^{B_{II}m_{II}} \frac{c^{i_2}}{i_2!}\right) \times A^{-1} \qquad (4.32)$$

The probability of handoff blocking is when all of the channels in the network are busy including the reserved channels is given by:

$$P_{HD_{II}} = P(0,0)\cdot\frac{c^{B_{II}m_{II}}}{B_{II}m_{II}!}\left[\left(\frac{a}{c}\right)^{N_{II}} - \frac{b}{N_{II}}\frac{N_{II}!}{c^{N_{II}}}\left(\sum_{\ell=0}^{N_{II}}\frac{a^{\ell}}{\ell!}\right)\right] + \frac{b}{N_{II}}\frac{N_{II}!}{c^{N_{II}}}\left(\frac{c^{B_{II}m_{II}}}{B_{II}m_{II}!}\right) \qquad (4.33)$$

The probability of delay of the new incoming calls is when all of the N_{II} channels are:

$$P_D = 1-\left(\sum_{\ell=0}^{N_{II}-1}\frac{a^{\ell}}{\ell!}\right)P(0,0) \qquad (4.34)$$

4.4 Effect of Queuing

Comparisons between WQRC and WRC are analyzed. Both schemes have the reserved channels for handoff calls where certain percentage of channels are reserved from the total channels. The percentage will depend on the number of population in the particular area. Also, both schemes use the VCT architecture as the network model. To make the comparison, some of the results in chapter III are used. The mathematical analysis in this chapter done using the Mathcad tool where the detail of this programmed had been explained in section 3.6. The programmed of probability of call block and handoff dropping for WQRC are as attached in Appendix D5.

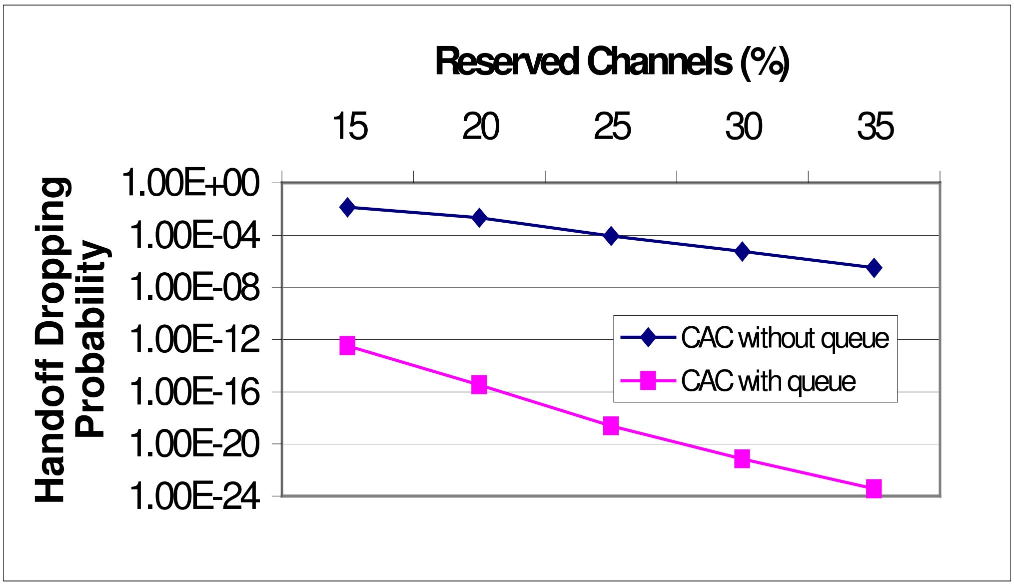


Figure 4.5: Handoff dropping probability against reserved channels.
(B_{II} = 7, m_{II} = 20, λ_{II} = 50, γ_{II} = 10, μ_{II} = 5)

Using equation (4.33), it is found that with 7 base stations per cell cluster the probability of handoff dropping for CAC with queuing is extremely low (see figure 4.5). The data of the graph is in appendix B1. Hence, to make a reasonable comparison, the number of base stations is reduced to 4 instead of 7. The handoff rate for this analysis is also changed to 20 calls per seconds. The new incoming call arrival will still vary from 50 to 75 calls per seconds. The call departure is assumed to be 2 calls per seconds in order to slow down the departing of calls. This is to balance with the calls, which are queued.

4.4.1 Effect of Queuing To Reserved Channels

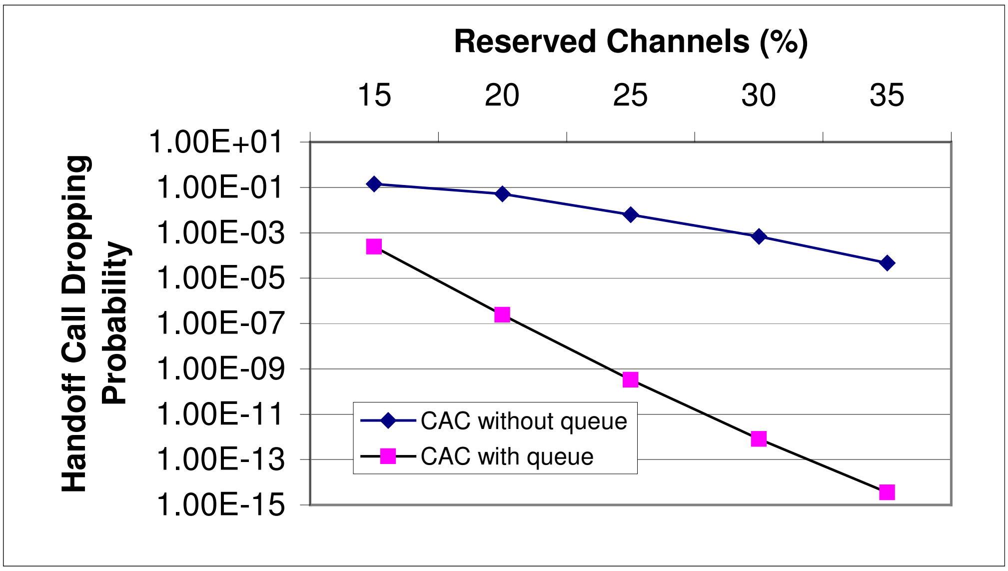


Figure 4.6: Handoff dropping probability against reserved channels. (B_{II} = 4, m_{II} = 20, λ_{II} =50 per seconds, γ_{II} = 30 per seconds, μ_{II} = 2 per seconds)

A graph is plotted in figure 4.6 showing the comparison of the handoff dropping probability for scheme with queuing of new call block and reserved channels (WQRC) and scheme with reserved channels (WRC), using 20 connections per base station (m = 20). The data is as appendix B2. Each call is assigned to a fixed bandwidth. Using the assumptions stated below figure 4.6 in equations (4.33) and (3.15), figure 4.6 was obtained. The handoff dropping probability for WQRC dropped as the reserved channels for handoff is increased. Using WRC, the handoff dropping probability also dropped with the increasing of reserved channels but with higher probabilities than WQRC. Since the inputs to the network, the new incoming call rate and the handoff call rate are fixed, expending reserved channels gave more bandwidth for the handoff call and this decreases the probability of blocking.

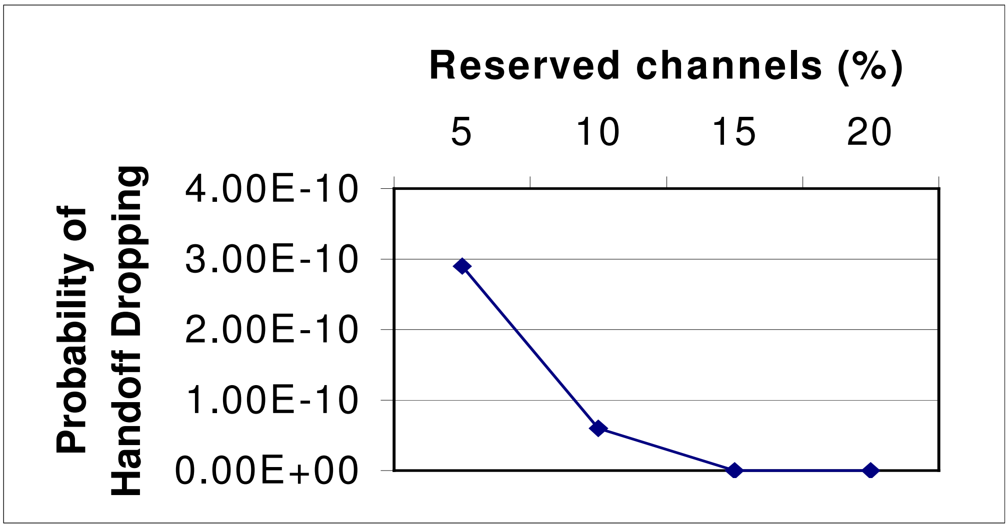


Figure 4.7: Probability handoff dropping against reserved channels.
(PH - probability of handoff dropping)
(B_{II} = 4, m_{II} = 20, λ_{II} = 50, γ_{II} = 30, μ_{II} = 2)

In figure 4.7 and 4.8, the effect of queuing to reserved channels is analyzed. For both graphs, the new incoming call arrival rate, λ_{II} is 50 calls per seconds and handoff call rate; γ_{II} is 30 calls per seconds with call departure ate of 2 calls per seconds. The data are in appendix B3 and B4. The number of channels for a base station is 20 with 4 base stations in a cell cluster. Using the above assumptions in equation (4.33), figure 4.7 shows that when more reserved channels are assigned to a network the probability of handoff dropping declined rapidly. This is due to the bandwidth capacity reserved for handoff calls during handover of access point to the backbone network.

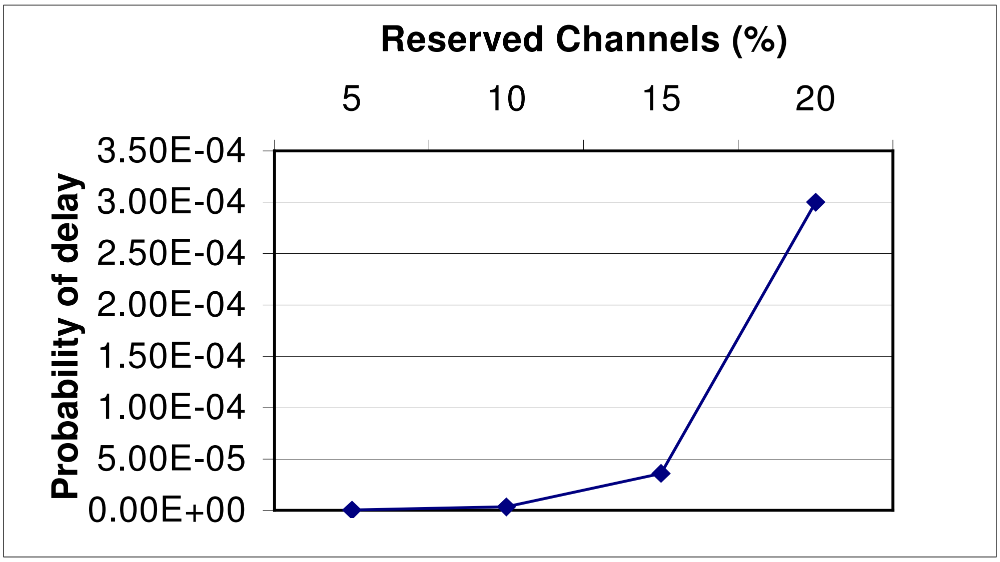


Figure 4.8: Delay probability against reserved channels.
(PD - probability of delay)
(B_{II} = 4, m_{II} = 20, λ_{II} = 50, γ_{II} = 30, μ_{II} = 2)

Using the assumptions stated below figure 4.8 in equation (4.34) figure 4.8 shows that with more reserved channels assigned to a network, the probability of delay of the new incoming calls rises. This is because when the amount of bandwidth reserved for handoff calls is expanded, the remained channels will be occupied and queue will build up. This means that as the reserved channels expanded, queue will be more but the probability of new call block will not increase. Hence, handoff dropping probability as in figure 4.7 will be less. This also shows that in deciding the capacity of the reserved channel, the maximum allowed delay would have to be taken into consideration.

4.4.2 Effect Of Queuing To Erlang Load Per Cell Site

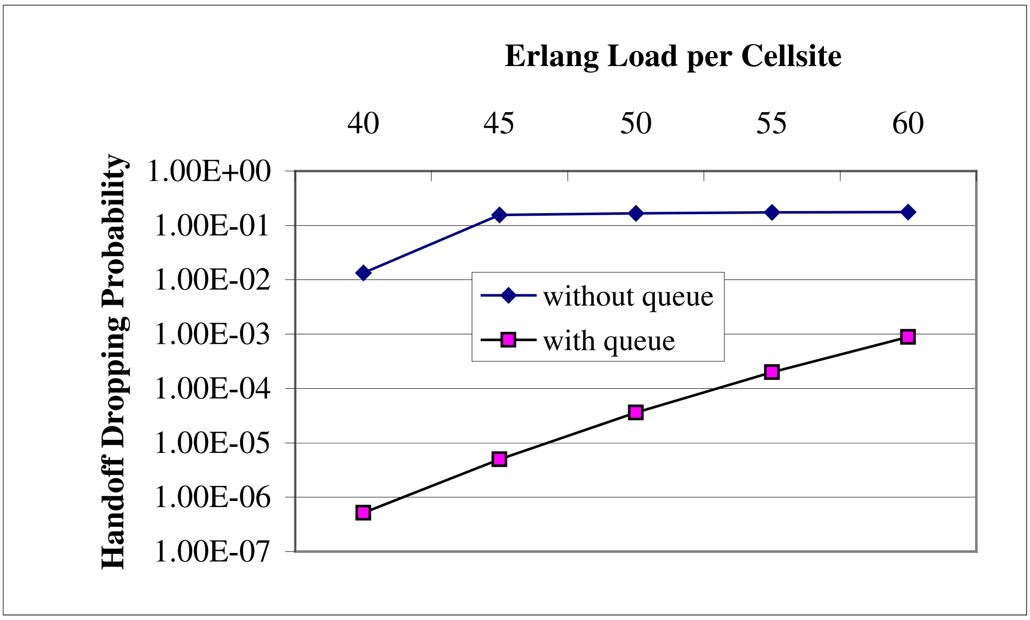


Figure 4.9: Handoff dropping probability against Erlang load per cell site.
(B_{II} = 4, m_{II} = 20, γ_{II} = 20, g = 0.85, μ_{II} = 2)

Figure 4.9 shows the comparison of handoff dropping probability for scheme with queuing of new call block and reserved channels (WQRC) and scheme with reserved channels (WRC). The data is as appendix B5. The assumptions stated below figure 4.9. Using these assumptions in equation (4.34), the probability a handoff call dropping increases as more load are injected into the network. WRC gave a steady value at the increment of load but WQRC seemed to have a much lower handoff dropping probability.

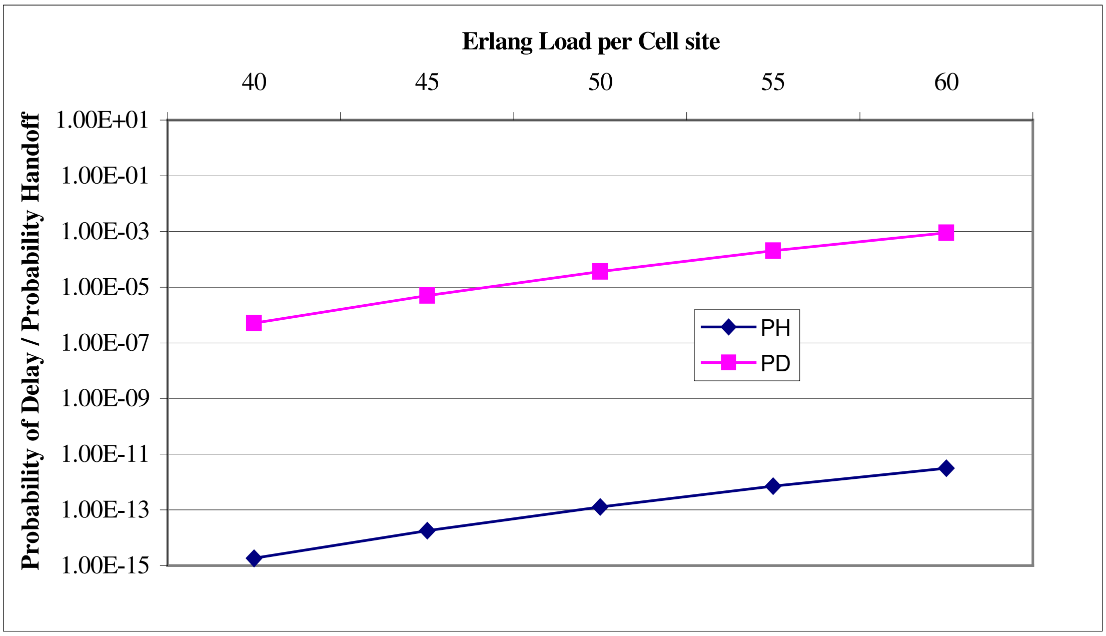


Figure 4.10:Probability of delay and handoff dropping probability against Erlang load per cell site
(PH; probability of handoff dropping, PD; probability of delay)
(B_{II} = 4, m_{II} = 20, μ_{II} = 2, γ_{II} = 30, g = 0.85)

Using the assumptions below figure 4.10 in equations (4.33) and (4.34), a graph of handoff dropping probability against the function of load per cell site is plotted in figure 4.10. The data is as appendix B6. It is found that when the load of the new incoming calls is increased, the probability of delay of the service increases. When probability of delay increases, the probability of new call blocked will decrease as no new incoming calls are blocked though the channels are busy. The handoff call arrival is maintained at 30 calls per unit time, the probability of handoff dropping increased. This is due to more channels, which is less than N_{II} are used by the new incoming calls in this particular case. The channels reserved for handoff calls have to cater for more handoff calls in a second as compared to when less new incoming call load are injected into the network.

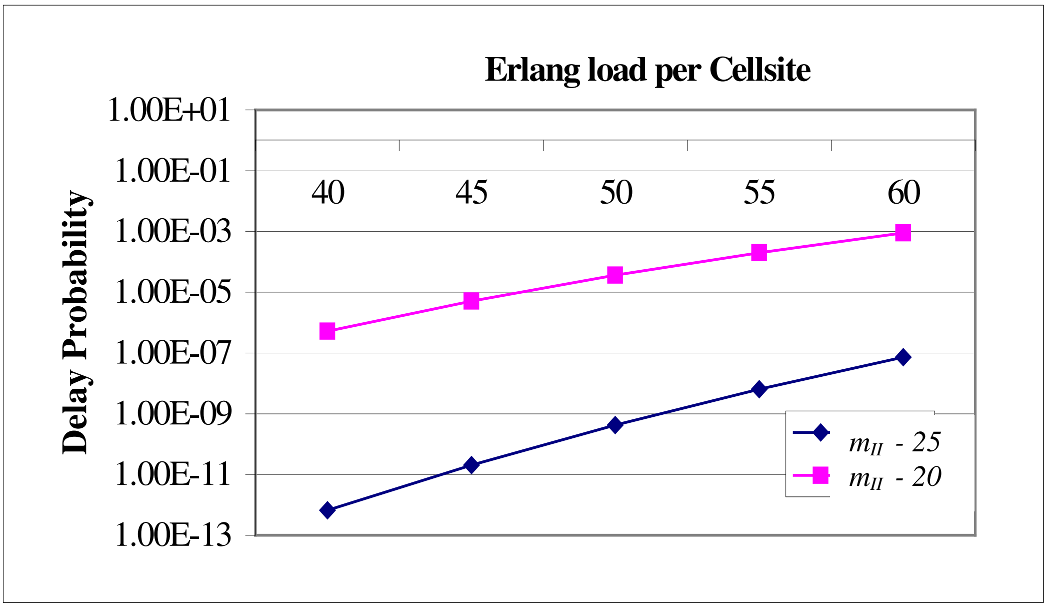


Figure 4.11: The delay probability against Erlang load per cell site (*m* -number of channels at a base station)
(B_{II} = 4, λ_{II} = 50, γ_{II} = 20, μ_{II} = 2, N_{II} = 2)

The effect of delay probability for different number of connection a base stations can handle, m_{II}, in network is analyzed in figure 4.11. The data of this graph can be obtained from appendix B7. Using the assumptions below figure 4.11 in equation (4.34) a reduction is obtained in delay probability with larger capacity of a base station. This is because the more channels assigned to a base station the more calls a base station can handle at a particular time and this will reduce the queuing of the new incoming calls.

4.5 Summary

In chapter IV, the effect of queuing to reserved channels and Erlang load per cell site had been analyzed using mathematical approach. From the results, it shows that as the percentage of reserved channels increases, the probability of handoff dropping decreases as well. This is due to the expanded capacity of reserved channels. However, this increases the probability of delay of the new incoming calls. When probability of delay increases, the probability of new incoming call block is eliminated. The Erlang load per cell site also affected the probability of handoff dropping and new call delay. It is found that the probability of delay and the probability of handoff dropping increases as Erlang load increase. This is because of the limited channels available when more traffic are injected into a network. However, the probability of delay can be reducing by expanding the capacity of base station.

The analysis done in chapter III found that scheme with reserved channels (WRC) decreases the rapid increment of probability of handoff call dropping in scheme without reserved channels (WORC) by reserving channels for handoff calls. However, by decreasing the probability of handoff call drop, WRC increased the probability of new call block when all of the channels in the network are not available. Queuing of new incoming is introduced to solve the above problem while maintaining the low probability of handoff call dropping. Eliminating call blocking probability however introduces delay.

CHAPTER V

ACCESS MECHANISM

5.1 Introduction

Maintaining quality-of-service (QoS) in a network is important to satisfy the network customers. Chapter III and chapter IV had improved QoS performance at call level in terms of probability of call block and probability of handoff dropping. Another parameters a connection must have in order not to violate QoS constraint at call-level is having the right bandwidth rate. Effective bandwidth makes a big different in allowing more connections to be accepted to a network. This does not only apply to data connection but also to speech connections. Integrated services supported by ATM, needs a mechanism to control multi services that is using the same line. This access mechanism can be the complete partitioning or the complete access. Each mechanism has its own advantage. Though effective bandwidth allows more connection in a network, the QoS performance has to be monitor to provide better service to the customer. Hence, the pre determined of bandwidth rate prior to connection is set to maintain the QoS requirement at call and cell level.

Asynchronous Transfer Mode (ATM) technology increase the efficiency of a connection using the fact that many sources do not offer a constant bit rate (CBR): the mean bit rate can be considerably lower than the peak bit rate. In other words ATM has put forward some flexibility in telecommunication, which is slot on demand concept. The Erlang theory deals with CBR sources that hold one unit of source for the whole duration of the connection. The activity of a CBR does not vary. While active, a CBR source produces, for example 64kb/s voice, one octet of information every 125μs. This can be compare with variable bit rate (VBR) sources of data from a computer, where for example in an interval of 46.1 seconds there is only activity during 1 second. During this active period, characters are transmitted at the rate of one character every 3333μs [34]. In this one connection, many active periods (bursts) may occur.

Figure 1.2 in section 1.5 illustrated the traffic generated by a VBR source at three levels. In this example the VBR source is an on/off source, which has an on/ off at burst level [35]. The important feature of this representation is the distinction of different levels each one with its own time scale. At the very top we have connection/ call level, where the line is high when the source is under consideration has a connection established and low otherwise. The duration of call is typically measured in minutes. For VBR sources, the faster time scale of the burst level is necessary, to describe the burst of activity that alternate with periods of silence. For VBR sources, the line at burst level is high for example, when a block of text is transmitted and low during the periods of silence. The fastest time scale is at cell level where at this level the transmission of ATM cells takes place. The function of CAC in determining the effective bandwidth is important in order to get higher number of connection at a base station. Moreover, it is to provide sufficient amount of bandwidth to the accepted calls and to maintain the QoS of the already connected calls.

These additional benefits of increased efficiency of ATM protocol does not come for free, somehow provisions must be made to avoid jeopardizing the QoS constraints [36]. Hence, this requires a method of setting a threshold bandwidth so that an accepted connection will be allowed to connect to the network provided that the bandwidth given is not more than the threshold bandwidth. This is done so that every connection in the network will have better QoS. Here, two types of access mechanism will be analyzed, which is the complete partitioning (CP) and the complete access (CA). The QoS of connections using both access mechanisms will be analyzed.

5.2 Statistical Multiplexing

In an ATM network, several sources will superimpose on a single link, e.g. a trunk line that may carry hundreds of videophone calls. While in a synchronous transfer mode (STM) network or network with fixed bit rate (FBR) coding, the required bandwidth on that trunk will simply be the mathematical sum of all individual fixed bit rate [37]. In ATM network, we gain on the efficiency by relying on the statistical multiplexing effect of sources, on the condition that enough sources are multiplexed and they are not correlated.

The issue of resource allocation to connection is a challenge due to the lack of physical separation between resources used by different connection in ATM based networks. The resources actually are the carrier system, which are for high capacity networks designed for the digital transmission of voice, video and data. The common carrier system used in Europe, Japan and North America is listed in table 1.5.

For real-time connection usually CBR are used and each connection is given a fixed bit rate (e.g.: 64kb/s for voice). For non real-time connections, effective bandwidth allocation can be applied to connections. Most of the ATM sources are of the on/ off type. When they are active, they transmit cells at their peak rate, which is for example determined by the processing speed of the sources or destination. This means that during a burst of activity, cells are transmitted less equally spaced in time. Clearly, this defined the mean bit rate, which is the bit rate averaged over the duration of a connection, and the peak rate is the bit rate averaged during a burst.

Table 5.1: Common carrier systems

North America	***Japan***	***Europe***
64kb/s	64kb/s	64kb/s
1.544 Mbit/s 24 voice channels	1.544 Mbit/s 24 voice channels	2.038 Mbit/s 30 voice channels
6.312 Mbit/s 96 voice channels	6.312 Mbit/s 96 voice channels	8.448 Mbit/s 120 voice channels
44.736 Mbit/s 672 voice channels	32.064 Mbit/s 480 voice channels	34.368 Mbit/s 480 voice channels
274.176 Mbit/s 4032 voice channels	97.728 Mbit/s 1440 voice channels	139.264 Mbit/s 1920 voice channels

Hence, burstiness, b_r of a connection can be calculated by:

$$b_r = \frac{b_p}{b_m} \qquad (5.1)$$

where b_p is the peak bit rate of a connection and b_m is the mean bit rate of a connection. In the section 5.1, the burst length had been discussed, which reflects the characteristics of a VBR connection - it is measured in seconds. In other words, if the burst length of a connection is long, the possibility of gaining bandwidth within the silence period is lower than if otherwise. However, in this section the effective bandwidth is determined without using the burst length since we are using the method of COST242 [38].

To find the effective bandwidth, we must first know the peak bit rate, B_p, and mean bit rate, b_m, of a connection. Moreover, it requires the parameters of the probability of cell loss, *Ploss* and the capacity of a base station, *C*. Assume that *Ploss* to be 10^{-9}.

Before pursuing on to the calculation, there are a few parameters that have to be considered in order to determine the right formula, which are:

$$x = \frac{b_m}{b_p} \qquad (5.2)$$

$$z = \frac{-2\log Ploss}{C/b_p} \qquad (5.3)$$

$$a = 1 - \frac{2\log Ploss}{100} \qquad (5.4)$$

If $z \leq 1$ and $x \leq 1/3z$, the effective bandwidth, b_E can be obtain by using the following equation,

$$b_E = ab_m\left(1 + 3z(1 - x)\right) \qquad (5.5)$$

If $z \leq 1$ and $x > 1/3z$, the effective bandwidth is

$$b_E = ab_p \qquad (5.6)$$

If $z > 1$ and $x \leq 1/3z^2$, the effective bandwidth is,

$$b_E = ab_m\left(1 + 3z^2(1 - x)\right) \qquad (5.7)$$

If $z > 1$ and $x > 1/3z^2$, the effective bandwidth is,

$$b_E = ab_p \qquad (5.8)$$

5.2.1 Effective Bandwidth

Finding the effective bandwidth in ATM connection is important in order to maintain the QoS of the connection and to make sure that the connections are used efficiently. The effective bandwidth can be found using equation (5.5) - (5.8) and depending on the values of parameter x and z. The effective bandwidth of different capacity assigned to a network is then compared. The capacities of a base station are measured in terms of units of bandwidth (UB) where 1UB is equivalent to 64kb/s. Probability of cell loss is assumed to be 1×10^{-9}.

Figure 5.1 compares three European standard capacities of networks (indicated in table 5.1), which are 32 UB, 125 UB and 531 UB (approximately 2Mb/s, 8Mb/s and 34Mb/s respectively). The data of this graph can be found in appendix C1. Assuming that the peak bandwidth rate is 64kb/s. If the mean bandwidth rate varies from 3.2kb/s to 28.8kb/s with increment of 3.2kb/s, the source utilized by the real-time connection will increase. It is found that the amount of effective bandwidth given to a connection increased as the source utilization for real-time connection increases. Also, the less capacity left for non real-time connection to share among them, the effective bandwidth tease towards the peak rate, 64kb/s. When less capacity available, more non real time connection will be rejected due to the insufficient bandwidth available. Moreover, the larger amount of capacity assigned to a base station, the less probability that a non real-time connection will be given the peak bit rate bandwidth. This means, more connections will be accepted.

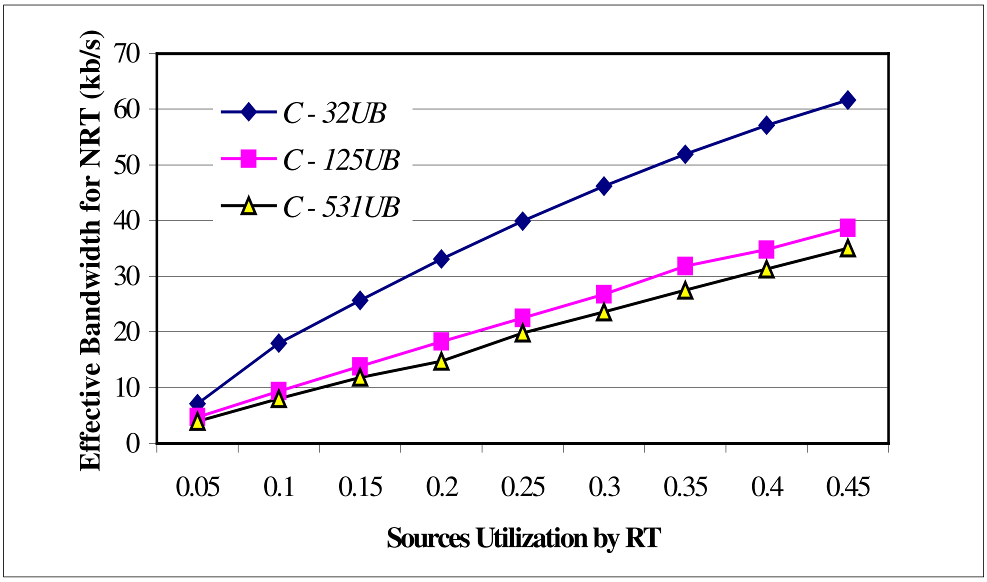


Figure 5.1: The effective bandwidth against sources utilized by real-time connections.
(*Ploss* = 1 x 10^{-9}, b_p = 64kb/s)

5.2.2 Effect of Burtiness to Statistical Multiplexing

Figure 5.2 shows the analysis done on the effect of burstiness to the number of connection that can be accepted to a network. Assumed that mean bit rate is 6.5kb/s with peak bit rate 64kb/s. Capacity of the real time connection varied from 0kb/s to 2Mb/s, the burstiness of different connections is analyzed. The total capacity will be used for both connections. The number of connection of NRT is found by dividing the total capacity of non real-time connection to the value of effective bandwidth at that particular burstiness.

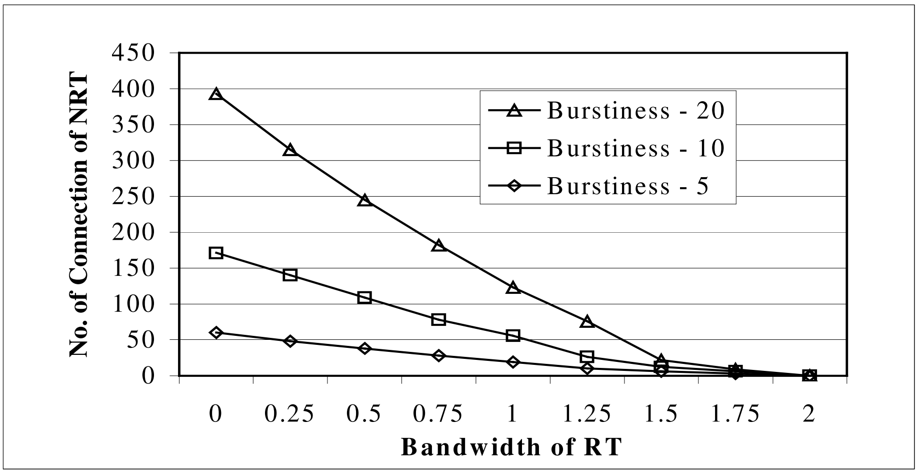


Figure 5.2: Number of non real-time (NRT) against bandwidth of real-time connection (RT) (*Ploss* = 1 x 10^{-9}, b_p = 64kb/s, b_m = 6.5kb/s)

When capacity for real-time connection is varied, the remaining bandwidth will be shared among the non real-time connection. Hence, this analysis is to observe the number of non real-time connections will be accepted with a certain level of burstiness. In figure 5.2, three different connections that have burstiness of 5, 10 and 20, are compared and are plotted to the function of number of non real-time connections. The data of this graph is in appendix C2. It is found that when more bandwidth is allocated for the real time connection, the number of non real-time connections allowed to be accepted to the network decreases. Further, it is found that burstiness affect the effective bandwidth, where with burstiness of 20 the number of non real-time connection allowed to the network is higher compare to burstiness of 5. The burstiness determines effective bandwidth. The effective bandwidth will be use as a parameter to decide the amount of bandwidth accepted to the non real time connections.

5.3 Complete Partitioning (CP)

WATM environment provides increase flexibility in supporting various services; its nature however, poses difficult traffic control problems when trying to achieve efficient use of network resources. One such problem is the bandwidth management and allocation. Complete partitioning is a mechanism that divides the total bandwidth allocated at a base station for real-time and non real-time connections. This is illustrated in figure 5.3. If C is the total capacity, 0 - $C1$ bandwidth will be allocated for real-time connections. Non real-time connection can use the remaining of C - $C1$ bandwidth.

QoS class in terms of cell delay, cell jitter and cell loss is essential in order to provide better performance for the network customers. However, if more connections are given less bandwidth than the connections required, these QoS parameters will be lowered. In real situation this have to be avoided. Here, the probability that a connection will be given less bandwidth than it required will be analyzed where the threshold or minimum bandwidth will be set to be less than 1 UB per connection. Assumed that the bandwidths are in terms of units of bandwidth (UB), where real-time connection fixed the amount of bandwidth, which equals to $BW1$ UB and bandwidth of $C1$ UB, is assigned to real-time connections in each base station. The total capacity of C UB will be assigned to each base station. Hence at any particular time, the bandwidth available to non real-time connections (C – $C1$) UB is shared equally between all of them [39].

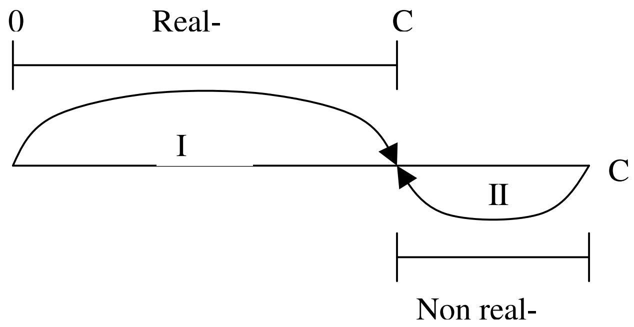


Figure 5.3: Complete Partitioning

In this access-sharing scheme the QoS of non real-time connections can be found independently from the real-time connections. The wireless bandwidth available to any mobile terminal is discrete random variable P_r where r is the available rate and belongs to the discrete space $\{C, C/2, C/3, \ldots, C/Nnrt\}$. The steady state probability of rate $r = C/s$ being available to a mobile terminal (MT) can be found by multiplying the probability s number of connection is connected to the network with number s and divide by the probability that all mobile terminals are connected to the network and this can be written as:

$$P_{r=C/s} = \frac{sP_{nrt}(s)}{\sum_{s=0}^{Nnrt} sP_{nrt}(s)} \tag{5.9}$$

where $P_{nrt}(s)$ is given in equation (3.11) and $Nnrt$ is the maximum number of connections allowed to the network. For real-time connection the admission control will allocate peak bit rate and for non real-time connections with variable bit rate, the admission control will give effective bit rate, b_E, where effective bit rate is in between peak bit rate, b_p, and mean bit rate b_m. The probability that the available bandwidth rate to any non real-time connection is less than a threshold $rmin$ is given by:

$$P[r < r\min] = \frac{\sum_{s=n}^{Nnrt} sP_{nrt}(s)}{\sum_{s=0}^{Nnrt} sP_{nrt}(s)} \tag{5.10}$$

Here, $r$min is defined by the effective bandwidth. n is the smallest integer greater than $C/r\min$ and $P_{nrt}(s)$ is given is equation (3.11).

5.3.1 The Effect Of Erlang Load Per Cell Site To Probability That The Available Bandwidth Rate To Non Real-Time Connection Is Less Than Threshold r*min*, P(r < r*min*)

An analysis was done on the affect of Erlang load of real-time connection to probability that the available bandwidth rate to any non real-time connection is less than a threshold *rmin*, P ($r < r$min). Assumed that the total capacity of a network is 40 UB (≈ 2.5Mb/s). The number of base stations, *B* in a cell cluster is assumed 7. The calls arrive at the network are exponentially distributed with rate varies from 35 to 55 calls per second and the calls departure are exponentially distributed at 2 call per seconds. The minimum bandwidth threshold *r*min is 0.8 UB. The real time connection is given 30UB of the total capacity.

Using the above parameter (assumptions) in equation (5.10), figure 5.4 shows the Erlang load per cell site did not affect the P ($r < r$min), but with lower value of maximum number of connection allow, the P ($r <$ rmin) is lowered. The data of this graph can be found in appendix C3. This result is obtain because when more number of non real-time connections is allowed to the network, the probability that a non real time connection will be allocated with less than *r*min bandwidth by the admission control is higher. Hence, statistical multiplexing has takes place. When the bandwidths for real-time and non real-time connections are partitioned, they are fixed. As the Nnrt vary with the same amount of bandwidth allocated for non real-time, the $P(r < r$min) changes. For example if $C1$= 30 UB, C-$C1$= 10 UB. For 10 UB remained for non real-time connection, if *Nnrt* =30 the effective bandwidth available for connection is 0.333UB (~ 21.333kb/s). In contrast if *Nnrt* =50, the effective bandwidth available for non real time connection is 0.2UB(~12.8kb/s). The mathematical analysis for this chapter was done using Mathcad and the programmes are as attached in Appendix D6 and D7.

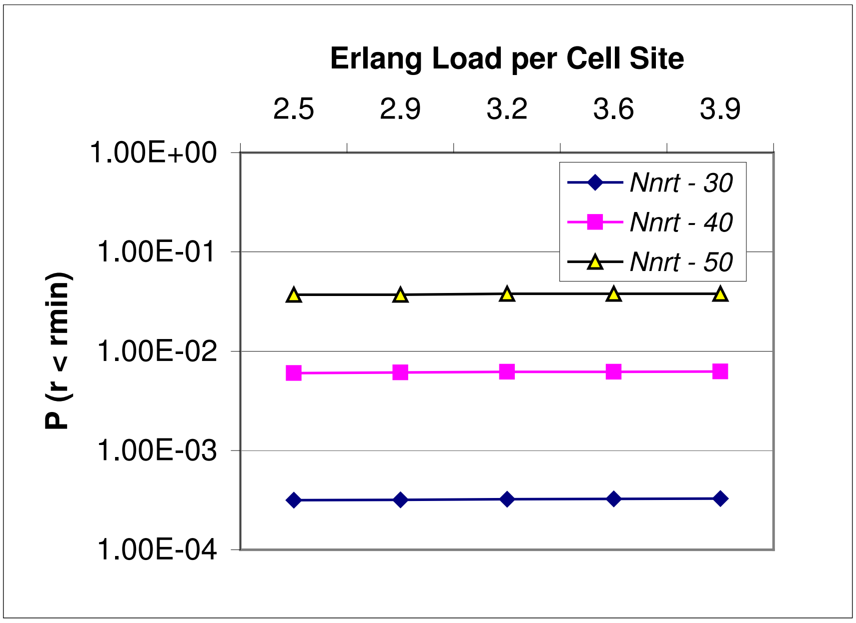


Figure 5.4: Probability of bandwidth is less than the minimum rate, P (r < $r$min) against Erlang load per cell site (C = 40UB, B = 7, μ_{nrt} = 2, $r$min = 0.8UB, C1 = 30)

5.3.2 The Effect of Maximum Number of Connection Allowed

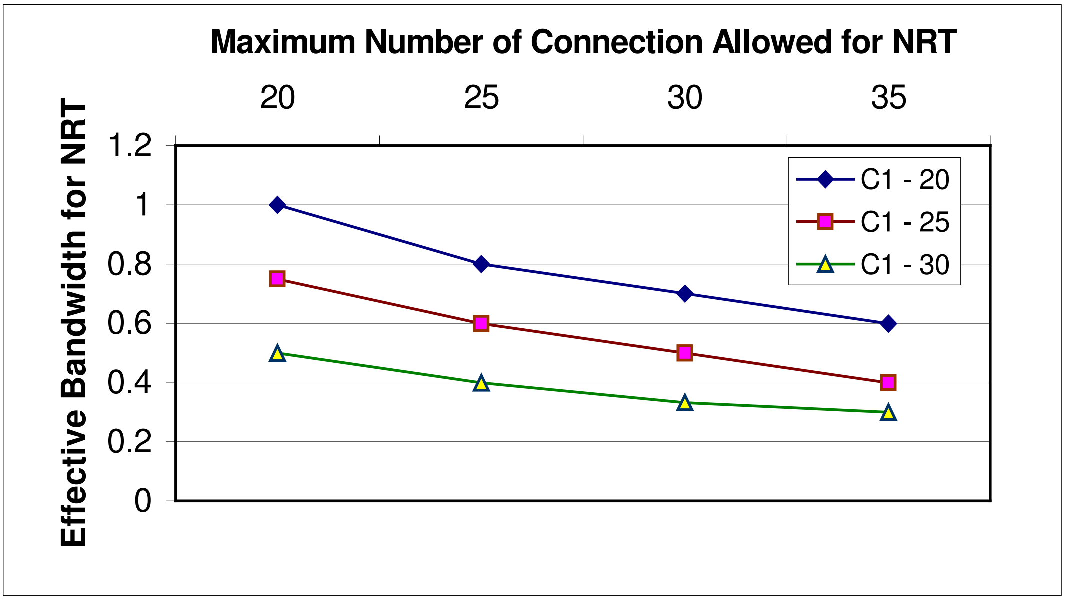


Figure 5.5: The effective bandwidth of non real-time against the maximum number of non real-time connections allowed, (*Nnrt*) for different C1.
(C = 40UB, B = 7, μ_{nrt} = 2, $r$min = 0.8UB)

Figure 5.5 shows the effective bandwidth of non real-time against the maximum number of non real-time connections allowed, (*Nnrt*) for different C1. The call arrival of the non real-time connection is varied from 35 to 55 calls per seconds. The data of this graph can be found in appendix C4. Using the above assumptions in equation (5.10), it is found that, using complete partitioning , CP, as the amount of bandwidth allocated for real-time increases, the amount of bandwidth remained for non real time decreases. Hence the effective bandwidth given to each non real-time connection will be less than the predetermined threshold, *r*min. This is because when less bandwidth remained for non real-time connection the available bandwidths are limited to be shared amount the requested real-time connection. Hence the effective bandwidth given to each non real-time connection is less than 0.8UB.

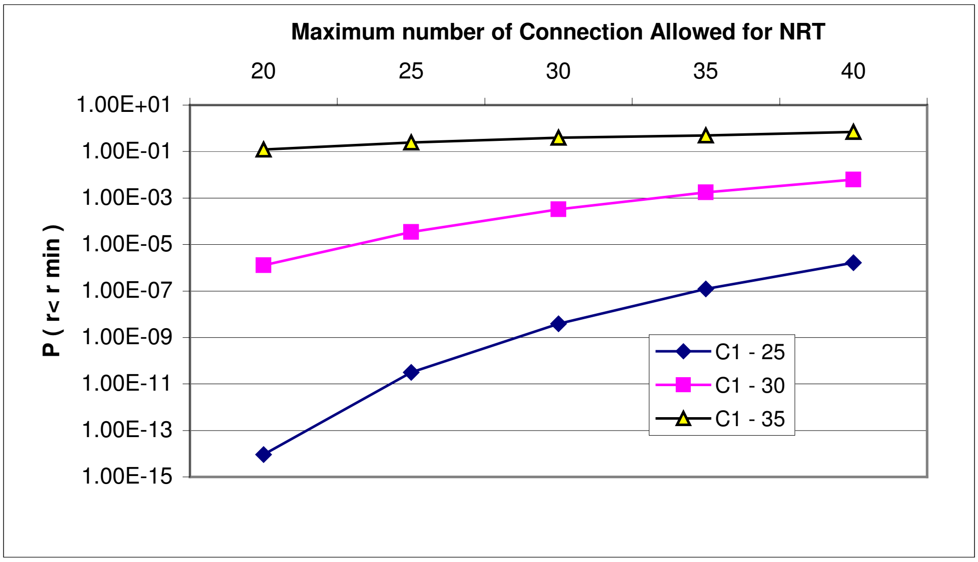


Figure 5.6: Probability that the bandwidth is less than the minimum rate, P (*r* < *r*min) against maximum number of connection allowed to a network for different size of C1.
(C = 40UB, B = 7, μ_{nrt} = 2, $r$min = 0.8UB, λ_{nrt} = 55)

Figure 5.6 plotted a graph of P(r < $r$min) against *Nnrt* for different size of partitioned real time connection, C1. The data of this graph is in appendix C5. By using equation (5.10) with the above assumptions, it is found that when *C*1 is large, the P (r < $r$min) is higher. This is because the bandwidth left for non real-time connection is smaller to cater a high number of connections and in the end the effective bandwidth given will not meet the required QoS parameters such as cell loss. However when C1 is less i.e. 25 UB, the probability that a non real-time connection will be given effective bandwidth of less rmin bandwidth was decreased. This is because when a generous amount of bandwidth remained for non real-time connection, a non real-time connection might be given 1-0.8UB and this will guarantee the QoS requirements.

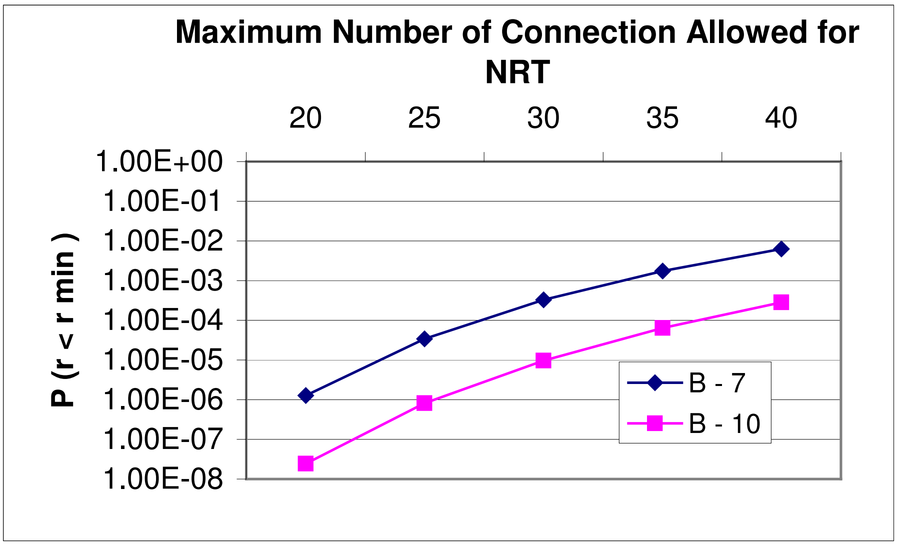


Figure 5.7: Probability that the bandwidth is less than the minimum rate, P (*r* < *r*min) against maximum number of connection allowed to a network for different size of cell cluster.
(C = 40UB, B = 7, μ_{nrt} = 2, $r$min = 0.8UB, λ_{nrt} = 55, $C1$ = 30)

Figure 5.7 the P (r < $r$min) against *Nnrt* for different size of cell cluster, B. The data of this graph can be found in appendix C6. The arrival rate is fixed at 55 calls per seconds and the amount of bandwidth allocated for real-time connections is 30UB. Using the above assumptions in equation (5.10), it is found that the size of cell cluster affects the P (r < $r$min) where the higher the number of base stations in a cell cluster, the higher the total bandwidth in the network. This will increase statistical multiplexing and hence the probability for a connection to get less than $r$min bandwidth is lowered. This is because when more base station are grouped into a cell cluster the capacities of the VCT is higher and this allows the network to cater for more connections.

5.4 Complete Access (CA)

In this scheme, the capacity of the network C will be shared among real-time and non real-time connections, $C1 = C$ and it is illustrated in figure 5.8. This means that real-time connections can use up to the total capacity of the network if the bandwidths are idle. At any particular time, the bandwidths that are not used by the real-time connections will be shared equally among the non real-time connections. The steady state probability of rate r being available to a non real-time connection can be found by summing up all the probabilities that i and s number of connections is connected to the network and divides it with the probability all of the bandwidth left for non real-time connection is used up.

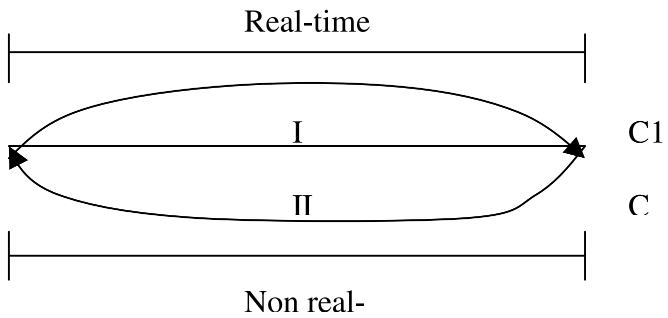


Figure 5.8: Complete Access

The probability of rate r being available to a non real-time connection, $P(r)$ can be written as:

$$P(r) = \frac{1}{\sum_{k=0}^{Nnrt} kP_{nrt}(k)} \sum_{i,s \in A_s} P_{rt}(i) s P_{nrt}(s) \qquad (5.11)$$

where $P_{rt}(i)$ is in equation (3.11) and $P_{nrt}(s)$ is in equation (3.22). s = 0, 1, 2…and i = 0, 1, 2…..

A_s defined all combination of i and s such that:

$$r = \frac{C - iBW1}{s} \qquad (5.12)$$

where $0 \le i \le C/BW1$, $0 \le s \le Nnrt$ and r is the bandwidth rate shared equally amount the non real-time connection. Hence, the probability that the available bandwidth rate to a non real-time connection is smaller than $r$min threshold is:

$$P(r < r\min) = \sum_{r \in A_r, r < r\min} P(r) \qquad (5.13)$$

where A_r is all the possible combinations of $\underline{i}$ and s.

5.4.1 The Effect Of Erlang Load Per Cell Site

In complete access, the QoS for non real-time connection will depend on the real-time connection because the whole capacity is shared between both types of connections. Hence, in this case we have to consider both classes in order to find the QoS of a non real-time connection. As mentioned earlier, the non real-time connections can use the available bandwidth left by the real-time connections and this bandwidth will be shared equally among the accepted non real-time connections.

Figure 5.9 shows the probability that the bandwidth is less than $r$min, P ($r < r$min) against Erlang load per cell site for different size of $C1$. Assumed that the total network capacity is 40 UB. For C1- 15, the real-time connections have used up to 15 UB of the total capacity and the network allows only 30 number of non real time connections. For C1- 30, the real time connections have used up to 30 UB of the total amount of network capacity and allow only 15 number of non real-time connection to the network.

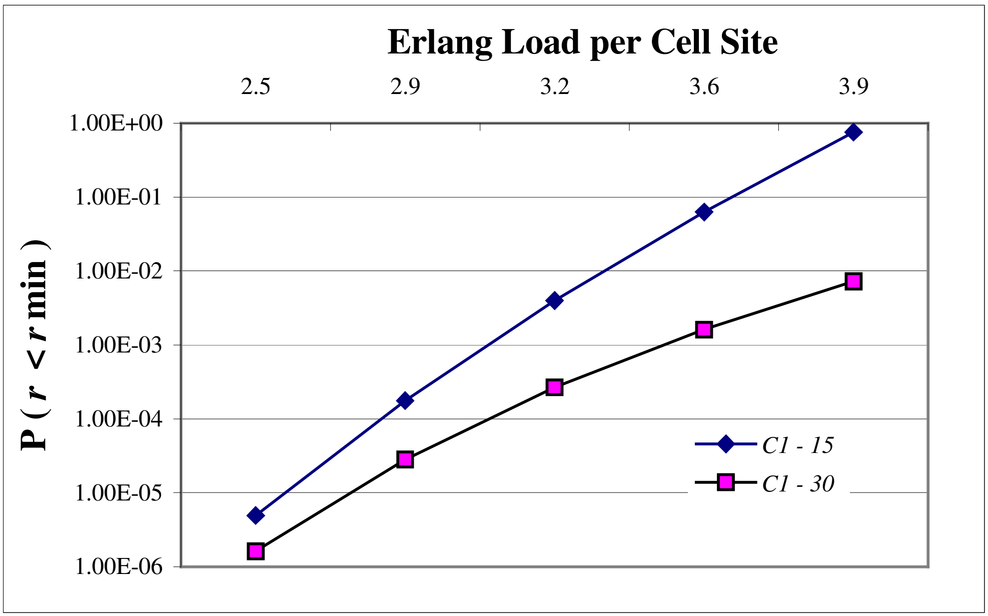


Figure 5.9: Probability that the bandwidth is less than *r*min, P (*r* < *r*min) against Erlang load per cell site for different size of *C*1.
(C = 40, Nnrt = 30, μ_{rt} = 2, B = 7, $r$min = 0.8)

For both situations the call arrival varies from 35 to 55 calls per seconds. The calls depart from the network exponentially at 2 calls per seconds. The threshold of minimum bandwidth is set to 0.8 UB. Using the above assumptions in equation (5.13), it is found in figure 5.9 that when more traffic arrives at the network the P $(r < r\text{min})$ is higher. This is due to limited bandwidth left for non real time connection to share among the accepted non real time connections. The data of this graph can be found in appendix C7. Further, when more bandwidths are used by the real-time connections, fewer bandwidths are left for non real-time connection and this will decrease the probability for a connection to get less than the threshold bandwidth. If more sources are utilized by the real-time connection then the effective bandwidth assigned to the non real-time connection will approach to 64kb/s (approx. 1UB). No statistical multiplexing takes place.

5.4.2 The Effect of Bandwidth Allocated to Real time Connection

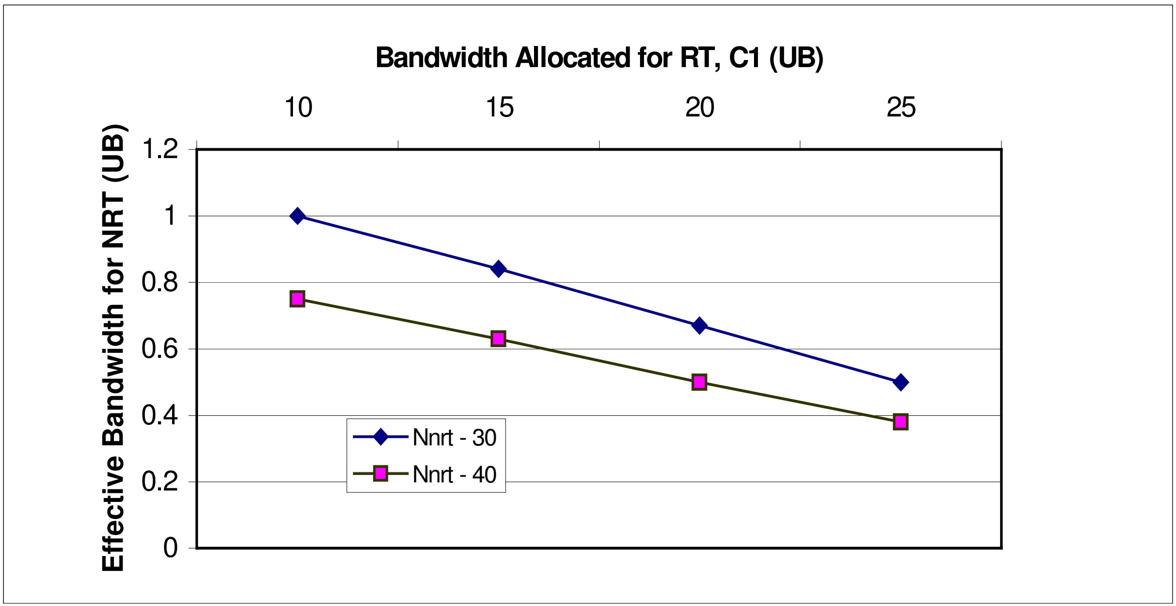


Figure 5.10: Bandwidth allocated for real-time connections against effective bandwidth available per non real-time connections. (*C* = 40, *C*1 = 15, *B* = 7, *r*min = 0.8)

The parameters or assumptions used in equation (5.13) are as follows. The call arrival rates vary from 35 to 55 calls per seconds and the call depart at 2 calls per seconds. The total capacity of the network is 40 UB. In the analysis, bandwidth used by the real time connections was varied from 10 to 25 UB with the maximum number of non real-time connection allowed to the network is fixed to 30 and 40. The minimum bandwidth rate, *r*min is set to 0.8 UB. Figure 5.10 show that as C1 increases, the effective bandwidth available for each non real time connection decreases. The data of this graph can be found in appendix C8. This is because the bandwidth left for non real time connection to share among the accepted connections is limited. Thus, the given effective bandwidth will be lesser but if less Nnrt is set, then each non real-time connection will be given better effective bandwidth. This is because the limit of Nnrt decides on the effective bandwidth, which will be allocated by the admission control.

5.5 The Effect of Erlang Load to CP and CA

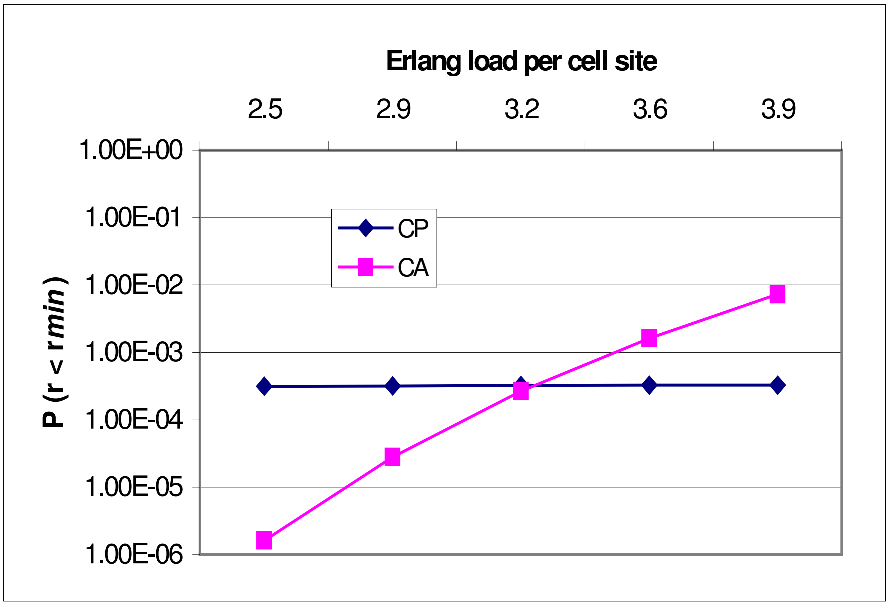


Figure 5.11: Probability that the bandwidth is less than *r*min, P (*r* < *r*min) against Erlang load per cell site for CA and CP
(C = 40, $Nnrt$ = 30, $r$min = 0.8, B = 7, $C1$ = 30, μ_{nrt} = 2)

Figure 5.11 shows probability that the bandwidth is less than the minimum rate, P (*r* < *r*min) against Erlang load per cell site of real-time connections. It compares access mechanism complete partitioning (CA) and complete access (CP). The data of this graph is in appendix C9. Using the assumptions stated below figure 5.11 in equation (5.10) and (5.13), it is found that access mechanism CP maintained at almost the same value of P (*r* < *r*min) as the Erlang load increases. CP increases with Erlang load and cross CP at Erlang load 3.2. Hence, it can be conclude that CA can be used at load below Erlang load per cell site of 3.2 and CP is better at Erlang load per cell site of 3.2 and above.

5.6 The Effect of Maximum Number of Connection Allowed to CP and CA

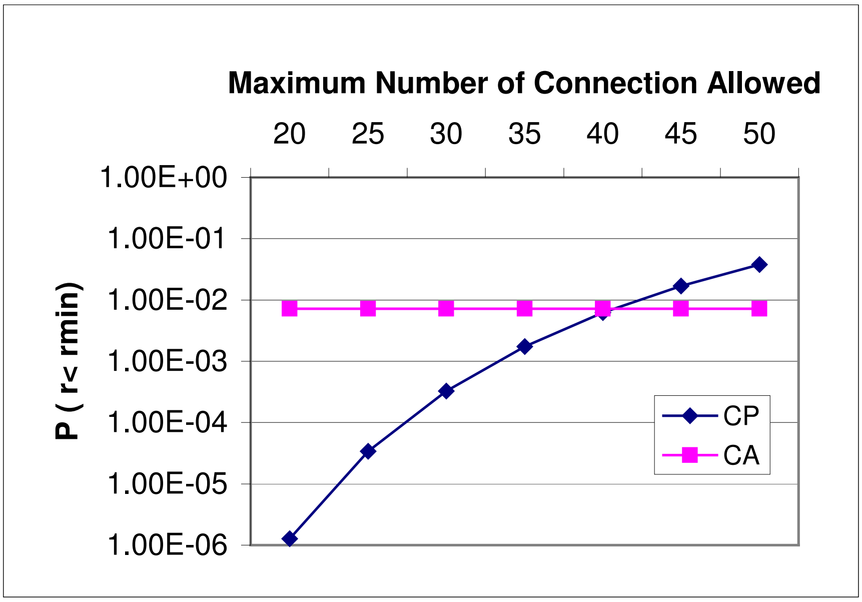


Figure 5.12: Probability that the bandwidth is less than rmin, P (r < $r$min) against maximum number of non real-time connection allowed for non real-time connections
(C = 40, $r$min = 0.8, B = 7, C1 = 30, μ_{nrt} = 2, λ_{nrt} = 55)

Figure 5.12 shows another comparison between access mechanism CA and CP. The data of this graph can be found in appendix C10. Figure 5.12 is for P(r < $r$min) against $Nnrt$. An interesting result is obtained where using access mechanism CP, it maintains at almost the same value of P (r < $r$min) despite of the increment of $Nnrt$. From the result, it is found that when the maximum number of connection allowed into the network is below than 40, access mechanism CP gives a better P (r < $r$min) than CA but at $Nnrt$ equals to 40 and above, CA can give better P (r < $r$min). This also shows that the interception of P (r < $r$min) for CP and P (r < $r$min) for CA depends on the amount of bandwidth partitioned for non real-time connection, which in this case is [C-C1=10UB]. Hence if C1 is smaller the interception will be more than $Nnrt$ 40. This is because in CP, the amount of bandwidth partitioned for non real time connection is fixed and if more connection are allowed into the network the system can not give a bandwidth that will satisfied the network QoS requirements.

5.7 Summary

The results show that the P (r < $r$min) of is a network were to use complete partitioning mechanism, CP it will be affected by the Erlang load, but will increase with the increment of Nnrt. This is because the amount of bandwidth for non real-time connection in CP is fixed and with more connection allowed into the network, a connection may not be given the bandwidth that will be acceptable to the QoS requirements in terms of probability that the bandwidth is less than the minimum rate, P(r < $r$min). However, the amount of bandwidth partitioned to the non real time connection is important to get low P(r < $r$min). Moreover, when the Erlang load into a network is heavy, the network should not accept more connection than when the Erlang load is normal, it all depend on the threshold maximum number of connections set by the network.

The P(r < $r$min) for complete access, CA, is very much affected by the Erlang load because the real-time connections are given priority to use the network bandwidth and the other idle bandwidth, which is not use by the real-time connection, will be given to the non real-time connections. The call departure rate had proved to decrease the P(r < $r$min) for CA, this is because when more connection terminates, other calls can be accepted to use the idle channels. In comparison of CA and CP access mechanism, it shows that CA can be use in cases where traffic are average or below 40 call per seconds but CP will maintain at a P (r < $r$min) value despite of the traffic load. However, CA is preferred than CP when more number of non real-time connections is allowed into a network. It can be conclude that CA and CP are preferred in certain cases depending on the traffic load and the maximum number of connection threshold.

CHAPTER VI

CONCLUSION AND RECOMMENDATIONS

6.1 CONCLUSION

Handoff priority scheme is essential in a mobile network to ensure connection continuity whenever a mobile device moves within or out of the network. In wireless ATM (WATM) network, call admission control (CAC) limits the incoming call in order to avoid congestion and unnecessary call termination. By using Virtual Connection Tree (VCT) model, an accepted connection will enjoy freedom of mobility under one Network Call Processor (NCP) where by handoff calls are guaranteed a good Quality-of-Service (QoS). When handoff calls are considered as a higher priority than the incoming calls, a scheme is needed to protect the handoff calls. The new call blocking probability and handoff dropping probability determine QoS performance provided by the WATM network. New call blocking probability determines QoS performance of the network, that it can provide the service whenever a mobile terminal (MT) request for a connection. Handoff dropping probability determines QoS performance of the network, that it can guarantee a smooth connection during the connection lifetime despite of the changing access point in the VCT by MTs. The overload probability is the QoS performance of the network, that the non real-time applications will not get thorough congested network.

The purpose of this project is to find a system that can serve better QoS performance in WATM network for real-time and non real-time connections. The QoS performance concerned is to maintain low probability of handoff dropping, low probability of new call blocking and low probability of overload. The method of finding the right system was first started with reserving a certain percentage of channels in the network for handoff calls in order to increase the continuity of handoff calls as users move from one cell site to another. The system is call system with reserved channels WRC system. The efficiency of the system in handling handoff call and new incoming calls was highlighted when compared to system without reserved channels, WORC system. A low probability of handoff dropping was able to achieved using WRC system. Further enhancement of WRC system depends very much on factors like size of the cell cluster, size of the reserved channels and the capacity a base station can handle. Despite of the improvement WRC system had obtained in contrast to WORC system, a few set backs was found. With the decrement of handoff dropping probability, the new call block probability had increased. Hence, further improvement was proposed by allowing queue for the new incoming calls in the system with reserved channels. The proposed system is call system with queuing of the new incoming calls and with reserved channels, WQRC system. With this system, it decreases the new call blockage at the network by the CAC when all channels in the network are used. Factors that provide better QoS performance using WQRC system are the size of the cell cluster, size of the reserved channels and the capacity a base station can handle. A set back obtained from the improvement of WRC system is the increment in probability of delay.

The probability of new call block can be reduced with the increment of base station in a cell cluster and expanding the capacity a base station can handle. This means that the VCT has the advantage to reduce initial call rejection. Even though a system has reserved channel but if is not use in VCT network the probability of new call lost would not reach as low as when a system with reserved channels is used in VCT networking. However, the new call block probability increase with the increment of reserve channels. Hence, to control the rapid growth of new call block probability when reserve channels increase, a certain amount of reserved channels has to be defined in order to maintain an acceptable increment of new call lost. The reduction of handoff dropping probability can be attained if reserved channels and cell cluster is large. VCT networks decreases the new call block probability as well as handoff dropping probability by grouping base station which directly enlarge the network capacity. Hence, when network capacity is large, the percentage for reserved channels can be increase. For non real-time homogeneous system the QoS performance is based on the overload probability. Maximum number of non real-time connection allowed into a network, *Nnrt* and the capacity assigned for base stations contribute to low probability of overload. Hence, CAC plays an important role in homogenous system of non real-time connection for WATM. In addition, VCT network decreases the overload probability of non real-time connection. The WRC was simulated using COMNET III and it was found comparable with the theory. The low probability of blocking by simulating shown that practical implementation of VCT can provide better QoS.

With WQRC system the handoff dropping probability as well as the new call block probability are reduce by using reserved channel for handoff calls and allowing queue for new incoming calls. Hence, this system has upgraded the previous system by avoiding rapid increasing of new call block probability when handoff dropping probability declining. The only set back obtained from the improved system is the delay where the probability that a call will delay increase with the decrement handoff call dropping. Such as the VCT network had help to reduce the handoff dropping probability, new call block probability and overload probability, here it too somehow reduce the rapid growth of delay probability in the queue. However, the delay will only affect the call set up protocol and not the data transmission.

A good WATM system would not be complete without an access mechanism for both real-time and non real-time connections. The effective bandwidth is one of the advantage features of ATM technology. Since ATM supports multi traffic, the numbers of real-time and non real-time connections that will be accepted to a network vary with the amount of bandwidth given to each application. Also, burstiness plays an important role in allowing more non real-time connection into a network. The number of non real-time connection depends very much on the capacity of a base station, amount of bandwidth assigned for real-time connections and burstinest of the connection. The two types of access mechanism concerned are the Complete Partitioning (CP) and Complete Access (CA). Using both access mechanisms the network will set a threshold minimum bandwidth rate so that the QoS requirements will not be violated. The type of access mechanism a network should be using will depend on the pattern of traffic it handles and the network threshold. CP is preferred for network that has to handle heavy non real-time traffic at most of the time because the increment of non real-time traffic load has little affect on the probability a connection will get less than the minimum bandwidth rate, $rmin$, P(r < $rmin$). The advantage of CP to real-time and non real-time connection is that the amount of bandwidth is partitioning prior to connection acceptance. Hence, both types of application are given space for connections. However when more real-time traffic arrives, the space are limited even though the partitioned part of non real-time connection is idle. CA is preferred for network that has to allowed a large number of non real-time connection. Or in other words, the change in non real-time connection has a little affect on the P ($r < rmin$) for CA. This is because of the amount of bandwidth assigned for non real-time connections are fixed but depends on the amount of bandwidth given to real-time connection. Real-time connection has the most advantage using this access mechanism as any number of connections accepted can be connected. In terms loss for real time connections, CA will give less probability because real-time connection has the priority to use the base station capacity until the maximum of the total base station capacity compare to CP where real-time connection has limited access.

The proposed WQRC system with either CA or CP access mechanism can upgrade QoS performance of WATM network using VCT model. With the proposed system handoff dropping probabilities maintained at low value and the new call block probabilities can be decrease. However, this can be improved even better with some recommendation, which will be discussed in the next section.

6.2 RECOMMENDATION

This work had shown that WQRC system with CP/ CA could enhance the VCT network model for WATM network. However, there is a lot of room for improvement for this project. Some of these recommendations include:

i) In this project the queue of the new incoming call is considered infinite. Finite queue can be obtain if the column of the two dimensional state diagram in figure 4.3 is limit to a certain figure. Hence the queue of the new incoming call is limited and the probability of new call block when queue is full can be obtained. In addition the average queue and delay time can be obtained.

ii) The effective bandwidth for the system can be calculated using other ways such as

iii) Comparing the QoS performance of the proposed system using VCT model with other models such as incremental extension, BAHAMA or two-tier architecture.

iv) The whole system can be verified with further analysis using discrete event simulation.

REFERENCES

1. Walke B., Petras D., Plassman D. (1996). "Wireless ATM: Air Interface and Network Protocols of the Mobile Broadband System". IEEE Personal Communication Magazine. August.
2. Bing K.T. (1998). "ATM for Wireless Network". ICATM 98'.
3. Raychaudhury D. (1996). "Wireless ATM Network: Architecture, System Design and Prototyping". IEEE Personal Magazine. August.
4. Shafi M. et. al. (1997). "Wireless Communication in the twenty First Century: A perspective". Proceedings of the IEEE. Vol. 85. No. 10. October.
5. Swain, R. S. "UMTS- A 21st Century System: A RACE Mobile Project Line Assembly Vision". http://www.vtt.fi/tte/nh/UMTS/umts. Html.
6. Acampora A. (1996). "Wireless ATM: A perspective on Issues and Prospects". IEEE Personal Communications. August.
7. Ayanoglu E, Eng K. Y and Karol M. J. (1996). "Wireless ATM: Limits, Challenges and Proposals". IEEE Personal Communication. August.
8. Rahnema M. (1993). "Overview of GSM System and Protocol Architecture". IEEE Communication Magazine. April.
9. Tekinay S. and Jabbari B. (1991). "Handover and Channel Assignment in Mobile Cellular Networks". IEEE Communication Magazine. November.
10. Lee, W. C. Y. (1991). "Smaller Cells for Greater Performance" IEEE Communication Magazine. November.
11. Handel R., Huber M. N. and Schreder S. (1998). "ATM Networks Concepts, Protocols and Applications". Addison Wesley, Harlow, England.
12. Kuehn P.J. (1989). "From ISDN to IBCN (Integrated Broadband Communication Network)." In Proceeding World Computer Congress IFIP.pg 476-486
13. Stallings W. (1996). "Data and Computer Communications". New Jersey. Prentice Hall.
14. Coover E. R. (1997). "ATM Switches". Artech House, London.
15. Schuler C. (1997). "A Quality of Service Concept for Wireless ATM". IEEE ATM Workshop Proceedings.

16. Acampora A. and Naghshineh (1994). “ Control and Quality-of-Service Provisioning in High Speed Microcellular Networks.IEEE Personal Magazine. Second Quarter.
17. Clark, M. P. (1996). “ATM Networks and Principles and Use”. New York. Wiley Tuebner
18. Black, U. (1995). “ATM: Foundation for Broadband Network”, New Jersey, Prentice Hall Series in Advanced Communications Technologies
19. Marsan M. A. et. al. (1997). “ Local and Global Handovers for Mobility Management in Wireless ATM Network” IEEE Personal Communication. October.
20. Eng, K. Y. et al. (1995). “ BAHAMA: A Broadband Ad-Hoc Wireless ATM Local Area Network” Proceedings ICC .
21. Acampora, A. S. and Nagshineh, M. (1994). “An Architecture and Methodology for Mobile Executed Handoff in Cellular ATM Network” IEEE Journal Selected Area in Communication. Vol. 12. No. 8. October.
22. Agrawal, P. et. al. (1996). " SWAN: A mobile Multimedia Wireless network" IEEE Personal Communication. April.
23. Sang-Ho, Lee and Sung, Woo Park (1998). " Handoff with Dynamic Guard Channels in ATM-Based Mobile Networks" Information Networking Proceedings (ICOIN-12).
24. Guerin, R. (1988). “Queueing-Blocking System with Two Arrival Streams and Guard Channels” IEEE Transaction on Communication. Vol. 36. No. 2. February.
25. Tekinay, S. and Jabbari, B. (1992). “ A Measurement-based Prioritization Scheme for Handovers in Mobile Cellular Networks” IEEE Journal on Selected Areas in Communication. Vol. 10. No. 8. October.
26. Gaasvik, P., Cornefjord, M. and Svenson, V. (1991). "Different Methods of Giving Priority to Handoff Traffic in a Mobile Telephone System with Directed Retry" Vehicular Technology Conference IEEE.
27. ATM User-Network Interface (UNI) Signalling Specification. http://www.atmforum.com/atmforum/spec/approved.html.
28. R.Yuan, S.K Biswas and D. Raychaudhuri. (1995). An Architecture for Mobility Support in an ATM Network. Proc. 5th WINLAB Worshop on Third Generation Wireless Network. April.
29. Kleinrock, L. (1975). “Queuieng System, Volume I: Theory” New York. John Wiley.

Appendix A1

Table of Quality-of-Service performance against Erlang load per cell site for WRC system and WORC system

Erlang load per cell site	P_{HD}	P_{AB}	$P_{AB\ (WORC)}$
0.71	1.91e-7	0	2.99e-5
1.43	1.53e-3	9.83e-9	9.80e-3
2.14	3.80e-2	1.20e-2	8.60e-2
2.86	7.20e-2	1.64e-1	2.09e-1
3.57	7.90e-2	3.18e-1	3.23e-1
4.29	8.20e-2	4.28e-1	4.14e-1
5.00	8.30e-2	5.09e-1	4.86e-1
5.71	8.30e-2	5.69e-1	5.43e-1
6.42	8.300e-2	6.170e-1	5.89e-9

P_{HD} - probability of handoff dropping
P_{AB} - probability of call block
$P_{AB\ (WORC)}$ - probability of both call block and handoff dropping for system without reserved channels

Call arrival rate, λ_{rt}, = 5 - 45
Handoff rate, γ_{rt} , = 2 per seconds
Departure rate, μ_{rt} = 1 per seconds
Number of base stations, B = 7
Channels for received calls, CRC = 0.85
Number of connection that a base station can support, m = 20

Appendix A2

Table of probability of call block against Erlang load per cell site for different capacity of base station

Erlang load per cell site	***m* =10**	***m* = 15**	***m* = 20**
1.43	1.79e-1	2.05e-3	9.83e-9
1.57	2.41e-1	1.20e-2	8.65e-7
1.71	2.97e-1	3.60e-2	2.85e-5
1.86	3.97e-1	7.50e-2	4.08e-4
2.00	3.91e-1	1.20e-1	2.91e-3
2.14	4.29e-1	1.66e-1	1.20e-2

m - number of connection that a base station can support

Call arrival rate, λ_{rt}, = 50 - 75
Handoff rate, γ_{rt} , = 10 per seconds
Departure rate, μ_{rt} = 5 per seconds
Number of base stations, B = 7
Channels for received calls, CRC = 0.85

Appendix A3

Table of probability of handoff dropping against Erlang load per cell site for different capacity of base stations

Erlang load per cell site	***m* =10**	***m* = 15**	***m* = 20**
1.43	1.24e-1	3.20e-2	1.50e-3
1.57	1.30e-1	5.10e-2	3.90e-3
1.71	1.34e-1	6.70e-2	8.32e-3
1.86	1.36e-1	8.00e-2	1.60e-2
2.00	1.38e-1	8.80e-2	2.60e-2
2.14	1.40e-1	9.40e-2	3.80e-2

m - number of connection that a base station can support

Call arrival rate, λ_{rt}, = 50 - 75
Handoff rate, γ_{rt} , = 10 per seconds
Departure rate, μ_{rt} = 5 per seconds
Number of base stations, B = 7
Channels for received calls, CRC = 0.85

Appendix A4

Table of probability of call block against Erlang load for different size of cell cluster

Erlang load per cell site	$B = 5$	$B = 7$
1.43	8.76e-7	9.83e-9
1.57	2.14e-5	8.68e-7
1.71	2.06e-4	2.85e-5
1.86	1.75e-3	4.08e-4
2.00	7.22e-3	2.91e-3
2.14	2.00e-2	1.20e-2

B - Number of base stations in a cell cluster

Call arrival rate, λ_{rt}, = 50 - 75
Handoff rate, γ_{rt} , = 10 per seconds
Departure rate, μ_{rt} = 5 per seconds
Channels for received calls, CRC = 0.85
Number of connection that a base station can support, m = 20

Appendix A5

Table of probability of handoff dropping against Erlang load for different size of cell cluster

Erlang load per cell site	***B* = 5**	***B* = 7**
1.43	1.53e-3	1.53e-3
1.57	3.87e-3	3.87e-3
1.71	8.31e-3	8.32e-3
1.86	1.50e-2	1.60e-2
2.00	2.50e-2	2.60e-2
2.14	3.60e-2	3.80e-2

B - Number of base stations in a cell cluster

Call arrival rate, λ_{rt}, = 50 - 75
Handoff rate, γ_{rt} , = 10 per seconds
Departure rate, μ_{rt} = 5 per seconds
Channels for received calls, CRC = 0.85
Number of connection that a base station can support, m = 20

Appendix A6

Table of probability of new call block against reserved channels

Reserved channels	$\boldsymbol{\lambda_{rt} = 50}$	$\boldsymbol{\lambda_{rt} = 60}$	$\boldsymbol{\lambda_{rt} = 70}$
10	1.66e-10	1.72e-6	5.13e-4
15	9.82e-9	2.85e-5	2.90e-3
20	3.87e-7	3.13e-4	1.10e-2
25	9.86e-6	2.24e-3	3.10e-2
30	1.57e-4	1.00e-2	6.50e-2
35	1.54e-3	3.2e-2	1.12e-1
40	9.00e-3	7.1e-2	1.68e-1

λ_{rt} - Call arrival rate

Number of base stations in a cell cluster, $B = 7$
Handoff rate, γ_{rt} , = 10 per seconds
Departure rate, $\mu_{rt} = 5$ per seconds
Number of connection that a base station can support, $m = 20$
Channels of received channels, CRC = 0.6 - 0.9

Appendix A7

Table of probability of handoff dropping against reserved channels

Reserved channels	λ_{rt} **= 50**	λ_{rt} **= 60**	λ_{rt} **= 70**
10	1.53e-3	8.31e-3	2.60e-2
15	1.53e-3	8.31e-3	2.60e-2
20	1.53e-3	8.31e-3	2.40e-2
25	1.53e-3	8.17e-3	2.10e-2
30	1.53e-3	7.63e-3	1.60e-2
35	1.51e-3	6.31e-3	1.10e-2
40	1.39e-3	4.37e-3	6.48e-3

λ_{rt} - Call arrival rate

Number of base stations in a cell cluster, B = 7
Handoff rate, γ_{rt} , = 10 per seconds
Departure rate, μ_{rt} = 5 per seconds
Number of connection that a base station can support, m = 20
Channels of received channels, CRC = 0.6 - 0.9

Appendix A8

Table of Probability of new call block against Erlang load per cell site for theoretical and simulation method.

Erlang load per cell site	Theoretical	Simulation
0.5	1.00e-4	3.70e-4
0.620	2.10e-3	5.98e-3
0.750	1.60e-2	2.71e-2
0.875	5.20e-2	8.80e-2
1.000	1.09e-1	1.30e-1
1.125	1.72e-1	1.80e-1
1.250	2.35e-1	2.25e-1
1.375	2.92e-1	2.64e-1
1.500	3.43e-1	0.28e-1

Call arrival rate, λ_{rt} = 40 - 75 per seconds (s)
Call departure, μ_{rt} = 2.5 per seconds
Hence, duration of a call is $\left(\frac{1}{2.5}\right)$ seconds = 0.4 s

$$60s - 1\,\text{min}$$

$$0.4s - 1x\left(\frac{0.4}{60}\right) = 0.007\,\text{min}$$

$\therefore$ 150 calls per minute
Number of channels that a base station can support, m = 32
Number of base station, B = 1

Appendix A9

Table of Probability of new call block against Erlang load per cell site for theoretical and simulation method.

Load / B/ m	Theoretical	Simulation
0.54	2.26e-4	3.21e-1
0.58	8.10e-4	3.71e-1
0.63	2.70e-3	4.14e-1
0.67	3.40e-3	4.52e-1

B - Number of base station

Call arrival rate, λ_{rt} = 120 -150 per seconds (s)
Call departure, μ_{rt} = 2.5 per seconds
Hence, duration of a call is $\left(\frac{1}{2.5}\right)$ seconds = 0.4 s

$60s - 1\,\text{min}$

$0.4s - 1x\left(\frac{0.4}{60}\right) = 0.007\,\text{min}$

$\therefore 150$ calls per minute
Number of channels that a base station can support, m = 32

Appendix A10

Table of overload probability against Erlang load for different number of maximum number allowed to a base station

Erlang load	***Nnrt* - 70**	***Nnrt* - 80**
10	1.58e-3	2.14e-4
15	2.52e-3	4.18e-4
20	2.98e-3	4.76e-4
25	3.16e-3	5.01e-4
30	3.26e-3	5.16e-4
35	3.32e-3	5.25e-4

Nnrt - number of maximum connection allowed to a network.

Number of base station, $B = 7$
Total capacity in a base station, $C = 20$ UB
Call departure rate, $\mu_{nrt} = 1$

Appendix A11

Table of overload probability against Erlang load for different number of maximum number non real-time connection allowed to a base station

Erlang load	***Nnrt* - 70**	***Nnrt* - 80**
10	6.84e-9	3.28e-10
15	3.54e-8	8.41e-10
20	4.63e-8	1.03e-9
25	5.12e-8	1.13e-9
30	5.40e-8	1.18e-9
35	5.57e-8	1.22e-9

Nnrt - maximum number of non real-time connection allowed to a network.

Number of base station, $B = 7$
Total capacity in a base station, $C = 30$ UB
Call departure rate, $\mu_{nrt} = 1$

Appendix A12

Table of overload probability against maximum number of non real- time connection allowed

Nnrt	***B* - 5**	***B* - 7**	***B* - 15**
60	4.26e-3	4.47e-5	1.19e-10
65	1.20e-2	1.66e-4	6.51e-10
70	2.70e-2	5.16e-4	2.98e-9
75	5.40e-2	1.38e-3	1.18e-8
80	9.60e-2	3.26e-3	4.08e-8

Nnrt - maximum number of non real-time connection allowed to a network.

Number of base station, $B = 7$
Total capacity in a base station, $C = 20$ UB
Call departure rate, $\mu_{nrt} = 1$
Call arrival rate, $\lambda_{nrt} = 30$

Appendix A13

Table of overload probability against Erlang load for system with control mechanism and system without controlled mechanism

Erlang load	System with control mechanism	System without control mechanism
10	6.84e-9	7.98e-8
15	3.54e-8	1.97e-4
20	4.63e-8	1.30e-2
25	5.12e-8	1.37e-1
30	5.40e-8	4.52e-1
35	5.57e-8	7.73e-1
40	5.69e-8	9.38e-1

Number of base station, $B = 7$
Total capacity in a base station, $C = 30$ UB
Call departure rate, $\mu_{nrt} = 1$
Maximum number of non real-time connection allowed to a network, $Nnrt = 80$

Appendix B1

Table of probability of handoff dropping against reserved channels for WRC system and WQRC system (*B*=7)

Reserved Channels (%)	WRC system	WQRC system
15	1.53E-03	3.41E-13
20	1.53E-03	3.39E-16
25	1.53E-03	2.33E-19
30	1.53E-03	7.08E-22
35	1.39E-03	3.58E-24

Call arrival rate, λ_{II} = 50 calls per seconds (s)
Call departure rate, μ_{II} = 5 calls per seconds
Handoff call rate, γ_{II} = 10 calls per seconds
Number of channels that a base station can support, m_{II} = 20
Number of base station, B_{II} = 7

Appendix B2

Table of probability of handoff dropping against reserved channels for WRC system and WQRC system.(B = 4)

Reserved Channels (%)	WQRC system	WRC system
15	2.41e-04	2.10e-02
20	2.39e-07	1.97e-02
25	3.33e-10	1.71e-02
30	8.08e-13	1.34e-2
35	3.58e-15	9.20e-03

Call arrival rate, λ_{II} = 50 calls per seconds (s)
Call departure rate, μ_{II} = 10 calls per seconds
Handoff call rate, γ_{II} = 20 calls per seconds
Number of channels that a base station can support, m_{II} = 20
Number of base station, B_{II} = 4

Appendix B3

Table of probability of handoff dropping against Erlang load per cell site for WRC system and WQRC system (B = 4).

Erlang load per cell site	WQRC system	WRC system
40	5.16e-07	1.34e-02
45	4.99e-06	1.56e-01
50	3.59e-5	1.67e-01
55	1.99e-04	1.74e-01
60	884e-04	1.77e-01

Call arrival rate, λ_{II} = 40 - 60 calls per seconds (s)
Call departure rate, μ_{II} = 10 calls per seconds
Handoff call rate, γ_{II} = 20 calls per seconds
Number of channels that a base station can support, m_{II} = 20
Number of base station, B_{II} = 4
Channels for received calls, CRC = 0.85

Appendix B4

Table of probability of handoff dropping and probability of delay against reserved channels

Reserved Channels (%)	PH	PD
5	2.90e-10	2.70e-07
10	6.02e-12	3.45e-06
15	1.26e-13	3.59e-05
20	2.69e-15	3.00e-04

PH - probability of handoff dropping
PD - probability of delay

Call arrival rate, λ_{II} = 50 calls per seconds (s)
Call departure rate, μ_{II} = 10 calls per seconds
Handoff call rate, γ_{II} = 20 calls per seconds
Number of channels that a base station can support, m_{II} = 20
Number of base station, B_{II} = 4

Appendix B5

Table of probability of handoff dropping and probability of delay against Erlang load per cell site

Erlang load per cell site	PH	PD
40	1.82e-15	5.16e-07
45	1.76e-14	4.99e-06
50	1.26e-13	3.59e-05
55	7.03e-13	1.99e-04
60	3.11e-12	8.84e-04

PH - probability of handoff dropping
PD - probability of delay

Call arrival rate, λ_{II} = 40 - 60 calls per seconds (s)
Call departure rate, μ_{II} = 10 calls per seconds
Handoff call rate, γ_{II} = 20 calls per seconds
Number of channels that a base station can support, m_{II} = 20
Number of base station, B_{II} = 4

Appendix B5

Table of probability of delay against Erlang load per cell site

Erlang load per cell site	m = 25	m = 20
40	6.67e-13	5.16e-07
45	2.03e-11	4.99e-06
50	4.24e-10	3.59e-05
55	6.38e-09	1.99e-04
60	7.16e-08	8.84e-04

m - Number of channels that a base station can support

Call arrival rate, λ_{II} = 40 - 60 calls per seconds (s)
Call departure rate, μ_{II} = 10 calls per seconds
Handoff call rate, γ_{II} = 20 calls per seconds
Number of base station, B_{II} = 4
Channels for received calls, CRC = 0.85

Appendix C1

Table of the effective bandwidth against sources utilized by real-time connections

Sources Utilized by RT	***C* -32 UB**	***C* - 125 UB**	***C* - 531 UB**
0.05	7.14	4.73	4
0.1	18	9.4	8
0.15	25.67	13.8	11.9
0.2	33.09	18.3	14.8
0.25	39.9	22.5	19.8
0.3	46.2	26.76	23.6
0.35	51.9	31.8	27.5
0.4	57.1	34.8	31.3
0.45	61.63	38.7	35.1

NRT - non real-time connection
RT- real time connection

Probability of cell loss, $Ploss = 1 \times 10^{-9}$
Peak bit rate , $B_p = 64$kb/s

Appendix C2

Table of the number of non real-time connection against bandwidth of real-time connections

Number of connection for NRT	Burstiness -5	Burstiness - 10	Burstiness -20
0	60	111	222
0.25	48	92	175
0.5	38	71	136
0.75	28	50	104
1.0	19	37	67
1.25	10	16	50
1.5	6	6	10
1.75	3	3	3
2	0	0	0

NRT - non real-time connection
RT- real time connection

Probability of cell loss, *Ploss* = 1×10^{-9}
Peak bit rate , B_p = 64kb/s
Mean bit rate, B_m = 6.5kb/s

Appendix C3

Table of Probability of bandwidth rate is less than *r*min P (*r* < *r*min) against Erlang load per cell site.(CP)

Erlang load per cell site	*Nnrt* - 30	*Nnrt* - 40	*Nnrt* - 50
2.5	3.17e-4	6.00e-3	3.70e-2
2.9	3.19e-4	6.10e-3	3.70e-2
3.2	3.22e-4	6.20e-3	3.80e-2
3.6	3.25e-4	6.21e-3	3.80e-2
3.9	3.27e-4	6.25e-3	3.80e-2

Nnrt - maximum number of non real-time connections allowed

Call arrival rate, λ_{nrt} = 35 - 55 calls per seconds (s)
Call departure rate, μ_{nrt} = 2 calls per seconds
Number of base station, B = 7
Minimum bandwidth rate, $r$min = 0.8 UB
Amount of bandwidth for real-time connection, $C1$ = 30
Total amount of bandwidth for the base station, C = 40 UB

Appendix C4

Table of probability of bandwidth rate is less than *r*min P (*r* < *r*min) against maximum number of connection allowed (CP)

Nnrt	**λ_{nrt} - 35/s**	**λ_{nrt} - 45/s**	**λ_{nrt} - 55/s**
20	1.23e-6	1.23e-6	1.28e-6
25	3.28e-5	33.36e-5	3.41e-5
30	3.15e-4	3.22e-4	3.27e-4
35	1.67e-3	1.71e-3	1.74e-3

λ_{nrt} - Call arrival rate
Nnrt - maximum number of non real-time connections allowed

Call departure rate, μ_{nrt} = 2 calls per seconds
Number of base station, B = 7
Minimum bandwidth rate, $r$min = 0.8 UB
Amount of bandwidth for real-time connection, $C1$ = 30
Total amount of bandwidth for the base station, C = 40 UB

Appendix C5

Table of probability of bandwidth rate is less than *r*min P (*r* < *r*min) against maximum number of connection allowed for different *C*1. (CP)

Nnrt	***C*1 - 25 UB**	***C*1 - 30 UB**	***C*1 - 35 UB**
20	9.15e-15	1.28e-6	1.21e-1
25	3.14e-11	3.41e-5	2.48e-1
30	3.92e-9	3.27e-4	3.96e-1
35	1.22e-7	1.74e-3	5.00e-1
40	1.67e-6	6.24e-3	7.00e-1

Nnrt - maximum number of non real-time connections allowed
*C*1 - Amount of bandwidth for real-time connection

Call arrival rate, λ_{nrt} = 55 calls per seconds
Call departure rate, μ_{nrt} = 2 calls per seconds
Number of base station, *B* = 7
Minimum bandwidth rate, *r*min = 0.8 UB
Total amount of bandwidth for the base station, *C* = 40 UB

Appendix C6

Table of probability of bandwidth rate is less than *r*min P (*r* < *r*min) against maximum number of connection allowed for different size of cell cluster.(CP)

Nnrt	***B* - 7**	***B* - 10**
20	1.28e-6	2.45e-8
25	3.41e-5	8.15e-7
30	3.27e-4	9.69e-6
35	1.74e-3	6.36e-5
40	6.24e-3	2.82e-4

Nnrt - maximum number of non real-time connections allowed
B - Number of base stations in a cell cluster

Call arrival rate, λ_{nrt} = 55 calls per seconds
Call departure rate, μ_{nrt} = 2 calls per seconds
Minimum bandwidth rate, *r*min = 0.8 UB
Amount of bandwidth for real-time connection, *C*1 = 30 UB
Total amount of bandwidth for the base station, *C* = 40 UB

Appendix C7

Table of probability of bandwidth rate is less than *r*min P (*r* < *r*min) against Erlang load per cell site for different *C*1. (CA)

Erlang load per cell site	***C*1 -15**	***C*1 -30**
0.58	4.95e-6	1.62e-6
0.67	1.78e-3	2.81e-5
0.75	3.99e-3	2.67e-4
0.83	6.30e-2	1.63e-3
0.92	7.53e-1	7.22e-3

*C*1 - Amount of bandwidth for real-time connection

Call arrival rate, λ_{nrt} = (35 - 55) calls per seconds
Call departure rate, μ_{nrt} = 2 calls per seconds
Minimum bandwidth rate, *r*min = 0.8 UB
Total amount of bandwidth for the base station, *C* = 40 UB
Maximum number of non real-time connections allowed, *Nnrt* = 30
Number of base stations in a cell cluster, *B* = 7

Appendix C8

Table of probability of bandwidth rate is less than *r*min P (*r* < *r*min) against Erlang load per cell site for different call departure rate. (CA)

Erlang load per cell site	μ_{nrt} **- 2**	μ_{nrt} **- 3**
0.58	4.95e-6	5.77e-11
0.67	1.78e-3	2.69e-9
0.75	3.99e-3	7.29e-8
0.83	6.30e-2	1.32e-6
0.92	7.53e-1	1.73e-5

μ_{nrt} - Call departure rate

Call arrival rate, λ_{nrt} = (35 - 55) calls per seconds
Minimum bandwidth rate, *r*min = 0.8 UB
Amount of bandwidth for real-time connection, *C*1 = 15
Total amount of bandwidth for the base station, *C* = 40 UB
Maximum number of non real-time connections allowed, *Nnrt* = 30
Number of base stations in a cell cluster, *B* = 7

Appendix C9

Table of Quality-of-Service of non real-time connection against Erlang load per cell site for CA and CP

Erlang Load per cell site	**CP**	**CA**
2.5	3.14E-04	1.63E-06
2.9	3.19E-04	2.81E-05
3.2	3.22E-04	2.67E-04
3.6	3.25E-04	1.63E-03
3.9	3.27E-04	7.23E-03

CA - Complete Access
CP - Complete Partitioning

Call arrival rate, λ_{nrt} = (35 - 55) calls per seconds
Call departure rate, μ_{nrt} = 2
Minimum bandwidth rate, $r$min = 0.8 UB
Amount of bandwidth for real-time connection, $C1$ = 30
Total amount of bandwidth for the base station, C = 40 UB
Maximum number of non real-time connections allowed, $Nnrt$ = 30
Number of base stations in a cell cluster, B = 7

Appendix C10

Table of Quality-of-Service of non real-time connection against maximum number of connection allowed for CA and CP

Maximum Number of Connection Allowed	CP	CA
20	1.27e-6	7.19e-3
25	3.41e-5	7.18e-3
30	3.27e-4	7.20e-3
35	1.74e-3	7.23e-3
40	6.24e-3	7.23e
45	1.70e-2	7.23e
50	3.80e-2	7.23e

CA - Complete Access
CP - Complete Partitioning

Call arrival rate, λ_{nrt} = (35 - 55) calls per seconds
Call departure rate, μ_{nrt} = 2
Minimum bandwidth rate, $r$min = 0.8 UB
Amount of bandwidth for real-time connection, $C1$ = 30
Total amount of bandwidth for the base station, C = 40 UB
Number of base stations in a cell cluster, B = 7

Appendix E1

--------------------------------MODEL PLAYBACK--------------------------------

Areas
  Backbones
    Backbone
      Animate = TRUE
      Application Types
        Other
      Background B =   60.0
      Background G =   60.0
      Background R =   60.0
      Call File Scale =   1.0
      Call Routing Protocol
        Minimum Hop
      Msg File Scale =   1.0
      Number Of Reps = 1
      Packet Routing Protocol
        RIP Minimum Hop
      Postrun Reports (all on) = Input Buffer Totals, Blocked Call Counts, Disconnected Call Counts, Preempted Call Counts, Message Delivered, Setup Counts, Message Delivered, Message Delivered, Setup Counts, Frame Delay by VC, Frame Counts by VC, Access Link Stats, Cloud Access Buffer Policy, Cloud VC Buffer Policy, EPD/PPD, Message Delivered, Message Delivered, Setup Counts
      Rep Length =   240.0
      Reset System = TRUE
      Session File Scale =   1.0
      Show Clock = TRUE
      Snapshot Reports (all on) = Disk Error(a), Variable Error(a)
      Step Size =   500.0
      Trace File Name = trace.xls
      Trace To Disk = TRUE
      Warmup Every Rep = TRUE
      Warmup Length =   10.0
      Zoom Y =   -15412.0
Arcs
  Source Arcs
    Between call1 @ MT1 and MT1
      Hidden Name = Arc1
      Joints
        (1 of 2) =   6648.0,   23942.5
        (2 of 2) =   11471.5,   23907.5

Node = MT1
Parent = mobile1
Source = call1 @ MT1
Port Arcs
Between MT1 and 2Mb/s(1)
Buffer Units = Packets
Comments = (none)
Call Routing Penalty Table
One Hop
Table
(1 of 1) = 0.000000: 1
Hidden Name = Arc24
Input Buffer Limit = 1.0
Input Buffer Policy
Default Parameter: DEFAULT
Joints
(1 of 2) = 11471.5, 23907.5
(2 of 2) = 16708.0, 23704.0
Link = 2Mb/s(1)
Node = MT1
Output Buffer Limit = 1.0
Output Buffer Policy
Default Parameter: DEFAULT
Packet Routing Penalty Table
One Hop
Table
(1 of 1) = 0.000000: 1
Parent = mobile1
Between base station 1 and 2Mb/s(1)
Buffer Units = Packets
Comments = (none)
Call Routing Penalty Table
One Hop
Table
(1 of 1) = 0.000000: 1
Hidden Name = Arc25
Input Buffer Limit = 1.0
Input Buffer Policy
Default Parameter: DEFAULT
Joints
(1 of 2) = 22533.0, 23766.0
(2 of 2) = 16708.0, 23704.0
Link = 2Mb/s(1)
Node = base station 1
Output Buffer Limit = 1.0
Output Buffer Policy

```
    Default Parameter: DEFAULT
   Packet Routing Penalty Table
    One Hop
     Table
      (1 of 1) = 0.000000: 1
   Parent = mobile1
Layouts
 Links
  Point-To-Points
   2Mb/s(1)
    Comments = (none)
    Icon Name = circuit.icn
    Icon Scale =           1.0
    Icon X Coord =        16708.0
    Icon Y Coord =        23707.0
    Parent = mobile1
    Parms
     Point-To-Point Parameter: 2048 kbps
      Comments = (none)
      Tot Pkt Circuits = 32
    Postrun Reports = (none)
    Snapshot Reports = (none)
    Specifics, Link
     --
      Number of bits correctable = 0
      Propagation Delay = (none)
      Frame Error Rate = (none)
      Use Bit Error Rate = FALSE
      End = (none)
    Statistics Request = (none)
    State = Up
    Trigger List Type = Simple
 Sources
  Message Sources
   call1 @ MT1
                                   Bytes Or Cycles = 1.0
    Comments = (none)
    Destinations
     base station 1
    Destination Type = Weighted list
    IAT = Exp(20.0)                      ;interarrival time
    Icon X Coord =         6648.0
    Icon Y Coord =        23942.0
    Msg Size Unit = Packets
    Parent = mobile1
    Postrun Reports = (none)
```

```
    Protocol
      ATM AAL5 example
        Comments = (none)
        Rate Control
          Constant Rate Control Parameter: DEFAULT
        Traffic Policy
          ATM Traffic Parameter: DEFAULT
            Comments = (none)
    Routing Class
      Standard
    Statistics Request = (none)
Nodes
  Routers
    base station 1
    Comments = (none)
    Icon Name = atm.icn.r11
    Icon Scale =          1.0
    Icon X Coord =        22533.0
    Icon Y Coord =        23766.0
    Parent = mobile1
    Parms
      Network Device Parameter: FORE Systems ForeRunner ASX-1000 (ATM Switch)
        Comments = (none)
        Files
          GENERAL STORAGE,              0 bytes           ;no buffer
        Node Buffer Units = Packets
        Protocol Processings
                    DEFAULT : 500.0                        ; service time
    Postrun Reports = Input Buffer Totals
    Snapshot Reports = Disk Error(a), Variable Error(a)
    Statistics Request = (none)
    State = Up
    Trigger List Type = Simple
  Processing Nodes
    MT1
    Comments = (none)
    Icon Name = phone.icn
    Icon Scale =          1.0
    Icon X Coord =        11473.0
    Icon Y Coord =        24178.0
    Parent = mobile1
    Parms
      Processing Node Parameter: MT1
        Comments = (none)
        Files
          GENERAL STORAGE,              0 bytes
```

Input Buffer Capacity = 1.0
Node Buffer Units = Packets
Output Buffer Capacity = 1.0
Packet Processing Uses Processor = FALSE
Protocol Processings
DEFAULT : (none)
Postrun Reports = Input Buffer Totals
Snapshot Reports = Disk Error(a), Variable Error(a)
Statistics Request = Call Bandwidth Level
State = Up
Trigger List Type = Simple
Miscellaneous
Call Routing Classes
Standard
Packet Routing Classes
Standard

Appendix E2

Point-to-Point Link Parameters

Number Of Circuits: Number of identical trunk circuits available to carry circuit-switched traffic.

Bandwidth/Circuit (kbps): Bandwidth per circuit. The total capacity for carrying circuit-switched traffic is the number of circuits multiplied by the bandwidth. A trunk could therefore be described as 30 circuits each at 64K, or 1 circuit at 1920K. When a circuit-switched call is routed over the link, the call has a bandwidth requirement. If the current free capacity on the link is greater than the call requirement then the call may be carried. The free capacity is therefore reduced by the bandwidth requirement for the call holding time.

Spare Call Channel Usage: This feature allows calls to be assigned to channels based upon how much bandwidth the calls require and what kinds of calls are already assigned on the channel. Typically, a call will take up a full circuit switched call channel, because each call will require the same bandwidth as the channel. If more bandwidth is needed, more channels are combined to support the call. In the newer technologies, the calls may be compressed so that a call (such as voice) does not fully occupy a call channel (Subchannel), or calls such as video calls may require a continuous bandwidth of more than one channel (Superchannel), but not need a full integer number of channels. The two types of channeling options are:

Subchannel: The case where a call takes up less bandwidth than the channel provides. The total call capacity is in terms of a number of circuits (channels) and an equal bandwidth per circuit.

Superchannel: The case where a call takes up more bandwidth than the channel provides. The total call capacity is in terms of a number of circuits (channels) and an equal bandwidth per circuit.

Appendix E3

Call Source

Scheduling

Application instances can be scheduled by either Delay Time, by Iteration Time or by Received Message.

> *Delay Time* scheduling uses the given times as the interval between the completion of one instance of the message and the start of the next instance of the message.
>
> *Iteration Time* scheduling uses the given times as the interval between the start of one instance of the message and the start of the next instance of the message. Overlapping instances of the same message can exist with this scheduling.
>
> If only one instance of a message is required in a replication, one way to achieve this is to use Iteration Time scheduling and to set the Interarrival Time Distribution to None, the Last Arrival to None, and the First Arrival to the time when the single instance is required.
> Received Message Scheduling requires that one or more messages must arrive at the node where the message prototype exists for new instances of the message to be created. In fact the scanning for received messages is based on the incoming Message Text. Multiple messages may be waited for, and message text may be scanned by wildcard.

Interarrival Time Distribution :

Used to make a statistical pick for the time between call instance creations. Any of the statistical distributions found in COMNET III may be used.

> It is typical for the interarrival time to be based on an exponential distribution. This results in call generation according to a Poisson process.

Appendix E4

Calls Duration

The time a call instance is to last. Any of the statistical distributions found in COMNET III may be used. A call instance is a demand to hold an amount of bandwidth between an origin and destination for a given duration.

It is common to quantify voice traffic in terms of Erlang. The Erlang load is not directly entered into COMNET III. Rather the call Interarrival Time and the call Duration are entered, which are related to the Erlang load by the formula:

Erlang Load = Mean Holding Time / Mean Interarrival Time

The Mean Holding Time and Mean Interarrival Time must be specified in the same time units. Note that the Erlang load and the Erlang distribution are not the same type of object and should not be confused. Voice call duration is usually based on a Normal distribution.

Appendix E5

Destination Type

The destination type may be:

Multicast List: Send a copy of the message to every destination in the list.

Random List: A list of destination names, one of which is picked at random when a destination is required. All destinations in the list have equal probability.

Random Neighbor: Automatically constructed for the originating node by COMNET III. In this case the nodes that are reachable within one hop from the origin in the same subnet are considered neighbors. When a destination is required one of the neighbors will be picked at random, all neighbors having equal probability.

Note: The list of nodes that are one hop away may change as the model is modified and, thus, the Random Neighbor may send to unexpected nodes in future modifications of the model. For more control over destinations, use a Random List.

Weighted List: A Weighted List destination is a list of destination names, one of which is picked at random when a destination is required. All destinations in the list have an assigned probability, so that you can specify which is the most likely destination. The total of all the probabilities must add up to 1.

Round Robin List: The Round Robin List forms an ordered sequence of destinations. The first time a destination is chosen from the list, the first member will be picked, the second time it will be the second member, and so on. After the last member, the first will be picked again. Ordering of the list is based on the order in which you add them to the list, unless you change it using the Set Order button.

Least Busy List: The destination whose processor is being used the least is chosen from the list of destinations. When the least busy destination turns out to be a group node, or if the only member of the list is a group node, the specific destination in the group will be chosen by the same Least Busy algorithm. If two or more destinations are equally busy (for example, both have 0% utilization), the tie is broken by round-robin selection. If all the nodes remain un-busy, the result will be identical to a round-robin list.

Edit Destination List: Allows the selection of a destination, or a group of destinations for every destination type except Random Neighbor.

Appendix F

1. Define the sequence of equation (4.15) into a differential equation:

$$q_k = \left(\overline{\alpha} - k\overline{\delta}\right)_{q_{k-1}} - \left(\overline{\beta} - k\overline{\delta}\right)_{q_{k-2}}$$

$$\frac{\delta P_k(t)}{\delta t} = \left(\overline{\alpha} - k\overline{\delta}\right)P_{k-1}(t) - \left(\overline{\beta} - k\overline{\delta}\right)P_{q_{k-2}}(t)$$

$$\frac{\delta P_o(t)}{\delta t} = P_o(t)$$

We then translate the different equation defining this sequence into a differential equation for the associated z transform:

$$P(z,t) \approx \sum_{k=0}^{\alpha} P_k(t) z^k$$

$$\sum_{k=1}^{\alpha} \frac{\delta P_k(t)}{\delta T} z^k = \left(\overline{\alpha} - k\overline{\delta}\right)\sum_{k=1}^{\alpha} P_{k-1}(t) z^k - \left(\overline{\beta} - k\overline{\delta}\right)\cdot \sum_{k=1}^{\alpha} P_{q_{k-2}}(t) z^k$$

$$\frac{\delta}{\delta t}\left[P(z,t) - P_o(t)\right] = \left(\overline{\alpha} - k\overline{\delta}\right)zP(z,t) - \left(\overline{\beta} - k\overline{\delta}\right)z^2 P(z,t)$$

$$\frac{\delta}{\delta t} P(z,t) - \frac{\delta}{\delta t} P_0(z,t) = \overline{\alpha} z P(z,t) - k\overline{\delta} z P(z,t) - \overline{\beta} z^2 P(z,t) + k\overline{\delta} z^2 P(z,t)$$

$$\delta P(z,t) - \left[P(z,0^+) - P_o(s)\right] = \left(\overline{\alpha} - \overline{\beta} z^2\right)P(z,s) - k\overline{\delta} z(1-z)P(z,t)$$

$$\left(1 - \overline{\alpha} z + \overline{\beta} z^2\right)P(z,s) + k\overline{\delta} z(1-z)P(z,s) = (1-z)P_o(s)$$

$$Q(z)(1-\overline{\alpha}z+\overline{\beta}z^2)+Q(z)\overline{\delta}z(1-z)=(1-z)q_o \qquad (C1)$$

where $\overline{\beta}=Bm_{II}c^{-1}, \overline{\alpha}=Bm_{II}c^{-1}+(r+1), \overline{\delta}=c^{-1}$

divide (C1) by $\frac{(1-z)}{\overline{\delta}z^2}$, we get:

$$Q(z)+Q(z)\frac{\left(1-\overline{\alpha}z+\overline{\beta}z^2\right)}{\overline{\delta}z^2(1-z)}=\frac{q_o}{\overline{\delta}z^2}$$

2. $$\frac{\Gamma(a+n,x)}{\Gamma(a+n}=\frac{\Gamma(a,x)}{\Gamma(a)}+e^{-x}\sum_{s=0}^{n-1}\frac{x^{a+s}}{\Gamma(a+s+1)}$$