Using Call Admission Control and Call Duration Control for Mobile Network Congestion Management

by

Oyeniyi Akeem Alimi

(215085115)

A dissertation submitted in partial fulfilment of the requirements for the degree of

MTech Electrical Engineering

in the

Faculty of Engineering and the Built Environment

at the

University of Johannesburg

Supervisor: Prof. Meera K. Joseph

July 2016

DECLARATION

I, **OYENIYI AKEEM ALIMI** hereby declare that the dissertation titled "Using Call Admission Control and Call Duration Control for Mobile Network Congestion Management" submitted for MTech is my original work. This dissertation has not been previously submitted to another University or higher educational institution.

16/11/2016

Signed ……………….. Date…………………

Abstract

Wireless communications have experienced massive development remarkably in the past decade. Networks such as global system for mobile communications (GSM) and universal mobile telecommunication service (UMTS) has enjoyed enormous patronization, hence leading to massive mobile network congestion. The problem of network congestion is a network managerial issue that affects the Quality of Service (QoS) rendered by a network. Hence for the sustainability of the system, there is a need to fully manage the radio resources during peak and off peak periods.

Call Admission Control (CAC) schemes are constantly being used extensively in managing mobile network congestion. CAC is an approach that can provide credible QoS by regulating the number of connections into the cellular network thereby allowing good use of the radio resources, reducing network congestions, interference and other QoS problems. Network resources cannot be available for all users at all times especially during busy hour traffic. Hence, these resources require effective and efficient allocation so that more subscribers are being allowed to use the network irrespective of the network traffic at a particular reference time.

Mobile network congestion can be controlled by suppressing the mobile network traffic demand. The traffic demand varies proportionally with average rate of call arrival and the average duration of the calls. Hence marginalizing the average call duration will minimize traffic load, thereby reducing network congestion. Some network users due to their lifestyle of being affluent or business demands, holds on to a particular channel during long duration of calls at the detriment of other network users. CAC scheme cannot solve congestion problems due to the selfishness of these particular sets of users. Here, the combination of a channel reservation CAC scheme with Call Duration Control (CDC) scheme was proposed, in which users that have stayed over a predetermined period in the network will be served a termination notice so as to make the channel available for new users. The motive is to provide the available channels to accommodate more users.

A simulation-based approach was used to model the combination of the two schemes and the combination produces good results in reducing the congestion menace. The network type deployed in the research is GSM. The result of the CAC/CDC combination schemes were compared to the result of the ordinary CAC schemes in order to verify the impact of the

combined schemes. An analytical model built on Markov decision process was used for the scheme modelling. QoS parameter metrics namely call arrival rates, call blocking probability and call dropping probability are used for the evaluation of how the network is being congested.
Assumptions are made that all the users respond positively to the termination notice once the call duration reaches the termination time. Simulations involving different termination time as well as different load situations are presented in this research. These were carried out for comparison purposes. The results obtained showed better system performance, lower congestion rate in terms of more users, getting access to the network, and most importantly a judicious use of the radio resources are obtained with the combined scheme.

Keywords

Call Admission Control (CAC), Call Duration Control (CDC), Simulation, Network congestion Handoff Call Dropping Probability, Termination Time.

List of publications

O. A. Alimi, M. K. Joseph and A. O. Akinlabi, "Network congestion management using call admission control" in Proceedings of the 7th International Conference on Latest Trends in Engineering and Technology 2015, page 199-203, November 2015.

Acknowledgement

I am grateful to God for His guidance, love and mercy.

I am deeply grateful to my supervisor, Dr. Meera K. Joseph who has been my supervisor for my thesis work. I strongly appreciate her support, inspiration and guidance during the period of the study.

I want to express my profound gratitude to my family for their financial, moral and spiritual support throughout my entire life.

Table Of Contents

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ACG	Automatic Call Gapping
AMPS	Advanced Mobile Phone Service
CAC	Call Admission Control
CDC	Call Duration Control
CDMA	Code-Division Multiple Access
FM	Frequency Modulation
GMSK	Gaussian Minimum Shift Keying
GSM	Global System for Mobile Communication
HCDP	Handoff Call Dropping Probability
HRPD	High Rate Packet Data
HSPA	High Speed Packet Access
ITU	International Telecommunication Union
ITU-R	International Telecommunication Union-Radio
LTE	Long Term Evolution
MANET	Mobile Ad Hoc Network
MSC	Mobile Switching Centre
MTSO	Mobile Telephone System Office
NMT	Nordic Mobile Telephony
OFDMA	Orthogonal Frequency Division Multiple Access
PB	Blocking Probability
PD	Dropping Probability
PSTN	Public Switched Telephone Network
QoS	Quality of Service
RAT	Radio Access Technology

RRM	Radio Resources Management
TRAI	Telecom Regulatory Authority of India
UMB	Ultra Mobile Broadband
UMTS	Universal Mobile Telecommunication System
WCDMA	Wideband Code-Division Multiple Access
WLAN	Wireless Local Area Network
WiMAX	Worldwide Interoperability for Microwave Access

1. INTRODUCTION

In recent years, the level of patronage of cellular devices is overwhelming. This is a direct result of the mobile network system being on a digital platform as well as mobility advantage of users irrespective of their location at any time. Network congestion increases as mobile users increases. Some subscribers carry two or three phones to ensure that they can be reached and also reach out, at all times. As a result of this congestion, Quality of Service QoS from the network providers have massively dropped. This makes it a priority that the congestion menace has to be handled appropriately in order to make communication easier and efficient.

Obviously, mobile network congestion can be solved naturally by providing more base stations thereby increasing the network capacity, but such an approach involves a lot of investments and it is not immediately possible everywhere. Also system capacity will be underutilized at off peak periods. Hence, it is very obvious to work for other forms of mobile network congestion management.

Some of the mobile network congestion control measures that have been taken include load balancing, Automatic Call Gapping (ACG) where attempt rates of calls are being reduced by allowing only one call attempt per a specified gap interval. Several other attempts that have been used to forestall and manage the congestion in mobile networks includes the Priority Token Bank, Channel borrowing and Channel allocation for mobile cellular networks.

The various techniques mentioned above are not effective and efficient enough to curb the mobile network congestion menace majorly due to the geometric surge in the number of mobile network users as well as underutilizing the resources during off peak period.

This study aims at reducing network congestion using the performance evaluation of a channel reservation based Call Admission Control scheme and its combination with Call Duration Control scheme throughput. The network type deployed in the research is Global System for Mobile Communication (GSM), a second generation mobile network which is the highest ranking network till date in terms of subscribers. This research focuses on voice call, hence the choice of GSM network. Results of the combination shows reduction in the percentage of blocking and dropping probabilities compared to conventional CAC scheme. Also, the thesis shows effective use of available radio resources and lower congestion turnout.

1.1 BACKGROUND AND RATIONALE

Cellular telecommunication network is growing rapidly whereby users or devices can communicate and access information anytime and anywhere [25]. The rate at which mobile networks has been fully accepted and its growth in South Africa is massive. From the survey conducted by Deloitte Digital Company which states the population of the country as approximately 51.8 million as of 2011, the survey billed the total number of active SIM cards at 66.1million which is approximately 1.35% higher than the active human population. Furthermore, the survey shows that the subscriber base covers 83% of the country landmass [68]. The most popular network in existence till date is the Global System for Mobile Communication. Billions of people rely on the GSM network for their daily businesses and living. Moreover, the use of the network is available to all humans irrespective of one’s financial status. This has made the use of the network to be growing exponentially with more users on daily basis

The escalation in users of telecommunication network leads to congestion in the network.

Congestion in GSM arises when there are no free resources to cater for new subscribers due to the fact that large number of subscribers are accessing the network or a situation whereby there is a malfunctioning in the network system at a particular point in time [18, 55]. Mobile network becomes frequently congested exceptionally at perilous disaster periods.

Congestion issue is a grave worry since it degrades the Quality of Service (QoS). Basically network congestion is due to unavailability and inappropriate use of the radio resources. Whenever the congestion rate is high, blocking of calls arises.

Blocking occurs due to the unavailability of accessible channels to allot to users in the base station [55]. Blocking in mobile telecommunication network are classified into two namely: New call blocking, which is explained as blocking new initiated calls while the other type is Handoff Call blocking which is explained as blocking of handoff calls as a result of the user’s mobility.

Managing mobile network has always been a serious concern to network providers. Several methods have been proposed and used to reduce the congestion, one of which is call admission control. However, due to the traffic pattern of mobile network and user, CAC scheme has not been able to manage the congestion issue effectively. Hence, providers have been sourcing for other methods to reduce the mobile traffic load. Call duration control scheme which marginally reduces the average call duration in such a way as to accommodate more users is being introduced in this thesis to effectively produce a better result.

1.2 PROBLEM STATEMENT

The telecommunications industry is rapidly growing, leading to massive mobile network congestion. This has led to vast mobile network congestion. The various techniques of controlling congestion employed so far in all systems have always in a general view, been focused on two principles. These ideologies are either to dismiss extreme load in order to avoid overload from happening or redirecting overloads when it happens.

Some of the proposed schemes include a buffer management procedure which involves accepting handoff calls temporarily using guard channel principle. Also, Call Admission Control schemes have been used for managing network congestion due to several of its characteristic features.

Some other schemes involves integrating pricing into CAC schemes in order to suppress the network load. The various pricing schemes basically request users to pay higher prices for their calls especially during peak period of mobile network traffic. Most of these schemes have their shortcomings and they are particularly beneficial to the providers in terms of added generated revenues, although at the detriment of users.

However, in this thesis, integration of a CDC scheme into a channel reservation based CAC was simulated and the results were compared with ordinary CAC scheme. The results from the combined scheme showed better performances in terms of managing network congestion.

1.3 RESEARCH OBJECTIVES

The key objective of this work is to use CDC and CAC scheme to manage mobile network congestion.

The sub-objectives are:

a. To explore how to integrate the Call Duration Control scheme into a channel reservation based Call Admission Control scheme. The performance of the combined scheme will be explored in a low traffic scenario and high load scenario.

b. To explore how the combined CAC/CDC scheme can be compared to the ordinary channel reservation based CAC scheme and use simulation to show how the combined schemes assist in managing mobile network congestion.

c. To determine the forecasted periods during the day when mobile network congestion is at its peak.

1.4 RESEARCH QUESTION

What is the performance of the combined scheme of call admission control and call duration control in managing mobile network congestion?

1.5 RESEARCH METHODOLOGY AND DESIGN

In this dissertation, a call duration control scheme was incorporated into a channel reserved based Call Admission Control scheme in order to reduce mobile network congestion. The choice of a reserved channel based CAC scheme was to prioritize handoff calls over new calls.

One of the main reasons why mobile network congestion has been prominent in recent years is due to long duration callers. Network traffic is proportional to average call duration and call arrival rates. Suppressing the average call duration will go a long way in minimizing the mobile network traffic load.

CAC has been used extensively for managing mobile network congestion, however, due to the mobile network traffic load that has been increasing geometrically in recent times, integrating a call Duration Control scheme into the CAC scheme in order to boost the performance in terms of managing the congestion is difficult.

Call duration control scheme involves the introduction of a termination notice time ***(T_u)*** to users once their call duration reaches a specific threshold, this is done in order to encourage long duration call users to free up channels for new users. The analytical modelling of the scheme is based on Markov decision process and assumptions were made that calls arrival rate follows poisson distribution.

Simulations are carried out for the CAC scheme alone and the result is compared with the combined CAC/CDC scheme. The flowchart in figure 1.1 illustrates the methodology stages in the dissertation.

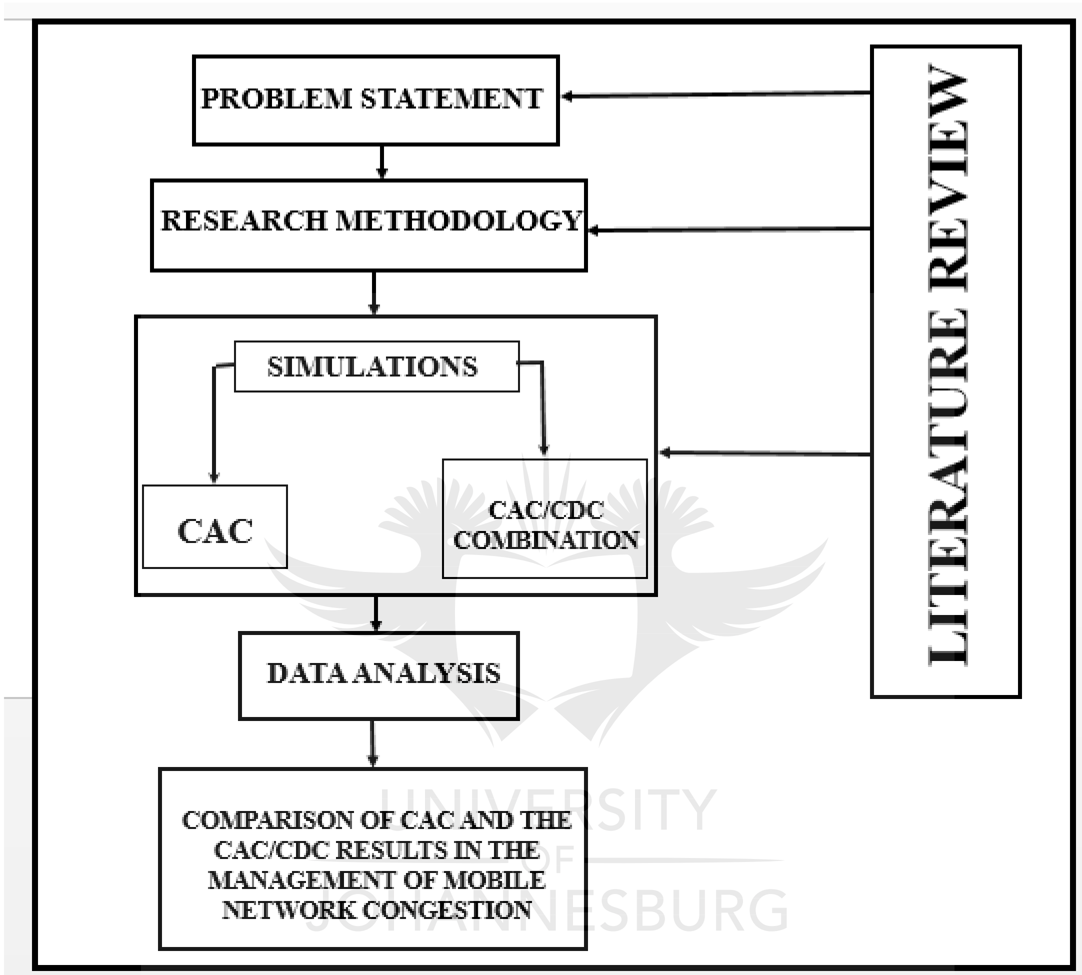

Figure 1.1 Flowchart of Managing Mobile Network Congestion

1.6 RELEVANCE OF THIS DISSERTATION

Cellular telecommunication especially the GSM has shown to be the fastest developing standard in the telecom world. Billions of people rely on mobile telecommunication especially the GSM network for their daily businesses and living. Moreover, the use of the network is available to all humans irrespective of one's financial status. This has made the use of the network to be growing exponentially with more users on daily basis.

The massive increase in the number of GSM telecommunication network users leads to congestion in the network. Mobile network congestion management has been a real issue to both users and providers. In recent times, managing mobile network congestion is an area where various researches are focusing on. Several management schemes have been proposed and used. However, they all have their shortcomings especially the schemes that involves integration of pricing. Although network providers generate more revenues from pricing integration to suppress network load, it constitutes a massive constraint on the users who have to pay more than their usual charging tariff.

In this dissertation, for the motive of suppressing mobile network traffic, price variation as added incentives was not integrated in the combined CAC/CDC scheme. The combined scheme only involves the marginalization of the traffic load through the introduction of termination notice time to reduce the average call duration of users, thereby accommodating more users into the network.

The results from the combined CAC/CDC schemes from the simulation show better performance when compared to ordinary CAC scheme without any financial constraint on the mobile users.

1.6.1 UNIQUE CONTRIBUTION

Individually, several Call Admission Control schemes [55, 82] have been singularly proposed and used extensively for managing mobile network congestion in virtually all type of networks due to its efficiency. Some of the CAC schemes already proposed and being used include non-prioritized CAC schemes [72] and channel reservation based CAC schemes [70, 71]. However, combining CAC scheme with the call duration control scheme whereby a termination notice time is sent to users in the network once they reach a predetermined call duration is the major contribution of the research in this thesis. With the integration of the call duration control into the channel reservation based CAC scheme, network congestion is maximally reduced when compared to the ordinary channel reservation based CAC scheme in terms of probability of calls being dropped or blocked with respect to call arrival rates.

1.7 THESIS LAYOUT

The thesis is arranged as follows:

Chapter 2 presents the overview to the history of Telecommunication, Mobile Networks as well as overview of mobile network congestion issues including its problems, effects and management techniques. Blocking, QoS, Radio Resources Management, Channel Allocation and handoffs are also discussed in the chapter. An overview on Call Admission control in managing Mobile Network Congestion is also discussed.

Chapter 3 presents the methodology and the procedures employed in the research work. Furthermore, the simulation tools and the simulation setup for the research work are also highlighted in the chapter.

The result of the research work and the discussion of the performance evaluation of simulation results are presented in chapter 4.

Chapter 5 presents the conclusion and envisaged future work.

2. LITERATURE REVIEW

2.1 INTRODUCTION TO TELECOMMUNICATION

Information dissemination is an integral part of human's daily activities [15]. Old ways of communication and information dissemination over long distance include the use of thick smokes, flashing mirrors, whistles, touch signaling etc. Furthermore, in order to relay the acquired information over thousands of kilometers, tall buildings and towers were built on mountains in order to relay the message [37]. However, in recent years, the desire to make communication effective and efficient is a target, motivating researchers to pursue better means of information dissemination on real time basis. Hence, new methods of telecommunication involves the use of radio devices such as telephones to communicate over long distances. Figure 2.1 shows the hierarchy of telecommunication.

New age telecommunication can either be wired or wireless depending on the mobility of devices and the users using the devices. Advancement of various telecommunication technology has resulted into vast benefits in various ways of life such as education, business and military. One of the main driving forces for the advancement is the increasing demand for the network services from users. Users expect to send and receive multimedia messages, have access to internet etc.

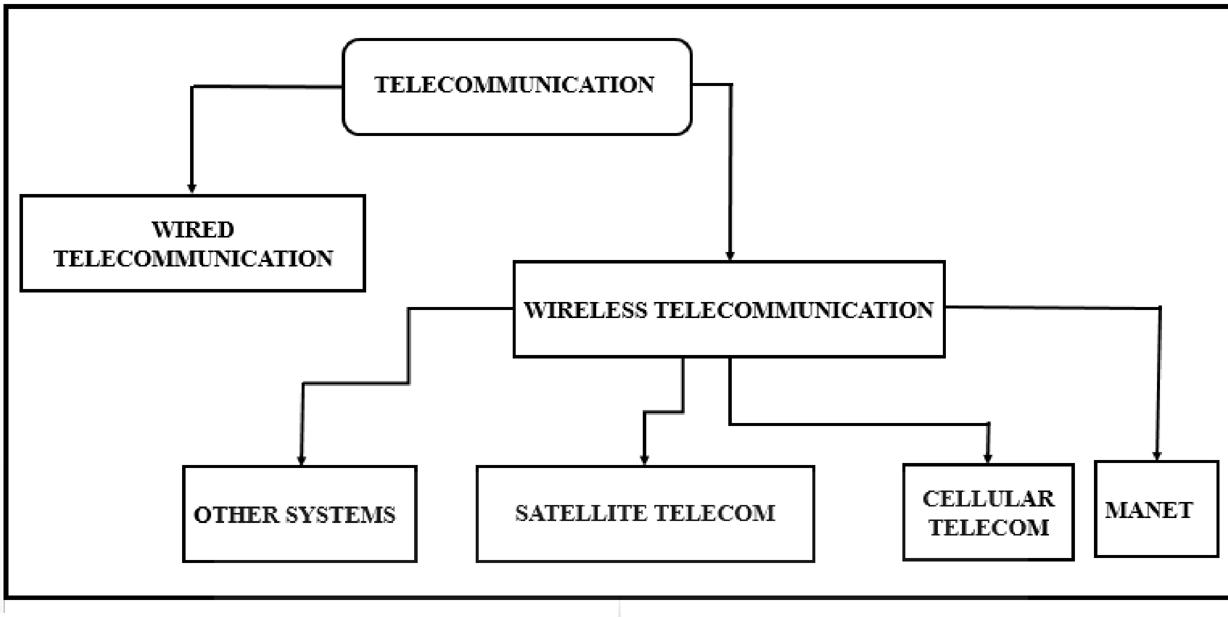

Figure 2.1 Telecommunication Classification

Wired telecommunication are the old generation type of telephony which does not allow mobility of users. It involves the use of cabled and landline phones for communication.

Whereas, wireless telecommunication has virtually taken over all forms of wired telecommunication in the present day world. The discovery of the possibility of transmission via electromagnetic waves led to the discovery of wireless telecommunication. It includes the various cellular telecommunication, satellite telecommunication, MANET telecommunication networks and several others systems.

2.2 CELLULAR TELECOMMUNICATION NETWORK

Cellular telecommunication networks has become prominent in this present day that it is no longer amazing seeing people everywhere holding a device or a mobile gadget and making conversations with it [40]. Hence, making it the most sought after type of communication. Cellular devices allow a person to make or receives phone calls in as much as the individual and the person he is reaching out to, is within the range of frequency spectrum of the mobile operator. Cellular telecommunication network is growing rapidly whereby users or devices can communicate and access information anytime and anywhere [25]. In Japan, the number of users has almost doubled within a span of a calendar year. In 1997, it was only an estimate of 33 million users and as at 2010, close to 100 million subscribers use the telecommunication network [24].

The use of cellular network could be followed back to as early as the 1900s when the United State law enforcement agency were researching radio communication to make their law enforcement profession easy. The introduction and advancement of Frequency Modulation (FM) during the 2^{nd} World War in 1940s assisted in the war [36].

Over the years, the advancement in researches have made the inventions of various cellular networks ranging from the first generation (1G) networks to the second generation (2G), 3G etc with new inventions far better and allowing better use and greater benefit to the previous inventions.

The evolution of cellular networks starts with mobile telephony in the late 1970s [36] and till date, more advancements and researches are being done to make communication easier, faster and better.

2.2.1 FIRST GENERATION (1G) NETWORKS.

Commercialization of the cellular communication takes to full swing with the First Generation (1G) system. 1G system allows frequency reuse and handoffs. AT&T at Bell Laboratories introduced and modified 1G model. Their intention was to provide adequate bandwidth for numerous users in a geographical region [41]. Furthermore, due to the unavailability of adequate frequency band for radio bandwidth, frequency reuse analogy has to be introduced. 1G network is analogue in nature. Also, 1G initiates the subdivision of coverage areas into cells. In early 1980s, the Nordic Mobile Telephony (NMT) system was launched in the northern European nations. During that period, North America also launched the analog Advanced Mobile Phone Service (AMPS) telephony [41]. As shown in figure 2.2, a typical 1G architecture contains a Mobile Telephone System Office (MTSO) which can also be called the Mobile Switching Centre (MSC), The Public Switch and the cell sites.

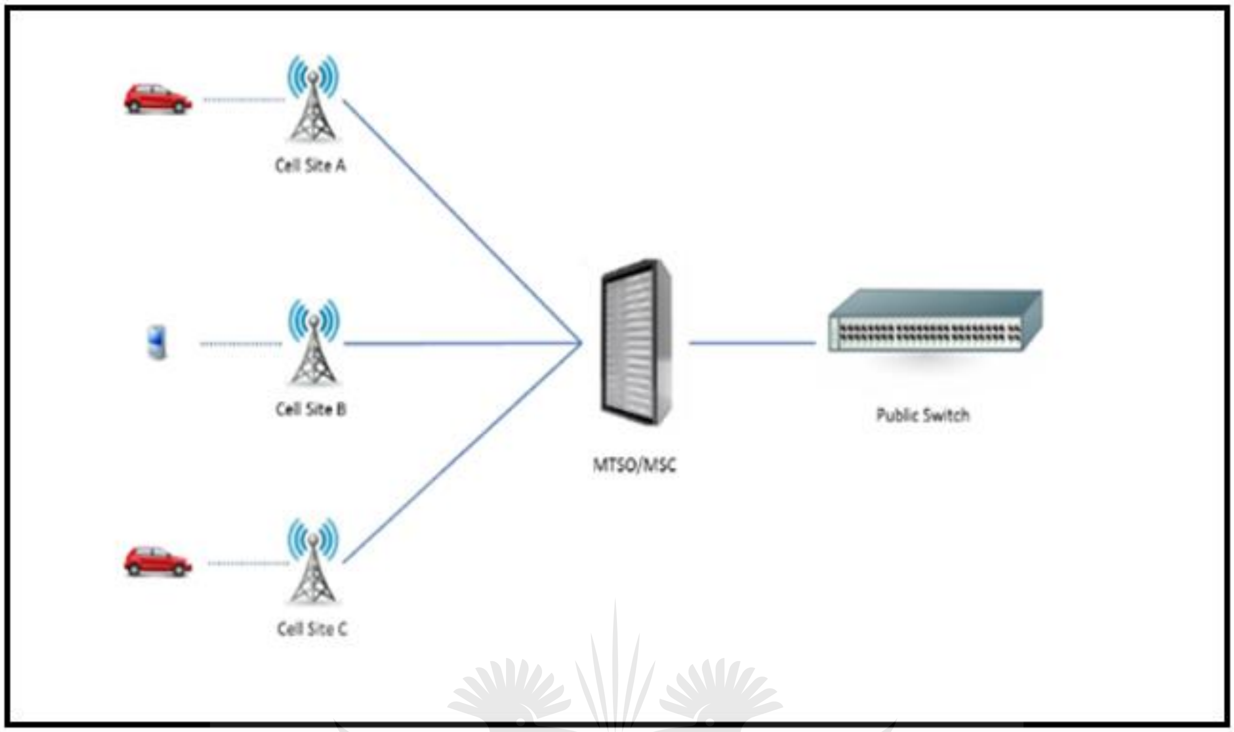

Figure 2.2 A Typical 1G Network Architecture [41]

2.2.2 SECOND GENERATION (2G) NETWORKS

The problem of inadequate bandwidth, unreliable and inflexible mobile communication associated with 1G networks led to the development of 2G networks [42, 43]. The network systems in 1G use analogue transmission technology unlike the 2G which provides a platform using digital transmission and circuit switching. Furthermore, 2G networks offers more services such as data capabilities, better speech quality and short call holding time [5].

2G network includes the IS-136 and IS-95 standards. The most popular 2G network is the GSM [41].

2.2.3 THIRD GENERATION NETWORKS (3G NETWORKS)

In the later years of the ‘80s, ITU-R launched Universal Mobile Telecommunications System (UMTS). The model was widely regarded as the Third Generation (3G) network system. Third generation networks are built on Wideband Code Division Multiple Access (WCDMA). In terms of bandwidth, they utilizes 5 MHz whereas the CDMA2000 (an advance version of 2G network in 3GPP2) utilizes 1.25 MHz. Afterwards 3GPP2 correspondingly developed a similar version with the extension of its frequency band to 5 MHz comprising of three 1.25 MHz. The standard is known as CDMA2000-3x [44]. In order to distinguish between the two standards, the 5 MHz CDMA standard is regarded as CDMA-3x whereas the standard containing the single carrier of 1.25 MHz CDMA is denoted as CDMA-1x or 3G-1x. Initial releases of the two standards never achieve their expectations. The expected data transmission rate was far lower when compared with the practical standards.

After a series of numerous researches in order to increase the data transmission rate, 3GPP2 came up with the invention of High Rate Packet Data (HRPD). The HRPD utilizes information optimizing procedures which are more advanced compared to the previous standards. However, High Rate Packet Data calls for extra 1.25 MHz subcarrier for it to transmit just data alone without any voice transmission. This particular standard is initially referred to as CDMA2000-1x EVDO (evolution data only).

3GPP improved on this invention and enhanced on the Wideband Code Division Multiple Access to produce High Speed Packet Access (HSPA). The HSPA uses the same procedure as the HRPD, although unlike the HRPD, the HSPA transmit both data and voice on the same 5MHz bandwidth. They are multiplexed in downlink. Further inventions from 3GPP2 includes the development of CDMA2000-1x EVDO from ordinary data transmission standard to CDMA2000-1x EVDV (evolution data and voice). This CDMA2000-1x EVDV standard uses the same subcarrier on 1.25MHz although it was not published for commercial usage. Subsequent inventions on 3GPP2 led to the invention and launching of HRPD Voice over IP (VoIP) which fully supports both data and voice on individual carrier. All recent technologies developed henceforth allows massive transmission of data for third generation networks and they are being used all over the world [81].

2.2.4 POST 3G PHASE

Apart from the exploration of the HSPA and HRPD which are just upgrades on the typical 3G versions. Infact HSPA is treated as 3.5G. However, IEEE802 LMSC launched new standard

IEEE 802.16E, an advanced standard of IEEE 802.16 [84]. The difference between the two standards is that IEEE802.16E has the advantage of mobility in terms of wireless broadband access. Also, IEEE802.16E uses OFDMA as well as providing improved data rates than the HSPA and HRPD. The new standard is better regarded as Worldwide Interoperability for Microwave Access (WiMAX). WiMAX provides the platform for both 3rd Generation Partnership Project (3GPP) and 3rd Generation Partnership Project 2 (3GPP2) to offer their own standards of Long Term Evolution (LTE) and Ultra Mobile Broadband (UMB) respectively [47]. LTE is an advancement over HSPA, they are superior to HSPA with attributes such as higher spectral efficiency of 2 to 4 times, and they can operate at a frequency of 900MHz to 2.6GHz. Furthermore, LTE makes provisions for a massive bandwidth ranging from 1.25MHz to 20MHz, which are quite higher than what HSPA offers. Hence, they are regarded as 4G [47].

Table 2.1 Evolution of Cellular Networks

Technology Standards	**Estimated Roll out year**
1G which includes NMT and AMPS	1980
2G which includes GSM, IS 136 and CDMA	1990s
3G which includes UMTS, HSPA and HRPD	2000-2010
4G which includes LTE, WIMAX and UMB	2011-

2.3 GSM ARCHITECTURE

From the time when GSM was introduced in the 1990s, it has experienced numerous progressions with the intentions of improving the innovation in terms of proficiency and efficiency. However, the issue of versatility and mobility limitation in the analog framework which was illuminated and resolved by the GSM framework likewise accompanied its own difficulties. Compared to the analog wired telephony whereby devices are static and immobile, devices used in GSM systems are not confined, which allows mobility advantage, as indicated in its name. The devices are not restricted in terms of mobility, hence users have the luxury to communicate over a long range of distance within the network coverage. First releases of mobile frameworks utilized powerful, centrally located transmitters for communication with powerful portable mobile stations on a low bandwidth of recurrence frequency channels. A couple of frequencies were utilized over a huge territory with numerous users contending to utilize the same channels. The accessible frequency band turned out to be clearly deficient for the ever-growing users [2]. The introduction of frequency re-use was subsequently embraced to boost and effectively allows more users to be attended to. For the sake of avoiding interference issues, cells utilizing the same frequencies are physically separated numerous kilometers. Frequency re-use is accomplished by partitioning the operator's geological territory into a few groups.

Currently, Global System for Mobile Communication is the most common and generally utilized mobile technology. As at 2012, it is accessible in over 219 nations and regions around the world, it has a market share of over 90% [11, 45]. Global System for Mobile Communication tolerates the voice transmission and data transmission on its network.

Global System for Mobile Communication networks architecture is analyzed into three sections, the network switching subsystem, the mobile station and lastly, base station subsystem [36, 45].

- The Mobile Station (MS), this comprises Mobile Equipment (ME) and Subscriber Identity Module (SIM).
- The Base Station Subsystem (BSS) maintains radio links between mobile station devices and other devices of the switching subsystem. The BSS is made up of various elements including Base Transceiver Station (BTS) and Base Station Controller (BSC).
- The network switching subsystem, this composes of the actual component of mobile network. Execution of call routing and other subscriber's related tasks are done by the network switching subsystem.

The Mobile Station can be attributed to gadgets and other electronic tools which have the capacity to connect to the central network wirelessly [46]. It is made up of ME and SIM. Mobile stations are the typical devices although they can be in form of other electronic gadgets whereas

Subscriber Identity Module consist of a chip built using ICs. Typical Subscriber Identity Module chips accommodates data such as identity of the subscriber.

The Base Station Subsystem (BSS) is the access channel of subscribers to Global System for Mobile Communication network, it comprises of the Base Station Controller (BSC) and the Base Transceiver Station (BTS).

Base Transceiver Station performs the task of sending and receiving information to and from ME. Transceivers are inclusive, which are responsible for creating cell coverage as well as handling radio links. BTS also performs data exchange cipher and decipher task alongside Mobile Stations. Base Transceiver Station allocates radio resources used for the establishment of radio links.

Base Station Controller as the name implies, controls the system and deals with the management of radio bandwidth [45].

Network switching subsystem (NSS), comprises the core of Global System for Mobile Communication network. It deals with the linking of MS to public switched telephone network (PSTN) and vice versa. NSS's function includes user's price charging as well as the withholding of data reserved for routing [47]. Furthermore, NSS is responsible for the management of handoffs between BSSs.

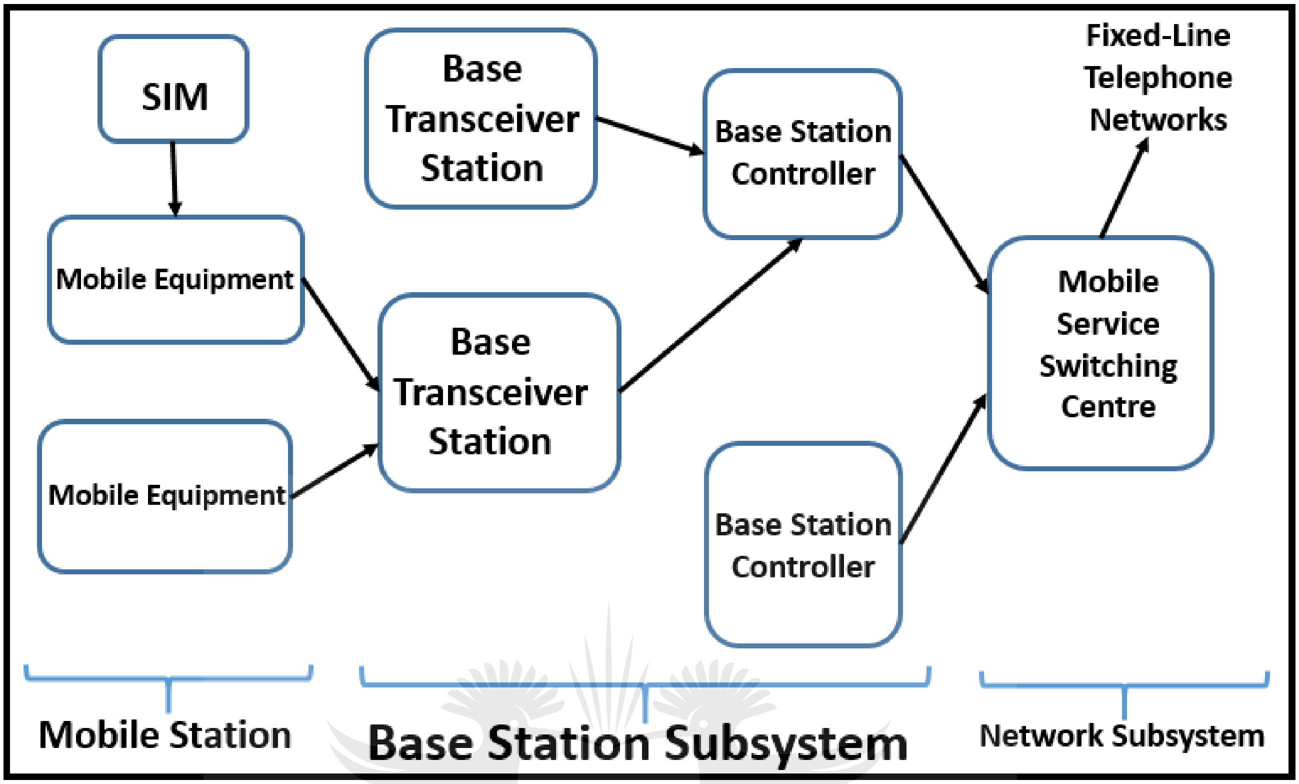

Figure 2.3 GSM Network Architecture

Being the most successful of the 2G standard, GSM's invention and advancement led to the development of the third generation (3G) technologies of Universal Mobile Telecommunication System (UMTS)

GSM is based on Frequency and Time Division Multiple Access (FDMA/TDMA) whereas IS95 standard uses Code Division Multiple Access (CDMA) technique. Second generation cellular networks are invented solely for voice call transmission. However, more researches led to the invention and rolling out of 2G standards which allows the transmission of both voice and data although the data transmission rate were far lower compared to dialups. Advancement on Global System for Mobile communication as well as Code Division Multiple Access standards lead to the formation of their 3G standards which are the 3G partnership projects (3GPP) and 3GPP2 respectively. The International Telecommunication Union-Radio (ITU-R) project on International Mobile Telecommunication IMT-2000 created an avenue for the creation of 3G network, the distinguished features includes data transmission rate in the region of 2 Mbps and vehicular mobility.

2.3.1 ISSUES AROUND MOBILE NETWORK CONGESTION IN GSM NETWORK

The most popular network in existence till date is the Global System for Mobile Communication. It all started in 1992, when the first GSM networks were launched [42]. Since then, it has been growing rapidly outside Europe to as far as Australia. By 1993, there were close to 35 Global System for Mobile Communication networks accessible in approximately 20 western countries [48]. This enormous growth continued massively ever since. Currently, it is the most widely used technology in mobile networks due to its simplicity, low charge rate, and ultimately good voice quality. A survey on world market share between existing standards towards the end of 2009 shows a result of the dominancy of the GSM network over the other similar networks as shown in figure 2.4 below.

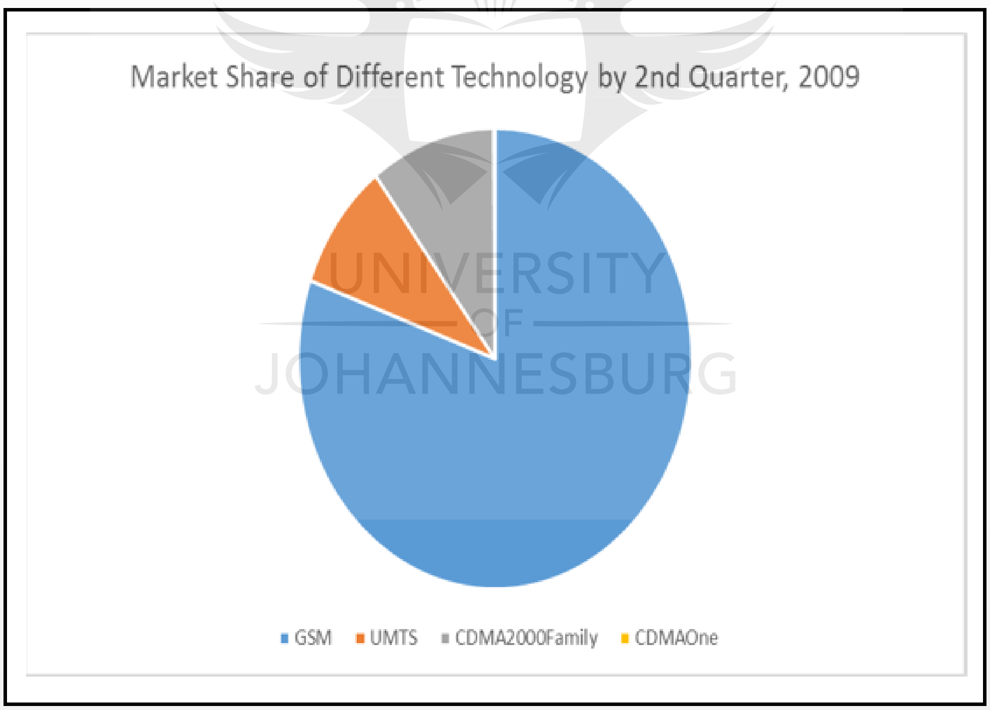

Figure 2.4 Survey on the Different Networks Dominancy in the World Market [36]

The rate at which mobile networks has been fully accepted and its growth especially that of GSM and other new generation network in South Africa is shown in figure 2.5. From the survey conducted by Deloitte Digital Company which states the population of the country as approximately 51.8 million as of 2011, the survey billed the total number of active SIM cards at 66.1 million which is approximately 1.35% higher than the active human population. Furthermore, the survey shows that the subscriber base covers 83% of the country landmass. This further shows the rate at which the mobile network has really grown since its introduction in the country in the early 1994 [68].

SOUTH AFRICAN POPULATION	51.8Million
TOTAL ACTIVE SIM CARDS	66.1Million
TOTAL UNIQUE SUBSCRIBERS	40.7Million
SUBSCRIBER BASE	83%
ESTIMATED ACTIVE SMARTPHONES	11Million

Analysis: Deloitte Digital South Africa
Source: Mobile Operator Financial Results, Stats SA Census 2011,

Figure 2.5 South African mobile network operator statistics result [68]

2.3.2 OVERVIEW OF MOBILE NETWORK CONGESTION IN GSM NETWORK

Cellular telecommunication especially the GSM has shown to be the fastest developing standard in the telecom world. New era mobile radio framework, allowing flexibility, versatility as well as advanced devices that support varieties of imaginative services to mobile users, are targets and primary interests of both telecommunication organizations, mobile users and providers as well. The idea of "whenever, anyplace" communications is exclusively turning into reality [49]. Billions of people rely on mobile telecommunication especially the GSM network for their daily businesses and living. Moreover, the use of the network is available to all humans irrespective of one's financial status. This has made the use of the network to be growing exponentially with more users on daily basis. Recalling that the radio spectrum allotted to mobile communication is inadequate, also the fact that users are growing at a geometric rate, and lastly, applications are turning out to be bandwidth intensive, the high population of users have led to massive congestion [34].

With the wide spread of various devices ranging from smart phones, androids and tablets, mobile networks easily get congested especially when lot of users converges at a particular place. The mobile network congestion is paramount specifically when a large scale event is held in a confined place which tends to make communication between participants and neighborhood occupants to increase [3]. The escalation in users of telecommunication network led to congestion in the network. Also, [18, 19, 20, 21, 55] ascertained that mobile congestion occurs when the inflow of connections is beyond the system bandwidth threshold.

Congestion issue is a grave worry since it degrades the Quality of Service (QoS). Basically network congestion is due to unavailability and inappropriate use of the radio resources.

2.4 RADIO RESOURCES MANAGEMENT

Two major determinants of the current state of mobile networks is the rate of new calls being blocked and handoff calls being dropped due to inaccessibility of channels. A base station can only allow a restricted number of users due to the limited amount of bandwidth available [50].

Congestion arises when cell nodes cannot offer enough bandwidth to accommodate new and handoff calls due to its bandwidth limit. Hence, creating issues of congestion. Radio resources management in mobile networks is a very intricate issue. RRM is an important factor in mobile networks which handles the scarce and inadequate radio resources. RRM management ensures that the required QoS levels are maintained [74].

Effective RRM is challenging majorly due to user mobility, scarce radio resources and physical properties of radio channels [7]. Since effective use of the available resources is vital, implementation of the maximal use of the available resources is very important [51]. Radio resource management procedures provides the effective management of the radio resources [35, 51]. RRM issue is intrinsic to cellular radio access networks and, systems and, thusly, it has been generally secured with the presentation of the distinctive mobile cellular frameworks [3].

Radio Resource Management in homogeneous networks embraces processes for optimizing the usage of available bandwidth [30, 52]. Functionalities including admission control, handover control, packet scheduling, and congestion control are well-known procedures in homogeneous networks. However, in heterogeneous networks, there is complications in the radio resource management processes due to the fact that varieties of RAT (Radio Access Technology) coexist in the same service area. Radio Access Technology RRM are anticipated to synchronize tasks which include power control as well as code management. However, other major tasks including admission control and Handoffs necessitate additional refined processes [52].

2.5 CHANNEL ALLOCATION

The enormous advancement of wireless telephony, coupled with the inadequate channels to handle all users demands for the effective reuse of channels [52, 53]. A resourceful channel allocation tactic is required which is expected to use the attribute of frequency reuse to intensify the accessibility of channels for maximal use from callers at any period in time [53]. Increase in request for cellular telecommunication coupled with inadequate frequency spectrum prompts the difficulty of channel allocation [53, 54]. Each cell has its own particular base station with Omni directional radiation design. Mobile stations demands base station for channel allotment [29, 53]. Base station will allot channels to mobile stations utilizing channel allocation methods. Part of channel allocation methods includes designation of channels in order to reduce call blocking probability and call dropping probability. Various channel allocation schemes are used in mobile communication system [54]. With the geometric rise in users, number of base station required to serve a vast region of land is a critical component. From the expense of administration perspective effective utilization of radio spectrum is highly imperative. Decline in the cost of base stations maintenance can be realized by judicious and effective reuse of the radio spectrum.

Several channel allocation schemes in existence are categorized namely [54];

1. Fixed Channel Allocation (FCA)
2. Dynamic Channel Allocation (DCA)
3. Hybrid Channel Allocation (HCA).

2.5.1 FIXED CHANNEL ALLOCATION

In Fixed Channel Allocation (FCA) arrangements, a secured quantity of channels are allocated to every cell as indicated by predetermined traffic demand and co-channel interference requirements. Fixed Channel Allocation system are exceptionally straightforward; be that as it may, they are uncompromising, as they do not adjust to varying load conditions. Keeping in mind the end goal to conquer these insufficiencies of FCA schemes, Dynamic Channel Allocation schemes were introduced [52].

2.5.2 DYNAMIC CHANNEL ALLOCATION

In this particular type of scheme, channels are positioned in a pool (typically brought together at Mobile Switching Center (MSC) or disseminated between different base stations) and doled out to users as required. Once the interference status level is maintained within its limit, any cell can make use of it. Immediately a call is terminated, the used channel is doled back to its position. At the expense of higher unpredictability and overhead controlling of message, Dynamic Channel Allocation provides adaptability and load versatility.

Nonetheless, DCA plans are less proficient than FCA especially during peak load situations [2]. In order to minimize its deficiency, some Dynamic Channel Allocation schemes reassign channels in such a way that ongoing calls have the luxury of switching in order to minimize the distance between cells [52]. Some other Dynamic Channel Allocation scheme borrow channels from nearby cells. In this kind of arrangement, channels are doled out to every cell as is typically done on account of Fixed Channel Allocation scheme. Nonetheless, in a situation whereby all channels are unavailable, borrowing another channel from adjacent cell is allowed provided it would not affect the interference level between channels [52].

2.5.3 HYBRID CHANNEL ALLOCATION

HCA scheme forms the combination of fixed channel allocation and dynamic channel allocation. This combination's end goal is to ensure taking favorable circumstances of both schemes. Hybrid Channel Allocation involves the partitioning of channels into disjoint cells namely, fixed and dynamic set. In this kind of partitioning, some set of channels are allotted to each cell on a fixed basis, and the rest of the channels are secured in a localized pool for dynamic assignment. The fixed set contains various channels that are doled out to cells peculiar to fixed channel allocation structure. At the point when a user requires a channel, and all ones in the fixed set are unavailable, channels in the dynamic sets will be consulted and used [52].

2.6 QUALITY OF SERVICE (QoS)

Quality of service, according to [1, 55] can be explained as an important key performance indicator that defines the effectiveness and productivity of an industry i.e. performance effects which determines how satisfied the user of a service is. Furthermore, in GSM mobile telecommunication system, the three main determinants for evaluating the quality of service of a network provider are accessibility, durability and good voice calling quality.

The mobile network users expect that they derive maximum satisfaction from the service they paid for [15]. Hence, providing a satisfactory QoS to mobile network users has hence turned into a tough mission for mobile network suppliers because of numerous reasons.

These reasons include the rapid increase in the number of subscribers and expansion of the network system. The mentioned incompatibility causes massive call dropping, call blocking especially at major metropolis, and hence poor QoS. Voice calls and date dissemination qualities of Global System for Mobile Communication are expected to be of high quality. In GSM networks, Gaussian Minimum Shift Keying Modulation (GMSK) method is employed for encoding voice data.

2.6.1 DEGRADATION OF QUALITY OF SERVICE

According to [14], degradation of QoS in mobile network has become rampant as a result of inadequate resources in the network. The resources include power, time slots and code which are being used and recycled by mobile subscribers. The moment a subscriber starts using the network, the QoS performance level gets degraded to co-subscribers. Generally, QoS are noticed to be extremely poor when there is a major congestion in the system. Furthermore, [2, 55] described congestion in GSM mobile network as a state whereby the quantity of connections in and out of a specific system is far greater than the limit the network system is programmed to accommodate, at that specific timeframe.

2.7 GENERAL PERSPECTIVE ON MOBILE NETWORK CONGESTION

There are two main ways that congestion can be viewed, which are listed;

- From operator's reference point, mobile network congestion is being interpreted as the inability to satisfy the traffic demand and hence some of the revenues are being lost.
- From user's reference point, mobile network congestion is being interpreted as the dropping and blocking of calls by the network providers.

2.8 BLOCKING

Blocking occurs due to the unavailability of accessible channels to allot to users in the base station [55]. Blocking in mobile telecommunication network are classified into two namely: New call blocking, which is explained as blocking new initiated calls while the other type is Handoff Call blocking which is explained as blocking continuing calls as a result of the user's mobility.

2.8.1 TYPES OF BLOCKING

- New call blocking: [49, 55] explained new call blocking as a situation whereby whenever a mobile subscriber decides to connect to a fellow subscriber, the user initiating the call will have to seek a free channel from a base station closest to his location. After getting an available channel, the mobile user is granted access. However, in an event whereby all channels are unavailable, the initiated call will not be allowed to gain access to the network and hence it will be blocked. The type of blocking is described as New Call Blocking.
- Handoff call blocking: [49, 55] explained further that a mobile user typically releases a channel it is using on two conditions, either when it is through with the call or a situation whereby the subscriber migrates to another cell while its call is still on. This process of migration between cells while the call is still on, is known as Handoff. If the channels in the cell it is migrating into, are unavailable, the call will be blocked. The type of blocking is described as Handoff call blocking. [19] described Handoff as a situation whereby an ongoing call migrates from its current cell into another cell and in the process of this transition, the call gets dropped. Thus the call can be dropped in a situation of insufficient bandwidth to support it during the inter-cell migration

Both new call blocking and handoff call blocking occurs due to channels unavailability.

Mobile network congestion management can be achieved naturally by increasing network capacity with more base stations but the solution is expensive and also the whole system capacity will be under-utilized at times of low traffic [55].

2.9 HANDOFFS

In wireless network, users' mobility tends to complicate the network system. Mobility results in active variations in connection quality as well as the level of interference in telecom network. Occasionally, it warrants the subscriber to switch from the base station it is using into another which has the better signal strength. The nature of this switching is pronounced as a handoff [55, 56]. Handoff is an important component of mobile networks which dictate the provision and maintenance of Quality of Service (QoS) [1] to the users and to support users' mobility. Forced termination of calls are the consequences of handoff failure [49].

Handoff happens when a current user in the network migrates from its cell into alternative cell due to the mobility of the mobile user [19]. Handoff enables mobile users to maintain their connectivity while migrating between cells. Managing handoff calls is a key issue in mobile networks. Cellular Network supports the handoff of users between various wireless technologies [58]. During this transition process, the call may get dropped. A current user of the network may be dropped due to inadequate bandwidth to support it in the cell it is migrating to. Handoff call blocking denotes the blocking of an ongoing call as a result of the mobility of the user. As shown in fig. 2.6, a user that is connected to Cell A will move to Cell B due to his/her mobility. However, if there is no available channel to accommodate the user in Cell B, the user will be blocked [55].

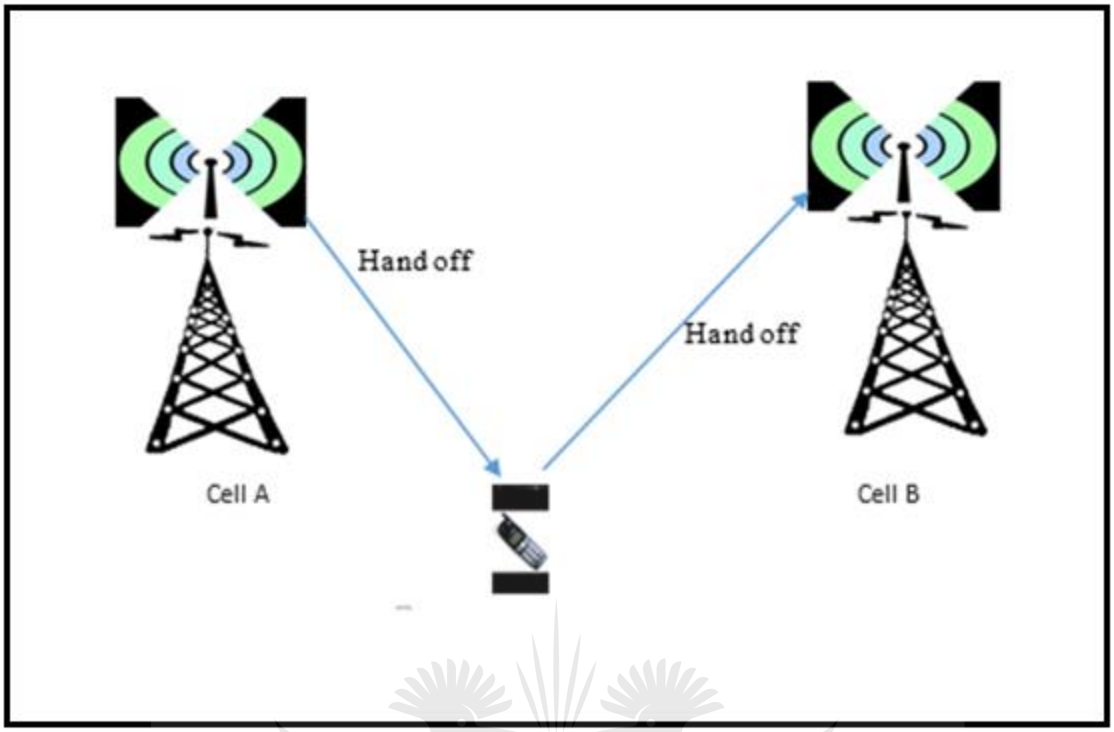

Figure 2.6 Schematic Illustration of a Handoff [55]

Handoff schemes can be classified according to the way the call is being handed off as well as the manner in which new channels are being arranged [55, 57]. At call level, Soft Handoff and Hard Handoff are the classes of handoff schemes.

2.9.1 SOFT HANDOFF

In this particular class, a mobile terminal uses multiple radio links to connect with the network via different base stations. The overlapping areas between several cells start the handoff process shortly before the actual handoff occurs. The old channel will be released only after the new channel has allocated a space for the ongoing call. Two channels are simultaneously used for a single call, hence reducing call dropping. Some of the mobile communication that uses soft handoff are the new generation CDMA-based system.

2.9.2 HARD HANDOFF

In the hard handoff class, the mobile terminal only communicates with a base station for the duration of the handoff process. The disengaging channel is disconnected before the new channel pick up the call. Hence, there is a probability of little disruption of ongoing calls during the transition [57]. Also there is a possibility of the call to be forcibly terminated if the old channel is disengaged earlier before the handoff transition. In a situation whereby there is an available channel, a handoff may fail once the link transfer response is slow. Second generation GSM mobile systems use the Hard Handoff scheme.

Furthermore, in terms of migration between various wireless access technologies, handoffs are explained into two types namely [49];

1. Horizontal Handoff
2. Vertical handoff

2.9.3 HORIZONTAL HANDOFF

Horizontal handoff is the type of migration between a single types of network. Based on this type of handoffs, Horizontal handoff can be categorized as either intracell handoff or intercell handoffs. In intracell handoffs [83], a user only migrates within a particular base station and this migration involves swapping radio channels to reduce interference. However, in intercell handoffs, a mobile station migrates into neighbouring cell which is totally away from its current base station [59].

2.9.4 VERTICAL HANDOFF

Vertical handoff is the type of migration that involves the user moving from one wireless technology into another. Vertical handoff mechanism allows the migration of mobile users to migrate between different networks including 2G, 3G, 4G networks without any hindrance. In the current age whereby new generation wireless network are in vogue, vertical handoff act pivotal to allow unbroken connectivity during the migration of users in an overlapping network setup. Downward Vertical Handoff and Upward Vertical Handoff forms the two types of vertical handoff [49]. In Downward Vertical Handoff, users only migrate to another network technology with higher bandwidth although at lower coverage whereas in Upward Vertical Handoff, users migrates into another technology of a bigger coverage which may tend to be of lower bandwidth. Figure 2.7 shows the classification of handoff based on migration between network technologies.

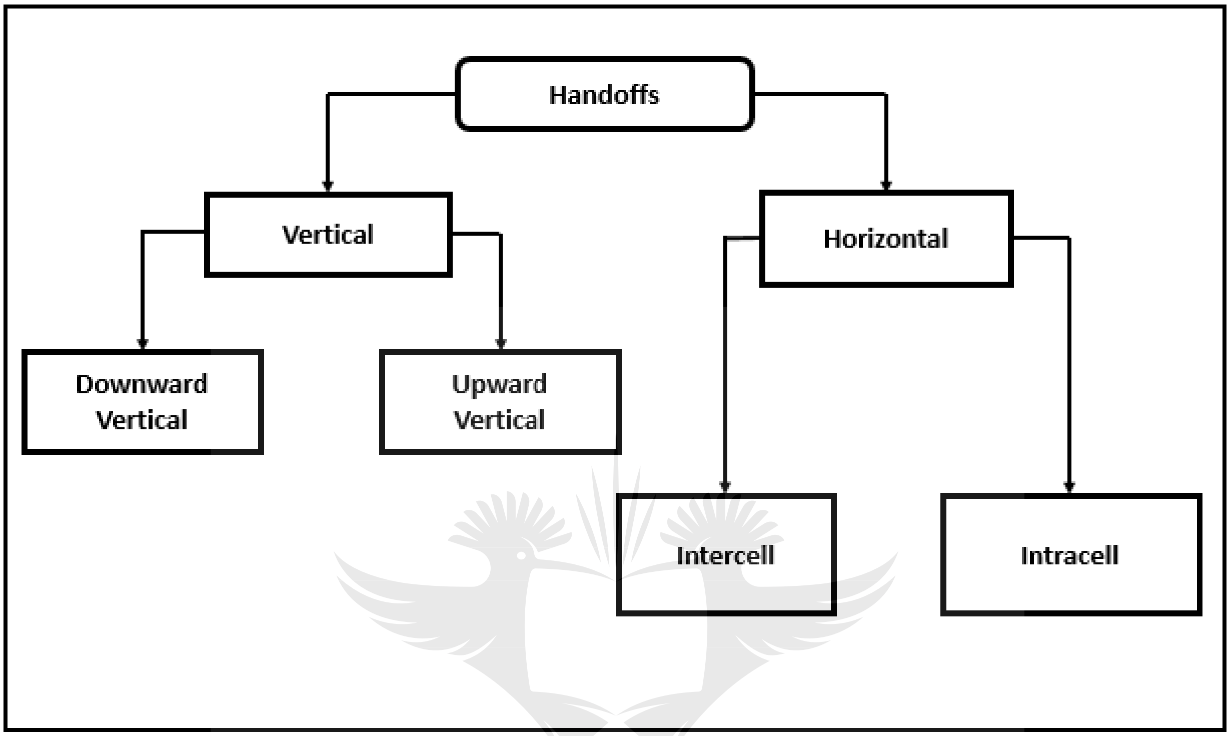

Figure 2.7 Handoff Classification Based on Migration between Networks and Cells

2.10 OVERVIEW ON MOBILE CONGESTION PROBLEMS.

Mobile network congestion being a major issue in mobile communication is yet to be properly managed with the various infrastructures and policies that operators offer, a fact that is continuously observed during critical situations and events e.g. torpedoes, earthquakes, New Year's Eve, public events, .etc. as well as hotspots like metropolis. Users strongly criticize the network providers on the failure to satisfy their communication [25].

According to [12], low voice quality of calls are as a result of numerous causes, one of which is network congestion. Network congestion in voice calls can leads to delay and jitter. In order to accomplish and satisfy users in voice calls, the end-to-end delay have to be maintained within a specific threshold.

It is observed that mobile network congestion problem is rampant within some specific periods (emergency, accident and hotspot periods) when the system resources are heavily overused or at a particular period of the day [8]. Mobile network providers usually design the radio networks facilities such as base stations in such a way as to satisfy load request of the location, putting into consideration, the load demand in those locations .i.e. the network providers build base stations based on the forecasted traffic requirements in that specific area. The population and the mobile network dictates and decides the load demand of the geographic locations. The traffic demand differs from one location to another [3].

Typically, there are some specific period of a typical day when the traffic demand swells up. This period is known as the peak period or busy hour period. Peak time load occurs in a situation whereby the network system are exhausted beyond their capacity threshold. The sole aim of network system dimensioning serves to ensure that the network delivers especially at peak periods. Expanding the system capacity is unnecessary in light of the fact that the system can be vacant at a specific time of the day and at other time of the day, it can be massively overloaded especially during the peak hour or busy hour period [14]. Thus, network system expansion is only required occasionally when the threshold for busy hours is being overreached frequently and congestion is becoming alarming.

2.11 BUSY HOUR PERIOD

The busy hour/peak hour of a system is the period in the day whereby network endures the maximum load. Peak hour measurements determine the performance and strength of a network system [2]. The busy hour measurement is used by network operators for dimensioning the equipment in GSM network especially in terms of the threshold level a geographical location base station can handle. The traffic load processed by a network system varies depending on a particular time of the day, week or month. Network systems tend to be massively congested only during some hours daily. This particular hour of the day is the period when there is a high tendency that new calls that are being initiated by users can be blocked and handoff calls can be easily dropped [14].

2.12 EFFECTS OF CONGESTION OF MOBILE COMMUNICATION NETWORKS

For many years now, massive increment in the number of mobile users has been continuously felt and this has led to massive congestion. Some of the impacts of mobile network congestion comprises undesirable delay in packet data, interference in access and selfish channel usage and most importantly, new call blocking and handoff call dropping [16]. Figure 2.8 shows a typical notification from network operator to a subscriber that is trying to access the network (make a new call) but could not due to network congestion

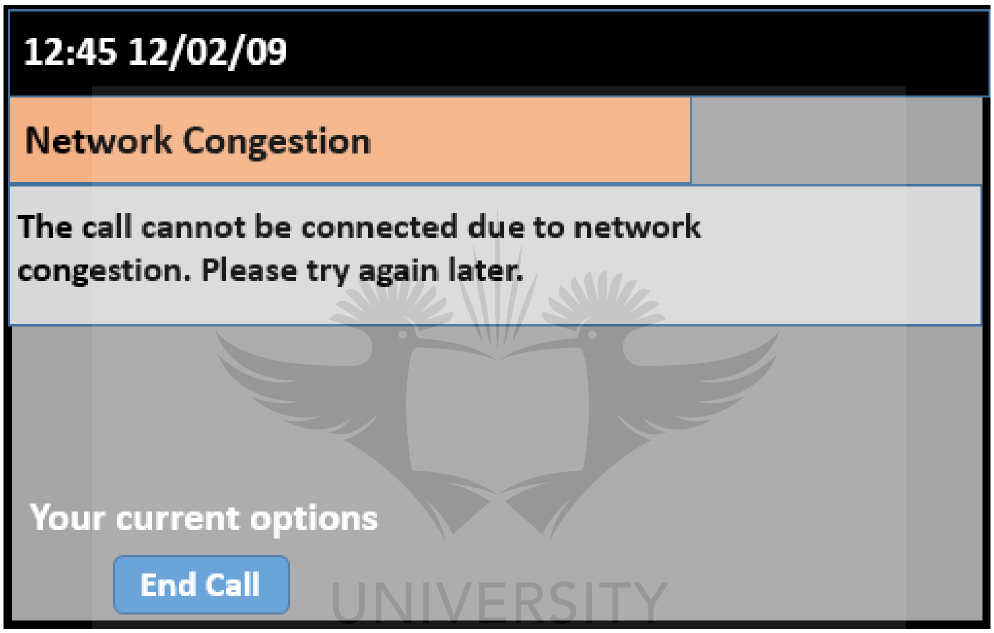

Figure 2.8 A Typical notification of call blocking due to network congestion [69]

Figure 2.9 shows the Telecom Regulatory Authority of India report for April –June 2015 on the high call dropping rate due to network congestion for some of the major cities in India.

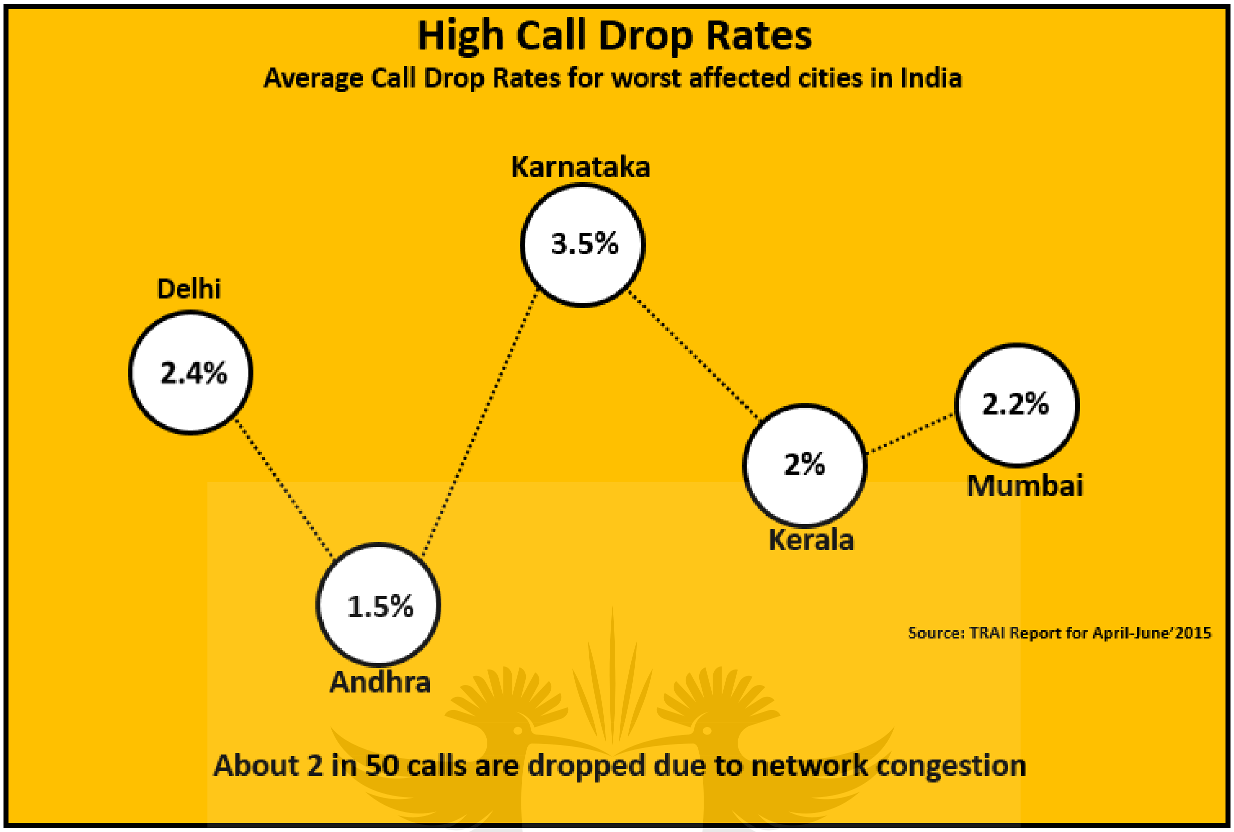

Figure 2.9 TRAI Report on Call Dropping Rate in Indian's Major Cities [67]

The problem of network congestion is majorly due to poor network resources management [24]. Furthermore, over-utilization of the network can lead to system malfunctioning.

So as to reduce New Call Blocking and Handoff Blocking which are consequences of mobile network congestion, it is essential that mobile network congestion should be managed for the sustainability of the system.

2.13 CONGESTION MANAGEMENT TECHNIQUES

Wireless communications have experienced massive development remarkably in the past decade. Networks such as global system for mobile communications and universal mobile telecommunication service are currently developing towards integrated networks, which will bolster various carrier administrations utilizing distinctive access technologies [35]. Several techniques of managing mobile network congestion have been widely researched and implemented. Mobile network congestion can be controlled by suppressing the traffic demand.

The traffic demand varies proportionally with average rate of call arrival and the average duration of the calls. Hence minimizing the average call duration will go a long way in minimizing traffic load, thereby reducing network congestion.

The problem of congestion according to [3], is typically experienced at some specific locations such as malls, markets and some specific time (peak period/busy hour) whereas the network system is underutilized at other times. Hence, it is appropriate to find a means to flatten the traffic demand curve. Making channels available for users and at the same time, discourages users from continuing non-urgent communications especially during busy traffic at the expense of pending users is a priority.

2.14 RELATED WORKS

Two of the most essential measures in cell systems using GSM technology include rates of new and handoff that are blocked and dropped respectively as a result of inaccessibility and unavailability of channels.

The various techniques of controlling congestion employed so far in all systems have always in a general view, been focused on two principles. These ideologies are either to dismiss extreme load in order to avoid overload from happening or redirecting overloads when it happens [2].

Lynn *et al.* [81] proposed the Automatic Call Gapping (ACG) and the Token Bank (TB). In Automatic Call Gapping (ACG), call attempt rates are reduced by allowing a single call attempt over a specified gap interval. When the congestion rate is high, the ACG is activated and call attempt reduction commences. This process last for a specified gap duration, specifically over the congestion period. The Automatic Call Gapping scheme is not adaptive and effective enough as a means for controlling mobile network congestion especially at peak period when there are numerous pending calls.

The Token Bank (TB) on the other hand is a scheme that involves regulating the inputs into the network system in order to protect the system from being overloaded. The scheme makes use of token bank capacity and the token bank rate. Tokens are produced intermittently into the bank. Each source controls the token bank which generates call requests. Hence, when the bank is saturated, the bank tokens are blocked and disengaged. However, in an event whereby the token bank is not empty, a call request is granted which will lead to a removal of a token from the token bank, otherwise the call request is rejected.

Oyebisi *et al.* in [31] proposed a congestion control scheme that involves a buffer management procedure which involves accepting handoff calls temporarily using guard channel principle. The scheme reduces handoff call dropping probability but increases blocking of incoming calls once there is unavailability of free channel. The scheme makes use of buffer management technique to accept handoff calls temporarily in case of unavailable free channel to minimize Handoff Call Dropping Probability (HCDP) under the notion of guard channel principle. However, the scheme blocks incoming calls when there is no free channel.

Nasser *et al.* [32] proposed a mobile network congestion scheme that involves the combination of call admission control with bandwidth adaptation. In this scheme, there will be blocking of a new call once ongoing calls numbers are the same or more than a stated threshold value. Also, there will be blocking of new calls once there is unavailable bandwidth in the cell. There is massive reduction in Handoff Call Dropping Probability in this scheme, although new call blocking probability rate is high.

Sajal *et al*. [33] proposed the use of channel borrowing in balancing loads. In channel borrowing scheme, there is an allocation of channels to every cells based on ration. When a new call finds all nominal channels occupied, borrowing of a channel from a neighboring cell once co-channel interference constrictions are not violated is allowed. But, if the call numbers are higher compared to channels numbers, excess calls would not be allowed to gain access. Channel borrowing aims at reducing interference, borrowed channels will be assigned to the co-channel cells of the lender and not that of the borrower. With this scheme, there is high tendency of interference especially during high traffic period.

Kumar *et al.* [22] proposed a dynamic mobile controlled handoff scheme. The work focuses on Quality of Service. In this scheme, there is an initiation of handoff technique whereby Received Signal Strength from serving base station drops below its threshold value using Dynamic Channel Assignment (DCA) information regarding the QoS parameters, already provided at the Base Stations. By considering the various parameters, Mobile System (MS) source for the most suitable base station to make a successful handoff, a base station that will provide a good clarity of voice. Unforeseen terminations of calls are generally perceived to cause degradation of QoS. Hence, handoff requests are given high priority of allocating channels more readily compared to new calls. This scheme reduces the hand off blocking probability although it has the disadvantage of increasing new call blocking probability.

Thakurta *et al.* [28] proposed a Dynamic Pricing Scheme called Priority Based Tree Generation (PTGM) to curb mobile network congestion. In this scheme, subscribers calls are divided into multiple priority levels and call requests are scheduled in form of a tree, hence establishing a path sequence that is unique. In this scheme, priority increment procedure is used to solve the problem of lower priority calls. Also more calls are admitted in this PTGM scheme, thereby reducing new call blocking probability although at a disadvantage of a higher price and handoff call probability.

Alarape *et al.* [2] proposed a scheme that involves the combination of dynamic load balancing technique with Call Admission Control. Load balancing is a technique for managing system performance. It is used in assigning tasks to processors in multiprocessing computing environments among various other applications. In this scheme, calls that are usually dropped by ordinary CAC schemes are rerouted to another cell within the base station controller (BSC) by this dynamic load balancing technique. The dynamic load balancing strategy has three phases which are information policy, selection policy and the migration policy. The selection policy phase of the CAC and dynamic load balancing combination finds the most appropriate cell within the base station to transfer new calls/hand off calls to. Also, this combined scheme allows the cells in the station to process more calls, thereby making the network performance to be more effective and efficient.

Al-kishrewo *et al.* [19] proposed a network congestion management scheme which involves integrating dynamic pricing into guard channel. Guard channel schemes (GCS) focus on moderating amid new call blocking probability and handoff call blocking probability with more reference on marginally escalating the new call blocking probability of the cells. From the proposed combined scheme, the price is not fixed and the current traffic load dictates the price. Hence, in this scheme, the price is adjusted dynamically when rate of call arrival is high and guard channels numbers cannot meet the probability threshold which will lead to a degradation of QoS.

Sharmal *et al.* [34] suggested an adaptive guard channel based CAC scheme. This scheme solves the problem of non-uniform traffic demand in various cells within the base station. In this scheme new call blocking probability and hand off call dropping probability becomes massively reduced in cells during high traffic. The guard channel policy scheme prioritizes handover calls over the new calls. Hand off call dropping probability reduction can be achieved by incrementing the guard channels.

Falowo [35] proposed a Joint Call Admission Control Algorithm JCAC whereby multiple Radio Access Technologies RATs are selected for a new call when none of the individual RATs that are available has sufficient basic bandwidth units to process the new call. The JCAC schemes make use of session splitting and multiple selections between several RATs to minimize both new call blocking probability and hand off call dropping probability in heterogeneous networks. This scheme involves admitting a new incoming call (that cannot be processed due to congestion in the network) into two or more RATs.

Jiongkuan *et al.* [17] also propose the integration of pricing into CAC scheme, whereby mobile network congestion is controlled by adjusting the price dynamically based on the current traffic load. At busy period, mobile subscribers are charged at a high rate so as to reduce the number of users. As load rate soars, the higher the price charged to users. In this proposed scheme, channels are available to all users and no control mechanism to phase out any user. Various base stations set their pricing using the present load condition as a determinant. Hence, it is the user decision to use the network once the user is comfortable with the price. Using this scheme, during heavy traffic period, both new call blocking and hand off call blocking probability are massively marginalised but at a serious disadvantage to the users due to the high price.

Manaffar *et al.* [4] also proposed a scheme that involves combining dynamic pricing scheme to Call Admission Control very similar to [17]. The dynamic pricing scheme is as flexible as [17]. The pricing also depend on the traffic load, it can be very low during off peak period so as to encourage more users into the network and the price soar during high traffic in order to reduce the congestion.

Gan Liu [60] *et al*. proposed guard channel scheme in which some channels are fixed and purposely reserved for handoff calls. This scheme tends more to handle handoff calls due to inadequate and unavailable channels in adjacent cells.

The aforementioned proposed schemes have the deficiency of adapting to varying load pattern due to their rigid nature.

2.15 MARKOV DECISION PROCESS (MDP)

Markov decision processes have been discovered and being used since 1950. They give a calculated framework for molding decision-making in circumstances whereby results are fairly irregular as well as being partly controlled by the decision maker. Markov decision processes have been highly valuable in the study and solution of various optimization problems through dynamic programming and reinforcement learning. In recent years, Markov decision processes have found application in research areas including robotics, mechanical autonomy, finances etc. MDPs offer analytical models for performance evaluation and reliability analysis. Specifically, MDPs are discrete time unpredictable control processes. At each point in time, the process is in a particular state ***s***, which leaves the decision maker the option of making a decision based on the available choices in the specified state ***s***.

The Markov decision process is also used to design analytical models that are state space [52]. A Markov chain therefore consists of a set of states and labelled transitions between the states.

2.16 *M/M/m/m* QUEUING MODEL

The *M/M/m/m* queuing model also known as m-server loss system model is used to define call arrival distribution in telecommunication systems. *M/M/m/m* queuing systems assume that arrival rate follows a poisson distribution while service rate follows exponential distribution [64]. In the *M/M/m/m* queuing model there are limited resources available in the system and there are a limit to the number of calls that the system can support. Calls arriving in excess of the maximum capacity are rejected. The figure 2.10 below shows the state transition diagram of a typical *M/M/m/m* queuing model, with λ representing arrival rate whereas the symbol μ represents the service rate [52].

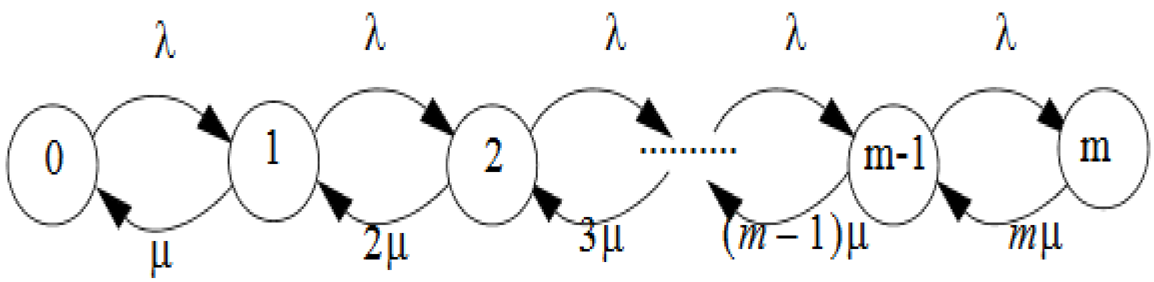

Figure 2.10 A typical *M/M/m/m* State Transition Diagram [63]

Steady state transition illustration for the *M/M/m/m* taking bandwidth reservation into consideration appears as shown in figure 2.11 [52, 78]. Where C represent the total capacity on the system and R the capacity reserved for handoff calls. With λn denoting the new calls arrival rate while λh represents the handoff call arrival rates, and μ representing the average service rate for both new calls and handoff calls.

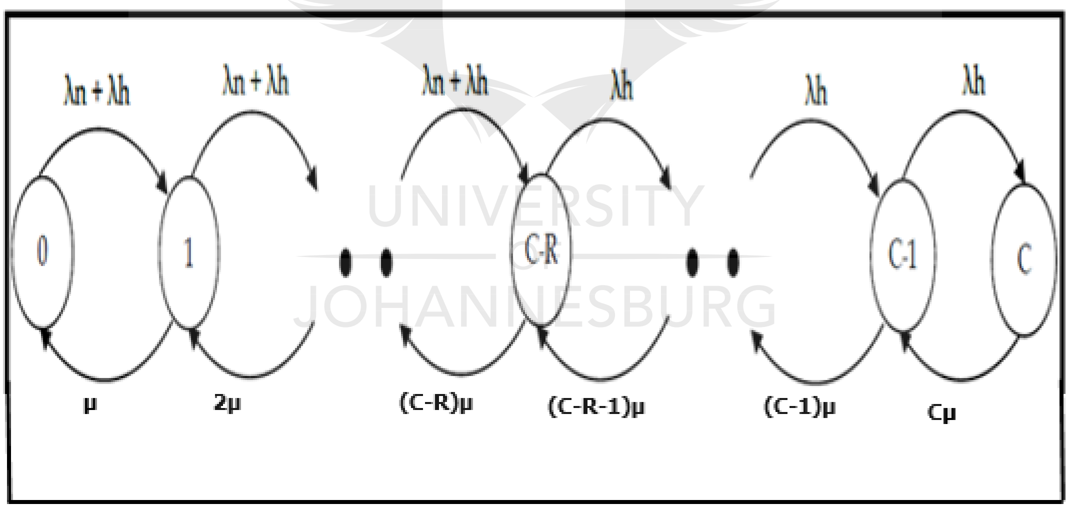

Figure 2.11 *M/M/m/m* Transition State Diagram with Bandwidth Reservation [52]

2.17 TWO-DIMENSIONAL MARKOV CHAIN BASED ON *M/M/m/m* QUEUING MODEL

In order to evaluate the CAC and CDC in GSM network, a two-dimensional Markov model is found desirable. The two-dimensional model is able to employ different channel holding times for new calls and handoff calls; which is not the case in the one-dimensional model [65]. Let C represent the total capacity on the system and R the capacity reserved for handoff calls. With λn and λn the arrival rate of new calls and handoff call correspondingly, and μ is the service rate for both new calls and handoff calls. The state diagram for a two-dimensional Markov model based on *M/M/m/m* queuing system is presented in Figure 2.12 [66]

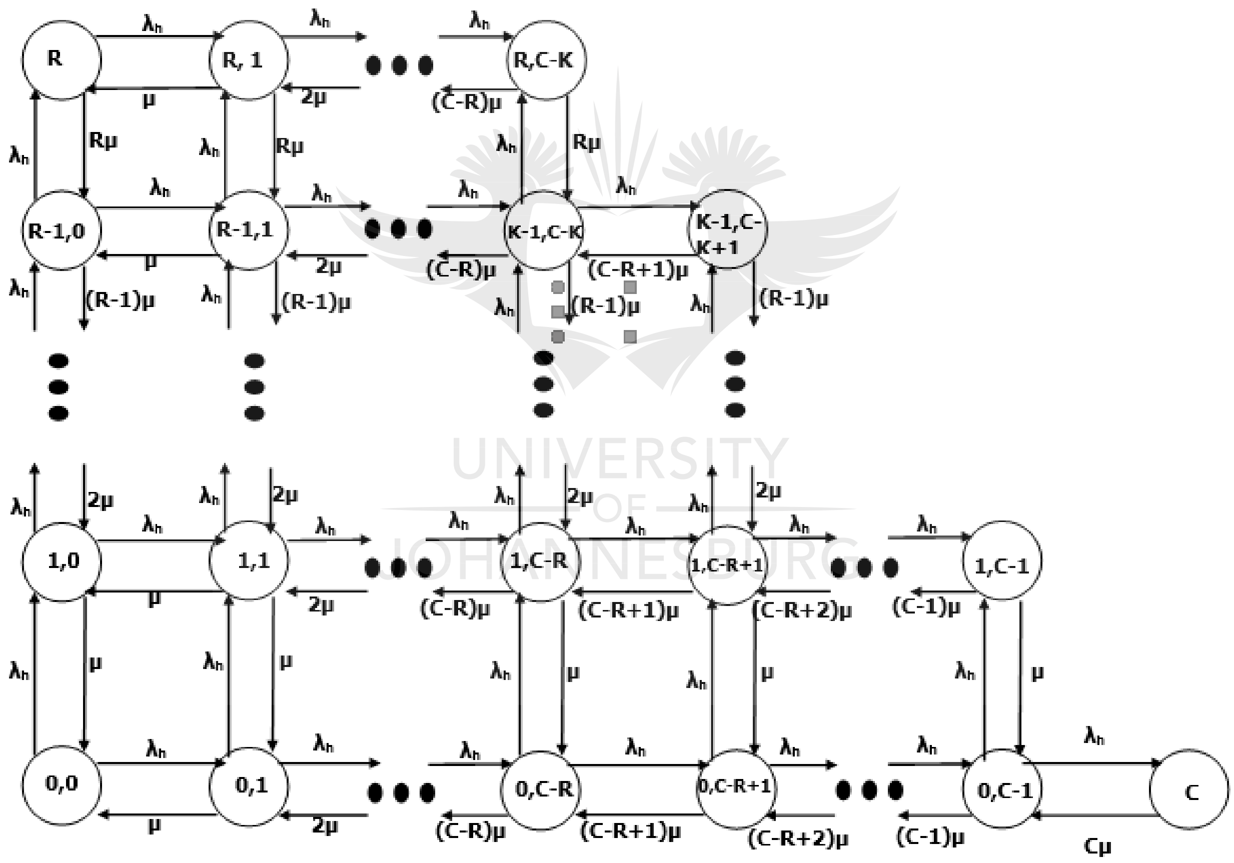

Figure 2.12 2-Dimensional Markov Chain Model with Bandwidth Reservation [52]

2.18 POISSON PROCESS

Poisson processes are unpredictable processes which are being utilized for demonstrating irregular large scale events that take place autonomously of each other. In tele-traffic hypothesis, poisson processes typically describe the calls and packets arrival rates. Furthermore, poisson process finds application in feasible models whereby a large number of users are making the calls and sharing the data.

2.19 AVERAGE SERVICE TIME / TRAFFIC INTENSITY

Average service time as well as the traffic intensity are parameters applied in Markov chain models. The service time is described as the channel holding time or channel occupancy. The average service time is usually expected to follow an exponential distribution [52]. In this analysis, assumption is made that the average service time relates to the duration of a call. Let D and μ denote the mean call duration/channel holding time and mean service rate respectively. The mean call duration D is inversely proportional to the service time μ by the equation [62];

$$D = 1/\mu \qquad (3.1)$$

Hence, $1/\mu n$ denote average service time for new calls

$1/\ \mu h$ denote average service time for handoff calls [55].

The traffic intensity ρ, is defined as the ratio of mean arrival rate to mean service rate. It can either be for new calls or in terms of handoff calls and it is denoted by ρn and ρh respectively [61, 62].

$$\rho n = \lambda n/\mu \qquad (3.2)$$

And

$$\rho h = \lambda h/\mu \qquad (3.3)$$

2.20 CHANNEL RESERVATION SCHEMES

2.20.1 RESERVED CHANNEL SCHEME

Since it is more desirable not to drop a handoff call than a new call, separation of the new calls and handoff calls are discussed in this subsection. There are some channel allocation schemes for accepting new calls and handoff calls. In the RCS scheme, there is a reservation of some channels purposely for the handoff calls. Blocking of handoff calls can only happen when all channels are unavailable, while blocking of a new call can still happen when there are several available channels. Hence, there is a higher priority for handoff calls when compared with new calls. Thus, reserved channel scheme allows the decrease in the call dropping probability.

2.20.2 NON PRIORITIZED SCHEME

Unlike in the RCS scheme whereby some channels are kept purposely for handoff calls whereby handoff calls are accorded higher priority. In the Non Prioritized Schemes, new calls and handoff calls are prioritized as equals. All incoming calls either new calls or handoff calls will gain access to the network only if there are available channels to accommodate them.

2.21 BANDWIDTH RESERVATION POLICY

For a good QoS requirement to be achieved and in order to cater for higher priority calls, a bandwidth reservation policy is applied. Prioritization of some particular types of calls may be required. To achieve this, some channels (resources) are reserved for the higher priority calls. This is indicated as guard channels. The bandwidth reservation policy follows after the guard channel approach explained in the previous chapter. For example, it is more annoying to customers to have a call dropped than blocked [55]. In view of the fact that customers are more sensitive to call dropping than to have their new call blocked, handoff calls are given higher priority over new calls. Therefore, more bandwidth is reserved for handoff calls than for new calls [61] [37]. Bandwidth reservation scheme is employed in the research work. The aim of using this policy is to set a threshold on the system that will support a particular category of calls [61].

Figure 2.13 illustrates the guard channel approach considering handoff and new calls. Given that total channel number is C_N and T_N is the threshold for new calls, and T_H is the threshold for handoff calls where $C_N=T_H$; then the guard channels are (C_N-T_N). A new call is accepted if the whole channel utilized by existing calls is less than T_N, handoff calls are allowed into the network only if there is available capacity.

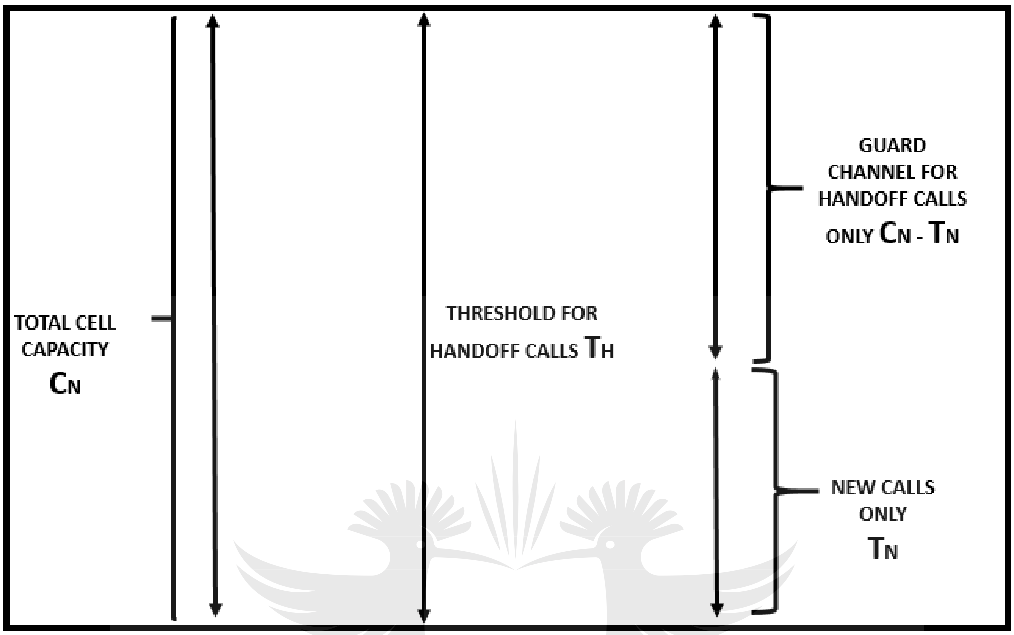

Figure 2.13 Bandwidth Reservation Channel Scheme Illustration

For the algorithms, a channel reservation based CAC scheme was employed, whereby handoff calls are prioritized over new calls due to users' sensitivity to handoff calls. In view of the fact that customers are more sensitive to call dropping than to have their new call blocked, handoff calls are given higher priority over new calls [5, 61]. Some bandwidths are reserved for handoff calls alone.

2.22 CALL ADMISSION CONTROL

Signal quality degradation, interference and network congestion has been a real issue for Global System for Mobile Communication (GSM) particularly with the fact that users increases often rapidly. It has been a real issue in providing a decent Quality of Service (QoS) [1] to the subscribers especially during periods of high network traffic. It is essential to maintain a certain level of quality in handling mobile network congestion. Fortunately, Call Admission Control is a tactic that can provide credible QoS by restraining connections into the cellular network thereby reducing network congestions, dropping of calls, interference and other QoS problems. Furthermore, CAC handles different types of services and it deals with resource allocation as well as resource management [73, 77]

Call Admission Control (CAC)[12, 27, 28, 55, 52, 73, 77] is a strategy that centers on limiting the number of mobile networks users so as to suppress traffic congestion, call blocking, call dropping as well as providing a good QoS to network subscribers [2]. It has been used widely for mobile network congestion management and improving the Quality of Service. Resources allocation and call admission control are essential radio resources management functions for proper QoS provision [74, 80]. CAC schemes consider the network users in the adjacent cells, as well as users in the particular cell being under consideration, so as to decide whether to admit or block the new call [6]. Advances in technologies has made it possible for handy computers and devices including notebooks and tablets to have wireless interfaces hence allowing massive networked communication and mobility. This trend has led to massive mobile congestion, hence making mobile network congestion a key issue.

Initiating new users into the system typically intensify the rate of interference in the network [10, 55, 77]. CAC scheme provide a suitable means of accommodating / blocking prospective callers based on the network condition. Resources allocation is a paramount feature of the CAC scheme. The core focus of Call Admission Control scheme is safeguarding ongoing calls by denying new users once there is no available channel to cater for the new calls as they approach the network. As a result of the ever-growing network subscribers coupled with relative inadequate resources in the network, Call Admission Control is essential so as to preserve a good QoS [75]. The QoS may be with respect to Connection Level, Packet Level and Packet Loss and the QoS may be with respect to packet level including Delay and Jitter experience by subscribers [79].

The basis of CAC is such that; assuming there are (N-1) calls already in the network, when N^{th} new user approaches the network, it considers the available resources in the network. If there is enough resources to admit the N^{th} user such that it will not compromise the QoS requirement status of the ongoing (N-1) users as well as the incoming N^{th} user, then N^{th} user will be admitted into the network. Otherwise N^{th} user will be blocked and denied connection to the network.

Literarily, the aim of CAC is to allocate resources to users so as to produce a reasonable QoS [55].

The admission decision in Call Admission Control is made in real time making the process efficient [55]. Furthermore, Call Admission Control scheme decreases handoff call dropping probability as well as system overload probability irrespective of the numbers of pending users and ongoing calls. However, Call Admission Control is however complicated in new generation wireless networks majorly as a result of the fact that users are mobile and fluctuating class of network link and nodes. Call admission control (CAC) is a means of minimizing New Call Blocking Probability (NCBP) and Handoff Call Dropping Probability (HCDP) [17, 55].

CAC schemes are simple and reliable means of managing network congestion. They yield less programming difficulties and allow multiple connection requests to be handled simultaneously. Also it tolerates less power consumption [13, 55].

Call Admission Control is an efficient method for quality management of network resource [8, 74, 75]. When a new user initiates a call in one cell, the new call will trigger a request for a channel. In a situation whereby there is no available channel, the call will relocate to neighboring cell, if it hits a gridlock in the neighboring cell, the call will be blocked.

As shown in figure 2.14, when a call (either new call or handoff call) tries to access the network, the request will be granted access only if there is available bandwidth to cater for the request. And if there is no available channel to cater for the call request, the call will be blocked/dropped [55]. The criteria for the network inability to accept the call arriving depends on the cell capacity in terms of the threshold for handoff calls *(Th)* and threshold for new calls *(Tn)*. Once it has reached and exceeded this threshold, the calls will be dropped and blocked. The process and the equations are well explained in section 3.8.

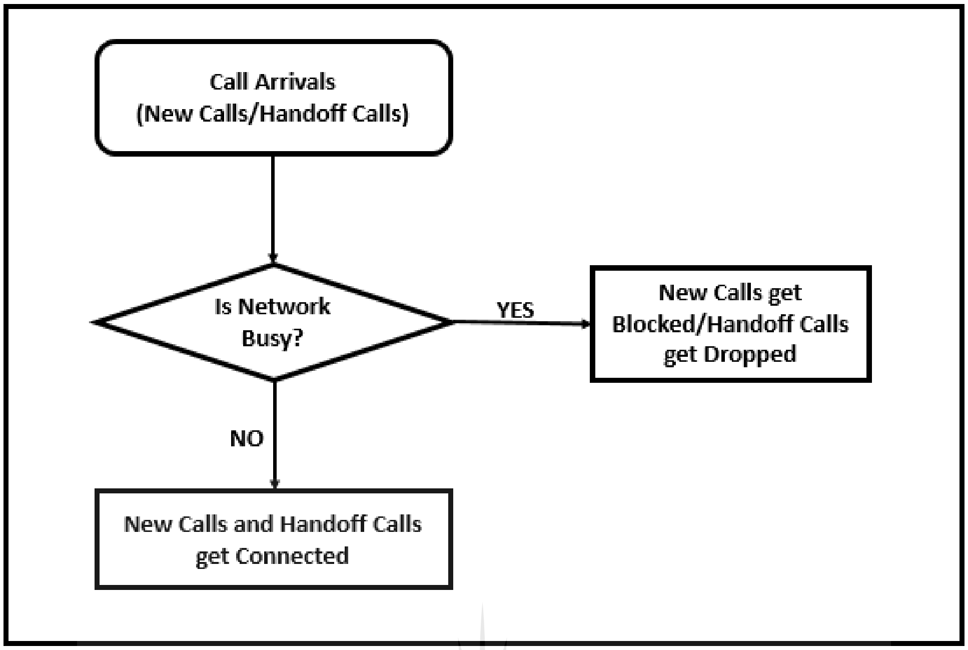

Figure 2.14 Call Admission Control Illustration [55]

Table 2.2 shows the new call blocking probability, handoff call blocking probability and network resources management for various mobile congestion management schemes.

Table 2.2 Comparing Various Mobile Congestion Management Schemes [55]

Congestion management methods	**Congestion Parameters**		
	New Call Blocking	***Handoff Blocking***	***Resources Management***
Cell splitting	Moderate	High	Low
Channel Allocation using channel borrowing	High	Moderate	Moderate
Priority Token Bank	High	High	Moderate
Call Gapping	High	Low	High
CAC	Low	Low	High

2.23 CONCEPT OF CALL DURATION CONTROL (CDC)

Long duration calls in the network has been adjudged as a major reasons for mobile network congestion. As mentioned earlier, average call duration is proportional to the traffic demand. A situation whereby a user is using a channel for a long period will automatically make the channel unavailable for subsequent users. Hence, there is a need to curb the excesses of long duration users. Call Duration Control (CDC) scheme involves controlling the duration of users of the network channel by sending a termination notice in advance to long duration callers in order to make channels available for pending new users that are unable to connect due to unavailable channels especially during peak periods. Thereby creating channels for calls that that would have been dropped or blocked.

2.24 THE COMBINATION OF CAC AND CDC

Although CAC scheme is a brilliant scheme in handling mobile network congestion due to its ability to suppress call blocking, call dropping and good resource management. However, with the rate of mobile users which is increasing at an alarming rate, call blocking and dropping occurrence have been more frequent in cellular networks, (CAC) scheme alone has been incapacitated in managing the congestion menace. The concept of a channel reservation based Call Admission Control (CAC) [81, 10, 12] and Call Duration Control (CDC) [3] are combined effectively to make the system flexible, efficient and effective. Also a principle of call on hold [3] in which when a new intending user finds that all nodes are busy, the user will not be blocked instantly but will be held in a line pending the time some nodes to be made available.

The integration of Call Duration Control (CDC) having call-on-hold principle scheme in the CAC process will produce efficient and effective use of the network. The combination of the two approaches will help in managing congestion even at peak periods in the major metropolis. Network congestion management using CAC and CDC combination was the main focus of the research work.

However, it should be noted that the combined scheme has some shortcomings. The termination of calls may not be generally accepted by some users who have to make high duration calls.

Some of these users may prefer to go for other options in terms of switching network operators and some will prefer to be charged at a higher rate rather than having their calls terminated.

2.25 CHAPTER SUMMARY

This chapter has given a brief discussion on telecommunication history till date. The various types of telecommunication techniques in existence with more focus on Global System for Mobile Communication. An overview was presented on mobile network congestion in GSM networks as well as network congestion problems, effects and issues around it. The subjects of Radio Resources Management, Channel Allocation was presented. Quality of Service (QoS) and its degradation was also discussed. Blocking and busy hour traffic was deliberated on as well. Handoffs and its types and classes were discussed in the chapter. Some mobile network congestion management techniques were compared. Some of the related work on mobile network congestion management schemes were identified including Call Admission Control Schemes that involves pricing. However, most of the related work and schemes are no longer appropriate to manage congestion as the number of users have increased massively over the years. This thesis involves the integration of Call Duration Control (CDC) having call-on-hold principle [3] scheme in an effective CAC scheme. The combination will produce efficient and effective use of the network and increase revenue generation. The combination of the two approaches will help in managing congestion even at peak periods in the major metropolis. New Call Blocking probability (NCBP) [17] and Handoff Call Dropping Probability (HCDP) [17] will be significantly reduced at peak period.

3. METHODOLOGY

3.1 INTRODUCTION

This section manages the methodology used in the research work. Cellular networks are characterized by static base stations which separate the geographical area into cells. All radio correspondence is between these base stations and the mobile equipment which are the clients' devices. Usually, mobile transceiver stations overlap in such a way that mobile users on different territories can typically choose which cells that have the highest reception signal to connect their call. Every call approaching the cell requires a definite range of bandwidth called a channel and each cell has a specific number of channels that they use to take calls.

As discussed in chapter two, all calls approaching mobile transceiver stations are classified into as either new calls or handoff calls. When there are numerous users in the network or approaching the network especially during period of high traffic, handling both new calls and handoff calls are major issue to network providers. Also, some network users due to their lifestyle of being affluent or due to the nature of their businesses which demands long duration of calls, holds on to a particular channel for a long duration at the detriment of new users. CAC scheme cannot solve congestion problem due to the selfishness of some of these mobile users who cling to the channel they are using for a long period at the expense of new users. Hence, the combination of a reserved channel based CAC scheme with a Call Duration Control to manage the congestion menace either during low load, medium or high load situations was deployed. The CDC algorithm serve as the introduction of termination notice to users of long duration calls in the network so as to free up the channel to new users. The performance of the combined scheme is compared to the ordinary channel reservation based CAC scheme, the results shows an improvement in the management of congestion.

3.2 SIMULATOR OVERVIEW

MATLAB 2012a was used for this research work. Nevertheless, this is an experimental simulation which does not involves feasible data in terms of users and cells.

3.3 ANALYTICAL MODEL

This research work made use of an analytical method. A two-dimensional Markov chain model based on the *M/M/m/m* queuing system explained in section 2.17 that takes into consideration both handoff and new calls is applied in the research work. Markov decision process [61] explained in section 2.15 and 2.16 is used to evaluate the system performance of the CAC algorithm and the CAC-CDC combination algorithm. Also, the assumption that call arrival rates follows poisson distribution explained in section 2.18 is adopted.

3.4 HOMOGENEOUS NETWORK ENVIRONMENT

It has been noted that GSM network is homogeneous in nature i.e. a single network architecture. A given coverage area will therefore contains a single RAT. RATs has maximum amount of radio resources that it can offer, which represents the available capacity of the RAT. In practice, radio resources can be offered in terms of timeslots, code sequence or frequency channels. In this analysis, the capacity offered is considered in terms of Basic Bandwidth Unit (BBU) [52] where each RAT has a maximum amount of BBU that it can offer, represented by Cell Capacity C.

3.5 LOADED CAC/CDC SCHEME

As explained in chapter 2.21, Call Admission Control is a scheme that takes exceptional consideration to both handoff and new calls by restraining the amount of incoming calls into the system thus providing good QoS. CAC is a good scheme to manage call connections and congestion problems such as interference. However, Call Admission control on its own cannot entirely avoid congestion especially during high network traffic.

Hence in this research work, CAC is combined with a scheme known as Call Duration and Control (CDC). As explained in chapter 2.22, CDC is a network managerial scheme in which termination notice is served to users once they have the threshold of a specific call duration in the network. The combined scheme takes into consideration both new calls and handoff calls, thus enabling credible QoS to be achieved. Hence, combining the two schemes will greatly marginalize the high congestion rate being experienced.

Also, the results of the ordinary CAC scheme is compared with the result of the combined CAC/CDC algorithm.

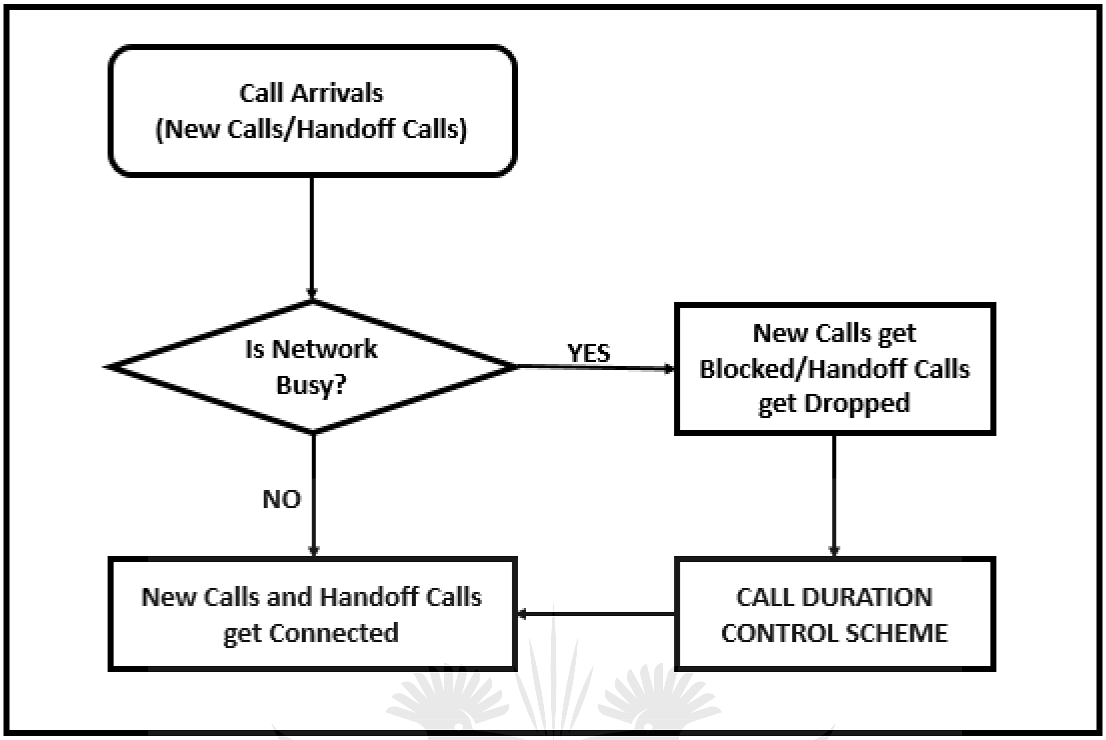

Figure 3.1 Block Diagram of the Combined CAC/CDC Schemes

As shown in figure 3.1, handoff calls and new calls that would have been dropped or blocked due to unavailable channels to accommodate them is accorded channels with the introduction of the CDC scheme into the CAC scheme. This is done by the introduction of the termination notice to long duration call users.

The CAC/CDC algorithm distributes calls to the RAT based on call arrival rate and the capacity of individual RAT. Call arrival rate is denoted by the symbol λ.

3.6 SIMULATION PARAMETERS

In the model as well as the simulation, as adapted in [58], assumption that 5% of the total bandwidth are solely reserved for hand-off calls is made. Also, all the users terminate their calls once the Termination Time elapsed is assumed. The path loss is modelled by approximately 40 + 44*log* (*s/*10) (in dB) where *s* represents the displacement between subscribers and the base station. Transmission power and bit error rate in the base station is set to 43 dB, 0.1%. In the

OFDM algorithm, users are usually allocated a specified number of subcarriers for a predetermined timeframe known as physical resource blocks (PRBs). For the research work, 5GHz band which contains 24 PRBs (Physical Resource Block) each having 10 subcarriers with 15 kHz band spacing is used. A frequency selective fading channel with six taps is assumed. Lastly, for all simulations, the time slot for each simulation is set at 2 seconds.

Table 3.1 shows the constant parameters that are used for all simulations involving the channel reservation based CAC and the combined CAC/CDC schemes.

Table 3.1 Simulation Parameters

SIMULATION PARAMETERS	
Path Loss	$40 + 44 log\,(s/10)$ (in dB)
Transmission power	43 dB
BER	0.1%.
PRB	24
Number of subcarriers	10
Time slot	2 seconds

Although, due to various advancements to GSM networks whereby they can now handle data and voice service classes in order to meet different users' requirements, this research is more particular to only voice service class GSM based network. In the research work, it is assumed that an approaching call to the RAT will require a particular amount of bbu represented by *b* in order to meet QoS requirement.

3.7 SIMULATION STAGES

In terms of load traffic, the simulations are classified into two stages basically in terms of how robust the system capacity is. A system capacity that can handle high load and a system capacity that can handle low load, these classifications are the High load scenarios and Low load scenarios respectively. For both stages, simulations were carried out for CAC alone and the combined scheme. The results of the two simulations are compared.

Also for the CAC/CDC algorithm, simulations are carried out for two different values of Termination notice time *(T_u)*. Figure 3.2 shows the flowchart of simulation stages

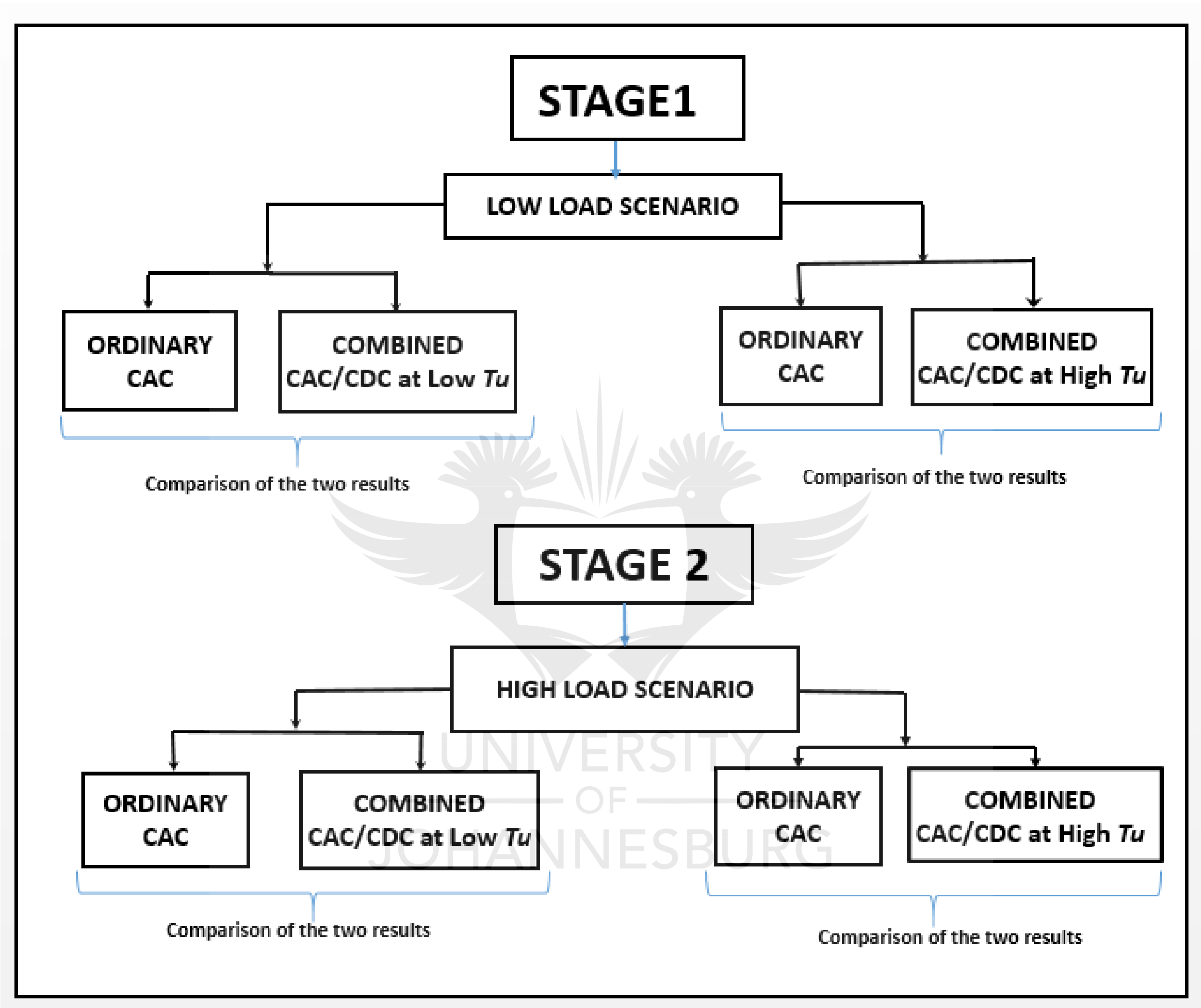

Figure 3.2 Flowchart of Simulation Stages

The result of the various simulations shows how congested the network is. This evaluation is shown in terms of the call blocking probability and call dropping probability against call arrival rates plots. As shown in the flowchart in figure 3.2, all simulations in both stage 1 and stage 2 is carried out in two parts. The first simulation part shows the result of the ordinary channel reservation CAC scheme. In the second simulation part (the combined CAC/CDC scheme), the termination notice *(T_u)* is introduced to the algorithm, whereby a T_u is sent to users immediately they reach the termination notice time set in the system. It should be noted that all the subplot graphs shows the comparison of the results of the first and second simulation results (the result of the CAC alone and the result of the combined schemes). The upper graphs in the subplots with green marker will indicates the CAC scheme result whereby the red marker lower graph in the subplot indicates the CAC/CDC combined scheme result.

It is noted that all the values used in the research work are experimental values and all the simulations does not involve real data in terms of users and cells. In the real world, a cell can handle thousands of subscribers.

Also, for comparison and measurement of the performance evaluation, analysis of the simulations in terms of load condition were performed, simulations using 10 subscribers as classified as low load scenario in stage 1 of the simulation whereby 30 subscribers are classified high load scenario for the stage 2 of the simulations. As discussed in the previous chapter, all users in the network are issued the termination notice once their call duration reaches the termination time and there call is disconnected. Also, for the various simulations, $\boldsymbol{T_u}$ is set at 10 seconds as low termination notice time while 240 seconds is attributed as high $\boldsymbol{T_u}$ time.

3.8 PERFORMANCE METRICS

3.8.1 CALL ARRIVAL RATE

Call arrival rate indicates user's demand for network resources. For a situation of high arrival rate, the demand is equally high. Assumption is made that the arrival rate obeys the technique of poisson distribution [62, 52]. The arrival of calls in the network is distributed to the different RATs based on the load that each RAT can handle. The calls arriving in the network are categorized as either new calls or handoff calls.

Considering a state space, which can be a set of possible values that can be taken. Denoting the state space as vector ***k*** which is given as [61, 62, 52, 55]:

$$\boldsymbol{k=(n, h : i= 1,2,3,4\ldots,X)} \quad (3.4)$$

Where *n* denotes the quantity of ongoing new calls whereas *h* represents that of handoff calls.
Let λn represent new calls arrival rate in the RAT system.

λh represent handoff calls arrival rate in the RAT system.

C denotes Cell Capacity

b denotes cell bandwidth

Tn denotes Threshold for new calls

Th denotes Threshold for handoff calls.

A threshold for new call *(Tn)* is set for the maximum capacity that can be occupied by new calls. Once the threshold is reached and exceeded, new call blocking will commence. The threshold for new calls is defined by the constraint [55, 61]:

$$b + \sum_{i=1}^{x} (b \times n) > Tn \tag{3.5}$$

Also a threshold *(Th)* is set for the maximum capacity possibly occupied by handoff calls. Once the threshold is reached and exceeded, handoff call dropping will commence. The threshold for handoff calls is defined by the constraint [55, 61]:

$$b + \sum_{i=1}^{x} (b \times h) > Th \tag{3.6}$$

Denoting *A* as the steady state of the system, ***k*** represents a subset of $A(\mathbf{k} \in A)$, if the following conditions are met [61]:

$$\sum_{i=1}^{x} (b \times n) \le Tn \Lambda \sum_{i=1}^{x} (b \times h) \le Th \Lambda \sum_{i=1}^{x} ((b \times n) + (b \times h)) \le C \tag{3.7}$$

Traffic intensity *(ρ)* is defined as the ratio of call arrival rate to service rate.

As defined in the previous chapter, new calls traffic intensity is expressed as [62]:

$$\rho n = \lambda n / \mu \tag{3.8}$$

and traffic intensity for handoff calls is expressed as:

$$\rho h = \lambda h / \mu \tag{3.9}$$

Where λh denotes arrival rate for handoff calls and μ denotes service rate.

Let *Pk* represent the probability that the network model is in a steady state ***k***. Hence, *Pk* is defined as [62]:

$$\prod_{i=1}^{x} \frac{(\rho h)^h}{h!} \frac{(\rho n)^n}{n!} \times P_O, \qquad (3.10)$$

$\mathbf{k} \in A$

Where *Po* is defined as:

$$P_O = \left[\sum_{k \in A} \prod_{i=1}^{x} \frac{(\rho h)^h}{h!} \frac{(\rho n)^n}{n!} \right]^{-1} \qquad (3.11)$$

3.8.2 NEW CALL BLOCKING PROBABILITY

Blocking of new call requests occur when there is no more capacity in a given coverage area to take on the new call request. This occurs when the maximum capacity of the coverage area is exceeded and is described with the following condition [61, 62]:

$$b + \sum_{i=1}^{x} ((b \times n) + (b \times h)) > C \qquad (3.12)$$

A threshold for new call *(Tn)* is set for the maximum capacity that can be occupied by new calls. Once the threshold is reached and exceeded, new call blocking will commence. The threshold for new calls is defined by the constraint [61, 62]:

$$b + \sum_{i=1}^{x} (b \times n) > Tn \qquad (3.13)$$

Assuming A_b denotes a set of states for which a new call is blocked, therefore [62]:

$$A_b = \left\{ k \in A(b + \sum_{i=1}^{X} (b \times n) > Tn \vee b + \sum_{i=1}^{X} ((b \times n) + (b \times h)) > C \right\} \qquad (3.14)$$

The New Call Blocking Probability (PB) is hence computed [61, 62]:

$$PB = \sum_{k \in A_b} P_k \qquad (3.15)$$

3.8.3 HANDOFF CALL DROPPING PROBABILITY

Dropping of a handoff calls occurs when there is no more capacity in a given coverage area to take on the handoff call. This occurs when the maximum capacity of the coverage area is exceeded and is described with the following condition [61, 62]:

$$b + \sum_{i=1}^{x}((b \times n) + (b \times h)) > C \tag{3.16}$$

A threshold *(Th)* is set for the maximum volume that can be occupied by handoff calls. Once the threshold is reached and exceeded, handoff call dropping will commence. The threshold for handoff calls is defined by the constraint [61, 62]:

$$b + \sum_{i=1}^{x}(b \times h) > Th \tag{3.17}$$

Assuming A_d denotes a set of states for which a handoff call is dropped, therefore [61, 62]:

$$A_d = \left\{ k \in A(b + \sum_{i=1}^{X}(b \times h) > Th \vee b + \sum_{i=1}^{X}((b \times n) + (b \times h)) > C \right\} \tag{3.18}$$

Handoff Call Dropping Probability (PD) is the probability that a handoff call request will be dropped once there is no more capacity in the cell the call is moving into. Hence, the Handoff Call Dropping Probability PD is hence computed as [61, 62]:

$$PD = \sum_{k \in A_d} P_k \tag{3.19}$$

The condition for both *PB* and PD is given as [62]:

$$b + \sum_{i=1}^{x}((b \times n) + (b \times h)) > C \tag{3.20}$$

3.8.4 TERMINATION TIME

Termination time is peculiar to the CAC/CDC combination scheme. It denotes the average time at which the termination noticed is served to users in the network. This notification is only sent to users of long duration calls when they reach the threshold of the termination time. This is done to allow them to free the channel for new users. It is denoted by $\boldsymbol{T_u}$. The $\boldsymbol{T_u}$ is set in seconds.

3.9 CHAPTER SUMMARY

This chapter has given a review of the analytical and experimental approach highlighting the parameters used for the reserved channel based CAC scheme and the CAC/CDC combination algorithm, the bandwidth reservation scheme and the Markov chain model. The chapter further discussed the system performance metrics that is used throughout the research work.

4. DATA ANALYSIS AND RESULT

4.1 INTRODUCTION TO SIMULATION SCENARIOS FOR PERFORMANCE EVALUATION

As discussed in the previous chapter, the modelling of the system is built on a GSM based network. All simulations were carried out using MATLAB R2012a. The result of the various simulations show how congested the network is. This evaluation is shown in terms of the call blocking probability and call dropping probability against call arrival graphs. In terms of the rate at which calls arrives and the rate at which there is a probability of the calls can be blocked or dropped, the graphs shows how congested the network is. Also, for comparison and measurement of the performance evaluation, the various simulations are analyzed in terms of load condition.

As explained in chapter 3, in terms of system capacity, the simulations are divided into two stages, the stage one is for the low load scenarios and the stage two is for the high load scenarios. For the two stages, all the simulations were carried out in two parts. The first simulation part shows the result of the ordinary channel reservation CAC scheme. In the second simulation part, the termination notice *(T_u)* is sent to users immediately they reach the termination notice time set in the system. It should be noted that all the subplot graphs shows the comparison of the results of the first and second simulation results (the result of the CAC alone and the result of the combined schemes). The upper graphs in the subplots with green marker indicate the CAC scheme result whereby the red marker lower graph in the subplot indicates the CAC/CDC combined scheme result.

As shown in table 4.1, for the two stages of the simulations, system capacity handling 10 subscribers is recognized as low load condition while system capacity handling 30 subscribers is classified as high load condition. As discussed in the previous chapter, all users in the network are issued the termination notice once there call duration reaches the termination time and there call is disconnected. Also, for the various simulations, setting T_u at 60 seconds is treated as low termination notice time whereas setting 240 seconds as the T_u is regarded as high T_u time.

Most importantly, it is noted that all the values used in the research work are experimental values and all the simulations does not involves real data in terms of users and cells . In the real world, a cell can handle thousands of subscribers.

Table 4.1 Simulation Stages Parameters

SCENARIO CONDITION	PARAMETER VALUE
HIGH LOAD	SUBSCRIBERS NUMBER IS SET AT 30
LOW LOAD	SUBSCRIBERS NUMBER IS SET AT 10
LOW TERMINATION NOTICE	T_u IS SET AT 60 SECONDS
HIGH TERMINATION NOTICE	T_u IS SET AT 240 SECONDS

4.2 SIMULATION STAGE ONE

4.2.1 LOW LOAD SCENARIO AT A HIGH TERMINATION NOTICE *(T_u)*

For the low load scenario, simulations is carried out using 10 subscribers. The result of the simulations shows how congested the network is. This evaluation is measured in terms of the call blocking probability and call dropping probability against call arrival graphs. The termination notice time is set a high value of 240 seconds. All users in the network got the termination notice immediately the call duration reaches 240 seconds and the call is disconnected to free up the channel for new users. Figure 4.1 shows the call dropping probability against call arrival result of the comparison of the ordinary CAC scheme and the combined CAC/CDC scheme while figure 4.2 shows the call blocking probability against call arrival result of the comparison between the ordinary CAC scheme and the combined CAC/CDC scheme.

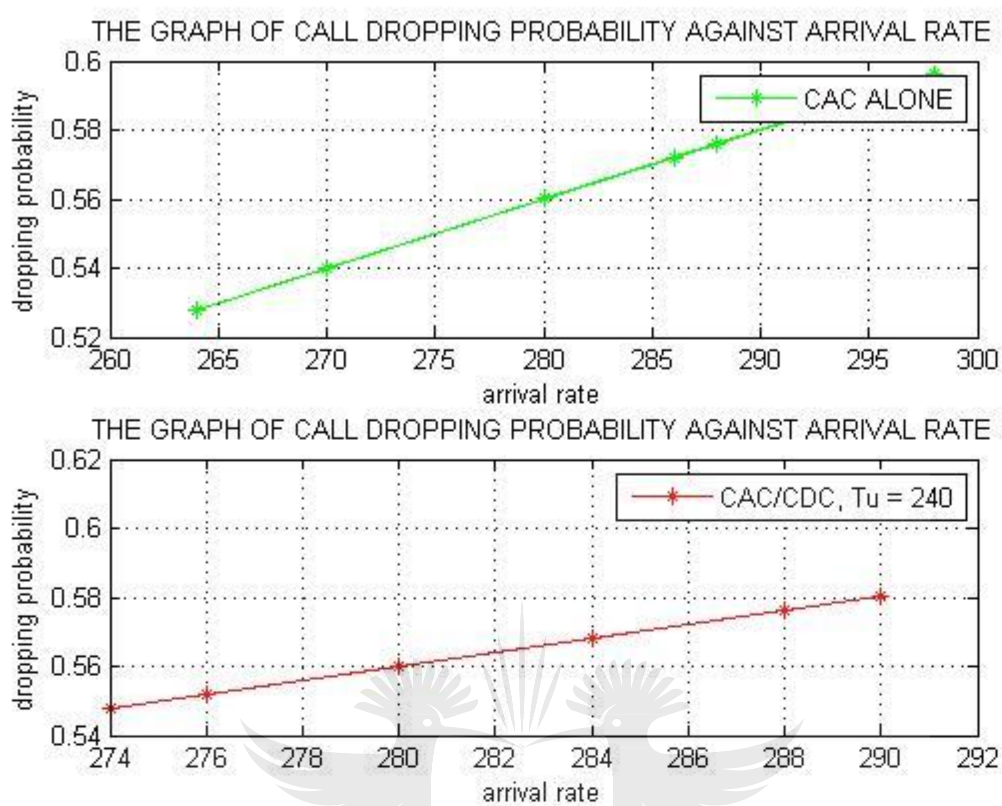

Figure 4.1 Call Dropping Probability against Call Arrival Rate for ordinary CAC and the CAC/ CDC combination for a low load scenario when T_u is set at 240 seconds.

From the subplots in figure 4.1, it was noticed that the probability of a call being dropped in the CAC scheme tends from approximately 0.52 at the lower limit to 0.6 in the upper limit for a low call arrival rates, whereas in the combined scheme, the values of the probability of a call getting dropped tends from 0.55 to 0.58 at a higher call arrival rates. This shows that there is a better ratio in the probability of a call being dropped with respect to the call arrival rates when the termination notice is set at 240 seconds compared to when all users are allowed to use the network as they wish i.e. ordinary CAC scheme without any termination notice to long duration users. Also, in figure 4.2, which shows the call blocking probability plot, a similar result was noted in the subplot, there is a better ratio in the probability of a call being blocked with respect to the call arrival rates when the termination notice is set at 240 seconds compared to when all users are allowed to use the network as they wish i.e. ordinary CAC scheme without any termination notice to long duration users.

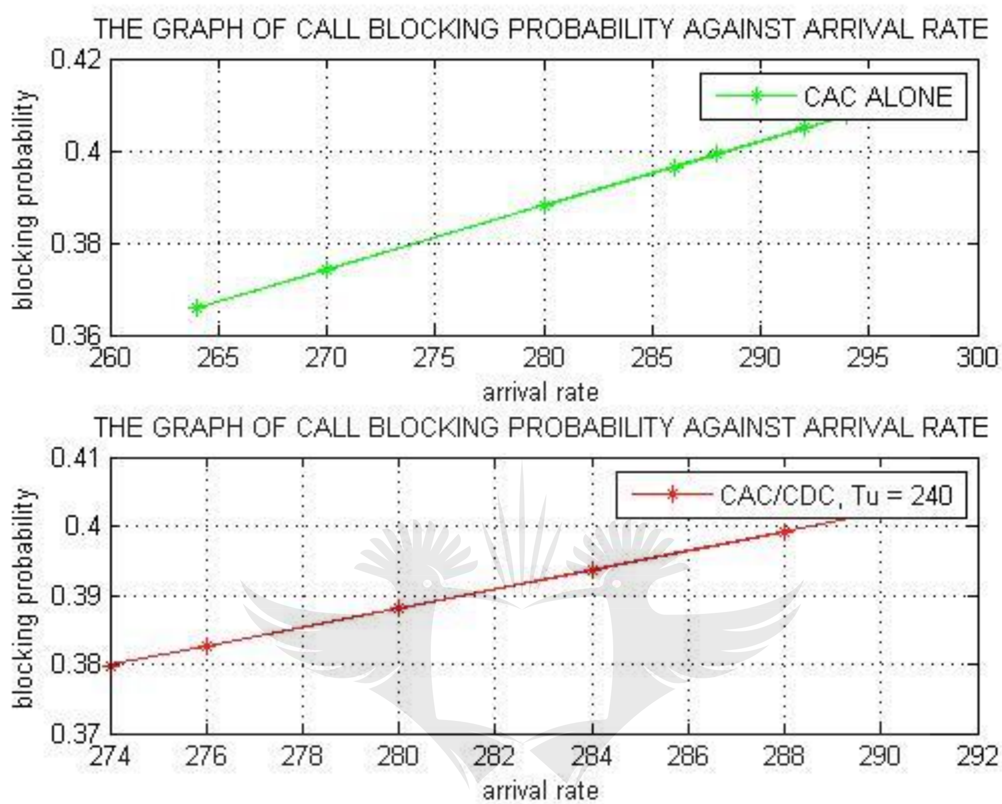

Figure 4.2 Call Blocking Probability against Call Arrival Rate for ordinary CAC and the CAC/ CDC combination for a low load scenario when *T_u* is set at 240 seconds.

The result evaluation of the low load condition is further explained at the call arrival rate axis when compared to the rate of call blocking probability, it is observed that the call arrival rate in the CAC scheme ranges from approximately λ=264 in the lower limit to λ=298 in the upper limit, this is far greater when compared with the combined scheme when the termination notice time is set at 240 seconds which yielded approximately λ=266 in the lower limit to λ=296 in the upper limit. Normally, in a congested network, higher call arrival rates is supposed to cause more congestion, but with the introduction of the termination notice $\boldsymbol{T_u}$ at 240 seconds, the increase in the call arrival rates didn't show much impact on the blocking probability graphs.

4.2.2 LOW LOAD SCENARIO AT A LOW TERMINATION NOTICE *(T_u)*

For this simulation, 10 subscribers is used whereby the termination notice is set to a lower value of 60 seconds. The result of the Call Dropping Probability against Call Arrival Rate for ordinary CAC and the CAC/ CDC combination for a low load scenario when T_u is set at 60 seconds is shown in figure 4.3 while Call Blocking Probability against Call Arrival Rate for ordinary CAC and the CAC/ CDC combination for a low load scenario when T_u is set at 60 seconds is shown in figure 4.4 below.

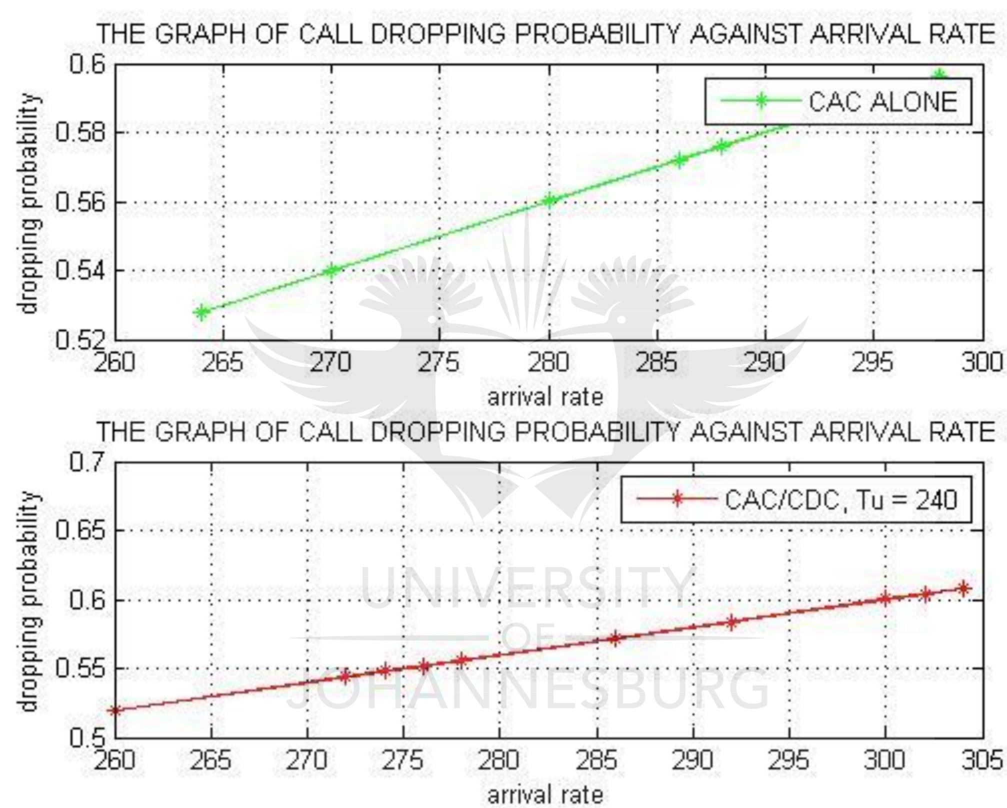

Figure 4.3 Call Dropping Probability against Call Arrival Rate for ordinary CAC and the CAC/ CDC combination for a low load scenario when T_u is set at 60 seconds.

As shown in figure 4.3, the simulation shows a better result when the T_u is introduced into the CAC scheme. There is more call arrival rate at a better proportion of call dropping rates when the termination notice T_u is introduced compared to the ordinary CAC scheme. This shows that there is a better result in the probability of a call being dropped with respect to the call arrival rates when the termination notice is set at 60 seconds compared to when all users are allowed to use the network as they wish i.e. ordinary CAC scheme without any termination notice to long duration users.

Furthermore, comparing the result in figure 4.3 with the result from figure 4.1, it was noticed that the call dropping probability graphs shows a difference when Tu changed to a lower value (from 240 seconds to 60 seconds).When the termination notice is set to a higher value, more users tends to hold on to the channel at the expense of new users unlike when the termination notice is set at a lower value whereby users tend to have shorter conversation before the termination of the calls. From the comparison of figure 4.1 and figure 4.3, the probability that a call can be dropped using a high termination notice (Tu =240 seconds) yielded 0.52 to 0.59 (lower limit to upper limit) at call arrival rates of λ=266 to λ=298 which is lower when compared to the scenario of termination notice of 60 seconds where the probability of a call being dropped is at 0.55 to 0.6 at a higher call arrival rates of λ=276 to λ=303 (lower limit to upper limit).

Figure 4.4 presents the call blocking probability against call arrival rate for ordinary CAC and the CAC/ CDC combination for the low load scenario when $\boldsymbol{T_u}$ is set at 60 seconds.

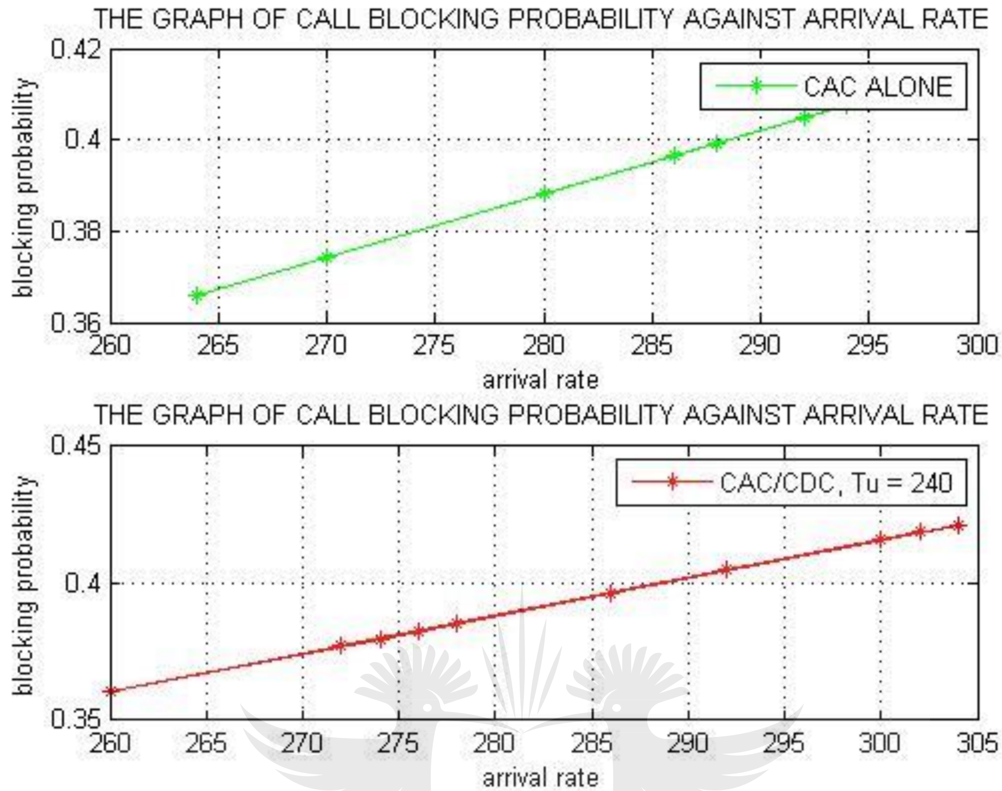

Figure 4.4 Call Blocking Probability against Call Arrival Rate for ordinary CAC and the CAC/ CDC combination for a low load scenario when *Tu* is set at 60 seconds.

From figure 4.4, a better result is shown using the combined scheme compared to the ordinary CAC scheme

As discussed earlier, the likelihood of another call being blocked is minimal with a lower T_u when contrasted with a higher T_u. This is better indicated when the graphs in figure 4.2 is compared with figure 4.4. The probability that a new call will be blocked using a high termination notice (T_u =240 seconds) yielded 0.37 to 0.41 (lower limit to upper limit) at call arrival rates of λ=266 to λ=296 which is lower when compared to the scenario of termination notice of 60 seconds where the probability of a new call being blocked is at 0.38 to 0.42 (lower limit to upper limit) at a high call arrival rates of λ=276 to λ=303 (lower limit to upper limit).

4.3 SIMULATION STAGE TWO

4.3.1 HIGH LOAD SCENARIO AT A LOW TERMINATION NOTICE *(T_u)*

For this simulation, 30 subscribers were used and interpreted as a high load scenario. The termination notice is set to a low value of 60 seconds. This is essential to be done, in order to get more users into the network especially at a period of high network load. With the introduction of a low termination notice time, more users easily and quickly have access to new channel. This scenario is simulated to evaluate the performance of the CAC/CDC combination at a high load condition. Comparison was made between the result of this simulation and the result of the scenario with low load condition.

The result of the simulation of the high load scenario (30 subscribers) whereby a $\boldsymbol{T_u}$ of 60 seconds is introduced into the model is shown in figure 4.5/4.6. The call dropping probability against call arrival of the CAC and the combined scheme when $\boldsymbol{T_u}$ is set at 60 seconds is shown in figure 4.5 while figure 4.6 shows the call blocking probability against call arrival rate graph. Comparing the result from figure 4.5, it was noticed that the combined scheme when the $\boldsymbol{T_u}$ is introduced into the channel reserved based CAC yielded a better result in terms of a call being dropped when compared to the equivalent ordinary CAC scheme. The probability of a call getting dropped tends from a value of 0.166 at a call arrival rate of λ=828 at the lower limit to 0.175 at a call arrival rate of 876 at the upper limit in the combined scheme, which is a better result when compare to the ordinary scheme that yielded a call dropping probability of 0.167 at call arrival rates of λ=837 at the lower limit to 0.177 at 878 call arrival rate in the upper limit. This shows that there is a better result in the probability of a call being dropped with respect to the call arrival rates when the termination notice is set at 60 seconds compared to when all users are allowed to use the network as they wish i.e. ordinary CAC scheme without any termination notice to long duration users.

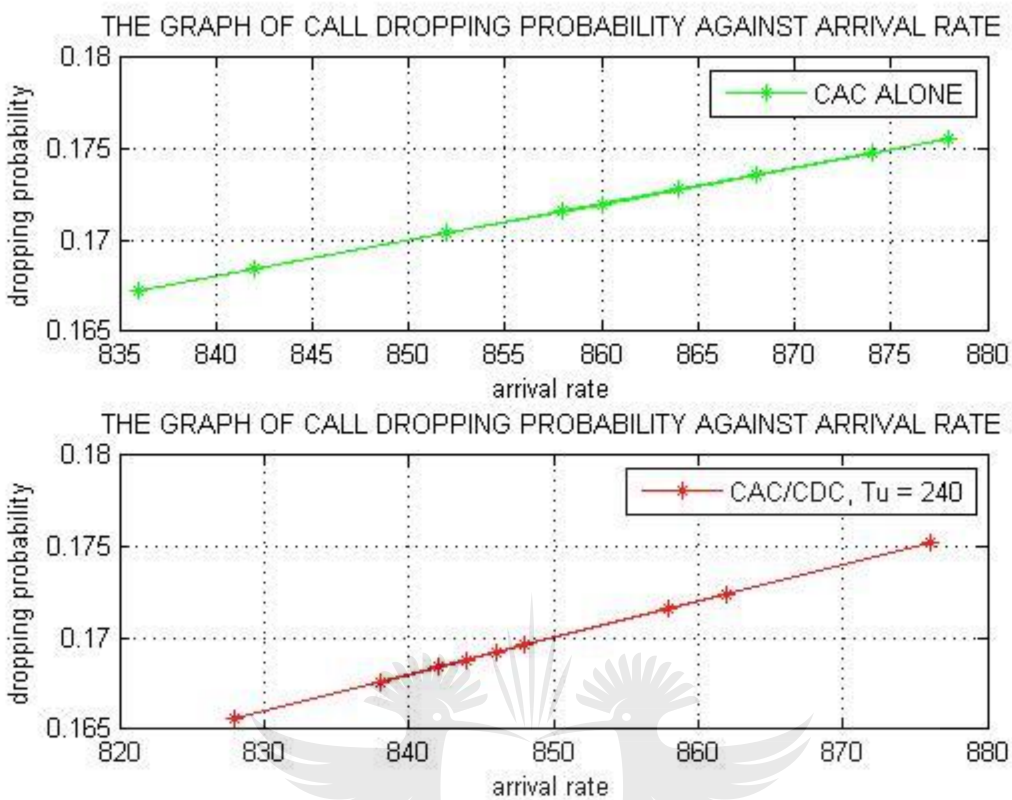

Figure 4.5 Call Dropping Probability against Call Arrival Rate for ordinary CAC and the CAC/ CDC combination for a high load scenario when *Tu* is set at 60 seconds.

With respect to the probability of new calls being blocked at a high call arrival rates, figure 4.6 also shows a better performance in terms of a new call being blocked with the introduction of 60 seconds T_u into the system model.

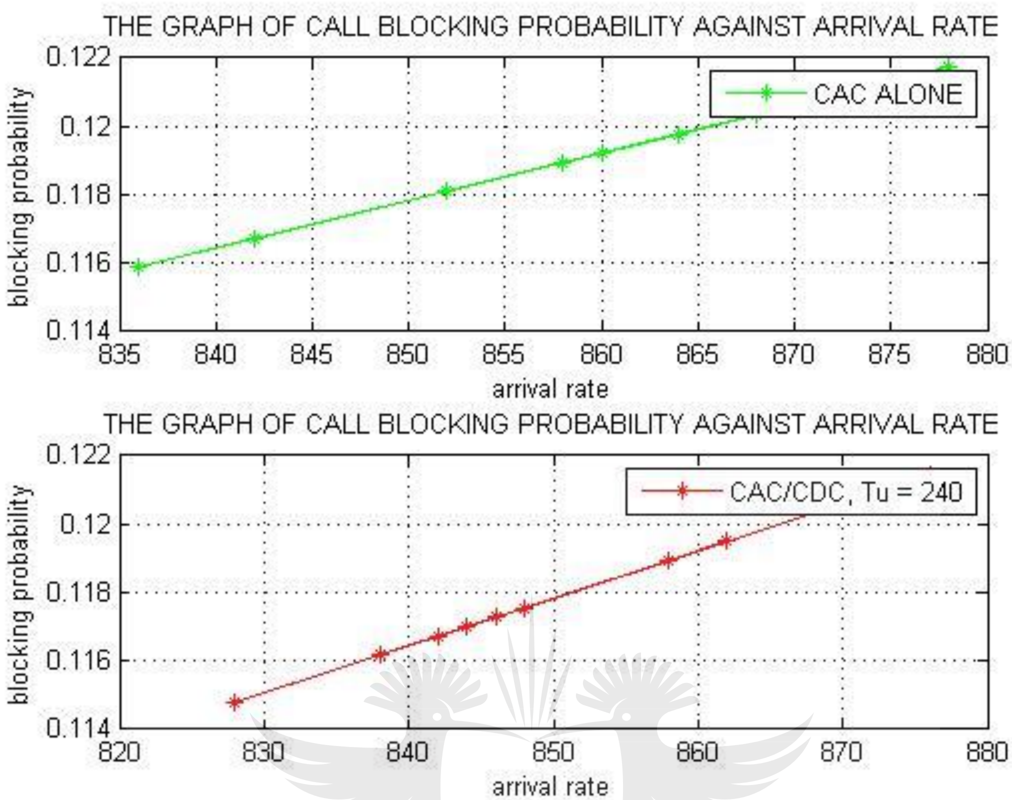

Figure 4.6 Call Blocking Probability against Call Arrival Rate for ordinary CAC and the CAC/ CDC combination for a high load scenario when *Tu* is set at 60 seconds.

Figure 4.6 shows that there is a better ratio in the probability of a new call being blocked with respect to the call arrival rates when the termination notice is set at 60 seconds compared to when all users are allowed to use the network as they wish i.e. ordinary CAC scheme without any termination notice to long duration users.

4.3.2 HIGH LOAD SCENARIO AT A HIGH TERMINATION NOTICE (T_u)

In this scenario, a high termination notice time (T_u = 240 seconds) was integrated into a high load condition whereby, the system capacity has been increased to a high load capacity which can accommodate 30 subscribers. With the introduction of a high value of T_u, users have more luxury of time to hold on to the channel at the detriment of new users, this tends to reduce congestion but on a minimal rate. Figure 4.7 presents the graph of the call dropping probability for the reserved channel Call Admission Control scheme and combined scheme of CAC and call duration control scheme against call arrival rates.

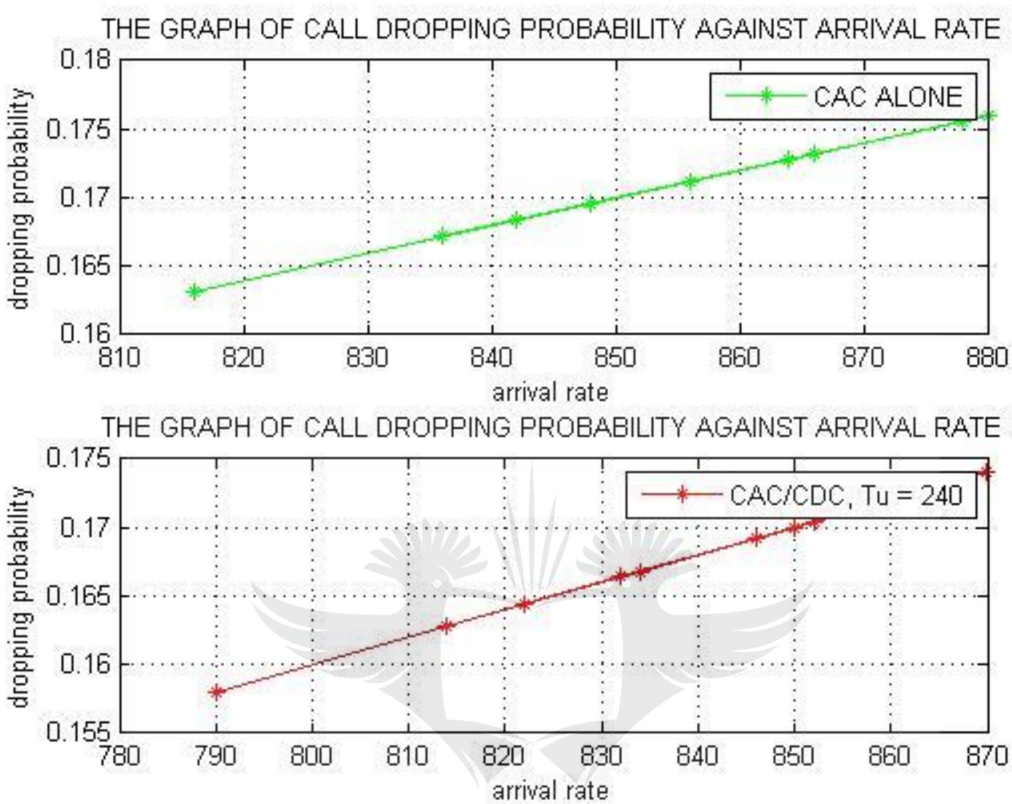

Figure 4.7 Call Dropping Probability against Call Arrival Rate for ordinary CAC and the CAC/ CDC combination for a high load scenario when *Tu* is set at 240 seconds.

As shown in figure 4.7, although the system capacity has been increased to a high load capacity, the combined scheme shows a better result when compared with the ordinary CAC scheme. With the introduction of the 240 seconds into the system model, the probability of a handoff call to be dropped reduced as indicated in the figure 4.7. It was noticed that the probability of a call being dropped in the CAC scheme tends from approximately 0.163 at the lower limit to 0.177 in the upper limit for a range of call arrival rates of λ=815 to λ=880 (lower limit to upper limit) respectively, whereas in the combined scheme, the values of the probability of a call getting dropped tends from 0.158 to 0.173 at a better call arrival rates. This shows that there is a better ratio in the probability of a call being dropped with respect to the call arrival rates when the termination notice is set at 240 seconds compared to when all users are allowed to use the network as they wish i.e. ordinary CAC scheme without any termination notice to long duration users.

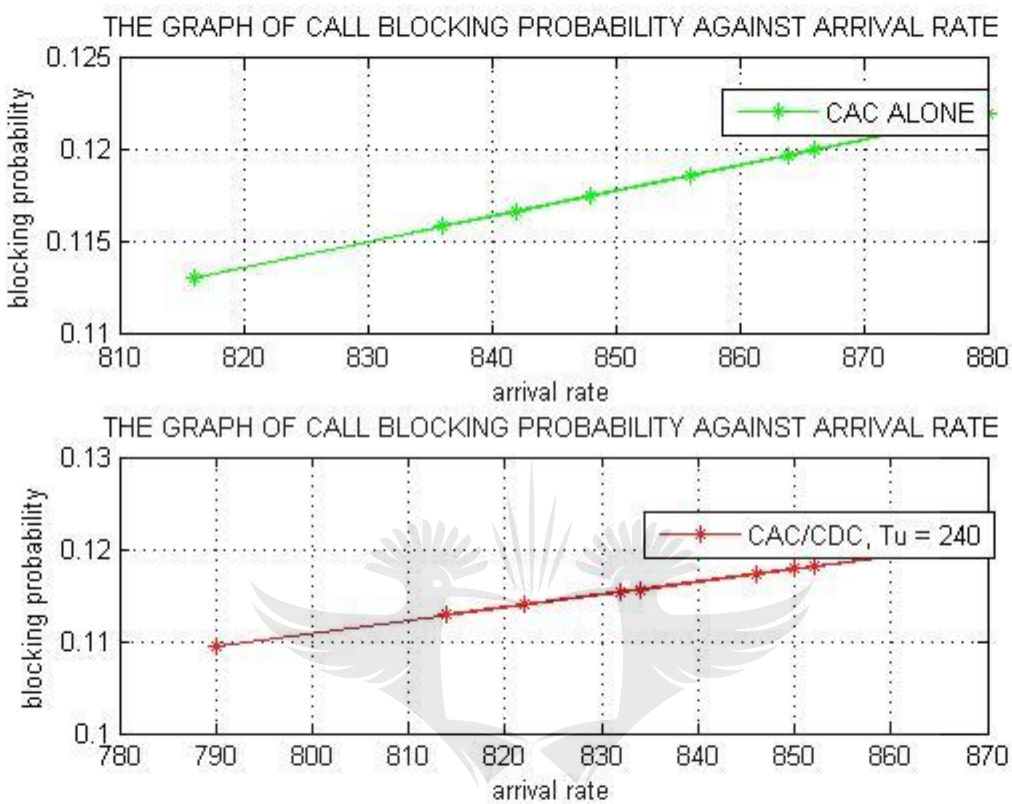

Figure 4.8 Call Blocking Probability against Call Arrival Rate for ordinary CAC and the CAC/ CDC combination for a high load scenario when *Tu* is set at 240 seconds.

Also, as shown in figure 4.8, the probability of a new call being blocked reduced further with the introduction of the termination notice (T_u= 240), the termination notice trigger the release of channels for new users to gain access to the network. Figure 4.8 shows that there is a better ratio in the probability of a new call getting blocked with respect to the call arrival rates when the termination notice is set at 240 seconds compared to when all users are allowed to use the network as they wish i.e. ordinary CAC scheme without any termination notice to long duration users.

4.4 HIGH LOAD SCENARIO AT DIFFERENT TERMINATION NOTICE *(T_u)*

In this scenario, comparison was made between the results of the introduction of different termination notice into a system capacity which can handle a high load. A high load situation was used for this evaluation in order to get a clearer result of the simulation. Introducing a high termination notice time (T_u = 360 seconds) into a high load condition of 30 subscribers, it was noticed that channels are not readily available compared to a situation of using a low termination time of 60 second whereby channels are being freed early for new users. This scenarios allows more users into the network and hence, reduce congestion more.

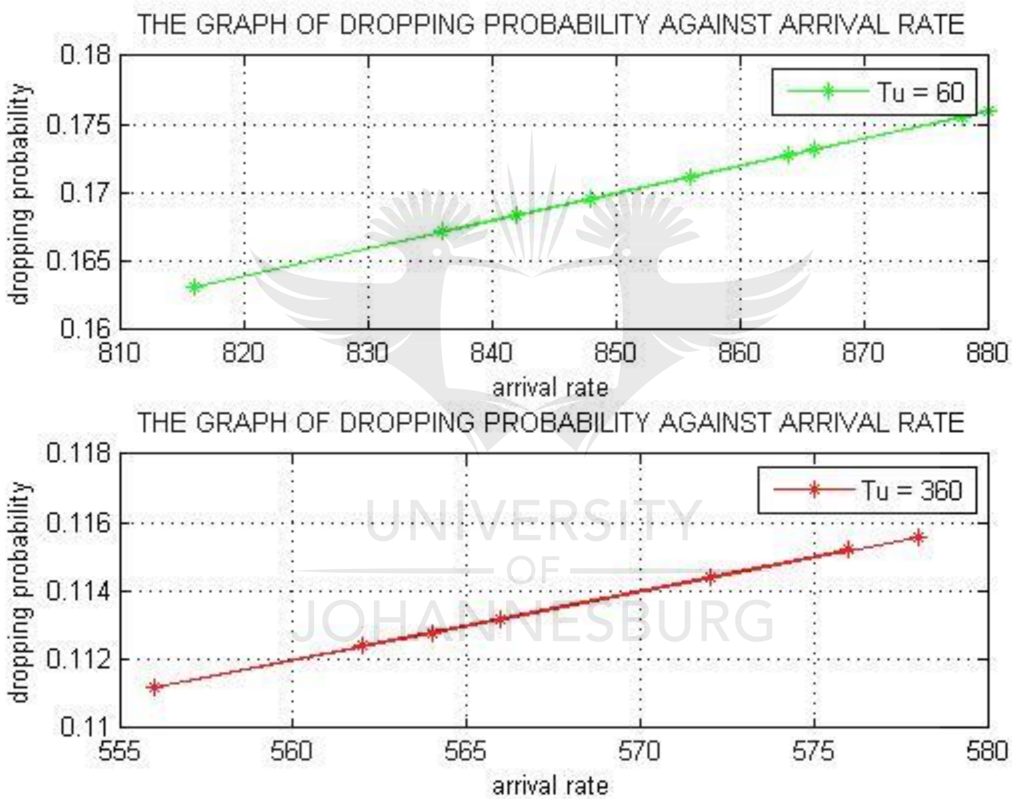

Figure 4.9 Call Dropping Probability against Call Arrival Rate for a high load scenario at different *Tu*.

Figure 4.9 shows the result of the dropping probability against time arrival of using a T_u of 60 seconds and T_u of 360 seconds. it was noticed that the probability of a call being dropped in the CAC scheme tends from approximately 0.163 at the lower limit to 0.177 in the upper limit for a range of call arrival rates of λ=815 to λ=880 (lower to upper limit) respectively, whereas in the scenario that involves T_u of 360 seconds, the values of the probability of a call getting dropped tends from 0.1115 to 0.1157 at a higher call arrival rates.

Also figure 4.10 shows the call blocking probability simulation result of a high load scenario but at different termination notice time. From figure 4.10, it was noticed that the probability of a call being blocked at a lower value of Tu =60 seconds tends toward approximately 0.115 in the lower limit to 0.123 in the upper limit at a call arrival rate of 815 at the lower limit to 880 in the upper limit, whereas for the scenario that involves a higher value of Tu =360 seconds, the values of the probability of a call getting blocked tends from 0.77 to 0.8 at a higher call arrival rates of 556 to 578 (lower to upper limit) respectively.

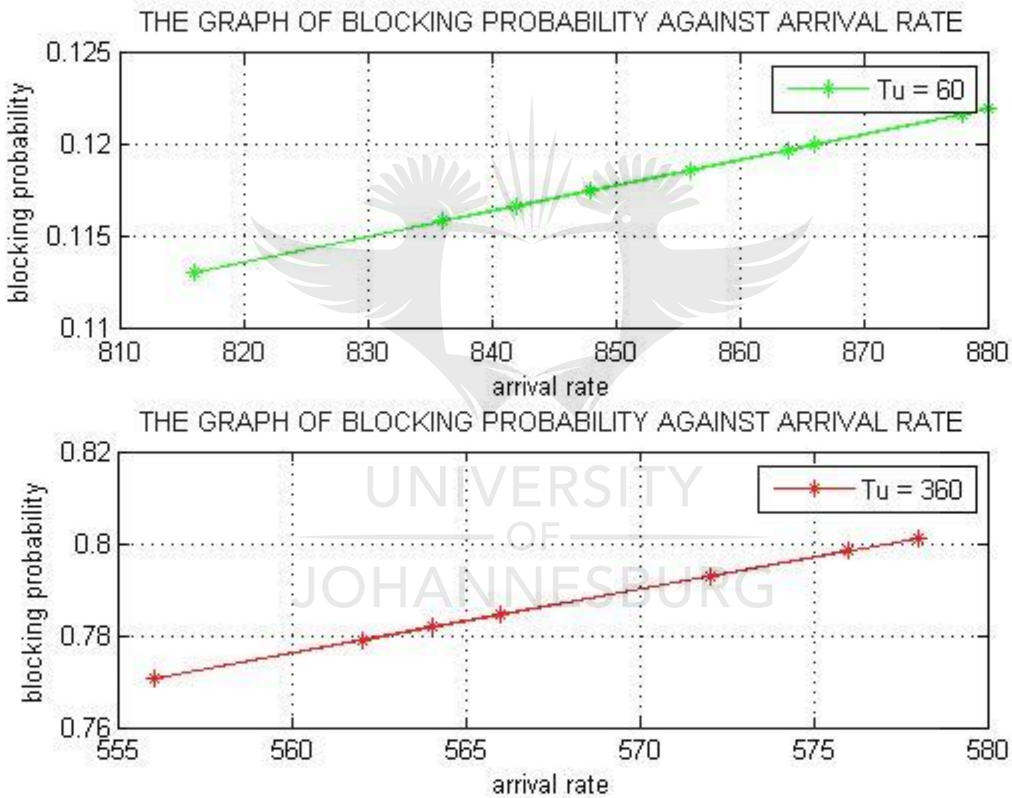

Figure 4.10 Call Blocking Probability against Call Arrival Rate for a high load scenario at different *Tu*.

From figure 4.9 and figure 4.10, at a lower termination notice time, (T_u set at 60 seconds), a far better result was achieved from both probabilities (call blocking and call dropping) rate with respect to the corresponding call arrival rates unlike in the case of a high T_u. Channels are readily and quickly available once users calls are terminated at a shorter call duration, compared to when users have the call terminated at a longer call duration. This explanation is also described in the transmission rate table of the two different $\boldsymbol{T_u}$ in a high load scenario. The transmission table of the comparison is obtained from the simulation and it is shown in table 4.2.

Table 4.2 shows the comparison of two different values of termination notice and when there was no termination notice (ordinary CAC scheme whereby users are allowed to use the channel for any call duration), each equivalent transmission rate values in adjacent roll shows the best result at $\boldsymbol{T_u}$ = 60. The transmission rate table is taken from the simulation and it virtually shows the relationship between the probability axis and the call arrival rate. Furthermore, from the transmission rate table, transmission rate values at $\boldsymbol{T_u}$ at 360 shows better results compared to the transmission rate values for the ordinary channel reservation based CAC scheme.

Table 4.2 Transmission Rate Table for 30 subscribers at different termination notice time

TRANSMISSION RATE TABLE FOR HIGH LOAD SCENARIO AT DIFFERENT $\boldsymbol{T_u}$		
CAC ALONE USING 30 SUBSCRIBERS (without $\boldsymbol{T_u}$)	T_u at 360 seconds	T_u at 60 seconds
0.814	0.860	0.8
0.844	0.834	0.81
0.85	0.844	0.834
0.844	0.84	0.838
0.83	0.815	0.836
0.872	0.828	0.834
0.85	0.864	0.846
0.85	0.808	0.86
0.874	0.87	0.846
0.872	0.836	0.842

It was noticed from table 4.2 that the average rate of tranmission when $\boldsymbol{T_u}$ is set at 60 seconds is faster with smaller values when compared to average transmision rate for $\boldsymbol{T_u}$ at 360 seconds. The channel reservation scheme CAC alone gives the worst average transmission rate.

4.5 FORECASTED BUSY HOUR TRAFFIC MODEL

Forecast of time arrangement is an essential information mining and it can be broadly utilized as a decision making tools of humans' day by day life activities. As discussed in chapter 2.11, the busy hour/peak hour of a system is the period in the day when the network system processes the highest traffic. Busy hour traffic can be used as a guage to identify the capability of telecom system, which in turn can be used for its dimensioning. Mobile operators care about the execution of anticipating the month to month and day-to-day activity burden. The study of the peak period and busy period on daily or monthly period assist the network operators in building the network infrastractures in locations. Busy hour period of the day is the period of the day when there is a high tendency that new calls will be blocked and handoff calls can be easily dropped.

4.5.1 SYSTEM MODEL

This work aims at the study of GSM traffic pattern over a period of 24 hours. The system is modelled and simulation is carried out using MATLAB. A bandwidth reservation based CAC scheme is used for the forecasted simulation. Higher priority is given to handoff calls than to new calls. The threshold set for rejecting handoff calls is higher than the threshold set for rejecting new calls.

In order to handle handoff calls, two cells (cell A and cell B) were used, both at the close distance are used for the system model and each cells have it own individual thresholds for handoff calls and new calls. Each cells have it own bandwidth, which is set at 2. Other parameter used for the simulation is shown in table 4.3.

As shown in table 4.3, the cell capacity for cell A is bigger compared to cell capacity B. Furthermore, in order to maintain higher priority for handoff calls, some channels are reserved mainy for handoffs only in both cells (cell A has 10 channels and cell B has 5 channels mainly reserved for handoff calls only).

Table 4.3 Simulation Parameter Table for Busy Hour Traffic over a period of 24 hours

CELL	**CELL A**	**CELL B**
BANDWIDTH	2	2
CELL CAPACITY	40	30
THRESHOLD FOR NEW CALLS	30	25
THRESHOLD FOR HANDOFF CALLS	40	30
SUMMATION OF CALL ARRIVALS	24	24

As explained in chapter 2.11, the peak period of the day is the period whereby there is the highest probability of having a call blocked or call dropped. Hence, the blocking and dropping probabilities against the time of the day were simulated. From the simulation results, assumption was made that the peak of the probability graphs against the period of the day typifies the peak period of traffic for the day. Figure 4.11 shows the result of the simulation. It contains the new call blocking probability and call dropping probabilty over a day period.

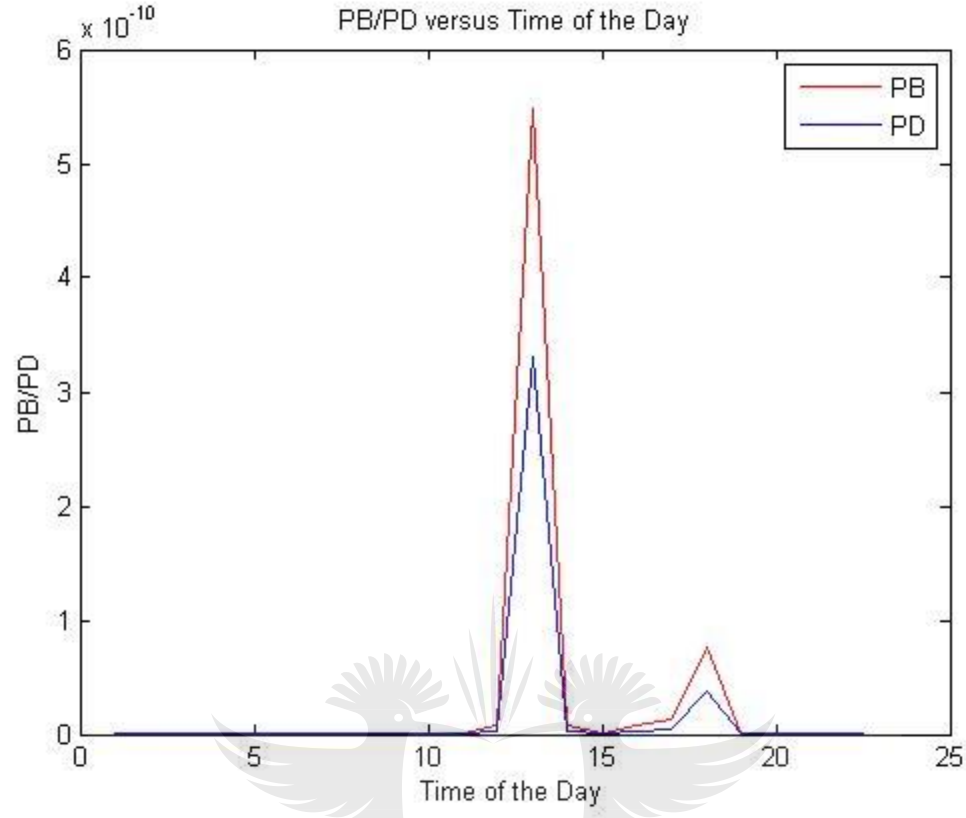

Figure 4.11 Call Probabilities against time of the day to show the busy hour traffic.

From figure 4.11, the peak of the probabilities were observed to have occurred at approximately 13:00hr

And the next peak probabilities occurs around 18:00hr. This can be assumed to be the busy hour period of the day. The 13:00hr can also be explained from humans' daily life activities as the period when people generally have time for lunch during office hour and the 18:00hr can be explained as the time people will not be at work and they have there own time to discuss personal issues with their loved ones etc. Furthermore, it was noticed from figure 4.11 that the blocking probability graph is slight lower when compared with the dropping probability graph. This is as a result of the channels purely reserved for handoff calls.

4.6 MOBILE NETWORK CONGESTION MANAGEMENT USING CAC AND CDC COMBINATION.

Mobile network congestion was managed using the combination of CAC AND CDC. Simulations were carried out in two stages. Referring to the flowchart in figure 3.2, Stage 1 shows the simuation for low load scenarios and stage 2 shows the simulation results for high load scenarios. For each stage, simulations were carried out for ordinary CAC scheme and the results of each simulation were compared to the result of the combined CAC and CDC scheme equivalent. The termination notice time T_u was either set at a high value of 240 seconds or a low value of 60 seconds for each simulation.

Figures 4.1, 4.3, 4.5 and 4.7 shows the simulation results of the call dropping probability against call arrival rates for each stages whereas figures 4.2, 4.4, 4.6 and 4.8 shows the call blocking probability against call arrival rates.

For each simulation, the combined scheme of CAC and CDC shows a better result compared to the ordinary CAC scheme especially in terms of the ratio of dropping and blocking probabilities against call arrival rates. from all the simulation results, the combined scheme gives a better and relatively lower dropping probability or blocking probability at a higher call arrival rates when compared to the ordinary CAC scheme.

4.7 CHAPTER SUMMARY

This chapter presents the results of the simulation software for the formulated problem. The introduction of the call duration control into the channel reservation based CAC scheme has been analyzed and evaluated for GSM based network. The application of the termination notice time causes rational and effective use of resources whereby more users are able to gain access to the network irrespective of the network condition either at low load or high load. Various load conditions are considered for simulation and the termination notice time is also varied to see the performance of the combined CAC/CDC scheme in managing network congestion. All the results shows improvement when compared with the ordinary CAC scheme. Furthermore, a system model for busy hour traffic forecasting was analyzed and evaluated in the concluding part of the chapter.

5. CONCLUSION AND RECOMMENDATONS FOR FUTURE WORK

This thesis has presented the combined scheme of Call Duration Control and Call Admission Control in managing mobile network congestion in GSM network. The call duration control scheme involves the introduction of termination notice time to long duration users in the network. Given that some network users, due to their affluent lifestyle can decide to make long duration calls, hence holding on to the channels for a long period, the excesses of such users have to be addressed and curbed in order to allow more users into the network and hence maximize the use of network resources.

However, it should be noted that the combined scheme has some shortcomings. The termination of calls may not be generally accepted by some users who have to make high duration calls. Some of these users may prefer to go for other options in terms of switching network operators and some will prefer to be charged at a higher rate rather than having their calls terminated.

In this research the evaluation of two different load scenarios (Low Load and High Load scenarios) using different termination notice time to end calls from long duration call users has been performed using an effective channel reservation based CAC scheme in GSM network. The simulation results obtained show that lower call blocking and call dropping probabilities are experienced during high load condition either at low or high the termination notice time, but better result is obtained at low termination notice time is high. These are the determinants of mobile network congestion that were focused on. Consequently, the combined scheme offers effective use of the channels and lowers the congestion rate, whereby more users are able to gain access to the network during low load conditions and overcomes network overload during peak periods.

Furthermore, the thesis presented a forecasted peak period situation of a mobile network traffic. A system model for busy hour traffic forecasting which showed the period of the day where there is a high tendency of high mobile network traffic was evaluated. Mobile networks are highly congested at periods whereby the blocking probability and dropping probability are at peak limits. Simulation results shows the peak of the blocking probabilities.

5.1 RECOMMENDATIONS FOR FUTURE WORK

Areas for future work are hereby discussed.

This thesis made a lot of assumptions which includes the notion that all users are in a homogeneous network environment. Flexibility in terms of integrating call duration control into the call admission control is an areas that can be explored in future research work. Flexibility whereby users can be notified about the option to terminate calls or having to pay higher tariff will go a long way in making the scheme convenient for the users especially in an heterogeneous network environment.

REFERENCES

[1] V. E. Idigo, C .O. Azubogu, C .O. Ohaneme and K. A. Akpado, “Real time assessment of QoS of mobile cellular networks in Nigeria,” *International Journal Of Engineering Inventions*, vol. 1, issue 6, ISSN: 2278-7461, [Published]; Retrieved from: www.ijeijournal.com [Accessed : June 2015]

[2] M. A. Alarape, A. T. Akinwale and O. A. Folorunso, “Combined scheme for controlling GSM network calls congestion,” *International Journal of Computer Applications*, (0975 – 8887) vol. 14, no. 3, January 2011.

[3] N. V. Marathe, G. S. Biradar and U. B. Desai, “A comprehensive call management strategy for congestion control in cellular networks,” 2nd Intl. Conf. on Communication Systems Software and Middleware 2007, pp. 1-7, Jan. 2007.

[4] M. Manaffar, H. Bakhshi and M. Pilevari, “A new dynamic pricing scheme with call admission control to reduce network congestion,” 22nd Intl. Conf. on Advanced Information Networking And Applications - Workshops 2008, pp. 347 – 352, March 2008.

[5] P. Zheng *et al.*, “GSM Versus CDMA,” in Wireless Networking Complete, Morgan Kaufmann, pp. 29, 2009.

[6] D. O. Dike, C. U. Iroh, G. N. Ezeh and B. C. Dike, “Minimization of call congestion in telecommunication system using OFDM optimization model,” 9th Intl. Conf. for Internet Technology and Secured Transactions 2014, vol. 1, pp. 288-294, Nov. 2013.

[7] Y. Shouyi, L. Xiaokang, “Adaptive load balancing in mobile ad hoc networks,” IEEE Wireless Communications and Networking Conference 2005, Vol. 4, pp. 1982-1987, March 2005.

[8] X. Li, W. Zhang, H. Zhang, W. Li, “A combining call admission control and power control scheme for D2D communications underlying cellular networks,” in China Communications, Vol. 13, Issue: 10, pp. 137-145, Oct. 2011.

[9] S. F Yang, J. S Wu, "Guard channel based call admission control schemes in hierarchical mobile IPv6 networks," in *Intl. Journal of Communication Networks and Information Security*, vol. 2, no. 2, pp. 68-76, August 2010.

[10] M. K. Luka, A. A. Atayero, O. I. Oshin, "Call admission control techniques for 3GPP LTE: A survey," in IEEE SAI Computing Conference, pp. 691-700, Sept. 2016.

[11] K. R. Poranki, Y. Perwej and A. Perwej, "The level of customer satisfaction related to GSM in India," in *Research Journal of Science & IT Management*, vol. 4, no. 3, pp. 30-36, 2015.

[12] Y. Fang, Y. Zhang, "Call admission control schemes and performance analysis in wireless mobile networks," in IEEE Transactions on Vehicular Technology, vol. 51, no. 2, pp. 371-382, August 2002.

[13] E. S. Mughele, W. Olatokun and T. Adegbola, "Congestion control mechanisms and patterns of call distribution in GSM telecommunication networks: The Case of MTN Nigeria," in *African Journal of Computing & ICT*, vol. 5, pp. 29-42, December 2012.

[14] K. B. Moses, A. B. Kayode, O. Fajuyigbe and O. S Adewale, "Development of models for managing congestion on Global System For Mobile Communication (GSM)," in *Nigerian Journal Of Wireless Networking And Communications*, vol. 1, no. 1, pp. 8-15, 2011.

[15] E. O. Oladeji, E. N. Onwuka and M. A. Aibinu, "Determination of voice traffic busy hour and traffic forecasting in Global System For Mobile Communication (GSM) in Nigeria," 11th IEEE Intl. Conf. on Communications, vol. 13, pp. 184-189, November 2013.

[16] S. Sahraei, D. Andgrigoras, "Multi-hop congestion control algorithm in mobile wireless networks," 8th Intl. Symposium on Parallel and Distributed Computing, pp. 21-28, 2009.

[17] J. Hou, J. Yang and S. Papavassiliou, "Integration of pricing with call admission control for wireless networks," in Proceeding of IEEE VTC conference, pp. 1344-1348, 2001.

[18] S. Kyriazakos, D. Drakoulis, G. Karetsos, K. Vlahodimitropoulos, "Enhanced capacity management for 2nd and 3rd Generation cellular networks," 52nd IEEE Intl. Conf. on Vehicular Technology, vol. 6, pp. 2701-2705, September 2000.

[19] O. A. Al-Kishrewo and M. M. Mousa, "Integration of dynamic pricing with guard channel scheme for uniform and non-uniform cellular network traffic," in Proceedings of the 2006 Intl. RF and Microwave Conf., vol. 6, pp. 363-366, September 2006.

[20] K. Ghanem, N. Z. Khan, A. Mitschele-Thiel, "Peak load reduction on the mobile networks by applying new pricing policies," 4th Intl. Symposium on Wireless Communication Systems 2007, pp. 446-450, 2007.

[21] S. Buruhanudeen, M. Othman, M. Othman, M. A. Borhanuddin, "Existing MANET routing protocols and metrics used towards the efficiency and reliability- An overview," IEEE Intl. Conf. on Communication and Telecommunications 2007, Malaysia, pp. 231-236, 14-17 May 2007.

[22] B.V. Kumar, G. Madhuri, M. Devadas, C. Kumar, "A mobile controlled handoff by using dynamic channel assignment," 2013 Intl. Multi-Conference on Automation, Computing, Communication, Control and Compressed Sensing, pp. 297-301, 2013.

[23] M. Oughdi, A. Caminada, S. Lamrous, and B. Morin, "Load regulation in mobile network with planned pricing model based on user behavior," 2005 IEEE Joint Intl. Conf. on Autonomic and Autonomous Systems and Intl. Conf. on Networking and Services 2005, pp. 53, October 2005.

[24]Press release: ITU sees 5 billion mobile subscribers globally in 2010, "Strong global mobile cellular growth predicted across all regions and all major markets," International Telecommunication Union, [Published], Retrieved from: http://www.itu.int/newsroom/ (Accessed: August 2015)

[25] D. M. Chiu and J. Raj, "Analysis of the increase and decrease algorithms for congestion avoidance in computer networks," *Computer Networks and ISDN Systems*, page 1-14, Elsevier Science Publishers.

[26] B. M. Kuboye, B. K. Alese, O. S. Adewale, "Multi-level access priority channel allocation strategies in Global System for Mobile communications (GSM) networks," 9th Intl. Conf. on Internet Technology and Secured Transactions, London, pp. 288-294, December 2014.

[27] S. F. Islam, Md. F. Hossain, “Comparison study of various call admission control scheme in WCDMA network,” Intl. Conf. on Electrical and Computer Engineering 2008, pp. 217 – 220, December 2008.

[28] P. K. G. Thakurta and S. Bandyopadhyay, “A new dynamic pricing scheme with priority based tree generation and scheduling for mobile networks,” 2009 IEEE Intl. Conf. on Advance Computing Conference 2009, pp. 1021-1025, March 2009.

[29] S. N. Ohatkar, D. S. Bormane, “Channel allocation technique with genetic algorithm for interference reduction in cellular network,” 2013 Annual India Conference 2013, pp. 1-6, 13-15 December 2013.

[30] L. B. Miguel, J. Gozalvez, “Common radio resource management algorithms for multimedia heterogeneous wireless networks, “IEEE Transactions on Mobile Computing, vol. 10, Issue 9, pp. 1201-1213, July 2011.

[31] T. O. Oyebisi and O. A. Ojesanmi, “Development of congestion control scheme for wireless mobile network, *Journal of Theoretical and Applied Information Technology*, pp. 966 – 972, 2005.

[32] N. Nasser and H. Hussanein, “Combined admission control algorithm and bandwidth adaptation algorithm in multimedia cellular networks for QoS provisioning,” Canadian Conf. on Electrical and Computer Engineering, vol. 2, pp. 1183 – 1186, 2004.

[33] S. K. Das, S. K. Sen and R. Jayaram, “A structured channel borrowing scheme for dynamic load balancing in cellular networks,” in Proceedings of the 17th Intl. Conf. on Distributed Computing Systems, page 116-123, May 1997.

[34] A. Sharma, S. Konai, and U. Bhattacharya, “New call and handoff cal1 management scheme for reuse partitioning based cellular systems,” IEEE Intl. Conf. on Recent Advances and Innovations in Engineering 2014, pp. 1-7, 2014.

[35] O. E. Falowo, ‘Joint call admission control algorithm for reducing call blocking/dropping probability in heterogeneous wireless networks supporting multihoming,” IEEE Intl. Workshop on Management of Emerging Networks and Services, pp. 611-615, 2010.

[36] T. J. Nel (2012), "Complex adaptive simulation of cellular network subscriber behavior," M.Sc., Information Technology [Unpublished], University of Johannesburg, Retrieved from: https://ujdigispace.uj.ac.za (Accessed: March 2016)

[37] M. Sumbwanyambe (2008), "Indoor mobility modelling for MANETS: an activity approach," M.Eng., Electrical Engineering [Unpublished], University of Johannesburg, Retrieved from: https://ujdigispace.uj.ac.za (Accessed: November 2015)

[38] B. R. Elbert, "Introduction to satellite communication" 1st edition Boston Hughes Communication Inc, 2008.

[39] M.Ilyas, "The handbook of adhoc wireless networks", CRC Press, 2003.

[40] R. Ganesh, K. Pahlavan, "Wireless network deployment," Springer publisher, 2000.

[41] C. Smith, D. Collins, 2001, "3G Wireless Networks," McGraw-Hill Professional, 2001.

[42] F. Hillebrand, "GSM and UMTS: The creation of Global Mobile Communication," John Wiley & Sons, 2001.

[43] H. Kaaranen *et al.*, "UTMS Networks: Architecture, Mobility and Services" 2nd edition, Wiley, 2005.

[44] E. Dahlman *et al.*, "3G Evolution, HSPA and LTE for mobile broadband, 2nd edition, Academic Press, 2008.

[45] O. A. Akinlabi, (2014) "Interference management in femtocells networks," M.Tech. Electrical and electronics engineering technology, [Unpublished], University of Johannesburg, Retrieved from: https://ujdigispace.uj.ac.za (Accessed: February 2016)

[46] J. H. Schiller, "Mobile Communications," Addison-Wesley, 2003.

[47] A. R. Mishra, "Fundamentals of cellular network planning and optimisation: 2G/2.5G/3G….Evolution to 4G," 1st edition, Wiley, 2004.

[48] GSM Association, "A brief history of GSM and the GSMA," [Published], Retrieved from: https://gsmworld.com/about-us/history.html (Accessed: February 2016)

[49] A. Sgora and D. D. Vergados, “Handoff prioritization and decision schemes in wireless cellular networks: a Survey,” IEEE Communications Surveys & Tutorials, 4th Quarter, vol. 11, no. 4, pp. 57-77, 2009.

[50] N. Verma, I. R. Chen, “Admission control algorithms integrated with pricing for revenue optimization with QoS guarantees in mobile wireless networks,” in Proceedings of the 10th Intl. Conf. on Parallel and Distributed Systems 2004, pp. 495- 502, 2004.

[51] I. Modeas, A. Kaloxylos, N. Passas and L. Merakos, “An algorithm for radio resources management in integrated cellular/WLAN networks,” IEEE Intl. Symposium on Personal, Indoor and Mobile Radio Communications 2007, pp. 1-5, September 2007.

[52] S. Kabahuma, (2010), “Joint call admission control incorporating pricing for congestion control to enhance QoS and ensure revenue for network operators in next generation wireless networks, M.Sc., Electrical engineering, [Unpublished], University of Capetown, Retrieved from: https://uct.ac.za (Accessed: September 2015)

[53] R. E. Ahmed, “A hybrid channel allocation algorithm using hot-spot notification for wireless cellular networks,” IEEE Canadian Conf. on Electrical and Computer Engineering 2006, pp. 891-894, May 2006.

[54] S. N. Ohatkar, D. S. Bormane, “Channel allocation technique with genetic algorithm for interference reduction in cellular network,” 2013 Annual India Conf. 2013, pp. 1- 6, December 2013.

[55] O. A. Alimi, M. K. Joseph and A. O. Akinlabi, “Network congestion management using call admission control” in Proceedings of the 7th IIE International Conference on Latest Trends in Engineering and Technology 2015, pp. 199-203, November 2015.

[56] D. Wong, T. J. Lim, “Soft handoffs in CDMA mobile systems,” IEEE Personal Communications, pp. 6-17, August 2002.

[57] S. Ghosh and A. Konar, “Call admission control in mobile cellular networks,” *Studies in Computational Intelligence*, vol. 437, pp. 1-62, April 2013.

[58] E. G. Sharma, “Handoff decision analysis for modern heterogeneous mobile networks to avoid ping-pong- a MATLAB approach,” in *Intl. Journal of Applied Engineering Research*, vol. 7, no 11, 2012.

[59] D. Saha, A. Mukherjee, I. S. Misra, and M. Chakaraborty, "Mobility support in IP: A survey of related protocols," IEEE Network, vol. 18, issue 6, pp. 34-40, November 2004.

[60] G. Liu, G. Zhu, Y. He, W. Lang and W. Wu, "Fair multi-services call admission in cellular networks using stochastic control," IEEE 60th Conf. on Vehicular Technology, vol. 4, pp. 2554-2558, September 2004.

[61] S. Kabahuma and O. E Falowo, "Analysis of network operators' revenue with a dynamic pricing model based on user behavior in NGWN using JCAC," in Proceedings of the Southern Africa Telecom. Networks and Application Conference 2010.

[62] N. Nasser and H. Hassanein, "Adaptive call admission control for multimedia wireless networks with QoS provisioning," in Proceedings of Intl. Conf. on Parallel Processing Workshops 2004, pp. 30-37, August 2004.

[63] M. Veeraraghavan, "Applications of queuing theory to circuit switched networks," March 31, 2004

[64] I. Chen and C. Chen, "Threshold based admission control policies for multimedia servers," *The Computer Journal*, vol. 39, no. 9, pp. 757-766, February 1996.

[65] H. Zeng and I. Chlamtac, "Adaptive guard channel allocation and blocking probability estimation in PCS networks," *Computer Networks*, vol. 43, no. 2, pp. 163-176, March 2003.

[66] S. Tzeng, "Call admission control policies in cellular wireless networks with spectrum renting," *Computer Communications*, vol. 32, no. 18, pp. 1905-1913, August 2009.

[67] Amitansh Gupta (2015), Telecom Sector Network Congestion Statistics for Q1-Q2FY15, Statistics for Telecom Network Congestion and how to proactively handle using intelligence machine,[Published]; Retrieved from: http://www.slideshare.net/AmitanshGupta/telecom-sector-network-congestion-statistics-for-q1q2fy15 (Accessed: March 2016)

[68] Deloitte Digital SA (2013), 2013 Forecast for the South African Telecoms Market, [Published]; Retrieved from: www.deloittedigital.co.za (Accessed: May 2016)

[69] Cisco (2006), Cisco TelePresence Phone Interface, Cisco TelePresence Phone Features, [Published] Retrieved from: www.cisco.com (Accessed: May 2016)

[70] Y. K. Sung, R. H. Hwang, M. X. Chen and J. M. Hsu, "Adaptive call admission control mechanism for DS-CDMA cellular system," in Proceedings of the 6th IEEE Intl. Circuit and

Systems Symposium on Emerging Technologies: Frontiers of Mobile and Wireless Communication, vol. 2, pp. 549-552, 2004.

[71] J. Hou, Y. Fang and A. N. Akansu, "Mobility-based channel reservation scheme for wireless mobile networks," IEEE Wireless Communications and Networking Conference 2000, vol. 2, pp. 527-531, September 2000.

[72] S. K. Emmadi and T. G. Venkatesh, "Call admission control schemes in cellular networks: a comparative study," in Proceedings of the 11th IEEE Intl. Joint Conf. on Computer Science and Software Engineering 2014, pp. 188-193, May 2014.

[73] A. Khare and Y. Raut, "Review of congestion control algorithms for performance evaluation of WCDMA," in *Intl. Journal of Emerging Technology and Advanced Engineering*, vol. 1, issue 1, pp. 20-23, 2011.

[74]G. I. Tsiropoulos, D. G. Stratogiannis, P. G. Cottis and T. D. Lagkas, "Adaptive resources allocation and dynamic call admission control in wireless networks," 2010 IEEE Globecom Workshop on Mobile Computing and Emerging Communication Networks, pp. 1217-1221, 2010.

[75] S. Valaee and B. Li, "Distributed call admission control for ad hoc networks," in Proceeding of the 2002 IEEE VTC Conference, vol. 2, pp. 1244-1248, 2002.

[76] S. Soundararajan and R. S. Bhuvaneswaran, "Multipath load balancing & rate based congestion control for mobile ad hoc networks (MANET)," in 2th IEEE Intl. Conf. on Digital Information and Communication Technology and it's Applications, pp. 16-18, 2012.

[77] W. Li and X. Chao, "Call admission control for an adaptive heterogeneous multimedia mobile network," in IEEE Transactions on Wireless Communications, vol.6, no.2, February 2007.

[78] J. Choi, T. Kwon, Y. Choi and M. Naghshineh, "Call Admission control for multimedia services in mobile cellular networks: a Markov decision approach," IEEE Intl. Symposium Computers Communications, pp. 594-599, 2000.

[79] L. Dobos and J. Goril, "Call admission control in mobile wireless," in *Radioengineering*, vol.2, no. 4, pp. 17-23, 2002.

[80] X. Li, L. Chen and B. Xu, “Performance analysis on call admission control in C3G-A system,” in Proceedings of the 2009 IEEE Wireless Communications and Signal Processing conference, China, pp. 276–279, November 2009.

[81] M. A. Lynn, J. M. Peha, “The priority token bank in a network of queues communications,” IEEE Intl. Conf. Towards The Knowledge Millennium, vol. 3, pp. 1387-1391.

[82] A. Ghosh, I.S. Misra, “An analytical model for a resource constrained QoS guaranteed SINR based CAC scheme for LTE BWA Het-Nets,” IEEE Intl. Conf. on Advances in Computing, Communications and Informatics, pp. 2186-2192, Sept. 2016.

[83] D. Lopez-Perez, A. Valcarce, A. Ladanyi, G. de la Roche, and J. Zhang, “Intracell handover for interference and handover mitigation in OFDMA two-tier macrocell-femtocell networks,” in *EURASIP Journal on Wireless Communications and Networking*, vol. 1, pp. 1-15, 2010.

[84] M. Hamdy, A. S. Abdel-Rahman, and N. Semary. "Evaluation of mobile WiMAX IEEE 802.16E handover load balancing trends." in Proceedings of the 9th IEEE International Conference on Informatics and Systems, pp. CNs 25-CNs 31, 2014.