Performance Evaluation of User Mobility on QoS Classes in a 3G Network

Norrozila Sulaiman

A thesis submitted for the degree of Doctor of Philosophy in Newcastle University

NEWCASTLE UNIVERSITY LIBRARY
207 32476 2
Thesis L8857

School of Electrical, Electronic and Computer Engineering
Newcastle University

Newcastle University

May 2008

Acknowledgements

I would first like to thank my supervisor Professor Rolando Carrasco, for his support and encouragement. He allowed me to develop my skills whilst always providing guidance and discussion whenever it was needed. I would also like to express my gratitude to my co-supervisor Dr. Graeme Chester for his help and valuable advice.

There are many people to thank within the School of Electrical, Electronics and Computer Engineering at Newcastle University. I want to thank all my colleagues for all their help, support, interest and valuable hints. I would like to thank Dr. Martin Johnston and Dr. Muhammad Kalimuddin Khan for the valuable advice and suggestions during the course. I would also like to thank Mr. Shaopeng Wu for helping me with the GSM model.

Lastly, I would like to acknowledge support from Mr. Gordon Jeffrey for his willingness to help me with the computers equipment. My deepest gratitude and appreciation to my mom, husband, children, brothers and sisters, for their help and support in difficult times. Your love, understanding and patience sustained me through to the end of my PhD.

Thank you.

Norrozila

Abstract

The popularity of IP services is increasing and the demand for managing traffic with different QoS classes has become more challenging. The stability of the system is affected by the rate of voice traffic. Mobility allows users to be connected at all time where handover may occur as it is not always possible to be connected to the same base station. Mobility and handover cause severe interference, which affects overall throughput and capacity of the system. The system requires greater capacity with more coverage area. This study deals with the impact of user mobility on voice quality in IP based application in a 3G Network. The aim is to improve the system performance in mixed traffic environment.

A mathematical model is used to analyse the impact of using different type of coder on packet end-to-end delay and packet loss. The simulation results indicate that types of coder affect the system performance. Application of scheduling based on weight and load balancing technique can improve the system performance. The deployment of scheduling based on weight and a load balancing technique have been investigated to reduce the end-to-end delay and to improve overall performance in mixed traffic environment. The results under different conditions are analysed and it is found that by applying scheduling scheme, the quality of voice communication can be improved.

In addition, load balancing technique can be used to improve the performance of the system. Apart from the decrease in delay, the technique can increase the capacity of the system and the overall stability of the system can be further improved. Finally, network security is another important aspect of network administration. Security policies have to be defined and implemented so that critical sections of the network are protected against unwarranted traffic or unauthorized personnel. The impact of implementing IPSec has been tested for voice communication over IP in a 3G network. Implementing the security protocol does not significantly degrade the performance of the system.

List of Tables

List of Figures

List of Symbols

α	Active state
β	Inactive state
λ	Arrival rate
m	Mean arrival rate
π_1	Voice in a queue
π_2	Data in a queue
t	Time
μ	Utilisation
v	Velocity
BS	Base station
C	Conversational service user
D	Data service user
H	Handover users
K	Path loss
N	Number of users
R	Bit rate
S	Service time
T	Total delay
Wi	Weight assigned
W	Waiting time

Abbreviations

1G – 1st generation

2G – 2nd generation

3G – 3rd generation

4G – 4th generation

3GPP – 3rd Generation Partnership Project

ACELP – Algebraic Code Excited Linear Prediction

AES – Advanced Encryption Standard

AH – Authentication Header

AMPS – Advanced Mobile Phone Service

AMR – Adaptive Multi-Rate

AMS – Advanced Mobile Phone Service

ATM – Asynchronous Transfer Mode

AUC – Authentication Centre

AUTN – Authentication Network

BER – Bit Error Rate

BLER – Block Error Rate

BS – Base Station

BSC – Base Station Controller

BSS – Base Station Subsystem

BTS – Base Transceiver Station

CBR – Constant Bit Rate

CDMA – Code Division Multiple Access

CM – Connection Management

CN – Core Network

CPICH – Common Pilot Channel

DECT – Digital Enhanced Cordless Telecommunications

DRNC – Drift RNC

DSS – Digital Signature Standard

ESP – Encapsulating Security Payload

ETSI – European Telecommunications Standards Institute

FDD – Frequency Division Duplex

FDMA – Frequency Division Multiple Access

GGSN – Gateway GPRS Support Node

GMSC – Gateway MSC

GPRS – General Packet Radio Service

GSM – Global system for mobile communications

GSM-EFR – GSM Enhance Full Rate

HLR – Home Location Register

IETF – Internet Engineering Task Force

IK – Integrity Key

IKE – Internet Key Exchange

IMT2000 – International Mobile Telecommunications 2000

IP – Internet Protocol

IPP – Interrupted Poisson Process

IPSec – IP Security

IS-95A – Interim Standard 95A

ISDN – Integrated Services Digital Network

ITU – International Telecommunications Union

ITU-T – ITU-Telecommunication Standardization Sector

MM – Mobility Management

MOS – Mean Opinion Score

MS – Mobile Station

MSC – Mobile services Switching Centre

NSS – Network Switching Subsystem

PCM – Pulse Code Modulation

PDP – Packet Data Protocol

PPP – Point Protocol

PSTN – Public Switched Telecomm Network

QoS – Quality of Service

RBS – Radio Base Station

RFCs – Request For Comments

RNC – Radio Network Controller

RNS – Radio Network Subsystem

RRM – Radio Resource Management

RSA – Rivest, Shamir and Adleman

RSCP – Received Signal Code Power

RSSI – Received Signal Strength Indicator(Signal Quality)

RTP – Real-Time Protocol

RTCP – Real-Time Control Protocol

SGSN – Serving GPRS Support Node

SIM – Subscriber Identity Module

SIP – Session Initiation Protocol

SMS – Short Message Service

SRNC – Serving RNC

SS7 – Signalling System Number 7

TDCDMA –Time Division CDMA

TDD –Time Division Duplex

TDMA – Time Division Multiple Access

UDP – User Datagram Protocol

UTRAN – UMTS Terrestrial Radio Access Network

UWC136 – Universal Wireless Communications 136

VAD – Voice Activity Detection

VBR – Variable Bit Rate

VLR – Visitor Location Register

VoIP – Voice Over Internet Protocol

VPN – Virtual Private Network

WCDMA – Wideband CDMA

Table of contents

Chapter 1

Introduction

1.1. Evolution of Mobile Communications Networks

One of the advanced forms of human communications is wireless communications. The development of mobile communications has been evolved over more than 20 years. Mobile communication technology is generally categorized into three generations. First generation (1G) appeared in Europe in the 1980s [1-3]. The 1G network used analogue circuit-switched technology, with Frequency Division Multiple Access (FDMA) which only supports voice service. The networks had a low traffic capacity and unreliable handover with poor voice quality. In order to enhance the capacity and improve quality, second generation (2G) of mobile communication systems was introduced. 2G uses a combination of FDMA and Time Division Multiple Access (TDMA) technology.

The 2G was the first purely digital technology that had made digital communications more economical than analogue technology. The system is based on circuit switching and it was designed for voice services. Hence, it is not efficient in handling packet-oriented services. Besides conveying speech, 2G provided better voice quality with high bit rates. It improved security and it has an ability to send data such as Short Message Service (SMS). 2G systems offer a higher spectrum efficiency and enhanced capacity. The most popular networks for 2G cellular mobile systems are the Global System for Mobile communications (GSM) and the Interim Standard 95A (IS-95A) system. GSM supports moderate bit-rate voice communications and limited multimedia communications. However, as the technology changed, there was a need for better services. The limits of 2G networks are the main reasons for considering third generation (3G) networks. Thus, 3G evolved and it can be seen as the evolution of 2G systems as explained in [4].

The 3G uses a new spectrum of about 2 GHz and a vast amount of money has been paid for the licences [5]. 3G systems like the Universal Mobile Telecommunications System (UMTS) have emerged to provide a high data rate transmission and to support heterogeneous services. It is capable of providing wider bandwidth and increased capacity. UMTS is developed by Third-Generation Partnership Project (3GPP) that is a standard body for its development [6].

The 3G systems use Code Division Multiple Access (CDMA), which is a packet switching technology. It is aimed to provide wider bandwidth and increased capacity. The system is capable to support few applications such as the Internet, multimedia application, voice, and data in a single network. It provides the ability to transfer both voice and data simultaneously. It was designed to support circuit and packet data at high data rates from 144Kbps to 2 Mbps. The mobile communications generations are as shown in Figure 1.1 [7].

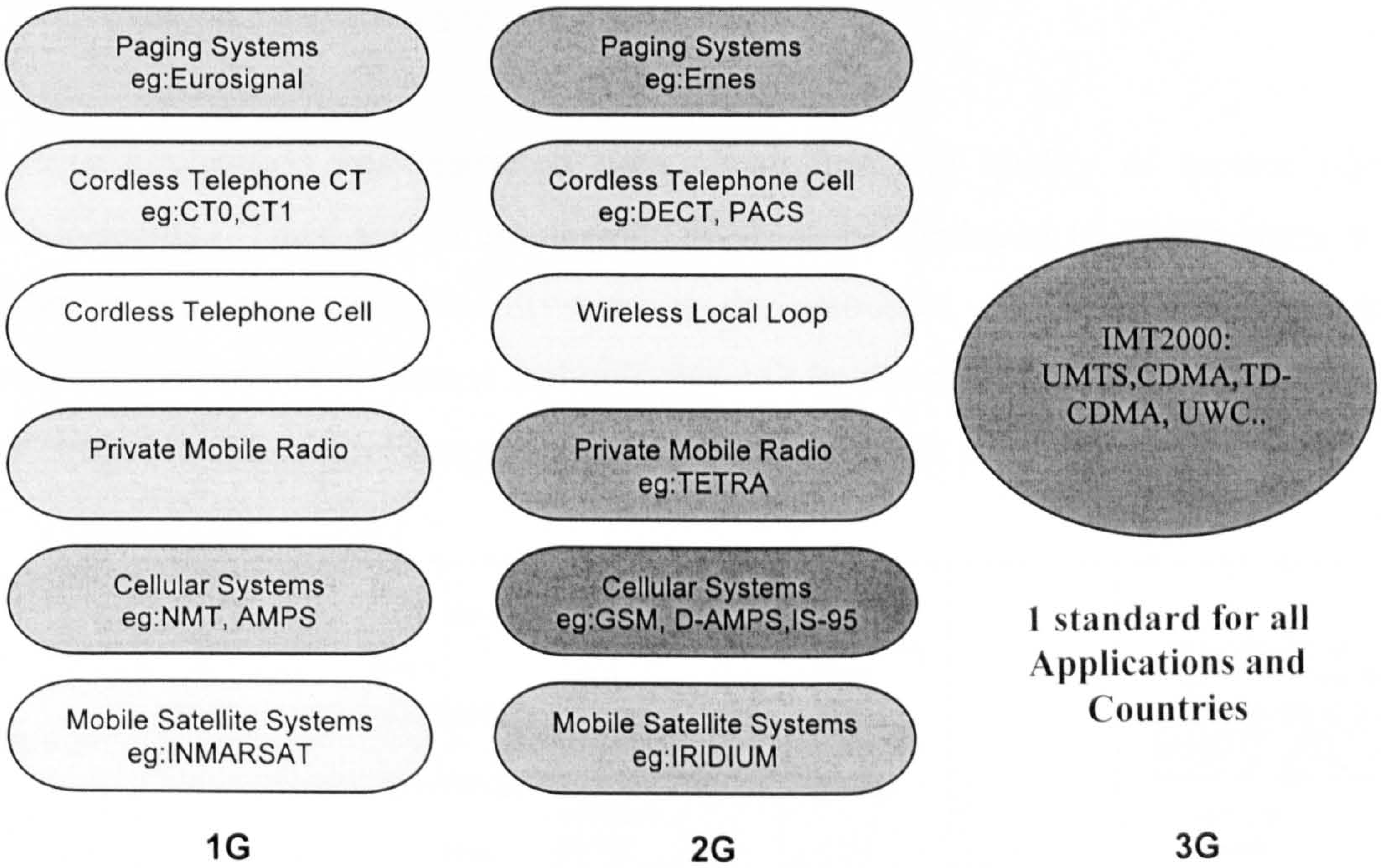


Figure 1.1: Generations of Mobile Networks [7]

International Mobile Telecommunications 2000 (IMT 2000) is a term that represents the family of technologies which include UMTS, CDMA2000, Digital Enhanced Cordless Telecommunications (DECT), Universal Wireless Communications 136 (UWC136) and 3G/IMT2000. It is a set of requirements which was defined by the International Telecommunications Union (ITU) [8]. The ITU specified the list of requirements of IMT2000 as follows: [9]

- High data rates requirements are up to at least 144Kbps
- Efficient support of asymmetric traffics, i.e. support bursty and asymmetric traffic in an efficient way by providing more flexibility in terms of FDD and TDD mode. This involves supporting ISDN and IP applications, by supporting high bit rate bearer services with QoS characteristic.

- Integrating circuit switching and packet switching
- High speech quality. There are many challenges in an IP network for instance, due to the inherent nature of the network such as delay, packet loss or bit error rate, and jitter.
- High spectrum efficiency. Improvement of spectrum efficiency can be achieved using several technologies such as such as the advanced channel coding schemes, diversity approach, digital modulation schemes, accurate channel estimation and antenna techniques.

Certain application has restricted time constraints and quality of service (QoS) requirements. Thus, a QoS framework needs to be deployed to enable users with different requirements efficiently sharing the resources. A flexible network that integrates various radio access technologies will be developed. Figure 1.2 shows the spectrum allocations in Europe, Japan, Korea and USA [1].

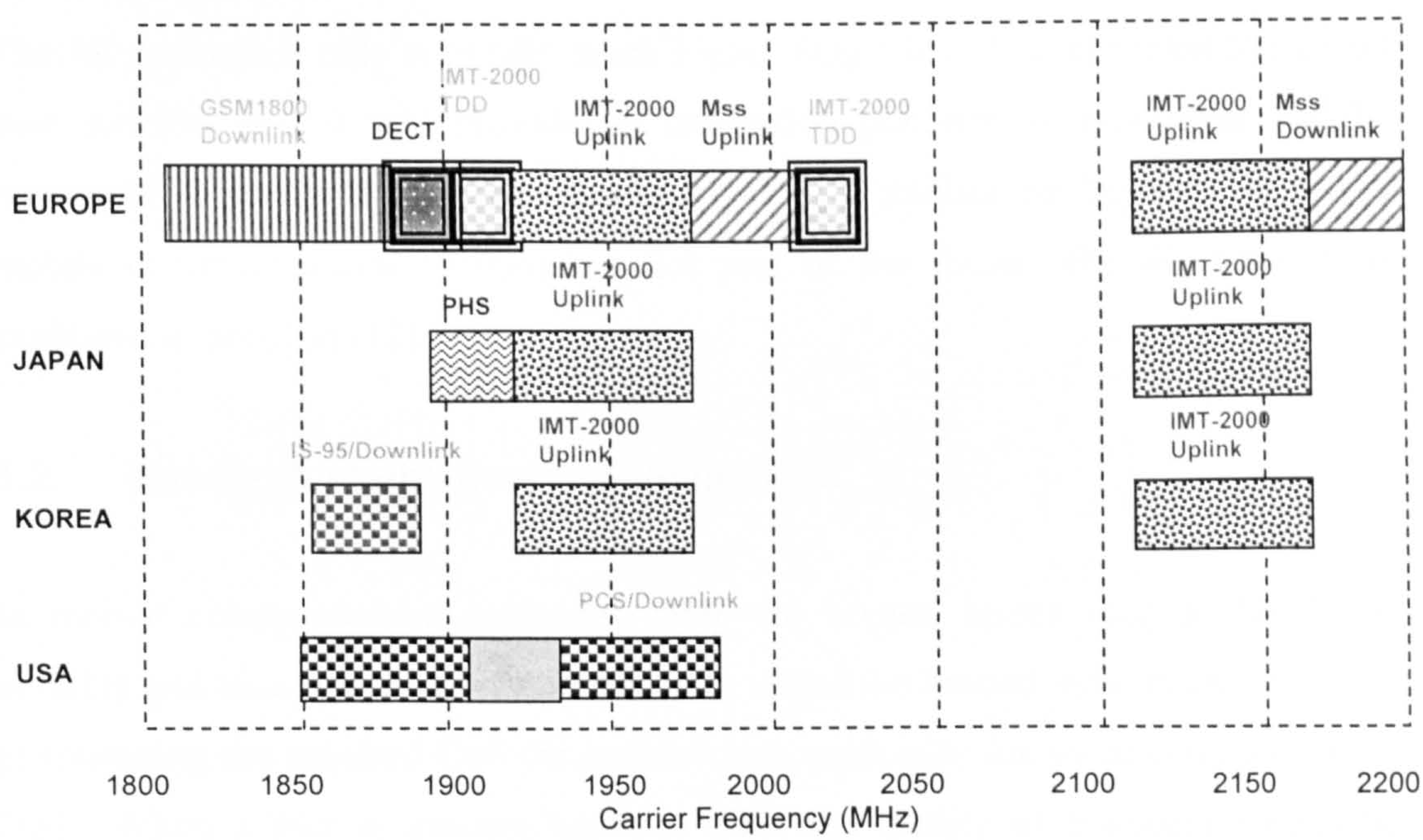


Figure 1.2: Spectrum Allocations in Different Countries [1]

Problems will arise in migrating from the current systems to 4G which includes mobile station, system and service. Thus, a new type of mobile terminals must be able to adapt to multiple wireless network that have different protocols and technologies seamlessly. The terminal should be able to choose a specific service from all the available wireless networks. It involves the networks specifications such as bandwidth, QoS supported and user preferences. There are several mobility issues such as location management and handover management. Roaming between different networks must be transparent to the user, for instance.

Managing vertical and horizontal handovers are necessary in order to maintain the communications while the user crosses the network boundaries. In addition, security seems to be another problem since there are many different types of mobile technologies. Technical aspects of different generation of mobile communications are explained in depth in [10].

3G system was designed to develop new technologies to further enhance the mobile practice. On the other hand, the fourth generation (4G) is trying to accomplish universal communication in which the networking infrastructure is more flexible. The main objectives are unlimited services, multi service platform and a low bit cost. The 4G network is able to handle much higher data rates. It is expected to support user mobility and it will provide an integration platform so that users will be connected to different networks seamlessly [11]. Studies on beyond 3G or 4G mobile communications systems are not part of the thesis. The 4G network is explained in detail in [12].

1.2. Research Motivation and Challenges

In mobile communication networks, there are several issues such as handover, mobility and security. It is very important to utilize the limited radio resources while guaranteeing the required QoS for mobile users especially for voice communication [13]. When a user is moving between cells, the quality of connection must be maintained. If handover does not occur quickly, the QoS may degenerate below an acceptable level. In order to ensure continuity of service and to avoid poor quality of services, mobiles in active calls that are leaving one base station coverage area need to be handed over to another base station. The handover should provide seamless services connectivity to the user.

Managing handovers in a mixed traffic environment is very important. Although considerable amount of research on voice over Internet Protocol (VoIP) research has been carried out in recent years, there are several weaknesses:

- Few studies have been performed on the handover of voice in Internet Protocol (IP) based cellular networks. Papers on VoIP mainly focused on voice traffic and seldom consider integrating them with other IP based applications.
- Lack of research on the impact of handover on voice quality in traffic mix environment. In UMTS system, bandwidth availability is limited. The bandwidth need to be used efficiently and effective method of optimising the resources is necessary.
- Few investigations have been carried out on the scheduling method for different types of application based on IP networks.
- Few studies on the impact of security on the QoS of voice in a mixed traffic environment.
- Few studies have been investigated on managing traffic mix. It involves managing how the system deals with mixed traffic since real-time services requiring high QoS

As these weaknesses may reduce the IP based application performance in UMTS system, the task of this thesis is to investigate the impact of traffic on performance of the system. This includes the QoS of voice such as packet end-to-end delay, packet loss and jitter in IP based cellular network. It also considers an efficient allocation bandwidth in traffic mix environment and the impact of security on the overall system performance.

1.3. Aims and Objectives

The project aim is to investigate a strategy for managing IP traffic in UMTS environment. It involves three main areas that are handover algorithms, mobility modelling and management, as well as the security issues. It is essential to evaluate the performance of handover in IP based 3G network in order to improve the load sharing and to enhance network capacity.

The objectives of this work are as follows:

- To investigate how the 2G and 3G systems work, specifically GSM and UMTS.
- To understand the technologies currently available such as 2G, 3G and 4G.
- To investigate the handover, security and mobility issues in 3G
- To analyse the impact of voice traffic in IP based 3G network.
- To investigate the way of optimizing the resources in mixed voice and data traffic environment.
- To recognise and look for ways to improve the overall system performance in handover process
- To investigate the impact of security protocol on the QoS of the system.

1.4. Scope and Statement of Originality

The following are believed to be the original points of this thesis:

- A model for analysing the voice quality in IP based UMTS network using different types of coders and comparing the performance with the queuing theory (Chapter 4).
- Evaluation of the impact of increasing voice traffic and number of users on system performance (Chapter 4 and 5).
- Analysis of the QoS of voice in mixed traffic environment and investigation of a method to manage traffic mix (Chapter 5).
- Estimation of the performance of voice handover and network capacity in a mixed traffic environment using a load balancing technique (Chapter 5).
- Evaluation of the impact of IP security on QoS and system performance (Chapter 6).

1.5. Organisation of Thesis

The remainder of this thesis is organized as follows. Chapter 2 outlines an overview of the GSM and UMTS systems. The differences and similarities will also be reviewed and discussed. After reviewing GSM and UMTS, mobility and different types of handover are presented in chapter 3.

Chapter 3 gives a general system model for GSM and UMTS networks. It includes a handover algorithm in GSM and UMTS environments. It also described the types of handover in UMTS system. Chapter 4 describes traffic modeling for the impact of coders on the voice quality. The benefits as well as challenges of VoIP are discussed and investigated in this chapter. Managing traffic mix is described in chapter 5. This chapter mainly investigates the impact of handover on mixed traffic and a way of improving the performance of traffic mix. Security issues are described in chapter 6 that discusses the impact of security on performance. Concluding remarks and future research work are summarized in chapter 7.

Chapter 2

Theoretical Background

2.1 Introduction

GSM has a predominant role in the wireless technologies and the historical background is shown in Appendix A [14]. GSM provides voice and low speed data services as it was designed primarily for a narrowband voice and circuit switched data traffic. There is a requirement for expansion of cellular services to deal with high speed mobile Internet, intranet, video and other data oriented services. UMTS provides high speed data, as well as multiple data rate services. It offers a higher user bit rate which is 384Kbps on circuit connections and it can reach up to 2Mbps on packet switched connections [1].

As GSM and UTMS operate on a different standard, it is necessary to keep mobility and connection without interruption. The purpose is to support mobility of users and interoperability of different network technologies, especially when the users move between different network areas. Both technologies offer a great opportunity to provide integrated services using a wireless network technology. 3GPP provides technical specification for GSM and UMTS standardization [15].

2.2. Related Work

Mobility, security and different in standards are likely to be the major problems in the handover. Although various investigations on different aspects of handover have been performed by many authors as in [16-27], only a few involving the handover performance of packetised voice in IP based cellular network has been reported as in [28-32]. One of the important aims in handover algorithm design is to decrease the unnecessary handovers. Therefore, a correct choice of handover conditions and parameters is essential. The call dropping rate, call quality, handover success and avoidance of unnecessary handovers, are the parameters that seem can improve the system performance. Some of the performance metrics used to evaluate the handover algorithms are the call blocking probability, handover blocking probability, handover probability, call dropping probability, probability of an unnecessary handover, rate of handover, duration of interruption and the delay [16].

Many handover algorithms proposed in the literature are classified according to the metrics used to decide whether a handover is necessary or not. It also proposed where these metrics are monitored and how these metrics are processed. Some studies were based on hard handover for time division multiple-access systems, where handover was performed at the cell boundary by switching the radio connection of the Mobile Station (MS) from one Base Station (BS) to the other. In [17], the author found that number of handovers and handover initiation delay were the factors that affect the performance measures. The decision to initiate a handover was based on the received signal and distance from the base stations. In [18] for example, an investigation of the handover algorithms for implementation in an urban mobile cellular was analysed. A method comprises of steps to determine a number of potential handover or reselection target cells was analysed in [19].

Another way to improve system performance of cellular networks is to use efficient handover schemes when users move between cells. Handover scheme was classified vertically and horizontally as shown in Figure 2.1. Vertical handover is based on the concepts of the handover scheme that are adopted whereas horizontal handover is based on how mobile hosts identify the potential base stations to be connected. Handover schemes based on prediction techniques have the potential to enhance the performance of cellular networks [20].

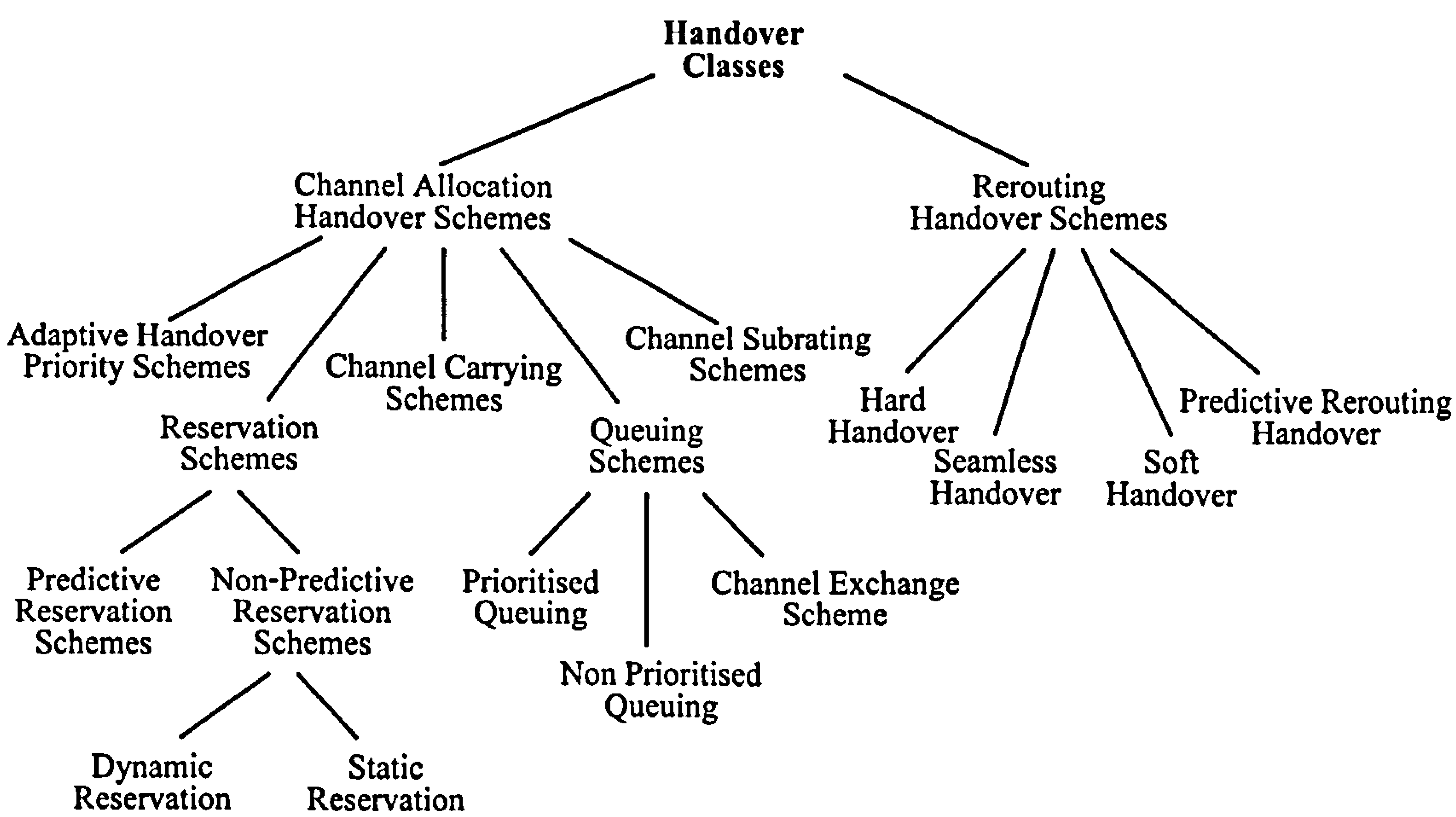


Figure 2.1: Handover Schemes [20]

One of the major issues to be considered is the effective handover. The communication link gets worse as the signal from the first base station weakens. Thus, the decision to make the handover should not be delayed too much. A few approaches were proposed to solve the handover decision problem. Some author considered the handover criteria in the order of user defined priority as in [21]. The author used context information regarding user devices, user location, networks environment and requested QoS in order to assist handover decision process.

In [22] a handover strategy for 4G wireless networks was proposed where handovers are classified as imperative and alternative handovers. The decision criteria were based on static and dynamic information. The handover decision was based on price of the service, QoS support, connections time in a cell, as well as battery consumption. In addition, users also played an important role in the making decision. For instance, users could indicate their preferences on certain criteria, or even dynamically change their preferences. However, various criteria in the decision process might oppose to each other. Thus, trade-offs are sometimes required. Fuzzy logic was applied to deal with imprecise information of some criteria and user preference.

There are several handover decision algorithms that use fuzzy logic, neural network and pattern recognition algorithms to process the collected metrics as in [23-27]. In [23], a fuzzy logic based handover scheme was used to evaluate the intersystem handover performance. The author focuses on evaluating the performance of a fuzzy algorithm for the handover process that was based on the received signal strength indicator (RSSI) criteria. It was claimed that the proposed scheme reduces the handover uncertainty as well as the number of unnecessary handovers.

Fuzzy logic was used in [24] to process handover metrics such as received signal strength measurements from the access points, the ratio of the used capacity to the total capacity for the access points, as well as relative directions and speeds of the mobiles to the access points.

A membership value is the degree of truth of a statement and the value is between 0 and 1. The threshold to trigger the handover decision algorithm was introduced. A handover would only occur when the membership value to its current access point is below a predetermined membership threshold.

It was reported in [25] that although the use of hysteresis and averaging window in classical handover techniques reduced unnecessary handovers, it caused delays that may result in the calls being dropped. Therefore, a fuzzy based handover algorithm was proposed. The performance of fuzzy based handover algorithm was claimed to be better, compared to that obtained using the classical handover algorithms.

Pattern recognition was used in [26] to determine the position of the mobile station for handover control. One of the approaches was based on neuro-fuzzy methods whereas the other was based on hidden Markov models. It was claimed that patterns are the sufficient characteristic for the proposed application. Neural network based on signal level measurement was used in [27] for the implementation of the handover initiation algorithm. The technique allows adaptability to traffic conditions as well as fast response compared to methods that are based on hysteresis margins.

Several studies on the performance of VoIP have been investigated. In [28] for instance, an analytical model was introduced to quantify the impact of the power budget and channel coding rate on the quality of VoIP calls, such as delay and packet loss probability. The performance of VoIP over 3G was studied in [29] where the VoIP network capacity under different packet delays was evaluated. A full rate GSM speech codec that uses silence compression techniques was used.

There are a few ways of managing VoIP handover which involves link layer, network layer and transport layer. An approach of handling mobility at the network layer by using Mobile IP was used in [30]. The disruption time is expected to be higher in an IPv6 environment as it has bigger message size. The author focuses on the network delay and the delay jitter effect on VoIP in [31]. The numerical result for investigation of the delay and delay jitter effect on a speech sample are presented.

In [32], the results of QoS for VoIP services over 3G networks were presented. The Adaptive Multi Rate (AMR) speech coder at 12.2Kbps was used. UMTS to GSM handover is based on an intersystem measurements, which are usually performed using the compressed mode technique defined in the 3GPP standard. A description of inter radio access technology handover procedure is described in [33, 34].

2.3 GSM Networks

GSM is a circuit-switched system which is used for transmitting voice and data services. The number of users continues to increase although the latest technology such as 3G has evolved in the market. The number of GSM connections until mid year 2007 as shown in Table 2.1 [35].

Table 2.1: Number of GSM Connections [35]

Market	Q1 2006 (~billion)	Q2 2006 (~billion)	Q3 2006 (~billion)	Q4 2006 (~billion)	Q1 2007 (~billion)	Q2 2007 (~billion)
World	1.83	1.93	2.04	2.17	2.28	2.38
Africa	0.15	0.16	0.17	0.19	0.21	0.22
Americas	0.16	0.18	0.20	0.22	0.23	0.25
Asia Pacific	0.67	0.72	0.76	0.81	0.87	0.92
Europe: Eastern	0.28	0.30	0.32	0.34	0.35	0.36
Europe: Western	0.38	0.38	0.39	0.39	0.39	0.39
Middle East	0.11	0.11	0.12	0.13	0.14	0.15
USA/Canada	0.08	0.08	0.09	0.09	0.09	0.09

A GSM network is composed of several functional entities, whose functions and interfaces are specified. GSM networks can be divided into three broad parts i.e. MS, Base Station Subsystem (BSS) and Network SubSystem (NSS). MS is the equipment that is carried by the subscriber. BSS composes of the Base Transceiver Station (BTS) and the Base Station Controller (BSC). It controls the radio link with the MS. The BTS contains the radio transceivers that define a cell and handles the radio link protocols with the MS [2, 36]. On the other hand, the BSC manages the radio resources for one or more BTSs. It handles radio channel setup, frequency hopping and handovers.

Mobile services Switching Centre (MSC) is the main part of NSS. NSS performs the switching of calls between the mobile and other fixed or mobile network users, as well as management of mobile services, such as authentication [37]. Figure 2.2 shows the layout of a generic GSM network.

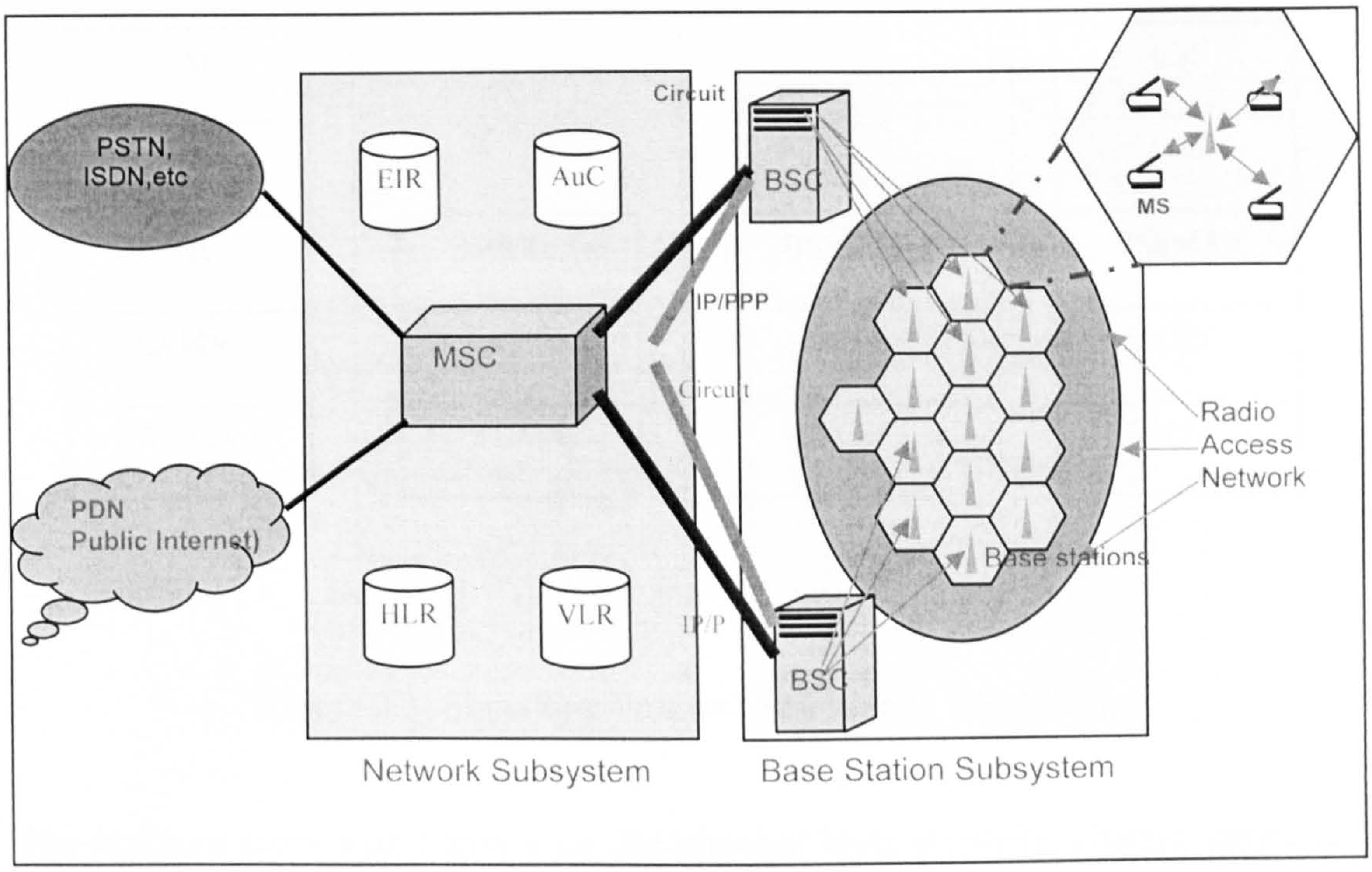


Figure 2.2: General Architecture of a GSM Network [35]

MS and the BSS communicate across the Um interface, which is also known as the air interface or radio link. BSS communicates with the MSC across the A interface. The MSC provides the connection to the Public Switched Telephone Network (PSTN) or Integrated Services Digital Network (ISDN), and signalling between functional entities uses the ITU-T Signalling System Number 7 (SS7). Home Location Register (HLR), Visitor Location Register (VLR) and MSC provide the call routing as well as roaming capabilities of GSM. A detail explanation of the architecture is described in [2, 36-38].

Part of the function of a cellular mobile network is assuring the quality of voice and data transmission over the network. The user mobile requires registration, authentication and location updating in the network. The area covered by the network is divided into cells which involve the implementation of handover mechanism. GSM uses signaling process to interface with the communication items. The signaling environment involves MS, BTS, BSC and MSC. A general view of the signaling protocol structure is as shown in Figure 2.3.

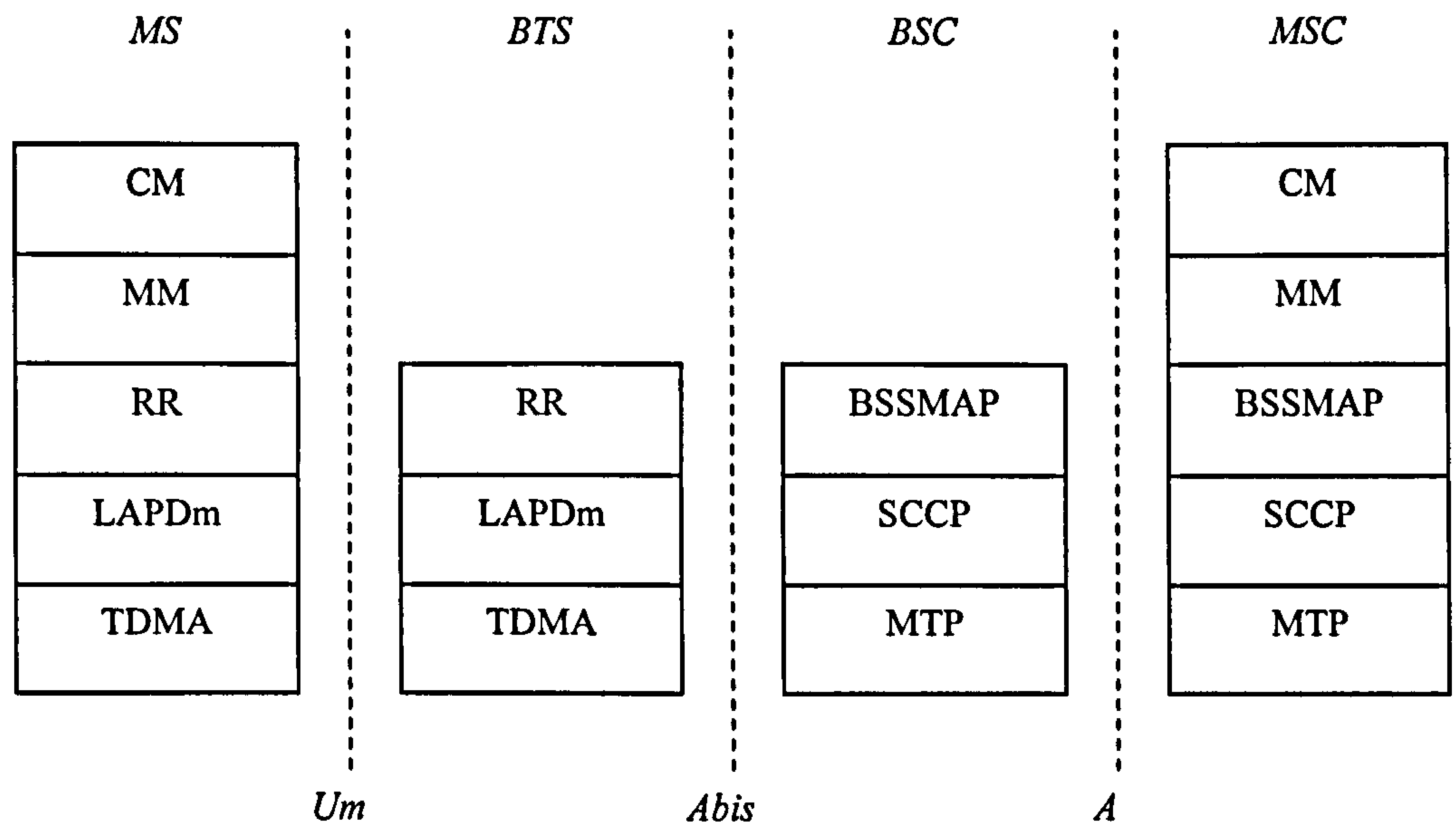


Figure 2.3: Signalling Protocol Structure in GSM [35]

The structure starts with Layer 1 i.e. the physical layer involving channel structure. It follows with Layer 2 that is data link layer. Layer 3 is the network management layer which includes Radio Resource Management (RRM), Mobility Management (MM) and Connection Management (CM). The RRM layer controls the setup, maintenance and termination of connection including handovers involving the mobile station and the MSC. The MM layer manages the location updating, registration, security and authentication. The CM layer handles call control and manages supplementary services [37].

2.4 UMTS Networks

A UMTS network consists of two main elements i.e. Core Network (CN) and the UMTS Terrestrial Radio Access Network (UTRAN). The CN is based on wired network that can be connected to IP backbone or a PSTN. The UTRAN is composed of Radio Network Controller (RNC) and Node B which is shown in Figure 2.4 [1-3, 8, 39, 40].

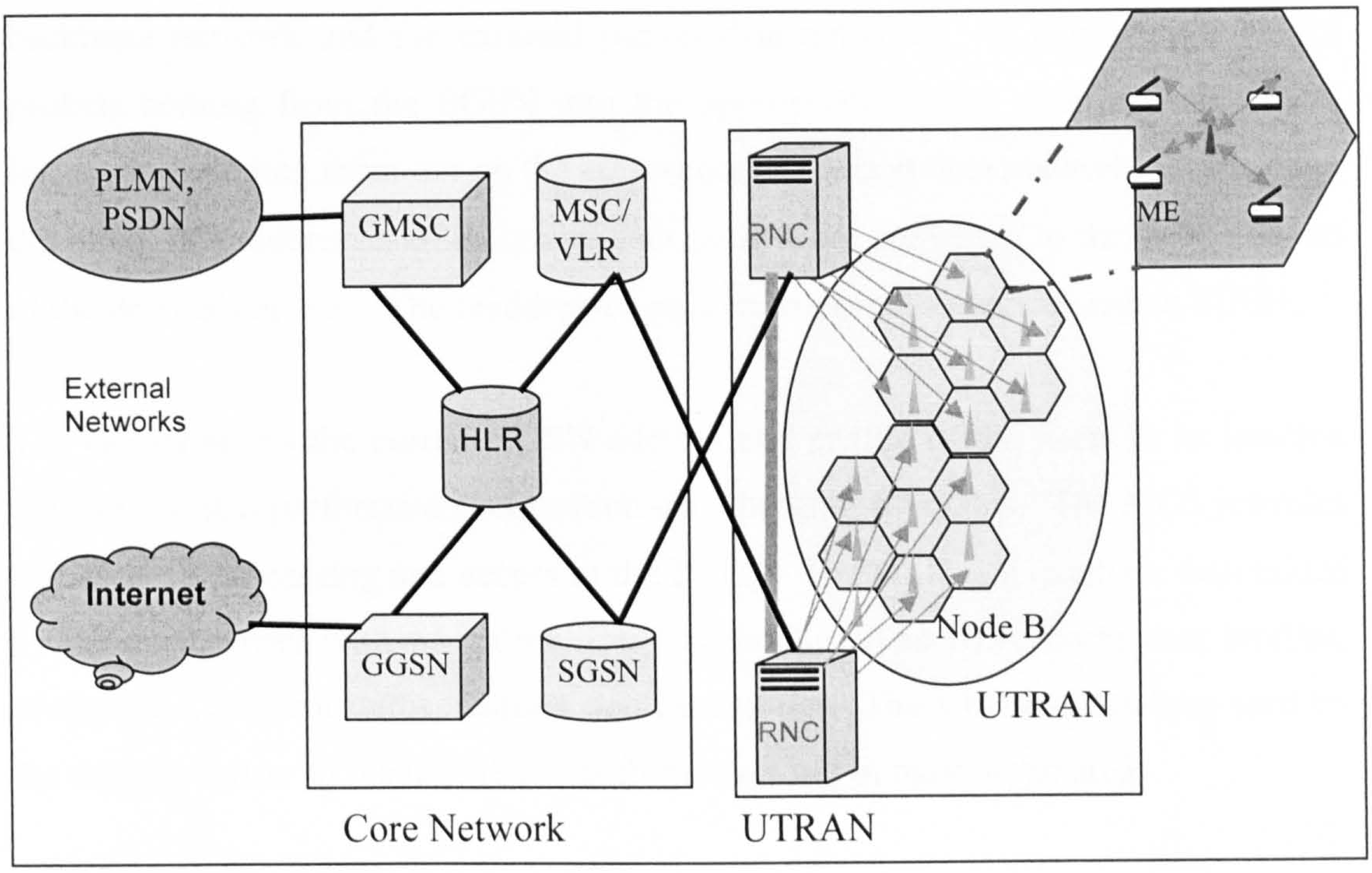


Figure 2.4: Layout of a Generic UMTS Network [37]

RNC is responsible for managing the radio resources including call set-up, the processing of voice and data traffic, as well as handover between cells. It enables RRM and provides central control for the Radio Network Subsystem (RNS) elements. RNS consists of the RNC and Node Bs. Node B is a base station which connects to the RNC. It is the physical unit for radio transmission or reception with cells. The Node B can support both Frequency Division Duplex (FDD) and Time Division Duplex (TDD) modes. The FDD mode is based on Wideband CDMA (WCDMA) and the TDD mode is based on Time Division CDMA (TDCDMA). The concept of RRM in WCDMA is discussed in [41].

The Core Network consists of the Serving GPRS Support Node (SGSN), Gateway GPRS Support Node (GGSN), HLR, MSC, VLR and Gateway MSC (GMSC). The SGSN is responsible for the delivery of data packets from and to the mobile stations within its service area. The tasks also include packet routing and transfer, mobility management, logical link management, as well as authentication and charging functions.

The GGSN acts as an interface between the General Packet Radio Service (GPRS) backbone network and the external packet data networks. It converts the GPRS packets coming from the SGSN into the appropriate packet data protocol (PDP) format and sending them out on the corresponding packet data network. In the other direction, PDP addresses of incoming data packets are converted to the GSM address of the destination user. The readdressed packets are sent to the responsible SGSN.

The GGSN stores the current SGSN address and profile of the users in its location register. It also performs authentication and charging functions. The MCS provides call control, processing and access to the PSTN. The HLR is a database maintained by the user carrier with which the user has service. The HLR stores user profiles, preferences, account status, features and capabilities. The VLR is a database used by the serving carrier to manage requests from users not in their home area.

2.4.1 Multiple Access Technologies

The amount of bandwidth is limited and the system has to accommodate as many users as possible. Therefore, effective sharing of the bandwidth is necessary. Allowing multiple users to simultaneously share the bandwidth is known as multiple access technology [8, 37, 39, 42, 43]. There are three types of multiple access technologies, i.e. FDMA, TDMA and CDMA as shown in Figure 2.5(a-c). The differences between these transmissions schemes can be seen in [44].

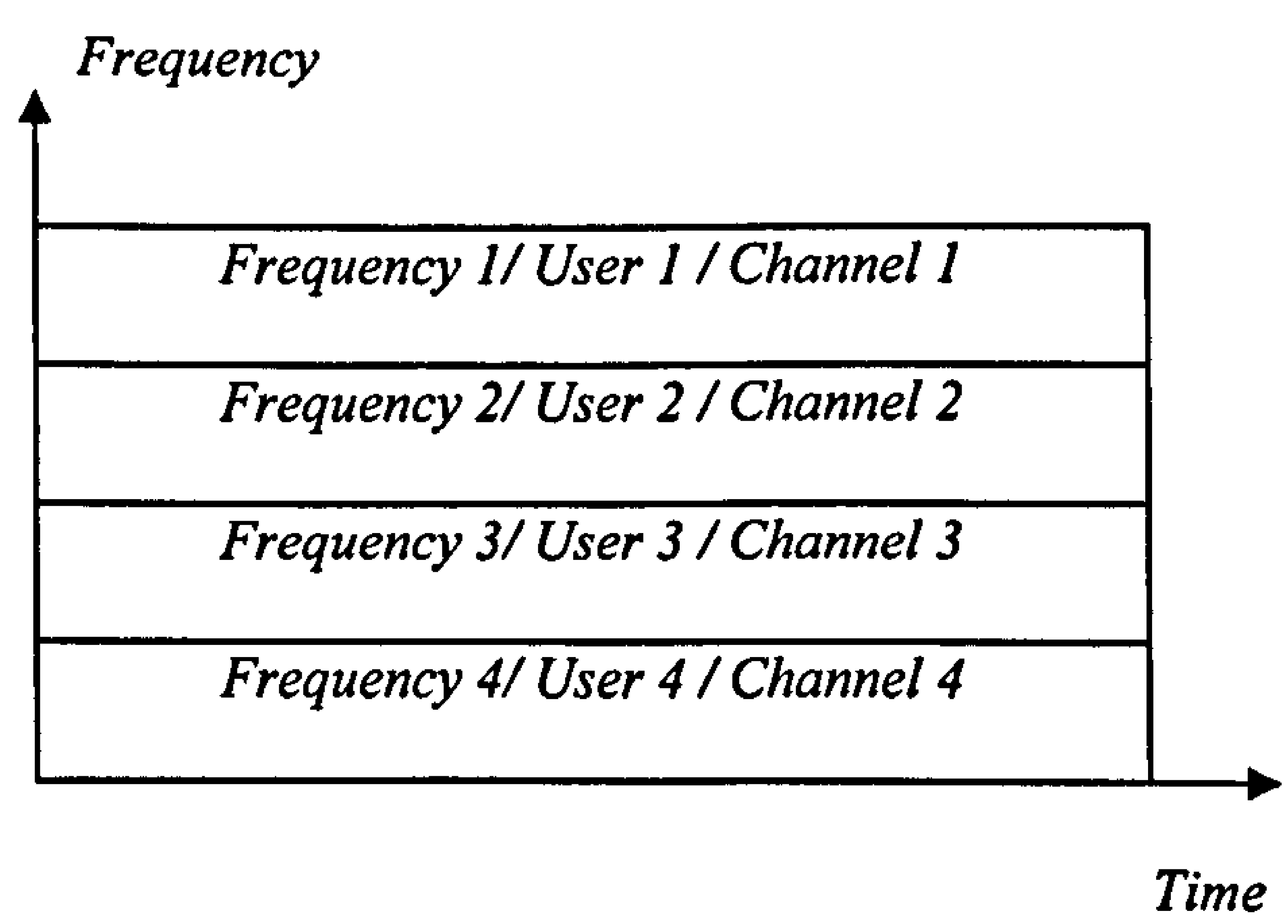


Figure 2.5a: Access Technologies: FDMA

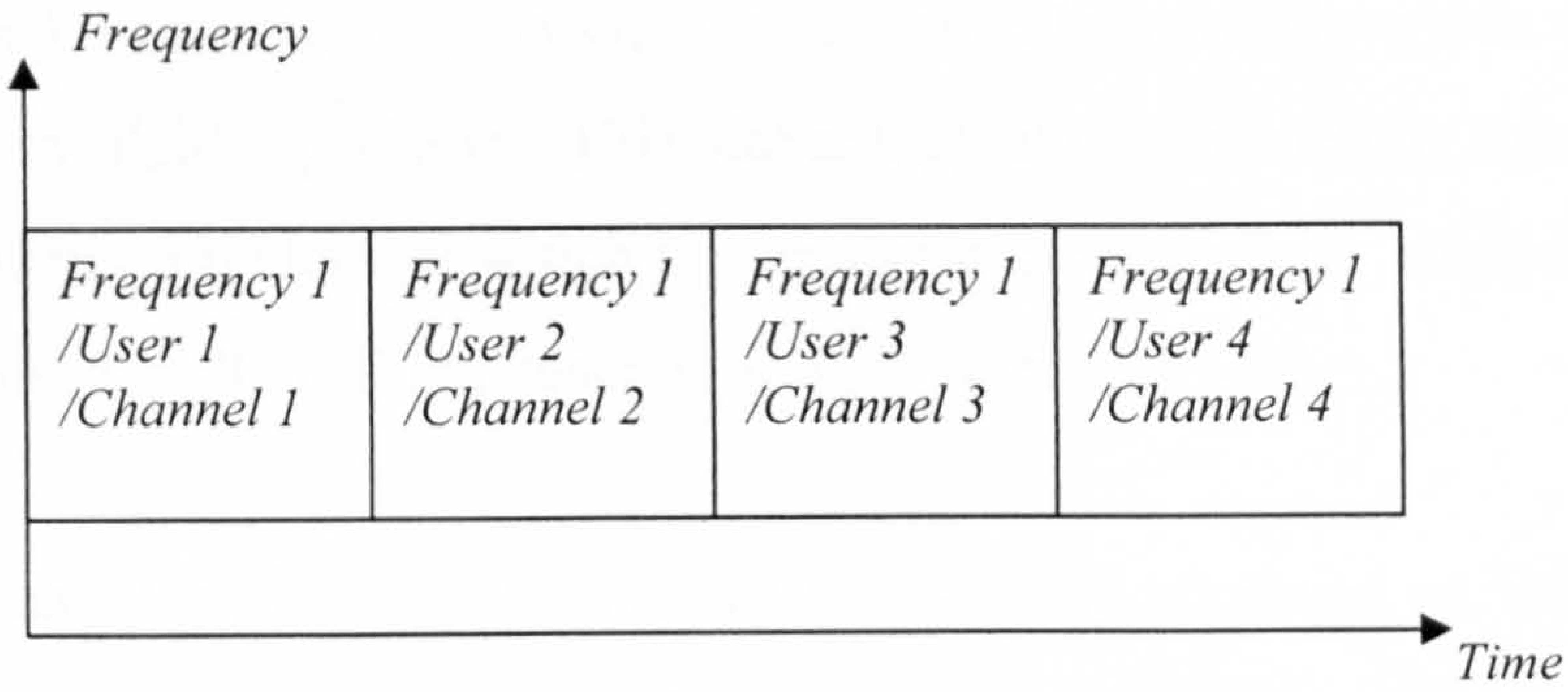


Figure 2.5b: Access Technologies: TDMA

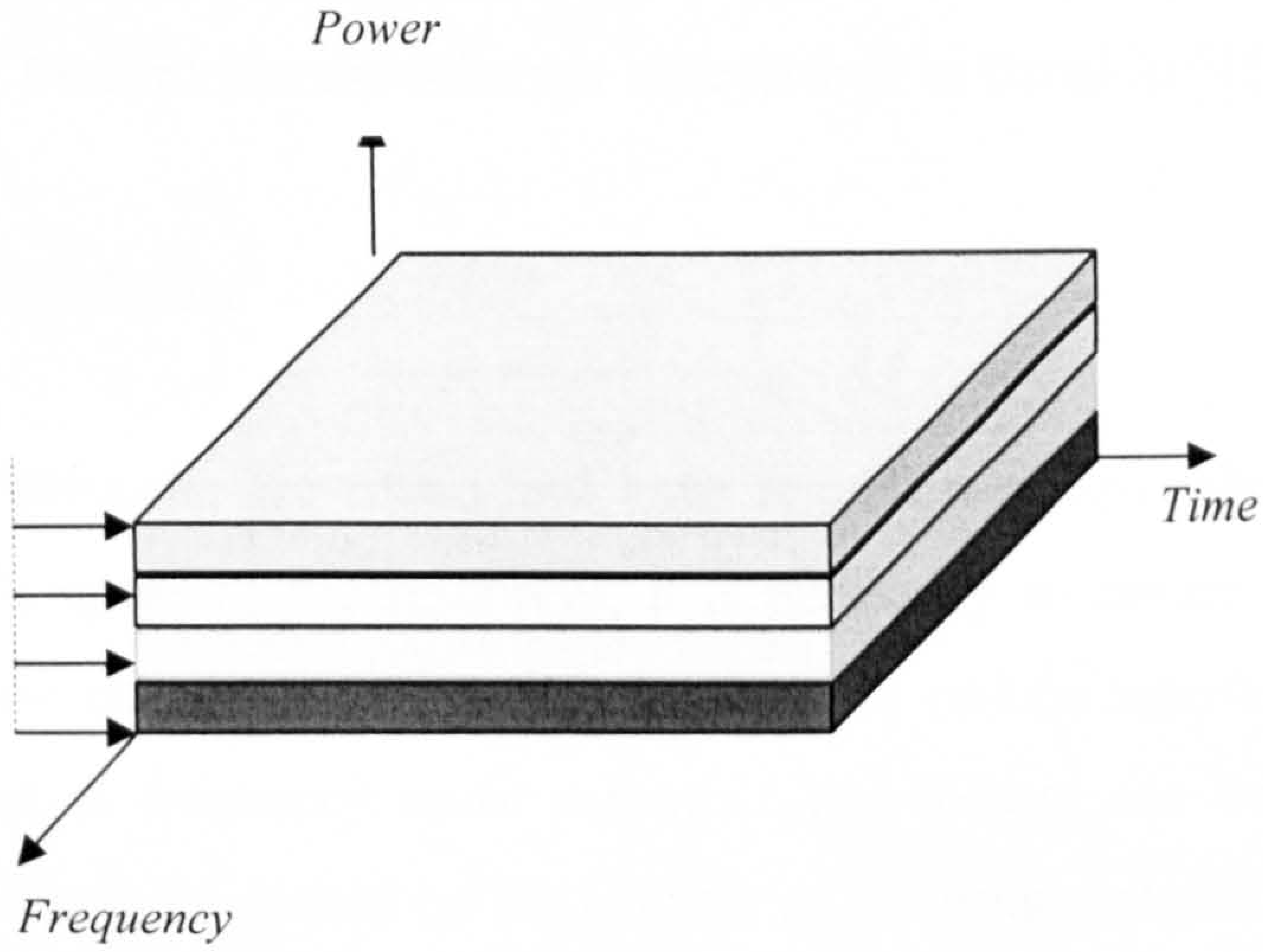


Figure 2.5c: Access Technologies: CDMA

The traditional way to transmit radio signals is to divide the frequency band into several frequency channels which are allocated to each transmission. Each channel can carry a voice conversation or digital data. This approach is known as FDMA. FDMA is a basic technology use in the analog Advanced Mobile Phone Service (AMPS). TDMA is a technology that is used in GSM. The technology divides each frequency channel into time slots and then allocates slots to multiple calls to increase the amount of data that can be carried. As a result, a single frequency can support multiple data channels at the same time [42].

CDMA uses spread spectrum techniques. Instead of assigning a specific frequency or time to each user, each transmission is separated by codes in which every channel uses the full available spectrum. This technology allows numerous signals to occupy a single transmission channel which can optimise the use of the bandwidth available. It is widely used in 2G cellular systems and 3G standards [42].

WCDMA technology has appeared as the most widely approved air interface. It is a high-speed mobile wireless technology which uses the CDMA multiplexing technique. It offers higher data speeds than CDMA. WCDMA can reach speeds of up to 2 Mbps for voice, video, data and image transmission. It consists of a complete set of specifications, a detailed protocol that defines how a mobile phone communicates with the tower, how signals are modulated and how datagram are structured [1]. Multiple access schemes is described in detail in [45].

2.4.2 Frequency Reuse

Theoretically, the cells are round and have regular overlaps where handovers can take place. To optimise the resources, it is necessary to ensure that neighbouring cells do not use the same frequencies in order to reduce interference. There are different model of frequency reuse patterns. The pattern can be repeated and the distance between cells should be far enough to minimise interference [43, 46]. A TDMA system uses a 7 cell reuse pattern and power control is used to deal with the co-channel interference.

Sectorisation concerns with increasing the number of sectors belonging to a site and normally 3 sectors per site are used. It is as a technique to enhance capacity and at the same time, service coverage is generally improved [42]. It is expected that higher sectorisation gives higher capacity. However, an increase of the capacity is not proportional to the number of sectors. This is due to overlap of the sectors and the influence of the environment affects the interference which reduces the capacity of the network. As a result, the gain in capacity is smaller than this expected value [47].

All cells share the same frequency band in CDMA as shown in Figure 2.6. To deal with the interference problem, fast power control and smart antenna are used.

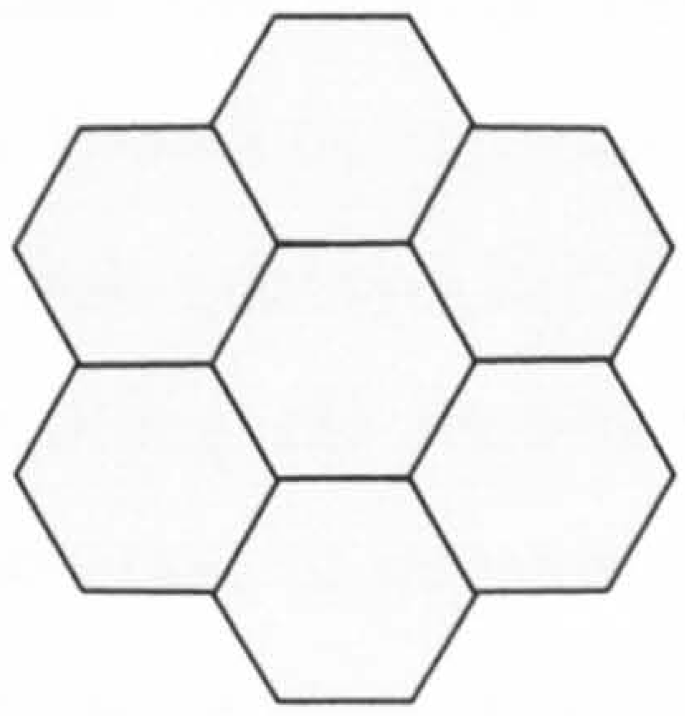


Figure 2.6: Cells Clusters using Same Frequency based on CDMA

2.4.3 Power Control

Power control is one of the most important elements in wireless cellular communication systems. In a wireless cellular system, a mobile station needs to connect to a base station when it wants to communicate with others. The base station is usually equipped with an omni-directional antenna, which can receive or transmit signals from all directions as shown in Figure 2.7.

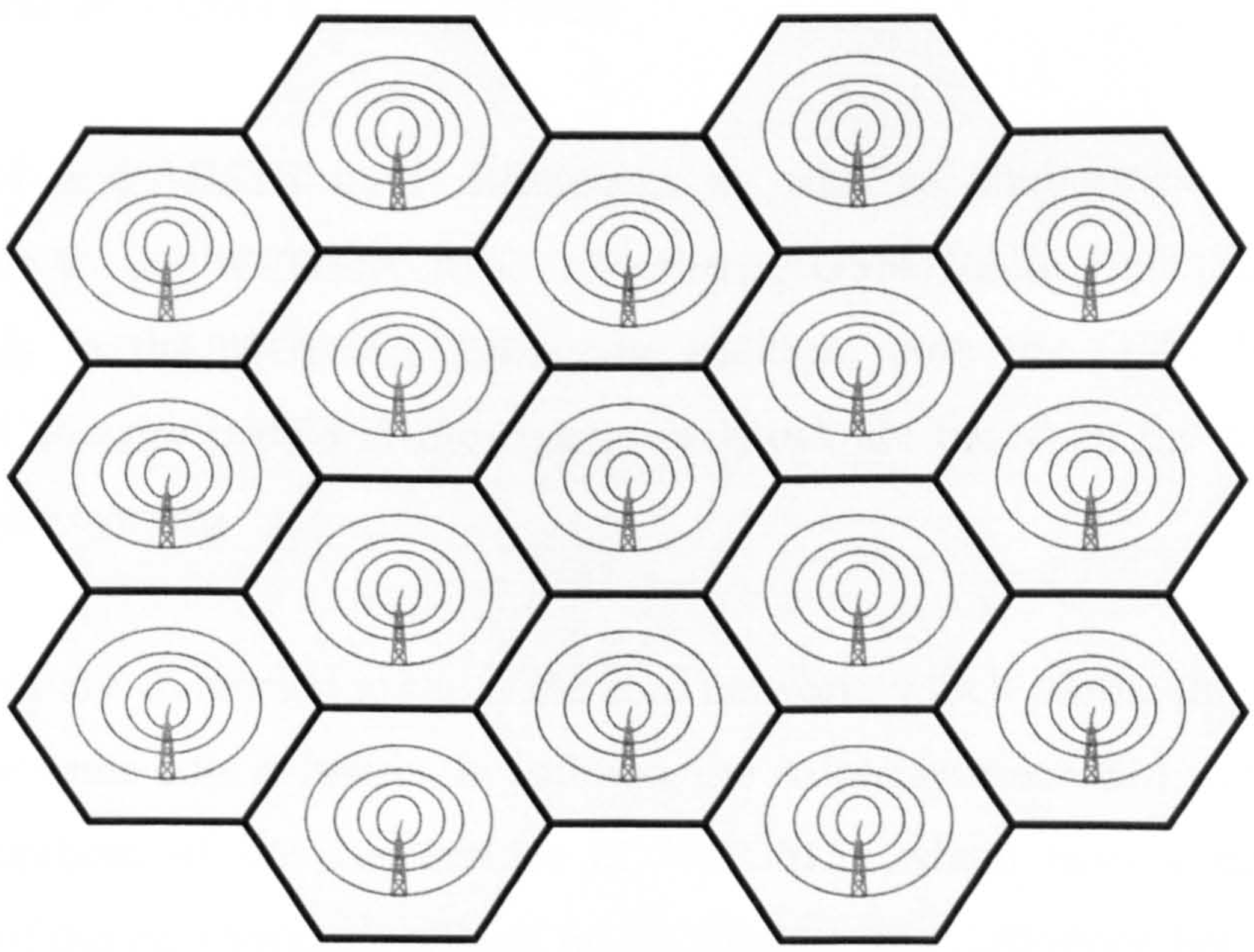


Figure 2.7: A Cellular System with all Base Stations Equipped with Antennas

From the mobile station perspective, the transmitted power level limits the battery life. The battery life can be increased by controlling the mobile transmitted power. Power allocation can play important role in improving the system performance and reducing interference. The power should be maintained at the minimum power level required to achieve the desired link quality [39]. On the other hand, from the cellular system perspective, controlling the transmitted power level can eliminate unnecessary interference. The reduction in interference can be traded for capacity, since it allows more users to share the same network.

Power control in wireless cellular systems can be applied to numerous communication architectures. In FDMA and TDMA, power control can be used to reduce co-channel interference and improve system capacity. In CDMA systems, power control is used to control the "near-far" problem [1]. When two mobile users MS1 and MS2 operate within the same frequency, the user which is far from the base station i.e. MS1, will suffer more path loss compare to the one which is near the base station. As a result, a large part of the cell could be blocked. To maximize the capacity, the received power of all mobile stations should be adjusted.

2.5 GSM and UMTS Comparison

Both GSM and UMTS have differences as well as similarities. GSM BSC corresponds to the WCDMA RNC. Similarly, GSM Radio Base Station (RBS) corresponds to the WCDMA RBS. In addition, both the GSM Base Station Subsystem and the UMTS Radio Access Network are based on the principle of a cellular radio system.

The systems are connected to the GSM core network, which enable the technologies to share the same core network. In addition, the A-interface of GSM was the basis of the development of the Iu-interface of WCDMA, which mainly differs in the inclusion of the new services offered by WCDMA. The difference between the two systems is shown in Figure 2.8 [48].

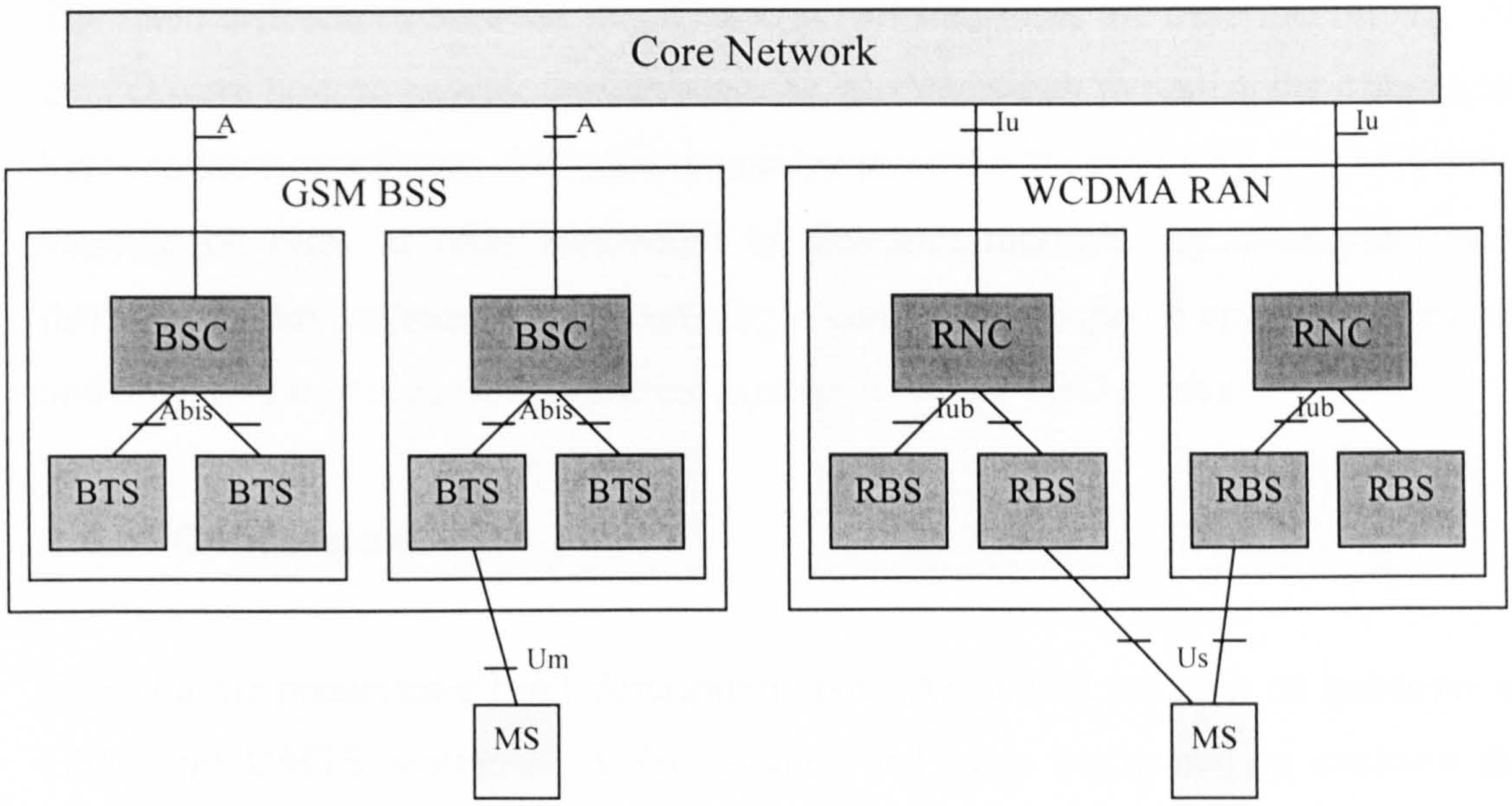


Figure 2.8: Difference between GSM and UMTS

There is lack of interface between the GSM BSCs and an insufficiently specified GSM A-bis-interface to provide multi-vendor operability [48]. The major difference is that the GSM system uses TDMA technology with a lot of radio functionality based on managing the timeslots. The UMTS system uses WCDMA that utilize CDMA scheme. This means that both the hardware and the control functions are different. Examples of WCDMA-specific functions are fast power control and soft handover. The major differences between the systems are shown in Table 2.2.

Table 2.2: Differences between WCDMA and GSM Air Interfaces [1]

	WCDMA	GSM
Carrier spacing	5 MHz	200 kHz
Frequency reuse factor	1	1–18
Power control frequency	1500 Hz	2 Hz or lower
Quality control	Radio resource management algorithms	Network planning (frequency planning)
Frequency diversity	5 MHz bandwidth gives multipath diversity with Rake receiver	Frequency hopping
Packet data	Load-based packet scheduling	Time slot based scheduling with GPRS
Downlink transmit diversity	Supported for improving downlink capacity	Not supported by the standard, but can be applied

The main differences between the 2G and 3G air interfaces are described in [1]. As the 2G were built to provide speech services, it is necessary to realize the differences between the two systems. The new requirements of the 3G are such as high bit rates; variable bit rates to offer bandwidth on demand; multiplexing of services with different quality of requirements on single connection; support asymmetric uplink and downlink traffic as well as coexistence of FDD and TDD modes.

2.6 Conclusion

This chapter presented a brief description about the related research on handover in GSM and UMTS systems. A few studies had been performed to evaluate the performance of handover in order to improve the load sharing and to enhance network capacity. However, there are not many studies on packetized voice in a 3G network. Various measurement parameters were used in making handover decision. It was found that these parameters affect the performance of handover in a 3G system which are likely to affect the performance in an IP based 3G network. A few schemes were implemented in the studies such as soft handover, reservation scheme as well as queuing techniques.

Different types of multiple access technologies that have evolved over the years and their use in the mobile communication generations were also discussed. A description about of GSM and UMTS network architecture was also presented. The architecture of both systems as well as the differences between the two were explained. In general, GSM and UMTS networks are similar as both systems are based on the principle of a cellular network. The major difference is the air interface of the two systems in which UMTS provides higher bit rate compared with GSM. In addition, UMTS is based on the CDMA technology which means the control function is different.

Chapter 3

Mobility and Handover

3.1 Introduction

Mobility is concerned with the ability to send and receive information without any constraints imposed by the user location. Mobility management is one of the issues in mobile communication systems and it is the major functions of a GSM or UMTS network. It involves monitoring the performance of the system including handover delay, termination probability as well as the blocking probability for new call. One of the benefits is that it improves the QoS by avoiding frequent handover request to the adjacent cell near the boundary. The aims are to track location of the subscribers, to allow contact to the subscribers to be established and to record the services that subscriber has paid for.

The main advantage of mobility is that users are allowed to have access to telecommunications services over a wide area as well as maintain communication while moving. Undoubtly, user satisfactions are the most important criteria in order to retain subscribers and it also give an impact to potential customer. To manage a high call density, the cell size is reduced which result in a high number of handover. Therefore, it is essential to improve the success rate of the handover algorithms in order to maintain reliable communication. A new generation of mobile communication systems should be able to satisfy various types of end-to-end services including voice and data integration. Voice service has symmetrical i.e. equal uplink and downlink bandwidth requirement. On the other hand, data transfer such as web service has asymmetrical requirements which may lead to waste of bandwidth.

A handover process consists of three main phases that are measurement phase, decision phase and execution phase. The measurement phase involves measuring signal strength as it may vary resulting from cell environment and user mobility. The overall QoS of the connection will be assessed and compared with the requested QoS at the decision phase before the execution phase is performed. The process is as shown in Figure 3.1.

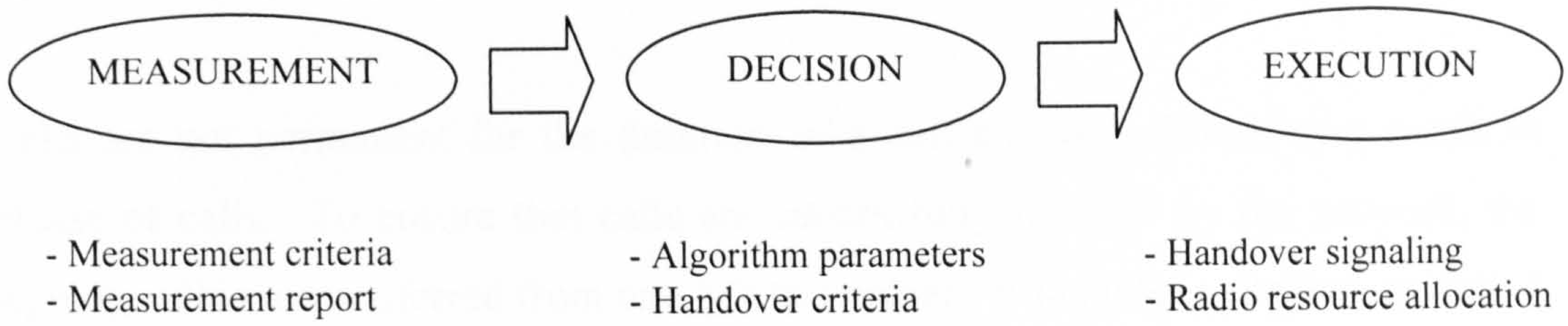


Figure 3.1: Handover Process

To improve performance such as less power emission, less interference and load balancing aspect, the user will be served by another cell. When a mobile station moves between cells, it may require the serving base station to be changed. When the user connectivity is changed from one base station to another, a handover occurs. Handover that occurs between base stations that are using the same type of cellular systems is called the horizontal handover. In contrast, a vertical handover occurs between base stations that are using different wireless interface technologies. Mobility in heterogeneous networks is seen to be one the issues in 4G networks [49]. The difference between horizontal and vertical handover is shown in Figure 3.2.

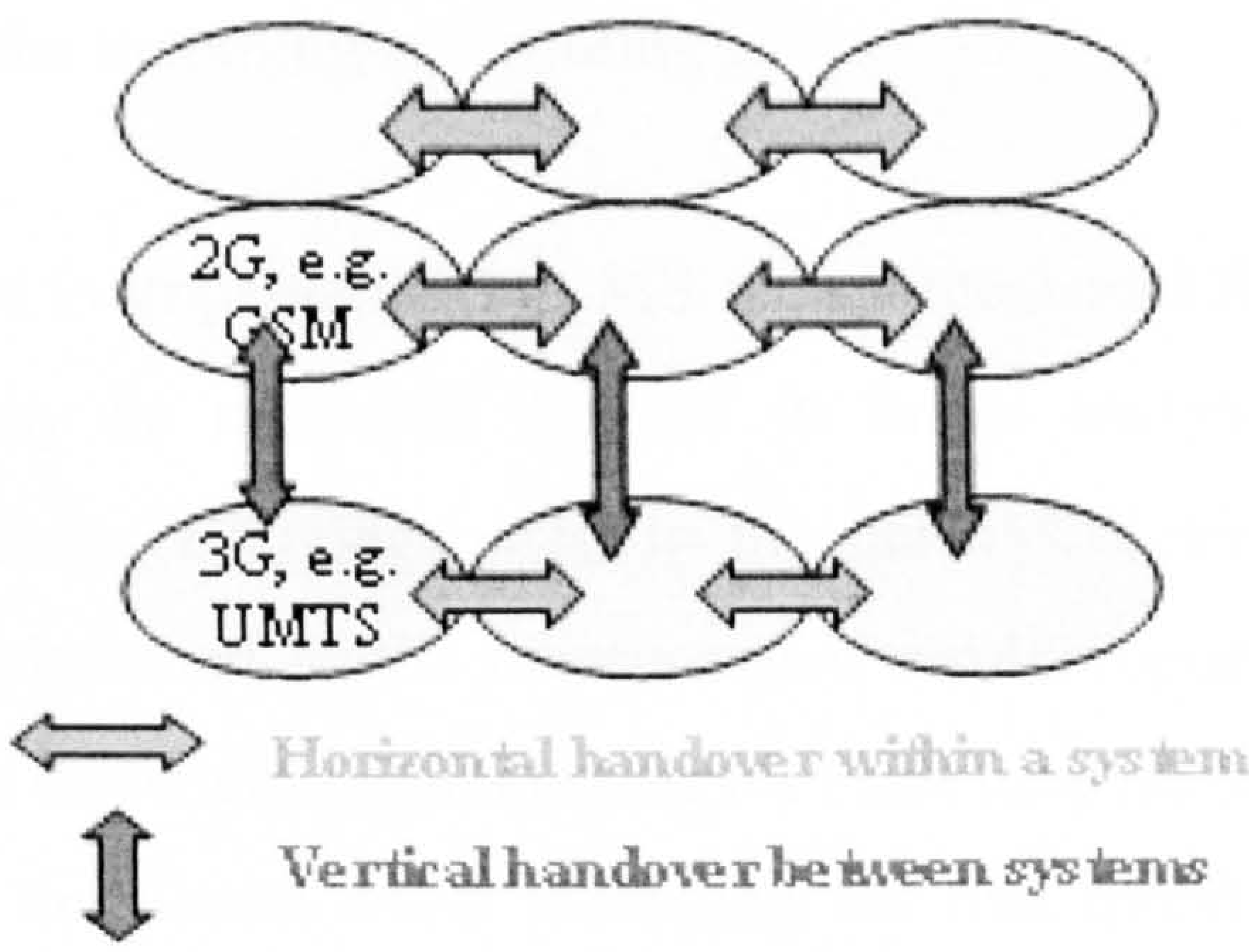


Figure 3.2: Horizontal versus Vertical Handover

The performance of RRM algorithms such as the handover, admission control and congestion control are the major issues to consider as they have consequences on the overall network capacity. Emphasizes have been given in optimizing handover because number of handover relates to dropped calls and network overload. Besides that, ability of a cellular network to perform an efficient handover is crucial as unnecessary and unfavourable handover are expected to happen.

3.2 Handover in GSM

Links are not permanent for the duration of a call and link failure may result in release of calls. To ensure that calls are successfully handled by the network, the ongoing calls are transferred from one cell to another. It includes calls that are failed from loss of coverage or unacceptable interference. Handover can be divided into a few categories. It involves transferring calls between:

- Channels i.e time slots in the same cell
- Cells i.e BTS with the same BSC
- Cells with different BSCs, but the same MSC, and
- Cells with different MSCs.

For traffic load balancing purposed, handovers can be initiated by MS or MSC. The MS forms a list of the best candidates for possible handover which is based on the received signal strength. The information will be updated every second to the BSC and MSC. The MM layer handles the mobility of the users including authentication and security purposes. The system will keep updating information of the current location to complete the incoming call routing [37].

The handover process is implemented in MS which measured the signal level. The interference level may be measured for use in initial assignment and handover assignment. The handover strategy used in the network determines the handover decisions that are made based on the measurement results reported by the MS. It is also based on various threshold parameter set for each base station. Power control is applied to minimize the power while managing the link quality. This will reduce interference to co-channel users.

Received signal quality (RxQual) is used as a criterion in power control and handover process. In addition, cell reselect hysteresis (RxLev) is used to prevent frequent handover and associated location update. The RxLev_UL and RxQual_UL measured values are used as the measurement parameters for base station selection. A base station is selected if the signal is better than the signal in the current location area [37].

In an interference situation, handover can be performed if the link measurements show a low RxQual but high RxLev on the serving base station. For a blockage scenario, handover can be performed if the link measurements show a low RxQual, and a low RxLev with high transmit power level.

A scenario of a successful handover is shown in Figure 3.3. During a handover, a message will be placed and exchanged. The message exchange can be composed of initiation, channel allocation, handover execution and deallocation of resources. A message HND_RQD is placed to inform the MSC that a handover is required. A list of possible cell to be handed over will be listed and the resources will be allocated for a cell in the new BSS. The handover procedure is described in detail in [50].

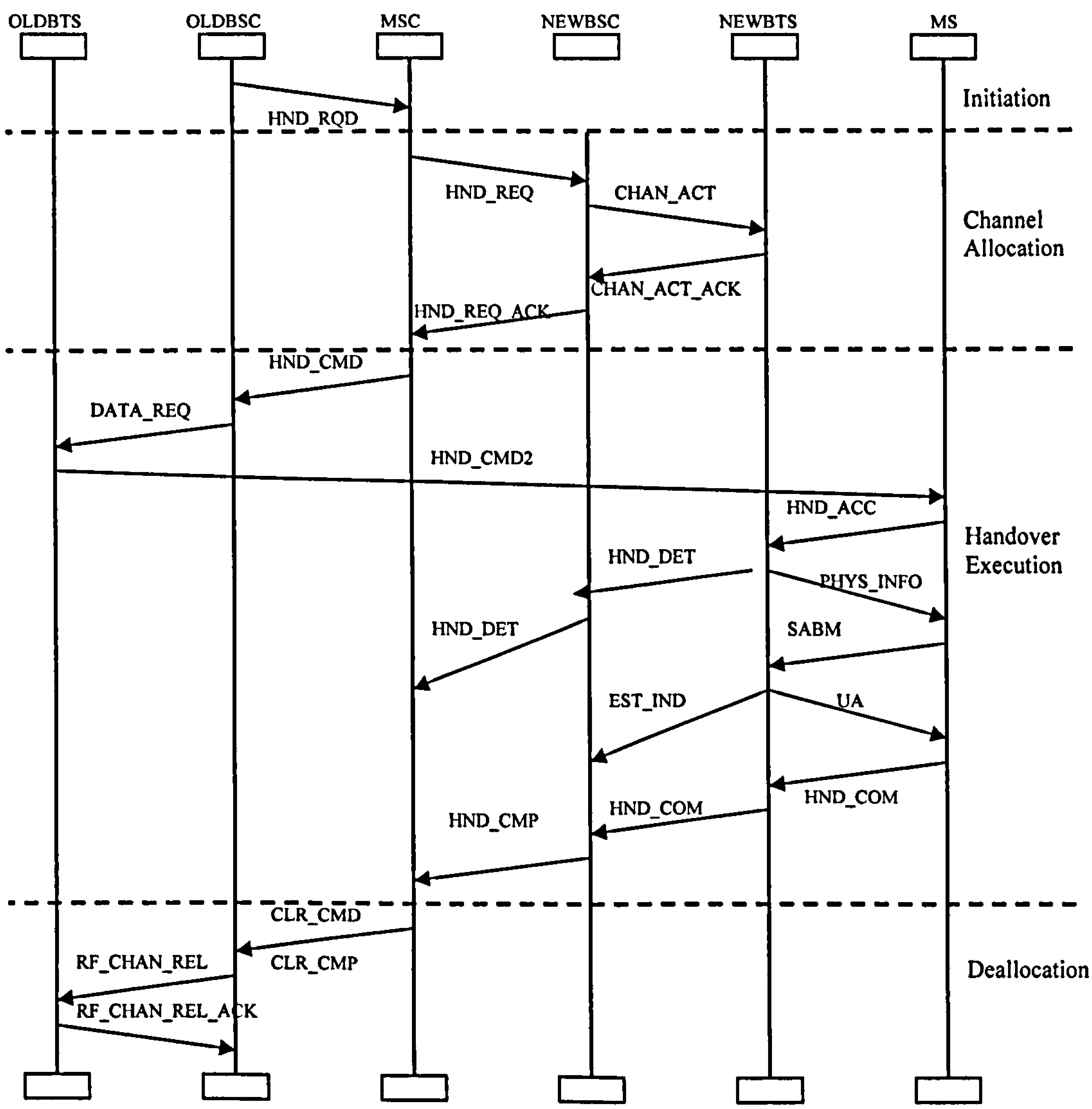


Figure 3.3: A successful handover scenario

3.2.1 Simulation Model

OPNET Modeler is used as the development tool. OPNET is the industry's leading network technology development environment, which is widely used for simulating, testing and evaluating network. It is a network simulator that is used to simulate the network environment as it has various built-in functions. OPNET contains extensive library of node models of different technologies such as Ethernet, ATM and wireless. It also consists of equipment models of specific manufacturers such as Cisco, Lucent, NEC, Nortel and 3 COM [51]. A brief description of OPNET network modeller is described in the Appendix.

OPNET does not provide a model for GSM. Hence, a contributed model was used to study how GSM works. The model was developed based on the previous version of OPNET and unfortunately, it did not work on the later version. A lot of time and effort was spent in debugging or fixing all the errors as well as trying to understanding the implementation of the whole system.

A few types of handover occur when a call is transferred from BTS1 to BTS5. It involves transferring calls to another BTS with the same BSC (i.e. process A, handover from BTS1 to BTS2), handover to BTS with different BSC (i.e. process B, BSC1 to BSC2) and handover with different MSC (i.e. process C, MSC1 to MSC2). The block diagram for the handover is as shown in Figure 3.4.

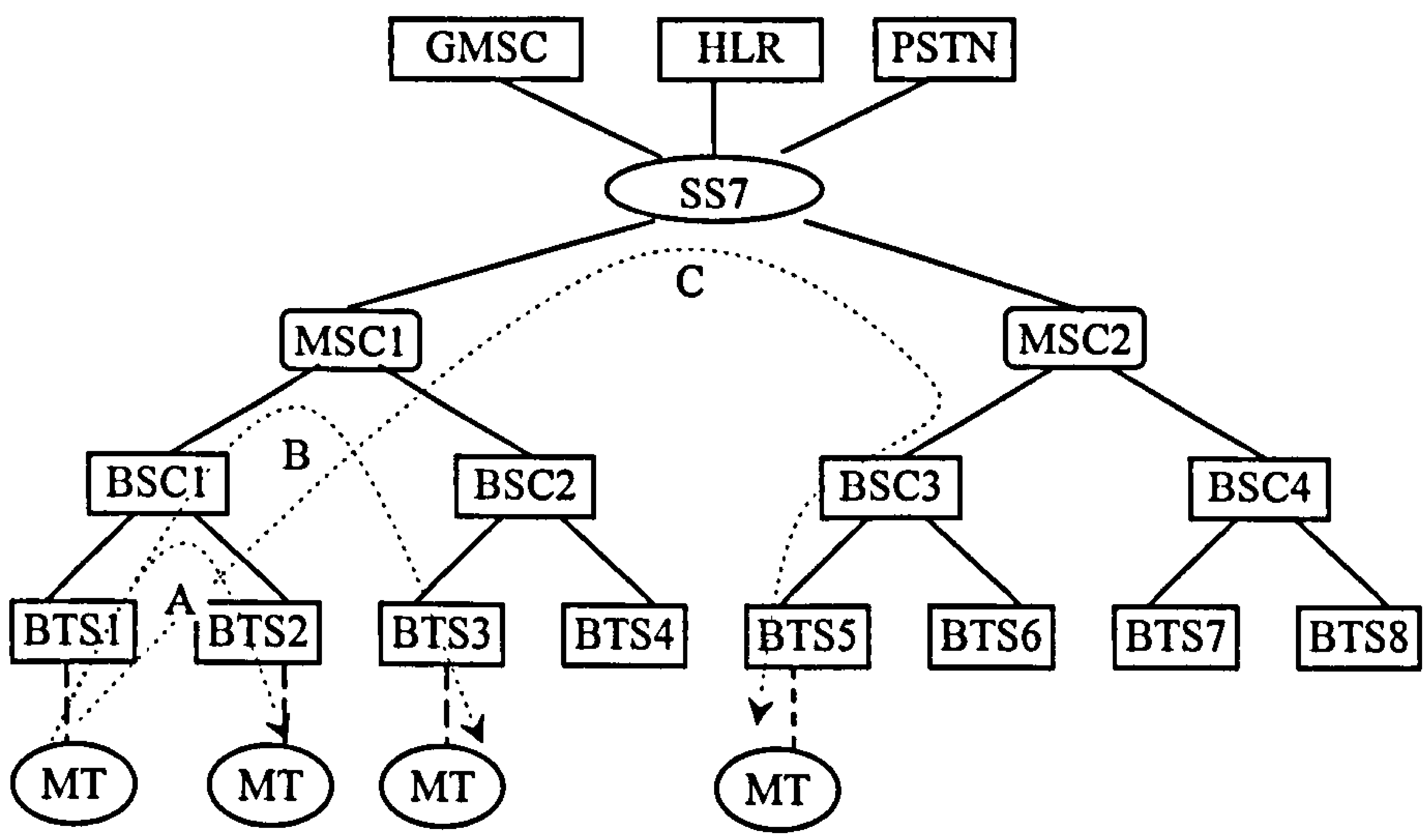


Figure 3.4: Block Diagram for GSM Handover

In the simulations it is assumed that the users are traveling along the trajectory at a constant speed that causes a small number of handovers to occur. The GSM model consists of 8 base stations with a few mobile users as shown in Figure 3.5. The path is located within the area formed by the base station. The BSC is expected to handover the users between different base stations along this path based on the measurement reports sent by the equipment. The power received by the mobile station from the base station and number of handover occurred were analysed.

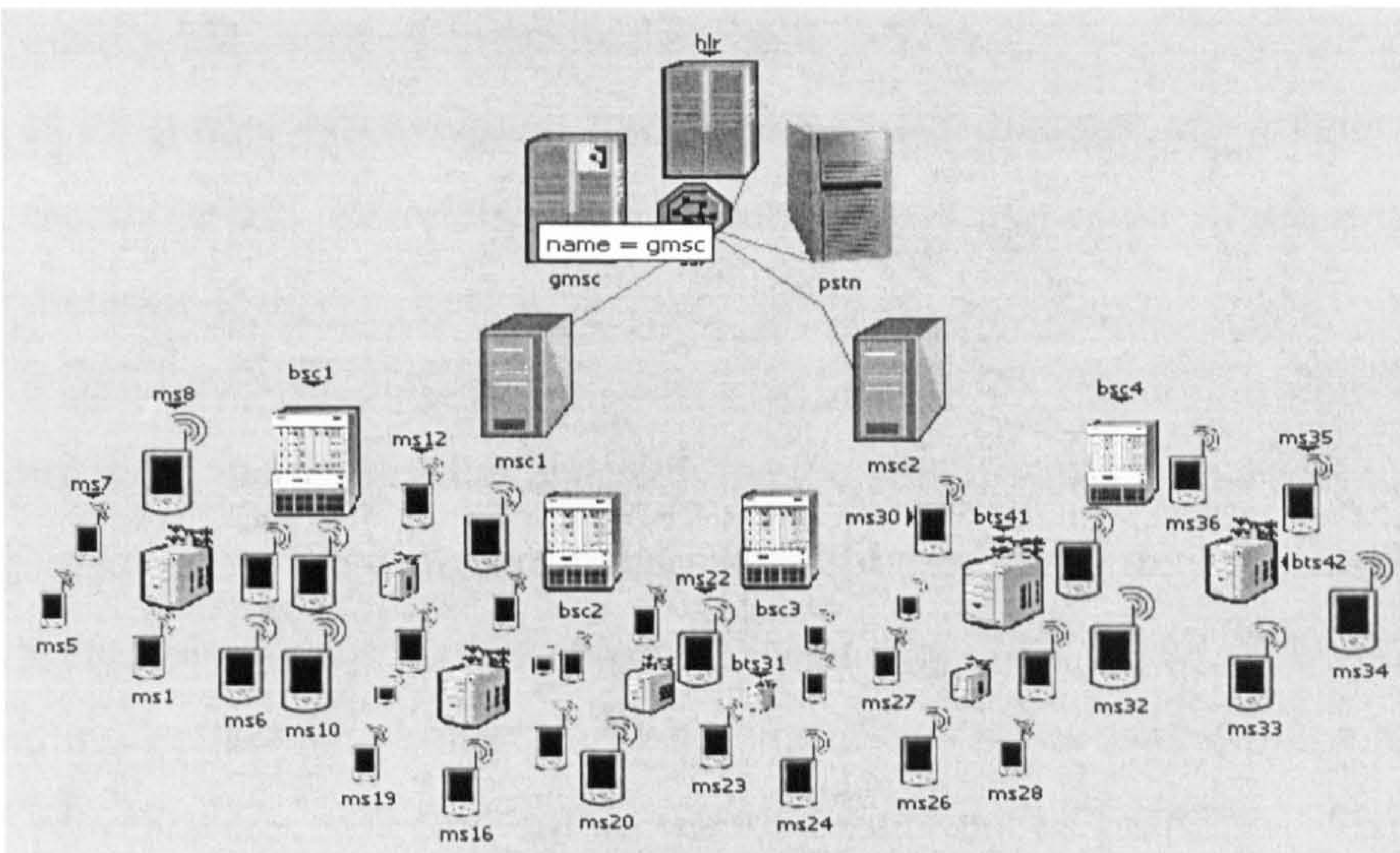


Figure 3.5: Screenshot of GSM Handover Simulation

The received signal strength and quality of the serving base station as well as the surrounding base stations are measured continuously. In addition, the base station measures the signal strength of the mobile station it is serving. 32 values of the sample measures are stored in the base station. P1 to P4 and N1 to N4 are the number of samples used in the threshold comparison processes for power control and an appropriate action will be performed. P1, P2, N1 and N2 are the parameters used to compare the RXLEV measurement whereas P3, P4, N3 and N4 are the parameters used for RXQUAL measurement.

Similarly, P5 to P8 and N5 to N8 are parameters used control the handover algorithm. A requirement for a handover is indicated by the value of Pi out of Ni. The average value has to be below or above the respective threshold for a handover to occur. The base station will decide if a handover is required and the decision will be based on the following conditions [52]:

Rule 1: If P5 out of N5 averaged values of the RXLEV are lower than the threshold level either on the downlink (DL) or on the uplink (UL), a handover is requested.

Rule 2: If P6 out of N6 averaged values of the RXQUAL are smaller than the threshold either on the DL or UL, a handover is requested.

Rule 3: If P7 out of N7 averaged values of the RXLEV on the UL are greater than the threshold, an intra cell handover is requested when in addition the signal quality has dropped under its threshold.

Rule 4: If P8 out of N8 averaged values of MS-BS distance are greater than the threshold MS_RANGE_MAX, a handover is requested which is caused by distance.

In the simulation, the reselection criterion for the signal strength is equal to -126dB. If the signal is below the criterion, handover will be performed. A handover should happen preferably when the distance exceed the MS_RANGE_MAX of the corresponding cells. The simulation was run a few times and the results for the received power were collected using built in function in OPNET. The average results were used to plot the graph. The assumptions that were used in the GSM system simulation is shown in Table 3.1.

Table 3.1: Parameters for GSM Handover

Hreqave(average number of measurement)	32
L_RXLEV_DL_H(Downlink signal level)	-70.5 dBm
L_RXLEV_UL_H(Uplink signal level)	-74.5 dBm
MS_RANGE_MAX(Maximum distance between MS and BTS)	17.20 km
RXLEV_MIN(Minimum received power level)	-73 dBm
Beacon frequency	935 dBm
TX channel power	100 dBm
Call time	30 sec
First call time	50 sec
Inter call time	20 sec

3.2.2 Result Analysis

The results for specification in Table 3.1 are shown in Figure 3.6 and Figure 3.7. Figure 3.6 shows the power received by all the users as they move along the trajectory. The power is measured in dBm, i.e. the power level in decibel is compared with 1 milliwatt. When the power received fall below the signal strength criteria, handover occurs. In the simulation, MS2, MS3 and MS4 perform handover because when the users move around a few base stations, the received power fluctuates around the threshold value. The received power for MS0 and MS1 seem to be constant because these users move between two different base stations. Therefore, handover is less likely to occur.

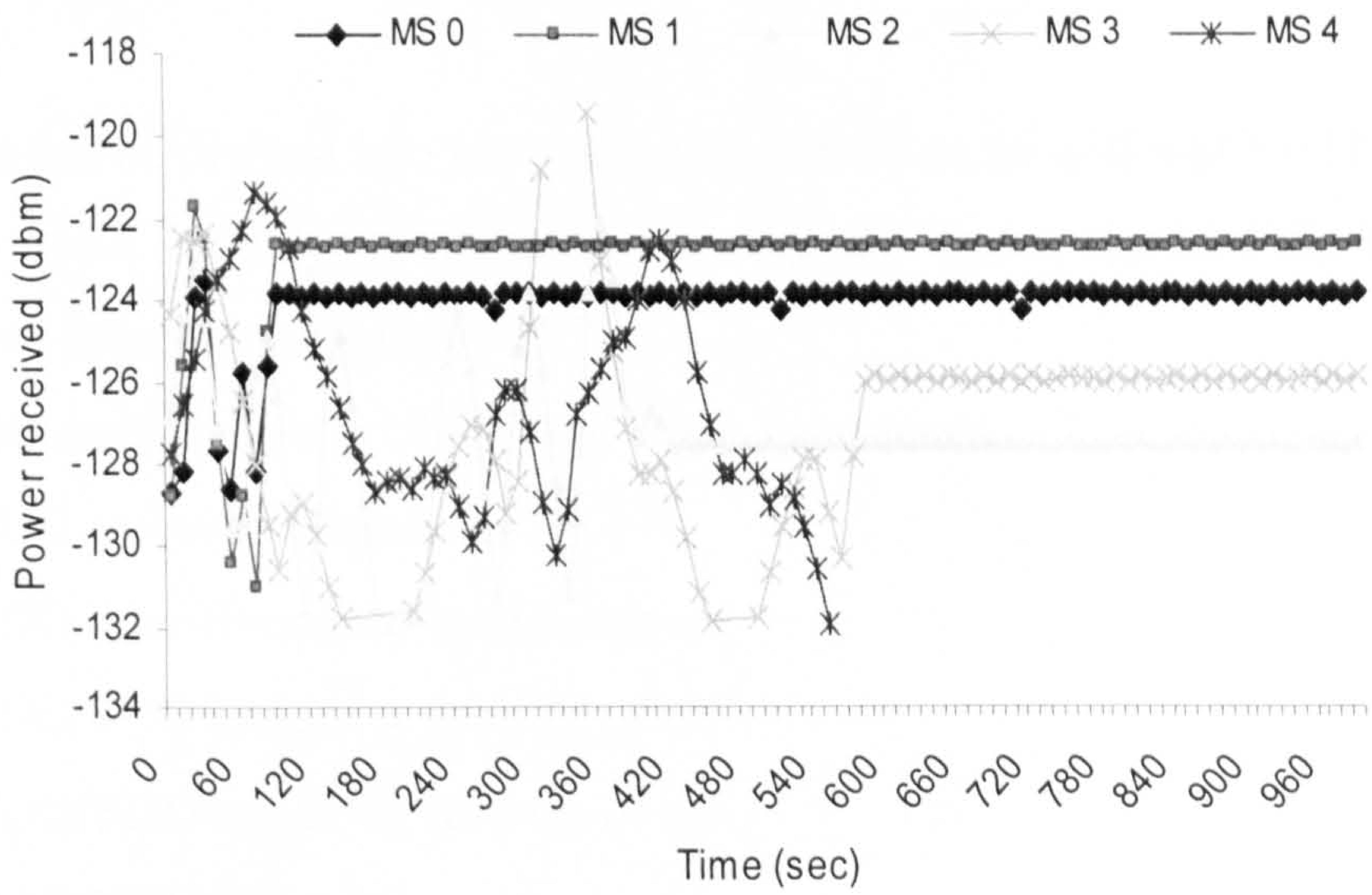


Figure 3.6: Power Receive against Time

The effect of hysteresis on the number of handover for all the mobile stations is shown in Figure 3.7. When the hysteresis is more than 3, the number of handover for MS0 and MS3 are reduced. On the other hand, the number of handover for MS1 is increased. In addition, MS4 will experience a handover. However, it is difficult to predict the handover due to the mobile trajectory and assumptions that have been made.

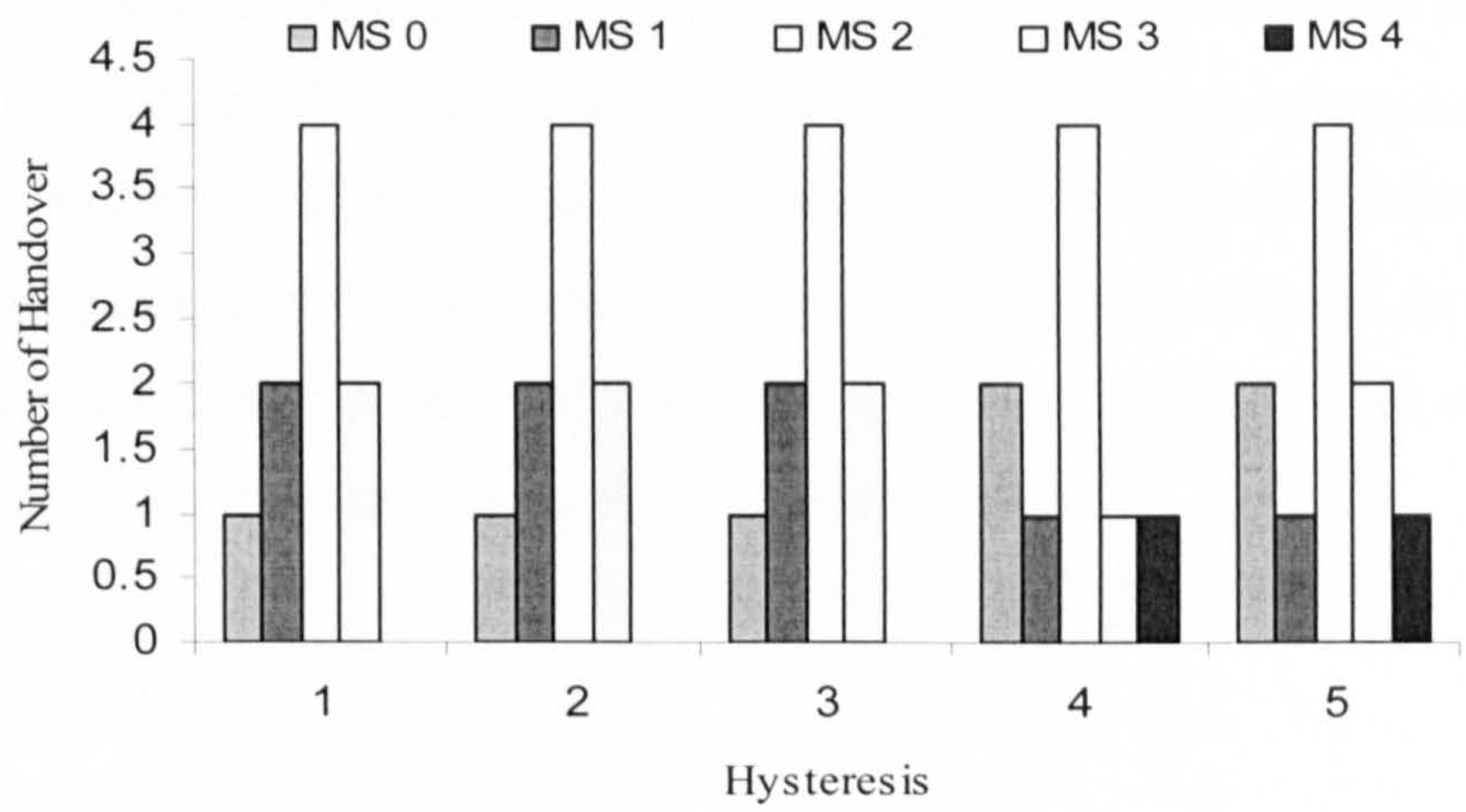


Figure 3.7: Number of Handover against Hysteresis

3.3 Types of UMTS Handover

Handover can be divided into two main categories that are soft and hard handover. It needs to occur quickly to maintain the required QoS. In UMTS, handover is categorised into following types: [53]

- Handover between UMTS and other 3G systems
- FDD soft/softer handover
- FDD inter-frequency hard handover
- FDD/TDD handover (change of cell)
- TDD/FDD handover (change of cell)
- TDD/TDD handover
- Handover 3G - 2G (e.g. handover to GSM)
- Handover 2G - 3G (e.g. handover from GSM)

3.3.1 Soft handover

Soft handover is described in [1, 3, 8, 42]. It is a situation where the links are added and removed where the mobile station will always keeps at least one link. There are several links active at the same time. Generally, handover occurs when cells operated on the same frequency are changed. During a soft handover process, there are at least two links activated for a certain period of time.

A mobile station is in the overlapping cell coverage area of two sectors belonging to different base station. Two power loops per connection are active, one for each base station. According to [1], soft handover occurs in about 20-40% of connections. The basic design parameters for this soft handover algorithm are add and drop margins, link preservation time, and size of the active set where the active set contains the cells that are currently connected to the mobile station.

3.3.2 Softer handover

Softer handover is part of soft handover situation. However for this situation, the links that are added and removed belong to the same Node B. In a softer handover case, the base station receives two separated signals. Due to reflections on buildings or other barriers, the signal sent from the mobile stations reaches the base station from two different sectors. In general, a mobile is in the overlapping cell coverage area of two adjacent sectors of a base station. According to [1], only one power control loop per connection is active during softer handover and softer handover occurs in 5-15% of the connections.

3.3.3 Hard handover

In a hard handover, the mobile uses only one channel at a time where the existing link is released before a new connection is made. It can be seamless i.e. significant, or insignificant to the users. Basically, hard handover is a handover that requires a change of the carrier frequency. Intersystem handovers can be seen as a type of hard handovers. Hard handover is chosen when soft or softer handover is impossible [1].

The main problem about hard handover in GSM system is the high blocking probabilities that sometimes experienced by users entering a new cell. Although this probability can be reduced by giving priority to handover users over new users, this results in a less efficient use of the capacity of the cellular systems or higher blocking probabilities for new users. For some reasons, hard handover exist together with soft and softer handover algorithms to increase overall system performance.

In general, hard handovers are only used for coverage and load reasons, whereas the main aims of soft and softer handover are to support mobility. The hard handover procedure for inter-frequency measurement quantities is described in [54].

3.4 Soft Handover Algorithm

UTRAN supports different types and of handovers procedures such as Intra-system handover, Inter-system handover, hand handover and soft handover as discussed in [47]. The soft handover is composed of acquiring and processing measurement, as well as executing the handover algorithm. The carrier energy to noise ratio (Ec/N0) of the Common Pilot Channel (CPICH) is necessary for making the handover decisions. The MS decides to add, delete or replace a Node B in the active cell based on the Ec/N0 measurement.

The parameters that could be used for the initiation of a handover process are as follows: [55].

- Uplink quality, e.g. Bit Error Rate (BER);
- Uplink signal measurements, e.g. Received Signal Code Power(RSCP) for TDD;
- Downlink quality, e.g. Transport channel Block Error Rate (BLER);
- Downlink signal measurements, e.g. CPICH RCSP, CPICH Ec/No, Path loss;
- Distance and Traffic load;

Figure 3.8 shows a sample pilot signal strength of different cell evolves in time. It illustrates the soft handover algorithm with the hysteresis and the time to trigger mechanism. A mobile user is connected to a base station in cell_1 which has the strongest pilot signal. The user starts to move in the direction where the received signal for the CPICH of cell_2 increases. At Event 1A, a new radio link connection to a BS in cell_2 will be established once the receive signal strengths for CPICH_1 and CPICH_2 differ by a maximum amount of the handover margin during the period T.

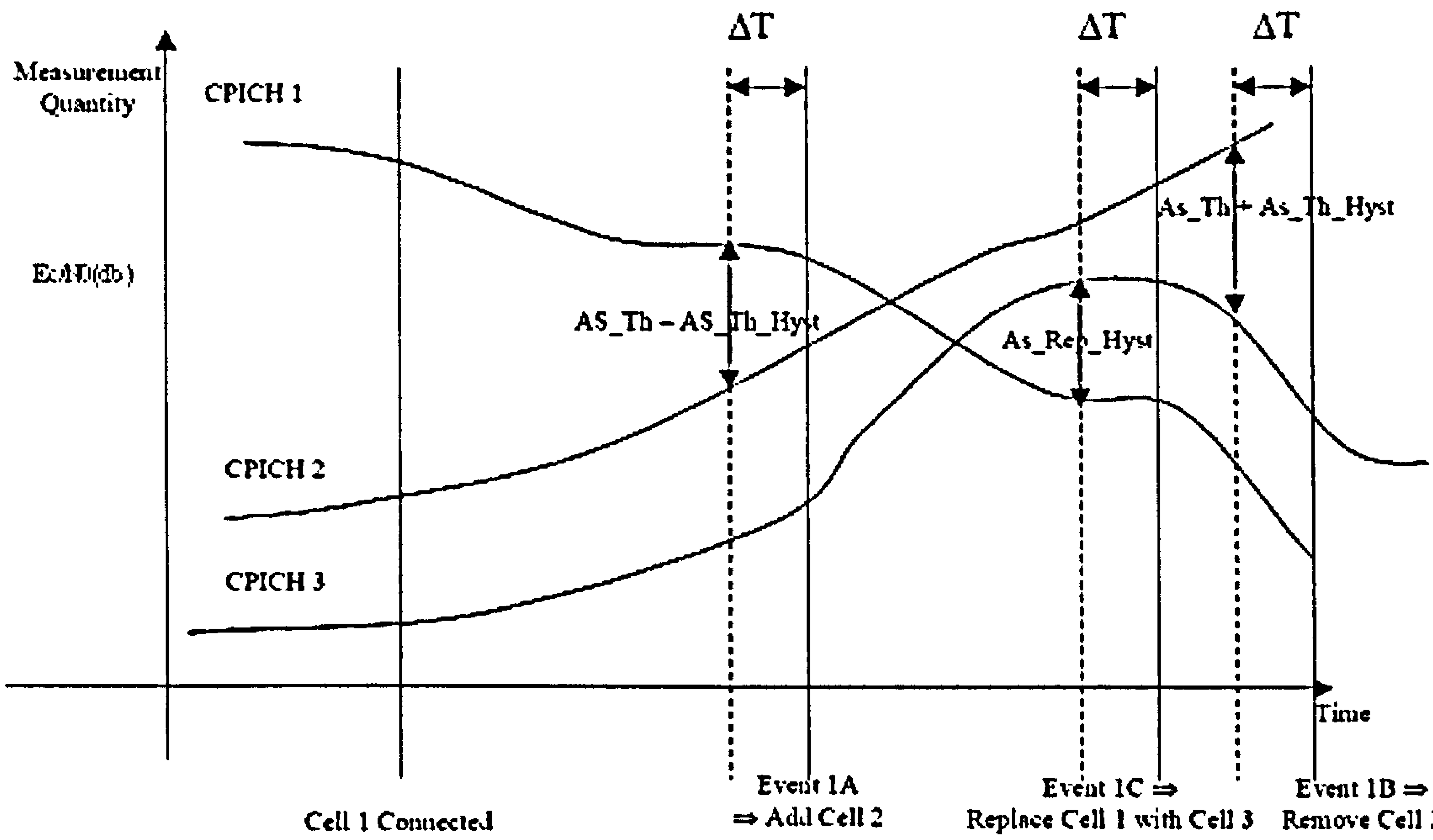


Figure 3.8: Soft Handover Algorithm [55]

Signal from cell_1 and cell_2 will be in the active set. A link to cell_1 will be removed when the received signal of CPICH_1 is smaller by a certain amount than that of CPICH_2 or if the received signal from another base station is better. At Event 1C, a connection to cell_1 is replaced with cell_3 and signal from cell_1 will not be in the active set. A link is removed only if the average receipt level remains below the specified level for a certain time period. At Event 1B, a connection to cell_3 is removed as the signal strength is below the specified level.

The actual algorithm described in Figure 3.8 is as follows.

- Radio Link Addition: Event 1A
  - If Signal_Ec/N0 is greater than (Best_Signal - As_Th + As_Th_Hyst) for a period of ΔT and the Active Set is not full, add Best cell outside the Active Set in the Active Set.
- Radio Link Removal: Event 1B
  - If Signal_Ec/N0 is below (Best_Signal - As_Th - As_Th_Hyst) for a period of ΔT, remove Worst cell in the Active Set.

- Combined Radio Link Addition and Removal: Event 1C
    - If Active Set is full and Best_Candidate_Signal is greater than (Worst_Old_Signal + As_Rep_Hyst) for a period of ΔT, add Best cell outside Active Set and Remove Worst cell in the Active Set.

Where:

- Best_Signal: the best measured cell present in the Active Set;
- Worst_Old_Signal: the worst measured cell present in the Active Set;
- Best_Candidate_Signal: the best measured cell present in the monitored set.
- Signal_Ec/N0: the measured and filtered quantity.

A flow-chart of the soft handover algorithm is shown in Figure 3.9.

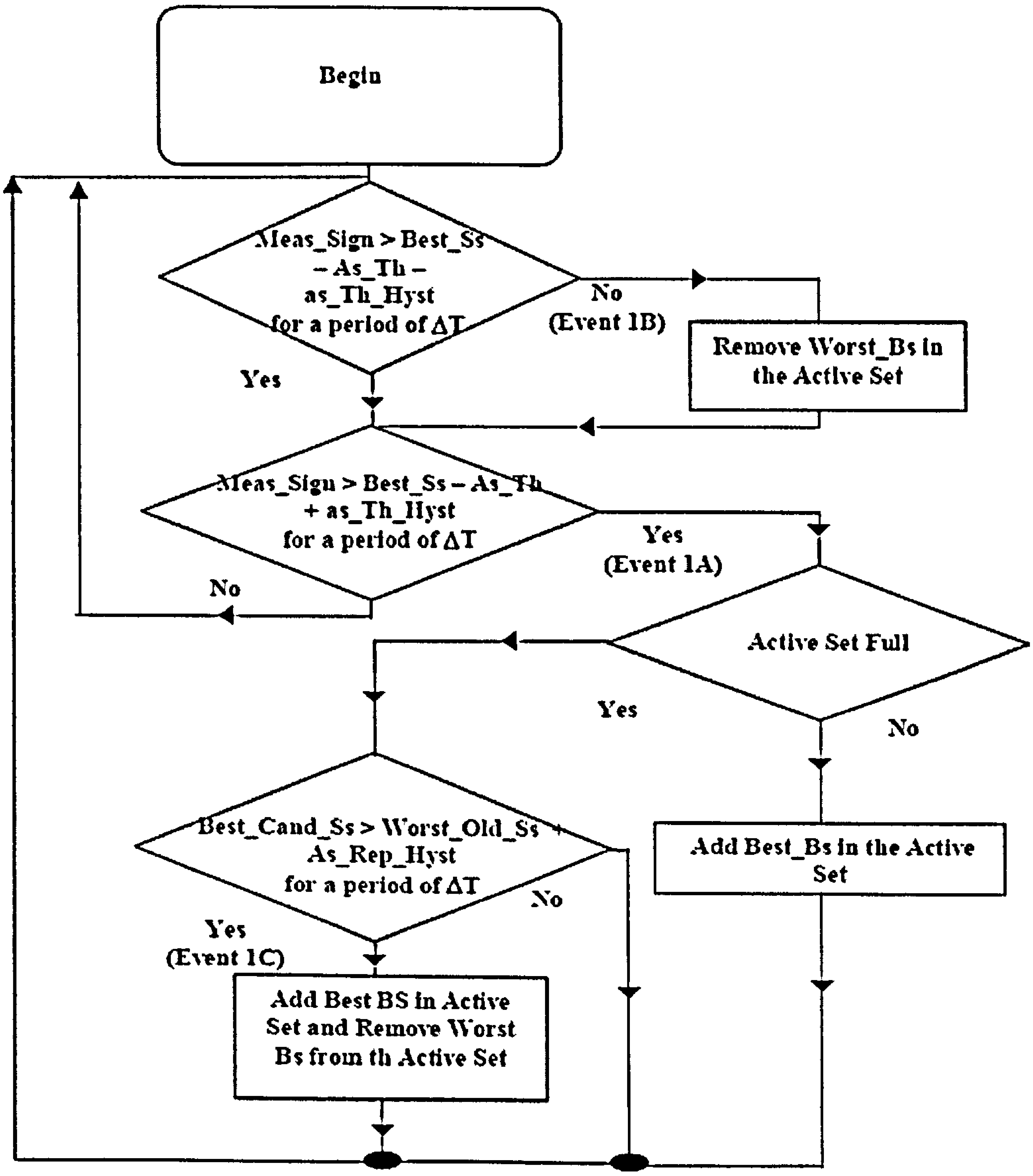


Figure 3.9: A Flow-chart of a Soft Handover Algorithm [55]

The following parameters are used in the algorithm [55]:

- AS_Th: Threshold for macro diversity (reporting range) for soft handover;
- AS_Th_Hyst: Hysteresis for the above threshold;(addition, removal);
- AS_Th - AS_Th_Hyst: Window Add;
- AS_Th + AS_Th_Hyst: Window Drop;
- AS_Rep_Hyst: Replacement Hysteresis;
- ΔT: Time to Trigger;
- AS_Max_Size: Maximum size of Active Set.

Figure 3.10(a-c) shows the block diagrams for process of relocating MS connection where the MS moves to a Node B which belongs to a different RNC [56]. The MS is in a soft handover situation where Node Bs belong to the same RNC as illustrated in Figure 3.10(a). The signals which are combined in the RNC will be sent to MSC.

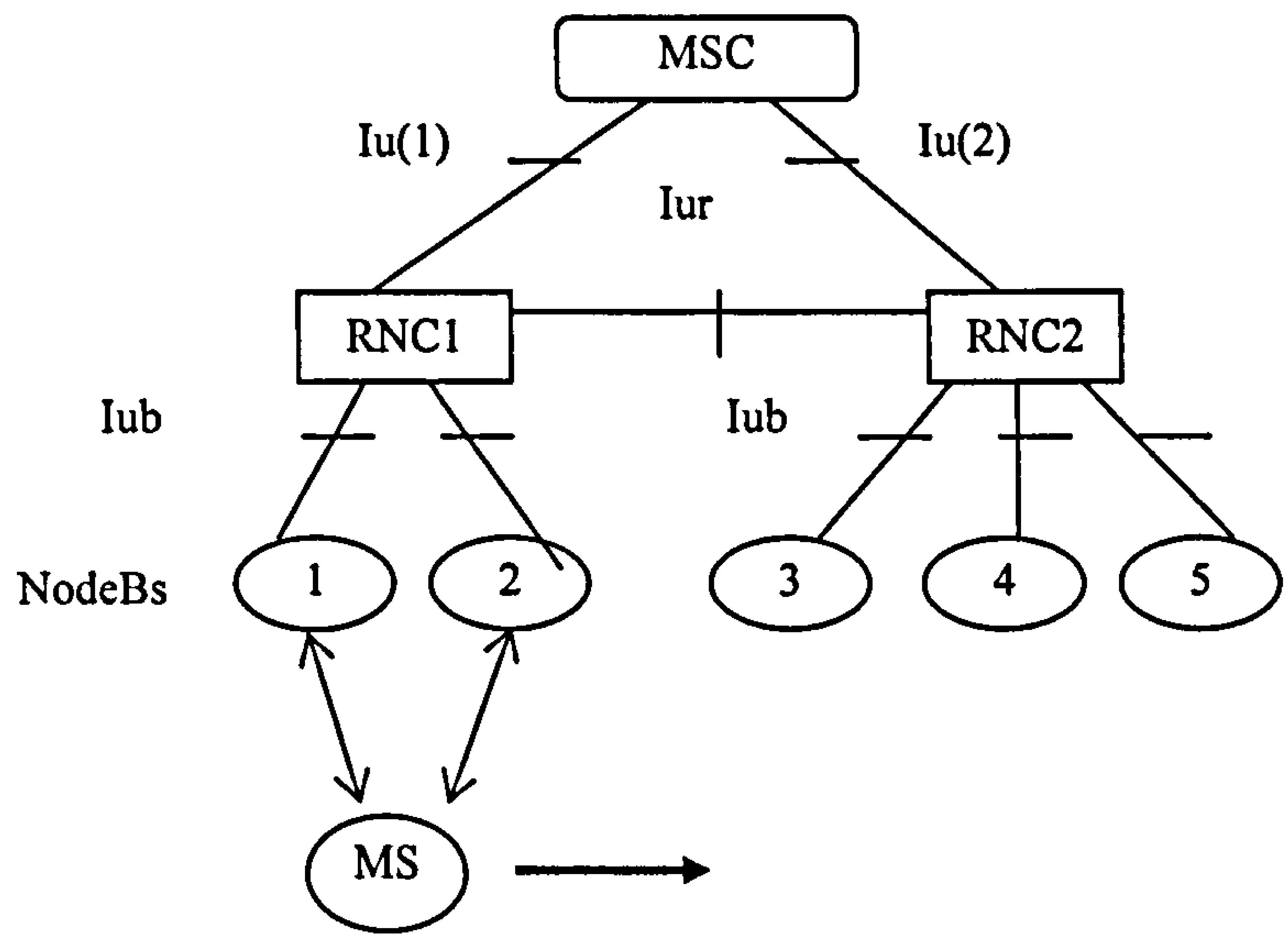


Figure 3.10(a): Process of Relocating MS during Handover

On the other hand, Figure 3.10(b) shows that the Node Bs belong to a different RNC. The signals will be transmitted to first RNC (RNC1) which is known as serving RNC (SRNC). The SRNC combines the signals before sending them to the MSC. The second RNC (RNC2) is known as the drift RNC (DRNC). The signals from cells 3 and 4 will be combined in the DRNC.

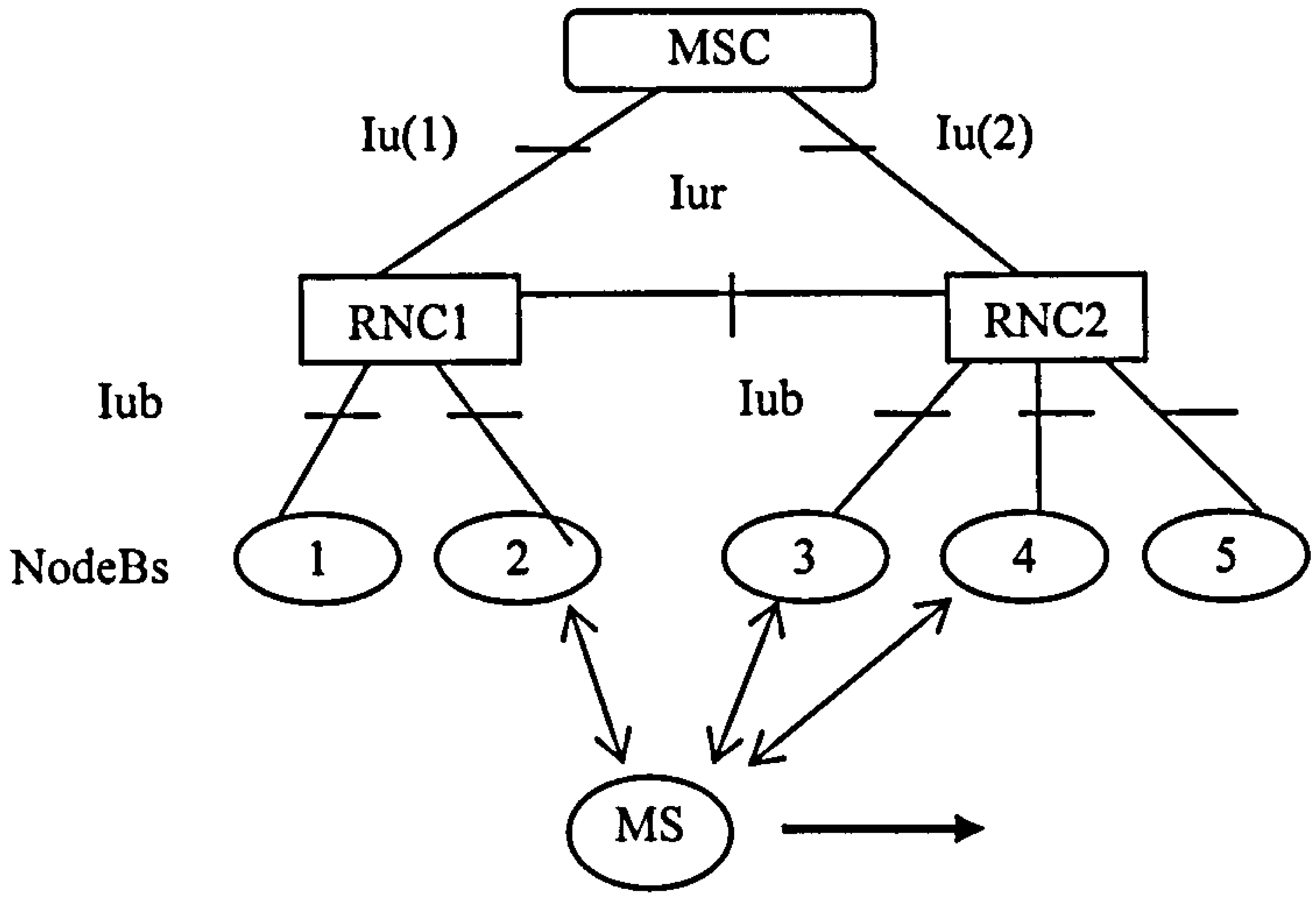


Figure 3.10(b): Process of Relocating MS during Handover

When the MS is completely controlled by RNC2, relocation process is needed where SRNC status will be changed from RNC1 to RNC2 as shown in Figure 3.10(c).

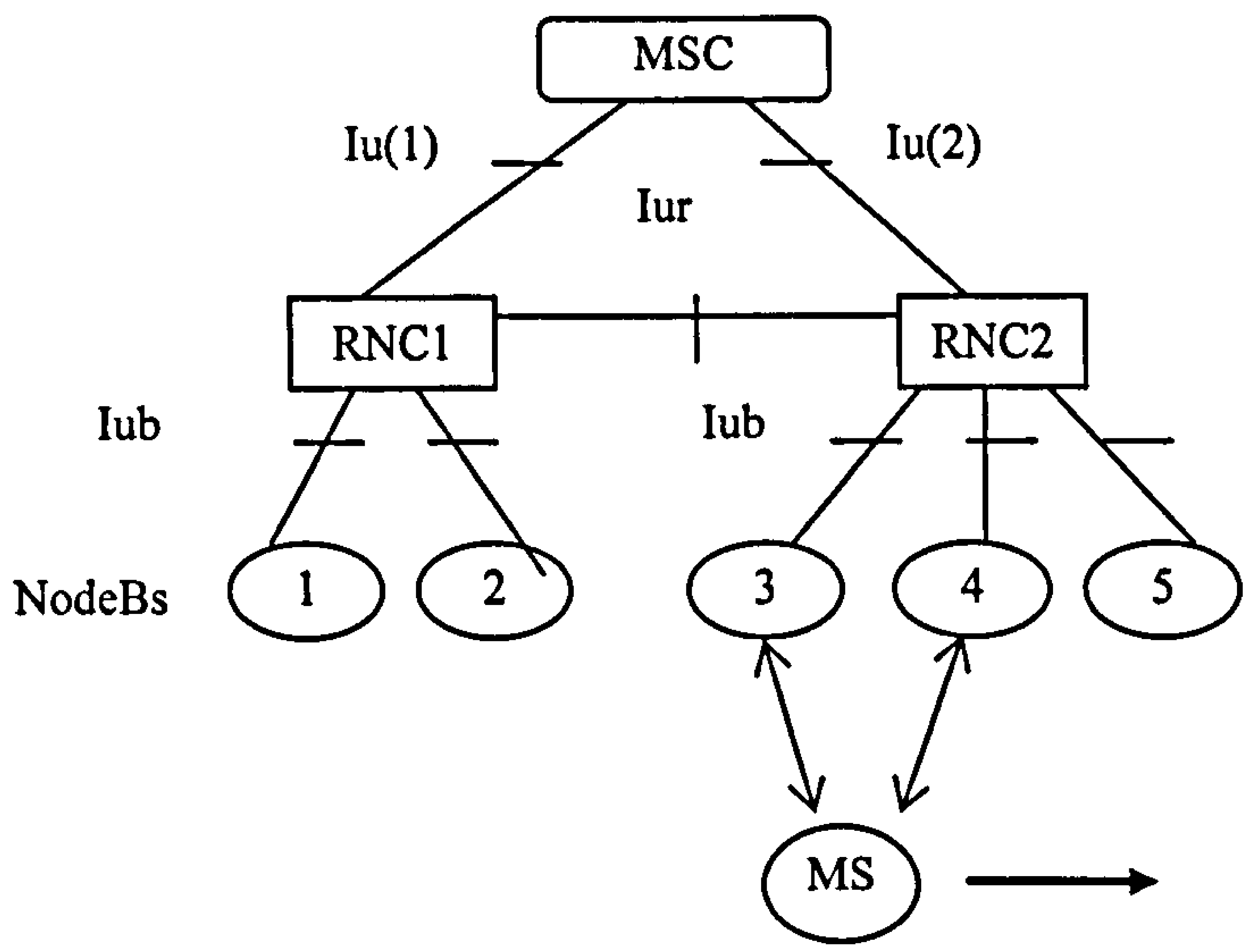


Figure 3.10(c): Process of Relocating MS during Handover

A number of organisations are involved in the ongoing standardizing of 3G and each of them pursuing their own goals and interests. Consequently, the standardization becomes extremely complex. Sometimes a political process is also involved where technical decisions are made in an environment which is full of different and contradictory interests [57].

3.5 Modelling of UMTS Handover

Various measurement parameters are used in making handover decision. There are a few parameters that affect performance in handover, which are likely to affect the handover in the 3G system such as threshold value for macro diversity, hysteresis, time to trigger, number of active sets, as well as distance between the user equipment and base station. The handover parameter was determined in order to achieve an optimal outcome. Delay under various conditions was calculated and simulation results were used to determine various measurements of parameters in making handover decision. In evaluating the performance of handover algorithms, the investigation was focused on the impact of soft handover and distance on system performance.

The handover system was modelled to allow various parameters of the system to be tested. The system parameters, including the number of base stations, mobile stations and station types, were changed to compare results. To analyse the results of various measurement quantities that were used in handover design, OPNET was chosen as the simulation tool. The UMTS system model was used to study the types of handover in 3G network. A model of the system was constructed where a few users moving around and connecting from one base station to another in the intra-system wireless environment. A simple block diagram for handover scenario in UMTS is shown in Figure 3.11.

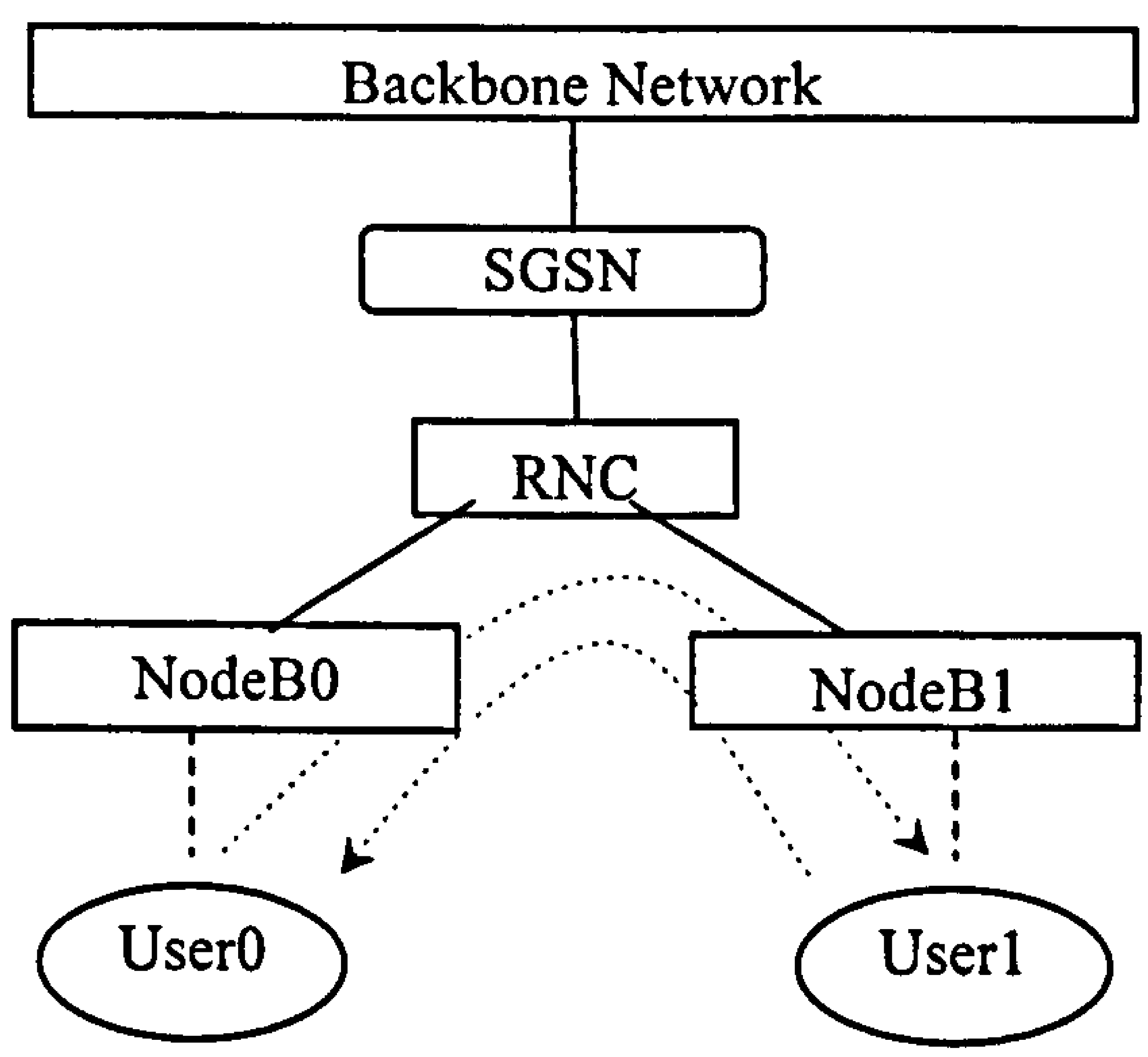


Figure 3.11: Block Diagram for UMTS Handover

The UMTS network consists of Node Bs, RNC and 20 mobile users is shown in Figure 3.12. The users are travelling along a zigzag path at a constant speed that will cause a number of handover to occur. The trajectories for the users have been specified. The path is located within the area formed by the Node Bs. The mobile users, "User0" and “User1” for instance, are initially connected to Node B0 and Node B1, respectively. The mobile users get closer to certain Node Bs at certain times. The RNC is expected to handover "User0" and "User1" between different Node Bs along this path based on the measurement reports sent by equipment. The Node Bs are connected to the RNC through an Asynchronous Transfer Mode (ATM) link. The RNC is connected to the core network (SGSN-GGSN) using Point to Point Protocol (PPP) links.

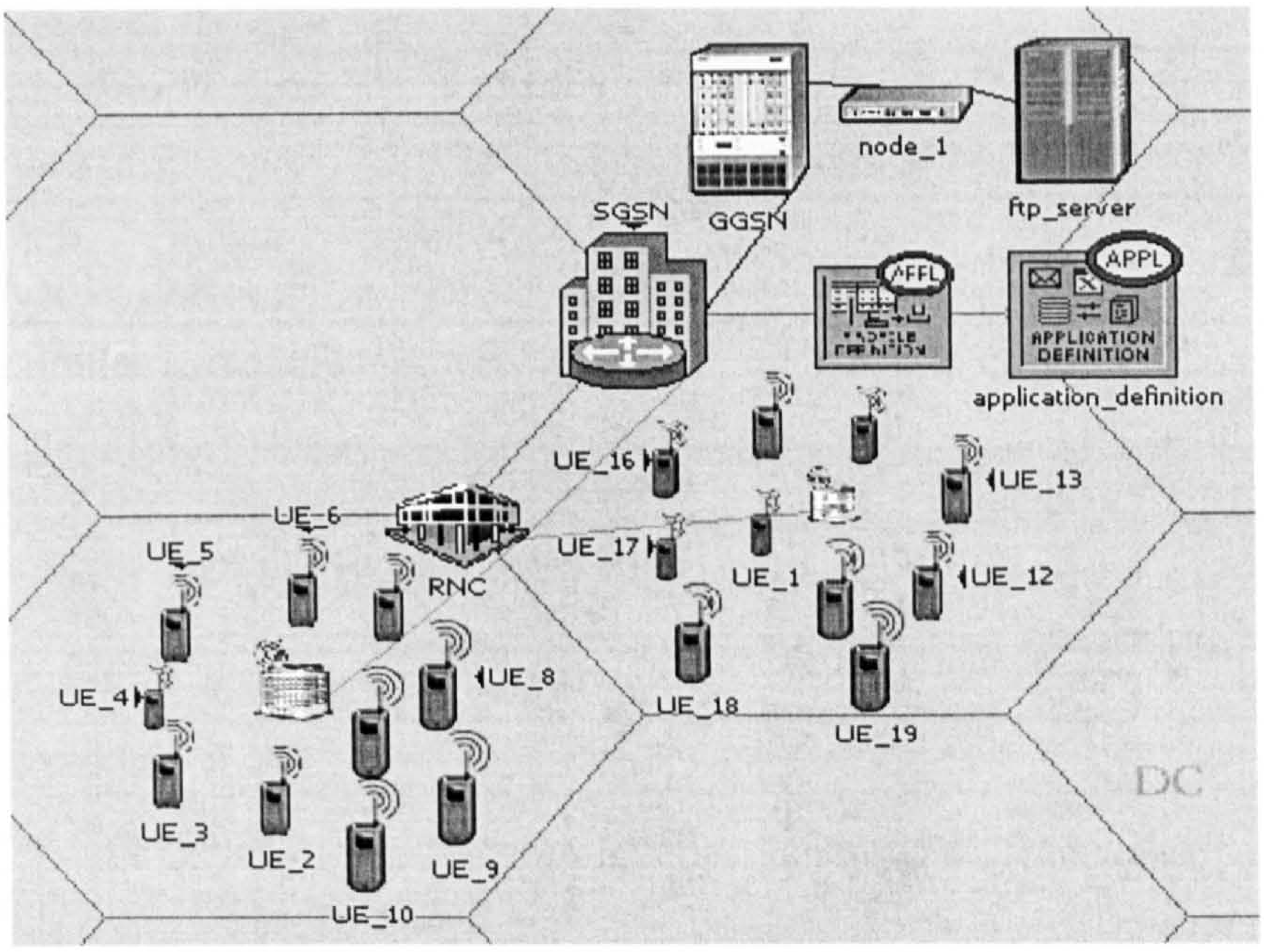


Figure 3.12: Screenshot of UMTS Handover

The impact of soft and hard handover on the system was studied. Two different scenarios were implemented where the networks and configurations of both scenarios were identical. The statistics for the transmission power, throughput, end-to-end delay as well as interference levels were collected while the mobile station follows a trajectory. The results were evaluated and compared. The average results were plotted using Excel. The first scenario focused on the statistics relates to the soft handover procedures.

The impact of distance on soft handover was also analysed and several different metrics were used to measure the quality of service. The soft handover scenario from the first model was used in this model. However, for the second scenario, the distance between the base stations was increased to approximately 3.5km and cell radius was increased to 2km. Table 3.2 shows the assumptions specifications that were used in the system simulation.

Table 3.2: UMTS Simulation Assumption Specification [51]

Parameters	Value
Chip rate	3.84Mcps
Frequency of system	1.9GHz
Shadowing variance	10dB
Voice activity factor	0.5
Service mix	100% voice users at 12.2Kbps
Uplink power control efficiency factor	0.85
Maximum transmit power per speech	0.5W
Mobile speed	50km/h
Minimum mobile threshold (Eb/N0)	-140dB
Node B antenna height	40m
Cell radius	1km
Node B spacing	1.73km
Number of cell in active set	3
Macro Diversity threshold	6dB
Macro Diversity Hysteresis	1.5 dB
Replacement Hysteresis	3.0 dB
Path loss model	Outdoor to Indoor and Pedestrian Environment
Time elapsed	5 minutes
Simulation time	350 seconds

3.5.1 Result Analysis for UMTS Handover

Figure 3.13 to Figure 3.16 shows the impact of handover on transmission power, actual Eb/N0, interference, and end-to-end delay with two different scenarios i.e. soft and hard handover scenario. Figure 3.11 shows the transmission power used by the MS while transmitting uplink packets during soft and hard handover. As can be seen, the transmission power increases as the user moves away from the base station and it decreases as the user move towards the base station to be connected. Since the mobile was set to move in a zigzag direction, the peaks of the uplink transmission indicates the maximum transmission power of the mobile and handover occurs at this point. The transmission power of the physical channel for both scenarios seems to be similar.

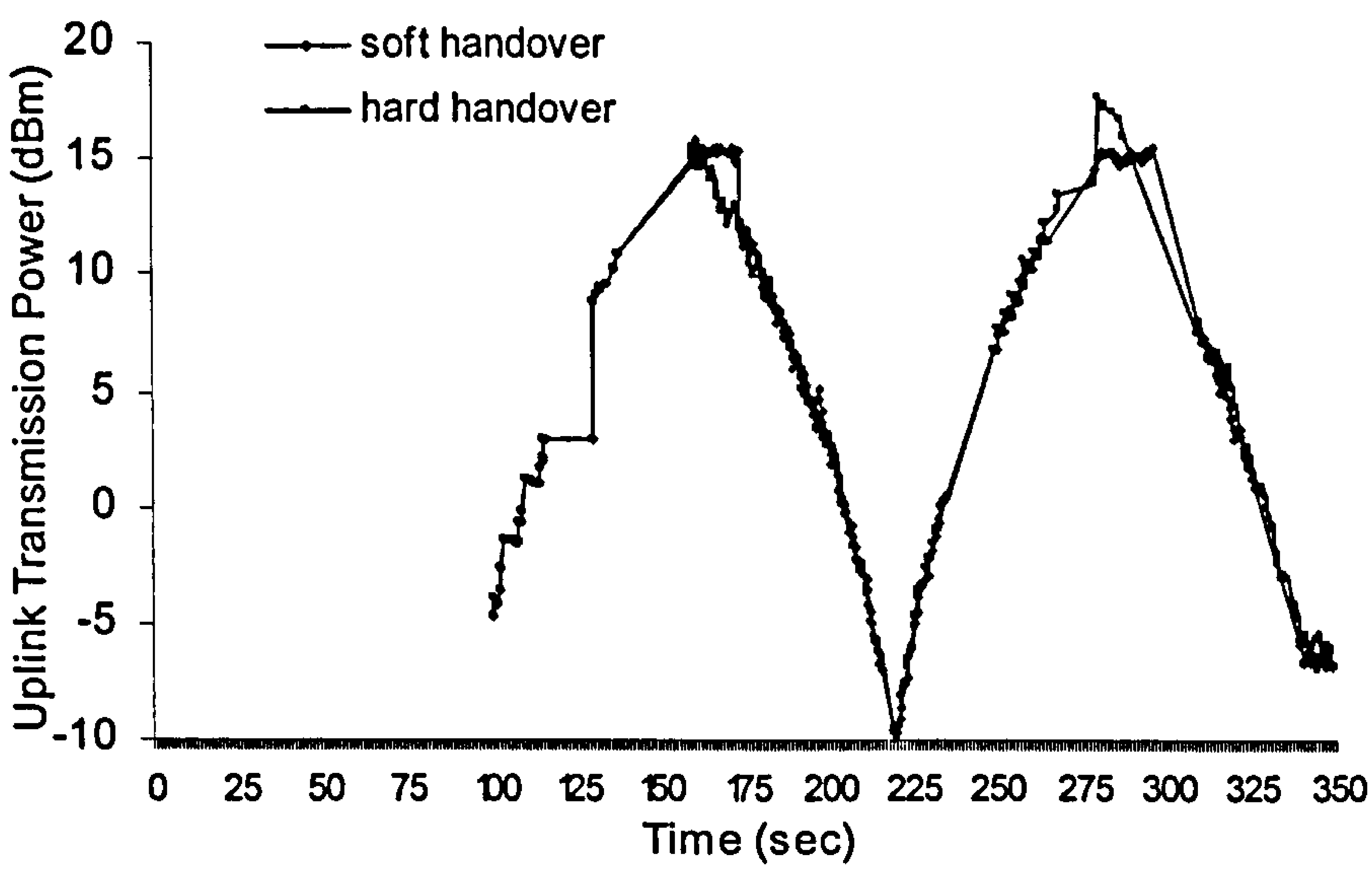


Figure 3.13: Uplink Transmission Power against Time

Figure 3.14 represents the actual Eb/No values of the uplink packets sent by this MS at the base station. Although multiple radio links is supported during the soft handovers, the results for both soft and hard handover show that the users perform handover at about the same power level.

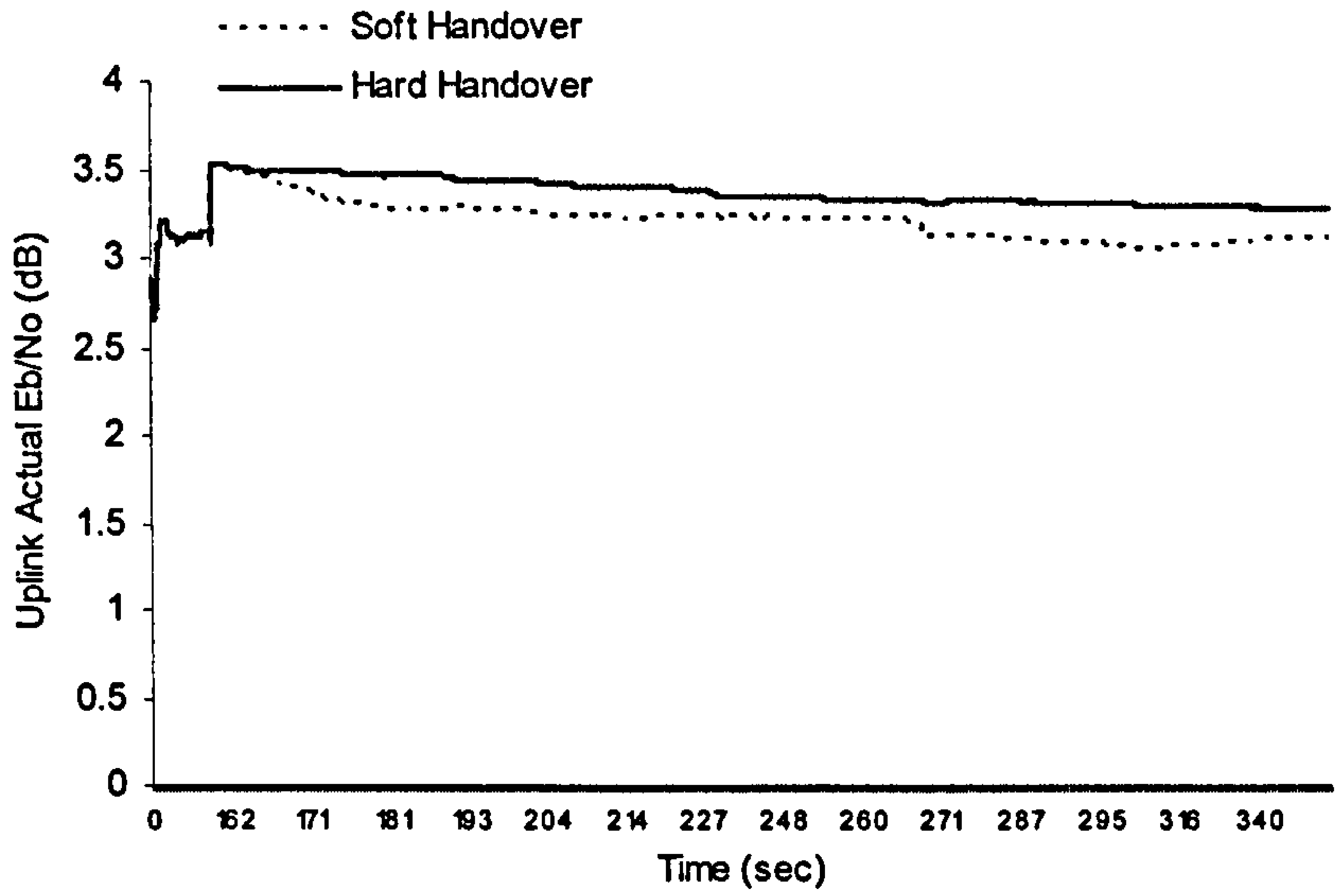


Figure 3.14: Actual (Eb/N0) against Time

Figure 3.15 represents the interference experienced by the uplink packets of this MS during the simulation time. The graph shows that the interference is slightly higher during the soft handover compared to the hard handover. This is because during the softer handovers the base station is receiving traffic through few cells simultaneously, while in the hard handover case, just one cell receives traffic at the same time. However, the average interference for both soft and hard handover is hardly noticeable in this study.

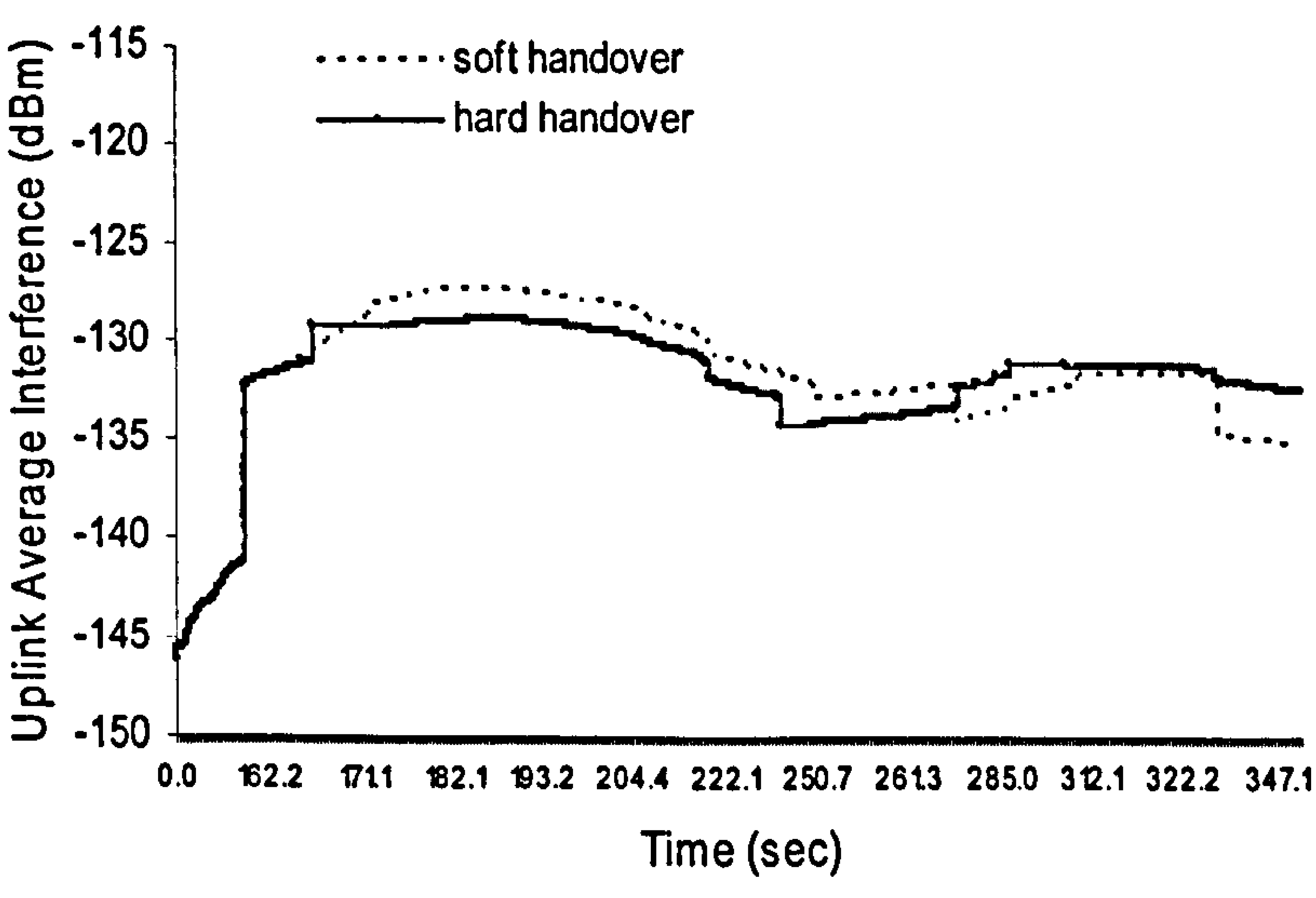


Figure 3.15: Uplink Interference against Time

Figure 3.16 shows the end-to-end delay for the soft handover compared to the hard handover. It can be noted that the soft handover has slightly lower end-to-end delay compared to the hard handover. The latency i.e. the average time it takes for a data packet to get to the final destination is low during the soft handover. However, the average delay for both soft and hard handover is about the same for this model.

Figure 3.16: End-to-End Delay against Time

Figure 3.17 to Figure 3.18 shows the impact of distance on transmission power and interference in soft handover scenario with two different distances. The statistics for the transmission power and interference at the MS were collected. Figure 3.17 shows the graph for the uplink transmission power against time. As can be seen from the graph, the transmission power for the larger cell size is slightly higher compared to the smaller cell size due to a small difference in the distance.

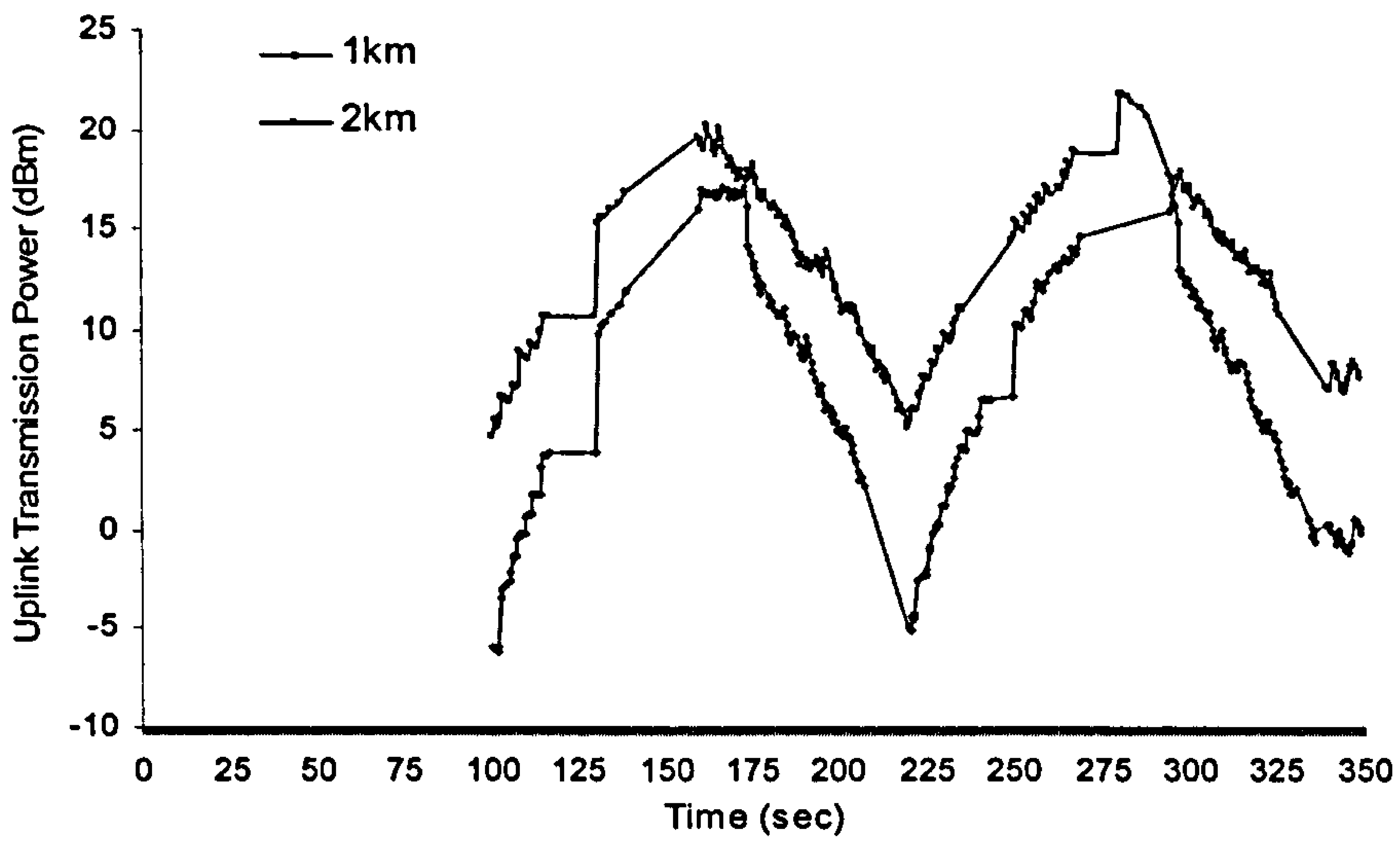


Figure 3.17: Uplink Transmission Power against Time for Soft Handover

Figure 3.18 shows the graph for the average uplink interference. The graph shows that the interference for the larger cell size is obviously lower.

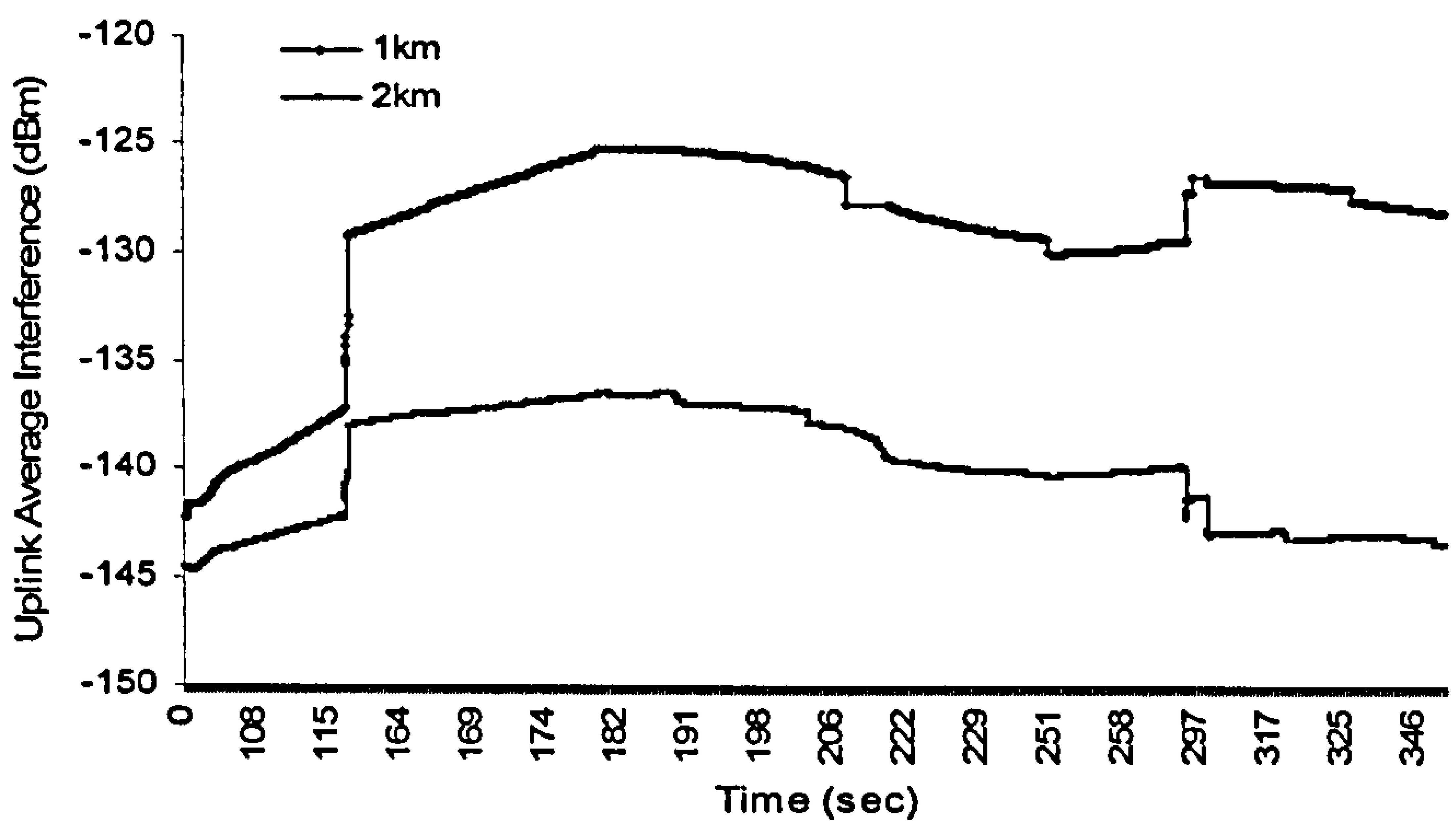


Figure 3.18: Uplink Interference against Time for Soft Handover

3.6 Conclusion

This chapter reviewed fundamentals handover process of GSM and UMTS systems. The handover model for both systems was presented. A GSM contributed model was used for the fundamental study of the GSM system. The model had a lot of errors and one of the reasons was because the model was developed based on the previous version of the software. After spending a lot of time on the model, it was realised that the system is not reliable. Thus, it is only suitable for understanding how the 2G system works but not for a research. There is no specific handover algorithm had been standardised for GSM. The GSM recommendations document only defined how the provided variables can be used. An example algorithm was given to demonstrate the usage of the parameters defined. A set of variables measured in the system were used to make a decision when a handover shall occur. The received signal strength was used as a basis for the handover decision. The measurements of the signal quality were made by the mobile station and the values were sent to their current base station.

The chapter also described the different between soft and hard handover. The UMTS model has several functionalities and it is almost a complete UMTS real system. However, some of the functionalities are not implemented in the model. In the simulation a simple soft handover algorithm was used. The parameters that affect the QoS estimation and system performance had been studied. The effect of the size of cells on soft handover was discussed. Some performance measures for handover algorithms were collected and the sampled values were averaged. The results showed that the cell size had an impact on the transmission power and interference, with the transmission power being higher in a large cell size of network. However, the larger the cell size, the lower the interference. On the other hand, smaller cell sizes lead to more frequent cell changes. In addition, types of handover affect the performance. However, for this study the results showed that the different in both soft handover and hard handover seems to be unnoticeable. This was because the model consists of two base stations only. Besides that, the assumptions that had been made affect the whole result.

Chapter 4

Traffic Modelling

4.1 Introduction

Mobile communication is one of the rapid growing markets in telecommunications. Recent technology offers service and equipment at an affordable cost to the end users. With the development of wireless packet service in 3G mobile communication systems, users will easily get connected anytime and anywhere. Phones can be integrated with other services over the Internet, such as video and data transfer alongside with the conversation. One of the features of 3G networks is IP based application such as VoIP where users can use the packet domain for making voice calls in order to reduce costs. Moreover, Internet Service Providers provide services for delivering voice packets across a network. They provide connectivity to the Internet where the users voice can be transferred using an IP.

Network traffic can be classified into time-based and non-time-based, i.e. real time and non real time. Non-time-based is insensitive to time such as image and data. Information is stored at the receiving points for later consumption. On the other hand, time based traffic requires information delivery for immediate consumption such as video and voice conversation. As the voice packet is categorised as a real time application, the need to guarantee the service is crucial. In the situation where the network suffers from overload due to congestion, the "best-effort" service is inadequate and it does not provide full reliability. The delivered data is not guaranteed. Thus, managing real time and non real time traffic is an issue [1].

4.1.1 Benefit and Challenges

Traditionally, voice communication has been carried over dedicated telephone networks and certain amount of resources is dedicated for each phone call. In the packet switching technique, several telephone calls are allowed instead of only one call in a circuit switched network. In addition, it offers services that are not available with a traditional phone. Most of the VoIP is used as an alternative to telephone call over a long distance, thus making a long distance telephone call using an IP network would be very cheap. The important feature of VoIP is the capability to facilitate voice and data convergent at an application layer in the same network infrastructure.

Since voice and data traffic can be integrated, the necessary infrastructure to provide both services is reduced. Implementing VoIP in a 3G network has a few challenges including the capability of the network to provide the required quality. Since VoIP uses an Internet connection, there are few factors that affect the quality. The challenges include delay, packet loss, jitter and security. Reliability is one of the major weaknesses as users would expect the system to be reliable and easy to use. The system frequently relies upon another service and no call can be made during a network failure or power outage. In addition, the voice quality must match the voice of conventional circuit based network. However, the quality of the conversation is usually lower than the normal telephone call.

Network congestion causes the loss of packets and an increase in the delay time. Hence, these features are likely to affect the quality of voice communication. In addition, the end-to-end delay is affected by the packet size and capacity of the service. Therefore, to ensure that the VoIP has the same voice quality as a traditional telephone call is a challenge. In fact the quality and overall reliability of the connection absolutely depend on the quality, reliability and speed of the Internet connection.

4.1.2 VoIP Process

VoIP allows voice conversations to be transmitted over a network using Internet technology. The data can be transmitted over the IP network which uses the IP protocol to transmit information. To be able to transmit a speech signal on the Internet, it must be digitised. VoIP is a method for transporting voice calls i.e. converting analog signals into digital data.

As speech is part of a real-time communication, the overall delay from source to destination should be low to avoid irritating long gaps of silence. When a conversation starts, a speech signal is sent from the sender to receiver. When the signal is detected, it will be converted into a digital representation. This process is called packetization and IP packets are sent across the IP network. The digitised information requires a certain amount of bandwidth for the connection.

Compression schemes are used to reduce the required bandwidth for voice communication. The speech is required to be compressed in order to reduce the size of each call. At the receiving end, the packets will be reconstructed into the speech signal. Once the compression block with speech data is received at the destination, it will be decompressed. The compressed signal will be transformed back into an analog signal. The technology converts the voice into digital unit and sends the packets over the network as shown in Figure 4.1.

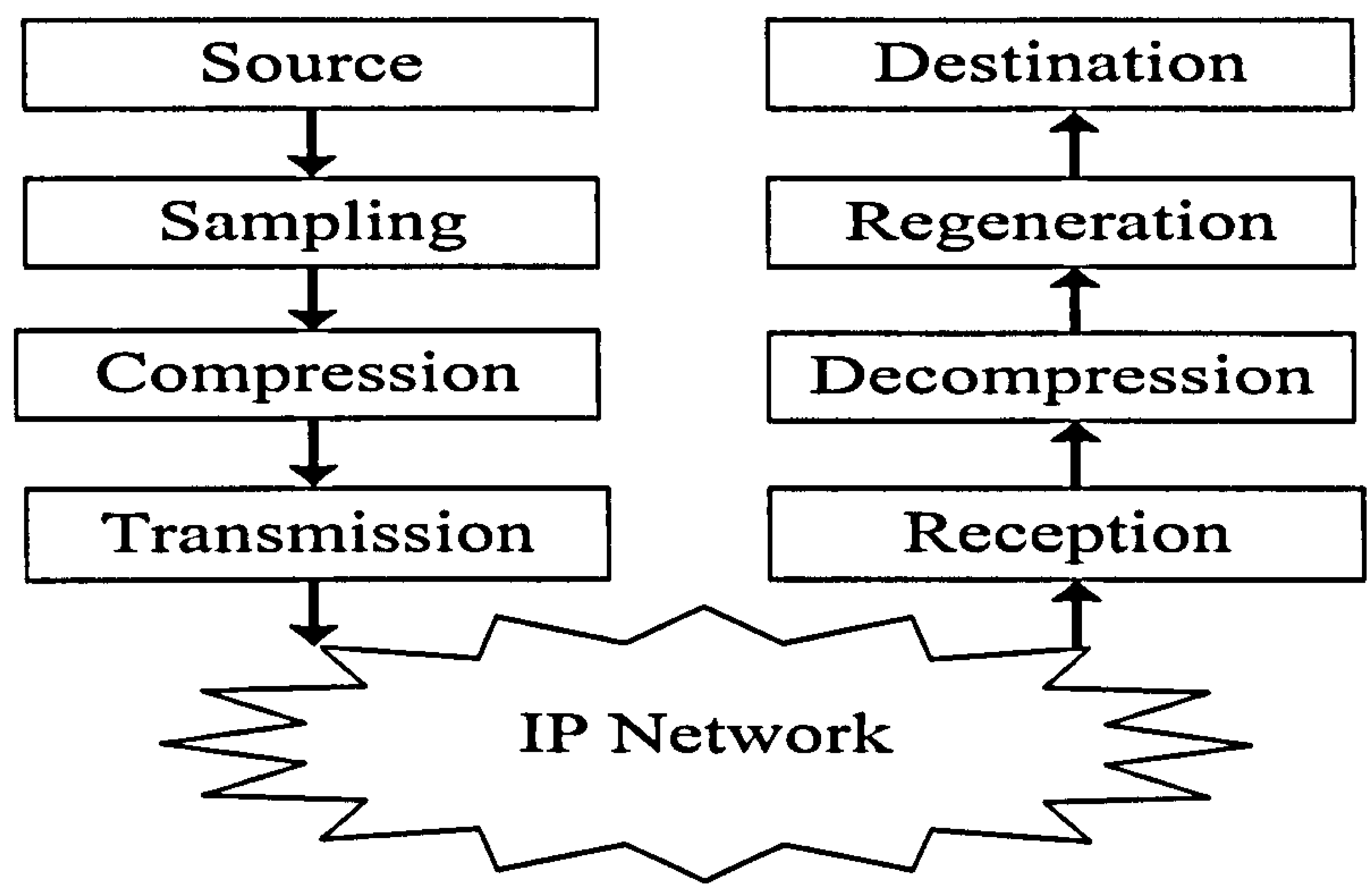


Figure 4.1: VoIP Process

A few problems were identified in voice transmission, such as the delay and voice quality. If the delay is too high, the quality of speech would be degraded below the acceptable level. The voice quality would also be affected by the load on the network due to packets lost in a heavy loaded network. If the number of lost packet is very large, it would cause an interruption in the communication between two parties.

4.2 Background

Several studies on the performance of VoIP have been investigated [28, 29, 31, 58-60]. VoIP performance was studied in [28] by analysing delay and packet loss. Research focusing on high-speed downlink channel was performed as in [29]. The VoIP network capacity under different packet delay budgets was evaluated. A full rate GSM speech codec that uses silence compression techniques was used.

The author focuses on the network delay and the delay jitter effect on VoIP in [31]. The numerical result for the investigation of the effect on delay and delay jitter to a speech sample were presented. The analysis of QoS had been the subject of much research in [58] where an approach to analyse the impact of UMTS service mix on capacity necessities was presented. In [59], the author derived an analytical expression to determine the packet size and dejiterring delay. The jitter behaviour was of greatest interest for Constant Bit Rate (CBR) sources of background traffic. It is the limitation for multimedia sources in high speed networks as traffic sources are very bursty.

In VoIP, the bandwidth needed can be adjusted according the requirement. However, the management of QoS and suitable protocol are needed. In [60], the end-to-end delay of each CBR and Variable Bit Rate (VBR) VoIP connection, their relationship with the number of "best effort" connections and the packet loss were evaluated. Similarly, the author evaluated the VoIP performance over high-speed downlink shared channel in [61]. However, the authors focused on system outage based on delay budget.

4.3 Voice Compression Techniques for IP Network

Communication system performance of VoIP application is a very important issue. Bandwidth limitation is one of the major issues concerning the VoIP network. As the number of users increase, the bandwidth requirement will increase. Compression technique is implemented in voice communication to conserve bandwidth and to assist the load. It offers the possibility of large compression ratios which result in more users being accommodated. As the available bandwidth is very limited, very often, the digitised speech is required to be compressed in order to reduce the amount of required bandwidth to transmit the signal and to maximize the number of users on a system. The bandwidth usage depends on the codec type and voice samples per packet [43].

There are a number of voice coding standards which have different rates, quality and complexity. The speech quality varies with the bit rate of the codes. In general, the type of codec, voice packet size and delay are some of parameters that have an impact on the VoIP call. These parameters affect the voice packet flow between the nodes in the call.

Compression technique will increase the packet size because it introduces an additional overhead. Figure 4.2 illustrates the additional header information that will be encapsulated in the digitised speech before it can be transmitted over an IP network. The total voice packet size is shown below:

IP Header	UDP Header	RTP Header	Payload
20 bytes	8 bytes	12 bytes	30.5 bytes

Figure 4.2: Overhead using Compression Technique

where IP Header represents Internet Protocol Header, UDP represents User Datagram Protocol Header, RTP represents Real-Time Protocol Header and Payload represents message to be sent.

Compression technique has proved to be effective in controlling the delay. Unfortunately, it also causes degradation of the signal and the target quality may not be achieved. Compressing and decompressing speech data introduces a certain amount of delay that could be unacceptable for real time information such as voice conversation. Hence, the overall delay has to be kept as low as possible to ensure a good communication quality. The quality of the speech resulting from the compression and the overall performance has to be balanced.

According to [43], the performance of speech coder determines the quality of the speech and the capacity of the system. Low bit-rate coding will enable more users to be accommodated with a limited allocated bandwidth. Implementation of speech coding must consume little power and provide tolerable speech quality. Speech coding aims to transmit the highest quality of speech with the use of the least possible channel capacity while maintaining a certain level of communication delay.

Different speech coders show different degrees of immunity to transmission errors and choice of the speech coder will affect the overall performance. AMR uses different techniques such as Algebraic Code Excited Linear Prediction (ACELP) and Voice Activity Detection (VAD). G.711, G.728, G.729 and GSM are among the most popular coding standards. G.729 uses ACELP and it is an algorithm that compresses voice audio in the size of 10msec. It is mostly used in VoIP due to its low bandwidth requirement and it is slightly higher rate codec compared to GSM. GSM Enhance Full Rate (GSM-EFR) is a codec that was developed to improve the quality of poor quality GSM full rate codec. It works at a bit rate of 12.2Kbps and it is compatible with the AMR mode. It provides good quality speech, although not as good as G729 [62].

G.711 is a standard to represent the Pulse Code Modulation (PCM) samples for signals of voice and it operates at 64Kbpsec. On the other hand, AMR has been targeted primarily at the GSM application and it was adopted as the standard speech codec by 3rd Generation Partnership Project (3GPP). It aims to be used as the default speech codec for the emerging UMTS standard.

AMR wideband speech codec has been introduced which bring considerable voice quality enhancement [63]. The coding rate changes rapidly in response to the conditions and the demand for capacity. The AMR rates are controlled by the radio access network. It improves the capacity and useful where signal levels are lower with higher interference. Eight source rates are defined for this multi-rate speech codec: 12.2, 10.2, 9.5, 7.4, 6.7, 5.9, 5.15, and 4.75 Kbps [64].

The speech quality varies with the bit rate of the codes as shown in Figure 4.3. Speech codecs can be categorised into waveform codec, vocoders and hybrid codec. The waveform codec are used at high bit rates and gives good quality speech whereas the vocoders operate at very low bit rates. As a result, it tends to product low quality speech. Hybrid codecs, on the other hand, use techniques from both coding scheme and give good quality speech an intermediate bit rates [65].

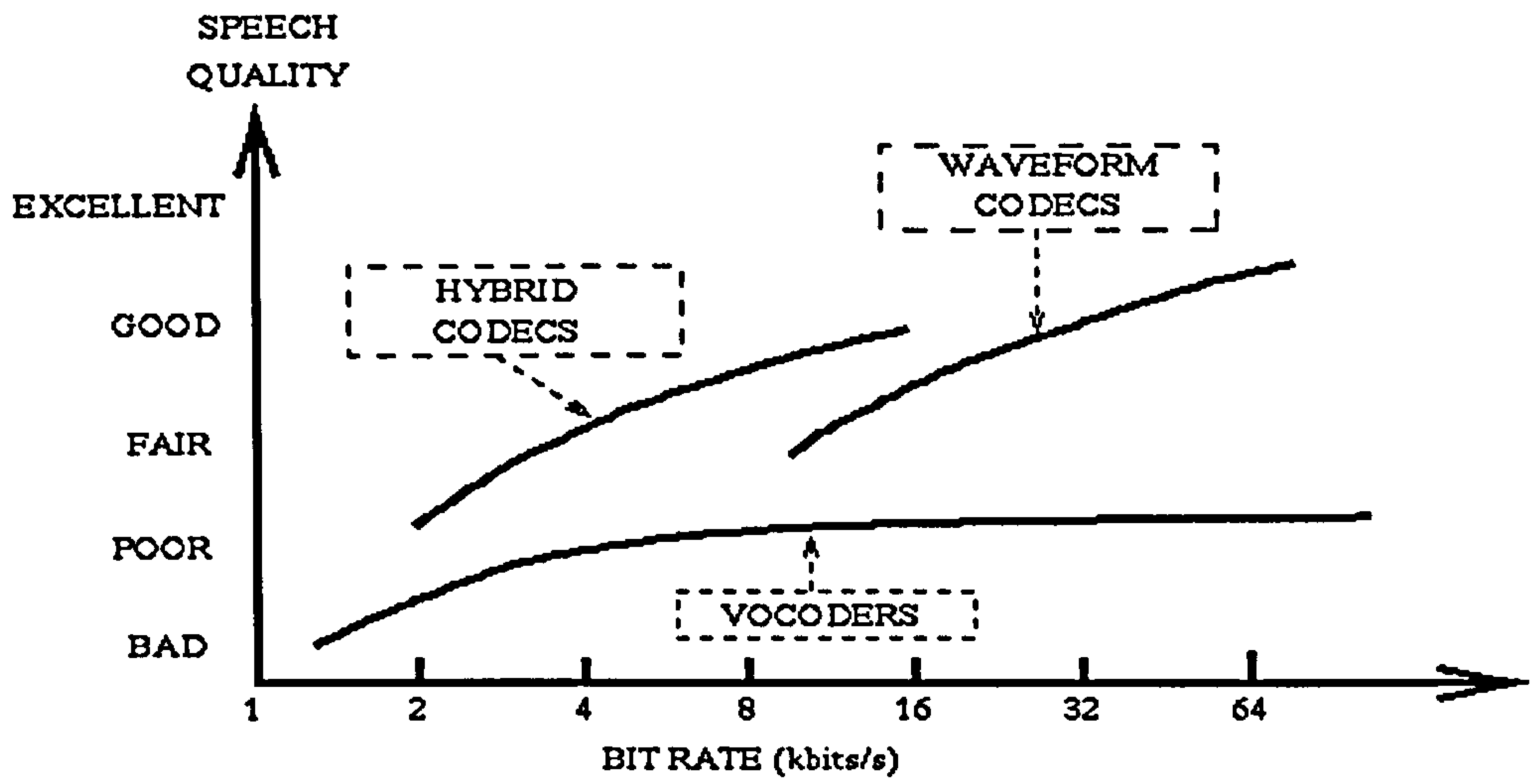


Figure 4.3: Speech Quality versus Bit Rate[65]

There are a number of voice coding standards, each associated with different rate, quality and complexity. A speech coding system is used to analyse the voice signal at the transmitter, transmit parameters derived from the analysis and synthesize the voice at the receiver. Each type of coder will produce a different size of effective bandwidth. Table 4.1 shows the effective efficiency of using different coder [66].

Table 4.1: Effective Efficiency of Different Types of Coder

Speech codec	# of bytes per frame	bit-rate	Total # of bytes	Effective efficiency
G.711	80	64 kb/s	120	67 %
G.723.1	24	6.4 kb/s	64	37 %
G.729	10	8 kb/s	50	20 %
G.729	10	8 kb/s	60	33 %
G.729	10	8 kb/s	70	43 %
GSM-EFR	31	12.4 kb/s	71	44 %

4.3.1 Voice Quality Assessment

Mean Opinion Score (MOS) is a standard used to ascertain the voice quality of speech produced by a compression technique. A wide range of listener judge the quality of a voice sample interrelated to a particular compression technique. The score are averaged to obtain a quantitative indicator of the systems performance, which also indicates the accuracy of voice quality to real life communication.

The voice quality is given a range of 1 to 5 i.e. (worst to excellent) as shown in Figure 4.4 [67]. The results are averaged across all the listeners.

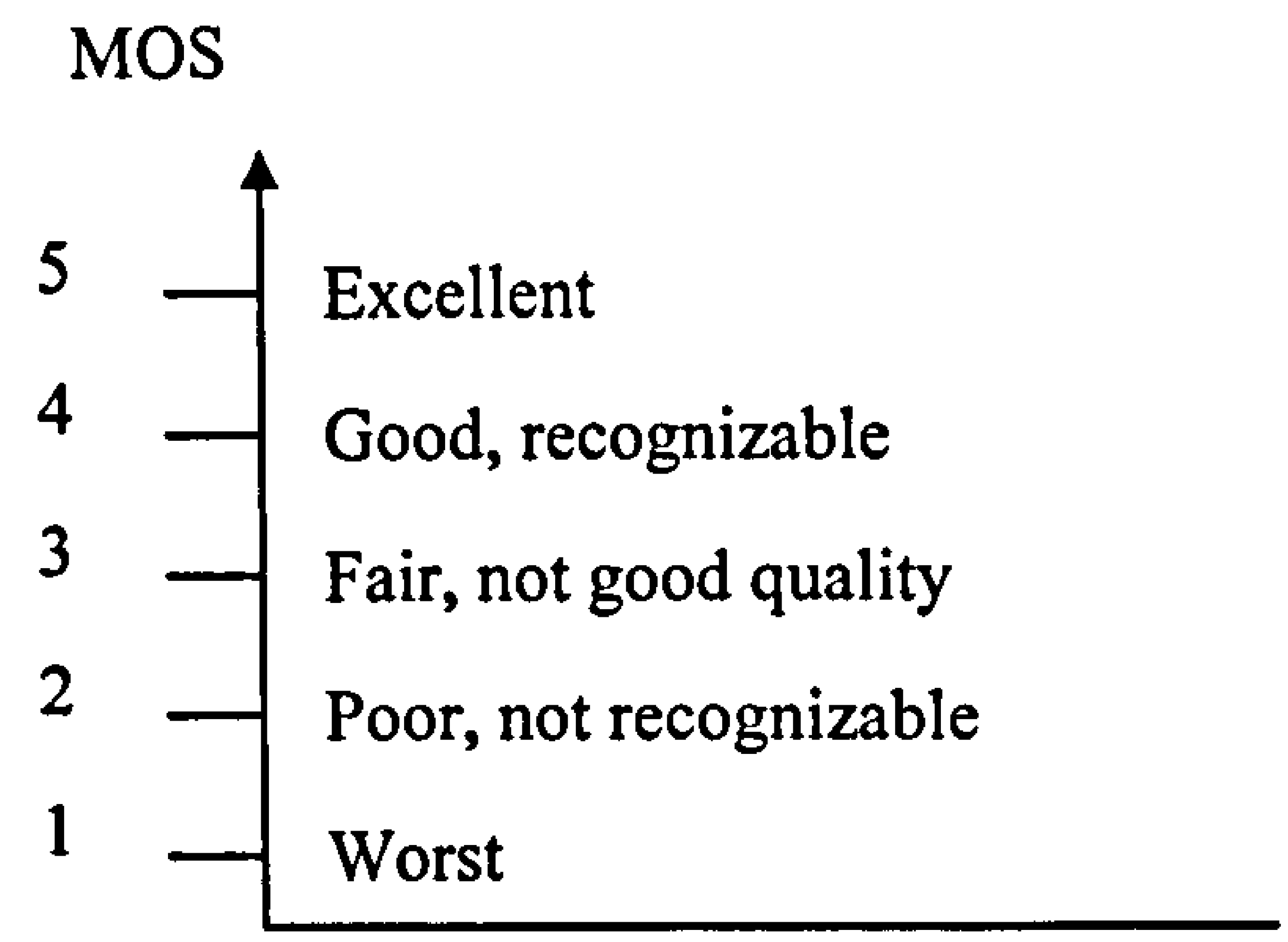


Figure 4.4: Quality Assessment

MOS testing is used to quantify the performance of different types of codec. A few standards voice compression techniques have established and most widely known standards are the G standards of the International Telecommunication Union Telecommunication Standardization Sector (ITU-T) and European Telecommunications Standards Institute (ETSI) GSM standards[68]. Table 4.2 shows a sample relationship between compression techniques with MOS scores [69].

Table 4.2: Compression Technique and MOS Score

Compression Method	Bit Rate (Kbps)	MOS Score
G.728 LD-CELP	16	3.61
G.729 CS-ACELP	8	3.92
G.729a CS-ACELP	8	3.7
GSM HR	5.6	3.5
GSM EFR	12.2	4.0
G.711 PCM	64	4.1

The voice quality based on compression schemes can be tested in two ways that are subjectively and objectively. Subjective result is collected based on human testing whereas objective result is collected using computers. Although it seems logical to convert all calls to low bit rate coders to save the infrastructure cost, there are some weaknesses. One of them is signal distortion, in which MOS score can drop to unacceptable level.

4.4 Voice Traffic Analysis

Traffic characteristics and the corresponding communication requirements can characterize an application. Its traffic generation process can formally specify the traffic characteristics of an application. Since the traffic generation process is basically a sequence of packets generated at arbitrary instances, two stochastic processes can characterise the traffic pattern [70]:

a) The packet arrival process
b) Packet length distribution function.

In general, there are three basic elements in a queuing system
a) arrival i.e. number of arrival per unit time
b) service mechanism i.e. capacity and duration of service time
c) service discipline i.e. the type of queue

In voice communication, the transmission of packet follows an Interrupted Poisson Process (IPP) with alternating active and inactive states. Voice packets are generated from a voice signal where the voice is assumed to be in two states that are "talkspurt" (active period) and "silent" mode (inactive period). Several models were introduced to model the burstiness and correlation characteristics of the packet arrival process from a voice source. Figure 4.5 shows a basic model of a voice call that is a periodic process alternating between the two states, where α and β represent the probability of active and inactive state, respectively.

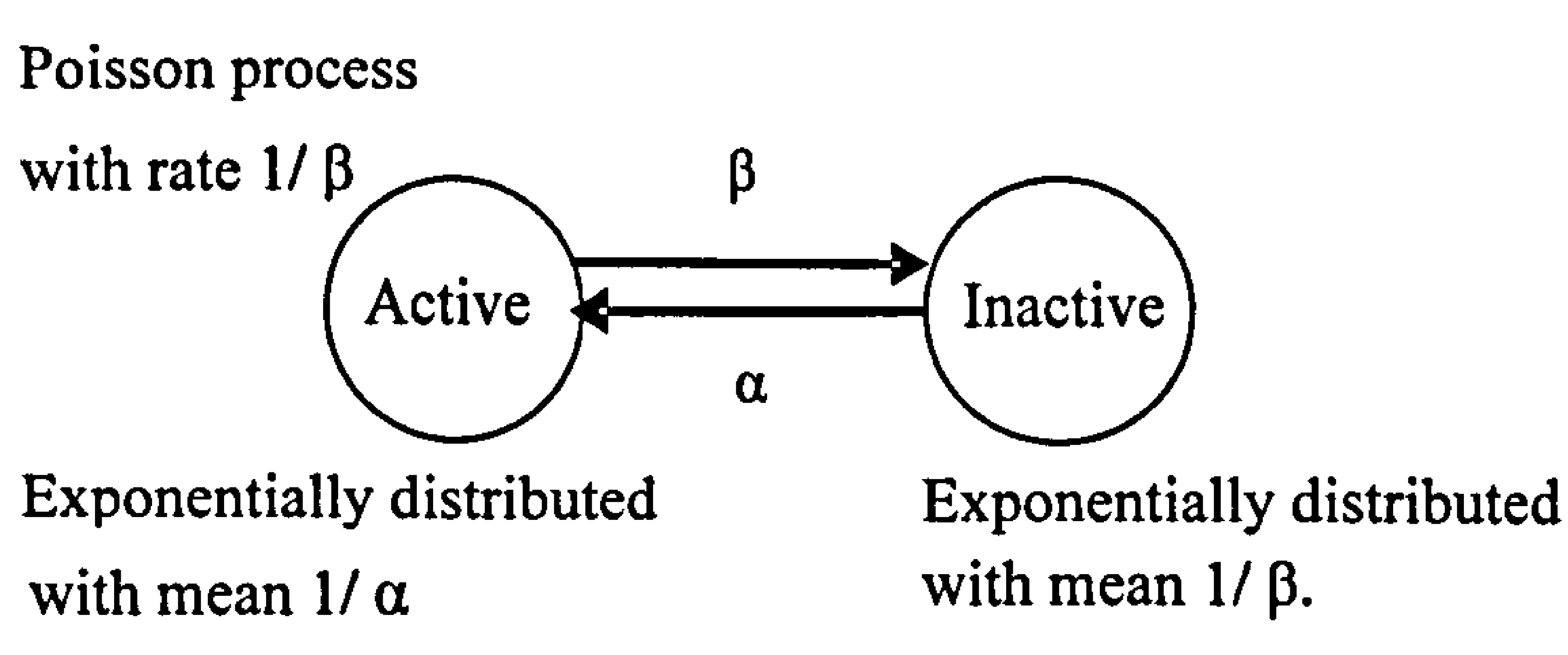


Figure 4.5: Behaviour of a Voice Source

When a user is actually speaking, the voice source will be in active state for a duration of time and it generates fixed length packets at periodic packet streams with a mean value of $1/\alpha$. It then goes to an inactive state listening with no packets being generated and stays inactive for a given time which is exponentially distributed with a mean value of $1/\beta$. The duration of active and inactive periods is assumed to be distributed exponentially.

A large portion of delay is due to the analogue-to-digital conversion, compression, packetization and overhead. The time taken for a voice from the talker's mouth to reach the listener's ear can be measured. For the real time services like voice, this end-to-end delay has to be very low. Figure 4.6 illustrates the block diagram of a mobile station that encodes and sends a voice stream.

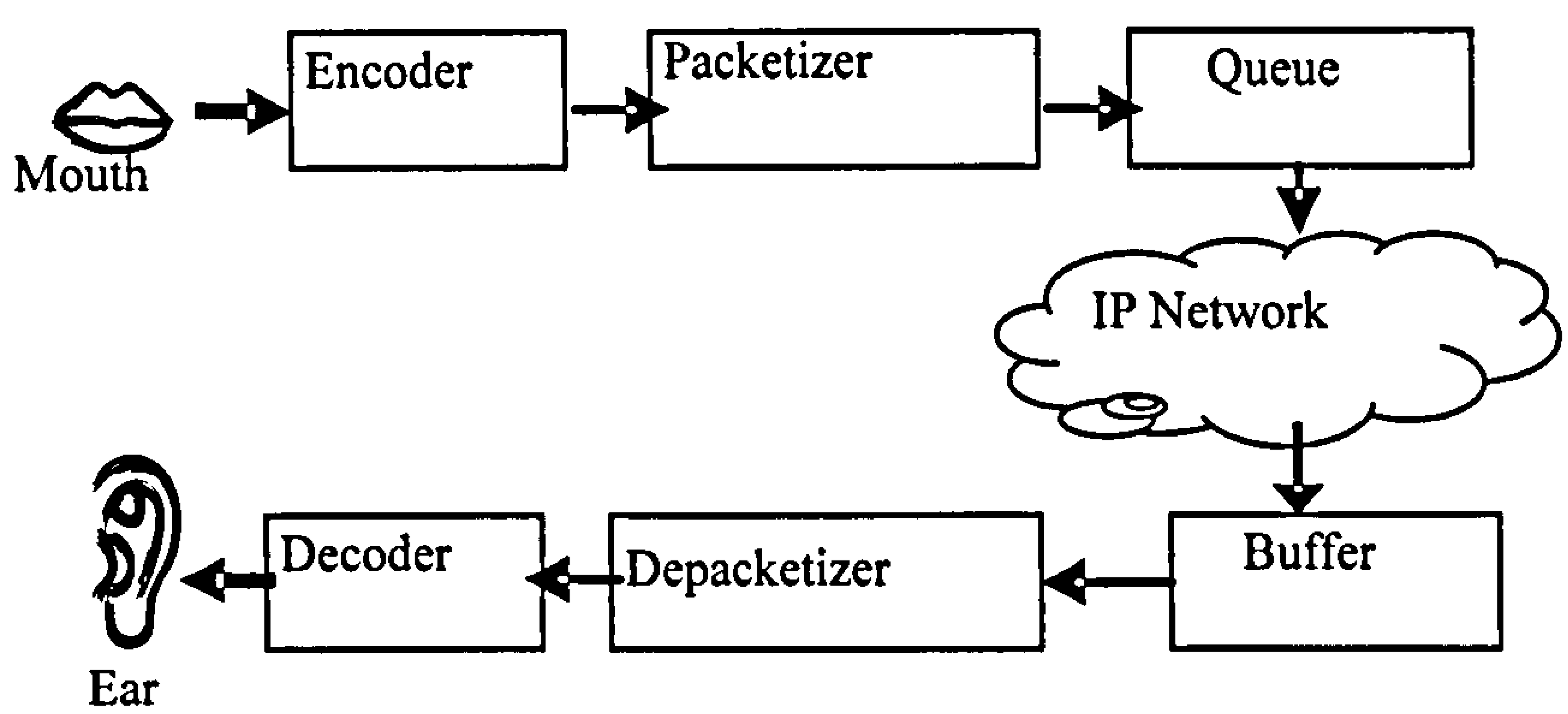


Figure 4.6: Voice Packet Encoded and Decoded

4.4.1. Voice Packet Arrival

To investigate the performance of Voice over IP, a theoretical model was developed. An analytical expression for the mouth-to-ear delay is derived and the expression is used to analyse the end-to-end delay and system performance. The packet arrival rate is part of the traffic model. Let N denote the number of users in the communication channel queue. For k customers, $N = (N_1, N_2, \ldots N_k)$, the aggregate packet arrival rate λ (packet), represents existing traffic in the queue with interarrival time τ (packets/sec) i.e. the time between two successive arrivals [70].

$$\text{Packet arrival rate, } \lambda = \left(\frac{N}{\tau}\right) \tag{4.1}$$

$$\text{Mean arrival rate, } m = \left(\frac{1}{\lambda}\right) \tag{4.2}$$

Since packets are only generated at active periods, the packet arrival rate depends on the number of active sessions. Assuming each active source produces packets with exponentially distributed lengths according to a Poisson process, the packet length and service rate were calculated based on the types of encoder.

In the calculation, the probability of active state $P(\alpha)$ was used to consider the traffic behaviour which represent the actual arrival rate. An active session has an arrival rate λ which is modified by the probability value. As a result, this modifies the overall packet arrival rate. As the period of active mode increases, $P(\alpha)$ also increases.

Thus, the probability of the arrival rate for an active session can be represented by the following [70]:

$$Pr(N = n) = P_n$$

$$= \frac{e^{-\lambda\tau}(\lambda\tau)^n}{n!}; \quad \text{for n=0, 1, 2,...} \tag{4.3}$$

Let $P_{on}(t)$ and $P_{off}(t)$ represent the probability the active and inactive session at time τ, respectively. The probability of transition states for active-to-active, active-to-inactive, inactive-to-active and inactive-to-inactive are given as:

$$1 - \alpha\delta t \tag{4.4}$$

$$\alpha\delta t \tag{4.5}$$

$$\beta\delta t \tag{4.6}$$

$$1 - \beta\delta t \tag{4.7}$$

If t is the time for the last active signal, the probability that it will be active at time $t + \delta t$ is:

$$P_{on}(t + \delta t) = P_{on}(t)(1 - \alpha\delta t) + P_{off}(t)(\beta\delta t + \delta t) \tag{4.8}$$

The probability that it will be inactive at time $t+\delta t$ is:

$$P_{off}(t+\delta t) = P_{on}(t)\alpha\delta t + P_{off}(t)(1-\beta\delta t) + \delta t \qquad (4.9)$$

It can be deduced that,

$$P_{on} = Limit_{\ t\to\infty}\ P_{on}(t) = \frac{\beta}{\beta+\alpha} \qquad (4.10)$$

Similarly,

$$P_{off} = Limit_{\ t\to\infty}\ P_{off}(t) = \frac{\alpha}{\alpha+\beta} \qquad (4.11)$$

For the steady state equilibrium solution,

$$P_{on}(t) + P_{off}(t) = 1 \qquad (4.12)$$

Each period can be represented by an exponential distribution of means $1/\alpha$ and $1/\beta$ respectively. Let τ represents the interarrival time, the arrival rate is given by:

$$\text{Peak arrival rate, } h = \frac{(1/\alpha)}{\tau(1/\alpha+1/\beta)} \qquad (4.13)$$

4.4.2. Service rate of Communication Channel

The number of packets in a "talkspurt" is geometrically distributed where the mean number of packets in a "talkspurt" depends on the encoder used. A GSM codec has a bit rate of 12.2Kbps. The GSM codec with a frame size of 20ms gives a mean of 32.5 packets per "talkspurt". The second coding scheme used is G729 which has a data transfer rate of 8Kbps.

The codec frame size of 10ms gives a mean of 10 packets per"talkspurt". The parameters for both coding schemes used in the simulation are coding rate, packet length, service rate and frame size. The first burstiness parameter set has 1/α=352 ms and 1/β=650 ms which corresponds to a 35 % activity factor. The average service rate (packet/sec) is calculated by dividing the average file size (packet/bit) with the available data rate (bit/sec). Thus, the service rate of channel is given by:

$$\mu = E[L]/R \qquad (4.14)$$

where $E[L]$ is the expected length of a file and R is the data rate.

4.4.3 Channel Utilisation

Utilisation of a communication network depends on the average packet arrival rate and the average service rate. The utilisation factor (ρ) varies with the average rate of request (λ) and the average amount of data per request i.e. the average service rate (μ). The utilisation can be calculated as [70]:

$$\rho = \lambda/\mu \tag{4.15}$$

Time-based applications are becoming widely used and they cause heavy traffic on mobile networks because the requested amount of data might be very high. If an average arrival rate is higher than the average service rate, the network becomes unstable. Therefore, the utilization factor should remain between zero and one. The utilization of a cell is calculated as the average load created by voice calls. The average load can be defined as average the call request and the average service rate for serving the request.

4.4.4 Voice Delay

One of the most important network performance measures is end-to-end delay. Each link consists of four delays components: processing delay, queuing delay, transmission delay and propagation delay. The total delay for a voice user in the system is equal to the sum of the waiting time in the queue and the service time. It is calculated using Little's formula [70].

Average packet delay,

$$\begin{aligned} \mathrm{E[T]} &= \mathrm{E[W]} + \mathrm{E[S]} \\ &= \frac{\lambda}{\mu((\mu-\lambda))} + \frac{1}{\mu} \\ &= \frac{1}{(\mu-\lambda)} \end{aligned} \tag{4.16}$$

where $\mathrm{E[T]}$ is the total delay, $\mathrm{E[W]}$ is the waiting time in a queue and $\mathrm{E[S]}$ is the service time.

The time taken for a voice from the talker mouth to reach the listeners ear can be measured. For the real time services like voice, this end-to-end delay has to be very low. According to [71], most users will not notice the delay if it is below 100ms. If it is between 100ms and 300ms, users tend to notice a low response from the other users. However, it is more that 300ms, the delay is obvious to the users and conversation seems to be impossible. The effect of end-to-end delay is as shown in Table 4.3 [56].

Table 4.3: Effect of End-to-End Delay on Voice

Delay	Effect on Voice Conversation
50ms	No audible delay
100ms	No audible delay if the link is good quality
150ms	Start to have an effect on voice communication
250ms	Significant disturbance
400ms	Upper limit of conversation delay
>600ms	No communication possible

4.5 Simulation Model for Voice Over IP

The properties of the voice packet arrival process differ from those of the burst non-voice traffic since active voice source generates a periodic packet stream. Therefore, a model needs to accurately reflect the statistical properties of packet voice systems in order to meet the acceptable voice delay constraints. In this study, simulations were carried out for both GSM and G.729 encoder to show their effects on the end-to-end delay, jitter, packet loss and throughput. The delay was analysed using different types of coding schemes to estimate the actual voice packet arrival.

As the speech signals are being transmitted, the delay in a one way process is analysed. The system was setup with one way communication and the performance was evaluated over a range of system parameters. It has been evaluated through extensive simulations for different values of various packet arrival rates with different types of encoders. The system performance was evaluated under certain IP conditions and the model is verified by the simulation result.

The model is shown in Figure 4.7 where the load on the sender was increased to represent an increase in traffic arrival. The model is then used to estimate the queue length and packet loss ratio. The performance is measured in terms of the end-to-end delay, jitter and packet loss. Assuming an On-Off voice source with actual packet arrival based on the probability of active state, the end-to-end delay is expected to increase by the same amount, if only conversational sources are present in the system. The types of encoder will affect the end-to-end delay because the packet size and service rate will be different.

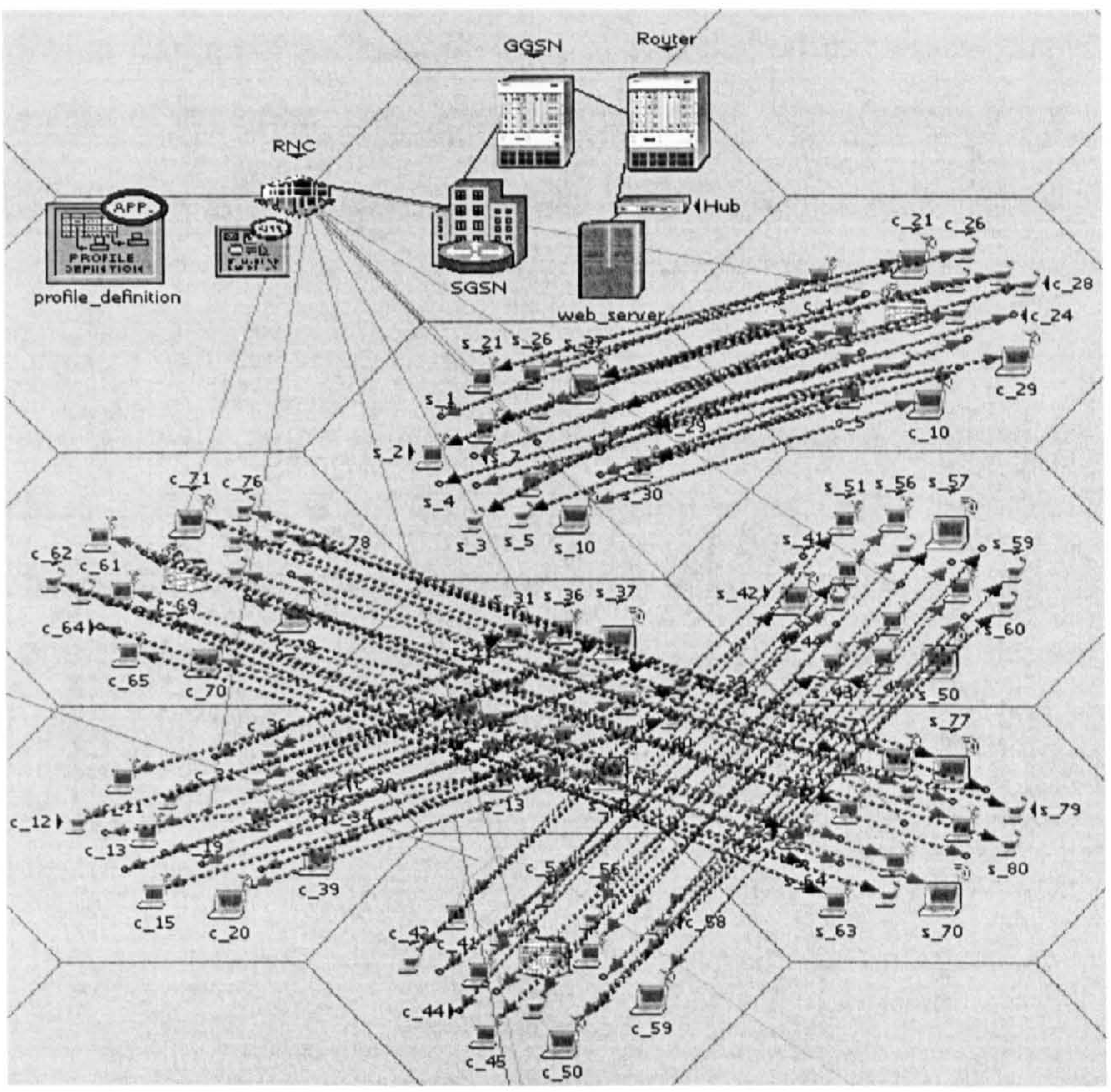


Figure 4.7: Screenshot of UMTS Handover

For both GSM and G.729 coder schemes, the end-to-end delay is calculated and compared against the simulation result. The delay considers all packets transmitted from a source node to its destination node and the subsequent delay incurred. The packet voice end-to-end delay includes the packetization time, compression time, queuing time, transmission time and network time. It is assumed that the compression and decompression time are equal. The parameters used in the simulation are as shown in Table 4.4. The packet size and the effective bit rate are computed based on the parameters.

Table 4.4: Setting Used for Voice Simulation Environment

Parameter	Value	Value
Voice Encoder Scheme	GSM	G.729
Talkspurt Length	Exp(0.35)s	Exp(0.35)s
Silence Length	Exp(0.65)s	Exp(0.65)s
Frame Size	20ms	10ms
Look ahead size	0ms	5ms
DSP Processing ratio	1	1
Coding rate	12.2Kbps	8Kbps
Voice frame per packet	1	1
Types of services	Interactive voice	Interactive voice

4.5.1 Voice Packet End-to-end Delay

In order to ensure that the voice quality is acceptable to the users, the packet end-to-end delay and jitter threshold should conform to the standard. Figure 4.8 shows the effect of voice traffic on packet end-to-end delay using GSM encoder. It is noted that as the traffic increases, the end-to-end delay is also increased consistently. After a certain point, the delay increases drastically. The result obtained from the theoretical model is compared with the simulation program [72]. The theoretical results seem to agree with the simulated results although it slightly differs as the simulation result shows that the delay fluctuates.

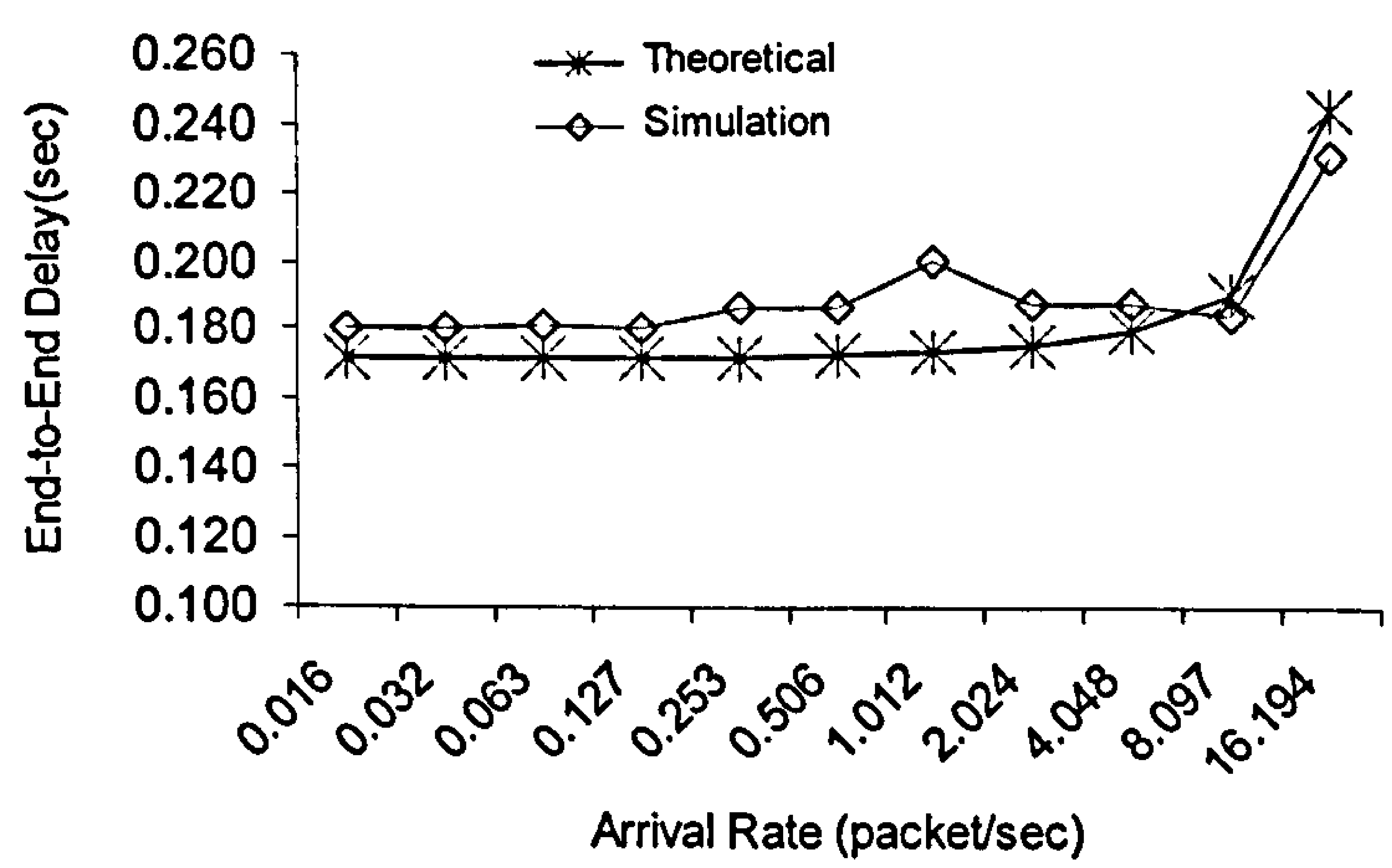


Figure 4.8: End-to-End Delay against Voice Traffic

Similar result on packet end-to-end delay using the G.729 encoder is shown in Figure 4.9. The result shows a significant increase in delay after a certain arrival rate. The theoretical model result indicates that it only matches closely to the simulation result [72]. This is an indication that the developed model is suitable to understand queuing performance with time-varying transmission capacity. However, after a certain amount of arrival rate, the simulation result starts to increase drastically.

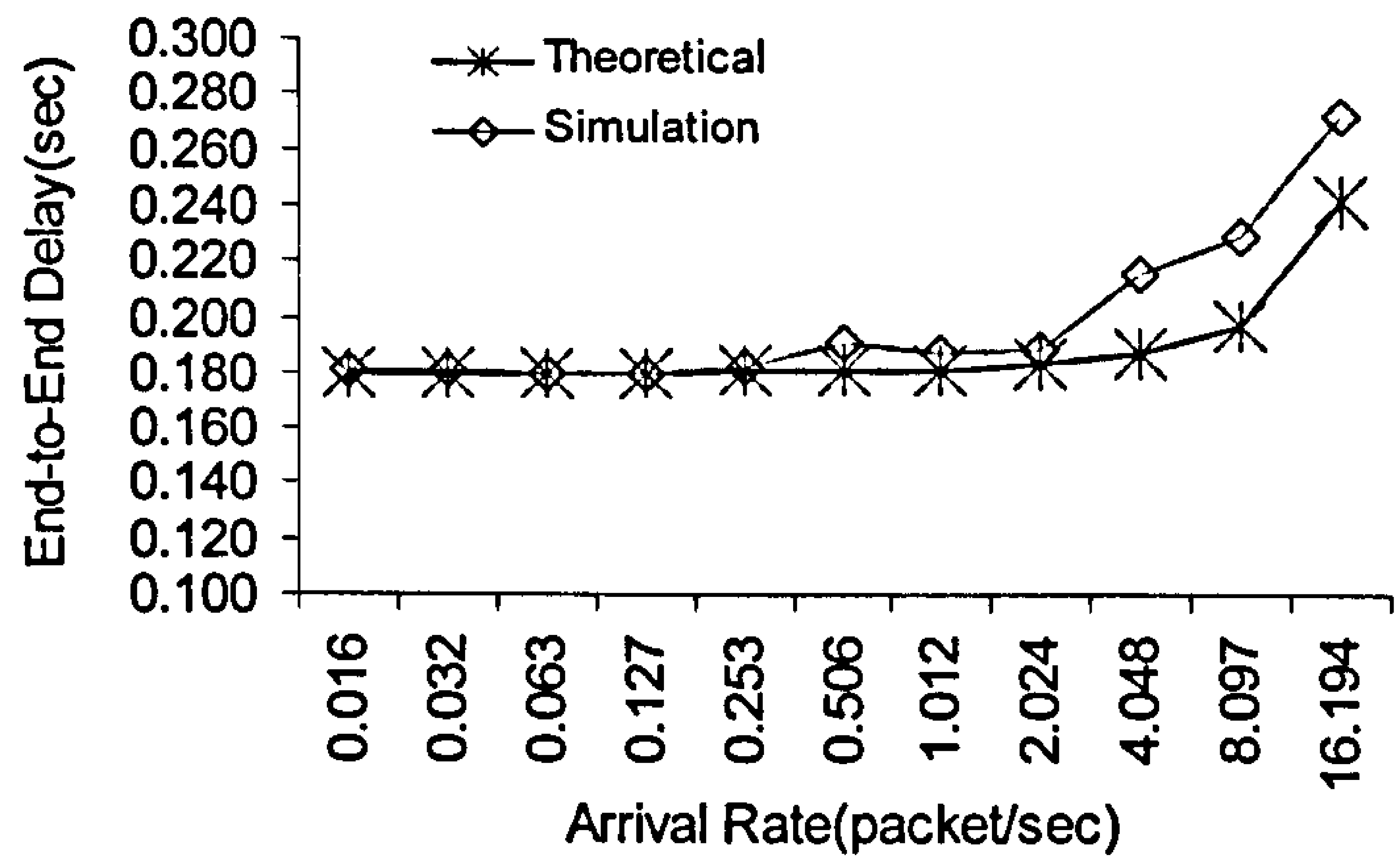


Figure 4.9: End-to-End Delay against Voice Traffic

Figure 4.10 shows the effect of number of users on the delay. The graph shows that when the number of user increases, the delay maintains a value of about 180msec which conforms to the standard. The simulation result is compared with the theoretical model and it seems to agree for low number of users [73]. However as the number of user increases, the delay increases drastically after the number of users is more than 8.

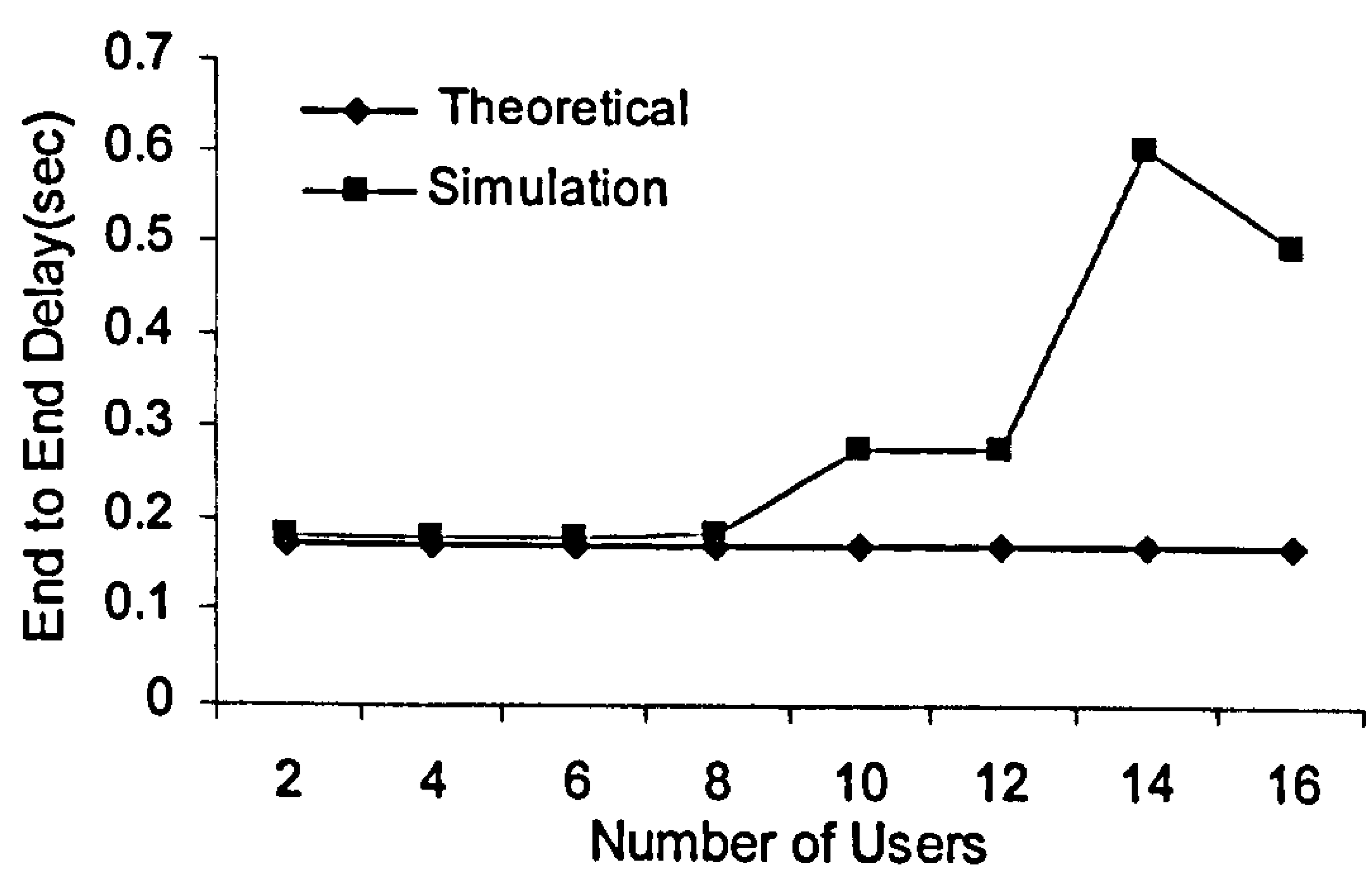


Figure 4.10: End-to-end Delay against Number of User

4.5.2 Jitter

It was specified that the jitter value should be close to 0 in order to conform to the standard. Figure 4.11 illustrates the effect of packet arrival rate on the jitter using different types of encoders. The jitter remains at about the same value and at a certain arrival rate, it increases sharply. The result shows that G729 encoder has more jitter compared to the GSM encoder.

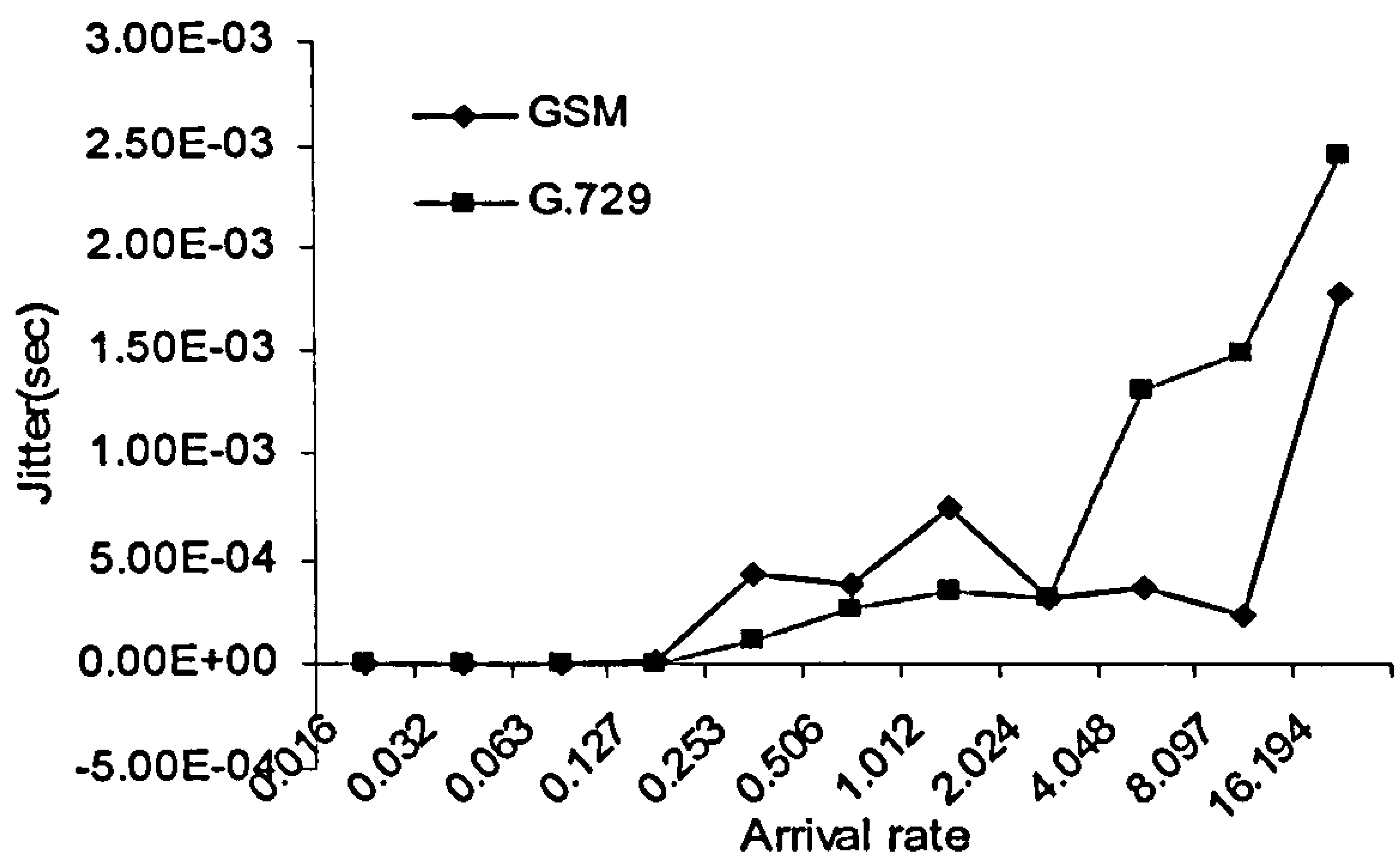


Figure 4.11: Jitter against Arrival Rate

Figure 4.12 shows that as the number of user increase, the jitter value remains at about the same value for low number of users and it increases as the number of users increase.

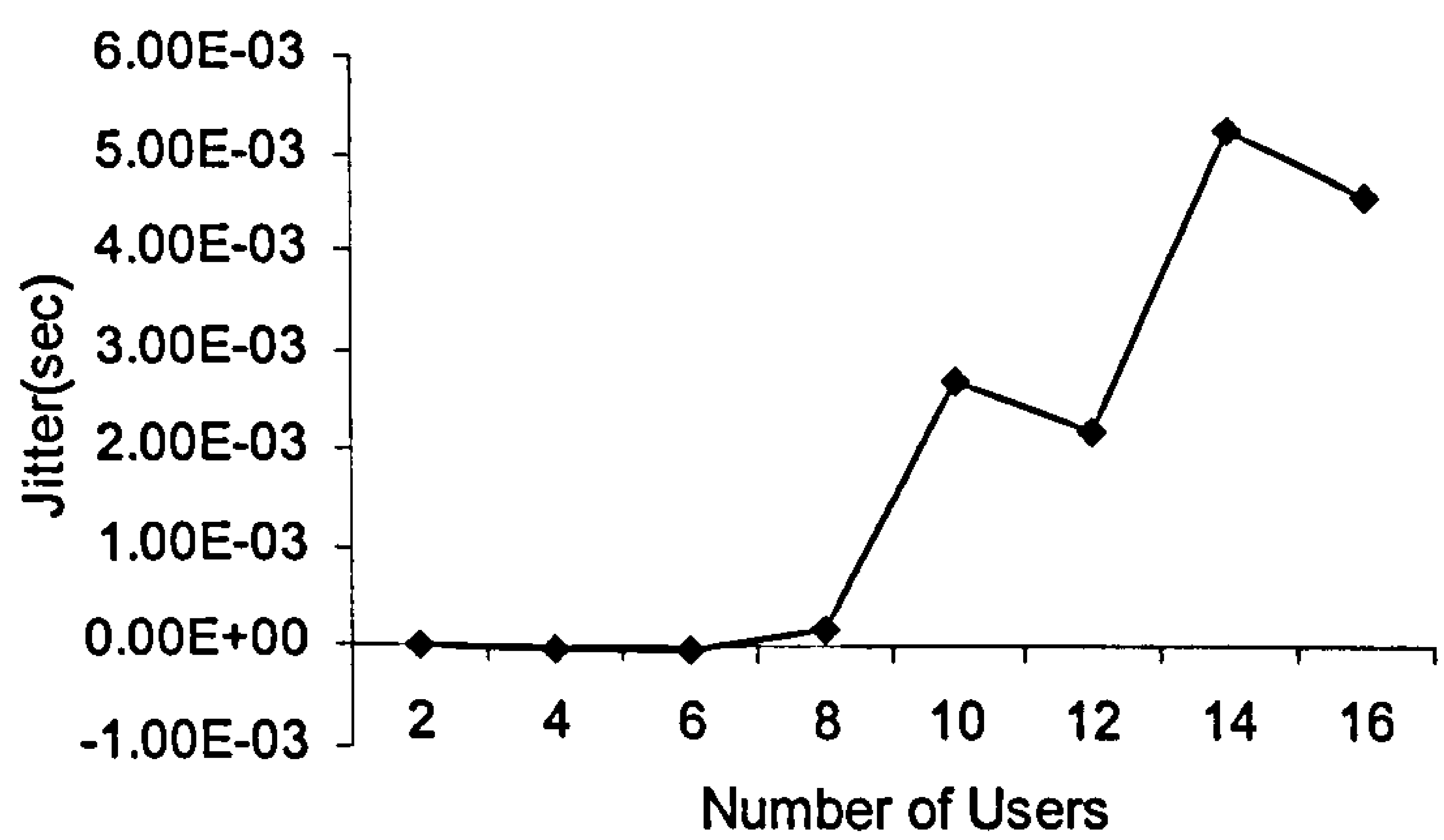


Figure 4.12: Jitter against Number of User

4.5.3 Packet Loss

The percentage of packet loss is shown in Figure 4.13. The result shows that the average percentage of packet loss is about 2%. As the packet arrival rate increases, the value is maintained up to a certain value of packet arrival rate. When the packet arrival rate continues increasing, the percentage of packet loss slightly increases. The packet loss is slightly noticeable using GSM encoder. It is noted that there is no significant loss of packet using the G729 encoder that affect the VoIP performance.

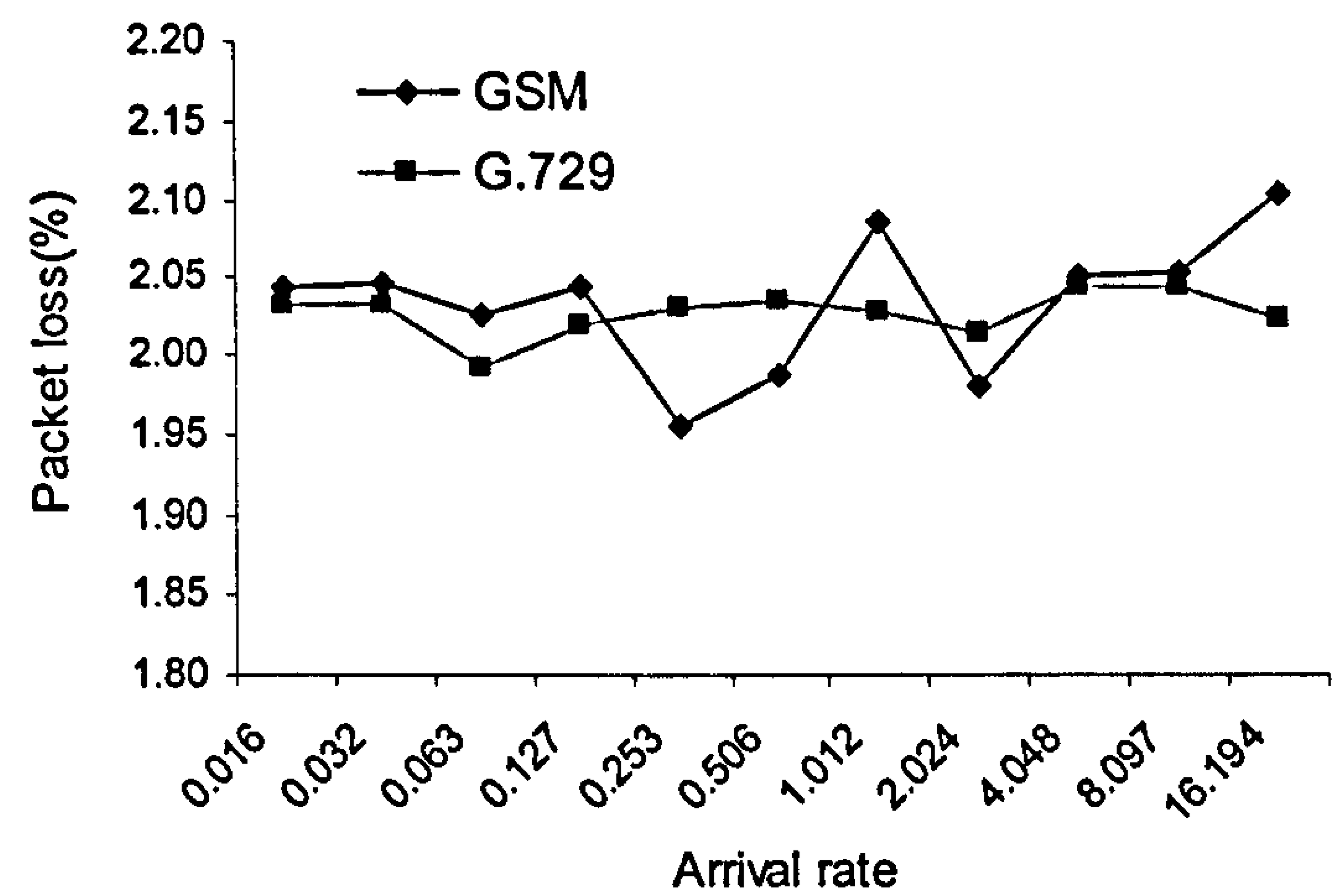


Figure 4.13: Packet Loss using Different Coders

Figure 4.14 illustrates the effect of increasing the number of users. The result shows that the average percentage of packet loss stays at about 2% as the number of users increase. However at certain number of users, it increases drastically.

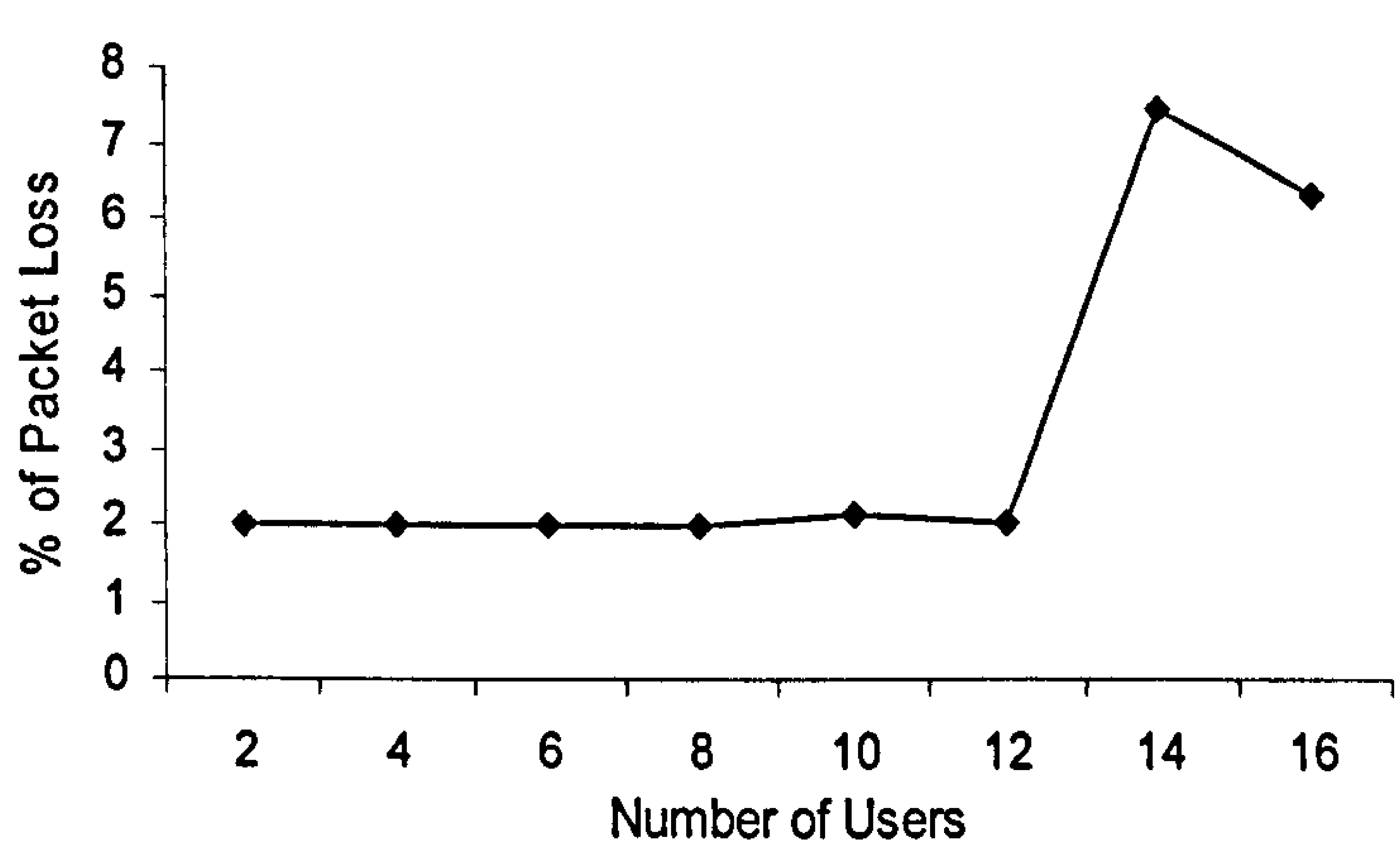


Figure 4.14: Packet Loss against Number of User

4.6 Conclusions

Delay is one of the factors that affect the performance of the VoIP. Since voice and data traffic can be integrated, the necessary infrastructure to provide both services is reduced. One of the ways to achieve higher quality is to select a specific coding system, monitor the system performance by reducing delay and packet loss. This chapter presented an analytical model for the performance of VoIP in a 3G network and the model was verified by the simulation result. The performance was measured in terms of the end-to-end delay, delay variance, traffic sent and traffic received. The model had been evaluated through extensive simulations for different value of traffic arrival rate with different types of coder and the results were collected. The influence of traffic on the voice in IP based network was analysed.

It was seen that the greater the jitter in an IP environment, the greater the packet loss which results in the lower voice quality. In addition, the greater the packet arrival rate the greater the percentage of packet loss. The end-to-end delay result showed that at the beginning of the graph, the simulation result match with the simulation result. However, the model can only support up to certain traffic and number of users. This was suspected due to the mobility pattern and assumptions that had been made. The stability of the system over a high rate of packet arrival is another challenge in assuring a high performance. Compression technique is one way of controlling the bandwidth however it introduces an additional overhead to the packet size. The results showed that types of encoder have a significant impact on the overall system performance. Different types of codec have different amounts of delay added to the compressed speech data. There is always a tradeoff between voice delay and system performance. However, deploying a suitable coder had proved to be effective in controlling the delay. From the results, it can be concluded that the end-to-end delay and jitter conform to the standard in certain condition only.

5.1 Introduction

The technology is moving towards convergence of services onto a single device, i.e. moving towards all IP based network. The 3G and the Internet are identified as the main drivers for the future generation mobile telecommunication market. Several concerns arise when Internet Protocol transport is introduced. Applications such as streaming and packet telephony will cause network congestion. The usage of voice and data services is growing in demand. This will affect the QoS of voice communication. When bandwidth is insufficient, different types of services, especially services with the demand of high service quality, need to be considered. In addition, the characteristic of this traffic need to be investigated. Thus, managing and optimizing the networks are necessary. This chapter discusses the influence of traffic aggregation on the system performance in IP based networks.

Mobile networks consist of a vast amount of network elements that should be monitored. These elements are controlled by a set of configuration parameters. Although packet voice can be delivered across network using the IP, lack of QoS guarantee is the crucial problem in the implementation of IP based application such as VoIP. Quality of service concerns with the process of delivering data packets in a reliable manner, which will include minimising data loss and reducing latency. It is necessary to guarantee a certain level of performance that can be achieved especially for real time applications as they may be delay sensitive.

Different users may need different QoS and a good QoS means that the probability that the requested level of performance will be provided is high. The network traffic and interference could cause delay and jitter. The proposed contribution involves investigation of QoS in intersystem handover such as packet loss, security and delay in data transfer that affect the performance. The QoS and security of the real time service are the major issues that need to be enhanced to the level of circuit switched networks. This includes the impact of mobility on QoS management.

As the number of users varies, mobility, call departures and call arrivals contributes various network management problems. New users should be admitted if the QoS of all the users can be guaranteed after admitting new call into the system. Otherwise, the new call request should be blocked. It is important to decide whether the new mobile should be accepted or rejected. Traffic management problem is necessary to guarantee the transmission quality of ongoing calls from dropping below a desired level. A good traffic management scheme should achieve high capacity utilization and guarantees QoS. The aim of traffic management is guarantee the QoS of the ongoing calls and at the same time manage the bandwidth efficiently. To satisfy users' requirements and guarantee high quality of service in the communication, users who are admitted to the system need to be given priority based on the application and needs. A heavy traffic scenario needs to be considered.

Various investigations on different aspects of IP based networks have been conducted [74-79]. An overview of VoIP and its transmission over mobile radio links with available packet-based access was explained in [74]. The UMTS all IP approach for 3G mobile systems was described as in [75] and it focused on the core network architecture. The concept of IP based networks and its architecture was explained where technical issues and challenges in VoIP were also discussed.

The limited upstream capability of 3G services is seen as a potential limit on VoIP over 3G data access. It can only deliver up to 64Kbps upstream which is however adequate for VoIP services [76]. Applying VoIP with mobile network involves technical issues as less attention has been paid to the application of VoIP in a mobile environment [77].

In recent mobile communications, the networks have been designed so that all services would use common facilities which result in efficiency and cost saving. The integration of voice as a component of multimedia IP based network allows a variety of services on a single infrastructure. Hence, traffic management is an issue as it involves managing real time and non-real time traffic [80].

The aggregation of VoIP calls was presented in [78]. It involved investigation of the relationship between the number of VoIP source, link rate and latency. In [30], mobility management issues regarding VoIP services were presented by comparing the Mobile IP and Session Initiation Protocol (SIP) approach. Similarly in [79], the author discussed mobility management issues to support VoIP services. Mobile IPv4 and low latency handover scheme was explained.

A scheme for different error and delay conditions was proposed in [81]. It was claimed that the scheme is efficient and robust for VoIP transmission in cellular environment. A scheduler that did not differentiate types of services was deployed in [73] to evaluate the performance aspects of streaming applications in a mixed streaming and "best effort" scenario. It was reported that reasonable streaming performance can be achieved by using fair scheduler.

5.2 QoS Classes

QoS is defined as the collective effect of service performance. It established a certain degree of users' satisfaction of a service. A large number of issues which affect user satisfaction with any network service are discussed in [82]. Different services may have different characteristics which require different of QoS. In order to satisfy the end-to-end service, a user is provided with certain QoS [83]. Therefore, the characteristic of this traffic needs to be investigated. There are four different traffic classes standardized by 3GPP and the classes have been defined to support a wide range of applications with different QoS requirements in UMTS [10].

5.2.1 Conversational Class

Conversational class is defined for real-time connection that requires low end-to-end delay. A number of applications require this class of service such as VoIP and Video telephony. To ensure a good quality video and audio conversation, the end-to-end delay has to be less than 400ms. In UMTS, the AMR codec is used for the speech codec.

5.2.2 Streaming Class

Multimedia streaming allows data to be transferred as well as processed as a steady and continuous stream. The Internet browser can start displaying the data without waiting for the entire file to be transmitted. This type of applications can accept more delay and jitter compared to the conversational services.

5.2.3 Interactive Class

Interactive class can be used for applications involving requesting data from remote equipment such as web browsing, server access, database retrieval and online games.

5.2.4 Background Class

Background class is used for data applications that do not require immediate action such as emails, Short Message Service (SMS), and downloading of data. The delay varies from a few seconds to a few minutes.

All the Internet applications such as Internet-browsing, emails, streaming video and audio, have different QoS requirements. Some of the applications are tolerant to packet errors, packet losses and delay while others are not. The main distinguishing factors among classes are the sensitivity to the delay required by each traffic classes. Table 5.1 shows how QoS requirements for different applications can be met using different QoS classes [3].

Table 5.1: Qualitative QoS Requirements for Different Applications

Error tolerant	Conversational voice and video	Voice messaging	Streaming audio and video	Fax
Error intolerant	Telnet, interactive games	E-commerce, WWW browsing	FTP, still image, paging	Email arrival notification
	Conversational (delay << 1sec)	Interactive (delay approx. 1sec)	Streaming (delay < 10sec)	Background (delay>10sec

Each UMTS QoS class can be described using the QoS attributes that have certain ranges. The fundamental characteristics of the classes and example applications are summarized in Table 5.2. Further details can be obtained from [39, 84]

Table 5.2: Types of Quality of Service

QoS Classes	Fundamental characteristics	Example of applications
a) conversational class	real-time connection, performed between human users, really low delay, nearly symmetric	voice
b) streaming class	real-time connection, transferring data as a steady and continuous, low delay, asymmetric	streaming video
c) interactive class	non-real-time packet data, response requested from other end-user, reasonable round-trip delay	web browsing
d) background class	non-real-time packet data, no immediate action expected, less sensitive to delivery time	emails

The conversational class has the most stringent QoS requirements. It is the most delay sensitive since it cannot tolerate with long delays. In contrast, the background class has very flexible QoS requirements in terms of delay and throughput. This class is the least delay sensitive. Table 5.3 gives examples of attributes and value ranges for UMTS QoS requirements [85]. The performance of a mobile station receiver is measured in terms of Frame Error Rate (FER), i.e. the percentage of frame that contains errors relative to the total number of frame received.

Table 5.3: UMTS QoS Requirements

Medium	Application	Degree of symmetry	Data rate	Key performance parameters and target values		
				End-to-end one way delay	Delay variation within call	Information loss
Audio	Conversation voice	Two-way	4-25Kb/s	<150ms preferred <400ms limit	<1ms	<3%FER
Video	Video phone	Two-way	32-384Kb/s	<150ms preferred <400ms limit		<1%FER
Data	Telemetry two-way control	Two-way	<28.8 Kb/s	<250ms	N/A	Zero
Data	Interactive games	Two-way	<1Kbps	<250ms	N/A	Zero
Data	Telnet	Two-way asymmetric	<1Kbps	<250ms	N/A	Zero

The QoS can be determined based on the following parameters.

a) Throughput: Throughput is the amount of the successful packet move and delivered from one place to another in a certain period of time. It is an important factor of QoS and it reflects the bandwidth of a network. To ensure the QoS, the networks should attempt to maximize service availability and throughput while minimizing the remaining measurements at the same time. The main goal of QoS is to reduce or eliminate delay of voice packets including packet loss that travels across a network. It can be defined as the effectiveness of a network to support better service.

b) VoIP Delay Elements: Delay is described as the total time for a packet to get to destination. It is the length of time taken to transmit data from one point to another. It is inherent in voice networking and it is caused by a number of different factors. The lower the delay, the better communication quality will be. Queuing delay is only one component of end-to-end delay and the delay is affected is through jitter.

VoIP communication encounters a few elements of delay such as: [68]

1. Coder delay : Analog-to-digital speech conversion and compression
2. Packetization delay : Time to fill a packet payload
3. Serialization delay : Time to push a packet payload onto the transmission link
4. Output queuing : Scheduling a voice packet out of devices queues
5. Network delay : Transmission delay across the network
6. Dejitter delay : Smoothing the inter-arrival time of voice packets

The average end-to-end delay, D_{Ave} for voice communication comprises of the following:

$$D_{Ave} = D_{Comp} + D_{Proc} + D_{Wait} \tag{5.1}$$

where D_{Comp} is the compression time; D_{Proc} is the processing time; and D_{Wait} is the waiting time.

Since voice traffic is a real-time traffic, speech will be unrecognizable if there is a long delay in voice packet delivery. If delay falls below an unacceptable level, it will degrade the voice quality. An acceptable level of end-to-end delay is less than 300 milliseconds. The ITU-T G.114 recommendation suggests that for good voice quality, the end-to-end delay should be less than 150 milliseconds (ms) [68].

c) Jitter for VoIP: The voice packet can be delayed as packets are not arrived at the same regular interval. Jitter is an issue that remains in packet based network. Jitter is variation in the delay of arrivals among similar packets traversing the same path in the network. It is the difference between expected and received time of the packets. Too much jitter could cause problem to voice quality as the end-to-end delay could be affected through jitter.

Each block of digitized voice signal arrives with a slightly different delay. As jitter occurs continuously, this causes a discontinuity of the voice stream and it will be very irritating to the talker and listener of the conversation. Let t_1 and t_2 be the time the two consecutive packets leave the source. The packets are played back at the destination node at time t_3 and t_4. Therefore,

$$\text{jitter} = (t_4 - t_3) - (t_2 - t_1) \tag{5.2}$$

d) Voice Packet Loss: The packets that are delayed must be discarded. Packet loss occurs when packets do not arrive or arrive too late to be processed and it is seen as gaps in the communication. IP packets may be lost in a highly congested network or a long distance call. Packet loss is measured based on the percentage of packet lost in the transmission. It is the ratio of packets received to packets sent.

Different application has different tolerance of packet loss. Packet loss is intolerant in VoIP as it may degrade the quality of service in VoIP. A packet loss of 2% is normally acceptable. However, if the losses are more than that, users' satisfaction will not be achieved. According to [72], the voice quality is intolerable if packet loss is more than 3%. Congestion is the main cause of packet loss. Therefore, congestion management and avoidance is crucial.

5.3 Managing Traffic Mix

Mobility will affect the overall system performance and for voice communication, maintaining the voice quality is necessary. When more users are connected to a system, the end-to-end delay will not be guaranteed. Handover which is based on incoming traffics may cause the system to be unbalanced. Emphasizes have been given in optimizing handover as it relates to dropped calls and network overload. Besides that, ability of a cellular network to perform efficient handover is crucial because irrelevant and improper handover are expected to happen.

As users move from one location to another, the QoS will be affected. The end-to-end delay must be kept at a certain level. Expanding services through the use of multiple applications create a challenge of managing traffic mix in IP based networks. Increasing the number of users to the system is not the primary aim as the ultimate goal is to provide high quality services to users. It is important to manage the traffic particularly during heavy traffic.

A good way of classifying different traffic is needed since there are a variety of applications, protocols and users in the network. It is assumed that two types of service present in which one of the service need high quality of service. One way of managing traffic mix is to classify the traffic based on the information types and not based on the time the packet arrives. To support users' communication, choosing the most effective traffic management schemes which can protect users' communication is necessary. Effective traffic management needs to be implemented and be able to deal with the problem of different services sharing the same resource.

To efficiently utilize radio resources, a scheduling scheme based on weight types and a load balancing technique were introduced in this section. These schemes were adopted for the two services. The aims of the schemes are to ensure the voice quality of every user is above the target quality and to make the end-to-end delay as low as possible.

The scheme involves managing traffic mix based on priority and type of packets. In the scheduling scheme, traffics are prioritised and priority is given to real time applications. The system will use the scheme to select types of packet and adjust the weight for the scheduling. Since different QoS are required for different type of services, all the traffic classes can get the services as required.

To optimize the network resources and improve the performance, a load balancing scheme is implemented at the base station controller for handling traffic mix in IP based network. The base station controller determines if a user's QoS can be satisfied and decides if it can be admitted into the base station. The types of traffic will be detected and the service time will be computed. The base station controller will decide the handover to another base station based on traffic types instead of incoming traffic.

The traffic loads were distributed based on the packet sizes in which handover priority will be based on the types. The handover process will be based on the handover algorithm parameter as described in [55]. Initial traffic load will be set and it will be changed after a new call is admitted to the system.

The signal strength of a new call must be above the target threshold in order to be admitted to the system otherwise, it will be denied. When traffic arrived to the IP network, the controller will decide whether to allow the call admission or reject a user with different traffic classes to the system based on the QoS that can be satisfied. It is assumed that all users are active in the system. The decision is made by comparing the traffic load with system capacity.

The call admission control schemes will allow new users to be admitted at the admission stage if the current load in the system has not exceeded the capacity, otherwise, new users will be denied. The maximum number of active users can be calculated based on the theoretical model presented in section 4.4. A new call will be admitted if the current traffic load is below the system capacity.

Assume that there are *E[N]* mobiles and the distribution of users is :

$$E[N] = E[C] + E[D] + E[H] \tag{5.3}$$

where E[C] is using conversational service; *E[D]* is using data service; and *E[H]* is new and handover users;

Traffic models which correspond to the conversational and background QoS classes are considered in the study. The conversational class is treated as real-time services while the background class is treated as non real-time service. It is assumed that all the active users in the system are uniformly randomly distributed. The traffic model depends on the service profile. The packet arrival is assumed to be according to Poisson process while the call duration is exponentially distributed.

Let λ represents the aggregate traffic in a queue and the packets are serviced at constant size. Let μ represents the average service rate and it is calculated based on the data rate as well as the average file size. Let π_1 and π_2 represent the voice and data in the queue. The packet arrival rate for voice and data are represented by λ_v and λ_d. Similarly, service rate for voice and data are represented by μ_v and μ_d respectively.

The average active number of users is equal to $(\lambda / \mu)(1 - P_{vd})$ [70]. It is assumed that:

a) The number of active voice calls is given by: λ_v / μ_v.

b) The number of active data calls is given by: λ_d / μ_d.

c) Traffic channels are divided into two distinct portions. N_1 channels are assigned for voice service and N_2 channels are assigned for data.

The blocking probability can be obtained from the Erlang B formula [45, 86, 87]. The blocking probability of voice and video can be obtained by the following formula.

The blocking probability for voice service is:

$$P_v = \frac{(\lambda_v / \mu_v)^{N_1} / N_1!}{\sum_{k_1=0}^{N_1} (\lambda_v / \mu_v)^{k_1} / k_1!} \tag{5.4}$$

Similarly, the blocking probability for data service is:

$$P_d = \frac{(\lambda_d / \mu_d)^{N_2} / N_2!}{\sum_{k_2=0}^{N_2} (\lambda_d / \mu_d)^{k_2} / k_2!} \tag{5.5}$$

The total blocking probability is given by:

$$P_{vd} = P_v + P_d \tag{5.6}$$

Utilisation of the network can be represented by the ratio of the channel occupied by users to the total resources or channel offered in the system. The network utilisation for voice users can be represented by:

$$\rho = (\lambda_1 / \mu_1)(1 - P_v) \tag{5.7}$$

Similarly, the network utilisation for data users can be represented by:

$$\rho = (\lambda_2 / \mu_2)(1 - P_d) \tag{5.8}$$

It is necessary that

$$\rho < 1 \tag{5.9}$$

If $\rho > 1$, it is clear that the system is unable to cope satisfactorily with the tasks.

Thus, a very long queue will be developed. The delay can be presented as follows:

$$D = \Sigma W_i / RW_i \tag{5.10}$$

where Wi is the weight assigned to queue i and R is the transmission rate.

Handover delay in cellular systems was analysed based on [88]. The signal levels received from two base stations are separated by a distance of $L\,(Km)$ at point P.

The mean signal levels can be derived as

$$BS_{Lev1}(d) = K_1 - K_2 \log_{20} d \tag{5.11}$$

$$BS_{Lev2}(d) = K_1 - K_2 \log_{20}(X - d) \tag{5.12}$$

where K_1 and K_2 represent path losses, BS_{Lev1} and BS_{Lev2} represent signal levels at the base stations, d represents the distance between $User1$ and $BS1$, and X represents the distance between $User2$ and $BS1$.

Handover occurs when the difference between the signals exceed certain level. For $i \in \{0,1\}$ which refer to the signal received at BS1 and BS2 respectively, handover is triggered when

$$BS_{Lev2} - BS_{Lev1} = K_2 \log_{20} \frac{d}{X - d} \tag{5.13}$$

The delay in handover decision can be calculated as

$$Delay, \hbar = \frac{d - X/2}{v} \tag{5.14}$$

where v represents speed of the mobile.

5.4 Simulation Model for Mixed Traffic Environment

A model with homogeneous cell was considered to analyse the impact of traffic mix and handover on voice quality. The coverage area was partitioned into cells of equal size where a fixed number of channels were assigned. The analysis was performed by running a few mobile nodes between a few base stations. The users are roaming between base stations that cause handover to occur a few times. It is assumed that the users are travelling along the predefined path at a constant speed that causes handover to occur a few times. The users can move between the base stations with a velocity of 50km/hr.

The devices were configured so that the system consists of a mixture of voice and data service users. A half of the mobile users were configured as voice users with VoIP in which they are communicating with the other mobile users using the IP. The other users were configured for non real time application. The traffic rates for both applications were increased and the impact on the packet call quality constraints in multi services networks was analysed.

A model consists of *M* voice traffic and *N* data traffic. $Aj,k(n)$ is the arrival process of voice j which is destined to output k at rate $\lambda j,k$. Similarly, $As,t(n)$ is the arrival process of data s which is destined to output t at rate $\lambda s,t$. Hence, $Aj(n)$ is the aggregate process of voice arrivals to input j at rate :

$$\lambda_j = \sum_{j=1}^{N} \lambda_{j,k} \tag{5.15}$$

Similarly, $As(n)$ is the aggregate process of data arrivals to input s at rate :

$$\lambda_s = \sum_{t=1}^{N} \lambda_{s,t} \tag{5.16}$$

A scheduling sequence according to weight will be generated to handle users with different QoS. Each user will be assigned a weight that indicates the priority. Users with higher weight receive faster response from the BTS than the one with less weight. The voice and data are assigned a weight denoted by wj, and ws respectively where $wj > ws$.

The scheduling produces better performance compared to the traditional queuing system. However, it may result in improper load balance among the base station if the voice requests are very highly. There is a possibility that the request may be handover to the same base station. Thus, a load balancing technique is used to reserve channel in advance for voice users that are predicted to make handover. The flowchart for balancing the load is shown in Figure 5.1.

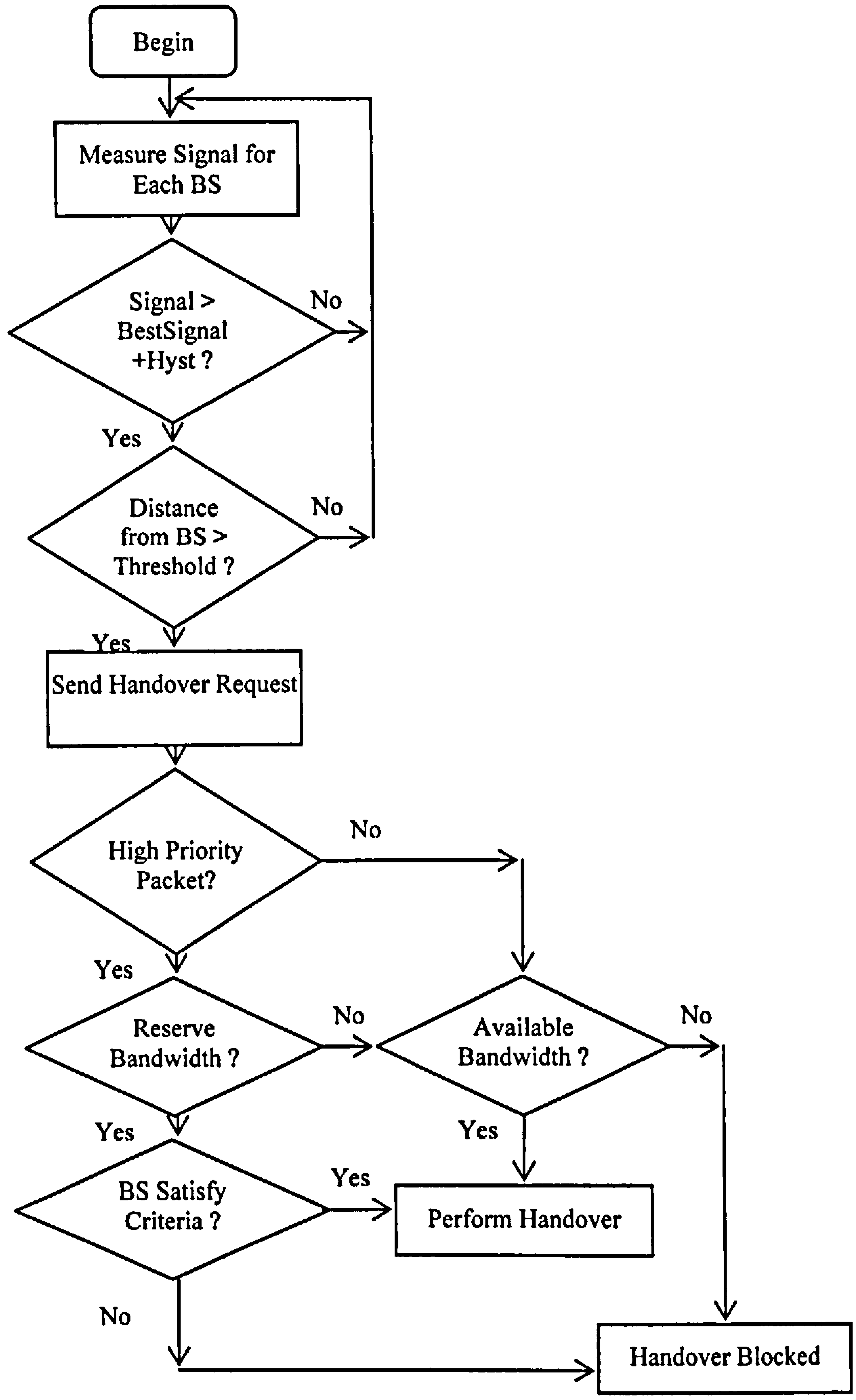


Figure 5.1: A Flowchart of prioritising handover

The input parameters involve traffic mix, handover mechanism and radio resource management strategy. The model produced variety of measurements such as packet end-to-end delay, throughput, packet loss rate, handover frequency and system capacity. To get the best results, the simulation was run extensively with a different number of users and different packet arrival rate. The parameters values that are used in the simulation are depicted in Table 5.4.

Table 5.4: System Parameters for simulation

Parameters	Value
Mobility	50km/hr
Shadow fading standard deviation	10dB
Maximum queuing time allowed	10s
Talkspurt	Exp(0.35)s
Silence	Exp(0.65)s
Frame size	20 ms
DSP Processing ratio	1
Data Rate	12.2 Kbps
Packet Size	32.5 bytes/packet
Voice frame per packet	1

5.5 Result Analysis for Voice in TrafficMix

The statistics for voice performance in the mixed traffic environment were collected and the results were compared with the performance obtained by deploying the schemes. Each simulation result is used as a data point on the graph. Figure 5.2 to Figure 5.13 show the impact of scheduling and load balancing technique on the system performance [89-91].

5.5.1 Packet end-to-end delay

A high packet end-to-end delay will introduce significant problems such as reduced call quality. Hence, minimizing one way end-to-end delay is a very important consideration. To ensure that voice application would not be significantly affected by delays, it is recommended that the delay should be kept below 150ms [68]. Figure 5.2 shows the impact of logarithmically varying arrival rate on end-to-end delay. The graph shows that at a low traffic rate, the end-to-end delay for voice is slightly high. As the traffic increases, the delay increases drastically. The increase in the delay is due to user movement between base stations and network congestion.

Figure 5.3 and Figure 5.4 illustrate the impact of increasing the number of users on packet end-to-end delay by deploying scheduling scheme and load balancing technique respectively. It can be seen that when the number of users or traffic rate increases, the end-to-end delay increases too. The result is compared with the theoretical model based on [72] and it match closely to the theoretical results at the beginning of the graph. However, when the number of users is more than 32, the delay will become unpredictable and the delay increase drastically. This is because in theory when $\rho > 1$, the systems become unstable as the traffic rate is greater than the service rate. By deploying the scheduling scheme or load balancing method, the graph shows that the packet end-to-end delay maintains at about the same value which is less than 0.4 seconds. The results show that the delay can be further reduced by deploying the load balancing technique.

5.5.2 Jitter for Mixed Traffic

Figure 5.5 presents the effect of traffic mix on the voice jitter when traffic arrival rate increase. It seems that at a low traffic rate, the jitter is lower. As the traffic increases, it seems to have no effect on jitter. Figure 5.6 and Figure 5.7 show the impact deploying scheduling scheme and load balancing technique respectively. When the number of users increases, the jitter maintain at a low value and it increases sharply at a certain point on the graph. However, as the number of user increases, the jitter increases too. By deploying the schemes, the jitter can be reduced. The load balancing scheme can further reduce the jitter.

5.5.3 Packet loss for Mixed Traffic

The number of packet dropped and corrupted are some of the cause of a high packet loss rate. The effect of traffic mix on the packet loss of voice packet was studied as shown in Figure 5.8. The graph shows that as the traffic increase, the average packet loss is also increase. The average packet loss maintains at about 2%. However, the packet loss rate increases drastically at a certain traffic rate.

Figure 5.9 and Figure 5.10 show the impact of increasing the number of users on the percentage of packet loss. The graph shows that after a certain number of users, the average percentage of packet loss starts to increase drastically. It is noted that by applying the schemes, the percentage of packet loss maintains at about 2%.

5.5.4 Throughput for Mixed Traffic

Throughput can be defined as the number of successful packet received per unit time. Figure 5.11 shows the impact of increasing the traffic rates on throughput. Similarly, Figure 5.12 and Figure 5.13 illustrate the impact of increasing the number of users on throughput. The results show that the throughput maintains at about the same value as the traffic rate or the number of user increases. At a certain point, it appears to decrease. By deploying the schemes, the throughput seems to only increase slightly.

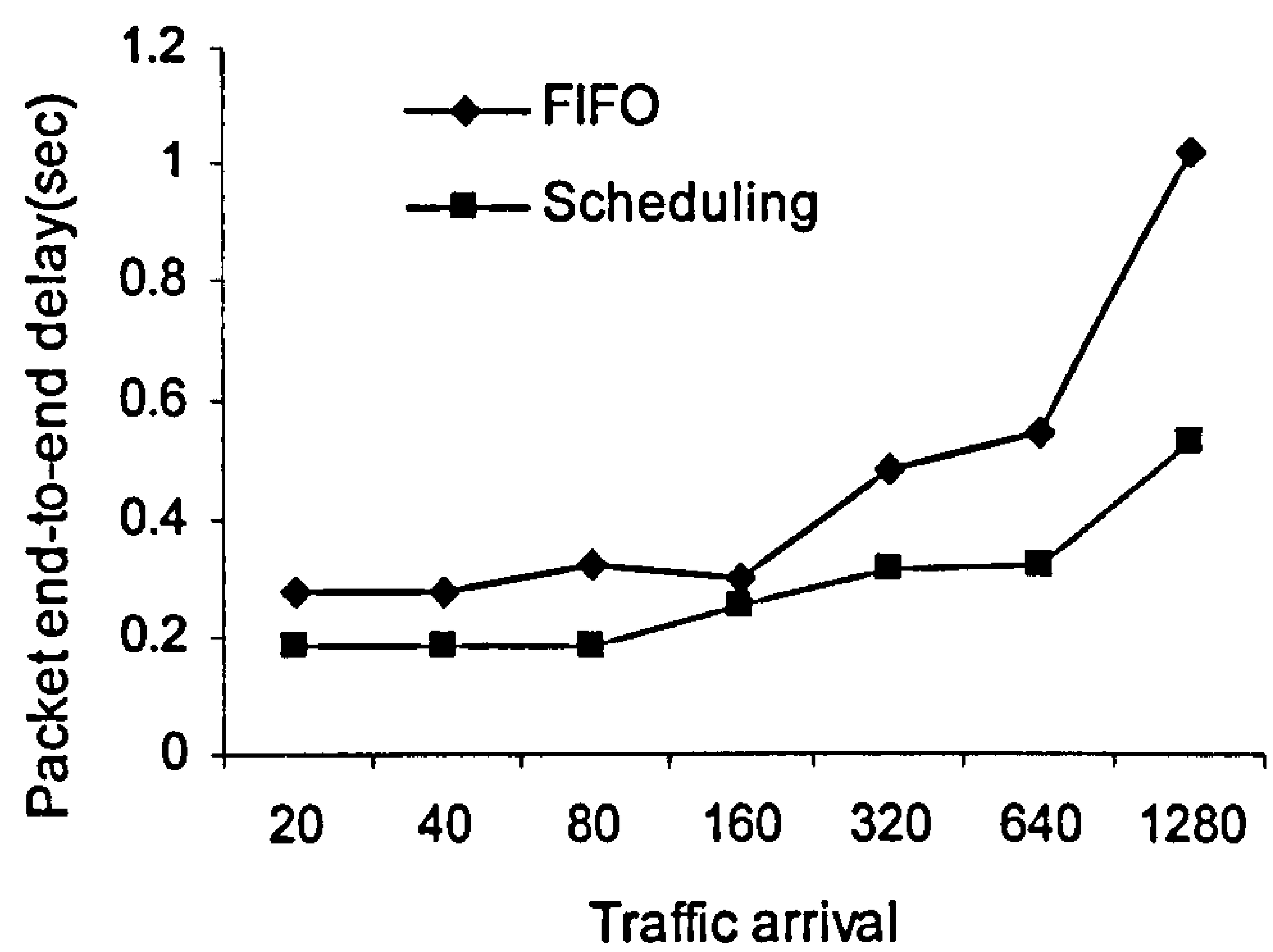


Figure 5.2: Packet End-to-End Delay using Scheduling Scheme

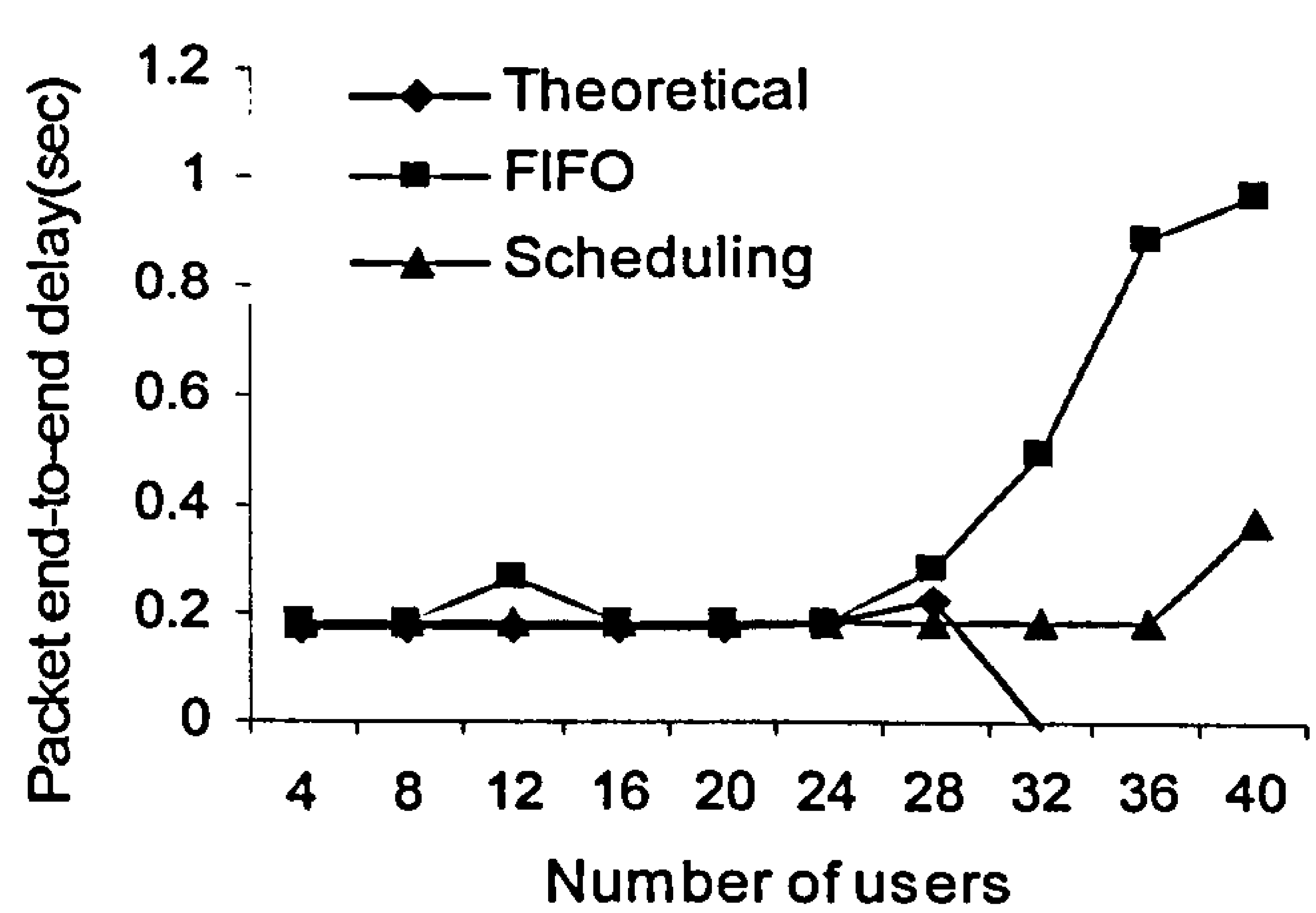


Figure 5.3: Packet End-to-End Delay using Scheduling Scheme

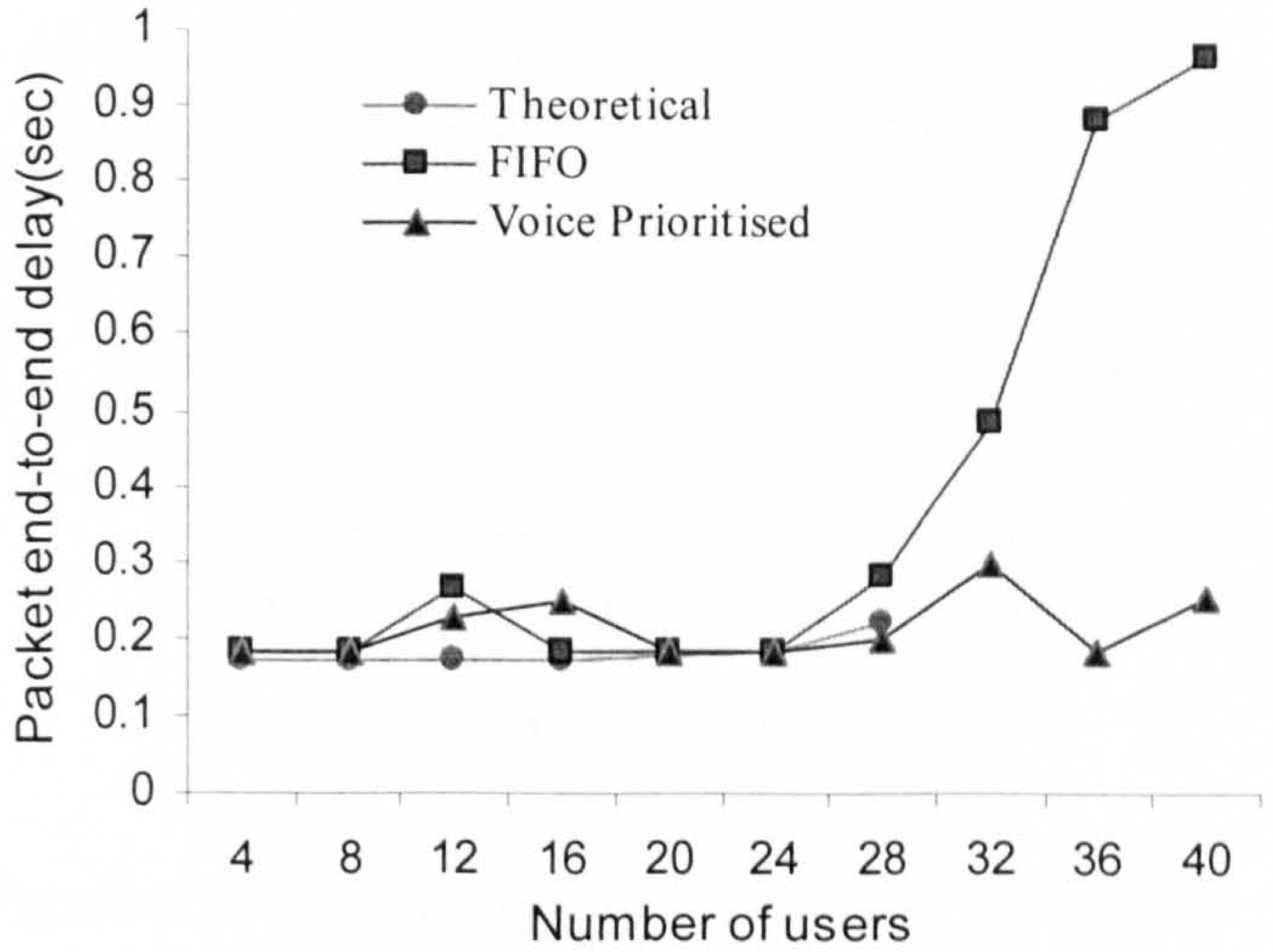


Figure 5.4: Packet End-to-end Delay using Load Balancing Technique

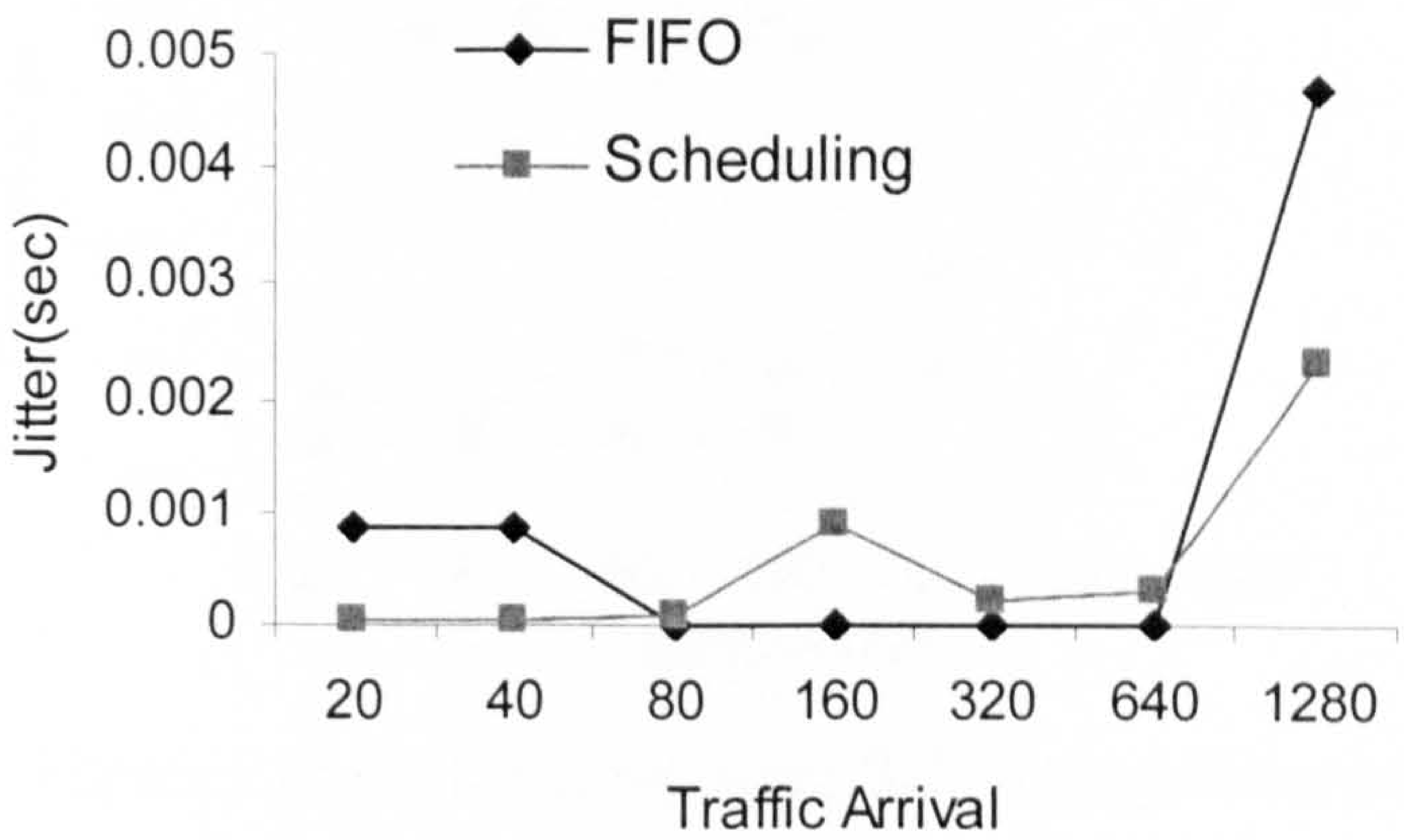


Figure 5.5: Jitter using Scheduling Scheme

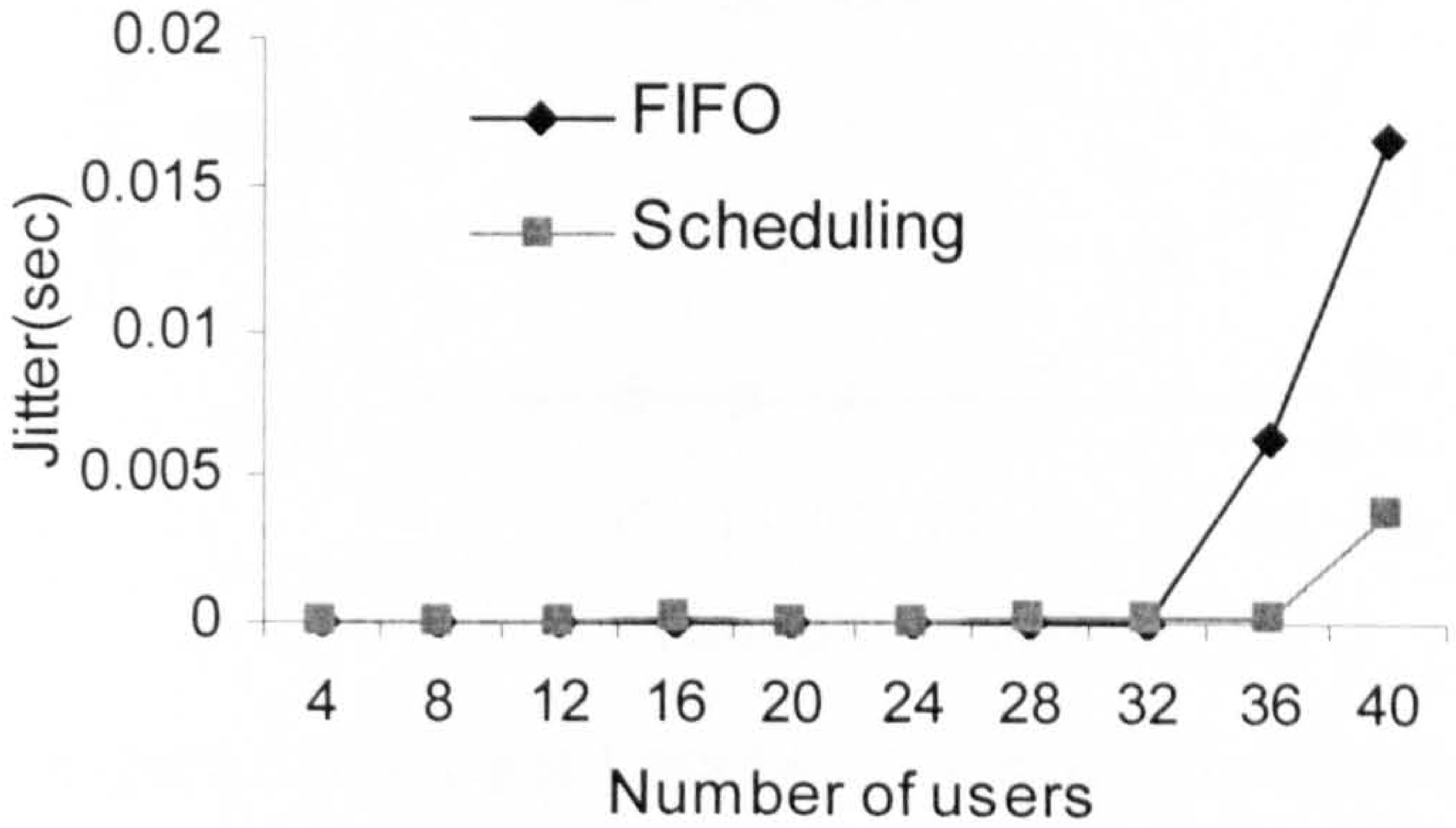


Figure 5.6: Jitter using Scheduling Scheme

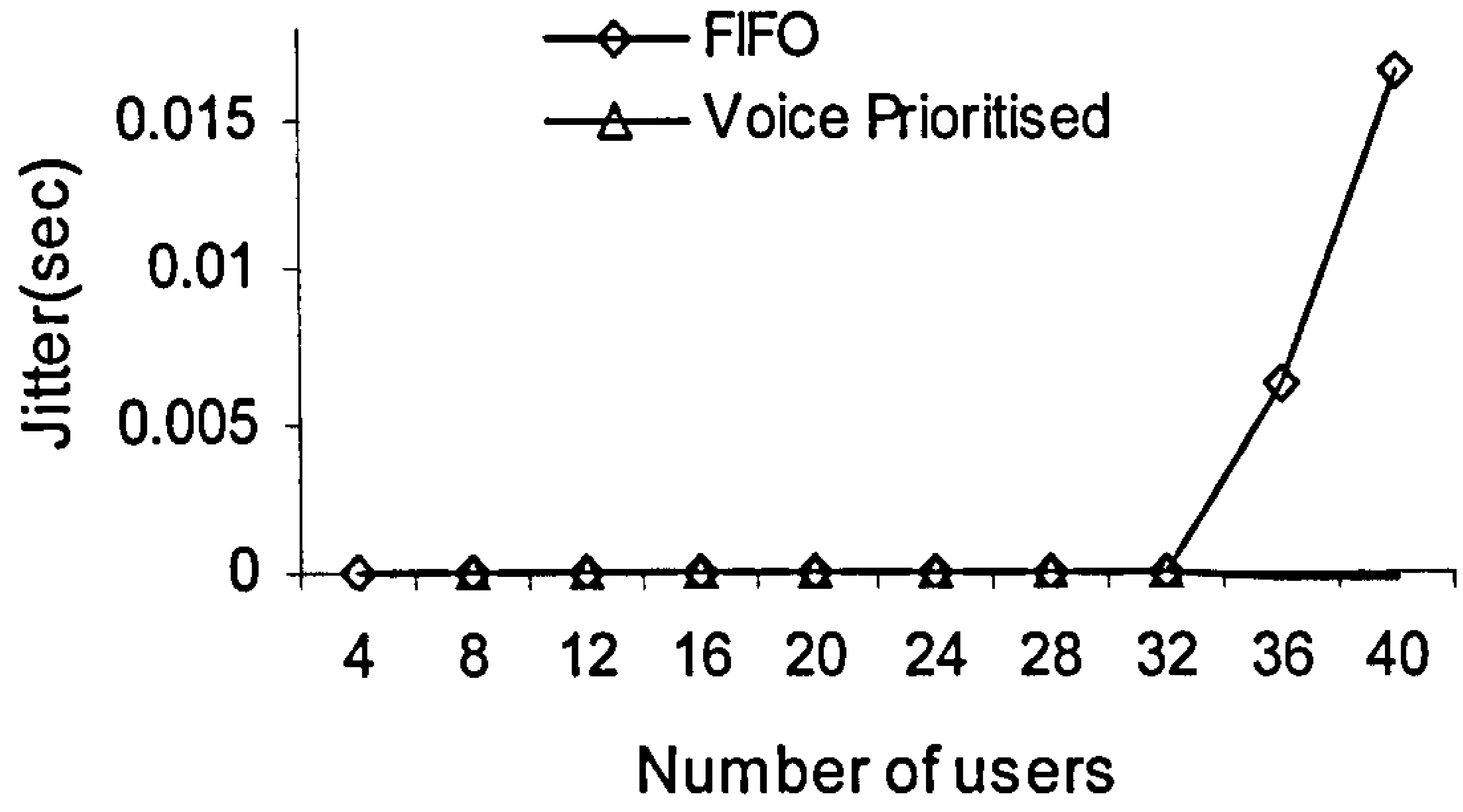


Figure 5.7: Jitter using Load Balancing Technique

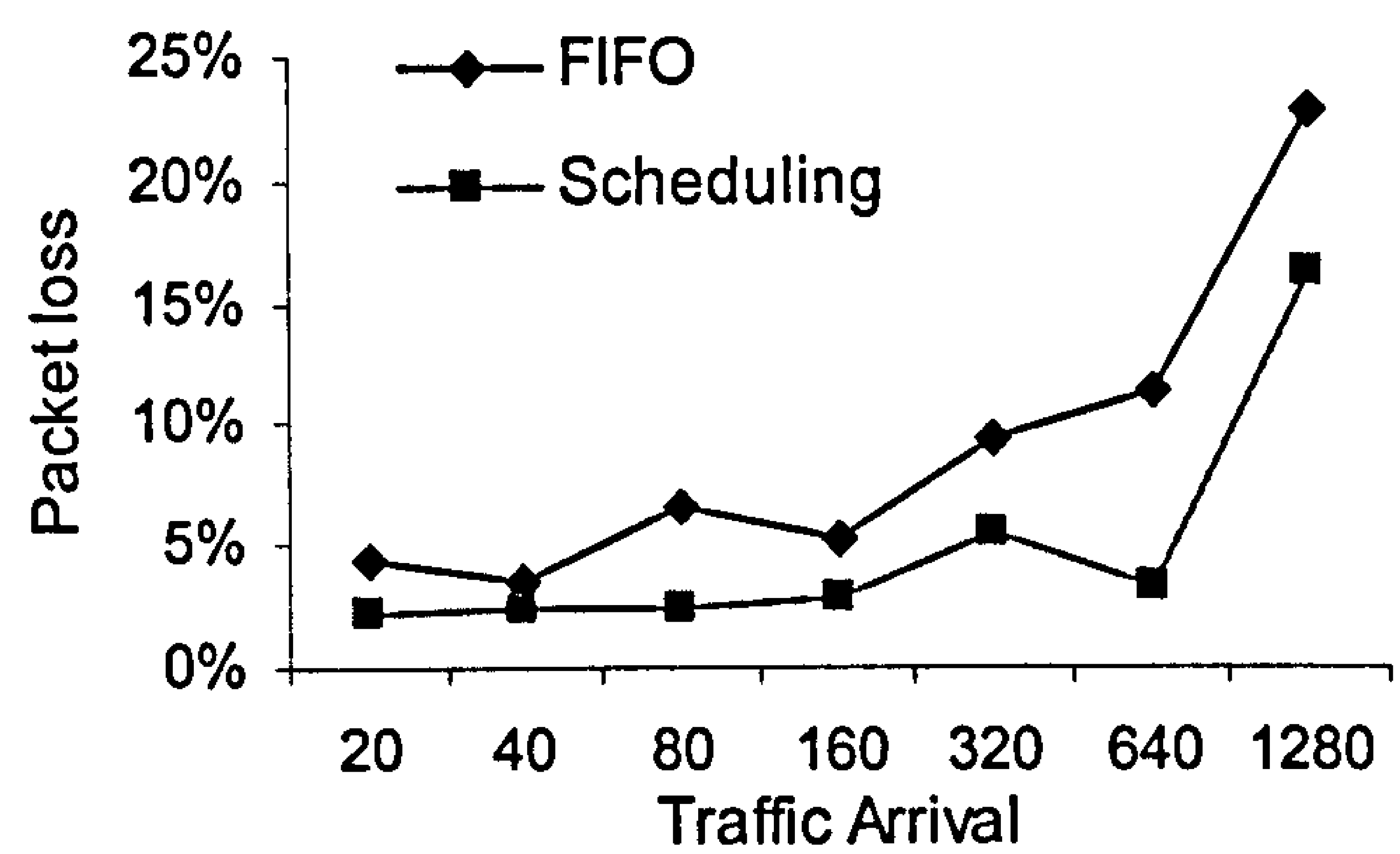


Figure 5.8: Packet Loss using Scheduling Scheme

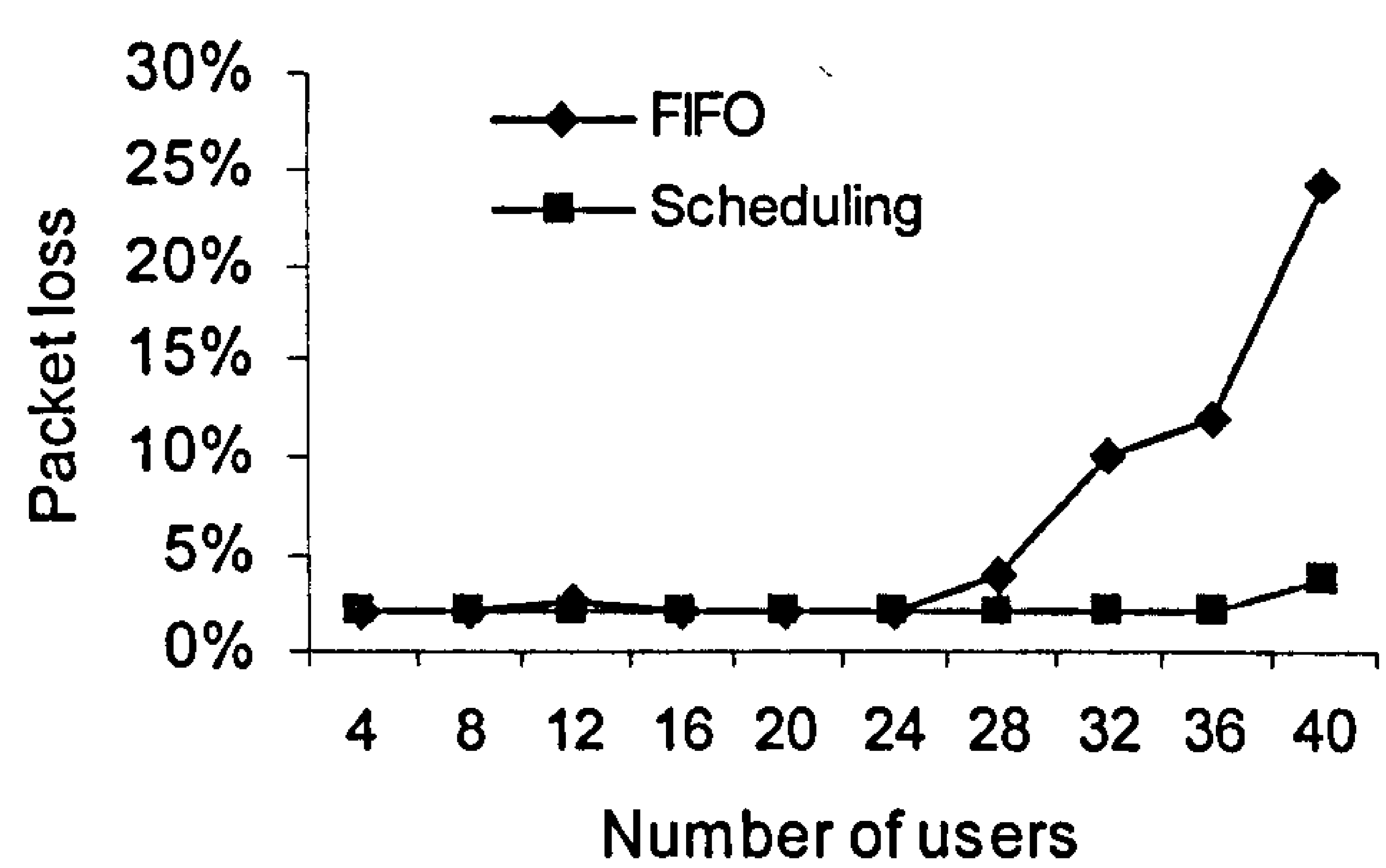


Figure 5.9: Packet Loss using Scheduling Scheme

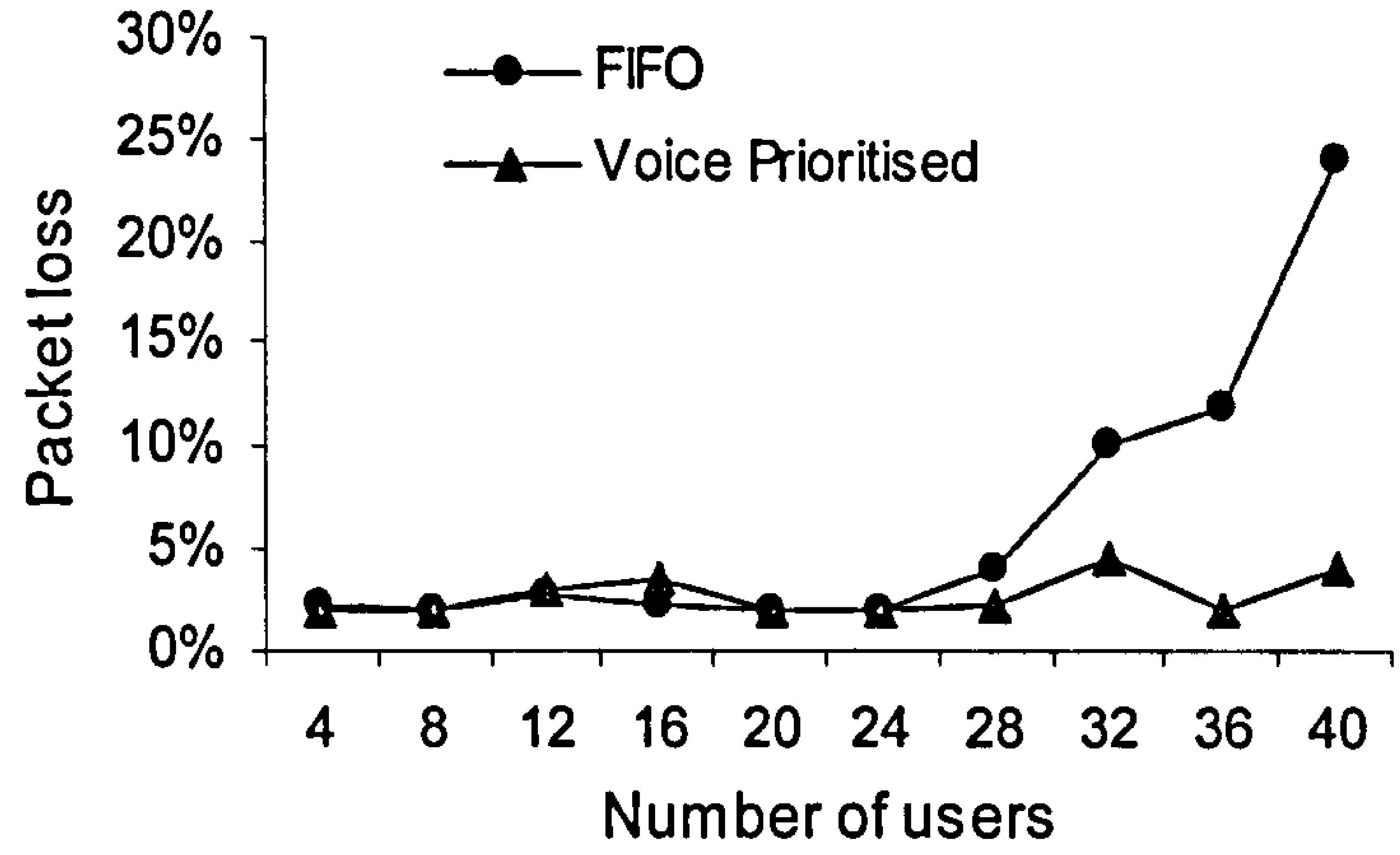


Figure 5.10: Packet Loss using Load Balancing Technique

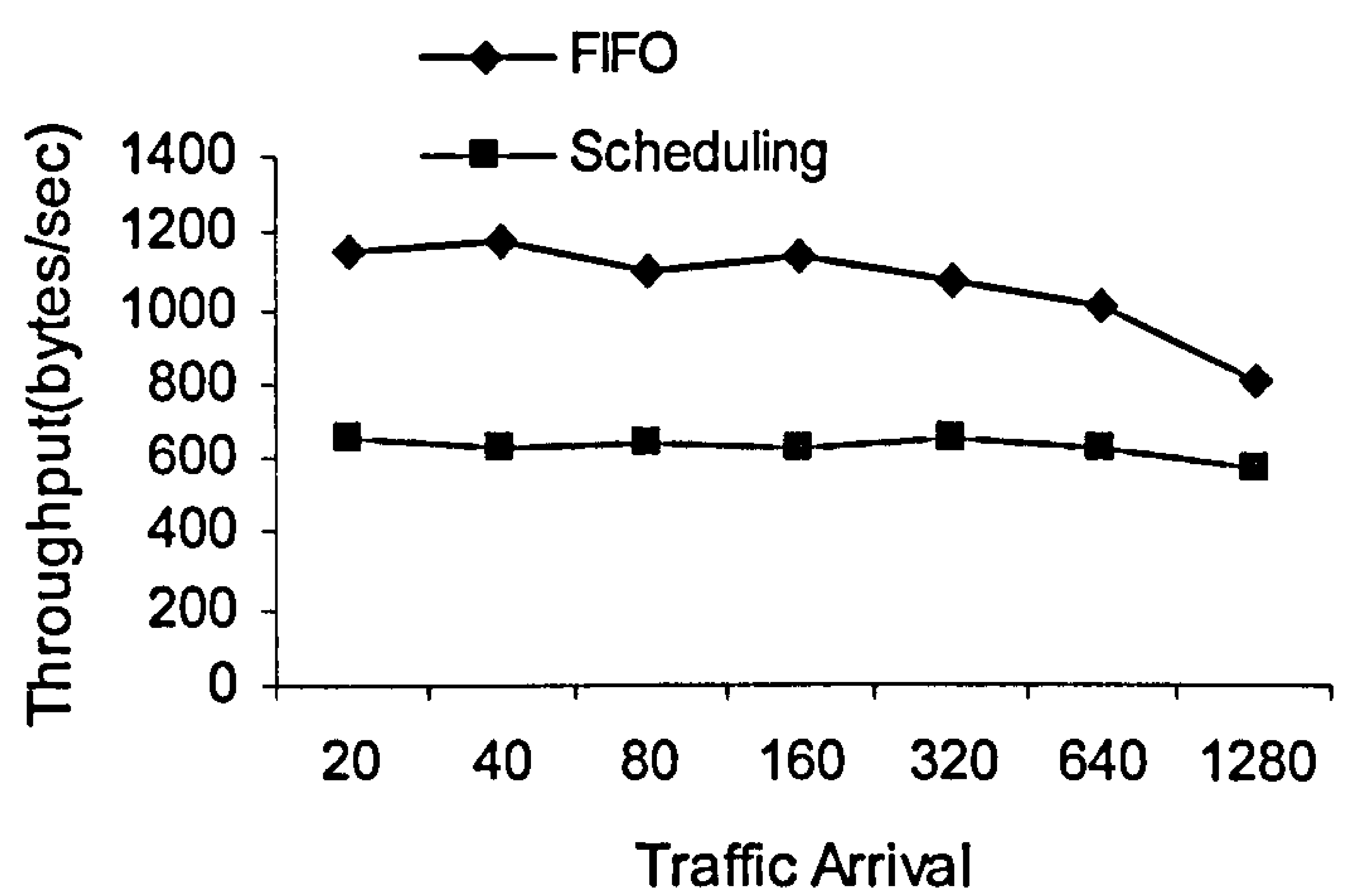


Figure 5.11: Throughput using Scheduling Scheme

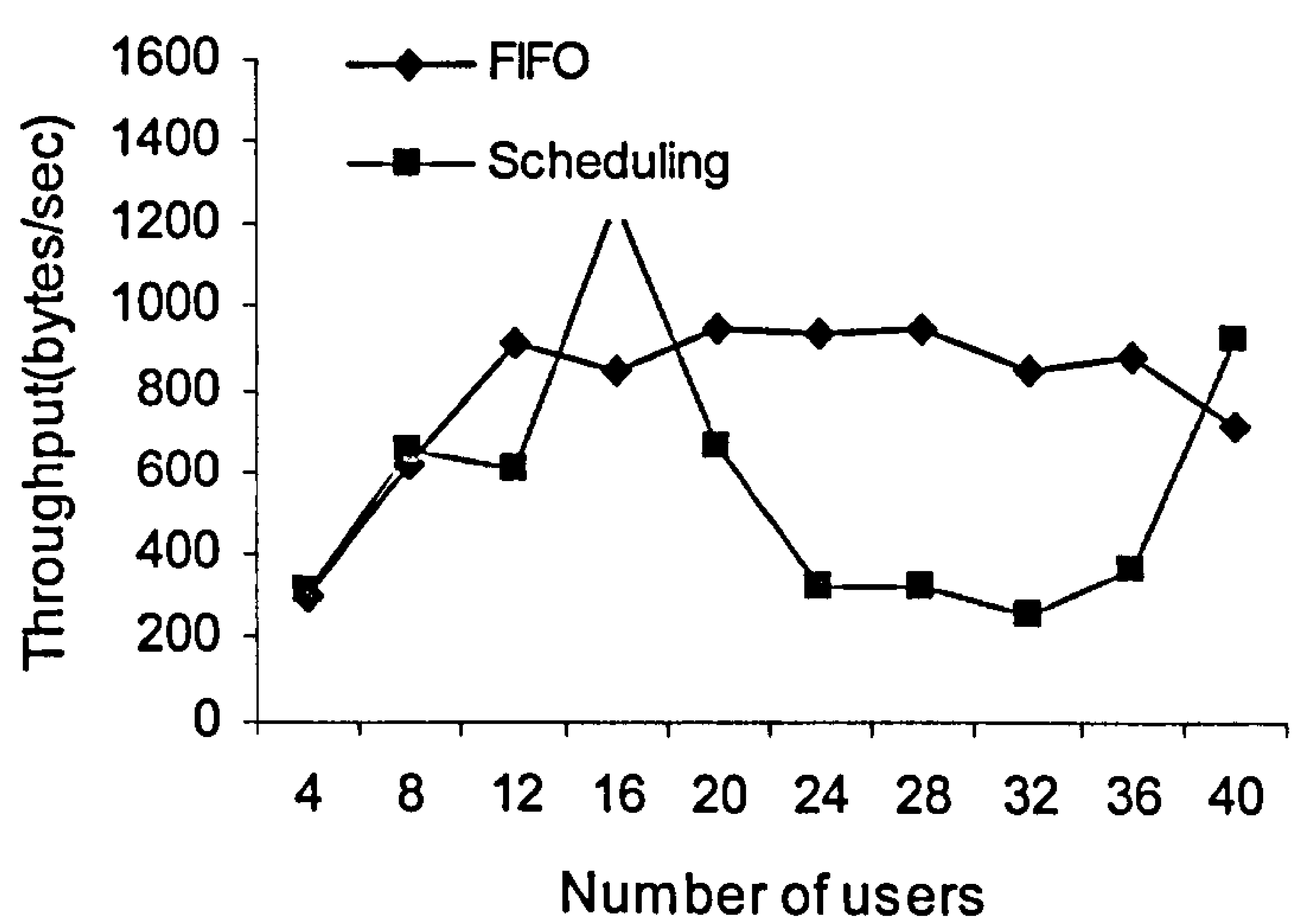


Figure 5.12: Throughput using Scheduling Scheme

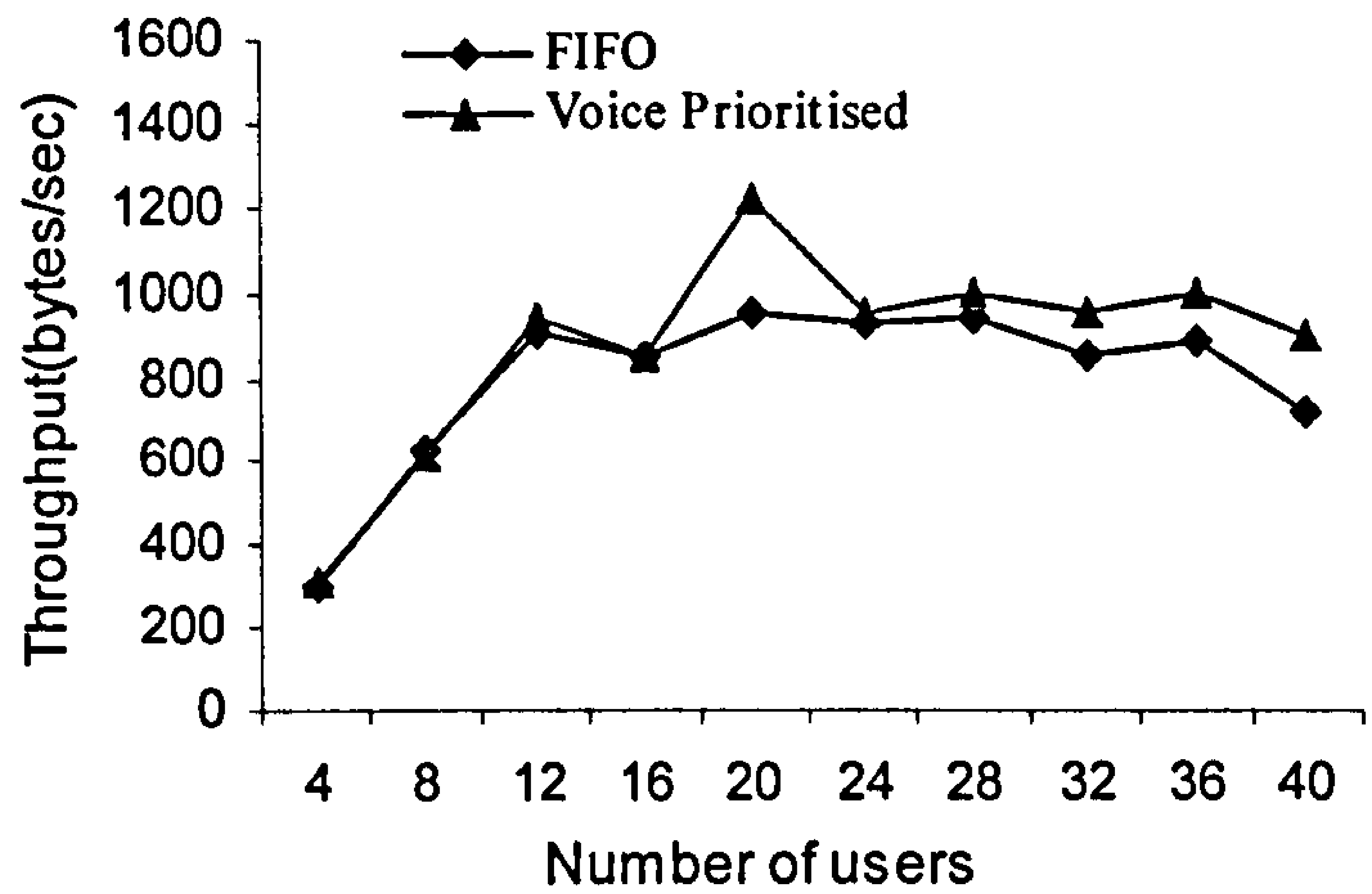


Figure 5.13: Throughput using Load Balancing Technique

5.6 Conclusion

This chapter presented the simulation results of traffic assessment based on the 3G traffic model that incorporates packetized voice traffic. As the number of users increase, the system performance decreases and the possibility of having unbalanced system could be increase due to the handover. Consequently, the system becomes instable if user mobility had not been managed properly. There are a variety of applications, protocols and users in the network. Hence, in this study, traffics are classified based on the types. A few schemes were implemented in the studies such as soft handover, predictive reservation scheme as well as prioritised queuing.

A scheduling based on weight was deployed to improve quality of voice. The results showed that the scheme can considerable improve the system performance while maintaining the end-to-end delay at the appropriate level. In addition, when the load balancing technique was deployed at the base station, the results showed that it considerably improved the system resource utilization, and reduced the delay, jitter as well as packet loss. The scheme seems to be an effective way of managing the traffic. The results indicated that a significant capacity improvement can be achieved. The system was able to support more number of connected users.

From the simulation, it can be seen that the schemes are practical and easy to implement. The theoretical results closely approximate the simulation and it showed that the system is capable of handling a certain number of users at a time. It seems that the system will be in a more stable condition despite the increase number of users in a mixed traffic environment. In conclusion, technology had improved the reliability of the system and voice quality over time. The technology will continue to improve VoIP performance as time goes on.

Chapter 6

Impact of Security Protocol on System Performance

6.1 Introduction

Security is another important issue that needs to be addressed and in some cases voice security could be more important compared to data networks. As protecting sensitive and financial data are crucial, ensuring the security of a conversation is desired. The popularity of IP applications had introduced serious security vulnerabilities. VoIP is one of the trends in telecommunications that introduces both opportunities and problems. With the introduction of VoIP, the need for security is necessary because data and voice must be protected. Securing the voice communication includes protecting the voice data and the identities of both parties. This chapter discusses the impact of implementing a security protocols on the performance of VoIP networks.

For a congested network, integrating a VoIP system would be a disaster to the infrastructure. Although it offers lower cost, it presents significant security challenges. Appropriate security measure is another issue in VoIP. Although VoIP serve the same purpose, the architecture is different from traditional telephony. In VoIP, voice traffic travels using the Internet protocol instead of using a conventional line. The speech is sampled and encoded before it is transferred over the network. IP telephony networks require little or no authentication to gain access. Hence, the security risks could be more complex and it may also introduce a higher perspective of attack due to the transmission voice data over the Internet.

A secure communication network provides: [92]

- Confidentiality ,
- Integrity,
- Authentication,
- Non-repudiation ,
- Service Reliability,

6.1.1 GSM and UMTS Security

GSM security is necessary to provide privacy and to ensure there is no interruption during the conversation. The security issues and the detailed description are discussed in [3, 8, 11, 37, 92]. The users will be identified and authenticated to ensure their privacy. They will not be interfered during the conversation. The user phone will be sent a 128-bit challenge and the Subscriber Identity Module (SIM) uses the A3 algorithm as well as the Individual Subscriber Authentication Key(Ki) to compute a Signed RESponse(SRES). The key is unique to every SIM and it is sent to the base station.

The SIM uses a different algorithm, A8 and Ki, to compute a Session Key(Kc) and this will be sent to the base station. The session key is used with the A5 algorithm to encrypt the data for over the air transmission. The A3 and A8 algorithms are used together as one algorithm (A38). The algorithm is implemented on the SIM to compute SRES and Kc in parallel. An example algorithm set can be found in [93-98]. Figure 6.1 below shows the security process in GSM [99]. Additional information about GSM security is explained in [100].

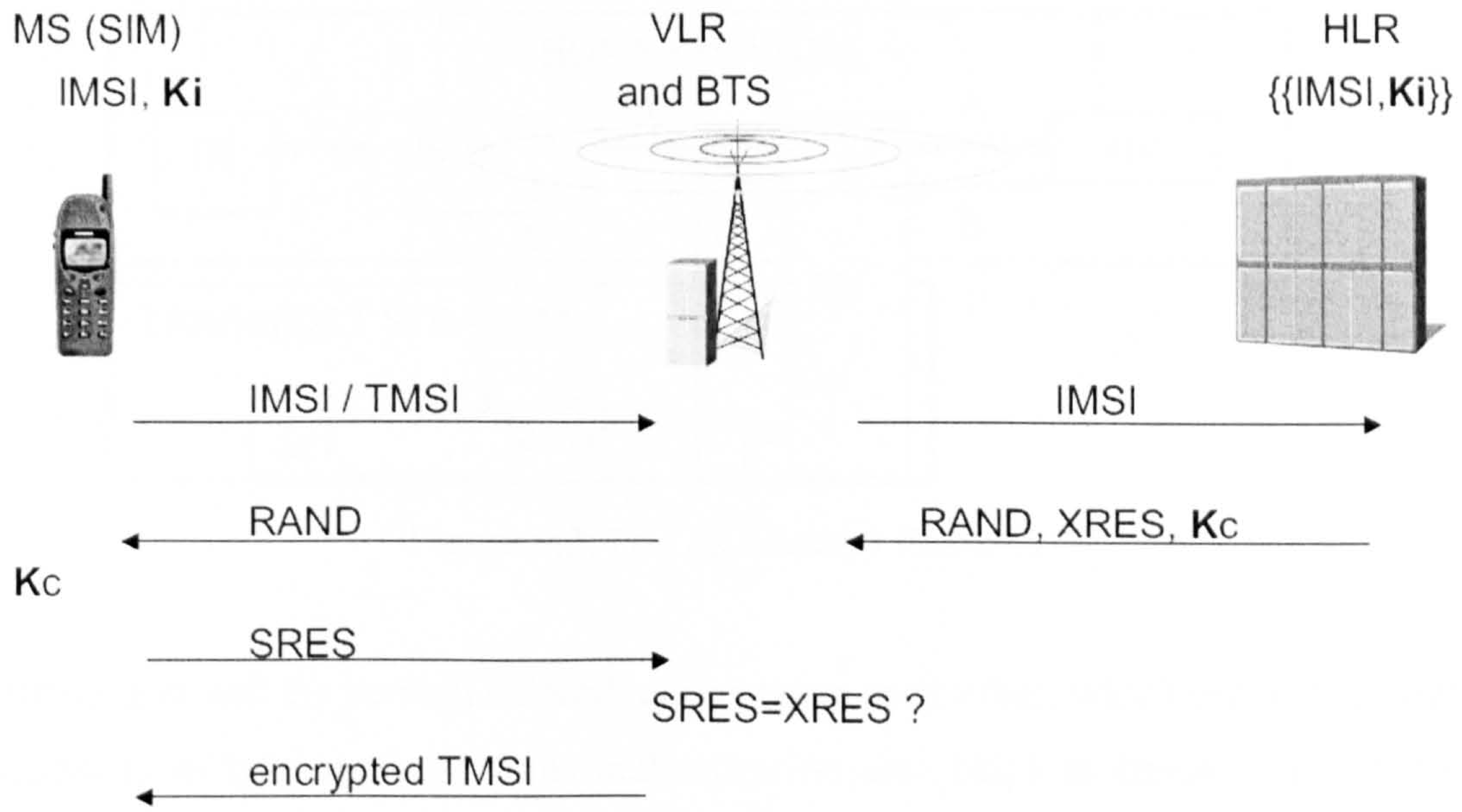


Figure 6.1: Identification and Authentication of a User (picture taken from [99])

UMTS is built based on the existing GSM infrastructure. Both packet and circuit data transmission are integrated in the system. Hence, UMTS can be used in parallel with GSM in the areas where the system has not fully implemented. This will allow a smooth transition into UMTS. GSM is expected to continue the operation for some years [101]. Good references for UMTS security can be found in [93, 94].

UMTS security is built by inheriting the GSM security features to ensure the compatibility between the two systems and it is discussed in detail in [3, 8, 11, 39, 92-94]. In addition, it improves the security by adding more features for 3G networks and services. UMTS consists of the following features as shown in Figure 6.2 [102]:

- Network Access Security (A)
- Network Domain Security (B)
- User Domain Security (C)
- Application Domain Security (D)
- Visibility and configurability

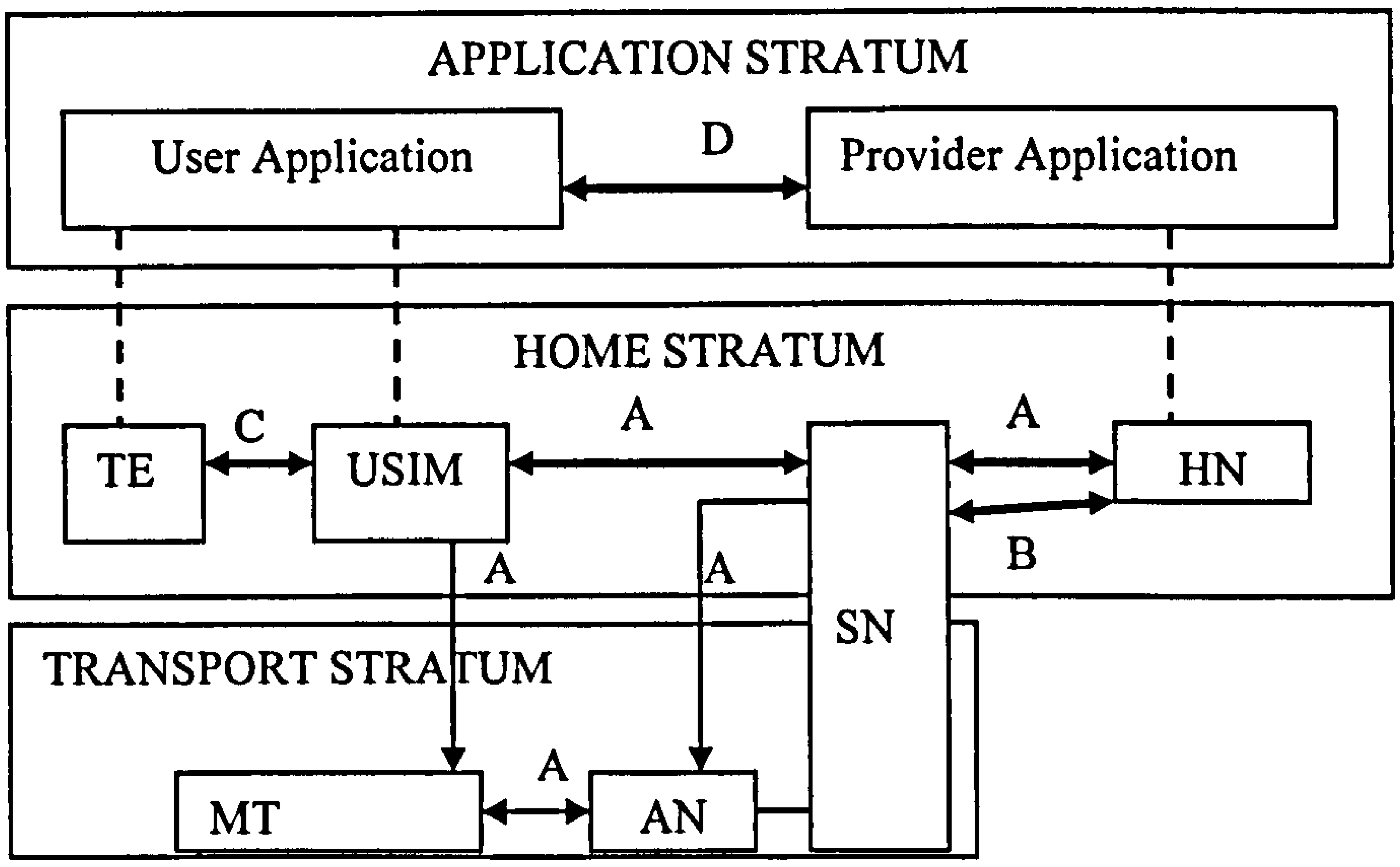


Figure 6.2: UMTS Security Features

Mobile user and the serving network authenticate each other, which provide security against false base stations. This authentication also helps to ensure that a bill is issued to the correct subscriber for example. The authentication consists of the user challenge (RAND), expected user response (X(RES)), the encryption key (CK), the integrity key (IK) and the authentication token for network authentication (AUTN).

The features ensure that messages are only available to the authorized users. The encryption is completed in the radio network controller whereas in GSM, the encryption is completed in the base station. UMTS also provides different security features for maintaining identity confidentiality such as user identity confidentiality, user location confidentiality and user untraceability [101]. Different level of security in GSM and UMTS are as shown in Table 6.1 [103].

Table 6.1: Different Security Levels of UMTS and GSM

GSM	UMTS
• No Serving Network (SN) Authentication • No Guarantee of Key Freshness • Length of KC 64 Bit	• Proof of Trust of HE by SN • Guarantee of Key Freshness • Integrity Protection of Signalling Messages • Length of CK, IK 128 Bit

The improved confidentiality use longer encryption key lengths, which is easier to upgrade compared to GSM. UMTS added a confidentiality algorithm which makes it more secure. The UMTS procedure is as shown in Figure 6.3 [104]. The authentication flowchart is as shown in Appendix C [99].

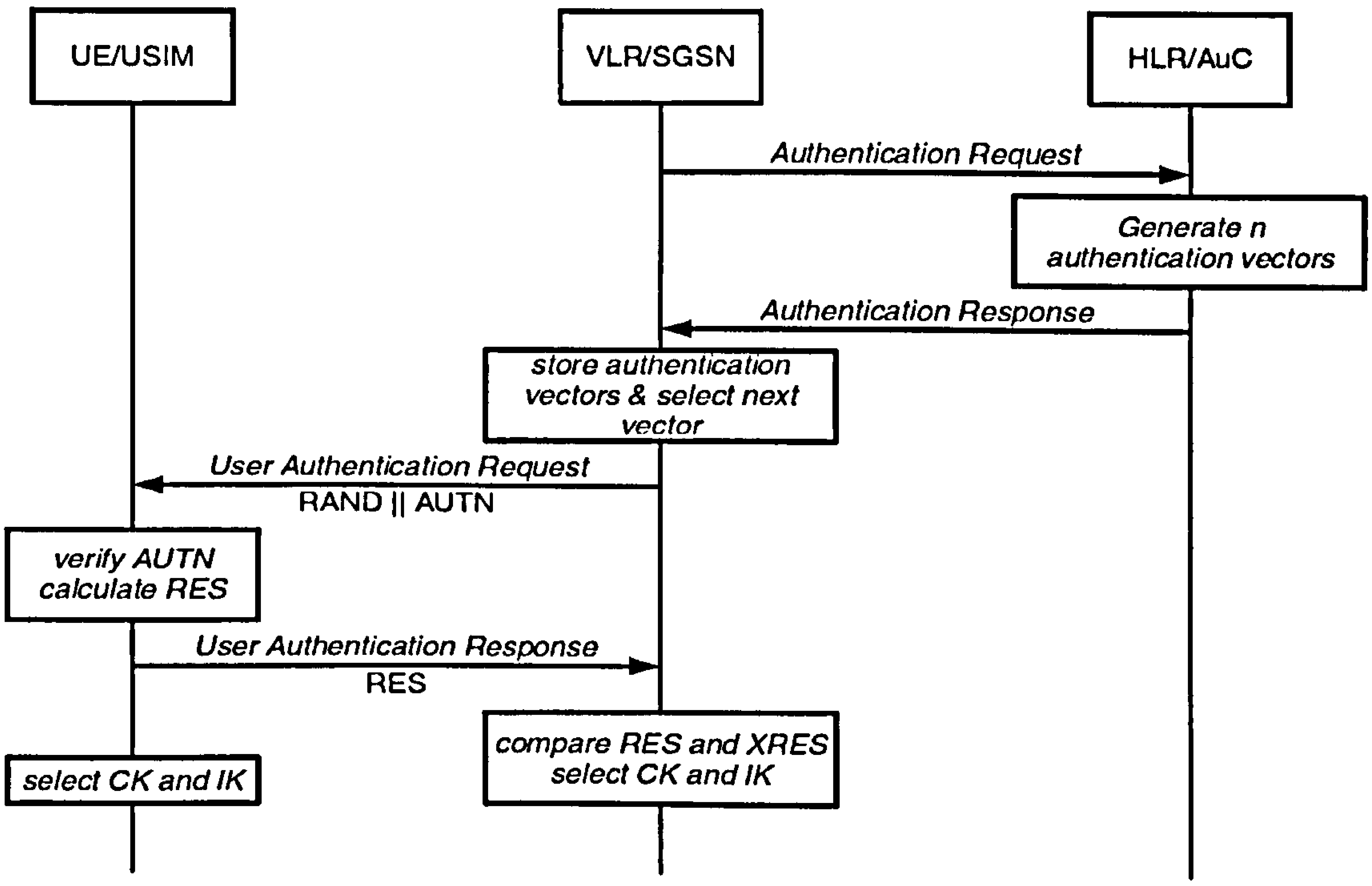


Figure 6.3: UMTS Authentication Procedure (picture taken from [104])

The cryptography methods used in GSM and UMTS are the conventional cryptography type. A transmission of secure VoIP can be implemented using IPSec. Cryptographic may be harmful to performance of VoIP and packet header compression for IPSec was introduced in [105]. A key is used for the processes related to data encryption, user authentication, data integrity, and digital signatures [106]. The introduction asymmetric cryptography has made the implementation of cryptography possible [107]. The detail specifications of the encryption algorithm have been prepared by [108].

There has been some research carried out on VoIP [29, 109-114]. Research on delay performance of VoIP in 3G systems was conducted in [109]. In this study, the network capacity was evaluated under different packet delay budgets associated with radio access channels. In [110], a new QoS control scheme that uses a simple protocol to detect the end-point CPU capabilities was proposed. A study which aimed to outline the potential security issues faced in transforming the traditional phone systems into VoIP systems was presented in [111]. In [22] a handover strategy for 4G wireless networks was proposed where handovers are classified as imperative and alternative handovers. Decision criteria are based on static and dynamic information.

A similar project has been carried out in [112] where the authors explained the challenges of VoIP security and steps for securing an organization's VoIP network. A model for carrying out simulations of the performance of secure session initiation protocol was described in [113] and the results of the performance analysis were presented. A quantitative analysis of the quality of service based on IPSec was evaluated in [114]. The aim was to analyse whether IPSec is good enough to transmit real time multimedia traffic. A study on the integration of standard security schemes with standard VoIP protocols was presented. The effects of firewall and virtual private network (VPN) techniques on the quality of a single Session Initiation Protocol (SIP)-based voice call were carried out. It also discussed issues related to implementing a secure and high quality VoIP network.

6.2 IP Security Architecture

VoIP threats can be eavesdropping, call recording, call modification, and voicemail forwarding or broadcasting. Unencrypted voice streams can be captured and reassembled using packet sniffers, for instance. A caller identity attack could result in redirect calls. As new IP telephony security standards and vendor functions continue to evolve, voice-oriented firewalls and data security techniques can used by service providers to increase voice security [115].

Threats to voice communication systems increase due to the IP communication. The threats are greater because voice files are kept on servers. IP Security (IPSec) is a standard framework for securing IP communication. It is a set of protocols developed by the Internet Engineering Task Force (IETF) to support the secure exchange of IP packets. It provides the capability to secure communications across a network. It is necessary to protect the data and it can be achieved by encrypting the packets at the IP level using IPSec. IPSec is a reliable, robust, and widely implemented method of protecting data and authenticating the sender. RSA is one of the network security company provide solution business and industry [116].

IPSec is a mechanism for protecting IP datagram and it can be used to provide data content confidentiality, data integrity authentication for the transmitted data, data source authentication, protection against traffic analysis and anti replay protection. It provides strong security that can be applied to all traffic crossing a firewall or router. It is transparent to applications as well as to the end users [117]. IPSec defines a method of specifying the traffic to protect, how that traffic is to be protected, and to whom the traffic is sent. It provides a standard, robust, and extensible mechanism in which to provide security to IP and upper-layer protocols such as UDP or TCP [118].

VoIP is time sensitive which requires a mechanism to ensure the QoS. Thus, usual security approach is not suitable. A security mechanism relevant to VoIP needs to be deployed to provide appropriate level of security. Several issues are crucial to security such as QoS, reliability, performance, scalability, authentication, availability as well as management.

Certain signaling protocol security recommendations exist such as the ITU-T H.235. This protocol provides authentication, privacy and integrity within the current H-Series protocol framework [115]. H.323 is an established ITU standard which has been used successfully for VoIP. It is designed to handle real-time voice and videoconferencing. The standard is based on the IETF Real-Time Protocol (RTP) and Real-Time Control Protocol (RTCP) [115]. There are a number of request for comments (RFCs) document related to IPSec which is published by the IP Security Protocol Working Group set up by the IETF. IPSec services and the security association concept are discussed in [119-122].

6.3 IPSec Modes

IP datagram (payload) or upper-layer protocols of an IP payload can be protected by using one of the IPSec protocols. There are two basic protocols used to provide security that are Encapsulating Security Payload (ESP) and Authentication Header (AH) as shown in Figure 6.4. The diagram illustrates the relationship among various components in IPSec and how the components interact with each other [118].

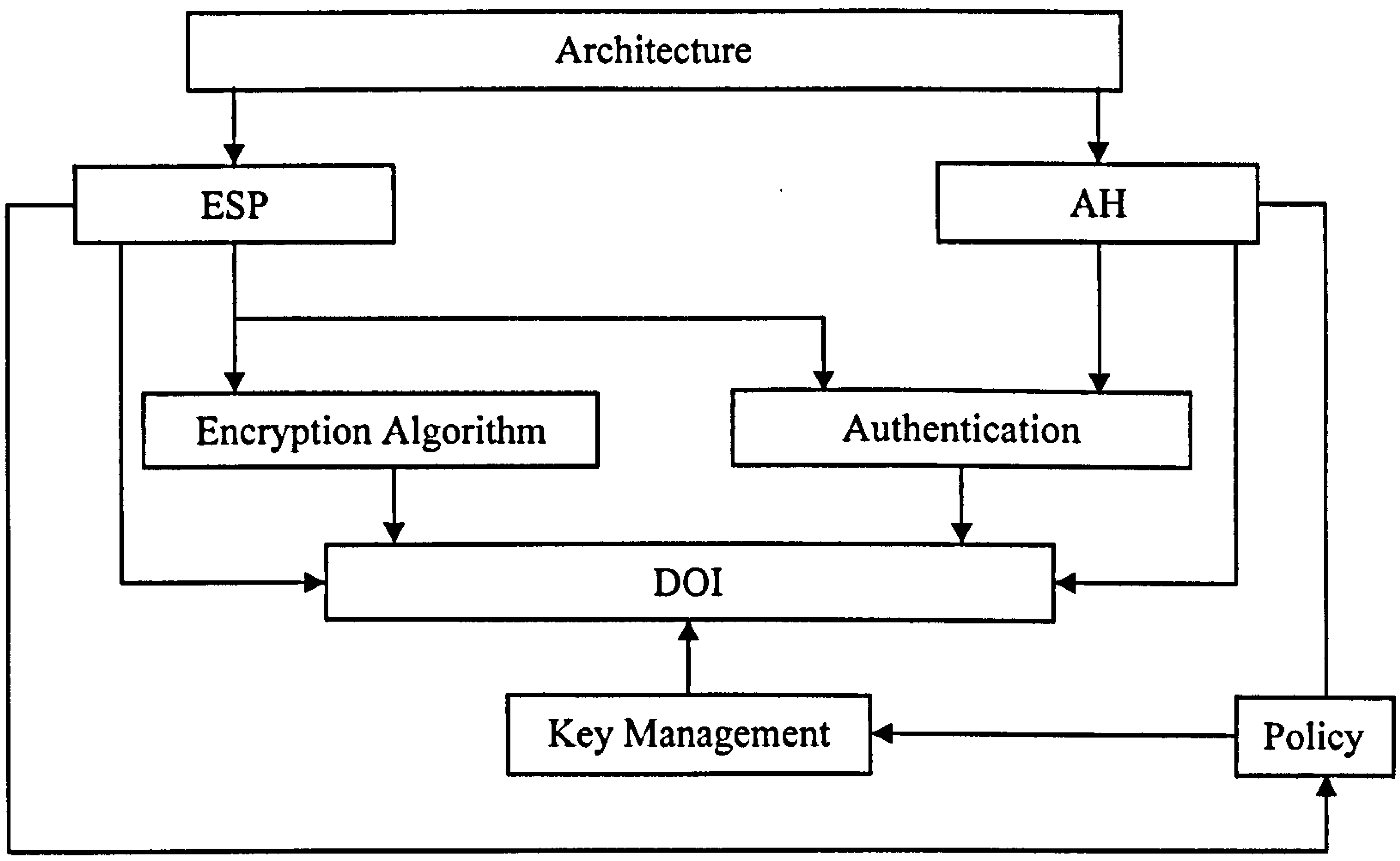


Figure 6.4: IP Security Components [118]

AH is the authentication protocol which is intended to guarantee connectionless integrity and data origin authentication. It provides data origin confirmation on received packets, and antiplay protection. However, AH does not provide any encryption services and confidentiality. It defines the method of protection, the format of the AH header, where that header is placed when doing transport mode or tunnel mode, output data processing, input data processing, and other information such as handling fragmentation and reassembly. But, it does not define the authentication algorithm to use [118].

AH can be used to protect an upper-layer protocol (transport mode) or an entire IP datagram (tunnel mode). The AH header immediately follows an IP header and it authenticates portions of the outer IP header of IPSec packet as shown in Figure 6.5.

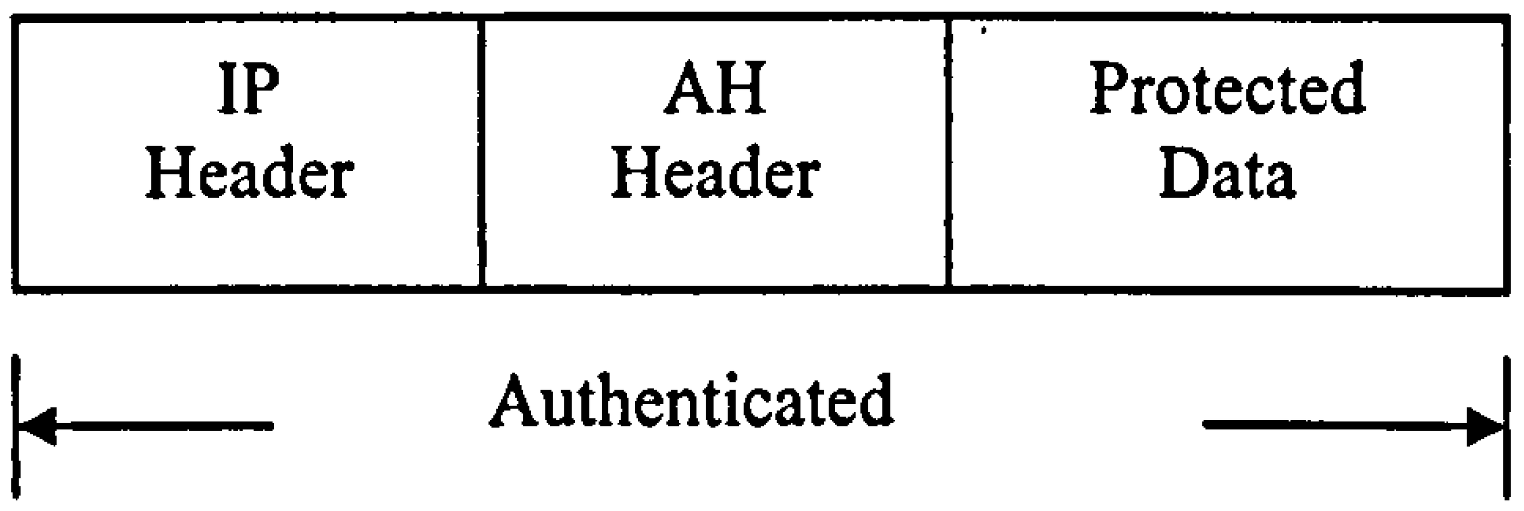


Figure 6.5: AH Protected IP Packet

Encryption is necessary for VoIP due to several techniques for capturing packet such as packet sniffing. Securing VoIP include protecting the content of the conversation and the receiver identity. This can be achieved by applying ESP using the tunnel method. By using IPSec for VoIP, it helps to reduce the threat and it makes VoIP more secure compared to phone line [115].

ESP may be applied in different modes. It is a combined encryption or authentication protocol. It is a protocol header inserted into an IP datagram to provide confidentiality of a packet, data origin authentication, antireplay, and data integrity services to IP packets it is protecting. An ESP header is inserted between the IP header and the upper-layer protocol header such as TCP or UDP header. It may be used to encapsulate an entire IP datagram. The ESP header and trailer in an ESP are as shown in Figure 6.6. [115].

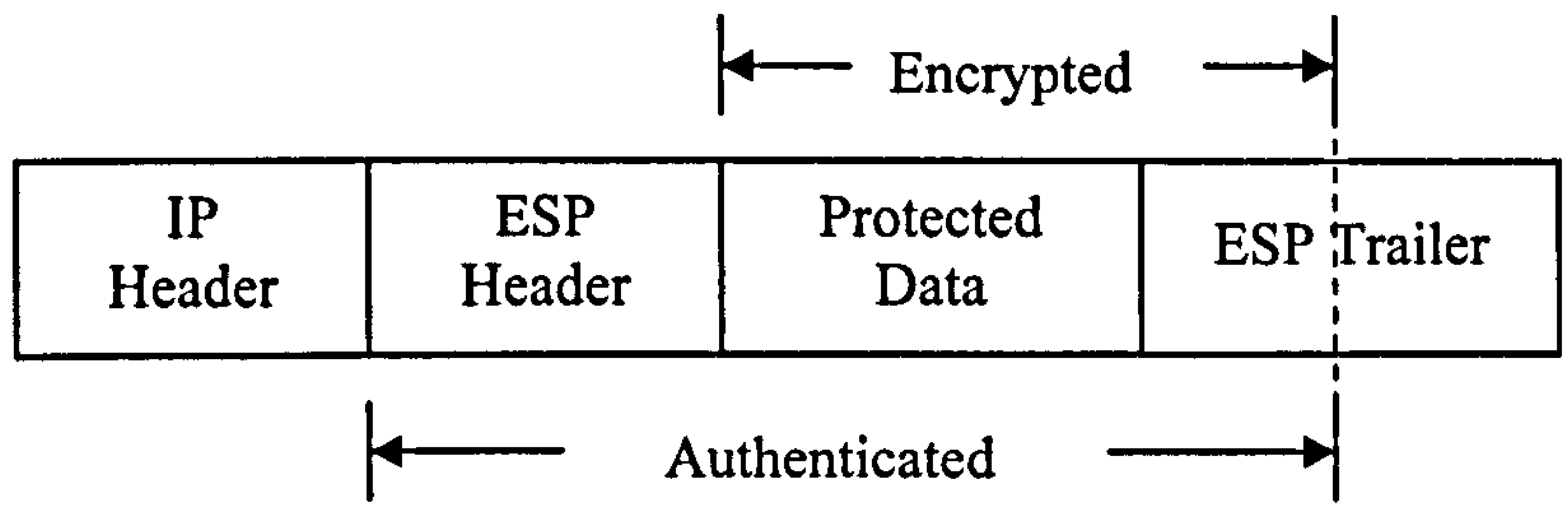


Figure 6.6: ESP Protected IP packet

Various aspects of IPSec are defined by RFCs such as architecture, key management and base protocols. IPSec can be deployed based on the security requirements of the users. It can be implemented and deployed in gateways or routers. When AH and ESP are protecting the same data, the AH header is always inserted after the ESP header. The AH header is much simpler than the ESP header. ESP provides all that AH provides with optional data confidentiality.

Internet Key Exchange (IKE) is a key management protocol that provides a method for authenticating IPSec, negotiating security services and generating shared keys. There are a few public key technologies such as Rivest, Shamir and Adleman (RSA) or Digital Signature Standard (DSS). However, these technologies, are too slow to operate on a packet-by-packet basis [118].

IPSec supports Transport and Tunnel modes of delivery. Different modes of IPSec can be deployed to protect any type of traffic carried over IP. Transport mode encrypts the data component (payload) of each packet and upper layer headers in the IP packet. In transport mode, an IPSec header is inserted between the IP header and the upper-layer protocol header. Transport mode is ideally suited for providing end-to-end security. On the other hand, in tunnel mode the entire IP packet to be protected is encapsulated in another IP datagram and an IPSec header is inserted between the outer and inner IP headers. Tunnel mode encrypts the entire IP datagram to be protected and places it in a new IP Packet. Both the payload and the IP header are encrypted. Tunnel mode is ideally suited for providing protection to transient traffic. The different between the two modes is as shown in Figure 6.7 [104].

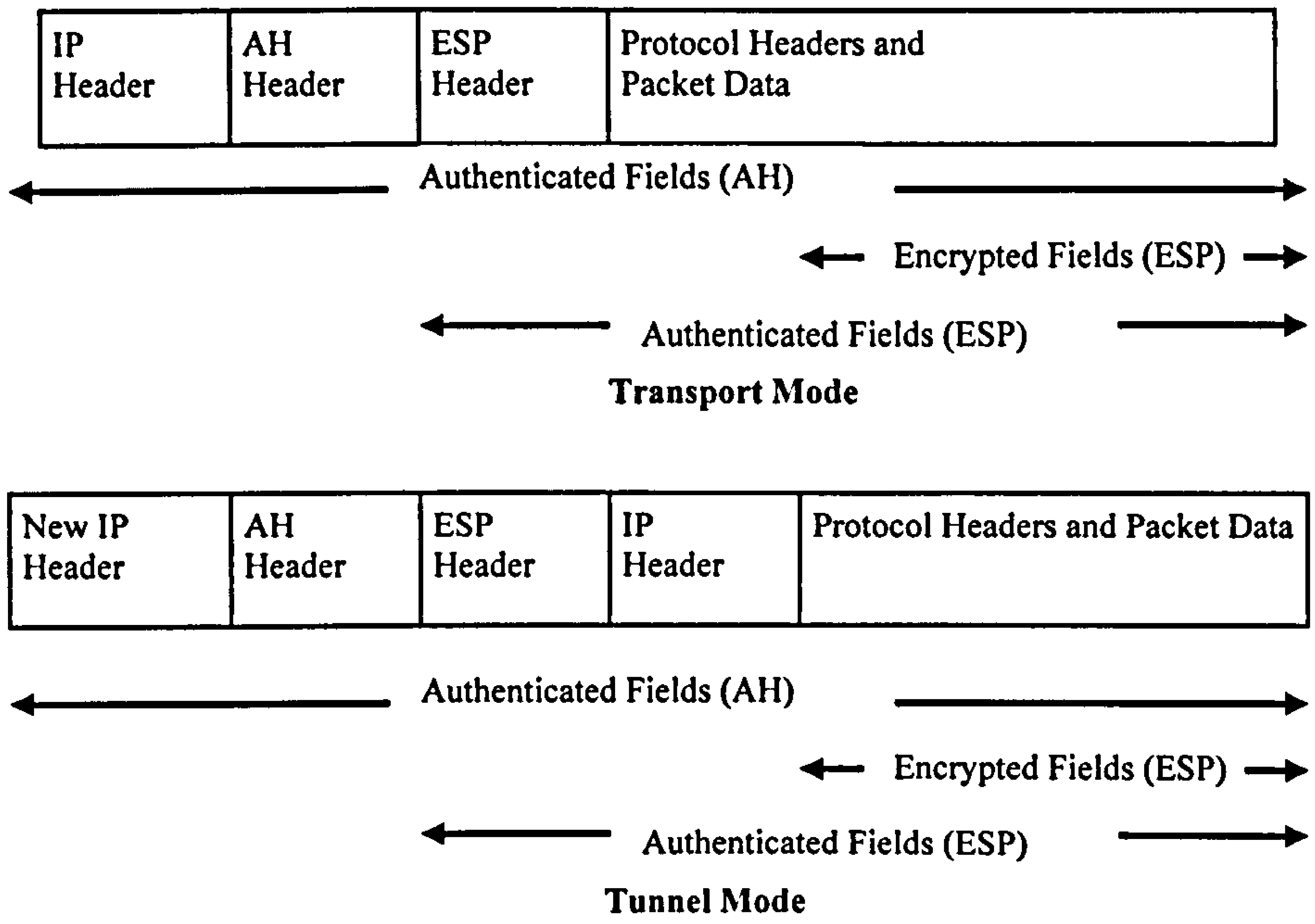


Figure 6.7: IP Packets in Transport Mode and Tunnel Mode

In transport mode, the content could not be determined by the attacker. They will not be able to intercept an IPSec packet. However, traffic analysis is allowed. The attacker will know who, when and how the conversation takes place between two parties. The unencrypted IP address of router or gateways are used at each node. Therefore, the source and destination address will not be acknowledged.

IPSec allows nodes to negotiate a security policy and a security association. This includes the encryption algorithm and algorithm key to be used. Different security deployments can be realized by placing IPSec at different points in the network such as routers, firewalls and nodes. End-to-end security can be achieved by deploying IPSec on nodes. However IPSec introduces an additional overhead to the packet size. As IPSec expands the packet size, encoding and decoding encryption can have a significant impact on delay. Thus, there must be a balance between the desired security and the desired recognised quality. A security solution should be implemented with a minimum delay to ensure the QoS conforms to the standard. The encryption process can be harmful to QoS and there will be trade-offs between convenience and security.

6.4 Simulation Model for IPSec Implementation

Handover could occur because of the signal strength falls below certain parameters specified in handover criteria. It could also occur when the traffic capacity of a certain cell has reached its maximum. Thus, the mobile users have to be handed over to neighbouring cells with less traffic load. The simulation was run to analyse the impact of IPSec protocol on a VoIP application and non-IPSec was used as a baseline.

The effect on the end-to-end delay and throughput was analysed. The performance was evaluated for a different packet arrival rate over a range of system parameters. The results of system simulations provide some insights into the conditions and parameters involved. In the study, users communicating using the IP were modelled to analyse the performance of packet voice communication using security protocol. A simple model for IPSec network is as illustrates in Figure 6.8.

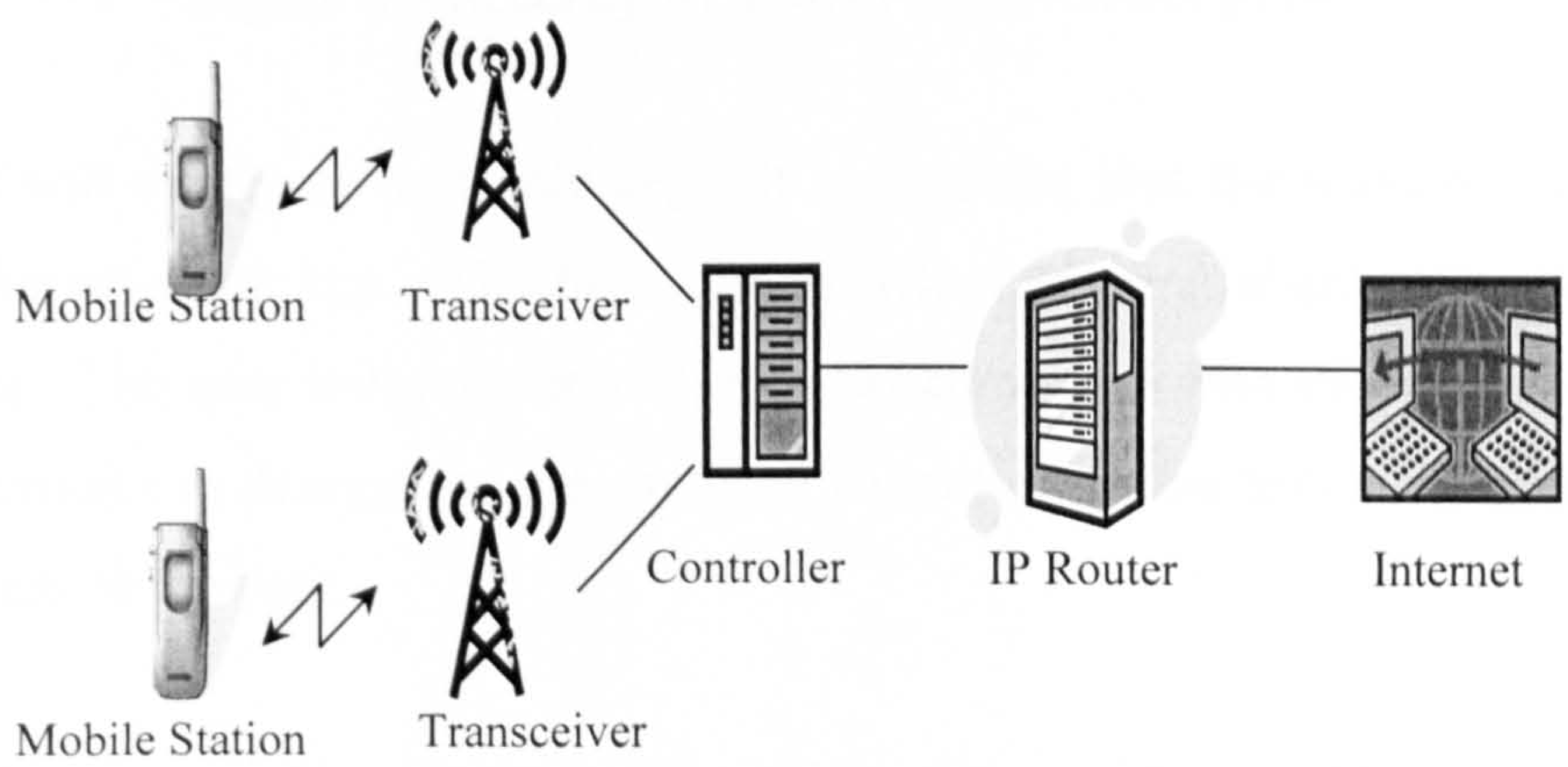


Figure 6.8: Packet Voice IP Model

IPSec algorithm was deployed at every mobile device for end-to-end data transfer. IPSec provides the shared key that is required to perform authentication and confidentiality. The 3DES encryption algorithm is widely supported and some implementations can make use of the AES encryption algorithm. The Advanced Encryption Standard (AES) based on the Rijndael encryption algorithm is used for the encryption scheme. Figure 6.9 shows the work flow for the security analysis process.

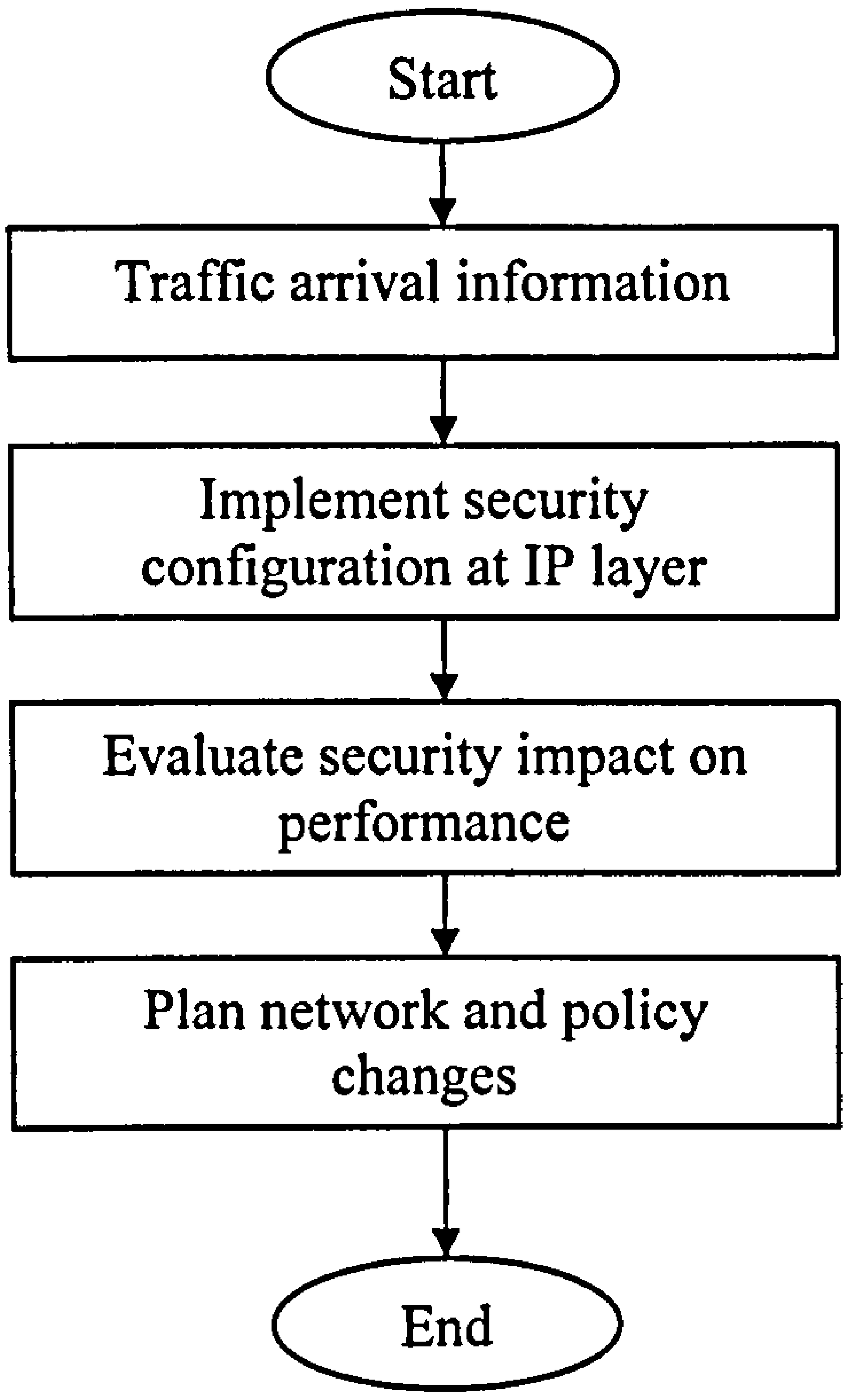


Figure 6.9 Security Analysis Process Work Flow

The users will configure a private key. It is assumed that the users of the system know a shared secret key. The base station will send an Authentication Response with a text. The user will answer with the text encrypted with the security key. If the base station can decrypt the message and validates that the text is the same, it will grants access to the user.

Rijndael has a variable number of rounds. To encipher a block of data, an AddRoundKey step, the ShiftRow step, and the MixColumn step with the AddRoundKey step will be performed. The ByteSub alters the bytes to be encrypted whereas the ShiftRow step ensures that different bytes of each row will involve permutation with the corresponding byte in other rows. The MixColumn step causes every byte in a column to affect every other byte. The step involves interaction of different bytes with each other. The AddRoundKey step, on the other hand, corresponds to the XOR of subkey. The following algorithm shows the process:

```
Round(State,RoundKey)
{
ByteSub(State);
ShiftRow(State);
MixColumn(State);
AddRoundKey(State,RoundKey);

Rijndael(State,CipherKey)
{
KeyExpansion(CipherKey,ExpandedKey) ;
AddRoundKey(State,ExpandedKey);
For( i=1 ; i<Nr ; i++ ) Round(State,ExpandedKey + Nb*i) ;
FinalRound(State,ExpandedKey + Nb*Nr);
}
```

6.5 Result Analysis for Secured Voice Environment

The impact of security protocol on system performance is as shown in Figure 6.10 to Figure 6.14 [123]. Figure 6.10 shows the impact of security on end-to-end delay. VoIP over 3G data access could be limited to only 64Kbps upstream [76]. The result shows that the end-to-end delay is unacceptable to the user when the rate is 32Kbps. The delay was very high when the security protocol was implemented.

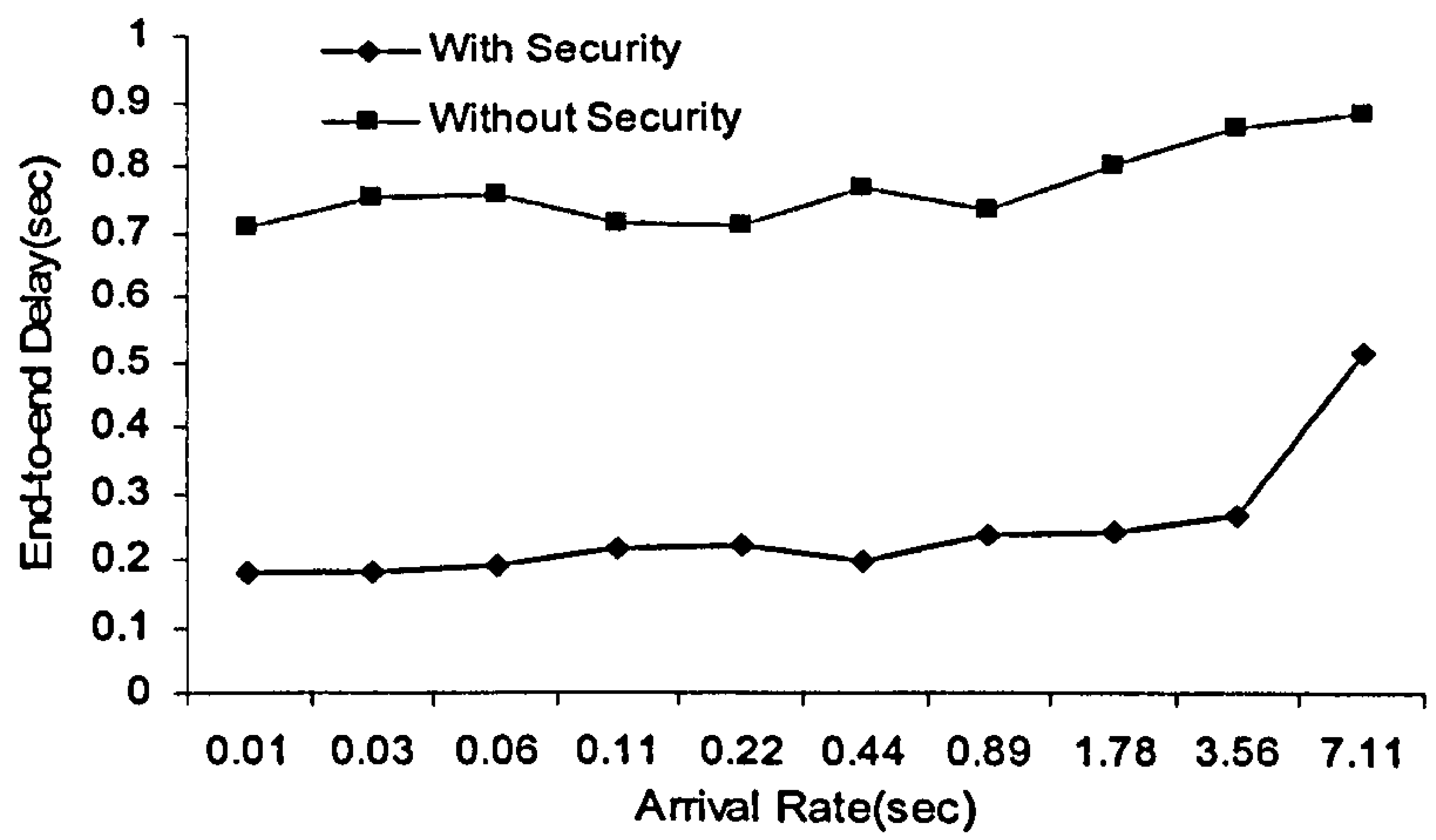


Figure 6.10: End-to-end delay when rate is 32Kbps

Similarly, Figure 6.11 illustrates the delay when the interface rate is 64Kbps. The delay is slightly noticeable and it is still falls within the recommended value. However, as the traffic increase, the delay increases drastically at certain point.

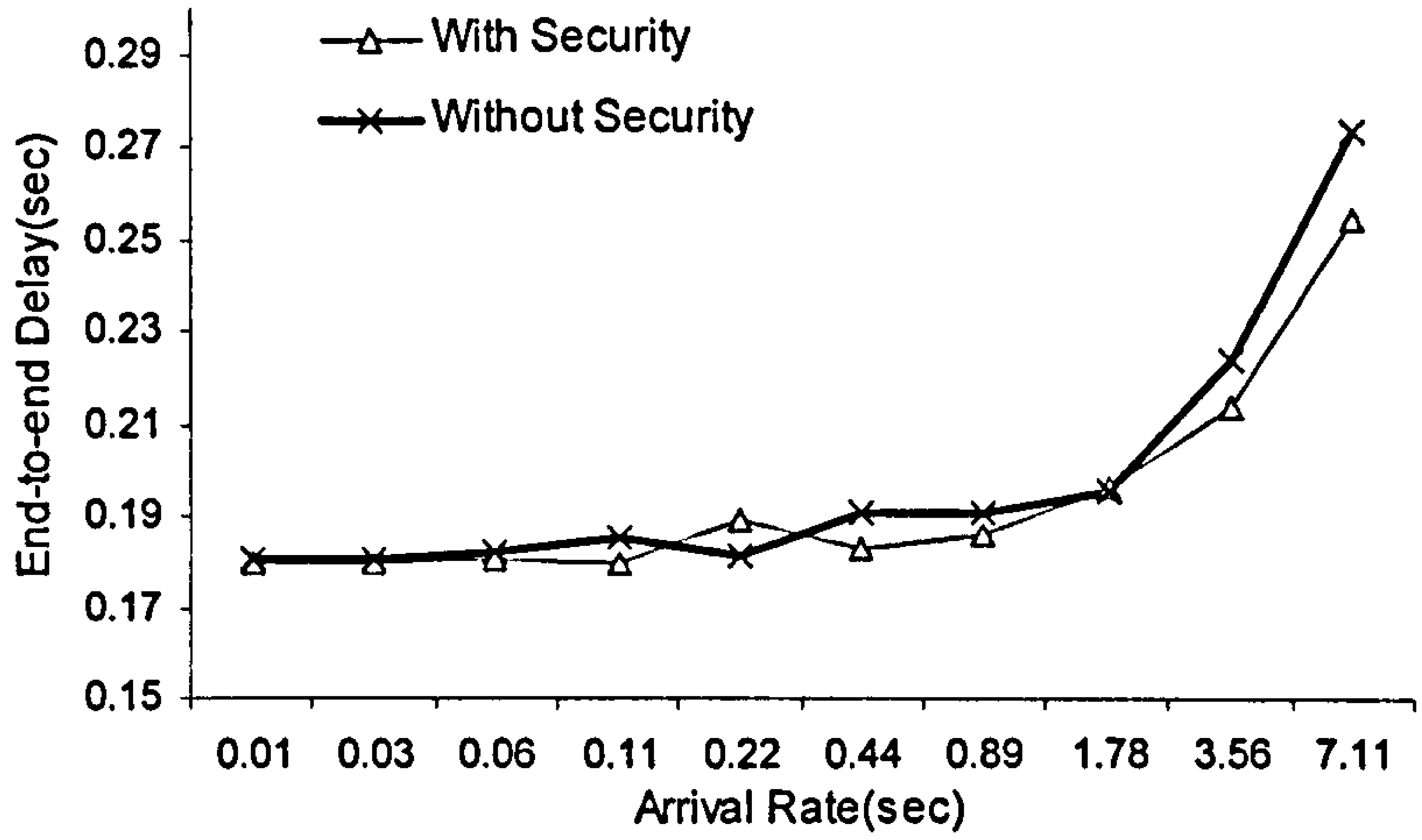


Figure 6.11: End-to-end delay when rate is 64Kbps

Figure 6.12 shows that the end-to-end delay of the simulation model is compared with the theoretical model as in [72]. As the traffic increase, the delay maintains at a certain level and at a certain point it increases drastically. The simulation result matches closely to the theoretical result.

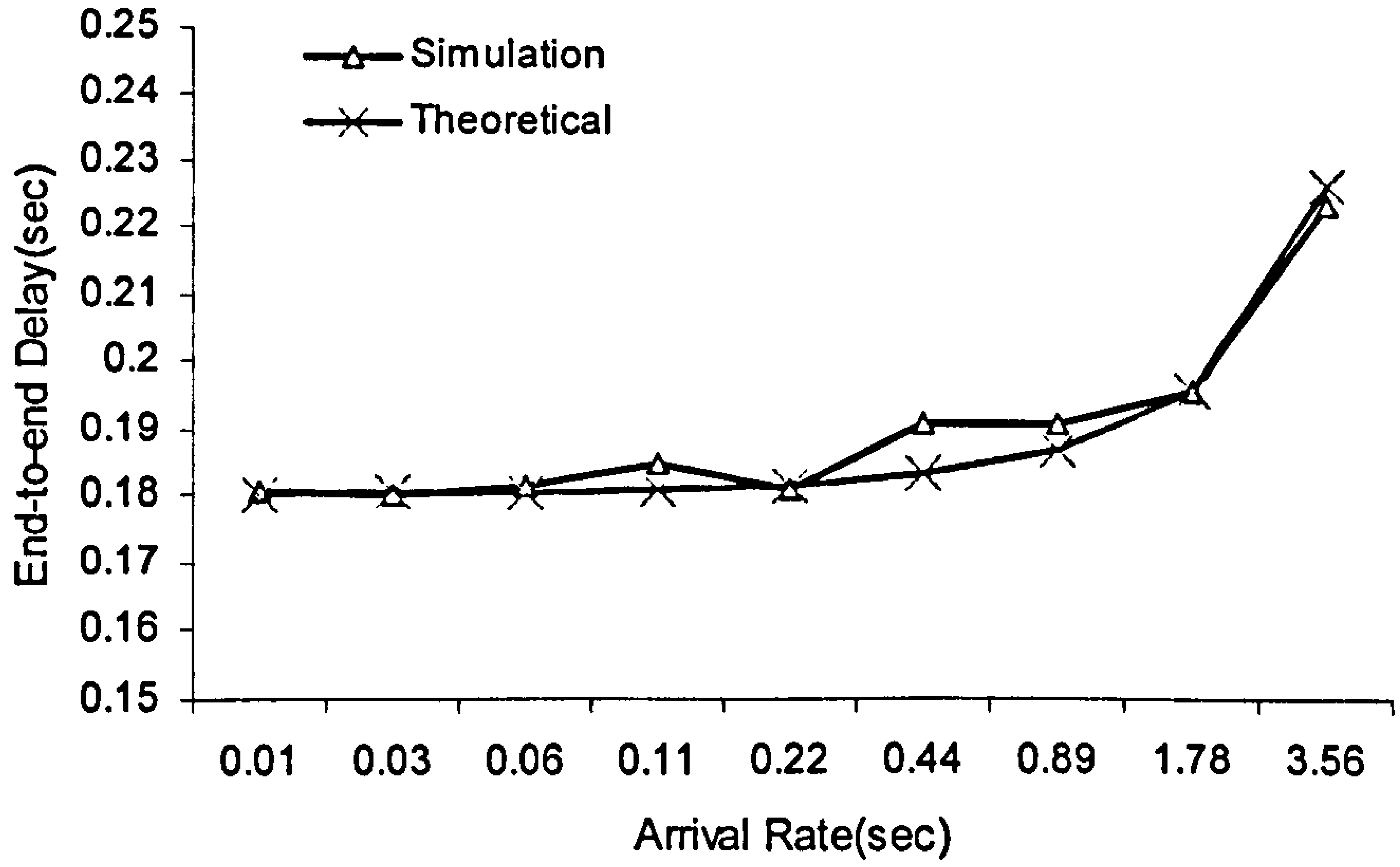


Figure 6.12: End-to-end delay against Arrival Rate

The impact of security on percentage of packet loss is shown in Figure 6.13. The graph shows that the packet loss is very noticeable and it does not conform to the recommended value.

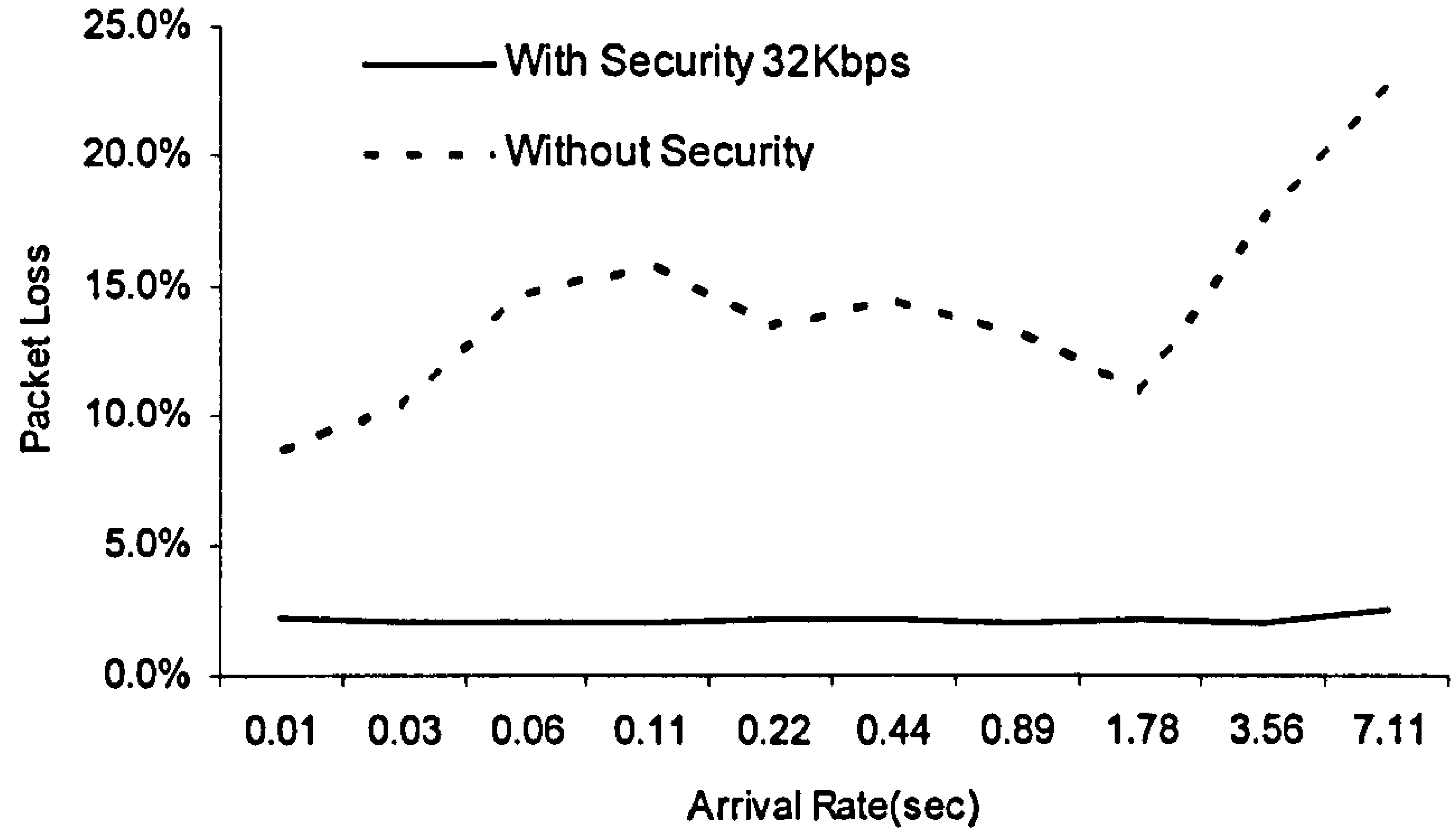


Figure 6.13 Packet loss against Arrival Rate

Likewise, Figure 6.14 shows the impact of security on percentage of packet loss. The percentage of packet loss fluctuates as the traffic increase. However, the difference is very small and is almost negligible.

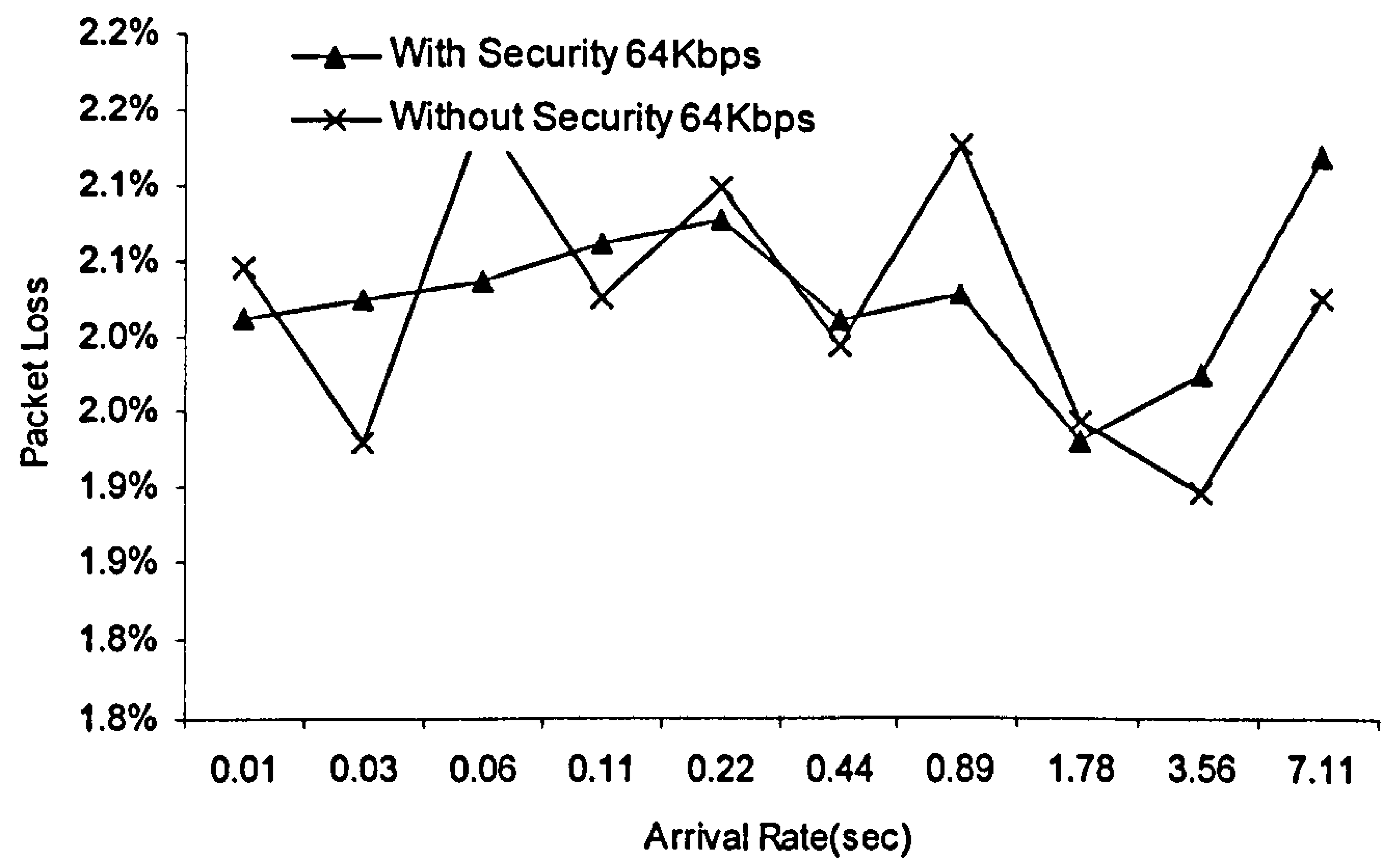


Figure 6.14: Packet loss against Arrival Rate

6.6 Conclusion

This chapter discussed the security issues in GSM and UMTS networks. Prevention methods associated with traditional voice security. The security issues are generally very specific to the technology in use. IP is the easiest way to obtain information and the future network seems to be based on IP. Thus, the architecture of IP networks is discussed. VoIP is being approved as a replacement for the traditional circuit switched infrastructure. Nevertheless, voice communication and its security are a relevant concern for the most critical infrastructure. Security aspects of VoIP have emerged as important as QoS.

Due to the nature of VoIP which is time-sensitive, most of the security measures in data networks could not be used in VoIP networks. The impact of IPSec on the quality of voice in IP environment was analysed. The basic QoS such as delay, jitter and packet loss were used in evaluating the impact of the IPSec. These issues arise in VoIP environment because the packet must arrive at the destination fast. If a packet is lost, there is no time to resent the packet. The result indicated that a security algorithm affect the overall performance of a voice service as it increased latency and contributed jitter which may degrade the voice quality to an unacceptable level. It was found that a few factors affect voice traffic over secure IP network such as encrypt payload and increased packet size. Encryption result in a larger header to payload ratio for each packet which reduced the effective bandwidth.

The increase in delay was due to a significant increase in packet size cause by the additional header for security. As VoIP uses small packets in a high volume, this may accumulate the total packet size. From the results, it can be concluded that the overall quality can be reduced when security protocol was implemented at a low bit rate. Although security results in a degradation of speech quality, there must be a compromise for securing the data. VoIP is a technology that would have wide scale of implementation in the future and secured communication in IP based networks is a key success for the VoIP evolution.

Chapter 7

Conclusion and Future Work

7.1 Conclusion

This thesis investigated the influence of traffic on the voice quality in IP based application in a 3G network. The importance of traffic management had been pointed out and an efficient way of managing traffic was analysed. The QoS becomes unacceptable when the desired criteria have not been fulfilled. Excessive call drops may occur during heavy traffic conditions that will cause unacceptable call quality. Therefore, the voice quality must be controlled so that it stays below an acceptable level. Handover is a mechanism that guarantees QoS by handed over signal to the target base station. To evaluate the handover algorithm, some performance measures were collected and compared with the threshold value.

A mathematical calculation was used to evaluate the performance. The impact of types of codes on the performance was analysed. By analysing the QoS, the model can be used to predict the overall performance such as queue length and end-to-end delay. From the numerical analysis, it was shown that the algorithm reduces the end-to-end delay. However, the mathematical model was closely matched only up to a certain value. This was due to several assumptions that were made in the simulation model. The dependency of the types of packet in the traffic mix was discussed. Appropriate scheduling for the handover can be used to improve the QoS. A handover algorithm based on weight was implemented in this study. The handover decision was based on the signal and distance from the base station as well as the types of application, where the handover will be performed based on the assigned weight.

In a mixed environment, the voice quality is not guarantee. To support multi services, traffic management is needed to match the ability of the system to support high transmission rates. The investigation was extended by deploying load balancing technique at the base station to estimate system capacity. To allocate the traffic efficiently, the traffics were classified into two types, high priority and low priority based on different types of services. The base station will reserve bandwidth to high priority traffic.

The type of traffics was considered where the mobile stations were handed over to the most appropriate base station. From simulation, the results showed that the schemes maintain the end-to-end delay. By classifying the packet types, the system capacity was enhanced. In addition, the results indicate that the network performance can be significantly improved by balancing the load caused by handover. This is mainly depended upon the type of traffic involved. The study showed that the end-to-end delay of the mathematical model match closely to the simulation when the traffic is low. This is due to assumption made in the simulation such as distance, type of path loss, and mobile trajectory.

A tradeoff between QoS and system capacity can be achieved by allowing a limited number of users with a guaranteed bandwidth. The influence of the balancing technique on the performance of voice in IP network can be seen in several ways. Due to the packetized voice were prioritised, less end-to-end delay were achieved. Furthermore, the load balancing scheme reduced jitter and packet loss, and therefore allowing more users to be served. Finally, when bandwidth was reserved to the voice packets, it reduced the end-to-end delay.

Providing a secure mobile communication in VoIP is the security challenges. Security prevention and security detection are two approaches in securing the networks. To ensure a complete security, the technical security features must be properly supported. The level of security protocol adaptation is necessary because at the same time it should not really affect the performance. IPSec was implemented in the study to analyse the impact of security on the packetized voice.

IPSec is introduced an alternative to Transport Layer Security to secure the packets. Congestion of the network contributes to the delay of the VoIP. Delay can be caused by the use of codec and encryption process. Thus, although security is one of the concerns, the aim of the study was to assure that delays caused by security devices are kept at a minimum level. By supporting QoS and improving the efficiency of bandwidth, it was hope to help minimizing delay in a secured VoIP network.

Complex security schemes may degrade performance. As voice is in data networks, it can be secured by encrypting the packets IPSec. However, it can cause an excessive amount of latency in the VoIP packet delivery which degraded voice quality. From simulation, it shows that the end-to-end delay is increase by deploying IPSec. The effect of IPSec on QoS issues and encryption issues is to handle encryption or decryption at the endpoints in the VoIP network. It is worth to establish a secure implementation of VoIP although there could be a significant delay due to the encryption and decryption.

Finally, it was shown that security is important for an IP environment. There should be a tradeoff between voice quality and security. A suitable balance between performance and security needs to be identified. In conclusion, mobility management and addressing are some of the issues that need to be clarified before the traffic mix in IP based service can be fully implemented.

It can be concluded from the whole thesis that these aims have been achieved:

- Understanding of different technologies in mobile communications generation.
- Understanding of how 2G and 3G system work, specifically GSM and UMTS.
- Evaluation on the impact of handover on the quality of service and system performance.
- Evaluation of the impact of different types of voice coder and voice traffic patterns over IP networks in 3G.
- Maximising the capacity and system performance in mixed voice and data traffic environment using scheduling based on weight and load balancing techniques.
- Evaluation of voice quality and system performance in a secured IP network environment by deploying IP security algorithm on packetized voice.

7.2 Suggestion for Future Research

There are some interesting possibilities for improvements and future work. In this study, the percentage of traffic types is constant. The next goal is to investigate the application of all QoS classes to achieve more accurate results. The simulation should be run in a more complicated environment. Focused should be more on the varying the percentage of traffic types. It can involve analysing how interference affects the data rate, lost packets and access delay factors in a congested area of the systems. The study will involve investigating security in terms of authentication trends. Differences and similarities of the different types of security algorithm can be implemented.

It would also be interesting to apply the Artificial Intelligence techniques to manage the mixed traffic. The mobile users behaviour such as speed and movement, need to be studied to get more accurate results. Hence applying Neural Network algorithm seems to be able to further improve the result. The quality of service provided to delay sensitive traffic can be decrease due to poor resource allocation. Therefore, further investigations are necessary to identify as well as quantify the nature and amount of capacity improvement possible. Mobile location prediction based on Neural Network seems to improve the resource management procedure. The movement history and type of application used by the mobile users will be monitored in the model.

The new wireless technologies have been designed to provide support for multimedia services, with different traffic characteristics and different QoS guarantees. Hence, deploying Genetic Algorithm technique for channel allocation reservation seems to be an effective method. Furthermore, the impact of different types of security was not studied. The study will involve investigating security in terms of authentication trends. Differences and similarities of the different types of security algorithm will be implemented. In addition, WLAN and UMTS technologies will have to coexist. Load can be transferred to WLAN so that UMTS congested cell can be freed for another connection. Thus, a study on the impact of performing load balancing between the two technologies is another interesting possibility.

References

References

[1] H. Holma and A. Toskala, *WCDMA for UMTS Radio Access for Third Generation Mobile Communications*, 3rd ed., West Sussex: John Wiley and Sons Ltd, 2004.

[2] G. L. Bodic, *Mobile Messaging Technologies and Services SMS, EMS and MMS*, 2nd ed., West Sussex: John Wiley and Sons", 2005.

[3] D. Wisely, P. Eardley, and L. Burness, *IP for 3G-Networking Technologies for Mobile Communications*, West Sussex: John Wiley and Sons Ltd, 2002.

[4] F. Hillebrand, *GSM and UMTS: The Creation of Global Mobile Communication*, West Sussex: John Wiley & Sons, 2001.

[5] http://www.3gnews.org/, retrieved June 2005.

[6] http://www.umtsworld.com/umts/faq.htm, retrieved June 2005.

[7] http://www.iec.org/online/tutorials/umts/topic01.html, retrieved June 2005.

[8] A. Richardson, *WCDMA Design Handbook*, Cambridge University Press, 2005.

[9] B. Walke, R Seidenberg, M. P. Althoff, *UMTS The Fundamentals*, West Sussex: John Wiley and Sons Ltd, 2003.

[10] R. J. Bates, *Broadband Telecommunications Handbook*, 2nd ed. United States: McGraw Hill, 2002.

[11] S. G. Glisic, *Advanced Wireless Networks 4G Technologies*, West Sussex: John Wiley and Sons Ltd, 2006.

[12] S. Hara and R. Prasad, *Multicarrier Techniques for 4G MobileCommunications*, London: Artech House, 2003.

[13] D. Zhao, X. S. Shen and J. W. Mark, "QoS Guarantee and Power Distribution for Soft Handoff Connections in Cellular CDMA Downlinks", *IEEE Transactions on Wireless Communications*, Vol. 5, No. 4, pp. 910-919, 2006.

[14] "GSM Historical Background", retrieved June 2005 from http://www.etsi.org/pressroom/articles/gsm-historical%20background.pdf.

[15] "GSM and UMTS Standard", retrieved June 2005 from http://www.etsi.org/services_products/freestandard/documents/gsmumts.pdf

[16] G. P. Pollini, "Trends in Handover Design", *IEEE Communications Magazine*, Vol. 34, No. 3, pp. 82-90, 1996.

[17] R. Vijayan and J. M. Holtzman, "A Model for Analyzing Handoff Algorithms", *IEEE Transactions on vehicular Technology*, Vol. 42, No. 3, pp. 351-356, 1993.

[18] Verdone, R and A. Nella, "Performance of Received Power and Traffic-driven Handover Algorithms in Urban Cellular Networks", *IEEE Wireless Communications*, Vol. 9, No. 1, pp. 61-70, 2002.

[19] D. C. Jose, et al, "A method of Performing a handover or reselection procedure", European Patent Application, retrieved 12 April 2005 from: http://l2.espacenet.com/espacenet/viewer?PN=EP1448010&LG=en&DB=EPD

[20] C. I. Bauer and S. J. Rees, "Classification of handover schemes within a cellular environment", *PIMRC*, Vol. 5, pp. 2199-2203, 2002.

[21] S. Balasubramaniam and J. Indulska, "Vertical handover supporting pervasive computing in future wireless networks", *Computer Communications*, Vol. 27, pp. 708-719, 2004.

[22] W. Zhang, J. Jaehnert and K. Dolzer, "Design and evaluation of a hand-over decision strategy for 4th generation mobile networks", *IEEE VTC*, Vol. 3, pp. 1969-1973, 2003.

[23] C. G. Gaytan, D. M. Rodri'gueza and R. Soto, "Uncertainty in cellular fuzzy handover", *Expert Systems with Applications*, Vol. 15, pp. 391-397, 1998.

[24] T. Onel, C. Ersoy, E. Cayirci and G. Parr, "A multicriteria handoff decision scheme for the next generation tactical communications systems", *Computer Networks*, Vol. 46, pp. 695-708, 2004.

[25] G. Edwards and R. Sankar, "Microcellular handoff using fuzzy techniques", *Wireless Networks*, Vol. 4, pp. 401-409, 1998.

[26] M. Junius and O. Kennemann, "Applications for Locating, Handover, and Network Management", *IEEE VTC*, pp. 338-342, 1994.

[27] G. Liodakis and P. Stavroulakis, "A novel approach in handover initiation for microcellular systems," *IEEE VTC*, Vol. 4, pp. 1820-1823, 1994.

[28] F. Poppe, D. De Vleeschauwer, G.H. Petit, "Guaranteeing Quality of Service to Packetised Voice over the UMTS Air Interface", *Eighth International Workshop on Quality of Service (IWQoS 2000)*, pp. 85-91, 2000.

[29] H. Zheng, G. Rittenhouse and M. Recchione, "The Performance of Voice over IP over3G Downlink Shared Packet Channels under DifferentDelay Budgets," *IEEE VTC*, Vol. 4, pp.2501–2505, 2003.

[30] H. Fathi and R. Prasad, "Mobility Management for VoIP in 3G Systems:Evaluation of Low latency Handoff Schemes", *IEEE Wireless Communications*, Vol. 4, pp. 96–103, 2005.

[31] L. Zheng, L. Zhang and D. Xu,. "Characteristics of network delay and delay jitter and its effect on Voice over IP (VoIP)", *IEEE International Conference on Communications (ICC 2001)*, Vol. 1, pp. 122-126, 2001.

[32] R. Cuny and A. Lakaniemi, "VoIP in 3G Networks:An End-to-End Quality of Service Analysis", *IEEE VTC Spring*, Vol. 2 pp. 930-934, 2003.

[33] 3rd Generation Partnership Project; "Universal Mobile Telecommunications System (UMTS); Physical layer; Measurements (FDD) (3GPP TS 25.215 version 6.2.0 Release 6)", ETSI TS 125 215, V6.2.0, March 2005.

[34] 3rd Generation Partnership Project; "Technical Specification Group Radio Access Network; Physical layer – Measurements (TDD) (Release 1999)", 3G TS 25.225, V3.2.0, March 2000.

[35] "What is GSM ?", retrieved June 2005 from http://www.gsmworld.com/technology/what.shtml.

[36] G. Heine, *GSM Networks:Protocol, Terminology and Implementation*, Artech House Publisher," 1998.

[37] S. M. Redl, M. K. Weber and M. W. Oliphant, *An Introduction to GSM*, Artech House Publishers 1995.

[38] http://www.gsmworld.com/index.shtml, retrieved June 2006.

[39] H. Kaaranen, A. Ahtiainen, L. Laitinen, S. k Naghian and V. Niemi, *UMTS Networks Architecture, Mobility and Services*, 2nd ed., West Sussex: John Wiley and Sons Ltd, 2005.

[40] http://www.umts-forum.org/servlet/dycon/ztumts/umts/Live/en/umts/What+is+UMTS_index, retrieved June 2006.

[41] J. Perez-Romero, O. Sallent, R. Agusti, M. Angel and D.-Guerra, *Radio Resource Management Strategies in UMTS*, West Sussex: John Wiley & Sons Ltd, 2005.

[42] M. J. Nawrocki, M. Dohler and A. H. Aghvami, *Understanding UMTS Radio Network Modelling, Planning and Automated Optimisation Theory and Practice*, West Sussex: John Wiley & Sons Ltd, 2006.

[43] T. Rappaport, *Wireless Communications: Principles and Practice, Design and Implementation*, Prentice Hall," 2003.

[44] "FDMA vs TDMA vs CDMA: What's the difference ?", retrieved February 2006 from http://www.owlnet.rice.edu/~elec301/Projects01/cdma/compare.html.

[45] A. Brand and H. Aghvami, *Multiple Access Protocols for Mobile Communications GPRS, UMTS and Beyond*, West Sussex: John Wiley and Sons Ltd, 2002

[46] G. Sanders, L. Thorens, M. Reisky, O. Rulik and S. Deylitz, *GPRS Networks*, West Sussex: John Wiley and sons, 2003.

[47] J. Laiho and A. Wacker, *Radio Network Planning and Optimisation for UMTS*, 2nd ed., West Sussex: John Wiley & Sons Ltd, 2006.

[48] "Basic Concepts of WCDMA Radio Access Network", retrieved 15 February 2005from:http://www.ericsson.com/products/white_papers_pdf/e207_whitepaper_ny_k1.pdf

[49] R. Prasad and L. Munoz, *WLANs and WPANs towards 4G Wireless*, London: Artech House, 2003.

[50] "GSM05.08 - Multi-Carrier Specification for Spread Spectrum 3 Systems on GSM MAP (MC-MAP) 4 (Lower Layers Air Interface)", retrieved 15 February 2006 from: http://www.3gpp.org/ftp/tsg_ran/tsg_ran/TSGR_07/Docs/Pdfs/RP-000013.pdf -

[51] The World's Leading Network Modeling and Simulation Environment, retrieved 3June 2005 from:http://www.opnet.com/products/modeler/home.html.

[52] ETSI TS 145 008 V6.13.0 (2005-06) Technical Specification, Digital cellular telecommunications system (Phase 2+); Radio subsystem link control (3GPP TS 45.008 version 6.13.0 Release 6).

[53] "UMTS Handover", retrieved 15 June 2006 from http://www.umtsworld.com/technology/handover.htm.

[54] 3rd Generation Partnership Project; "Universal Mobile Telecommunications System (UMTS);Radio Resource Control (RRC) protocol specification (3GPP TS 25.331 version 6.5.0 Release 6)", ETSI TS 125 331, V6.5.0, March 2005.

[55] 3rd Generation Partnership Project; "Technical Specification Group Radio Access Network; Radio resource management strategies (Release 6)", 3GPP TR 25.922, V6.0.1, April 2004

[56] J. Korhonen, *Introduction to 3G Mobile Communications*, 2nd ed., London: Artech House, 2003.

[57] M. Taferner and E. Bonek, *Wireless Internet Access Over GSM and UMTS*, Springer, 2002.

[58] L. Q.Liu, S.Berley and M. Witwit, "QoS aware queuing analysis for UMTS," *The 5th International Symposium on Wireless Personal Multimedia Communications*, Vol.2, pp. 526-530, IEEE 2002.

[59] F. Poppe, D. De Vleeschauwer and G.H. Petit, "Choosing the UMTS Air Interface Parameters, the Voice Packet Size, and the Dejittering Delay for a Voice-Over-IP Call between a UMTS and a PSTN Party", Vol. 2, pp. 805-814, *IEEE INFOCOM*, 2001.

[60] R. Rajavelsamy, V. Jeedigunta, B. Holur, M. Choudhary and O. Song, "Performance evaluation of VoIP over 3GWLAN Interworking system", WNWC 2005 *IEEE Communications Society*, Vol. 4, pp 2312-2317, 2005.

[61] G. Rittenhouse and H. Zheng, "Providing VoIP service in UMTS-HSDPA with frame aggregation", *Acoustics, Speech, and Signal Processing*, Vol. 2, pp. 1157-1160, 2005.

[62] T. Halonen, J. Romero and J. Melero, *GSM, GPRS and EDGE Performance, Evolution Towards 3G/UMTS*, 2nd ed., West Sussex: John Wiley & Sons, 2003.

[63] 3GPP, Technical Specification Group Services and System Aspects, Speech Coded speech processing functions, AMR Wideband Speech Codec, General Description (3G TS 26.171 version 5.0.0), 2001.

[64] I. Poole, *Cellular Communications Explained From Basics to 3G*, Elsevier, 2006.

[65] http://www-mobile.ecs.soton.ac.uk/speech_codecs/common_classes.html, retrieved December 2006.

[66] Support for Multimedia Session Management, retrieved 20 June 2006 from:http:/www.3gpp.org/ftp/tsg_cn/WG1_mm-cc-sm/TSGN1_05/Tdocs/N1-99670.doc, 2006.

[67] C. Perkins, *RTP: Audio and Video for the Internet*, Pearson Education, 2003.

[68] The International Telecommunication Union Telecommunication Standardization Sector (ITU-T) Recommendation G.114, "One-way transmission time", 2003.

[69] http://www.vianix.com/comparison.html, retrieved January 2007.

[70] E. Arnold, *An Introduction to queuing theory Brian D Bunday*, Edward Arnold Ltd, 1996.

[71] http://www.iec.org/online/tutorials/voice_qual/topic04.html, retrieved August 2005.

[72] N. Sulaiman, R. Carrasco and G. Chester, "Analyzing the Performance of Voice over Internet Protocol in a 3G Network", *International Conference in Innovations on Information Technology (IIT)*, 2006.

[73] N. Sulaiman, R. Carrasco and G. Chester, "Performance Evaluation of Voice Call over an IP based Network", *41st Annual Conference on Information Sciences and Systems (CISS)*, pp. 392-395, 2007.

[74] J. Yang and I. Kriaras , "Migration to All IP Based UMTS Networks", 3G Mobile Communication Technologies, *IEE Conference Publication*, No. 471, pp. 19-23, 2000.

[75] YB. Lin, AC. Pang and YR. Haung, "An All-IP Approach for UMTS Third-Generation Mobile Networks", *IEEE Network*, 2002.

[76] D. Minoli, *Voice Over IPv6: Architectures for Next Generation VoIP Networks*, Oxford: Elsevier, 2006.

[77] R. Komolafe and O. Gardner, "Aggregation of VoIP streams in a 3G mobile network: A teletraffic perspective", *PMC Conference*, Vol. 492, pp. 545-549, 2003.

[78] T. Taekyoung Kwon and M. Gerla, "Mobility Management for VoIP Service:Mobile IP vs SIP',. *IEEE Wireless communications*, Vol. 9, pp. 66-75, October 2002.

[79] M. Sief and Y. Dakroury, "Guaranteed end-to-end QoS for VoIP over cellular links based on IPv6 compression", *International Conference on 3G Mobile Communications Technologies*, pp. 296-301, 2004.

[80] http://www.wirtel.co.uk/article_globalcarriers_2005q2_003_nms.htm, retrieved August 2005.

[81] M. Lundevall, B. Olin, J. Olsson, J. Eriksson and F. Eng, "Streaming Applications over HSDPA in Mixed Service Scenarios", IEEE *VTC Fall*, Vol. 2, pp. 841-845, 2004.

[82] http://www.umtsworld.com/technology/qos.htm, retrieved August 2006.

[83] S. Morris, *Network Management, MIBs and MPLS: Principles, Design and Implementation*, Upper Saddle River: Pearson Education, 2003.

[84] "3GPP, Quality of Service (QOS) concept and architecture, TS 23.107, retrieved February 2007, from http://www.3gpp.org/ftp/Specs/html-info/23107.htm.

[85] M. Chuah, and Q. Zhang, *Design and Performance of 3G Wireless Networks and Wireless LANs*, United States: Springer, 2006."

[86] A. Sampath, N. B. Mandayam and J. M. Holtzman "Erlang Capacity of a Power Controlled Integrated Voice and Data CDMA System", *IEEE VTC*, Vol. 3, pp. 1557-1561, 1997.

[87] J. P. Castro, *The UMTS Network and Radio Access Technology Air Interface Techniques for Future Mobile Systems*, West Sussex: John Wiley & Sons Ltd, 2001.

[88] M. Zonoozi and P. Dassanayake, "Handover Delay in Cellular Systems," IEEE International Conference on Personal Wireless Communications (ICPWC), pp. 24-27, 1997.

[89] N. Sulaiman, R. Carrasco and G. Chester, "Impact of Traffic on Multi Service IP Based Applications", *International Conference on Innovative Computing, Information and Control (ICICIC)*, 2007.

[90] N. Sulaiman, R. Carrasco and G. Chester, "Performance Evaluation of Traffic Mix in IP Based Networks," *32nd IEEE Conference on Local Computer Networks (LCN 2007)*, pp. 267-268, 2007.

[91] N. Sulaiman, R. Carrasco and G. Chester, "Estimating Quality of Service of Mixed Traffic in 3G", *International Conference on Intelligence and Advance Systems (ICIAS)*, 2007.

[92] P. Chandra, *Bulletproof Wireless Security-GSM, UMTS, 802.11 and AdHoc Security*, Oxford: Elsevier," 2005.

[93] "UMTS Standard 3G T3G TS 33.102 Security Architecture", retrieved May 2007 from ftp://ftp.3gpp.org/Specs/2000-12/R1999/33_series/33102-370.zip.

[94] "UMTS Standard TS 33.900 A Guide to 3rd Generation Security", retrieved May 2007 from ftp://ftp.3gpp.org/TSG_SA/WG3_Security/_Specs/33900-120.pdf.

[95] "3rd Generation Partnership Project (3GPP); 3G Security; Security Threats and Requirements", 3G TS 21.133 version 2.0.0, retrieved June 2006 from http://www.3gpp.org/ftp/tsg_sa/TSG_SA/TSGS_03/Docs/PDF/SP-99145.pdf.

[96] "3rd Generation Partnership Project; 3G security; Security principles and objectives", 3GPP TS 33.120 V4.0.0 (2001-03) (Release 4), retrieved February 2007 from http://www.3gpp.org/ftp/tsg_sa/WG3_Security/_Specs/33120-400.pdf.

[97] "Cryptographic Algorithm Requirements", 3G TS 33.105, retrieved February 2006 from http://www.3gpp.org/ftp/Specs/html-info/33105.htm.

[98] http://www.gsmworld.com/using/algorithms/index.shtml, retrieved February 2006.

[99] V. Niemi and K. Nyberg, *UMTS Security*, West Sussex: John Wiley & Sons Ltd, 2003.

[100] http://www.gsm-security.net/gsm-security-faq.shtml, retrieved March 2007.

[101] D. M. Keon, C. Brewer, J. Carter and M. M. Taggart, "GSM and UMTS Security", accessed February 2006 from http://www.3gpp.org/ftp/tsg_sa/WG3_Security/_Specs/33120-300.pdf

[102] http://www.umtsworld.com/technology/security.htm, retrieved October 2006.

[103] P. Rysavy, "Data Capabilities for GSM evolution to UMTS", Nov 19 2002, white paper developed for 3G Americas, retrieved 5 June 2005 from: http://www.rysavy.com 1-542-386-7474.

[104] J. Bannister, P. Mather and S. Coope, *Convergence Technologies for 3G Networks IP, UMTS, EGPRS and ATM*, West Sussex: John Wiley & Sons Ltd, 2004."

[105] R. Barbieri, D. Bruschi, and E. Rosti, "Voice over IPsec:Analysis and Solutions," *IEEE Annual Computer Security Application Conference*, pp. 261–270, 2002.

[106] C. F. Grecas, S. I. Maniatis and I. S. Venieris, "Mobile Networks and Applications" Kluwer Academic Publishers, 2003.

[107] "3G Security, Security threats and requirements (Release 4), 3GPP TS 21.133 V4.1.0 (2001-12)", retrieved March 2007 from http://www.3gpp.org/ftp/Specs/html-info/21133.htm.

[108] "3rd Generation Partnership Project Technical Specification Group Services and System Aspects; 3G Security; Specification of the A5/3 Encryption Algorithms for GSM and ECSD, and the GEA3 Encryption Algorithm for GPRS; Document 1:A5/3 and GEA3 Specifications (Release 6)."

[109] A. Elbayoumy and S. Shepherd, "QoS control using an endpoint CPU capability detector in a secure VoIP system", *IEEE Symposium on Computers and Communications*, pp175-181, June 2005.

[110] F. Cao and S Malik," Security Analysis and Solutions for Deploying IP Telephony in the Critical Infrastructure," *Security and Privacy for Emerging Areas in Communication Networks*, pp. 171-180, 2005.

[111] T. J. Walsh and D.R. Kuhn, "Challenges in securing voice over IP", *IEEE Security & Privacy*, Vol. 3, pp. 44-49, 2005.

[112] M. K. Raganathan and L. Kilmartin, "Performance analysis of secure session initiation protocol based VoIP networks," *Computer Communications*, Vol. 26, pp. 552–565, 2003.

[113] J. A. Pérez, V. Zárate and Á. Montes, "Quality of Service Analysis of IPSec VPNs for Voice and Video Traffic," *Proceedings of the Advanced International Conference on Telecommunications and International Conference on Internet and Web Applications and Services (AICT/ICIW)*, 2006.

[114] E.T. Aire, B.T. Maharaj and L.P. Linde, "Implementation considerations in a SIP-based secure voice over IP network", *Africon Conference in Africa*, Vol. 1, pp. 167-172, 2004.

[115] H. F. Tipton and M. Krause, *Information Security Management Handbook*, 5th ed., United States: Auerbach Publications, 2004.

[116] http://www.rsa.com, retrieved June 2006.

[117] W. Stallings, *Cryptography and Network Security Principles and Practices*, 4th ed. Prentice Hall, 2005."

[118] N. Doraswamy and D. Harkins, *IPSec: The New Security Standard for the Internet, Intranets, and Virtual Private Networks*, 2nd ed., Prentice Hall, 2003.

[119] "RFC 2401: An overview of a security architecture", retrieved February 2007 from http://www.ietf.org/rfc/rfc2401.txt."

[120] "RFC 2402: Description of a packet authentication extension to IPv4 and IPv6", retrieved February 2007 from http://www.ietf.org/rfc/rfc2402.txt."

[121] "RFC 2406: Description of a packet encryption extension to IPv4 and IPv6", retrieved February 2007 from http://www.ietf.org/rfc/rfc2406.txt."

[122] "RFC 2408: Specification of key management capabilities", retrieved February 2007 from http://www.ietf.org/rfc/rfc2408.txt."

[123] N. Sulaiman, R. Carrasco and G. Chester, "Impact of Security on Voice Quality in 3G networks", The IEEE Conference on Industrial Electronics and Applications (ICIEA), 2008.

Appendix

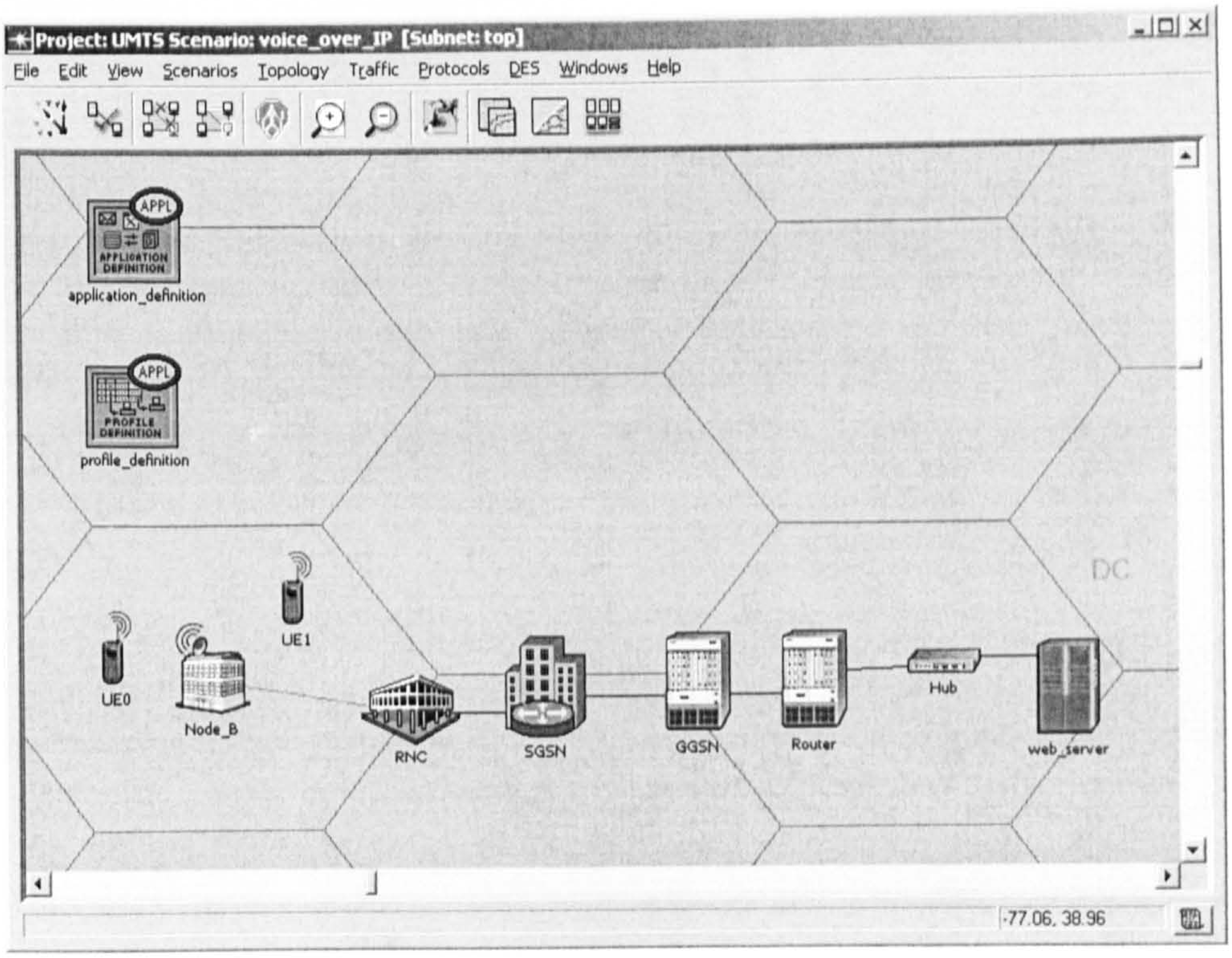
Appendix

Introduction

OPNET Modeler is a software package that enables network modelling on different level of details. It provides a graphical user interface and it is based on object oriented modelling technique where nodes as well as protocols are modelled as classes with inheritance. It consists of three main sections, the project editor, nodes level and links. The model has a detailed description of the model, broken down into the network, node and process level implementations. Every link and nodes has its attributes that can be changed.

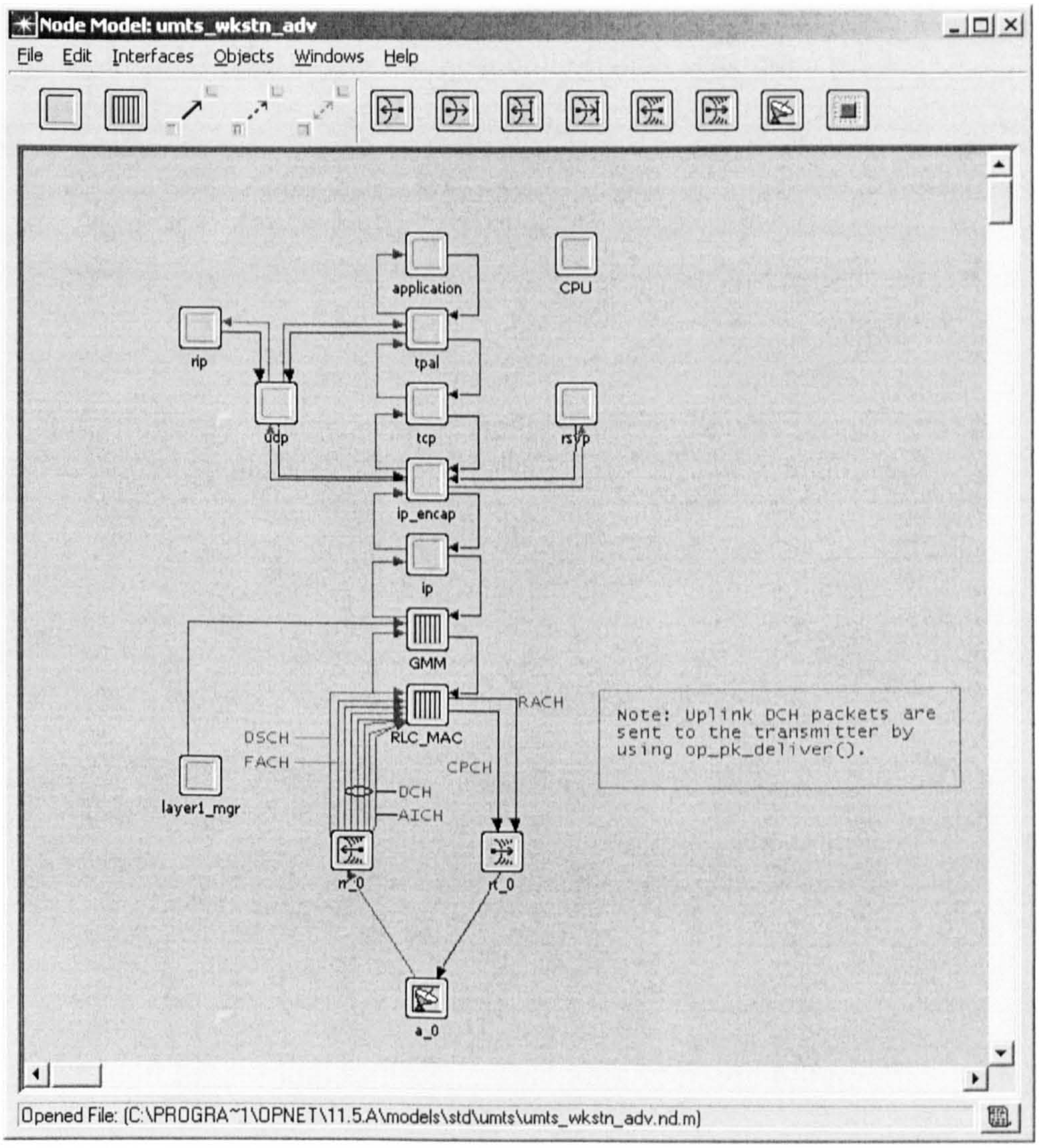
Project Layer

A project enables a basic overview of a network and it can have several scenarios with different attribute in order to compare result. The project layer enables the user to define the general network topology on a geographical map. The network is defined in terms of the scale of the network (e.g. world, enterprise, campus, office, etc), size of the network (x and y span in degrees, metres, kilometre, etc), technologies to be used (e.g. WLAN, Ethernet, UMTS, etc), and nodes and links. A simple network for UMTS model is as shown below.



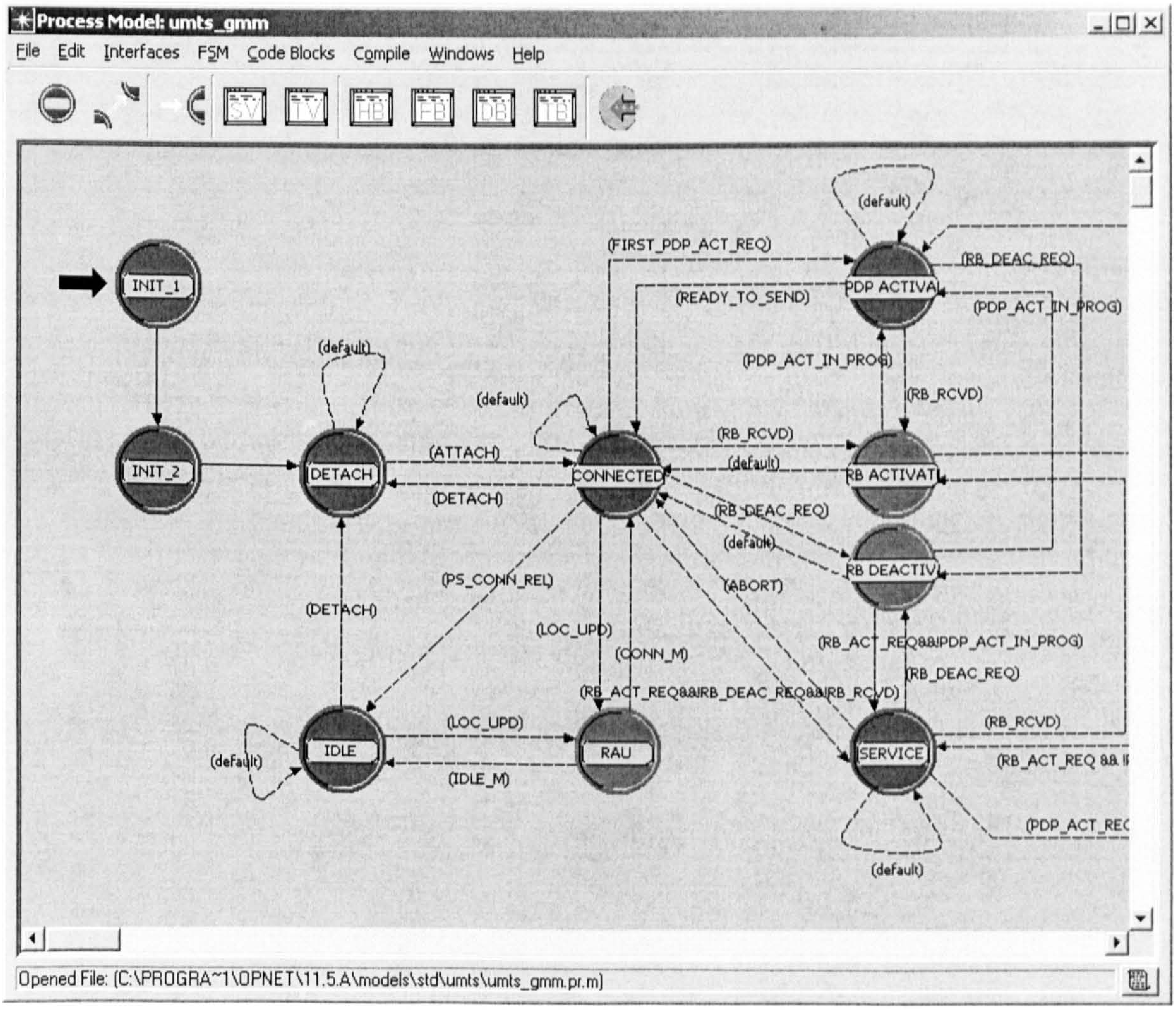
Node Level

The node editor illustrates all the protocols where the arrows between the protocols simulate incoming and outgoing packet streams. Nodes layer provides functionality to define the structure and internal configuration of the network elements. It defines the behaviour of each network object defined in the project model. Behaviour is defined using different modules, each of which models some internal aspects of node behaviour such as data creation, data storage etc. Nodes are build out of processors, queues, transmitters and receivers where there are interconnected with packet streams, statistic wire or logical associations. The node layer for a mobile station is as shown below.



Process Layer

The process layer enables configuration of process inside the node and direct implementation of communication protocols. The process modelling defines the underlying functionality of the node model. Finite state machine is used to define the actions of different protocols where each protocol is always in one specific state at any instant of time. Each state consists of several lines of C or C++ code. Transitions describe the possible movement of a process from state to state and the conditions allowing such change. The state transition in handover process will be modelled and simulated. A discrete-time approach will be used to evaluate handover probability and its performance. In addition, OPNET had a specific UMTS model built in which is based on 3GPP specification. It provides the UTRAN parts such the Node B and the RNC. The Mobility Management and Session Management (GMM) layer for the mobile station is as show below.



A process can remain in a static state, waiting for an event to occur. Figure below shows a sample of OPNET source code.