جامعة الشرق الأوسط
MIDDLE EAST UNIVERSITY

Secure Authentication Protocol and Key Agreement for Mobile Networks Using Public-Key Cryptography

A Thesis submitted in partial fulfilment of the requirements for Master in Computer Science

By

Haneen Mohammed Al-Fayoumi

Department of Computer Information System
Faculty of Information Technology

Supervisor

Prof. Nidal Shilbayeh

Middle East University

(March 2011)

تفويض

أنا حنين محمد الفيومي أفوض جامعة الشرق الأوسط بتزويد نسخ من رسالتي للمكتبات أو المؤسسات أو الهيئات أو الأفراد عند طلبها.

التاريخ: ٢٢/٦/ ٢٠١١

التوقيع: حنين الفيومي

Authorization statement

I, **Haneen Mohammad Al-Fayoumi,** Authorize the Middle East University to supply a copy of my thesis to libraries, establishments or individuals upon their request.

Date: 06/22/2011

Signature:

Committee Decision

This Thesis (**Secure Authentication Protocol and Key Agreement for Mobile Networks Using Public-Key Cryptography**) was successfully defended and approved on April 19th, 2011

Examination Committee signatures:

Prof. Nidal Shilbayeh
Professor, Computer Science Department
(Middle East University)

Dr. Bilal Abu Alhaija
Assistant Professor, Computer Science Department
(Middle East University)

Dr. Mohammad Ahmad Alia
Assistant Professor, Computer Information System Department
(Al-Zaytoonah University)

Dedication

To my lovely mother, who gave me endless love, trust, constant encouragement over the years, and for her prayers.

To my Family, for their patience, support, love, and for enduring the ups and downs during the completion of this thesis.

This thesis is dedicated to them.

Acknowledgements

In the Name of Allah

I would like to thank my supervisor, prof. Nidal Shilbayeh for his trust and faith in me, and for constantly inspiring me and for lecturing us in Wireless Communications and showing keen interest in the subject matter and reviewing the thesis. His constant encouragement and enthusiasm for my work resulted in the successful completion of my thesis. I would also like to thank Prof. Caroline Strange for serving in my supervisory committee and for her guidance and assistance during my research.

My sincere thanks are also extended to my father Prof. Mohammad Al-Fayoumi and my uncle Dr. Mustafa Al-Fayoumi for the precious help they gave and for being a light on the dark path of my study. Their trust and encouraging comments on this research enabled me to reach the end.

Finally, I would like to express my special thanks to my husband for his support and patience during my research.

Abstract

One of the considerable difficulty facing the mobile networking is security. Thus, a secure and an efficient authentication scheme is required for mobile communication systems. The authentication protocol encounters overheads on the transmission process. These overheads influence the mobile network performance such as delay and bandwidth. The main objective of this thesis is to improve authentication scheme in mobile networks by generating a complete solution for the authentication scheme in mobile networks to enhance the security level and to improve the efficiency.

The analytical result and a simulation program were employed in this thesis, to consider the existing and proposed methods for authentication scheme in mobile networks. An improvement authentication scheme is proposed for reducing the number of messages between authentication users in the network. Therefore, the bottleneck at the authentication centre is avoided by reducing the number of messages between mobile station and the authentication centre. As a result , the authentication time delay, call setup time, signaling traffic and the number of transmissions between the home network and visited networks for roaming authentication are minimized. Also, the proposed scheme was considered to be secure against network attacks.

الملخص بالعربية:

يعتبر الامان احدى المشاكل الكبيرة التي تواجه شبكات الموبايل حيث يعتبر الامان واثبات كفاءة النظم قضية اساسيه لنظم اتصالات الموبايل ، حيث ان برتكول الاثبات يقاوم العبئ الزائد من المشاكل التي تظهر اثناء عملية معالجة الارسال للنطاق الترددي ، ومثل هذه المشاكل توثر على اداء شبكة الموبايل منها مثلا التاخير في معالجة الارسال ضمن النطاق الترددي ، حيث ان الهدف الرئيسي لهذه الرسالة هو تحسين نظام الاثبات في شبكات الموبايل من خلال استحداث حل كامل لنظام الاثبات في شبكات الموبايل لتعزيز مستوى الحماية وتحسين الكفاءة .

ان نتائج التحليل وبرنامج المحاكاة الذي طبق في هذه الرسالة هو لدراسة الطرق الموجودة ولاستخدامهما في الطريقة المقترحة لنظام الاثبات في شبكات الموبايل ، ان نظام تحسين الاثبات الذي اقترح هو لتقليل عدد الرسائل المتداولة بين مستخدمي الاثبات في الشبكة ، لذا فان الاختناق في مركز الاثبات تم تجنبه من خلال تقليل عدد الرسائل المرسلة بين محطة الموبايل ومركز الاثبات . ونتيجة لذلك فان التاخير في وقت الاثبات ، ودعوة اشارة الاستعداد للعمل ، واشارة حركة السير ، وعدد التنقلات بين الشبكة الداخلية والشبكات المزارة لاثبات عملية التجوال تم تقليلها .. كذلك فان النظام يوفر الامان ضد الهجمات على الشبكة .

Table of Contents

Chapter 1 Introduction.. 1

1.1 Background.. 1

1.2 Statement of problem.. 4

1.3 Research Objectives 5

1.4 Motivations.. 5

1.5 Significance of research 5

1.6 Limitations.. 6

1.7 Organization of the thesis... 6

Chapter 2 Literature review & Related Work.. 8

2.1 Introduction 8

2.2 Security attacks in mobile system 8

2.2.1 Kinds of attacks in mobile system 9

2.3 Security Services for Mobile Communications 10

2.3.1 Mutual authentication 10

2.3.2 Confidentiality (location, privacy and confidentially) 11

2.3.3 User anonymity 11

2.3.4 Data integrity 12

2.3.5 Non-repudiation... 12

2.3.6 Authorization.. 12

2.3.7 Minimize the resources utilization in a scheme 13

2.4 Security mechanisms.. 13

2.4.1 Symmetric-key cryptography (Single-key cryptography)................................. 14

2.4.2 Asymmetric-key cryptography (Two-key cryptography) 16

2.5 The Authentication Protocol 19

2.5.1 The UMTS authentication protocol.. 19

2.5.2 UMTS architecture 20

2.5.3 The authentication scheme 23

2.5.4 The confidentiality scheme... 31

2.5.5 The integrity protection.. 34

2.5.6 The subscriber identity/location confidentiality 36

2.6 Security considerations and threats for UMTS... 37

2.7 Related Work 38

2.8 Summary.. 41

Chapter 3 An Enhancement of Authentication Protocol and Key Agreement (AKA) for 3G Mobile Networks...43

3.1 Introduction 43

3.2 Authentication framework.. 45

3.3 The Operation modes.. 50

3.4 Description of the proposed protocol.. 51

3.4.1 Initial authentication procedure 52

3.4.2 Subsequent authentication procedure 57

3.5 Achieved goals 59

3.6 Security analysis of the proposed protocol 60

3.7 Summary 64

Chapter 4 Analysis of the Proposed Authentication and Key Agreement (AKA) Protocol........ 65

4.1 Analysis of UMTS authentication protocol 66

4.1.1 Signalling load of the UMTS protocol 68

4.1.2 Authentication delay of the UMTS protocol 73

4.1.3 Bandwidth requirement for the UMTS protocol 74

4.2 Analysis of the proposed authentication protocol 78

4.2.1 Signalling load of the proposed protocol 78

4.2.2 Authentication delay of the proposed protocol 80

4.2.3 Bandwidth requirement of the proposed protocol 81

4.3 Discussion of analytical and simulation results 85

4.4 Summary 91

Chapter 5 Conclusions and Future Work 92

5.1 Conclusions 92

5.2 Future work 94

References 96

Appendices 108

List of Figures

Figure 2.1: Diffi-Hellman Key Exchange. 18

Figure 2.2: UMTS architecture. 23

Figure 2.3: Authentication and Key Agreement (AKA). 29

Figure 2.4: Generation of authentication vectors. 30

Figure 2.5: User authenticatin function in the USIM. 31

Figure 2.6: Encryption in radio access networks (UMTS) 32

Figure 2.7: Ciphering of user and signaling traffic in UMTS. 34

Figure 2.8: Message authentication code. 35

Figure 3.1: Proposed Initial auhtentication procedure. 55

Figure 3.2: Subsequent authentication procedure. 57

Figure 4.1 : Registration Area (RA) and UTRAN Registration (URA) Layout. 67

Figure 4.2: UMTS challenge/response signalling messages flow for registration 70

Figure 4.3: UMTS call origination authentication signalling messages 70

Figure 4.4: Signalling message flow for UMTS authentication delay 74

Figure 4.5: Signalling messages flow for the proposed scheme 80

Figure 4.6: The proposed call origination/termination authentication signalling meaasges80

Figure 4.7: Comparison of the authentication delay between UMTS and the proposed protocol when $T_{DB} = 1ms$ 87

Figure 4.8: Load transaction messages per second between entities. 88

Figure 4.9: Total signalling messages/second for all activity in current and proposed protocol. 89

Figure 4.10: Network signalling traffic with different mobility rate. 90

Figure 4.11: The relationship between authentication delay and the number of MS (simulation results) 90

List of Tables

Table 2.1: AKA functions with their outputs....................24
Table 3.1: Software speeds of RSA....................52
Table 3.2: Comparison of the recommendation of NIST and RSA Laboratory 61
Table 4.1: Assumptions parameters 69
Table 4.2: Total Authentication request per VLR (serving area) and HLR....................72
Table 4.3: Signalling messages per Authentication request for each activity....................73
Table 4.4: Total signalling traffic and load transaction messages per each activity 73
Table 4.5: Authentication parameters for UMTS protocol....................75
Table 4.6: the Bandwidth between entities for UMTS protocol 77
Table 4.7: Signalling messages per Authentication request for each activity....................79
Table 4.8: Total signalling traffic and load transaction messages per each activity 79
Table 4.9: Authentication parameters for proposed protocol 81
Table 4.10: The bandwidth between entities for proposed protocol....................85
Table 4.11: Comparison of signalling messages between current and proposed authentication protocol....................86
Table 4.12: Comparison of the total signalling traffic and load messages/s between entities for each activity. 87
Table 4.13: Comparison of the total signalling traffic and load messages/sec between entities. 87
Table 4.14: Comparison of the bandwidth for each activity between databases and MS-VLR. . 87
Table 4.15: Network signalling traffic with different mobility rate. 89

Abbreviations

1G	First Generation
2G	Second Generation
3G	Third Generations
3GPP	Third Generation Partnership Project
A3	Authentication algorithm A3
A3/A8	A single algorithm performing the functions of A3 and A8
A5/1	Encryption algorithm A5/1
A5/2	Encryption algorithm A5/2
A8	Encryption key generating algorithm A8
AES	Advance Encryption Standard
AK	Anonymity Key
AKA	Authentication and Key Agreement Protocol
AMF	Authentication Management Field
AP-AKA	Adaptive Protocol-AKA
AuC	Authentication Centre
AUTN	Authentication Token
AV	Authentication Vector
BS	Base Station
BSC	Base Station Controller
BSS	Base Station System
BTS	Base Transceiver Station
CDMA	Code Division Multiple Access
CK	Confidentiality Key
CN	Core Network
CS	Circuit Switch
EIR	Equipment Identification Register
ESN	Electronic Serial Number
ETSI	European Telecommunication Standard Institute
FIFO	First In First Out
GGSN	Gateway GPRS Support Network
GPRS	General Packet Radio Switching
GSM	Global System for Mobile Communications
HE	Home Environment

HLR	Home Location Register
HMAC	Hash Message Authentication Code
HN	Home Network
ID	Identifier
IK	Integrity Key
IMEI	International Mobile station Equipment Identity
IMSI	International Mobile Subscriber Identity
IMT-2000	International Mobile Telecommunication for year 2000
ITU	International Telecommunication Union
Iu	UMTS Interface between 3G-MSC/SGGN and RNC
Iub	Interface between Node B and RNC
Iups	Interface between RNC and CN
Iur	UMTS Interface between RNCs
LAC	Location Area Code
LAI	Location Area Identifier
MAC^1	Message Authentication Code
MAC^2	Media Access Control
ME	Mobile Equipment
MIN	Mobile Identification Number
MNC	Mobile Network Code
MS	Mobile Station
MSC	Mobile-services Switching Centre, Mobile Switching Centre
OSI	Open System Interconnection
PLAU	Periodic Location Area Update
PLMNs	Public Land Mobile Networks
PS	Packet Switching
PSTN	Public Switched Telephone Network
RAN	Radio Access Network
RAND	Random Number
RAs	Registration Areas
RLC	Radio Link Layer
RNC	Radio Network Control
RNS	Radio Network System
RRC	Radio Resource Control
RSA	Rivest Shamir Adleman
SGSN	Serving GPRS Support Network

SHA-1	Secure Hash Algorithm-1
SIM	Subscriber Identity Module
SMS	Short Message Service
SN	Serving Network
SQN	Sequence Number
SRES	Signed Response
TMSI	Temporary Mobile Subscriber Identity
UE	User Equipment
UMTS	Universal Mobile Telecommunications System
USIM	Universal Subscriber Identity Module
UTRAN	UMTS Terrestrial Radio Access Network
Um	Interface between MS (Mobile Station) and BSS
Uu	radio interface for UTRA
VLR	Visitor Location Register
VN	Visited Network
XMAC	Expected Message Authentication Code
XRES	Expected Response
X-AKA	Extension Protocol-AKA

Chapter 1 Introduction

1.1 Background

With the wide spread of wireless communication and computer technology, mobile communication gives more versatile, portable and affordable networks (Passerini, et al, 2007) (Yuan, et al, 2007) (Wadekar & Fagoonee, 2006) than ever. As a result, the number of users using mobile communication networks has increased quickly. The modification of communication not only brings a new collection of technical problems, but also raises a new category of exciting applications. This is because of the change in communication from single-medium oriented into multimedia oriented such as image, Internet services, e-commerce (Fitzek, et al, 2002) (Iftikhar, et al, 2007), and so on. At the same time, to decrease the possibility of masquerading, and protect privacy on the radio channels are a very important matters (3GPP, 1999a) (3GPP, 1999b) (3GPP, 1999c).

Since the transmission of information through insecure communication channels are unprotected, security will be the most important requirement for the exchange of user's or systems' private information. Therefore, preventive security measures for mobile communication systems must be provided. As a solution to prevent the unauthorized access of frauds and eavesdroppers, authentication and confidentiality are important security services to subscribers and the service provider (Stalling, 2003), (Kaaranen, et al., 2005). Take cellular mobile communication systems for instance, entities of a cellular mobile communication system include:

- A mobile station (MS), which is on behalf of a user.

- A home network (HN), with which the MS contracts.
- A foreign network, which is called serving network (SN). An MS can connect to an HN or an SN.

Commercial mobile communication systems can be track back to late 1970’s. Today, the mobile communications networks are commonly divided into three generations, these are first, second and third generations. Currently, the third generation system is deployed in many nations and the fourth generation of mobile communication systems is now under development. The main difference between the generations is the construction technology (analogue or digital) and the services they offer.

The first generation (1G) was presented in 1980 (Safavi-Naini, et. al, 2001), where the construction technology used in this generation is analogue system, which transited straight from original wire-typed telephone system into mobile system. This generation has several options such as: Nippon Telephone and Telegraph Corporation (NTT), Total Access Communication System (TACS), and Advanced Mobile Phone System (AMPS). However, the low cost of its equipment and its security problems are not correctly tackled. This gives an opportunity to impostor to be able to listen in or intercept user traffic through radio interface or even change the identity of mobile phone to get unauthorized services.

To fix the security difficulties in the 1G, the second generation (2G) cellular mobile communication was presented in 1990 (Safavi-Naini, et. al, 2001). The second generation mobile using the construction technology of digital system; this generation afforded reliable voice communication, and has several types such as; United States Digital Cellular (USDC) using Time Division Multiple Access (TDMA), IS-95 Code Division Multiple Access CDMA using Direct Sequence Code Division Multiple

Access, and Direct Sequence CDMA (DS-CDMA) and Global System for Mobile (GSM).

The third generation (3G) introduced in 1995 (Safavi-Naini, et. al, 2001) while the International Telecommunication Union (ITU) started developing IMT-2000 (International Mobile Telecommunication for the year 2000). The main requirements of the IMT-2000 involve support for a data rate of 144 Kbit/s for users in fast-moving vehicles over large areas, and for pedestrians at a rate of 384 Kbit/s and 2.048 Mbit/s operations for office use (Safavi-Naini, et. al, 2001). To conquer the security difficulties in GSM an emerging standard for 3G digital cellular systems the Universal Mobile Telecommunication System (UMTS), adopts an improved authentication and key agreement protocol (AKA) recommended by the Third Generation Partnership Project (3GPP) (3GPP, 2001d) (3GPP, 2001e).

The UMTS authentication protocol has the framework of the GSM and add a new improvement characteristics such as mutual authentication, agreement on an integrity key between the users and the serving network (SN), and guaranteed freshness of agreed encryption key and integrity key. According to the characteristics of the Message Authentication Code (MAC), the mobile station and the home location register (HLR) in the home network can achieve mutual authentication by sharing the same private key in advance.

A new authentication scheme will proposed to satisfy the security requirements of the third generation mobile systems and enhance performance by reducing the communication times, and by creating few authentication messages and data sizes during the process of authentication. The suggested protocol considerably will decrease the communication overhead between the home network and the visited network mainly for roaming authentication. Also, the suggested protocol will be secured against network attacks, such as the replay attack, guessing attack, substitution attack, and impersonating attack.

1.2 Statement of Problem

The UMTS protocol has some drawbacks such as:

1. The bandwidth consumption between Serving Network (SN) and Home Network(HN)
2. Storage space of SN
3. The sequence number (SQN) synchronization (Huang & Li, 2005)
4. The International Mobile Subscriber Identity (IMSI) that uniquely identifies a user, is still disclose to the visited network (Gódor & Imre, 2006) and can still be demanded by a hacker who impersonates a base station, as there is no network authentication in this case
5. The non-denial services requirement which give the protection for the subscribers from incorrect bill charging, and the service providers with legal evidence when collecting the bills, are two important points in the non- denial requirement. Since, the true non- denial service among MS, Visit Location Register (VLR) and HLR can only accomplished by a public-key scheme using digital signatures (Harn &Hsin, 2003), then both GSM and UMTS ignore the requirement.

According to these concepts, security for wireless networks is increasingly needed. Thus, security system is proposed to protect communications, but this will add more overheads on the transmission.

1.3 Research Objectives

The objectives of this thesis are as follows:

i. Investigate the Authentication and Key Agreement protocols for the Universal Mobile Telecommunications System (UMTS).

ii. Suggest a new authentication protocol to conquer the security problems in the present protocol and enhance the performance of mobile networks.

1.4 Motivations

1. Due to the fast growth of wireless technology and wireless services, a detailed observe the issue security is needed. Mobile networks are protected by using authentication security systems

2. The authentication protocol incurs overheads on the transmission. These overhead involves the mobile network performance such as delay, bandwidth allocation efficiency. This needs intensive research and improvement in order to reach the satisfaction of the mobile user.

3. Several authentication protocols have been proposed to improve the security of mobile schemes, but none of them can satisfy the security requirements of third generation systems (Cheng, et al. 2005).

4. Few of researchers studied the relationship between security and performance, and few of them introduce a new cryptography scheme to be fitted for mobile network.

1.5 Significance of Research

1. Wireless communication is a technology that is becoming a feature in many aspects of daily life.

2. Mobile phone systems have been enhanced by other applications such as e-commerce, e-learning, e-voting and e-business. The radio signal transmitted by the mobile phone is accessible to everyone.
3. The authentication process in mobile network provides a reasonable security level against fraudulent and eavesdropping. The authentication protocol incurs overheads on the transmission process. The overheads have an effect on the mobile network performance in terms of the signaling traffic, time delay and the bandwidth. The signaling load and the authentication delay are of particular importance and have become the subject of widespread research interest.

1.6 Limitations

Testing by analytical model and software simulation has proved that the proposed protocol is efficient and robust. However, real network validations are still required. The results of real experiments would support the effectiveness and robustness of the proposed protocol.

1.7 Organization of the Thesis

In addition to the introduction, there are four other chapters. Chapter 2 describes the security of mobile networks and major terminologies relevant to the contents of the contributions which are made. Related work relevant to mobile network security is investigated in order to assist the conduct of this research. The authentication protocols designed for mobile networks are described. Chapter 3 proposes a secure authentication mechanism for mobile communication systems by integration the proposed of pubic key cryptography and hash chaining function. The new approach proposed is able to remedy the failings of the UMTS. The proposed protocol reduces the network traffic and signalling messages between entities, and consequently the bottleneck at the authentication centre is avoided. It includes the security analysis of the proposed protocol. Chapter 4 includes the efficiency analysis of the proposed protocol and the performance comparison of the proposed scheme with the protocol of UMTS. A fluid mobility model is used to investigate the performance of signalling traffic and load transaction messages between mobile databases such as Home Location Register (HLR) and Visitor Location Register (VLR). Finally, chapter 5 draws conclusions and suggests future work in this research area.

Chapter 2 Literature review & Related Work

2.1 Introduction

Secure mobile networks will help to increase productivity and efficiency. However, this technology leaves sensitive information open to attack, since wireless networks increase security risks.

It is necessary to consider the security requirements of any firm to evaluate and choose security policies. Then, to define the needs of security and the methods to satisfy those needs. One method is to consider three aspects of information security according to OSI (Open System Interconnection) (Stalling, 2003).

1. **Security attacks** are categorized as either passive attacks, which include unauthorized reading of a message and traffic analysis, or active attacks, such as modification of messages, and denial of service.
2. **Security services** are a processing or communication service that is provided by a system to give a specific type of protection to system resources. Security services employ security policies and are executed by security mechanisms. They involve authentication, access control, data privacy, data integrity, and non-denial.
3. **Security mechanism** is any process that is considered to detect, prevent, a security attack. Examples of mechanisms are encryption algorithm, digital signature, and authentication protocol.

2.2 Security attacks in mobile system

A security system is a method for protecting data in computer and communication system. Therefore, in order to design a security system it is necessary to take into account three main considerations. Firstly, vulnerability which is a point where a system is subject to attack. Secondly, threat by an unwanted event that causes damage or trouble to information systems or services. Thirdly, countermeasures, the techniques for protecting the system.

2.2.1 Kinds of attacks in mobile system

i. **Eavesdropping (Interception):** an unauthorised party gains access to a data. This is an attack on confidentiality like tapping a conversation between parties. This attack is a passive attack. Here the attacker could be eavesdropping on network traffic between the transmitter and reciever to capture data in a network without altering the information itself. The countermesure against this attack is encryption(Stalling, 2003).

ii. **Modification (Tampering):** an unauthorised party alters the content of a message which is transmitted between entities. In other words, the information is altered and then sent to the recipient. This is an attack on integrity like changing the content of message being transmitted. The contermesure against this attack is cryptographic technique (checksums or digital signature) (Stalling, 2003).

iii. **Fabrication (Impersonation):** an unauthorised party inserts counterfeit objects into the system, or pretends to be some other party. This is an attack on authenticity. Examples include the insertion of false messages (e.g. signalling messages in the GSM) in a network and the addition of records to a file. The countermesure against this attack is cryptographic technique (Ford, 1994) (Stalling, 2003).

iv. **Data transmission interruption (jamming):** the action of preventing a message from reaching its intended recipient. It can also occur when an asset of a system is destroyed or becomes unavailable or unusable. This is an attack on availability. An asset of a system becomes unavailable. The attacker may cut the communication line or use jamming to interrupt wireless communications. These are all denial of service attacks which are difficult to prevent (Stalling, 2003).

2.3 Security Services for Mobile Communications

There is a need to incorporate security techniques into communication systems to protect information from passive eavesdropping and active tampering (Boyd & Park 1998) (Lin, 1999) (Stach, et al., 1998). Otherwise payment may be made for calls made by an impostor. Air interfaces expose the content of communication so that exposure makes it difficult to maintain confidentiality and control fraud (Liang &Wang 2004) (Stalling, 2003). It is worth considering security requirements for mobile communication systems.

2.3.1 Mutual authentication

Mutual authentication enables communicating parties to verify the identity of the other and to exchange a session key. One important tool to achieve authentication is the digital signature (Ford, 1994) (ISO, 1996). The existing GSM only provides a unilateral authentication scheme that ensures only authorized devices gain access to the network. So, the user side is unable to verify the visited network and its home network. In addition, there is no mutual authentication between visited network and home network because it assumes the communication path between them is fully secure. Consequently, impersonation attacks may occur in this unilateral authentication scheme such as a false base station. The UMTS authentication protocol remedies these weaknesses. This eliminates the chance that illegal users make fraudulent phone calls and thwarts fraud (Menezes, et al., 1997).

2.3.2 Confidentiality (location, privacy and confidentially)

This is the property of ensuring that information is only disclosed to authorized individuals, entities or processes. In other words, confidentiality is the protection of transmitted data from passive attack, or anyone who is not authorised to access it. Encryption provides confidentiality (Gollmann, 1999). For user traffic confidentiality, privacy is only achieved in the radio portion (MS ↔.base station) of current GSM systems. UMTS extends the user traffic confidentiality to the Radio Network Controller (RNC). After leaving the Base Station BS or RNC, data will be decrypted and transmitted in a plaintext form over the networks.

2.3.3 User anonymity

This user's identity and his location are valuable information, and identity should be concealed from potential eavesdroppers when the true identity is compromised when roaming. The disclosure of this sensitive information may result in serious consequences, especially if the subscriber is a very important person (VIP). Anonymity is currently provided by use of temporary identities for communication. However, in the case of new registration and roaming, the genuine identity is necessary. The existing mechanisms of GSM and UMTS have experienced some security breaches, which may cause the protection to this sensitive information to fail (Ford, 1994) (ISO, 1996).

2.3.4 Data integrity

The integrity service addresses both message stream modification and denial of service. In general, data integrity services are complementary to data confidentiality services. A variety of mechanisms is used to assure the integrity of a data unit or stream of data units. It basically means that the information exchanged in an electronic data transfer is not alterable without detection. Modification types include writing, changing, deleting, etc. Integrity is achieved through several techniques such as: checksums, message digests, or digital signature (Putz, et al., 1998) (Gollmann, 1999) (Stalling, 2003).

2.3.5 Non-repudiation

Non-repudiation prevents the sender or receiver from denying a transmitted message. Such a service can be used for a secure billing service to ensure that users cannot deny having requested a certain value added service. It provides protection against denial by one of the entities involved in a communication of having participated in all or part of the communication, neither sender or receiver can deny a transmitted message. In other words, non-repudiation is the ability to prove that an action or event has taken place (ISO, 1996). There are two types of non-repudiation. Firstly, non-repudiation with proof of origin, which proves that the message was sent by the specified party. Secondly, non-repudiation with proof of destination, which proves that the message was received by the specified party (Stalling, 2003).

2.3.6 Authorization

Authorization is the process which determines if a person has permission to conduct a particular action. Authorization is related to the existence of a security policy, which is a set of rules that specifies which action is permitted and which is prohibited

(Gollmann, 1999) (ITU, 1991). Unauthorized access includes unauthorized use, disclosure, modification, destruction, and issuing of commands. It requires that access to protected resources is controlled.

2.3.7 Minimize the resources utilization in a scheme

The resources of a mobile device are limited. Therefore, minimization of resource utilization is important when designing a security scheme.

These services are basic requirements for the safety of mobile communication, but they do not guarantee perfect security for the system.

Many services and applications are not standardized. In general, it is difficult to predict their exact nature. Therefore, European Telecommunication Standard Institute ETSI draws up the specification (3GPP, 1999a) for authentication protocol design. The security analysis performed relies on previous experience with second generation systems (in particular GSM) and takes into account known problems from that area. The security requirements listed in that specification are used as input for the choice of security features and the design of authentication protocols. In this thesis, the specification and the general objectives for 3G security features (3GPP, 1999a) (3GPP, 1999c) will be taken into account to design a secure new authentication protocol.

2.4 Security mechanisms

There is no single mechanism that can provide all the security services that are mentioned in section 2.3. However, cryptographic mechanisms underly most of the security mechanisms in use (Al-Muhtadi, et al., 2002) (Stalling, 2003). This section describes the security mechanisms and highlights important concepts which are important to this thesis.

Cryptography is used as the basis for much computer security as it can keep information confidential and can also preserves the integrity of data particularly when begin stored or transmitted. It is the science of secret writing (Stalling, 2003).

Cryptography is used to transform original information, called plaintext or clear-text into transformed information, called ciphertext or code-text, or simply cipher, and the process of producing this cipher text is known as encryption or enciphering. It sends messages in a masked form so that only the intended receiver can remove the mask and get the message. Any one not in possession of the proper cipher algorithm and keys cannot read the information. On the other hand, cryptanalysis is the study of how to attack cryptosystems. In general, cryptography is classified according to the number of keys used:

i. Symmetric key cryptography (Private-Key cryptography).

ii. Asymmetric key cryptography (Public-Key Cryptography).

2.4.1 Symmetric-key cryptography (Single-key cryptography)

In the symmetric-key system an algorithm uses a key to convert information into what looks like random bits before it is sent. Then, the same algorithm uses the same key to recover the original data when it is received. Symmetric-key encryption can keep your secrets safely, but because you need your keys to recover encrypted data, you must also keep them safely. So key-management in symmetric-key cryptography systems is very important.

The structure has a venerable history related to conventional cryptosystems. Some examples of symmetric algorithms are DES (Data Encryption Standard), AES (Advanced Encryption Standard), Twofish, Serpent, Blowfish, CAST5, RC4, Triple

DES, and IDEA (International Data Encryption Algorithm). The original intelligible message, referred to as plaintext, is converted into apparently random nonsense, referred to as cipher text (Menezes, et al., 1997). Encryption and decryption with a secret-key cryptosystem are denoted by:

$$E_K(M) = C$$

$$D_K(C) = M$$

This notation indicates that the ciphertext, C, is produced by using encryption algorithm E as a function of the plaintext, M, with the specific function determined by the value of the key, K. The intended receiver, in possession of the key, is able to invert the transformation (Stalling, 2003).

Symmetric-key cryptography systems are also classified into two types along the way in which the plaintext is processed (Stalling, 2003):

i. Stream Cipher: this operates on the plaintext a single bit or sometimes a byte at a time. With a Stream Cipher, the same plaintext bit or byte will encrypt to a different bit or byte every time it is encrypted (Stalling, 2003).

ii. Block Cipher: this operates on the plaintext in groups of bits. The groups of bits are called blocks. Examples may be 64 bit, 128 bit or 256 bit encryption. This technique is very fast, because little processing is required. This technique does, however, have the drawback that identical blocks of data will produce the same cipher text under the same block cipher key (Stalling, 2003).

The symmetric key cryptosystem has some drawbacks as follows:

i. In order to use a secure channel, it requires prior communication of the key between sender and receiver before any ciphertext is transmitted. In practice, this may be very difficult to achieve, because there is no security channel in wireless communication systems.

ii. It requires a large amount of keys. For a cryptosystem with n users, since each user has to possess $n-1$ keys, the required total number of keys are $n(n-1)/2$. Thus, when the number of users increases, the risk of revealing the secret information was drastically increased.

2.4.2 Asymmetric-key cryptography (Two-key cryptography)

The first major development of the asymmetric-key (public key cryptography) occurred in 1976 when Diffie-Hellman published their well-known paper on new directions in cryptography (Diffie & Hellman, 1976). Diffie-Hellman suggested a great concept for public key cryptography and developed a scheme without a secure communication, but able to provide secret communication. Diffie-Hellman suggested a technique for distributing the private key to be employed in the classical schemes in an insecure communication channel (Bruce, 1996). The concept of asymmetric-key cryptography evolved from an attempt to solve two of the most difficult problems associated with conventional encryption. The first problem is that of key distribution. The second problem is that of "digital signatures".

There are many asymmetric cryptosystems, such as RSA (Rivest, et al. 1978), Rabin (Rabin, 1979), ElGamal (ElGamal, 1985), and Elliptic curve cryptography (Caelli, et al., 1999). Their security bases are dissimilar. The factoring problem (RSA) and the discrete logarithms problem (ElGamal, Elliptic curve cryptography) are the two important problems in cryptography.

2.4.2.1 Asymmetric encryption

In this scheme two mathematically related keys are used, one key (public key) to encrypt, and the other (private key) to decrypt. Although they are related to each other and they are significantly different. An attacker can have access to a cryptography public key and still not be able to decrypt the data. If the owner of the private key keeps it private, plaintext encrypted with the public key will remain secure. Asymmetric encryption can be used to provide confidentiality.

The essential steps of an asymmetric-key are as follows:

1. Each entity in a network generates a pair of keys (a public and a private key) to be used for encryption and decryption of messages respectively.
2. Each entity publishes its encryption key by placing it in a public domain or file. This is the public key. The private key is kept private by the owner.
3. If user A wishes to send a message to user B, then A encrypts the message by using B's public key.
4. When user B receives the encrypted message, then B decrypts it by using B's private key. No other recipient can decrypt the message because only B knows B's private key.

The sender and receiver each use a different key, therefore the system is referred to as an asymmetric cryptosystem. The cryptosystem is called "public-key" because the encryption key can be made public.

Encryption and decryption with the public-key cryptosystem are denoted by:

$$E_{K_{UB}}(M) = C,$$

$$E_{K_{PB}}(C) = M.$$

There is some source A for a message, which generates a message in plaintext M. Along with the message M, the encryption public key K_{UB} for the user destination B, these will be as input parameters to perform the encryption algorithm. A produces the cipher text C. The intended receiver B, in possession of the corresponding private key K_{PB}, is able to recover the original message M by using the decryption algorithm.

2.4.2.2 Diffie-Hellman Key Exchange

The Diffie-Hellman key Exchange mostly describers the algorithm that enables two users to exchange a key securely, and the shared key is used for encryption of messages. Next we detail introduce the procedure as following: Figure 2.1 (Stalling, 2003)

1. We define a primitive root of a prime number $\boldsymbol{q}$, and $\boldsymbol{\alpha}$, a primitive root of the prime number $\boldsymbol{q}$.
2. User **A** selects a random integer $X_A < \boldsymbol{q}$, and computes $Y_A = a^{XA} \bmod \boldsymbol{q}$. And user **B** selects a random integer $X_B < \boldsymbol{q}$ and computes $Y_B = \boldsymbol{\alpha}^{XB} \bmod \boldsymbol{q}$.
3. User A and User B are the same of the compute the session key: $K = (Y_A)^{XB} \bmod \boldsymbol{q} = (Y_B)^{XA}\ mod\ \boldsymbol{q} = \boldsymbol{\alpha}^{XAXB} \bmod \boldsymbol{q}$.

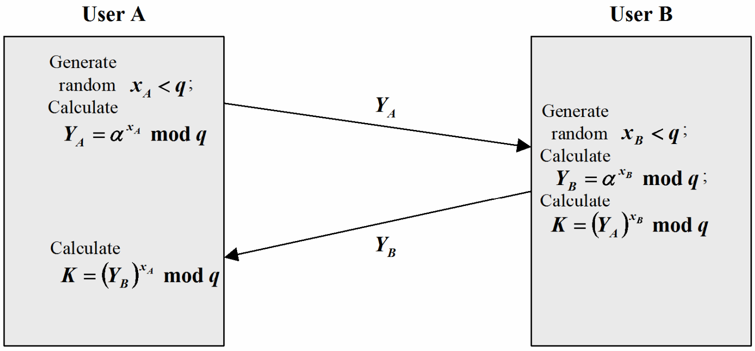

User A
User B
Generate random $x_A < q$;
Calculate $Y_A = \alpha^{x_A}$ mod q
Y_A
Generate random $x_B < q$;
Calculate $Y_B = \alpha^{x_B}$ mod q;
Calculate $K = (Y_A)^{x_B}$ mod q
Y_B
Calculate $K = (Y_B)^{x_A}$ mod q

Figure 2.1: Diffie-Hellman Key Exchange (Stalling, 2003).

2.5 The Authentication Protocol

This section gives a review of authentication protocol of Universal Mobile Telecommunication System (UMTS), as an example of third generation (3G).

2.5.1 The UMTS authentication protocol

The Universal Mobile Telecommunications System (UMTS) is one of the new ‘third generation’ (3G) mobile cellular communication system developed within the framework defined by the International Telecommunication Union (ITU) and known as IMT-2000 (ITU, 1997). The ITU proposed three authentication techniques for International Mobile Telecommunications-2000 (IMT-2000), which is the global standard for third generation wireless communications (ITU, 1997). UMTS builds on the capability of today’s mobile technologies by providing increased capacity, data capability and a greater range of services using a new radio interface standard called UMTS Terrestrial Radio Access (UTRA) (Richardson, 2000).

The worldwide harmonization and globalization process for the 3G radio network and service parameters introduced place in early 1988. So, the standardisations of the UMTS system were defined by the European Telecommunications Standards Institute (ETSI) which involved in the third Generation Partnership Project (3GPP). The 3GPP security group was established in early 1999 to define the UMTS security. The standardisation of UMTS security within 3GPP has now reached a reasonably stable state. UMTS aims to give a broadband, packet-based service for transmitting video, text, digitized voice, and multimedia at data rates of up to 2 megabits per second while remaining cost effective. For simplicity, the following sections give a brief description of the UMTS architecture, authentication mechanisms, the data confidentiality schemes, the integrity protection, the subscriber identity/location, Security Considerations and

Threats for UMTS and the comparisons between GSM and UMTS (3GPP, 2004b) (3GPP, 2004c), (Salkintzis, 2004).

2.5.2 UMTS architecture

In order to demonstrate the security scheme of an UMTS network, the elements of network are introduced as the architecture of UMTS. Figure 2.2 illustrates the UMTS's architecture (3GPP, 1999b). UMTS is divided into three major parts: the User Equipment (UE), the UMTS Terrestrial Radio Access Network (UTRAN), and the UMTS core network. The base stations and the Radio Network Controllers (RNCs) are collectively known as the UTRAN. From the UTRAN to the core network, the RNC will decide to where the traffic will be transmitted. Packet traffic is sent to a new component, the Serving GPRS Support Node (SGSN), and then to the Gateway GPRS Support Node (GGSN). The functions of the GGSN are very similar to the normal IP gateway, which transfer the receiving packets to the appropriate Internet. On the other hand, if there is a voice call from a subscriber, the RNC will transmit the traffic to the Mobile Switching Center (MSC). If the subscriber is authenticated before, the MSC switches the phone call to another MSC (if the called end is another mobile subscriber), or the call will be switched to the Gateway MSC (GMSC) (if the called end is in the public fixed phone network). Therefore, UMTS architecture is arranged into the following components (Boman, et. al, 2002) (Niemi & Nyberg, 2003):

i. User Equipment (UE) : is the user end of the UMTS system. It contains two separate components:

a. Mobile Equipment (ME) is the hardware device itself. The device alone can not use any UMTS services. For example, ME may be a mobile phone, a personal digital assistant (PDA), or notebook.

b. UMTS Service Identity Module (USIM) contains all of the authentication functions and necessary data needed to identify and authenticate the user and getting access to the UMTS network. This card is equivalent to the SIM-card in GSM.

ii. UMTS Terrestrial Radio Access Network (UTRAN) : is a conceptual term identifying that part of a UMTS network which consists of Radio Network subsystems $(RNSs)$ Each RNS contains two components.

a. Node B is equivalent to the Base Transceiver Station (BTS) in GSM. This makes the physical connection to the UE. It also performs some basic Radio Resource Management operation such as checking the power control received from the different terminals. Most of the Node Bs manage three cells. However, a group of Node Bs is connected with the Iub interface to one RNC via the ATM network.

b. The RNC is responsible for one or more Node Bs (BTS in GSM) and controls their radio resources. It is also the service access point for the services the UTRAN provides to the CN. Another important job of the RNC is confidentiality and integrity protection. After the authentication and key agreement procedures have taken place, the subscriber's integrity and confidentiality keys are placed in the RNC . These are then used together with the 'built-in' security functions, f_8 and f_9 (refer to Table 2.1).

iii. Core Network (CN) : the functionality of this component is to provide switching, routing and transit for user traffic. Core network also contains the databases and network management functions. The CN is divided into two parts, the packet switching (PS) and circuit switched (CS) domains. The PS domain offers data services for the user by connections to the Internet and other data networks, and the CS domain offers 'standard' telephone services to other telephone networks. In the CN CS domain are two basic network elements which can be physically combined. These elements serve the Mobile Switching Centre/Visitor Location Register (MSC/VLR) and Gateway Mobile Switching Centre (GMSC). The CN PS domain has two basic network elements.

a. The Serving GPRS Support Node / Visit location register (SGSN/VLR) is the main node of the packet switched domain. It is connected to UTRAN by the Iu PS interface and to the GGSN by the Gn interface. The SGSN is responsible for all packet switched connections for the subscriber. Also it stores in VLR two types of subscribe data such as, International Mobile Subscriber Identity (IMSI), Temporary identities (P-TMSI) which is used in the authentication process. The other type of data is used in the mobility management such as Routing Area of the subscriber (RA), VLR number, and GGSN addresses of every GGSN that are active in the connection.

b. Gateway GPRS Support Node (GGSN) is a SGSN that is interconnected to other data networks. All data communications go through a GGSN between the subscriber and external networks. As with the SGSN, it holds both types of data, subscriber information, such as IMSI, and location information such as the address of the current SGSN the subscriber is connected to.

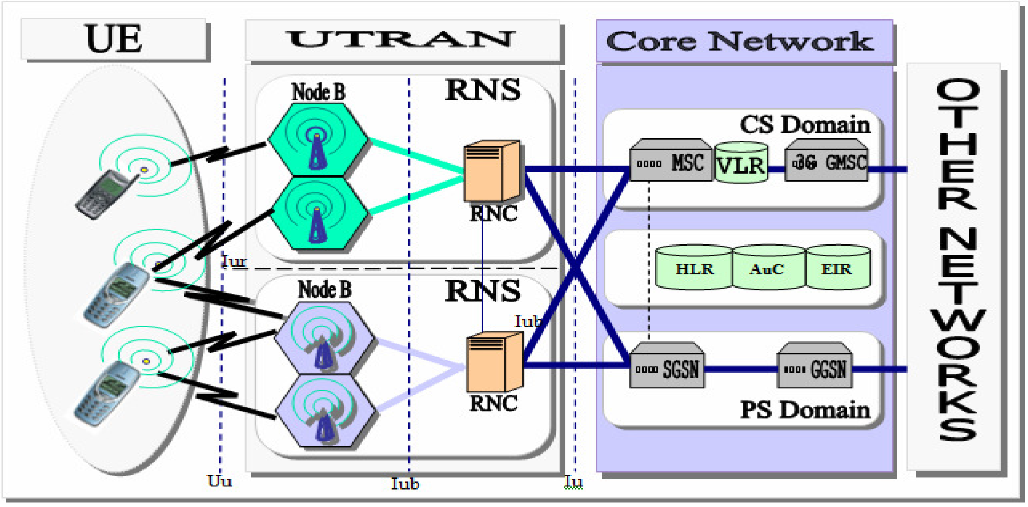


Figure 2.2 UMTS architecture.

2.5.3 The authentication scheme

There are three key principles behind UMTS security.

a. 3G is built on the security of 2G systems. Security elements within GSM that proved to be needed and robust shall be adopted for 3G security.

b. They improve the security of 3G beyond the security of 2G. 3G security will address and correct real and perceived weaknesses in 2G systems.

c. 3G security will offer new security features and will secure new services offered by 3G.

In UMTS the key agreement protocol involved entities such as, a Home Network (HN) with MS related to it, and a Serving Network (SN) which the MS visits. This mechanism uses a shared secret key K_i and a certain cryptography algorithm that are shared between MS and the HLR/AuC in the HN. This is known as authentication and key agreement. The cryptography functions ($f_1 \ldots f_5$) used in the UMTS AKA are executed only within the User Services Identity Module (USIM) and AuC in the HN of that user. Moreover, there are another three operator-specific cryptography functions f_o, f_1^* and f_5^*. The f_o is used to generate a random challenge $RAND$. The $RAND$ then becomes the basic input of functions $f_1 \ldots f_5$. Therefore, the shared cryptography algorithms between MS and its HN include three message authentication code functions f_1, f_1^* and f_2, and four key generation functions f_3, f_4, f_5 and f_5^* (3GPP, 2007a) (3GPP, 2007b) (3GPP, 2007c). The f_1^* is a message authentication function used to provide data origin authentication for the synchronization failure information sent by the USIM to the HLR/AuC. The f_5^* is a key generating function used to compute AK in order to provide user identity confidentiality during re-synchronization. Table 2.1 summarises the functions needed to perform the UMTS AKA protocol.

Function	Description	Output
f_0	**The random challenge generating function**	**RAND**
f_1	**The network authentication function**	**MAC/XMAC**
f_1^*	**The re-synchronization message authentication function**	**MAC-S/XMAC-S**
f_2	**The user authentication function**	**RES/XRES**
f_3	**The cipher key derivation function**	**CK**
f_4	**The integrity key derivation function**	**IK**
f_5	**The anonymity key derivation function**	**AK**
f_5^*	**The anonymity key derivation function for the re-synchronization**	**AK**
f_8	**Ciphering of users and signalling traffic**	**Cipher text**
f_9	**The integrity function**	**Integrity message**

Table 2.1: AKA functions with their outputs.

In addition, UMTS AKA employs a complicated sequence number *(SQN)* technique to accomplish the network authentication. Only when the SQN between AuC and USIM is synchronous, the network authentication is accepted. Therefore, f_1^* and f_5^* are used when the synchronization fails. They are allocated to the AuC and the USIM. UMTS operators can choose these cryptographic functions freely according to the function input/output specification given in (3GPP, 2004b). Since the HLR/AuC control these cryptographic functions totally, any variation of them will not influence the operations of the VLR/SGSN.

In order to execute the synchronization process, the HN must maintain a sequence number (SQN_{HN}) for each individual subscriber, and the MS must maintain a sequence number (SQN_{MS}) for itself. This sequence number can help the MS resist the replay attack, for more details refer to (3GPP, 2007a). On the other hand, the SN and MS's HN transmit data to each other via a secure channel mechanism, which is described in the network domain security of UMTS (3GPP, 1999a).

The UMTS AKA protocol is performed in two procedures, which are depicted in Figure 2.3 (Koien, 2004) (Niemi & Nyberg, 2003). First, are the registration and distribution vectors. The MS registers with its HN and then the HN distributes authentication information to the SN. Second, the authentication and key agreement phase runs between SN and MS. The SN uses the authentication information of the first procedure to carry out the mutual authentication between SN and MS and then an agreed key and a cipher key are provided. The UMTS AKA protocol can be separated into two procedures.

The First Procedure

This procedure is executed when a subscriber roams into a new network or turns on its mobile equipment, the user needs to identify itself using an old $TMSI$ ($TMSI_o$) that was assigned by the previous visited network (VLR_o). Therefore, this procedure is separated into eight steps refer to Figure 2.3.

1. The MS sends a registration request to visitor location register/serving GPRS support node (VLR/SGSN) in the SN. The registration request includes an $MS's$ IMSI.

2. Upon receiving the registration request, the $VLR / SGSN$ in the SN passes the registration authentication request that is generated in step 1 to $MS's$ HLR in the HN.

3. Upon receipt of a request from the $VLR / SGSN$, the HLR / AuC in the HN generates an ordered array of n authentication vectors AV(1..n) whose order is based on SQN_{HN}. It sends an authentication data response as an ordered array of n authentication vectors to the VLR/SGSN in the SN by a secure channel. The AVs is generated by using the secret key K_i which is pre-shared with the subscriber. Each AV consists of RAND, XRES (Expected Response), CK (Cipher Key), IK (Integrity Key) and AUTN (Authentication Token).

4. Upon receipt of AVs, the $VLR / SGSN$ in the SN stores the authentication vectors for performing the subsequent authentication and key agreement procedure. Each AV is used for one authentication and key agreement between the VLR/SGSN in the SN and the MS (Boman, et al., 2002) (Vanneste, et al., 1997) (Xenakis & Merakos, 2004).

The Second Procedure

5. SN generates challenge information for MS and sends it to MS. In the i^{th} performance of the second procedure, $VLR / SGSN$ in the SN initiates an authentication and key agreement to authenticate the MS, the $VLR / SGSN$ selects the next authentication vector from the ordered array. It selects the $i^{th} AV$ to run this procedure on a first-in/first-out ($FIFO$) basis. SN sends the parameters $RAND_i$ and $AUTH_i$ to the mobile station (MS). Therefore, one authentication vector is needed for each authentication instance. This means that the signalling between the VLR and AuC is not needed for every authentication event.

6. MS authenticating network, and generating a response information. Upon receipt of $RAND_i$ and $AUTH_i$ where ($\mathrm{AUTN}=\mathrm{SQN}\oplus\mathrm{AK}\|\mathrm{AMF}\|\mathrm{MAC}$), the MS performs six steps.

- **i.** Computes the anonymity key $AK = f_K^5(RAND)$.
- **ii.** Retrieves the sequence number $\mathrm{SQN} = (SQN \oplus AK) \oplus AK$.
- **iii.** Computes expected message authentication code $XMAC = f_K^1(SQN, RAND, AMF)$ and compares this with MAC which is included in $AUTN$.
- **iv.** If they are different, the MS sends the user authentication rejection back to the VLR/SGSN with an indication of the cause and the user abandons the procedure
- **v.** Otherwise, the MS confirms whether the freshness that the received sequence number SQN is in the correct range (i.e., $SQN_{HN} \rangle SQN_{MS}$) or

not. If the result is negative, the MS sends synchronization failure back to the VLR/SGSN including an appropriate parameter, and abandons the procedure.

vi. Otherwise, the , the MS computes $RES = f_K^2(RAND)$ and sends it back to VLR/SGSN in the SN. Meanwhile MS computes the cipher key $CK = f_K^3(RAND)$ and the integrity key $IK = f_K^4(RAND)$.

7. The VLR/SGSN makes verification by comparing the received RES with $XRES$. If they match, the mutual authentication and key agreement between the MS and the VLR/SGSN is completed successfully. The established keys CK and IK are stored in the VLR/SGSN and will then be transferred to the RNC when needed.

8. If a VLR/SGSN runs out of AVs, it can request another ordered array of AVs from the HLR/AuC.

The VLR/SGSN is assumed to be trusted by the subscriber's HLR / AuC to handle the authentication information securely. It is also assumed that the communication links between the VLR/SGSN and the HLR / AuC are adequately secure. Also, it is assumed that the subscriber trusts its HLR / AuC.

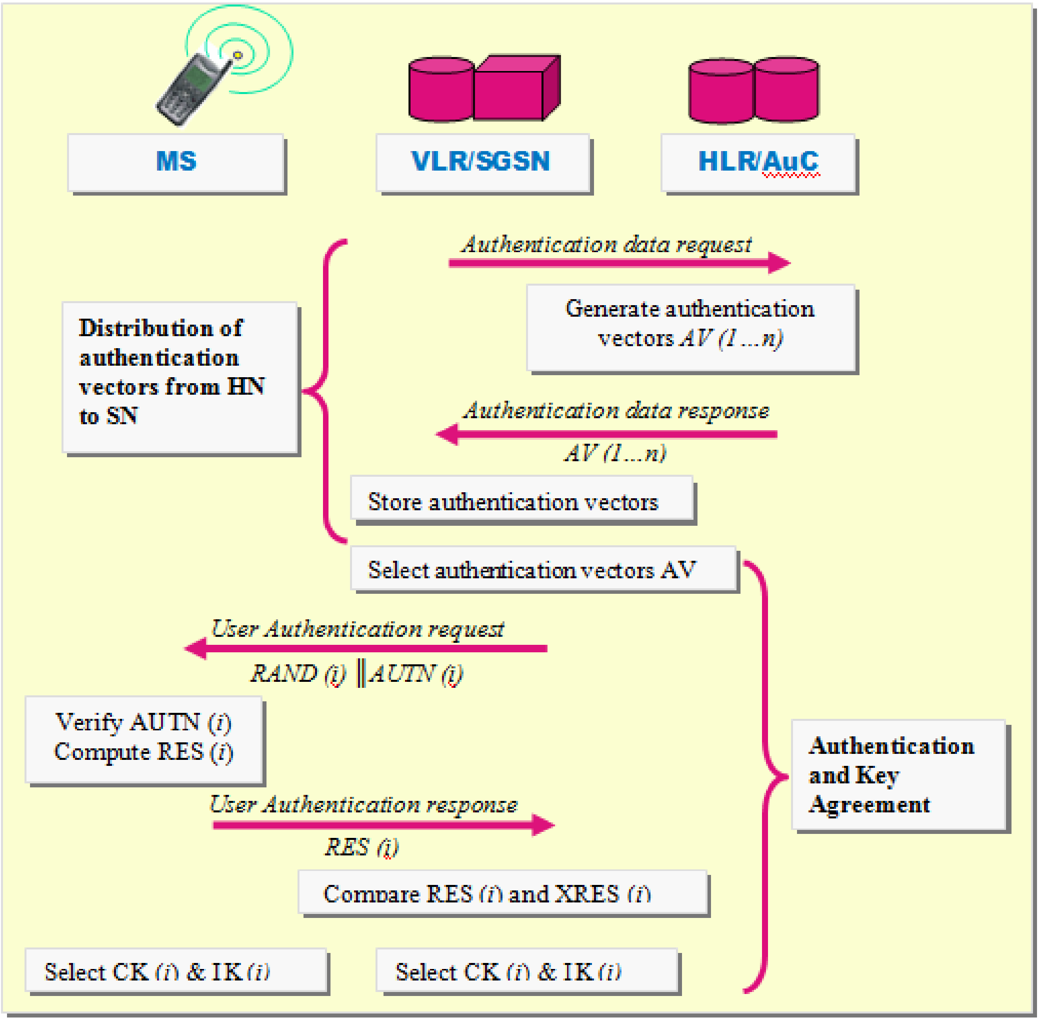

MS
VLR/SGSN
HLR/AuC
Authentication data request
Generate authentication vectors AV (1…n)
Distribution of authentication vectors from HN to SN
Authentication data response
AV (1…n)
Store authentication vectors
Select authentication vectors AV
User Authentication request
RAND (i) ‖ AUTN (i)
Verify AUTN (i)
Compute RES (i)
Authentication and Key Agreement
User Authentication response
RES (i)
Compare RES (i) and XRES (i)
Select CK (i) & IK (i)
Select CK (i) & IK (i)

Figure 2.3: Authentication and Key Agreement (AKA) (Koien, 2004).

When the subscriber roams to a newly visited $VLR_n / SGSN_n$ within one serving network domain, the $VLR_n / SGSN_n$ will send the $TMSI_o$ of the subscriber to its previous visited $VLR_o / SGSN_o$. If the subscriber is found, the $VLR_o / SGSN_o$ will send a response back to the $VLR_n / SGSN_n$. The *user identity request* shall include the *IMSI* of the subscriber, and a number of unused *AVs* ordered on *FIFO* basis. The $VLR_o / SGSN_o$ subsequently deletes the *AVs* which have been sent. If the $VLR_o / SGSN_o$ cannot identify the subscriber, it will indicate that the subscriber cannot be retrieved in the user identity request message. Figure 2.4 and Figure 2.5 describe authentication and key agreement (AKA) in USIM and HE/AuC in an UMTS network (3GPP, 1999a) (3GPP, 1999b) (Kaaranen et al., 2005) (Constantinos et al., 2001).

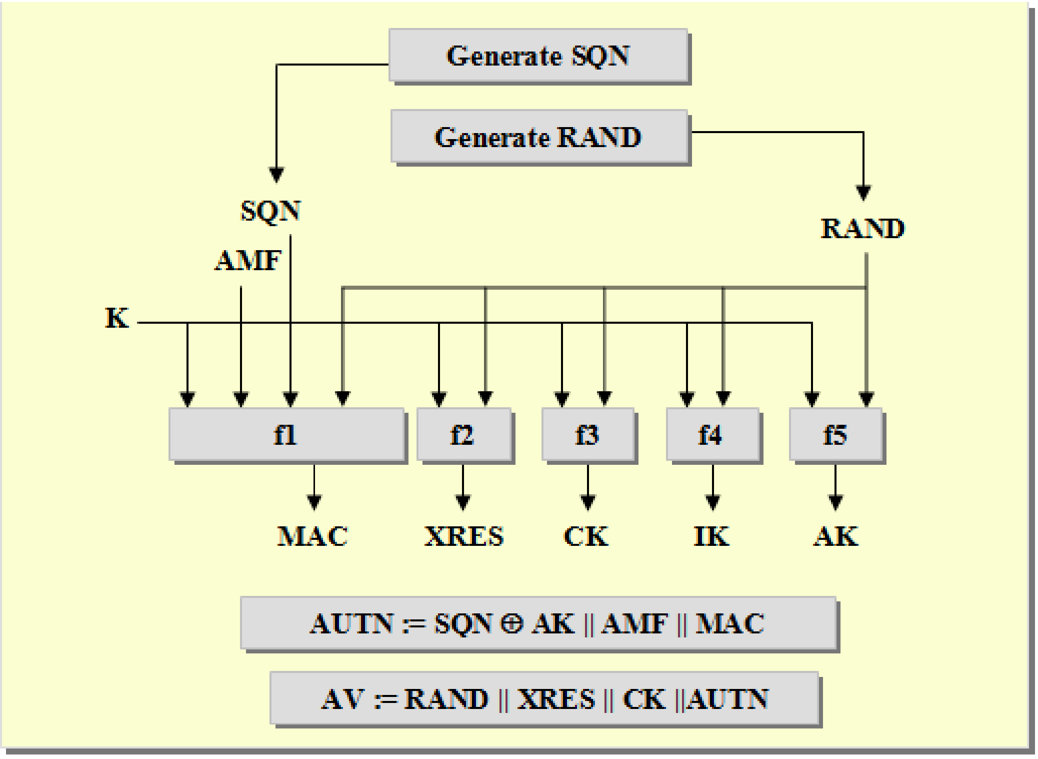

Generate SQN
Generate RAND
SQN
AMF
RAND
K
f1
f2
f3
f4
f5
MAC
XRES
CK
IK
AK
AUTN := SQN ⊕ AK || AMF || MAC
AV := RAND || XRES || CK ||AUTN

Figure 2.4: Generation of authentication vectors (Kaaranen et al., 2005).

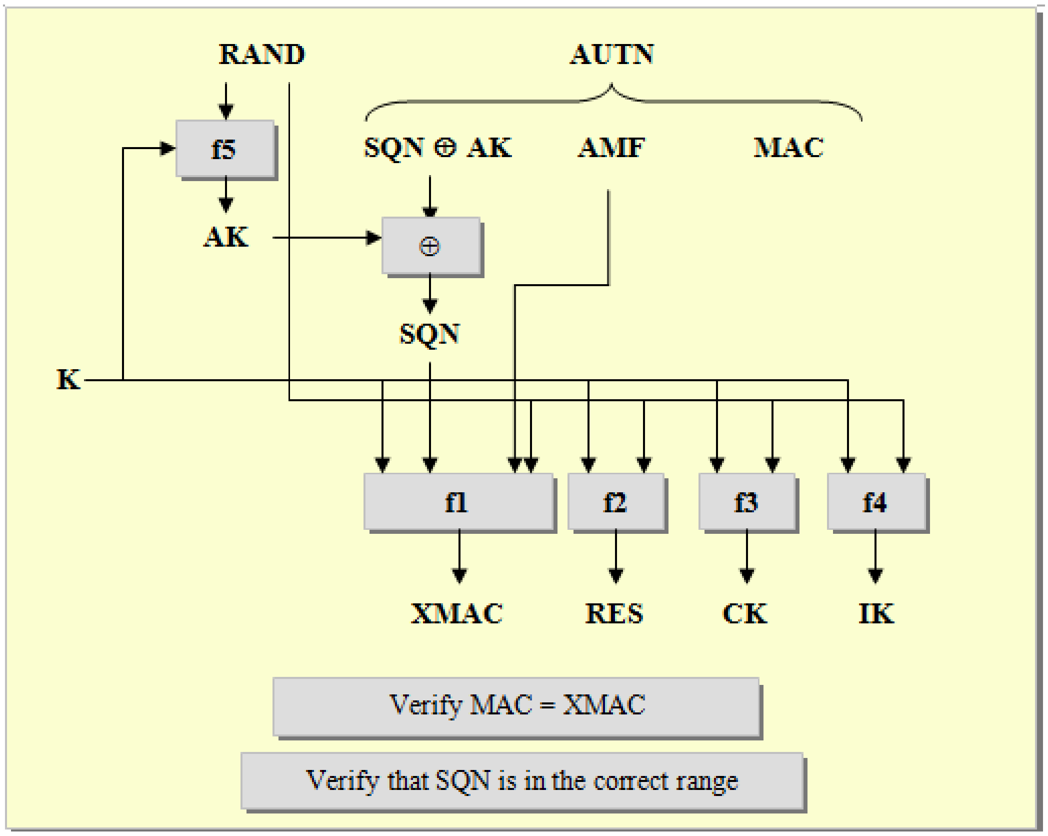


Figure 2.5: **User authentication function in the USIM (Kaaranen et al., 2005)**.

The major differences in authentication and key agreement between UMTS and GSM are: (1) GSM allows only subscriber authentication, while UMTS provides mutual authentication for both subscriber and network, and (2) UMTS creates new integrity keys *(IKs)* in the AKA in order to provide the data integrity protection to the system signalling.

2.5.4 The confidentiality scheme

In GSM, confidentiality is normally terminated in the base station. However, in UMTS the coverage of data confidentiality is extended to the Radio Network Control (*RNC*). The main reason is because there is a substantial amount of connections between base stations and controllers which are based on unsecured radio link hops (Koien, 2004). Therefore, data confidentiality between the base stations and the *RNC* are covered in UMTS (Figure 2.6).

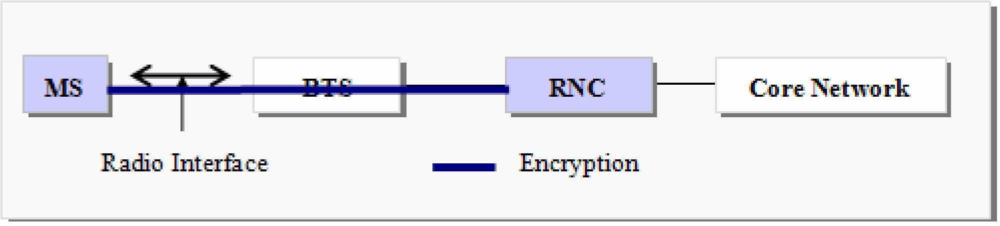


Figure 2.6: Encryption in radio access networks (UMTS)

The authentication vectors (AVs) contain sensitive data, e.g. cryptographic keys and challenge-response authentication data. Therefore, they should be protected against eavesdropping and modification while transferring between the $VLR/SGSN$ and the HLR/AuC. In the age of GSM, there are no cryptographic security mechanisms available for inter-network communication.

The Mobile Application Part (MAP) is an SS7 protocol which provides an application layer for the various nodes in GSM and UMTS mobile core network and GPRS core networks to communicate with each other in order to provide services to mobile phone users. The Mobile Application Part is the application-layer protocol used to access the Home Location Register, Visitor Location Register, Mobile Switching Center, Equipment Identity Register, Authentication Centre, Sort message service center and Serving GPRS Support Node (SGSN).

However, this is no longer the case. The situation now is changing for two main reasons.

1. There is a trend to replace the SS7 (Signalling System number 7) with MAP (Mobile Application Part).
2. The number of different operators and service providers is increasing. There are many ready-made hacking tools in IP network, and they become applicable to the telecommunication networks.

For these reasons, the actual mechanism used for the AVs transmission, MAP protocol, is the essential part to be protected. The original MAP protocol does not contain security functionality, but a security extension to MAP named MAPsec has been developed (3GPP, 2004b) (Liu, 2004).

The MAP protocol can be run on top of IP, that is, MAP over IP. In this case, there are two alternative methods to protect MAP: either to use MAPsec (3GPP, 2004b) or IPsec (3GPP, 2004c). The latter method has the advantage that the protection also covers lower layer headers as it is clone in the IP layer.

It is imperative that the system operator provides full confidentiality and integrity protection to this communication link (Iu interface). If an operator chooses to use IP as the preferred transport protocol over the Iu interface, Network Domain Security for IP (NDS/IP) (3GPP, 2004c) can be used to solve this problem.

The encryption mechanism in UMTS is based on a stream cipher concept as described in Figure 2.7 (Kaaranen, et al, 2005). This means the plaintext data is added bit by bit to random looking mask data, which are generated based on cipher key CK and a few other parameters. The advantage of this encryption is that the mask data can even be generated before the actual plaintext is known (Sutton, 2002). Therefore, the final encryption is executed rapidly. The decryption is performed in a similar way as encryption. Since the product of the data mask is not dependent on the text, another input parameter is required so that the mask is different for different key streams.

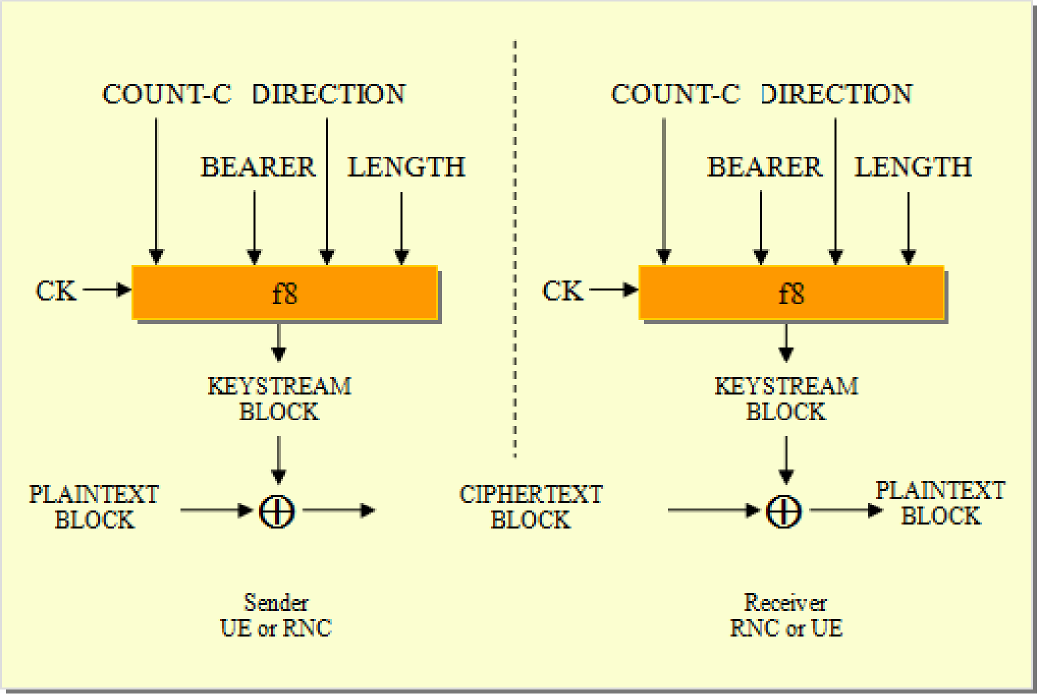


Figure 2.7: Ciphering of user and signaling traffic in UMTS (Kaaranen et al., 2005).

2.5.5 The integrity protection

Sometimes a message's origin or contents have to be verified. Even though it might come from a previously authenticated party, the message may have been altered. To avoid this, integrity protection is necessary.

The integrity security service is a new feature in UMTS systems. But in the GSM system the integrity service is not provided. The integrity protection range in UMTS is the same as the physical range in the confidentiality protection. However, the integrity service in UMTS is only limited to signalling information between *MS* and the *RNC* , but not the user traffic. The reason for providing the signalling information with a dedicated integrity service in UMTS is most of this information is considered sensitive and must be integrity protected (Liu, 2004). For example, UMTS networks must be able to order the MS to use the connection without ciphering for many reasons. The use of ciphering cannot be made compulsory. In this case, the message's origin or the user communication contents will be compromised. This will result when an attacker impersonates a network to establish a connection with a user without ciphering. To avoid this, integrity protection is necessary.

Since the user traffic is not protected by this integrity security service, it represents a problem under certain circumstances. For example, there will be the situation where the data encryption is unavailable. Thus the integrity protection will become the only defence against the insertion, modification, deletion and replay attacks. Furthermore, the message itself might not even have to be confidential; the important thing is that it is genuine.

The method for integrity protection in UMTS is to generate a message authentication code (MAC) to be added to message. The MAC can only be generated at the nodes that know the keys derivate of the pre-shared secret key K_i. They are stored in the USIM and the AuC. It is very important to offer integrity protection, especially since the serving network is often operated by an operator other than the subscriber's own operator. Figure 2.8 describes the integrity function (Kaaranen, et al., 2005).

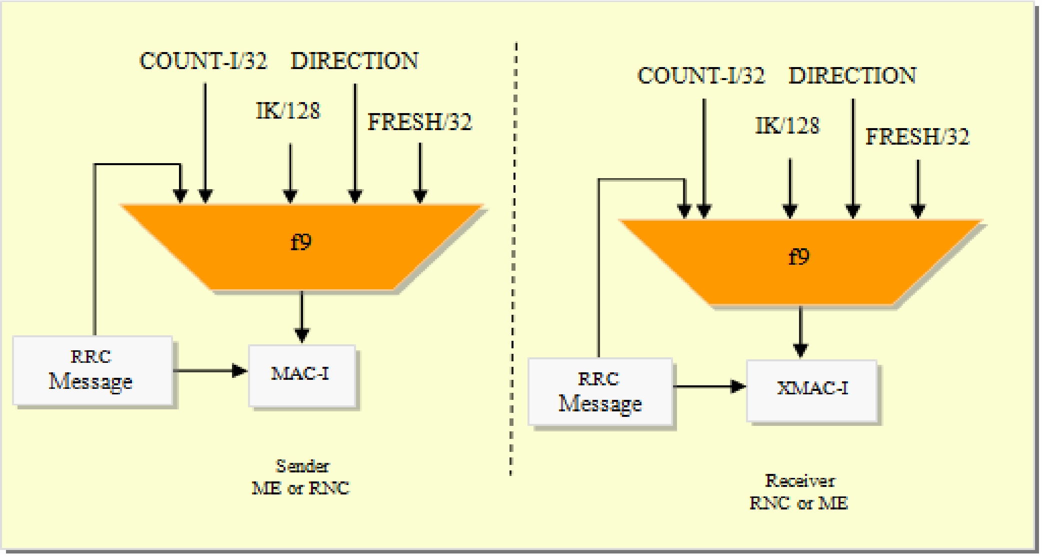


Figure 2.8: Message authentication code (Kaaranen et al., 2005).

2.5.6 The subscriber identity/location confidentiality

The UMTS provides the security features related to the Subscriber Identity/location Confidential can be summarised as follows:

i. User identity confidentiality: the permanent user identity (IMSI) of a subscriber to whom services are delivered cannot be eavesdropped on the radio access link.

ii. User location confidentiality: the location of a subscriber cannot be determined by eavesdropping on the radio access link. Thus, the presence or the arrival of a user in a certain area is secure.

iii. User untraceablity: an attacker is unable to deduce whether different services are delivered to the same user by eavesdropping on the radio access link.

The identification process in UMTS uses the same permanent identity (IMSI) as in GSM. To achieve these objectives, the UMTS has Temporary Mobile Subscriber Identity (TMSI) as in GSM to make the user identifiable on the radio access link. However, UTRAN identification is performed using the temporary (TMSI) in CS traffic and P-TMSI in PS traffic.

Once the user has been identified by the SN, a TMSI and P-TMSI is provided by the SN ensuring that each user has a different identity and maintains the relationship between the IMSI and TMSI (Sutton, 2002). The provided TMSI is transferred to the user after the encryption has been activated. This TMSI is used until a new TMSI is allocated by the network (Sutton, 2002). When a new TMSI is allocated and acknowledged by the network, the old temporary identity is removed from the VLR/SGSN . If an acknowledgment has not been received, the VLR/SGSN keeps both TMSI and can accept either of them (Liu, 2004).

To make the user untraceable, the TMSI should not be identified for a long period by means of the same temporary identity (TMSI). In addition it is required that any signalling or user data that might reveal the user's identity is ciphered on the radio access link. These features are identical to those provided in GSM. They protect against passive, but not against active attacks.

2.6 Security considerations and threats for UMTS

According to above analysis, UMTS Release'99 security is quite similar to GSM security that shown as follows (3GPP, 1999b).

1) The home AuC generates challenge-response vectors.

2) The challenge-response vectors are transmitted unprotected via the signalling network to a visited network that needs to check the authenticity of a mobile station.

3) Unlike in GSM, the network also authenticates itself to the mobile.

4) The security model still assumes trust between all network operators.

5) The IMSI which uniquely identifies a user:

   - Is still revealed to the visited network.
   - Can still be demanded by an attacker that impersonates a base station, as there is no network authentication in this case.

6) Confidentially is only provided on the radio link.

Security poses a much wider set of issues, which need to be considered by all the players in the mobile communication systems. The UMTS system requires mutual authentication, data integrity, data confidentiality, user anonymity and non-repudiation, et cetera. Therefore, ETSI draws up these specifications "Security Threats and Requirements” (3GPP, 1999a) for the authentication protocol designer. In this thesis, the specifications and the general objectives for UMTS security features (3GPP, 1999c) (Liu, 2004) are used to design new authentication protocols.

2.7 Related Work

Several authentication schemes have been proposed to enhance the security of mobile communication systems, but, none of them can fulfil the security requirements of third generation mobile systems (Cheng, et al. 2005). Those designed by Brutch and Brutch (1998), Horn, et al. (2002), Harn and Lin (1995), Lee, et al. (1999), Lin and Shieh (2000), Looi (2001), and Molva, et al. (1994) were not based on third generation mobile systems and may incur much computational overhead. Some authentication schemes for third generation mobile systems were proposed by Dell’Uommo and Scarrone (2001), Putz, et al. (1998), Putz and Schmitz (2000), to solve some security issues but these schemes did not address other security issues, such as end-to-end security, anonymity and confidentiality.

The International Telecommunication Union (ITU) proposed three authentication techniques for International Mobile Telecommunications-2000 (IMT-2000), which is the global standard for third generation wireless communications (ITU, 1997). These techniques only provide some security features and have some weaknesses. The first technique is based on the use of symmetric key cryptosystems and a challenge-response exchange. This technique requires too many authentication messages and does not ensure end to end security. Also, assuming the connection between the network operator and service provider is secure; the messages that are transmitted through the connection are vulnerable.

The second authentication technique is based on the unilateral use of a digital signature scheme and a challenge-response exchange. This technique did not provide mutual authentication or end to end security and created high computation costs.

The third technique is also based on the use of a digital signature scheme. The public key certificates and timestamps are combined to provide user identity confidentiality and unilateral entity authentication in a single mechanism. The third technique is the same as the second technique; but did not provide mutual authentication, end to end security and created high computation costs.

In all of these techniques the authentication technique is fixed, such as that the network operators of visited domains must be involved in the authentication procedure between roaming users and home service providers. This represented a common weakness in the three techniques.

Harn and Hsin (2003) proposed an enhanced registration and AKA scheme for UMTS. By introducing a combination of hash-chaining and keyed HMAC techniques, in their proposed protocol they claim it can provide strong periodically mutual authentication, strong key agreement, and a non-repudiation service in a simple and elegant way. Compared with the standardized AKA procedure a lot of data is being sent to the HLR. This could have a negative impact on the performance of this protocol.

Huang and Li (2005) feel that 3G-AKA has three weaknesses: bandwidth consumption between a VLR in the serving network and HLR in the home network; the space overhead of the VLR in the serving network, which is required to store AVs, and the need for SQN synchronization. In 2005 they proposed an extension of the UMTS-AKA protocol, called UMTS X-AKA protocol, as a solution to overcome these weaknesses. Zhang and Fang (2005) proposed a new authentication and key agreement protocol, which overcomes redirection attack and drastically lowers the impact of network corruption. The protocol is called adaptive protocol AKA (AP-AKA). They solve these problems by eliminating the need for synchronization between a mobile station and its home network and providing a way for the MS to verify if AV's are indeed coming from SN and have not been used before (i.e., The MS can verify for itself if an AV used by the VLR is a fresh one).

Jun and Chen (2005) presented a novel mutual authentication and key agreement protocol based on the Number Theory Research Unit (NTRU) public key cryptography. The symmetric encryption, hash and “challenge-response” techniques were adopted to build their protocol.

Gódor and Imre (2006) suggested a novel authentication algorithm (GSZV) based on public key infrastructure by using digital signatures, certifications, and two different sequence numbers. The main goal of that algorithm is to guarantee a secure and confidential communication between the users and the network. In that algorithm all information, including the IMSI of MS is encrypted on the air interface which is needed since if the IMSI gets known an attacker can misuse it.

Yeh and Lee (2007) suggested a dual-purpose signature for authentication on UMTS which provides valuable improvements to UMTS by using the digital signature technique to reduce the storage needed at the HLR and guarantees the access rights of the mobile station (MS). The Dual-Purpose signature concept provides an alternative application for signature technique in an efficient way. With the suggested method, the

UMTS will benefit from the elimination of bulky storage and face fewer security threats.

Yan et al. (2009) proposed a n improved authentication protocol wit less Delay for UMTS Mobile Networks which require a little modification based on UMTS protocol. Their protocol reduced the number of messages and requires less traffic, bandwidth between MS and CN. In addition, their protocol out-performs the UMTS authentication process by reducing the signalling overhead and authentication delay time

Naveed, et al. (2010) proposed a new security algorithm as Airam, for authentication on UMTS which resolves the problem of sending IMSI as a plaintext message over the air interface by using a hybrid approach between Symmetric and Asymmetric cryptography.

Chouldury, et al (2010) presented End-to-End User Identity Confidentiality (E2EUIC), an extension of 3GPP-AKA. The main goal of that solution is to guarantee a secure and confidential communication between the users and the network. In that solution all information, including the IMSI of MS is not only encrypted on the air interface, but goes one step ahead to ensure the same over the wired network as well.

Hassan, et al (2010) presented an analysis and evaluation of the security of UMTS which points out a major weakness in UMTS is IMSI in plaintext. Moreover, it provides simulation scenarios for an attack on IMSI of MS when it sends a registration request to the serving network.

2.8 Summary

In this chapter, the security of mobile networks and relevant terminologies are explored. The security attacks, the services which are defined requirements for the safety of mobile communication and the approaches to satisfying those requirements were presented.

Also in this chapter the research related to UMTS security is reviewed. There are some protocols that are relied on symmetric key techniques, asymmetric key techniques or hybrids. Many of them paid attention to how to provide a secure, robust, efficient security scheme to existing UMTS.

Chapter 3 An Enhancement of Authentication Protocol for 3G Mobile Networks

This chapter proposes a secure and flexible authentication mechanism which can satisfy the security requirements of the third generation mobile communication systems. By integrating the public key with the hash-chaining technique, the security of the 3G protocols in network access is improved to provide the protection of the subscriber's true identity (IMSI) to ensure subscriber un-traceability, key refreshment periodically and dynamically, strong key management and a new non-repudiation service in a simple and elegant way. In addition, the bi-unilateral and mutual authentication among MS, VLR/SGSN and HLR/AuC resulted in a more secure protocol than the other available authentication protocols. The proposed protocol improves performance by reducing communication times, and by creating fewer authentication messages and data sizes during the process of authentication. To avoid the complicated synchronization as in UMTS the proposed protocol does not use *SEQ*, the management of a hash chain is simple and elegant compared to that of *SEQ*. However, the proposed protocol demonstrate better performance in terms of latency and storage space, compared to the 3G network approach of home network transporting authentication vector to visited network. This proposed protocol is secure against network attacks, such as replay attacks, guessing attacks and other attacks.

3.1 Introduction

In Chapter 2 research related to UMTS security is reviewed. Much of it is concerned to add a secure, robust, efficient security scheme to existing UMTS. Most of current authentication schemes for mobile systems have some weaknesses. Leakage of UE identities in which the IMSI is revealed to the visited network and can be demanded by an attacker that impersonates a base station, as there is no network authentication in this process. The design of the current 3GPP authentication and key agreement protocol (AKA) has some shortcomings in the performance and complexity of operations.

Firstly, the bandwidth consumption between visitor serving network and home network is inadequate when *MS* requests to authenticate itself for the *VLR/SGSN* in serving network (SN) and no authentication vectors are available the *VLR*/*SGSN* must turn back to *HLR* in the Home Network (*HN*) to make a registration request to generate another array of *n* authentication vectors. Moreover, when the subscriber roams to a newly visited *VLR/SGSN* within a different serving network domain the authentication vectors in the old *VLR/SGSN* are deleted, which is called an unused authentication vectors problem. Subsequently, as a lot of data being sent between *VLR/SN* and *HLR*/HN, this has impact on the performance of AKA protocol.

Secondly, as the generation of authentication vectors *(AV)* is expensive and which require the generation of five records in each *AV* will increase the delay time in home network.

Thirdly, the storage space overhead occurs if there are *m* subscribers, and an array of *n* authentication vectors for each subscriber in SN, then the SN must wastefully store $n \times m$ authentication vectors.

Fourthly, the management of sequence number (*SEQ*) which is needed for synchronization between mobile station and its home network, and the periodical authentication is achieved by comparing a *SEQ* countervalue between an MS and a

VLR/SGSN periodically. The SEQ is susceptible to synchronization failure. Therefore, both MS and HN are needed to maintain SEQ to accomplish that process between them.

The proposed scheme integrates the public key with the hash-chaining technique. The security of the 3G protocols is improved to protect the subscriber's true identity (IMSI) and to provide, key refreshment periodically and dynamically, strong key management and a new non-repudiation service in a simple and elegant way. In addition, the mutual authentication among MS, VLR/SGSN and HLR/AuC has resulted in a more secure protocol than other available authentication protocols. The proposed protocol fulfils the security requirements of the third generation mobile systems and improves performance by reducing the communication times, and by creating fewer authentication messages and data sizes during the process of authentication. With the design goal of achieving security by lower cost, the proposed protocol significantly reduce the communication overhead between home network and visited network especially for roaming authentication. To avoid complicated synchronization as in UMTS, the proposed protocol does not use *SEQ*, the management of a hash chain is simple and elegant compared to that of *SEQ*.

3.2 Authentication framework

In third generation mobile systems, many emerging services, such as the World Wide Web, stock quotes, e-mail account and multimedia, can be accessed through a wireless link. When a mobile user roams far from his home domain and wants to access these services, the user may intend to use the servers in a visited domain instead of the ones in his home domain. To meet today's needs for wireless communication the protocols need to be highly secure, and require low computational overhead and thus low power.

To acquire mobile services in visited domains, mobile users must be authenticated. Normally, the VLR/SGSN in a visited domain is unable to solely perform the authentication procedure without any prior knowledge of the roaming user; hence, the visited domain requires the participation of the home domain to authenticate the user.

To enhance the 3G AKA protocol, the proposed authentication protocol has adopted two major techniques: public key, and hash chaining. Public key cryptography has not previously been used in mobile communication environments due to performance constraints. It was not consider suitable for second generation systems because of the resulting length of messages and the necessary computational loads. New protocols for authentication between user and network have been developed to overcome these problems.

One-way function is a variation of the message authentication code as with the message authentication code, a hash function accepts a variable size message M as input and produces a fixed size output, referred to as a hash code $H(M)$. The hash code is a function of all the bits of the message and provides an error detection capability. When it changes any bits in the message result in a change to the hash code, a hash function H has some properties, (Burnett & Pause, 2002), (Simmons, 1992) (Stallings, 2003).

1. H can work with any size of data.
2. H can output a fixed-length message.
3. It is easy to compute the hash value $H(x)$ for a given message x.
4. It is computationally infeasible to retrieve the x from a given value $H(x)$. This property is called one-way property.

5. Given x and $x^{'} = H(x)$, it is computationally infeasible to find a value $x^{'}$ where $x^{'} \neq x$, such that $H(x^{'}) = H(x)$. This property is called weak collision resistance.
6. It is computationally infeasible to find two values $x^{'}$ and x, such that $H(x^{'})$ = $H(x)$. This property is called strong collision resistance.

The proposed protocol uses a one-time password/hash-chaining technique which was proposed by Lamport (1981). It used a hash function with a property to construct a sequence of hashing value. They designed it in a remotely accessed computer system. One of the aims of the one-way hash function is to prevent eavesdroppers discovering the password and to reduce the computing time, which this technique has used in many applications (Al-Fayoumi, et al., 2007) (Anderson, et al., 1996), (Gennaro & Rohatgi, 1997), (Harn & Lin., 2001), .

In this method, let the user (claimant) and the server (verifier) deal with the secret (M) as a seed of hash value and $f(M)$ be a one-way function, when a user (i.e., the one wishes to be authenticated) wants to register or log in the system, then the user should construct $f^{n}(M) = f(f(\ldots(f(M)\ldots)))$, where n represents the maximum number of services that the user can request after the registration phase (i.e., the composition of *nfs)*, and sends $f^{n}(M)$ to the server (i.e., the one decides whether the user is who it is). Then the server uses it to compute a sequence of passwords $f^{n-1}(M), f^{n-2}(M), \ldots, f(f(f(M))), f(f(M)), f(M)$ and the server stores those. The user holds $f^{n}(M), \ldots, f(f(f(M))), f(f(M))$.

After the registration is completed, each hash chain can be used by the claimant to prove itself to the server N times. In the j^{th} session, the user provides $f^{n-1}(M)$ to ask for a connection to prove itself. The server can verify the correctness of $f^{n-1}(M)$by means of the one way function by computing $f\left(f^{n-1}(M)\right)$ and the server needs to store $f^{n-1}(M)$ as the last value of user to authenticate the next visit. So, the user reveals $f^{n-1}(M)$, $f^{n-2}(M)$, $\cdot\cdot\cdot$, $f(M)$, and $M = f^{0}(M)$ in sequence to prove itself n times. In this way $f^{n-j}(M)$ can be used as a proof of the j^{th}connection. In the proposed protocol the dynamic concept achieved through management the dynamic keys between the MS ⇔ HLR/AuC and MS ⇔ SGSN/VLR. In other words the dynamic mechanism works at two levels by the keys refreshment are used whether at MS/HLR and MS/VLR for each Initial and subsequent authentication session respectively. Moreover, this property provides service providers with the ability to develop proprietary authentication mechanisms and adjust the keys in run time. Where in the dynamic key agreement the HLR/AuC is determine number of subsequent authentication procedure that will be executed for each time the initial authentication procedure starts which will discusses in the following section. When MS makes a service contract with his/her home network HLR generates the public and private keys and subscribes public key to MS and keeps it in its database and save KHU, IMSI and CertM in the SIM/USIM of MS. In the initial authentication procedure the MS encrypt an Authentication Request Message between MS to HLR (***AUTHM***$_{\mathrm{H}}$) by HLR's public key that has been saved in the SIM. HLR decrypt it by its private key, and then refresh the new public key and send back to MS within Authentication Data Response Message between HLR to MS (***RAUTHM***$_{\mathrm{H}}$) to use it in the second time. Consequently, Dynamic key management is achieved at the level of MS and HLR/AuC.

Meanwhile, the MS generates a session key K_{VM} = f (K_{VM}, IK ⊕ CK) as a shared key between MS and SGSN/VLR, where IK and CK are a nonce numbers and f^{n}(M) where

f^n(M) is a one-way hash chaining function and n represents the maximum number of services that the MS can request after initial authentication, then send it within ***AUTHM***$_H$ to SGSN/VLR through HLR. The VLR save K_{VM}, CK and f n(M) under the ID of that user and sends Authentication Data Response Message between VLR to MS (***RAUTHM***$_V$) encrypted by the session key K_{VM} of its response message by using Advance Encryption Standard AES algorithm.

In the subsequent authentication procedure the MS generates a new session key K_{VM}' $= f(K_{VM}, IK \oplus CK)$, where IK is a new generated nonce and K_{VM} is the shared key. CK is in the messages sent by the MS to the VLR in the initial authentication procedure. Meanwhile, the MS produces f^{n-i}(M), where (i) is the number of services that have been requested, and M is the secret key generated in the initial authentication. VLR generates a new session key K_{VM}' using the same function used by the MS and then encrypts ***RSAUTH***$_{UV}$ = (IK+1 $\oplus$ TMSI) with K_{VM}', and then sent is sent back to the MS. After that upon receipt of the response message, the MS decrypts the ***RSAUTH***$_{UV}$ using K_{VM}'.

The subsequent authentication procedure only contains two message exchanges. The nonce number IK transmitted between the MS and VLR is used to refresh the session key. In this way, the encryption key used for every session is different. Except for the first session key, key generation is performed by both the MS and VLR. The first session key is generated by the MS and sent to the VLR. After that, the VLR can generate the following session key by itself. By using K_{VM} and IK $\oplus$ CK as two inputs, the MS and VLR can generate the same new session key if the inputs are identical. Consequently, Dynamic key management is achieved at the level of MS and SGSN/VLR.

Therefore, using the refreshment keys concept in both MS, SGSN/VLR and HLR/AuC according to the n and t value which has determined by MS and HLR/AuC respectively, the subscribed service period (t) is used to determine whether the service request is out-of-date or not, and (*n*) is used to determine the number of times and the ith session to perform the subsequent authentication procedure dynamically, without transfer any clear parameters. In the i-th session, the user provides f^{n-i} (M) to ask for a connection to proof of the i-th connection. Consequently, Dynamic key management is achieved at the level of MS and SGSN/VLR.

3.3 The Operation modes

In this proposed protocol, VLR/SGSN in SN maintains the profiles and privileges of the registered MS. Thus, only MS's home network (HLR) can initially authenticate MS. Another entity, VLR/SGSN in SN, is responsible for forwarding MS's authentication request to HLR in HN. In the proposed protocol, after initial authentication has been performed, VLR/SGSN in SN is then capable of authenticating MS when it is required. The proposed authentication protocol contains two operation modes for initial and subsequent authentication.

1. Registration and distribution of authentication information (Initial Authentication): This procedure is used when a mobile user (MS) leaves his home domain and roams to a visited domain. The user may request services from the network operator of the visited domain. In this case, the initial authentication shown in Figure 3.1 is performed between the three parties. First, the request message is generated by MS and sent to the authentication VLR/SN in the visited domain. Since VLR/SN is unable to authenticate MS by itself, it forwards it to HLR in *MS*’s home domain. The verification procedure is performed by HLR. A response message is generated corresponding to the authentication result as authentication vector (AV). The VLR/SN forwards the response message to MS and decides whether or not to provide the service to MS according to the authentication result. Here, VLR/SN caches some authentication information, which can be used in subsequent authentication. The response message lets MS know whether the authentication was successful or not. After the initial authentication, both VLR/SN and MS obtain the authentication result from HLR/HN and share some secret information without intervention of HLR/HN.

2. Authentication and Key Agreement (Subsequent Authentication): After initial authentication, VLR/SGSN has the ability to authenticate MS in subsequent communication. If MS remains in the same visited domain and requests services, then the user should ask for subsequent authentication. MS similarly generates an authentication request message, which should contain the information shared between MS and VLR/SN; VLR/SN then uses this information to authenticate MS. As mentioned above, VLR/SN has cached information needed to authenticate MS. After authenticating MS, VLR/SSN sends a response message containing the authentication result to MS. The MS receives the response message and learns whether the authentication was successful or not.

3.4 Description of the proposed protocol

In this section, the proposed protocol shows how the proposed framework can be applied to improve the performance of authentication in call setup services. The proposed protocol satisfies the security requirements of third generation mobile systems and has the advantages of a dynamic framework. The proposed schemes involve the use of a public key of HLR and VLR, which is used for a legitimate MS to encrypt an authentication key that generated by MS himself/herself and passes it to VLR.

Moreover, we simply employ a challenge response to resist the replay attacks. The proposed authentication protocol is divided into two procedures; the first one is called the initial authentication procedure, which flows from MS⇔VLR⇔HLR. The second one is limited between MS⇔VLR and is called the subsequent authentication procedure.

3.4.1 Initial authentication procedure

To mitigate the computation burden of mobile equipment, the encryption is clone in MS side since the public key operation takes O (K2) complexity but private key operation takes O (K3) complexity with the typical modular exponentiation algorithms used to implement the RSA algorithm (ITU_T, 1993), where K is the number of bits in the modulus. Table 3.1 gives the software speeds of RSA (Lacy, et al., 1993). RSA goes much faster if we choose a value of e carefully. Therefore, we suggest the exponent value e should be smaller. In order to make use of public key cryptography on the low-computation mobile equipment, the related research can be found in (Belier, Chang & Yacobi, 1993).

RSA Speeds for Different Modulus Lengths with an 8-bits Public Key(on a SPARC II)			
	512 bits	768 bits	1,024 bits
Encrypt	0.03 *sec*	0.05 sec	0.08 *sec*
Decrypt	0.16 *sec*	0.48 sec	0.93 *sec*
Sign	0.16 *sec*	0.52 sec	0.97 *sec*
Verify	0.02 *sec*	0.07 sec	0.08 *sec*

Table 3.1: Software speeds of RSA (Lacy, et al., 1993).

In 2003, RSA Laboratories recommends the minimum key length for general data is 1024 bits without any specifying lifetime (kaliski, 2003). NIST recently recommends 1024 bits for RSA, which is taking into account the lifetime of the data (NIST, 2003). For security concerns and the execution speeds of the public key encryption, we suggest the value of public key length is optimally 1024 bits.

Before we describe the common registration phase of the proposed mechanism, we assume the following operations are performed when MS makes a service contract with his/her home network HLR:

- HLR generates the Public and Private Keys.
- HLR subscribes (Public Keys) to MS.
- HLR produces a certificate CertM to public keys and keeps it in its database. HLR writes K_{HU}, IMSI and $Cert_{M}$ in SIM/USIM of MS.

At first, we consider the scheme that consists of four messages exchange among MS, VLR and HLR. The message flows are indicated in Figure 3.1. The notations in Figure 3.1 are defined as follows:

NOMENCLATURE

IMSI	*International* Mobile Subscriber Identity
TMSI	Temporary Mobile Subscriber Identity generated by HLR/AuC
K_{HU}, K_{HP}	Public/private key pair of HLR
K_{VU}, K_{VP}	Public/private key pair of SGSN/VLR
K_{HU}'	The new HLR's public key
K_{VM}	Session key shared by the MS and SGSN/VLR
K_{VM}'	f(IK, CK): session key between the MS and VLR, the function f may be a simple function. e.g. the XOR of IK and CK. $h\ (K_{VM}, IK \oplus CK)$
IK, CK	Nonce numbers
T	Subscribed Service Period
$f(\ )$	A one way Hash function;
$AUTH_{MH}$	Authentication Request Message between MS to HLR $ID_{M}, ID_{H}, E_{KHU}(\ IMSI \| IK \| CK \| K_{VM} \| f^{n}(M) \| n \| ID_{V})$
$AUTH_{VH}$	Authentication Request Message between VLR to HLR $ID_{V}, E_{KVP}(ID_{V}, R_{V})$
$RAUTH_{MH}$	Authentication Data Response Message between HLR to MS $E_{KHP}(ID_{H} \| IK+1 \oplus TMSI \| T \| K_{HU}' \| K_{VU})$
$RAUTH_{MV}$	Authentication Data Response Message between VLR to MS $E_{KVP}(E_{KVM}(IK+1), TMSI\ ))$
$RAUTH_{HV}$	Authentication Data Response Message between HLR to VLR $E_{KVU}(R_{V} \| IK \| CK \| K_{VM} \| f^{n}(M) \| n \| T \| TMSI)$
ID_{M}	The identity of the MS
ID_{V}	Identity of VLR
ID_{H}	Identity of HLR
$\oplus$	Logical computation, bit-wise exclusive or operation

This section describes how that can be applied to enhance the authentication procedure. It is known that the authentication process is achieved by all the authentication entities of 3G mobile network. The proposed authentication protocol is divided into two procedures; the first one is named Initial authentication procedure, which flows from MS⇔VLR⇔HLR. The second one is limited between MS⇔VLR is a subsequent authentication procedure.

In the proposed authentication protocol, we assume that MS⇔HLR/AuC and SGSN/VLR⇔HLR/AuC have the public/private key pair and use Public-Key cryptosystems. In addition, there is public key infrastructure so that public keys can be correctly and efficiently distributed. This enables all entities of network (3G) to mutually authenticate each other easily. MS can obtain the public key of VLR to be sent by HLR. At first, MS sends secret message to challenge VLR and HLR, and VLR also sends secret message to challenge HLR. However these secret messages are encrypted with its public and private key respectively. After that VLR and HLR send a message to response MS that decrypt by its private key. The HLR also decrypts the secret messages to response VLR based on VLR's public key. If the processes are finished, they can achieve mutual authentication between all participants, and refresh HLR's public key.

At first, we consider the scheme that consists of four messages exchange among MS, SGSN/VLR and HLR/AuC. The messages flowing procedure in the initial authentication procedure are indicated in Figure 3.1.

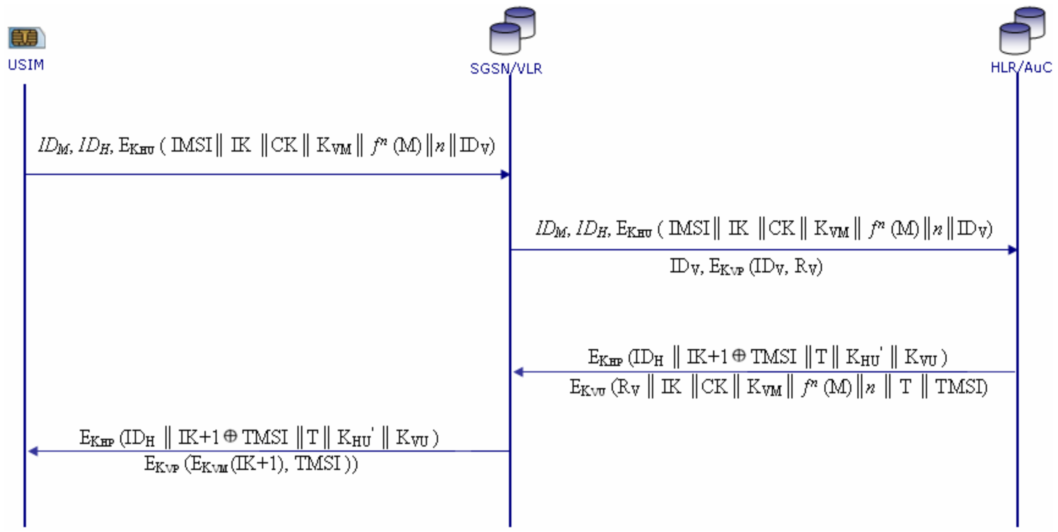

USIM
SGSN/VLR
HLR/AuC
$ID_M, ID_H, E_{K_{HU}}$ (IMSI ∥ IK ∥CK ∥ K_{VM} ∥ f^n (M) ∥n ∥ID_V)
$ID_M, ID_H, E_{K_{HU}}$ (IMSI ∥ IK ∥CK ∥ K_{VM} ∥ f^n (M) ∥n ∥ID_V)
$ID_V, E_{K_{VP}}$ (ID_V, R_V)
$E_{K_{HP}}$ (ID_H ∥ IK+1 ⊕ TMSI ∥T ∥ K_{HU}' ∥ K_{VU})
$E_{K_{VU}}$ (R_V ∥ IK ∥CK ∥ K_{VM} ∥ f^n (M) ∥n ∥ T ∥ TMSI)
$E_{K_{HP}}$ (ID_H ∥ IK+1 ⊕ TMSI ∥T ∥ K_{HU}' ∥ K_{VU})
$E_{K_{VP}}$ ($E_{K_{VM}}$(IK+1), TMSI))

Figure 3.1: Proposed Initial authentication procedure

Step 1: M1 Authentication Request Message

When an MS needs to authenticate itself to all entities of network to access or utilize of network services, MS invokes the distribution of authentication procedure by sending the Authentication Request messages to HLR/AuC (***AUTH***$_{\text{MH}}$) through VLR. Authentication between MS and his HLR/AuC relies on the use of its public key KHU. This process is achieved as follows:

The MS generates the following:

1. The Nonce Numbers IK, CK

2. The Session Key K_{VM}

3. M is secret information

4. f^n(M) where f^n(M) is a one-way hash function and n represents the maximum number of services that the MS can request after initial authentication. Here f^n ( ) = $f(f^{n-1}$ ( )), f^1( ) = f( ).

The MS sends $\boldsymbol{AUTH_{HM}}$ to VLR:

$$\mathrm{AUTH_{HM}} = ID_M, ID_H, \mathrm{E}_{\mathbf{KHU}}\ (\ \mathrm{IMSI} \,\|\, \mathrm{IK} \,\|\mathrm{CK}\|\, \mathrm{K_{VM}} \|\, f^n\ (\mathrm{M}) \,\|\, n \,\|\, \mathrm{ID_V})$$

Where: $\mathrm{ID_M}$ is the identification of the MS that HLR can verify his signature.

Step 2: M2 Authentication Request Message

When the VLR receives the message from the MS, it passes the message to the HLR, and sends the $\mathrm{AUTH_{VH}}$ of its challenge message to the HLR. However, the $(\mathrm{ID_V} \| \mathrm{R_V})$ $\mathrm{K_{VP}}$ is encrypted by VLR private key using Encryption/digital signature RSA algorithm. After receiving these messages, the HLR decrypts by their corresponding private and public key them access the database to obtain the CertM and CertV, respectively.

$$\mathrm{ID_V},\ \mathrm{E_{KVP}}\ (\mathrm{ID_V},\ \mathrm{R_V})$$

Step 3: M3 Authentication data Response Message

The HLR sends RAUTHHM encrypted by HLR's private key and $RAUTH_{HV}$ encrypted by VLR’s public key of its response messages to the VLR, respectively. After receiving these messages, the VLR decrypts $\mathrm{RAUTH_{HV}}$ by his secret key to get $\mathrm{R_V}$, TMSI, IK, CK, $\mathrm{K_{VM}}$, f^n (M), n and T. Then, the VLR saves f^n(M), n, and CK for subsequent authentication and session key generation.

$$\mathrm{RAUTH_{HM\,=}\,E_{\mathbf{KHP}}}\ (\mathrm{ID_H} \,\|\, \mathrm{IK{+}1} \oplus \mathrm{TMSI} \,\|\mathrm{T}\|\, \mathrm{K_{HU}}' \,\|\, \mathrm{K_{VU}}\,)$$

$$\mathrm{RAUTH_{HV}} = \mathrm{E_{KVU}}\ (\mathrm{R_V} \,\|\, \mathrm{IK} \,\|\mathrm{CK}\|\, \mathrm{K_{VM}} \|\, f^n\ (\mathrm{M}) \,\|\, n \,\|\, \mathrm{T} \,\|\, \mathrm{TMSI})$$

Step 4: M4 Authentication Response Message

The SGSN/VLR sends $\boldsymbol{RAUTH}_{\mathrm{HM}}$ encrypted by HLR’s private key and $\boldsymbol{RAUTH}_{\mathrm{VM}}$ encrypted by VLR’s private key of its response to the MS by using Encryption/digital signature RSA algorithm. After receiving these messages, the MS decrypt ***RAUTH***$_{HM}$ by HLR's public key to get K_{VU}, K_{HU}’, IK+1, TMSI and ID_H. After getting K_{HU}’, it knows that key refreshment is successfully. Filially, it gets KVU that has sent from HLR to decrypt ***RAUTH***$_{VM}$ to obtain E_{KVM} (IK+1), TMSI and then encrypt E_{KVM} (IK+1) to get IK, then MS verifies the value of (IK + 1) if it is correct, then the authentication is successful, and the MS gets new temporary identities, TMSI’. Also, K_{VM} becomes the shared key used by the MS and SGSN/VLR, the authentication process is finished.

$$RAUTH_{HM} = E_{KHP} (ID_H \parallel IK+1 \oplus TMSI \parallel T \parallel K_{HU}' \parallel K_{VU})$$

$$RAUTH_{VM} = E_{KVP} (E_{KVM} (IK+1), TMSI))$$

3.4.2 Subsequent authentication procedure

After the initial authentication, SGSN/VLR gets a secret key K_{VM} that it shares with the MS and subsequently can accomplish the authentication by itself. That is, subsequent authentication only happens between the MS and SGSN/VLR using two message exchanges. Figure 3.2 exhibits the subsequent authentication procedure, and the authentication steps are described as follows.

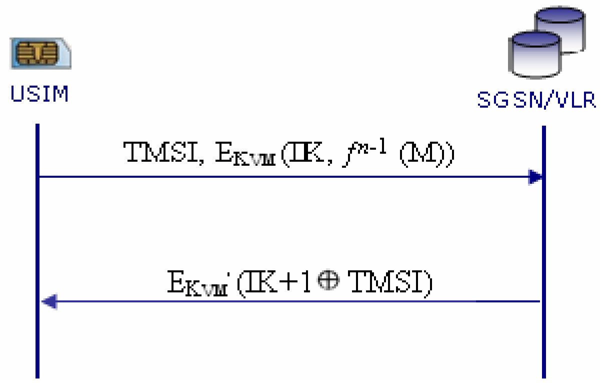


Figure 3.2: Subsequent authentication procedure

The notations in Figure 3.2 are defined as follows:

- ***$SAUTH_{MV}$*:** TMSI, E_{KVM} (IK, f^{n-1} (M))
- ***$SRAUTH_{VM}$*:** $E_{KVM}{}'$ (IK+1, $\oplus$ TMSI)

Step 1: The MS generates a new session key $K_{VM}' = h\ (K_{VM}, IK \oplus CK)$, where *IK* is a new generated nonce and K_{VM} is the shared key. *CK* is in the messages sent by the MS to the VLR in the initial authentication procedure. Meanwhile, the MS produces $f^{n-i}(M)$, where *i* is the number of services that have been requested, and *M* is the secret key generated in the initial authentication. The MS sends ***$SAUTH_{MV}$*** that encrypted the session key ***K_{VM}*** to the VLR, which contain *IK*, $f^{n-i}(M)$.

$SAUTH_{UV}$: TMSI, E_{KVM} (IK, f^{n-1} (M))

Step 2: SGSN/VLR first checks the subscribed service period of the mobile user for the requested service. If the service request is not made within the valid subscribed service period, the service request is rejected. The procedure then restarts from step 1.1. Else SGSN/VLR decrypt ***$SAUTH_{MV}$*** by the shared session key K_{VM} and compares the result with *IK*. Moreover, SGSN/VLR computes $f(f^{n-i}(M))$ to verify whether it is the same as the number, $f^{n-i+1}(M)$, which SGSN/VLR saved in the last authentication. If they are identical, the MS has been authenticated successfully. SGSN/VLR send ***$RSAUTH_{MV}$ = (IK+1 $\oplus$ TMSI***) encrypted by the new session key K_{VM}'**,** which was generated using the same function that used by the MS were the encryption type is AES algorithm. Upon the MS receipt of the response message, the MS decrypts the authenticator using K_{VM}'.

$SRAUTH_{UV}$: $E_{KVM}{}'$ (IK+1, $\oplus$ TMSI)

3.5 Achieved goals

The achieved goals of the proposed protocol are described as follows:

(1) Entity authentication of the MS to VLR:

The MS sends challenge random numbers IK,CK to VLR. After receiving $\boldsymbol{RAUTH}_{\mathrm{VM}}$ encrypted by VLR's private key and the session shared key, MS use K_{VM} and K_{VM} to decrypt it to get IK and verify legitimacy of VLR's identity.

(2) Entity authentication of the MS to HLR:

First, MS sends $\boldsymbol{AUTH}_{\mathrm{HM}}$ to challenge HLR under HLR's public key. After that HLR sends $\boldsymbol{RAUTH}_{\mathrm{HM}} = E_{\mathbf{KHP}}$ ($ID_H \parallel IK+1 \oplus TMSI \parallel T \parallel K_{HU}' \parallel K_{VU}$) to response the MS based on HLR's private key. When MS gets the ID_H and K+1 from response message, it can verify legitimacy of HLR's identity.

(3) Entity authentication of the VLR to MS:

The MS send ID_V to HLR, and VLR wait the response message from HLR to gets R_V of the Challenge random number to verify that HLR is legal. Consequently, VLR decide this message is response to the request message from legal user. While if HLR reject that request message from that user then the VLR decide the user is illegal user.

(4) Entity authentication of the HLR to MS

By verifying the challenge message ID_M, ID_H, $E_{\mathbf{KHU}}$ ($IMSI \parallel IK \parallel CK \parallel K_{VM} \parallel f^n(M) \parallel n \parallel ID_V$), the HLR knows that the message comes from the MS. This is because that encrypted based on HLR's public key.

(5) Entity authentication of the HLR to VLR:

By decrypting $(ID_V \| R_V)_{KVP}$ the HLR knows that the challenge message comes from the VLR. This is because that encrypted by VLR's private key.

(6) Entity authentication of the VLR to HLR:

By decrypting E_{KVU} $(R_V \| IK \| CK \| K_{VM} \| f^n (M) \| n \| T \| TMSI)$, VLR gets R_V of the Challenge random number to verify that HLR is legal.

(7) Assignment of a new session key is assigned to MS and VLR.

(8) Confidentiality: The TMSI is used to protect MS true identity.

3.6 Security analysis of the proposed protocol

In accordance with the proposed scheme, it is assumed that a VLR has powerful computation ability and has no worry about power supply, which means it can handle more complex calculations. Since we consider the low computation ability and low power of mobile equipments, we make the RSA encryption be clone in MS side. Table 3.1 indicates that the RSA encryption provides far superior performance than decryption.

Furthermore, the session key K_{VM} is generated by MS. The VLR will use the key to verify MS again when he requires a service, e.g. making a call etc. We assume that the authentication key is generated through a secure random number generator and kept securely for each related parties. Based on the recommendation of RSA Laboratories (Kaliski, 2003) and NIST (NIST, 2003), the public key length that we use is 1024 bits. For NIST, 1024 bits public key length is appropriate for protecting data through the year 2015, which means it can hardly be broken using today's computer technology. Table 3.2 gives the comparison of the recommendation of NIST and RSA Laboratory.

Since the security of public-key cryptography depends on the key length and assumes that factoring this large numbers is very hard. From the previous suggestion, we assume the public key pair (*e, n*) of VLR is secure, and the private key d is safe and known only by VLR. Furthermore, KHP is supposed to be kept secretly by HLR. Base on these hypothesis, we make the following security analysis to prove that modification is robust and against to replaying, guessing and substitution attacks. Note that the analysis is based on the UMTS.

Units	Year	Minimum Key Sizes
NIST	Present-2015	1024 bits
RSA Laboratories	Present-2010	1024 bits

Table 3.2: Comparison of the recommendation of NIST and RSA Laboratory

In this section, we will discuss the security of the proposed protocol. In order to ensure that the proposed protocol is secure, we will analyze and discuss the attack methods. The security requirements of third generation mobile systems are mutual authentication, MS anonymity, end-to-end security, non-repudiation, and data integrity and data confidentiality (3GPP, 1999a) (3GPP, 2001b). The proposed scheme can fulfil all of these requirements.

First, consider the authentication requirement. It is clear that the proposed authentication protocol can authenticate MS, HLR/AuC and SGSN/VLR. Because the message sent to the HLR/AuC is encrypted using its public key K_{HU}, there is no one except for the home HLR/AuC can decrypt the message. Therefore, authentication between the MS and the HLR/AuC can be achieved using K_{HU}. Consequently, mutual authentication is achieved.

Next, consider MS anonymity. To provide MS anonymity, the permanent identity IMSI of MS is never exposed in the plain-text mode. A cracker cannot get the real identity of MS by eavesdropping on the authentication messages on both wireless and wired networks.

Consider the requirement of non-repudiation; our protocol can also satisfy this requirement. By using the one-way function, we can achieve non-repudiation. In the i-th session, the user provides f^{n-i} (M) to ask for a connection. The SGSN/VLR can verify the correctness of f^{n-i} (M) by means of the one way function, but it cannot derive f^{n-i} (M) from f^{n-i+1}(M). In this way, f^{n-i} (M) can be used as a proof of the i-th connection. Whenever a random challenge occurs, the SGSN/VLR can be required to show f^{n-i} (M).

The requirement of end-to-end security is also addressed in the proposed protocol. When a MS makes a call, the caller and callee negotiate a common encryption key to encrypt the data flowing over the channel. Since the full communication path is protected, data confidentiality and integrity are both achieved. The communication between the two parties in both wired and wireless paths is protected with the encryption key K_e. Therefore, end-to-end security is achieved in this way.

The proposed authentication protocol achieves all the requirements shown above. The proposed scheme is superior to other published schemes. The proposed protocol can prevent common attacks as follows:

i. **Replay attacks** It can repulse replay attack, a form of network attack in which a valid data transmission is maliciously or fraudulently repeated or delayed. Hackers capture old messages and replay them at later times. By replying to the message it appears to be legal. Suppose *MS* wants to prove its identity to the

HLR. The *HLR* requests its *IMSI* as proof of identity, which *MS* dutifully provides (possibly after some transformation like a hash function). Meanwhile, the hacker is eavesdropping on the conversation and keeps the *IMSI.* After the interchange is over, the hacker connects to the *HLR* posing as the first *MS*. When asked for proof of identity, the hacker sends the first MS's *IMSI* read from the last session, which the HLR must accept.

The proposed authentication protocol can prevent the replay attack by the freshness properties. The MS refreshes the session key by using the nonce to ensure the freshness of authentication sessions. Since the MS and SGSN/VLR must input IK to generate a new session key; the session key can be refreshed for each authentication process. If an attacker replaces the TMSI of an intercepted authentication message and replays the authentication request, the attack will not succeed because the $\boldsymbol{AUTHM}_{\mathrm{H}}$ is encrypted using K_{HM}, so the replay attack is infeasible.

ii. Guessing Attacks Password authentication is widely used by many security systems. However a password is vulnerable under dictionary attack in which an attacker can guess the password successfully. In the proposed scheme, the key refreshment method prevents the guessing attack. In each authentication process, the subscriber uses different keys to request registration and authentication. Public key cryptosystems provide a means for preventing the guessing attack. Since the public key digital signature is used to sign the message, the guessing attacks fail. It has been shown that the signatures are distributed so that the attacker is unable to guess a signature value which is shared by different messages.

iii. **Substitution Attacks** If an attacker replaces some fields of the authentication request, then the authentication will fail. For example, if an attacker replaces any parameter in the initial authentication, the SGSN/VLR will find the nonce *IK'* is different from the nonce *IK* encrypted along with the authentication status sent by HLR/AuC. Because the nonce used by SGSN/VLR is the same as the one used by HLR/AuC, SGSN/VLR can compare them to verify MS. Therefore, our protocol can resist the substitution attack. Moreover, even if an attacker gets the session key, he will not have the ability to generate new session keys. This is because session key generation involves *CK*, and because only the MS and SGSN/VLR know the *CK*. The above analysis shows that our protocol can successfully prevent this kind of substitution attack.

From the above security analysis, we can find that our scheme is secure and fully satisfies the security requirements. Furthermore, the authentication key $\boldsymbol{K_{VM}}$ is only aware by MS and VLR. It can assure that in the later services request phase, only the legitimate MS will be allowed to use the service provided by VLR. Since MS has no reason to compromise the $\boldsymbol{K_{VM}}$ to a third party, therefore, our scheme ensures that except the related participants, no one can harm the rights and interests of MS

3.7 Summary

The proposed authentication for UMTS has been developed with the aim of keeping the complexity of this function as low as possible, and providing a high level of security and efficiency for the bandwidth used. A detailed analysis has been made of how the proposed scheme meets the security requirements. The proposed scheme deters intruders from replaying attack, guessing attack and substitution attack.

Chapter 4 Analysis of the Proposed Authentication Protocol

This chapter investigates and analyses the Authentication and Key (AKA) protocol for Universal Mobile Telecommunications System (UMTS) mobile networks, and the proposed authentication protocol. In the proposed protocol, the number of messages between authentication entities of the network is reduced to four messages instead of five in the initial authentication procedure during registration procedure. In addition, the subsequent authentication procedure only contains two message exchanges during call origination/termination. Reducing the number of messages will impact on the network signalling overhead and authentication delay proportionally. Consequently, the call setup time is minimized without compromising the UMTS security. However, the proposed protocol reduces the communication overhead significantly between the home network and the visited network, especially for roaming authentication, by reducing the number of authentication vectors to one (AV) instead five which decreases the cost.

In the previous chapter, detailed analysis was made of the proposed scheme achieves the security requirement and the proposed scheme deters intruders from replaying attacks, guessing attack, substitution attack and impersonating attack. In this chapter, two analyses of the proposed scheme are given. A fluid mobility model is used to investigate the performance of signalling traffic and load transaction messages between mobile databases, such as the Home Location Register (HLR) and the Visitor Location Register (VLR), for both the current protocols and the proposed protocol. The simulation results show that the authentication delay and current load transaction messages between entities and bandwidth are reduced compared with the current protocols. The performance and authentication delay time have been improved significantly.

4.1 Analysis of UMTS authentication protocol

In a UMTS network, a VLR/SGSN manages one or more Radio Network Controller (RNCs) and a RNC manages a set of Node Bs. To track the location of an MS, some geographical groups are defined within the Terrestrial Radio Access Network (UTRAN) (Holma & Toskala, 2000):

- Location Area (LA): A *LA* covers the area of one or more Radio Network Subsystems (RNS), if the corresponding RNCs are managed by the same SGSN.
- Routing Area (RA): A *RA* is a subset of a *LA*. It only covers one RNS or a subset of a RNS.
- UTRAN Registration Area (URA): An URA is a subset of an RA. It only covers some Node Bs of one RNS.

UMTS service area is partitioned into several groups by cell Node Bs. To deliver services to an MS, the cells in the group covering the MS will page the MS to establish the radio link. In order to track the location change of an MS, the cells broadcast their cell identities. The MS periodically listens to the broadcast cell identity, and compares it with the cell identity stored in the MS's buffer. If the comparison indicates that the location has been changed, then the UE sends the location update message to the network [Holma & Toskala, 2000].

In the CS-domain, the location Areas (*LAs*) are used: this is out of the scope of this thesis. While in the PS-domain the cells are portioned into RAs, which this work focuses on. The RA of an MS is tracked by the Core Network (CN) (that means by the SGSN). In case of an active Radio Resource Control (RRC) connection the current URA and UE is located in and is tracked by the UTRAN. If the UE is also cell connected, the UTRAN even tracks the cell the UE is located in. Figure 4.1 illustrates an example of RA and URA layout. When the MS moves from one RA to another, a location update is performed, which informs the VLR/SGSN of the MS's current location. Note that a crossing of two RAs within a VLR/SGSN area requires an *intra-SGSN* location update, while a crossing of two RAs of different SGSN areas requires an *inter-SGSN* location update. Details of location update and mobility management modelling can be found in (Akyildiz, et al., 1999) and (Fang & Chlamtac, 1999).

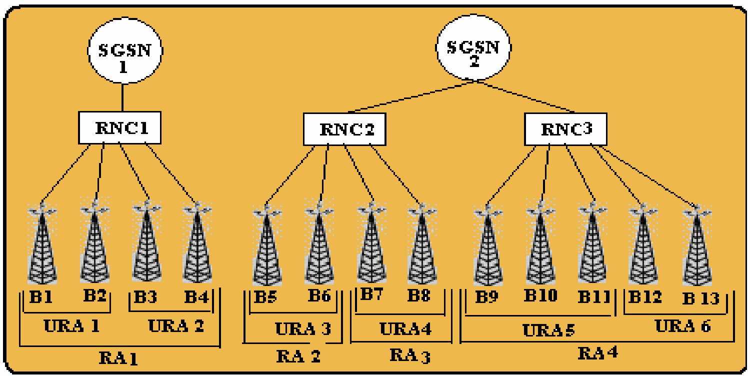


Figure 4.1: **Registration Area (RA) and UTRAN Registration (URA) Layout**

In UMTS, authentication function identifies and authenticates an MS, and validates the service request type to ensure that the user is authorized to use the particular network services. Specifically, authentication is executed for every registration, call origination, and call termination. The registration process triggers will operate in two different circumstances. First, when the mobile is turned on. Second when the mobile moves from one registration area (RA) to another. The MS has to register in the new registration area and deregister from the old registration area. While the mobile movement is within registration area (RA) then the registration process will be not necessary. While this mechanism gives a reasonable level of security it generates a large amount of signalling which increases the call setup time.

4.1.1 Signalling load of the UMTS protocol

This analysis investigates, analyses, and compares the performance of the current protocol (UMTS) and the proposed protocol. A fluid flow mobility model is adopted which is employed in Mohan and Jain, (1994), Mohan (1996), Thomas, et al. (1988) to gauge the impact of the enhancement on the signalling traffic, load, and bandwidth that are generated by these protocols and the delay in the call setup time. This model assumes the following parameters:

1. Mobile user mobility at an average velocity v.
2. Mobility direction of movement is uniformly distributed over $[0,2\pi]$.
3. Mobile users are uniformly distributed in a registration area with the density ρ.
4. The registration area (RA) boundary is given as length L.

The rate of registration area crossing R, is given by Mohan (1996):

$$R = \frac{\rho \times v \times L}{\pi} \quad \text{...} \quad (4.1)$$

The above model is used with the following assumptions to analyze the traffic involved in the authentication process Table 4.1 (Mohan, 1996):

Parameter	**Value**
Total registration area (RA)	128
Square registration area size	8.65km^2=74.8225sq km
Border length L	32.45 km
Mean density of MS ρ	328 km^2
Total number of MS users	3.141*10^6 million
Average call origination rate	1.4/h/user
Average call termination	1.4/r/user
Average speed of user who carrying mobile ***v***	5.6 km/h

Table 4.1: Assumptions parameters (Mohan, 1996).

The traffic due to registration is generated by MS moving into a new registration area (*RA*). A fluid model is assumed with local balance within registration areas (*RAs*). In the steady state, the rate at which users move into an *RA* is equal to the rate of deregistration which MS move out of that *RA* (registration cancellations). The traffic computations are as follow:

Since every registration area is handled by one *VLR*, the rate of registration area crossing, *R*, is given by:

$$R_{Reg,RA} = \frac{\rho \times v \times L}{\pi}$$

$$R_{Reg.,RA} = \frac{\rho \times v \times 32.45}{3600 * \pi} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (4.2)$$

$$R_{Reg.,RA} = \frac{328 \times 5.60 \times 32.45}{3600 * \pi} = 5.27 / sec$$

According to above assumption, the rate of deregistration area crossing *R* is

$$R_{Dereg.,RA} = \frac{\rho \times v \times 32.45}{3600 * \pi} = 5.27 / \mathrm{sec}$$

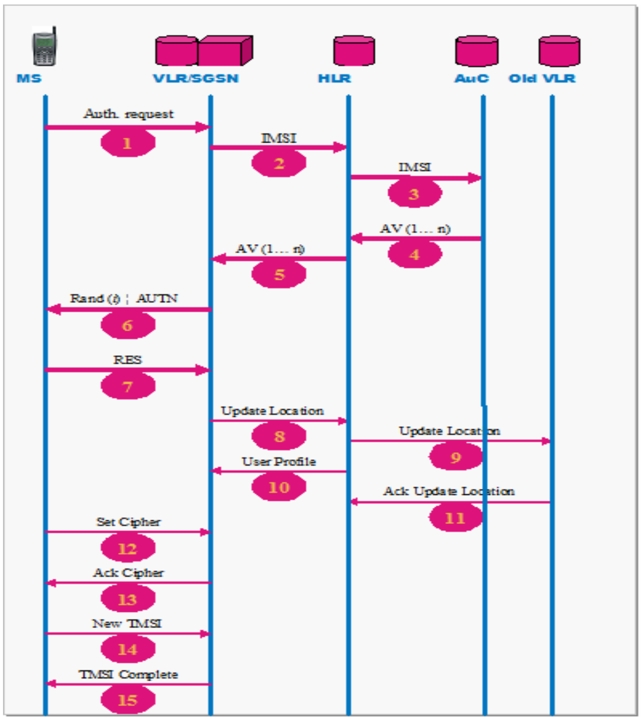


Figure 4.2: UMTS challenge/response signalling messages flow for registration

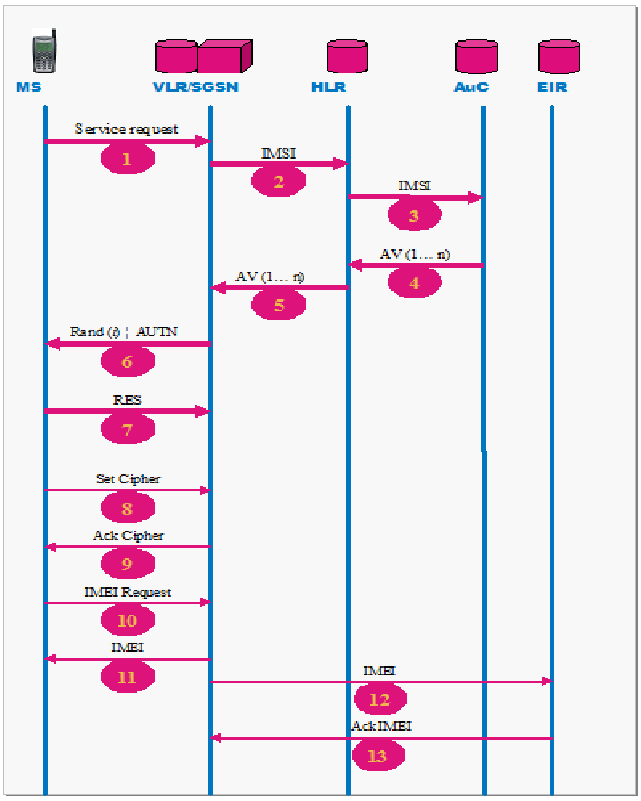


Figure 4.3: UMTS call origination authentication signalling messages

Thus the total number of registration messages per second arriving at the *HLR* (i.e., this is equivalent to registration in the serving network when the user will enter to new serving area) is:

R

$\boldsymbol{R_{Reg.,,HLR} = R_{Reg.,RA}} \times$ ***Total number of registration areas*****,**

$$R_{Reg.,HLR} = R_{reg.,RA} * 128 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(4.3)}$$

$$R_{Reg.,HLR} = 5.27 * 128 = 674.58 / \sec$$

From the equations (4.2), (4.3) we have the rate of registration per registration area and registrations in serving area, authentication requests per each ***RA*** and ***HLR*** respectively. The total number of authentication requests for call origination is computed in the entire serving network area (*SN*) as follows:

The total number of call originating per serving area is equal to:

$$R_{call.Origination/SN} = \frac{\boldsymbol{call\ origination\ rate \times \rho \times 74.82 \times 128}}{3600} \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(4.4)}$$

$$R_{call.Origination/SN} = \frac{\boldsymbol{1.4 \times 328 \times 74.82 \times 128}}{\boldsymbol{3600}} = \boldsymbol{1221.64} / sec$$

The total number of authentication requests due to call origination per serving network (*SN*) and the total number of authentications due to call termination per serving network are symmetric with respect to the number of messages. Thus the $R_{\boldsymbol{call.Termination/SN}}$ is calculated as follows:

$$R_{\boldsymbol{call.Termination/SN}} = \frac{\boldsymbol{call\ termination\ rate \times \rho \times 74.82 \times 128}}{3600} \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(4.5)}$$

$$R_{call.Ter\min ation/SN} = \frac{1.4 \times 328 \times 74.82 \times 128}{3600} = 1221.64 / \sec$$

The number of calls origination per registration area ($R_{\text{Call origination}/RA}$) is calculated as follows:

$$R_{Call.Originartion\,/\,RA} = \frac{\boldsymbol{call\ origination\ rate \times \rho \times 74.82}}{3600} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(4.6)}$$

$$= \frac{\boldsymbol{1.4 \times 328 \times 74.82}}{\boldsymbol{3600}} = \boldsymbol{9.54}\,/\,sec$$

The number of calls terminating per registration area ($R_{Call\ Termination/RA}$) and the number of calls originating per registration area, $R_{Call\ Termination/RA}$ are symmetric with respect to the number of messages. Thus the $R_{Call\ Termination/RA}$ is calculated as follows:

$$R_{Call.Ter\,min\,ation\,/\,RA} = \frac{\boldsymbol{call\ termination\ rate \times \rho \times 74.82}}{3600} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(4.7)}$$

$$= \frac{\boldsymbol{1.4 \times 328 \times 74.82}}{\boldsymbol{3600}} = \boldsymbol{9.54}\,/\,sec$$

Table 4.2 summarises the total authentication request rate per VLR and HLR for each type of activity as computed above.

Activity	VLR/sec	HLR/sec	Total
Registration(*Reg*)	5.27	674.58	**679.85**
Call origination (*Orig*)	9.54	1221.64	**1231.18**
Call termination (*Term*)	9.54	1221.64	**1231.18**
Total/Network	**24.35**	**3117.86**	**3142.21**

Table 4.2: Total Authentication request per VLR (serving area) and HLR

Figure 4.2 and Figure 4.3 present the signalling messages flow between all entities for UMTS networks. This information is used to compute the number of signalling messages per authentication request for registration, call origination and call termination. Table 4.3 summarises the signalling messages flow for each activity.

Activity	AuC	HLR	VLR	Old VLR	Total
Registration(*Reg*)	2	4	5	1	**12**
Call Origination (*Orig*)	2	4	5	0	**11**
Call Termination (*Term*)	2	4	5	0	**11**
Total/Network	**6**	**12**	**15**	**1**	**-**

Table 4.3: Signalling messages per Authentication request for each activity

The total signalling traffic and load transaction messages between mobile database (VLR and HLR) for each activity are illustrated in Table 4.4. It is calculated from the values in Tables 4.2 and 4.3.

Activity	AuC	HLR	VLR	Old VLR	Total
Registration(***Reg.***)	1349.16	2698.32	26.35	5.27	**4079.1**
Call origination (***Orig.***)	2443.27	4886.54	47.72	0.00	**7377.53**
Call termination (***Term.***)	2443.27	4886.54	47.72	0.00	**7377.53**
Total/Network	**6235.7**	**12471.4**	**121.79**	**5.27**	**-**

Table 4.4: Total signalling traffic and load transaction messages per each activity

From the above equations, and calculations, it has been found the relationships between velocity of movement of users and the total authentication requests per VLR and HLR for UMTS authentication process is directly proportional.

4.1.2 Authentication delay of the UMTS protocol

To compute the time delay, it is assumed $\boldsymbol{T}_{\mathbf{Delay}}$, the $\boldsymbol{T}_{\mathbf{Delay}}$ is divided into two parts. First, the time delay due to network database (DB) messages exchange and to access VLR and HLR database is assumed to be T_{DB}. Second, the time delay between MS and VLR is assumed to be $T_{\mathrm{MS\text{-}VLR}}$. So, the time delay can be computed as follows:

$$\boldsymbol{T}_{\mathbf{Delay}} = \textit{\#of messages between } \mathrm{MS} \Leftrightarrow \mathrm{VLR} * T_{\mathrm{MS\text{-}VLR}} + \textit{\# of messages between databases} * T_{\mathrm{DB}}$$

From the equation above, it has been shown that the authentication delay is linearly dependent on the $T_{\mathrm{MS\text{-}VLR}}$. Figure 4.4 depicts the authentication delay $\boldsymbol{T}_{\mathbf{Delay}}$ for the UMTS protocol, which demonstrates the two types of time delay (T_{DB}, $T_{\mathrm{MS\text{-}VLR}}$).

There are three number messages between MS⇔VLR, (1, 6, and 7), and there are four messages between databases (VLR, HLR), which are numbered (2, 3, 4, 5). Thus the total time delay T_{Delay} can be computed as follows:

$\boldsymbol{T_{Delay} = 3 * T_{MS\text{-}VLR} + 4 * T_{DB}}$ ………………………...……………………….……… (4.8)

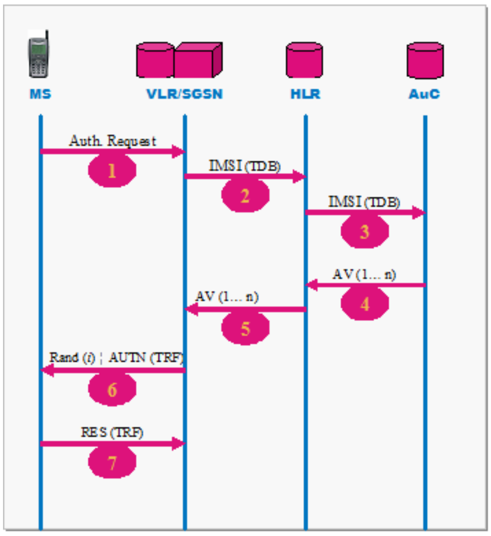


Figure 4.4: Signalling message flow for UMTS authentication delay

4.1.3 Bandwidth requirement for the UMTS protocol

For the UMTS protocol, the Table 4.5 depicts the authentication parameters which are used to calculate the bandwidth for each activity. Thus, the equation of the bandwidth is as follows:

The total Bandwidth = the size of messages * authentication request/sec ………. (4.9)

The size message is assumed to be *S*, the bandwidth between MS⇔VLR and between databases is computed as follows: Firstly, the size messages between MS⇔VLR are computed. Secondly, the size messages between databases are computed.

Parameter	Definition	Bit Length
IMSI	International Mobile Subscriber Identity	128
K	Pre-shared secret key	128
RAND	Random challenge	128
SQN	Sequence number	48
AK	Anonymity key	48
AMF	Authentication Management Field	16
MAC	Message Authentication Code	64
CK	Cipher key	128
IK	Integrity key	128
RES	Authentication response	32
X-RES	Expected response	32
AUTN	Authentication token	128
AV	Authentication vector one record	544
AVs	Standard number of records in authentication vector *K*	5
LAI	Location area identifier	40
Service request	Service request	8

Table 4.5: Authentication parameters for UMTS protocol (3GPP, 2007b).

1. The computation of the messages size between MS ⇔VLR

There are three messages between MS⇔VLR: 1, 6, and 7. So, the total size of messages between MS⇔VLR will take the sum of (S (M1) +S (M6) +S (M7)).

i) The first message (M1) sent through the air from/to the network is the service request, which contain the parameters IMSI/TMSI, service request, and *LAI*. So, the size of M1 computed as follows:

$$S(\mathrm{M1}) = S(\mathrm{IMSI/TMSI}) + S(\text{service request}) + S(LAI)$$

$$= 128+8+40=176 \text{ bits}$$

ii) The sixth message (M6) which is *RAND, AUTHN*. The size of that message is computed as follows:

$$AUTHN = (\mathrm{SQN} \oplus \mathrm{AK} \| \mathrm{AMF} \| \mathrm{MAC}),$$

$$S(AUTHN) = 48 + 16 + 64 = 128 \text{ bits}$$

$S(\text{M6}) = S(RAND) + S(AUTHN)$

$= 128 + 128 = 256$ bits

iii) The seventh message (M7) which is *RES*.

$S(\text{M7}) = S(RES) = 32$ bits

The total size of the authentication messages (M1, M6, M7) between MS⇔VLR can be computed as follows:

$S_{\text{MS-VLR}} = S(\text{M1}) + S(\text{M6}) + S(\text{M7})$

$= 176 + 256 + 32 = 464$ bits

<u>2. The computation of the messages size between databases can be computed as follows</u>

i) The second message (M2), contains the parameters IMSI/TMSI, service request, and *LAI*. So, the size of M1 computed as follows:

$S(\text{M2}) = S(IMSI/TMSI) + S(\text{service request}) + S(LAI)$

$= 128 + 8 + 40 = 176$ bits

ii) The third message (M3), contains the parameters IMSI/TMSI, service request, and *LAI*. So, the size of M1 is computed as follows:

$S(\text{M3}) = S(IMSI/TMSI) + S(\text{service request}) + S(LAI)$

$= 128 + 8 + 40 = 176$ bits

iii) The fourth message (M4), contains the only vector array (*AVs*). The authentication vector (*AV*) is RAND, XRES, CK, IK, and AUTHN. So, the size of *AV* is computed as follows:

$S(AV) = S(Rand) + S(XRES) + S(CK) + S(IK) + S(AUTN)$

$= 128 + 32 + 128 + 128 + 128 = 544$ bits.

Since the *AuC* generates AV array, where the number of *AVs* obtained from the AuC is 5, the total size of M4 is:

$S(AVs) = 5 * 544 = 2720$ bits. So, $S(\mathrm{M4}) = 2720$ bits

iv) The fifth message (M5), contains the same parameters as M4. Therefore, the size of M5 is:

$S(AV) = 2720$ bits.

The total size of the authentication messages (M2, M3, M4, and M5) between databases can be computed as follows:

$S_{\mathrm{DB}} = S(\mathrm{M2}) + S(\mathrm{M3}) + S(\mathrm{M4}) + S(\mathrm{M5})$

$= 176 + 176 + 2720 + 2720 = 5792$ bits

The total size of messages (M1…M7) in the authentication process is computed by taking the sum of $S_{\mathrm{MS\text{-}VLR}} + S_{\mathrm{DB}}$. Therefore, $S_{\mathrm{Auth}} = 464 + 5792 = 6256$ bits $= 782$ bytes.

Table 4.6 summarises the maximum bandwidth for each activity between MS⇔VLR and between databases.

Activity	**Bandwidth between** MS⇔VLR **(B/sec)**	**Bandwidth between Databases (B/sec)**	**Total**
Registration(***Reg***)	305.67	884463.94	**488702.38**
Call origination (***Orig***)	553.55	884463.94	**885017.49**
call termination (***Term***)	553.55	2257324.6	**885017.49**
Total/Network	**1412.77**	**884463.94**	**2258737.36**

Table 4.6: the Bandwidth between entities for UMTS protocol

4.2 Analysis of the proposed authentication protocol

As previously explained, it is proposed to change the number of messages between authentication entities of the network to four messages instead of five in the initial authentication procedure (registration procedure). The subsequent authentication procedure only contains two message exchanges during call origination/termination. Reducing the number of messages will impact on the network signalling overhead and authentication delay proportionally. Consequently, the call setup time is minimized without compromising the UMTS security. Significantly, the communication overhead between home network and visited network is slashed, particularly for roaming authentication, because the number of authentication vectors are cut to one (AV) instead of five. This leads to decreased cost.

Consequently, this approach is secure and practical as it can satisfy the security requirements of third generation mobile communication systems. In this section the analysis of the proposed scheme evaluates the impact of enhancement features on a set of measurements, which are:

1. Signalling traffic load.
2. Time delay.
3. The bandwidth of the proposed protocol.

4.2.1 Signalling load of the proposed protocol

The same model (fluid mobility model) and analysis which are used in investigation and analysis of the UMTS protocol will be used to analyse the impact of signalling messages for each database register (AuC, VLR and HLR) using the rate of authentication requests per second for each activity (registration, call termination and origination) on the performance of proposed protocol. The time delay and the bandwidth of the proposed protocol are analyzed to evaluate their effects on the proposed protocol in each activity.

For the proposed protocol, Figure 4.5 and Figure 4.6 depict the corresponding signalling messages flow between network data bases (registers) for registration, and call origination/termination, respectively. Using these figures, the number of signalling messages per authentication request for the registration, call origination and call termination activities is found as illustrated in Table 4.2. Table 4.7 shows the total signalling messages generated per VLR and HLR for each type of activity. The total signalling traffic and load transaction messages between mobile databases (VLR and HLR) are shown in Table 4.8 and are calculated from the values in Tables 4.2 and 4.7.

Activity	AuC	HLR	VLR	Old VLR	Total
Registration(***Reg***)	2	4	4	1	**11**
Call Origination (***Orig***)	0	0	2	0	**2**
Call Termination (***Term***)	0	0	2	0	**2**
Total/Network	**2**	**4**	**8**	**1**	**-**

Table 4.7: Signalling messages per Authentication request for each activity

Activity	AuC	HLR	VLR	Old VLR	Total
Registration(*Reg*)	1349.16	2698.32	21.08	5.27	**4073.83**
Call origination (*Orig*)	0.00	0.00	19.09	0.00	**19.09**
call termination (*Term*)	0.00	0.00	19.09	0.00	**19.09**
Total/Network	**1349.16**	**2698.32**	**59.26**	**5.27**	**-**

Table 4.8: Total signalling traffic and load transaction messages per each activity

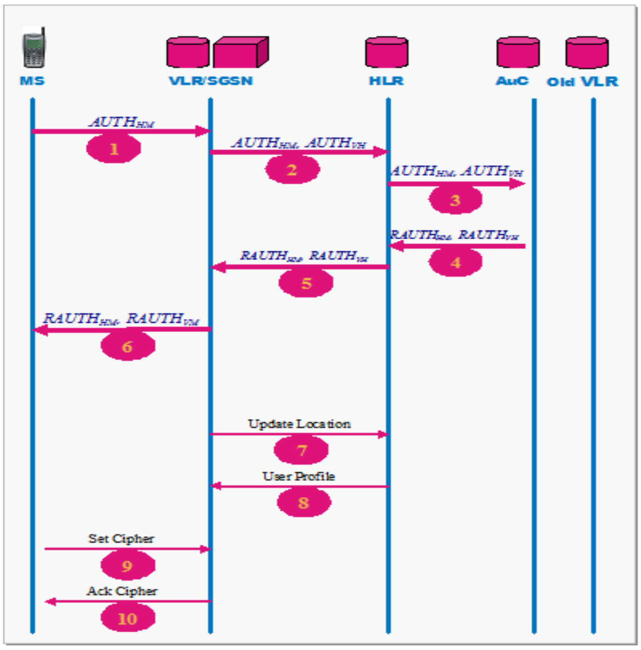


Figure 4.5: Signalling messages flow for the proposed scheme

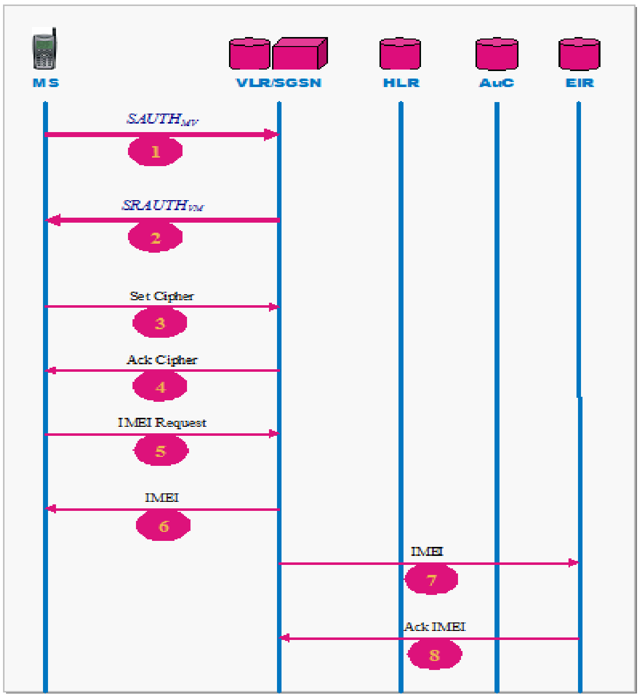


Figure 4.6: The proposed call origination/termination authentication signalling messages

4.2.2 Authentication delay of the proposed protocol

The authentication delay of the proposed scheme is computed with the same assumptions used for the UMTS protocol. T_{DB} is assumed the time delay between network databases message exchanges, and the time that the message takes from MS to

VLR is denoted by $T_{MS\text{-}VLR}$. There are two messages between MS⇔VLR and four messages between databases (VLR, HLR). Thus the total time delay T_{Delay} can be computed as follows:

$$\boldsymbol{T_{Delay} = 2 * T_{MS\text{-}VLR} + 4 * T_{DB}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (4.10)$$

4.2.3 Bandwidth requirement of the proposed protocol

For the proposed protocol, Table 4.9 depicts the authentication parameters which are used to calculate the bandwidth for each activity. Thus, the bandwidth of the proposed protocol is computed using the same equations of the bandwidth for UMTS used in the previous section for computing the size of messages and using the data in Table 4.8.

To compute the bandwidth, there are six messages to authentication; two of them are between MS⇔VLR and the other four are between databases. The size of these messages can be computed as follows.

Parameter	Definition	Bit Length
IMSI	International Mobile Subscriber Identity	128
TMSI	Temporary Mobile Subscriber Identity	128
K_{VM}	Pre-shared secret key(AES)	128
ID_M, ID_V,ID_H	Identity of Mobile Station, SN and HN	20
NK, IK, R_V	Nonce Number s	48
M	Message Authentication Code	512
$E_x(\ )$	Encryption and Decryption Function (AES)	128
AV	Authentication vector one record	544
H( ), $f^n(M)$	Hash Chaining Function(SHA1)	160
LAI	Location area identifier	40
Service request	Service request	8
N	The maximum number of hash chaining composition	8
E_{Kxx}	Length of cipher code and signature (RSA)	1024
N	The Modula (RSA)	1024
E	The Public Key	17

Table 4.9: Authentication parameters for proposed protocol

1. *The computation of the messages size between MS ⇔VLR*

There are two messages between MS⇔VLR which are (1, 6). So, the total size of messages between MS⇔VLR will take the sum of (S (M1) +S (M6))

i) The first message (M1) sent through the air from/to the network is the service request (AUTH_{HM}), which contains the parameters ID_M, ID_H, $\text{E}_{\mathbf{KHU}}$ (IMSI‖ IK ‖CK‖ K_{VM}‖ f^n (M)‖n‖ID_{V}), service request, and *LAI*. So, the size of M1 is computed as follows:

$S(\text{M1}) = S(\text{ID}_{\text{M}}) + S(\text{ID}_{\text{H}}) + S(\text{E}_{\text{KHM}}) + S(\text{service request}) + S(LAI)$

$= 20 + 20 + 1024 + 8 + 40 = 1112$ bits

ii) The sixth message (M6) is $\text{RAUTH}_{\text{HM =}}$ $\text{E}_{\mathbf{KHP}}$ (ID_{H} ‖ IK+1⊕ TMSI ‖T‖ K_{HU}' ‖ K_{VU}) and RAUTH_{VM} = E_{KVP} (E_{KVM} (IK+1), TMSI)). The size of that message is computed as follows:

$S(\text{M6}) = S(\text{RAUTH}_{\text{HM}}) + S(\text{RAUTH}_{\text{VM}})$

$= 1024 + 1024 = 2048$ bits

The total size of the authentication messages (M1, M6) between MS⇔VLR can be computed as follows:

$S_{\text{MS-VLR}} = S(\text{M1}) + S(\text{M6})$

$= 1112 + 2048 = 3160$ bits

2. *The computation of the messages size between Databases can be computed as follows*

There are four messages between databases: 2, 3, 4, and 5. So, the total size of messages between databases will take the sum of (S (M2) +S (M3) + S (M4) + S (M5))

i) The second message (M2), which contain the parameters ID_M, ID_H, $\mathrm{E}_{\mathbf{KHU}}$ (IMSI$\|$ IK $\|$CK$\|$ $\mathrm{K}_{\mathrm{VM}}\|$ f^n (M)$\|n\|$ID$_{\mathrm{V}}$), service request, *LAI, and* ID$_{\mathrm{V}}$, $\mathrm{E}_{\mathrm{KVP}}$ (ID$_{\mathrm{V}}$, R$_{\mathrm{V}}$) . So, the size of M2 is computed as follows:

S(M2) = *S*(ID$_{\mathrm{M}}$) + *S*(ID$_{\mathrm{H}}$) + *S*($\mathrm{E}_{\mathrm{KHM}}$) + *S*(service request) + *S*(*LAI*) + S(ID$_{\mathrm{V}}$) + S($\mathrm{E}_{\mathrm{KVP}}$)

= 20 + 20 + 1024 + 8 + 40 + 20 + 1024 = 2156 bits

ii) The third message (M3), contains the parameters ID_M, ID_H, $\mathrm{E}_{\mathbf{KHU}}$ (IMSI$\|$ IK $\|$CK$\|$ $\mathrm{K}_{\mathrm{VM}}\|$ f^n (M) $\|n\|$ID$_{\mathrm{V}}$), service request, *LAI*, and ID$_{\mathrm{V}}$, $\mathrm{E}_{\mathrm{KVP}}$ (ID$_{\mathrm{V}}$, R$_{\mathrm{V}}$) .

So, the size of M3 is computed as follows:

S(M3) = *S*(ID$_{\mathrm{M}}$) + *S*(ID$_{\mathrm{H}}$) + *S*($\mathrm{E}_{\mathrm{KHM}}$) + *S*(service request) + *S*(*LAI*) + S(ID$_{\mathrm{V}}$) + S($\mathrm{E}_{\mathrm{KVP}}$)

= 20 + 20 + 1024 + 8 + 40 + 20 + 1024 = 2156 bits

iii) The fourth message (M4), which contains only one vector (*AV*). The authentication vector (*AV*) is RAUTH$_{\mathrm{HM}\,=}$ $\mathrm{E}_{\mathbf{KHP}}$ (ID$_{\mathrm{H}}$ $\|$ IK+1$\oplus$ TMSI $\|$T$\|$ $\mathrm{K}_{\mathrm{HU}}'$ $\|$ K_{VU}) and RAUTH$_{\mathrm{HV}}$ = $\mathrm{E}_{\mathrm{KVU}}$ (R$_{\mathrm{V}}$ $\|$ IK $\|$CK$\|$ $\mathrm{K}_{\mathrm{VM}}\|$ f^n (M)$\|n$ $\|$ T $\|$ TMSI). So, the size of *AV* is computed as follows:

S(*AV*)= *S*(RAUTH$_{\mathrm{HM}}$) +*S*(RAUTH$_{\mathrm{HV}}$)

= 1024 + 1024=*2048* bits.

iv) The fifth message (M5) contains the same parameters as M4. Therefore, the size of M5 is:

S(*AV*) = 2048 bits.

The total size of the authentication messages (M2, M3, M4, and M5) between databases can be computed as follows:

$S_{\mathrm{DB}}= S(\mathrm{M2}) + S(\mathrm{M3}) + S(\mathrm{M4}) + S(\mathrm{M5})$

$= 2156 + 2156 + 2048 + 2048 = 8408$ bits

To compute the bandwidth for call origination/termination, there are two messages in subsequent authentication, the first message from MS to VLR it is denoted as message (7). The second one from VLR to MS is denoted as message (8)*;* the sizes of these messages can be computed as follows:

i) The seventh message (M7), which contain the parameters TMSI, $\mathrm{E_{KVM}}$ (IK, $f^{n\text{-}1}$ (M)), service request, and *LAI*. So, the size of M7 computed as follows:

$S(\mathrm{M2}) = S(\mathrm{TMSI}) + S(\mathrm{E_{KVM}}\,(\mathrm{IK}, f^{n\text{-}j}\,(\mathrm{M})) + S(\text{service request}) +S(LAI)$

$= 128+ 128 + 8 + 40 = 304$ bits

ii) The eighth message (M8), which contain the parameters $\mathrm{E_{KVM}}'$ (IK+1, ⊕ TMSI). So, the size of M8 is computed as follows:

$S(\mathrm{M8}) = S(\mathrm{E_{KVM}}'\,(\mathrm{IK+1}, \oplus\ \mathrm{TMSI}))$

$= 128$ bits

The total size of the authentication messages (M7, M8) between MS-VLR for call origination/termination can be computed as follows:

$S_{\mathrm{MS\text{-}VLR}}= S(\mathrm{M7}) + S(\mathrm{M8})$

$= 304 + 128 = 432$ bits

The total size of messages (M1…M6) in the authentication process for registration activity is computed by taking the sum of $S_{\mathrm{MS\text{-}VLR}} + \mathrm{S_{DB}}$.

Therefore, $S_{\text{Auth}} = 3160 + 8408 = 11568$ bits = 1446 bytes. Table 4.10 summarises the maximum bandwidth for each activity between MS⇔VLR and between databases.

Activity	Bandwidth between MS⇔VLR (B/sec)	Bandwidth between Databases (B/sec)	Total
Registration(*Reg*)	2081.72	708984.73	**711066.45**
Call origination (*Orig*)	515.38	0.00	**515.38**
Call termination (*Term*)	515.38	0.00	**515.38**
Total/Network	**3112.47**	**708984.73**	**712097.20**

Table 4.10: The bandwidth between entities for proposed protocol

4.3 Discussion of analytical and simulation results

The simulation study has been carried out in order to analyse signalling traffic performance, load transaction messages and bandwidth consumption between mobile network entities. The simulation is carried out using different mobility rates. The analytic results were validated against the simulation experiments. Therefore, the merit of the proposed protocol can be evaluated with respect to many criteria.

The first criterion is security where the proposed protocol adopts the architecture of UMTS AKA which is a very important issue in this analysis study. In UMTS AKA protocol, security issues include (i) entity authentication between an MS and an SN, (ii) signalling data integrity, (iii) user traffic confidentiality, and (iv) protection against various attacks. The proposed authentication protocol addressed the same security issues as do the current UMTS. The proposed protocol is examined in relation for additional security issues with less signalling traffic and better call set up time. Also, eliminates a few drawbacks of 3GPP AKA which are; the bandwidth consumption between VLR and HLR, storage space in VLR, and the synchronization problem. Additionally it provides an effective method for non-repudiation. In a billing dispute between a user and SN, the hash chains can provide proof of previous visits from the user. The second criterion is the important reduction in signalling messages between the various network elements. Tables 4.11, 4.12, and 4.13 illustrate the differences in total network traffic between the current UMTS authentication protocol and the proposed protocol. Table 4.11 outlines the difference between UMTS protocol and the proposed protocol in terms of signalling messages. The current protocol needs 12 messages between mobile network entities to perform registration and 11 messages to perform the call origination/termination, but the proposed protocol needs only 11 messages to perform registration and 2 messages for call origination/termination. As shown in table 4.13, VLR has 121.79 signalling messages per second in the current protocol and 59.26 in the proposed scheme. The HLR handles around 12471.41 signalling message per second when the current protocol is used and 2698.32 messages per second when the proposed protocol is used. The percentage of improvement is more than 78%. The third determinant factor in this analysis is the authentication delay. From equations (4.8) and (4.10), where it is assumed that T_{DB} = 1ms, the proposed protocol has less delay than the current UMTS protocol as shown in Figure 4.7. The results show that the authentication delay and current load transaction messages between entities and bandwidth are minimised when compared with the current protocol, as illustrated in Figures 4.8, 4.9, 4.10, and 4.11. So, the authentication delay time have been improved significantly.

Activity	Current Protocol				Proposed Protocol			
	AuC	HLR	VLR	Old VLR	AuC	HLR	VLR	Old VLR
Registration(*Reg*)	2	4	5	1	2	4	4	1
Call origination (*Orig*)	2	4	5	0	0	0	2	0
Call termination (*Term*)	2	4	5	0	0	0	2	0

Table 4.11: Comparison of signalling messages between current and proposed authentication protocol.

Activity	Current Protocol				Proposed Protocol			
	AuC	HLR	VLR	Old VLR	AuC	HLR	VLR	Old VLR
Registration	1349.16	2698.32	26.35	5.27	1349.16	2698.32	21.08	5.27
Call origination	2443.27	4886.54	47.72	0.00	0.00	0.00	19.09	0.00
Call termination	2443.27	4886.54	47.72	0.00	0.00	0.00	19.09	0.00

Table 4.12: Comparison of the total signalling traffic and load messages/s between entities for each activity.

Entity	Current Protocol	Proposed Protocol	% improvement
AuC	6235.70	1349.16	**78**
HLR	12471.41	2698.32	**78**
VLR	121.79	59.26	**51**
Total	**18828.90**	**4106.74**	**78**

Table 4.13: Comparison of the total signalling traffic and load messages/sec between entities.

Activity	Bandwidth of Current Protocol			Bandwidth of Proposed Protocol		
	MS/VLR	DB	Total	MS/VLR	DB	Total
Registration	305.67	884463.94	**488702.38**	2081.72	**708984.73**	**711066.45**
Call origination	553.55	884463.94	**885017.49**	501.06	0.00	**515.38**
Call termination	553.55	2257324.6	**885017.49**	501.06	0.00	**515.38**
Total	**1412.77**	**884463.94**	**2258737.36**	**3112.47**	**708984.73**	**712068.57**

Table 4.14: Comparison of the bandwidth for each activity between databases and MS-VLR.

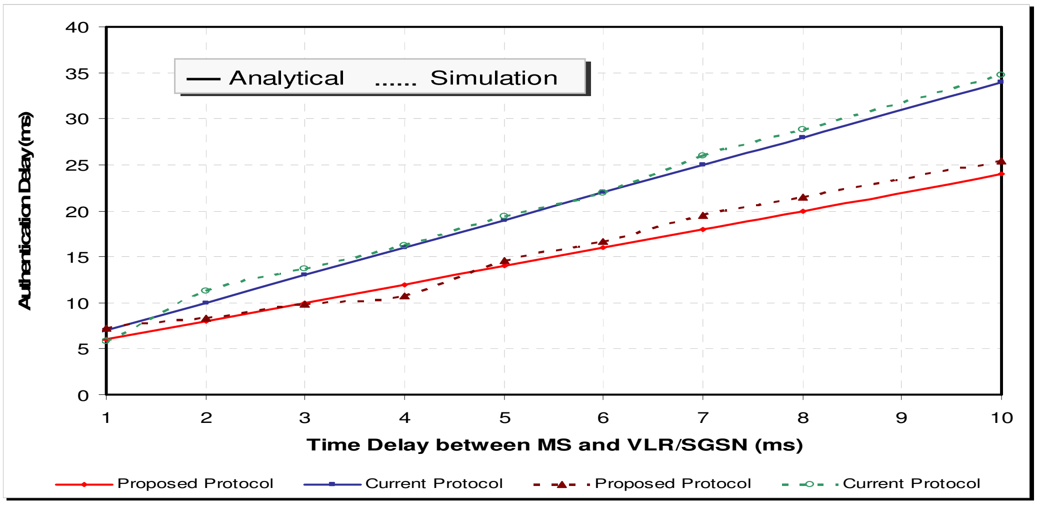


Figure 4.7: Comparison of the authentication delay between UMTS and the proposed protocol when $T_{DB} = 1ms$

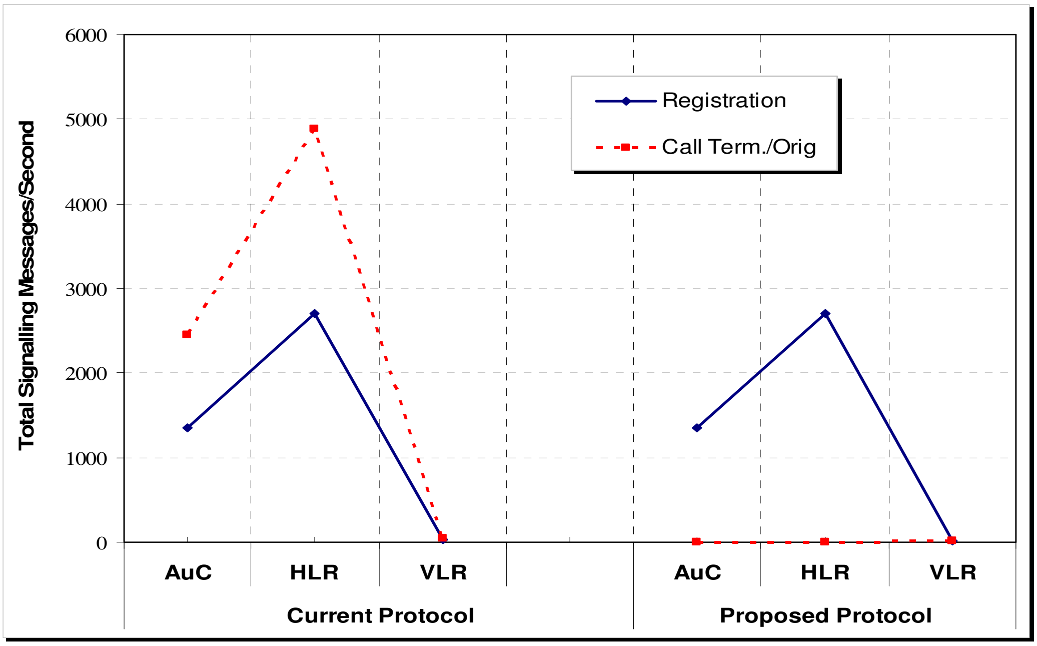


Figure 4.8: Load transaction messages per second between entities.

The analysis adopts a range of velocities: slow pedestrian (5.6 k/h), fast pedestrian (11.2 k/h), slow vehicle (44.8 k/h) and fast vehicle (89.6 k/h). These parameters are used for the Table 4.15 which illustrates the total network signalling traffic with different mobility rates. By varying MS mobility rate (the speed of movement) and call origination rate, it can be seen in Table 4.15 that the advantage of the proposed protocol over the current protocol ranges from over 18 per cent to 78 percent. This conclusion is generally valid, though the percentage of improvement may differ with a different set of assumptions. For example, for a slow pedestrian with high call origination call rates, the performance for the proposed protocol rose to about 78%. This means a user stays in a visited network for long time, which indicates a roaming user makes more calls than normal during it residence in a visited network (intra-network). For a fast vehicle with low call origination call rates, the performance for the proposed protocol decreases to about 18% which indicates a roaming user makes more registrations than normal because the user moves to a new serving network frequently (inter-network).

Velocity (Km/h)	Rate	Current Protocol					Proposed Protocol				
		AuC	HLR	VLR	Old VLR	Total	AuC	HLR	VLR	Old VLR	Total
5.6	5.27	6235.70	12471.41	121.79	5.27	18834.17	1349.16	2698.32	59.26	5.27	4112.01
11.2	10.54	7584.87	15169.73	148.14	10.54	22913.28	2698.32	5396.65	80.34	10.54	8185.85
44.8	42.16	15679.84	31359.68	306.25	42.16	47387.92	10793.3	21586.6	206.82	42.16	32628.88
89.6	84.32	26473.14	52946.27	517.05	84.32	80020.78	21586.6	43173.19	375.47	84.32	65219.57

Table 4.15: Network signalling traffic with different mobility rate.

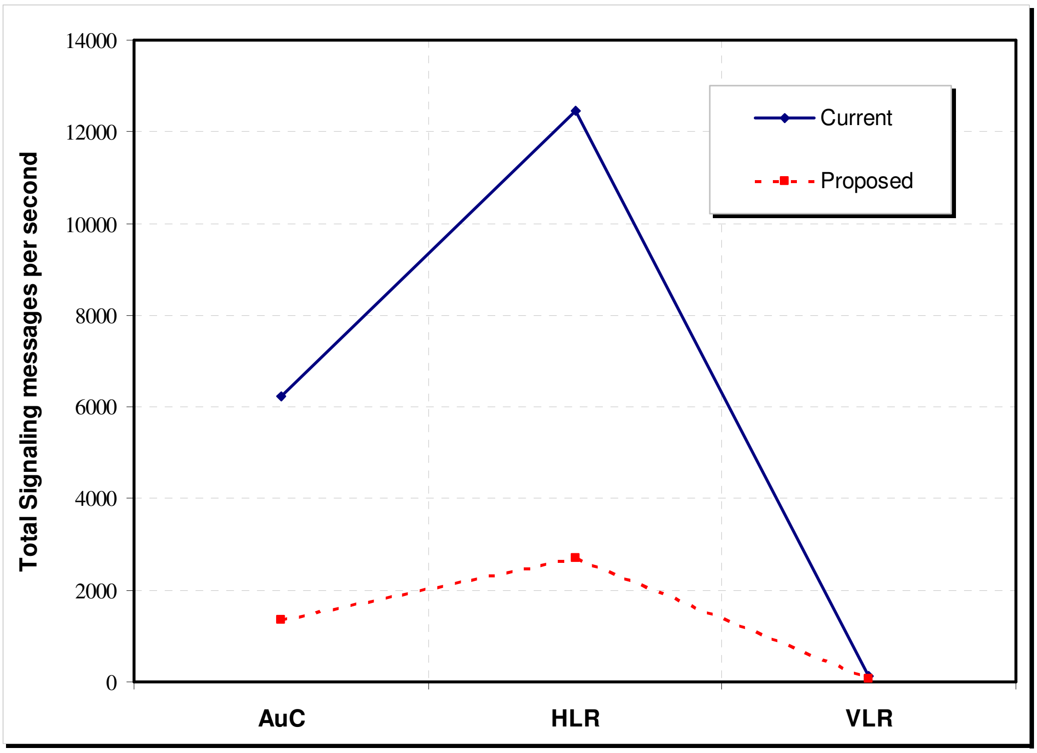


Figure 4.9: Total signalling messages/second for all activity in current and proposed protocol.

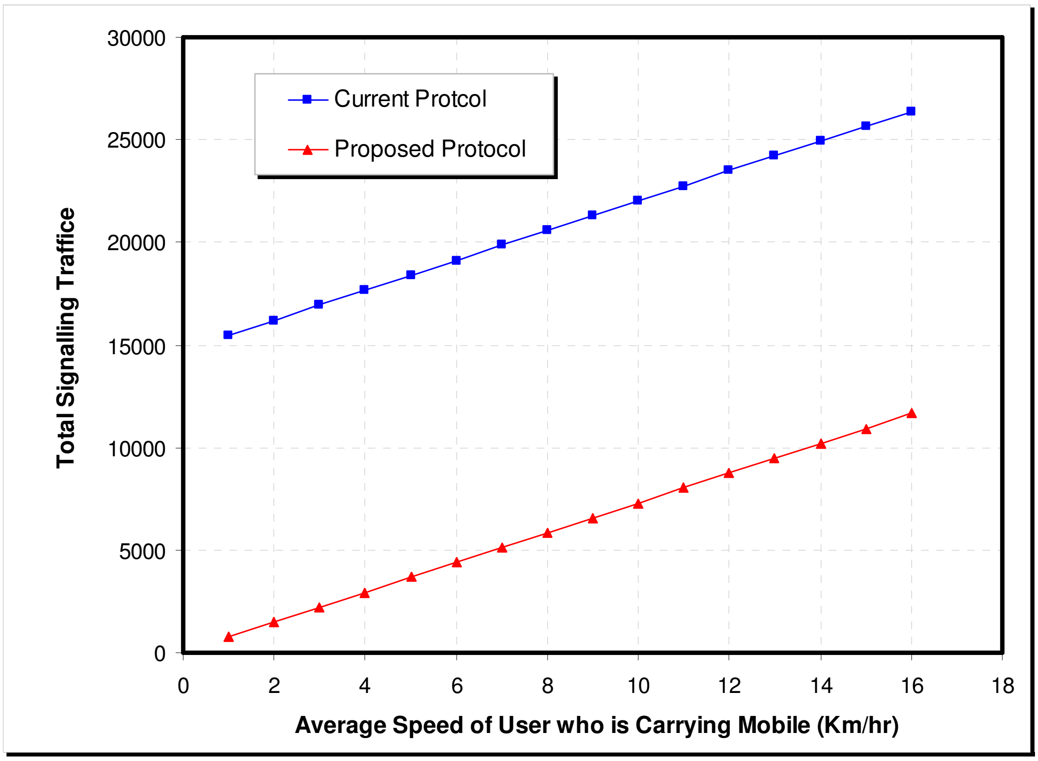


Figure 4.10: Network signalling traffic with different mobility rate.

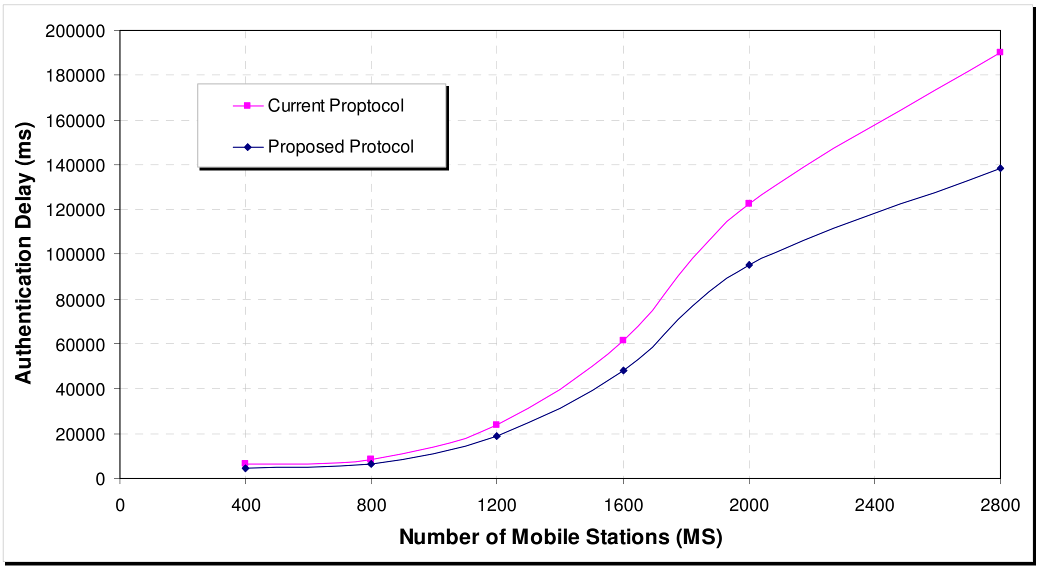


Figure 4.11: The relationship between authentication delay and the number of MS (simulation results)

4.4 Summary

In this chapter, the UMTS authentication and key agreement protocol and the signalling traffic that is generated by registration, call origination and termination have been studied and analysed. The bandwidth that is used between MS and VLR, and between database registers has been is analysed. The analysis based on a simple mobility model: the fluid flow mobility model. The signalling traffic overhead and the delay incurred in accessing the service are the key traffic performance parameters for a mobile network. The proposed authentication protocol has improved the performance by reducing the authentication times. The results have shown that at different speed of movement and call origination rate, the performance of the proposed protocol outperforms the current protocol (UMTS- AKA) in terms of the number of signalling messages and authentication delay.

Compared with the current protocol the storage space overhead of the SN was reduced. Moreover, the number of transmissions between the HN and VN for roaming authentication was reduced. Experimental results have shown that the proposed protocol can reduce the authentication traffic overhead for the network operators, the authentication latency from an end user's point of view, and energy consumption of a mobile terminal. Generally, the proposed protocol has the best performance.

Chapter 5 Conclusions and Future Work

This chapter presents the main conclusions of this thesis and suggestions for areas in the fields of mobile network security that need further research and improvement.

5.1 Conclusions

A literature review covered research related to mobile network security. The following issues were reviewed: authentication, privacy, integrity, identity/location anonymous and non- repudiation services in UMTS for 3G. Extensive investigation into the existing authentication mechanisms for mobile communications was performed as a step to improve mobile authentication.

In this thesis, by integrating the proposed public key cryptography with the hash-chaining technique, the security of the 3G protocols in network access is improved to provide key refreshment periodically, strong key management and a new non-repudiation service in a simple and elegant way. In addition, this mechanism has provided a new feature to let the encryption switches turn on before the authentication process commences and protect the subscriber's true identity. The bi-unilateral and mutual authentication among MS, VLR/SGSN in the serving network and HLR/AuC in the home network has been adopted in the proposed scheme and result in a more secure protocol than the other available authentication protocols.

The protection of mobile network communications by the authentication security mechanism incurs overheads on the transmission process. These overheads affect the mobile network performance in delay and bandwidth allocation efficiency. The authentication and key agreement protocol for home users and roaming users in mobile networks were studied and analysed. A new authentication protocol has been suggested to fulfil the security requirements of the third generation mobile systems and improve performance by reducing the communication times, and by creating fewer authentication messages and data sizes during the process of authentication. The proposed protocol significantly reduces the communication overhead between the home network and the visited network especially for roaming authentication. To avoid the complicated synchronization found in UMTS, the proposed protocol does not use SEQ, the management of a hash chain in the proposed protocol is simple and elegant compared to that of SEQ. The proposed protocol demonstrates better performance in terms of latency and storage space, compared to the 3G network approach of the home network transporting an authentication vector to the visited network. This proposed protocol is also secure against network attacks, such as the replay attack, guessing attack, substitution attack, impersonating attack, and redirection attack.

In chapters 3 and 4, the security of the proposed protocol is analysed. Traffic analysis indicates that the signalling traffic performance, load transaction messages and bandwidth are minimised during registration, origination and termination call phases. This study was performed under various cell densities and traversal velocities ranging from pedestrian to vehicular speeds. In comparing with the current authentication protocol, the percentage of improvement of proposed authentication protocol in terms of the total signalling and load transaction between entities ranges from 18% to 0.78%. A simulation was made by varying velocity and call origination rate to determine their impact on an estimate of total control traffic. Its results have proved that the proposed protocol has a level of improvement ranging from around 18% to 0.78%, compared to the UMTS AKA mechanism. The analytical and simulation results have proved that the average authentication delay and the average bandwidth for the proposed protocol are minimised when compared with the current protocol.

Key management related to MS, SN and HN generates a challenge for wireless roaming authentication. For instance, in 3G/UMTS the Authentication and (AKA) protocol, the initial and intra-network roaming authentication relies on HLR/AuC in the home networks to generate the authentication vectors. This feature means that a visited network needs to frequently communicate with a home network to bring back the authentication vectors. The proposed protocol solved this problem. It eliminates the sequence number (SEQ) which is used in 3G/UMTS roaming authentication. The proposed protocol employs a key authentication as a temporary key, which is generated during the execution of the registration process and caches in MS and visited network for the subsequent authentication process (intra-network).

5.2 Future work

There are many research issues relevant to the verification protocols for mobile networks which still need further research. New problems are likely to appear with the rapid growth in the technology of mobile networks. This thesis may be a possible starting point for further work and research in different areas. Based on the research performed in this thesis, three areas of future research into mobile network security are recommended:

1. Due of the nature of battery-powered mobile devices, energy consumption is an important issue for mobile networks. Therefore, investigation of power consumption and how the battery life is affected by use of the proposed protocol is desirable.

2. Testing by analytical models and software simulation has proved that the proposed protocol is efficient and robust. However, real network validations are still needed. The results of real experiments would corroborate the effectiveness and robustness of the proposed protocol.

In conclusion, this study contributes significantly to mobile network security. This study provides an important research reference for understanding the relationship between authentication and security. It provides a framework for the development of future mobile network security and the art of cryptographic science.

References

3GPP, 1999a. Security Threats and Requirements: *3rd Generation Partnership Project (3GPP), Technical Specification Group*, 3G TS 21.133, [online]. 2.0.0 (1999-04), Available at: www.3gpp.org/ftp/tsg_sa/TSG_SA/TSGS_03/Docs/PDF/SP-99145.pdf [accessed 22 December 2010].

3GPP, 1999b. 3G Security, Security Architecture: *3rd Generation Partnership Project (3GPP), Technical Specification Group*, 3G TS 33.102, [online]. 2.0.0 (1999-04), Available at: www.3gpp.org/ftp/tsg_sa/TSG_SA/TSGS_03/Docs/PDF/SP-99144.pdf [accessed 22 October 2010].

3GPP, 1999c. 3G Security; Security Principles and Objectives: *3rd Generation Partnership Project (3GPP), Technical Specification Group Services and System Aspects*; 3G TS 33.120, [online]. 3.0.0, (1999-05), Available at: www.3gpp.org/FTP/tsg_sa/WG3_Security/_Specs/33120-300.pdf [accessed 22 October 2010].

3GPP, 2001d. 3G Security; Security Threats and Requirements: *3rd Generation Partnership Project (3GPP), Technical Specification Group Services and System Aspects*, 3GPP TS 21.133, [Online]. 4.1.0 (2001-12) (Release 4), Available at: www.arib.or.jp/IMT-2000/V660Jun08/2_T63/ARIB-STD-T63/Rel4/21/A21133-410.pdf [Accessed 25 September 2010].

3GPP, 2001e. Network architecture: *3rd Generation Partnership Project; Technical Specification Group Services and Systems Aspects*; 3G TS 23.002, [Online]. 3.5.0

(2002-01) (Release 1999), Available at: www.arib.or.jp/IMT-2000/V310Sep02/S3g/R99/23/23002-350.pdf [Accessed 25 September 2010].

3GPP (2004b), " GPRS Tunnelling Protocol (GTP) across the Gn and Gp interface" *3^rd^ Generation Partnership Project (3GPP), Technical Specification Group* Core Network; General Packet Radio Service (GPRS), 3GPP TS 29.060, [Online]. 3.19.0 (2004-03) Release 1999, Available at: www.arib.or.jp/IMT-2000/V700Sep08/5_Appendix/R99/29/29060-3j0.pdf [Accessed 12 November 2010].

3GPP, 2004c. General Packet Radio Service (GPRS); Service Description: *3rd Generation Partnership Project (3GPP), Technical Specification Group*, Service (GPRS), 3G TS 23.060, [Online]. 6.0.0(2004-09), Available at: www.3gpp.org/ftp/tsg_sa/tsg_sa/TSGS_24/Docs/PDF/SP-040313.pdf [Accessed 12 January 2011].

3GPP, 2007a. 3G Security, Specification of the MILENAGE Algorithm Set: An Example Algorithm Set for the 3GPP Authentication and Key Generation Functions f1, f1*, f2, f3, f4, f5 and f5*, document 1: General : *3rd Generation Partnership Project (3GPP), Technical Specification Group Services and System Aspects*, 3GPP TS 35.205 [Online]. 7.0.0 (2007-06), Release 7, Available at: www.arib.or.jp/IMT-2000/V700Sep08/5_Appendix/Rel7/35/35205-700.pdf [Accessed 20 January 2011].

3GPP, 2007b. 3G Security, Specification of the MILENAGE Algorithm Set: An Example Algorithm Set for the 3GPP Authentication and Key Generation Functions f1, f1*, f2, f3, f4, f5 and f5*; Document 2: Algorithm Specification, *3rd Generation Partnership Project (3GPP), Technical Specification Group Services and System Aspects*, 3GPP TS 35.206, [Online]. 6.0.0 (2007-06) (Release 7), Available at:

http://www.arib.or.jp/IMT-2000/V700Sep08/5_Appendix/Rel7/35/35206-700.pdf [Accessed 12 February 2010].

3GPP 2007c. 3G Security, Specification of the MILENAGE Algorithm Set: An Example Algorithm Set for the 3GPP Authentication and Key Generation Functions f1, f1*, f2, f3, f4, f5 and f5*, Document 3: Implementors Test Data, 3rd *Generation Partnership Project (3GPP), Technical Specification Group Services and System Aspects*, 3GPP TS 35.206, [Online]. V7.0.0 (2007-06) (Release 7), Available at: http://www.arib.or.jp/IMT-2000/V700Sep08/5_Appendix/Rel7/35/35207-700.pdf [Accessed 12 February 2010].

3GPP TS 29.061. 2005. 3GPP: Technical specification group core network; interworking between the Public Land Mobile Network (PLMN) supporting packet based services and Packet Data Networks (PDN). Technical Report Release 6, 3GPP, Mar. 2005.

Akyildiz, I., Mcnair, J., Joseph, S., Uzunalioglu, H., and Wang, W., 1999. Mobility management in next-generation wireless system. Proceeding of the IEEE, 87(8), pp(s): 1347-1384.

Al-Fayoumi, M., Nashwan, S., Yousef, S. and Alzoubaidi, A., 2007. A New Hybrid Approach of Symmetric/Asymmetric Authentication Protocol for Future Mobile Networks. In: Third IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, WiMob, pp.29.

Al-Muhtadi, J., Mickunas, D., and Campbell, R., 2002. A Lightweight Reconfigurable Security Mechanism for 3G/4G Mobile. IEEE Wireless Communications, 9(2), pp.60-65.

Al-Tawil, K., krami, A., and Yousef, H., 1998. A New Authentication Protocol for GSM Networks. In: Local Computer Networks, 198 LCN'98, Proceedings of 23rd

Anderson, R., Manifavas, C. and Southerland, C., 1996. NetCard - A Practical Electronic Cash System. In: Proc. International Workshop on Security Protocols, Cambridge, April 10-12, 1996, pp. 49-57, UK.

Belier, M., Chang, L., and Yacobi, Y., 1993, Privacy and authentication on a portable communications system. *IEEE Journal on Selected Areas in Communications,* 11(6) :821 829. Aug. 1993.

Boman, K., Horn, G., Howard, P., and Niemi, V., 2002. UMTS Security. Electronics Communication Engineering Journal, 14(5), pp.191-204.

Boyd, C. and Park, D. G., 1998. Public Key Protocols for Wireless Communications. In: Proceedings of ICISC'98, Korea Institute of Information Security and Cryptology, pp. 47-57.

Bruce, S., 1996. Applied Cryptography. 2nd John Wiley and Sons, Inc.

Brutch, T., and Brutch, P., 1998. Mutual authentication, confidentiality and key Management (MACKMAN) system for mobile computing and wireless communication. In: Proceedings of the 14th Annual Computer Security Applications Conference, pp.308-317.

Burnett, S. and Pause, S., 2002. RSA Security's Official Guide to Cryptography. McGrawn, Hill, 2002.

Caelli, W., Dawson, E., and Rea, S., 1999. PKI, elliptic curve cryptography and digital signatures. Computer & Security, 18(1), pp.47-66.

Cheng, S., Shieh, S., Yang, W., Lee, F., and Luo, J., 2005. Designing Authentication Protocols for Third Generation Mobile Communication Systems. Journal of Information Science and Engineering, 21, pp.361-378.

Choudhury, H., Roychoudhury, B., and Saikia, D.K., 2010. End-to-End User Identity Confidentiality for UMTS Networks. 3rd IEEE International Conference on Computer Science and Information Technology (ICCSIT), pp. 46-50.

Constantinos, F., Sotirios, I., and Iakovos, S., 2001. Towards the Introduction of the Asymmetric Cryptography in GSM, GPRS, and UMTS Networks. Computers and Communications.

Dell'Uommo, S., and Scarrone, E., 2001. The mobility management and authentication authorization mechanisms in mobile networks beyond 3G. In: Proceedings of the 12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, pp.44-49.

Diffie, W., and Hellman, E., 1976. New Directions in Cryptography. IEEE Transactions on Information Theory, Volume IT22, November 1976, pp.644-654.

ElGamal, T., 1985. A public-key cryptosystem and a signature scheme based on discrete logarithms. IEEE Transactions on Information Theory, 31(4), pp.469-472.

ETSI, 1993. Recommendation GSM 03.20: Security related network functions: Technical report. European Telecommunications Standards Institute, ETSI.

Fang, Y. and Chlamtac, I., 1999. Teletraffic analysis and mobility modeling for PCS networks. IEEE Trans. Communication, 47, pp. 1062–1072, July 1999.

Fitzek, F., Schulte, G. and Reisslein, M,. 2002. System Architecture for Billing of Multi–Player Games in a Wireless Environment using GSM/UMTS and WLAN Services. proceeding by ACM 1–58113–493–2, NetGames, Braunschweig, Germany.

Ford, W., 1994. Computer Communications Security: Principles, Standard Protocols and Techniques. Prentice Hall PTR.

Gennaro, R., and Rohatgi, P., 1997. How to Sign Digital Streams. Advances in Cryptography - Crypto'97. pp.180-197.

Gódor, G., and Imr,e S., 2006. Novel Authentication Algorithm – Public Key Based Cryptography in Mobile Phone Systems. IJCSNS International Journal of Computer Science and Network Security, 6(2), February.

Gollmann, D., 1999. Computer Security. 1st Edition, John Wiley & Sons, Inc

Harn, L., and Hsin, W.-J., 2003. On the security of wireless network access with enhancements. In Proceedings of the p003 ACM workshop on Wireless security, pp.88 -95, San Diego, CA, USA.

Harn, L., and Lin, H., 2001. A Non-Repudiation Metering Scheme. IEEE Communications Letters, 5(12).

Harn, L., and Lin, H.Y., 1995. Modification to enhance the security of the GSM protocol. In: Proceedings of the 5th National Conference on Information security, Taipei, Taiwan, pp. 41-20.

Hassan M., Razzaq M., and Shahzad A. 2010. Comprehensive analysis of UMTS Authentication and Key Agreement. (IJCNS) International Journal of Computer and Network Security, 137 Vol.2, No. 2.

Holma, H., and Toskala, A., 2000. WCDMA for UMTS. proceeding by John Wiley & Sons, NewYork.

Horn, G., Martin, K., and Mitchell, C., 2002. Authentication Protocols for Mobile Network Environment Value-Added Services. IEEE Transactions on Vehicular Technology, 51(2), pp.383-392.

Huang, C. M., Li, J. W., 2005. Authentication and Key Agreement Protocol for UMTS with Low Bandwidth Consumption. In: Proceedings of the 19th International Conference of Advanced Information Networking and Applications (AINA'05), 1, Issue: 28-30, pp.392-397.

Iftikhar, M., Landfeldt, B., and Caglar, M., 2007. Traffic Engineering and QoS Control between Wireless DiffServ Domains Using PQ and LLQ. In: proceeding by ACM 978-1-59593-809-1, MobiWac'07, Chania, Crete Island, Greece.

ISO, 1996. Information Technology - Security Techniques - Key Management, Part 1: Framework: International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC), Geneva. ISO/IEC 11770-1: 1996. [Online]. Available at: *http://webstore.ansi.org/ansidocstore/subscriptions.* [Accessed 03 November 2010].

ITU, International Telecommunication Union, 1991. Data Communication Networks: Open Systems Interconnections (OSI): Security, Structure and Applications, Security Architecture for Open Systems Interconnection for CCITT Applications. , *International Telecommunication Union*, 1991. Geneva, [online]. Available at: *fag.grm.hia.no/IKT7000/litteratur/paper/x800.pdf.* [Accessed 03 November 2010].

ITU, International Telecommunication Union, 1997. Evaluation of Security Mechanisms for IMT-2000", *Recommendation: 1223, ITU Radiocommunication Sector*

Security, [online]. Available at: http://www.itu.int/rec/R-REC-M.1223-0-199702-I/en. [Accessed 17 October 2010].

Jun, J., and Chen, H., 2005. A novel mutual authentication and key agreement protocol based on NTRU cryptography for wireless communications. Journal of Zhejiang University Science, 6(5), pp.399-404.

Kaaranen, H., Ahtiainen, A., Laitinen, L., Naghian, S., and Niemi,V., 2005. UMTS Networks: Architecture, Mobility and Services. 2nd Edition, John Wiley & Sons, Inc.

Kaliski, B.,2003. TWIRL amid RSA key size. Technical report, RSA Laboratories. May 2003.

Koien, G., 2004, An Introduction to Access Security in UMTS. IEEE Transactions on Wireless Communications, 11(1), pp.8-18.

Lacy, J., Mitchell, D., and Schell, W., 1993. Cryptolib: Cryptography in software. In *Proceedings of the UNIX Security Symposium IV,* pages 1 17, 1993.

Lamport, L., 1981. Password authentication with insecure communication", Communication of ACM, 24(11), pp.770-772.

Lee, C., Hwang, M., and Yang, W., 2003. Extension of Authentication Protocol for GSM. IEE Proceeding Communication, 150(2), pp.91-95.

Lee, C.H., Hwang M. S., and Yang, W.P., 1999. Enhanced privacy and authentication for the global system for mobile communications. Wireless Network, 5, pp.231-243.

Liang, W., and Wang, W., 2004. An Analytical Study on the Impact of Authentication in Wireless Local Area Network. , Computer Communications and Networks, ICCCN 2004, Proceedings of 13th International Conference, pp.361-366.

Lin, C., and Shieh, S.,2000. Chain authentication in mobile communication systems. Journal of Telecommunication Systems, vol. 13, pp.213-240.

Lin, H. -Y., 1999. Security and Authentication in PCS. Computer & electrical Engineering, 25(4), pp.255-248.

Liu, Z., 2004. Security in 3G Wireless Network", [online]. Available at: http://plaza.ufl.edu/zipliu/ [Accessed 12 October 2010]

Lo, C.C., and Chen, Y.J., 1999a. A secure communication architecture for GSM networks. In: Proceedings of IEEE Pacific Rim Conference on Communications, computers and signal processing, pp.221-224.

Lo, C.C., and Chen, Y.J., 1999b. Secure communication mechanisms for GSM networks. IEEE Transactions on Consumer Electronics, 45(4), pp.1074-1080.

Looi, M., 2001. Enhanced authentication services for internet systems using mobile networks. In: IEEE Global Telecommunications Conference, 6, pp.3468-3472.

Menezes, A., Oorschot, P., and Vanstone, S., 1997. Handbook of Applied Cryptography, CRC Press.

Mohan, S., 1996. Privacy and authentication protocols for PCS. IEEE Personal Communication, pp.34-38.

Mohan, S., and Jain, R., 1994. Two User Location Strategies for Personal Communication Services. IEEE Personal Communications, 1(1), pp.42-50.

Molva, R., Samfat, D., and Tsudik, G., 1994. Authentication of Mobile Users. IEEE Network, 8(2), pp.26-34.

National Institute of Standards and Technology (NIST), 2001. Advanced encryption standard. Technical Report FIPS 197, NIST, US Department Commerce, Nov. 2001.

National Institute of Standards and Technology (NIST), 2003. Special publication 800-57: Recommendation for key management, part 1: General guideline. Technical report, National Institute of Standards and Technology. Jan. 2003.

Naveed M., Minhas A. and Jehanzeb Ahmad J. 2010. A Novel Security Algorithm for Universal Mobile Telecommunication System, International Journal of Multimedia and Ubiquitous Engineering, Vol. 5, No. 1

Niemi, V., and Nyberg, K., 2003. UMTS Security. John Wiley & Sons.

Passerini, K., Patten, K., Bartolacci, R., and Fjermestad, J., 2007. Reflections and Trends in the Expansion of Cellular Wireless Services in the U.S. and China. In: proceeding by communications of the ACM. 50(10).

Putz, S., and Schmitz, R, 2000. Secure Interoperation between 2G and 3G Mobile Radio Networks. In: 3G Mobile Communication Technologies, 2000, First International Conference on (IEE Conference Publication No. 471), London, UK, pp.28-32.

Putz, S., Schmitz, R., Tonsing, F., 1998. Authentication Schemes for Third Generation. Mobile Radio Systems. In: Personal, Indoor and Mobile Radio Communication., The 9th IEEE International Symposium on Personal, 1, pp.126-130.

Rabin, M., 1979. Digitalized signatures and public-key functions as intractable as factorization. MIT/LCS/TR-212, MIT Laboratory for Computer Science, Cambridge.

Rahnema, M., 1993. Overview of the GSM System and Protocol Architecture. IEEE Communication Magazine, 31(4), pp.92-100.

Richardson, K., 2000. UMTS overview. Elec. Communication Engineering, 12(3), pp.93–100.

Rivest, R.L., Shamir, A., and Adleman, L., 1978, A method for obtaining digital signatures and public key cryptosystems. Communications of the ACM, 21, pp.120-126.

Safavi-Naini, R., Susilo, W., and Taban, G., 2001. Towards securing 3G mobile phones. In: Proceedings Ninth IEEE International Conference on Networks, 2001, pp.222-227.

Salkintzis, A., 2004. Interworking Techniques and architectures for WLAN/3G Integration Toward 4G Mobile Data Networks. IEEE Wireless Communications, 2004, 11(3), pp.50-61.

Simmons, G., 1992. Contemporary Cryptology: The Science of Information Integrity. Piscataway, NJ: IEEE Press.

Stach, J.-S., Park, E.-K., and Su, Z., 1998. An Enhanced Authentication Protocol for Personal Communication Systems. Application-Specific Software Engineering Technology, 1998 ASSET-98, Proceeding, 1998 IEEE Workshop on, pp.128-132.

Stalling, W., 2003. Cryptography and Network Security: Principles and Practice. 3rd Edition, Prentice Hall.

Sutton, R., 2002. Secure Communications: Applications and Management . John Willey and Sons.

Telecommunication Standardization Sector of ITU.1993.ITU-T Recommendation E.212: Identification plan for land mobile stations. Technical report, ITU-T, 1993

Thomas, R., Gilbert, H., and Mazziotto, G., 1988. Influence of the Mobile station on the Performance of the Radio Mobile Cellular Network. Proc. 3rd Nordic Seminar, Paper 9.4, Copenhagen, Denmark, Sept. 1988.

Vanneste, G., Degraeve, J., and Franco, B., 1997. Authentication for UMTS: introduction and demonstration. In: 2nd International Distributed Conference on Network Interoperability, pp.715-721.

Wadekar, R., and Fagoonee, L., 2006. Beyond Third Generation (B3G) Mobile Communication: Challenges, Broadband Access and Europe. proceeding by ACM, 1-59593-519-3, Mobility 06, Bangkok, Thailand.

Xenakis, C., and Merakos, L., 2004. Security in third Generation Mobile Networks. www.elsevier.com/locate/comcom, Computer Communications 27, pp. 638–650.

Yan-zhi, H., Da-wei, M., and Xiao-fei, L., 2009. "An Improved Authentication Protocol with Less Delay for UMTS Mobile Networks," icnds, vol. 2, pp.111-115, 2009 International Conference on Networking and Digital Society, 2009.

Yeh, C.-K., and Lee, W.-B., 2007. A Dual-Purpose Signature For Authentication On UMTS. Journal of the Chinese Institute of Engineers, 30(2), pp. 343-347.

Yuan, S., Elizabeth, B., Gao, X., and James, K., 2007. Real-time traffic support in heterogeneous mobile networks. proceeding by Springer Science + Business Media, Wireless Network13, pp.431–445.

Zhang, M., and Fang, Y., 2005. Security Analysis and Enhancements of 3GPP Authentication and Key Agreement Protocol. IEEE Transactions on wireless communications, 4(2), pp.734-742.

Appendices

Appendix A

A.1 Useful code of the RSA simulator

The following function is considered as constructor, Execute automatically when begin the running the main program. Constructor use to generate the private key, public key

```
public void Rsa_Mainn (int bitlen)

{

        //Bitlen: indicate the length of key

        g1 = System.nanoTime ();

        //to calculates the execution time, which it store the beginning time of
        execution.

        SecureRandom r = new SecureRandom ();

        //This class provides a cryptographically strong pseudo-random number
         generator (PRNG).

        BigInteger p = new BigInteger (bitlen/2, 1, r);
        BigInteger q = new BigInteger (bitlen/2, 1, r);

        // P and q two large prime number typed as big integers, which the length of
         each half key length.

        modulus = p.multiply (q);

        //Returns a BigInteger whose value is p multiple q

        BigInteger m = (p.subtract (BigInteger.ONE))
                     .multiply (q.subtract (BigInteger.ONE));

      //That means multiple p minus one by q minus one and stores the values in m

      if (q.hashCode () == p.hashCode())

      {
        // to compare if p and q are equal

        m = (p.subtract (BigInteger.ONE)).multiply (q);
```

```
    }




else
    {

     m = (p.subtract (BigInteger.ONE))
                 .multiply (q.subtract (BigInteger.ONE));
    }
    PublicKey = new BigInteger ("65537");

     //Default value for the public key

    //Find the public key

    while (m.gcd (publicKey).compareTo(new BigInteger("1"))!= 0)

    PublicKey = publicKey.add (new BigInteger ("2"));

    //Generate the private Key

    BigInteger pp = p.subtract (BigInteger.ONE);
    BigInteger qq = q.subtract (BigInteger.ONE);
    BigInteger gcd = pp.gcd (qq);
    BigInteger thi = m.divide (gcd);
    PrivateKey = publicKey.modInverse (thi);

     //modInverse Returns a BigInteger whose value is public key inverse mod thi

     g2=System.nanoTime ();

    //to calculates the execution time, which it store the end time of execution.

    gall+= g2-g1;

    //gall indicate to a mount of execution time

}

    //The following code examine the input message to be want to encrypt, when it
    large than a key, split it into piece equal the Key

    public void test ()
    {
     String temp="";

    //The following loop execute until to finish the split a message

    while ((rsa.tot_msg.length ())>=this.modulus.toString()
    .length ()/2)

    {
```

```
    temp = rsa.tot_msg.substring (0,(int)(modulus.
    toString ().length ()-2)/2);

//this does the splitting all message into a message equal the Key

   blocks.add (temp);

   //That is a queue contain each of a message after splitting

   rsa.tot_msg=rsa.tot_msg.substring ((modulus.toString ()
   .length ()-2)/2);
   }

   blocks.add (rsa.tot_msg);
   }

   // this is signing function

   public BigInteger encrypt (BigInteger message)
   {

   return message.modPow (publicKey, modulus);

   // Returns a BigInteger whose value is message power public key that result mod
    modulus

  }

// this is verification function

    public BigInteger decrypt(BigInteger message)
    {
    return message.modPow (privateKey, modulus);
    }

    //The following function coded the message before singing it

    public BigInteger normal_encrypt(String msg)
    {

    rsa.RsaMessage1(msg);

    // above function coded the message

    return encrypt (new BigInteger(rsa.getMessage()));

    //Encrypt the message after coded and return cipher message

    }

    ///The following function decrypts a message after that decodes it and returns a
     plain message

    public String normal_decrypt (BigInteger enc_msg)
```

```
{
return rsa.decode (decrypt(enc_msg).toString());
}
```

A.2 Example of the RSA digital signature

The message =

983867660444149806437584738755315751294592321342

p =

1196233420820853325534279135252246854516526443900 94

9354670455786776766812637750637941672979803838287 16

5306412855505044921692529799547962057625894508226 58

49

q =

130084969516727140392554988115736245733936567700312

146484755861762146599277187358265678996931724884813

946437358424913799495814113611285017645977281707335

59

n =

155611988082370934188436997303428961272835420317212

461192915518386064157517732487438386497323404867846

428347471230712137358131316144508227543044153296779

021451000428899143445882601426295621821156307457931

494926458037572778437018789250775086223254522987033958461653508379527322746368415393910836871868183926591

m =

155611988082370934188436997303428961272835420317212461192915518386064157517732487438386497323404867846428347471230712137358131316144508227543044153296778996480169269017896151284311262199528702597386248890786731277893528796108964693008569101593763312115680910753789110837698156239659058785788496015194930927184

e=

65537

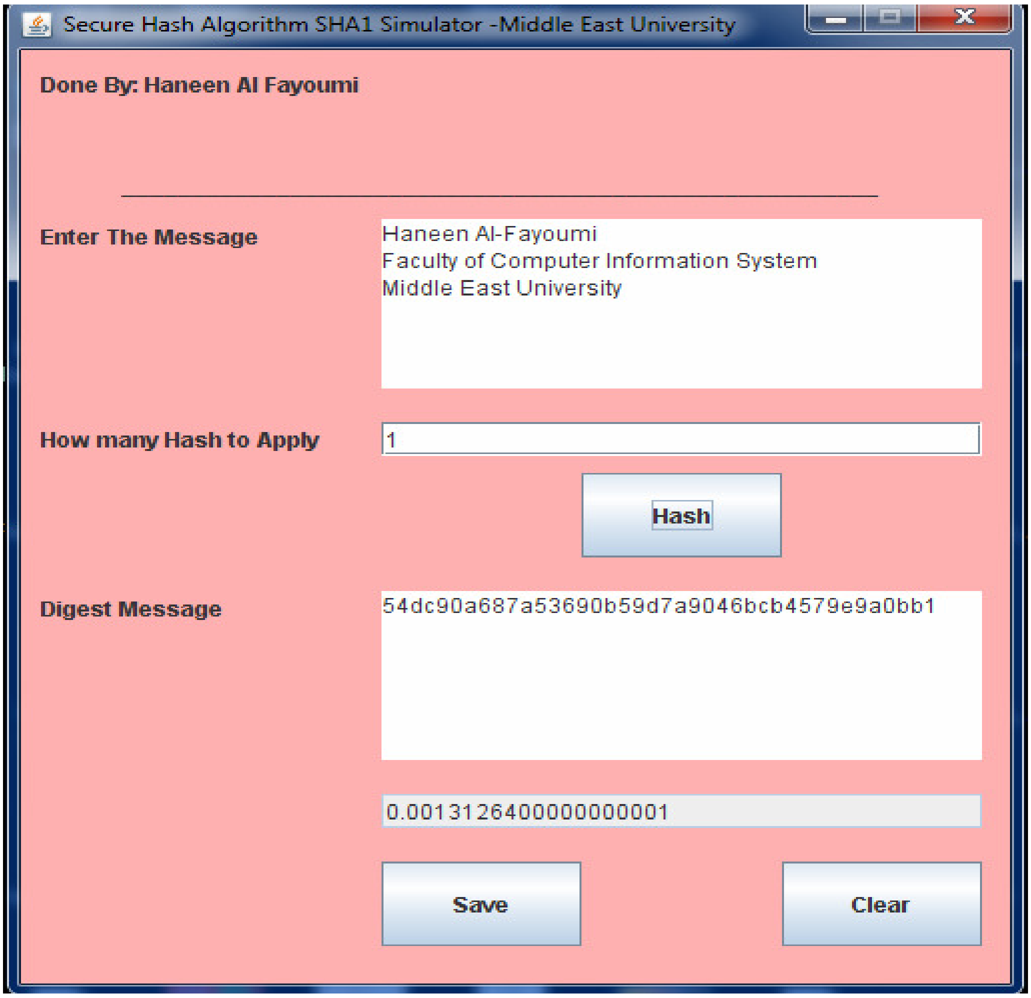
d =

8477846154814920129389723345162169556504857839733388661203669047979646221478917960827608659127166038202426333918691085744498647547095024588285077884952765719096760227480632439089878414993320301874782209508430718948362674008683010152846422284213984556953151530378320952225479757292867683822235149949700182729

Signature=

1375018321287642168683600297844059523864293254816801983483857652498477274709330649327574618723210156338473427512162914444468059925842535973709689917465509019117505698502295273554540516634155774677301570170707439575206257672820632989062706107374734867180606848763818903586214594693915946322172366060175365561 47

A.3 Output SHA1 algorithm



Secure Hash Algorithm SHA1 Simulator -Middle East University

Done By: Haneen Al Fayoumi

Enter The Message

Haneen Al-Fayoumi
Faculty of Computer Information System
Middle East University

How many Hash to Apply 1

Digest Message

54dc90a687a53690b59d7a9046bcb4579e9a0bb1

0.0013126400000000001

Appendix B

B.1 Network simulator

The simulator provides a framework for building a network model, specifying data input and analysing output data. It is a discrete event simulator targeted at networking research. It is a widely used simulation tool for simulating inter-network topologies to test and evaluate various networking protocols.

The simulation tool supports a trace file that used to trace and analyse the packets for both wireless and wired networks.

B.2 Cell structure

```
package UMTS_elements;
public class Base_Station
{
             private int Base_Station_id;
            private int capacity;
            private int coverage_area;//used in future for Rncs
             public  int rx_axis;
             public int ry_axis;
 public int r;

 public int r_green;
  ////////////////////
            int n_imax=0;
             int ki;
            int cri=0;
            int csi=0;

  public void set_ki(int val)
            {
              ki=val;
            }

  public void set_n_imax(int val)
            {
                        n_imax=val;
              }
  public void set_cri(int val)
             {
             cri=val;
             }
  public void set_csi(int val)
  {
      csi=val;
  }
  public int get_ki()
  {
      return ki;
  }
  public int get_n_i_max()
  {
      return n_imax;
  }
  public int get_cri()
  {
      return cri;
  }
  public int get_csi()
  {
      return csi;
  }
/////////////////////////////////////////
 //begin to set value to variable

public void set_Base_Station_id(int baseid)
  {
       Base_Station_id=baseid;
  }

public void set_capacity(int c)
 {
   capacity=c;
 }

public void set_coverage_area(int c)
 {
 coverage_area=c*2;
 }

public void set_rx_axis(int x)
  {
      rx_axis=x;
  }

public void set_ry_axis(int y)
  {
      ry_axis=y;
  }
 /////////////////////////////
//end to set value to variable
///////////////////////////
  // begin to get value of variable

public int get_Base_Station_id()
```

```
  {
       return Base_Station_id;
  }

public int get_capacity()
{
  return capacity;
}

public int get_coverage_area()
{
   return coverage_area;
}

public int get_rx_axis()
{
  return rx_axis;
}

public int get_ry_axis()
{
   return ry_axis;
}


/////////////////////////////
//end to get value of variable
//////////////////////////////



}
```

B.3 RNC structure

```
public class RNC
{

   int w;//size of window must be determined
   int csrnc=0; //to check packets overflow
   int total_packet;
   boolean ack=true;
   public RNC(int w,int total_packet)
   {
   this.w=w;
   this.total_packet=total_packet;
   }

When the rnc sent  a packet execute the following function to guarantee
Non-over flow may be occur

public void send_ok()
{
   csrnc+=1;
if (csrnc%w == 0)
  ack=false;
   else
  ack=true;

}
create number of BS depend on passed number for function

   set of each base station your properties such that
   *base id
   *and set for each base station x axis and y axis
   *for radius
   *must pass to function thier parameter :
   *1)number of base station required.
   *2)distribution for vertical of base Station
   *3)x ,y overlap
   *and finally path to store result


public void Create_base_station(int no_bs,int no_ver,int raduis,
        int xoverlap,int yoverlap,String path)
{
  Distribution_convegare_bs m=new Distribution_convegare_bs();
        try {
            m.write_distributed(no_bs,no_ver,raduis,xoverlap,yoverlap,path);
        } catch (Exception ex) {
            ex.printStackTrace();
        }

b=new Base_Station[no_bs];

for(int i=0;i<no_bs;i++)
{
b[i]=new Base_Station();
b[i].set_Base_Station_id(i);
b[i].set_rx_axis(m.rx_range[i]);
b[i].set_ry_axis(m.ry_range[i]);
b[i].set_coverage_area(m.get_rcoverage_area_bs());
b[i].r=(int)((raduis)*0.50);
```

to find coverage area by each BS

```
public void distribution(int no_vertical,int xoverlap,int yoverlap)
{

    int rx,ry;


     xoverlap=(int)((rcoverage_area_bs*2)*xoverlap)/100;
     yoverlap=(int)((rcoverage_area_bs*2)*yoverlap)/100;
    rx_range=new int[no_sections];
    ry_range=new int[no_sections];
    int m=0;
```

```
   /////
   rx=rcoverage_area_bs;
   ry=rcoverage_area_bs;
  for(int i=0; i < no_sections; i++)
   {

     if(i==0)
     {
      rx_range[i]=rx;
      ry_range[i]=ry;
     }
     else
     {
       if(i%no_vertical==0 && i!=0)
        {
           rx+=(rcoverage_area_bs*2);
         ry=(rcoverage_area_bs);
         rx_range[i]=rx;
         ry_range[i]=ry;
         }


       else

       {
        ry+=(rcoverage_area_bs*2);
        rx_range[i]=rx;
        ry_range[i]=ry;
       }



       }



  }
```

B.4 Mobility

After period of time change the state of mobile and choice random number; that is indicate the new state of mobile.

```
  cases=rand.nextInt(6);
switch(cases)
{
case 0:ue.cases=0;break;
case 1:ue.cases=1;break;
//case 2:ue.cases=2;break;
case 3:ue.initial_state=true;ue.cases=3; ue.subsequent=0; break;
case 4:ue.finished=true;ue.cases=4;break;
case 5:ue.finished=true;ue.cases=5;break;
}


     case 0 :temp=" UTRAN hand-over "
     case 1 :temp=" UE request service "
     case 3 :temp=" SGSN-VLR hand-over "
     case 4 :temp=" Out of service "
     case 5 :temp=" Turn-off "
```

B.5 Distribution of area

```
package UMTS_elements;
public class Distribution_convegare_bs
{//class
int shifting_area[]=new int[100];
int shift=0;
public Distribution_convegare_bs()
{
   for(int i=0;i<100;i++)
    this.shifting_area[i]=-1;
}
             private int all_coverage_area;
             private int rcoverage_area_bs;//(radius/2   )
            public int rx_range[];
             public int ry_range[];
            private int no_sections;

            public void set_no_sections(int s)
            {
             no_sections=s;
            }
            public int get_no_sections()
            {
              return no_sections;
            }

            public void set_rcoverage_area_bs(int s)
            {
            rcoverage_area_bs=s;
            }
            public int get_rcoverage_area_bs()
            {
             return rcoverage_area_bs;
            }
            public int get_all_coverage_area()
            {
            int n=get_no_sections()*(rcoverage_area_bs*2);
            return n;
            }
```

```
//to find coverage area by each BS
public void distribution(int no_vertical,int xoverlap,int yoverlap)
{
       int rx,ry;
              xoverlap=(int)((rcoverage_area_bs*2)*xoverlap)/100;
              yoverlap=(int)((rcoverage_area_bs*2)*yoverlap)/100;
              rx_range=new int[no_sections];
              ry_range=new int[no_sections];
              int m=0;
    /////
              rx=rcoverage_area_bs;
             ry=rcoverage_area_bs;
             for(int i=0; i < no_sections; i++)
              {
                   if(i==0)
                          {
                           rx_range[i]=rx;
                          ry_range[i]=ry;
                           }
              else
                         {
                           if(i%no_vertical==0 && i!=0)
                           {
                                        rx+=(rcoverage_area_bs*2);
                                       ry=(rcoverage_area_bs);
                                        rx_range[i]=rx;
                                        ry_range[i]=ry;
                           }
             else
                     {
                          ry+=(rcoverage_area_bs*2);
                          rx_range[i]=rx;
                         ry_range[i]=ry;
                           }
        }
    }

   //process for x overlap
    int counter=1;
             for(int k=no_vertical;k<no_sections;k++)
             {
                   if((k%no_vertical)==0 && k!=no_vertical)
                   counter++;
                   rx_range[k]=rx_range[k]-xoverlap*counter;
             }
///end process for x overlap
//process for y overlap
             counter=0;
              for(int k=1;k<no_sections;k++)
              {
                 if(((k%no_vertical)==0))
              {
                     counter=0;
```

```
                    ry_range[k]=ry_range[k];
                }
                else
                        {
                            counter++;
                           ry_range[k]=ry_range[k]-yoverlap*counter;
                        }
             }
this.shifting(no_sections,no_vertical);
            for(int q=0;shifting_area[q]!=-1;q++)
            rx_range[shifting_area[q]]=rx_range[shifting_area[q]]+(rcoverage_area_bs);

            }
//used for write of previous function to file and uses
//for bind_BS class
public void shifting(int bs_no,int no_vertical)
{
              int temp=bs_no/no_vertical;
                for(int i=1;i<no_vertical;i++)
                  if(i%2!=0)
                  {
                    for(int j=i;j<temp+i;j++)
                         if(j==i)
                        {
                         shifting_area[shift]=i;
                        shift++;
                         }
                        else
                         {
                               shifting_area[shift]=shifting_area[shift-1]+no_vertical;
                              shift++;
                        }
                  }
 }
 public void write_distributed(int no_bs,int no_ver,int raduis,int xoverlap,int yoverlap,String
path)
{
             rcoverage_area_bs=raduis;
             no_sections=no_bs;
             distribution(no_ver,xoverlap,yoverlap);
             CreateFile m1=new CreateFile();
             m1.creater(path);
            int j=get_no_sections();
            for(int i=0;i<get_no_sections();i++)
            {
             m1.writetofile(i);
            m1.writetofile((int)(rx_range[i])/5);
            m1.writetofile((int)(ry_range[i])/5);
            m1.writetofile((int)get_rcoverage_area_bs()/5);
            m1.writetofile();
}
m1.closefile();
}
 }
```