

Master's thesis

**GSM-Security: a Survey and Evaluation
of the Current Situation**

by

Paul Yousef

LITH-ISY-EX-3559-2004

2004-03-05

Master's thesis

GSM-Security: a Survey and Evaluation of the Current Situation

by **Paul Yousef**

LITH-ISY-EX-3559-2004


2004-03-05

Supervisor and examiner:

Viiveke Fåk

ISY, Linköping Institute of Technology

Linköping, 5th Mars 2004

 LINKÖPINGS UNIVERSITET	Avdelning, Institution Division, Department Institutionen för systemteknik 581 83 LINKÖPING	Datum Date 2004-03-05						
Språk Language Svenska/Swedish X Engelska/English	Rapporttyp Report category Licentiatavhandling X Examensarbete C-uppsats D-uppsats Övrig rapport _____	<table border="1"> <tr> <td colspan="2">ISBN</td> </tr> <tr> <td colspan="2">ISRN LITH-ISY-EX-3559-2004</td> </tr> <tr> <td>Serietitel och serienummer Title of series, numbering</td> <td>ISSN _____</td> </tr> </table>	ISBN		ISRN LITH-ISY-EX-3559-2004		Serietitel och serienummer Title of series, numbering	ISSN _____
ISBN								
ISRN LITH-ISY-EX-3559-2004								
Serietitel och serienummer Title of series, numbering	ISSN _____							
URL för elektronisk version http://www.ep.liu.se/exjobb/isy/2004/3559/								
<table border="1"> <tr> <td>Titel Title</td> <td>GSM-säkerhet: En Översikt och evaluering av nuvarande situation GSM-Security: A Survey and Evaluation of the Current Situation</td> </tr> <tr> <td>Författare Author</td> <td>Paul Yousef</td> </tr> </table>			Titel Title	GSM-säkerhet: En Översikt och evaluering av nuvarande situation GSM-Security: A Survey and Evaluation of the Current Situation	Författare Author	Paul Yousef		
Titel Title	GSM-säkerhet: En Översikt och evaluering av nuvarande situation GSM-Security: A Survey and Evaluation of the Current Situation							
Författare Author	Paul Yousef							
<table border="1"> <tr> <td> Sammanfattning Abstract <p>The Global System for Mobile Communications (GSM) is the most widely used cellular technology in the world. For GSM, like many other widely used systems, security is crucial. The aspects of security that this report covers are mainly anonymity, authentication and confidentiality.</p> <p>It appears that many of the very valuable aspects of GSM can be attacked. Anonymity, authentication mechanism and confidentiality can be attacked and compromised if the attacker possesses the right equipment. In order to break the protection, the attacker needs to utilise active attacks, i.e. base station functionality is needed. However, if the attacker is able to decrypt GSM traffic, i.e. break A5/1 and A5/2, passive attacks are sufficient.</p> <p>The cryptographic algorithms used to encrypt GSM traffic and data are cryptographically weak and can be cryptanalysed in real-time, resulting in compromised confidentiality. Cryptanalysis of A5 is however nontrivial and often requires huge amounts of computation power, mainly for the one time pre-computation step.</p> <p>GSM does not provide sufficient security for users with very valuable information to communicate. These users are advised to use an additional layer of security on top of GSM.</p> </td> </tr> </table>			Sammanfattning Abstract <p>The Global System for Mobile Communications (GSM) is the most widely used cellular technology in the world. For GSM, like many other widely used systems, security is crucial. The aspects of security that this report covers are mainly anonymity, authentication and confidentiality.</p> <p>It appears that many of the very valuable aspects of GSM can be attacked. Anonymity, authentication mechanism and confidentiality can be attacked and compromised if the attacker possesses the right equipment. In order to break the protection, the attacker needs to utilise active attacks, i.e. base station functionality is needed. However, if the attacker is able to decrypt GSM traffic, i.e. break A5/1 and A5/2, passive attacks are sufficient.</p> <p>The cryptographic algorithms used to encrypt GSM traffic and data are cryptographically weak and can be cryptanalysed in real-time, resulting in compromised confidentiality. Cryptanalysis of A5 is however nontrivial and often requires huge amounts of computation power, mainly for the one time pre-computation step.</p> <p>GSM does not provide sufficient security for users with very valuable information to communicate. These users are advised to use an additional layer of security on top of GSM.</p>					
Sammanfattning Abstract <p>The Global System for Mobile Communications (GSM) is the most widely used cellular technology in the world. For GSM, like many other widely used systems, security is crucial. The aspects of security that this report covers are mainly anonymity, authentication and confidentiality.</p> <p>It appears that many of the very valuable aspects of GSM can be attacked. Anonymity, authentication mechanism and confidentiality can be attacked and compromised if the attacker possesses the right equipment. In order to break the protection, the attacker needs to utilise active attacks, i.e. base station functionality is needed. However, if the attacker is able to decrypt GSM traffic, i.e. break A5/1 and A5/2, passive attacks are sufficient.</p> <p>The cryptographic algorithms used to encrypt GSM traffic and data are cryptographically weak and can be cryptanalysed in real-time, resulting in compromised confidentiality. Cryptanalysis of A5 is however nontrivial and often requires huge amounts of computation power, mainly for the one time pre-computation step.</p> <p>GSM does not provide sufficient security for users with very valuable information to communicate. These users are advised to use an additional layer of security on top of GSM.</p>								
<table border="1"> <tr> <td> Nyckelord Keyword </td> <td>GSM, security, attacks, cryptanalysis, protocols, flaws, resources</td> </tr> </table>			Nyckelord Keyword	GSM, security, attacks, cryptanalysis, protocols, flaws, resources				
Nyckelord Keyword	GSM, security, attacks, cryptanalysis, protocols, flaws, resources							

Abstract

The Global System for Mobile Communications (GSM) is the most widely used cellular technology in the world. Approximately 800 million people around the world are using GSM for different purposes, but mostly for voice communication and SMS. For GSM, like many other widely used systems, security is crucial. The security involves mechanisms used to protect the different shareholders, like subscribers and service providers. The aspects of security that this report covers are mainly anonymity, authentication and confidentiality.

The important aspects of the system that need protection are described, along with the implementation of mechanisms used for the protection. It appears that many of the very valuable aspects of GSM can be attacked.

The anonymity of a GSM user is compromised resulting in the attacker being able to observe the time, rate, length, sources or destinations of e.g. calls. Even tracking a subscriber's movements becomes possible. However, a passive attack is not sufficient to perform these attacks. The attacker needs to mount an active attack using equipment offering base station functionality.

Authentication is a crucial aspect of a wireless communication system due to the nature of the medium used, i.e. the radio link that is available to every one and not only the legitimate entities. Even the authentication mechanisms are attacked. It is possible to clone a subscription either by having physical access to the smart card or over the air interface. Cloning a subscription over the air requires base station functionality.

The most obvious threat against communication systems is eavesdropping on conversations. The privacy of GSM conversations is protected using some version of the A5 algorithm. There are several impressive cryptanalytical attacks against these algorithms, that break the encryption and make it possible to eavesdrop in real-time. Most of these algorithms require, however, extensive computation power and unrealistic quantities of known plaintext, which make it difficult to use them in practice. Difficulties using cryptanalytical attacks to break the confidentiality of GSM calls does not mean that conversations are well protected. Loopholes in the protocols used in GSM make it possible for an outsider, with access to sufficient equipment, to eavesdrop on conversations in real-time.

In the presence of these threats and vulnerabilities it is justified to wonder whether GSM provides sufficient security for users with very valuable information to communicate. These users may be military organisations, senior management personnel in large companies etc. GSM's current security model does not provide sufficient protection for these entities. An additional layer of security should be added to the current security model.

Acknowledgements

Many people have supported me, in different ways, during the work with the thesis. I'd like to thank my supervisor and examiner Viiveke Fåk for the help during my work. My family has, as always, offered me their unconditional support, thank you! Then of course, I want to thank my wonderful girlfriend Carolina for her continuous support and encouragement.

Linköping, Mars 2004

Paul Yousef
(paul.yousef@comhem.se)

Table of Contents

Abstract	vii
Acknowledgement	ix
Definitions	xiii
Abbreviations	xv
1 Introduction	1
1.1 Background.....	1
1.2 Purpose	1
1.3 Reading Instructions	2
Part I	5
2 Security Requirements of Wireless Networks	7
2.1 Requirements for End-User Privacy	7
2.1.1 Protection of Call-Setup Information	7
2.1.2 Protection of Speech	7
2.1.3 Privacy of User-Location.....	7
2.1.4 Privacy of Calling Patterns	7
2.1.5 Privacy of User-ID.....	7
2.2 Integrity Protection of Data	8
2.3 Requirements for Preventing Theft of Service or Equipment	8
2.3.1 Cloning and Clone Resistant Design	8
2.3.2 Equipment Identifiers	9
3 Security Attacks and the Use of Cryptography for Protection	11
3.1 Security Attacks.....	11
3.2 Cryptographic Protection Methods.....	13
3.2.1 Secret Key Cryptography.....	13
3.2.2 Public Key Cryptography	16
3.2.3 Hash Algorithms/Functions	18
3.3 Attacking the Cryptographic Protection	19
Part II	21
4 Layers, Channels and Signalling Principles in the GSM System	23
4.1 The Layers of GSM	23
4.2 The Physical Layer – Layer 1	24
4.2.1 Frequency-Division Multiple Access and Time-Division Multiple Access	24
4.2.2 The Radio Channel	25
4.2.3 The Frequencies.....	25
4.2.4 Transmission on the Radio Channels.....	25
4.2.5 Logical Channels	27
4.2.6 Frame Structures	28
4.2.7 Examples of How a Mobile Station Behaves	30
4.2.8 From analog to digital.....	34
4.2.9 Frequency Hopping	35
4.3 The Data Link Layer – Layer 2	35
4.4 The Network Layer – Layer 3	36
4.4.1 Sublayers of Layer 3	36
4.4.2 Structure of a Layer 3 Message	38
4.4.3 A Layer 3 Signalling Trace.....	39
5 GSM Architecture and Security	41
5.1 An Overview of the GSM Network.....	41

5.1.1	The Mobile Station (MS).....	41
5.1.2	The Base Transceiver Station (BTS)	42
5.1.3	The Base Station Controller (BSC)	42
5.1.4	Mobile Services Switching Center (MSC)	42
5.1.5	Home Location Register (HLR)	42
5.1.6	Authentication Center (AuC).....	42
5.1.7	Visitor Location Register (VLR)	43
5.1.8	Equipment Identity Register (EIR)	43
5.2	The Security Implementation – Protecting Valuable Assets	43
5.2.1	Anonymity	44
5.2.2	Authentication.....	44
5.2.3	Confidentiality	46
5.2.4	Preventing Theft of Service or Equipment	50
PART III	51
6	Attacks on GSM.....	53
6.1	Capturing One or Several Mobile Stations	53
6.2	Attacks on the Anonymity of GSM Users	54
6.2.1	Passive Monitoring	54
6.2.2	Active Monitoring	55
6.3	Attacks on the Authentication Algorithm.....	59
6.3.1	Cloning with Physical Access to the SIM Module	60
6.3.2	Cloning over the Air	62
6.4	Attacks on the Confidentiality of GSM	63
6.4.1	Brute-Force Attacks.....	63
6.4.2	Cryptanalytical Attacks against GSM.....	64
6.4.3	Attacks Using Loopholes in the Protocols.....	67
6.5	Attacks on the Equipment Protection Mechanism.....	79
6.6	Denial of Service (DoS) Attacks	80
6.6.1	Denial of Service – Physical Intervention	80
6.6.2	Denial of Service – Logical Intervention.....	80
7	Evaluation of the Suitability of GSM for Special Users	83
7.1	Security Threats	83
7.1.1	Unauthorised Access to Data.....	83
7.1.2	Unauthorised Manipulation of Sensitive Data	84
7.1.3	Denial of Service Attacks	85
7.1.4	Unauthorised Access to Services.....	85
7.1.5	Threats Associated with Attacks on the Terminal (ME) and SIM.....	85
7.2	Risk Assessment	86
7.3	Results of the Threat Analysis	87
PART IV	89
8	Discussion and Conclusions	91
8.1	Cryptanalytical attacks.....	91
8.2	Attacks based on protocol weaknesses	92
8.2.1	Anonymity	92
8.2.2	Authentication.....	93
8.2.3	Confidentiality	95
8.3	Conclusion.....	96
9	Future Work.....	99
References	101

Definitions

A–Interface

On the physical level the A-interface consists of one or more pulse code modulation (PCM) links between the MSC and the BSC. Each one has a transmission capacity of 2 Mbps.

Abis-Interface

The Abis-interface¹ is the interface between the BTS and the BSC. It is a pulse code modulation (PCM) 30 interface. The transmission rate is 2 Mbps which is partitioned into 32 channels of 64 Kbps each. The compression techniques that GSM utilises packs up 8 GSM traffic channels into a single 64 Kbps channel.

Authentication

The provision of assurance of the claimed identity of an entity.

Confidentiality

The property that information is not made available or disclosed to unauthorised individuals, entities or processes.

Data integrity

The property that data has not been altered in an unauthorised manner.

Data origin authentication

The corroboration that the source of data received is as claimed.

Handover

The GSM user movements can produce the need to change the channel or cell, specially when the quality of the communication is decreasing. This procedure of changing the resources is called handover.

Octet

In many places in the ETSI specification of GSM, a message is described as a succession of octets. An **octet** is generally a succession of 8 bits.

One Time Pad

An unbreakable cipher, where the key used is truly random and as long as the message to be encrypted.

Pseudo Random Number generator

An algorithmic technique for random number generation. These algorithms are deterministic but still generate a sequence of numbers that passes many reasonable tests of randomness.

¹ Abis is a French term meaning ‘the second A-interface’.

Roaming

Roaming is defined as the ability for a cellular customer to automatically make and receive voice calls, send and receive data, or access other services when travelling outside the geographical coverage area of the home network, by means of using a visited network.

Signalling

The exchange of information, in telephony between involved parties in the network, that sets up, controls, and terminates each telephone call or other services offered by the network

Um

The Um interface, also known as the air interface or radio link.

Abbreviations

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
BSC	Base Station Controller
BSS	Base Station System
BTS	Base Transceiver Station
CEIR	Central Equipment Identity Register
DES	Data Encryption Standard
EIR	Equipment Identification Register
ETSI	European Telecommunications Standards Institute
FDMA	Frequency Division Multiple Access
GSM	Global System for Mobile Communications
GSM MoU	The GSM Memorandum of Understanding, an agreement signed between all the major European operators to work together to promote GSM. The precursor of the GSM Association
HLR	Home Location Register
IE	(signalling) Information Element
IMEI	International Mobile station Equipment Identity
IMSI	International Mobile Subscriber Identity
ISDN	Integrated Services Digital Network, a data network usually provided by public carriers (BT, AT&T etc) providing digital communication at 56k (US & Japan) or 64k (rest of world)
LAC	Location Area Code
LAI	Location Area Identifier
LFSR	Linear Feedback Shift Register
ME	Mobile Equipment
MNC	Mobile Network Code
MS	Mobile Station
MSC	Mobile-services Switching Centre, Mobile Switching Centre

MSCM	Mobile Station Class Mark
MTC	Mobile-terminated call
PLMN	Public Lands Mobile Network
PSTN	Public Switched Telephone Network
SIM	Subscriber Identity Module
SMS	Short Message Service
SS7	Signalling System 7
TDMA	Time Division Multiple Access
TMSI	Temporary Mobile Subscriber Identity

This report is a Masters's Thesis at the Computer Science Program at Linköping Institute of Technology – Linköpings Tekniska Högskola. It has been conducted at the Division of Information Theory at the Department of Electrical Engineering – Institutionen för Systemteknik (ISY).

This chapter gives a short introduction to the thesis. It describes the background to the thesis, presents the questions to be answered, and describes the organisation of the thesis.

1.1 Background

Security plays a more important part in wireless communication systems than in systems that use wired communication. This is mainly because of the ubiquitous nature of the wireless medium that makes it more susceptible to security attacks than wired communications. In the wireless medium, anyone can listen to whatever is being sent over the network. Also, the presence of communication does not uniquely identify the originator (as it does in the case of a pair of coaxial cables or optical fibers). To make things worse, any tapping or eavesdropping cannot even be detected in a medium as ubiquitous as the wireless medium. Thus security plays a vital role for the successful operation of a mobile communication system. GSM is a 2G system that is used daily by hundreds of millions of people. Can it withstand today's high-tech-equipped hackers?

1.2 Purpose

This document aims to give an introduction to the security mechanisms used to protect GSM², and present the attacks possible to mount on the system, mainly on the anonymity, authentication and confidentiality aspects of security, along with the resources needed. This will include:

- Describing how the very complex GSM system works. Components used to build the system are introduced and the techniques used to provide the functionality are described. This will answer the question: How does GSM work?
- Introducing the requirements on the security of a wireless communication system along with the mechanisms used by GSM to meet these requirements. This will answer the question: What are the valuable assets of GSM and how are these assets protected?

² GSM was formerly acronym for Groupe Spéciale Mobile (founded 1982). Now is acronym for Global System for Mobile Communication.

- Presenting attacks on GSM security, which include recent cryptanalytical attacks on the cryptographic algorithms protecting the confidentiality of GSM user traffic as well as other types of attacks, especially those making use of weaknesses in the GSM protocols, and examining the resources needed in order to mount these attacks successfully. This will answer the question: How can valuable aspects of GSM be attacked and what resources are needed in order to realise these attacks?
- Drawing conclusions about the suitability of GSM as a communication infrastructure for different user groups. Finally, this will answer the question: Is GSM suitable for providing communication services for users with very valuable information to protect?

1.3 Reading Instructions

The thesis is divided into four parts:

Part I

Part 1 introduces the reader to the security requirements of wireless networks, highlights the types of threats that face communicating parties in general and gives an introduction in cryptographic methods that can be used to protect against attacks. It contains the following chapters:

Chapter 2 highlights the important aspects of the system that the security mechanisms should protect, with emphasis on wireless systems.

Chapter 3 provides a general overview of various types of attacks that can be mounted against computer systems and networks along with cryptographic paradigms that are commonly used in practice to protect against these attacks.

Part II

In part 2 we start to look at GSM. This includes examining the technology used to make the system a digital wireless communication system by means of layers, channels, frequencies etc. Next we look at the architecture of the GSM network, by means of the components that build up the network. Further on the security mechanisms that are used to provide anonymity, authentication and confidentiality are introduced. Part 3 contains the following chapters:

Chapter 4 gives an introduction to the technology that makes the GSM system work, describing the architecture at several layers.

Chapter 5 gives an overview of the architecture of the GSM network, including description of the components that build up the system, and presents the mechanisms used in order to implement security.

Part III

Now that we know how GSM works, what the valuable aspects of GSM are and how these aspects are protected, we try to break the protection. Different attacks on anonymity, authentication and confidentiality are described and evaluated.

Further, a risk analysis is made to examine whether users with high requirements for security should trust GSM with their valuable information. Part 3 contains the following chapters:

Chapter 6 presents several attacks against the security implementation of GSM.

Chapter 7 examines whether GSM is suitable to be used by entities with higher security requirements than private persons, e.g. the military.

Part IV

In part 4 the attacks against GSM, presented above, are discussed and conclusions are drawn.

Chapter 8 contains a discussion of the subjects that the report has addressed along with conclusions.

Chapter 9 discusses future work in the subject.

Part I

Requirements for Security, Attacks and Cryptography Protection

- Chapter 2 Security Requirements of Wireless Networks
- Chapter 3 Security Attacks and the Use of Cryptography for Protection

GSM, like many other large systems with large numbers of users, contains many valuable assets that need protection against misuse and deliberate attacks. This chapter will highlight the valuable assets that, in general, exist in a wireless communication system, and that are crucial to protect for the best of the system's shareholders (subscribers and service providers) .

2.1 Requirements for End-User Privacy

A subscriber to a mobile communication system needs protection in the following areas:

2.1.1 Protection of Call-Setup Information

During the call-setup process, the mobile terminal will communicate important call-setup information to the network. Some of the information that could be sent is: calling party number, calling card number, service type requested, etc. This information must be protected and secured from eavesdroppers. [1]

2.1.2 Protection of Speech

All spoken communication and other communication services must be properly encrypted by the cryptographic system, so that it cannot be intercepted by any eavesdropper listening to the radio interface or other interfaces of the system. [1]

2.1.3 Privacy of User-Location

Any leakage of specific signalling information on the network may enable an eavesdropper to approximately locate the position of a subscriber, which will jeopardize the subscriber's privacy. Hence the subscriber must be protected from such attacks on his/her privacy of location. [1]

2.1.4 Privacy of Calling Patterns

Information related to traffic generated by a particular user and his/her calling patterns should not be made available to eavesdroppers. Typical information is: caller-id, frequency of calls to some particular number, etc. [1]

2.1.5 Privacy of User-ID

All mobile communication systems use some sort of user-ID to identify their subscribers. This subscriber identification information (or the user-ID) must be

protected from hackers. Transmission of this information in the clear either over the radio interface or over the network must be avoided as far as possible. [1]

2.2 Integrity Protection of Data

In addition to securing the data (system data or traffic data) against eavesdroppers, there must be a provision in the network and the terminal to detect or verify whether the data it receives has been altered or not. This property is called *Data Integrity*. System and user data that are considered to be sensitive must be protected by using this method. [1]

2.3 Requirements for Preventing Theft of Service or Equipment

Theft of service and equipment is a very serious problem in mobile personal communications. The network subsystem doesn't care whether a call has originated from a legitimate or from a stolen terminal (the mobile equipment/phone) as long as it bills the call to the correct account (the legitimate user cares, though!). There are two kinds of theft that could be possible here, namely the theft of personal equipment and theft of the services offered by the service provider. The cryptographic protection must be designed to make the reuse of stolen terminals as difficult as possible. Further, it should block theft of services made possible by techniques such as cloning. Note that e.g. cloning can be done both by the hackers using stolen equipment, as well as *legitimate users*. [1]

The following sections will present important requirements for preventing theft.

2.3.1 Cloning and Clone Resistant Design

Cloning is a serious problem in mobile communication systems. Cloning refers to the ability of an intruder to determine information about a personal terminal and clone, i.e. create a duplicate copy, of that personal terminal using the information collected. This kind of fraud can be easily accomplished by legitimate users of the network themselves, since they have all the information they need to clone their own personal terminal stored in the Subscriber Identity Module (SIM) in the terminal. In this way, multiple users can use one account by cloning personal equipment. It could even be done by a stranger who wants to use services on the expense of legitimate users or sell the cloned devices. This is where equipment cloning causes problems. The cryptographic protection for the mobile network must incorporate some kind of clone-resistant design. The most obvious requirement for this design is the security of personal equipment information. This security must be provided for the radio-interface, the network databases, and the network interconnections such that personal equipment information is secure from impostors.

Since the terminal can be used by anyone, it is necessary to identify the correct person for billing purposes, i.e. the *user* must be *identified* to the *network*. This

may take the form of a *smart-card* or a *plug-in* that plugs into a terminal and is unique to each user. The process by which the network identifies the user is called the *authentication* process, where information about the identity of the user is transmitted to the network and verified using some cryptographic technique.

2.3.2 Equipment Identifiers

In systems where the account information is separated (both logically and physically) from the terminal, e g GSM, stolen personal equipment and its resale could be an attractive and lucrative business. To avoid this, all personal equipment must have unique identification information that reduces the potential of stolen equipment to be re-used. This may take the form of tamper-resistant identifiers permanently plugged into the terminals. [1]

This chapter provides a general overview of various types of attacks that can be mounted against computer systems and networks, and cryptographic methods that are commonly used in practice to protect against these attacks. Cryptographic concepts that are relevant for wireless communications, in particular GSM, are emphasised where necessary.

3.1 Security Attacks

Attacks on the security of computer systems and networks are best characterised by viewing the function of the computer system or network to be providing information. The attacker is an entity trying to disturb the normal flow of information in the system (Figure 1).

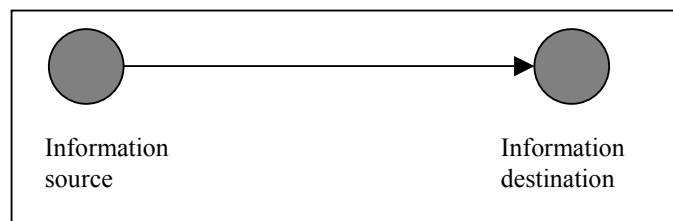


Figure 1. Normal flow of information [25]

Attacks can be categorised as follows:

- **Interruption:** An asset of a system is either destroyed or it becomes unavailable or unusable (Figure 2). This is an attack on *availability*. The attacker may e.g. cut a communication line or use jamming to interrupt wireless communications. [24]

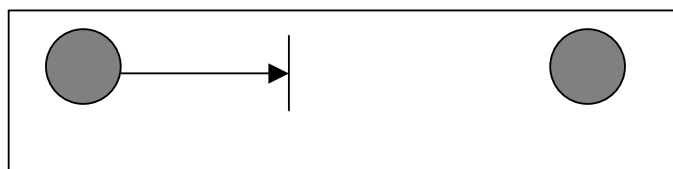


Figure 2. Interruption [25]

- **Interception:** An unauthorised party gains access to an asset (Figure 3). This is an attack on *confidentiality*. The unauthorised party could be a person or a computer process. Examples include wiretapping/eavesdropping to capture data in a network. [24]

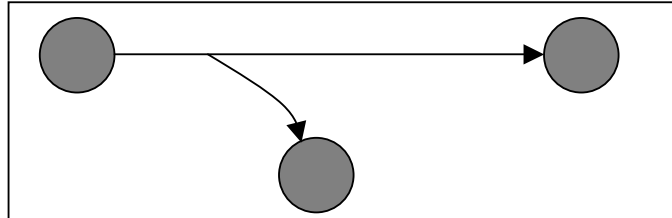


Figure 3. Interception [25]

- **Modification:** An unauthorised party not only gains access to but also tampers with an asset (Figure 4). This is an attack on *integrity*. Examples include changing values in a data file, altering a program so that it performs differently, and modifying the content of messages being transmitted between communicating entities. [24]

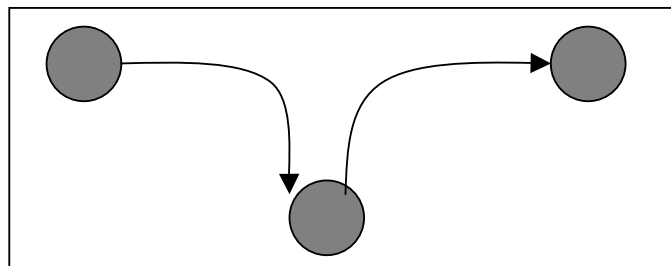


Figure 4. Modification [25]

- **Fabrication:** An unauthorised party inserts counterfeit objects into the system, or claims to be some other party (Figure 5). This is an attack on *authenticity*. Examples include the insertion of spurious messages (e.g. signalling messages in the GSM) in a network and the addition of records to a file. [24]

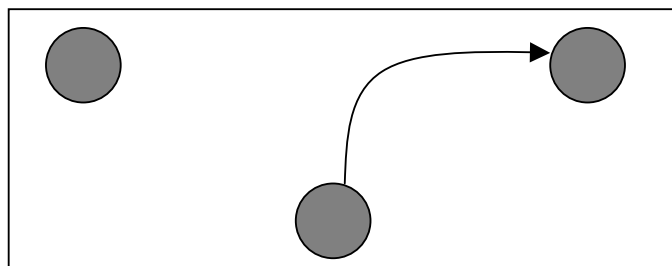


Figure 5. Fabrication [25]

3.2 Cryptographic Protection Methods

In traditional cryptography, a message in its original form is known as *plaintext* or cleartext. The encrypted information is known as *ciphertext* and the process of producing this ciphertext is known as *encryption* or *enciphering*. These two terms will be used interchangeably in this report and will refer to the same process. The reverse process of encryption is called *decryption* or *deciphering*. Cryptographic systems tend to involve an algorithm and a secret value. The secret value is known as the *key*. The reason for having a key in addition to an algorithm is that it is difficult to keep devising new algorithms that will allow reversible scrambling of information.

There are three types of cryptographic paradigms:

3.2.1 Secret Key Cryptography

Secret key cryptography involves the use of a *single key* that is shared by the communicating parties (Figure 6). This is the method used in GSM for providing confidentiality. Given a message (plaintext), encryption produces the ciphertext, which is of the same length as the plaintext. Decryption retrieves the plaintext, using the *same key* used for encryption. This kind of encryption is also called *conventional* or *symmetric cryptography*.

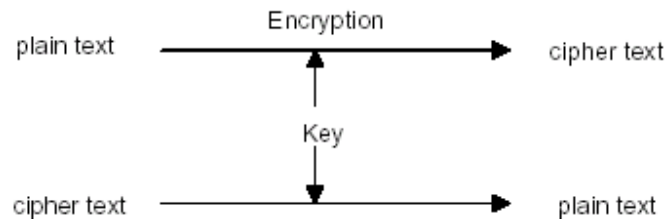


Figure 6 A Secret key cryptographic system [1]

Secret key systems also provide strong authentication functionality. This implies that someone can prove knowledge of a secret without revealing it, a functionality that is essential for wireless systems. [24]

Authentication can be implemented using a *Challenge-Response* mechanism (Figure 7). For example, suppose A and B wish to communicate with each other and they decide upon a key K_{AB} to verify each other's identity. Each of them picks a random number, which is known as a *challenge* and send it to each other. The value of the random number, say x , encrypted with the key K_{AB} , using a common algorithm, is known as the *Response* to the challenge x . [1]

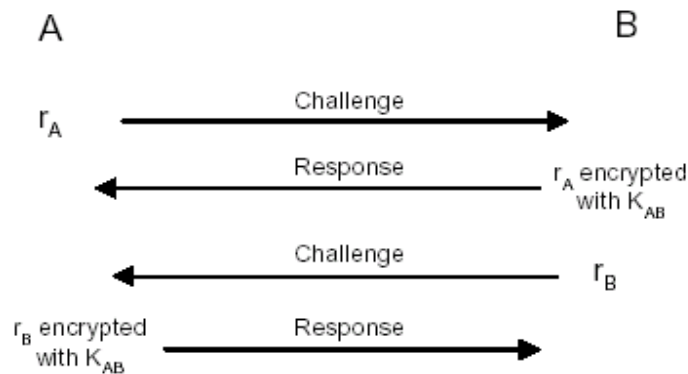


Figure 7 Challenge-Response mechanism in secret key systems [1]

Thus, if A and B complete this exchange, they have proved to each other that they know K_{AB} without revealing it to an impostor or an eavesdropper. Of course this is also accomplished if A, who sends the challenge to B, computes the correct response to the challenge and compares it to the response from B. If the responses are equal B has proved its identity. [1] This kind of Challenge-Response mechanism is used in GSM for authenticating a mobile user. One apparent flaw in these kind of systems is that an eavesdropper can form Challenge-Response pairs, since he/she can pose a challenge to either A or B and store the responses. To avoid this situation it is essential that the challenges be chosen from a large enough space, say 2^{128} values, so that there is no significant chance of using the same challenge twice. Another scenario is an attacker intercepting a challenge and its response and later challenging A with the captured challenge. This is called a *replay attack*. The attack can be avoided by attaching a timestamp to the challenge. A receiver of an replayed challenge can easily discover the attack by realising that the timestamp is outdated.

Further it should be noticed that the key K_{AB} can also represent an algorithm A_{AB} , that uses the random number x and produces an encrypted value. This algorithm is only known to A and B (for example the GSM A3/A8 algorithm, see Section 5.2.2, is one such algorithm). This means that the security of the system not only relies on the secrecy of the key, which should be the case, but also the algorithm. This has been the case in GSM, and is called *security by obscurity*, an approach to security that has been widely criticised due to the fact that it has been shown that secret algorithms tend to be cryptanalysed which jeopardises the security of the system.

Secret key cryptography can further be divided in two categories:

3.2.1.1 Block Ciphers

As the name suggests, block ciphers encrypt or decrypt data in blocks or groups of bits. The most popular block cipher historically, and a widely used one, has been Data Encryption Standard (DES). DES uses a 56-bit key and processes data

in 64-bit blocks, producing 64-bits of encrypted data for 64-bits of input, and vice-versa. Block algorithms are further characterised by their mode of operation, such as electronic code book (ECB), cipher block chaining (CBC), and cipher feedback (CFB). CBC and CFB are examples of modes of operation where the encryption of successive blocks is dependent on the output of one or more previous encryptions. These modes are desirable because they break up the one-to-one correspondence between ciphertext blocks and plaintext blocks (as in ECB mode). Block ciphers may even be implemented as a component of a stream cipher. [24]

3.2.1.2 Stream Ciphers

Stream ciphers operate on a bit-by-bit basis, producing a single encrypted bit for a single plaintext bit. Stream ciphers are commonly implemented as the exclusive-or (XOR) of the data stream with the keystream. The security of a stream cipher is determined by the properties of the keystream. A completely random keystream would effectively implement an unbreakable one-time pad encryption, and a deterministic keystream with a short period would provide very little security. [27]

Linear Feedback Shift Registers (LFSRs) are a key component of many stream ciphers. LFSRs are implemented as a shift register where the vacant bit created by the shifting is a function of the previous state. With the correct choice of feedback taps, LFSRs can function as pseudo-random number generators. The statistical properties of LFSRs make them useful for other applications such as pseudo-noise (PN) sequence generators in direct sequence spread spectrum communications, and for distance measurement in systems such as the Global Positioning System (GPS). LFSRs have the additional advantage of being easily implemented in hardware. [27]

The maximal length sequence (also called *period*) is equal to $2^n - 1$ where n is the degree of the shift register. An example of a maximal length LFSR is shown in Figure 8 below. This LFSR will generate the periodic sequence (also called *m-sequence*) consisting of the following states (1111, 0111, 1011, 0101, 1010, 1101, 0110, 0011, 1001, 0100, 0010, 0001, 1000, 1100, 1110). [27]

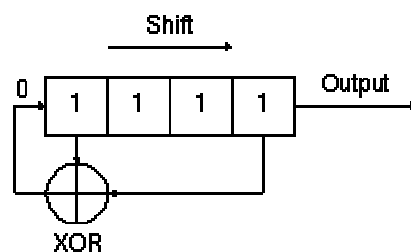


Figure 8. An example LFSR [27]

In order to form an m-sequence, the feedback taps of an LFSR must correspond to a primitive polynomial modulo 2 of degree n . A number of stream cipher designs consist of multiple LFSRs with various interconnections and clocking schemes. The GSM A5 algorithm, used to encrypt voice and signalling data in GSM, is a stream cipher based on three clock-controlled LFSRs. [27]

3.2.2 Public Key Cryptography

Public key cryptography is not used in the current GSM security model. It is still an important technology to present in this report due to the many proposals for increased security in GSM that make use of public key protocols.

In public key cryptography, the keys are not shared. Instead, each individual user has two keys: a *Private Key* (that is not revealed to anyone) and a *Public Key* (that is open to the public). This kind of cryptography is also commonly called *Asymmetric Cryptography* and was invented by Diffie and Hellman in 1975. In these systems, encryption is done using the public key and decryption is done using the private key (Figure 9).

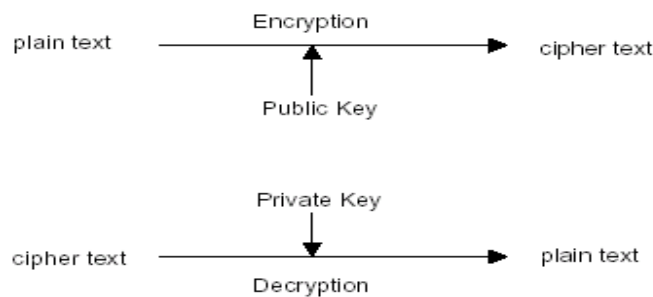


Figure 9. A public key cryptographic system [1]

An example of public key cryptography is described in the following paragraph.

Consider two people A and B wishing to communicate over an insecure channel (say, a wireless channel). Suppose that A's <public key, private key> pair is $\langle e_A; d_A \rangle$ and B's pair is $\langle e_B; d_B \rangle$. Moreover assume that the public keys are known to both A and B (and the public). Figure 10 explains the procedure to be followed by A and B for communication. It is clear that each person encrypts the data using the other person's public key, which can be decrypted by the other person using his/her own private key. [1]

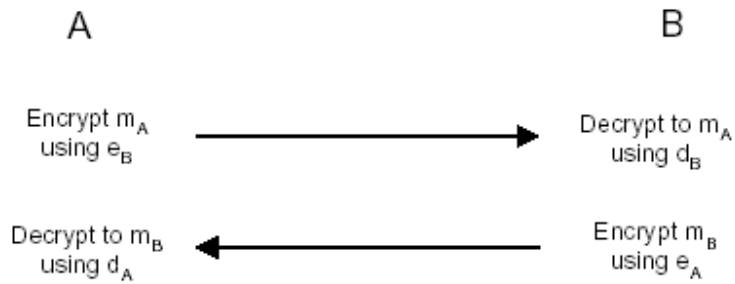


Figure 10. Information transfer in a public key cryptographic system [1]

This kind of encryption/decryption is not much different from secret key systems, but the biggest benefit of public key systems over secret key systems comes from the authentication mechanism. In the case of authentication in secret key systems, if A and B want to communicate with each other, they have to share a secret (key K_{AB} or algorithm A_{AB}) among themselves. If one wants to communicate with many entities he/she must remember many secret keys each corresponding to every entity he/she wishes to communicate. Public key cryptography avoids this problem by the use of public keys. In this case, the entities wishing to communicate with each other have to remember only their private key. To communicate with another entity, they have to look up the public key of the other entity (from a Directory Server) and use it to encrypt the messages to be communicated to this entity. For example, suppose A wants to verify (authenticate) B's identity. A chooses a random number r , encrypts it using B's public key e_B and sends the result to B. Now, B can prove his/her identity by decrypting the encrypted message (the Challenge) using his/her private key and sending the decrypted random number r (the Response) back to A. [1] (Figure 11)

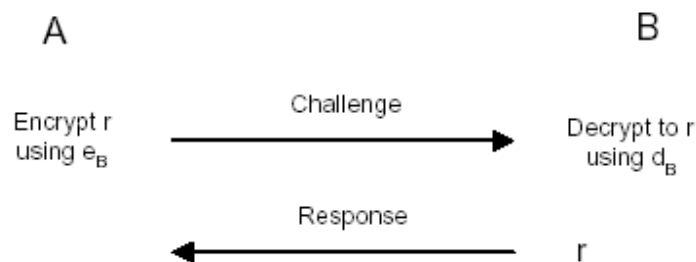


Figure 11. Authentication mechanism in a public key system [1]

Though public key systems provide a highly efficient authentication mechanism, they are orders of magnitude slower than secret key systems. In the case of communication networks, these public key systems require excessive computations and transfer of large numbers of bits along power/bandwidth-limited channels. Thus, these systems were not initially recommended for

wireless/mobile communications where bandwidth and power³ are at a premium. This is one of the main reasons that the 2nd Generation GSM systems are primarily secret key systems. However, since higher capacities have been introduced with the introduction of 3rd generation systems, public key systems will begin to play an important role in providing confidentiality and authentication mechanisms. [1]

Public key cryptography also facilitates digital signatures, whereby a person can sign plaintext using his/her private key and anyone can verify the person's identity by using the public key of that person. Further, others cannot forge the signature of the person since it involves the use of his/her private key. An illustration of digital signatures is presented in Figure 12. [1]

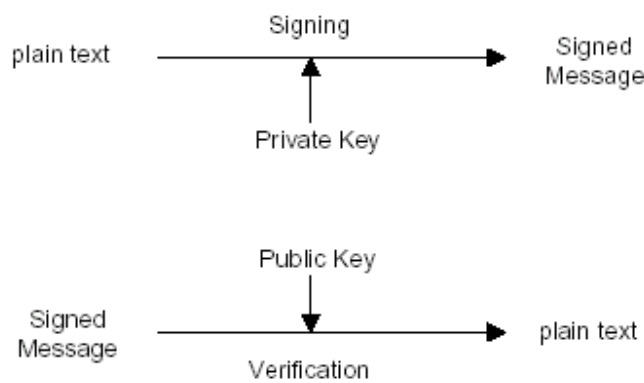


Figure 12. Digital signatures in a public key system [1]

3.2.3 Hash Algorithms/Functions

Hash Algorithms are also called *message-digests* or *one-way* transformations. The hash function h takes as input a message of arbitrary length and produces as output a message digest of fixed length. Certain properties should be satisfied:

1. Given a message m , the message digest $h(m)$ can be calculated very quickly.
2. Given a message digest y , it is computationally infeasible to find an m with $h(m) = y$ (in other words, h is a one-way, or pre-image resistant, function).
3. It is computationally infeasible to find messages m_1 and m_2 with $h(m_1) = h(m_2)$. This condition is requiring h to be strongly collision-free. [2]

A typical example of message-digesting is password authentication in personal computer systems. For security reasons, the system does not store the actual (unencrypted) password, but a hashed or digested value of it. When a password is supplied, the system computes the hashed or digested value of the supplied password and compares it the stored hash value. If the hash values match, then the supplied password is deemed correct. Hashing can also be used for other

³ And hence battery life of portable devices

functions such as message fingerprinting, digital signatures, message integrity checking etc. [1]

The algorithm A3/8 used in GSM for authentication and session key generation is another example of a hash function.

3.3 Attacking the Cryptographic Protection

The security of cryptographic algorithms is a difficult property to measure. As mentioned earlier, most algorithms employ keys, and hence the security of the algorithm is strongly related to how difficult it is for an attacker to determine the key⁴. The process of attempting to discover the plaintext or the key is known as *cryptanalysis*. The strategy used by the cryptanalyst depends on the nature of the encryption scheme and the information available. The most obvious approach to acquiring the key is to try every possible key and see which ones yield meaningful decryptions. Such attacks are called *brute force attacks* or *exhaustive key search*. [24] In a brute force attack the length of the key is directly related to how long it will take to search the entire keyspace (see Table 1).

Key Size (bits)	Number of Alternative Keys	Time required at 1 encryption/ μ s	Time required at 10^6 encryptions/ μ s
32	$2^{32} = 4.3 \times 10^9$	$2^{31} \mu\text{s} = 35.8$ minutes	2.15 milliseconds
56	$2^{56} = 7.2 \times 10^{16}$	$2^{55} \mu\text{s} = 1142$ years	10.01 hours
128	$2^{128} = 3.4 \times 10^{38}$	$2^{127} \mu\text{s} = 5.4 \times 10^{24}$ years	5.4×10^{18} years
26 characters (permutation)	$26! = 4 \times 10^{26}$	$2 \times 10^{26} \mu\text{s} = 6.4 \times 10^{12}$ years	6.4×10^6 years

Table 1 Average time required for exhaustive key search [24]

With the increasing amount of computing power available at lower and lower costs, today's cryptosystems must be able to withstand brute-force attacks that would have been unthinkable in the relatively recent past. However, long keys are not guaranteed to make an adversary's task difficult. The algorithm itself plays a critical role. Some algorithms might be able to be attacked by means other than brute force, and some algorithms just don't make very efficient use of their key's bits. Cryptanalysts often exploit the fact that traces of structure or pattern in the plaintext may survive encryption and be discernible in the ciphertext. This weakness can make it possible to discover the plaintext or even the key. [24] COMP128⁵ has this weakness, which makes it possible to find the secret key of a GSM subscriber.

⁴ This holds at least for systems employing public algorithms.

⁵ This is the algorithm many GSM operators use as A3/A8

Part II

GSM Layers, Architecture and Security Implementation

- Chapter 4 Layer, Channels and Signalling Principles in GSM
- Chapter 5 GSM Architecture and Security Implementation

Signalling is required to establish, maintain and terminate connections or communication links and to make sure that the provision of services is taking place, by the use of defined procedures. Therefore many of the attacks are focused on the signalling system in order to make it fail to work properly. Thus understanding how the signalling is done, how the signalling information is provided and how it looks is necessary for getting a clear and comprehensive view of the system's vulnerabilities and how to attack them.

The European Telecommunications Standards Institute (ETSI) is the official European organisation for standardisation of telecommunications. It is not vested with the powers of an authority but is a private organisation, formed in 1988 and assigned the task of creating standards for the European common market. Members are administrations, network operators, service providers, manufacturers and users, and all these categories now have direct influence on the standardisation work. ETSI is the organisation that has standardised GSM and will be referred to in many places in this report.

This chapter will introduce to the reader the technology that makes the GSM system work. It will describe the architecture of its complex signalling system by presenting the logical layers used in GSM, the functional entities in the signalling system, how signalling is done and the channels used for signalling and traffic. The goal is to get a feeling and a basic understanding of the protocols and functions necessary to establish, maintain and terminate mobile connections. This information will later be used in attacking the system.

4.1 The Layers of GSM

The Open Systems Interconnection (OSI) model (Figure 13) divides the tasks involved with moving information between networked computers into seven smaller, more manageable task groups. A task or group of tasks is then assigned to each of the seven OSI layers. Each layer is reasonably self-contained so that the tasks assigned to each layer can be implemented independently. This enables the solutions offered by one layer to be updated without adversely affecting the other layers. [26]

The signalling between all of the interfaces from a GSM mobile station to the MSC takes place in the lower three layers (i.e., layers 1 to 3 in Figure 13).

To illustrate the functionality, as used in GSM, we could say that Layer 1 is the freight train, switches, lights, and tracks. Layer 2 is the pallets, boxes, drums, and

carefully labelled envelopes in the train. Layer 3 is the valuable contents of all the containers and envelopes themselves.

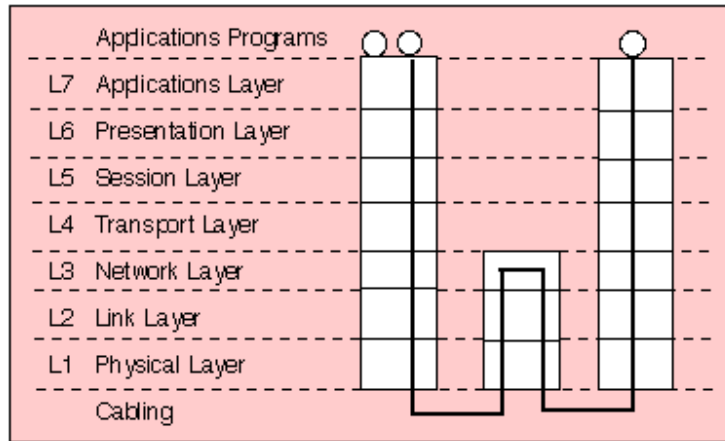


Figure 13 An overview of the GSM network [26]

In the following sections the functionality of these layers will be described.

4.2 The Physical Layer – Layer 1

All of the schemes and mechanisms used to make communications possible on the mobile radio channel with some measure of reliability between a mobile and its base station are called the *physical layer* or the *Layer 1* procedures. These mechanisms include modulation, power control, coding, timing, and other details that manage the establishment and maintenance of the channel.

The following sections will introduce several of these mechanisms:

4.2.1 Frequency-Division Multiple Access and Time-Division Multiple Access

GSM uses Time-division multiple access (TDMA) on top of Frequency-division multiple access (FDMA) in order to provide users with access to the radio resources in GSM. With FDMA, users are assigned a channel from a limited set of channels ordered in the frequency domain (Section 4.2.3). Usually, the initial assignments to channels are made from a common control channel, to which all radios tune for instructions when they first try to use the system. Since there is a limited number of frequency bands, TDMA is used on top of FDMA in order to further divide the use of each channel between several users. In TDMA, users share a physical channel where they are assigned time slots. All the users sharing the physical resource have their own assigned repeating time slot within a group of time slots called a frame. So in GSM, users are sorted onto a physical channel in accordance with simple FDMA techniques. Then the channel's use is divided up in time into frames, during which eight different users share the channel. A

GSM time slot is $577 \mu\text{s}$, and each user gets to use the channel for $577 \mu\text{s}$ every 4.615 ms ($577 \mu\text{s} \cdot 8 \text{ slots} = 4.615 \text{ ms}$). [42]

4.2.2 The Radio Channel

Cellular radio uses the word channel in many ways. It is a pair of radio frequencies, used by two entities to communicate with each other. There are two sources of trouble in the channel: noise and interference. Channel coding (see Section 4.2.8) is applied to the channel in order to minimise the influence of these destructive forces on the transmitted signal. [42]

4.2.3 The Frequencies

The frequencies used in GSM are defined in the FDMA part of the physical layer. GSM uses three different frequency bands, 900 MHz, 1800 MHz and 1900 MHz. The frequency bands used within each of the three ranges are similar and therefore only the frequency usage in the 900 MHz range will be described (Figure 14). In the 900 MHz GSM, two 25- MHz frequency bands are used. The mobile station transmits in the 890- to 915-MHz range, and the base station transmits in the 935- to 960-MHz. [42]

The end points within the physical layer are the mobile station and the BTS. The MS -to-BTS direction is referred to as the *uplink* (ul) and the BTS-to-MS direction as the *downlink* (dl). [42]

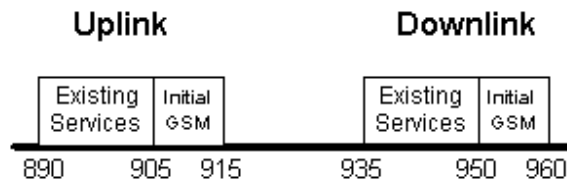


Figure 14 The GSM 900 MHz frequency band [29]

The frequency bands are divided into 125 channels with widths of 200 kHz each. These channels are numbered from 0 to 124. Channel number 0 is used as a guard band between GSM and other services on lower frequencies. Any frequency may be assigned to a mobile station by the base station from a selection of between 1 and approximately 16 frequencies. The number of channels a base station may have at its disposal depends on network planning considerations and the traffic density expected in the base station's coverage area. [42]

4.2.4 Transmission on the Radio Channels

As mentioned in Section 4.2.1, TDMA is used to make additional allocations in the time domain. This means that each frequency channel is further subdivided into eight different time slots numbered from 0 to 7. Each of the eight time slots is assigned to an individual user. A set of eight time slots is referred to as a *TDMA*

FRAME (Figure 15), and all of the users of a single frequency share a common frame.

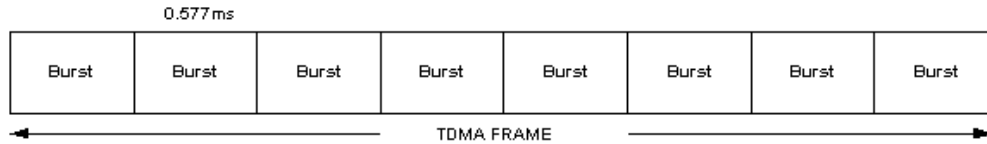


Figure 15 The structure of the TDMA frame [28]

If a mobile, for example, is assigned time slot number 1, it transmits only in this time slot and stays idle for the remaining seven time slots with its transmitter off. The mobile's regular and periodic switching (on and off) of its transmitter is called *bursting* and results in a so called *burst*. The length of a time slot, which is equivalent to a burst from a mobile, is as already mentioned $577 \mu\text{s}$, and the length of a TDMA frame is 4.615 ms ($8 \cdot 577 \mu\text{s} = 4.615 \text{ ms}$). [42]

Information is moved between mobiles as data (ones and zeros) that are confined to time slots. Each slot contains a burst, which is the information. Depending on the sort of information to be transmitted, different burst structures are used. GSM uses four different burst structures:

- The normal burst
- The "F" or frequency control burst
- The "S" or synchronous control burst
- The access control burst. [42]

A fifth type of bursts is the *dummy burst* which is to be sent downlink continuously in order to make the detection of a base station easier.

The *normal burst* is the most common burst in GSM and will therefore be described below.

The normal burst is used to carry data and most signalling. It has a total length of 156.25 bits, made up of two 57 bit information bits, a 26 bit training sequence used for equalisation, 1 stealing bit for each information block, 3 tail bits at each end, and an 8.25 bit guard sequence (Figure 16). The 156.25 bits are transmitted in $0.577 \mu\text{s}$, giving a gross bit rate of 270.833 Kbps. [29]

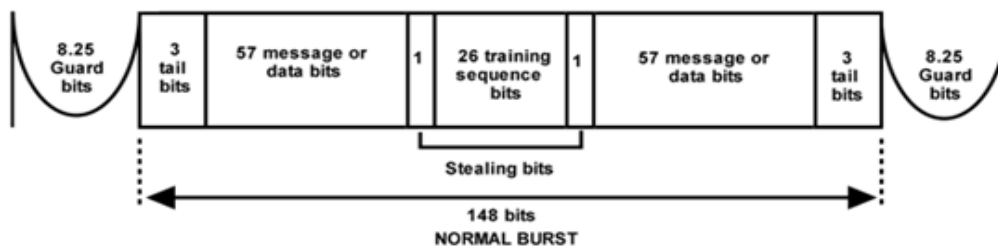


Figure 16 Structure of a normal burst [29]

This burst carries the conversation content in digital form. That's what the two 57 information, message, or data bits are for. The normal burst also carries signalling information needed to manage, e g call processing, which is data for setting up, maintaining, and ending a call. The different types of bits are described below:

Training sequence bits: These bits get the BTS and the MS in “tune” with each other.

Stealing bits: These bits are used to keep the mobile terminal linked to the base station even when there is no connection, e g when entering a tunnel or possibly when a large truck gets in the way.

Tail bits: These bits are always set to zero and are used as *guard time*.

Guard bits: These bits are empty time spaces separating data packets to make sure one burst does not run into another. [29]

4.2.5 Logical Channels

With the concept of logical channels, we are getting farther away from the physics of the signals in GSM and closer to the information carried. The way information is moved depends on the type of information. Different types of information can exist in the system on different types of logical channels. The contents of the different logical channels can appear in any physical channel (frequency and time slot). [42]

A logical channel carries signalling data, or a user's data. The data, of whatever kind, are mapped onto a physical channel. The manner in which the data are mapped onto the physical resource depends on the data's content. One should be more careful with important data than the more trivial data. [42]

GSM distinguishes between *traffic channels*, which are reserved for user data (speech and data), and *control channels*, which are used for network management messages and some channel maintenance tasks. The signalling (using the control channels) is the most important here and will be described more closely. [42]

The control channels are divided into four different classes:

- broadcast channels
- common control channels
- dedicated control channels
- associated control channels. [42]

Table 2 lists the control channels and gives a description about their use.

	Control Channels	Channel Types	Usage
The broadcast channels (BCH)	Broadcast Control Channel (BCCH)	Broadcast BTS → MS	Continually broadcasts, on the downlink, information including LAC ⁶ , MNC ⁷ , the information on which frequencies the neighbouring cells may be found, different cell options, and access parameters.
	Frequency Correction Channel (FCCH)	Broadcast BTS → MS	Used to synchronise the mobile to the time slot structure of a cell by defining the boundaries of burst periods, and the time slot numbering.
	Synchronisation Channel (SCH)	Broadcast BTS → MS	Every cell in a GSM network broadcasts exactly one FCCH and one SCH, which are by definition on time slot number 0 (within a TDMA frame).
The common control channels (CCCH)	Random Access Channel (RACH)	BTS ← MS	Slotted Aloha channel used by the mobile to request access to the network.
	Paging Channel (PCH)	BTS → MS	Used to alert the mobile station to an incoming call.
	Access Grant Channel (AGCH)	Broadcast BTS → MS	Used to allocate an SDCCH to a mobile for signalling (in order to obtain a dedicated channel), following a request on the RACH.
Dedicated/Associated control channels (DCCH)	Standalone dedicated control channel (SDCCH)	BTS ↔ MS	Used for the transfer of signalling information between a mobile and a base station.
	Slow associated Control Channel (SACCH)	BTS ↔ MS	Located in every traffic channel. Used for low rate, non critical signalling.
	Fast Associated Control Channel (FACCH)	Uplink and downlink BTS ↔ MS	A high rate signalling channel, used during call establishment, subscriber authentication, and for handover commands.

Table 2 Control channels in GSM [31]

Transmission of speech is done using the *traffic channel/full-rate speech* (TCH/FS). The net speech rate is 13 Kbps.

4.2.6 Frame Structures

In a manner similar to the TDMA frame structure that allows time slots to be ordered on a carrier, there are also some *multiframe* structures made of a fixed number of TDMA frames that allow logical channels to be ordered into time slots. There is a big difference between the logical channels that carry speech data and those that carry signalling data. A 26-*multiframe* structure is used for the traffic,

⁶ Location Area Code – Uniquely identifies a Location Area (LA) within a Public Land Mobile Network (PLMN).

⁷ The Mobile Network Code is part of the International Mobile Subscriber Identity (IMSI) and is used to uniquely identify a given network from within a specific country.

and a 51-multiframe structure is used for the signalling. To combine both structures onto the radio interface, a new frame format is introduced: the *superframe*. The superframe has a length of $51 \cdot 26 = 1,326$ TDMA frames. Superframes are used to build *hyperframes*, which consists of 2,048 superframes [42] (Figure 17).

The system sometimes refers to frame numbers within a hyperframe context, and the hyperframe represents the most comprehensive structure in the system and lasts for nearly 3,5 hours before it is repeated. This organisation of frames and frame types makes it easy to determine what sort of information communicating entities expect to find in a given period of time. [42]

When speaking about signalling, it is important to know exactly which frame is currently being transmitted. To remove the possibility of ambiguity, the frames are numbered in a special way: there are three counters, which will be called T1, T2 and T3. Counter T1 counts the superframes. [42]

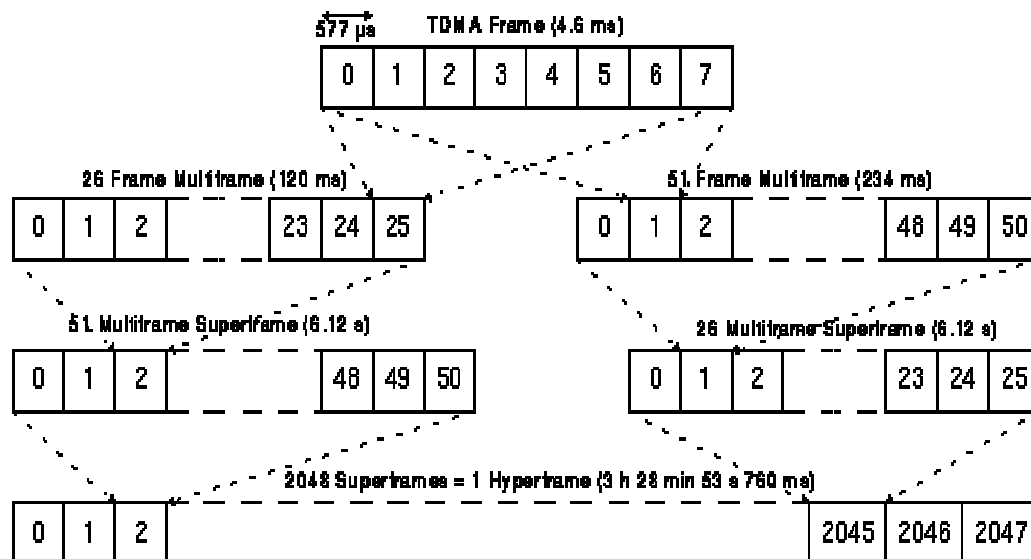


Figure 17 Frame Structures in GSM [30]

Whenever a superframe is completed, T1 is incremented by 1. T1 has values between 0 and 2,047; there are 2,048 superframes in a hyperframe. T2 counts the speech frames, which only occur in 26-multiframe structures. T2's value, therefore, ranges from 0 to 25. Finally, T3 counts the signalling frames, which are 51-multiframe structures. Similarly to the traffic counter, T3's contents can be anything from 0 to 50. At some starting time, all three counters are set to 0, and then the frames start to be transmitted. Whenever a speech or a signalling multiframe structure is finished, its respective counters (T2 and T3) are reset to 0 and start again. After 1,326 TDMA frames, both T2 and T3 are finally reset together and start counting again from 0 at that time. This marks the duration of

one superframe. When the first superframe is completed, T1 increments by 1 count. T1 only resets after 2,047 counts, which takes exactly 3 h 28 min 53 s 760 ms to do, and this is the duration of a hyperframe. If one knows the values in the T1, T2 and T3 counters, then one knows exactly what is in each and every time slot at that instant, provided one knows what kind of multiframe was assigned to each of the eight available time slots in the TDMA frame. An entity knowing T2 and T3 easily finds the BCCH and the system information. This fact makes it easy for MSs entering a new area to find the frequency of the specific area and start tuning to the new cell. The counters mentioned above make up the frame number (F_n), which is used together with the session key as input to the encryption algorithm used for voice encryption in GSM. [42]

4.2.7 Examples of How a Mobile Station Behaves

In this section a number of useful scenarios for understanding how a MS behaves on the radio interface are introduced [42]. The logical channels (described in Table 2) used to send different types of messages will be indicated here while they will be left out when attacks are described, since the same logical channel will be used. All of the information presented here will be used in later chapters when attacks against GSM security are described.

Of particular interest are three scenarios:

- synchronising with the network,
- location updating,
- call establishment.

These are described below:

4.2.7.1 Synchronisation with the Network

When a mobile station is turned on, it has to orient itself within the network. The mobile does this in three steps. First, it synchronises itself in frequency, then in time. Finally, it reads the system and cell data from the BCCH. This procedure is purely passive; no messages are exchanged.

The first task is to find the frequency where the FCCH, SCH, and BCCH are being transmitted. In the GSM system, a base station must transmit something in each time slot of the base channel. The base channel is the broadcast carrier. It contains the FCCH, SCH, and BCCH and is the network beacon. Even if certain time slots are not allocated to communication with any terminal, the base station has to transmit predefined *dummy bursts*, especially defined for this purpose, in all idle time slots of the base channel. If the base station, taxed with the responsibility of broadcasting the base channel, fills all of its timeslots, then the power density for this frequency is higher than that for any of the other channels in the cell, which may have only a few time slots out of eight allocated. This peculiarity of the base channel makes it easy for a mobile (or an intruder) to find the right frequency. It even enables an outsider to make a mobile think it is

communicating with a legitimate base station. The mobile simply scans for the physical channels with the highest apparent power levels. After finding one of them, the mobile searches for the FCCH. The FCCH is easy to find once the base channel is located. After the mobile synchronises with the system in the frequency domain, it proceeds to do the same in the time, or data, domain. The mobile uses the SCH for this second step, but it has already found the FCCH, so it already knows that the SCH will follow in the next TDMA frame (FCCH and SCH come always in consecutive slots in the 51-multiframe)

With this information available on the SCH, the BCCH is an open book for the mobile station (or an intruder), and it reads about the location of the cell, any cell options of interest, and how to access this particular base station. All of these steps are passive and take somewhere between 2 and 5 seconds.

4.2.7.2 Location Updating

The location updating procedure is always initiated by the mobile station e.g. when it finds itself in a different location area from the one in which it was registered before [22]. The network (or an intruder) can, however, force a mobile station to perform a location update when it is switched on. This is accomplished with a flag set in the system information transmitted on the BCCH. If all mobile stations have to register themselves after being turned on, then the network has exact knowledge of which mobile stations are currently active, as well as in which cell they can be found.

If the mobile is switched on in a different area from that stored on the SIM card (where it was last switched off), or if it enters a new area (roaming), the mobile station initiates a location updating procedure (Figure 18) to inform the network about its new location, which the network needs e.g. if a call has to be routed from the public network to the mobile station.

The principle of location updating is illustrated in Figure 18 along with the logical channels that are used during the procedure. Before the location update messages can be exchanged, the mobile has to request a signalling channel on which to exchange the messages. The mobile starts its channel request with a RACH, which it places on a random access burst. After it has sent the burst, the mobile listens to the AGCHs from the base. If there is no response within a certain period of time, the random access burst is repeated.

Upon receipt of the AGCH (in which there is a description of a dedicated channel the MS will have to go to), the mobile moves onto the new channel, which is now a dedicated channel between the mobile and the base station.

On the new channel – the SDCCH – the mobile station tells the network that it wishes to perform a location update. Before the network processes this request any further, it demands that the authentication procedure be performed. If the authentication is okay, the network assigns the new location area and makes note of the mobile's new location as it enters this information into the relevant registers (databases), namely the VLR and the HLR.

If necessary, the network assigns a temporary identity (TMSI) to the mobile, or it renews the old one. Now that the location update procedure is performed, the signalling channel is no longer needed, and the dedicated SDCCH is released for others to use.

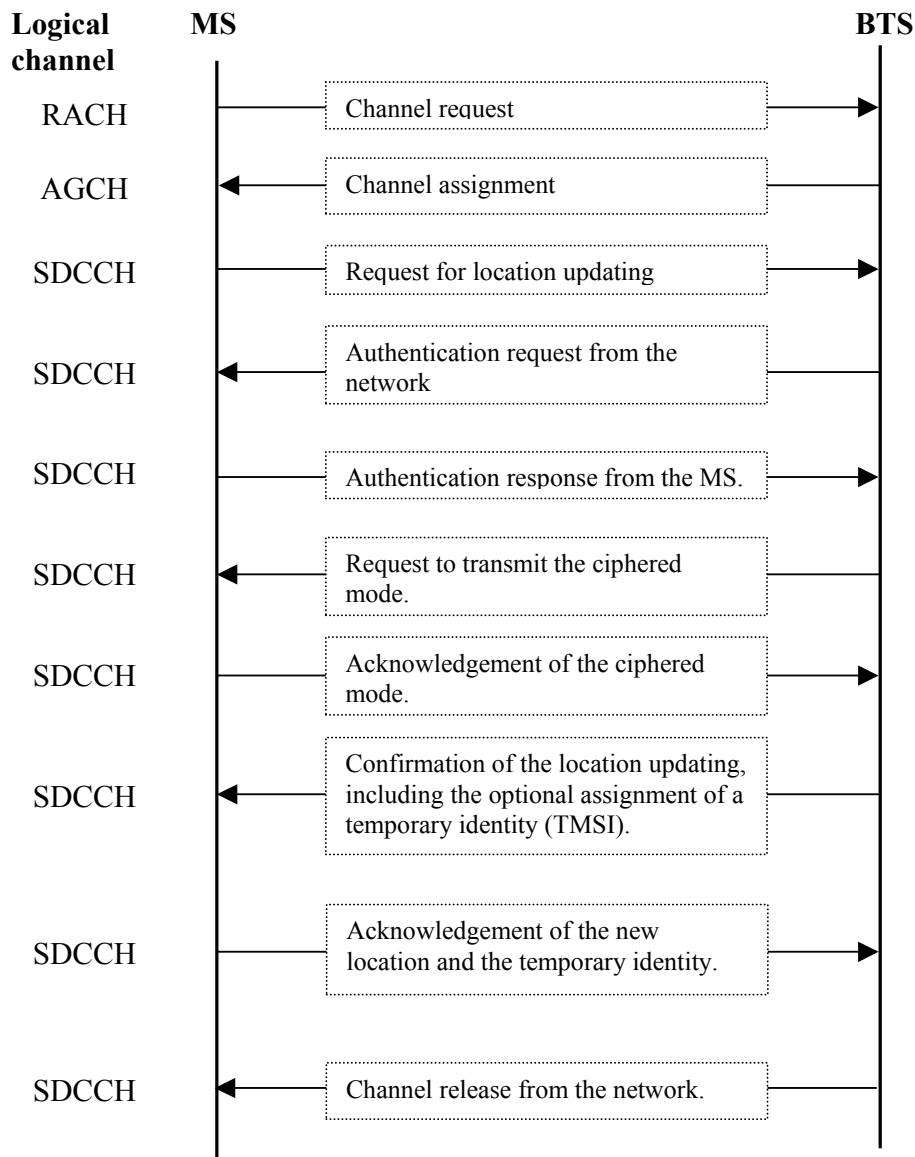


Figure 18 Principle of location updating

4.2.7.3 Call Establishment (Mobile-Terminated Call - MTC)

If a mobile station is switched on and already updated, it is in a state called *idle updated*. In this state the mobile passively monitors the BCCH and the CCCH, which is the PCH.

If the mobile is called from the public network, the base station will issue a paging message on the PCH to which a channel request from the mobile is the appropriate response. From now on, the MTC procedure (Figure 19) follows nearly the same rules as already described for location updating (Figure 18).

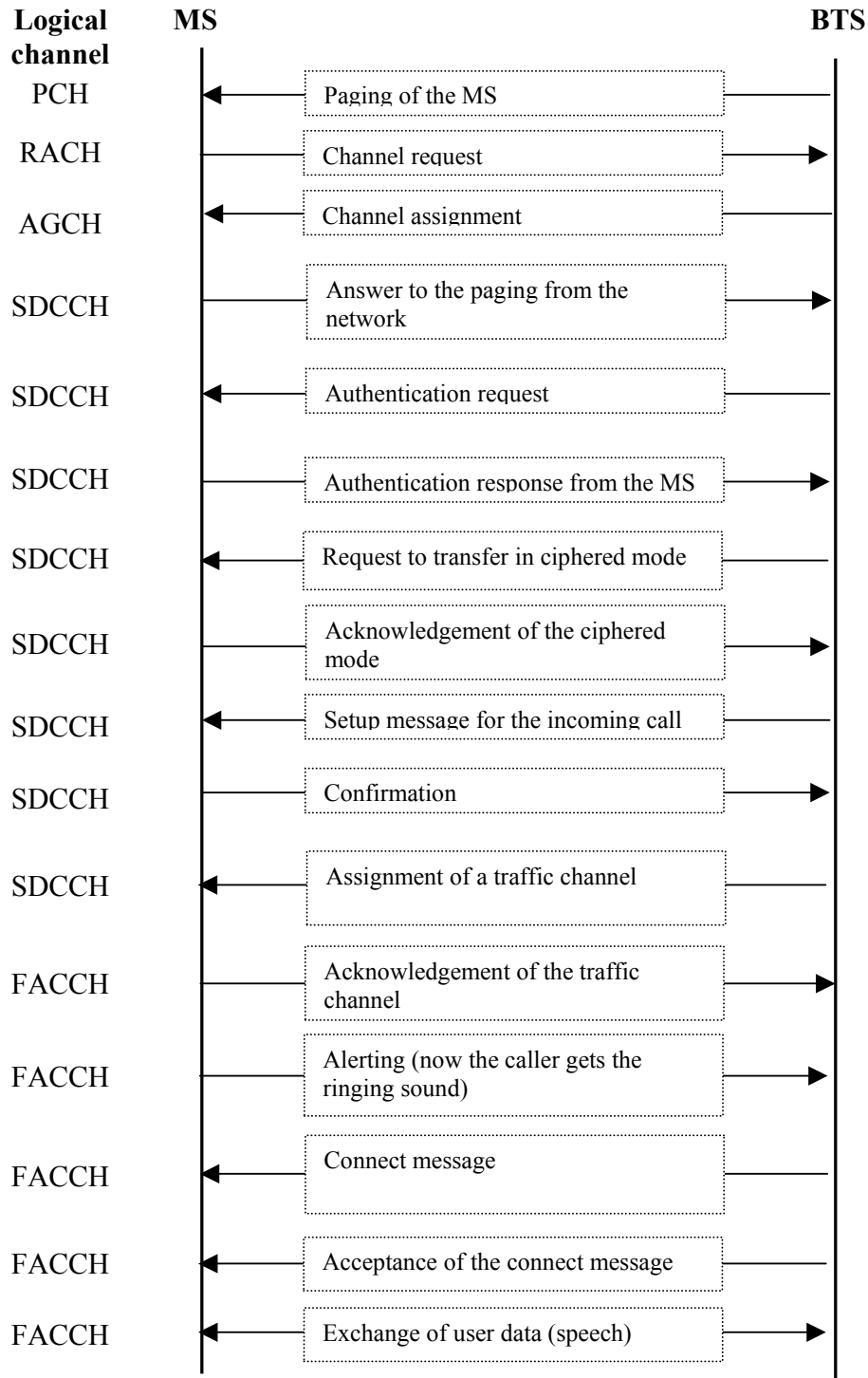


Figure 19 Principle of call establishment – mobile terminated

One of the differences is the first message on the assigned SDCCH signalling channel, which in the MTC case is the “answer to a page” message. Then some more messages follow on the SDCCH to set up the call until a traffic channel is finally assigned. From the instant the mobile and the base switch to the traffic channel, the remaining signalling messages are transmitted on a FACCH. Since the signalling is not finished yet and speech data are not yet transmitted, the FACCH is not yet displacing any traffic data, as the FACCH often has to do. When the call is finally connected, no further dedicated signalling messages need to be exchanged, and the traffic channel assumes the routine purpose for which it is intended.

4.2.8 From analog to digital

Before the voice data are transmitted on the radio channel the signal is transformed through several processing steps (Figure 20).

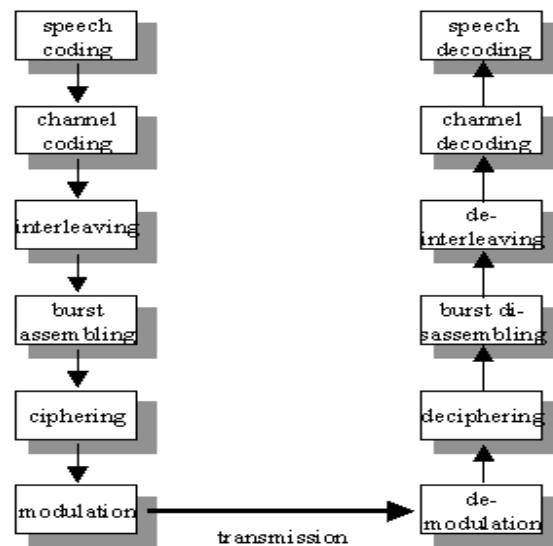


Figure 20 Transformation from speech source to radio waves [4]

The first step is speech coding. GSM is a digital system, therefore speech signals, inherently analog, have to be digitised. The GSM group studied several voice coding algorithms on the basis of subjective speech quality and complexity (which is related to cost, processing delay, and power consumption once implemented) before arriving at the choice of a Regular Pulse Excited - Linear Predictive Coder (RPE-LPC) with a Long Term Predictor loop. Basically, information from previous samples, which does not change very quickly, is used to predict the current sample. The coefficients of the linear combination of the previous samples, plus an encoded form of the residual, the difference between the predicted and actual sample, represent the signal. Speech is divided into 20 millisecond samples, each of which is encoded as 260 bits, giving a total bit rate of 13 Kbps. [31]

Due to natural or man made electromagnetic interference, the encoded speech or data transmitted over the radio interface must be protected as much as is practical. The GSM system uses convolutional encoding and block interleaving to achieve this protection. The exact algorithms used differ for speech and for different data rates. [32]

At the 900 MHz range, radio waves bounce off of everything - buildings, hills, cars, airplanes, etc. Thus many reflected signals, each with a different phase, can reach an antenna. Equalisation is used to extract the desired signal from the unwanted reflections. Equalisation works by finding out how a known transmitted signal is modified by multipath fading, and constructing an inverse filter to extract the rest of the desired signal. This known signal is the 26 bit training sequence transmitted in the middle of every time slot burst (Figure 16). The actual implementation of the equaliser is not specified in the GSM specifications. [32]

4.2.9 Frequency Hopping

The mobile station already has to be frequency agile, meaning it can move between a transmit, receive, and monitor time slot within one TDMA frame, which may be on different frequencies. GSM makes use of this inherent frequency agility to implement slow frequency hopping, where the mobile and BTS transmit each TDMA frame on a different carrier frequency. The frequency hopping algorithm is broadcast on the BCCH. Since multipath fading is (mildly) dependent on carrier frequency, slow frequency hopping helps alleviate the problem. In addition, co-channel interference is in effect randomised. [32] Some GSM officials have responded to claims that GSM security is on its way to be broken by referring to the frequency hopping as a sort of defence. Frequency hopping can be turned off by the base station. [35]

4.3 The Data Link Layer – Layer 2

The previous section gave a description about how the physical layer, i.e. the layer responsible for physically transmitting the digitised information over the radio link, generally works. The data link layer is responsible for the correct and complete transfer of *information blocks* between Layer 3 entities over the GSM radio interface. Layer 2 forms the envelopes that will contain the data to be transmitted. The protocol contains the following functions:

- Organisation of the valuable Layer 3 information into frames
- Peer-to-peer transformation of signalling data in defined frame formats
- Recognition of frame formats
- Establishment, maintenance (supervision), and termination of one or more (parallel) data links on signalling channels
- Acknowledgement of transmission and reception of numbered information frames (I-frames)

- Unacknowledged transmission and reception of unnumbered information frames (UI-frames). [42]

The BTS passes signalling messages between the mobile station and the BSC or MSC. The BTS seldom takes part in the conversations except when it has to respond to commands for adjustments in its operations.

4.4 The Network Layer – Layer 3

Up to now we have described the techniques that are used to generate the information and the infrastructure that is used to transmit it. Now we will take a look at what is transmitted, i.e. the information contained in the time slots and frames etc. This is where things become interesting, because the information generated at this layer is the information determining the services that the subscribers will get. [42]

The network layer in the GSM architecture, also referred to as the *signalling layer*, uses a protocol that contains all the functions and details necessary to establish, maintain, and then terminate mobile connections for all of the services offered within a GSM PLMN. The network layer also provides control functions to support additional services such as supplementary services and short message services. There are thoroughly defined procedures and structures for the protocol. [42]

This section will describe the structure of Layer 3, which can be divided further into three *sublayers*. The general procedures of the Layer 3 protocol will be introduced, as well as the parameters and the elements of a Layer 3 message. [42]

4.4.1 Sublayers of Layer 3

There exist three sublayers defined for Layer 3:

- Radio resource management RR
- Mobility management MM
- Connection management CM. [42]

These sublayers will be further explained in the following three subsections:

4.4.1.1 Radio Resource Management Sublayer (RR)

The tasks covered in this segment of the network layer are closely connected to the physical layer. The RR sublayer is responsible for the management of the frequency spectrum, the GSM system's reaction to the changing radio environment, and everything related to maintaining a clear channel between the PLMN and the mobile station. These responsibilities include channel assignment, power-level control, time alignment, and handover from one cell to another. The RR sublayer handles all the procedures necessary to establish, maintain, and release dedicated radio connections. Appropriately, the radio connections are also

called RR connections. An RR connection is necessary in order to provide a path for further signalling traffic, and then eventually a suitable traffic channel to carry the user's data. Within the RR sublayer, there are some carefully defined procedures used to cover these tasks. [42]

4.4.1.2 Mobility Management Sublayer (MM)

The MM sublayer has to cope with all of the effects of handling a mobile user that are not directly related to radio functions. These tasks include the kinds of things a fixed network would do to authorise and connect a user to a fixed network, which are modified to account for the fact that the user may not remain in the same place:

- Location update procedure
- Periodic updating
- Authentication
- IMSI attach procedure. A mobile station will perform (when indicated by the attach flag in the base station's BCCH) a location update procedure after power on, present its IMSI to the network, and get a TMSI in return
- IMSI detach procedure. A mobile station will perform (when indicated by the detach flag in the base station's BCCH) a *detach* procedure just after power off, telling the network that it is no longer in service
- TMSI reallocation procedure
- Identification procedure (authentication). [42]

4.4.1.3 Connection Management Sublayer (CM)

The CM sublayer manages all the functions necessary for circuit-switched call control in the GSM PLMN. These functions are provided by the call control entity within the CM sublayer. There are other entities within the CM sublayer to cope with providing supplementary services and SMSs. The call control responsibilities are almost identical to those provided in a fixed network; the CM sublayer is virtually blind to the mobility of the user. Again, there are some specific procedures defined for this purpose:

- Call establishment procedures for *mobile-originated cal*,
- Call establishment procedures for *mobile-terminated calls*
- Changes of transmission mode during an ongoing call (in-call modification)
- Call reestablishment after interruption of an MM connection
- Dual-Tone Multi-Frequency (DTMF) control procedure for DTMF transmissions. [42]

A call control connection is always based on an existing MM connection. The call control entity uses the MM connection to exchange information with its peer.

4.4.2 Structure of a Layer 3 Message

The layer 3 signalling at the radio interface may be divided into so-called structured procedures which consist of specific combinations of elementary procedures (Figure 21); furthermore each procedure consists of an exchange of signalling messages. A signalling message is, however, not the smallest unit we can consider when dividing signalling procedures into their constituent parts. A signalling message consists of information elements (IE), having a specific purpose, function, and information content. [22]

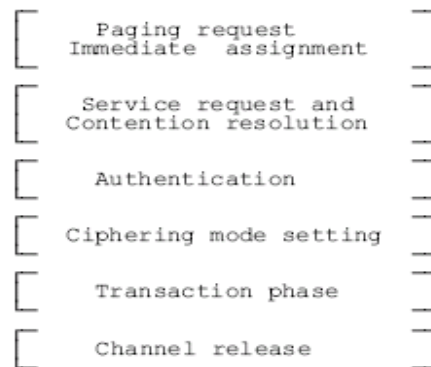


Figure 21 Components of structured procedures [22]

Figure 22 shows a schematic view of the structure of a Layer 3 message. The information contained in the message, and its binary representations (each octet of eight bits), are the Layer 3 part that goes into the information field of a Layer 2 frame. [42]

Only the *optional* and *mandatory* information elements of the Layer 3 message have a variable length, indicated with the dotted lines in Figure 22. All of the elements of the message are explained below:

Protocol Discriminator

Bits 1 to 4 of the first octet of a standard L3 message contain the protocol discriminator (PD) information element. The PD identifies the L3 protocol to which the standard layer 3 message belongs. The correspondence between L3 protocols and PDs is one-to-one. [42]

Transaction Identifier

The transaction identifier (TI) is a pointer with a length of four bits. It is used to distinguish between (possible) multiple parallel CM connections and between the various transactions taking place over these simultaneous CM connections. For RR and MM connections, the TI is not relevant; the TI field is always coded with 0000 in these cases. [42]

Message Type

The message type (MT) indicates the function of the Layer 3 message. It uses the lower six bits of the second octet in the Layer 3 message. Six bits are sufficient to address 64 different message types in a protocol indicated by the PD. The MT is part of a set of messages in a protocol. Bit 8 is reserved and is always set to 0. Bit 7 is a send sequence variable and may be used for MM and CM messages. [42]

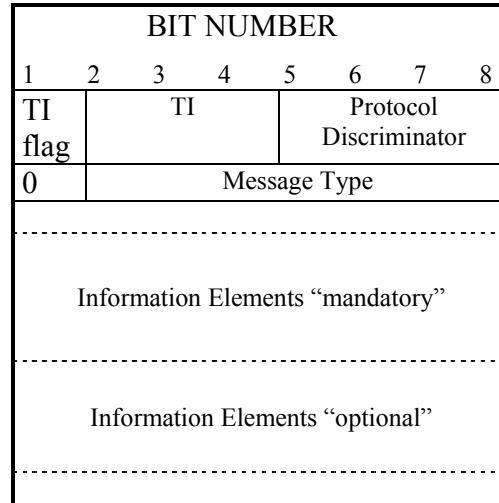


Figure 22 Structure and elements of a Layer 3 message [42]

4.4.3 A Layer 3 Signalling Trace

A call-setup recorded by a GSM test system will illustrate the exchange of signalling data over the radio interface (see Figure 23 on the next page). In this trace the Layer 2 envelopes are ignored and only the content is viewed. The procedure is a mobile originated call-setup that a GSM mobile station has performed in a network simulated by the test system. The information that is useful includes frame numbers, time stamps, the direction of transmission, the logical channels, time slot number, etc.

Layer 3 Trace	
MS transmits (uplink)	BS transmits (downlink)
FN : Frame Number	
FN:0004610: Umdl1,rach, TS(0),TI(0),pd(rr), 0003:08:20 chan_req, chan_req(= 7h,0Ah);	FN:0004622: Umdl1,agch,TS(0),TI(0),pd(rr), 0003:20:32 imm_ass, pag_mod(= 0h), chan_desc(= 8h,3h,0h,0i,0h,02h), req_ref(= EAh,03h,14h,08h), tim_adv(= 00h), mob_alloc(=);
FN:0004656: Umdl1,sdcch(0), TS(3),TI(0),pd(mm), 0003:02:15 cm_serv_req, cm_serv(= 1h), ciph_key_seq(= 0h), mob_class2(= 0h,0h,2h,0i,0h), mob_id(= 4h,0i,12345678h);	FN:0004743: Umdl1,sdcch(0),TS(3),TI(0),pd(mm), 0003:11:00 auth_req, ciph_key_seq(= 0h), rand(= 0123456789ABCDEF 0123456789ABCDEFh);
FN:0004809: Umdl1,sdcch(0), TS(3),TI(0),pd(mm), 0003:25:15 auth_res, sres(= 90DFD9F0h);	FN:0004845: Umdl1,sdcch(0),TS(3),TI(0),pd(rr), 0003:09:00 ciph_mode_cmd, ciph_mode_set(= 0i);
FN:0004911: Umdl1,sdcch(0), TS(3),TI(0),pd(rr), 0003:23:15 ciph_mode_com;	
FN:0004962: Umdl1,sdcch(0), TS(3),TI(0),pd(cc), 0003:22:15 setup, bear_cap(= 1i,1h,0i,0i,0h), cd_p_bcd(= 0h,1h,089996410h);	FN:0004998: Umdl1,sdcch(0), TS(3),TI(8),pd(cc), 0003:06:00 call_proc;
	FN:0005100: Umdl1,sdcch(0), TS(3),TI(0),pd(rr), 0003:04:00 assign_cmd, chan_desc(= 01h,4h,0h,1i,00h,00h), pow_cmd(= 0Fh),

Figure 23 A Layer 3 trace, call establishment [43]

Now we have an idea about how radio signals are generated and used to transmit different kinds of information across the air. In this section we will continue to study GSM and take a look at the architecture of the GSM network. This will include a description of the components that build up the system, and how these are interconnected, in addition to introducing the mechanisms used in order to protect the system's valuable assets, as outlined in Chapter 2.

5.1 An Overview of the GSM Network

A GSM network (Figure 24) is comprised of the Mobile Equipment (ME), the Subscriber Identity Module (SIM), the Base Station Transceiver (BTS), the Base Station Controller (BSC), the Transcoding Rate and Adaptation Unit (TRAU), the Mobile Services Switching Center (MSC), the Home Location Register (HLR), the Visitor Location Register (VLR), and the Equipment Identity Register (EIR). Together, they form a Public Land Mobile Network (PLMN). [3]

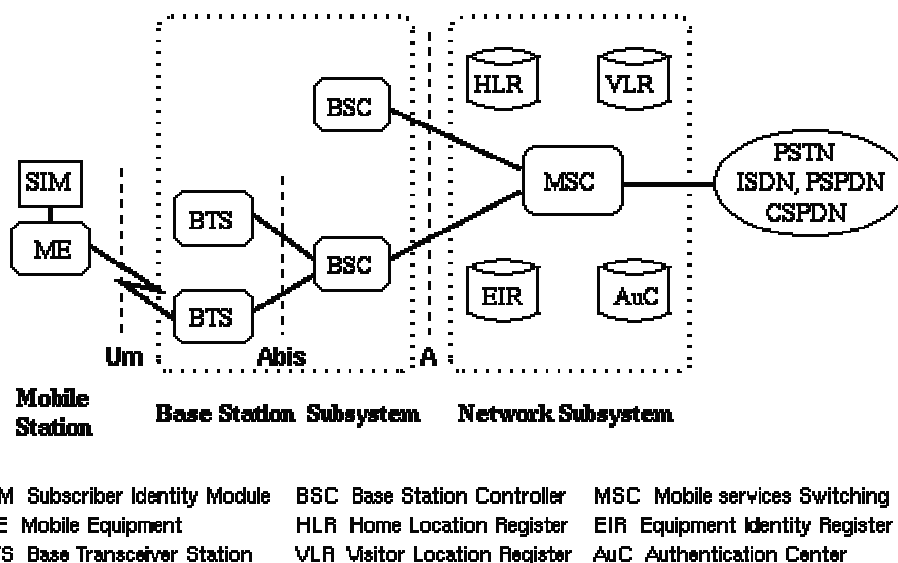


Figure 24 An overview of the GSM network [4]

These different components are described below:

5.1.1 The Mobile Station (MS)

The MS is carried by the subscriber. It is made up of the ME, also known as the terminal, and a smart card known as the Subscriber Identity Module (SIM).

The SIM, which is basically a smart card, determines the directory number and the calls billed to the subscriber. The SIM contains the following subscriber related information:

- The International Mobile Subscriber Identity (IMSI) , which uniquely identifies a subscriber and without which the GSM service is not accessible. IMSI is only used by the network.
- A secret subscriber authentication key K_i and a cryptographic algorithm A3/A8 which provide security functions for authenticating the SIM, and generating session keys.
- Temporary network related data like the Temporary Mobile Subscriber Identity (TMSI), Location Area Identifier (LAI), K_c , forbidden PLMNs, etc.
- Service related data like Language Preference and Advice of Charge.
- Card Holder Verification Information, that authenticates the user to the card and provides protection against the use of stolen cards [5]. A Personal Identification Number (PIN) is used. If the wrong PIN is entered three times in a row, the card locks itself, and can only be unlocked by providing a Personal Unblocking Key (PUK).

5.1.2 The Base Transceiver Station (BTS)

The BTS controls all of the radio related tasks and provides connectivity between the network and the Mobile Station (MS) via the radio interface.

5.1.3 The Base Station Controller (BSC)

The BSC takes care of all the central functions and controls a set of BTSs. The BSC and the controlled BTSs form the Base Station Subsystem (BSS).

5.1.4 Mobile Services Switching Center (MSC)

The MSC controls a large number of BSCs. It is very similar to a digital telephone exchange or a switch and it handles the routing of incoming and outgoing calls and the assignment of user channels on the A-interface.

5.1.5 Home Location Register (HLR)

The HLR is a data repository that stores the subscriber specific parameters of a large number of subscribers. The most important parameters of a subscriber, like the K_i and IMSI are stored in the HLR. Every PLMN requires at least one HLR and every user is assigned to one specific HLR.

5.1.6 Authentication Center (AuC)

The AuC has as a key component a database of identification and authentication information for each subscriber, and is in most cases an integral part of the HLR.

Attributes in this database include the subscriber's IMSI, secret key K_i , LAI, and TMSI. The AuC is responsible for generating triplets of values consisting of the RAND, SRES (Signed RESponse), and session key K_c which are stored in the HLR for each subscriber. [1]

5.1.7 Visitor Location Register (VLR)

The VLR network element was devised to off-load the HLR of user database related functions. The VLR, like the HLR, contains subscriber information, but only information for those subscribers who roam in the area for which the VLR is responsible. When a subscriber roams away from the network of his/her own service provider, information is forwarded from the subscriber's HLR to the VLR of the serving network, in order to complete the authentication process. When a subscriber moves out of a VLR area, the HLR takes care of the relocation of the subscriber information from the old to the new VLR. A VLR may have several MSCs, but one MSC always uses one VLR. [1]

5.1.8 Equipment Identity Register (EIR)

Since the subscriber identity (SIM) and the ME are treated independently by GSM, it is possible to operate any GSM ME with any valid GSM SIM. This makes cellular terminal theft an attractive business and probably starts a possible black market for stolen GSM terminals. To protect against such thefts, the Equipment Identity Register (EIR) was introduced in the GSM system. Every GSM terminal has a unique identifier, called the International Mobile Station Equipment Identity (IMEI), which (according to the GSM organisation) cannot be altered without destroying the terminal. It contains a serial number and a type identifier [6]. The EIR maintains three lists:

- *The White list*: is composed of all number series of equipment identities that are permitted for use
- *The Black list*: contains all equipment identities that belong to equipment that need to be barred
- *The Grey list*: MEs on the grey list are not barred (unless on the black list or not on the white list), but are tracked by the network (for evaluation or other purposes). [1]

Equipment Identification can be done by the network operator by requesting the IMEI from the ME. [6]

5.2 The Security Implementation – Protecting Valuable Assets

The mechanisms used in GSM to provide anonymity, authentication and confidentiality to shareholders are described in the following subsections.

5.2.1 Anonymity

Anonymity is provided by using temporary identifiers. When a user switches on his/her mobile terminal, the real identity (IMSI) is used to identify the MS to the network and then a temporary identifier Temporary Mobile Subscriber Identity (TMSI) is issued and used for identifying the MS to the network in future sessions. According to the ETSI specification the network should always encrypt TMSI before transmitting it to the MS.

A LOCATION UPDATE REQUEST results in the MS receiving a TMSI [22]. The TMSI has significance only within a location area. Outside the location area it has to be combined with the LAI to provide for an unambiguous identity. Usually the TMSI reallocation is performed at least at each change of a location area, as a LOCATION UPDATE REQUEST is issued by the MS to the network (Such choices are left to the network operator). [20]

From then on the temporary identifier is used. Only by tracking the user is it possible to determine the temporary identifier being used. [7]

5.2.2 Authentication

Since the radio medium can be accessed by anyone, authentication of users to prove that they are who they claim to be is a very important element of a mobile network. Authentication involves two functional entities, the SIM card in the mobile, and the Authentication Center (AuC). One of the primary security functions of the SIM is to authenticate the subscriber to the network. This process assures the network that the MS requesting service is a legitimate subscriber and not some intruder. A GSM network verifies the identity of a subscriber through a *challenge-response* process similar to the mechanism described in 3.2.1. When a MS requests service, the network sends a mathematical challenge to the MS (RAND), which it must answer correctly before being granted access. [7]

The challenge sent by the network to the MS consists of a 128 bit number called RAND. It is very important that RAND is unpredictable and has a very slim chance of being repeated, otherwise an attacker could easily make a codebook of (RAND, SRES) pairs and use the information to gain access to services. When the MS receives RAND it passes it into the SIM for processing. The SIM sends RAND and the secret 128-bit key K_i^8 through the A3 algorithm to produce a 32-bit "signed response". The response, called SRES, is transferred out of the SIM into the terminal, where it is then transmitted to the network. This is the MS's response to the network's challenge. Meanwhile the network (the AuC) has performed the same set of operations (Figure 25). Using the same value of RAND and an identical copy of K_i , the network has computed its own value for SRES. When the network receives SRES from the MS it compares it to its own SRES. If the two values are identical, the network assumes the MS is legitimate and allows service to proceed. If the two values are not the same, the network assumes the

⁸ The GSM specifications do not specify the length of K_i , thus it is left for the choice of the operator, but usually it is a 128-bit key.

SIM does not have the proper secret key K_i and therefore denies service to the MS. [5]

Since the RAND value changes with (almost) every access attempt, an eavesdropper recording the SRES response will not be able to successfully reuse it later. Even if by chance a particular RAND challenge happens to be reused (and an attacker manages to impersonate a legitimate subscriber to the network), a GSM network has the flexibility to authenticate the MS as often as it wishes; perhaps several times throughout the duration of a call. The next challenge the MS (SIM) receives from the network will probably be a new one for the attacker, impossible for him/her to compute the right SRES for [45]. It should be noticed that a cornerstone of the GSM security protocols is that a subscriber's secret key, K_i , remains secret. While stored in both the SIM and the AuC, K_i is never transmitted over the network. Figure 25 illustrates the authentication process.

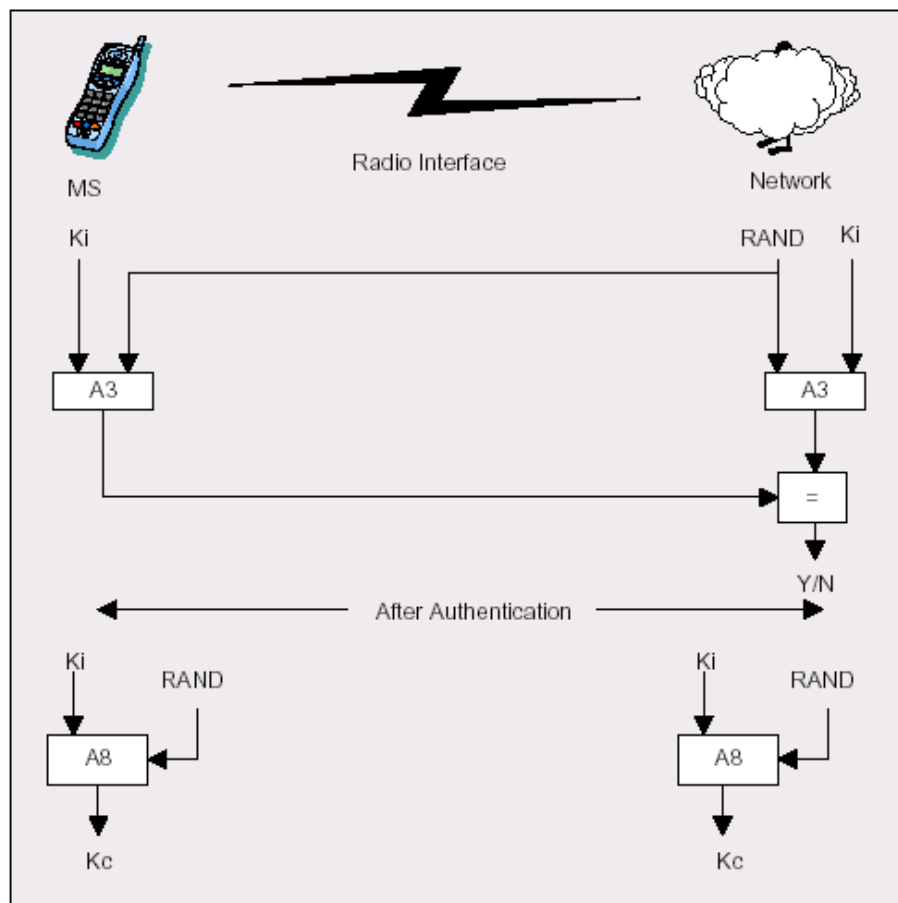


Figure 25. Authentication and session key generation in GSM

A3 and A8 aren't actually algorithms, but simply placeholders [5]. The COMP128 algorithm (Figure 26) is almost exclusively used for A3 and A8 throughout the world. This algorithm was designed to be a reference model for

GSM implementation but for various reasons has been adopted by almost all GSM providers world-wide. COMP128 was cracked in April 1998 and a new stronger version, COMP128-2 was developed. However, due to the huge amount of cost involved in replacing COMP128 (or maybe ignorance in some cases), it is believed that most operators are still using the old flawed algorithm. [44]

5.2.3 Confidentiality

The SIM also provides information needed to encrypt the radio connection between the MS and the BTS. More specifically it computes the session key K_c , which is later used in some version of A5 for encrypting the voice or data before transmission on the radio path. The algorithm used for computing the 64 bit K_c , is called A8 and is invoked according to Figure 25 [5]. A3/A8 is, as mentioned in the previous section, often realised in practice using the initial design specification given by the GSM MoU, which is a single algorithm called COMP128 (Figure 26).

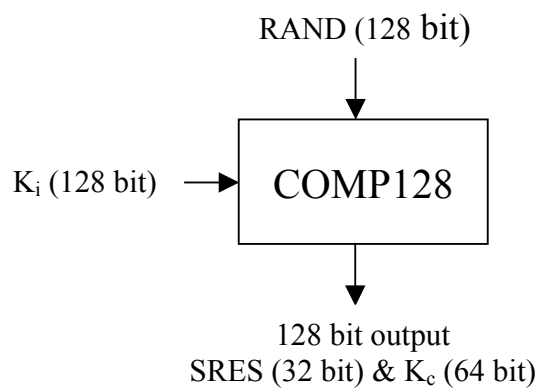


Figure 26 A popular A3/A8 implementation (COMP128)

Recall that GSM uses a technique called *time division* to share the radio channel with up to eight other users (see Section 4.2.1). Each user takes turns using the common radio channel, sending and receiving information only during one of the eight available time slots in every frame. Each frame is very short, lasting only about 4.6 milliseconds, and is identified by a *frame number*. A GSM conversation uses two frames, one going from the base station to the MS and another going from the MS back to the base station. Each of these frames (time slots) contains 114 bits of user information, which is often digitised and compressed speech. Thus, every 4.6 milliseconds the MS receives 114 bits of information from the base station and transmits another 114 bits to the base station. It is these 228 bits that require encryption to protect it from eavesdroppers.

Using RAND and the secret key K_i , the SIM runs the A8 algorithm (or COMP128) to produce a 64-bit long session key called K_c . K_c is transferred out

of the SIM and into the MS, where it is used by a third algorithm (A3 and A8 are the other two) called A5.

A5 uses K_C and the current, publicly known, frame number to produce a key stream of 228 bits, half of which encrypts the dl and the other half encrypts the ul. For each new frame to be transferred a new 228 bit key stream is produced by the A5 to be used to encrypt (and decrypt) the frame. A5 resides in hardware in the terminal, not in the SIM, and must operate quickly and continuously to generate a fresh set of 228 bits every 4.6 milliseconds. Also, because GSM terminals are designed to operate in different networks, the A5 algorithm must be common to all GSM networks.

There are presently at least two known versions of A5. The first, called A5/1, is only used in countries that are members of CEPT⁹, provides the strongest level of encryption across the air link (Actually A5/3 is stronger but it is to be used in future networks). Although officially using 64 bit keys, in actual practice the keys are no more than 54 (!) bits long, the last ten bits are forced to be zeros. The second algorithm, A5/2, is considered to be much weaker than A5/1 and is designed for export to countries outside CEPT where presumably there is an interest in easily cracking encrypted conversations.

Since encryption requires additional hardware in each base station, raising the cost and complexity of the network, a third “encryption” option is to employ what’s called A5/0, that is, no encryption at all. The network asks the MS to start encrypting using a certain encryption algorithm each time information will be transmitted on the radio link.

The two algorithms A5/1 and A5/2 will be described in the following subsections:

5.2.3.1 Description of the A5/1 Stream Cipher

A5/1 is built from three short linear feedback shift registers (LFSR) of lengths 19, 22, and 23 bits, which are denoted by $R1$, $R2$ and $R3$ respectively (Figure 27). The rightmost bit in each register is labelled as bit zero. The taps of $R1$ are at bit positions 13,16,17,18; the taps of $R2$ are at bit positions 20,21; and the taps of $R3$ are at bit positions 7,20,21,22. When a register is clocked, its taps are XORed together, and the result is stored in the rightmost bit of the left-shifted register. [8]

The three registers are maximal length LFSRs with periods $2^{19} - 1$, $2^{22} - 1$, and $2^{23} - 1$, respectively. They are clocked in a stop/go fashion using a majority rule as follows. Each register has a single “clocking” tap (bit 8 for $R1$, bit 10 for $R2$, and bit 10 for $R3$); each clock cycle, the majority function of the clocking taps is calculated and only those registers whose clocking taps agree with the majority bit are actually clocked, then at each step either two or three registers are clocked, and each register moves with probability 3/4 and stops with probability 1/4. [8]

⁹ CEPT – Conférence Européenne des administrations des Postes et des Télécommunications.

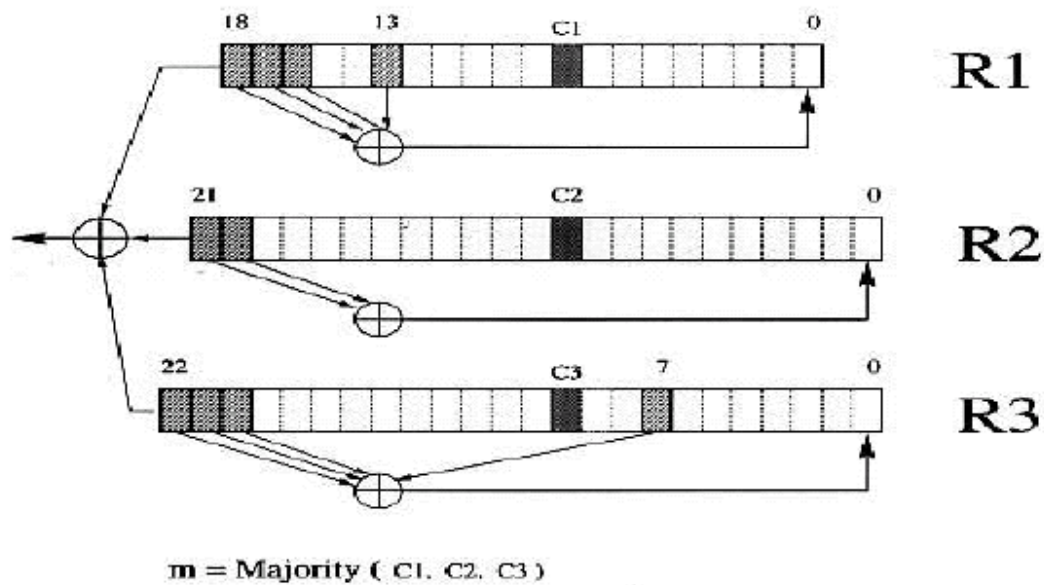


Figure 27 The internal structure of the A5/1 Stream cipher [8]

The process of generating pseudo random bits (the keystream) from the session key K_c and the frame counter F_n is carried out in four steps:

1. The three registers are zeroed, and then clocked for 64 cycles (ignoring the stop/go clock control). During this period each bit of K_c (from lsb to msb) is XORed in parallel into the lsbs of the three registers.
2. The three registers are clocked for 22 additional cycles (ignoring the stop/go clock control). During this period the successive bits of F_n (from lsb to msb) are again XORed in parallel into the lsbs of the three registers. The contents of the three registers at the end of this step is called the *initial state* of the frame.
3. The three registers are clocked for 100 additional clock cycles with the stop/go clock control but without producing any outputs.
4. The three registers are clocked for 228 additional clock cycles with the stop/go clock control in order to produce the 228 output bits. At each clock cycle, one output bit is produced as the XOR of the msbs of the three registers. [8]

Each 4,6 ms a fresh set of 228 bits are generated and XORed with 228 bits of plaintext to produce the ciphertext, before transmitting it on the radio link.

5.2.3.2 Description of the A5/2 Stream Cipher

The A5/2 stream cipher is very similar to A5/1. It consists of four LFSRs: R1, R2, R3 and R4 with the lengths 19, 22, 23 and 17 bits respectively (Figure 28).

Each register has taps and a feedback function. At each step of A5/2 R1, R2, and R3 are clocked according to a certain clocking mechanism that is determined using R4. After the three registers are clocked R4 is clocked. Post clocking, one output bit is ready at the output of A5/2. The output bit is, as in the case of A5/1, a non-linear function of the internal state of R1, R2 and R3. [35]

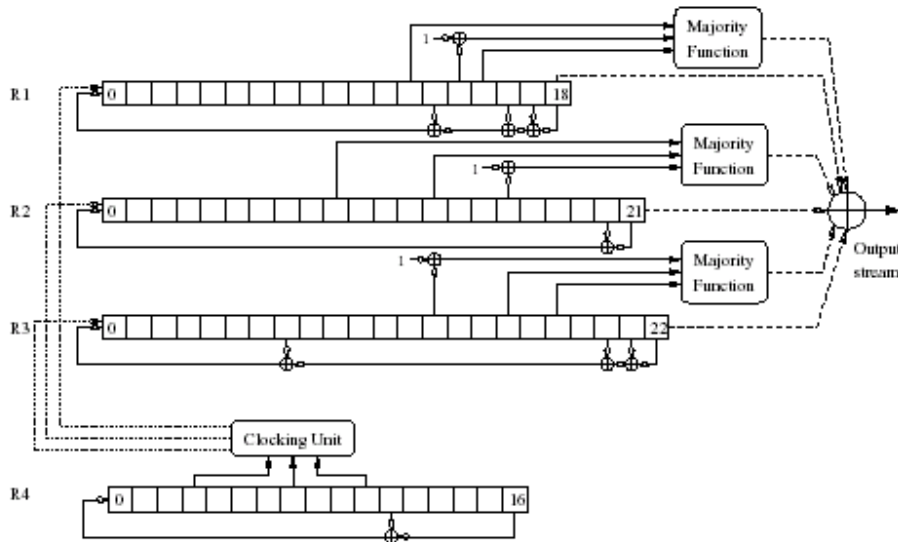


Figure 28 The internal structure of the A5/2 Stream cipher [36]

The process of generating pseudo random bits from the session key K_c and the frame counter F_n is carried out in the same way as in the case of A5/1 with a few differences:

1. Same as in the A5/1 case except that we have 4 registers
2. Force the bits R1[15], R2[16], R3[18], R4[10] to be 1
3. Run A5/2 for 99 clocks and ignore the output. [35]

After the first clocking is performed the first output bit is ready at the output of A5/2. The clocking mechanism works as follows: R4 controls the clocking of R1, R2 and R3. When clocking of R1, R2 and R3 is to be performed, bits R4[3], R4[7] and R4[10] are the input of the clocking unit. The clocking unit performs a majority function on the bits. R1 is clocked if and only if R4[10] agrees with the majority. R2 is clocked if and only if R4[3] agrees with the majority. R3 is clocked if and only if R4[7] agrees with the majority. After these clockings R4 is clocked. [35]

Once the clocking is performed, an output bit is ready. The output bit is computed as follows: in each register the majority of two bits and the complement of a third bit is computed; the results of all the majorities and the rightmost bit from each register are XORed to form the output. Note that the majority function is quadratic in its input: $maj(a,b,c) = a.b + b.c + c.a$. [35]

The difference between A5/2 and A5/1 is that A5/2 also initialises R4, and that one bit in each register is forced to be 1 (!) after initialisation. Then A5/2 discards 99 bits of output while A5/1 discards 100 bits of output. The clocking mechanism is the same, but the input bits to the clocking mechanism are from R4 in the case of A5/2, while in A5/1 they are from R1, R2, and R3. The designers meant to use similar building blocks to save hardware in the mobile terminal [11].

This algorithm outputs 228 bits of key-stream. The first block of 114 bits is used as a key-stream to encrypt the link from the network to the customer, and the second block of 114 bits is used to encrypt the link from the customer to the network. Encryption is performed as a simple XOR of the message with the key-stream.

5.2.4 Preventing Theft of Service or Equipment

As mentioned earlier, in GSM the customer subscription and authentication capability is contained within a smart card (SIM). Any mobile will take on the identity of a subscriber by insertion of a smart card. The mobiles now become attractive items to steal, as they can be used with another SIM card.

To prevent this, GSM has specified an International Mobile Equipment Identifier (IMEI). Although at first evaluation to an operator, it may seem as the stolen mobiles have no effect since they do not affect a subscription, there will be problems with an increase in customer facing staff to handle esquires and a possibility that GSM terminals are expensive to insure. [1]

An Equipment Identity Register (EIR) exists in each network, with Black, White and Grey Lists (see Section 5.1.8) for stolen or non type approved mobiles, valid mobiles and mobiles that need tracking, respectively.

GSM has defined a procedure so that approved, lost or stolen mobile IMEIs can be communicated to all other operators. Type approval authorities issue white list numbers (random ranges of valid IMEIs) to mobile manufacturers, and manufacturers inform the Central Equipment Identity Register (CEIR) when the mobiles are released to market. All operators are able to post their black lists to the CEIR, and in return collect a consolidated list of all operators' black and white lists. [50]

Part III

Breaking the Security of GSM, and Evaluating the Consequences for Users with High Security Requirements

Chapter 6 Attacks on GSM

Chapter 7 Evaluation of the Suitability of GSM for Special Users

This chapter will describe attacks on the important aspects of anonymity, authentication and confidentiality.

6.1 Capturing One or Several Mobile Stations

In many of the attacks which will be described in this chapter, the attacker is required to impersonate the network to the MS or impersonating the MS to the network or combining both in a so called *man-in-the-middle* attack. An attacker impersonating both entities to each other is able to eavesdrop, modify, delete, re-order, replay, spoof and even function as a repeater relaying signalling and user data between the two communicating parties. The required equipment is a modified BTS *in conjunction with a modified MS*. The modified BTS impersonates the network to the MS, whereas the modified MS impersonates the MS to the network. The term rogue base station (RBTS) will be used to refer to this device. BTS will refer to a legitimate base station. Figure 29 illustrates the situation.

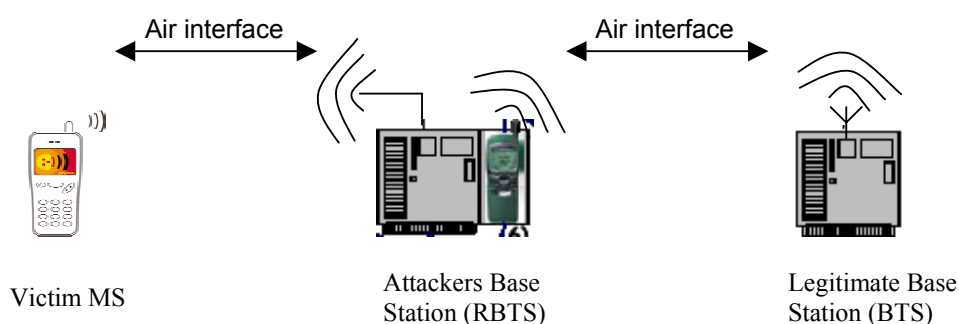


Figure 29 Mounting a man-in-the-middle attack [10]

Before an active attack can be conducted the attacker may have to capture the MS. As described in Section 4.2.7.1, a legitimate base station continuously broadcasts dummy bursts on the base channel to make it possible for MSs to find the serving cell's BCCH. When the BCCH is found all the information needed about the network in order to request services will be found on this channel. Recall that this information includes cell identity, network identity, control channel structure, list of channels in use and details of the access protocol. An attacker equipped with a RBTS residing between the MS and the legitimate BTS, providing higher power levels than the BTS (by transmitting dummy bursts more frequently) is able to make the MS use the RBTS's BCCH, thereby having total control over which system information the MS gets and even which messages

(and with which content) will reach the legitimate network from this MS. The MS is captured by the attacker who controls what messages go between the MS and the legitimate network as well as messages flowing in the other direction. The captured MS identity will then be used to provide fabricated messages on behalf of a legitimate subscriber.

Capturing MSs not only gives the attacker the ability to mount several attacks on GSM users, it is in itself a quite severe attack (a denial of service attack) . Captured MSs have no contact with the network (apart from information that the attacker relays) and are therefore unable to get services. It is fully possible and quite easy to do, though it is also easy to detect. A denial of service attack does not, however, require that the attacker use a RBTS; the attacker only needs to jam the radio signals.

6.2 Attacks on the Anonymity of GSM Users

Recall from Section 5.2.1 that anonymity in GSM is provided by using the temporary identifier TMSI instead of IMSI, which uniquely identifies a subscriber. An attacker may want to track some subscriber's movements and/or calling patterns and thus needs to know the IMSI and the TMSI of the MS. This information, if compromised, may also be used to launch attacks on other aspects of the system than the anonymity, where the victim has to be a specific person using GSM and not just any user, for instance eavesdropping on a specific person. If the attacker can get the IMSI of a subscriber or associate a TMSI currently being used in the cell with a specific IMSI then the anonymity of the user of the system owning that IMSI is compromised. The compromised anonymity leaves, as already mentioned, the door open for the attacker to perform traffic analysis, i.e. observe the time, rate, length, sources or destinations of messages on the radio interface or other system interfaces in the network.

Attacks on the anonymity of GSM users can be passive or active, as described below:

6.2.1 Passive Monitoring

Every time a MS is powered on an IMSI attach is performed. IMSI attach is performed by the MS in order to indicate the IMSI as active in the network. IMSI attach is performed by using the location updating procedure. The location updating type information element in the LOCATION UPDATING REQUEST message is sent by the MS and in this case indicates IMSI attach. Since the IMSI is not registered in the network it is not associated with a K_i and encryption cannot be applied. Therefore the IMSI is transmitted in the clear. An attacker listening to the traffic is able to extract the IMSI and can then register that user in the current area.

It is not only when IMSI attach is performed that the IMSI can be captured. Each time TMSI cannot be used by the HLR to retrieve user specific parameters, e.g.

when the database is not functional for some reason, the IMSI is requested from the service requesting MS and is then transmitted in the clear.

Before a location updating procedure is terminated the subscriber is assigned a TMSI to be used in the future when communicating with the network. The TMSI, as mentioned earlier, should according to the GSM specifications be encrypted prior to transmission and used in future communication sessions with the network. An operator following the ETSI specifications, assigning TMSI to the subscriber after IMSI attack is performed, makes life difficult for an attacker who is trying to track a subscriber. It is very difficult for an attacker lacking decryption capabilities to perform tracking. Only knowledge of the IMSI is not sufficient.

The message used by the network to request a MS's IMSI or other identification parameters is the IDENTITY REQUEST message. This very message will be used in the next section to perform an active attack on the anonymity of GSM users.

Trying to passively track GSM users and eavesdropping on the users' permanent identity (IMSI) is possible and quite easy, although the passive nature of the attack limits the possibilities. It provides the attacker with a functional IMSI, i.e. an identity that can be used, and the attacker knows that the owner of that IMSI is in the present area. Passive monitoring is however inefficient and time-consuming because the attacker needs to either wait for MSs to perform IMSI attach when it is powered on or for a database failure to occur in the network, which probably does not happen so frequently. If the service provider follows the GSM specifications and encrypts TMSI before transmission on the radio link, it will become very difficult to capture, hence tracking will become impossible¹⁰.

[15] offers "a passive monitoring system for the encrypted GSM networks". This equipment is supposed to be completely transparent in operation to the operator network, capable of monitoring in excess of two hundred targets simultaneously providing target and corresponding call data TMSI, IMSI, IMEI, dialed number, time and date and two way speech recording of one target. Furthermore, this device is, according to the manufacturer's homepage, able to decrypt GSM conversations.

6.2.2 Active Monitoring

As mentioned in the previous section, passive monitoring is considered inefficient for tracking GSM users and performing traffic analysis. Much better efficiency is reached if the attacker is able to communicate with the target MSs. This requires however that the attacker has more advanced equipment than in the passive monitoring case. In the passive monitoring case scanning for GSM radio frequencies is sufficient, whereas in the active case the attacker is in need of base station functionality that provides him/her with the ability to e.g. fabricate messages.

¹⁰ This assumes that the attacker lacks decryption capabilities.

To track a subscriber the attacker can make use of the identification procedure. The network may initiate an identification procedure, e.g. if the network is unable to identify the MS using its TMSI. The identification procedure is initiated by the network by transferring an IDENTITY REQUEST message to the mobile station (Figure 30) thereby asking it to transmit a specified identification parameter. The parameter can be specified using the IDENTITY TYPE information element (Figure 31). As seen in Figure 31 the network can choose to request IMSI, IMEI or TMSI. [22]

IEI	Information element	Type / Reference	Presence	Format	Length
	Mobility management protocol discriminator	Protocol discriminator 10.2	M	V	1/2
	Skip Indicator	Skip Indicator 10.3.1	M	V	1/2
	Identity Request message type	Message type 10.4	M	V	1
	Identity type	Identity type 10.5.3.4	M	V	1/2
	Spare half octet	Spare half octet 10.5.1.8	M	V	1/2

Figure 30 IDENTITY REQUEST message content [22]

Since GSM does not use message authentication to check message origin on the radio link, an attacker with sufficient base station functionality is able to use these messages, in an active attack, to retrieve the same information as a legitimate base station.

The attacker starts by capturing the MSs as described in Section 6.1. Every MS in the area will request a dedicated channel and initiate a location update procedure with the RBTS. This is a man-in-the-middle attack and the messages transmitted by the MS are relayed by the attacker to the legitimate network, if necessary. Relaying the whole session between the MS and BTS is necessary if the channel (frequency and time slot) used between the MS and the RBTS is not the same as the one used between the RBTS and BTS. This is probably the case in most cases.

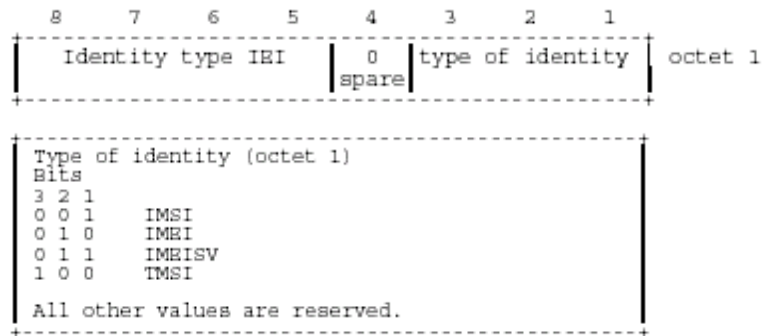


Figure 31 IDENTITY TYPE information element [22]

The attacker first ensures that he/she gets the identity information by transmitting IDENTITY REQUEST message to the MS, where the IDENTITY TYPE IE indicates that identification with IMSI is required. The MS will respond with the IDENTITY RESPONSE message of the type shown in Figure 32, providing the attacker with the requested information.

IEI	Information element	Type / Reference	Presence	Format	Length
	Mobility management protocol discriminator	Protocol discriminator 10.2	M	V	1/2
	Skip Indicator	Skip Indicator 10.3.1	M	V	1/2
	Identity Response message type	Message type 10.4	M	V	1
	Mobile identity	Mobile identity 10.5.1.4	M	LV	2-10

Figure 32 IDENTITY RESPONSE message content [22]

Now the attacker has the IMSI of the subscriber and is able to uniquely identify him/her. The next step is to capture the TMSI that the network allocates the specific MS, so that the attacker is able to associate the IMSI and the TMSI. This will enable the attacker to track the MS's movements and the sort of traffic/services that the subscriber utilises, since the TMSI is used in communications when it has been issued. The TMSI is, however, encrypted before transmission on the radio link, therefore the attacker needs to suppress the encryption somehow. This can be done in several ways. E.g the attacker can create a situation where the two legitimate entities communicating believe that they have incompatible encryption capabilities.

Figure 33 shows the messages exchanged between the entities in this attack and gives a good idea about what it means to mount a man-in-the-middle attack. As

can be seen in the figure, the attacker inserts his/her own messages (Identity Request), discards, and fabricates responses (Ciphering Mode Command and Ciphering Mode complete, respectively). The attacker can even relay modified messages. These big holes in the security of GSM are a due to lack of message integrity assurance and the non-presence of network-authentication.

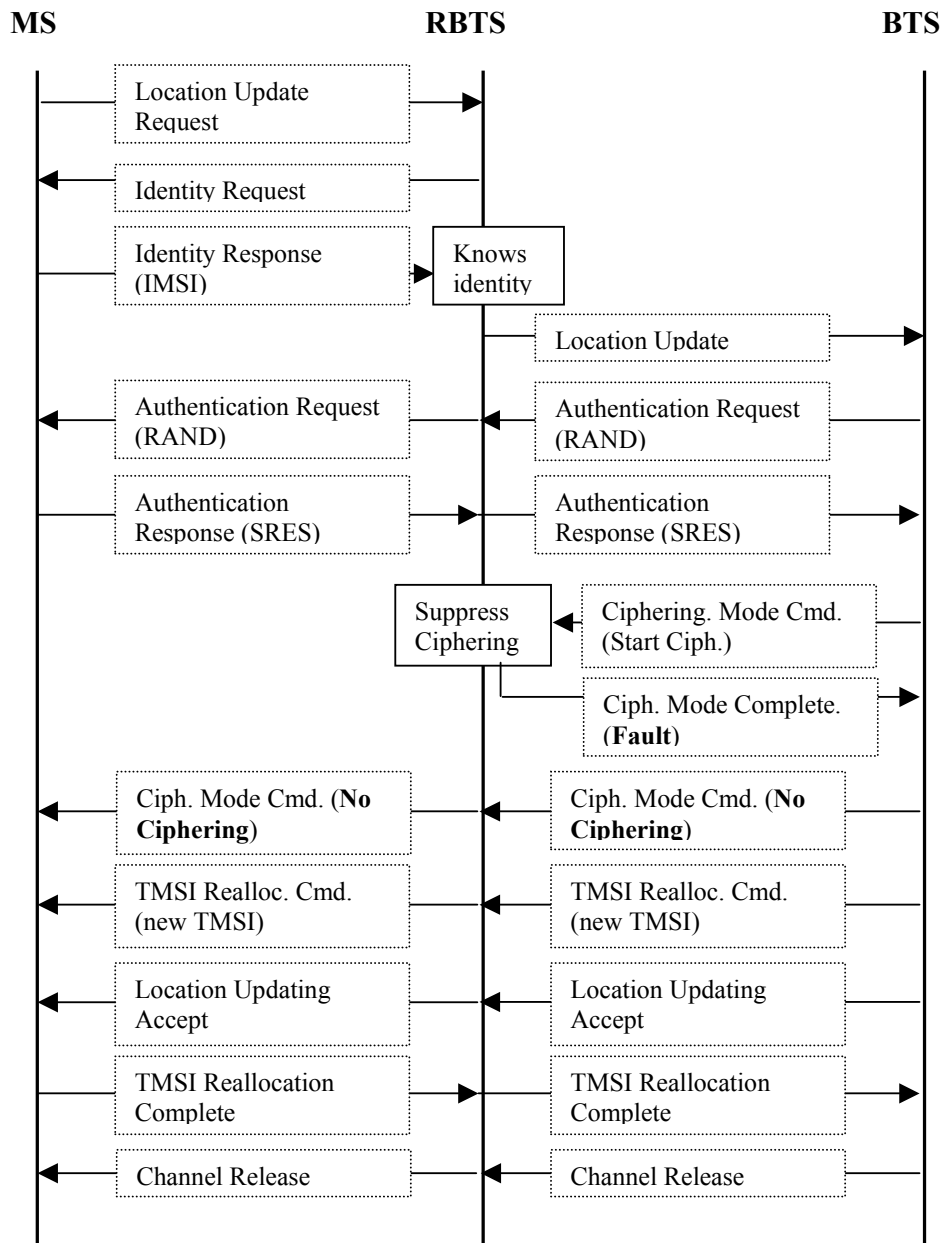


Figure 33 Principle of capturing GSM user identities

The reader may get confused about the Ciphering Mode Complete (**Fault**) message and may ask why it results in the network deciding on “No Ciphering” mode. This is simply because the ETSI specifications state that

Upon receipt of the CIPHERING MODE COMPLETE message or any other correct layer 2 frame which was sent in the new mode, the network starts transmission in the new mode. [22]

When this statement does not hold the network proceeds in the previous mode, that is no ciphering. A simple reason for this could be that operators fighting for market shares prioritise service on security. A customer whose call is transmitted in the clear does not even know about it (GSM users do not know what sort of encryption is used when they are calling), and may even not care. A customer not able to establish a call for “no reason” may go to another operator.

It should also be possible to query a subscriber, whose IMSI is known to the attacker, for his/her TMSI using the identification procedure.

When the attacker knows the IMSI/TMSI it is possible to locate a specific subscriber without having to continuously track him/her. The attacker simply pages the MS with the specific IMSI/TMSI.

6.3 Attacks on the Authentication Algorithm

Many GSM operators use the design specification given in the GSM MoU, COMP128, instead of designing their own algorithm for authentication and session key generation. The difficulty in starting using a different algorithm is because the algorithm resides inside the SIM, and subscribers who bought subscriptions (SIMs) before the eventual introduction of a different (presumably stronger) algorithm are forced to use their SIMs with the old algorithm. Other reasons may be the overhead costs associated with changing the software in databases etc. It is, however, possible to use more secure versions of COMP128¹¹ in new SIMs that are handed to new subscribers.

The design of COMP128 was never made public, but the design has been reverse engineered and cryptanalysed [16]. The author of this document could easily find the software implementation of COMP128 by a simple search on the Internet. Since the GSM specification for SIM cards is widely available, all that is needed to clone a SIM card is the 128 bit COMP128 secret key K_i and the IMSI which is embedded in the card.

COMP128 is basically a keyed hash function (see Section 3.2.3) that takes a 128 bit key K_i , and 128 bit of data RAND, to output a 96 bit hash value. The input RAND is the random challenge supplied by the BTS. The first 32 bits of the hash are used as the response SRES to the challenge and sent back to the network. The remaining 64 bits are used as the session key K_c for voice encryption using some version of the A5 algorithm.

¹¹ There exists newer more secure versions of COMP128, e.g. COMP128-2. There are claims that even COMP128-3 and COMP128-4 exist. [44]

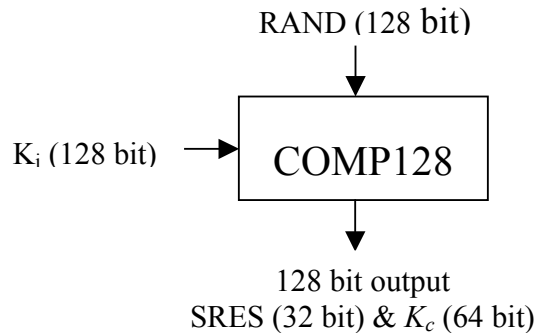


Figure 34 The A3/A8 implementation (COMP128)

By copying K_i and IMSI into an empty SIM (easy to buy) the attacker can authenticate himself to the network as the legitimate subscriber and thus call for free. The attacker can even, instead of using the subscription, use the captured key K_i for decrypting all the calls from and to the subscriber.

Cloning can be done either by having physical access to the SIM to be cloned, or over the air. The following subsections will examine the two cases:

6.3.1 Cloning with Physical Access to the SIM Module

If the attacker has *physical access* to the SIM module, several attacks can be launched in order to clone the module. Some of these attacks concentrate on using flows in the cryptographic algorithm resided on the smart card, whereas others use vulnerabilities in the smart card itself.

The most popular attack on GSM SIM modules attacks the cryptographic algorithm (COMP128) itself. It is a *chosen-challenge* attack and use flows in the hashing function to deduce the secret key K_i . The attacker forms a number of specially-chosen challenges and queries the SIM for each one; the SIM applies COMP128 to its secret key and the chosen challenge, returning a response back. By analysing the responses, the attacker is able to determine the value of the secret key K_i . The result of this attack is thus that the attacker gains access to the secret key K_i of the MS. The attack exploits a lack of diffusion, which means that some parts of the output hash depend only on some parts of the input to the algorithm [17]. Mounting this attack requires, apart from having *physical* access to the target SIM, an off-the-shelf smartcard reader [18], and a computer to direct the operation. The attack requires one to query the SIM about 150,000 times; An average SIM reader can issue 6.25 queries per second, so the whole attack takes approximately 8 hours. By overclocking the SIM or using a higher frequency oscillator on the SIM card reader the processing time could be reduced considerably. This increases however the risk of failure and damage to the original SIM.

This is, as already mentioned, the most common attack on GSM SIM modules, and the simple countermeasure against it is to change the hash function used for

authentication to a stronger one [41]. It should be noted that COMP128-2, a new version of the COMP128 algorithm, has remedied the issues present in the original, and as such requires more exotic methods of attack in order to gain K_i from newly issued SIMs. There are claims that even COMP128-3 and COMP128-4 exist. It is however not known to what extent these algorithms have been adapted by operators. [44]

Since the SIM module is implemented on a smart card, any discovery of a common vulnerability in smart cards immediately affects the security of the information stored in the SIM [39], e.g. IMSI and K_i . A relatively new class of attacks on smart cards called *optical fault induction*, revealed by Skorobogatov and Anderson [38], can be used to break the security of GSM SIM modules. Skorobogatov and Anderson discovered the attack after Skorobogatov found that he could interrupt the operation of the smart card's microprocessor by exposing it to an electronic camera flashbulb. They have carried them out using a camera flashgun (bought second-hand for \$30) and a microscope to extract secret information from smart cards. According to them, illumination of a target transistor causes it to conduct, thereby inducing a transient fault. They were able to expose the circuit to the light by scraping most of the protective coating from the surface of the microprocessor circuit embedded in each smart card. With more study, they were able to focus the flash on individual transistors within the chip by beaming the flash through a microscope; and then by sequentially changing the values of the transistors used to store data, they were able to reverse engineer the memory address map, which allowed them to extract the secret data from the smart card. Nevertheless the aforementioned asserted that they have developed a technology to block these attacks.

Another relevant smart card weakness was lately exposed in an IBM paper [17]. The IBM team launched a new class of side channel attacks¹² called *partitioning attacks*. The partitioning attack exploits vulnerabilities in the execution of COMP128 table lookups. They have launched a version of the attack on several implementations of COMP128 in various types of SIM cards. The results reveal that K_i can be recovered from a SIM card with less than 1000 invocations (challenges) with random inputs, or 255 chosen inputs. The number of challenges is reduced to only eight if the challenges are chosen adaptively. This means that if the attacker has possession of a SIM card for only a minute, K_i can easily be extracted. The new IBM approach seems to be a much more useful method than either breaking the cryptographic algorithms (COMP128) used by the SIM card or by intrusive attacks, such as the optical fault induction [43]. IBM has, however, been working with SIM providers in order to provide protection for new cards. The hardware required to perform this kind of attack is generally found in university laboratories rather than in people's garages and high street shops – at least for the moment.

¹² I.e. timing of operations, power consumption, electromagnetic emanations, etc.

6.3.2 Cloning over the Air

The attacker can even launch the attack over the air, using a rogue base station. Apart from the equipment (RBTS), the attacker needs to know the target IMSI or TMSI. (see Section 6.2 for information about how these parameters can be obtained). When these resources are available the attacker starts by capturing the/some MS. The captured MS(s) will immediately make a location update request which is conducted as described in Section 4.2.7.2. After the channel assignment is completed the attacker initiates an authentication process. Immediately after the attacker has a challenge-response pair he/she initiates a new authentication procedure. The MS is required to respond to every challenge made by the GSM network. This process continues until the attacker has got the required number of pairs to be able to initiate the cloning procedure.

The layer 3 trace depicted in Figure 23 gives us an idea of how much time the operation (challenging the MS) will take before the attacker has enough information to be able to clone the subscribers SIM. We assume that the channel establishment stage only has to be done once. The number of frames exchanged between the network and an MS, for one authentication process, are approximately 66 frames (4809 – 4743 see Figure 23). Since the duration of one TDMA frame is 4.610 ms, the duration of the whole signalling sequence is $4.615 \text{ ms/frame} \times 66 \text{ frames} = 0.30459 \text{ s}$. The time it takes to get the number of challenge-response pairs needed for the attack can be calculated. We know that the cryptographic attack requires approximately 150 000 challenge-response pairs. This means that the attack takes approximately 45 689 seconds ($150\,000 \text{ challenges} \times 0.30459 \text{ s}$), that is approximately 13 hours. This means that the MS has to be “available” to the attacker over the air for the whole time it takes to gather the information. This is quite unrealistic, because people use their mobiles e.g. to make calls or receive calls in addition to the fact that such a bombardment with challenges may cause the battery of the MS to run out, which would make the victim suspicious. To get around these problems the attack can be performed in parts; instead of performing an 13-hour attack, the attacker could interrogate the MS for 30 minutes every day. That way the battery would not run out and there would be a slighter risk of making the owner or the serving network suspicious. This attack can be mounted simultaneously against as many phones in radio range as the rogue base station has channels.

The cloning of a SIM not only makes it possible for the attacker to make money selling it, but even gives him/her the possibility to eavesdrop on the victim’s calls. Recall from Section 5.2 that the session key K_c that is used to encrypt user voice and data is derived by applying the A8 (or COMP128) algorithm to the secret key K_i and the challenge RAND. If the attacker has K_i (by cloning the SIM) and intercepts the RAND value over the air during the call establishment procedure, all requirements for calculating K_c are met (given that the algorithm used is COMP128 or some other known algorithm). The attacker can easily calculate K_c and listen to the whole conversation in real-time.

The defence against cloning over the air is to limit the number of times a SIM can be authenticated to a number significantly smaller than 150 000. The SIM locks up if the limit is exceeded. [44] The drawback with this solution is that a new SIM has to be issued and distributed to the subscriber, which involves costs both for the subscriber and the operator.

6.4 Attacks on the Confidentiality of GSM

As described in Section 5.2.3, the over-the-air privacy of GSM telephone conversations is protected using the A5 stream cipher. This algorithm has two main variants: A5/1 is the “strong” export-limited version used by CEPT-countries, and A5/2 is the “weak” version that has no export limitations. The approximate design of A5/1 was leaked in 1994, and the exact design of both A5/1 and A5/2 (briefly described in Section 5.2.3) was reverse engineered by Briceno from an actual GSM telephone in 1999 [20].

There have been some proposals on how to attack the confidentiality protection of GSM.A5. In the following subsections attacks are divided into brute-force attacks, cryptanalytical attacks, and non-cryptanalytical attacks.

6.4.1 Brute-Force Attacks

The confidentiality of GSM is protected by the secrecy of K_c . As mentioned earlier, K_c is 64 bits although we know that the last 10 bits are set to zero. This reduces the key space from 2^{64} to 2^{54} (1024 times smaller key space). According to information published by Bruce Schneier [25], A5/2 was developed with assistance from the NSA, and can be broken in real time with a work factor of approximately 2^{16} . A5/1, the stronger of the two variants, is however susceptible to attacks that can break it with a work factor of 2^{40} .

If we have a Pentium 4 chip with approximately 60 million transistors [33] and the implementation of one set of LFSRs (A5/1) would require about 2000 transistors [34], we could have a set of 30.000 parallel A5/1 implementations on one chip. If the chip was clocked to 3.2 GHz (a rather ambitious assumption) and each A5/1 implementation would generate one output bit for each clock cycle then we would need to generate 100+114+114 output bits, hence we could try approximately 10M keys per second per A5/1 implementation. A key space of 2^{54} would thus require about 18 hours, using all of the parallel implementations on the chip. If the attack in the average case succeeds after searching half of the key space, the key is found in about 9 hours. Further optimisation by e.g. giving up on a specific key after the first invalid key stream bit and distributing the computation between multiple chips will decrease the computation time by several magnitudes. This, still in the worst case, means several hours/many minutes of processing and is far away from a real-time attack. Bear in mind that the complexity of the attack is even greater due to the fact that it is quite difficult to determine when the key is found due to the nature of the plaintext.

To conclude, it is too difficult to succeed in a brute-force attack in real-time, but it is fully possible to find a key given a couple of hours. Entities with enough resources (computation power) can probably cut the processing time greatly.

Even though a brute-force attack may not be used as a real-time attack on the A5 algorithm, it could easily be used to find the key used in a specific conversation “offline”. The attacker intercepts and records the interesting conversation and decrypts it at a later time.

6.4.2 Cryptanalytical Attacks against GSM

In this section recent cryptographic attacks on the algorithms protecting the confidentiality of GSM calls are presented. The attacks will be explained along with the resources, both available data and computation power, they demand in order to break the security of the algorithms used for protection of confidentiality.

6.4.2.1 Real Time Cryptanalysis of A5/1 on a PC

A Biryukov, A Shamir, and D Wagner present two cryptanalytic attacks on A5/1, in which a single PC can extract the conversation key K_c in real time from a small amount of generated output. The attacks are related, but each one of them optimises a different parameter: The first attack (called the Biased Birthday attack) requires two minutes of data (known key stream) and one second of processing time, whereas the second attack, called the random subgraph attack, requires two seconds of data and several minutes of processing time. The authors claim there are many possible choices of tradeoff parameters in these attacks. Three of them are summarised in Table 3.

Attack Type	Pre-processing steps	Known plaintext	Number of 73GB disks	Attack time
Biased Birthday attack (1)	2^{42}	2 min	4	1 sec
Biased Birthday attack (2)	2^{48}	2 min	2	1 sec
Random Subgraph attack	2^{48}	2 sec	4	Minutes

Table 3 Three possible tradeoff points in the attacks on A5/1 [8]

The main idea in the Biased Birthday attack is to try to find out the initial internal state of the algorithm, assuming that the attacker has complete knowledge of the outputs of the A5/1 algorithm (the generated key stream) during some initial period of the conversation (the “Known plaintext” column in table 3). Since GSM terminals send a new frame every 4.615 milliseconds, each second of conversation contains about 2^8 frames. When the attacker has found the initial state of any frame, running the algorithm in the reverse direction and with knowledge of the (publicly known) frame number, the session key K_c can be derived.

The attack requires a huge pre-computation step (see Table 3). During precomputation a large set A of precomputed states is stored on a hard disk along with their output prefixes. A5/1 has a relatively small number of internal states,

since it has $n = 2^{64}$ states defined by the $19+22+23 = 64$ bits in its three shift registers. The basic idea is to keep a large set A of precomputed states on a hard disk, and to consider the large set B of states through which the algorithm progresses during the actual generation of output bits (the output prefixes). Any intersection between A and B will make it possible to identify an actual state of the algorithm from stored information.

The output prefix is the first $\log(n)$ bits in a state's output sequence. The pairs (prefix, state) are sorted into increasing prefix values, thus allowing the use of the prefix as an index for efficient disk probing. Given actual outputs of the A5/1 algorithm, extract all their (partially overlapping) prefixes, and define B as the set of their corresponding (unknown) states. Searching for common states in A and B can then be efficiently done by probing the sorted data A on the hard disk with prefix queries from B .

The attack requires, as mentioned earlier, that the attacker knows some pseudo random bits generated by A5/1 in some of the frames. The problem is how to get access to these bits. Eavesdropping a conversation is quite "easily" done using a scanner or a modified MS, but the conversation is encrypted and the intercepted data will not reveal any of the pseudo random bits used in the encryption process (at least not as many bits as required for the attack to be successful). The attacker can not derive the needed pseudo random bits and the attack fails. Depending on the amount of known plaintext needed, the attacker could try to guess bits that are needed. The problem is that the space and time complexity would grow very rapidly making the attack impractical and it would definitely not be a real-time attack.

The minimum known plaintext requirement is two seconds. This is the same as requiring that the precise sequence of bits in approximately 433 encrypted consecutive frames be known. The attacker may be able to derive this amount of the key stream if he/she is able to mount a man-in-the-middle attack by asking the MS, after encryption is enabled, to respond to certain signalling requests that yield responses with content that the attacker can guess with high probability. If we assume that the attacker succeeds in deriving the required amount of the key stream then we can move to the next step in the attack; that is calculating the session key itself. Using this attack this step takes several minutes in the best case. This is however too much time to satisfy real-time requirement, especially if it is compared to the Biased Birthday attack which is able to find the key within a second. The problem with the Biased Birthday attack is the amount of known plaintext required. It is very unrealistic that 2 minutes of the key stream will be available to the attacker.

Thus the conclusion is that these attacks are fine pieces of cryptanalytic work, which may have theoretical value but in reality would probably not be practical.

6.4.2.2 Other Attacks on A5/1

Barkan E, Biham E and Keller N present in [35] cryptanalysis of both A5/1 and A5/2. The methods and ideas described in the paper are patent pending and are

therefore not presented here, however the resources needed in order to mount the attacks are examined.

They start off with a known plaintext attack on A5/2. The attack is then converted to a ciphertext only attack on A5/2. This attack requires, according to the paper, only a few dozen milliseconds of *encrypted conversation* data. The average time complexity is approximately 2^{16} dot products, the memory complexity is about $2^{28.8}$ bytes (less than 500 MBs), and the pre-computation time complexity is about 2^{47} bit-XORs. The authors' implementation of the attack on a personal computer (taking advantage of their 32-bit XOR) recovers the session key in less than a second, and it takes about 320 minutes to complete the one-time pre-computation. The attack on A5/2 is extended to a more complex ciphertext-only attack on A5/1.: The attack against A5/1 is more complex and requires much more resources. Table 4 summarises the needed resources.

Available data (Ciphertext)	Pre- processing steps	Number of PCs to complete pre- processing in one year	Number of 200 GB disks	T	Number of PCs to complete attack in real- time
2^{12} (appr 5 min)	2^{52}	140	22	2^{28}	1
$2^{6.7}$ (appr 8 sec)	2^{41}	5000	176	$2^{32.6}$	1000
$2^{6.7}$ (appr 8 sec)	2^{42}	5000	350	$2^{30.6}$	200
2^{14} (appr 20 min)	2^{35}	35	3	2^{30}	1

Table 4 Three possible tradeoff points in the attacks on A5/1 [35]

T is the number of calculations during the real-time phase of the attack and 2^{20} such calculations are assumed doable every second on a modern personal computer. The authors even claim that using their ideas in this attack they have been able to make the attack on A5/2 proposed by Goldberg, Wagner and Green (se Section 6.4.2.3) a ciphertext only attack!

Another attack on A5/1, which is even called *Another attack on A5/1*, was proposed by Ekdahl P and Johansson T [36]. This attack, in contrast to the previous attacks, is not a time-memory trade-off attack, but uses ideas from correlation attacks. It exploits the bad key initialisation in A5/1, the fact that the key and the frame counter are initialised in linear fashion. This “bad property” makes it possible to launch a type of correlation attach which makes it possible to separate the session key from the frame number in binary linear expressions.

The complexity of the previous attacks increases exponentially when the length of the LFSRs increases, due to the fact that the number of initial states increases exponentially. This new attack is however only linear in the length of the shift registers and depends instead on the number of irregular clockings before the keystream is produced.

This is a known plaintext attack. The implemented attack requires the 40 first bits from about 2^{16} (possible non-consecutive) frames, which corresponds to about five minutes of GSM conversation. The complexity of the attack is quite low and

it requires very little pre-computation time and memory. In the presence of the required known plaintext the attack takes about five minutes on a modern home PC. This makes it hard to use the attack in real-time. It is however fully possible to record the conversation and decrypt in at later time.

6.4.2.3 Attacking A5/2

Back in 1999 Goldberg, Wagner, and Green presented the first attack on A5/2. It took them less than a day to crack the algorithm. Information about the methods used was never published, therefore the information available on homepages of the authors has been used.

The time complexity of the attack is very low. The session key is found within milliseconds and demonstrates that A5/2 provides weak security. A problem with this attack is that it requires the knowledge of the XOR of the key stream used for encryption of two frames that are exactly 2^{11} frames apart (approximately six seconds apart). If this knowledge is provided the key can be found in approximately 10 milliseconds [9, 20]. This is a softer requirement than the attack on A5/1 in the pervious section since it requires much less known plaintext. This in addition to the claim made by the authors of [35] that they have optimised the attack to a ciphertext-only attack makes this attack more practical than the previous one. (It remains a problem how to obtain the required amount of known plaintext, especially as it has to be exactly 2^{11} frames apart, if the ciphertext-only optimisation is not possible to implement)

6.4.3 Attacks Using Loopholes in the Protocols

In this section several attacks that result in the attacker being able to eavesdrop on GSM users' conversations will be presented. One of the attacks assumes that the attack on A5/2 (see Section 6.4.2.3) is practical. The rest of the attacks use flaws or weaknesses in the GSM architecture and/or flaws in the protocols used in the communication between the GSM networks and the subscribers.

The following facts will be used in the coming attacks:

- It is common knowledge that most GSM mobile phones can communicate with most different base stations and networks. This is possible because all of the different manufacturers follow the specifications and standards of how GSM should function. These specifications are developed by the European Telecommunications Standards Institute (ETSI). It is possible to study the specifications on how the communication between the network and the MS is conducted, and get detailed information on the communication protocols and mechanisms (Layer 3 messages) used e g when a MS is to be authenticated by the network.
- The same key K_c is used for the different encryption algorithms A5/1, A5/2, and A5/3 (A5/3 is, as mentioned before, an algorithm even stronger than A5/1). This means that breaking one of this three algorithms and

retrieving the session key threatens the confidentiality of the conversation even when the stronger versions of the algorithm are used later.

- A base station does not need to authenticate itself to the MS it is communicating with. Furthermore messages are not authenticated and their integrity is not protected.
- [20] states that it is mandatory for A5/1, A5/2 and A5/0 (non encrypted mode) to be implemented on mobile stations. The reason is to make roaming between different networks and operators (that potentially uses different encryption algorithms) possible.

Implementation of other variations of the A5 algorithm (Figure 35), than A5/1 and A5/2 are optional.

algorithm identifier		
If SC=1 then:		
bits		
4 3 2		
0 0 0	cipher with algorithm A5/1	
0 0 1	cipher with algorithm A5/2	
0 1 0	cipher with algorithm A5/3	
0 1 1	cipher with algorithm A5/4	
1 0 0	cipher with algorithm A5/5	
1 0 1	cipher with algorithm A5/6	
1 1 0	cipher with algorithm A5/7	
1 1 1	reserved	
If SC=0 then bits 4, 3 and 2 are spare and set to "0"		
SC (octet 1)		
Bit		
1		
0	No ciphering	
1	Start ciphering	

Figure 35 The coding of the different A5 algorithms [22]

The rest of this section will present several attacks on GSM that will result in compromised confidentiality.

Attack Scenario 1 – Cryptanalysing A5/2 to Find the Session Key

Recall that the confidentiality of GSM speech depends on the secrecy of the session key K_c . An attacker that has discovered the session key is able to decrypt and listen to the conversation. Assuming A5/2 is a weak cipher, much easier to break than A5/1 and A5/3 [7, 8, 20], it is much more practical to try to derive the session key K_c by attacking A5/2. When the session key is retrieved, the attacker is able to decrypt and listen to the conversation even if it is encrypted using A5/1 or A5/3.

The attack is mounted as follows:

The victim MS has been captured. The MS listens and responds to the signal with greatest strength. Recall from Section 4.2 that the MS finds the BTS by tuning to the frequency with the highest power density. Thus a mobile base station residing between the MS and the legitimate BTS, able to provide a base channel with higher power density than the legitimate BTS, will force the MS to synchronise with it and read the system information it provides on its BCCH. This procedure makes it possible for the RBTS to mount a man-in-the-middle attack. Of course for this to be possible the attacker has to have access to the equipment (hardware and software implementation of some of the channels and protocols used in the communication between a MS and a BTS) needed for the communication, according to the ETSI specifications in addition to the implementation of the cryptanalytical attack against A5/2.

After capturing the MS, the attacker waits until the MS requests a service, e.g. making a call. Then the RBTS impersonates the network to the calling subscriber, and the subscriber to the network until the session key is found (Figure 38). We know that, in the standard case, before a conversation starts the network requires the caller to be authenticated (see Section 4.2.7.3). The network sends RAND to the attacker. The attacker forwards RAND to the subscriber, who computes SRES and returns it to the attacker believing that the attacker is the network. Recall that the attacker is not able to compute SRES even if RAND (which is sent in the clear) is known, because the attacker neither knows the exact authentication algorithm (if other than COMP128) nor the secret key K_i . The attacker relays the signalling messages between the network and the MS until SRES is received from the MS. Instead of forwarding SRES to the network, he/she transmits the CIPHERING MODE COMMAND message to the caller MS (Figure 36) asking it to encrypt using A5/2 (Figure 37).

IEI	Information element	Type / Reference	Presence	Format	length
	RR management Protocol Discriminator	Protocol Discriminator 10.2	M	V	1/2
	Skip Indicator	Skip Indicator 10.3.1	M	V	1/2
	Cipher Mode Command Message Type	Message Type 10.4	M	V	1
	Ciphering Mode Setting	Cipher Mode Setting 10.5.2.9	M	V	1/2
	Cipher Response	Cipher Response 10.5.2.10	M	V	1/2

Figure 36 CIPHERING MODE COMMAND message content

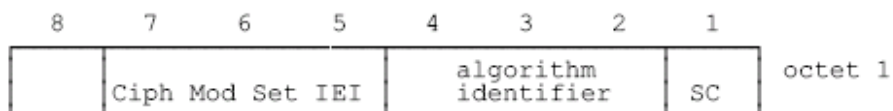


Figure 37 CIPHER MODE SETTING information element

The victim MS, who now believes it is authenticated by the network, starts encryption using A5/2 and responds, according to the protocol, with an encrypted acknowledgement of the ciphered mode.

Now the attacker can use cryptanalysis of A5/2 to retrieve the encryption key K_e using the data he/she has received from the subscriber. When the key is retrieved (within 6s +10 ms or by possibly using the improved ciphertext-only attack on A5/2), the attacker sends the authentication information (SRES) to the waiting network. It is important that the attacker succeeds in finding the session key before the assigned timer expires (timers are used to ensure that responses are received within defined time intervals so that the network will not wait forever for an answer), because if the timer expires the network may become suspicious and abort the transaction. Now when the network receives the SRES, it continues with the rest of the call establishment procedure, until the call is established.

The attacker now has the session key that will be used to generate the key stream. He/she examines the message exchange between the network and the MS to determine which encryption algorithm the network will ask the MS to use, in order to engage the right encryption algorithm when relaying messages between the communicating parties. Even if the network now tells the MS to use A5/1 the attacker is able to decrypt it since no matter which of the A5 algorithms is chosen, the same session key will be used as input to that algorithm, together with the frame number (F_n) of each transferred frame. Recall that the F_n is publicly known. Now the attacker receives frames from the MS, encrypted using A5/2. These frames are decrypted and saved by the attacker, then the attacker encrypts them using A5/1 and relays them to the network. In the other direction the attacker receives frames encrypted using A5/1; these are decrypted, saved and then encrypted using A5/2 and sent to the MS.

This attack is useful in illustrating the loopholes in the protocol and how the man-in-the-middle attack works, it may however not be successful in practice because of the known plaintext requirement (see Section 6.4.2.3). It is however possible to obtain the required amount of known plaintext if the attacker can make the MS send signalling messages with content that is known or almost known to the attacker that are on the required distance from each other. Another potential reason for failure may also be that cryptanalysing A5/2 may take too much time resulting in the expiring of the assigned timer, which will result in the network denying service to the legitimate MS. An implementation of the ciphertext-only attack on A5/2 presented in [35] would however find the session key without the need of known plaintext.

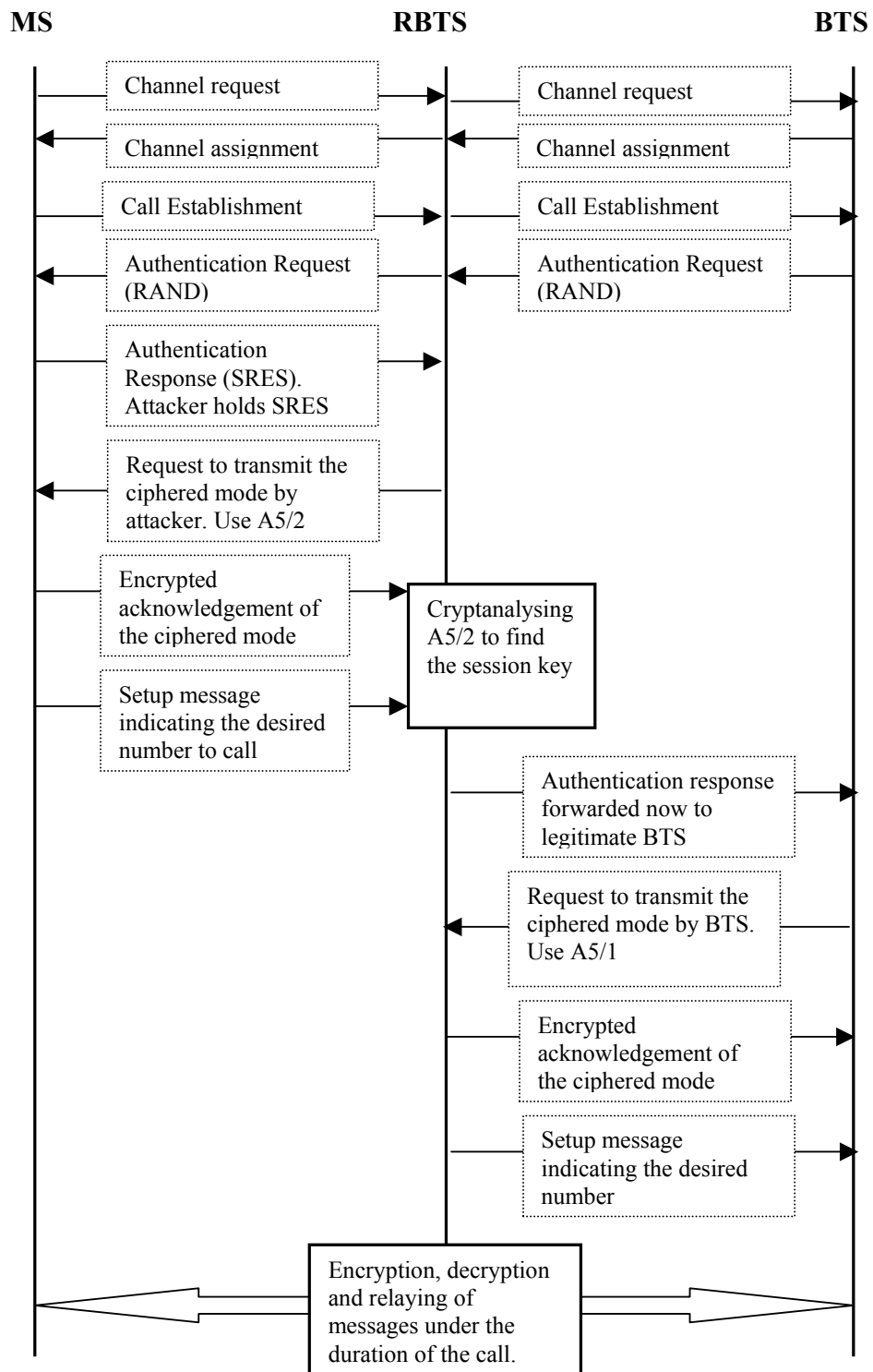


Figure 38 Breaking A5/2 and eavesdropping in real-time

An attacker capable of mounting a man-in-the-middle attack can listen to the conversation even without cryptanalysing the encryption algorithm being used. This can be achieved by using one of the two following attacks:

Attack Scenario 2 – Using the Attacker’s SIM

This attack is based, as in the previous attack, on the attacker’s capability of impersonating a MS and a BTS and results in the attacker being able to eavesdrop on the caller’s conversation. The attacker uses his own SIM when communicating with the network on behalf of the subscriber. The attacker simply uses a compromised K_i together with its IMSI. It is important that the attacker ensures that the rightful owner of the parameters is not active in the same area to avoid the network getting suspicious.

The attacker (RBTS) assures that the MS to eavesdrop on, whose IMSI/TMSI is assumed to be known, is captured and all messages/service requests that go between the MS and the network are relayed by the attacker. Once the attacker detects that the captured MS is demanding service (e.g. a call-setup) the attack starts. In the (important) case the victim MS demands that a call will be set, the RBTS (acting, to the network, as a legitimate customer with a valid SIM) makes the exact same request to the network using *his/her* stolen IMSI/TMSI (Figure 39).

The attacker assigns the victim MS a channel and authenticates it exactly the same way the network would. The difference is that the SRES from the MS is useless for the attacker and is discarded. Next the attacker asks the MS to use no encryption (A5/0) and the MS sends to the RBTS the call-setup message with the number it wants to call. Now the attacker has the number and can initiate a call-setup procedure with the network on his/her behalf to the number that the legitimate subscriber wishes to call, using the *attacker’s* identity. The network suspects nothing and uses the K_i of the attacker in authentication and session key generation. This way the attacker will have the session key and will be able to decrypt the call. The BTS demands the RBTS (which is acting as a MS demanding service) to encrypt the communication and RBTS does so. Now the call is transmitted unencrypted from the MS to the RBTS. The attacker, who operates the RBTS, eavesdrops on the call, encrypts it with the session key obtained from the RAND received from the BTS, and transmits it to the BTS. The unencrypted call from the victim MS to the RBTS is connected to the encrypted call from the RBTS to the BTS, so it seems to the legitimate participants that they have the call they requested. And since the call between the RBTS and the BTS is an encrypted genuine call, the network does not see that anything is wrong. Figure 39 shows the steps to be followed in order to mount this attack.

One effect of this attack is that the call is made on the attacker’s subscription and not that of the MS meaning that the attack can be detected later if an itemised bill is checked. Another effect is that the MS who receives the call will see the number of the attacker’s subscription and not that of the legitimate calling party in the display of the terminal when the call is connected. This problem can

however easily be solved by assuring that the number of the attacker's subscription is secret. This way, the MS receiving the call will not see any number in the display. It will only view a message that the calling number is not available.

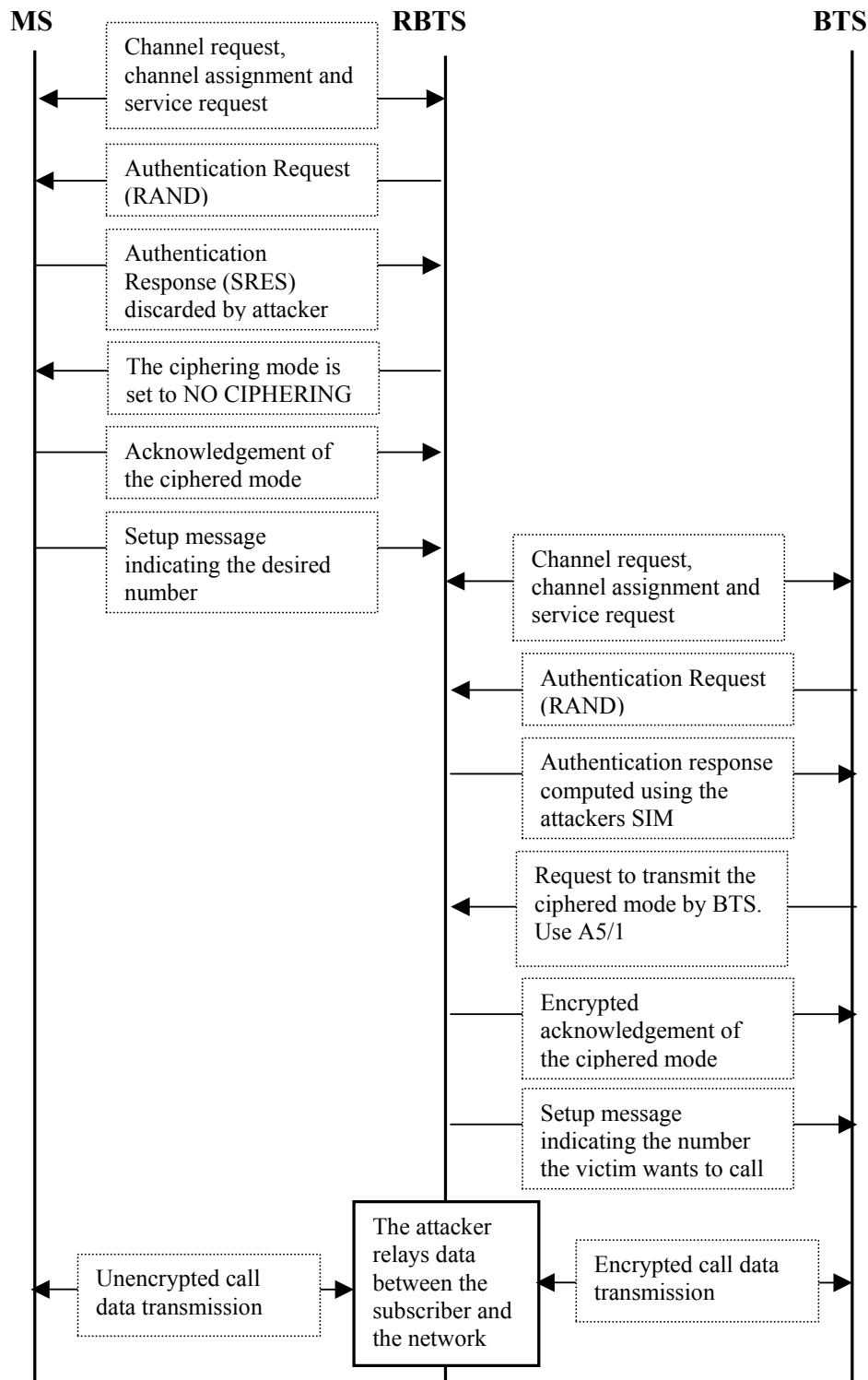


Figure 39 Eavesdropping using the attackers SIM

Attack Scenario 3 – Using the Classmark Information:

This attack makes use of the classmark information that the MS sends to the network to inform about, among other things, its ciphering capabilities. The goal of the attack is to make the MS send a message to the network, indicating that it can only encrypt and decrypt using A5/2 and A5/0 (no encryption). Later when a call is made, with this specific MS involved, the encryption method chosen will be one of these two (preferably A5/0). In the case of A5/0, no encryption will be used, meaning it is straightforward to eavesdrop on the conversation. In the case of A5/2 the attacker has an implementation of one of the attacks against A5/2 and is able to cryptanalyse the encryption.

The attack is active in the start. The attacker functions as a repeater between the MS and the BTS. When signalling messages containing classmark IE is observed, the attacker modifies it in such way that the network thinks that this specific MS has no encryption capabilities or only A5/2. When this is done the active part of the attack is done and the attacker goes over to passive monitoring of the traffic. Every message the specific MS sends and receives will unencrypted or possible to decrypt.

The classmark information is transferred on several occasions, e.g. when the MS performs location updating or when the network pages the MS for an incoming call. Figure 40 shows the content of the PAGING RESPONSE message, that is sent by the MS in response to a paging request from the network. It includes four octets of classmark information. The content of the classmark information is shown in figure 41. Since GSM does not provide integrity control of the transmitted messages, the attacker is able to alter the classmark IE. The attacker only has to change a few bits of information, thereby making the network think that the victim MS is not able to use A5/1 or A5/2.

IEI	Information element	Type / Reference	Presence	Format	length
	RR management Protocol Discriminator	Protocol Discriminator 10.2	M	V	1/2
	Skip Indicator	Skip Indicator 10.3.1	M	V	1/2
	Paging Response Message Type	Message Type 10.4	M	V	1
	Ciphering Key Sequence Number	Ciphering Key Sequence Number 10.5.1.2	M	V	1/2
	Spare Half Octet	Spare Half Octet 10.5.1.8	M	V	1/2
	Mobile Station Classmark	Mobile Station Classmark 2 10.5.1.6	M	LV	4
	Mobile Identity	Mobile Identity 10.5.1.4	M	LV	2-9

Figure 41 PAGING RESPONSE message content [23]

Using location updating and paging responses for changing the classmark information means that the attacker has to wait until these messages are used in order to modify the content (or actively trigger these messages). It is however

possible for the attacker to force the MS to send its classmark information to the network. This can be done using the CLASSMARK ENQUIRY message.

IEI	Information element	Type / Reference	Presence	Format	length
	RR management Protocol Discriminator	Protocol Discriminator 10.2	M	V	1/2
	Skip Indicator	Skip Indicator 10.3.1	M	V	1/2
	Classmark Enquiry Message Type	Message Type 10.4	M	V	1

Figure 41 CLASSMARK ENQUIRY message content [23]

The attacker, impersonating the network to the MS, asks the MS to inform about its capabilities by sending it a CLASSMARK ENQUIRY message (Figure 41). The answer from the MS will be a CLASSMARK CHANGE message (Figure 42)

IEI	Information element	Type / Reference	Presence	Format	length
	RR management Protocol Discriminator	Protocol Discriminator 10.2	M	V	1/2
	Skip Indicator	Skip Indicator 10.3.1	M	V	1/2
	Classmark Change Message Type	Message Type 10.4	M	V	1
	Mobile Station Classmark	Mobile Station Classmark 2 10.5.1.6	M	LV	4
20	Additional Mobile Station Classmark Information	Mobile Station Classmark 3 10.5.1.7	C	TLV	3-14

Figure 42 CLASSMARK CHANGE message content [23]

containing a mobile station classmark 2 information element. The CLASSMARK CHANGE message is used by the MS either to answer a CLASSMARK ENQUIRY message it has received from the network or to inform the network about changed conditions. Upon reception of the CLASSMARK CHANGE message the attacker alters the parts dealing with the MS's encryption capabilities, in the Classmark 2 IE (Figure 41), in such a way that the network thinks that A5/1 and A5/2 are not available to the MS. This is achieved by altering the bits used to indicate encryption algorithm availability, for A5/1, A5/2 and A5/3 to be 0. 1 indicates that the algorithm is available for use.

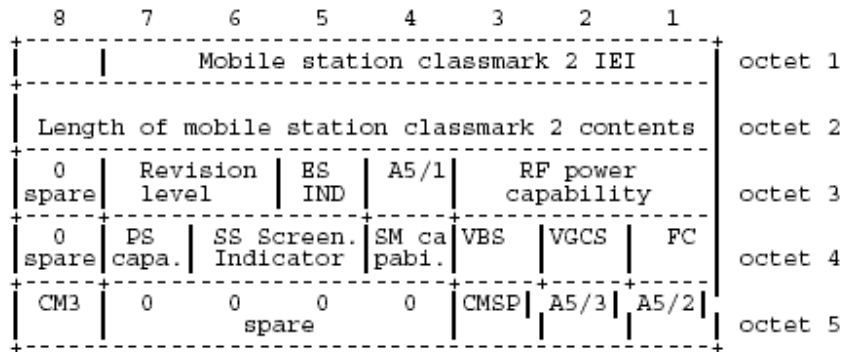


Figure 43 Mobile station Classmark 2 information element [23]

Now the attacker relays the altered classmark information to the network, which believes that the MS is trying to inform about its capabilities. The network does not know that this message is an altered response to a CLASSMARK ENQUIRY message that the MS received from RBTS. From now on all the messages transferred between the MS and the BTS (both signalling and traffic) will be unencrypted and can be eavesdropped by the attacker. The attacker aborts the active attack and goes to passive monitoring of the information flow. Now the call will be set up with encryption disabled, enabling the attacker to listen to the conversation in real-time.

Attack Scenario 4 – Accessing the Signalling Network:

All of the attacks until now have concentrated on the radio link between the MSs requesting a service and the BTS which provides the link to the network. The radio link is, however, not the only vulnerable point in the GSM system. There is a possibility in which the attacker does not even need to cryptanalyse the A5 algorithm used for encryption to disclose the information transmitted between GSM entities. GSM does not provide end-to-end protection of the transmitted information. The transmission, of both signalling and user data, is encrypted only between the MS and the BTS. After the BTS (often) the traffic is transmitted in plaintext within the operator's network [7, 10].

This opens up new possibilities. If the attacker can access the operator's signalling network, he/she will be able to listen to everything that is transmitted, including the actual phone call as well as the RAND, SRES and, K_c . The SS7 signalling network used in the operator's GSM network is completely insecure if the attacker gains direct access to it. Having access to the unencrypted call and data makes the eavesdropping very convenient for the attacker.

Accessing the signalling network is not very difficult. Although the BTSs are usually connected to the BSC through a cable, some of them are connected to the BSC through a microwave or even a satellite link. This link would be relatively easy to access with the right kind of equipment. The microwave link might be encrypted, however, depending on the hardware manufacturer, thus making it

slightly more difficult to monitor. It is, however, really a question of whether the attacker wants to crack the A5 encryption protecting the session of a specific MS or the encryption between the BTS and the BSC and thus gaining access to the backbone network. The possibility of accessing the cable leaving the BTS should not be ruled out either. To illustrate how an attack on the network residing behind the BTS can be mounted, we give an example.

An illustration of a common situation in many GSM networks is depicted in Figure 44. The connection between BTS and the rest of the network is simply an unencrypted microwave link.

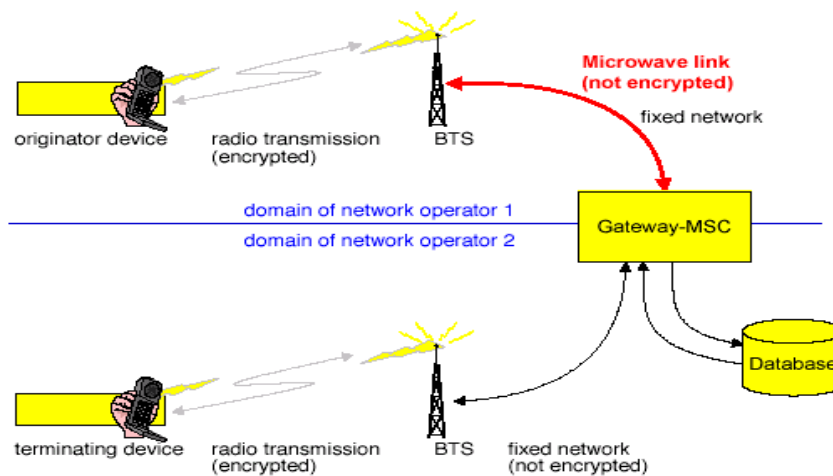


Figure 44 Internal links of a GSM network [41]

Assume the attacker wants to make a telephone call for free (or eavesdrop on a conversation). He/She has in a previous attack (see Section 6.2) captured the user parameters (IMSI, TMSI) of other legitimate subscriber(s). The attacker establishes a communication channel with the BTS and requests to make a call. The BTS will respond with an authentication request (RAND). Prior to challenging the attacker the BTS has contacted the VLR to get a triplet (RAND, SRES, K_c) to use in authenticating the subscriber requesting a service (Figure 45). The VLR contacts the HLR of the subscriber (which is the only entity knowing the K_i of the subscribers) to get the information. HLR asks AuC to calculate a set of triplets. Usually five triplets are returned to the VLR to be used in future requests from the same subscriber. This is done to reduce the traffic to HLR. Recall from Section 5.1.6 that the AuC, which sometimes is implemented as an integral part of the HLR, is responsible for the generation of these triplets to be used in authenticating subscribers before service is provided. Now five triplets are generated and transmitted from the HLR to the BTS via the VLR, MSC, and BSC, over the A- and the Abis-interfaces (see Figure 24). We assume that at least one of these interfaces is an unencrypted microwave link. In the same time that the attacker is communicating with the BTS he/she even eavesdrops on the microwave link. Since the triplets are transmitted in the clear it is easy for the

attacker with a microwave scanner to intercept the triplets. The attacker has the triplets before the BTS. Now the BTS challenges the attacker with the RAND. The attacker retrieves the SRES and is able to respond correctly to the issued challenge. Now the attacker is authenticated and the call is established. The BTS does not suspect anything and asks the attacker to start encryption. The attacker, who knows K_c can easily perform encryption and decryption. The rightful owner of the TMSI/IMSI that the attacker used in establishing the link with the serving network will be billed for the call. The four unused triplets can be used in the future when the attacker wishes to request services from the network.

This might be a very real threat and an attack could go undetected for a long time, if implemented carefully. The ability to tap into the data transmitted between the BTS and BSC even enables the attacker to eavesdrop on calls without worrying about cryptanalysing the encryption. Figure 45 below outlines the attack.

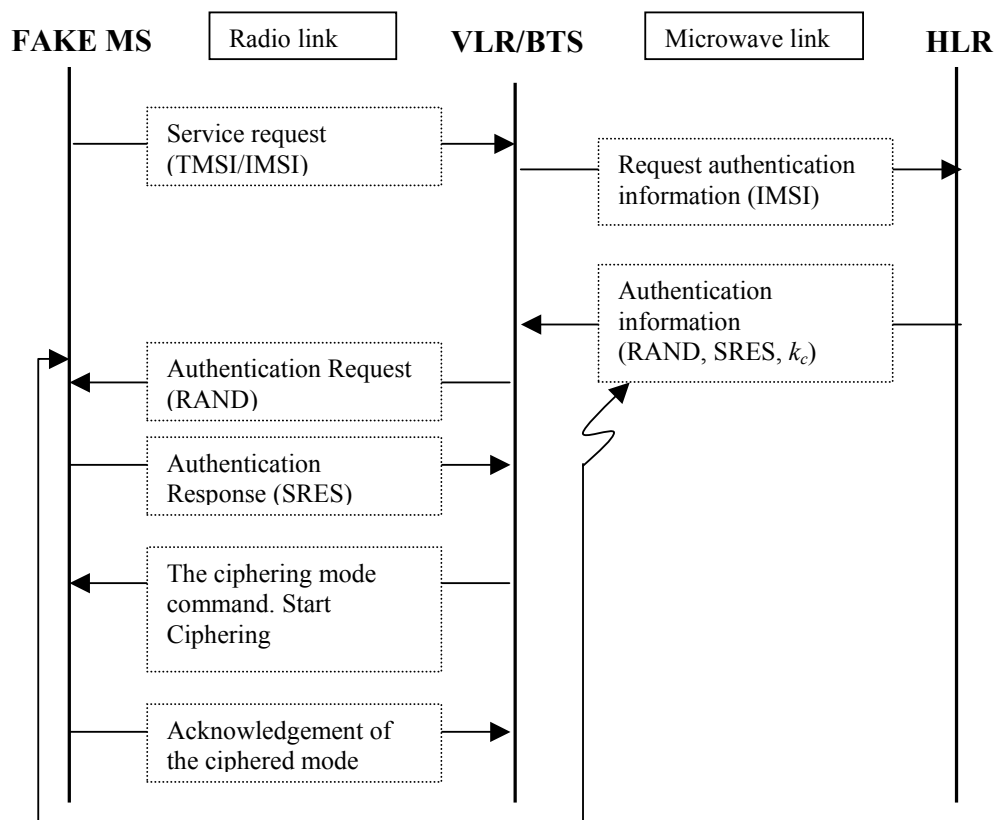


Figure 45 Capturing authentication triplets and calling for free.

In another scenario, the attacker could attack the HLR of a particular network. If the attacker can access the HLR, he/she will be able to retrieve the K_i s for all the subscribers of that particular network. Luckily the HLR is usually a bit more secure than the rest of the network, thus making it a slightly less probable point of

entry, yet not completely improbable either keeping in mind the potential gain involved. [45]

Other Scenarios:

Another possibility for the attacker is to intercept and record the encrypted conversations and try, at a later time (off-line), to retrieve the session key used to encrypt each of the conversations. What the attacker needs in order to accomplish this, apart from the TMSI/IMSI of the subscriber that is used to identify the calls to be recorded, is to know the RAND value that was used in the authentication session prior to each conversation. Recall from Section 5.2.2 that many operators use a single algorithm to implement A3/A8 (COMP128), and that the output of that algorithm forms the response to the challenge from the network and the session key used in encryption of data and voice. Thus challenging the MS with the same RAND value as one used in previous authentication yields the same encryption key. Hence if the attacker has recorded conversations along with their respective RAND values, then it would be possible to challenge the MS and ask it to start encrypting using A5/2. Cryptanalysing the encrypted transmission would give the session key used for that particular RAND value and the conversation associated with it. Since GSM uses the same session key format for all encryption algorithms it is possible to obtain the key stream used in encrypting the former call even if it was encrypted using a different algorithm than A5/2 (e.g. A5/1). Now the attacker is able to decrypt and listen to the recorded call.

A second interesting scenario is a subscriber on a trip in a foreign country. The subscriber wants to make a call and a call establishment process starts. The serving VLR in the foreign network does not have the K_i of the subscriber so authentication information is requested from the home network of the subscriber (HLR). A set consisting of up to five triplets of (RAND, SRES, K_c) are sent to the hosting network's VLR, and the VLR authenticates the visiting subscriber. What guarantees has the subscriber that the hosting network, or personnel administrating the concerned databases do not sell that information or even use it for personal purposes? This scenario is of course possible in all countries, but should be more

Another approach to breaking the security of GSM goes, like many other systems, through *social engineering*. This approach should not be underestimated although it sounds ludicrous. The attacker might pretend to be a repair man or such, enter a suitable building and install a wire tap. He/she might also bribe an engineer to do it for him or to give him all the K_i s for all the subscribers of that particular operator. The possibilities are countless and real.

6.5 Attacks on the Equipment Protection Mechanism

As mentioned in Section 5.1.8, the GSM terminals are protected against theft through the use of the IMEI. The IMEI uniquely identifies each terminal and makes it possible for the operators to block stolen terminals. The robustness of

this mechanism depends entirely on the difficulty of changing a terminal's IMEI. This has however proven to be quite easily done. There exist several off the shelf devices that can be used for changing the IMEI and thereby making the stolen terminal fully legal again, easily sold for half the price of a legitimate terminal. There even exists software which can be downloaded from the Internet that can be used to change the IMEI of a terminal. The terminal is connected to a PC, where the software is installed and the software does the job.

6.6 Denial of Service (DoS) Attacks

DoS attacks can be performed by physically disturbing radio signals or by logical means. These two possibilities will be explained further in the two sections below.

6.6.1 Denial of Service – Physical Intervention

The physical attacks are the most straight forward attacks. The attacker prevents user or signalling traffic from being transmitted on any system interface, whether wired or wireless, by physical means. An example of physical intervention on a wired interface is wire cutting. The attacker could for example cut the wire leaving a base station. An example of physical intervention on a wireless interface is jamming. Having the equipment that jam GSM radio signals is sufficient. The equipment is placed in the area where traffic is to be disturbed and the GSM equipment within the device's range will not function properly. Note that the frequency hopping makes the jamming more difficult than usual.

There are examples of jamming causing problems for GSM operators. Recently a GSM operator in Moldova suffered heavily from jamming activities that effectively caused a drop rate of lost calls of about 7 %. The operator and the authorities had major problems in stopping the attacks. [46]

6.6.2 Denial of Service – Logical Intervention

An attacker can perform DoS attacks by logical means also as the following examples show:

- The attacker spoofs a *de-registration request* (IMSI-detach) to the network. The network de-registers the subscriber from the visited location area and instructs the HLR to do the same. The user is subsequently unreachable for other subscribers. The attacker needs a modified MS and the IMSI of the user to de-register.
- The attacker spoofs a *location update request* in a different location area from the one in which the subscriber is roaming. The network registers the subscriber in the new location area and the target user will be paged in that new area. The user is subsequently unreachable for mobile terminated services. [47]

- An attacker in possession of a modified base station, transmitting the base channel with higher signal strength will force the MSs in the area to camp on the radio channels of the false base station, making them unreachable for the serving network. [47]

In this chapter we will examine whether GSM is suitable to be used by entities with higher security requirements than private persons, e.g. the military and emergency service. [48] and [49] give a more thorough description of the threats and risk assessment.

The chapter starts by summarising the security threats, it then continues with risk assessment and finally conclusions will be drawn.

7.1 Security Threats

In this section possible security threats to users of GSM, who we assume have high requirements on security, are presented.

It is possible to classify security threats in many different ways. Here we assume that the following categories are very critical for our specific users.

- Unauthorised access to sensitive data (violation of confidentiality and/or anonymity)
- Unauthorised manipulation of sensitive data (violation of integrity)
- Disturbing or misusing network services (leading to denial of service or reduced availability)
- Unauthorised access to services

Even though these threats are associated with attacks on the radio interface, they can even, as explained in Scenario 4 in Section 6.4.3, be associated with attacks on other parts of the system, both wired and wireless.

7.1.1 Unauthorised Access to Data

Unauthorised access to data is a very serious threat. It can be achieved either by gaining access to the data (confidentiality) or by performing analysis using information about what services the victim uses or has used (Anonymity - traffic analysis). We have the following categories of threats:

- **Eavesdropping user traffic:** Intruders may eavesdrop user traffic on the radio interface or by accessing the signalling network. Inside the signalling network the traffic and signalling is transmitted in the clear.
- **Eavesdropping signalling data:** Intruders may eavesdrop signalling data or other control data on the radio interface or accessing the signalling network. This may be used to access security management data or other

information which may be useful in conducting active attacks on the system. E g capturing authentication triplets from the signalling network by intercepting a microwave link or obtaining the IMSI/TMSI of target MSs.

- **Masquerading as a communications participant:** Intruders with access to a modified base station may masquerade as a network element to intercept user traffic, signalling data on the radio interface or on any other system interface. This has been thoroughly examined in chapter 6.
- **Passive traffic analysis:** Intruders may observe the time, rate, length, sources, or destinations of messages on the radio interface to obtain access to information about the activity of the user. This information can be used to deduce additional information, e g by noticing unusual traffic between certain users and/or destinations. We have reached the conclusion that such attacks are quite inefficient if the attacker lacks decryption capabilities (see Section 6.2.1).
- **Active traffic analysis:** Intruders may actively initiate communication sessions and then obtain access to information through observation of the time, rate, length, sources or destinations of associated messages both on the radio interface and other interfaces of the network. This sort of attack is very efficient in tracking GSM users (see Section 6.2.2)

7.1.2 Unauthorised Manipulation of Sensitive Data

We saw in several attacks in chapter 6 that an attacker mounting a man-in-the-middle attack is able to modify, insert, replay or delete messages while they are transmitted between legitimate entities of the network. This gives rise to the following threats:

- **Manipulation of user traffic:** Intruders may modify, insert, replay or delete user traffic on any system interface, whether wired or wireless. Note that replayed data which cannot be decrypted by an intruder may still be used to conduct attacks against the integrity of user traffic and signalling traffic.
- **Manipulation of signalling data:** Intruders may modify, insert, replay or delete signalling data on any system interface, whether wired or wireless.
- **Manipulation by masquerading as a communications participant:** Intruders may masquerade as a network element to modify, insert, replay or delete user traffic, signalling data or control data on any system interface, whether wired or wireless. Intruders may even masquerade as a legitimate subscriber and thereby get access to sensitive information about the system and the users.
- **Manipulation of data stored by system entities:** Intruders may modify, insert or delete data stored by system entities.

7.1.3 Denial of Service Attacks

In Section 6.1, we saw that captured MSs were totally dependant on the attacker in order to get access to the services of the network. The nature of the radio interface enables actually an attacker to seriously disturb the function of the network solely by jamming the radio signals. This disturbs the service provided to the users. The following threats are to consider:

- **Physical intervention:** Intruders may prevent user or signalling traffic from being transmitted on any system interface, whether wired or wireless, by physical means. An example of physical intervention on a wired interface is wire cutting e.g. cutting the wire connecting BTSs to the rest of the network. An example of physical intervention on a wireless interface is jamming. Physical intervention involving interrupting power supplies to transmission equipment may be conducted on both wired and wireless interfaces.
- **Protocol intervention:** Intruders may prevent user or signalling traffic from being transmitted on any system interface, whether wired or wireless, by inducing protocol failures. GSM utilises many different timers, which are used to ensure that the timing limits on responses are not exceeded. Delaying responses may lead to denial of service.
- **Denial of service by masquerading as a communications participant:** Intruders may deny service to a legitimate user by preventing user traffic and/or signalling data from being transmitted by masquerading as a network element (e.g. BTS)

7.1.4 Unauthorised Access to Services

The following threats may seem less severe than those mentioned above, but they could however easily lead to severe attacks on anonymity, confidentiality, etc.

- **Masquerading as a user:** Intruders may impersonate a user to utilise services authorised for that user. The intruder may have received assistance from other entities such as the serving network, the home network or even the user himself.
- **Masquerading as a serving network:** Intruders may impersonate a serving network, or part of a serving network's infrastructure, perhaps with the intention of using an authorised user's access attempts to gain access to services himself.
- **Misuse of serving network privileges:** Serving networks may abuse their privileges to gain unauthorised access to services. The serving network could e.g. misuse authentication data for a user to allow an accomplice to masquerade as that user, falsify charging records to gain extra revenues from the home environment, or even eavesdrop on a user's calls.

7.1.5 Threats Associated with Attacks on the Terminal (ME) and SIM

The following threats are against the equipment used by the users:

- **Confidentiality of certain user data in the terminal or in the SIM:** Intruders may wish to access personal user data stored by the user in the terminal or the SIM.
- **Confidentiality of authentication data in the SIM:** Intruders may wish to access authentication data stored by the service provider, e.g. the user's secret key that is stored in the SIM.

7.2 Risk Assessment

In this section, relevant threats will be analysed and evaluated with regard to the likelihood of occurrence and severity of impact. The list of the threats will follow below (Table 5) and be evaluated as being of major or medium impact. The threats of major impact will be indicated.

Threats	Major impact
Eavesdropping user traffic	×
Eavesdropping signalling data	
Masquerading as a communications participant	×
Passive traffic analysis	
Active traffic analysis	×
Manipulation of user traffic	
Manipulation of signalling data	
Manipulation by masquerading as a communications participant	×
Manipulation of data stored by system entities	
Physical intervention	
Protocol intervention	
Masquerading as a user	×
Masquerading as a serving network	
Misuse of serving network privileges	
Confidentiality of certain user data in the terminal or in the SIM	×
Confidentiality of authentication data in the SIM	×

Table 5 Evaluation of threats

7.3 Results of the Threat Analysis

Users of wireless communication systems with special requirements for security are more sensitive to vulnerabilities in the system than other categories of users. A country's military organisation probably has information that is regarded as very sensitive, which if exposed to outsiders, in the worst case, would cause severe damage to the organisation and even the country. Almost the same reasoning can be applied in the case of large companies fighting for market shares. For such entities confidentiality is very important. However, confidentiality and integrity are quite tightly connected; if unauthorised entities are able to easily alter information even confidentiality can be threatened. We saw in the previous chapter several examples on how the lack of integrity protection results in compromised confidentiality. Other organisations, e.g. emergency services may have somewhat lower demands on confidentiality though they are totally dependent on the availability of the communication services. For these users, threats realised into actual attacks would in the worst case lead to a catastrophe.

In the previous chapters, we have shown that the GSM system is vulnerable in many different ways. The threats are there and some of them, if realised into attacks, would result in major damage to the users. The damage could be so severe that it can result in a loss that is invaluable. The potential attacker is supposed to have enough resources, like time, computer power, money, knowledge etc in order to carry out the attacks. The threats exist indeed; when the attacker believes the information gain or the damage to the user is worth the investment made, he/she will exploit the vulnerabilities of the system to make the attack successful. The result of the attack will be serious damage to the victim users. (Figure 46)

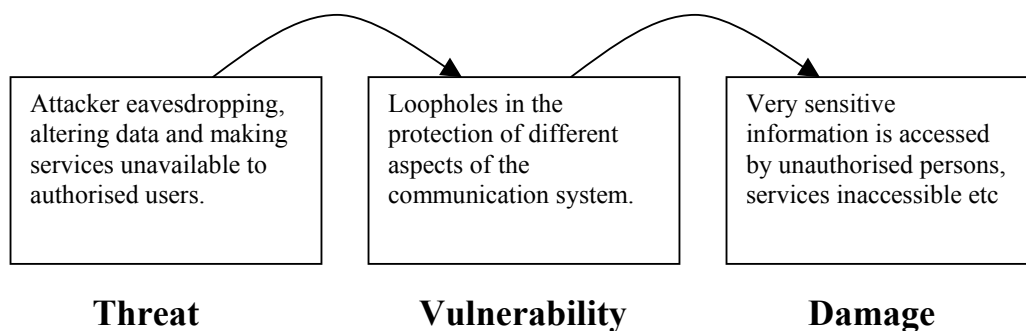


Figure 46 Threat-Vulnerability-Damage [43]

Since many of the threats have major impact and the probability that they will be realised into attacks is quite high, it may be warranted for users with very high security requirements to look for some system that is more secure. The conclusion is thus that the system is not suitable to use in communicating very sensitive information. This does not, however, necessarily mean that GSM should not be used by user groups with very high security requirements. GSM is a very complex

and impressive system, that provides the infrastructure for wireless communications almost everywhere in the world. Not using such a system would be a big waste of resources. What these users need is to add an additional layer of security to the current security model of GSM. In practice this could mean encrypting user data before standard GSM encryption is applied. This probably means an algorithm has to be chosen and implemented in hardware in the terminals that members of these organisations will use when using GSM services. In addition to the encryption of user data, smartcards specially designed to withstand partitioning attacks should be used to store sensitive user parameters.

Part IV

Discussion and Conclusions

Chapter 8 Discussion and Conclusions

Chapter 9 Future Work

In this chapter results will be presented and discussed.

8.1 Cryptanalytical attacks

There exist several *cryptanalytical* attacks against the algorithms protecting different aspects of GSM. The algorithm used by many operators to authenticate subscribers (COMP128) is broken due to flaws in the design of the hash function (see Section 6.3). The result is the ability of intruders to clone subscriptions either by having physical access to the target SIM or over the air. The most popular attack requires physical access to the SIM to clone and is completed in about 8 hours. It can be speeded up with the risk of damaging the SIM. The most efficient way to clone a GSM smartcard is a partitioning attack proposed by a team from IBM. It requires challenging the target SIM only 8 times in the best case, which means that cloning can be done in minutes or even seconds. The equipment needed to mount this attack (a specially designed smartcard reader and software) is however only available in laboratories yet. Newer versions of COMP128 has been developed and distributed. It is however not known to what extent these stronger versions have been adapted by operators. A guess is that many operators still use the old algorithm due to the costs involved in upgrading. What is known for sure is that users who had COMP128 inside their SIMs when they bought a subscription are still using COMP128.

There exist several cryptanalytical propositions on how to attack the encryption algorithms used for confidentiality protection (see Section 6.4.2), that break these algorithms in real-time. Several attacks against A5/1 and A5/2 exist (see Section 6.4.2 and Section 6.4.2.3), although most of them have only theoretical value. Most of the attacks require that the attacker knows portions of the key stream. It is possible to obtain small portions of plaintext because the attacker often knows the structure and content of the signalling messages (especially if the attacker is impersonating the network to the victim MS and is thereby able to query the MS for information) in addition to the fact that channel coding is applied to the data *before* encryption.

An attacker mounting a man-in-the-middle attack may ask the victim subscriber to transmit certain signalling messages (of which the content is known or almost known) after encryption has started. The attacker then has access to the ciphertext in addition to the known portions of the plaintext and can thereby derive portions of the key stream used in the encryption process. It is, however, hard to obtain the amounts of the known plaintext that some of these attacks require. The attack that requires least known plaintext, is an attack against A5/2. It requires that the attacker knows the plaintext of two frames approximately six seconds apart from each other and finds the session key in about 10 ms. The known plaintext

requirement may be possible to satisfy using the method mentioned earlier, therefore this attack on A5/2 has been used in one of the attacks on confidentiality (see Section 6.4.3). It is worth mentioning that it only took a couple of hours to crack A5/2, which illustrates the weaknesses of this algorithm.

The most recent attack on A5/1 is a ciphertext-only attack. This is an impressive attack only requiring knowledge of a small number of encrypted frames, enabling the attacker to listen to the “encrypted” conversation data, in real-time. Further the authors (of [35]) propose a ciphertext-only attack on A5/2 that improves the previous attack on A5/2 to a ciphertext-only attack. The problem of known plaintext is no longer a concern using this attack. This is however an attack requiring huge amounts of computation power. Table 6 gives an idea of the computation requirements in the proposed ciphertext-only attack on A5/1.

Available data (Ciphertext)	Pre- processing steps	Number of PCs to complete pre- processing in one year	Number of 200 GB disks	T	Number of PCs to complete attack in real- time
2^{12} (appr 5 min)	2^{52}	140	22	2^{28}	1
$2^{6.7}$ (appr 8 sec)	2^{41}	5000	176	$2^{32.6}$	1000
$2^{6.7}$ (appr 8 sec)	2^{42}	5000	350	$2^{30.6}$	200
2^{14} (appr 20 min)	2^{35}	35	3	2^{30}	1

Table 6 Three possible tradeoff points in the attacks on A5/1 [35]

Since this is a ciphertext only attack, no plaintext is required in order to find the session key in real-time. However, as seen in Table 4 the computation and storage requirements for this attack are very high making it very unlikely that an individual hacker would have the needed resources to mount the attack. The requirements for the ciphertext-only cryptanalysis of A5/2 are however fulfilled by most personal computers of today.

8.2 Attacks based on protocol weaknesses

Not having the resources needed to mount a cryptanalytical attack on GSM is however no guarantee of security. There exist several “flaws” in the GSM protocols and other vulnerabilities in GSM architecture that enable attackers with relatively modest resources to break the protection of valuable assets of GSM.

The following section will summarise these attacks.

8.2.1 Anonymity

It is obviously possible to build a device that is able to track a specific MS and extract its IMSI [15]. All that is needed is to buy the device and choose a location. In Section 6.2.1 we could see that it is possible to, capture the IMSI of a MS, only by listening to GSM traffic and hoping to catch a MS performing IMSI attach. This can be performed using a GSM scanner. Having access to a number of

IMSI doesn't necessarily mean that the owners are damaged in any way. The damage is done when the attacker can associate an IMSI with its owner and the TMSI it is assigned by the network, because the attacker is then able to locate the subscriber, track the subscriber's movements, know the sort of services he/she requests from the network etc. This however can not be done by passively listening to the radio signals (given that the attacker lacks decryption capabilities). Therefore the conclusion is that passive attacks do not form a real threat against the anonymity of GSM users.

In order to mount more serious attacks against the anonymity of GSM users the attacker needs to mount active attacks. Having access to a modified base station in conjunction with a modified MS makes active attacks possible. The intruder is able to interrogate users, using flaws in the protocol design (see Section 6.2.2), and get their IMSI, TMSI and their IMEI, among other parameters. This way traffic analysis can be performed very easily and the conclusion is that many of the requirements in Section 2.1, concerning user anonymity, are not met in today's GSM, given that the attacker is able to mount active attacks. To be able to mount active attacks the attacker needs a device hosting base station and MS functionality. The specifications for how this sort of equipment works is available and used base station are not that expensive to buy.

8.2.2 Authentication

The lack of subscriber anonymity in GSM does not, directly, form a monetary threat to the operator (the users of the system could though be damaged in many different ways because of the lost anonymity). What could seriously threaten the billing security is an attacker cloning a SIM and making business of selling it. Cloning a SIM, which is using COMP128 for authentication and session key generation, is quite easy if the attacker has *physical access* to it. All that is needed is a cheap off-the-shelf device and the SIM to be cloned. The most popular attack against COMP128 requires that the SIM is challenged about 150 000 times and takes about 8 hours to complete. Another attack is a partitioning attack requiring only 8 challenges in the best case, and about 1000 challenges in the worst case. This means that a SIM could be cloned in a few minutes or even seconds using this method. It requires however equipment that is only available in university laboratories – at least for the moment. This is however a serious threat; not only because it is very fast, but also because it does not attack a certain algorithm that can easily be exchanged with another stronger algorithm, but the smartcard itself.

Cloning using the physical SIMs is quite easy. It does not however form a serious threat to the security of GSM, since the attacker has to physically have the SIM to begin with. If the attacker gets hold of a couple of SIMs a month and clones them, that does not form a big threat to the credibility of a GSM service provider, because only the subscriber whose SIM has been cloned is concerned. This is, somewhat, similar to the fact that it is easy to clone a SIM used to decode digital TV signals, but people still buy them from the legitimate service provider. The situation could be different if the attacker is able to clone SIMs by using the radio interface (see Section 6.3.2) without having physical access to the card. The

attacker is then able to clone a considerably larger number of SIMs and make a hack of that, at the expense of the legitimate subscriber to begin with and later on the credibility of the service provider (operator), seriously harming its reputation, in addition to the costs for the services provided to users with cloned SIMs that the operator will not get paid for. In these times of hard competition between different operators, such incidents could easily make subscribers opt for a different operator.

Cloning over the air consists of two steps. Firstly the attacker needs to get the RAND-SRES-pairs needed in the cryptanalysing step. This is done by interrogating a MS for the information. This step takes about 13 hours to complete and could be done in parts. As many SIMs can be cloned as there are channels available. The second step consists of processing the data and is the same operation as in the case of cloning when the SIM to be cloned is available. The resources needed to mount an over the air attack exceed greatly those needed when the attacker has physical access to the SIM, but the damage (and profit for the attacker) is greater. The attacker needs the equipment needed to mount an active attack, that is a modified base station, or a device implementing the functionality needed which is a subset of the functionality of a legitimate base station. Since the implementation of COMP128 is available and there already exist devices that can extract the key, it should be possible to do the processing step work on a PC. Hard disks are cheap, and the attacker needs not more than approximately 24 MB of space on the hard disk for each SIM to be cloned ($150' \text{ challenges} \times 128 \text{ bits} + 150' \times 32 \text{ bits}$). The conclusion here is that the requirement of Section 2.3 is absolutely not fulfilled 100 percent.

Being able to challenge MSs using the radio link suddenly gives the attacker access to many cloned SIMs, that can be sold to people who want to call more cheaply or want to be anonymous to authorities. Since the attacker uses the radio link the fraud cannot be traced and the “customers” are totally anonymous. It is possible for the operator to track a specific MS but it is unrealistic to believe that a buyer (or the person doing the cloning) of a cloned SIM can be caught. The attacker makes initially a big investment acquiring the equipment needed, but is able to produce large quantities of cloned SIMs. But how big is the investment? [14] gives a hint that it is possible to mount an active attack on a MS. Providing a stronger signal to a MS than the legitimate BTS that is serving a cell makes the MS think that it has entered a new cell and that way the attacker is able to communicate with the MS. For this to be successful the device mentioned in [15] does not suffice. Some of the protocols and channels that are needed in order to do the challenging, specified and publicly available by ETSI, have to be implemented. To get an estimate on the effort needed to engineer or buy a device that can be used to make an attack on GSM, the company that develops the device mentioned in [15] was asked about the cost and delivery conditions. The device is a passive system, not able to communicate with the MS and therefore active attacks are not possible using it, but it provides some of the needed functionality. The answer was that the device can only be sold to law enforcement, government agencies or persons/organisations that have government approval. The device

costs approximately US\$180.000 and delivery can be made within eight weeks from the ordering date. This sounds like an enormous investment and probably not too many “hackers” can afford it. It is even questionable whether such an investment will be profitable. It would be very interesting to see how much it costs to build a device with the minimum functionality needed to act like a (limited) base station. Probably the cost would be much lower than the sum mentioned above. A trustworthy guess is US\$10.000 [37].

8.2.3 Confidentiality

The ETSI specifications reveal that a MS has to have support for different encryption algorithms to support roaming. Many western countries use the “strong” encryption algorithm A5/1, while other countries are “forced” to rely on the much weaker A5/2. A MS which in default mode encrypts using A5/1 is forced to use A5/2 if it is communicating with another MS that only supports A5/2. This fact can be used in a man-in-the-middle attack to eavesdrop on the users’ conversations. The attacker asks the calling MS to start encrypting using A5/2, and extract the session key, and then the conversation can be decrypted in real time by the attacker. Such an attack requires that the attacker is able to intercept and actively communicate with the MS and the legitimate BTS. This means that the attacker needs the equipment hosting the functionality of a modified BTS and a modified MS enabling the attacker to mount a man-in-the-middle attack. Another requirement is that the plaintext needed to cryptanalyse A5/2 to find the session key is available. This is one of many attack scenarios that can brake the confidentiality protection of GSM (see Section 6.4.3). Other attacks do not even require cryptanalysing the encryption algorithm.

An attacker that wants to get hold of a specific person’s (or organisation’s) secret business plans etc can mount one of the attacks mentioned in Section 6.4.3. The difference between attacks on the confidentiality aspect and those in the cloning case is that the attacker needs to know the identity of the SIM of the person that is to be eavesdropped on. For cloning it suffices to challenge *any MS*, for meaningful eavesdropping the attacker is supposed to know the IMSI of the interesting SIM. When the IMSI is known the next step is to locate the target MS and place the intercepting equipment near the target. Since no operator informs the subscriber about the type of encryption that is used for each call (this can be done according to the ETSI specifications) the subscriber does not suspect anything.

Being eavesdropped on is of course not acceptable for most of us, whether using mobile communications or not. However if a “normal” conversation is eavesdropped, the participants often do not lose anything but some privacy and probably nobody is ready to invest many thousands in buying equipment for intercepting and decrypting my private mobile calls (the reward does not motivate the investment), but certainly there are people/organisations who could be seriously damaged (e g financially) by being eavesdropped. Users of mobile communications that want to exchange very sensitive and valuable information

should be careful using the system, and at least be aware that the second generation GSM does not provide perfect confidentiality.

There are some easy ways to considerably increase the security of the GSM system. The most obvious way is to start using the full length of the session key K_c instead of setting the last 10 bits to zeros. A second “solution” that should not be too complicated to implement is to force the BTS (network) to authenticate itself to the MS when an authentication session is conducted. Even message-origin authentication and message integrity protection would be large steps in the right direction. That way an attacker would find it considerably harder to mount the man-in-the-middle attack meaning that many of the attacks are warded off. Another action in the right direction would be to inform users about what sort of protection they are using in real-time. It should not be too hard to indicate for a calling customer whether A5/1; A5/2 or A5/0 is being used to protect the conversation.

8.3 Conclusion

Given the strong belief in the security community that only protocols that can be tested should be trusted (that security should depend on the secrecy of keys and not of algorithms), some believe that it was inevitable that GSM would be attacked for its dependency on the proprietary authentication and confidentiality algorithms. These algorithms are viewed as cryptographically weak by many security analysts and this is proved by the increasing number of propositions for how to break these algorithms. It is a fact that COMP128, the algorithm used for authentication and session key generation, only required a couple of hours to crack (Wagner and Goldberg) and it has been broken for a couple of years, making the process of cloning SIMs using COMP128 trivial and cheap. Another fact is that it only took a couple of hours for the same team to crack A5/2. Their attack only requires a few cycles to crack the algorithm.

There are propositions on how to break the algorithms protecting the privacy of GSM conversations as well. Many of these propositions demand however unrealistic portions of known plaintext and/or huge amounts of computation power (especially for the one time pre-computation part of the attacks). The latest cryptographic attack on A5/1 is however a ciphertext-only attack requiring only a small number of encrypted frames in order to find the session key in real-time. This attack requires however very large amounts of computation power both in the pre-computation stage and in the real-time part of the attack.

Looking at the history of the cryptographic protection of GSM the picture becomes clear. Although it is obvious that secret algorithms make it harder to break the protection in the short run, it often fails in the long run. Designers of security reason that before cryptanalysing the algorithm the potential attacker has to know the algorithm, which will make the task much harder. This kind of reasoning fails; history shows that COMP128, A5/1 and A5/2 was reverse-engineered and cracked in a short period of time by individual researchers, for

some reasons. Firstly, the public crypto community is not given a chance to examine the algorithm to find eventual flaws. Secondly, some entities may have interest in deliberately build in flaws in the algorithms to make it easy for them to crack when they need to. Limiting the key bits to 54 instead of 64 may be an indicator of this. This reasons will make the device in [15] possible to build even in the future, meaning that authorities and other entities that can buy and use the device will be able to perform illegal tracking, eavesdropping etc, violating the personal integrity of the concerned users of the system. Note that law enforcement agencies are able to perform this actions in a legal way asking for permission to perform tracking and/or eavesdropping.

Not having access to the resources required to break the cryptographic algorithms protecting GSM does not however mean that GSM is secure. Certain flaws in the protocols that are used to manage the system make it possible for people with relatively modest resources to listen to GSM conversations in real-time without breaking the encryption algorithms. An attacker with access to a modified base station (easy to buy a used one) can mount active attacks on the system enabling the attacker to break the anonymity and confidentiality aspects of GSM and even clone SIMs using the radio link.

This means that the security provided by GSM is quite weak and users with very high demands on communication security should be more careful when the phone rings. However, both cryptanalytical attacks and attacks against protocols are considered nontrivial, except in the case of operator abuse.

The author of this report has presented some attacks on the second generation GSM that are believed to be possible in presence of the needed resources. However, the presentation has omitted the details of how things are done and the attacks have not been verified through practical experiments . A future work could go deeper into:

- examining which functionality needs to be implemented in a base station in order to be able to mount an active attack on a MS, e g to be able to mount a man-in-the-middle attack. Of course an estimation of how much a device that can act like a (limited) base station costs may be interesting knowledge. Of course it is essential to examine also whether such a device can operate according to the outlined attacks without making operators and authorities suspicious.
- performing practical experiments using equipment hosting base station functionality in conjunction with mobile station functionality to examine to what extent theory hold in practice.
- look into what work is needed to make the attacks presented in this report impossible or at least harder to mount. What is the situation in the UMTS systems?
- investigating the security situation for one of the large GSM operators in Sweden. This would take the form of a qualitative study consisting of interviews with security personnel in charge of the operators security implementation. Interesting questions to answer could be:
 - To what extent is COMP128 used? What version of COMP128 is used in newly issued subscriptions? Is cloning considered as a problem?
 - The authentication process is of essential significance within the GSM security model. Not only it ensures that only legitimate users get access to services, it also generates a fresh session key to be used in the subsequent encryption of signalling and data transmission. ETSI specifications recommend that users should be authenticated before service is granted. This means that a subscriber could be authenticated several times every day. Is this how it works in practice or is the authentication process performed rarely? How often does the network initiate an authentication process without a user requesting a service?
 - Is it common that signalling and user data are transferred in the clear? Which cases result in the network deciding on “no ciphering”-mode? Does the network always decide on “no

ciphering” when incompatibilities arise in the ciphering capabilities of the entities taking part in the communication session.

- To what extent are unencrypted wireless links used in the network behind the BTSs? Are satellite link used? To what extent?

References

1. Vijaya C, *Security, Authentication and access control for mobile communications*,
http://www.ittc.ku.edu/~rvc/documents/865/865_securityreport.pdf
2. Trappe W, Washington L C, *Introduction to Cryptography with coding theory*, Prentice Hall 2001
3. GSM 01.02, European Telecommunications Standards Institute (ETSI), European digital cellular telecommunications system (Phase 2+) (GSM), *General description of a GSM Public Land Mobile Network (PLMN)*, 1997, <http://www.etsi.org>
4. Javier Gozalvez Sempere, *An Overview of the GSM System*,
<http://www.comms.eee.strath.ac.uk/~gozalvez/gsm/gsm.html>, March 5, 2000
5. GSM 02.17 (ETS 300 509): European Telecommunications Standards Institute (ETSI), European digital cellular telecommunication system (Phase 2); *Subscriber identity modules (SIM), Functional characteristics*,
<http://www.etsi.org>
6. GSM 02.16, European Telecommunications Standards Institute (ETSI), European digital cellular telecommunications system (Phase2); *International Mobile station Equipment Identities (IMEI)*, <http://www.etsi.org>
7. Brookson C, *Security and Cryptography Applications to Radio Systems, IEE Colloquium on GSM security: a description of the reasons for security and the techniques* 1994
8. Biryukov A, Shamir A, Wagner D, *Real Time Cryptanalysis of A5/1 on a PC*
<http://www.technojunkie.gr/gsm/data/gsmsec/a51-bsw.htm>
9. Wagner D, *Cellphone Security*,
<http://www.cs.berkeley.edu/~daw/talks/SAC02.ppt>
10. GSM 03.20 (TS 100 929), European Telecommunications Standards Institute (ETSI), European digital cellular telecommunications system (Phase 2+); *Security related network functions*, <http://www.etsi.org>
11. Security Algorithms Group of Experts (SAGE), *Report on the specification and evaluation of the GSM cipher algorithm A5/2*
12. J. Vales-Alonso, F. Isasi de Vicente, F. J. González-Castaño, J. M. Pousada-Carballo, *Real-Time Detector of GSM Terminals*
13. Manuel J. Fernandez Iglesias, Francisco J. Gonzalez-Castano, Jose M. Pousada Carballo, Martin Llamas Nistal, Alberto Romero Feijoo, *From Complex Specifications to a Working Prototype. A Protocol Engineering Case Study*

14. Francisco J. González-Castaño, Javier Vales-Alonso, José M. Pousada-Carballo, Fernando Isasi de Vicente, Manuel J. Fernandez-Iglesias *Real-Time Interception Systems for the GSM Protocol*,
15. Endoacustica, Security and Surveillance Products,
http://www.endoacustica.com/english/gsm_interceptor_en.htm
16. M Briceno, I Goldberg, D Wagner, *GSM Cloning*,
<http://www.isaac.cs.berkeley.edu/isaac/gsm.html>
17. J R Rao, P Rohatgi H Scherzer, *Partitioning Attack: Or How to Rapidly Clone Some GSM Cards*, IBM Watson Research Center,
<http://www.research.ibm.com/intsec/gsm.ps>.
18. Endoacustica, security and surveillance products,
http://www.endoacustica.com/sim_card_reader.htm
19. M Briceno, I Goldberg D Wagner, *A pedagogical implementation of A5/1*,
<http://www.scard.org>, May 1999
20. GSM 04.07 (TS 100 929), European Telecommunications Standards Institute (ETSI), European digital cellular telecommunications system (Phase 2+), *Security related network functions*, <http://www.etsi.org>
21. I Goldberg, D Wagner L Green, *The (Real Time) Cryptanalysis of A5/2*, presented at the Rump Session of Crypto '99, 1999
22. GSM 04.08 (TS 100 929), European Telecommunications Standards Institute (ETSI), European digital cellular telecommunications system (Phase 2+), *Security related network functions*, <http://www.etsi.org>
23. GSM 01.04, European Telecommunications Standards Institute (ETSI), European digital cellular telecommunications system (Phase 2+); *Abbreviations and acronyms*, 1999, <http://www.etsi.org>
24. Stallings W, *Cryptography and network security*, 2:nd edition, Prentice-Hall 1999
25. Bruce Schneier, *European Cellular Encryption Algorithms*, from Cryptogram, December 15, 1999, <http://kiwibyrd.chat.ru/gsm/cr-gram.htm>.
26. Fairhurst G, *Lecture notes in Communications Engineering*,
<http://www.erg.abdn.ac.uk/users/gorry/course/intro-pages/osi.html>
27. Margrave D, *GSM Security and Encryption*,
<http://spyhard.narod.ru/phreak/gsm-secur.html>
28. Azizi N, *GSM 900*, <http://www.eecg.toronto.edu/~nazizi/gsm/ma/>
29. <http://www.privateline.com/PCS/GSM06.html>
30. <http://www.privateline.com/PCS/GSM07.html>
31. Scourias J, *A Brief Overview of GSM*,
<http://kbs.cs.tu-berlin.de/~jutta/gsm/js-intro.html>

32. GSM 05.02, European Telecommunications Standards Institute (ETSI), European digital cellular telecommunications system (Phase 2+), *Multiplexing and multiple access on the radio path*, 1998, <http://www.etsi.org>
33. Intel Corporation, <http://www.intel.com>
34. *Technical information: GSM System Security Study*, <http://jya.com/gsm061088.htm>
35. Barkan E, Biham E, Keller N, *Instant Ciphertext-Only Cryptanalysis of GSM Encrypted Communications*, 2003, <http://www.cs.technion.ac.il/~biham/publications.html>
36. Ekdahl P, Johansson T, *Another Attack on A5/I, Abstract*, Proceedings of International Symposium on Information Theory (ISIT), Washington, 2001, <http://www.it.lth.se/patrik/publications.html>
37. Schneier B, *European Cellular Encryption Algorithms*, <http://www.schneier.com/crypto-gram-9912.html#EuropeanCellularEncryptionAlgorithms>
38. S Schmitz, "SHAMAN Deliverable D02 – Intermediate Report: Results of review, Requirements and reference Architecture", Information Society Technologies, 08 November 2001,
39. Schneier B, A Shostack, *Breaking Up Is Hard To Do: Modeling Security Threats for Smart Cards*, <http://www.counterpane.com/smart-card-threats.pdf>, 1999.
40. Briceno M, Goldberg I, *GSM Cloning*, <http://www.isaac.cs.berkeley.edu/isaac/gsm-faq.html>
41. Federrath H, *Security functions in mobile communication systems*, <http://www-sec.uni-regensburg.de/publ/2001/NECHidelberg20011106.pdf>
42. S M Redl, M K Weber, M W Oliphant, *An introduction to GSM*, Artech House 1995
43. Fåk V, *Lecture notes in Computer Security*, http://www.it.isy.liu.se/studentinfo/TSIT84/Risk_analysis.pdf
44. Brookson C, *Can You Clone a Smart Card (SIM)?*, <http://www.brookson.com/gsm/clone.pdf>
45. Pesonen L, *GSM Interception*, <http://www.dia.unisa.it/professori/ads/corso-security/www/CORSO-9900/a5/Netsec/netsec.html>
46. http://www.cellular.co.za/news_2003/120903-bizarre_jamming_of_moldova_gsm_n.htm
47. Gadaix E, *GSM and 3G Security*, <http://opensores.thebunker.net/pub/mirrors/blackhat/presentations/bh-asia-01/gadiax.ppt>

48. 3G TS 21.133, 3rd Generation Partnership Project, *Security Threats and Requirements*, <http://www.3gpp.org>
49. ETSI ETR 332, *Security Techniques Advisory Group; Security requirements capture*, <http://www.etsi.org>

På svenska

Detta dokument hålls tillgängligt på Internet – eller dess framtida ersättare – under en längre tid från publiceringsdatum under förutsättning att inga extra-ordinära omständigheter uppstår. Tillgång till dokumentet innebär tillstånd för var och en att läsa, ladda ner, skriva ut enstaka kopior för enskilt bruk och att använda det oförändrat för ickekommersiell forskning och för undervisning. Överföring av upphovsrätten vid en senare tidpunkt kan inte upphäva detta tillstånd. All annan användning av dokumentet kräver upphovsmannens medgivande. För att garantera äktheten, säkerheten och tillgängligheten finns det lösningar av teknisk och administrativ art.

Upphovsmannens ideella rätt innefattar rätt att bli nämnd som upphovsman i den omfattning som god sed kräver vid användning av dokumentet på ovan beskrivna sätt samt skydd mot att dokumentet ändras eller presenteras i sådan form eller i sådant sammanhang som är kränkande för upphovsmannens litterära eller konstnärliga anseende eller egenart.

För ytterligare information om Linköping University Electronic Press se förlagets hemsida <http://www.ep.liu.se/>

In English

The publishers will keep this document online on the Internet - or its possible replacement - for a considerable time from the date of publication barring exceptional circumstances.

The online availability of the document implies a permanent permission for anyone to read, to download, to print out single copies for your own use and to use it unchanged for any non-commercial research and educational purpose. Subsequent transfers of copyright cannot revoke this permission. All other uses of the document are conditional on the consent of the copyright owner. The publisher has taken technical and administrative measures to assure authenticity, security and accessibility.

According to intellectual property law the author has the right to be mentioned when his/her work is accessed as described above and to be protected against infringement.

For additional information about the Linköping University Electronic Press and its procedures for publication and for assurance of document integrity, please refer to its WWW home page:

<http://www.ep.liu.se/>

© Paul Yousef
Linköping, 5th mars 2004