A study of security in wireless and mobile payments

By
Ahmed Ali
LiTH-ISY-EX--10/4157--SE
2010-06-30
A study of security in wireless and mobile payments

By
Ahmed Ali
LiTH-ISY-EX--10/4157--SE
2010-06-30
Nyckelord
WLAN, 3G, GSM, Security, Mobile payments, Cryptography,
Abstract

Mobile payments are increasing in popularity in recent years. New mobile solutions are being developed in the form of new Internet capable mobile devices such as the IPhone and new wireless networks such as the LTE and WiMAX networks.

This report will present, explain and compare some of the most popular wireless networks that enable mobile payments, from a security point of view. The chosen networks are 3G with connection to GSM and WLAN networks. The main security mechanisms involved in each network, and how they work will be studied. Security requirements and some of the most important threats each network faces will be presented and discussed. The main purpose of the report is to examine if mobile payments offer an acceptable level of security to the average user.
Acknowledgements

I would like to express gratitude to my examiner Viiveke Fåk for understanding and supporting. I would also like to thank my family for their support and love. Finally, I would also like to thank my friend Marc for providing encouragement and motivation.
Contents

Abstract ..............................................................................................................................................

Acknowledgements ..........................................................................................................................

Contents .............................................................................................................................................

Chapter 1 – Introduction .....................................................................................................................1
  1.1 Background ..................................................................................................................................3
  1.2 Purpose and goal ..........................................................................................................................4
  1.3 Limitations ....................................................................................................................................4
  1.4 Target audience ...........................................................................................................................4
  1.5 Reading instructions .....................................................................................................................5

Part I

Chapter 2 – Fundamentals ..................................................................................................................7
  2.1 What is a mobile payment system ...............................................................................................9
  2.2 Short presentation of different payment solutions .......................................................................11
  2.3 Use areas ......................................................................................................................................13
  2.4 A buying example .......................................................................................................................15
  2.5 Real world examples and security concerns .............................................................................17
  2.6 The mobile unit – a short introduction .......................................................................................18
    2.6.1 2G/3G networks ....................................................................................................................18
    2.6.2 WLAN .....................................................................................................................................19
  2.7 An overview of the GSM network structure .............................................................................20
  2.8 3G network overview ................................................................................................................22
  2.9 An overview of the WLAN network structure ..........................................................................26
    2.9.1 The WLAN main components .............................................................................................27
    2.9.2 WLAN network overview ....................................................................................................28
    2.9.3 802.11x protocols ................................................................................................................29
  2.10 Air-interface ...............................................................................................................................30
  2.11 Market situation ..........................................................................................................................31
Part II

Chapter 3 - Security requirements of wireless networks

3.1 Security requirements of wireless networks

3.2 3G/GSM security implementation

3.2.1 Confidentiality

3.2.2 Authentication

3.2.3 Availability and reliability:

3.2.4 Anonymity

3.4 WLAN security implementation

3.4.1 WEP

3.4.2 WPA

3.4.3 WPA2

3.5 Cryptography and encryption

3.5.1 Introduction

3.5.2 Some goals of cryptography

3.5.3 Symmetric key cryptography

3.5.4 Examples of symmetric-key ciphers

3.5.5 Asymmetric or public key cryptography

3.5.6 Symmetric-key versus public-key cryptography

3.5.7 Hybrid cryptosystems

3.5.8 Cryptographic hash function

3.5.9 Key management

3.6 Other aspects

3.6.1 Security Policy

Chapter 4 - Security attacks

4.1 Security attacks

4.2 General attacks

4.3 Attacks on 2G/3G networks

4.3.1 Types of attacks

4.3.2 Examples of attacks

4.4 Attacks on WLAN networks

4.4.1 Types of attacks

4.4.2 Examples of Attacks

4.4.3 Attacking WLAN security mechanisms

4.5 Cryptanalytical attacks
Chapter 1 – Introduction
This report is a Thesis done at the Computer Science Program at Linköping Institute of Technology – Linköpings Tekniska Högskola. It has been conducted at the Division of coding Theory at the Department of Electrical Engineering – Institutionen för Systemteknik (ISY).

This first chapter will introduce the reader to the subject covered in this report. The background, motivation and purpose for this study will be presented. In addition a view over the rest of the report and its disposition will be presented.

1.1 Background

“Mobile commerce (also known as M-commerce) is any transaction, involving the transfer of ownership or rights to use goods and services which is initiated and/or completed by using mobile access to computer-mediated networks with the help of an electronic device” [4].

Today it’s a common thing to see people with several mobile devices all connected to the Internet or some other network. We are connected all the day at all times for better or worse. Today we can’t imagine a mobile device whether it is a cellular phone or a laptop without connectivity functionality. New use areas with many possibilities and new business models have emerged and still do to accommodate this new way of using our devices. For instance, it is possible to conduct our daily business and errands on the move and even if we were in a different country.

This new lifestyle presents us with many new ways of doing things and it wouldn’t be to exaggerate to say that our everyday life is easier in many ways. However as with everything else in life there are costs to this wonderful technology and lifestyle. Not only the costs of equipment, hardware and software which are obvious. There are costs that are unseen to the average user but at the same time play a major role in making the whole thing work properly. We are talking about security here, and though some might argue that security is not so vital or important and that the average user has nothing to hide, the opposite will be shown further in this document with some real-life examples of how bad or ugly it could be with such a device or system with no security mechanisms present.

Some businesses exist today with the sole goal or business model of conducting online frauds, identity thefts and other various attacks. Their motivation is that there are huge revenues in these business models. I don’t think that it is difficult to understand that information is valuable especially when it has to do with companies.
Chapter 1

Introduction

1.2 Purpose and goal

My goal with this document is to study mobile payments in general and focus on wireless security. Two popular wireless standards, 3G and 802.11, will be studied extensively and compared from a security point of view. Within this study different payment methods and the security mechanisms present in those methods and their requirements will be presented. To better understand security and how it works possible attacks and threats will be studied. Some questions that this document will try to answer are:

- What is a mobile payment system and what is its market situation today?
- What are the requirements of a secure mobile payment system?
- How secure are mobile payments and what are the most important security attacks?
- Can a mobile payment system be considered to be secure? And for whom?
- Is the level of technology and available features in mobile devices satisfactory for an average user?
- What can be said about the future of mobile payment systems?

1.3 Limitations

Although this report focuses on a small part of the information security field, it covers enough of the fundamental principles so that the reader will be prepared for further study in the field. This document is a thesis work done at the university and will therefore be an introduction to the field of information security in wireless networks. For in-depth study of the mechanisms of wireless networks there are many books and reports better suited for the purpose. This document requires only minimal background knowledge in computer science, cryptography and mathematics. However despite these limitations the report covers many of the most important topics in the field.

The economic aspects introduced in this report are included to build a bridge of understanding between the theoretical technologies used in such systems and the practical use of these systems in the real world. For further reading about the economic situation and impact there are other materials to continue on.

1.4 Target audience

The target audience of this document is mainly other students in the field of information technology and computer science who are familiar with the basics of information security on a graduate level. It can also be read by others focusing on other aspects of this document, mainly the economic aspects. It can also be viewed as an introduction to mobile and online payment systems for further reading.
1.5 Reading instructions

Part I

Chapter 2: in this chapter the concept of a mobile payment system is introduced. The different types of payments will be presented and compared. An introduction to wireless networks and some real world scenarios of online and mobile payments are also included.

Part II

Chapter 3: here we will highlight the security requirements of wireless networks in general and present some specific security measures for the two types of networks studied in this document.

Chapter 4: here we provide a view of the most important attacks and security vulnerabilities of wireless networks in general and the GSM and Wi-Fi networks in specific. We also present possible solutions and/or preventive actions that could be taken against the presented attacks when possible.

Part III

Chapter 5: in this final chapter our main findings will be presented and discussed. It will also include answers to the questions presented in the purpose of this document and a discussion to those answers. A discussion about the current market situation of mobile payment systems and their financial aspects will be presented. Finally some word will be said on future work in this field.
Part I

Chapter 2 – Fundamentals
In this chapter we will discuss and explain some of the basics of the topics covered in the rest of the report. This chapter will lay the foundation for the remaining chapters. Being an introduction chapter it will sometimes be out of scope of the report. The aim is to give the reader a “big picture” view in order to more easily understand the basics. We will begin and define the main topic and then work my way downwards to define the building blocks of each part.

2.1 What is a mobile payment system

“Electronic commerce, commonly known as e-commerce or eCommerce, or e-business consists of the buying and selling of products or services over electronic systems such as the Internet and other computer networks” [35].

A mobile payment system, sometimes referred to as M-commerce, is any system where transactions with monetary value are conducted via a mobile network or other wireless networks using a mobile handset or some other device with capability to connect to a network wirelessly [23]. In this report a mobile payment system refers to the mobile device and the network it operates on.

According to the definitions above, M-commerce is a subset of e-commerce, meaning that there are other ways of conducting electronic commerce other than the types discussed in this report. For instance e-commerce needs not to be conducted wirelessly or by using a mobile handset. This report limits itself to discuss the business-to-customer model leaving out the business-to-business part as it will be out of scope and requires us to present some complex business models existing in different businesses.

Historically the most common payment method has been paying with physical money, cash. Physical money can easily be transferred and there are no transaction charges attached to payments, which is a favourable feature in the case of small payments. However, as banking businesses develop and transaction amounts became larger, problems take place with cash payments. Large transactions involve much security and maintenance related concerns. Some of the concerns are replacing old and worn out currency with new ones, printing, maintaining and transferring them. [14]

The Internet as a marketplace differs from conventional markets in some aspects. The Internet provides the ability to effectively make many transactions that can be geographically spread over great distances. It also provides a virtual marketplace for tangible and intangible goods, such as information, electronic applications, images, music and videos. The advantage of using the Internet for intangible goods is that they can be delivered electronically. [14]

The business model covered in this report will be rather simple and easy to follow because the aim of this report is to discuss and explain the security mechanisms included which will be the perspective of studying in this report. The business model is included to better understand how the security mechanisms are implemented and how they are used. In short our model is that a customer wants to purchase a product or a
service from a retailer or service provider over a network doing all this wirelessly. The main components of our principal mobile payment system will be shown later in chapter 2. We will limit the report to data traffic meaning that voice traffic will not be included. Ordering a product or service by making a simple phone call or leaving a voice message is therefore out of scope. [23]

M-commerce is a worldwide phenomenon and a growing market in many countries. Our primary geographic focus for this report will be Western Europe and specifically Sweden. This is because in the area of mobile communications and data traffic infrastructure Europe has a clear lead in terms of usage, application development and a high market penetration of mobile devices. [23]

M-commerce can be done in different ways. This report will focus on solutions where commerce and transactions are done over the internet using the 3G\textsuperscript{1} network or WLAN\textsuperscript{1} network. Other solutions will only be shortly overviewed.

Some benefits of a mobile payment system

- Transactions can be made on the go
- Providing wider reach
- Reducing transaction cost
- Very large market penetration ability. There are approximately 4.6 billion people with mobile phones globally. This means that the number of potential customers exceeds any other device used today
- Competitive pricing.
- Reducing time to order
- High anticipation and demand on new technologies and new products.

Some disadvantages of a mobile payment system

- Less functionality for mobile Internet done on mobile devices compared to wired solutions
- Limited bandwidth
- Technology constraints of mobile devices (memory, processing power, display capabilities, input methods)
- Limited battery life
- Connectivity can be interrupted
- Security of data moved across some mobile and wireless networks
- Businesses investment in hardware and infrastructure is seen as riskier as rapid evolution of mobile and wireless technologies continues
- Data throughput speeds can still be limited compared to wired solutions
- Mobile devices can be lost
- Higher risk of getting broken

\textsuperscript{1} Definition and explanation will follow in chapter 2
2.2 Short presentation of different payment solutions

There are many ways to make financial transactions or to purchase goods. Ranging from the oldest way of using money as something you physically have with you, to using electronic money consisting of electronic signals. Below follows a brief presentation of the most common methods.

Cash:

As mentioned above cash is one of the oldest ways of making financial transactions and it needs no extensive presentation. The main advantages of using cash is that it is universally accepted and doesn’t require the user to have any other dedicated software, hardware or any identification measures to use. However there are many disadvantages to using cash:

- Making large transactions requires large amounts of money transporting from one place to another with many risks
- Counting large amounts of money is time consuming
- Having large amounts of money in one place like a bank or a store makes it an attractive target for criminals
- This payment method requires everyone to have money with them at all time
- The amounts needed to make everyday transactions are unknown beforehand. Which makes the whole buying and selling business difficult
- The banking business would be much more difficult to conduct
- Everyone carrying large amounts of money would be a victim to theft
- Different countries have different currencies
- Managing physical money is associated with huge costs for our society.

Cheque:

The main advantages of using cheques are that money doesn’t need to be carried physically and that transactions can be maid easier. However, this payment method has several disadvantages:

- They are not secure because cheques are easily forged
- They could be returned by banks if the account associated to the cheque has insufficient funds. This is what is referred to as a bouncing check
- They require an exhausting process by banks and retailers
- Cashing cheques across international boundaries is associated with high costs
- Cheque books can be lost or stolen. [20]

Debit card

Debit cards such as Visa or MasterCard were the next step in making financial transactions easier. They address some of the disadvantages of using cash and cheques. The main advantage is that they offer a secure
way of carrying large amounts of money at any time. They also provide the possibility of withdrawing money as cash from any ATM\textsuperscript{2} system available. In early days when debit cards first entered the market there were problems of availability of both ATM systems and payment systems in stores, but that problem is less apparent today and the debit card payment method is the most spread and used method after using cash [22]. Credit cards have almost the same functionality and differ in that the bank having your account can give you a limited credit to make transactions with money that you don’t have, counting on your promise to pay it back adding a cost of interest.

Electronic money:

Money has been transferred and exchanged electronically long before the electronic or online payment methods were made available to the public. All major banks have been using it since many years to transfer money between banks and to make large transactions. Transactions between countries and large corporations have been made electronically for a long period of time too. However those systems are not disclosed to the public and its difficult to determine what security mechanism is used and if there is any standard method. This report discusses electronic payment systems using public networks, such as the Internet, to make online transactions or online payments. [21]

Mobile payments

Using electronic signals as means of payment as described in the previous chapter. What we will compare in this report is making mobile payments with a mobile device using the GSM/3G network with a mobile device using the WLAN network. We include GSM because the 3G network has much in common and is based on the same architecture of GSM. WLAN was chosen because of its popularity and that in recent years many of the devices supporting 3G functionality also support WLAN networks.

\footnote{See Abbreviations}
Chapter 2  
Fundamentals

2.3  Use areas

In this chapter we will present a short overview of some use areas for online payments. In recent years there has been considerable interest in developing various mobile commerce services. Some of these services are a new way of doing an existing business model and others are completely new business and payment methods.

Mobile ticketing

Buying tickets for bus and train fares has and is still one of the most important use areas of mobile payments. Tickets can be sent to users mobile devices using a variety of technologies. Purchased tickets are usually ready to use immediately or on a later occasion.

Mobile vouchers and coupons

Mobile ticketing technology can also be used for the distribution of vouchers, coupons and loyalty cards. The voucher, coupon, or loyalty card is represented by a virtual token that is sent to the mobile phone. Presenting a mobile phone with one of these tokens at the point of sale allows the customer to receive the same benefits as another customer who has a loyalty card or other paper coupon/voucher.

Content purchase and delivery

Mobile content purchase started primarily with selling ringtones, wallpapers and games for mobile phones. However in recent years where almost every mobile phone has the capability of internet connectivity, many other business models have appeared. The convergence of mobile phones, mp3 players and video players into a single device will undoubtedly result in an increase in the purchase and delivery of full-length music tracks and video content. Future network standards which increase download speeds to higher levels, will make it possible to buy full-length movies on a mobile device in a couple of seconds, while on the go.

Location based services

Unlike a home PC, the location of the mobile phone user is an important piece of information used during mobile commerce transactions. Knowing the location of the user allows for location based services such as local maps, local offers and local weather. [3]

Information services

A wide variety of information services: such as news services, sports results and traffic information, can be delivered to mobile phone users in much the same way as it is delivered to personal computers. Particularly, more customized traffic information, based on users' travel patterns, will be multicast on a differentiated basis, instead of broadcasting the same news and data to all users. This type of multicasting will be suited for more bandwidth-intensive mobile equipment. [3]
Mobile banking

Mobile banking allows the use of mobile phones to access account information and also make transactions such as purchasing stocks and remitting money.

Auctions

One of the use areas is a mobile reverse auction. Unlike traditional auctions, the reverse auction (or low-bid auction) bills the consumer's phone each time they place a bid. Many mobile PSMS (Premium Short Message Service) commerce solutions rely on a one-time purchase or one-time subscription; however, reverse auctions are high return applications as they allow the consumer to transact over a long period of time.

Mobile purchase

Mobile purchase allows customers to shop online at any time in any location. Customers can browse and order products while using a cheap, secure payment method. Instead of using paper catalogues, retailers can send customers a list of products that the customer would be interested in, directly to their mobile device or consumers can visit a mobile version of a retailer’s e-commerce site. Additionally, retailers will also be able to track customers at all times and notify them of discounts at local stores that the customer would be interested in. [3]

Mobile marketing and advertising

Mobile marketing is an emerging concept. Mobile marketing is a highly responsive sort of marketing campaign, especially from brands experience point of view. And almost all brands are getting higher campaign response rates. Corporations are now using m-commerce to expand everything from services to marketing and advertisement.
2.4 A buying example

As mentioned before our model is composed of many parts or entities and it is why we call it a system. Here we will define the main components of our principal mobile payment system model.

The main components

The main components of our model are depicted in figure 2.1 below:

1. Here we have the information sources where a potential customer finds products and services of interest. They can be in printed or physical form as magazines and newspapers or in electronic form as from the Internet. The product of interest can be anything from physical products to services, bus tickets etc.

2. Here is our typical buying customer or user.

3. A mobile device or handset. Most of the time we refer to a mobile phone. We require that the phone has Internet connectivity functionality. There is however no restraint on the type of device used, as long as it has wireless connectivity functionality on WLAN and 3G/GSM networks.

4. Here is a part of the network used to place an order. It is limited to the part between the customer and service provider. We require this part to be a wireless network.

5. The service provider or mobile operator, which manages authorized users identification, balance check, etc and forwards the order to the designated online store.

6. Another part of the network in our model between the service provider and online store. This part can be a wired network or wireless.

7. The online store receives the order placed by our customer from the service provider.

8. A confirmation message will be sent to our customer confirming that the order has been received. Some online stores send additional information with the confirmation message such as estimated time of deliverance and alternative payment methods.

The scope

This report will limit itself to a small part of our model depicted as the cloud in figure 2.1 below. See 1.3 for explanation. The cloud includes objects ②, ③ and ④ from the list above. Our viewing and studying perspective will be from the security point of view. Additional information about economic aspects of the model will be presented in some chapters to give the reader a “big picture” and how our model is supposed to work in practice.
A buying example

Here we present a simple buying scenario that helps to understand how our model works. This scenario is best observed by studying figure 2.1 below. A user finds an interesting product in a magazine, newspaper or on the internet from an online store. After examining the product the user decides to buy it. The user places an order by using his/her mobile phone and fills in with personal and product information required for ordering. The order will first arrive to the mobile operator which will determine identification and balance checking then resend the order to the designated store. The store confirms receiving the order to the mobile operator and the user and then package and sends the order.

**Figure 2.1:** Principal model of a mobile payment system
2.5 Real world examples and security concerns

When we presented the use areas of M-commerce above and a typical buying example, one can ask why we need security and how are all the additional costs and overhead in performance created by security measures justified. Here we examine two hypothetical scenarios, where there are no security measures implemented, to study what happens and what the consequences are of having no security at all. This is done to better explain the need of security.

Scenario 1: Hijack an account

Without a proper security system implemented in the mobile network where there would be no measures to authenticate users, it would be possible for an outsider to hijack someone’s account and impersonate that person or simply make that person pay for the services and goods purchased. This is a form of identity theft where the hijacker, by having access to a victim’s account, can take various actions to discredit or financially harm the victim. [6]

Scenario 2: Phishing e-mails

Another scenario can be of an attacker sending an e-mail to an unsuspecting user, disguising the message as an official e-mail from a legitimate bank as seen in the figure below. The victim will most certainly be alarmed by receiving this message and most likely follow the link. The attacker’s, in this case the sender of this message, purpose by sending this message is tricking the victim into revealing his or her personal information such as account information and password or pin-number. The attack is carried out by the victim clicking on the link which, instead of leading to the legitimate bank site, leads to the attackers site. On the illegitimate site all information typed in by victims will be stored for later use or sold on to other entities, thus taking the victims money or making other transactions that the victim is unaware of.

Figure 2.2: A phishing example [46]
Scenario 3: A system with inadequate security

As opposed to the two scenarios above there are many systems available today that have some security measures available, but these measures are inadequate or they are used in an improper way, for instance purchasing products on the Internet using debit and credit cards. If the data provided by the customer, such as card number or pin code, is not transferred securely, it would be possible for an attacker to intercept the traffic and making use of that information for personal goals. Many of these sites have some kind of security involved in their connections and data storing, but sometimes these security measures are used improperly and/or involve errors made by humans.

2.6 The mobile unit – a short introduction

Different devices can be used to make online purchases on the various networks available today. Some of these devices were early introduced for other purposes, but have undergone some evolution steps to make them compatible with new networks, and others are new devices developed specifically to work with new networks. Here we will present a short overview of some of the most popular devices.

2.6.1 2G/3G networks

Here there are two main categories of devices used on the 2G and 3G networks. The first is mobile phones of different types and the second is laptops.

Mobile phones

There are some early examples of mobile units used before the emergence of the GSM network and phones. These early units were primarily used by the military and/or other special groups. The first commercial fully automated cellular network, 1G, was launched in Japan in 1979. The first modern network technology on digital 2G cellular technology was launched 1991 in Finland on the GSM standard. In 2001 the first commercial launch of 3G was again in Japan by the WCDMA standard. [35]

The units used in 3G networks are similar in functionality and design to the units used with 2G networks except the ability to function with the new network. Almost all 3G units are compatible with the older 2G network, and for the purpose of making online purchases there are only minor differences in the devices.

Modern phones are usually composed of a plastic casing housing a circuit board. A keypad and a display are integrated in the plastic casing. Some new phones have touch screens instead of keypads. Older phones, typically 2G phones, had a visible antenna but those are also integrated in the plastic housing in newer units.
There are several categories of mobile phones, from basic phones to feature phones such as music-phones and camera-phones. Another category is smartphones which in many ways resemble a personal computer or laptop but with much smaller format and performance levels. Today almost all new mobile phones have features beyond sending text messages and making voice calls, including call registers, GPS\(^3\) navigation, music and video playback, document editing, memory card reader, USB, infrared, Bluetooth and WLAN connectivity functionality.

**Laptops**

Laptops are portable personal computers developed primarily for mobile use, first introduced in the early 1980s. They are typically notebook shaped and have a keyboard on one side and a screen on the other side when opened. They weigh between 1 to 5 kg and have one or more batteries attached in addition to the ability to use plugged to an electricity outlet for normal function and for charging the batteries. Laptops’ battery life can vary between minutes to several hours depending on model and the way they are used. Typically a laptop has most of its components integrated.

Laptops have had the ability to use 2G networks using modems and PCMCIA\(^4\) cards as a way to give them mobile online functionality. This was before the introduction of WLAN networks (802.11). Typical speeds of 2G modems varied between 9.6 Kbit/s and 14.4 Kbit/s.

When 3G networks were first introduced, they were primarily developed to use with mobile phones. This made the laptop market fall behind regarding online functionality. However to counter one of WLAN networks major problems which is range and mobility, some service providers have begun to offer a new alternative to use the 3G networks. A new market is growing where laptops connect to the 3G network using new types of modems.

**2.6.2 WLAN**

Similar devices are used in the case of a WLAN network as in a 3G network and often we have devices capable of using both network types. During the period of introduction of WLAN as a network, the primary device used was a laptop. A wireless network’s primary purpose is mobility, and therefore the mobile laptop was a natural choice. However, recently we have seen a development to widen the range of devices using WLAN as a wireless network. Many cellular and/or smart phones have functionality to support WLAN networks.

A cellular phone is much more mobile than a laptop and the average user carries his/her cellular phone at all time. This makes the cellular phone as a device a much more practical option to users. Manufacturers and service providers are realizing this matter and the potential that WLAN capable cellular phones provide. However there is a rather big obstacle for mobile phones using WLAN. The range of a WLAN network is

---

\(^3\) Global Positioning System

\(^4\) Personal Computer Card International Association
very limited when compared to that of a 3G network. This problem makes it very difficult for service providers to take full advantage of the new technology. Today, there are some attempts to build large public WLAN networks in big cities around the world, but this development is yet to be implemented.

2.7 An overview of the GSM network structure

The network behind the GSM seen by the customer is large and complicated in order to provide all of the services which are required. It is divided into a number of sections and these are, according to figure 2.3:

- **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 Ki and a cryptographic algorithm A3/A8 which provides 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), Ke, etc.
    - Service related data like Language Preference and Advice of Charge.
    - Card Holder Verification Information, authenticates the user to the card and provides protection against the use of stolen cards. 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). [12]

- **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. [12]

- **The Base Station Controller (BSC)**
  The BSC takes care of all the central functions and controls for a set of BTSs. The BSC and the controlled BTSs form the Base Station Subsystem (BSS). [12]
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. [12]

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 Ki and IMSI are stored in the HLR. Every PLMN requires at least one HLR and every user is assigned to one specific HLR. [12]

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 Ki, LAI, and TMSI. The AuC is responsible for generating triplets of values consisting of the RAND, SRES (Signed RESponse), and session key Kc which are stored in the HLR for each subscriber. [12]

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. [12]

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. 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 needs to be barred
- The Grey list: ME:s 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).
- Equipment Identification can be done by the network operator by requesting the IMEI from the ME. [12]
2.8 3G network overview

GSM (2G) networks provided only modest data rates which were considered sufficient in time of its introduction. However the growing need for greater bandwidth and faster rates for primarily multimedia content made it necessary to further develop and research into some new technology without having to revolutionize and rebuild the entire network.

2G networks offered its customers high-quality voice services; however they were not optimized for high-speed data. Service providers realised the needs of customers for more content in form of wireless data. They wanted a new method and/or design to add high-speed data without requiring massive investments in new technology on their part. A new specification was developed around the 1990s for high-speed wireless data called General Packet Radio Service (GPRS), also referred to as 2.5G. GPRS main advantages were:

- GPRS packet based technology\(^5\) made it compatible with the Internet which is also packet based, as opposed to GSM which was based on circuit switching\(^6\)

\(^5,6\) For explanation see definitions
End-users were able to be connected at all time because GPRS didn’t require a physical link to be opened for data transfer which allowed users to receive information only when they needed. Packet switching technology made the network work in a more efficient way by redirecting communications over optimal network paths bypassing bottlenecks. Data rates were increased. [26]

However GPRS still had its limits and the growth in the need for greater bandwidth was rapid. Efforts were made by some entities, notably Ericsson, to enhance data rates utilizing the existing radio spectrum and network. The result was the Enhanced Data rates for GSM Evolution (EDGE)⁷, also referred to as 2.75G. EDGE only required some minor software upgrades to the existing base stations, and therefore the costs of introducing it were only a fraction of what a pure 3G network deployment would cost. [26]

The EDGE technique provided the GSM network with enhanced data rates. New sophisticated coding methods were used over the Internet to increase data rates up to 384 kbps. [11]

In EDGE faster data rates were made available, however the air-interface was still limited due to limitations in the technique which made it poor in terms of efficiency. Another problem was that there were different standards to develop this technique around the world. There was a need for a new technology that made it possible to have a network providing services independent of the platform technology used and use the same network design standards globally. This new technology was 3G, which is short for third generation. In Europe it was called Universal Mobile Telecommunication System (UMTS). The American 3G variant was called CDMA⁸2000. [11]

3G technology is therefore an evolution of GSM using much of 2G technology. 3G offers users increased data rates, while retaining area coverage as well as the established user subscriber base of 2G systems. See table 2.1 below for a short comparison. [10]

<table>
<thead>
<tr>
<th>Technology</th>
<th>Real world (avg)</th>
<th>Theoretical (max)</th>
<th>Availability</th>
</tr>
</thead>
<tbody>
<tr>
<td>2G</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GSM</td>
<td>Download (Kbps)</td>
<td>Upload (Kbps)</td>
<td>Download (Kbps)</td>
</tr>
<tr>
<td></td>
<td>9.6</td>
<td>-</td>
<td>9.6</td>
</tr>
<tr>
<td>2.5G</td>
<td>GPRS</td>
<td>32-48</td>
<td>15</td>
</tr>
<tr>
<td></td>
<td>EDGE</td>
<td>175</td>
<td>30</td>
</tr>
<tr>
<td>3G</td>
<td>UMTS</td>
<td>226</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>HSPA 3.6</td>
<td>650</td>
<td>260</td>
</tr>
<tr>
<td></td>
<td>HSPA 7.2</td>
<td>1400</td>
<td>700</td>
</tr>
<tr>
<td>Pre-4G</td>
<td>WIMAX</td>
<td>3-6 Mbps</td>
<td>1Mbps</td>
</tr>
<tr>
<td></td>
<td>LTE</td>
<td>5-12 Mbps</td>
<td>2-5Mbps</td>
</tr>
</tbody>
</table>

Table 2.1: data rates for different networks and technologies [48]

⁷ See definitions
⁸ See 2.10
High Speed Packet Access (HSPA) was a minor evolution to 3G networks and provided increased performance by using improved modulation schemes and by refining the protocols by which handsets and base stations communicate. These improvements lead to a better utilization of the existing radio bandwidth provided by WCDMA (see 2.10).

Here we describe the main parts of the 3G network architecture, as seen in figure 2.4:

- **UTRAN**: UMTS Terrestrial Radio Access Network, which includes the node Bs and RNC. It allows connectivity between the user device and Core Network.

- **RNC**: Radio Network Controller, is responsible for controlling the Node Bs that are connected to it and similar to BSC in GSM. It is the point where encryption is done before user data is sent to and from the mobile. It connects to the Core network through MGW.

- **RNS**: It controls the allocation and release of specific radio resources to establish a connection between a user device and the UTRAN. Each UTRAN can contain more than one RNS

- **Iub**: a logical interface between the RNC and Node Bs.

- **CBC**: the Cell Broadcast Centre is the functional entity within the mobile network that is responsible for the generation of cell broadcast information

- **Core Network**: is the central part of the network and it routes calls across the public switched telephone network. Primary functions are user authentication and charging mechanisms.

- **Node B**: is a UMTS base station and serves one or more users. It is more complex than the base station of a GSM network. Some of its functionality includes handover channel management, base band conversion, channel encoding and decoding.

- **SGSN**: The Serving GPRS Support Node keeps track of the location of an individual MS and performs security functions and access control. The SGSN also exists in a UMTS network, where it connects to the RNC

- **MSC**: A Mobile Switching Centre is a telecommunication switch or exchange within a cellular network architecture which is capable of interworking with location databases
MGW: The media gateway manages data transfers in both circuit and packet switched networks. A MGW operates in conjunction with the MGCF (Media Gateway Control Function) in order to support interworking between IP based transport networks and the circuit switched PSTN (Public Switched Telephone Network) or PLMN (Public Land Mobile Network) domains. The MGW will translate media arriving within a TDM⁹ based timeslot to an RTP¹⁰ (Real time Transport Protocol) stream or vice versa. [13]

3G networks are undergoing some evolution (as in the case of GSM) since they were first introduced. One of these major changes is the shift to an all-IP network structure. This new network structure treats all information whether it is voice, data or video in the same way, which is by encapsulating everything into packets. The new network will be built around the Internet Protocol (IP), hence the name all-IP. This shift is important for our study since a network based on IP introduces many threats that, although they are known in the PC world, are new to voice communication networks.

⁹ See definitions chapter
Chapter 2  
Fundamentals

4G
Further development to make data rates even faster and address some network cost issues has been made into a new generation known as 4G. The principal difference between 4G and 3G is that 4G is based on all-IP packet switched networks and that the functionality of the RNC and BSC is distributed to the BTS and a set of servers and gateways, which in practice means that networks become less expensive and that data rates will be much faster. The reason for the transition to the all-IP is to have a common platform for all the technologies that have been developed so far, and to harmonize with user expectations of the many services to be provided. However this new technology is in its first stages and we have only seen some limited deployment in specific countries, therefore our discussion will be limited to 2G and 3G. [11]

The 4G standard can be divided into two main and competing air-interface technologies:

- Long Term Evolution (LTE): developed by 3rd Generation Partnership Project (3GPP). It changes two fundamental aspects of current 3G networks. LTE introduces all-IP flat network and therefore finally leaves behind the circuit switched network. This means that everything will be transmitted as data by the new network, even voice. The second change is the use of multiple-input and multiple-output (MIMO) technology at the transmitter and receiver end to improve communication performance. This can be used either to increase data rate throughput or to decrease interference. [11]

- WiMAX: WiMAX is a wireless broadband access standard developed and maintained by the IEEE under the 802.16 designation. As its name suggests, WiMAX can be thought of as an extension of Wi-Fi designed to enable pervasive, high-speed mobile Internet access on a wide range of devices, from laptops to smartphones. WiMAX also supports MIMO technology as with LTE. [11]

2.9 An overview of the WLAN network structure

Wireless local area networks (WLAN) were first introduced to the market around the 1990s [25], primarily as an alternative to cabled LAN in places where cabling was difficult or impossible. WLAN networks usually operate in unlicensed frequency bands to simplify its operation and expand the deployment base. Among these bands are the 2.4 GHz and the 5.8 GHz bands that are the most widely used frequencies among WLAN communications, where the 2.4 GHz band is the most widespread [24].

Efforts to specify a standard from the beginning were important to ensure product compatibility and reliability among manufacturers. One of the most important and widespread standards developed was the IEEE 802.11 and its later evolutions (802.11b, 802.11g and 802.11n). When referring to a WLAN network in this report it is the IEEE 802.11x standard that is studied.

26
2.9.1 The WLAN main components

The 802.11x uses spread spectrum technology to broadcast. Spread spectrum is a method of dividing data and sending it over a wide band of different frequencies. Multiple frequencies are used and they appear to be radio noise to narrowband devices. This noise can be filtered out easily, which enables the coexistence of narrowband devices. Some common spread spectrum methods are FHSS (frequency hopping spread spectrum), DSSS (direct sequence spread spectrum) and IR (infrared). For explanation on FHSS, IR and DSSS, see the definitions chapter. [26]

![Diagram of 802.11x protocol architecture](image)

**Figure 2.5:** 802.11x protocol architecture

The 802.11 defines the interface between wireless clients and their network access points (AP). This includes the PHY and MAC layer and also defines the security mechanisms, such as WEP (see 3.2.1), and an outline of how roaming between AP:s should work. The PHY defines the wireless transmission and there are three types of transmissions that can be used: FHSS, DSSS and IR. The most commonly used is the DSSS radio transmission, because its resistance to signal jamming, ability to share a single channel among multiple users and determination of relative timing between transmitter and receiver. [26]

In the FHSS the transmitter and the receiver hop from one frequency to another in an arranged synchronized pattern. Hops occur frequently and each hop consumes little time on any specific frequency. This reduces the possibility of interference with other devices and enables several FHSS systems to be operational at the same time. [26]

In the DSSS data is pushed through a binary encoding process that spreads data by combining it with a multibit pattern. This results in the data being hidden and inflated. For instance if the bit pattern is 11 bits long, then 1 bit of data would be 11 bits long. This data is modulated and sent out over multiple frequencies at the same time. Since the original data bit was encoded into 11 bits, the data is more resilient to air loss because the data has a big amount of redundancy [26]. The IR uses infrared radiation for short range transmissions.

The MAC layer in 802.11 controls access to the PHY and performs error recovery, roaming functionality and power conservation. These functions are not normally provided by a MAC used in wired networks. The
MAC layer also hides the physical characteristics of the wireless medium from the higher networking layers. The MAC layer has two main standards of operation:

- Distributed mode: in this mode it uses basically the same methods that wired Ethernet networks use (carrier sense multiple access with collision detection) to share the same wire.

- Coordinated mode: uses a centrally coordinated polling mechanism to provide support for applications that require support for real-time traffic. [26]

### 2.9.2 WLAN network overview

The IEEE 802.11 supports communication of terminals via both structured and ad-hoc network architectures. In the structured mode, an access point aggregates traffic from multiple mobile stations (STA), also called client stations, onto the wired network via the distribution system. The configuration that consists of at least one access point connected to the wired network infrastructure and a set of wireless end stations is called basic service set (BSS). Wireless connections between the access points are supported via a special frame format that effectively tunnels original frames over the 802.11 wireless networks. The set of two or more BSS:s is called extended service set (ESS). [24]

![Structured WLAN network architecture](image)

**Figure 2.6:** Structured WLAN network architecture
Chapter 2

Fundamentals

Ad-hoc is the simplest type of network where two STA:s communicate only peer-to-peer (P2P). Here the WLAN is typically created and maintained without prior administrative arrangement for specific purposes, such as transferring a file from one personal computer to another.

2.9.3 802.11x protocols

The letters following the 802.11, such as 802.11b, define different subgroups of the 802.11 that have been formed and given specific areas of wireless networking. Some involve higher-speed technologies; others involve specific areas that need an implementation solution such as security mechanisms. Below is a list of the most important 802.11 protocols:

- **802.11b**: uses the 2.4 GHz radio band and provides data rates up to 11 Mbps, which is the data rate of the physical interface. Data throughput rates are less than that of the physical interface because of the MAC layer overhead, errors and collisions. It uses Complimentary Code Keying (CKK)\textsuperscript{10} modulation technique and DSSS spread spectrum method. Operation in the 2.4 GHz radio band creates some interference problems with other devices, such as microwave ovens and Bluetooth devices which operate in the same range.

- **802.11a**: is a high-speed interface definition which uses the 5 GHz radio band and OFDM modulations scheme. Data throughput is approximately 54 Mbps. The 5 GHz radio band operation provides better transmissions without interference, since the 5 GHz band is relatively unused. However, the a standard is not compatible with devices using the b standard or older devices and there are some license problems of using the 5 GHz radio band in some countries where it is restricted. [26]

- **802.11g**: this standard uses the same radio band of 2.4 GHz as the b standard and the OFDM modulation scheme of the a standard. The g standard combines the advantages of both a and b standards providing high data rates of 54 Mbps and compatibility with older devices. It has however the same disadvantage as the b standard of frequency interference.

\textsuperscript{10} See definitions
2.10 Air-interface

The radio communication link between a mobile device and an active base station is called the air-interface. The different networks that we study in this report have different channel access methods. “A channel access method allows several devices connected to the same multi-point transmission medium to transmit over it and to share its capacity” [35].

TDMA

*Time Division Multiple Access* (TDMA) is the air-interface used by GSM (2G). It is a channel access method that digitizes the voice signal and turns the signal into short packets. A frequency channel is then used for a very short time period before migrating to another channel. The voice packets can occupy different time slots in different frequency ranges at the same time. At the receiving end sent packets will be recognized to recreate the conversation. [28]

WCDMA (CDMA)

*Wideband Code Division Multiple Access* (WCDMA) is the air-interface technology used by the UMTS (3G). WCDMA added some improvements to the air-interface compared to TDMA and at the same time allowed the use of older GSM/EDGE devices. This technique allows several users to share the same frequency by dividing the signal into different time slots. WCDMA uses *spread-spectrum*\(^{11}\) technology and a special coding scheme, where each transmitter is assigned a code to allow multiple users to use the same physical channel\(^{12}\). [28]

OFDM

*Orthogonal Frequency Division Multiplexing* (OFDM) is “a multicarrier modulation method that divides a communication channel into a number of equally spaced frequency bands. Each band is then used to transmit a portion of the user information and each band is independent of or orthogonal to every other band” [26]. OFDM in different variations is used for the WLAN 802.11x standard.

---

\(^{11}\) See definitions

\(^{12}\) For further studies on air-interface technologies see [26] and [28]
2.11 Market situation

This report’s main purpose is to study the security aspects of a mobile payment system. However, it is important to give some additional context to the subject and study how a concept mobile system is used and by whom.

According to a fresh report from WII (World Internet Institute) named “Svenskarna och Internet” [7], it shows that Internet users, who use Internet daily, are increasing in numbers every year, from a 25% in 2003 to circa 60% in 2009 of the total Swedish population. Internet is not only being used more often, but also the way we use it has changed. From being a place to simply view web pages in the early days, the Internet has become a platform for many other activities today, such as chatting, making calls, interacting in communities, making online transactions etc.

As opposed to making transactions or payments with conventional methods such as buying a product in a store using cash or a credit card, online payments have increased in popularity in recent years [14]. The increase is largely based on the fact that people use Internet on a daily basis compared to how it was a decade ago and that Internet availability has become widespread in most countries around the globe. Note that when we mentioned online payments earlier, it also included online payments done using personal computers and other devices which don’t connect wirelessly.

Almost everyone in Sweden, 93%, has a mobile phone, and many of them with Internet capability but only 18% of them use this opportunity. In contrast to other countries, such as Japan, mobile devices have yet not evolved into a device that is naturally used to access the Internet and/or do other online activities. According to the same report it is people between the ages 26-35 years that dominate the use of mobile devices to do online activities. [7]
Part II

Chapter 3 - Security requirements of wireless networks

Chapter 4 - Security attacks
In this chapter we look at the security requirements in 3G/GSM networks and the WLAN network. We explain what kind of security is needed on different layers of both networks and how they work. Later in the chapter we examine cryptography in detail and explain why it is needed.

The CIA security model will be used in several chapters within this report. It is a model that covers the most important aspects of security and its requirements and it is applicable on information security in general. The security of the networks that we discuss in this report is a subgroup of information security, since it is information that we desire to secure and protect. The acronym CIA stands for: Confidentiality, Integrity and Availability which are the traditional areas of security. There are some additions to this model that are important in some cases, for instance: accountability, nonrepudiation and reliability.

- Confidentiality: is the prevention of unauthorized disclosure of information
- Integrity: is the prevention of unauthorized modification of information
- Availability: is the prevention of unauthorized withholding of information or resources
- Accountability: is that users should be held responsible for their actions
- Nonrepudiation: is a way to provide unforgeable evidence that a specific action or event occurred
- Reliability: a system must perform properly in adverse or unpredicted conditions. [5]

### 3.1 Security requirements of wireless networks

The requirements listed below are set up to apply to general wireless communications and therefore can be used for our purpose. When we review security requirements in the 3G network we must begin with explaining security requirements in GSM. In previous chapters we showed how 3G network architecture is built upon the GSM network and has many parts in common. This is also the case with the security mechanisms involved.

#### End user privacy requirements

What kind of security does the user need to protect his/her privacy and sensitive information from being lost or stolen by other entities. We limit our study to security requirements needed to make online mobile payments and leave the remaining requirements for further study, see [14].

- **User location privacy:**
  
  Any leakage of specific signalling information on the network may enable an eavesdropper to approximately locate the position of a subscriber, which will risk the subscriber’s privacy. Hence the subscriber must be protected from such attacks on his/her privacy of location. [12]
User-ID privacy:
Wireless networks are in contrast to traditional or wired networks accessible to everyone. Illegitimate use of the network and gaining unauthorized access is an immense problem for wireless networks. Service providers use a user-ID to identify their subscribers. User-ID information can be used by attacker in various ways and therefore this information must be protected primarily by encrypting it before it is sent on the network. [12]

Data traffic privacy:
Data traffic between the end-user and the service provider must be protected, usually by encrypting all communications and proper authentication mechanisms. [12]

Integrity protection of data:
System data and traffic data must have some functionality to verify its integrity. Their must be a way to detect if received information has been altered. [12]

Requirements for Preventing Theft of Service or Equipment:
Since mobile devices are mobile per definition they are at great risk of being stolen. The network does not have the ability to physically identify a legitimate user. Use of stolen devices is made even easier if there is no password protection on the device or if the attacker has access to the specific password. To avoid stolen devices being sold or transferred to other users, it is required that each device has a unique identifier. This unique identifier is connected to the user identifier in some sort of database accessed by service providers. Devices with unique identifiers can, if stolen, be reported to service providers where adequate steps can be taken to block a certain device making it unusable. [12]

There are two kinds of thefts possible here. First there is the theft of personal equipment as described above, secondly we have theft of the services offered by the service provider. Stolen devices must have means of making them unavailable or non-functioning when legitimate users report them stolen. [12]

Cloning and Clone Resistant Design:
Cloning refers to the ability of an attacker to determine information from a specific user and/or device and make a clone, i.e. create a duplicate copy, of that personal device using the information collected. Cloning is a serious problem in mobile communication systems. This kind of fraud can be easily accomplished by legitimate users of the network, since they have all the information they need to clone their own personal device stored in the Subscriber Identity Module (SIM). In this way, multiple users can use one account by cloning personal equipment. This kind of attack can also be carried out by an attacker who wants to use services at the expense of legitimate users or sell the cloned devices. [12]

Clone resistant design can be done by using cryptographic protection on the different layers of the network to provide protection. There has to be adequate means to correctly identify legitimate users.
This can be done by using unique tokens each user has or to use unique passwords or PIN codes for identification. This process of identification is called the authentication process. [12]

### 3.2 3G/GSM security implementation

Here we explain how the security requirements discussed above are implemented. We can use the CIA security model here, to better see how the security implementation meets the requirements.

#### 3.2.1 Confidentiality

"Confidentiality is the prevention of unauthorised disclosure of information" [5]. To fulfil the requirements of end-user ID privacy, user location privacy and data privacy there must exist some kind of functionality to ensure confidentiality.

The SIM provides information needed to encrypt the radio connection between the MS and the BTS. It contains a key, $K_i$, which is later used in some version of A5 (see 3.5.4) for encrypting the data before transmitting on network. An algorithm is used for computing the 64-bit session key, $K_c$ (see [27] for further study) [12]. COMP128-4 for the 3G network is based on AES.\(^{13}\)

\[
\begin{align*}
\text{RAND (128 bit)} & \quad \downarrow \\
K_i (128 \text{ bit}) & \quad \rightarrow \\
\text{COMP128} & \\
\downarrow & \\
128 \text{ bit output} & \\
SRES (32 \text{ bit}) \text{ and } k_c (64 \text{ bit}) & \\
\end{align*}
\]

**Figure 3.1: COMP128**

GSM uses TDMA to share the radio channel with up to seven other users. Each user takes turns using the shared radio channel, sending and receiving information only during one of the eight available time slots in every frame. Each frame is identified by a “frame number”. A GSM conversation uses two frames, one

\(^{13}\) See definitions
going from the base station to the MS and another going from the MS back to the base station. Each of these frames contains bits of user information and it is this information that needs to be encrypted [12]. 3G networks use a similar technique, to the one used in GSM, called WCDMA and explained in 2.10.

The session key produced from the COMP128 algorithm described above is later transferred to the MS where it is used by another algorithm, the A5. A5 uses the session key and the frame number to produce a key stream of bits. For each new frame to be transferred a new key stream is produced by the A5 to be used to encrypt (and decrypt) the specific frame. The A5 algorithm resides in hardware in the device, not in the SIM, and must operate quickly and continuously to generate new key streams. [12]

3G networks use the KASUMI block cipher (see 3.5.4) instead of the A5/1 stream cipher. They are used in a similar way as A5/1 in GSM with some differences. The session key in KASUMI is, for instance, composed of 128-bit instead of 64-bit.

### 3.2.2 Authentication

2G and 3G networks are designed to be accessed by anyone and therefore authentication of users is very important. End-users must be able to identify themselves and validate that they are who they claim to be.

Authentication in 2G and 3G networks involves two functional entities, the SIM card in the mobile device, 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 an intruder. The network verifies the identity of a subscriber through a challenge-response process similar to the mechanism described in 3.5.3 and figure 3.3. When a MS requests service, the network sends a mathematical challenge to the MS, which it must answer correctly before being granted access. 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 small chance of being repeated, or it will be easy for an attacker make a codebook of (RAND, SRES) pairs and use the information to gain access to services. [12]

When the MS receives RAND it passes it into the SIM for processing. The SIM sends RAND and the secret 128-bit key $K_i$ 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 networks AuC will perform the same set of operations. 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 SIM does not have the proper secret key $K_i$ and therefore denies service to the MS. [12]
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 any particular RAND challenge is being 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, sometimes several times throughout the duration of a call. The next challenge the MS and SIM receives from the network will probably be a new one for the attacker, impossible for him/her to compute the right SRES for. It should be noticed that a cornerstone of the GSM security protocols is that a subscriber’s secret key, $Ki$, remains secret. While stored in both the SIM and the AuC, $Ki$, is never transmitted over the network. [12]

The COMP128 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 compromised in April 1998 and a new stronger version, COMP128-2 was developed. However, due to the huge amount of cost involved in replacing COMP128, it is believed that most operators are still using the old flawed algorithm. [12]

The authentication process for GSM and 3G networks are identical, because the GSM authentication was considered to have a sufficient security level. However the algorithm used in GSM (see A5/1 in 3.5.4) was weak, and for that reason it was replaced with KASUMI (see 3.5.4).

### 3.2.3 Availability and reliability:

GSM has a periodic location updating procedure to maintain reliability of service. If an HLR or MSC/VLR fails, to have each mobile register simultaneously to bring the database up to date would cause overloading. Therefore, the database is updated as location updating events occur. The enabling of periodic updating, and the time period between periodic updates, is controlled by the operator, and is a trade-off between signalling traffic and speed of recovery. If a mobile device does not register after the updating time period, it is deregistered. A procedure related to location updating is the IMSI attach and detach. A detach lets the network know that the mobile station is unreachable, and avoids having to needlessly allocate channels and send paging messages. An attach is similar to a location update, and informs the system that the mobile is reachable again. The activation of IMSI attach/detach is up to the operator on an individual cell basis. [12]

### 3.2.4 Anonymity

Anonymity in 3G and GSM is provided by using temporary identifiers. When a user turns on his/her device, a unique number associated to each network user (IMSI) is used to identify the MS to the network and then a 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 [12].
A location update request results in the MS receiving a TMSI. The TMSI has significance only within a location area. Outside the location area it has to be combined with the LAI (location area identifier) 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. From then on the temporary identifier is used. Only by tracking the user is it possible to determine the temporary identifier being used. [12]

3.4 WLAN security implementation

There are different existing methods and solutions in the standard 802.11x to address the security requirements of wireless networks. Most of the solutions rely on encryption, as it is proven to be a good way to provide adequate security if implemented properly. Cryptography is explained in the next chapter, therefore, to fully understand this chapter it is advised to review subchapter 3.5 first.

3.4.1 WEP

WEP (wired equivalent privacy) is a security mechanism included in the IEEE 802.11x standard. It was designed to provide confidentiality and authentication services.

Confidentiality

WEP provides confidentiality using a stream cipher called RC4 (see 3.5.4) and the CRC-32 checksum (see 3.5.8) for integrity. WEP was compromised in 2004 and is no longer considered to be secure [29]. The first goal of WEP is to prevent casual eavesdropping; the second is to protect the wireless network by discarding improperly encrypted messages. A third goal is to prevent tampering with transmissions by using integrity checksums. [30]

The RC4 cipher operates by expanding the public IV and the secret key into a long key stream of pseudorandom bits. Encryption is done by performing an XOR\(^\text{14}\) operation (marked with the symbol \(\oplus\) in the figure below) on the plaintext and the generated key stream. Decryption is done by generating the identical key stream by performing an XOR operation on the IV, secret key and the ciphertext. Encrypting two messages with the same IV and key can reveal information about both messages [30]. It is considered that WEP completely failed to meet its security goals [5].

---

\(^{14}\) See definitions
Chapter 3

Security requirements of wireless networks

Standard 64-bit WEP uses a 40 bit key (also known as WEP-40), with a 24-bit initialization vector (IV) to generate a stream of encrypted packets used by the RC4 cipher. Later versions of WEP have increased the key size to 104-bit in a 128-bit protocol called (WEP-104). [29]

In a 128-bit protocol, a 128-bit key is entered by the user as a string of 26 hexadecimal characters. Each character represents four bits of the key. 26 digits of four bits each give us 104 bits; adding the 24-bit IV produces the final 128-bit WEP key. In a 256-bit WEP, 24 bits are reserved for the IV and the remaining 232 bits are entered as 58 hexadecimal characters. [29]

WEP should not be used for security purposes anymore because of its many weaknesses. WEP keys can be compromised using software available on the Internet in a few hours [26]. Another problem with WEP is the key managing problem. WEP uses a common key for all users on a given wireless network, which makes it difficult to protect the key. WEP keys are there needed to be updated frequently. Increasing the key size alone is not sufficient to make WEP secure. Longer keys require interception of more packets but there are active attacks that are able to circumvent this. [26]

**Authentication**

WEP has two ways of authenticating users: Open System authentication and Shared Key authentication. Open System authentication is a null authentication algorithm meaning that the user doesn’t need to provide credentials to the Access Point during authentication. Any user, regardless of used WEP keys, can authenticate with the Access Point and then attempt to associate. After the authentication and association, WEP can be used for encrypting the data frames. At this point, the client needs to have the right keys. [29]
In Shared Key authentication, a four-way challenge-response handshake is used:

- The user sends an authentication request to the Access Point
- The Access Point sends back a clear-text challenge
- The user has to encrypt the challenge text using the configured WEP key, and send it back in another authentication request
- The Access Point decrypts the material, and compares it with the clear-text it had sent. Depending on the success of this comparison, the Access Point sends back a positive or negative response.

It is possible to derive the key stream used for the handshake by having knowledge of a plaintext/ciphertext pair of the necessary length. It is relatively easy to obtain this information by monitoring a legitimate authentication sequence. [30]

3.4.2 WPA

The failure of WEP made it necessary to develop a new security protocol quickly. WPA (WiFi Protected Access) was designed as a temporarily solution to address the major flaws of WEP. It was also required that WPA would run on existing hardware. [5]

WPA provides improved authentication to the network using The Extensible Authentication Protocol (EAP). The CRC-32 used for checksum integrity in WEP was replaced by a message integrity code called Michael (MIC). The length of the IV increased to 48-bits instead of 24-bit in WEP and The Temporal Key Integrity Protocol (TKIP) further increased security [5]. Other improvements are: better encryption by using AES. See table 3.1 below for a short comparison. [24]

For private use and for small businesses WPA can be used without requiring the complexity of an authentication server. This is called Pre-shared key mode (PSK), where the shared key is a secret key shared between clients and the AP. Each wireless network device encrypts the network traffic using a 256 bit key. This key may be entered either as a string of 64 hexadecimal digits, or as a passphrase of 8 to 63 printable ASCII characters. PSK is often based on symmetric-key methods and therefore reducing the overall security level of WPA.

Extensible Authentication Protocol (EAP)

The EAP is a protocol used in the authentication process in WLAN networks. EAP supports several authentication methods called EAP methods, which include authentication based on passwords, certificates and tokens. It can also include a combination of authentication techniques, for instance a certificate followed by a password. It can therefore be considered as an authentication framework instead of a specific authentication mechanism. [41]
Temporal Key Integrity Protocol (TKIP)

The TKIP protocol was designed to increase security and fix all known weaknesses in WEP. One of its most important design goals was that it would be compatible with hardware using the older WEP. TKIP uses the new Message Integrity Check (MIC) which is generated by the keyed cryptographic algorithm Michael. TKIP also implements some countermeasures to handle security concerns because of the design constraints to be backward compatible. TKIP has protection against replay attacks (see 4.4.3). TKIP also uses a cryptographic per-packet key mixing function and de-correlates the public initialization vector (IV) from weak keys to counter weak-key attacks (see 4.4.3). [40]

<table>
<thead>
<tr>
<th>Standard</th>
<th>WEP</th>
<th>WPA</th>
<th>WPA2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cipher algorithm</td>
<td>RC4</td>
<td>RC4</td>
<td>AES</td>
</tr>
<tr>
<td>Encryption key length</td>
<td>40 bit and 104 bit</td>
<td>128 bit</td>
<td>128 bit</td>
</tr>
<tr>
<td>IV length</td>
<td>24 bit</td>
<td>48 bit</td>
<td>48 bit</td>
</tr>
<tr>
<td>Packet key</td>
<td>Concatenated</td>
<td>Mixing function</td>
<td>Not needed</td>
</tr>
<tr>
<td>Data integrity</td>
<td>CRC-32</td>
<td>MIC</td>
<td>CCMP</td>
</tr>
<tr>
<td>Key management</td>
<td>Static</td>
<td>EAP</td>
<td>EAP</td>
</tr>
<tr>
<td>Header integrity</td>
<td>None</td>
<td>MIC</td>
<td>CCMP</td>
</tr>
<tr>
<td>Replay attack</td>
<td>None</td>
<td>IV sequence</td>
<td>IV sequence</td>
</tr>
</tbody>
</table>

Table 3.1: 802.11 security protocols comparison [24]

3.4.3 WPA2

WPA was only an intermediate solution that was needed to address the WEP protocol security weaknesses. The complete protocol was called WPA2 or Robust Security network (RSN). WPA2 replaces the stream cipher RC4 used in WEP with AES in counter mode with CCMP. WPA2 requires new hardware. [5]
Robust security network (RSN)

The RSN authentication and encryption algorithms is a step in developing better security for the 802.11x WLAN standards. It includes security enhancements that address flaws of WEP and TKIP and provide robust protection for the wireless link, including data integrity and confidentiality. RSN is a wireless security network that that allows the creation of Robust Security Network Associations (RSNA), which are wireless connections that provide assurance against security threats in the WLAN network by using a variety of cryptographic techniques. [41]

The main security features of RSN are:

- Enhanced user authentication mechanisms
- Cryptographic key management
- Data confidentiality
- Data origin authentication and integrity
- Replay protection.

What RSN does is that it adds new algorithms when a new threat is detected and is therefore flexible by its design. [41]

**CBC-MAC (CCMP)**

Counter mode\(^{15}\) CBC-MAC, called CCMP, is an algorithm that includes the AES encryption used in WPA2. It uses the 128-bit AES for data protection, where data is encrypted in 128-bit packets using the block chaining\(^1\) (CBC) mode and provides data integrity via a MAC\(^{16}\) (Message Authentication Code). CCMP uses a 48-bit IV. The input to CCM mode includes three elements:

- A payload consisting of authenticated and encrypted data
- A header, which is associated data, for authentication but not encryption
- A unique value called a nonce, assigned to the payload and the associated data. [24]

---

\(^{15}\) See definitions

\(^{16}\) See 3.5.8
3.5 Cryptography and encryption

Many of the security mechanisms discussed in previous chapters are based on cryptography. In this chapter we explain the main concepts of cryptography and how they work, however we limit our study to aspects of cryptography that are used in our security mechanisms. For further studies in cryptography, see [1], [2], [8], [15] and [16].

3.5.1 Introduction

Cryptography is the study of hiding information. It is a technique in which algorithms provide a security service to protect the integrity of data and guarantee the authenticity of the source of data and also provide confidentiality for data.

Encryption schemes provide confidentiality. They transform data, which sometimes is called a message or plaintext, into coded text, called ciphertext, and make ready to securely transmit over a network. This way the message can not be revealed by an unauthorized entity. The receiver of the message must have access to secret information that is not known to any unauthorized person. This secret is known as a key. [15]

There are different kinds of encryption algorithms for different purposes, however there are two important encryption schemes that many other encryption algorithms are built upon. These schemes are symmetric-key ciphers and public-key (also called asymmetric) ciphers. The study of breaking an encryption scheme or message is called cryptanalysis.

![Figure 3.1: An encryption scheme family tree](image)

In both symmetric and asymmetric schemes, an encryption scheme consists of three algorithms:
- A key generation algorithm
- An encryption algorithm to encrypt the plaintext message to ciphertext
- A decryption algorithm to retrieve the plaintext from the ciphertext. [15]
3.5.2 Some goals of cryptography

- Confidentiality: is a way to ensure that information is accessible only to those authorized to have access [16]. Secrecy or privacy is another term used for confidentiality. There are many approaches to providing confidentiality, from physical protection to mathematical algorithms that render data incomprehensible. For example, a transaction on the Internet requires some personal information (such as account number and personal number, ID) to be transmitted from the buyer to the merchant over a public network. The system attempts to enforce confidentiality by encrypting this information during transmission, by limiting the places where it might appear (in databases, log files, backups, printed receipts, and so on), and by restricting access to the places where it is stored. If an unauthorized party obtains the card number in any way, a breach of confidentiality has occurred. Sensitive information must be encrypted to avoid unauthorized access. [16]

- Integrity: data integrity addresses the unauthorized modification of data. To ensure data integrity there must be some mechanisms to detect data manipulation by unauthorized entities. For example integrity is violated when a computer virus infects a computer and manages to modify or delete some files. Encryption helps to validate the origin of data; modified data has a different source than the origin. [16]

- Authentication: is when we want to identify the origin of information or an entity. In a transaction or communication the two parties involved must be able to identify themselves correctly to each other or a third entity. Information transferred over a channel should be authenticated to its origin, date of origin, data content, time sent, etc. For these reasons this aspect of cryptography is usually subdivided into two main classes: entity authentication and data origin authentication. Data origin authentication implicitly provides data integrity as we mentioned above. [16]

3.5.3 Symmetric key cryptography

Symmetric-key cryptography is a method of encryption where both the sender and the receiver share the same key. Two well known examples of symmetric-key encryption are the DES\(^\text{17}\) and the AES (which are described in short in the definitions chapter).

Symmetric-key encryption can be described with a simple example using the figure below. Alice wants to send a message to Bob. She doesn’t want any unauthorized person to read it, so she encrypts the message using her secret key, \(K_s\), and sends it over a public or unsecure channel. For Bob to read the message he has to have the same key, \(K_s\), that Alice used to encrypt the message. For an adversary the message can be intercepted but not read. The message cannot be easily decrypted without the secret key. Using a secure channel makes the data more resistant to intercepting and tampering.

\(^{17}\) See definitions
Another simple way of authentication is the challenge-response mechanism described in the figure below. Alice and Bob decide to communicate with each other and they want to encrypt their communications, so that others can’t read the information. They decide to use symmetric-key encryption since it is simple to implement and use. Alice and Bob construct a shared key, $K_{AB}$, to use for encryption and decryption. Alice sends a random number, known as a Challenge, to Bob. The value of the sent random number encrypted with the shared key, $K_{AB}$, using an encryption algorithm which is known by both Alice and Bob, is known as the Response. The same procedure is done again, this time with Bob sending the Challenge. When this exchange of challenge and response is done, then both sides have proven to each other that they know $K_{AB}$ and therefore they are who they claim to be, since they alone knew the secret key, $K_{AB}$. It would be difficult for an adversary to compute the secret key, $K_{AB}$, and therefore will be easily detected.

Figure 3.2: an example of a two party symmetric-key communication [15]

Figure 3.3: challenge response mechanism
Private-key or symmetric-key cryptography is designed to minimize requirements on computation when encrypting and decrypting data [16], enabling operation in high speeds, and therefore mainly used in mobile devices and other devices with limited computation power. There is however an important drawback in symmetric ciphers and that is that both the sender and receiver have to know the secret key which is a key shared by both to enable encryption and decryption. In order to ensure secure communication between everyone in a population of n people a total of n(n-1)/2 keys are needed, which is the total number of communications channels [16]. To limit the impact of a potential discovery by a cryptographic adversary, they should be changed regularly and kept secure during distribution and in service. The process of selecting, distributing and storing keys is known as key management. Symmetric ciphers can be categorized in two main classes, block ciphers and stream ciphers.

**Block ciphers:**

A block cipher is a symmetric-key encryption scheme which breaks up the plaintext message to be transmitted into bits or fragments (called blocks) of a fixed length and encrypts one block at a time. Each block must satisfy the predetermined size, and leftover fragments are padded to the appropriate block size. For example, if the predetermined block size is 16-bytes per block, and a 38-byte frame is to be encrypted, the block cipher fragments the frame into two 16-byte blocks and one 6-byte block. The 6-byte block is padded with 10-bytes of padding\(^{18}\) to meet the 16-byte block size. [43]

Many well known symmetric-key encryption techniques are block ciphers. An early highly influential block cipher design was the Data Encryption Standard (DES), developed by IBM and published as a standard in 1977. Two important classes of block ciphers are substitution ciphers and transposition ciphers and product ciphers which combine these two. [1]

---

\(^{18}\) Extra nonsense (in this context: information)
Stream ciphers

A Stream cipher is a symmetric-key encryption scheme which can be considered to be very simple block ciphers with a block length equal to one. A stream cipher encrypts data by generating a key stream from the secret key and performs an XOR function on the key stream with the plaintext data. The key stream can be of any size necessary to match the size of the plaintext frame to encrypt. [43]

![Diagram of a stream cipher](image)

**Figure 3.5:** An example of a stream cipher [43]

Stream ciphers can be viewed as approximating the action of a proven unbreakable cipher, the *one-time pad* (OTP)\(^{19}\). However since an OTP system is very difficult to implement, a stream cipher uses a much smaller and more convenient key size and stream ciphers use pseudorandom key generation as opposed to an OTP true random key. A stream cipher generates a pseudorandom key stream which can be combined with the plaintext digits in a similar manner to the OTP. That does not imply that stream ciphers have the same level of security as an OTP encryption scheme or that they have a good security level at all.

---

\(^{19}\) See definitions
3.5.4 Examples of symmetric-key ciphers

This chapter will describe some symmetric-key encryption ciphers that are implemented in devices used for mobile payments.

**A5/1 (used in GSM)**

A5/1 is a stream cipher algorithm used to provide over-the-air communication privacy in GSM networks by encrypting voice and data transmissions after a successful authentication. [26]

A GSM transmission is organized as sequences of frames. Each frame is sent every 4.615 milliseconds and contains 114 bits available for information which represents the digitized A to B communication and an equal 114 bits for the B to A communication. Each communication will be encrypted by a new session key $K$ of 64 bits, which is then mixed with a publicly known frame counter $F_n$ of 22 bits. The result serves as an initial state of a generator that produces 228 pseudorandom bits. These bits are then XORed by the two parties with the 114+114 bits of plaintext to produce the 114+114 bits of ciphertext. [44]

Three linear feedback shift registers (LFSR\(^{20} \)) of lengths 19, 22 and 23 bits which are denoted by R1, R2 and R3 respectively in the figure below construct the A5/1. The first bit from the right in each register is labelled as bit zero. The taps, which are bits to be XORed together to provide the next bit when the LFSR is shifted, shown in light grey of $R_1$ in figure 3.6, are at bit positions 13, 16, 17 and 18; the taps of $R_2$ are at bit positions 20 and 21; and the taps of $R_3$ are at bit positions 7, 20, 21 and 22. [44]

When a register is clocked, its taps are XORed (represented as a circle with a cross inside in the figure below) together, and the result is stored in the rightmost bit of the left-shifted register. The three registers are maximal length LFSR's with periods $2^{19} -1$, $2^{22} -1$, and $2^{23} -1$, respectively. They are clocked in a stop/go fashion using the following majority rule: Each register has a single "clocking" tap (bit 8 for $R_1$, bit 10 for $R_2$ and $R_3$ represented in dark grey in figure); 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. Note that at each step either two or three registers are clocked, and that each register moves with probability 3/4 and stops with probability 1/4. [44]

\(^{20}\) See definitions
The process of generating pseudorandom bits from the session key $K$ and the frame counter $F_n$ is carried out in four steps:

- The three registers are zeroed, and then clocked for 64 cycles (ignoring the stop/go clock control). During this period each bit of $K$ (from LSB$^{21}$ to MSB) is XORed in parallel into the LSB's of the three registers.
- 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 LSB's of the three registers. The contents of the three registers at the end of this step is called the initial state of the frame.
- The three registers are clocked for 100 additional clock cycles with the stop/go clock control but without producing any outputs.
- 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 MSB's of the three registers. [44]

---

$^{21}$ LSB is least significant bit and MSB is most significant bit.
Chapter 3  Security requirements of wireless networks

KASUMI (used 3G)

KASUMI is a block cipher used in 3G networks to provide confidentiality and integrity. The cipher is based on the cipher algorithm MISTY1\(^{22}\). The name KASUMI is the Japanese word for mist. KASUMI uses a 128-bit key and a 64-bit block size. The core of the algorithm is an 8-round Feistel\(^{22}\) network. The round functions in the main Feistel network are irreversible Feistel-like network transformations. In each round the round function uses a round key which consists of eight 16-bit sub keys derived from the original 128-bit key using a fixed key schedule.

RC4 (WEP)

RC4 is an encryption algorithm used in WLAN networks using WEP. It is a simple and fast stream cipher and has a key length of 8-2048 bits or 1-256 bytes. The algorithm’s internal state is a permutation of values ranging from 0 to 255 cyclically. RC4 is comprised of two algorithms:

- A key scheduling algorithm (KSA) that uses the encryption key to initialize the RC4 permutation
- A pseudorandom generating algorithm (PRGA) which is used to modify the RC4 state and outputs a byte of the key stream. [2]

RC4 is easy to implement in software and does not use LSFRs which are used in other stream ciphers, such as A5/1. However it is vulnerable to insertion attacks. [2]

3.5.5 Asymmetric or public key cryptography

Public-key cryptography is a relatively new cryptographic approach whose distinguishing characteristic is the use of asymmetric key algorithms instead of or in addition to symmetric key algorithms. The main difference between asymmetric-key and symmetric-key methods is that asymmetric methods use two keys for their function and only one key needs to be secret.

Each user has a pair of cryptographic keys, a secret or private key used for decryption and a public key which is used for encryption. The public key can be sent over networks without having fear that the security of the private key may be compromised. The private key must be kept secret and it is never sent over networks. This makes key managing much simpler than the case for symmetric-key methods. The keys are mathematically related, but the private key cannot be derived from the public key in feasible time. In an asymmetric-key scheme, anyone can encrypt messages using the public key, but only the holder of the paired private key can decrypt. Security depends on the secrecy of that private key. [2]

\(^{22}\) See definitions
Chapter 3  Security requirements of wireless networks

The asymmetric-key scheme provides message authentication by creating a digital signature of a message using the private key, which can be verified using the public key. It also allows protection of the confidentiality and integrity of a message, by public key encryption, encrypting the message using the public key, which can only be decrypted using the private key. [2]

![Principal public key scheme](image)

**Figure 3.7:** Principal public key scheme [51]

Asymmetric-key schemes have some disadvantages however.

- The algorithms are difficult to implement
- The algorithms tend to require much from hardware, and therefore are slow
- The algorithms are vulnerable to chosen cipher text attacks (see 4.5). [2]

Because of these disadvantages it is common to use asymmetric methods to encrypt session keys which are short and use a symmetric-key method to encrypt the messages. [2]

“An analogy to public-key encryption is that of a locked mailbox with a mail slot. The mail slot is exposed and accessible to the public; its location (the street address) is in essence the public key. Anyone knowing the street address can go to the door and drop a written message through the slot; however, only the person who possesses the key can open the mailbox and read the message.” [35]

**Digital signatures**

A message signed with a sender's private key can be verified by anyone who has access to the sender's public key, thereby proving that the sender had access to the private key (and therefore is likely to be the person associated with the public key used), and the part of the message that has not been tampered with. On the question of authenticity, see also message digest.

“An analogy for digital signatures is the sealing of an envelope with a personal wax seal. The message can be opened by anyone, but the presence of the seal authenticates the sender.” [35]
3.5.6 Symmetric-key versus public-key cryptography

In this chapter we present some of the advantages and disadvantages to both encryption schemes.

Advantages of symmetric-key cryptography

- It provides high data throughput rates due to its small requirements on hardware
- Keys can be kept relatively short
- Symmetric-key ciphers can be composed to produce stronger ciphers
- Symmetric-key schemes have an extensive history and large knowledgebase.

Disadvantages of symmetric-key cryptography

- For a two-way communication, the key must be kept secret at both ends
- For a two-way communication, keys should be changed frequently and preferably for each session to maintain a good security level
- In large networks key management becomes a problem, since many key pairs have to be managed

Advantages of public-key cryptography

- Only private keys must be kept secret
- Key administration and management in large networks is done easier, since fewer keys are needed
- A private key/public key pair doesn’t require frequent changes

Disadvantages of public-key cryptography

- Data throughput is less and slower than that for symmetric-key schemes, since public-key schemes require more from the hardware
- Key sizes are typically much larger than that for symmetric-key schemes
- Public-key schemes are relatively new in comparison being introduced in the 70s. [16]

3.5.7 Hybrid cryptosystems

Hybrid or asymmetric cryptosystems combine some of the advantages from both symmetric-key and public-key encryption schemes. They do not require using a secret key for secure communications as in symmetric-key schemes and by using short session keys of 8-16 bytes, they do not have large demands on hardware as some public-key encryption schemes. It therefore combines the easy way of managing keys of public-key schemes and the efficiency of symmetric-key schemes. [2]

54
A hybrid cryptosystem can be constructed using any two separate cryptosystems. A key encapsulation scheme which is a public-key cryptosystem and a data encapsulation scheme which is a symmetric-key cryptosystem.

If we use the Alice and Bob example (see Figure 3.8), a hybrid crypto system can work as follows:

- Alice gets Bobs public key, $X$
- Alice uses an efficient symmetric method to create a secret key, $S$, and encrypts the message, $M$, that is to be sent to Bob
- Alice uses Bobs public key, $X$, to encrypt the secret key, $S$, and then sends the encrypted message along with the encrypted secret key to Bob
- Bob uses his secret or private key, $P$, to decrypt the secret key, $S$, encrypted with, $X$, and recovers $S$. After decrypting the secret key, the encrypted message can be decrypted into plaintext. [2]

**Figure 3.8:** Hybrid method for transmitting encrypted messages [2]
Chapter 3  Security requirements of wireless networks

TLS/SSL (Internet)

Transport Layer Security (TLS) and Secure Socket Layer (SSL) are not ciphers but cryptographic protocols used to provide security for communications over networks such as the Internet. TLS is the successor to SSL. Since our study handles online transactions done on the Internet it is important to view at least one protocol and how it works. We choose to study SSL because it is the protocol that is used for the majority of secure online transactions [45]. SSL is transparent, which means that the sent data arrives at the destination unchanged by the encryption and decryption process. This makes SSL very useful and suitable for many applications [26].

Consider this simple example of a buying situation. You see an interesting book on your favourite online retailer’s, $R$, Internet site that you want to buy using your favourite mobile device. Before providing any private information about yourself you want to be sure that you are dealing with $R$. Put in other words you want to authenticate $R$. As a company do not need to authenticate you, since it is sufficient for them that you are a customer with enough money. After you verify that it is actually $R$ you are dealing with, you continue by providing your personal information such as account number, your name and address. You don’t want this information to fall in the wrong hands or in anybody else’s hands other than $R$. You want both confidentiality and integrity protection.

A basic example of how SSL works can be illustrated as a number of messages sent between a client and server:

- **Message 1:** Alice, $A$, informs the online retailer, $R$, that she wants to conduct a secure transaction, and sends a list of ciphers that she supports along with a nonce $R_A$.
- **Message 2:** $R$ responds with its certificate. $R$ selects on of the ciphers from the list of ciphers sent by $A$ in the first message along with a nonce $R_B$.
- **Message 3:** $A$ sends the “pre-mastered secret” $S$ that she generated along with a hash that is encrypted with the key $K$ generated from $S$. The hash includes “msgs”, which are all previous messages sent and CLNT which is a literal string. The hash is used to verify that the previous messages have been received correctly.
- **Message 4:** $R$ responds with a similar hash including “msgs”, SRVR which is a literal string and $K$. $A$ can thereby verify that $R$ received the messages correctly and she can authenticate $R$, since only $R$ could have decrypted $S$ which is required to generate the key $K$. At this point $A$ has authenticated $R$ and they both have a shared symmetric key $K$, which they can use to encrypt and integrity protect subsequent messages. [45]
Different keys are used in each direction to protect against certain types of replay attacks. SSL is based on a combination of symmetric and asymmetric algorithms. The authentication is accomplished by using public-key cryptography and is referred to as a *handshake*[^23], while the communication uses symmetric-key encryption to maintain performance. [26]

### 3.5.8 Cryptographic hash function

*A cryptographic hash function is a deterministic procedure that takes an arbitrary block of data and returns a fixed-size bit string, the (cryptographic) hash value, such that an accidental or intentional change to the data will change the hash value. The data to be encoded is often called the "message" and the hash value is sometimes called the message digest or simply digests." [35]

A cryptographic hash function $h(x)$ must have the following properties:

- **Compression**: for any input size the output length is small
- **Efficiency**: $h(x)$ must be easy to compute for any input
- **One-way**: it should be computationally infeasible to find the value $x$ for any value $y$ such that $h(x) = y$. It is difficult to invert the hash
- **Weak collision resistance**: for $x$ and $h(x)$, it should be infeasible to find $y$, with $y \neq x$, such that $h(y) = h(x)$
- **Strong collision resistance**: it should be infeasible to find any $x$ and $y$, with $y \neq x$, such that $h(y) = h(x)$. [45]

Collisions exist because the input space is much larger than the output space. Hash functions provide very good security and can be used to compute digital signatures. [45]

[^23]: See definitions
CRC-32 checksum (used in WEP)

“Cyclic Redundancy Check (CRC), is a non-secure hash function designed to detect accidental changes to raw computer data, and is commonly used in digital networks.” [35].

It is an algorithm based on polynomial division. It takes input data as a long binary number and divides it by a constant divisor. Every division provides a rest which is called the CRC checksum. For instance $8/3 = 2$ and a rest of 2, this means that the checksum of 8 is 2. CRC-32 produces a checksum with a length of 32 bits, hence the name CRC-32. CRC-32 is used for data integrity in the WEP standard described in chapter 3.4.1.
3.5.9 Key management

Key management is fundamental to every security mechanism based on cryptography. A cryptographic protocol uses keys to authenticate entities and grant access to protected information to those who exhibit their knowledge of the keys.

Key management is “the process of ensuring that both the sender and receiver have access to the key required to encrypt and decrypt a message, while making sure that the key does not fall into enemy hands.” [24].

There are two ways to distribute secret keys: the pre-established secure channel and the open channel. Key distribution was a major problem in terms of logistics and security before the invention of public-key cryptography. [24]

According to a report from NIST (National Institute of Standards and Technology) on recommendations for key management, the following requirements need to be met:

- Keys should be randomly generated to reduce the probability that they could be determined by an attacker or that they will be reused
- Keys should be changed frequently to reduce the possibility of being discovered through cryptanalysis
- Keys should be protected during transmissions over networks
- Keys should be erased when no longer needed
- Keys should be protected while in storage to prevent deciphering of previous communications. [42]
Chapter 3  Security requirements of wireless networks

3.6 Other aspects

In previous chapters we discussed security requirements for each specific network technology and also cryptographic security. However there are other aspects to security which are not related to a specific network or system. In this chapter we will present some general security considerations that are important when building a secure system.

A secure system is no more secure than its least secure component. For instance if a system has several software security layers implemented in different stages, such as password protection and encryption, without considering the hardware layer or the risks bound to network connectivity then that system will most likely be unsecure. Building a secure system requires a design with security in mind at all time.

3.6.1 Security Policy

Implementing a good security policy before continuing on with the design and implementations stages is advised. A security policy "is a set of rules that state which actions are permitted and which actions are prohibited. The domain of a security policy is the set of entities, i.e. users, data object, machines, etc., that are governed by the policy" [5]

Implementing a good security policy is not easy. The difficulties arise from these simple reasons:

- There is no standard security policy that applies on all systems and situations
- It is difficult to identify all the threats that a system faces
- It is difficult to identify all the vulnerabilities the designed system has
- It is relatively difficult to enforce a security policy on the systems users, since this typically requires the users to invest time and effort.

There is however some security policy models than can be used as a frame or reference system to start with. The Bell-La Padula model is a confidentiality policy model, and the Biba model is an integrity policy model. Sometimes a combination of several policies can be used to achieve the desired security level.

There are some simple steps that could be taken to build a simple conceptual security policy:

- Identify the security needs and the threats that the desired system faces
- Identify the consequences of the threats if they breach the designed security policy
- Define a set of rules that must be implemented for the identified risks and threats
- Monitor and detect system vulnerabilities
- Define actions to be taken in case a threat is detected
- It is desired that the designed system is relatively flexible in its design. This makes it easier to encounter new and unexpected threats which were not considered during the design and implementation.
- It should be possible to update the security policy to better encounter new threats.

In addition to building a good security policy, there has to be good means to enforce the security policy. Users of the specific system should be at least familiar with the main security mechanisms involved and how they work. Users should be aware of the threats and risks that exist to enable a better understanding and motivation for using the applied security policy in the proper way.

Measures must be taken and weak points must be considered all the way from designing the hardware, implementing software, and networking. Considerations should be made about what way and by whom the designed system is to be used. The building process can be viewed as a cycle called the security life cycle as shown in the figure below.

![Conceptual Security Lifecycle Model](image)

**Figure 3.10:** Conceptual security lifecycle model

### Passwords

Passwords are the most often used form of authentication, primarily because passwords are free and not because they are the most secure way of authentication. A typical password has fixed size and is composed by a number of characters. The more constraints there are on what characters that can be used, the weaker a password gets. For example, if only lower-case letters from the English alphabet, 26 letters, are allowed and the password length is composed of 8 possible characters, then a user has $26^8$ available choices. Making
numbers and special characters available and enabling case sensitive characters provides stronger passwords.

Password characters should be chosen at random, preferably without any connection to the user. Choosing a name, birthday date, consecutive numbers from 1 – x (depending on size), only zeroes or age as password characters make weak passwords. The main problem facing users is that random passwords are very hard to remember. This becomes particularly difficult when each user has to remember several passwords for different devices and/or services. Using the same password for several or all services and devices is not advised, even if the password is very strong.

A simple guideline for choosing strong passwords:

- It should have no connection to the user
- It should be something hard to guess (preferably random)
- It should make use of all character types available i.e. mix numbers, case sensitive characters and special characters
- It should be something easy to remember

In addition to choosing a strong password, passwords should never be revealed to other users in any way. Some attacks, such as a phishing attack, have had great success only exploiting this simple mistake which is made by many users. There are some other aspects of using passwords which do not involve the user. Passwords must be stored in a secure way, preferably encrypted. Passwords should not be sent over networks. However, if that is required, then they should always be sent encrypted.
Chapter 4 - Security attacks
In the previous chapter we presented security requirements and implementations for our different solutions. In this chapter we continue with presenting some of the most important security attacks done on the security model presented previously. When possible, solutions and remedies to attacks will be discussed.

4.1 Security attacks

Many security attacks if not most take advantage of existing design flaws within the security model. Security attacks may occur or be possible for different reasons:

- Some security systems are not developed to be used in ways or environments where they are being used
- It is impossible to design a completely flawless system
- It is difficult to foresee technology advancement and the new kinds of attacks it makes possible
- Many security systems depend on the user having either knowledge of security threats or good judgement when using the system/device, which is not always the case
- There is big money involved in security attacks today and the trend is that this business is growing. Selling stolen and private information is lucrative.
- There is no global security standard and/or policy, nor one standard platform or application. Each company, service provider and country have their own implementation of security.
- Wireless technology is relatively new
- Internet is a global “arena”, making the number of potential attackers and their target groups very large and spread over several countries.

If we continue to use our CIA security model discussed earlier in chapter 3, we can divide or categorise security attacks based on the same model:

- Attacks on confidentiality: common for this kind of attacks is the desire to access or illegally obtain information. This is achieved with various methods and techniques. Examples: eavesdropping, phishing

- Attacks on integrity: here we have attacks with a purpose to modify data making it unusable. Examples: computer virus attack, human factor (unintentional attacks)

- Attacks on availability: these attacks usually have a simple purpose of making a service or resource unavailable to its legitimate users. Examples: DOS attacks, hardware failure (unintentional attack)

- Special attacks: there are many other attacks done on specific parts of a security system. Examples: attacks on cryptography, attacks on the users, attacks on software and hardware etc.
A further division of the kinds of attacks is that some are intentional attacks and others are unintentional. Hardware failure can for instance count as an attack on availability of a certain service. It is not an attack in the sense we are presenting in this report, but the result is the same for users depending on that specific service. However this report will focus on intentional attacks, since they are considered to do more damage and that unintentional attacks can generally be avoided with a good security policy which we will discuss in chapter 5. Some of the unintentional attacks which have direct effect or a close relationship to our security model will be mentioned and shortly overviewed.

Another classification of attacks is a division into active and passive attacks. Passive attacks, such as eavesdropping, suggest that an adversary by listening or reading gets access to information. Active attacks, such as man-in-the-middle attack, on the other hand have a slightly different approach and goal. An active attacker can with unauthorized modification of data deceive other users and making them using or relying on false information. Here the attacker is actively involved and making decisions to succeed with the attack.

[18]

In some cases passive attacks can be followed by active attacks. For instance an attacker may obtain valuable information, such as a password, by eavesdropping and then using that information to initiate another attack. The second attack can in this case be a man-in-the-middle attack, which is an active attack.

Some attacks are more common and used on different kinds of networks and platforms, whereas others have specific platforms to attack and may only work on a specific network. This report will divide attacks in four primary classes or categories which are directly connected to the reports main subject, with some exceptions for the purpose of giving a background on attacks or to make it easier to understand other attacks. Our four main categories are:

- General attacks
- Attacks on 2G/3G networks
- Attacks on WLAN networks
- Attacks on cryptography (cryptanalysis)
4.2 General attacks

Here we present some attacks that are common to both cellular networks and WLAN (Wi-Fi) networks.

Physical theft of mobile devices

This is considered to be one of the simplest and most obvious attacks. They are very common and have a great success rate, since mobile devices are carried along, small and easy to forget. Physical access to mobile devices makes it much easier for an attacker to carry out other attacks, such as obtaining private information to use for different purposes, or to simply sell the device to an unsuspecting user or criminal organisation. The success of these other attacks depends on how long it takes for the victim to realise that his or her device has been stolen and how long it takes the service provider to counteract the problem. [26]

Hacking Subscriber Information

An attack done on confidentiality with the main goals of getting access to private information for further use in other attacks or as a case of identity theft. The attack itself can be done in different ways. One method can be to make a successful attack on an information database containing subscriber information. Compromising the database can also be done in several ways. For further studies in the topic of database security see [31].

An interesting case concerning this attack is the rather significant breach of T-Mobile (a large service provider in the United States) subscriber databases in 2004. A hacker was able to access information on any of T-Mobile’s 16 million customers in Washington, including Social Security numbers and dates of birth, voicemail PINs and passwords providing customers with Web access to their T-Mobile email accounts. [32]

Overbilling Attack

Another type of possible attack is called an overbilling attack. Overbilling can be done when, for example, an attacker hijacks a subscriber’s IP address and then use that connection to download premium content or to simply use the hijacked internet connection for personal purposes without actually paying for the connection. In this case the legitimate user is billed for activity which he/she did not authorize. [19]
4.3 Attacks on 2G/3G networks

After more than one decade from the introduction of public 2G and 3G networks, service providers have developed experience and techniques to counter different attacks done on their networks. At the same time new methods and techniques to attack 2G/3G networks are constantly being developed. The two sides tend to balance themselves against each other, as in any other system. However introducing new technologies in the networks from the service provider part makes their networks vulnerable for several reasons:

- Service providers and mobile operators are building high speed wireless networks that are based on the Internet Protocol (IP) which allows users to do more while connected.
- Cellular networks are being opened up to the public Internet and to other data networks, making their 2.5G/3G networks more vulnerable to attacks.
- Mobile networks are evolving to IMS (IP Multimedia Subsystem), enabling interconnected networks all running on IP.
- New technologies always introduce new vulnerabilities, especially in security. Until the new technology is better understood and security flaws patched, attackers will do their best to exploit these vulnerabilities. [19]

New technology and development combined with an expanding network and new trends of purchasing and using multimedia content over the network have created new problems for service providers. An expanding user base with varied data-capable devices accessing content, communicating and making online business across multiple networks create new security implications. One of the main implications is that there will be much more traffic on cellular networks, which implies a higher possibility of attacks occurring from different sources. The more traffic there is, the harder it is to identify the threats. [19]

This new development can be summed in two important points:

- Cellular networks are being connected to other networks, such as the public Internet, other mobile operator networks, private networks (including company LANs), content servers, etc.
- Customers are using multiple device types – Smartphone’s with different operating systems and platforms, notebook computers and data-capable feature phones.

From a security perspective, this new market trend of making networks more open and available is a problem because there are now far more elements which have to be accounted for and which may be at risk. [19]

This new development of cellular networks has given a rise to new opportunities for attackers. Exploiting vulnerabilities in the new technology in cellular networks and the systems connected to those networks has created new market segments for attackers and criminals.
Historically attackers were usually individuals with some knowledge of the cellular network architecture and good knowledge in programming with the primary motivation of “bragging”. However in recent years we are witnessing a different development. Criminal organizations have begun to realise that great amounts of money can be made in trading information. For instance, criminals might be interested in acquiring subscriber information to either steal identities or to exploit billing and credit card information. [19]

4.3.1 Types of attacks

In chapter 2 we studied the 3G network architecture and mentioned how it was built upon the existing 2G networks. The main reasons for doing this were to minimise network deployment costs and also to ensure compatibility with the older 2G networks. We also mentioned that the 3G network was a move to an IP based network. These two important design decisions introduced new problems for service providers.

Because 3G networks were not all built from the ground up, they were not necessarily built with IP data security in mind. Service providers or mobile operators during the 2G era faced other types of attacks dealing with voice-centric security threats. IP data networks security was a relatively new concept to mobile operators. [19]

Attacks could originate from outside the mobile network, such as the public Internet or private networks. In this case the attackers are usually not users of the attacked network. Alternatively attacks can originate from within the mobile network. In this case attackers can be legitimate users of the attacked network or an attacker that impersonates a legitimate user. The origin of an attack can have some effect on the kind of attack done.

<table>
<thead>
<tr>
<th>Type of attack</th>
<th>Target</th>
<th>Purpose</th>
<th>Targets in CIA model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Malware: Worm, Virus, Trojan, SMS/MMS spam</td>
<td>Other users, network elements (content servers)</td>
<td>Harassment, denial of service/service interruption, information theft</td>
<td>C,I,A</td>
</tr>
<tr>
<td>Denial of service, buffer overflows, SYN flood, application layer attacks</td>
<td>HLR, AAA, content servers, signalling nodes</td>
<td>Attack the ability to provide service</td>
<td>A</td>
</tr>
<tr>
<td>Overbilling attack</td>
<td>Operator’s management elements (AAA, HLR, VLR, etc.)</td>
<td>Fraud</td>
<td>I</td>
</tr>
<tr>
<td>Signalling-level attacks (SIGTRAN, SIP) which involve modification, interception, DoS</td>
<td>Signalling nodes</td>
<td>Attack ability to provide service</td>
<td>I,A</td>
</tr>
</tbody>
</table>

Table 4.1: summarizes the various types of attacks to which GSM/3G operators are now vulnerable. [19]
Mobile operators and service providers are many times aware of existing and/or potential attacks and that their networks are vulnerable to them. However, many service providers choose to keep information about threats and attacks from public knowledge. There are several reasons for this kind of behaviour and it is closely combined with a company’s financial strategy and public image. [19]

### 4.3.2 Examples of attacks

In this chapter we will present and explain some of the most common attacks. Connection to our CIA security model will be made when possible.

**Malware attacks**

Malware is software with malicious purposes. Computer viruses, worms and Trojans are all considered to be a form of malware. Malware attacks target almost every aspect of our CIA security model.

According to a recent report from Kaspersky Labs, mobile malware evolved rapidly during its first two years of existence (2004 – 2006). In that period a wide range of malware programs targeting mobile devices appeared. These programs were very similar to older and existing malware targeting computers, such as computer viruses, worms and Trojans. For malware programs to be successful, an attacker needs to know beforehand what platform the targeting system is using. For instance, a Nokia device typically uses the Symbian[^24] operating system while some other device model may use the Google Android operating system etc. [33]

According to the same report from Kaspersky Labs, mobile malware families and modifications have increased by 200% during the years 2004-2009. Several malware types that have some aspects in common, such as a unique signature, are called a family and a modification is when a malware code modifies itself on each infection. [33]

The existence of different platforms make the attackers business difficult to conduct, therefore attempts were made to create cross-platform[^25] malicious code to make it easier to target large groups of users independent of the platform used. Creating cross-platform code requires that the included platforms have some common component that can be exploited. One of these common components is Java 2 Micro Edition. Almost every modern mobile device supports or uses Java, or Java applications which can be downloaded from the Internet. Creating malware targeting Java applications, attackers have found a solution to their problem of limiting their attacks to a specific platform. [33]

[^24]: An open source operating system designed primarily for Nokia smartphones.
[^25]: See definitions chapter
Some typical damage done by malware is for instance, deleting or copying files and/or contact lists from target devices, disabling an application, enabling remote control of the targeted device, sending SMS messages and calling paid services.  

An example of how Internet malware can impact a mobile operator’s operations is the Slammer/Sapphire worm outbreak in 2003, which inflicted great damage and destruction. Almost 20 percent of global Internet traffic was lost; 13,000 cash machines shut down; emergency services in Washington DC were lost for a short time; and commercial airline flights were delayed. Slammer’s impact was felt as far away as South Korea, where 27 million South Korean wireless subscribers lost their cellular service. [19]

Denial of Service Attacks (DoS)

A denial of service attack is “the prevention of authorized access to resources or the delaying of time-critical operations” [5]. It is an attack on availability in our CIA security model. One example of this kind of attack on cellular networks is “flooding” the network with SMS messages. A report published by two computer science professors at Pennsylvania State University in 2005, detailed how text messages originating from the public Internet could be used to flood or overwhelm a mobile operator’s short message services centre and, potentially, their ability to provide cellular voice services. [19]

DoS attacks target a scarce resource to make a service using that resource unavailable. An example of a scarce resource in a wireless network is the wireless spectrum. For instance available bandwidth of a wireless spectrum can be overwhelmed by excess data traffic. This attack has negative effects on both service providers in form of lost revenue and for the subscribers in form of no available service. An alternative is to create radio interference in the wireless spectrum, which can be considered as a DoS attack against the radio resources of the network. DoS attacks can also be launched from outside the network, such as from the public Internet into the operator’s mobile core to interrupt subscribers’ access to the Internet. [19]

Another form of DoS attack is known as distributed denial of service (DDoS) attack. DDoS attacks use brute-force methods to overwhelm the targeted network with data making it function “slower” or stopping it from functioning completely. This kind of attack requires huge amounts of data to be sent over the network and is difficult to achieve with only one computer or other compatible devices. A network of compromised computers, Botnets, can easily and effectively achieve this. From the networks point of view the attack will look like its originating from the botnet, whereas it in reality is remotely controlled by an attacker commencing a DDoS attack. [19]

Marketing harassment can in some cases be considered a DoS attack. Marketing harassment is when an attacker sends data traffic typically in the form of advertisement to a large group of subscribers. This advertisement can easily be sent as SMS. As we mentioned above this attack's purpose can be to overflow a network with useless data or having some other negative impact on its operation. However the attacks

---

26 See reference [34] for the full report.
27 See 4.5
28 See definitions
purpose can also be a way to attack the subscribers of the target network by creating inconveniences and possibly introducing some extra charges on their monthly bills. [19]

**Passive monitoring**

Passive monitoring is a technique used to capture traffic from a network by generating a copy of that traffic. This is an attack on anonymity of the specific network subscribers. Passive monitoring can provide attackers with a specific user’s location and give the ability to eavesdrop on voice conversations or data traffic from that user. Passive monitoring can be done easily, but it is inefficient and time-consuming because the attacker has to wait for the mobile device to be powered on or for a database failure to occur in the network which does not occur frequently. There are complete and capable hardware for passive monitoring easily available to anyone with some money and knowledge to use them. [12]

**Active monitoring**

The difference between active and passive monitoring is that active monitoring injects test packets into the network or sends packets to servers and applications, thus creating extra traffic on the target network. This method is more efficient than passive monitoring, but it requires the attacker to have advanced equipment that is more difficult to obtain. Here an attacker needs to have a device that enables base station functionality to fabricate messages to be sent on the network. This attack is also done on user anonymity and it provides an attacker with the victim’s current location and identity. [12]

**SIM Cloning**

SIM cloning is a method to steal a legitimate subscriber’s identity and use that identity in another mobile device for malicious purposes, such as making fraudulent calls where the bills will be sent to the victim [35]. SIM cloning can be achieved using easy methods when the attacker has physical access to the SIM card, and it is possible to find guides to how it could be done on the Internet\(^\text{29}\).

SIM cloning can also be done over the air-interface of the target network. This requires the attacker to have some advanced equipment that either imitates a legitimate base station or by using a rogue base station. This attack is more difficult in the sense that it requires advanced equipment and good knowledge of the target network architecture. However it is easier in the sense that obtaining a SIM is not an easy matter and cannot be done on large groups of subscribers at the same time.

\(^{29}\) See [36] for one such guide
4.4 Attacks on WLAN networks

In the previous chapters we studied the architecture of a principal WLAN network and explained the security mechanisms involved. Here we will study threats and attacks on our WLAN model network. We will explain types of attacks and present some examples of the most common attacks.

4.4.1 Types of attacks

Attacks on WLAN networks have some similarities with attacks done on 3G networks, studied in the previous chapter. However, there are some differences. WLAN networks that we study in this report are in many cases public networks which connect to the public Internet. Public networks, such as an Internet café, imply that there is less security involved. In some cases the service providers of a public network don’t require user authentication, meaning users can be anonymous.

The history of WLAN and its development is a little different from that of GSM and 3G networks. Users of WLAN networks, in its introduction period, were primarily PC users adopting the new wireless technology. Those users, including the malicious users, were well oriented in security mechanisms involved in the PC world, since it had existed for some years and security flaws were known. The use of wireless interface instead of wired connections introduced only some minor obstacles for attackers learning to exploit the new technology for personal purposes. In fact the new wireless technology introduced many new opportunities for attackers.

4.4.2 Examples of Attacks

In this chapter we present attacks done on WLAN networks. We will study the most common attacks and connect them to our CIA model in the same way as the previous chapter. This way it will be easier to compare both networks and their advantages and disadvantages concerning security.

Accidental and malicious association

This situation occurs when a user turns on his/her mobile device and it connects to some other geographically close wireless network than the one intended. The user could be unaware of what network the device is connected to. Accidental association can be considered an attack on confidentiality, since valuable information on the connected device can be transferred over the wrong network giving unauthorized users access to confidential information.

Malicious association is when attackers make wireless devices actively connect to a network. For instance an attacker can make his/her devices act as a legitimate access point (AP) on the target network. These types
of devices are known as soft AP:s and are easily created by an attacker. If the attack is successful the attacker can steal passwords, launch attacks on the wired network, or plant Trojans.

Man-in-the-middle (MITM) and injection attacks

“This attack is a form of active eavesdropping in which the attacker makes independent connections with the victims and relays messages between them, making them believe that they are in direct communication with each other when in fact the communication is controlled by the attacker” [35]. It is an attack that targets confidentiality and integrity of a communication session. MITM attacks can be done when an attacker impersonates a network resource and require the attacker to have great knowledge about the network. [26]

When \( A \) wants to communicate to \( B \) it initiates a connection and sends message \( m_1 \). The attacker, \( M \), in the figure above must be able to intercept all messages going between the victims, \( A \) and \( B \), and inject new messages. \( M \) will therefore intercept message \( m_1 \) and continue by sending message \( m_2 \) making \( B \) believe that it was sent by \( A \). The same thing will be done from the other end, from \( B \) to \( A \) with \( m_3 \) and \( m_4 \). \( M \) will complete an initiated connection from \( A \) to its intended destination \( B \) and proxy all communications of \( A \) and \( B \). A successful attack gives the attacker ability to modify communications, inject data, or eavesdrop on communications gaining access to unauthorized information. A man-in-the-middle attack can only be successful when the attacker can impersonate each endpoint to the satisfaction of the other. For example, an attacking entity within reception range of an unsecure WLAN AP, can insert itself as a man-in-the-middle. [26]

Modern cryptographic protocols, such as SSL, include some form of endpoint authentication specifically to prevent MITM attacks.

Passive eavesdropping and traffic analysis

Traffic analysis is a passive attack with the main purpose of gathering sufficient information about the network to use later in an active attack. Wardriving is a process where an attacker physically moves between different areas (typically by car) to gather information about active AP:s of networks to be attacked. Passive eavesdropping is similar to traffic analysis where an attacker can gather information about a network and

---

This attack is a form of active eavesdropping in which the attacker makes independent connections with the victims and relays messages between them, making them believe that they are in direct communication with each other when in fact the communication is controlled by the attacker” [35]. It is an attack that targets confidentiality and integrity of a communication session. MITM attacks can be done when an attacker impersonates a network resource and require the attacker to have great knowledge about the network. [26]

![Figure 4.1: conceptual figure of a man-in-the-middle attack](image)

When \( A \) wants to communicate to \( B \) it initiates a connection and sends message \( m_1 \). The attacker, \( M \), in the figure above must be able to intercept all messages going between the victims, \( A \) and \( B \), and inject new messages. \( M \) will therefore intercept message \( m_1 \) and continue by sending message \( m_2 \) making \( B \) believe that it was sent by \( A \). The same thing will be done from the other end, from \( B \) to \( A \) with \( m_3 \) and \( m_4 \). \( M \) will complete an initiated connection from \( A \) to its intended destination \( B \) and proxy all communications of \( A \) and \( B \). A successful attack gives the attacker ability to modify communications, inject data, or eavesdrop on communications gaining access to unauthorized information. A man-in-the-middle attack can only be successful when the attacker can impersonate each endpoint to the satisfaction of the other. For example, an attacking entity within reception range of an unsecure WLAN AP, can insert itself as a man-in-the-middle. [26]

Modern cryptographic protocols, such as SSL, include some form of endpoint authentication specifically to prevent MITM attacks.

Passive eavesdropping and traffic analysis

Traffic analysis is a passive attack with the main purpose of gathering sufficient information about the network to use later in an active attack. Wardriving is a process where an attacker physically moves between different areas (typically by car) to gather information about active AP:s of networks to be attacked. Passive eavesdropping is similar to traffic analysis where an attacker can gather information about a network and

---

See definitions
also read data transmitted in a session to get various information about network packet characteristics. This attack is particularly easy to employ when no encryption exists. [37]

Denial of service

DoS attacks on WLAN networks are done in a similar way as done in 3G networks, with only minor differences in respect to the different architecture of the WLAN network. A DoS attack can aim to “bring down” the target system making it unavailable to its users. This can be done by sending vast amount of traffic to the networks AP:s making it unable to respond. [37]

4.4.3 Attacking WLAN security mechanisms

Here we present some of the attacks targeting the WEP security mechanism described in 3.4.1 and attacks on WPA security mechanism which was described in 3.4.2.

Replay attack (WEP)

The WEP security mechanism of WLAN networks has many flaws as we described in previous chapters. One of these flaws is that WEP does not add timestamps to packets sent over the network. The same Ethernet packet is always converted into the same WEP packet, giving an attacker the opportunity to exploit this weakness. An attacker can for instance delay messages or retransmit them.

Eavesdropping on WEP

Eavesdropping on the target network using WEP can provide an attacker with the RC4 key and is considered a passive attack. Depending on the amount of network traffic, and the number of packets available for inspection, a successful key recovery could be possible in only one minute. If an insufficient number of packets are being sent, there are ways for an attacker to send packets on the network and thereby stimulate reply packets which can then be inspected to find the key. This attacks ease of use made it a good method for attackers and criminal organisations, which implemented it and developed automated tools to make this attack even better. It is possible to perform the attack with a personal computer with free software available on the Internet, such as aircrack-ng\textsuperscript{31}, to crack any WEP key in minutes. [39]

Fragmentation-only attack (WEP)

After eavesdropping a single packet, an attacker can immediately transmit arbitrary data on the network. The eavesdropped packet can then be decrypted one byte at a time (by transmitting about 128 packets per byte to decrypt) to discover the local network IP addresses. In the case that the target network is connected to the Internet and the attacker has access to the routers MAC address and the networks IP prefix, it is possible for the attacker to capture data and retransmit it to the Internet, where it will arrive in “clear text”. The attacker can use 802.11 fragmentation\textsuperscript{31} to replay eavesdropped packets while crafting a new IP header onto them.

\textsuperscript{31} See definitions
The AP can then be used to decrypt these packets in real-time within a minute of eavesdropping the first packet. [39]

**Password cracking (WPA)**

Shared-key WPA is vulnerable to password attacks. Protection against brute-force attacks are provided by choosing strong passwords, see 3.6.1 for password guidelines and 4.5 for brute-force attacks.
4.5 Cryptanalytical attacks

Many of the security mechanisms and schemes involved in the networks studied make use of encryption. This is why there are many attacks that target the encryption of a security mechanism.

Cryptanalytic approaches

- Differential cryptanalysis is a chosen plaintext attack that studies the propagation of input differences to output differences in iterated transformations [2]

- Linear cryptanalysis is a known plaintext attack that is based on “effective” linear approximate relations between the plaintext, the ciphertext, and the key. [2]

Ciphertext-only attack

The key or plaintext is revealed exclusively by means of the ciphertext. This method is one of the most difficult cryptanalysis methods. If little is known of the rules of the ciphertext to be able to exploit them, only one obvious thing remains: trying every possible key. This is called brute-force attack (exploiting the key space; exhaustion method). Often however it is sufficient to try just a few keys. [2]

The attack is completely successful if the corresponding plaintexts can be deduced, or even better, the key. The ability to obtain any information at all about the underlying plaintext is still considered a success. For example, if an adversary is sending ciphertext continuously to maintain traffic-flow security, it would be very useful to be able to distinguish real messages from nulls. Even making an informed guess of the existence of real messages would facilitate traffic analysis. This attack proved to have some successful results on the stream cipher A5/1 used in GSM networks, although requiring great computational power. [2]

Known-plaintext attack

Part of the plaintext is known in addition to the ciphertext, and used to reveal the remaining plaintext, normally by means of the key. This is perhaps the most important cryptanalytic method because it is much more powerful than a ciphertext-only attack and normally possible: the attacker guesses certain words in the text; the beginning of the text is fixed; known, uncritical plaintexts are encoded with the same key as confidential plaintexts.

Encrypted file archives such as ZIP are also very prone to this attack. For example, an attacker with an encrypted ZIP file needs only one unencrypted file from the archive which forms the "known-plaintext". Then using some publicly available software they can quickly calculate the key required to decrypt the entire archive (unless a secure algorithm such as AES is used). To obtain this unencrypted file the attacker could
search the website for a suitable file, find it from another archive they can open, or manually try to reconstruct a plaintext file armed with the knowledge of the filename from the encrypted archive. [2]

**Chosen-plaintext attack**

This is also a plaintext attack, except that the attacker can choose the plaintext so that the attack becomes possible in the first place, or it will become easy. In this case, the cryptanalyst is active himself: he or she needs some clever guessing to deliberately introduce some text. [2]

**Chosen-ciphertext attacks**

Here an attacker deliberately introduces a certain ciphertext and gains access to the plaintext generated from that ciphertext. The attacker can then use this information to calculate other plaintexts. This attack is used against digital signatures described in 3.5.5

**Adaptive-chosen-plaintext attack**

This is a repeated attack with selected plaintext, where the plaintext deliberately introduced is selected dependent on the current state of the cryptanalysis. Algorithms used in ciphering devices with permanently burnt-in keys have to be resistant against this method. [2]

**Ciphertext-ciphertext attack**

In this case the plaintext is encrypted with two different methods. The attacker can exploit this in different ways. In general, a method is already broken so that everything boils down to a plaintext attack. Such an attack is always based on a ciphering error. Good cryptographers use a different plaintext for each method. [2]

**Brute-force attacks**

“A brute force attack is a strategy used to break the encryption of data. It involves traversing the search space of possible keys until the correct key is found.” [35]

The success rate of a brute-force attack depends on the key size of the targeted encryption scheme. Large key sizes make brute-force attacks infeasible to compute with conventional hardware and in feasible time.

**Attacks on symmetric key**

The process of encryption previously described for stream ciphers and block ciphers is known as Electronic Code Book (ECB) encryption mode. With ECB, the same plaintext input always generates the same ciphertext output. As Figure 4.2 below illustrates, the input text of "FOO" always produces the same ciphertext. This is a potential security threat because eavesdroppers can see patterns in the ciphertext and start making
Chapter 4  Security attacks

educated guesses about what the original plaintext is. There are two encryption techniques to overcome this issue namely initialization vectors and feedback modes which will not be discussed in this thesis [17].

![Diagram](keystream_xor_plaintext_ciphertext.png)

**Figure 4.2:** ECB encryption mode [52]

**Insertion attacks**

A relatively easy attack done on stream ciphers. An attacker intercepts the ciphertext without initially knowing any bits of the key stream and the plaintext. If the attacker successfully introduces a few bits or bytes into the plaintext and in intercepting the ciphertext encrypted with the same key stream, then it may be possible to compute the key stream and the plaintext from the insertion point onwards. [2]
Part III

Chapter 5 - Discussion and conclusions
In this final chapter our findings will be presented and discussed. Using the discussion the questions posed in the first chapter will be answered in the best way possible. There will also be a discussion about the current market situation of mobile payment systems and their financial aspects will be presented. Finally some word will be said on future work in this field.

5.1 Mobile payments

Mobile devices are getting more advanced rapidly and the number of total users with capable devices is growing equally rapidly. However investors and service providers alike are not seeing the “boom” they have hoped for. There are many reasons for this that can be discussed extensively, but some reasons are more important than others.

New technology takes time to get “mature” and user acceptance. That is the case for almost any new technology unless it is something revolutionary that will change our lives forever, or it is something with no other established alternatives. This is one of the biggest obstacles for mobile and wireless payments and transactions. There are alternatives that are established and well known. If we only consider online payments, and neglect other types of transactions such as paying with physical money, mobile payments’ greatest competitor is online transactions made by wired devices. Online transactions itself is relatively new whether done wired or wirelessly. Doing online business with wired devices at the comfort of our homes is difficult to beat. Wired solutions provide many other advantages too:

- They are cheaper in terms of costs connected to having access to the Internet (which is where online business is made)
- There are less security threats for wired devices than for wireless. Wireless solutions face all the threats of wired solutions as well as several threats that are only attached to wireless
- Wired technology has existed for a longer period of time
- Security mechanisms of wired connections and devices are well known by the average user when compared to wireless.

There has to be some motivation for users and customers to change the way they do their online business. What does mobile or wireless online transactions provide that cannot be found in the wired solution?

In subchapter 2.1 we viewed some of the benefits of a mobile solution. The most important criteria for mobile payments is that transactions can be made “on the go”, and it is an important point depending on the type of the user. If it is a user only paying bills once a month, then mobile payments is not an attractive option. However, if it is someone who travels often or have several transactions to make daily, then mobility is important. Looking back at subchapter 2.3, we notice that current mobile payment use areas are limited to small payments and transactions. This contributes to the fact that the security level mobile payments should provide for these small payments should be balanced to the value we assign to these small transactions.
Mobile payments in their current situation are not suitable for larger transactions due to inadequate security level.

A great problem facing online business, both wired and wireless is that the Internet is a global network. This means that criminal organisations that are in the information trading business have much greater liberties than other types of criminals. For example you could be making an online transaction in Sweden and be a victim for a cyber criminal based in the USA trying to steal your private information. There is no global law enforcement authority, at least when we consider cyber criminals, which can act in an effective way. There are other problems connected to this matter too. Something that is illegal in one country can be perfectly legal in another. Perhaps there should be some global standard on what is and what is not allowed to do on the Internet. However keeping the Internet "free" has its advantages too.

5.2 Technology

What is the current state of the mobile solutions used for mobile payments? Do mobile solutions provide enough features to satisfy the typical customer?

In terms of data throughput speed, the technology offered by 3G networks is still limited when compared to other solutions. Recall table 2.1 from chapter 2, 3G networks offered a download rate of 226 Kbps and an upload rate of 30 Kbps. There has been some upgrades to the 3G network making them better in terms of data throughput, such as HSPA 7.2 offering 1.4 Mbps download and 700 Kbps upload on average. These numbers are considered satisfactory for Internet browsing and making online purchases. In the case of GSM, the numbers look much worse. Even mobile handsets using EDGE technology do not provide satisfactory speeds. In the case of WLAN, data throughput has always been one of its greatest advantages, considering the 802.11g standard with download speeds of 54 Mbps (theoretical maximum), which even in its worse case is several magnitudes of what 3G networks offer.

In terms of service range and coverage, we have a very different situation. GSM networks which have been in service for a long time, have the best service coverage and base stations geographic distribution globally. 3G networks are close to GSM in terms of coverage in many countries where they are available. Upgrading the GSM networks to 3G is expensive and in many countries this upgrade has not yet been done. WLAN networks have the worst coverage compared to the other networks. There are only few cities in the world with WLAN coverage, and even then it’s done on an experimental level. When range of service is considered the situation is almost the same. GSM and 3G networks have a service range of several kilometres, whereas WLAN networks suffer in network speed and connectivity after short distances from the nearest access point. Using WLAN indoors affects service quality and the presence of other devices, such as microwave ovens, interfering with its network frequency, can negatively affect it too.

In terms of service and device costs, we have an interesting situation. GSM, which is the oldest network, provides low cost services. However its low speed connection makes it unsuitable for making online
purchases. The same situation applies for GSM devices. 3G devices are a little more expensive than GSM devices and service costs are a little higher too. They provide however good connection speeds. When considering the case of WLAN networks, the service costs are comparable to that of the 3G network. The interesting thing here is the type of device to use. As we mentioned in previous chapters, many cellular phones provide WLAN functionality as well as GSM and 3G. We could also use a laptop for our purpose, but the question is if they can be considered to be truly mobile. It depends on the definition of mobility. Using a laptop device “on the go” is rather difficult considering their size and weight. If “on the go” means sitting in a bus, then a laptop is preferable. 3G devices supporting WLAN functionality tend to cost a little more than regular 3G devices and are comparable to the cheapest laptops available.

In terms of features each device has to offer, there are some interesting points. As we mentioned, there are cellular phones with functionality supporting GSM/3G and WLAN. We therefore compare these devices to laptops in terms of features. However, remember that laptops can also make use of GSM and 3G networks using a modem. The most important two features that are interesting to compare are screen size and battery life. In terms of battery life then cellular phones have longer battery life than laptops. In terms of screen size then laptops are superior. Laptops have larger screens and with better resolution, making Internet browsing much easier to manage. Recent cellular phone models have been gifted with larger screens and better screen resolution. However, they will always have the smaller format to maintain their mobility property. Small screens and buttons which are relatively cumbersome to manage browsing the Internet on the smallest devices difficult. One possible development in this area is the customization of Internet to better fit smaller screens.

5.3 Security

For an average user one important security requirement when considering a mobile payment system is the user’s privacy. We need to feel confident that our personal information sent over a network does not fall into the wrong hands. Many threats and attacks have come to public knowledge in recent years making the average user suspicious to online payments in total and more so towards wireless online payments, because it’s new and still not well known.

The problem here is not that wireless solutions are completely unsecure or only provide a low level of security. If we look back to previous chapters, it is relatively easy to identify a pattern in threats and attacks done on our studied networks. No solution is completely secure. There is no such thing as a system with perfect security and this applies to all systems irrespective of if it is a mobile payment system, a medical or a military system. A medical system would however require a very high level of security and fail tolerance for it to be acceptable. Not loosing our private information can be considered as an acceptable level of security for mobile payments. Considering that small payments dominate in mobile payment systems today, our private information has greater value than the total value of the individual transactions made. Note that a higher level of security perhaps makes the average user more confident in using our mobile payment system; however the huge costs attached to it are not something that customers appreciate. To maintain (among other
things) user privacy, mobile networks provide several security mechanisms, many of which depend on encryption.

Confidentiality

3G/GSM and WLAN networks provide user confidentiality by using encryption. GSM used the A5/1 algorithm which is a stream cipher (described in 3.2.1), 3G networks use a block cipher called KASUMI (described in 3.5.4) and WLAN networks can use the WEP standard (described in 3.4.1) or the more secure WPA standard (described in 3.4.2).

The A5/1 proved to have many flaws in its design enabling many attacks against GSM networks, such as the ciphertext-only attack (described in 4.5) and also brute-force attacks. It is worth mentioning that these attacks are relatively difficult and require advanced hardware. Having knowledge about the target to be attacked can be useful to guess secret information making the attack simpler by reducing the search space.

For 3G networks a new block cipher was developed with the flaws of A5/1 in mind. The block cipher, KASUMI, provided confidentiality and integrity algorithms. This new block cipher proved to be somewhat vulnerable against chosen-plaintext attacks (described in 4.5), however this attack can be considered impractical due to large time complexity.

The WEP standard used in WLAN networks also had many flaws which were soon to be exploited by attackers. Replay attacks which exploited the absence of timestamps on sent messages over networks using WEP and fragmentation-only attacks (described in 4.4.3), which is a form of eavesdropping on WLAN networks. This attack can be efficient and practical in terms of time complexity, requiring only some knowledge about the target.

As we mentioned before confidentiality, which is the prevention of loosing information to unauthorized entities is one of the most important security concerns for the average user. Comparing GSM/3G networks with WLAN networks in terms of confidentiality, GSM and 3G networks provide a better alternative. One important point in this context is that most of GSM and 3G networks’ security mechanisms are not made public when introduced. This secrecy has advantages in that it becomes more difficult for attackers to mount their attacks without fully understanding the security architecture. There are disadvantages to this secrecy though. When a security mechanism, such as an encryption algorithm, is made public there will be many more people testing this specific algorithm and detecting weaknesses. In other words these “public algorithms” receive much more feedback from people testing them, thus making their final versions much more secure.

Authentication

GSM and 3G networks provided authentication by using a SIM which authenticates itself to an AUC through a challenge response process similar to that of a symmetric-key method (see 3.5.3).
authentication process in GSM was considered to be relatively secure and was reused in 3G networks. However some of the algorithms used in GSM, such as the A5/1, were replaced with better ones.

WLAN networks provided authentication through the WEP standard (see 3.4.1) by using a shared key authentication, which uses a four-way challenge response handshake.

Authentication is an equally important aspect of any secure system. Authentication of users to their specific networks and vice versa can be considered a necessary requirement. An overbilling attack (described in 4.2) is one type of attack exploiting weaknesses in an authentication process. It would be unacceptable if attackers could easily impersonate their targets and send bills to them. A simple example of this is a “casual” attacker, which could be your neighbour, piggybacking on your WLAN using WEP and make purchases using your personal information, which can be easily accessed because the attacker might know who you are.

All networks also provide means to maintain availability of service. GSM and 3G networks periodically update their databases of active users and have redundancy in the amount of base stations to maintain service availability. WLAN networks use similar methods, having redundancy in access points and recovery mechanisms to maintain availability. There are some known attacks targeting the availability aspect of networks such as denial of service attacks, which are relatively widespread and easy to mount. Maintaining availability is not as crucial as maintaining confidentiality and authentication provided that they do not occur often. We have all experienced web sites and services that are unavailable for a period of time. As long as they return to normal functionality after a short period of time and that they are functioning correctly most of the time, then they will satisfy our needs. Note that service could be unavailable due to hardware and software failure and need not to be a result of an attack.

Although most security mechanisms we have studied are already broken or will be soon, it has to be mentioned that this is a process than can be viewed as an arms race between security developers and attackers. Each time a new security mechanism is developed there will be attackers studying them in detail to exploit design flaws. Each time a security mechanism is broken, there will be a new one to replace it. For instance WEP was replaced by WPA. This can also be viewed as a cycle of products replacing each other. When a new security mechanism or a new encryption scheme is released it has good security because it takes time for attackers and others to understand the new technology. After some time the new technology will be well understood and detected security flaws will be exploited. This is when it is time to release new products again, and this cycle continues.
5.4 The future

Mobile devices and their networks are growing in popularity as we have mentioned in previous chapters. New networks and devices are developed to satisfy public demands. 4G networks are to be deployed very soon, promising higher data rates, better security and richer content. Two competing technologies are currently being developed for 4G networks, the LTE and WiMAX. These new technologies and their networks mean that huge investments are made in development and it also means that wireless networks, and mobile payment systems as a subgroup, are here to stay. User acceptance to mobile payments will hopefully “mature” by the time 4G networks have arrived.

Considering the future of WLAN networks, there are also some interesting points. If WLAN networks are to be a serious competitor to the upcoming 4G network, then large scale deployment is necessary. This is a big challenge technically and economically and has yet not been done other than some experimental attempts in some large cities. Another major problem here is spectrum availability, which is something limited and has to be licensed (by the government) to use.

For widespread deployment to occur, WLAN devices and service providers must meet consumer demands for:

- Easier access: It would be desirable to have public WLAN networks free to access by, for instance, paying a download/upload size fee
- Better quality of service: service range and availability could be increased
- Flexible pricing: for example, the option to be charged per transaction rather than for the duration of a connection as is currently the case
- Speed: data rates could be further increased to support high-end applications and multimedia content.

A possible future is the merging of the most popular networks in one “super network”. This will provide a better standardization process for future products and an easier situation for device manufacturers. It will also provide a better environment for developers of security mechanisms, because more people will be actively engaged in working with one or some few standards, in contrast to today’s situation. From previous chapters we have seen that 3G and WLAN networks have many things in common and even though they have many differences it is possible to build something that makes use of the best from each existing network without being victim of too many security flaws.

Considering the future of involved security mechanisms, it is still unknown what 4G networks will introduce in terms of new security mechanisms and/or if that information will come to public knowledge anytime soon. New encryption methods are and will be developed each time it is considered necessary to replace a broken encryption method. Note that developing new encryption methods that provide good security is not an easy matter. “A cryptographer designing a cryptographic algorithm has to cover himself/herself against everything, which is impossible. On the other hand it is sufficient for a cryptanalyst or an attacker to find one design flaw, making the algorithm considered insecure” [2].
5.5 Conclusions

After studying security requirements with attacks done on each network and discussed the findings, it is time to draw some conclusions:

Mobile payments’ current market situation is somehow less than what was expected by investors and service providers prior to network deployment. There are several reasons for this, among them: user acceptance, mobile transactions costs and the existence of other established alternatives.

Does a mobile payment system provide a good level of security to the average user? Yes but in connection to user awareness. Many of the threats that were studied in previous chapters, such as malware attacks, can be avoided if the user is aware of the risks and has some knowledge of security mechanisms involved. Investment in time and effort to learn the most important aspects of security involved in a specific system provides a better overall experience and future savings in form of time, effort and money.

From chapter 3 we can conclude that the studied networks satisfy security requirements according to the CIA model presented in the same chapter. Some components were vulnerable to specific attacks, but new security mechanisms with better security are developed frequently to address the most important flaws.

From chapter 4 we viewed security threats and attacks done on the networks of our study. From this chapter we can conclude that some attacks impose great threats to the security of a mobile payment system. However, when the nature of the transactions is taken into account and balanced to the costs of implementing further security mechanisms, we conclude that mobile payment systems provide a good level of overall security to the average user, again provided that the user has some knowledge of the security involved.

The use of a mobile payment system can be considered secure for small transactions and average users/customers. Larger transactions require a better level of security and better standards regarding service providers’ ability to provide higher security in their networks and also a better user acceptance level.

The current state of mobile devices, regarding level of technology and available features, still has room for improvement to satisfy customers’ demands. New devices with better properties, such as longer battery life and better screens, are being developed frequently, and network technology, such as service range and data transmission speeds, is being improved too.

Regarding the future of mobile payments, the conclusion is that wireless networks, their devices and new mobile payment solutions have good future prospects. Huge investments are already been made in the upcoming LTE and WiMAX technologies (which will be 4G networks) and network deployment.
5.6 Future work

This report’s purpose was to introduce the reader to some wireless networks and the security mechanisms involved. Security attacks were only viewed, and there was no opportunity to test these attacks in practice. Future work on this subject can be done considering different aspects and in different directions:

- Further studies can be done on each solution introduced in this report in addition to the possibility of adding new solutions not viewed in this report
- A possible approach could be to concentrate on one network and explain its components, security mechanisms and threats in detail
- Another approach is to only concentrate on security attacks, explain them in detail and make practical experiments to verify their results
- The study of mobile payments systems can also be done from another point of view. It could, for example, be studied from an economic perspective.
Definitions

**802.11 fragmentation**: the optional fragmentation function enables an 802.11 station to divide data packets into smaller frames. This is done to avoid needing to retransmit large frames in the presence of radio frequency interference (RFI). The bits errors resulting from RFI are likely to affect a single frame, and it requires less overhead to retransmit a smaller frame rather than a larger one.

**AES**: Advanced Encryption Standard, is a symmetric-key encryption standard. It comprises three block ciphers, AES-128, AES-192 and AES-256. Each of these ciphers has a 128-bit block size, with key sizes of 128, 192 and 256 bits, respectively. For detailed information see [2]

**aircrack-ng**: is a network software suite consisting of a detector, packet sniffer, WEP and WPA/WPA2-PSK cracker and analysis tool for 802.11 WLAN. [35]

**Botnet**: botnets are computers that have been compromised by attackers, generally through the use of Trojans (malware disguised as or embedded within legitimate software).

**Certificate**: is an electronic document which uses a digital signature to bind together a public key with an identity information such as the name of a person or an organization, their address, and so forth. The certificate can be used to verify that a public key belongs to an individual. [35]

**Circuit switching**: in telecommunications, a circuit switching network is one that establishes a circuit (or channel) between nodes and terminals before the users may communicate, as if the nodes were physically connected with an electrical circuit. [http://en.wikipedia.org/wiki/Circuit_switching]

**Counter mode**: is one of the modes of operation used with block ciphers in cryptography that turns a block cipher into a stream cipher.

**Cipher block chaining**: one of the modes of operation used with block ciphers in cryptography that performs an XOR operation between each block of plaintext and the previous block of ciphertext.

**CKK**: Complementary Code Keying is a modulation scheme used with WLAN networks that employ the IEEE 802.11b specification. [35]

**Cross-platform**: in computing, cross-platform, or multi-platform, is an attribute conferred to computer software or computing methods and concepts that are implemented and inter-operate on multiple computer platforms. [35]

**DES**: Data Encryption Standard, is a block cipher based on a symmetric-key algorithm that uses a 56-bit key size. DES is considered insecure because of its small key size and has been replaced with AES. For detailed information see [2]

**DSSS**: Direct sequence spread spectrum is a spread spectrum method where data is pushed through a binary encoding process that spreads the data by combining it with a multi-bit pattern. The resulting data is hidden and inflated. [26]

**EDGE**: is a backward-compatible digital mobile phone technology that allows improved data transmission rates, as an extension on top of standard GSM. [35]
Definitions

**Entropy:** in computing, entropy is the randomness collected by an operating system or application for use in cryptography or other uses that require random data. This randomness is often collected from hardware sources, either pre-existing ones such as mouse movements or specially provided randomness generators. [35]

**FHSS:** Frequency Hopping Spread Spectrum is a spread spectrum method where the transmitter and receiver hop from one frequency to another in prearranged synchronized patterns. Hops occur frequently with very little time being spent on any single frequency. This reduces the possibility of interference with other devices and enables several overlapping FHSS systems to operate at the same time. [26]

**Feistel:** is a symmetric structure used in the construction of block ciphers. The Feistel structure has the advantage that encryption and decryption operations are very similar, even identical in some cases, requiring only a reversal of the key schedule. Therefore the size of the code or circuitry required to implement such a cipher is nearly halved [35]. For more information see [2].

**Handshake:** is an automated process of negotiation that dynamically sets parameters of a communications channel established between two entities before normal communication over the channel begins. It follows the physical establishment of the channel and precedes normal information transfer. [35]

**IR:** infrared radiation is an electromagnetic radiation that can be used to transmit data among devices in short-range wireless communications. [35]

**LFSR:** Linear Feedback Shift Register, is a shift register whose input bit is a linear function of its previous state.

**MISTY1:** a secret key cryptosystem developed by Mitsubishi Electric Corporation with a 128-bit key size and a 64-bit block with variable number of rounds. They are designed to have provable security against differential and linear cryptanalysis. They are designed to have high speed encryption and decryption properties on hardware platforms and software environments. [53]

**OTP:** a one-time pad uses a key stream of completely random digits. The key stream is combined with the plaintext digits one at a time to form the cipher text. However, the key stream must be at least the same length as the plaintext, and generated completely at random. OTP systems are very difficult to implement in practice, and as a result the OTP has not been widely used, except for the most critical applications.

**Packet based technology:** also called packet switching is a digital networking communications method that groups all transmitted data into blocks, called packets. [http://en.wikipedia.org/wiki/Packet_based]

**RTP:** Real-time Transport Protocol defines a standardized packet format for delivering audio and video over the Internet. [35]

**Spread-spectrum:** is a technique where signals generated in a particular bandwidth are deliberately spread over the frequency domain, resulting with a signal with wider bandwidth.

**Soft AP:** a software access point
Definitions

**TDM:** time-division multiplexing  is a type of digital or analogue multiplexing in which two or more signals or bit streams are transferred apparently simultaneously as sub-channels in one communication channel, but are physically taking turns on the channel. [35]

**WiFi:** trademark of the Wi-Fi Alliance that manufacturers may use to brand certified products that belong to a class of wireless local area network (WLAN) devices based on the IEEE 802.11 standards. [35]

**XOR:** also called *exclusive or*, is the logical operation exclusive disjunction of two operands that results in a value of true if exactly one of the operands has a value of true.
Abbreviations

3GPP: 3rd Generation Partnership Project
AES: Advanced Encryption Standard
AP: Access point
ATM: Automated Teller Machine
CDMA: Code Division Multiple Access
CKK: Complementary Code Keying
EDGE: Enhanced Data rates for GSM Evolution
GPRS: General Packet Radio Service
GPS: Global Positioning System
GSM: Global System for Mobile Communications
HSPA: High Speed Packet Access
IEEE: Institute of Electrical and Electronics Engineers
IMS: IP Multimedia Subsystem
IP: Internet Protocol
LTE: Long Term Evolution
MITM: man-in-the-middle attack
PCMCIA: Personal Computer Memory Card International Association
PGP: Pretty Good Privacy
PIN: Personal Identification Number
PRNG: Pseudo-Random Number Generator
RNG: Random number generator
SIM: Subscriber Identity Module
SMS: Short Message Service
SSL: Secure Socket Layer
TLS: Transport Layer Security
Abbreviations

**UMTS**: Universal Mobile Telecommunications System  
**VPN**: Virtual Private Network

**WAP**: Wireless Application Protocol

**WEP**: Wired Equivalent Privacy

**WiMAX**: Worldwide Interoperability for Microwave Access

**WLAN**: Wireless Local Area Network

**WPA**: WiFi Protected Access
References


7. Findahl, Olle, Svenskarna och Internet, 2009, World Internet Institute, PDF [7]


References


23. Mobile commerce report, Durlacher research Ltd, PDF [23]


28. Shepard, Steven, Telecom crash course, 2005, Book [28]


30. Borisov, Nikita & Goldberg, Ian & Wagner, David

31. Thuraisingham, Bhavani M, Database and applications security : integrating information security and data management, 2005, Book [31]


33. Gostev, Alexander, Mobile Malware Evolution: An overview Part 3, Kaspersky Labs, 2009,

34. William Enck, Patrick Traynor, Patrick McDaniel, and Thomas La Porta, Exploiting Open Functionality in SMS-Capable Cellular networks, Systems and Internet Infrastructure Security Laboratory, Department of Computer Science and engineering Pennsylvania University, 2005, PDF [34]


37. Passive & Active Attacks Against Wireless LAN’s, Mohteshim Hussain, University of Hertfordshire, England, UK, PDF [37]

38. The Final Nail in WEP’s Coffin, Bittau, Handley and Lackey, 2006, PDF [38]

References

40. Halvorsen, Finn Michael & Haugen, Olav, Cryptanalysis of IEEE 802.11i TKIP, 2009, PDF [40]


53. Matsui, Mitsuru, New Block Encryption Algorithm MISTY, Information Technology R&D Center, Mitsubishi Electric Corporation, PDF [53]
Upphovsrätt

Detta dokument hålls tillgängligt på Internet – eller dess framtidiga ersättare – från publiceringsdatum under förutsättning att inga extraordinära 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 garantiera äktheten, säkerheten och tillgängligheten finns lösningar av teknisk och administrativ art.

Upphovsmannens ideella rätt innefattar rätt att bli nämd 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 fölagets hemsida http://www.ep.liu.se/

Copyright

The publishers will keep this document online on the Internet – or its possible replacement – from the date of publication barring exceptional circumstances.

The online availability of the document implies permanent permission for anyone to read, to download, or to print out single copies for his/hers own use and to use it unchanged for non-commercial research and educational purpose. Subsequent transfers of copyright cannot revoke this permission. All other uses of the document are conditional upon 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/

© Ahmed Ali.