LTE SECURITY

Dan Forsberg
Security Consultant, ISECure.fi Oy, Finland

Günther Horn
Senior Security Specialist, Nokia Siemens Networks, Germany

Wolf-Dietrich Moeller
Senior Security Specialist, Nokia Siemens Networks, Germany

Valtteri Niemi
Nokia Fellow, Nokia Corporation, Switzerland

A John Wiley and Sons, Ltd., Publication

LTE SECURITY

LTE SECURITY

Dan Forsberg
Security Consultant, ISECure.fi Oy, Finland

Günther Horn
Senior Security Specialist, Nokia Siemens Networks, Germany

Wolf-Dietrich Moeller
Senior Security Specialist, Nokia Siemens Networks, Germany

Valtteri Niemi
Nokia Fellow, Nokia Corporation, Switzerland

A John Wiley and Sons, Ltd., Publication

This edition first published 2010

Registered office
John Wiley & Sons Ltd, The Atrium, Southern Gate, Chichester, West Sussex, PO19 8SQ, United Kingdom

Library of Congress Cataloging-in-Publication Data

LTE security/Dan Forsberg, Günther Horn. . . [et al.].
p. cm.
Includes bibliographical references and index.
ISBN 978-0-470-66103-1 (hardback)
1. Long-Term Evolution (Telecommunications) 2. Global system for mobile communications. I. Forsberg, Dan.
II. Horn, Günther.
TK5103.48325.L74 2010
621.3845′6–dc22

2010022116

A catalogue record for this book is available from the British Library.

Print ISBN: 9780470661031 (hb)
ePDF ISBN: 9780470973288
oBook ISBN: 9780470973271

Set in 10/12pt Times by Aptara Inc., New Delhi, India

Contents

Foreword

The early to mid 1980s saw the commercial opening across Europe of public-access mobile communications systems. These cellular systems all used analogue technology, but outside of the Nordic countries no attempt was made to standardize the systems – so the technology adopted differed from country to country. Unfortunately, one thing they did have in common was a total absence of adequate security features, which made them open to abuse by criminals, journalists and all manner of opportunists. User's calls could be eavesdropped on the air using readily available and comparatively inexpensive interception devices, and there were celebrated cases of journalistic invasion of privacy. A well-known example was the "squidgy" tapes, where mobile telephone calls between members of the British royal family were recorded. Mobile telephone operators and their customers became very concerned.

The operators also had another problem with serious financial consequences. When a mobile phone attempted to connect to a network the only check made on authenticity was to see that the telephone number and the phone's identity correctly corresponded. These numbers could be intercepted on the air and programmed to new phones creating clones of the original. Clones were used by criminals to run up huge charges for calls which had nothing to do with the legitimate owner. Cloning became very widespread, with criminals placing their "cloning" equipment in cars parked at airports to capture the numbers from business people announcing their arrival back home to their families. It represented a serious financial problem for operators who ended up covering the charges themselves. The problems caused by lack of security in European analogue systems were a significant factor in accelerating the creation and adoption of GSM.

GSM is a standard for digital mobile communications, designed originally for Europe but now adopted all over the world. Being an international standard it brings economy of scale and competition, and it enables users to roam across borders from one network to another. Being digital it brings transmission efficiency and flexibility, and enables the use of advanced cryptographic security. The security problems of the original analogue systems are addressed in GSM by encryption on the air interface of user traffic, in particular voice calls, and authentication by network operators of their customers on an individual basis whenever they attempt to connect to a network, irrespective of where that network may be. From both a technical and a regulatory perspective the use of cryptography in GSM was ground-breaking. Initially manufacturers and operators feared it would add too much complexity to the system, and security agencies were concerned that it may be abused by criminals and terror organizations. The legitimate fears and concerns constrained what was possible, especially with the encryption algorithm, which was designed against a philosophy of "minimum strength

to provide adequate security". Despite this, and the continuing efforts of organized hackers, eavesdropping on the air of GSM calls protected using the original cipher has still to be demonstrated in a real deployment, and with a stronger cipher already available in the wings, any future success will be largely pointless. This doesn't mean that GSM is free from security weaknesses – the ability to attack it using false base stations is very real.

GSM is the first in an evolving family of technologies for mobile communications. The second member of the family is 3G (or UMTS as it is often referred to in Europe) and the third, and most recent, is LTE (EPS to give it its proper title which is used throughout the main body of this book). With each technology evolution the security features have been enhanced to address learning from its predecessor, as well as to accommodate any changes in system architecture or services. The underlying GSM security architecture has proved to be extremely robust, and consequently has remained largely unchanged with the evolving technology family. It has also been adapted for use in other communications systems, including WLAN, IMS and HTTP. It is characterized by authentication data and encryption key generation being confined to a user's home network authentication center and personal SIM, the two elements where all user-specific static security data is held. Only dynamic and user session-specific security data goes outside these domains.

3G sees the addition to the GSM security features of user authentication of the access network – to complement user authentication by the network, integrity protection of signalling and the prevention of authentication replay. Start and termination of ciphering is moved from the base station further into the network. Of course the false base station attack is countered. A new suite of cryptographic algorithms based on algorithms open to public scrutiny and analysis is introduced, and changes of regulation governing the export of equipment with cryptographic functionality make their adoption easier for most parts of the world.

LTE heralds the first technology in the family that is entirely packet-switched – so voice security has to be addressed in an entirely different way from GSM and 3G. LTE is a much flatter architecture, with fewer network elements, and is entirely IP-based. Functionality, including security functionality, is migrated to the edge of the network, including encryption functionality which is moved to the edge of the radio network, having been moved from the base station to the radio network controller in the evolution from GSM to 3G. While maintaining compatibility with the security architecture developed for GSM and evolved for 3G, the security functionality has been significantly adapted, enhanced and extended to accommodate the changes that LTE represents, as well as security enhancements motivated by practical experience with 3G. Much of this plays back into 3G itself as new security challenges arise with the advent of femto cells – low-cost end nodes in exposed environments that are not necessarily under the control of the operator of the network to which they are attached.

The book takes the reader through the evolution of security across three generations of mobile, focusing with clarity and rigor on the security of LTE. It is co-authored by a team who continue to be at the heart of the working group in 3GPP responsible for defining the LTE security standards. Their knowledge, expertise and enthusiasm for the subject shines through.

Professor Michael Walker
Chairman of the ETSI Board

Acknowledgements

This book presents the results of research and specification work by many people over an extended period. Our thanks therefore go to all those who helped make LTE possible through their hard work. In particular, we thank the people working in 3GPP, the standardization body that publishes the LTE specifications, and, especially, the delegates to the 3GPP security working group, SA3, with whom we were working to produce the LTE security specifications over the past years.

We would also like to express our gratitude to our colleagues at Nokia and Nokia Siemens Networks for our longstanding fruitful collaboration. We are particularly indebted to Wolfgang Bücker, Devaki Chandramouli, Jan-Erik Ekberg, Silke Holtmanns, Jan Kåll, Raimund Kausl, Christian Markwart, Kaisa Nyberg, Martin Ötttl, Jukka Ranta, Manfred Schäfer, Peter Schneider, Hans-Jürgen Schwarzbauer, José Manuel Tapia Pérez, Janne Tervonen, Robert Zaus and Dajiang Zhang who helped us improve the book through their invaluable comments.

Finally, we would like to thank the editing team at Wiley whose great work turned our manuscript into a coherent book.

The authors welcome any comments or suggestions for improvements.

Copyright Acknowledgements

The authors would like to include additional thanks and full copyright acknowledgement as requested by the following copyright holders in this book.

1

Overview of the Book

Mobile telecommunications systems have evolved in a stepwise manner. A new cellular radio technology has been designed once per decade. Analogue radio technology was dominant in the 1980s and paved the way for the phenomenal success of cellular systems. The dominant second-generation system GSM was introduced in the early 1990s, while the most successful third-generation system 3G – also known as UMTS, especially in Europe – was brought into use in the first years of the first decade of the new millennium.

At the time of writing, the fourth generation of mobile telecommunications systems is about to be introduced. Its new radio technology is best known under the acronym 'LTE' (Long Term Evolution). The complete system is named 'SAE/LTE', where 'SAE' (System Architecture Evolution) stands for the entire system, which allows combining access using the new, high-bandwidth technology LTE with access using the legacy technologies such as GSM, 3G and HRPD. The technical term for the SAE/LTE system is Evolved Packet System (EPS), and we shall be using this term consistently in the book. The brand name of the new system has been chosen to be LTE, and that is the reason why the title of the book is *LTE Security*.

With the pervasiveness of telecommunications in our everyday lives, telecommunications security has also moved more and more to the forefront of attention. Security is needed to ensure that the system is properly functioning and to prevent misuse. Security includes measures such as encryption and authentication, which are required to guarantee the user's privacy as well as ensuring revenue for the mobile network operator.

The book will address the security architecture for EPS. This is based on elements of the security architectures for GSM and 3G, but it needed a major redesign effort owing to the significantly increased complexity, and new architectural and business requirements. The book will present the requirements and their motivation and then explain in detail the security mechanisms employed to meet these requirements.

To achieve global relevance, a communication system requires world-wide interoperability that is easiest to achieve by means of standardization. The standardized part of the system guarantees that the entities in the system are able to communicate with each other even if they are controlled by different mobile network operators or manufactured by different vendors. There are also many parts in the system where interoperability does not play a role, such as the internal structure of the network entities. It is better not to standardize

LTE Security Dan Forsberg, Günther Horn, Wolf-Dietrich Moeller, and Valtteri Niemi

wherever it is not necessary because then new technologies can be introduced more rapidly and differentiation is possible among operators as well as among manufacturers, thus encouraging healthy competition.

As an example in the area of security, communication between the mobile device and the radio network is protected by encrypting the messages. It is important that we standardize how the encryption is done and which encryption keys are used, otherwise the receiving end could not do the reverse operation and recover the original content of the message. On the other hand, both communicating parties have to store the encryption keys in such a way that no outsider can get access to them. From the security point of view, it is important that this be done properly but we do not have to standardize how it is done, thus leaving room for the introduction of better protection techniques without the burden of standardizing them first. The emphasis of our book is on the standardized parts of EPS security, but we include some of the other aspects as well.

The authors feel that there will be interest in industry and academia in the technical details of SAE/LTE security for quite some time to come. The specifications generated by standardization bodies only describe *how* to implement the system (and this only to the extent required for interoperability), but almost never inform readers about *why* things are done the way they are. Furthermore, specifications tend to be readable only by a small group of experts and lack the context of the broader picture. This book is meant to fill this gap by providing first-hand information from insiders who participated in decisively shaping SAE/LTE security in the relevant standardization body, 3GPP, and can therefore explain the rationale for the design decisions in this area.

The book is based on versions of 3GPP specifications from March 2010 but corrections approved by June 2010 were still taken into account. New features will surely be added into these specifications in later versions and there will most probably also be further corrections to the existing security functionality. For the obvious reason of timing, these additions cannot be addressed in this book.

The book is intended for telecommunications engineers in research, development and technical sales and their managers as well as engineering students who are familiar with architectures of mobile telecommunications systems and interested in the security aspects of these systems. The book will also be of interest to security experts who are looking for examples of the use of security mechanisms in practical systems. Both readers from industry and from academia should be able to benefit from the book. The book is probably most beneficial to advanced readers, with subchapters providing sufficient detail so that the book can also be useful as a handbook for specialists. It can also be used as textbook material for an advanced course, and especially the introductory parts of each chapter, when combined, give a nice overall introduction to the subject.

The book is organized as follows. Chapter 2 gives the necessary background information on cellular systems, relevant security concepts, standardization matters and so on. As explained earlier, LTE security relies heavily on security concepts introduced for the predecessor systems. Therefore, and also to make the book more self-contained, Chapters 3, 4 and 5 are devoted to security in legacy systems, including GSM and 3G, and security aspects of cellular–WLAN interworking.

Chapter 6 provides an overall picture of the EPS security architecture. The next four chapters provide detailed information about the core functionalities in the security architecture. Chapter 7 is devoted to authentication and key agreement which constitute the cornerstones for

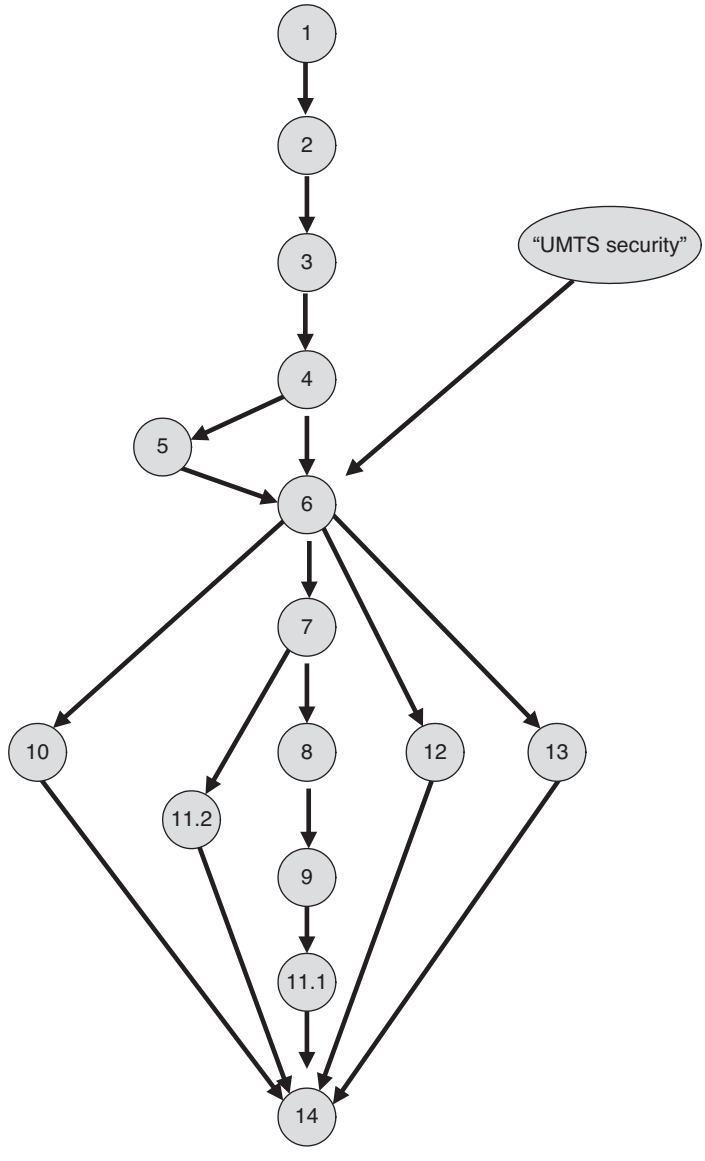

Figure 1.1 Major dependencies among chapters

the whole security architecture. Chapter 8 shows how user data and signalling data is protected in the system, including protecting confidentiality and integrity of the data. A very characteristic feature in cellular communication is the possibility of handing over the communication from one base station to another. Security for handovers and other mobility issues is handled in Chapter 9. Another cornerstone of the security architecture is the set of cryptographic

algorithms that are used in the protection mechanisms. The algorithms used in EPS security are introduced in Chapter 10.

In the design of EPS, it has been taken into account already from the beginning how interworking with access technologies that are not defined by 3GPP is arranged. Also, interworking with legacy 3GPP systems has been designed into the EPS system. These two areas are discussed in detail in Chapter 11.

The EPS system is exclusively packet-based; there are no circuit-switched elements in it. This implies, in particular, that voice services have to be provided on top of IP packets. The security for such a solution is explained in Chapter 12.

Partially independently of the introduction of EPS, 3GPP has specified solutions that enable the deployment of base stations covering very small areas, such as in private homes. This type of base station may serve restricted sets of customers (e.g. people living in a house), but open usage in hotspots or remote areas is also envisaged. These home base stations are also planned for 3G access, not only for LTE access. Such a new type of base station may be placed in a potentially vulnerable environment not controlled by the network operator and therefore many new security measures are needed, compared to conventional base stations. These are presented in detail in Chapter 13.

Finally, Chapter 14 contains a discussion of both near-term and far-term future challenges in the area of securing mobile communications.

Many of the chapters depend on earlier ones, as can be seen from the above descriptions. However, it is possible to read some chapters without reading first all of the preceding ones. Also, if the reader has prior knowledge of GSM and 3G systems and their security features, the first four chapters can be skipped. This kind of knowledge could have been obtained, for example, by reading the book *UMTS Security* [Niemi and Nyberg 2003]. The major dependencies among the chapters of the book are illustrated in Figure 1.1.

2

Background

2.1 Evolution of Cellular Systems

Mobile communications were originally introduced for military applications. The concept of a cellular network was taken into commercial use much later, near the beginning of the 1980s, in the form of the Advanced Mobile Phone System (AMPS) in the USA and in the form of the Nordic Mobile Telephone system (NMT) in northern Europe. These first-generation cellular systems were based on analogue technologies. Simultaneous access by many users in the same cell was provided by the Frequency Division Multiple Access (FDMA) technique. Handovers between different cells were already possible in these systems and a typical use case was a phone call from a car.

The second generation of mobile systems (2G) was introduced roughly a decade later, at the beginning of the 1990s. The dominant 2G technology has been the Global System for Mobile (GSM) communications, with more than three and a half billion users worldwide at the time of writing. The second generation introduced digital information transmission on the radio interface between the mobile phone and the base station. The multiple access technology is Time Division Multiple Access (TDMA).

The second generation provided an increased capacity of the network (owing to more efficient use of radio resources), better speech quality (from digital coding techniques) and a natural possibility for communicating data. Furthermore, it was possible to use new types of security feature, compared to analogue systems.

Again roughly one decade later, the third-generation technologies (3G) were introduced at the beginning of the twenty-first century. Although GSM had become a phenomenal success story already at that point, there were also other successful 2G systems, both in Asia and in North America. One of the leading ideas for 3G was to ensure fully global roaming: to make it possible for the user to use the mobile system services anywhere in the world. A collaborative effort of standards bodies from Europe, Asia and North America developed the first truly global cellular technologies in the 3rd Generation Partnership Project (3GPP). At the time of writing, there are almost half a billion 3G subscriptions in the world.

The third generation provided a big increase in data rates, up to 2 megabits per second (Mbps) in the first version of the system that was specified in Release 99 of 3GPP. The multiple-access technology is Wideband Code Division Multiple Access (WCDMA).

LTE Security Dan Forsberg, Günther Horn, Wolf-Dietrich Moeller, and Valtteri Niemi

Both GSM and 3G systems were divided into two different domains, based on the underlying switching technology. The circuit-switched (CS) domain is mainly intended for carrying voice and short messages while the packet-switched (PS) domain is mainly used for carrying data traffic.

One more decade passed and the time was ripe for taking another major step forward. In 3GPP the development work was done under the names of 'Long Term Evolution' (LTE) of radio technologies and 'System Architecture Evolution' (SAE). Both names emphasized the evolutionary nature of this step, but the end result is in many respects a brand new system, both from the radio perspective and from the system perspective. The new system is called Evolved Packet System (EPS) and its most important component, the new radio network, is called Evolved Universal Terrestrial Radio Access Network (E-UTRAN).

The EPS contains only a packet-switched domain. It offers a big increase in data rates, up to more than 100 Mbps. The multiple-access technology is again based on FDMA, namely Orthogonal Frequency Division Multiple Access (OFDMA) for the downlink traffic (from the network to the terminal) and Single Carrier FDMA (SC-FDMA) for the uplink traffic (from the terminal to the network).

2.1.1 Third-generation Network Architecture

In this section we give a brief overview of the 3GPP network architecture. A more thorough description of the 3G architecture can be found elsewhere [Kaaranen *et al.* 2005].

A simplified picture of the 3GPP Release 99 system is given in Figure 2.1.

The network model consists of three main parts, all of which are visible in Figure 2.1. The part closest to the user is the terminal that is also called the User Equipment (UE). The UE has a radio connection to the Radio Access Network (RAN), which itself is connected to the Core Network (CN). The core network takes care of coordination of the whole system.

The core network contains the PS domain and the CS domain. The former is an evolution of the GPRS domain of the GSM system and its most important network elements are the Serving GPRS Support Node (SGSN) and the Gateway GPRS Support Node (GGSN). The CS domain is an evolution from the original circuit-switched GSM network with the Mobile Switching Centre (MSC) as its most important component.

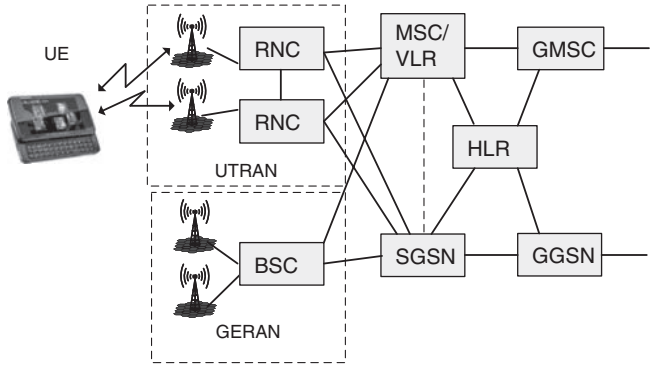

Figure 2.1 The 3G system

In addition to the various network elements, the architecture defines also interfaces or, more correctly, reference points between these elements. Furthermore, protocols define how different elements are able to communicate over the interfaces. Protocols involving the UE are grouped into two main strata: the Access Stratum (AS) contains protocols that are run between the UE and the access network, while the Non-Access Stratum (NAS) contains protocols between the UE and the core network. In addition to these two, there are many protocols that are run between different network elements.

The core network is further divided into the home network and the serving network. The home network contains all the static information about the subscribers, including the static security information. The serving network handles the communication to the UE (via the access network). If the user is roaming, then the home and the serving network are controlled by different mobile network operators.

2.1.2 Important Elements of the 3G Architecture

The user equipment consists of two parts: the Mobile Equipment (ME) and the Universal Subscriber Identity Module (USIM). The ME is typically a mobile device that contains the radio functionality and all the protocols that are needed for communications with the network. It also contains the user interface, including a display and a keypad. The USIM is an application that is run inside a smart card called Universal Integrated Circuit Card (UICC) [TS31.101]. The USIM contains all the operator-dependent data about the subscriber, including the permanent security information.

There are two types of radio access network in the 3G system. The Universal Terrestrial Radio Access Network (UTRAN) is based on WCDMA technology, and the GSM/EDGE Radio Access Network (GERAN) is an evolution of GSM technology.

The radio access network contains two types of element. The base station (BS) is the termination point of the radio interface on the network side, and it is called Node B in the case of UTRAN and Base Transceiver Station (BTS) in GERAN. The base station is connected to the controlling unit of the RAN, which is the Radio Network Controller (RNC) in UTRAN or the Base Station Controller (BSC) of GERAN.

In the core network, the most important element in the circuit-switched domain is the switching element MSC that is typically integrated with a Visitor Location Register (VLR) that contains a database of the users currently in the location area controlled by the MSC. The Gateway MSC (GMSC) takes care of connections to external networks, an example being the Public Switched Telephone Network (PSTN). In the packet-switched domain, the role of MSC/VLR is taken by the SGSN, while the GGSN takes care of connecting to IP services within the operator network and to the outside world, such as the Internet.

The static subscriber information is maintained in the Home Location Register (HLR). It is typically integrated with the Authentication Centre (AuC) that maintains the permanent security information related to subscribers. The AuC also creates temporary authentication and security data that can be used for security features in the serving network, such as authentication of the subscriber and encryption of the user traffic.

In addition to the elements mentioned here and illustrated in Figure 2.1, there are many other components in the 3G architecture, an example being the Short Message Service Centre (SMSC) that supports storing and forwarding of short messages.

2.1.3 Functions and Protocols in the 3GPP System

The main functionalities in the 3GPP system are:

- Communication Management (CM) for user connections, such as call handling and session management;
- Mobility Management (MM) covering procedures related to user mobility, as well as important security features;
- Radio Resource Management (RRM) covering, for example, power control for radio connections, control of handovers and system load.

The CM functions are located in the non-access stratum while RRM functions are located in the access stratum. The MM functions are taken care of by both the core network and the radio access network.

The division into user plane and control plane (also called signalling plane) defines an important partition among the protocols. User-plane protocols deal, as the name indicates, with the transport of user data and other directly user-related information, such as speech. Control-plane protocols are needed to ensure correct system functionality by transferring necessary control information between elements in the system.

In a telecommunication system, in addition to the user and control planes, there is also a management plane that, for example, keeps all elements of the system in operation. Usually, there is less need for standardization in the management plane than there is for the user plane and the control plane.

The most important protocols for the Internet are Internet Protocol (IP), User Datagram Protocol (UDP) and Transmission Control Protocol (TCP). In the wireless environment there is a natural reason to favour UDP over TCP: fading and temporary loss of coverage make it difficult to maintain reliable transmission of packets on a continuous basis. There is also a 3GPP specific protocol that is run on top of UDP/IP. This is the GPRS Tunnelling Protocol (GTP). It has been optimized for data transfer in the backbone of the PS domain.

The interworking of the different types of protocol can be illustrated by a typical use case: a user receiving a phone call. First the network pages for the user. Paging is an MM procedure; the network has to know in which geographical area the user could be found. After the user has successfully received the paging message, the radio connection is established by RRM procedures. When the radio connection exists, an authentication procedure may follow, and this belongs again to the MM. Next the actual call set-up (CM procedure) occurs during which the user may be informed about who is calling. During the call there may be many further signalling procedures, such as for handovers. At the end of the call, the call is first released by a CM procedure and after that the radio connection is released by the RRM.

2.1.4 The EPS System

The goals of the EPS are [TS22.278]:

- higher data rates;
- lower latency;

- high level of security;
- enhanced quality of service (QoS);
- support for different access systems with mobility and service continuity between them;
- support for access system selection;
- capabilities for interworking with legacy systems.

The main means to achieve these goals are:

- the new radio interface and the new RAN based on it (E-UTRAN);
- a flat IP-based architecture that has only two network elements on the user plane (evolved NodeB and Serving Gateway).

Figure 2.2 (adapted from [TS23.401]) illustrates the EPS network architecture in a case where the UE is not roaming into a different network than where it has its subscription. Note that the legacy radio access networks UTRAN and GERAN are included in the system together with the legacy core network element SGSN.

The new core network element is called Mobility Management Entity (MME). The HLR of the original GSM and 3G architecture is extended to the Home Subscriber Server (HSS). The core network element for user-plane handling is called Serving Gateway (S-GW). The PDN Gateway (PDN GW) handles the traffic towards packet data networks. It is also possible that S-GW and PDN GW are co-located. The core network of the EPS is called Evolved Packet Core (EPC).

The architecture of E-UTRAN is depicted in Figure 2.3 (see also [TS36.300]). The base station eNB is the only type of network element in E-UTRAN. On the other hand, there is an interface between two eNBs facilitating fast handovers between different base stations.

2.2 Basic Security Concepts

It is not easy to define 'security' even though people tend to understand quite well what is meant by it. Protection methods against malicious actions lie at the core of security. There is also a clear distinction between security, on one hand, and fault-tolerance and robustness, on the other.

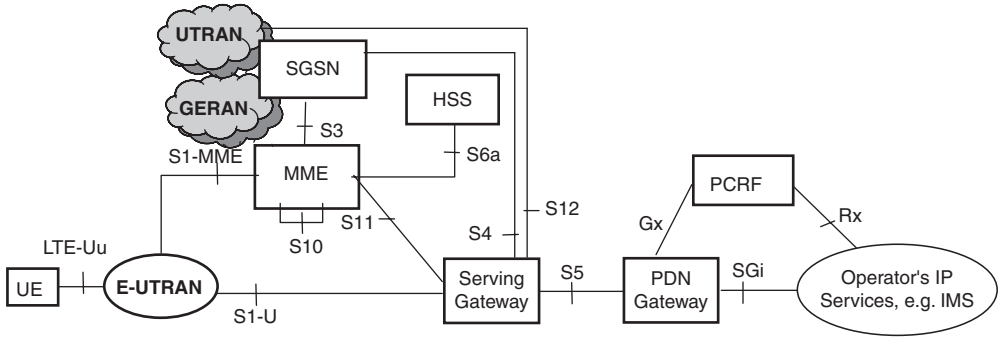

Figure 2.2 The EPS architecture (non-roaming case). Adapted with permission from © 2010, 3GPP™

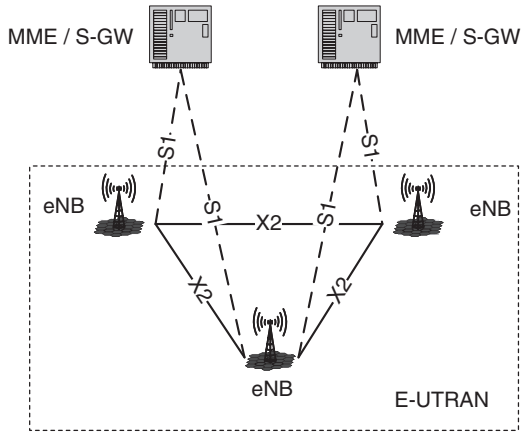

Figure 2.3 The E-UTRAN architecture

Many aspects of security are relevant for a communication system. There are physical security aspects and information security aspects. The former include issues such as locked rooms, safes and guards: all these are needed when operating a large-scale network. Another property that belongs to the area of physical security is tamper-resistance. Smart cards play a major role in the system we describe in this book, and tamper-resistance is a key property of smart cards. Sometimes guaranteed tampering evidence is a sufficient protection method against physical intrusion: if tampering can be detected quickly enough, corrupted elements can be cut out of the network before too much damage is caused.

Biometric protection mechanisms are examples of methods between physical security and information security. For example, checking of fingerprints assumes both sophisticated measurement instruments and a sophisticated information system to support the use of these instruments as access control devices.

In this book we concentrate mainly on aspects belonging to the broad category of information security. In particular, we put focus on communication security. But physical security is also important for EPS security and will be covered to some extent as well.

2.2.1 Information Security

In the context of information security, the following areas can be studied fairly independently of each other:

- *System security*. An example is trying to ensure that the system does not contain any weak parts. Attackers typically try to find a point weak enough to be broken.
- *Application security*. Banking over the Internet, for example, typically uses security mechanisms that are tailored to meet the application-specific requirements.
- *Protocol security*. Communicating parties are, for example, able to achieve security goals by executing well-defined communication steps in a certain well-defined order.

- *Platform security*. The network elements and mobile terminals depend on the correct functionality of the operating system that controls them. Physical security, too, has an important role in platform security.
- *Security primitives*. These are the basic building blocks on top of which all protection mechanisms are built. Typical examples are cryptographic algorithms, but also items like a protected memory can be seen as a security primitive (thus bringing physical security also into the picture).

In this book we put the main emphasis on system security, protocol security and security primitives. Platform security is covered only briefly, and application security is seen as more or less orthogonal to the purposes of this book.

In the design of a practical security system there are always tight constraints. The cost of implementing protection mechanisms must be balanced with the amount of risk mitigated by these mechanisms. The usability of the system must not suffer because of security. These trade-offs depend also on the intended use of the system: in a military system, for example, trade-offs between security, cost and usability are done on a different basis from in a public or a general-purpose communication system.

2.2.2 Design Principles

The design process of a security system contains typically the following phases:

- *Threat analysis*. The intention is to list all possible threats against the system, regardless of the difficulty and cost of carrying out an attack to materialize a particular threat.
- *Risk analysis*. The weight of each threat is measured quantitatively or, at least, in relation to other threats. Estimates are needed for both the probability of various attacks and the potential gain for the attacker and/or damage to the attacked side caused by them.
- *Requirements capture*. Based on the earlier phases, it is now decided what kind of protection is required for the system.
- *Design phase*. The actual protection mechanisms are designed in order to meet the requirements. Existing building blocks, such as security protocols or primitives, are identified, possibly new mechanisms are created, and a security architecture is built. Here the constraints have to be taken into account, and it is possible that not all requirements can be met. This may cause a need to re-visit earlier phases, especially the risk analysis.
- *Security analysis*. An evaluation of the results is carried out independently of the previous phase. Usually, automatic verification tools can be used only for parts of a security analysis. There are often holes in the security system that can be revealed only by using creative methods.
- *Reaction phase*. While planning of the system management and operation can be seen as part of the mechanism design phase, reaction to all unexpected security breaches cannot be planned beforehand. In the reaction phase it is vital that the original design of the system is flexible enough and allows enhancements; it is useful to have a certain amount of safety margin in the mechanisms. These margins tend to be useful in cases where new attack methodologies appear faster than expected.

We have listed here only the phases that can be considered part of the design process. In addition, implementation and testing are also important in building a secure system.

One factor that affects several phases is the fact that often the security system is part of a much larger system that is under design at the same time. This has been the case for EPS specification work also. An iterative approach is needed because the general system architecture and requirements are changing in parallel to the security design. Although these iterations seem to slow down the process, it is important that the security for the system be designed at the same time as the system itself is designed. Trying to add security to an existing completed system typically leads to impractical and inefficient solutions.

2.2.3 Communication Security Features

Although security as an abstract concept is hard to define, its ingredients or features are typically easier to grasp in definitions. In the following we list the most important features in communication security:

- *Authenticity*. In a classical scenario where parties *A* and *B* are communicating over some channel, both typically want to begin with identifying each other. Authentication is the process of verifying the identities.
- *Confidentiality*. In the same classical scenario, parties *A* and *B* may want to limit the intelligibility of the communication just to the two parties themselves, to keep the communication confidential.
- *Integrity*. If all messages sent by the party *A* are identical to the ones received by the party *B*, and vice versa, then integrity of the communication has been preserved. Sometimes the property that the message is indeed sent by *A* is called 'proof-of-origin', while the term 'integrity' is restricted to the property that the message is not altered on the way.
- *Non-repudiation*. It is often useful for the receiving party *B* to store a message received from the sending party *A*. Now non-repudiation of the message means that *A* cannot later deny having sent it.
- *Availability*. This is an underlying assumption for the classical scenario of *A* and *B* communicating with each other. The communication channel must be available for parties *A* and *B*.

Typical attacks and attackers against these features are as follows:

- Authentication – an imposter tries to masquerade as one of the communicating parties.
- Confidentiality – an eavesdropper tries to get information about at least some parts of the communication.
- Integrity – a third party tries to modify, insert or delete messages in the communication channel.
- Non-repudiation – it may sometimes give a benefit for the sender of a certain message if he can later deny sending of it. For example, the message may relate to a financial transaction, or a commitment to buy or sell something.
- Availability – a Denial of Service (DoS) attack tries to prevent access to the communication channel, at least for some of the communicating parties.

The main emphasis in this book is on the first three features: authenticity, confidentiality and integrity. The whole point of introducing LTE and EPS is to improve the availability of the cellular access channel. The non-repudiation feature is still of less importance in EPS networks; it is much more relevant for the application layer.

2.3 Basic Cryptographic Concepts

Cryptology is sometimes defined as the art and science of secret writing. The possibility to apply cryptology for protecting the confidentiality of communications is obvious. Additionally, it has been found that similar techniques can be successfully applied to provide many other security features, such as for authentication.

Cryptology consists of two parts:

- *cryptography* – designing systems based on secret writing techniques;
- *cryptanalysis* – analysing cryptographic systems and trying to find weaknesses in them.

The twofold nature of cryptology reflects a more general characteristic in security. As explained earlier, it is very difficult to find testing methods that can be applied to reliably assess whether a designed system is secure. The reason for this is that the true test for a system begins when it is deployed in real life. Then attackers may appear who use whatever ways they can find to break the system. What makes the situation even more difficult is that these real-life attackers typically try to hide their actions and methods as far as possible. Cryptanalysis (and security analysis more widely) tries to anticipate what attackers might do and is constantly searching for novel ways of attacking systems. In this manner, cryptanalysis (and security analysis) contributes indirectly to achieving a better security level.

The role of cryptanalysis in modelling attackers is a complex issue. It is perfectly fine to find weaknesses in systems that are still under design and not deployed in practice. This is because then it is still easy and relatively cheap to take corrective action. However, when the system is already in wide use the role of cryptanalysis may become controversial. A clever attack found by a researcher may be reproduced by a real-life attacker who would not have invented it by himself. In this case, the attack found by the researcher seems to cause a decrease in the level of security rather than an increase.

One obvious solution to this dilemma is to keep the cryptanalytic result confidential until corrective action has been done to remove any real-life vulnerabilities. After these vulnerabilities have been removed, publishing the results helps to avoid similar vulnerabilities in future implementations. Note that there are similar debates on how to handle vulnerabilities discovered in, for example, operating systems and browsers. There seems to be no general agreement on the appropriate handling of vulnerabilities in the security community.

Another solution to the problem is to be secretive even in the design phase. If real-life attackers do not know what kind of cryptographic algorithms are in use in the real-life systems it is difficult for them to apply any cryptanalytic results in their attacks. In fact, this approach was widely used until the 1970s. Before that time, academic published results in cryptology were scarce, and their potential relation to real-life systems was not known in public. The big disadvantage of the secretive approach, sometimes called ‘security by obscurity’, is that feedback from practical experience to academic research is completely missing, which slows down progress on the academic side.

As long as cryptography is used in closed and tightly controlled environments, such as for military communications or protecting databases of financial institutions, there is no need to open up the used systems to academic cryptanalysis. But the situation changes when cryptographic applications are used in commercial systems involving consumers. First, there could be potential attackers among the users of the system and, therefore, the design of the system could leak out to the public through various reverse-engineering efforts. Second, it is harder to build trust in the system among bona fide users if no information is given about how the system has been secured. This trend towards usage of cryptology in more open environments is one reason for the boom in public cryptologic research since the 1970s.

Another, perhaps bigger, reason was the introduction of novel, mathematically intriguing cryptologic concepts, most notably the public key cryptography [Diffie and Hellman 1976].

2.3.1 Cryptographic Functions

Let us next present formal definitions of some central cryptographic notions.

- *Plaintext space P* is a subset of the set of all bit strings (denoted by $\{0,1\}^*$); we assume here, for simplicity, that everything is coded in binary.
- *Cryptotext* (or *Ciphertext*) *space C* is also a subset of $\{0,1\}^*$.
- *Key space K* is also a subset of $\{0,1\}^*$. Often $K = \{0,1\}^k$ where k is a fixed security parameter.
- *Encryption* function is E: $P \times K \rightarrow C$.
- *Decryption* function is $D : C \times K \rightarrow P$.
- *Cryptosystem* consists of all of the above, i.e $(P;C;K;E;D)$.
- *Symmetric* encryption is defined by $D(E(p, k), k) = p$.
- *Asymmetric* encryption is defined by $D(E(p, k_1), k_2) = p$, where keys k_1 and k_2 are not identical, and moreover k_2 cannot be derived easily from k_1.

Modern cryptography is based on mathematical functions that are non-trivial from the point of view of computational complexity. This means that either the function as such is complex to compute or the function can only be computed once a certain piece of information – a key – is available. Randomness is another fundamental notion in modern cryptography. A pseudorandom generator is an algorithm that takes a truly random bit string as an input (called a 'seed') and expands it into a (much) longer bit string that is infeasible to be distinguished from a truly random bit string of the same length.

One important function type is a one-way function. Roughly speaking, a function has the one-way property if

- it is easy to compute $f(x)$, if x is given; but
- for a given y, it is infeasible to find any x with $f(x) = y$.

A more accurate definition could be given using terminology from complexity theory [Menezes *et al.* 1996], but we do not need it for the purposes of this book.

Another important function type is a trapdoor function. It is similar to the one-way function with one important difference: if a certain piece of information (a secret key) is known then