Project Number	**:**	**AC095**
Project Title	**:**	**ASPeCT : Advanced Security for Personal Communications Technologies**
Deliverable Type	**:**	**P (Public)**

CEC Deliverable Number	**:**	**AC095/ATEA/W21/DS/P/02/B.2**
Contractual Date of Delivery to the CEC	**:**	**Y01 / M06 (February 1996)**
Actual Date of Delivery to the CEC	**:**	**Y01 / M06 (February 1996)**
Title of Deliverable	**:**	**Initial report on security requirements**
Work packages contributing to the Deliverable	**:**	**WP2.1/**WP2.2/WP2.3/WP2.5
Nature of the Deliverable	**:**	**R (Report)**
Author (s)	**:**	**Geneviève Vanneste, Johan Degraeve (Editors)**

Abstract :

This report collates and analyses initial requirements for fraud management, UMTS security migration, trusted third parties and security of charging and billing.

Fraud scenarios and fraud indicators have been identified and described. Details of fraud scenarios are restricted and do not appear in the report.

UMTS and migration towards UMTS are briefly summarized. The security mechanisms in use by GSM and DECT are compared with 3 proposed authentication mechanisms for UMTS. The more specific requirements for the different roles on the migration path are summarised.

Trusted third parties are studied. An overview of the basic role of a TTP is provided, in order to identify possible UMTS security services that require their support. The corresponding functions to be provided by the TTP are identified.

The study of the requirements on secure billing started with an overview of existing methods in GSM. The UMTS defined principles and security issues were studied. Billing scenarios suitable for demonstrations were evaluated, resulting in the TTP services needed being identified.

Keyword List :
ACTS, ASPeCT, UMTS, security, migration, evolution, fraud, billing, TTP, charging

0. Executive Summary

The mobile telecommunications world is undergoing a continuing transformation as increasing numbers of services are being offered to a growing number of users by more and more operators. It is essential for the continuing success of this process that the evolving security requirements of users and service providers are addressed in an appropriate and timely way. ASPeCT aims to ensure that this happens by implementing and running trials of advanced security features to prove their feasibility and acceptability.

This deliverable, the first technical one, contains an initial report on security requirements, collating and analysing initial requirements on UMTS security migration, security services and fraud management. These requirements are based on the views of users, service providers, network operators, regulatory bodies and manufacturers.

The deliverable contains contributions from 4 work packages: Detection and management of fraud in UMTS networks, Migration towards UMTS security, Trusted third parties and Security and Integrity of billing in UMTS.

Fraud scenarios in mobile telecommunication networks have been identified and categorised. This has been done based on the partner Operators' experience of analysing fraud in existing analogue and GSM networks. These scenarios were extended towards the next generation of networks. Fraud indicators have been identified and defined, enabling the detection of important fraud scenarios. A selection has been made of those scenarios which cannot be easily detected using existing tools, but which could be identified using the rule-based or neural network-based approach. (Note that due to the sensitive nature of the fraud-related information, part of the information cannot be distributed outside the ASPeCT project).

Currently work is under way to define in ETSI, a third generation mobile telecommunications system, UMTS (universal mobile telecommunications system). The main objectives of UMTS are summarised in this contribution, with particular focus on those relevant to security and on the ideas relating to migration/evolution. The security mechanisms provided in currently available mobile systems (such as GSM and DECT) are described and compared with the security mechanisms proposed for UMTS. At the moment only mechanisms proposed for authentication are available, three of them (one based on symmetric keys, two based on asymmetric keys) were studied and described in a unified notation. This part was also contributed to SMG SG, facilitating the ongoing work on the authentication framework.

As a start for the real work on the definition of a migration scenario for the security mechanisms and security features, a requirements capture has been done. The requirements on the migration path from the different roles have been gathered. The following roles were studied: manufacturers, network operators, service providers, regulators, users and subscribers.

The general role of TTPs (Trusted third parties)was studied. This was enhanced by identifying the role a TTP can play in supporting the provision of security services in future mobile telecommunications systems. Not all UMTS security services gain from the introduction of a TTP. The core functions which a TTP has to perform in order to support the relevant UMTS security services, have been identified and described. The necessary internal operations (e.g. cryptographic calculation, key storage) and external interfaces have been identified for all the TTP-defined functions. The final part of the work on requirements capture for TTP security services has focused on analysing precisely what the functional requirements are for the ASPeCT initial and final implementations of TTP services.

The requirement capture for security and integrity of billing in UMTS started with studying the important aspects of charging and billing in GSM which may be relevant also for UMTS, including GSM charging principles for mobile-originating and mobile-terminating calls and the transfer of accounting and billing data between GSM operators. The principles and security issues for charging and billing in UMTS were derived from information already available from ETSI technical reports.

ASPeCT will further investigate methods, which are suitable to increase confidence in the correctness of the billing process in a situation where trust among network operators and service providers and trust between network operators / service providers and users can no longer be taken for granted. The requirements on secure billing services in different scenarios were derived. Finally the TTP services which are likely to be needed to support the demonstration of secure billing services were identified.

Table of Contents

1. Introduction

This report collates and analyses initial requirements for fraud management, UMTS security migration, trusted third parties and security of charging and billing.

Fraud scenarios (Chapter 4) and fraud indicators (Chapter 5) have been identified and described. However due to the sensitive nature of this information, Chapter 4 is ASPeCT confidential.

In Chapter 6 UMTS and migration towards UMTS is are briefly summarized. The security mechanisms, in use by GSM and DECT are compared with 3 proposed authentication mechanisms for UMTS in Chapter 7. The more specific requirements for the different roles on the migration path are summarised in Chapter 8.

In Chapter 9 Trusted third parties were studies. An overview of the basic role of a TTP is provided, in order to identify possible UMTS security services that require their support. The corresponding functions to be provided by the TTP are identified.

In Chapter 10 the study of the requirements on secure billing started with an overview of existing methods in GSM. The UMTS defined principles and security issues were studied. Billing scenarios suitable for demonstrations were evaluated, resulting in the TTP services needed being identified.

1.1 Contributors

This is a list of all project managers involved in the ASPeCT project plus the principal editors (Johan Degraeve and Geneviève Vanneste) whose contact details are included.

Bart Preneel	ESAT/COSIC KU Leuven K. Mercierlaan 94 B 3001 Heverlee Belgium	Phone: +32 16 32 1148 Fax: +32 16 32 1986	bart.preneel@esat.kuleuven.ac.be
George Vorvis	PANAFON 2 Mesogeon Avenue 11527 Athens Greece	Phone: +30 1 6407268 Fax: +30 1 7483322	pfeng1@compulink.gr ahad@mail.hol.gr
Chris Mitchell	Royal Holloway, University of London Egham Surrey TW20 0EX England	Phone: +44 1784 443423 Fax: +44 1784439786	cjm@dcs.rhbnc.ac.uk
Günther Horn	Siemens AG ZFE T SN 3 D-81730 München Germany	Phone: +49 89 636 41494 Fax: +49 89 636 48000	Gunther.Horn@zfe.siemens.de
Geneviève Vanneste	Siemens Atea Atealaan 34 B-2200 Herentals Belgium	Phone: +32 14 252937 Fax: +32 14 253339	p82586@vnet.atea.be
Johan Degraeve	Siemens Atea Atealaan 34 B-2200 Herentals Belgium	Phone: +32 14 252431 Fax: +32 14 253339	p82953@vnet.atea.be

Eric Johnson	GIESECKE & DEVRIENT GMBH Prinzregenstr. 159 D-81607 München Germany	Phone: +49 89 4119 944 Fax: +49 89 4119 905	X400: c=de; a=cwmail; p=g+d; s=johnson; g=eric Internet: 100277.1206@compuserve.com
Nigel Jefferies	Vodafone Ltd The Courtyard 2-4 London Road Newbury Berks RG14 1JX England	Phone: +44 1635 503883 Fax: +44 1635 31127	Nigel.Jefferies@ vodafone.gold-400.gb

1.2 Document History

Revision	Date	Changes
A	05/12/95	preliminary version
B	15/02/96	Draft for review by ASPeCT
BA1	23/02/96	Version including comments on version B from all the partners, with revision marks
C	26/02/96	Version C, equal to version BA1, without revision marks
1	28/02/96	Final Version

2. References

2.1 References for WP2.1 (Chapters 6,7,8)

[1] GSM 03.20 Security related network functions, version 4.3.1

[2] GSM 02.09 Security aspects, version 4.2.4

[3] ETSI/SMG/ETR 050101, Objectives and overview, version 2.1.0, SMG5 TD373/95

[4] ETSI/SMG/ETR 050102, Vocabulary, version 2.0.0, SMG5 TD388/94

[5] ETSI/SMG/ETR 050103, System requirements, version 3.1.0, SMG5 TD276/95

[6] ETSI/SMG/ETR 050104, Scenarios and considerations for the introduction of the Universal mobile telecommunication system (UMTS), version 0.3.0, SMG5 TD334/95

[7] ETSI/SMG/ETR 050901, Security principles, version 2.3.0, SMG SG 24/96 rev1

[8] Former ETSI/SMG/ETR 050902, Security studies for the universal mobile telecommunications system (UMTS), version 0.3.1, 19/09/1995, SMG SG 81/95

[9] ETSI/SMG/ETR 030101, Functional model and architecture of the Universal mobile telecommunication system (UMTS), version 0.0.3, SMG5 309/95

[10] ETSI/NA/61301 IN/UMTS Framework document, version 8.1.0, September 1995, SMG5 TD 279/95

[11] Document 8-1/TEMP/207-E Draft new recommendation: Evaluation of security mechanisms for FPLMTS (FPLMTS.ESM), Document 8-&/TEMP/207E + Corrigendum 1

[12] Recommendation ITU-R M.1078: Security principles for future public land mobile telecommunication system (FPLMTS). Remark: currently revised, first part in Rev1 to Document 8-1/TEMP/169E, second part of Document 8-1/197-E

[13] Draft report on evolution and migration FPLMTS, Document 8-1/TEMP/218E

[14] CEC Deliverable, R2066/BT/PM2/DS/P/113/b1 : Monet : UMTS System structure document

[15] LINK PCP, 3GS3, Technical Report 2 (Version 2), Security Mechanisms for Third Generation Systems

[16] Technical Annex Part B to ACTS Proposal No. 00144 : ASPeCT

[17] Radio Equipment and Systems (RES); Digital European cordless telecommunications (DECT): Common Interface Part 7: Security features, ETS 300 175-7, October 1992

[18] A public-key based protocol for UMTS providing mutual authentication and key agreement, Siemens AG, ETSI SMG SG DOC 73/95

[19] Public Key authentication for UMTS, KPN, ETSI SMG SG DOC 36/95 revision 1

[20] ETSI/SMG/ETR 050301 UMTS: Framework of Network Requirements, Interworking and Integration for the Universal Mobile Telecommunications System (UMTS), version 2.0.0, section 4 : Requirements for UMTS Specific Procedures.

[21] UK 3GMG, Draft Industry Advisory Document on Third-Generation P.C.S., Policy Document, section 6.9 : Security.

2.2 References for WP2.3 (Chapter 9)

[1] ASPeCT/DOC/VOD/005/WP2.3/c, Identification of security services requiring trusted third parties.

[2] ASPeCT/DOC/RHUL/006/WP2.3/b, Functions required of trusted third parties.

[3] ISO/IEC 10181-1, Information Technology - Security frameworks in open systems - Part 1: Frameworks overview.

[4] ISO/IEC JTC1/SC27 TR 14516, Guideleines for the use and management of trusted third parties Part 1: General model, WD 1.0, April 1995.

[5] ISO/IEC JTC1/SC27 TR 14516, Guidelines for the use and management of trusted third parties Part 2: Technical aspects, WD 1.0, April 1995.

[6] ETSI/SMG/ETR 050103, Special mobile group - UMTS system requirements, Version 3.1.0.

[7] ISO/IEC 10181-4, Information technology - Security frameworks in open systems - Part 4: Non-repudiation.

[8] ISO/IEC CD 13888-1, Information technology - Security techniques - Non-repudiation - Part 1: General Model.

[9] ETSI/SMG/ETR 050901, Special mobile group - Security principles for the UMTS, Version 2.3.0.

[10] ISO 7498-4, Information technology - Open systems interconnection - Basic reference model - Part 4: Management framework.

2.3 References for WP2.5 (Chapters 10)

[1] GSM MoU PDR BA.07, Charging and accounting principles, Version 3.0.4.

[2] GSM MoU PDR BA.12, Transferred account procedure and billing information, Version 3.6.1.

[3] GSM MoU PDR TD.04, The Use of FTAM in the Transferred Account EDI Procedure, Version 3.1.0.

[4] GSM 01.78 Requirements for the Customised Application for Mobile networks Enhanced Logic (CAMEL) feature.

[5] ETSI/SMG/ETR 050201, Framework for services to be supported by the UMTS, Version 3.2.0.

[6] GSM 02.24, Description of Charge Advice Information (CAI).

[7] ETSI/SMG/ETR 050901, Security principles for the UMTS, Version 2.3.0.

[8] R. Grimm: "Non-repudiation in open telecooperation", 16th Nat. Comp. Sec. Conf., Baltimore (MD), Sep 1993.

3. Abbreviations

AoC	Advice of Charge
APM	Alternative Payment Methods
ASPeCT	Advanced Security for Personal Communications Technologies
AuC	Authentication Centre
BACF	Bearer Access Control Function
BARG	Billing Administration and Roaming Group
BCF	Bearer Control Function
BCFr	Bearer Control Function radio
BSS	Base Station Subsystem
CAMEL	Customised Application Mobile-network Enhanced Logic
CCF	Call Control Function
CMIP	Common Management Information Protocol
CP	Content Provider
CRL	Certificate Revocation List
DAM	DECT authentication Module
DCCH	Dedicated Control Channel
DCS1800	Digital Cellular System on 1800 Mhz band
DECT	Digital European Cordless Telecommunications
DSAA	DECT Standard Authentication Algorithm
EIR	Equipment Identification Register
ETR	ETSI Technical Report
ETSI	European Telecommunications Standards Institute
FPLMTS	Future Public Land Mobile Telecommunications System
FTAM	File Transfer Access and Management
GSM	Global System for Mobile Communications
HLR	Home Location Register
HPLMN	Home PLMN
IMEI	International Mobile Equipment Identity
IMSI	International Mobile Subscriber Identity
IMTI	International Mobile Terminal Identity
IMUI	International Mobile User Identity
IN	Intelligent Network
ISDN	Integrated Services Digital Network
ITU	International Telecommunication Union
LAI	Location Area Identification
MCCF	Mobile Call Control Function
MCF	Mobile Control Function
MO	Mobile Originated
MoU	Memorandum of Understanding
MS	Mobile Station
MSC	Mobile Services Switching Centre
MSF	Mobile Storage Function
MT	Mobile Terminated
NO	Network Operator
PABX	Private Automatic Branch Exchange
PAC	Privilege Attribute Certificates
PLMN	Public Land Mobile Network
PMR	Public Mobile Radio
PO	Payment Organisation
PPC	Pre-personalisation Centre
PRD	Permanent Reference Document
PSTN	Public Switching Telephone Network

QoS	Quality of Service
RACF	Radio Associated Control Function
RBCF	Radio Bearer Control Function
RCF	Roaming Call Forward
S-PCN	Satellite Personal Communications Network
SCAF	Service Control Agent Function
SCF	Service Control Function
SDF	Service Data Function
SIM	Subscriber Identification Module
SMG	Special Mobile Group
SNMP	Secure Network Management Protocol
SP	Service Provider
SRES	Signed Response
SSF	Service Switching Function
TACAF	Terminal Access Control Agent Function
TADIG	Transfer Account Data Interchange Group
TAP	Transferred Account Procedure
TCH	Traffic Channel
TMSI	Temporary Mobile Subscriber Identity
TMSIn	new TMSI
TMSIo	old TMSI
TTP	Trusted Third Parties
UIM	User Identity Module
UMTS	Universal Mobile Telecommunications System
UPT	Universal Personal Telecommunication
VAS	Value Added Service
VASP	Value Added Service Provider
VHE	Virtual Home Environment
VLR	Visitor Location Register
VLRn	new VLR
VLRo	old VLR
WP	Work Package
xor	Bit-wise exclusive or operation

4. Fraud scenarios in mobile telecommunication networks (WP2.2)

In this section, a brief overview of the most common types of fraud is presented, together with an outline for each of their high level characteristics as seen by the mobile network operator. This list is not intended to be exhaustive, but does appear to represent the vast majority of frauds being perpetrated at the present time on cellular mobile telecommunication networks. The various types of fraud are classified according to whether they are technical frauds which are typically operated for financial gain or whether they are related to the personal use of mobiles and thus are typically not employed for profiteering.

This section has been withheld from any public dissemination due to considerations of commercial confidentiality. Accordingly it has been classified as ASPeCT Confidential. The contents are moved to Annex A.

5. Fraud indicators in mobile telecommunication networks (WP2.2)

5.1 Introduction

This section identifies and classifies potential indicators of fraud in cellular mobile telecommunications systems. This section is largely based upon internal Vodafone studies concerned with characterising *technical* fraud (i.e. usage of cloned mobile phone equipment) on the Vodafone UK TACS (analogue) network. This should be remembered when reading the document. Some, but not all, of the indicators referenced may also be useful for the detection of *commercial* fraud (i.e. that committed using false subscriptions) and for the detection of fraud in general in other cellular systems (such as the Vodafone UK GSM (digital) network). It is entirely possible, however, that other indicators not referenced in this document may be useful for the detection of fraud in these other cases.

There are many different types of technical and commercial fraud e.g. direct call selling, PABX fraud, personal use of clones. Each has its own characteristics and can thus be identified by particular combinations of indicators. Groups of indicators for each known type of technical fraud are identified in this document. The reader is referred to Appendix A *Fraud Scenarios in Mobile Telecommunication Networks.* (Note that this document has been withheld from public dissemination).

The document then discusses those indicators which cannot be measured in existing networks. Future work should address the measurement of these indicators by placing requirements on Third-Generation systems e.g. UMTS and FPLMTS, and developing mechanisms to facilitate the measurements.

The document contains a brief section outlining how the set of indicators identified in previous sections can be extended to accommodate features (such as roaming) of Second-Generation and Third-Generation digital systems which are generally not found in First-Generation analogue systems.

The document is concluded with a section detailing the various fields which are contained within a typical GSM toll ticket, together with a description of their purpose. A list of the different types of call for which toll tickets are generated is given.

5.2 Identification of Potential Fraud Indicators

The objective of this section is to build upon the high-level discussion of fraud scenarios in mobile telecommunication networks discussed in Sections 2 - 4 of Appendix A, in order to identify and classify potential indicators of fraud in cellular mobile telecommunication systems.

Two classifications of potential indicators and their interrelationship are considered :-

- Classification by type of indicator (Section 5.2.1).
- Classification by use of indicator (Section 5.2.2).

5.2.1 Classification of Indicators by Type

In terms of a classification of indicators by type, three categories are considered :-

- Usage indicators
- Mobility indicators
- Deductive indicators

Each of these three categories is considered in turn. The significance of the majority of the indicators can be appreciated by an understanding of Sections 2 - 4 of Appendix A.

5.2.1.1 Usage Indicators

A usage indicator is defined by some criterion relating to the way in which a mobile telephone is used. For example, the number of mobile originating calls made in a defined time interval by an individual subscriber may potentially be such an indicator. As discussed in Sections 2 - 4 of Appendix A, most types of fraud are characterised by unusually high usage, the notable exception being personal use of fraudulent mobiles.

In this document, only usage characteristics of mobile-originating calls (or other mobile-originating transactions) are considered. This is reasonable in First-Generation systems such as AMPS and TACS because incoming calls do not in general incur any charges (for the receiving party), and hence mobile terminating calls on their own cannot be used to commit fraud, except in the case of a reverse-charge (operator connected) call. This approach is harder to justify in Second and Third-Generation systems where, in particular, roaming can lead to charges for a receiving mobile party. Thus indicators based upon the usage characteristics of mobile terminating calls may also be appropriate in these newer systems.

The actual number of basic usage criteria is extremely small. As illustrated in Table 5-1 (below), most can be represented in *absolute* and *differential* form. For subscriber specific analysis, only the absolute form is required; the differential form only complicates the analysis. For blanket or global analysis, however, there is a possible advantage in using differential analysis.

Absolute Usage Criteria for Mobile-Originating Calls and Other Transactions (per subscriber)
Number of transactions within a defined time interval
Total time usage within a defined time interval
Total time required for a defined number of transactions or total time usage to occur (from the inception of the first call within the defined time interval)
Duration of individual transactions
Differential Usage Criteria for Mobile-Originating Calls and Other Transactions (per subscriber)
"Differences" between the number of transactions within corresponding defined time intervals
"Differences" between the total time usage within corresponding defined time intervals
"Differences" between the total time required for a defined number of transactions or total time usage to occur within corresponding defined time intervals

Table 5-1 : Basic Usage Indicators for the Detection of Fraud

Suppose for the purposes of argument that the usage (however this is defined) corresponding to any individual genuine subscriber over a given time period is constant; with differences in the usage existing from subscriber to subscriber. Then the variation in usage of a subscriber from one time period to the next i.e. the difference, is zero for all subscribers irrespective of whether they are low or high usage subscribers. Thus, in such a scenario, the respective usage levels of different genuine subscribers have been effectively mapped into a into a single profile - any deviations from this unique profile imply suspicious activity. This is far preferable to the use of absolute criteria in which the genuine subscriber with the highest usage determines the threshold level for all subscribers and a fraudulent mobile associated with a low usage subscriber can operate for a significant time period without detection. In practice, however, genuine subscribers can exhibit large variations in their usage from one defined time period to the next. A possible subscriber usage profile is shown in the appendix to this chapter. In a worst-case scenario, the use of differentials could lead to significantly poorer fraud detection capability relative to absolutes for a blanket analysis.

Thus this document concentrates on the use of absolute usage criteria for fraud detection. The adoption of a differential form of usage criteria is also complicated by the fact that there are a number of ways that "differences" can be formed (the standard first-order difference, formed by subtracting the current measure from the corresponding previous measure, may not always be the optimum choice).

1. Classification by Geographic Source of Mobile-Originating Transactions
Classification by cell site(s) or switch area(s)
2. Classification by Destination of Mobile-Originating Transactions (B-Number Analysis)
Geographic
Classification by national, regional or international calls
Non-geographic
Classification by repeated calls to a single number
Classification by calls to fixed or mobile numbers
Classification by calls to premium rate or non-premium rate numbers
Classification by calls to freefone or non-freefone numbers
3. Classification by Temporal Factors
Classification by time of day (e.g. 0700-1900hrs, 1900hrs-0700hrs)
Classification by time of week (e.g. week day, weekend)
Classification by special periods or events (e.g. Christmas, bank holidays)
4. Classification by Type of Mobile-Originating Transaction
Classification by mobile originating call, enquiry calls or flash requests, conference calls, activation/deactivation of supplementary services or operation of call diverts

Table 5-2 : Four Possible Generic Classifications of Basic Usage Indicators

The basic usage criteria can be classified into distinct categories in a number of different ways to yield a large array of possible indicators. It is useful to allocate these classifications into one of four groups, as illustrated in Table 5-2 (above). So, for example, in addition to the number of calls made by an individual subscriber in a defined time interval being a potential indicator, there are also potential indicators such as the number of calls made by an individual subscriber in a defined time interval to national destinations, and the number of calls made by an individual subscriber in a defined time interval to international destinations.

It is also possible to combine two or more of these classifications to yield hybrid classifications. For example, the usage criteria may also be classified into the following categories :-

- Mobile-originating calls to national destinations from inside switch area A.
- Mobile-originating calls to international destinations from inside switch area A.
- Mobile-originating calls to national destinations from outside switch area A.
- Mobile-originating calls to international destinations from outside switch area A.
- All other mobile-originating transactions.

5.2.1.2 Mobility Indicators

A mobility indicator is defined by some criterion relating to the mobility of a mobile. For example, the number of successful "rescue" handovers[1] performed during all mobile originating calls made in a defined time interval by an individual subscriber may potentially be such an indicator. As discussed in Sections 2 - 4 of Appendix A, most types of fraud are characterised by unusually low mobility, the notable exception being personal use of fraudulent mobiles.

Measurement of mobility in a terrestrial cellular network is difficult because there is no precise position location capability. The resolution is limited to cells during calls and to location areas (i.e. sets of contiguous cells) between calls. Mobility indicators must, therefore, be defined differently, or at least in different contexts, for the in-call and out-of-call cases. Table 5-3 (below) illustrates an example set of basic mobility indicators for these two different scenarios. Note that it is possible to define many other basic mobility indicators as a consequence of the fact that mobility is difficult to measure precisely, although the indicators illustrated are probably the most useful.

[1] In general terms, there are several different types of handover. A "rescue" handover is the most common type and is one which is mandatory in order to maintain acceptable signal quality during a conversation usually because of the mobility of a subscriber. There are, however, other generic types:-"confinement" (to reduce local interference and thus improve QoS) and "traffic" (to redistribute local traffic in order to improve capacity) are two examples. Not all these types may relate to mobility to the same extent.

In-Call Mobility Criteria
Number of successful handovers within a defined time interval
Number of successful handovers within a defined time interval weighted by cell size
Number of successful handovers which do not return a mobile to the previous cell within a defined time interval
Number of successful handovers which do not return a mobile to the previous cell within a defined time interval weighted by cell size
Number of distinct cells visited within a defined time interval
Number of distinct cells visited within a defined time interval weighted by cell size
Geographical separation of cells visited within a defined time interval
Number of successful handovers for individual transactions
Number of successful handovers for individual transactions weighted by cell size
Number of successful handovers which do not return a mobile to the previous cell for individual transactions
Number of successful handovers which do not return a mobile to the previous cell for individual transactions weighted by cell size
Number of distinct cells visited for individual transactions
Number of distinct cells visited for individual transactions weighted by cell size
Out-of-Call Mobility Criteria
Number of successful mobility-based location updates within a defined time interval
Number of successful mobility-based location updates within a defined time interval weighted by location area size
Number of successful mobility-based location updates which do not return a mobile to the previous location area within a defined time interval
Number of successful mobility-based location updates which do not return a mobile to the previous location area within a defined time interval weighted by location area size
Number of distinct location areas visited within a defined time interval
Number of distinct location areas visited within a defined time interval weighted by location area size
Geographical separation of location areas visited within a defined time interval

Table 5-3 : Basic Mobility Indicators for the Detection of Fraud

The potential weighting of many of these indicators by cell or location area size, as appropriate, is a reflection of the fact that cell and location area sizes fluctuate significantly with geographical position in accordance with the expected offered traffic. Exactly how such weightings can be applied in practice requires further study. For the out-of-call case, location updates can, in principle, occur for a variety of reasons e.g. on movement from one location area to an adjacent one, after a defined time interval or as commanded by the appropriate base station. Only those which correspond to a movement from one location area to an adjacent one need to be considered since these are the only type which are a measure of mobility.

For the in-call case, there are at least two reasons for considering indicators other than the number of successful handovers (whether weighted by cell size or not) :-

- To take account of the fact that cells can be revisited.

- To take account of the mobility characteristics of subscribers who make or receive calls in many cells during a defined time interval, but who do not tend to move across cell boundaries while making calls.

Analogous arguments exist for considering indicators other than the number of successful mobility-based location area updates (whether weighted by location area size or not) for the out-of-call case.

Clearly, the values of in-call and out-of-call mobility criteria depend not only upon mobility considerations, but also upon the relative times for which a mobile is in the in-call and out-of-call states. The basic in-call mobility indicators can be classified as illustrated in Table 5-4 (below).

1. Classification by Type of Transaction
Classification by mobile-originating or mobile-terminating transaction
2. Classification by Type of Handover (Handover-related indicators only)
Classification by rescue, confinement or traffic handovers etc.

Table 5-4 : Two Possible Classifications of Basic In-Call Mobility Indicators

5.2.1.3 Deductive Indicators

A deductive indicator is an indicator which arises as a by-product of fraudulent behaviour. For example, the presence of overlapping calls other than enquiry or conference calls (i.e. between two or more clones or between a genuine mobile and one or more clones all with the same identity) may potentially be such an indicator.

Table 5-5 (below) illustrates the set of basic deductive indicators.

Control and traffic channel congestion in cells
Overlapping calls or overlapping calls and location updates
Velocity checks
Recent history of alarms associated with mobile
Calls to and/or from other mobiles which have a recent history of alarms or have subsequently raised alarms

Table 5-5 : Basic Deductive Indicators for the Detection of Fraud

5.2.2 Classification of Indicators by Use

In terms of a classification of indicators by use, three categories are considered :-

- Primary indicators
- Secondary indicators
- Tertiary indicators

Primary indicators are those which, in principle, can be employed in isolation to detect fraud.

Secondary indicators are those from which, in principle, useful information can be gained if they are considered in isolation, but which should not be employed in isolation to detect fraud.

Tertiary indicators are those from which no useful information can be gained if they are considered in isolation, but which can, in principle, be used to provide ancillary information in connection with the detection of fraud.

Note that this classification of indicators by use into primary, secondary and tertiary indicators does not imply anything about the value or reliability of a particular indicator at a particular point in time. In fact, this classification is independent of the value or reliability of indicators and is purely conceptual. An indicator classed as primary may well be completely unreliable at a particular time and, therefore, should not be employed. Conversely, an indicator classed as tertiary may be very reliable at a particular time; the fact that it is tertiary simply implies that it must, by definition, be employed in conjunction with other (reliable) indicators.

Table 5-6 (below) illustrates a classification of all the (absolute) indicators which are considered in section 5.2.1, according to use.

Some of the assignments in Table 5-6 may need some clarification. In particular, it may not be clear why some of the indicators have been assigned to the secondary rather than to the primary category. For example, all usage criteria classified according to the geographic source of transactions are considered to be secondary indicators. In order to explain why this is so, suppose that an arbitrary subscriber is allowed to make a total of 300 minutes of mobile originating calls within any period of 24 hours before raising an alarm. Now, it would be quite reasonable to suggest, for example, that only 90 minutes of these calls might be made to international destinations. This is why usage criteria classified according to destination are assigned a primary status. It would be quite unreasonable to suggest, however, that 90 minutes (for example) of all calls would be made from an arbitrary switch area. The mobility patterns of genuine subscribers will fluctuate a great deal on a day-to-day basis and it is quite unlikely that a subscriber will make a large proportion of calls from a single switch area, particularly since a switch area is a very artificially defined region. Thus, although usage criteria classified according to geographic source do provide useful information about the likelihood of fraud (since fraudulent activity does tend to occur in certain geographical areas), it would be very unwise to use them in a primary capacity.

Further comment is deserved regarding the deductive indicators which are assigned to the *secondary* category. One reason why, for example, *detection of overlapping calls* and similar indicators cannot be assigned to the *primary* category is because of the presence of extension mobiles (i.e. clones of a subscriber's own mobile for personal (“legitimate”) use of that subscriber) on the network. Thus, because these indicators cannot distinguish between use of “legitimate” extension mobiles and illegitimate clones, they cannot be given primary status. Switch boundary effects are also influential in this decision.

The provision for extension mobiles within GSM is supported by certain network operators, however it is not supported within the GSM specifications and it is not supported by Vodafone. The use of extension mobiles within UMTS is still being debated. Irrespective of the fact that the functionality supported by extension mobiles may be provided either by multiple-SIM subscription or by network services, it is crucial (from the point of view of fraud detection) that if SIMs are duplicated, they are identifiable by some individual data field. Otherwise it would be impossible to

distinguish between legitimate (multi-SIM) usage and a definite fraud where MS-related identity has been cloned.

Primary Indicators
Number of mobile-originating transactions within a defined time interval (unclassified or classified by the destination of mobile originating transactions i.e. b-number analysis)
Total time usage of mobile-originating transactions within a defined time interval (unclassified or classified by the destination of mobile originating transactions i.e. b-number analysis)
Duration of individual mobile-originating transactions (unclassified or classified by the destination of mobile originating transactions i.e. b-number analysis)
Secondary Indicators
Number of mobile-originating transactions within a defined time interval (classified by the geographic source of mobile originating transactions, temporal factors or type of MO transaction)
Total time usage of mobile-originating transactions within a defined time interval (classified by the geographic source of mobile originating transactions, temporal factors or type of MO transaction)
Duration of individual mobile-originating transactions (classified by the geographic source of mobile originating transactions, temporal factors or type of MO transaction)
Total time required for a defined number of transactions or total time usage to occur (from the inception of the first call within the defined time interval) (classified or unclassified)
Overlapping calls or overlapping calls and location updates (with or without geographic source comparison)
Velocity checks
Tertiary Indicators
All mobility indicators (classified or unclassified)
Control and traffic channel congestion in cells
Recent history of alarms associated with mobile
Calls to and/or from other mobiles which have a recent history of alarms or have subsequently raised alarms

Table 5-6 : Classification of Indicators by Use into Primary, Secondary and Tertiary Categories

5.2.3 Summary of Useful Indicators

Table 5-7 (below) summarises indicators which appear to be useful in principle for detecting each type of fraud discussed in Sections 2 - 4 of Appendix A. Note from Table 5-7 that it is extremely difficult to universally characterise and hence devise indicators to reliably detect the personal use of fraudulent mobiles as distinct from the legitimate use of mobiles. This is probably because such fraudulent mobiles are used by different people for a wide range of reasons including calling friends and relatives abroad, in addition to use normally associated with fraudulent mobiles *e.g.* such as communications with other members of a drug ring etc. This problem could be solved by sub-classifying what has previously been described as personal use of fraudulent mobiles into a number of different types of fraud and then establishing indicators for each type.

5.3 Measurement of Indicators

Many of the indicators mentioned in Section 5.2 are measured as a normal part of cellular system operation. For instance, all usage indicators can be derived from billing information. There is no mechanism in existing cellular systems, however, capable of measuring mobility criteria on a per-subscriber basis. This needs to be addressed by placing requirements on future systems e.g. UMTS and FPLMTS, and developing mechanisms to facilitate the measurements.

5.3.1 PABX Fraud

Another area of interest is whether or not the dial-on digits for calls made through dial-on PABXs and dial-on freefone services can be obtained for analysis by the original cellular network operator. With reference to Sections 2 - 4 of Appendix A, this would aid in diagnosing PABX and freefone fraud. There are a number of questions to be answered in this context. For instance, not only is the question of technical feasibility of concern, but also whether it will be legally possible for a network operator to intercept dial-on-digits which are transmitted after a connection has been set up and hence constitute user traffic.

PABX fraud is not limited to just mobile-originated calls, however it is largely attributable to such calls because of the mobile (non-traceable) nature of the calling party. Irrespective of this, PABX fraud is a significant factor of mobile-associated fraud and is, therefore, to be considered by the network operators. Note that certain types of (potential) PABX fraud may be detected without any co-operation with the PABX operators, and that a co-operation between the network operator and the PABX operators would act to supplement the data types available to the network operator for the detection of potential fraud.

Basic Indicator	Classification/Type	Type of Fraud					
		Direct Call Selling	PABX Fraud	Freefone Fraud	Mobile-to-Mobile Fraud	Premium Rate Line Fraud	Personal Use
Primary							
Total MO time usage	Unclassified	√√	√√	√√	√√	√√	√
	Geographic destination	√√					√
	Repeated calls to a single b-number		√√	√√	√√	√√†	
	Calls to UK freefone numbers			√√			
	Calls to UK mobile numbers				√√		√
	Calls to UK premium rate numbers					√√	√
MO call duration	Unclassified	√√	√√	√√	√√	√√	√
Secondary							
Total MO time usage	Geographic source	√√	√√	√√	√√	√√	
	Temporal factors	√√	√√	√√	√√	√√	
	Type of transaction					√√††	
Time to surpass thresholds	All thresholds	√√	√√	√√	√√	√√	√
Overlapping `calls'	All types	√√	√√	√√	√√	√√	√√
Velocity checks	Unclassified	√√	√√	√√	√√	√√	√√
Tertiary							
Mobility	All types and classifications	√√	√√	√√	√√	√√	
Cell congestion	Unclassified	√√	√√	√√	√√	√√	
Recent alarms	Unclassified	√	√	√	√	√	√
Calls to/from other fraudulent mobiles	Unclassified	√	√	√	√	√	√

Table 5-7 : Summary of Useful Indicators at the Present Time as a Function of Type of Fraud

(√√ = useful in the majority of cases, √ = useful in isolated cases)
† only for the first type of premium rate line fraud discussed in Section 5.2.
†† only for the second type of premium rate line fraud discussed in Section 5.2.

5.3.2 Mobility Indicators

The mobility information referred to in Table 5-3, such as cell size, area size and geographical location may act to supplement the detection of fraud by extending the set of possible indicators, for example, in overlapping calls and velocity checking. Note, however, that this information is not currently used by Vodafone in the detection of fraud and that toll tickets with a meaningful inclusion of such data are not currently available. It is suggested, therefore, that these mobility indicators are not used in the development of a rule-based or adaptive analysis of toll tickets in the detection of fraud. Refer to section 5.5 for a discussion of GSM toll tickets.

5.3.3 Thresholds

The determination of thresholds for incorporation into fraud indicators is not feasible on a personal basis, due to the fact that the usage patterns of any one individual may be expected to fluctuate greatly from time to time. It is suggested, therefore, that thresholds be set (initially) according to a “rolling average”, based upon usage figures for each particular subscription tariff and each particular type of call.

5.3.4 Legality Issues

Legal problems associated with the protection of subscriber-related signalling data are not perceived, since such data would normally be dealt with in-house. For the purposes of developing either a rule-based or an adaptive fraud detection mechanism, any data provided by Vodafone would be sanitised to remove any real relation to actual calling-party or called-party identifier, or to their equipment. Long-term analysis of user behaviour should create no problems with regard to legality, since such practices are common on the basis that the information relating to such analysis is properly protected and is not divulged outside the any company which collects it. Note, however, that Vodafone is not currently in possession of any such data sets related to individual users, since it sells its network services via Service Providers, who re-sell the services to the individual mobile subscribers.

5.4 Fraud Detection in GSM and UMTS/FPLMTS

As discussed in the introduction, the set of fraud indicators identified in Section 5.2 is specific to the case of technical fraud on First-Generation analogue networks, in particular the Vodafone UK TACS network. In this section, the features of Second and Third-Generation systems which will cause additional indicators and/or fraud detection mechanisms to arise are examined.

5.4.1 The GSM System

The two most significant differences with respect to fraud detection between a Second-Generation system, such as GSM, and First-Generation systems are :

- The ability to differentiate easily between different tele-services and bearer services in GSM.
- The ability to roam between different networks in GSM.

The former feature allows usage indicators to be classified according to the exact tele-service or bearer service.

The latter feature implies that the fraud engines (i.e. the systems that perform fraud management) of different network operators must co-operate in order to facilitate efficient fraud detection. This is not simply a case of the visited network operator monitoring a subscriber's behaviour in isolation. For example, if cloning of SIM cards were possible, this approach would not detect simultaneous use of a subscription on different GSM networks.

5.4.2 UMTS and FPLMTS

Three fundamental system requirements of UMTS and FPLMTS are that they will be :-

- Multi-service
- Multi-environment
- Multi-operator

The multi-service aspect of UMTS and FPLMTS allows usage indicators to be classified according to the exact service being used.

The multi-environment feature of UMTS and FPLMTS should allow greater resolution to the process of monitoring subscribers mobility in urban/office environments where microcells/picocells are in existence.

As far as the third requirement is concerned, a multitude of operators may be in existence be they small, large, localised or nation-wide. The fraud detection architecture for such a scenario requires careful consideration. It is unlikely that the smaller localised operators will be able to afford to operate dedicated fraud engines and, in any case, a multitude of fraud engines in any arbitrary region is bound to be counter-productive in terms of signalling and effectiveness.

5.5 Toll Tickets

The toll ticket is a set of details stored electronically or in printed format about any individual call being made on a mobile telephone. The toll ticket can be analysed to provide information about customer usage and to facilitate the detection of any possible fraudulent use. Different formats are used according to whether the telephone is running on the TACS (First-Generation) system or on the GSM (Second-Generation) system. This document focuses only on GSM toll tickets, since these will be used to develop neural-net and rule-based mechanisms for fraud detection within the Third-Generation UMTS.

5.5.1 Toll Ticket Statistics

A toll ticket is prepared for every call made on a mobile telephone and, to give an idea of the size of the data collected as a whole, an example is given here of a recent analysis of the Vodafone UK GSM network. There are approximately one million toll tickets generated per weekday, plus approximately one million toll tickets generated per weekend. Thus the weekly toll ticket count is in

the region of six million. If an assumed 200 bytes of information is contained in each Toll Ticket, this equates to an upper bound of approximately 1 Gigabyte of information per week.

5.5.2 Contents of a Toll Ticket

The toll ticket contains details such as the base station used in the caller's initial connection, the a-number, equipment serial number (IMEI) and subscriber number (IMSI), the date time and duration of the call and the b-number. Other details refer to the call class, such as international calls or premium-rate calls.

Depending on the nature of the call being made, data might not be recorded regarding certain subscriber-related information. For example, if the call is to a fixed terrestrial (PSTN/ISDN) system or to a communications network abroad, then no mobile subscriber number exists for the recipient.

Note that the data types in the various fields of a GSM toll ticket are either of fixed length or have a fixed maximum length.

A failed call attempt does not cause a toll ticket in GSM. The "Cause for Termination" field is used to indicate causes for termination of a call, such as "terminated by calling party", "terminated by called party", "time-out" or "network failure" etc.

A "partial call" is not a result of a dial-on call; it is used to refer to a toll ticket which is generated during a call which exceeds a given time duration threshold. For example, a call which exceeds a time duration of 6 hours may cause the generation of a toll ticket for a partial call. In this case, a "partial record number" would be assigned to the toll ticket together with the appropriate "Cause for Output", allowing the toll ticket to be uniquely associated with any other toll tickets caused as a result of that particular call.

5.5.3 GSM Toll Ticket Fields

For the GSM communications system, there are eight types of data sets depending on what type of call is being made. The eight call-type classes are :-

- Mobile-Originated (MO)
- Land-to-Land (LL)
- Mobile-Terminated (MT)
- Forwarded (FWD)
- Roamer Call Forward (RCF)
- Supplementary Service (SS)
- Short Message Service (point-to-point) Mobile-Originated (SMS/MO)
- Short Message Service (point-to-point) Mobile-Terminated (SMS/MT)

In all, there are a total of 52 different fields within a GSM Toll Ticket, excluding any data specific to the particular billing system. It is intended that a description of the position, (maximum) length and purpose of each (relevant) field within a GSM toll ticket will be provided by Vodafone, specific to any sets of GSM toll tickets which are provided for analysis. Also a suggestion will be made of the relevance of each individual toll ticket field, with respect to the analysis of the toll tickets for the purposes of detecting fraud. Note that GSM specification 12:05 gives a comprehensive description of the data types included in toll tickets, produced by the MSC for analysis by the Billing Centre.

Table 5-8 (below) shows the format of GSM Toll Tickets for Mobile-Originated (MO) Calls, where data is collected in 34 fields. It is intended that Vodafone will provide a comprehensive list of the data fields which are included in the different types of toll ticket, specific to the eight different types of call.

With respect to Table 5-8, the **Record Type** identifies whether it is a Mobile-Originated call, Land-to-Land, etc., and the **Record Number** is a sequential number of generated toll tickets. The **Call Reference** is unique to a Mobile Services Switching Centre (MSC), and the **Partial Record Number** is a sequential number of a partial record. The **Cause For Output** field records whether or not this is a partial call. The **Number of SS Records** refers to the number of supplementary service records relating to one call.

The **Calling IMEI** (International Mobile Equipment Identity) and **Calling IMSI** (International Mobile Subscriber Identity) identify the subscriber and the telephone being used to make the call. The **Calling Number** identifies from which directory number the call is being made (the a-number) and the **B-Type Of Numbe**r notes whether (for example) the dialled number is an international call. The **Called Number** is the directory number being dialled and the **First Location Area ID** and **Last Area Location ID** indicate from which area code the call is being made and to where it is being made. The **First Cell ID** and **Last Cell ID** identify the cells relating to this call. The **Last Exchange ID** is the MSISDN (Mobile Station ISDN) of the Exchange.

Field	Contents	Field	Contents
1	Record Type	18	Charging Start Time
2	Record Number	19	Chargeable Duration
3	Call Reference	20	Cause For Termination
4	Partial Record Number	21	Data Volume
5	Cause For Output	22	Chargeable Service Type
6	Number Of SS Records	23	Chargeable Service Code
7	Calling IMSI	24	Secondary Service Type
8	Calling IMEI	25	Secondary Service Code
9	Calling Number	26	Tariff Class
10	B Type Of Number	27	Call Type
11	Called Number	28	Half-Rate Indicator
12	First Location Area ID	29	Non-Transparency Indicator
13	First Cell ID	30	Category
14	Last Exchange ID	31	IMS Class Mark
15	Last Location Area ID	32	Out Circuit Group NR
16	Last Cell ID	33	DTMF Sender Indicator
17	Channel Allocation Time	34	AOC Indicator

Table 5-8 : Summary of Useful Indicators at the Present Time as a Function of Type of Fraud

The **Channel Allocation Time** is the time of seizure of the channel. The **Charging Start Time** notes when the call starts, and the **Chargeable Duration** is used to determine how long the call was. The **Cause For Termination** notes the reason for the success or failure of a call attempt. The **Data Volume** is the number of 64-Octet packets transmitted on a bearer service. The **Chargeable Service Code** is the primary service for which the customer is charged. The **Chargeable Service Type** determines whether the Chargeable Service Code is a bearer service or a tele-service code.

The **Secondary Service Code** is given if a tele-service has been used. This then contains the associated bearer service code. The **Secondary Service Type** qualifies the Secondary Service Code if

a tele-service has been used. The **Tariff Class** identifies the charging attributes for the call. The **Call Type** notes what type of call record is being generated by the MSC. The **Half-Rate Indicator** notes whether the service is at full-rate or half-rate. The **Non-Transparency Indicator** indicates whether a call is "transparent". The **Category** field records the priority or the function of the call (ordinary/priority, testphone/payphone).

The **NS Class Mark** is the power capability of the mobile equipment. The **Out Circuit Group NR** notes the number of the trunk circuit group from the Public Switched Telephone Network. The **DTMF Sender Indicator** records whether the Dual Tone Multi-Frequency (DTMF) indicator is on or off. Finally, the **AOC Indicator** indicates if an Advice Of Charge is associated with this call.

5.6 APPENDIX: Possible Subscriber Profile

A possible customer profile based on network-provided data to be used for the detection of fraud. Typical data fields may include the following :-

MSISDN :
IMSI :

SUBSCRIPTION TYPE:
ADD. SERVICES:
ADDITIONAL MSISDNs:

BARS (OPERATOR):
BARS (SUBSCRIBER):
HLR ID:
VLR ID (CURRENT):

CALLS/DAY :
AVG CALL DURATION :
CELLS USUALLY USED:

CALLS/DAY INTERNATIONAL :
DESTINATIONS OF INTERNATIONAL CALLS (COUNTRIES):

CALLS/WEEK ROAMING:
AVG CALL DURATION ROAMING:
COUNTRIES USUALLY ROAMED:

AVG MONTHLY BILL:
CREDIT LIMIT:

CALL FORWARDING AVAILABLE:
NUMBER OF DIVERTS USUALLY APPLIED /TYPE :
CALL FORWARDED NUMBERS USED (frequently-used numbers):

OF CALLS LAST HOUR:
INTERNATIONAL CALLS LAST HOUR (to a specific country):
FROM CELL:

AVG WEEKEND USAGE:
HOURS WEEKENDS (busy period):
HOURS WEEKDAYS (busy period):

6. UMTS : the third generation mobile telecommunications system & migration towards UMTS (WP2.1)

6.1 Introduction

The Universal Mobile Telecommunications System (UMTS) is the third generation mobile telecommunications system, scheduled to start service in Europe around 2000-2005 to provide a range of telecommunications services to mobile and stationary users in a variety of environments.

Here are some general objectives for the third generation mobile telecommunications system (General Objectives are listed in [3]):

- Integration of existing telecommunication services into a single system.
- Personal mobility by using an identification module on any mobile station.
- Universal Personal Telecommunications (UPT)
- Control of theft, fraud and abuse
- Security : The same security must be given as in the fixed network. Due to the vulnerability of the radio interface in mobile communications, special precautions will be taken to guarantee security.

6.2 Role Model

The concept of a Role Model is introduced where the various actors (persons, legal entities, or machines as their delegates) are assigned a role and relationships are defined between the roles. A simplified representation of the model is depicted in Figure 6-1.

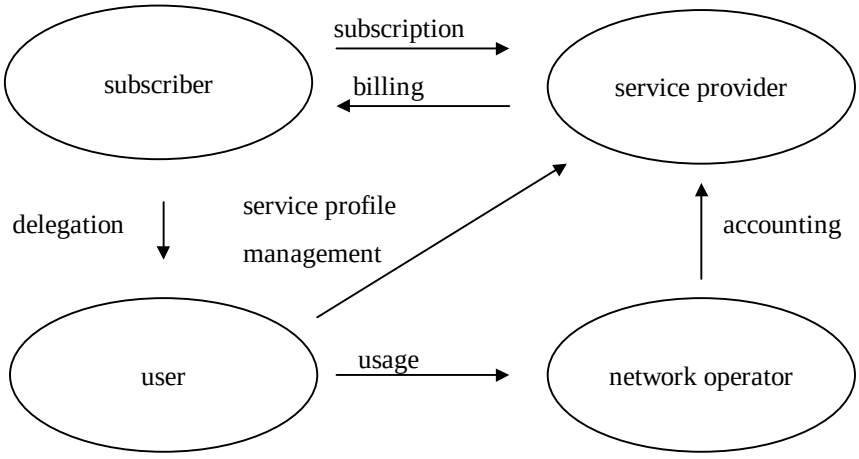


Figure 6-1 : Simplified Role Model

- The UMTS role model consists of 4 actors, the user, the subscriber, the service provider and the network operator.
- Between these actors relations can be defined:

- subscription : a contractual relationship between a subscriber and a service provider, defining among other things, the services subscribed to and the tariffs that will be applied
- delegation : the subscriber authorises a user to use (a part of) the telecommunication services subscribed to; this authorisation is implemented in a user service profile
- service profile management : the relationship between the user and the service provider, devoted to modification of the relevant user service profile data, as agreed at subscription
- usage : the use of telecommunication services as agreed at subscription; the user uses network capabilities provided by the network operator to access telecommunication services
- accounting : the network operator notifies the service provider of the use of telecommunication services invoked by the users, who are using a subscription to the service provider; this is a prerequisite to billing
- billing : the subscriber receives notification of the monetary counterpart of the use of telecommunication services by its users.

6.3 Functional Model

In [9] a draft functional model is defined (See Figure 6-2)

In the end-to-end functional model a distinction is made between

a. home network of user A (or user A related service provider)
b. home network of user B (or user B related service provider)
c. originating network
d. terminating network

The home networks are understood to act in their role as service provider only, hence the FEs SSF / CCF, SCAF and RACF are not shown. In the originating network SSF / SCF and RACF are shown since these FEs are used in call / connection related signalling.

Intermediate networks are not shown in the figure. A box drawn between the originating network and the terminating network shows that it must be possible to establish a relation between the terminating SCF and originating SCF, and between the terminating SSF / CCF and the originating SSF / CCF. In principle the intermediate network(s) should be transparent to the originating and terminating networks.

As it is an assumption for UMTS that the connection control is separated from the call control, it is not necessary to show service control or call control entities in intermediate networks. Connection control is deemed to be available in intermediate networks, but the only requirement is that the connection control related functionality can inter-work with the connection control functionality in the originating and terminating networks.

In the following sections some mobile specific entities and functions are described.

The following actions require relations among instances of SDF FEs

- registration
- Location update
- possibly call set-up

The following actions require relations among instances of SCF FEs

- split charging
- call set-up

- handover provision of identical VPN functionality across networks of different network operators
- provision of VHE

On the terminal side a functional entity is inserted that represents the handover control functionality necessary for a terminal initiated / controlled handover. This FE is called TACAF.

The TACAF may include the following functionality with respect to handover control:

- determination of the transmission mode and, if needed, determination of the relevant connection type to support the required service;
- indication of relevant data needed for a radio resource allocation request;
- co-ordination of link establishment and release (e.g. for terminal controlled handover);
- processing control of data having a local significance (e.g. radio measurement samples used to monitor and decide on terminal controlled handover);

The TACAF's functionalities regarding paging include:

- reply at the paging request from the SCF.

The functions of the BCAF are

- establishing the radio link
- maintaining the radio link (power control, timing advance)
- monitoring of link quality (bit error rate, frame error rate etc.)
- releasing the radio link

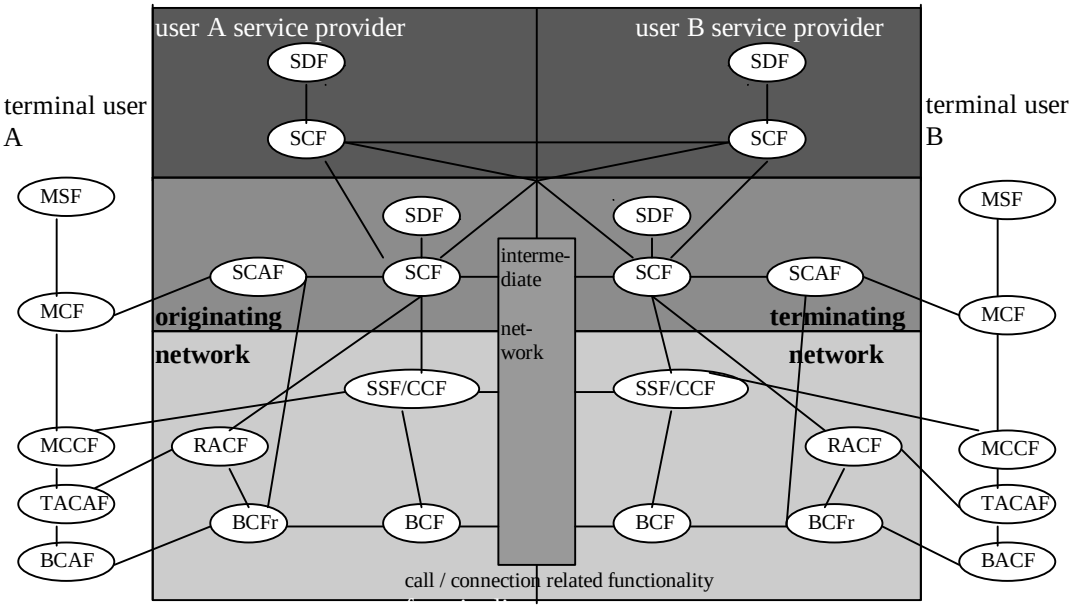


Figure 6-2 : End to End functional model

Definitions of the IN functional entities used in Figure 6-2.

SDF - Service Data Function (Mobile) : This function handles storage and access to service related data and network data and provides consistency checks on data. It hides from the SCF the real data implementation and provides a logical data view to the SCF.

SCF - Service Control Function (Mobile) : For UMTS this function contains the overall service and mobility control logic and handles service related processing activity. It supports all mobile specific functions and provides overall service control. Service logic is invoked by service requests from other functionalities to support location management, mobility management, security management and user service control as defined.

SCAF - Service Control Agent Function : This functional entity interacts with the MCF and SCF.

MSF - Mobile Storage Function : This is a pure data storage function at the mobile side of the radio interface.

MCF - Mobile Control Function : This function contains the service logic and service related processing required at the mobile side of the radio interface. It supports all mobile specific functions (e.g. location management, mobility management, identity management) and provides local service control.

CCF - Call Control Function :

In general the CCF includes the following functionalities:

analyse and process service requests;

establish, maintain and release calls;

request for allocation of network bearer resources;

execution of some types of handover

provide information relevant to charging

SSF - Service Switching Function :
The SSF is the service switching function, which interacts with the CCF and the Service Control Function (SCF).

BCF - Bearer Control Function :
This function controls the bearer connection elements in the fixed network in order to provide the bearer capabilities requested by the CCF.

BCAF - Bearer Control Agent Function :
This function is responsible for the control of the radio bearer connection at the terminal side.

MCCF Mobile Call Control Function :
The MCCF is the agent between the user and the network call control functions.

TACAF - Terminal Access Control Agent Function :
This function is responsible for the control of terminal initiated handover at the terminal side.

RACF - Radio Associated Control Function :
This functional entity provides the control functionality required in the radio access subsystem. The role of the RACF is to drive the actions which have to be performed by other FE's for radio resource allocation/de-allocation (e.g. RRC) and radio bearer connection establishment/release (e.g. RBCF). The RACF is also responsible of some local supervisory actions like processing control needed for the handover monitoring and decision phases and co-ordination between the RBCF and BCF in order to maintain the context of the end-to-end connection (e.g. during handover execution).

6.4 Migration

6.4.1 What is migration towards UMTS ?

Both UMTS and FPLMTS give definitions for migration and evolution towards third generation [6,13] :

Evolution towards UMTS :

A process of change and development of a telecommunications system towards the capabilities and functionalities of UMTS.

Migration to UMTS :

Movement of users and/or service delivery from existing telecommunications systems to UMTS.

In the ACTS Proposal one can read that the objectives of Work Package 2.1, to which this document will be the first contribution, include both evolution and migration to UMTS (See Section 3.3.3 in [16])

In this document "migration" will include both migration and evolution as defined above.

6.4.2 Why migration towards UMTS

The requirement for migration is based on the UMTS system requirements defined in ETSI ETR /SMG 050103 [5], which are themselves based on the UMTS objectives defined in ETSI ETR/SMG 50101 [3].

Some of the system requirements are listed here:

R092 : UMTS aims at the convergence of current mobile and fixed network services with cost efficient re-use and development of functionality for maximum efficiency.

R096 : The UMTS standard must enable the provision of cost effective, efficient interworking with second generation mobile systems and fixed systems. Ideally, seamless service delivery is required, enabling services to be provided transparently across several networks.

R098 : Provision shall be made for the support of terminal roaming between second generation systems and UMTS. This shall not constrain the UMTS standards.

R099 : Provision shall be made for SIM/UIM roaming between second generation handsets and UMTS handsets.

6.4.3 Standardisation of migration

This section summarises some documents describing migration.

6.4.3.1 ETSI : SMG 050104 Scenarios and considerations for the introduction of the UMTS [6]

In this document some candidates for migration are discussed:

- GSM/DCS1800/DCS1900
- S-PCN (Satellite Personal Communications Network)
- DECT (Digital European Cordless Telecommunications)
- IN and fixed network

Two important examples for migration towards UMTS are mentioned :

- *"GSM/DCS, DECT/GSM, IN. The further evolution can comprise the introduction of a new (UMTS) air interface, increased bit-rates and enhancements of the MAP to fulfil UMTS system requirements. Potentially also new base station system technologies might be introduced. Gradually GSM evolves and in the end UMTS services can be offered."*
- *"Adding IN-capabilities with mobility support and radio access subsystems to the fixed networks forms the second main track towards UMTS. CTM is one such example."*

A "four level and three step model" is defined :

Each level contains new technology and is a step to UMTS. Level 1 is the current situation (GSM Phase 2, IN CS-1, B-ISDN CS-1, D-AMPS, PDC/PHS). The new technologies introduced in the new steps are :

- step A : GSM Phase 2++, Advanced PDC/PHS, IN CS-2, B-ISDN CS-2.1 and CS-2.2
- step B : IN CS-3, B-ISDN CS-3, new frequency band added, new air interface added
- step C : to UMTS.

A modular approach is defined :
The whole system is divided into modules with, as an example of different modules : user data (user-id, subscription,...), terminal, access subsystem, transport subsystem, service subsystem, mobility subsystem, security subsystem. Each module can be maintained, enhanced or changed as it complies with the UMTS standards and, if necessary, still being backwards compatible with the previous version.

The importance of backwards and forward compatibility are emphasised :
Backwards compatibility means that a user should be able to use his second generation terminal after the network is upgraded to third generation. Third generation terminals should also be unable in to operate in a second generation network.
Forward compatibility means that a second generation terminal should be prepared for network capabilities that will be introduced in third generation networks.

The "Virtual Home Environment" (VHE) is considered as an important system concept : "*The primary aim is to provide the user (subscriber, service provider or network operator) with a comprehensive set of services, features and tools, which have the "same look and feel" whether they are used "at home or abroad" -hence the term "Virtual Home Environment".*

Enhancements to this document can be expected. There are already two proposals : CTM evolution scenario towards UMTS written by Telenor and a GSM evolutionary scenario written by Telecom Finland. The last proposal refines the four level and three step model.

6.4.3.2 Draft report on evolution and migration FPLMTS, Document 8-1/TEMP/218E [13]

This document has been a base for ETSI SMG 050104 [6] (See Section 6.4.3.1) so most of the ideas described are the same. This document [13] is more extensive because FPLMTS will start from a wide range of technologies from different countries. Some differences are :

- An Evolution Framework is defined which concludes with the same modular approach as UMTS
- Standardisation of a Radio Bearer Adaptation Functionality
- Standardisation of a User Identity Module (UIM) : with evolution towards a common UIM.
- Software-Defined Radio Technologies in FPLMTS

7. Security in second and third generation mobile systems (WP2.1)

7.1 Introduction

The following sections describe security in second generation and security in third generation systems.
Section 7.2 describes security in GSM and DECT.
Section 7.3 describes the security features in UMTS for which mechanisms are proposed until now.
The only security features for which mechanisms are available is authentication.

7.2 Security in second generation systems

7.2.1 GSM

GSM security features are standardised in GSM 02.09 [2].
The following security features are considered :
– subscriber identity (IMSI) confidentiality
– subscriber identity (IMSI) authentication
– user data confidentiality on physical connections
– connectionless user data confidentiality
– signalling information element confidentiality

7.2.1.1 Subscriber identity (IMSI) confidentiality

The purpose of this function is to avoid an intruder being able to identify which subscriber is using a given resource on the radio path. The provision of this function implies that the IMSI (International Mobile Subscriber Identity) should not normally be transmitted in clear text in any signalling message on the radio path.
The means used to identify a mobile subscriber on the radio path consists of a TMSI (Temporary Mobile Subscriber Identity). The TMSI is allocated by the VLR (Visitor Location Register) where the MS (Mobile Station) is registered.
The very first time that a MS registers in the network or after a software failure in the VLR, the fixed part of the network can require the MS to send the IMSI in clear. A new TMSI is allocated and sent to the MS in ciphertext.
Figure 7-1 demonstrates the use of the TMSI : “Location Updating in a new VLR; old VLR not reachable”. It is an example where the IMSI has to be used because no information can be obtained using the TMSI, since the old VLR can not be contacted.

Management of means for new ciphering :
The MS and BSS/MSC/VLR agree on means for ciphering signalling information elements, in particular to transmit TMSIn

7.2.1.2 Subscriber identity (IMSI) authentication

Subscriber identity (IMSI) authentication is the corroboration by the land-based part of the system that the subscriber identity (IMSI or TMSI), transferred by the mobile subscriber within the identification procedure at the radiopath, is the one claimed.

The authentication of the GSM PLMN subscriber identity may be triggered by the network when the subscriber applies for:

- a change of subscriber-related information element in the VLR or HLR (including some or all of: location up-dating involving change of VLR, registration or erasure of a supplementary service); or
- an access to a service (including some or all of: set-up of mobile originating or terminated calls, activation or deactivation of a supplementary service); or
- first network access after restart of MSC/VLR;

or in the event of cipher key sequence number mismatch.

The authentication procedure consists of :

- The fixed subsystem transmits a non-predictable number RAND to the MS
- The MS computes the signature of RAND say SRES, using algorithm A3 and some secret information: the Individual Subscriber Authentication Key, denoted below by Ki.
- The MS transmits the signature SRES to the fixed subsystem.
- The fixed subsystem tests SRES for validity.

Note that this procedure is also used to set the ciphering key (see section7.2.1.3.2).
The general procedure is schematised in Figure 7-2

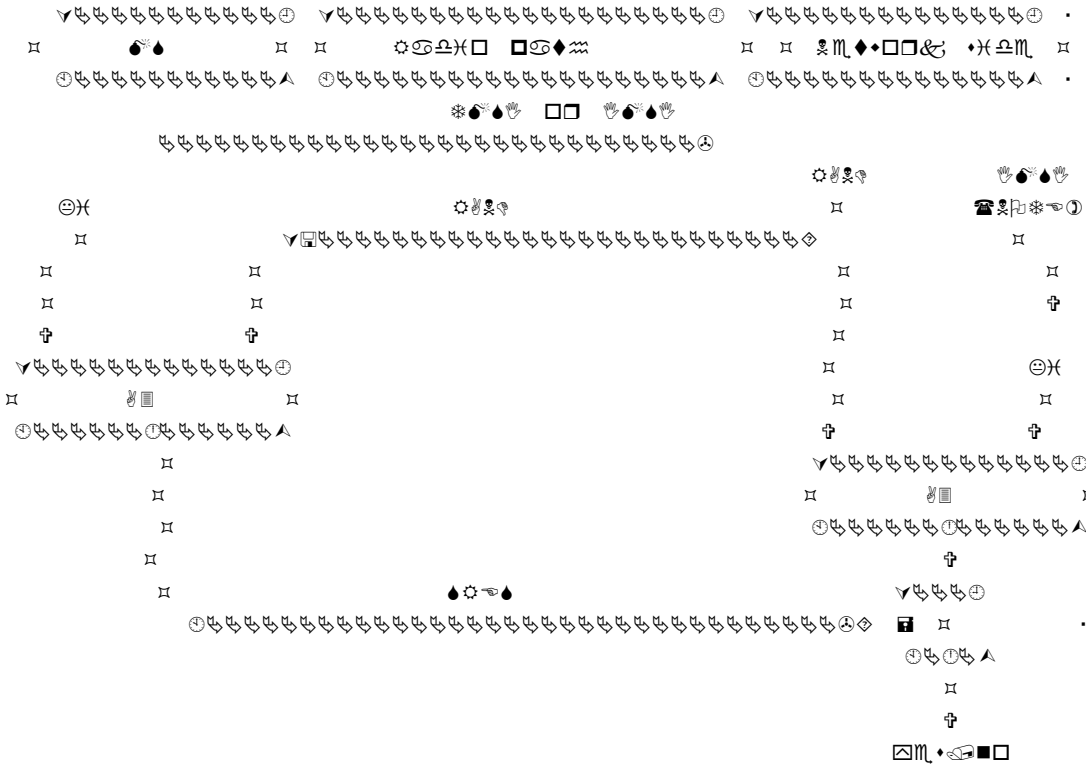


Figure 7-2 : The authentication procedure

NOTE: IMSI is used to retrieve Ki in the network.

Algorithm A3 is PLMN operator specific. Only the formats of their inputs and outputs must be specified.

The Subscriber Authentication Key Ki is allocated, together with the IMSI, at subscription time. Ki is stored on the network side in the Home Public Land Mobile Network (HPLMN), in an Authentication Centre (AuC). A PLMN may contain one or more AuC. An AuC can be physically integrated with other functions, e.g. in a Home Location Register (HLR).

Several scenarios are possible when a VLR wants to perform an authentication, depending on whether the TMSI or IMSI is used for identification, and whether the TMSI can be used to retrieve security information already available in the old VLR. This security related information consists of a Random (RAND), a Signed Response (SRES) and a ciphering key Kc which will be used later for ciphering. These three together are called a triplet.

When no triplets are available in the VLR or can't be retrieved from the old VLR (in case the MS performs a location update in a new VLR), then the triplets are requested to the HLR/AuC.

7.2.1.3 Confidentiality of signalling information elements, connectionless data and user information elements on physical connections.

The following information is considered sensitive and must be protected against eavesdropping:

- Signalling information elements
- To ensure identity confidentiality, the Temporary Mobile Subscriber Identity must be transferred in a protected mode at allocation time and at other times when the signalling procedures permit it.
- The confidentiality of connection-less user data requires at least the protection of the message part pertaining to OSI layers 4 and above.
- user data

Confidentiality is achieved by one mechanism and is situated in the OSI layer 1.
Four points are specified:

- the ciphering method;
- the key setting;
- the starting of the enciphering and deciphering processes;
- the synchronisation.

7.2.1.3.1 The ciphering method

Stream ciphering is done using the key Kc (ciphering key) generated during authentication (See section 7.2.1.2) and the ciphering algorithm A5. Algorithm A5 is specified in Annex C of GSM 03.20 [1].
A5 is not PLMN specific. However several A5 versions are possible and negotiation between the MS and the network is carried out to decide on which A5 to use.

7.2.1.3.2 Key setting

The key Kc is generated during the authentication procedure (See section 7.2.1.2). Kc is calculated at the same time as SRES and is transmitted from the AuC to the VLR together with SRES. Algorithm A8, which is PLMN specific is used for the generation of Kc.
Key setting is schematised in Figure 7-3.

Figure 7-3 : Key setting

7.2.1.3.3 Starting the ciphering and deciphering processes

A distinction is made between data on a Dedicated Control Channel (DCCH) and data on a Traffic Channel (TCH).
On a DCCH the start of enciphering is under control of the network. See Figure 7-4.
The BSS sends in clear text a message “Start cipher” and deciphering is started in the BSS. The MS starts enciphering and deciphering and sends its next message (any message) enciphered. When this message is deciphered correctly in the BSS, enciphering is started in the BSS.
On a TCH enciphering and deciphering are started as soon as a key is present, unless “Null Cipher” mode is selected.

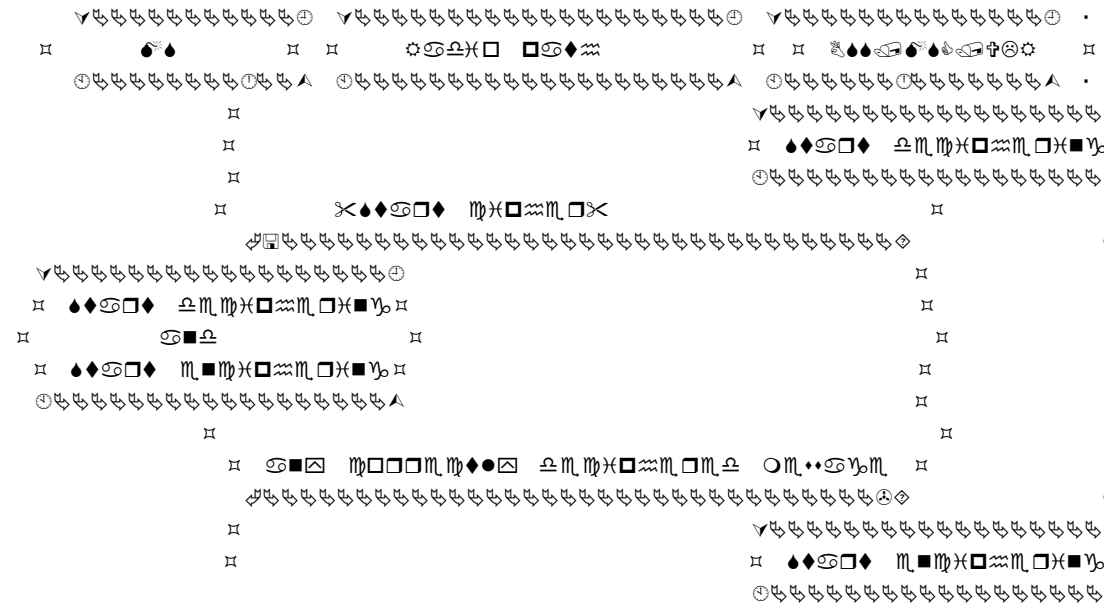


Figure 7-4 : Starting of the enciphering and deciphering processes

7.2.1.3.4 Synchronisation

The enciphering stream at one end and the deciphering stream at the other end must be synchronised, for the enciphering bit stream and the deciphering bit streams to coincide. The underlying Synchronisation scheme is described in Annex C of [1].

7.2.2 DECT security

DECT security features are standardised in [17].

DECT defines following security features:
- authentication of the DECT terminal
- authentication of the DECT network operator
- mutual authentication of DECT terminal and DECT network operator
- data confidentiality on the common air interface
- User authentication

Mutual authentication may be provided by combining the two other authentication security features.

The user authentication service allows the network or fixed part to authenticate a user of a DECT by checking a User Personal Identity (UPI) value associated with that user. This service is similar to on-line personal identity number (PIN) verification provided by banking systems.

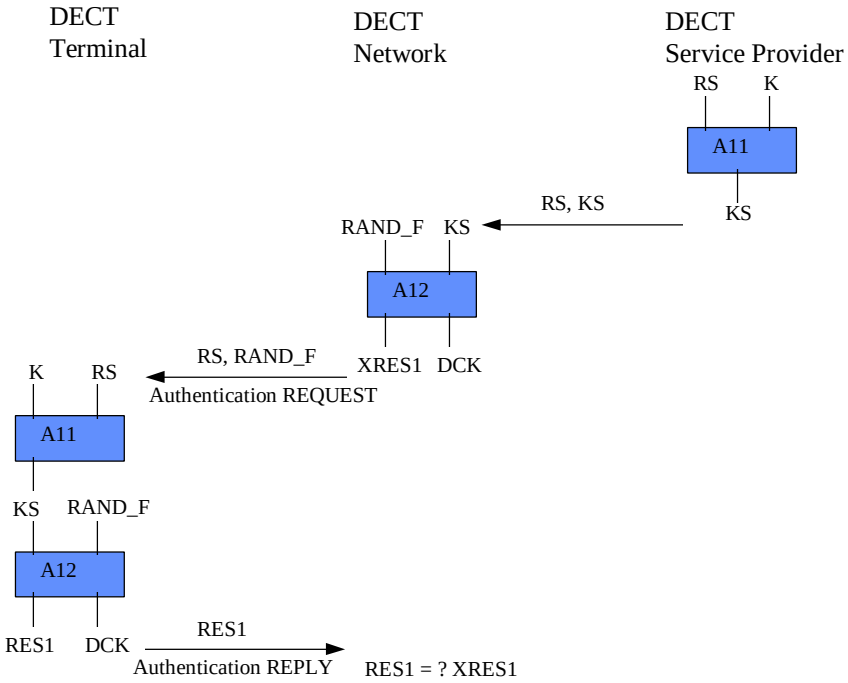
7.2.2.1 Authentication

In the following sections the authentication of the DECT terminal and the DECT network operator are described. The DECT standard leaves multiple options on how to implement the authentication mechanism. The mechanism described below uses the session key for roaming with mutual authentication, with the possibility of using different algorithms.

In DECT the user is normally associated with a specific terminal, since the DECT Access Module, is currently still under standardisation and not yet commonly used. As a result, the user's identity is contained in the DECT terminal.

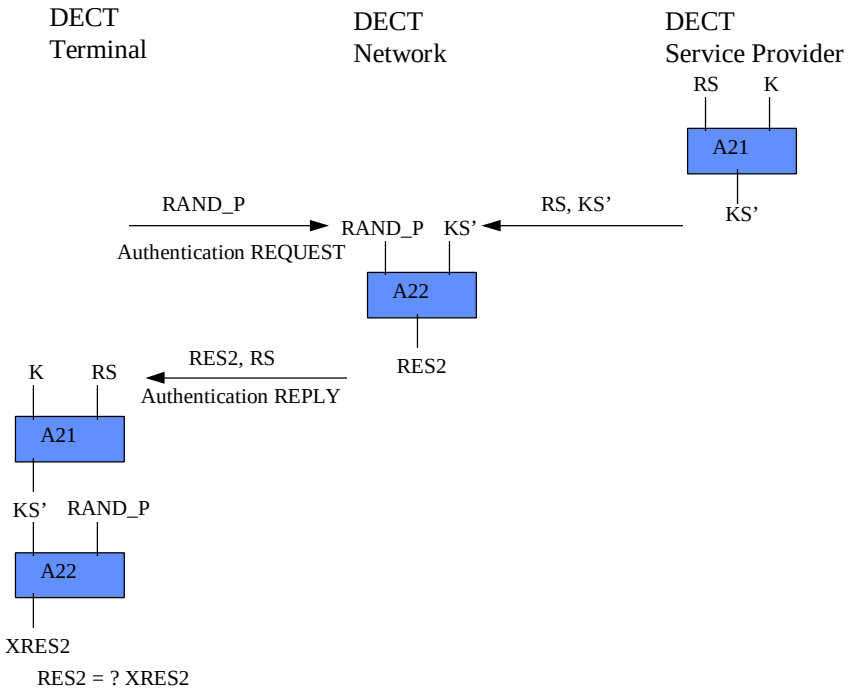
Following symbols are used:
K secret user key, uniquely assigned to one user, only known by the terminal/user and the service provider
RS random number for generation of session key
KS session key for authentication of DECT terminal
KS' session key for authentication of the DECT network operator
RAND_P random number generated by the mobile terminal
RAND_F random number generated by the network
XRES1 response calculated by authenticating entity, for terminal to network authentication
RES1 response calculated by entity to be authenticated, for terminal to network authentication
XRES2 response calculated by authenticating entity, for network to terminal authentication
RES2 response calculated by entity to be authenticated, for network to terminal authentication
A11 session key generation algorithm for mobile to network authentication
A12 authentication algorithm for mobile to network authentication
A21 session key generation algorithm for network to mobile authentication
A22 authentication algorithm for network to mobile authentication

7.2.2.1.1 Authentication of the DECT terminal



The DCK (derived cipher key) will be used to encrypt the air interface, by means of a stream cipher algorithm. An algorithm proposed by DECT is DSC (Dect standard cipher).

7.2.2.1.2 Authentication of the DECT network operator



For the algorithms A11, A21, A12 and A22 DECT proposes to use the DSAA (DECT standard authentication algorithm).

If session keys are used for roaming, A11 and A21 can be chosen to be service provider specific.

An established session key KS/KS' can remain valid during longer periods, this avoids signalling to the home network for future authentications.

7.3 Security in third generation systems

7.3.1 Security requirements

The Security requirements for UMTS are listed in ETSI SMG 050901 [7]. They are divided into :

- customer security requirements
- provider security requirements
- supplementary service security requirements
- UPT security requirements
- satellite security requirements
- PMR security requirements

Only customer and provider security requirements are listed in ETSI SMG 050901. The other sections are still open.

7.3.2 Mechanisms available

Currently detailed proposals are only available for authentication mechanisms.
All proposed authentication mechanisms do more than just authentication of the mobile user to the network and vice versa.

Security features realised by all of them are:

- mutual authentication of the user and the network
- key agreement for ciphering of the air interface
- confidentiality of the user identity on the air interface

In addition some of them realise:

- authentication of the service provider towards the user
- non-repudiation of part of the transmitted data
- assurance that certificates are not revoked
- ...

In the following sections 3 currently proposed authentication mechanisms are described in a unified notation.
Each mechanism still has some proprietary symbols and notation, depending on the specific needs.
For each mechanism a list of provided security features is given and a description of how the security feature is provided.

The three mechanisms are:

- A challenge-response mechanism using symmetric key techniques (Royal Holloway)
- A public-key based mechanism (Siemens)
- A public-key based mechanism (KPN)

Currently in SMG SG the definition of a general framework is under study.

7.3.2.1 Used Symbols and Notations

This section lists the symbols and notations used in the following protocols (section 7.3.2.2 to 7.3.2.4).

$\|\|$	concatenation
A_K	session key generation algorithm : This algorithm takes as input a secret key and a data string and outputs a session key K_S.

	A_K (K,data) denotes the encipherment of data with algorithm A_K and key K.
A_N	user-network operator key generation algorithm : This algorithm takes as input a secret key and a data string and outputs a user-network operator secret key K_{NU}. A_N (K,data) denotes the encipherment of data with algorithm A_N and key K.
A_S	Service provider authentication algorithm : This algorithm takes as input a secret key and a data string and outputs a check value AUTH. A_S (K,data) denotes the encipherment of data with algorithm A_S and key K.
A_U	User authentication algorithm : This algorithm takes as input a secret key and a data string and outputs a check value AUTH. A_U (K,data) denotes the encipherment of data with algorithm A_U and key K.
$AUTH_N$	Value to authenticate the network operator to the user, mostly this will be a challenge response value.
$AUTH_S$	Value to authenticate the service provider to the user, mostly this will be a challenge response value.
$AUTH_U$	Value to authenticate the user to the network operator, mostly this will be a challenge response value.
CA	Certification Authority
CertN	a valid certificate, issued by a certification authority CA, on the public key of the asymmetric signature system of N, available at N
CertU	a valid certificate, issued by a certification authority CA, on the public key of the asymmetric signature system of U, available at U
$CIPH_N$	Data used to conceal $TMUI_N$
$CIPH_S$	Data used to conceal $TMUI_S$
CS	Certificate Server, may coincide with the service provider of the user
C_U	anonymity algorithm : This algorithm takes as input a secret key, a data string, and possibly a key offset, and outputs a string CIPH used to conceal a temporary user identity. C_U(K, data, KO) denotes the encipherment of data with algorithm C_U, key K and key-offset KO (optional).
data1, data2, data3	optional data fields
Dec	A decryption algorithm, corresponding with the encryption algorithm Enc
Enc	A symmetric encryption algorithm. Enc(K,data) means data encrypted with encryption algorithm Enc and key K
expon	modular exponentiation
f	may be the identity function or a compressing function. The requirement is that f(x) depends on all bits of x.
g	generator g, known by UIM/terminal, Network Operator, and Service Provider. g is a generator of a finite group G, e.g. the multiplicative group of a finite field or a subgroup of an elliptic curve, in which the Discrete Logarithm Problem is hard.
g^s	public key agreement key of the Network Operator.
g^u	public key agreement key of the Certificate Server.
h1	one-way function
h2	hash function
h3	hash function
id_{ca}	identity of the Certification Authority
id_{cs}	identity of the certificate server
id_{no}	identity of the network operator
IMUI	International Mobile User Identity
K_{NU}	Secret authentication key shared between user and network operator
KO	Key Offset

K_s	Secret session key shared between user and network operator
K_{SU}	Secret authentication key shared between user and service provider
L	length of session key K_s
NO	Network operator
NOID	Network Operator Identification
p	a prime
PK_CA	public key of Certification Authority used to verify signatures from the Certification Authority.
PK_CS	Public key of the Certificate Server used to verify signatures from the Certificate Server.
PK_NO	Public key of the Network Operator used to verify signatures from the Network Operator.
PK_SP	Public key of the Service Provider used to verify signatures from the Service Provider.
PK_U	Public key of the user used to verify signatures from the UIM/Terminal.
q	the order of the generator g, with q | p - 1 a large prime
RND_N	Random challenge, generated by the network operator
RND_U	Random challenge, generated by the user
s	secret key agreement key of the Network Operator.
Sig_{cs}	A secret signature transformation owned by the certificate server.
Sig_{no}	A secret signature transformation owned by the network operator.
Sig_u	A secret signature transformation owned by the user.
SK_CA	Secret key of the Certification Authority used to generate signatures.
SK_CS	Secret key of the Certificate Server used to generate signatures.
SK_NO	Secret key of the Network Operator used to generate signatures.
SK_SP	Secret key of the Service Provider used to generate signatures.
SK_U	Secret key of the user used to generate signatures.
$TMUI_N$	Temporary user identity assigned by a network operator, used for identification of the user towards the network. The notation $TMUI_N$' is a new identifier.
$TMUI_S$	Temporary user identity assigned by a service provider, used for identification of the user towards the service provider. The notation $TMUI_S$' is a new identifier.
TS1	Time Stamp 1
u	secret key agreement key of the Certificate Server.
UIM/Terminal	The combination of the UIM and the terminal denoting the user. No decision is made here whether the key storage or algorithm calculation are done in the UIM or in the terminal.
Ver_{cs}	A verification algorithm, corresponding with the signature algorithm Sig_{cs}
Ver_{no}	A verification algorithm, corresponding with the signature algorithm Sig_{no}
Ver_u	A verification algorithm, corresponding with the signature algorithm Sig_u

7.3.2.2 A challenge-response mechanism using symmetric key techniques (Royal Holloway)

In SMG 050901, Security principles for the UMTS [7], clause 9, a security mechanism is defined.
It is a challenge - response mechanism using symmetric key techniques, providing mutual authentication between user and network operator/service provider, and incorporating user identity confidentiality and session key generation. The mechanism was developed at Royal Holloway, University of London.
In clause 11 of [7] this mechanism is mapped into the UMTS functional model.

There are two cases : *Current Registrations* where the user is already registered with the network operator where it is currently roaming. The user and the network operator already share a temporary identity $TMUI_N$ and secret key $K_{NU.}$
New Registrations where user and network operator don't share any secret information.

7.3.2.2.1 Current registrations

7.3.2.2.1.1 PROTOCOL GOALS

- mutual authentication between Network Operator and UIM/Terminal.
- establishment of a new session key K_S with mutual key authentication, mutual key freshness assurance
- assignment of a new temporary identity $TMUI_N$'.
- user identity confidentiality on the air-interface.
- user identity confidentiality to the Network Operator.

7.3.2.2.1.2 PREREQUISITES ON MECHANISM

- The UIM/Terminal has a temporary identifier $TMUI_N$ which is assigned by the Network Operator during a previous registration. The $TMUI_N$ is unique to the registering UIM/Terminal and is only known to the Network Operator and the UIM/Terminal.
- The UIM/Terminal and the Network Operator share a unique secret key K_{NU} which is assigned by the user's Service Provider during a previous new registration.

7.3.2.2.1.3 DESCRIPTION OF THE PROTOCOL

The message flow is shown in Figure 7-5 .

The mechanism consists of three messages exchanged between the user and the network operator. The service provider is not involved. The three messages are indicated in the figure with M1, M2 and M3.

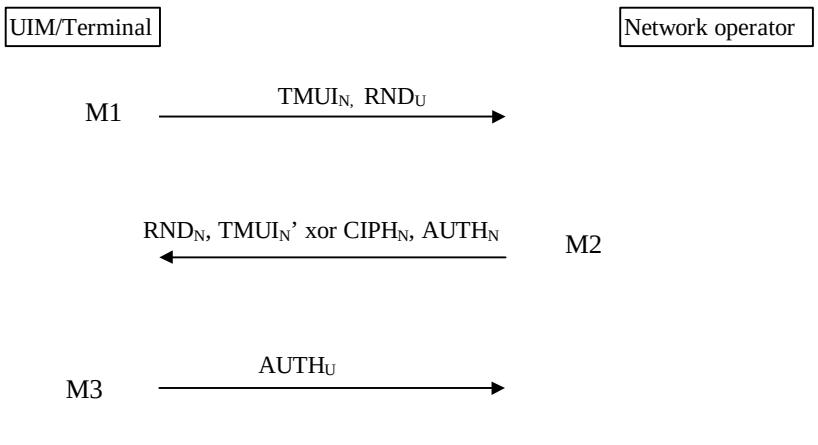


Figure 7-5 : Royal Holloway, Current registrations

The UIM/Terminal generates a random number RND_U.

Message M1:

The UIM/Terminal sends his temporary identity $TMUI_N$ and RND_U to the network operator.
The Network operator generates RND_N and a new temporary identity $TMUI_N$' for subsequent use between the UIM/Terminal and the network operator, and calculates :

- $AUTH_N = A_U (K_{NU}, RND_N \| RND_U \| TMUI_N')$
- $CIPH_N = C_U (K_{NU}, RND_U)$. $CIPH_N$ will conceal the new temporary identity $TMUI_N$' when it is transmitted the first time to the UIM/Terminal (message M2).
- a session key $K_S = A_K (K_{NU} , RND_U \| RND_N \| TMUI_N')$.

Message M2:

The network operator sends RND_N , $TMUI_N$' xor $CIPH_N$ and $AUTH_N$ to the UIM/Terminal. While the Network Operator waits for message M3 it can calculate $AUTH_U = A_U (K_{NU}, RND_U \| RND_N)$.
The UIM/Terminal calculates

- $CIPH_N$ in the same way as the network operator did.
- $TMUI_N' = (TMUI_N' \text{ xor } CIPH_N) \text{ xor } CIPH_N$
- $AUTH_U$, K_S and $AUTH_N$ in the same way as the network operator did.

The UIM/Terminal compares the received $AUTH_N$ with the calculated one.

Message M3:

The UIM/Terminal sends $AUTH_U$ to the Network operator. The Network operator compares the received $AUTH_U$ with the previously calculated one.

7.3.2.2.1.4 ACHIEVED GOALS

Session Key authentication of the UIM/Terminal to the Network operator :
No third party can calculate K_S because K_S is calculated with the secret information $TMUI_N$' and secret key K_{NU}.

Session Key confirmation of the UIM/Terminal to the Network operator :
Not achieved.

Session Key authentication of the Network Operator to the UIM/Terminal :
No third party can calculate K_S because K_S is calculated with the secret information $TMUI_N$' and secret key K_{NU}. Verification of $AUTH_N$ assures the UIM/Terminal that $TMUI_N$' is assigned by the Network Operator.

Session Key confirmation of the Network Operator to the UIM/Terminal :
Not achieved.

Assurance to the UIM/Terminal that the Session key is fresh:
K_S is based on RND_U , generated by the UIM/Terminal.

Assurance to the Network Operator that the Session key is fresh:
K_S is based on RND_N, generated by the Network Operator.

Entity authentication of the UIM/Terminal to the Network Operator:
By verifying $AUTH_U$ which is based on data RND_N and secret key K_{NU}.

Entity authentication of the Network Operator to the UIM/Terminal:
By verifying $AUTH_N$ which is based on data RND_U and secret key K_{NU}.

Non-repudiation of data sent by the UIM/Terminal:
Not achieved.

Non-repudiation of data received by the UIM/Terminal:
Not achieved.

Confidentiality of the user identity
The temporary identity $TMUI_N$ hides the real user identity on the air-interface and also to the Network Operator. A new temporary identity $TMUI_N$' is assigned by the Network Operator to avoid tracing of the user. $TMUI_N$' becomes the new $TMUI_N$ which will be used in cleartext during the next registration or other communications (like call-setup).

7.3.2.2.2 New registrations

7.3.2.2.2.1 PROTOCOL GOALS

In addition to the goals for "Current registrations" (section 7.3.2.2.1)

- authentication from the Service Provider to the UIM/Terminal.
- establishment of a new user-network operator secret key K_{NU}
- assignment of a new temporary identity $TMUI_S$'.

7.3.2.2.2.2 PREREQUISITES ON MECHANISM

- The UIM/Terminal has a temporary identifier $TMUI_S$ which is assigned by the Service Provider during a previous registration. The $TMUI_S$ is unique to the registering user and is only known to the Service Provider and the UIM/Terminal.
- The UIM/Terminal and the Service Provider share a unique secret key K_{SU}.
- The identification of the Network Operator (NOID) is used by the UIM/Terminal and the Service Provider to calculate the secret key K_{NU}. They must both know this identification in the same format. If necessary it should be added in one of the messages.

7.3.2.2.2.3 DESCRIPTION OF THE PROTOCOL

The message flow is shown in Figure 7-6.

The mechanism consists of five messages exchanged between the UIM/Terminal, the network operator, and the service provider of the user. The five messages are indicated in the figure with M1, M2, M3, M4 and M5.

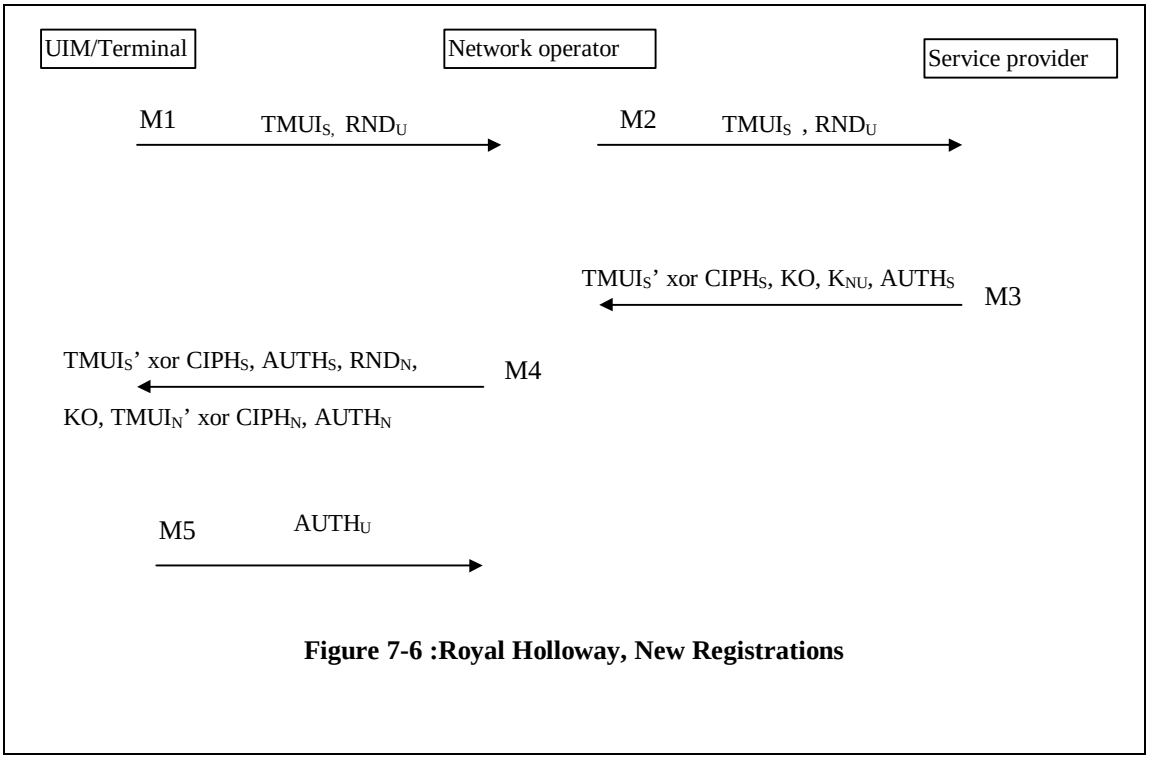


Figure 7-6 :Royal Holloway, New Registrations

The UIM/Terminal generates RND_U.

Message M1:

The UIM/Terminal sends $TMUI_S$ and RND_U to the Network operator .

Message M2:

The Network operator sends $TMUI_S$ and RND_U to the Service Provider.

The Service Provider calculates :

- $CIPH_S = C_U(K_{SU}, RND_U, KO)$. $CIPH_S$ will conceal the new temporary identity $TMUI_S$' when it is transmitted the first time to the UIM/Terminal (message M3 and M4).
- The service provider generates a new temporary identity $TMUI_S$'
- $AUTH_S = A_S (K_{SU}, RND_U \| TMUI_S')$
- $K_{NU} = A_N(K_{SU}, NOID, KO)$

Message M3:

The Service Provider sends $TMUI_S$' xor $CIPH_S$, KO, K_{NU} , $AUTH_S$ to the Network operator. The Network operator calculates $AUTH_N$, $AUTH_U$, $CIPH_N$ and K_S as in section 7.3.2.2.1.3 (Current registrations), and generates a new temporary identity $TMUI_N$'.

Message M4:

The Network operator sends $TMUI_S$' xor $CIPH_S$, $AUTH_S$, RND_N, KO, $TMUI_N$' xor $CIPH_N$, $AUTH_N$ to the UIM/Terminal :

The UIM/Terminal calculates
- $CIPH_N$ and $CIPH_S$ in the same way as the network operator and the Service Provider did, respectively.
- $TMUI_S$' = ($TMUI_S$' xor $CIPH_S$) xor $CIPH_S$
- $TMUI_N$' = ($TMUI_N$' xor $CIPH_N$) xor $CIPH_N$
- $AUTH_S$ and K_{NU} in the same way as the Service Provider did.
- $AUTH_U$, $AUTH_N$ and K_S in the same way as the Network Operator did.

The UIM/Terminal compares the received $AUTH_N$ and $AUTH_S$ with the calculated values.

Message M5:

The UIM/Terminal sends $AUTH_U$ to the Network operator.
The Network operator compares $AUTH_U$ with the calculated value.

7.3.2.2.2.4 ACHIEVED GOALS

In addition to the goals achieved for "Current registrations " (section 7.3.2.2.1) :

Entity authentication of the Service Provider to the UIM/Terminal:

By verifying $AUTH_S$ which is based on RND_U generated by the user and secret key K_{SU}

Establishment of a new key K_{NU} :

The new user-network operator secret key K_{NU} can only be calculated by the UIM/Terminal and the Service Provider and is sent to the Network Operator in message M3.

Assignment of a new temporary identity $TMUI_S$':

Verifying $AUTH_S$ assures the UIM/Terminal that $TMUI_S$' is generated by the Service Provider. $TMUI_S$' is sent encrypted to the UIM/Terminal.

7.3.2.3 A public-key based mechanism (Siemens)

This mechanism is defined by Siemens AG in a contribution to SMG sg (SMG SG Doc 73/95) [18]. There are three versions of the protocol, called protocol A, B and C

7.3.2.3.1 PROTOCOL A

This is the case where authentic copies of public keys of the UIM/Terminal and the network operator are already available at the Network Operator and the UIM/Terminal respectively and, hence, are not exchanged in the course of the protocol.

7.3.2.3.1.1 PROTOCOL GOALS

- mutual explicit authentication of User and Network operator
- agreement between the user and the Network operator on a shared secret key K_S with mutual implicit key authentication
- mutual key confirmation of the User and the Network operator
- mutual assurance of key freshness
- non-repudiation by the User of data sent by the User to the Network operator
- confidentiality of the identity IMUI of the User on the air interface

7.3.2.3.1.2 PREREQUISITES ON MECHANISM

- The identity of the Network Operator is assumed to be known to the UIM/Terminal at the start of the protocol.
- There is a finite group G with generator g, e. g. the multiplicative group of a finite field or a subgroup of an elliptic curve, in which the Discrete Logarithm Problem is hard.
- The Network Operator has secret and public key agreement keys s and g^s respectively.
- The UIM/Terminal possesses an asymmetric signature system with secret signature transformation Sig_u. In the case of a signature with message recovery, $Sig_u(M)$ denotes the signature itself. In the case of a signature with appendix, $Sig_u(M)$ denotes only the appendix.
- An authentic copy of the public key PK_U of the asymmetric signature system of the UIM/Terminal is available at the Network Operator . (It may have been distributed to the Network Operator in an earlier run of protocol B or protocol C.)
- An authentic copy of the public key agreement key g^s of the Network Operator is available at the UIM/Terminal . (It may have been distributed to the UIM/Terminal in an earlier run of protocol B or protocol C.)
- It is assumed for protocol A that a compromise of the secret signature transformation Sig_u of the UIM/Terminal need not be taken into account for non-repudiation.

7.3.2.3.1.3 DESCRIPTION OF THE PROTOCOL

The message flow is shown in Figure 7-7.

The mechanism consists of three messages exchanged between the user and the network operator. The service provider is not involved. The three messages are indicated in the figure with M1, M2 and M3.

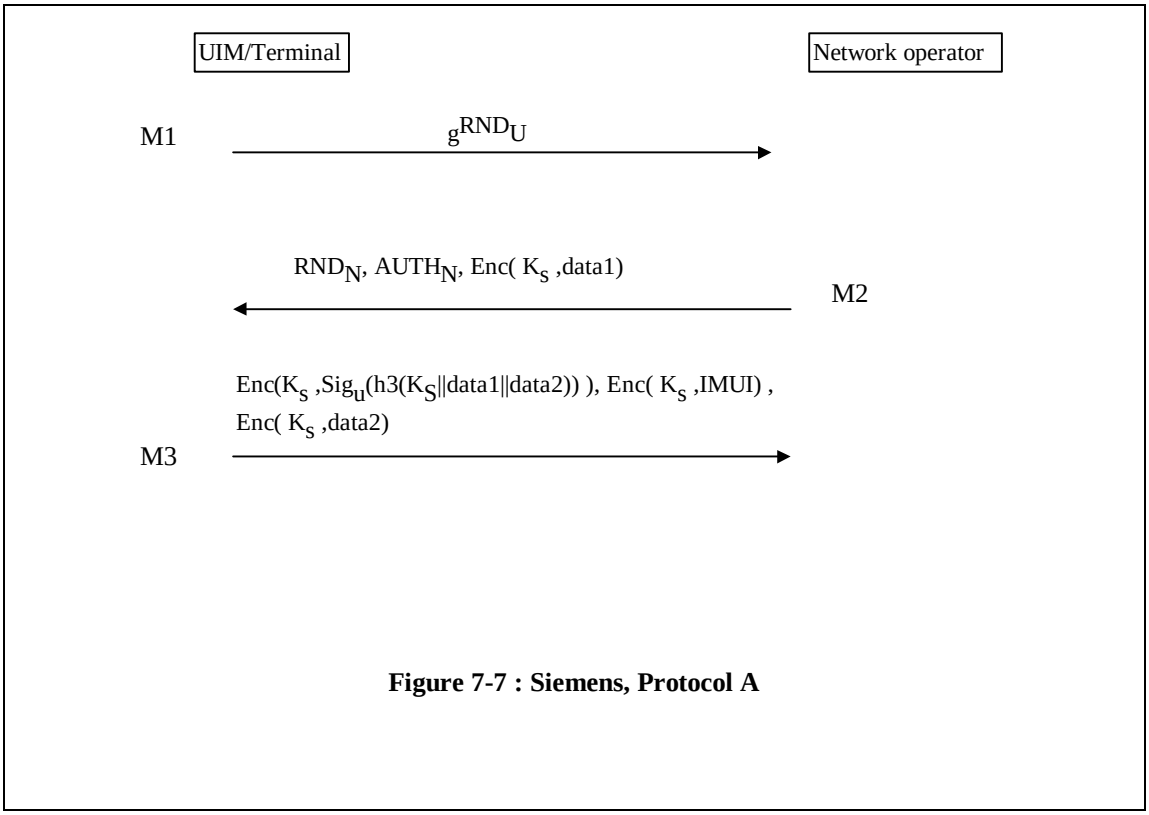


Figure 7-7 : Siemens, Protocol A

The UIM/Terminal calculates g^{RND_U}

Message M1:

The UIM/Terminal sends g^{RND_U} to the Network operator.
The Network operator calculates :

- $(g^{RND_U})^{s}$
- the session key K_S = h1($(g^{RND_U})^{s}$ || RND_N)
- $AUTH_N$ = h2(K_S)

Message M2:

The Network operator sends RND_N, $AUTH_N$ and Enc(K_S , data1)to the UIM/Terminal.
The UIM/Terminal calculates:

- $(g^{s})^{RND_U}$
- the session key K_S = h1($(g^{s})^{RND_U}$ || RND_N)
- $AUTH_N$ = h2(K_S)
- Enc(K_S ,Sig_u(h3(K_S||data1||data2)))
- Enc(K_S ,IMUI)
- Enc(K_S ,data2)

$AUTH_N$ is compared with the one received from the Network operator.

Message M3:

The UIM/Terminal sends Enc(K_S ,Sig_u(h3(K_S||data1||data2))), Enc(K_S ,IMUI) and
Enc(K_S , data2)to the Network operator.

The Network Operator
- decrypts every part in the message with decryption algorithm Dec and session key K_S
- thus learns the IMUI and knows which public key (PK_U) he has to retrieve from his database in order to verify the signature
- knows K_S, data1 and data2 and calculates $h3(K_S||data1||data2)$
- retrieves $h3(K_S||data1||data2)$ from $Sig_u(h3(K_S||data1||data2))$ with verification algorithm Ver_u and key PK_U.and compares the two values.

7.3.2.3.1.4 ACHIEVED GOALS

Session Key authentication of the UIM/Terminal to the Network operator :
Implicitly by including a hash value in the third message : $h3(K_S||data1||data2)$

Session Key confirmation of the UIM/Terminal to the Network operator :
by including a hash value in the third message : $h3(K_S || data1 || data2)$

Session Key authentication of the Network Operator to the UIM/Terminal :
K is derived from the Network Operator secure key s

Session Key confirmation of the Network Operator to the UIM/Terminal :
by including a hash value h2(K) in the second message.
Instead of using a hash function for key confirmation, one could encrypt with K a quantity known to both sides after the reception of M1.

Assurance to the UIM/Terminal that the Session key is fresh:
The session key is derived from the random value RND_U.

Assurance to the Network Operator that the Session key is fresh:
The session key is derived from the random value RND_N.

Entity authentication of the UIM/Terminal to the Network Operator:
By including a signature on $h3(K_S)$ in message M3, knowing that K_S is based on RND_N.

Entity authentication of the Network Operator to the UIM/Terminal:
Implicitly by providing key authentication, confirmation and freshness, since the key is based on the secure key s and RND_U, which is generated by the UIM/Terminal.

Non-repudiation of data sent by the UIM/Terminal:
data2 sent signed ($Sig_u(h3(K_S||data1||data2))$) and encrypted ($Enc(K_S,data2)$) to the Network Operator

Non-repudiation of data received by the UIM/Terminal:
data1 received from the Network Operator is sent back signed ($Sig_u(h3(K_S||data1||data2))$) to the Network Operator.

Confidentiality of the user identity
by encrypting the IMUI in the third message. Also the signature in the third message is encrypted because an attacker may be able to detect the identity of the UIM/Terminal by verifying the signature. Whether this is possible - and hence, whether encryption of the signature is truly needed - depends on the particular signature system and padding method chosen.

7.3.2.3.2 PROTOCOL B

The protocol is executed between the UIM/Terminal and the network operator if a valid certificate on the public verification key PK_U of the UIM/Terminal is available at the UIM/Terminal, but not at the Network Operator, and a valid certificate on the public key agreement key g^s of the Network Operator is available at the Network Operator, but not at the UIM/Terminal.

7.3.2.3.2.1 PROTOCOL GOALS

In addition to the protocol goals defined for protocol A :
- exchange of certified public keys between U and N

7.3.2.3.2.2 PREREQUISITES ON MECHANISM

The same as for protocol A, except that

- The UIM/Terminal has no authentic copy of the public key agreement key g^S of the Network Operator.
- The Network Operator has no authentic copy of the public verification key PK_U of the UIM/Terminal.
- There is a valid certificate CertU, issued by a certification authority CA, on the public key PK_U of the asymmetric signature system of the UIM/Terminal, available at the UIM/Terminal.
- There is a valid certificate CertN, issued by a certification authority CA on the public key agreement key g^S of the Network Operator , available at the Network Operator.
- It is assumed for protocol B that a revocation of the certificate CertU need not be taken into account for non-repudiation.
- The UIM/Terminal and the Network Operator possess the public key necessary to verify certificates issued by CA (PK_CA)

7.3.2.3.2.3 DESCRIPTION OF THE PROTOCOL

The message flow is shown in Figure 7-8.

The difference with protocol A is that the UIM/Terminal does not know the public key of the Network Operator and the Network Operator does not know the public key of the UIM/Terminal. Therefore the UIM/Terminal will include in the first message (M1) the identification of the certification authority of which the UIM/Terminal can verify signatures (id_{CA}). The Network Operator will include in the second message (M2) his certificate signed by the corresponding certification authority (CA). The UIM/Terminal can verify this certificate (CertN) and retrieves the public key agreement key g^S of the Network Operator which is used for calculation of $(g^S)^{RND}U$. In the third message (M3) the certificate of the UIM/Terminal (CertU) is encrypted (Enc(K_S ,CertU)) in stead of the IMUI. After receipt of message M3 the Network Operator retrieves the public key of the UIM/Terminal (PK_U) from the user's certificate and uses it for the other calculations.

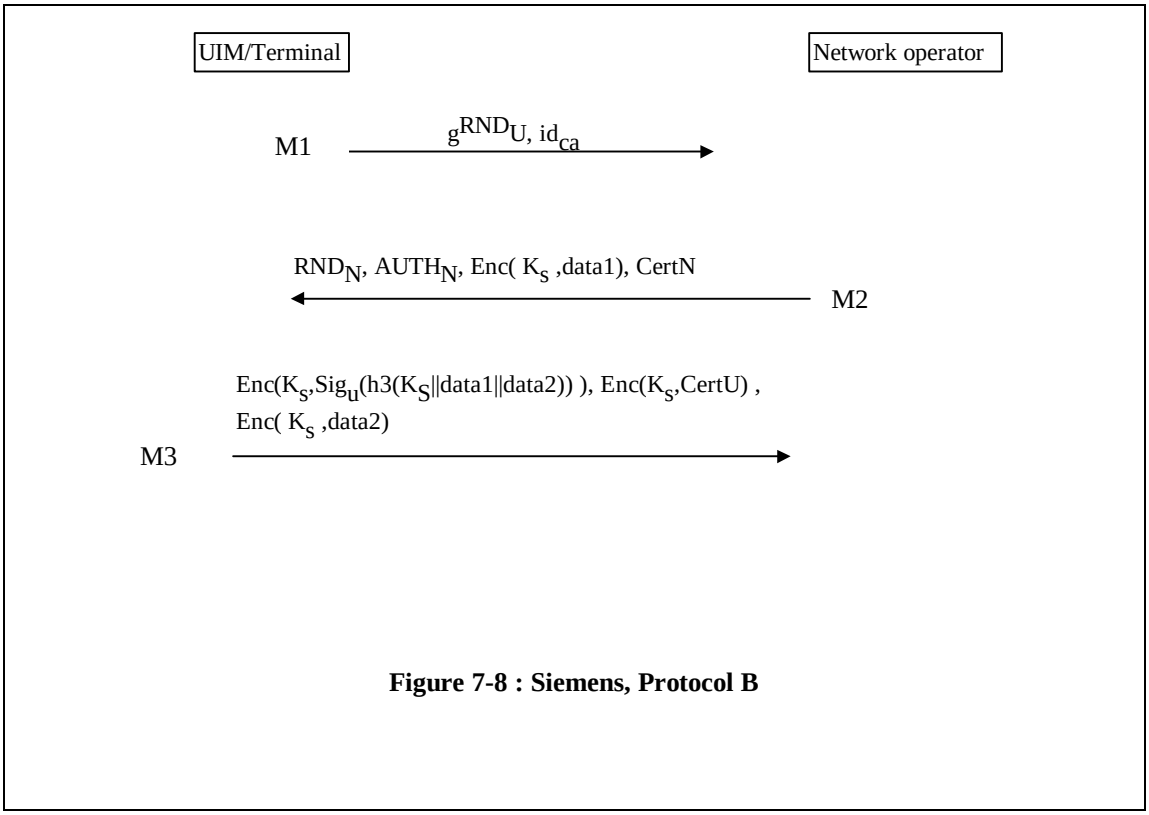


Figure 7-8 : Siemens, Protocol B

7.3.2.3.2.4 ACHIEVED GOALS

The same goals are achieved in the same way as for protocol A except for :
Confidentiality of the user identity:

by encrypting the user certificate CertU in stead of the IMUI in the third message.

Exchange of certificates:

id_{ca} is sent in message M1 to indicate to the Network Operator which certificates can be verified by the UIM/Terminal. The Network Operator sends a certificate on his public key CertN to the UIM/Terminal in message M2. the UIM/Terminal sends a certificate on his public key CertU to the UIM/Terminal in message M2.

7.3.2.3.3 PROTOCOL C

There is no authentic copy of the public key of the UIM/Terminal available at the network operator. There is no authentic copy of the public key of the network operator available at the UIM/Terminal.

7.3.2.3.3.1 PROTOCOL GOALS

In addition to the protocol goals defined for protocol A :

- distribution of public key PK_U of the UIM/Terminal, certified by a certification authority (CA), from the certificate server (CS) to the Network Operator
- distribution of the public key agreement key g^S of the Network operator, certified by the Certificate Server, from the Network operator to the UIM/Terminal
- assurance for the Certificate Server that the public key it certifies is indeed the public key of the Network operator
- assurance for the UIM/Terminal and the Network operator that the certificates of the Network operator and the UIM/Terminal respectively have not been revoked

7.3.2.3.3.2 PREREQUISITES ON MECHANISM

In addition to A :

- the Certificate Server has access to up-to-date revocation lists relevant for the public keys of the Network Operator and the UIM/Terminal
- Both the Certificate Server and the Network Operator are capable of creating and checking time-stamps TS
- Both the Network Operator and the UIM/Terminal possess the public key PK_CS needed to verify signatures by the Certificate Server
- The Certificate Server possesses a private-public key agreement key pair (u, g^u)
- The UIM/Terminal has an authentic copy of the public key agreement key g^u of the Certificate Server.
- The Certificate Server possesses the public verification key PK_NO needed to verify signatures made by the Network Operator using his private signature key SK_NO.

7.3.2.3.3.3 DESCRIPTION OF THE PROTOCOL

The message flow is shown in Figure 7-9.

The mechanism consists of five messages exchanged between the UIM/Terminal, the network operator, and a Certificate Server trusted by the user. The certificate server CS has access to a certificate on the public key of the UIM/Terminal, issued by a certification authority CA. CS and CA may coincide. CS may be identical with the Service Provider of the UIM/Terminal.
The five messages are indicated in the figure with M1, M2, M3, M4, M5.
The flow of the messages exchanged between the UIM/Terminal and the Network Operator is identical to that in protocols A and B. Over the air interface, the protocol itself is also very similar to protocols A and B. The major difference is that now a certified public key of the Network Operator is distributed from the Network Operator to the UIM/Terminal, but no certified public key is distributed from the UIM/Terminal to the Network Operator.
In addition to the messages exchanged between the UIM/Terminal and the Network Operator, there is a two-pass exchange of messages between the Network Operator and the Certificate Server in which Certificate Server distributes public keys of the Network Operator and the UIM/Terminal to the Network Operator which are signed by CS.

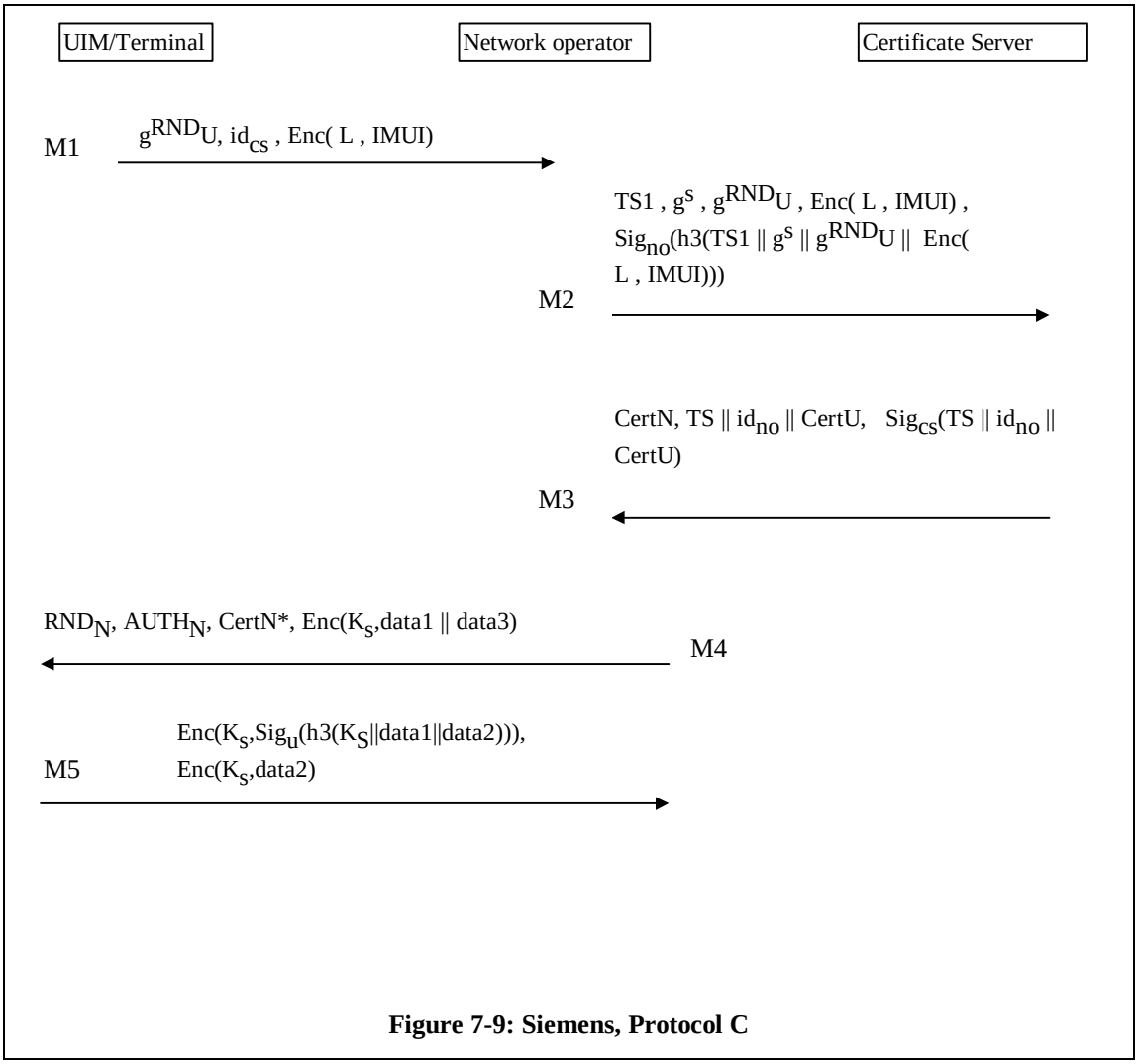


Figure 7-9: Siemens, Protocol C

The UIM/Terminal generates $g^{RND}U$ and calculates :

- $g^{RND}U$
- $L = (g^u)^{RND}U$
- Enc(L , IMUI)

Message M1:

The UIM/Terminal sends $g^{RND}U$, id_{cs} and Enc(L , IMUI) to the Network Operator.

id_{cs} is the identification of the certificate server of which the UIM/Terminal can verify signatures.

The Network Operator retrieves a (possibly new) public key g^s from storage and creates a time-stamp TS1. In the following steps, a certificate on the new key g^s is obtained from the Certificate Server. This facilitates changing private-public key pairs of the Network Operator.

The Network Operator calculates a signature on TS1 || g^s || $g^{RND}U$ || Enc(L , IMUI) with the hash function h3, signature algorithm Sig_{no} and key SK_NO

Message M2:

The Network Operator sends TS1 , g^s , $g^{RND}U$, Enc(L , IMUI) , Sig_{no}(h3(TS1 || g^s || $g^{RND}U$ || Enc(L , IMUI))) to the Certificate Server.

The Certificate Server:

- calculates h3(TS1 || g^s || g^{RND_U} || Enc(L , IMUI)) and verifies the received signature with verification algorithm Ver_{no} and public key PK_NO.
- checks the Time Stamp TS1
- $L = (g^{RND_U})^u$
- decrypts Enc(L , IMUI) with decryption algorithm Dec and key L
- retrieves CertU associated with the obtained IMUI from its database.
- checks the (eventually new) key g^s of the Network Operator and the certificate CertU against revocation lists.
- creates the credentials = g^{RND_U} || g^s || id_{no} || data3 and calculates a certificate on the Network Operator's public key agreement key i.e. CertN which is a signature on the credentials. CertN = credentials, Sig_{cs}(h3(credentials)). data3 is an optional field transmitted to the UIM/Terminal in an authentic way.
- creates a time stamp TS and calculates a signature on TS || id_{no} || CertU.

Message M3 :

The Certificate Server sends CertN, TS || id_{no} || CertU, Sig_{cs}(TS || id_{no} || CertU) to the Network Operator.
The Network Operator:
- verifies the signature on TS || id_{no} || CertU and the CertN with verification algorithm Ver_{cs} and key PK_CS.

The Network Operator calculates :
- a shortened CertN named CertN* = g^s || Sig_{cs}(h3(credentials)).
- $(g^{RND_U})^s$
- the session key K_S = h1($(g^{RND_U})^s$ || RND_N)
- $AUTH_N$ = h2(K_S)
- Enc(K_s ,data1 || data3) encrypted with algorithm Enc and key K_S

data1 is an optional data field sent from the Network Operator to the UIM/Terminal in an authentic way.

Message M4 :

The Network Operator sends RND_N, $AUTH_N$, CertN* and Enc(K_s ,data1 || data3) to the UIM/Terminal.

The UIM/Terminal reconstructs the credentials = g^{RND_U} || g^s || id_{no} || data3 and verifies the signature on CertN (which is Sig_{cs}(h3(credentials)) and is part of CertN*) with verification algorithm Ver_{cs} and key PK_CS.
The UIM/Terminal calculates:
- $(g^s)^{RND_U}$
- the session key K_S = h1($(g^s)^{RND_U}$ || RND_N)
- $AUTH_N$ = h2(K_S)
- data1 || data3 = decryption of Enc(K_s ,data1 || data3) with decryption algorithm Dec and key K_S
- Enc(K_s ,Sig_u(h3(K_S||data1||data2)))
- Enc(K_s ,data2)

$AUTH_N$ is compared with the one received from the Network operator.

Message M5 :

The UIM/Terminal sends Enc(K_S ,Sig_u(h3(K_S||data1||data2))), Enc(K_S ,data2) to the Network Operator.
The Network Operator

- decrypts every part in the message with decryption algorithm Dec and session key K_S
- knows K_S, data1 and data2 and calculates h3(K_S||data1||data2)
- retrieves h3(K_S||data1||data2) from Sig_u(h3(K_S||data1||data2)) with verification algorithm Ver_u and key PK_U.and compares the two values.

7.3.2.3.3.4 ACHIEVED GOALS

The same goals are achieved in the same way as for protocol A except for :
Confidentiality of the user identity :

by encrypting the user identity IMUI in the first message with secret key L which can only be calculated by the UIM/Terminal and the Certificate Server.

Exchange of certificates :

id_{CS} is sent in message M1 to indicate to the Network Operator which certificates can be verified by the UIM/Terminal. The Certificate Server sends CertN and CertU to the Network Operator. The Network Operator sends a shortened certificate CertN* to the UIM/Terminal in order to reduce the message length on the air interface. The message length is of no concern in the fixed network.

Other remarks on this protocol :

- In addition, in order to achieve non-repudiation, it may be required that the Network Operator submits Sig_{no}(h3(...)) to a time-keeper who is trusted by the UIM/Terminal and who signs Sig_{no}(h3(...)) together with a timestamp and returns it to the Network Operator. Otherwise, the UIM/Terminal could repudiate a signature claiming that the signature was generated by an impostor after the certificate had been revoked. Whether this additional measure should be implemented, however, depends on the security policy and on a trade-off between a higher security level and additional effort. A corresponding exchange of messages is not shown here.
- The time-stamp TS1 gives assurance to CS that the signature was generated recently, and hence that g^s is indeed the public key agreement key of the Network Operator. This gives more flexibility to the Network Operator in changing its key agreement key pair. It may be desirable for the Network Operator to change its key agreement key pair frequently as the compromise of the Network Operator's secret key agreement key also compromises all session keys generated with it in the past.
- The inclusion of g^t in the credentials allows the UIM/Terminal to check that CertN is fresh, i.e. that the public key g^s of the Network Operator is valid
- An optional field data3 is included in the credentials formed by CS. In this field, CS can transmit information to the UIM/Terminal in an authentic way. Furthermore, by including the freshness parameter g^t in the signature, CS gives assurance to the UIM/Terminal that this information is fresh. This may be information which the UIM/Terminal is unable to generate and/or verify (e.g. a time-stamp). This feature may be useful e.g. in cases where the UIM/Terminal is to include such information in the field data2 which is signed in message M3 for non-repudiation purposes
- CS shall send M3 to the Network Operator only if CertU has not been revoked. The signature on the time-stamp TS and on CertU assures that CS cannot deny later having told the Network Operator that CertU had not been revoked at a certain time.

7.3.2.4 A public-key based mechanism (KPN)

This is a proposal from KPN Research (SMG SG Doc 36 rev 1/95) [19].
There are two cases:

- New Registration
  For the New Registration version there are again two versions:
  - The user sends his certificate (CertU) as identification
  - The user sends his identity (IMUI), the network operator retrieves a certificate (CertU) from a certificate server.
- Current Registration.

7.3.2.4.1 New Registration - use of CertU

7.3.2.4.1.1 PROTOCOL GOALS

- mutual authentication of network operator and mobile user
- establishment of a session key shared by network operator and mobile user
- confidentiality over the air interface of the IMUI of the user
- exchange of certificates between the UIM/Terminal and the Network Operator on each others public key needed to verify signatures.

The session key is implicitly verified.

7.3.2.4.1.2 PREREQUISITES ON MECHANISM

- p is a prime, g a generator of order q, with q | p-1 a large prime
- Both the UIM/Terminal and the Network Operator possess the public key PK_CS of the Certificate Server needed to verify certificates signed by the cs.

7.3.2.4.1.3 DESCRIPTION OF THE PROTOCOL

The message flow is shown in Figure 7-10.

The mechanism consists of three messages exchanged between the UIM/Terminal and the Network Operator.
The three messages are indicated in the figure with M1, M2, M3.

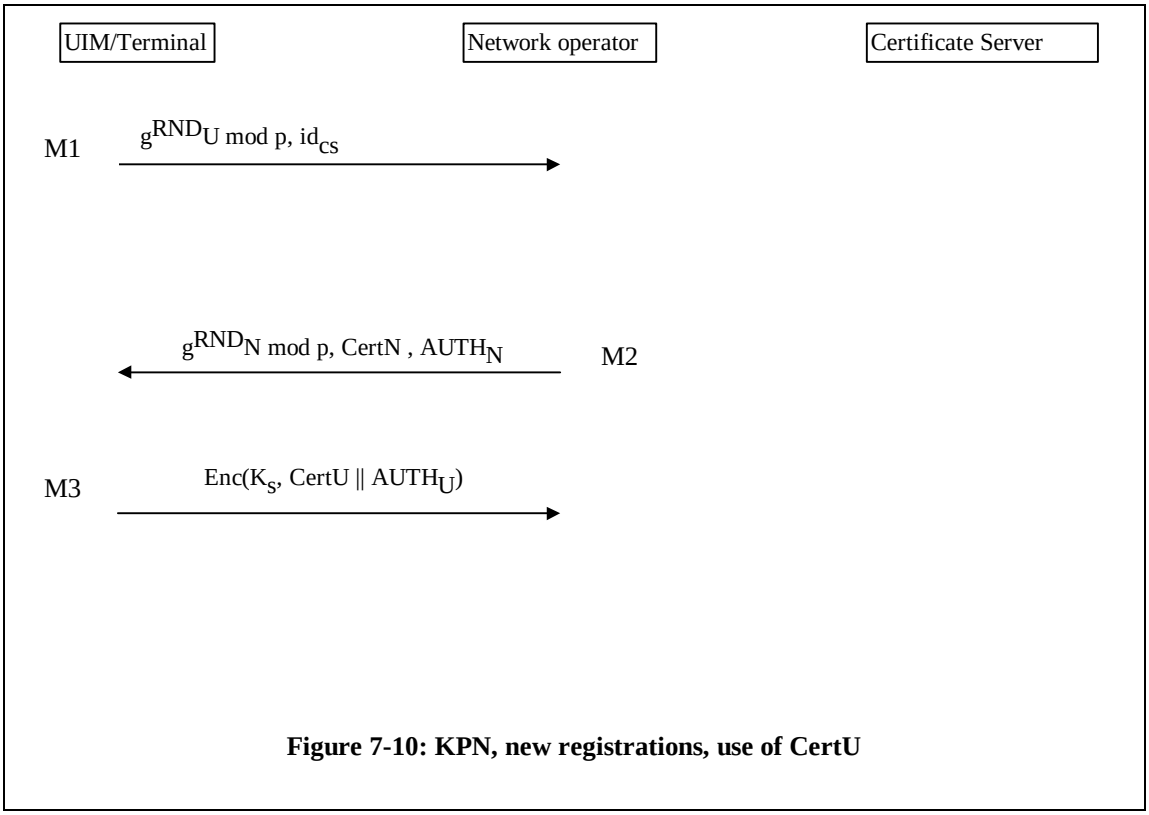


Figure 7-10: KPN, new registrations, use of CertU

The UIM/Terminal generates RND_U with $RND_U < q$ and calculates $(g^{RND_U}) \bmod p$

Message M1:

The UIM/Terminal sends (g^{RND_U}) mod p and id_{cs} to the Network Operator.
id_{cs} is the identification of the Certificate Server of which the UIM/Terminal can verify signatures.
The Network Operator generates RND_N with $RND_N < q$ and:

- calculates $K_S = (((g^{RND_U})^{RND_N}) \bmod p) \bmod 2^L$
- selects a certificate CertN signed by the Certificate Server with identity id_{cs}
- calculates $AUTH_N = Sig_{no}(f(g^{RND_U}) \| f(g^{RND_N}))$

Message M2 :

The Network Operator sends g^{RND_N} mod p, CertN and $AUTH_N$ to the UIM/Terminal.
The UIM/Terminal:

- calculates $K_S = (((g^{RND_N})^{RND_U}) \bmod p) \bmod 2^L$
- verifies CertN with verification algorithm Ver_{cs} and public key PK_CS of the Certificate Server.
- retrieves the public key of the Network Operator (PK_NO) from CertN
- calculates $f(g^{RND_U}) \| f(g^{RND_N})$
- retrieves $f(g^{RND_U}) \| f(g^{RND_N})$ from $AUTH_N$ using the public key of the Network Operator (PK_NO) and the verification algorithm Ver_{no}.and compares the value with the calculated value.
- calculates $AUTH_U = Sig_u(f(g^{RND_N}) \| f(g^{RND_U}))$
- calculates Enc(K_S , CertU || $AUTH_U$)

Message M3:

The UIM/Terminal sends Enc(K_S , CertU || $AUTH_U$)

The Network Operator decrypts this message with decryption algorithm Dec and key K_S :

– verifies CertU with the public key of the Certificate Server.
– retrieves the public key of the UIM/Terminal (PK_U) from CertU
– calculates $f(g^{RND}N) \| f(g^{RND}U)$
– retrieves $f(g^{RND}N) \| f(g^{RND}U)$ from $AUTH_U$ using the public key of the UIM/Terminal (PK_U) and the verification algorithm Ver_u.and compares the value with the calculated value.

7.3.2.4.1.4 ACHIEVED GOALS

Session Key authentication of the UIM/Terminal to the Network operator :

By sending a signature in $AUTH_U = Sig_u(f(g^{RND}N) \| f(g^{RND}U))$ in message M3, the UIM/Terminal proves that $g^{RND}U$ is generated by himself and thus that no third entity can calculate K_S.

Session Key confirmation of the UIM/Terminal to the Network operator :

The UIM/Terminal calculates $f(g^{RND}N)$ and sends this value signed to the UIM/Terminal in message M3. Verification ensures the Network Operator that the UIM/Terminal received the correct value of $g^{RND}N$.Session Key authentication of the Network Operator to the UIM/Terminal :

By sending a signature in $AUTH_N = Sig_{no}(f(g^{RND}U) \| f(g^{RND}N))$ in message M2, the Network Operator proves that RND_N is generated by himself and thus that no third entity can calculate K_S.

Session Key confirmation of the Network Operator to the UIM/Terminal :

The Network Operator calculates $f(g^{RND}U)$ and sends this value signed to the UIM/Terminal in message M2. Verification ensures the UIM/Terminal that the Network Operator received the correct value of $g^{RND}U$..

Assurance to the UIM/Terminal that the Session key is fresh:

The session key is derived from the random value RND_U.

Assurance to the Network Operator that the Session key is fresh:

The session key is derived from the random value RND_N.

Entity authentication of the UIM/Terminal to the Network Operator:

By sending $AUTH_U$ which includes a signature on RND_N

Entity authentication of the Network Operator to the UIM/Terminal:

By sending $AUTH_N$ which includes a signature on RND_U

Non-repudiation of data sent by the UIM/Terminal:

Not achieved.

Non-repudiation of data received by the UIM/Terminal:

Not achieved

Confidentiality of the user identity

By encrypting the certificate of the user CertU in the third message M3.

Exchange of certificates:

The Network Operator sends a certificate on his public key CertN in message M2, the UIM/Terminal sends a certificate on his public key CertU in message M3.

7.3.2.4.2 New Registration - use of IMUI

7.3.2.4.2.1 PROTOCOL GOALS

The same goals as for “New Registration - use of CertU” (section 7.3.2.4.1) except exchange of certificates :

- The Network Operator sends a certificate on his public key to the UIM/Terminal.

- The Network Operator obtains a certificate on the user's public key from the Certificate Server.

7.3.2.4.2.2 PREREQUISITES ON MECHANISM

The same prerequisites as for "New Registration - use of CertU" (section 7.3.2.4.1)

7.3.2.4.2.3 DESCRIPTION OF THE PROTOCOL

The message flow is shown in Figure 7-11.

The mechanism consists of five messages exchanged between the UIM/Terminal and the Network Operator.
The five messages are indicated in the figure with M1, M2, M3, M4 and M5.

The calculations are the same as in the case where the certificate of the UIM/Terminal is used (section 7.3.2.4.1) except that not the user's certificate is sent encrypted in message M3 but some data uniquely identifying this certificate (for example the IMUI). An extra step is then required for the Network Operator to obtain the corresponding certificate from a server.

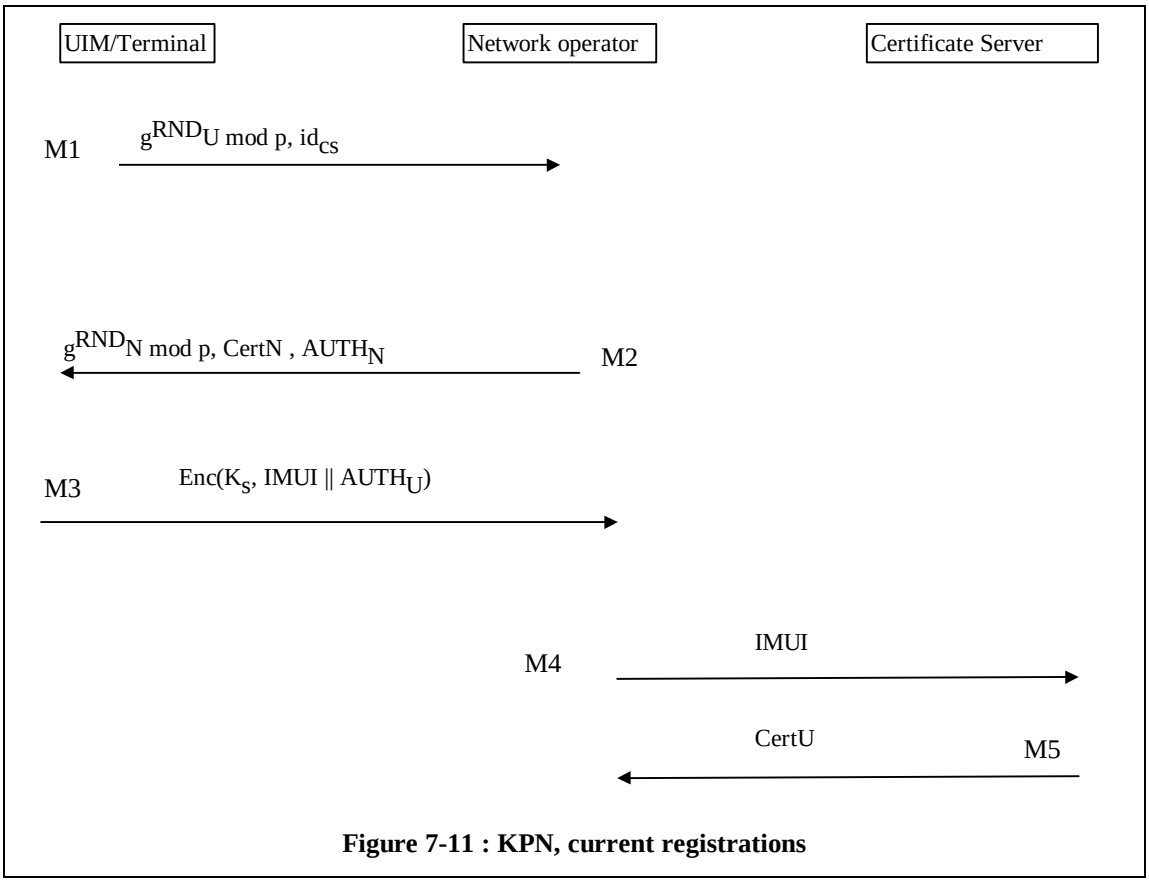


Figure 7-11 : KPN, current registrations

(Only the differences with the previous protocol are mentioned)
Message M3:

The user sends Enc(K_S , IMUI || $AUTH_U$) to the Network Operator.

Message M4:

The Network Operator sends the identity of the UIM/Terminal (IMUI) to the Service Provider or a Certificate Server.

Message M5:

The Service Provider or Certificate Server sends back a user certificate CertU.

7.3.2.4.2.4 ACHIEVED GOALS

The same goals as for “New Registration - use of CertU” (section 7.3.2.4.1) except exchange of certificates.

7.3.2.4.3 Current Registrations

7.3.2.4.3.1 PROTOCOL GOALS

- mutual authentication of network operator and mobile user
- establishment of a session key shared by network operator and mobile user
- confidentiality over the air interface of the IMUI of the user

The session key is implicitly verified.

7.3.2.4.3.2 PREREQUISITES ON MECHANISM

- The UIM/Terminal and the Network Operator have valid certificates on each other’s public key.

7.3.2.4.3.3 DESCRIPTION OF THE PROTOCOL

The message flow is shown in Figure 7-12.
Because the certificates of the other party’s public keys are already available, they are not exchanged in the protocol. CertN is omitted in message M2. The user encrypts his identity IMUI (or some other data uniquely identifying his certificate) in stead of CertU and sends this to the Network Operator.

(Only differences with “New Registration - use of CertU” in section 7.3.2.4.1 are mentioned).

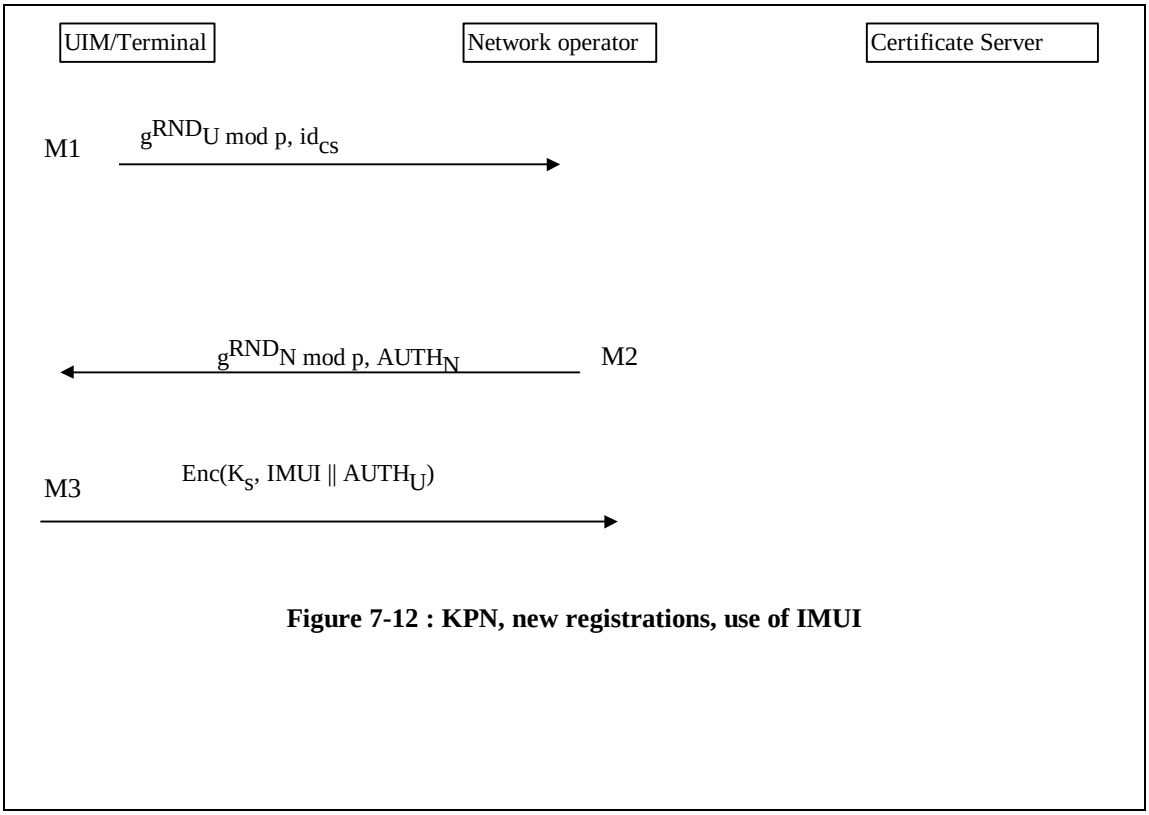


Figure 7-12 : KPN, new registrations, use of IMUI

7.3.2.4.3.4 ACHIEVED GOALS

The same goals as for "New Registration - use of CertU" (section 7.3.2.4.1) except exchange of certificates.

7.3.2.5 General Authentication Framework

The authentication framework will provide a common description for all proposed mechanisms, resulting in one set of requirements to the signalling procedures. This will facilitate the work in the other groups (e.g. network services, ...), there are still open points which require bilateral interaction. ASPeCT will follow the work on the authentication framework.

8. Requirements on the migration path (WP2.1)

8.1 Introduction

Chapter 8 presents a high-level overview of the security-related requirements for migration of Network Operators, Service Providers, Regulators, Users, Subscribers and Manufacturers from Second-Generation PLMNs (typically GSM900/DCS1800) to Third-Generation PLMNs (UMTS/FPLMTS). It addresses the work carried out as part of A2.1.2.

Note that the term "migration" is used in this document to represent both the migration of "users" from a present system to a next-generation system, and also the evolution of technology from supporting a present system to a next-generation system.

The desired security features which are anticipated to be included within the Third-Generation are taken from ETR 050901 [7], and are listed below. Refer to ETR 050901 (chapters 7 and 8) for an explanation of these security features.

- Authentication
- Data Confidentiality
- Party Anonymity
- Access Control
- Data Integrity
- Message Non-Repudiation
- Supplementary Security Features

This chapter is subdivided in different sections according to the different points of view : manufacturers (Section 8.2), network operators and service providers (Section 8.3), regulators (Section 8.4), users and subscribers (Section 8.5).

8.2 Requirements from the point of view of manufacturers

8.2.1 Interworking requirements

8.2.1.1 Open Interfaces

The different entities should be designed with open interfaces so that different suppliers can be used for the different entities.

8.2.1.2 Interworking between second and third generation systems

Requirements on third generation systems should not constrain the interworking with second generation systems. Entities satisfying the third generation requirements should still be able to execute security related functions needed to interwork with second generation PLMN's or operators.

8.2.2 Satisfaction of Users

The security mechanisms should be designed in such a way that they fulfil the users' needs and are friendly to use (e.g. not increasing too much call-setup time). The user should not be affected by any upgrade in the security mechanisms of their home-network or the networks where they can roam in during the migration/evolution towards third generation systems. This means users of second

generation systems should still be able to use their equipment (SIM or DAM and terminal) during the migration/evolution without any extra costs for the user, and without a degradation to the second generation security features, within those upgrading networks (home and visited).

8.2.3 Technically realizable mechanisms

8.2.3.1 Cost-effective

The manufacturers should make sure that the designed protocols are realizable, in a cost-effective way with current available technology.

8.2.3.2 Intellectual Property Rights

Intellectual property rights on mechanisms should not restrict manufacturers in their possibilities to choose for realizing different options.

8.2.3.3 Broad Standardisation

Standardisation of security mechanisms should be broad enough to allow the implementation of different protocols.

8.3 Requirements from the point of view of Network Operators and Service Providers

The requirements of Network Operators and Service Providers for security-related migration from Second-Generation PLMNs to Third-Generation PLMNs are summarised in the following sub-sections. Refer to the UMTS objectives stated in ETR 050101, from which the following requirements are derived.

Note that, for several of these security-related requirements, there is no clear distinction between those which are attributable purely to the Service Provider, those which are attributable to the Network Operator and those which are attributable to both. These security-related requirements have therefore been grouped together as being attributable to the Service Provider and/or the Network Operator.

Note also that these security-related requirements of Network Operators and Service Providers cannot be considered in isolation, since the security-related requirements of all other interested parties (particularly the subscriber/user) have indirect implications for the requirements of Network Operators and Service Providers, since they have to cater for the security-related requirements of the other interested parties.

Section 8.3 has been derived with reference to ETR 050103 [5], ETR 050104 [6] and ETR 050301 [20].

8.3.1 Integration of Infrastructure

In order to facilitate the transition towards UMTS deployment, it shall be possible to provide UMTS services in a system which contains components of fixed networks and Second-Generation PLMNs, forming an integrated system infrastructure. This means that the various services and (security-related) features which are to be supported within UMTS must allow for both forwards compatibility and backwards compatibility between UMTS and Second-Generation PLMNs and PSTNs.

8.3.2 Access and Interworking

The UMTS operational and service environment should allow for maximum flexibility concerning interworking with services of existing systems. The standards of UMTS and GSM are not fully compatible, and interoperability cannot be provided. Interworking is therefore a major requirement regarding the interactions between 2nd and 3rd generation systems.
Clearly UMTS must be capable to provide voice and data services to and from other existing networks. Completely new UMTS-specific service capabilities will be offered according to market needs, however migrated pre-UMTS system components should, to the edge of their limitations, be used to provide UMTS services. UMTS security features should not therefore prevent or complicate the communication interaction between 2nd and 3rd generation systems.

8.3.3 Prevention of Misuse

The UMTS standard shall ensure the prevention of misuse (or misappropriation) of data, resources or services through the impersonation of Users/Subscribers or Service Providers/Network Operators. This security requirement is expanded into further security-related requirements, including the authentication of Users/Subscribers, the authentication of Service Providers/Network Operators, the non-repudiation of messages and access control at the mobile terminal. Note that these security-related requirements are equally required by the Users/Subscribers.

8.3.4 Control of Resources and Services

The UMTS standard shall allow authorised entities to control the use of resources and services by other entities. This requirement includes the secure billing exchanges which occur between Network Operator and Service Provider, and between Service Provider and Subscriber. It also has implications for the techniques employed to combat fraud within cellular communication systems.

8.3.5 Integrity of Signalling Information

The UMTS standard shall ensure that the integrity of signalling information is guaranteed. This requirement includes that location management procedures be performed in a secure manner, and that the integrity of signalling information is unaffected by handovers. Note that security is a mobility-related procedure and, as such, should be an integrated part of call handling.

8.3.6 Confidentiality of Data

The UMTS standard shall provide for the confidentiality of network signalling information, to support location confidentiality and party anonymity over at least the air interface, in addition to the protection of any data related to party authentication or to the generation of traffic ciphering keys. Note that this security-related requirement would normally be extended to provide for the (User/Subscriber) security-related requirement of privacy of user data, at least over the air interface.

8.3.7 Access Control

Unauthorized attempts to access information and resources should be prevented. This is a result of the UMTS access control requirements of the providers. Resources in UMTS should be protected because

serious complications may arise in case of unrestricted access. From the provider perspective these resources include UMTS services provision, UMTS (sub)network use, UMTS equipment use and management of stored data. The relevant access control security mechanisms should provide for protection of entities by means of verification of access rights combined with authentication of the initiator.

8.3.8 Fraud management

The fraudulent use of network resources by means of stolen or cloned UIMs and mobile terminals, or intruder disguise as users, service providers or network operators is identified in the security requirements described above and will be confronted by the relevant security mechanisms. However, other fraud cases that are grouped under the general categories of technical or airtime fraud and roaming fraud and involve misuse of a subscription, mostly by means of call-forwarding and roaming capabilities, are required to be detected and prevented by the appropriate standardised security mechanisms within UMTS.

8.3.9 Non-repudiation of transaction

The expected vast expansion of parties and services involved in UMTS, will increase the need for the provider to be able to prove that a particular entity at a particular instance will have been involved in an action related to the network. This will allow the provider to solve possible disputes with the involved party who will try to deny involvement in the certain action. This may be needed in the billing process on disputes over charges of entities.

8.3.10 Standardisation and Intellectual Property Rights

The UMTS standard shall incorporate security features that can be incorporated in a general standardisation of UMTS functionality. The UMTS standard shall also incorporate security features that are provided via the use of security mechanisms which have minimal intellectual property rights associated with them.

8.4 Requirements from the point of view of regulators

The security-related requirements of regulators include certain aspects which are not (necessarily) included in the security-related requirements of any of the other parties associated with (public) telecommunication systems; namely the Network Operators/Service Providers, the Users/Subscribers or the Equipment Manufacturers.

The security-related requirements of regulators for migration from Second-Generation PLMNs to Third-Generation PLMNs are summarised in the following sub-sections. Such requirements may (where indicated) involve certain Supplementary Security Features as mentioned in Section .

Section 8.4 has been derived with reference to the following document :
[21] UK 3GMG, Draft Industry Advisory Document on Third Generation P.C.S., Policy Document, section 6.9: Security

8.4.1 Lawful Interception

Certain (government) authorities may require the ability to intercept communications over any public telecommunications network. Such interception may be required, for example, over the air interface, at a switch or at an exchange. Note that such interception would typically require the availability of non-ciphered user data and/or network signalling information. If (transparent) end-to-end ciphering of user data were to be supported on a public telecommunications network, a requirement to support lawful interception of user data may be the enforcement of a key escrow mechanism, allowing the recovery of user data via a "Trusted Third Party". For this reason, certain Supplementary Security Features may be required to be incorporated within UMTS, such as the provision of (User/Subscriber) end-to-end security and the provision of key escrow (via a Trusted Third Party).

8.4.2 Use of Encryption Technology

Certain (government) authorities may require the restriction of the use of or the export of particular types of cryptographic algorithm. Such restriction may be required, for example, to prevent the dissemination of cryptographic-related technology to organisations or countries whose communications are required to be recoverable for intelligence purposes.

8.4.3 Standardisation and Type-Approval

For the purposes of standardisation, a Regulator may require that security-related functionality (for example, to support ciphering or authentication) be type-approved. This may result in a requirement within UMTS that mechanisms be provided to permit network-based approval checking of mobile terminals. Similarly, a regulator may require that any network-based approval checking in general (for example, to support verification of mobile terminal type/parameters) is provided for in a secure manner.

8.4.4 Emergency Calls

A Regulator may require that UMTS provides the facility to make emergency calls from mobile terminals with the incorporation of a minimal security-related functionality. For example, the provision of ciphering of user data may be suppressed during an emergency call to reduce the probability of such a call being dropped due to a ciphering malfunction. Similarly, the security-related provisions to support location confidentiality and party anonymity (over the air interface) may be relaxed during an emergency call. Note, however, that party authentication (and message non-repudiation) would still be likely to be required in order to combat abuse by hoax callers.

8.4.5 Legal Evidence

A regulator may require that certain UMTS security features are provided to at least a guaranteed minimum performance standard. This is to ensure that data derived from network activity associated with Users/Subscribers is acceptable for the purposes of legal evidence. Typical security-related data useful for the purposes of providing legal evidence is that derived from the non-repudiation of messages.

8.5 Requirements from the point of view of users and subscribers

8.5.1 Speech Quality

The UMTS standard will allow high-quality speech connections to be established to a Second-Generation user or between a fixed user and a UMTS subscriber. This has implications for the nature of the security mechanisms employed to support the desired security functions, since the implementation of the security mechanisms must not have any detrimental effect upon speech quality, thus there will be associated constraints imposed upon the processing and signalling overheads of any cryptosystem. Similarly, there may be implications for the nature of any error correction/detection mechanisms employed within UMTS.

8.5.2 Mobility in UMTS

UMTS should allow for mobility on the terminal and the user. Specifically, UMTS should enable the same terminal to be used in all co-existing environments as well as in all different (competitive or not) UMTS networks, subject only to constrains imposed by the terminals capabilities and by agreements on service provision. Furthermore, UMTS users should be allowed to move between UMTS terminals and UMTS networks (by means of a personal User Identity Module) and for the purpose of accessing different telecommunication services. Therefore, additional security features should be offered due to the increased risks resulting from the mobility of the user and terminal.

8.5.3 Requirements for the User Identity Module

The UIM will contain data and functions needed to identify and authenticate the user when accessing UMTS services and will be managed by the user's (subscriber's) Service Provider. The UIM will play the major role in the authentication process that should be applied between users, network operators and service providers. Apart from the advanced authentication requirements, a number of additional provider requirements are related to the UIM. It will be necessary to protect against unauthorised modification of data stored in the card. It will also be necessary to detect and prevent the use of UIMs that will fall in one of the 'non-healthy' categories, i.e. stolen, cloned, expired lost and faulty UIMs.

The UMTS standard shall provide for SIM/UIM roaming between Second-Generation handsets and UMTS handsets. Support for SIM/UIM roaming is required to allow personal mobility between different terminals (both UMTS and Second-Generation). This is an essential aspect of terminal-associated functionality for supporting migration to a Third-Generation system. Users/Subscribers must be able to authenticate themselves in a seamless manner as they roam between different network systems. In this direction, there is a requirement for flexible or even multiple security and authentication mechanisms. Additionally. there may be security-related requirements associated with multi-vendor use of a SIM/UIM.

8.5.4 Requirements for the Mobile Terminal Equipment

The UMTS standard will distinguish users from their terminals. This will allow user mobility independant of terminal mobility but will also generate the need for independant treatment of user and terminal security issues. Thus, UMTS will enable the same terminal to be used in all environments and all UMTS networks, as well as the use of terminals from different terminal manufacturers in the same network.

The security requirements specifically related to the use of mobile terminal equipments are relevant to the requirements associated with the use of the UIM, i.e. to ensure the protection of unauthorised modification of data stored in the mobile terminal and to be able to detect and prevent the use of terminals belonging to one of the categories of not-type-approved, faulty, stolen or cloned terminals.

The security mechanisms that will realise these requirements should not deter the roaming/mobility capability of the UMTS terminal. Such terminal mobility is required to support the freedom of use of different manufacturers' terminals. The advent of large numbers of Network Operators and Service Providers will require the provision of mechanisms to support secure on-line roaming. Additionally, there will be need to provide mechanisms/procedures to negotiate the level of security and the nature of security-related functionality afforded to roamers as they move between different networks, however the level of security should not be affected by handovers.

Furthermore, as for the UIM's, terminals should be able to roam between second generation and third generation systems. As this should be possible without constraints to the UMTS standards, this will lead to the production of multimode terminals supporting both second and third generation systems.

8.5.5 Privacy of information

In general terms, the UMTS standard must ensure that data will not be made available or disclosed to unauthorised parties. The term 'data' in UMTS includes a whole range of information elements, however it could be grouped under two major categories, namely transferred and stored data:
- Transferred data: it includes speech, user (fax, files, etc), signalling and management (identifiers, location, charging/billing, etc) data that is transferred over the radio interface or a signalling channel respectively - however we should mention that security requirements over speech and user data on the air interface is basically a user concern. The relevant confidentiality services are required to be applied continuously during each traffic, signalling or management transmision.
- Stored data: it includes signalling and management data, data stored in the UIM and the mobile terminal as well as charging and billing data; the appropriate confidentiality services are required to to be applied continuously when data is stored permanently or temporarily.
This requirement should not deter the ability of UMTS to provide UMTS services to multiple environments and users.

8.5.6 Integrity of information

In the context of UMTS it will be required to ensure that data (as defined previously) will not be altered or destroyed on the transfer path or location of storage. This security requirement concerns the ability to detect such attempts. The appropriate security mechanisms will be applied continuously during any transfer or storage of data. The ability of UMTS to provide UMTS services to multiple environments/users should not be detered by this requirement.

8.5.7 Standardisation

The security to be provided by UMTS should be adequately standardised to provide secure interworking and and roaming. However, they should allow the maximum freedom for all parties involved in UMTS to set their own security policies.

9. Requirements for Trusted Third Parties (WP2.3)

9.1 Introduction

The ASPeCT project aims to propose and implement a solution to the management of complex trust relationships involved in the provision of end-to-end security services for UMTS. In the multi-operator Europe of the future, the management of these relationships will be a major issue and a suitable Europe wide architecture based on trusted third parties (TTP) is required.

This Section will address the work carried out as part of Activity A2.3.1. The objective of this activity is to define the security services involving TTPs that are to be developed within the project. An overview of the basic role of a TTP is provided in order to identify possible UMTS security services that require their support. The functions provided by the TTPs for these services are also identified. The component internal operations and external interfaces for each function are then described. The Section concludes with a list of the requirements for TTPs with respect to work to be carried out within ASPeCT.

The work in this Section is compiled from two documents that have been prepared as part this activity; namely, a Vodafone document describing TTP Services [1] and a Royal Holloway document describing TTP Functions [2].

9.2 The general role of TTPs

9.2.1 The requirement for TTPs

The concept of a trusted third party (TTP) with respect to communications security is defined in ISO/IEC 10181-1 [3]. It is described as a security authority or its agent, trusted by other entities with respect to security-related activities.

The widespread public use of digital signature and confidentiality services, and the need to conform with national legislation, implies the availability of TTP services to provide essential functions.

TTP services can be considered as value-added communication services available to users wishing to enhance the trust in the services used. Therefore, TTPs have to be able to offer value with regard to availability, integrity, confidentiality and assurance.

The role of TTPs includes providing assurance that:

- messages and transactions are being transferred to the right person at the right location;
- messages are received in a timely and secure manner;
- messages will be acted upon; and
- for any business dispute that arises, there are appropriate mechanisms for establishing what happened.

9.2.2 Establishing trust

The use of TTPs is dependent on the fundamental requirement that the TTP is trusted by the entities that it serves. Trust may need to be established between two entities in a communication relationship. In general, this trust can be established in three ways:

- either on a unilateral basis, that is, an entity, which needs to authenticate another entity, must know some specific authentication information about that entity;
- or, when the authentication must be mutual, the trust is established on a bilateral basis, that is each entity must know some specific authentication information about the other entity;
- or by trusting a third party, it is possible to assure one entity of the trustworthiness of another entity.

The third method involves the use of a TTP. The advantage of this method is that individual bilateral agreements between entities are not necessary.

In practice, a number of TTPs will exist at a national and international level. TTPs may have trust agreements arranged with other TTPs to form a network, which may or may not be hierarchical. The simplest TTP network would be a hierarchical tree structure with one common TTP at the root of the tree. However, in practice an organised structure will probably not exist. Instead, subsets of TTPs may group together to form an organised hierarchy within the network of TTPs. Such organised hierarchies may be formed by certain groups for specific purposes (e.g. the banking community).

A TTP network should allow any two entities, which have trust agreements with any two TTPs in the same network, to find at least one path of mutual trust between each other. Such a feature may be required by some TTP supported services.

9.2.3 The assurance of trust in a TTP

A TTP will typically be an organisation, licensed by a national authority, which provides security services to a wide range of bodies, including those within the telecommunications, finance and retail sectors. TTPs will provide a variety of security related services to users on a commercial basis. The user should be able to choose from a number of TTPs available to him, thus increasing the trust in the service he receives.

The user's trust in a third party will result from the assurance of an assumed level of security. This assurance is obtained from evidence regarding:

- trust in the organisation and its regulation
- trust in the standardised operations and practices
- the existence of a legally binding contract between the user and the third party
- the correct selection and implementation of security mechanisms
- a clear statement of the security policy
- a clear definition of roles and responsibilities
- correct interfaces and procedures with user.

Clearly, the management and operation of a TTP must be organised such that an appropriate level of security is assured.

9.2.4 Management and operation of a TTP

Within ISO/IEC, work is currently under way to produce guidelines for the use and management of TTPs. A draft technical report is available, which proposes guidelines for the use and management of TTPs. The report is divided into two parts; Part 1 is a general overview [4], whilst Part 2 deals with technical aspects [5].

The commitment of a TTP to providing a security related service should take the form of a documented Security Policy. The Policy shall identify all relevant targets, objects and threats related to the services provided. It should provide the rules, directives and procedures regarding how the specific security services and the associated security are assured. Elements of a TTP Security Policy are given in [5].

A TTP shall assume responsibility of liability within defined limits as stated in a formal contract. The contract shall also cover legal aspects regarding the operation of a TTP (such as its use in providing a key escrow service).

A regulatory body should be set up to carry out accreditation of TTPs. This should involve the approval and authorisation of the use of a TTP service according to the Security Policy within the operating environment. Note that although TTPs may be set up on a national basis within national law, they will typrically need to be trusted internationally.

A TTP should also have implemented a Quality Policy and a System of Quality in accordance with the ISO 9000 standards series.

9.2.5 Location of a TTP

TTPs can be categorised according to their communication relationships with the entities they serve. The location of the TTP will influence the services that it will be capable of fulfilling.

On-line TTP
An on-line TTP does not lie in the communication path between the two entities. However, it is requested by one or both entities in real-time to provide, or register, security-related information.

In-line TTP
An in-line TTP is positioned in the communication path between the two entities. Such an arrangement allows the TTP to offer a wide range of security services directly to the users. Since the TTP interrupts the communication path, different security domains can exist on either side of it.

Off-line TTP
An off-line TTP does not interact with the entities during the process of the given security service. However, interaction is carried out off-line in order to manage the data associated with the service.

9.3 The role of TTPs in UMTS

9.3.1 UMTS role model

The UMTS role model, as defined in ETSI/SMG/ETR 050103 [6], is shown in Figure 9-1 below. It illustrates the interrelationship between users, subscribers, service providers and network operators. The TTP(s) are shown as external entities to UMTS, which will provide security related services to entities within UMTS.

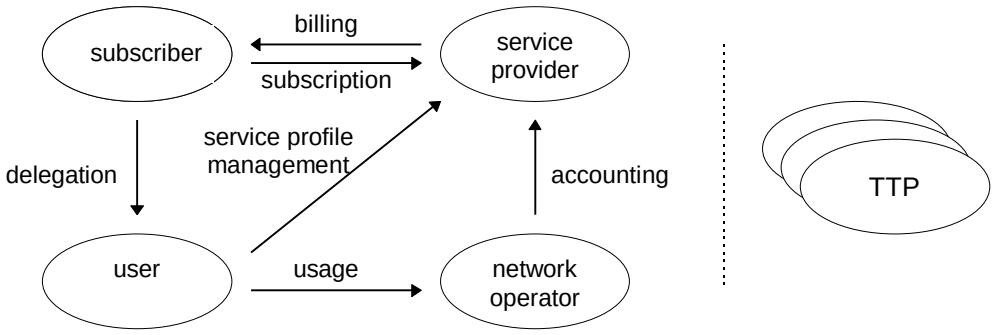


Figure 9-1 : UMTS role model

Within UMTS the entities requiring a TTP based service may be categorised as follows:

- users;
- subscribers;
- service providers;
- network operators;
- national authorities;
- regulators.

National authorities and regulators interact with all the other parties involved in enforcing the legal and fair operation of the telecommunication service.

One of the advantages of security techniques based on TTPs is that they can resolve conflicts of requirements between different entities. For example, conflicts exist between the requirements of users

for privacy, the requirements of providers for commercial security and the legitimate demand of national authorities for eavesdropping capabilities.

A TTP network, which could provide services to UMTS, would have clear benefits in the provision of end-to-end security services. With the emergence of new operators and service providers in Europe, there will be a requirement for well defined trust relationships between the entities. A Europe wide TTP network would allow this requirement to be met more easily by eliminating the need for entities to form individual trust agreements. A TTP network is also likely to resolve problems associated with inter-operator roaming, which is likely to become more common with UMTS.

9.3.2 The requirement for TTPs in UMTS

In general, security services can be provided either by the UMTS provider alone, or by the UMTS provider working with the support of a TTP. A particular service can be categorised according to its dependence on a TTP. A possible categorisation is given below:

Category	Description
A - essential	the security service can only be provided with the support of a TTP
B - optional (advantageous)	the security service need not be provided with the support of a TTP, however its use would be advantageous
C - optional (disadvantageous)	the security service need not be provided with the support of a TTP, and its use is disadvantageous
D - inapplicable	the security service cannot be provided with the support of a TTP

Table 9-1 : Categorisation of the dependence of a security service on a TTP

Initially, the motivation to use TTPs in UMTS will be brought about by their use to support security services for which the use of a TTP is essential or at least highly advantageous. A notable example would be the use of a TTP to support the provision of end-to-end encryption services to users, with key escrow functionality to support the demand for warranted interception. The introduction of TTPs to support services like this will increase the perceived benefit of other TTP services and make them more commercially viable.

Many uses of a TTP to support a security service will be optional. To decide whether its use is advantageous or disadvantageous, several factors can be considered. Some of these factors are discussed below.

Level of trust

One or more TTP may be used since they may enhance the user's trust in the security services which the TTP supports. A TTP should be able to assure an assumed level of security. In general, the TTP is viewed by the service user as trustworthy because of the commercial/contractual relationship with the user, and because of the legal and regulatory controls within which the TTP operates. The factors affecting the level of security are listed in Section 9.2.3.

Level of independence

For some security services it is preferable that certain operations are carried out by an independent and trusted third party.

Security mechanisms available

A TTP may be able to offer a 'higher level of security' by supporting more powerful cryptographic mechanisms.

Location of security service provider

A TTP may be able to act as a centralised server to provide security services. This may be a more appropriate architecture for the provision of particular security services in UMTS.

9.3.3 TTP services, functions and components.

A TTP can be used to support a range of security services in UMTS. In order to provide these services the TTP must carry out a set of functions (e.g. key generation, certification, etc.). For the purposes of the work on TTPs we distinguish between TTP services, TTP functions, and TTP components.

A TTP service is a collection of elementary functions that together perform a duty which fulfils a particular requirement in the provision of UMTS security services. A TTP function is an elementary part of a TTP service. A particular TTP function may constitute a part of several TTP services. TTP functions will themselves be composed of internal operations and external interfaces, which are termed TTP components.

TTP services, TTP functions and TTP components for UMTS are discussed, in turn, in the following Sections.

9.4 TTP services for UMTS

9.4.1 Identification of TTP services for UMTS

In this Section eight general categories of TTP service are identified and discussed. The TTP services are defined according to the roles they perform in supporting UMTS security services.

9.4.1.1 Symmetric key management

A user could request a TTP to generate and distribute symmetric keys. A TTP should offer separate generation and distribution services depending on the user's requirements. For example, the user may request a TTP to distribute a key generated by some other entity. Alternatively, a TTP may be asked to generate a key but not distribute it.

The symmetric keys could be used to provide a variety of security services including entity authentication, integrity protection, and encryption of user traffic and signalling information.

Symmetric key management using TTPs would be especially useful in providing end-to-end security services, such as end-to-end encryption, since a TTP architecture would facilitate the implementation of warranted interception (key escrow) schemes. A TTP could also offer a disaster recovery service based on key escrow. A user could retrieve lost or corrupted keys from a TTP using a similar mechanism to that used to uncover keys for warranted interception.

9.4.1.2 Asymmetric key management

A user could request a TTP to generate and distribute asymmetric keys. As with symmetric key management, a TTP should offer separate generation and distribution services depending on the user's requirements. For example, users may choose to generate their own asymmetric key pairs, and then use a TTP to help distribute their public key.

The asymmetric keys could also be used to provide a variety of security services including entity authentication, integrity protection, and encryption of user traffic and signalling information.

If a user requests the distribution of asymmetric keys, its identity is first verified. The TTP will then respond by securely distributing the private key to the user and by publishing the certified public key in a directory.

For asymmetric encryption the sender uses the certified public key of the recipient to encrypt the data and the recipient uses its corresponding private key for decryption. The sender must be sure that the recipient's public key is genuine, otherwise an intruder may have substituted the key for its own public key, thereby allowing the intruder to retrieve the data. To guarantee the authenticity of the public key certificate itself, the sender can check the current revocation list to ensure that the certificate has not been revoked.

In asymmetric integrity protection, a public key cryptosystem is used to digitally sign user data. There are many ways of doing this, but essentially the sender must obtain a valid public key certificate from the directory server, and the recipient must have in their possession the corresponding private key.

Asymmetric key management using TTPs would be especially useful in providing end-to-end security services, such as end-to-end encryption, because a TTP architecture would facilitate the implementation of warranted interception (key escrow) schemes.

9.4.1.3 Identification and authentication services

A TTP can be used as an authentication server in authentication schemes that involve third parties. The server can be located in any of the communication relationships identified in Section 9.2.5.

On-line authentication services

In symmetric authentication schemes there is a requirement for every verifier to maintain a secret symmetric key with every claimant. This may not be practical for end-to-end user authentication in UMTS. Instead, a trusted on-line authentication server could be introduced. This would share a secret key with every claimant and verifier.

Two general approaches to this scheme are outlined below:

- In the first approach the claimant encrypts or seals a message with its secret key and sends it to the verifier. Because the verifier does not share the claimant's key it must obtain it through a separate exchange with the server. This exchange must be secured using a shared key between the server and the verifier.
- In the second approach the claimant first obtains a ticket from the server which contains a secret key to be used by the claimant to authenticate itself. The communication with the server is protected using a secret key shared between the server and the claimant. After the exchange the claimant authenticates itself by encrypting or sealing a message with the secret key it received from the server and sending it to the verifier for checking.

Off-line authentication services

With asymmetric authentication the need for the authentication server to be on-line is removed. Instead the verifiers can obtain certified public keys for claimants and certificate revocation lists from an off-line server, prior to or during authentication. This information can be cached and reused to avoid having to communicate with the server each time authentication is initiated.

An off-line authentication server may just act as a directory service for certificates generated by another TTP called a Certification Authority (CA). In this case the server does not need to be trusted since the certificates are integrity protected public information. However, the entity which creates the certificates i.e. the CA must, of course, be trusted.

In-line authentication services

In-line authentication involves an authentication server positioned in the communication path between claimant and verifier. The authentication is, in effect, split into two sets of interactions. Firstly, the claimant attempts to authenticate itself to the server, which vouches the identity of the verifier. Secondly, the verifier attempts to authenticate the server, which vouches the identity of the claimant. This scheme has the advantage that the claimant and verifier may belong to different security policy domains. The server may apply different mechanisms to each domain to realise the different security policies of each domain.

9.4.1.4 Access control management

A TTP can be responsible for the management of access control information. In this scheme a TTP would issue users with privilege attribute certificates (PACs) which would enable them to access

various resources in the UMTS network[2]. The PAC will contain a list of resources that the user may wish to access and the associated privilege level that the user has been assigned.

A user will send a request for a given privilege attribute to a TTP, which will then identify and authenticate the user. If the current security policy states that the user is authorised to make the requested access, then the TTP will generate and certify the privilege attribute. The privilege attribute certificate is then distributed by the TTP by publishing it in a directory.

The veracity of privilege information can be verified on demand by anyone who obtains the TTP's public certification key. The PAC may need to be revoked if changes in access control privileges are required, or if a compromise of sensitive information is suspected.

9.4.1.5 Non-repudiation services

TTPs have an important role to play in the provision of non-repudiation services. In general, two main forms of non-repudiation service can be identified; non-repudiation of origin and non-repudiation of delivery. With non-repudiation of origin, one or more TTPs may act together with the originator of the data in evidence generation. The recipient of the data may also act with TTP(s) in evidence verification. With non-repudiation of delivery, evidence generation involves the recipient and, in some cases, one or more TTPs. In this case the originator's role is one of evidence verification which may involve support from one or more TTP(s).

ISO/IEC 10181-4 [7] describes the application of non-repudiation services in open systems. It identifies non-repudiation mechanisms which involve TTPs. These are listed below:

- Non-repudiation involving a TTP security token
- Non-repudiation using a digital signature (the TTP supports digital signatures)
- Non-repudiation using time-stamping (the TTP carried out time-stamping)
- Non-repudiation using an in-line TTP
- Non-repudiation using a notary (an in-line TTP acts as a notary)

A non-repudiation service model has been defined in ISO CD 13888-1 [8]. The model is illustrated in Figure 9-2 below, which also shows the placement of TTPs, which will interact with the originator or recipient.

[2] This is one of the two most common approaches to providing access control services. The other approach is where the resource checks the identity of the user against a privilege list held at the resource. The best approach depends on particular circumstances. However, the use of a TTP only seems applicable to the privilege attribute approach.

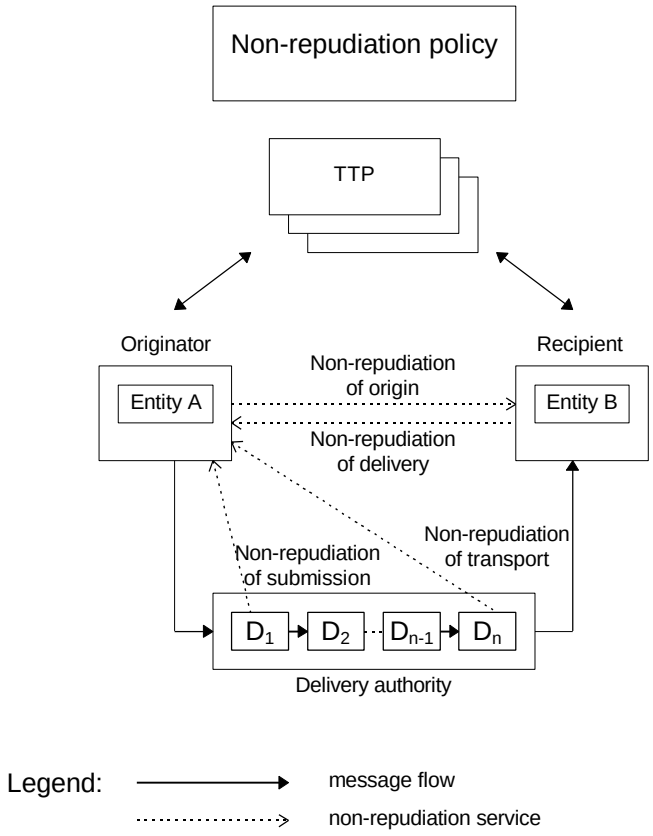


Figure 9-2 : The use of TTPs in supporting non-repudiation services

In general, the non-repudiation service involves the generation, verification and recording of evidence, and the subsequent retrieval and re-verification of this evidence in order to resolve disputes. Its purpose is to provide evidence about a particular event or action. There are four distinct phases to non-repudiation services:

- evidence generation;
- evidence transfer storage and retrieval;
- evidence verification; and
- dispute resolution.

TTPs may be involved in all of these phases as described in ISO CD 13888-1 [8]. However, their involvement depends on the mechanism used. For example, the use of asymmetric cryptographic techniques to generate and verify evidence may or may not require the involvement of a TTP, whereas the use of symmetric cryptographic techniques in evidence generation and verification always requires a TTP to generate and verify certificates. Each phase is dealt with in turn below.

Evidence generation follows an invocation of the service by a non-repudiation service requester. Relevant evidence may include the identities of the entities involved, the communicated data, and the date and time. Additional information that may be required could comprise mode of transfer, location of entities or creator of data[3]. The evidence may be generated by the evidence subject, perhaps in conjunction with a TTP, or by a TTP alone.

Evidence transfer, storage and retrieval involves the transfer of evidence either between entities or to or from storage. In this phase the TTP may carry out the following functions: evidence recording, certification, directory, revocation, time-stamping, audit, and delivery authority.

[3] A more complete list of information that can be used as evidence is given in [8]

The purpose of evidence verification is to provide the evidence user with the confidence that the supplied evidence will be adequate in the event of a dispute arising. The evidence verifier may be the evidence user, or a TTP trusted by the evidence user.

In dispute resolution an adjudicator is responsible for collecting evidence from the disputing parties and/or a TTP in order to make a decision which will resolve the dispute.

9.4.1.6 UMTS user accreditation

A TTP can support UMTS user accreditation services. In this role a TTP could be responsible for correctly identifying UMTS users on their initial registration with a service provider. This may involve the user presenting its credentials to the TTP, which will carry out user identification and credit checking functions.

9.4.1.7 Pre-personalisation of UIMs

In this role a TTP will be required to securely generate and load specific user information onto a User Identity Module (UIM). Some of this information may be confidential, so a TTP must take measures to protect it. Once in the UIM, the generated information should be physically and logically secure. However, the information may need to be transported to the network operator or service provider, or archived in the TTP. In both of these cases the information must be protected. In this role a TTP would commonly be referred to as a Pre-personalisation Centre (PPC).

9.4.1.8 Fraud control services

A TTP could act as an entity which can alert various parties to possible fraudulent behaviour based on real-time event sequence monitoring. The advantage of using an external TTP for this purpose, is that the TTP can independently monitor *all* entities in the system to identify fraud scenarios that occur between different entity domains.

9.4.2 Fulfilling UMTS security requirements using TTPs

The TTP services outlined above can be used to fulfil the UMTS security requirements identified in ETSI ETR 050901 [9].

In the tables below the UMTS security requirements are listed and matched against the relevant TTP supported security service(s). The requirements are grouped together in accordance with the classification used in ETSI ETR 050901, where they are identified. Note that ETSI ETR 050901 is still in development. As a result, the requirements listed in the following Sections are liable to vary with those identified in the most recent version of ETR 050901.

In general, the TTP service(s) will only provide a supporting role in fulfilling the requirement. The dependence on a trusted third party will vary and the categorisation outlined in Section 9.3.2 is used to indicate the level of dependence in each case. The categories are denoted in the tables as follows:

- Essential ★
- Optional (advantageous) ✓✓
- Optional (disadvantageous) ✓
- Inapplicable <blank>

9.4.2.1 Customer security Requirements

Security Requirements \ TTP Services	Symmetric key management	Asymmetric key management	Identification and authentication	Access control management	Non-repudiation	UMTS user accreditation	Pre-personalisation of UIMs	Fraud control
User identity module								
Access to the UIM restricted to the user or a party authorised by the user only				✓				
Privacy of certain data stored on the UIM should be ensured							✓	
Protect against unauthorised modification of certain data stored on the UIM								
Access to telecommunication services								
Prevent an intruder masquerading as a user and accessing telecommunication services			✓			✓		
Prevent an intruder masquerading as a service provider or network operator			✓			✓		
Provision of telecommunication services								
Privacy of user traffic, location and identity over the air interface	✓	✓						
Protect against unauthorised modification of user data or signalling over the air interface	✓	✓						
Protect against intruders taking over a service already provided to a user			✓			✓		
Access to service profiles								
Privacy of service profile data				✓				
Protect against unauthorised modification of service profile data				✓				
Ensure users, subscribers and service providers cannot subsequently deny having accessed a service profile					✓			
Support for supplementary security services								
End-to-end privacy, integrity and non-repudiation services	✓✓	✓✓						
End-to-end authentication			✓✓					
Data protection								
Privacy of user activity monitoring								
Ensure that the transmission of user signalling or management data by service providers to network operators cannot subsequently be denied					✓			
Protect against unauthorised modification of user signalling or management data which is stored or processed by a service provider or network operator, or sent between service providers and network operators	✓	✓						
Privacy of signalling or management data which is stored or processed by a service provider or network operator, or sent between service providers and network operators	✓	✓						
Monitoring and registration of all attempted interceptions								

Table 9-2 : Meeting customer security requirements using TTPs

9.4.2.2 Provider Security Requirements

Security Requirements \ TTP Services	Symmetric key management	Asymmetric key management	Identification and authentication	Access control management	Non-repudiation	UMTS user accreditation	Pre-personalisation of UIMs	Fraud control
User identity module								
Deter, detect and prevent the use of stolen, cloned, expired and lost UIMs			✓			✓		
Prevent the use of faulty UIMs			✓			✓		
User access to telecommunication services								
Prevent an intruder masquerading as a user and accessing telecommunication services			✓			✓		
Ensure that a user cannot subsequently deny having used a telecommunication service					✓✓			
Prevent an intruder masquerading as a service provider or network operator			✓			✓		
To be able to restrict the services available to a specific user								
Provision of telecommunication services								
Privacy of signalling information transmitted over the air interface	✓	✓						
Protect against unauthorised modification of signalling information over the air interface	✓	✓		✓				
Protect against intruders restricting the availability of services by logical means	✓	✓						
Detect and prevent the fraudulent use of telecommunication services								✓
Service profiles								
Ensure that users and subscribers cannot subsequently deny having accessed a service profile					✓			
Protect against unauthorised modification to service profile data	✓	✓						
Protection of provider resources								
Ensure privacy of data stored by service provider or network operator	✓	✓		✓				
Prevent unauthorised modification to data by service provider or network operator				✓				
Protection of communications within or between provider domains								
Prevent an intruder from masquerading as one network entity to another			✓			✓		
Ensure the privacy of signalling and management data sent within or between UMTS provider domains	✓	✓						
To detect unauthorised modification to signalling and management data sent within or between UMTS provider domains	✓	✓						
To ensure the origin of signalling and management data sent within or between UMTS provider domains cannot subsequently be denied by the originator					✓			
To ensure the delivery of signalling and management data sent within or between UMTS provider domains cannot subsequently be denied by the recipient					✓			

Table 9-3 Meeting provider security requirements using TTPs

9.4.2.3 Miscellaneous security requirements

Security Requirements \ TTP Services	Symmetric key management	Asymmetric key management	Identification and authentication	Access control management	Non-repudiation	UMTS user accreditation	Pre-personalisation of UIMs	Fraud control
Mobile terminal equipment								
Access to the mobile terminal restricted to the user or a party authorised by the user only				✓				
Privacy of certain data stored on the mobile terminal should be ensured							✓	
Protect against unauthorised modification of certain data stored on the mobile terminal								
Detect, deter and prevent the use of mobile terminal equipment which is not type approved, cloned or stolen			✓			✓		
Prevent the use of faulty mobile terminal equipment			✓			✓		
Charging and billing and accounting								
Privacy of charging, billing and accounting information	✓	✓						
Protect against unauthorised modification of charging, billing and accounting information	✓	✓		✓				
Ensure that entities having incurred charges cannot subsequently deny this					✓✓			
Alternative payment methods								
Subscriber and user information should be protected against disclosure to unauthorised parties	✓	✓		✓				
It should be an open system i.e. enable participation of several financial institutes, service providers, retailers and users without the need for a relation of trust between them. Trust is needed between service provider and subscriber, and between service provider and retailer			✓✓			✓		
Anonymous payment to the service provider may be necessary	✓	✓						
The subscriber may require some insurance against theft or loss (e.g. a maximum payout may be set)								
Adaptive terminals								
Ensuring the integrity, authenticity, and possible confidentality of software and any other control data	✓	✓						
Legal interception								
Ensure that facilities exist for legal interception by goverments of all communications both within their own domain and possibly in other domains by agreement with the relevant goverment.	★	★						
PMR								
Group authentication and confidentiality	✓✓	✓✓						
Satellite								
Satellite specific requirements relating to: Access control, Denial of service, Handover, Limitation of service domain, Operation of mechanisms, Other issues								
Extention to other services, including: Location privacy, Legal interception, Impersonation, Key distribution								

Table 9-4 : Meeting miscellaneous security requirements using TTPs

9.5 TTP functions for UMTS

9.5.1 Functions required to support TTP services

The TTP services outlined in Section 9.4 require TTPs to perform a core set of support functions. The set of core functions are identified in Table 9-5 below and matched against the TTP services they support.

TTP Functions \ TTP Services	Symmetric key management	Asymmetric key management	Identification and authentication	Access control management	Non-repudiation	UMTS user accreditation	Pre-personalisation of UIMs	Fraud control
Key generation	✗	✗	✗				✗	
Key distribution (on-line)	✗	✗	✗				✗	
Key archiving	✗	✗	✗				✗	
Key escrow functionality	✗	✗						
Privilege attribute generation				✗				
Privilege attribute distribution				✗				
Privilege attribute archiving				✗				
Certification functions		✗	✗	✗	✗			
Directory functions		✗	✗	✗	✗			
Revocation functions		✗	✗	✗	✗			
Alert management functions		✗	✗	✗	✗			
Claimant and verifier functionality			✗		✗			
Evidence generation					✗			
Evidence recording					✗			
Evidence verification					✗			
Time-stamping functions					✗			
Audit functions					✗			
Delivery authority					✗			
User identification						✗		
Credit checking						✗		
Fraud detection and management functions								✗

Table 9-5 : Matching TTP functions against services

9.5.2 Description of TTP functions

The functions discussed in this section correspond to the set of core functions identified above.

1. **Key generation**
   This may involve the generation of symmetric or asymmetric keys for a variety of entities in the UMTS environment.

2. **Key distribution (on-line)**
   A TTP may be required to distribute keys (generated either by some external entity or by the TTP itself) for subsequent use by UMTS entities, e.g. UMTS users, service providers, network operators, or other TTPs.

3. **Key archiving**
   A TTP may be responsible for archiving cryptographic keys that it has generated or distributed. In addition a user may request a TTP to archive keys that he has generated himself.

4. **Key escrow functionality**
   In a warranted interception scheme a TTP might combine the roles of a key distribution agent for its clients, and as a supplier of user keys (under warrant) to an interception agency. In addition key escrow may provide the user with a form of disaster recovery where lost or corrupted cryptographic keys can be retrieved.

5. **Privilege attribute generation**
   A TTP may have to generate privilege attributes as part of an access control management service.

6. **Privilege attribute distribution**
   As with public keys, privilege attributes should be integrity protected by publishing them in the form of a certificate. Certification could be carried out by a TTP which would generate a Privilege Attribute Certificate.

7. **Privilege attribute archiving**
   A TTP may be responsible for archiving privilege attributes that it has generated or distributed.

8. **Certification functions**
   A TTP can be used to protect the integrity of published information. Typical information that may be protected might include public keys for asymmetric cryptography and privilege attributes for access control management. The information is published in the form of a certificate which contains the information itself together with other essential information such as the ID of the owner, the certificate's expiry date and the ID of the TTP that issued it. To provide integrity protection the certificate has a digital signature appended, which is computed on the contents of the certificate using the TTP's private certification key.

   Provided that all participants have an authentic copy of the TTP's public certification key, they can check the integrity of the issued certificates at any time by checking the validity of the signature on the certificate.

9. **Directory functions**
   Directories will contain information regarding a variety of different types of entity, much of it not security-related. For security purposes, directories can store signed certificates, for example to enable one entity to obtain a verified copy of another entity's public key. The directory will typically also contain a copy of a current *Certificate Revocation List (CRL)* which will indicate which current certificates have been revoked (i.e. are no longer valid). The public key certificates and the CRL held on the directory server must be updated regularly.

10. **Revocation functions**
    Certificates will generally contain an expiry date after which they no longer guarantee the authenticity of the certified information. This feature is provided as a safeguard against cryptanalysis. Before the expiry date the certificate must be revoked by the TTP and a new one generated and issued. The TTP may also have to revoke certificates before the expiry date as a result of alarms created by alert management functions. This means that a list of revoked certificates must be published, so that users can check the validity of certificates before using them. It must be the responsibility of users to check current revocation lists in order to guarantee the validity of a certificate. The certificate revocation list (CRL) will be maintained by the TTP.

It should have a time-stamp so that any entity requiring it can be sure that it is up-to-date. (The CRL will also have a signature itself so that its integrity can be checked)

11. **Alert management functions**
There must be an adequate mechanism for informing the TTP of security alerts, such as compromises in private keys, changes in security policy, etc. The mechanism must keep the revocation list and the current list of certificates up-to-date.

12. **Claimant and verifier functionality**
In its role as an in-line authentication server, a TTP may act as a claimant or a verifier in the authentication scheme. It must therefore be able to provide the functionality for taking on these roles.

13. **Evidence generation**
In the context of non-repudiation, a TTP may cooperate with an originator or recipient to generate evidence in the form of digital signatures.

14. **Evidence recording**
A TTP may also have to record evidence in a non-repudiation role so that it can be retrieved by an evidence user or adjudicator.

15. **Evidence verification**
A TTP may act as an on-line authority which is trusted by the evidence user to verify evidence by checking digital signatures.

16. **Time stamping functions**
A TTP may be required to certify that it has affixed a signature to a data item at a particular time. This is an important function in non-repudiation services.

17. **Audit functions**
A TTP may act as an on-line authority which would carry out audit functions in line with ISO 7498-4 [10]. In this role it would monitor the transfer of data and provide evidence about what was monitored.

18. **Delivery authority functions**
A TTP may act as an on-line authority that interacts with the intended recipient of data and releases the data if, and only if, correct proof of delivery is provided by the recipient.

19. **User identification functions**
A TTP can be responsible for identifying users prior to offering them any kind of service.

20. **Credit checking**
A TTP may be involved in a credit checking role as part of the accreditation of new UMTS users.

21. **Fraud detection and management functions**
A TTP providing this function will act as an on-line authority which will monitor fraud related data. The TTP will be able to detect fraudulent behaviour and report this to relevant authorities and/or take appropriate action to stop the fraud. The analysis of fraud may use adaptive techniques so the TTP may have to keep adaptive profiles of 'normal' behaviour.

9.6 TTP components for UMTS

9.6.1 Internal operations

The following types of internal operation will be required to support the complete list of functions identified in Section 9.5.1. Note that this list is not intended to imply that all TTPs should provide all these operations. Rather the selection of functions to be supported by a TTP will determine what types of operation it must be capable of performing.

I1. *Cryptographic key generation*, including:

- the generation of secret keys for symmetric algorithms, and
- the generation of public/private key pairs for asymmetric algorithms.

I2. Note that there may be a need for the use of specially 'certified' software for the key generation process, since there is typically no way of checking whether keys have really been generated in the way claimed by the authors of the software.

I3. *Cryptographic computation*, including:

- encipherment/decipherment for symmetric algorithms,
- encipherment/decipherment for asymmetric algorithms,
- signature generation/verification,
- MAC and other (symmetric) cryptographic check function calculation,
- hash-function computation.

I4. *Cryptographic key storage*, including:

- private signature key(s) for the TTP,
- secret keys shared with other TTPs (e.g. used for establishing shared session keys between users).

I5. *User information storage*, including:

- key escrow information for a user (possibly including secret or private keys belonging to a user),
- account/credit information for a user,
- policy information for a user,
- access control information for a user.

I6. *Event information storage*, e.g. as used to support audit, fraud detection and alert management functions.

I7. *Access control information generation*, e.g. constructing user privilege attributes (perhaps using the user information storage operation).

I8. *Certificate generation*, i.e. assembling information to be certified into an appropriate form (e.g. as specified in ITU-T X.509), and then adding the TTP's signature.

I9. *Event analysis*, i.e. examining events as notified to the TTP and generating alarms if potentially fraudulent behaviour is detected.

I10. *Time-stamp generation*, which will require the TTP to maintain an accurate real-time clock.

9.6.2 External Interfaces

The following types of external interface will be required to support the complete list of functions identified in Section 9.5.1. Note that this list is not intended to imply that all TTPs should provide all these types of interface. Rather the selection of functions to be supported by a TTP will determine what types of interface it must provide.

I1. *User interface*, i.e. the means by which a UMTS user interacts with a TTP. In a simple implementation of a TTP this might actually mean a direct user interface between the user sitting at a keyboard and screen and the TTP software. In a practical implementation this probably means some kind of Applications Program Interface (API), by means of which application software, perhaps providing an interface with the user as one of its functions, can interact with the TTP. It might be possible to extend this API to also provide the Security manager interface (I5 below).

I2. *Directory interface*, as used to send or receive certificates and CRLs for storage and distribution.

I3. *Network interface*, e.g. as used to support authentication and key distribution, or to signal alarms to a Network Operator or Service Provider. The TTP may either be directly accessible to its clients via some kind of network ('on-line'), or it may only be responsible for performing functions off-line (as in the traditional role of a Certification Authority). Note that, to support end-to-end encryption with an escrow capability will typically require a TTP to be on-line. The use of existing network management standards (e.g. SNMP or CMIP) might be of assistance in defining this interface.

I4. *Interception authority interface*, e.g. as used to receive requests for escrowed keys and to supply key escrow information as authorised. The information supplied by a TTP to an interception agency (which might, for example, be the police) could take various forms, including:

- a secret or private key (or keys) required to decipher an enciphered message (or messages),
- the deciphered version of an enciphered message supplied by the interception agency.

I5. *Security manager interface*, by which the TTP can be given instructions as to how to operate and when to perform certain functions.

I6. Financial institution interface, which can be used by a TTP to obtain information relating to a user's credit status from a financial institution.

As far as the project is concerned, these interfaces will need to precisely defined (or taken from the outside literature) before implementation. Any such interface design will need to be done in close co-operation with other WPs (notably WPs 2.4 and 2.5). The definition of an interception authority interface is something of potential long term value, and standardisation of this interface might be a possibility.

9.6.3 Matching functions to operations/interfaces

In the list below we consider each of the 21 TTP functions, and indicate which internal operations and external interfaces might be required to support them. It is important to note that inclusion of an operation or interface under a particular function does not mean that it *must* be used to support this function.

9.6.3.1 Key generation

TTPs may be required to generate keys for a variety of entities, including:

- individual users,
- Network Operators,
- Service Providers,
- secret keys for the TTP's own use,
- secret keys to be shared with other TTPs, and
- private/public key pairs for the TTP's own use.

These keys may either be asymmetric (public/private) key pairs or symmetric secret keys. When a TTP is required to generate asymmetric key pairs, care will need to be taken in the specification and design of the key generation software/hardware.

It will potentially involve the following internal operations and external interfaces:

- **O1**. Cryptographic key generation,
- **O3**. Cryptographic key storage,
- **O4**. User information storage,

- **I1**. User interface,
- **I5**. Security manager interface.

9.6.3.2 Key distribution (on-line)

Keys for subsequent on-line distribution may be obtained by the TTP in one of four ways:

- via the User interface (e.g. for receiving user-generated keys),
- via the Network interface (e.g. for receiving keys generated by network entities),
- by using the TTP's own Cryptographic key generation facility, or
- from the TTP's Cryptographic key storage facility.

Note that if keys are sent to the TTP via the Network interface then they may need to be cryptographically protected, which might involve use of other TTP internal operations (e.g. Cryptographic computation and Cryptographic key storage).

The (on-line) key distribution process will clearly involve use of the Network interface. Given that this will involve keys being sent across potentially insecure network components, the Cryptographic computation and Cryptographic key storage facilities will almost certainly be required to protect the keys being distributed.

The Security manager interface may be required to enable a supervisory entity to authorise the key distribution process. The User information storage capability may be required to enable the TTP to work out what types of key to generate for a user, and how to charge the user for services rendered.

Thus this TTP function will potentially involve the following internal operations and external interfaces:

- **O1**. Cryptographic key generation,
- **O2**. Cryptographic computation,
- **O3**. Cryptographic key storage,
- **O4**. User information storage,
- **I1**. User interface,
- **I3**. Network interface,
- **I5**. Security manager interface.

9.6.3.3 Key archiving

The key archiving process will involve three main stages:

- the initiation of the archiving process for a key or keys,
- obtaining the keys to archive, and
- the archiving process itself.

The archiving process may be triggered in a number of ways, for example by:

- the receipt of an archiving request from a UMTS user via the User interface,
- the receipt of an archiving request from a UMTS entity via the Network interface,
- an instruction from a supervisory entity via the Security manager interface.

Keys for archiving may be obtained from:

- a UMTS user or other UMTS entity either via the User interface or the Network interface,

- the TTP's User information storage facility (where user keys are stored), or
- the TTP's own Cryptographic key storage (in the case where the TTP's own keys are to be archived).

The archiving process itself may involve the following processes:

- adding a time-stamp to the key information to be archived, generated using the Time-stamp generation facility,
- operating cryptographically on the key information to be archived, e.g. signing it, using the Cryptographic computation and Cryptographic key storage facilities,
- adding other user information to the key information to be archived, obtained from the User information storage facility, and
- storing the archive information, as part of the Event information storage facility.

Thus this TTP function will potentially involve the following internal operations and external interfaces:

- **O2**. Cryptographic computation,
- **O3**. Cryptographic key storage,
- **O4**. User information storage,
- **O5**. Event information storage,
- **O9**. Time-stamp generation,
- **I1**. User interface,
- **I3**. Network interface,
- **I5**. Security manager interface.

9.6.3.4 Key escrow functionality

The 'obvious' approach to supporting end-to-end encryption is to make use of precisely the same certification structure as is commonly proposed to support digital signatures (see Section 9.6.3.8 below). Instead of requiring the TTP to sign the user public verification key, the TTP is asked to sign the user's public encryption key (for some asymmetric encryption scheme), acting as a type of *Certification Authority*.

However, in the case of end-to-end encryption, there is a potential need to support key escrow. That is, in many countries, government agencies, or other legally supported bodies, may need the means to decrypt some users' traffic (both incoming and outgoing). Typically, a TTP will be required to supply the means to decrypt a particular user's messages to an interception agency, given that the interception agency has the appropriate legal authority (e.g. a warrant); moreover this will typically have to happen in such a way that the user is not made aware of the interception agency's interest.

A conventional Certification Authority does not have the means to supply all the information likely to be needed by such an interception agency. As a result a number of schemes to support warranted interception have been devised, many of them making use of some kind of TTP. Such a TTP might combine the roles of a key distribution agent for its clients, and as a supplier of user keys (under warrant) to an interception agency).

It is important to note that, unlike the Certification Authority used to support digital signatures, TTPs used to support end-to-end encryption may need to be 'on-line' in certain cases.

It will potentially involve the following internal operations and external interfaces:

- **O3**. Cryptographic key storage,
- **O4**. User information storage,

- **I3**. Network interface,
- **I4**. Interception authority interface,
- **I5**. Security manager interface.

9.6.3.5 Privilege attribute generation

The Privilege attribute generation process has two main stages:

- the authorisation step,
- the generation step.

The authorisation step may involve the Security manager and/or the User interfaces. The generation step will require the collection of relevant information about the user (from the User information storage function), the possible addition of a current time-stamp (from the Time-stamp generation facility), and the actual generation of the access control information itself (using the Access control information generation facility). The generated privilege attribute may be stored in the User information storage of the TTP.

Thus this TTP function will potentially involve the following internal operations and external interfaces:

- **O4**. User information storage,
- **O6**. Access control information generation,
- **O9**. Time-stamp generation,
- **I1**. User interface,
- **I5**. Security manager interface.

9.6.3.6 Privilege attribute distribution

Privilege attribute distribution will involve four main steps:

- authorisation of the distribution process,
- obtaining the privilege attribute,
- generating the privilege attribute certificate (PAC), and
- passing the PAC to the relevant entity for distribution.

The authorisation process will be carried out using either the User interface or the Security manager interface. The privilege attribute to be distributed may be obtained from the TTP's own User information storage, or may be passed to the TTP via the User interface, Security manager interface or the Network interface. Generating the PAC will require use of the Certificate generation function in conjunction with the Cryptographic key storage function, and the PAC can then be passed for further distribution via one or more of the Directory interface, the Network interface and the User interface.

Thus this TTP function will potentially involve the following internal operations and external interfaces:

- **O3**. Cryptographic key storage
- **O4**. User information storage,
- **O7**. Certificate generation,
- **I1**. User interface,
- **I2**. Directory interface,
- **I3**. Network interface,

- **I5**. Security manager interface.

9.6.3.7 Privilege attribute archiving

Like, the key archiving process, the privilege attribute archiving process will involve three main stages:

- the initiation of the archiving process for a privilege attribute or attributes,
- obtaining the privilege attributes to archive, and
- the archiving process itself.

The archiving process may be triggered in a number of ways, for example by:

- the receipt of an archiving request from a UMTS user via the User interface,
- the receipt of an archiving request from a UMTS entity via the Network interface,
- an instruction from a supervisory entity via the Security manager interface.

Privilege attributes for archiving may be obtained from:

- a UMTS user or other UMTS entity either via the User interface or the Network interface, or
- the TTP's User information storage facility.

The archiving process itself may involve the following processes:

- adding a time-stamp to the privilege attribute to be archived, generated using the Time-stamp generation facility,
- operating cryptographically on the privilege attribute to be archived, e.g. signing it, using the Cryptographic computation and Cryptographic key storage facilities,
- adding other user information to the privilege attribute to be archived, obtained from the User information storage facility, and
- storing the archive information, as part of the Event information storage facility.

Thus this TTP function will potentially involve the following internal operations and external interfaces:

- **O2**. Cryptographic computation
- **O3**. Cryptographic key storage
- **O4**. User information storage,
- **O5**. Event information storage
- **O9**. Time-stamp generation
- **I1**. User interface,
- **I3**. Network interface,
- **I5**. Security manager interface.

9.6.3.8 Certification functions

An 'obvious' TTP function which is necessary to support the use of digital signatures is *Public Key Certificate Generation*. A widely accepted format for such certificates is specified in the 1993 version of ITU-T Recommendation X.509 (known as X.509 version 2). A revised version of this specification (known as X.509 version 3) currently has draft amendment (DAM) status. The TTP responsible for

generating certificates is often referred to as a *Certification Authority (CA).* This CA will need to have its own digital signature key pair.

The generation of a user public key certificate requires the following main steps:

- The user identity needs to be verified.
- The CA needs to be supplied with a public signature verification key for the user. This may be part of a key pair generated in advance by the user, in which case the user will pass the CA its public key, or it may be part of a key pair generated by the CA for the user, in which case the CA also needs to pass both parts of the key pair to the user.
- The CA will need to pass its public signature verification key to the user.
- The CA can then generate the *user certificate*, which will be a string containing the user name, user verification key, certificate expiry date, and other information, all signed using the CA's private signature key.
- The user certificate will then be passed to the user, as well as possibly being distributed by other means (e.g. via an X.500 directory).

Thus the certification function will potentially involve the following internal operations and external interfaces:

- **O2**. Cryptographic computation,
- **O3**. Cryptographic key storage,
- **O4**. User information storage,
- **O7**. Certificate generation,
- **O9**. Time-stamp generation,
- **I1**. User interface,
- **I2**. Directory interface,
- **I5**. Security manager interface.

9.6.3.9 Directory functions

The performance of the directory function will involve three main steps:

- receipt of certificates for storage in the directory,
- storage of certificates,
- providing entity information (including certificates) on request.

Certificates will be received either directly from the TTP generating them (using the Directory interface), or via the User interface. Certificates will be stored using the User information storage facility. Certificates will be distributed via the Network interface. The Security manager interface can be used to change the configuration of the directory, e.g. to change access rights to parts of the directory database.

Thus this TTP function will potentially involve the following internal operations and external interfaces:

- **O4**. User information storage,
- **I1**. User interface,
- **I2**. Directory interface,
- **I3**. Network interface,
- **I5**. Security manager interface.

9.6.3.10 Revocation functions

When acting as a CA, a TTP will be responsible for regularly generating and distributing *Certificate Revocation Lists (CRLs)*, containing a list of certificates which, although not expired, are no longer valid for some reason (for example, the user private key may have been compromised).

Deciding whether or not to revoke a certificate may require invoking the User information storage facility. The CRL will be a signed structure containing a time-stamp, and thus generating a CRL will require use of the Cryptographic computation, Cryptographic key storage and Time-stamp generation facilities. The CRL may be provided to a directory using the Directory interface, or may be sent directly to UMTS entities using the Network interface. The entire revocation function can be managed using the Security manager interface.

It will potentially involve the following internal operations and external interfaces:

- **O2**. Cryptographic computation,
- **O3**. Cryptographic key storage,
- **O4**. User information storage,
- **O9**. Time-stamp generation,
- **I2**. Directory interface,
- **I3**. Network interface,
- **I5**. Security manager interface.

9.6.3.11 Alert management functions

Alert management involves four main processes:

- receiving alert information,
- processing and storing alert information,
- triggering management actions, and
- providing input to the Revocation function.

Alerts will be received via the Network interface. Alerts will be stored using the Event information storage function. Management actions will be triggered via the Security manager interface. Input to the revocation function will be handled internally to the TTP.

Thus this TTP function will potentially involve the following internal operations and external interfaces:

- **O5**. Event information storage,
- **I3**. Network interface,
- **I5**. Security manager interface.

9.6.3.12 Claimant and verifier functionality

Whilst acting as a claimant or verifier, as part of in-line authentication service, the TTP will need to engage in authentication exchanges with UMTS users and/or UMTS network entities. This will involve the Cryptographic computation and Cryptographic key storage facilities, and, if keys are to be generated as part of this process, then it may also involve the Cryptographic key generation facility. Some authentication protocols involve sending and checking time-stamps, for which the Time-stamp generation facility will be essential. Messages will be exchanged with UMTS entities via the Network interface. The Security manager interface may be required to configure the operation of this function.

Thus this TTP function will potentially involve the following internal operations and external interfaces:

- **O1**. Cryptographic key generation,
- **O2**. Cryptographic computation,
- **O3**. Cryptographic key storage,
- **O9**. Time-stamp generation,
- **I3**. Network interface,
- **I5**. Security manager interface.

9.6.3.13 Evidence generation

Evidence generation will involve the following main processes:

- initiation of evidence generation,
- receipt of information for which evidence is to be generated (e.g. a message), and
- generation of evidence (in the form of a digital signature).

The evidence generation step will normally be initiated via the Network interface, and the information for which evidence is to be generated is also likely to be received via the Network interface. Evidence generation will potentially require the addition of a time-stamp to the data (using the Time-stamp generation facility), followed by use of the Cryptographic computation and Cryptographic key storage facilities. The Security manager interface may be required to manage the process.

Thus this TTP function will potentially involve the following internal operations and external interfaces:

- **O2**. Cryptographic computation,
- **O3**. Cryptographic key storage,
- **O9**. Time-stamp generation,
- **I3**. Network interface,
- **I5**. Security manager interface.

9.6.3.14 Evidence recording

Evidence recording can be divided into two main steps:

- receipt of the evidence to be recorded, and
- the actual recording of the evidence itself.

The evidence to be recorded will typically be received via the Network interface, although the User interface and Security manager interface may also be involved. The recording process may involve the addition of a time-stamp (using the Time-stamp generation facility) and certain user-related information (using the User information storage facility), and the evidential information will actually be stored using the Event information storage facility.

Thus this TTP function will potentially involve the following internal operations and external interfaces:

- **O4**. User information storage,
- **O5**. Event information storage,
- **O9**. Time-stamp generation,
- **I1**. User interface,
- **I3**. Network interface,

- **I5**. Security manager interface.

9.6.3.15 Evidence verification

Evidence verification will involve the following three main steps:

- receipt of evidence to be verified,
- the verification process itself, and
- passing information about the success or failure of the verification process to interested entities.

The evidence to be verified will typically be received via the Network interface, although the User and Security manager interfaces might also be involved. The verification process will involve the verification of signatures and/or check-values, requiring use of the Cryptographic computation and Cryptographic key storage facilities. Information regarding the success or failure of the verification will again typically be passed using the Network interface.

Thus this TTP function will potentially involve the following internal operations and external interfaces:

- **O2**. Cryptographic computation,
- **O3**. Cryptographic key storage,
- **I1**. User interface,
- **I3**. Network interface,
- **I5**. Security manager interface.

9.6.3.16 Time-stamping functions

The time-stamping function will involve three main stages:

- the receipt of information to be time-stamped,
- the addition of a time-stamp and a signature, and
- the transmission of the signed, time-stamped information back to the requester.

The information to be time-stamped will typically be received via the Network interface, although the User and Security manager interfaces might also be used (the same will apply to the transmission of 'time-stamped' data). The time-stamping process itself will obviously involve use of the Time-stamp generation facility, as well as the Cryptographic computation and Cryptographic key storage facilities to generate the signature. Finally, the Certificate generation facility will be required to generate the certificate data structure for the time stamp.

Thus this TTP function will potentially involve the following internal operations and external interfaces:

- **O2**. Cryptographic computation,
- **O3**. Cryptographic key storage,
- **O7**. Certificate generation,
- **O9**. Time-stamp generation,
- **I1**. User interface,
- **I3**. Network interface,
- **I5**. Security manager interface.

9.6.3.17 Audit functions

The audit function will have three main stages:

- the reception of audit information,
- the storage of audit information, and
- the supply of audit information when required.

The input and output of audit information will use the Network interface. The storage of audit information will use the Event information storage facility. The Time-stamp generation facility may be used to add time-stamps to audit information prior to storage. Finally, the entire audit function will be controlled by the Security manager interface.

Thus this TTP function will potentially involve the following internal operations and external interfaces:

- **O5**. Event information storage.
- **O9**. Time-stamp generation,
- **I3**. Network interface,
- **I5**. Security manager interface.

9.6.3.18 Delivery authority functions

The delivery authority function will involve three main steps:

- receipt of information to release once appropriate evidence is received,
- the receipt and verification of evidence, and
- the release of information.

The receipt and release of information and the receipt of evidence will involve use of the Network interface. The verification of evidence will require use of the Cryptographic computation and Cryptographic key storage facilities, as well as the Time-stamp generation facility (to check the currency of any time-stamps in the evidence). The Security manager interface may be required to authorise and control the release of information. The Event information storage facility may be used to record parts of the overall procedure.

Thus this TTP function will potentially involve the following internal operations and external interfaces:

- **O2**. Cryptographic computation,
- **O3**. Cryptographic key storage,
- **O5**. Event information storage.
- **O9**. Time-stamp generation,
- **I3**. Network interface,
- **I5**. Security manager interface.

9.6.3.19 User identification

One fundamental function for TTPs will be to identify users for the purposes of generating user certificates, supplying users with keys, etc. The TTP will typically be responsible for generating information by which other entities identify a user, and hence will need to take some care in properly identifying a user before issuing that user with newly generated keys, key certificates, etc.

It will potentially involve the following internal operations and external interfaces:

- **O4**. User information storage,
- **I1**. User interface,
- **I5**. Security manager interface.

9.6.3.20 Credit checking

Part of the user accreditation process may involve performing checks on the user's creditworthiness. The User Interface will be required in order to communicate with the user requesting the accreditation. The Security Manager Interface may be required to enable a supervisory entity to authorise the credit checking procedure. The User Information Storage may be required to store user credit information and other information relating to the user. The Financial Institution Interface may be required to enable the credit checking TTP to request information about the user from financial institutions.

Thus this TTP function will potentially involve the following internal operations and external interfaces:

- **O4**. User information storage,
- **I1**. User interface,
- **I5**. Security manager interface,
- **I6**. Financial institution interface.

9.6.3.21 Fraud detection and management functions

Fraud detection and management will involve two main stages:

- the processing of received and stored event information to detect fraudulent behaviour,
- the generation and distribution of alarms when potentially fraudulent behaviour is detected.

Event information will be received via the Network interface, and stored event information will be obtained by using the Event information storage facility. Relevant information may also be obtained from the User information storage facility. The Event analysis facility can be used to process the received and stored information, detect potential fraudulent behaviour, and generate alarms. Alarms may need to have time-stamps attached (using the Time-stamp generation facility). Alarms can be distributed via either the Network interface or the Security manager interface.

Thus this TTP function will potentially involve the following internal operations and external interfaces:

- **O4**. User information storage,
- **O5**. Event information storage,
- **O8**. Event analysis,
- **O9**. Time-stamp generation,
- **I3**. Network interface.
- **I5**. Security manager interface.

9.7 TTP requirements for ASPeCT

9.7.1 Evaluation of TTP based security services

The TTP based security services listed in Section 9.4.1 were categorised according to their dependence on a TTP. This categorisation can be used to help identify the most important services

required from a TTP in order to define and prioritise the TTP supported UMTS security services to be developed within ASPeCT.

It has been agreed that the TTP supported UMTS security services, listed in order of priority in Table 9.7-1 below, are relevant to the ASPeCT project.

UMTS security services	TTP services probably involved
Secure billing	Non-repudiation services Symmetric key management Asymmetric key management
End-to-end confidentiality with support for warranted interception	Symmetric key management Asymmetric key management
Other end-to-end services, e.g. integrity, non-repudiation, origin authentication	Symmetric key management Asymmetric key management Non-repudiation services Identification and authentication
Other security services requiring the support of a TTP, e.g. • authentication servers, • access control to service profiles, • key escrow for disaster recovery.	all identified services

Table 9.7-1 List of prioritised security services for development within ASPeCT

It is initially anticipated that the demonstration and trial implementations will only support the first two security services, together with some additional end-to-end services from the third category.

The full set of identified TTP functions will probably not be implemented within ASPeCT. Instead, only those functions which are essential in supporting the selected services will be implemented. For example, providing time-stamping and notarisation for signatures was a function which is thought to be of value in the selected services, and which we should attempt to include. On the other hand, support for privilege attribute generation/distribution/archiving is probably not needed.

We should, at this stage, recognise the value and importance to practical systems of various functions of a TTP which may not be essential for our selected services. These may include provision of audit trails, arbitration in non-repudiation disputes, providing archives of signed documents, etc.

9.7.2 TTP security services to be developed within ASPeCT

In this Section, the four TTP supported UMTS security services identified within ASPeCT are discussed. Each security service is taken in turn and the ETSI UMTS security requirements which are fulfilled in each case are listed. This is largely based on the discussion in Section 9.4.2.

9.7.2.1 Secure billing

The secure billing service fulfils the following UMTS security requirements:

Miscellaneous requirements, including:

- Privacy of charging, billing and accounting information,
- Protect against unauthorised modification of charging, billing and accounting information,
- Ensure that entities having incurred charges cannot subsequently deny this.

In addition, this security service fulfils other requirements laid out by WP2.5 in Section 10 of this deliverable.

Secure billing includes overcoming the problems associated with the lack of trust between the participants involved in the billing process. A typical requirement is that a user must be able to find the charging for the services he uses incontestable. This will typically involve the service provider/network operator producing evidence to show the correctness of a bill.

A TTP supported non-repudiation service will be used in this application to generate evidence so that a user cannot deny that he invoked a given chargeable service. Other TTP services that may be used in the secure billing application may include symmetric and asymmetric key management services for supporting the integrity protection and confidentiality of billing information.

9.7.2.2 End-to-end confidentiality with support for warranted interception

This service fulfils the following UMTS security requirements:

User requirements, including:

- End-to-end privacy

Miscellaneous requirements, including:

- Ensure that facilities exist for legal interception by governments of all communications both within their own domain and possibly in other domains by agreement with the relevant government.

End-to-end confidentiality is seen as a increasingly important service as users are becoming more likely to transfer commercially sensitive information via telecommunications networks. The conflicting requirement of governments to be able to intercept communications demands a suitable key escrow mechanism. Such a mechanism would allow the retrieval of decryption keys under warrant from one or more TTPs, so that protected information can be intercepted by legitimate authorities.

9.7.2.3 Other end-to-end services

Other end-to-end services fulfil the following UMTS security requirements:

User requirements, including:

- End-to-end integrity, non-repudiation and origin authentication.

End-to-end integrity could be used to protect user traffic between sender and receiver. Similarly non-repudiation services could be offered over the end-to-end link between sender and receiver. However, these services are not seen as being as important as end-to-end confidentiality so have a lower priority.

End-to-end origin authentication would allow end users to authenticate the origin of messages. Again this is seen as an interesting service to develop, but time and manpower constraints are likely to mean that this service will not be developed within ASPeCT.

9.7.2.4 Other security services requiring the support of a TTP

The above identified security service areas involve only part of the full range of services that could be supported by a TTP. Other TTP services could potentially be used to offer a wider range of security services in UMTS. These services would fulfil various other UMTS security requirements, which were not mentioned above.

Examples of services would include (in no order of priority):

- trusted authentication servers for ‘ad-hoc’ authentication of various entities to each other within the UMTS environment, including authentication of users to each other,
- TTP supported access control to user service profiles, network management data, etc.,
- accreditation of UMTS users by TTPs,

- pre-personalisation of UIMs by TTPs,
- use of a TTP in fraud detection and management,
- key escrow based on TTPs for disaster recovery services,
- ...etc.

9.7.3 Functions required to support the selected security services

In this section the TTP functions required to support the two main security services we intend to implement are listed. We consider each security service in turn and briefly describe the role of each function.

9.7.3.1 Secure billing

To support secure billing, the provision of the following TTP functions is required:

- *Key generation*, notably for the generation of private/public keys for asymmetric algorithms. These keys will be used primarily in the provision of digital signatures to be used to prove that certain actions relating to billing have taken place.
- *Key distribution*, for the distribution of private keys.
- *Certification*, signed certificates are required to allow one entity to obtain a verified copy of another entity's public key. This will principally be used to allow entities to check digital signatures generated by other entities.
- *Directory*, for the storage and distribution of certified public keys.
- *Revocation*, for regularly generating and distributing Certificate Revocation Lists (CRLs). Note however, that it is still questionable whether on-line distribution of CRLs is required as part of the demonstration.
- *Evidence generation/recording/verification*, these functions may also be required in helping a TTP to prove that a user has carried out certain actions, notably those which initiate (and possibly terminate) chargeable services.
- *Time-stamping functions*, for certifying that the TTP has affixed a signature to a data item at a particular time. This will be used so that a user cannot deny initiating (or terminating) certain chargeable services at certain times.

9.7.3.2 End-to-end confidentiality with support for warranted interception

To support this security service, the provision of the following TTP functions is required:

- *Key generation*, primarily for generating private/public keys for end-to-end confidentiality using asymmetric algorithms.
- *Key distribution*, for distributing encipherment/decipherment keys to users that require end-to-end confidentiality.
- *Key escrow functionality*, here the TTP will act as a key distribution agent for the end users (as above), but will also support warranted interception by acting as a supplier of user keys to an interception agency.

9.7.4 Components required to support the selected security services

In this section the TTP components required to support the two main security services we intend to implement are listed. These lists are derived from the components of the required TTP functions for the two main security services, which are identified above in Section 9.7.3. Again, we consider each security service in turn.

This list of TTP components will serve as one of the main inputs to the design and specification of the TTP demonstration and trials. Decisions as to precisely which of these components actually need to be implemented will be made during the design and specification process. Final decisions will very much depend on the precise requirements of the charging and billing activity, and these will only fully emerge during the design and specification of the charging and billing security services.

9.7.4.1 Secure billing

To support secure billing, the provision of the following internal operations and external interfaces may be required:

- **O1**. Cryptographic key generation,
- **O2**. Cryptographic computation,
- **O3**. Cryptographic key storage,
- **O4**. User information storage,
- **O5**. Event information storage
- **O7**. Certificate generation,
- **O9**. Time-stamp generation,
- **I1**. User interface,
- **I2**. Directory interface,
- **I3**. Network interface,
- **I5**. Security manager interface.

9.7.4.2 End-to-end confidentiality with support for warranted interception

End-to-end confidentiality with support for warranted interception may require the provision of the following internal operations and external interfaces:

- **O1**. Cryptographic key generation,
- **O3**. Cryptographic key storage,
- **O4**. User information storage,
- **I1**. User interface,
- **I3**. Network interface,
- **I4**. Interception authority interface,
- **I5**. Security manager interface.

10. Requirements for security and integrity of billing (WP2.5)

10.1 Introduction

Relating to requirements on secure billing, three areas have been addressed:

- Overview of charging and billing in GSM;
- Principles and security issues for charging and billing in UMTS;
- Requirements on secure billing for UMTS to be investigated in ASPeCT.

Overview of charging and billing in GSM

The purpose of this activity was to ensure that important aspects of charging and billing in GSM which may be relevant also for UMTS are taken into account. These include GSM charging principles for mobile originating and mobile terminating calls and the transfer of accounting and billing data between GSM operators.

Principles and security issues for charging and billing in UMTS

In this activity, information already available from ETSI technical reports was exploited and four main security issues were identified which are expected to be crucial for charging and billing in UMTS. They are:

- problems resulting from lack of trust among network operators and service providers and between network operators / service providers and users;
- secure alternative methods of payment and advice of charge;
- electronic commerce;
- transfer of billing information between network operators and service providers.

Requirements on secure billing for UMTS to be investigated in ASPeCT

It was agreed that the security issues identified in the previous activity could not all be fully supported in the demonstrations within ASPeCT due to a lack of manpower. Therefore, the issues had to be prioritized. The prioritization is reflected in the order the issues were presented above, i.e. the highest priority for the work in ASPeCT has the investigation of methods which are suitable to increase confidence in the correctness of the billing process in a situation where trust among network operators and service providers and trust between network operators / service providers and users can no longer be taken for granted.

To this end, certain *billing scenarios* were evaluated which appeared to be suitable for the demonstrations in ASPeCT. A scenario which involves the provision of premium rate services by a Value Added Service provider was selected as the preferred scenario.

Threat and risk analyses were performed for the scenarios, taking into account the views of all the parties involved.

A semi-formal concept of *obligations* in open telecooperation was introduced in order to capture the requirements on non-repudiation services needed to secure the scenarios.

Trust relations between the involved parties were examined in more detail.

From the above, the *requirements* on secure billing services in the considered scenarios were derived. They are stated in terms of the *evidence* which needs to be provided to the involved parties in the course of the provision of a service so as to make the billing of the service secure.

Finally, the *Trusted Third Party services* which are likely to be needed to support the demonstration of secure billing services were identified.

10.2 Background

A service provider offering services to a user/subscriber has to ensure that the billing of the user/subscriber for the telecommunication services that he has used is done in a correct and transparent way acceptable to the parties involved. This is not new and applies to existing networks. However, that the current situation is not satisfactory was highlighted by a number of widely publicised fraud scandals in some telecommunications networks, which led to incorrect bills, contributed to an increased awareness of the problem with the general public, and resulted in an increased number of disputes between users/subscribers and network operators/service providers.

While it could be argued that these problems pertain to existing telecommunications networks, which implement only relatively weak security measures, a number of foreseeable technical developments show that secure billing will be a crucial issuenot only for existing networks, but also for UMTS. The most important of these developments are:

- an increasing variety of services, especially value added services or information services offered over telecommunications networks;
- an increasing use of telecommunications networks for the order and delivery of goods in electronic form, which may require support from security services provided by the network;
- the expected emergence of many new network operators and service providers, which may have serious implications for the trust relations among them.

The technical measures investigated in this project should concentrate on the additional risks introduced by the new developments listed above. For example, such measures are expected to be most

useful when certain value added services are requested. Another case where such measures might be useful are calls from origins or to destinations which are fraud-sensitive.
The introduction of measures for secure billing in UMTS is facilitated by the following facts:

- in UMTS every user is equipped with a smart card, the UIM, which makes it possible to provide additional security measures based on cryptographic procedures involving users;
- UMTS is still in the process of being defined so measures related to secure billing can be taken into account in the system architecture from the start.

Existing networks address the issue of secure billing to various degrees, none of them, however, address secure billing in a way that would be satisfactory for future systems like UMTS. A weak form of ensuring the integrity of billing, that may help to settle some disputes, is itemized billing. Another measure that may help to instil the confidence of a subscriber in a bill is advice of charge. In GSM more sophisticated security measures (e.g. authentication based on cryptographic procedures), which are used to counter fraud, may further increase the confidence of subscribers in the correctness of the bill.
However, none of these measures produces hard evidence that may allow a third party (such as a judge) to settle a dispute between a subscriber and a service provider. Also, these measures do not address the aspects of secure billing between network operators and service providers. For these purposes, stronger forms of protection are required in the future. These depend, however, on the availability of adequate legal frameworks, which cannot be the subject of the technical studies undertaken in this project. Stronger forms of protection may also contribute to increased confidence in the proper working of the system by all parties involved and may greatly reduce the number of cases where disputes arise.

10.3 Charging and billing in GSM

10.3.1 Overview of GSM billing

Billing involves the collection of revenue relating to calls made and services invoked on a telecommunications system. In GSM, charges for calls and other services are compiled from data recorded by the traffic handling part of the network. Each MSC/VLR will generate individual call records for each call containing all the information necessary to calculate the charge for the call. The call record will include the IMSI of the subscriber to be billed, which will uniquely identify the subscriber to its home PLMN.
Since GSM permits roaming between networks, the IMSI will be used as a basis for forwarding the bill to the correct home network. Roaming between networks is only permitted where the home and visited networks have arranged a bilateral roaming agreement.
Because GSM allows international roaming, billing and accounting principles are regulated by the signatories of the GSM MoU. When roaming is possible, the MoU states that charges made to a subscriber are to be collected by the operator with whom he holds a subscription. Therefore, the responsibility for paying the visited network operator lies with the home network operator of the subscriber; this is the concept of the Transferred Account Procedure (TAP) and is discussed in more detail in Section 10.3.3.
A visited network will sort its call records before forwarding them to the correct home networks for internal accounting. A global bill will be compiled for each network that the visited network has a roaming agreement with.
Call records for home subscribers, whether internally generated or coming from visited networks, must be sorted on a subscriber basis. If GSM service providers are used, a global bill will be sent to each one for calls corresponding to their subscribers. If the operator bills subscribers directly, then individual subscriber bills are established and sent out by the operator. Other variants exist where, for example, the operator compiles individual subscriber bills on behalf of the service provider.

10.3.2 GSM charging principles

A general charging principle has been developed by the signatories of the GSM MoU. These principles are laid out by the MoU Billing Administration and Roaming Group (BARG). This group deals with all commercial and administrative issues required to support international roaming. These issues include charging, billing and accounting.

MoU BARG Permanent Reference Documents (PRDs) contain the regulations applicable to the relationships between PLMN operators in different countries. These regulations are only appropriate to those operators who have arranged bilateral roaming agreements.
The general tariffing principle laid down in BARG is that the calling party pays the total charge for the calls that he initiates [1]. The charging principles relating to Mobile Originated (MO) and Mobile Terminated (MT) calls are dealt with in the following sub-sections.

10.3.2.1 Mobile originated calls

According to the general tariffing principle, the calling party pays the total charges for all calls that he initiates. The only exceptions to this general principle are:

- Reverse or transfer charge calls.
- Call re-routing charges[4] in respect of calls to a mobile station that has roamed away from its home PLMN.
- Forwarded calls: The forwarded part of a call is treated and charged as if it were a separate mobile originated call initiated by the forwarding subscriber.
- Where the terminating PLMN operator charges his subscribers for mobile terminating calls, and chooses to similarly charge visiting subscribers. (see Section 10.3.3).

Where a subscriber roams internationally and originates a call on a visited PLMN, the visited operator will bill the home operator through the Transferred Account Procedure (TAP) in respect of the use of the visited PLMN. The charges raised will be at the visited PLMN operator's normal network tariff computed as if the visiting subscriber were a home subscriber of the visited PLMN.
The visited PLMN operator may apply a multiplier (the visited PLMN multiplier) to the normal network tariff for MO calls. The visited PLMN multiplier may be a maximum of 1.15. All visiting subscribers to a PLMN are treated equally by the visited PLMN operator in terms of tariff and visited PLMN multiplier.
The home PLMN operator will use the charging record supplied through the TAP as the basis for subscriber billing.

10.3.2.2 Mobile terminated calls

According to the general tariffing principle, the calling party, whether a mobile or fixed subscriber, calling towards a local PLMN subscriber pays a charge incorporating the revenue necessary to support the called PLMN structure.
However, under some circumstances, PLMN operators may be required to charge their subscribers for mobile terminated calls. There are two general approaches:

- **Home PLMN charging** Where a subscriber has roamed away from his home PLMN, the charges for the call re-routing part from the subscriber's nominal address to his actual address may be charged to that subscriber.
- **Visited PLMN charging** If a PLMN operator charges his own subscribers for Mobile Terminated calls received while in the home PLMN, he may charge visiting subscribers for the same type of calls at the same rates.

Where a subscriber receives a call whilst roaming internationally, the visited PLMN operator will raise a charging record and transmit this to the home PLMN operator through the Transferred Account Procedure (TAP) [2]. According to the general tariffing principle, the charging record will be zero priced. However, this charging record is still used by the home PLMN to raise a charge for the re-routing of the call to the visited PLMN.
Mobile terminated calls are charged on the basis of a call originating in the home PLMN of the roaming subscriber and terminating in the visited country in which the call is received. The call re-routing charge to the called party is consequently dependent on where the call is received rather than where it originated. However, if optimal routing is used this may not be the case (see Section 10.3.4).

[4] If a call is directed towards a MS, which has roamed from its home PLMN, the call will need to be re-routed to the visited PLMN. The call re-routed part covers the part of the call from the point at which it is re-routed up to the point at which the call reaches the terminated PLMN.

10.3.3 Transfer of accounting and billing data between GSM operators

The Transferred Account Procedure (TAP) is described in the GSM MoU regulations [2]. The TAP involves the transfer of charging records of visiting subscribers between operators. Issues relating to the TAP are dealt with in MoU Transfer Account Data Interchange Group (TADIG). This group also deals with issues relating to the transfer of other data between PLMN operators, including data between EIR administration systems.

A visited PLMN, with which a roaming subscriber interconnects, must receive compensation for the receipt and delivery of calls. In the TAP, the visited PLMN operator agrees that the charge for the GSM telecommunication services used will be paid by the home PLMN operator instead of by the user.

Charging records for roaming subscribers making MT and MO calls are transferred between operators on a regular basis. The aim is to exchange data within 36 hours of call completion and there is nominally at least one exchange a day. If no data is due, an electronic notification is sent on a daily basis. However, there is a requirement to receive the charging records in near real-time.

The method of transfer of charging records is electronic file transfer using FTAM over X25. The protocol stack for the transfer is given in MoU PRD TD 04 [3].

Instead of transferring call records directly to the appropriate operator, they could be sent to a clearing house. In the clearing house scenario, all network operators could have their liabilities to other operators settled, whilst minimising the number of financial transactions. In addition, the clearing house allows billing information for roamers to be made more accessible to network operators.

10.3.4 Optimal routing in GSM

Optimal routing is a GSM feature by which calls directed to a mobile subscriber, who has roamed away from his home PLMN, are routed directly, instead of via his home PLMN, to the mobile subscriber's actual location or the forwarded-to destination. The visited network can obtain the indirect route to the roamer's HLR from the dialled number, and the direct route to the called party's location by interrogating the roamer's HLR. The two routes can then be compared and a decision made as to which one is the preferred option. Optimal routing eliminates the "tromboning effect" and thus reduces the cost of certain traffic cases involving roaming subscribers.

The reasons for directing all calls to a subscriber's home PLMN in the first implementations of GSM are three fold:

- administrative and operational problems;
- limitations of technical functionality;
- ease of charge computation.

The first two limitations have been largely overcome. The issue of charge computation still exists, however, and will only be fully resolved with the introduction of an intelligent routing decision node in the home network. The intelligent node will determine and compare the charges for the direct route and the route via the home network.

In a first phase of implementation, calls shall only be optimally routed in the case where the decision whether to directly route the call is easy and where the subscribers can be charged adequately on the basis of the TAP as defined in TADIG. These cases exist when:

- the optimal routed leg originates and terminates within the same country;
- the optimal routed leg terminates in the country to where the leg would have been routed without optimal routing.

In particular, the first phase of optimal routing will solve tromboning problems for mobile-to-mobile calls, where both mobiles are in the same country, and for calls with conditional call forwarding towards the home country of the forwarding party.

The intermediate phase of implementation would deal with more complex charging scenarios. This would require use of enhanced IN functionality which is due to come on-line with the introduction of CAMEL (Customised Application Mobile-network Enhanced Logic) [4] The final phase would cover all scenarios including the optimal routing of mobile terminated calls that originate in the PSTN/ISDN. This will require a significant level of enhancement to the fixed network.

The charging principles relating to optimal routing and the transfer of accounting and billing information in the optimal routing case are summarised in the sub-sections below. These principles and standards are taken from the MoU regulations [1].

10.3.4.1 The charging principles applicable to optimal routing

TAP charging records transferred between PLMN operators when optimal routing is invoked must contain sufficient information to allow for the correct billing of subscribers, inter-network accounting, and the handling of customer care requirements. This information must include the originating network for MT calls and the terminating network for MO call records. The revenue generated by an optimally routed call should be enough to cover the inter-network accounting costs of each network handling that call.

With optimal routing, international Signalling System No 7 (SS-7) may be used to communicate with the home network with no corresponding traffic. If a signalling charge is demanded by an international operator, it could be absorbed by the PLMN operator or passed on to subscribers. If it is passed on to subscribers, it is essential that this charge should not wipe out the benefits of optimal routing.

10.3.4.2 The transfer of billing information in the optimal routing case

With optimal routing two charging records are generated. These are passed to the home network using the TAP.

A MO call record is generated in the originating PLMN of the optimally routed leg of the call. This MO call record shall include information identifying the terminating PLMN. The charges raised will be at the originating PLMN operator's normal network tariff for a MO call. The visited PLMN may apply a multiplier (the visited PLMN multiplier) to its normal network tariff. The visited PLMN multiplier may be a maximum of 1.15 (see Section 10.3.2.1).

A MT call record is generated in the terminating PLMN of the optimally routed leg of the call. A MT call record shall include information identifying the terminating PLMN of the optimally routed leg of the call. This call record shall be zero-priced except in the 'visited PLMN charging' case (see Section 10.3.2.2). The home PLMN of the subscriber will identify that the MT call record corresponds to an optimally routed leg, because the originating PLMN identified in the MT call record does not correspond to the home PLMN of the subscriber.

Note that with optimal routing, the called subscriber's home network has no record of the call. Instead, it relies on the appropriate charging records being raised by the originating and terminating networks. Without optimal routing, a Roaming Call Forward (RCF) charging record is generated by the called subscriber's home network. This accompanies the MO call record generated in the originating PLMN and the MT call record generated in the terminating PLMN.

10.4 Principles for billing in UMTS

10.4.1 Charging principles

From the subscriber's point of view, UMTS must fulfil the following general requirements which are taken from a draft ETSI Technical Report developed by ETSI SMG, cf. [5]:

- a UMTS subscriber shall be billed, for the UMTS services subscribed to, by their home UMTS service provider only.
- the subscriber should generally find charging understandable and incontestable in relation to the services offered. The user should be able to estimate any charges that may result before making use of services.
- UMTS must generally ensure that the subscriber is not charged for the usage of services by other parties. Other parties may be allowed to transfer charges for use of services to the subscriber only if specifically authorised to do so.

Recall, there is an important difference between the GSM and UMTS role models, which is that the service provider entity in UMTS holds all subscriber related information, including information that would have been held by the home network operator's HLR in GSM. As a result, there is no home or visited network as such in UMTS. Instead, all calls are directly routed, with the originating network interrogating the service provider for routing information.

10.4.2 Charging based on optimal routing

Route optimisation in UMTS reduces the involvement of the service provider to pure data enquiry; requiring the bearer capabilities to exist only between the originating and visited networks and the involved transit networks [5].
It should be possible to pass the reduced charges, based on the optimal route, onto the subscribers. The involvement of the originating and terminating party's service providers may incur an extra component of charging for signalling enquiries and related activity. As with GSM, the cost of extra signalling in UMTS must not wipe out the original cost benefits of optimal routing. In addition, a mechanism must be developed for the secure transfer of billing information.

10.4.3 Supplementary services

In UMTS, supplementary services relating to billing may be offered. These may include, Advice of Charge (AoC) and Alternative Payment Methods (APM).

10.4.3.1 Advice of charge

An Advice of Charge (AoC) service may be provided in UMTS. This would provide the user with a reliable estimate of the cost of the service used. When required to indicate the total accumulated charge, the Mobile station would be able to display the running cumulative unit charge[5], which may be stored in the User Identity Module (UIM).
In GSM, the charge is estimated by the Mobile station (MS) [6]. At the charging point, the MS is informed of the charging rate in home units. The MS then uses its independent internal clock to time the call from the charging point to the end of call. The MS can then derive the number of home units used in order to advise the user of the charge incurred.

10.4.3.2 Alternative methods of payment

Within UMTS, secure Alternative Payment Methods (APM) may be defined [6]. These are ways of paying other than the conventional way where the subscriber is billed after he has used services. APM could be realised with User Identity Modules (UIMs) in the form of prepaid cards or reloadable smart cards.

10.5 Security issues for billing in UMTS

10.5.1 Problems resulting from lack of trust between the participants

10.5.1.1 Relation between users and service providers

A user must be sure that his service provider will correctly process the billing information he receives from network operators. The level of trust offered by a service provider influences the user's choice in deciding which service provider he should subscribe to. However, in order for the user to find charging "incontestable" (cf. charging principles in 10.4.1above), additional measures may be required in order to increase the user's confidence. On the other hand, service providers do not necessarily trust users. Users may challenge correct bills. Thus, it may be useful for service providers to be able to produce evidence showing the correctness of the bill.

10.5.1.2 Relation between users and network operators

A user does not have as much flexibility in choosing his network operator as he does in choosing his service provider, since the choice is limited to the networks that have roaming agreements with the user's service provider and to those that can provide coverage at the user's location. Although there is no direct relation between users and network operators with respect to billing, the user must be sure that the network operator generates correct charging records for services, which the user invoked while using the network. This is because the user's service provider will base his bill on those

[5] Translating this unit charge into an actual cost, in the correct home currency, is a difficult problem

charging records. As in 10.5.1.1 above, it may be helpful if evidence was available supporting the correctness of charging records.

10.5.1.3 Relation between service providers and network operators

In GSM, transfer of billing information between different organisations is only required when roaming outside the home network. However, with UMTS, charging records are generated by network operators and must be forwarded to the appropriate service provider for all calls. The service provider must be sure that the network operators with whom it holds a roaming agreement will supply correct billing information.

A GSM MoU-like arrangement may not be possible with UMTS, since the number of network operators and service providers will increase. Without an MoU-like body, network operators and service providers may not be bound to a set of regulations. As a result, the required level of trust may not be obtainable.

The equivalent of a GSM roaming agreement in UMTS will be an agreement between a service provider and the network operator that provides the bearer capabilities for the service. Because the number of operators is likely to increase dramatically, individual bilateral agreements are likely to become unworkable in their existing form. As a result, some other mechanism must be provided so that the service provider can have confidence in the billing information supplied by the network operator.

10.5.2 Security requirements relating to supplementary services

Security requirements relating to Advice of Charge may include:

- Ensure that advice of charge information is a reliable estimate of the actual charge incurred.
- Privacy of advice of charge information.
- Protection against unauthorised modification of advice of charge information.

Security requirements relating to Alternative Payment Methods may include:

- Protection against unauthorised disclosure of subscriber information.
- It should be an open system, i.e. enable participation of several financial institutions, service providers, retailers and users without the need for a relation of trust between them. Trust is needed between service provider and subscriber, and between service provider and retailer.
- Anonymous payment to the service provider may be necessary.
- The subscriber may require some insurance against theft or loss (e.g. a maximum amount may be set).
- If during a (reload) transaction the connection fails, the procedure must be terminated properly.

10.5.3 Electronic commerce over UMTS

Electronic commerce is described here as the order and delivery of goods in electronic form using UMTS as the communication network linking the customer and the merchant. This involves two billing relations: Firstly, the UTMS service provider will bill the user and the merchant in their roles as UMTS users for the provision of the communication link between them. This is identical to the billing of basic service UMTS calls and is therefore of no interest here. Secondly, the merchant will bill the customer for the delivered goods. The corresponding payment may be handled by a payment organization (e.g. a credit card organization). This second billing process between customer and merchant is of course independent of UMTS. It is included here, however, because the exchange between customer and merchant may benefit from the support of end-to-end security services provided by the UMTS network. Therefore, it would be worth while to investigate, which end-to-end security services would be needed for this application, and then to demonstrate the usefulness of these end-to-end security services in the application.

10.5.4 Transfer of billing information between service providers and network operators

The transfer of information between service providers and network operators involves at least two issues.These issues are discussed below.
A reduction in the transfer time of billing information between organisations is required to allow service providers to reduce their risks of fraud by monitoring, in near real-time, the charges being accrued by any particular service. In principle, call records could be recorded in participating networks and transmitted to the appropriate service provider in real-time on completion of a particular call sequence. However, in practice technical functionality may limit this.
The transfer mechanism itself should be made secure to provide confidentiality, integrity, and origin authentication of billing information. Non repudiation of billing information transfer between organisations may also be required[6].

10.6 Requirements on secure billing for UMTS to be investigated in ASPeCT

10.6.1 Selection of security issues

It was agreed that the security issues identified in the previous Section could not all be fully supported in the demonstrations within ASPeCT, due to a lack of manpower. Therefore, the issues had to be prioritized. The prioritization is reflected in the order the issues were presented above. Thus, the highest priority for work to be done within ASPeCT is the investigation of suitable methods for increasing the confidence of users in the correctness of the billing process in a situation where trust among operators and providers and trust between operators / providers and users can no longer be taken for granted. This choice was made since it was felt that this work has the greatest potential for innovation.
Further investigation is required into how far the mechanisms used for increasing confidence between users and operators can be combined with a secure advice of charge supplementary service. Depending on the outcome of this investigation, work on secure advice of charge could be included as well in WP 2.5.
Electronic commerce is seen as an interesting application as it may benefit from end-to-end security services offered by UMTS, such as end-to-end confidentiality and integrity, authentication and non-repudiation. However, it is unlikely that the resources will permit to work on this issue as well.

10.6.2 Billing Scenarios

In order to discuss secure billing in UMTS, we describe two scenarios in which certain services are requested for which a user is consequently billed. The purpose of the discussion is to find a scenario which covers typical security requirements and is thus suitable for a demonstrator in the later phases of the project. It is not claimed that the scenarios cover all cases in which billing may occur. The second scenario which includes the provision of value added service is an extension of the first one which involves the provision of basic services only. It is the preferred one as, on the one hand, the billing relations are more complex there and, on the other hand, the potential for fraud is greater.
In addition to the two scenarios relating to billing for UMTS services, we also describe a scenario relating to electronic commerce in order to show in more detail in how far a certain end-to-end security service, namely non-repudiation, may be beneficial for such an application over UMTS.
For each scenario, we describe the participants involved and their relationship in the actual provision of service and in the billing process.

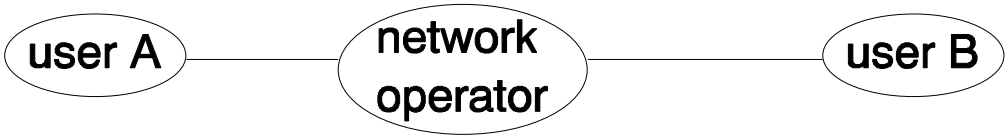
10.6.2.1 Scenario including the provision of basic services only

This is the scenario where a user may roam in a network other than his home network. He uses only basic services, no value added services. For simplicity, we only consider the case of mobile originating calls. (Mobile terminating calls are treated in a similar way.) Hence the roles in the provision of the service and in the billing for the use of the service are the following:
Participants
1. User *U*

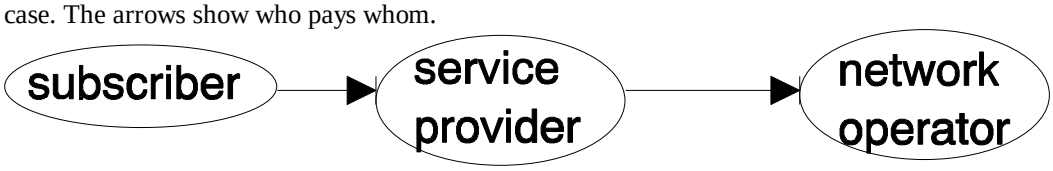
[6] This is separate to the non-repudiation of the use of a chargeable service.

2. Subscriber *S*
3. Service provider *SP*
4. Visited UMTS network operator *NO*
   This is the network operator in whose domain the user requesting a service is roaming.
5. Other network operators
   There may be a number of other network operators involved. These include the network operator in whose domain the destination of the call lies and any number of transit network operators.
6. An intruder *Z* masquerading as legitimate user.

Relations in provision of service
The following figure depicts the communication relations among the participants involved in the provision of service in the simplest case where only one network operator is involved.



Billing relations
The following figure depicts the relations among the participants in the billing process in the above case. The arrows show who pays whom.

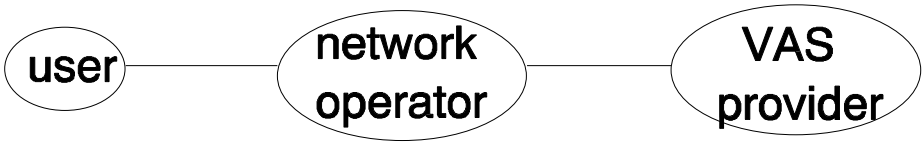


10.6.2.2 Scenario including the provision of value added services

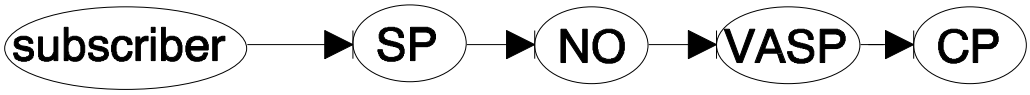
The user is charged a particular rate for this type of calls. The charges are collected by the user's UMTS service provider who in turn forwards part of the money payed by the user to other parties involved.

Participants
1. User *U*
2. Subscriber *S*
3. Service provider *SP*
4. Visited UMTS network operator *NO*
   This is the network operator in whose domain the user requesting a service is roaming.
5. Other network operators
   There may be a number of other network operators involved. These include the network operator in whose domain the destination of the call lies and any number of transit network operators.
6. Value added service provider *VASP*
7. Content provider *CP*
8. An intruder *Z* masquerading as legitimate user.

Relations in provision of service
The following figure depicts the communication relations among the participants involved in the provision of service in the simplest case where only a network operator is involved.



Billing relations
The following figure depicts the relations among the participants in the billing process in the above case. Other billing relations may be possible. The arrows show who pays whom.



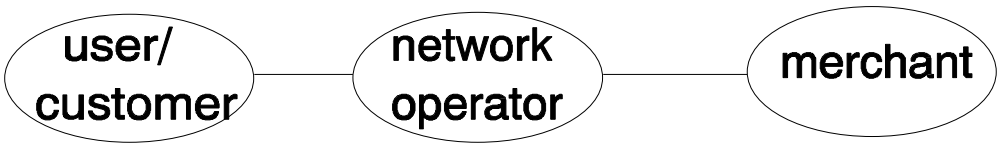
10.6.2.3 Scenario including electronic commerce over UMTS

Here, UMTS is only used to provide a communication path between a customer and a merchant selling electronic goods. The UMTS service provider is only responsible for collecting the charges for the call, not for the payment by the user /customer for the delivered goods. This payment is assumed to be handled by a payment organization (like e. g. a credit card organization).
The scenario is included in order to investigate what UMTS security services (including end-to-end security services) may be required or desirable to support this kind of transactions.
Participants (in addition to those listed in section 10.6.2.1 above**)**

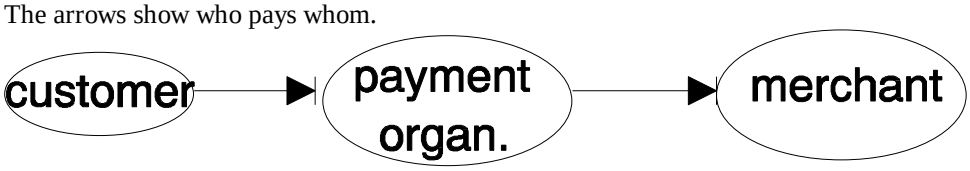
7. Customer *C*, acting as a normal user from a UMTS point of view.
8. Merchant *M* of electronic goods ("infogoods" in short) acting as a normal user from a UMTS point of view.
9. Payment organisation *PO*, possibly acting off-line.

The transactions among C, M and PO depend on the type of protocol used for electronic commerce. As an example, they may be as follows: *C* calls *M* by using UMTS and requests an offer for infogoods. *M* sends back a detailed offer, containing, at least, the type of goods, the amount, and the price. *C* might then send an order to *M* and *M* returns the infogoods. *C* then sends a commitment to pay (cheque) to *M*.

Relations in provision of service
The following figure depicts the communication relations among the participants involved in the provision of service in the simplest case where only one network operator is involved.



There are two billing processes involved here: One is for the call, the other is for the delivered infogoods.
Billing relations (call)
The relations among participants in this case are as in 10.6.2.1 above.
Billing relations (infogoods)
The following figure depicts the relations among the participants in the billing process in this case. The arrows show who pays whom.



10.6.3 Threats

10.6.3.1 Introduction

In sections 10.6.3.2- 10.6.3.4 below, threats are described in an abstract way. In this introduction, we briefly want to mention (in a certainly non-exhaustive manner and for illustrative purposes only) a few concrete ways in which threats have been realized in existing (fixed or mobile) networks[7]. Some of these threats can be countered by already envisaged security measures for UMTS (other than secure billing), some cannot.

[7] Section 4 contains a more comprehensive discussion on fraud scenarios in UMTS

1. Fraudsters obtain a valid subscription under a false identity. They then set up a very high number of calls, requesting particular VASs. They act in collusion with the corresponding VAS providers or content providers and share the profit with them later. When the bill arrives the fraudster has disappeared. The VAS providers and content providers receive their legal share of the charges incurred nevertheless, unless their involvement in the scheme can be proven.

   This type of fraud can only be prevented by organizational means like thorough checking of identities and credit records, not by cryptographic techniques. It can often be detected by fraud detection means such as the ones being developed in WP 2.2.
2. Fraudsters gain access to telecommunications services by masquerading as legitimate users. This is typically done by tapping access lines in fixed networks or using cloned phones in (typically) analogue mobile networks. The fraudsters then proceed as in 1.
   This type of fraud can be effectively prevented by authenticating the user before granting service and by preventing cloning. Consequently, this type of fraud is unlikely to occur in UMTS (assuming that UIM smart cards cannot be cloned).
3. Fraudsters have accomplices among the employees of the network operator or service provider. These employees have access to the switch and/or can manipulate the billing data in such a way that no user or the wrong user is charged for the call. This type of fraud cannot necessarily be prevented by authentication. It calls for different measures. It is also potentially dangerous for UMTS. For UMTS users and service providers the situation is aggravated by the fact that a roaming user may gain access to services from any network in which he is allowed to roam. The UMTS service provider has no control over the level of security in the visited network. Nevertheless, he is reponsible for paying the visited network operator.
4. In addition, disputes over bills may arise from the fact that users are correctly billed, but do not agree with the bill (e.g. because they were unaware of the particular rates which applied to the selected type of service) and are therefore not willing to pay - although it should be mentioned that this does not necessarily constitute fraud.

In the last two cases, additional features providing secure billing are expected to be useful.
In order to reduce the threat analysis to the essential relations in our context, we do not distinguish between the user and the subscriber. Of course, a subscriber can be defrauded by a user he has trusted in. But handling the relation between user and subscriber is outside the scope of the communication system.
Similarly, we do not take into account here the relation between the Value Added Service Provider and the Content Provider.

10.6.3.2 Scenario including the provision of basic services only

From U's point of view
In general, *U* fears an unjustified bill, either too high for a service he has used or for services he has not used at all. From the user's point of view the threats that exist are:
T1 *NO* defrauds *U*
A dishonest *NO* falsely claims that at a given point *U* has made a call of a certain duration to a certain person.
T2 SP defrauds U
SP falsely pretends that some *NO* has charged *SP* on *U*'s account.
T3 Z masquerades as U
An intruder *Z* requests services on *U*'s account so that *U* is going to be charged for that.

From SP's point of view
T4 U does not pay for his accounts
U is either not willing or not able to pay for his accounts.
T5 U defrauds SP
A dishonest user *U* rejects a correct bill by falsely claiming that this bill is incorrect.
The crucial point for *SP* is to distinguish between this case and one of the cases below.
T6 NO defrauds SP
A dishonest *NO* presents a forged bill for a user *U* who is subscriber of *SP*
T7 Z defrauds SP by masquerading as U
An intruder *Z* requests services on *U*'s account.

One could argue that threats T6 and T7 are more a threat to *U* than to *SP* as *SP* may not even recognize that he was defrauded and may be able to recover the charges from *U*. But they definitely become financial threats to *SP* if later on *U* can prove that he was not responsible for that billing. But even if this proof is not possible this threat might lead to financial losses. For example, *U* could cancel the subscription, or disputes about bills could damage the reputation of *SP* so that customers might prefer other service providers.

From NO's point of view
U defrauds NO
U rejects a correct billing by claiming that either he did not make that call at all or that some parameters (duration time, tariff etc.) are wrong.
T8 Z defrauds NO by masquerading as U
An intruder *Z* requests services from *NO* on *U*'s account and *U* can prove that he is not responsible for that.
T9 SP defrauds NO
A dishonest service provider *SP* charges the users, but falsely claims against *NO* that the billings are incorrect.

10.6.3.3 Scenario including the provision of value added services

In addition to the threats listed in 10.6.3.2, we have to consider the following threats:

From U's point of view
T10 VASP defrauds U
- A dishonest *VASP* falsely claims that *U* has used a certain type and a certain amount of his info service.
- *VASP* does not provide the service he announced. For example, *VASP* promises to provide some information, *U* gives a call to *VASP*, listens some minutes to the tape but does not get the desired information. Since *U* has initiated the call, he will have to pay for it without getting any benefit.

T11 SP defrauds U
SP falsely pretends that some *VASP* has charged *SP* on *U*'s account.

From SP's point of view
T12 VASP defrauds SP
A dishonest *VASP* presents a forged billing for subscribers of *SP*.
(For this threat the same remark holds as for the threats T6 and T7.)

From NO's point of view
T 13 VASP defrauds NO
A dishonest *VASP* presents against *NO* an incorrect billing on *U*'s account and *U*'s service provider is convinced that *U* is not responsible for that billing.

From VASP's point of view
T14 U defrauds VASP
U rejects a correct billing by claiming that either he did not make that call at all or that some parameters (duration time, tariff etc.) are wrong.
T15 Z defrauds VASP by masquerading as U
An intruder Z requests info services from VASP on U's account and U can prove that he is not responsible for that.
T16 NO defrauds VASP
A dishonest network operator *NO* charges the service provider of the user but falsely claims against *VASP* that the billings are incorrect.

10.6.3.4 Scenario including electronic commerce over UMTS

In addition to the threats listed in 10.6.3.2, the following threats are possible. We assume here that fraud committed by the *PO* is unlikely.

From C's point of view
T17 M defrauds C
M does not stick to his offer, i.e.
1. *M* does not send the ordered infogoods
2. after sending the infogoods *M* wants to get more money than he mentioned in his offer.

From M's point of view
T18 C defrauds M
After accepting the offer, ordering and delivering of the goods *C* is not willing to pay the price.

From PO's point of view
T19 C does not pay for his accounts
U is either not willing or not able to pay for his accounts. (This is a general risk a *PO* faces which is not particular to electronic commerce.)
T20 C defrauds PO
A dishonest user *U* rejects a correct billing by falsely claiming that this billing is incorrect.
T21 M defrauds PO
A dishonest *M* presents a forged bill for a customer *C*.

10.6.4 Obligations

The concept of formalizing obligations in open telecooperation was investigated by Rüdiger Grimm and is published in [i] . Describing the obligation structures helps to find out which sort of non-repudation service is actually necessary in the underlying scenario.
Interesting in our case are the so-called *conditional obligations*, i.e. obligations which exist if one or more preconditions are fulfilled. Thus conditional obligations are described by implications
if event(s) ... happen(s) *then P* is obliged to ...
An obligation state is *true* if either the precondition does not hold or if the obligation has been fulfilled. An obligation state is *false* if the precondition holds and the obligation is not fulfilled. Each honest participant will keep his obligations states true. On the other hand, he has to pay attention to false obligation states of his partners.

10.6.4.1 Scenario including the provision of basic services only

In the billing scenario there are the following two obligation expressions:
O1 *If U* has used the service of *NO*
then U is obliged to pay *SP* for this service
and SP is obliged to pay *NO* for this serivce.

10.6.4.2 Scenario including the provision of value added services

O2 *If U* has used the service of *VASP*
then U is obliged to pay *SP* for this service
and SP is obliged to pay *NO* for this serivce
and NO is obliged to pay *VASP* for this serivce.
Note that if one of the participants does not fulfill his obligations, the other partners still have to fulfill their obligations. If, for example, *U* has used the service of *NO* and he does not pay *SP*, *SP* is still obliged to give *NO* the money.

10.6.4.3 Scenario including electronic commerce over UMTS

In addition to O1 we have the following obligations
O3 *If M* has made an offer to *C*
and C has ordered these goods
then M is obliged to deliver the goods according to his offer.
O4 *If C* has ordered these goods
and M has delivered these goods
then C is obliged to pay the price according to *M*'s offer to *PO*
and PO is obliged to forward the money to *M*.

10.6.4.4 Balance Properties

There are two cases which cause problems:

B1 Some participant does not fulfill his obligations although the precondition holds, i.e., his obligation state is false. Then the participant to whom the first is obliged needs a proof that the precondition has actually been fulfilled.

B2 Some participant falsely claims that another participant has not fulfilled his obligation, i.e., that his obligation state is false. Therefore, each participant needs a proof for the fact that he actually has fulfilled his obligations.

Thus a participant needs both a proof for the true preconditions of obligations of other participants to him and a proof for the obligations he has fulfilled. The first one serves the legal enforcement of a justified claim and the latter serves the defense in case of an unjustified accusation.

The system is said to be in balance if each change of the obligation state involves the provision of the corresponding proofs in the above sense.

10.6.5 Trust Relations

As already mentioned in 10.5.1, trust relations are expected to influence the choice of security mechanisms and, in particular, the kind of proof a participant accepts. These relations can be represented in a trust matrix. So, before security mechanisms are specified, the assumptions on the entries of the trust matrix have to be made clear.

Trust relations for a UMTS system in operation are, of course, difficult to predict and generalize. In some cases a certain *SP* completely trusts a certain *NO* (maybe they are identical or they belong to the same company), in other cases *SP* and *NO* will not trust each other.The following trust matrices are therefore to be taken as examples.

10.6.5.1 Scenarios including the provision of basic services only and of value added services

In the scenarios under consideration the trust relations between the User, the UMTS Service Provider, the Network Operator and the Value Added Service Provider are important.

↓ trusts →	*U*	*SP*	*NO*	*VASP*
U	-	partly	partly	no
SP	no	-	partly	no
NO	no	partly	-	no
VASP	no	no	partly	-

Figure 1: Trust matrix 1

The entries "partly" may be justified as follows:

In general, there may be trust in a company and that the company will not defraud deliberately. But this does not exclude that there are malicious employees within in the company who will defraud to their own benefit, maybe in cooperation with a third party, e.g. some *VASP*. And this does not exclude malfunctioning of the system. Software errors might result in wrong billings.

A few remarks on the nature of trust relations:

1) *U* has chosen *SP* as his service provider out of a set of competing companies. Thus we can assume that he trusts at least "partly". On the other hand *SP* has a lot of customers and there might be some dishonest users so that, in general, *SP* will not trust his clients. In particular, we see that the trust relations are not symmetric.

2) We may assume that other participants will not trust *NO* concerning jurisdiction on other participants. For example, *NO* has a contract with a certain *VASP* and trusts *VASP* to be an honest company. But *SP* does not trust *VASP*. Thus *SP* may trust *NO* not to defraud him but *SP* might feel that *NO* does not carefully check its contract partners. In particular, trust relations are not transitive.

10.6.5.2 Scenario including electronic commerce over UMTS

In this case the trust relations between the client, the merchant and the payment organisation are important.

↓ trusts →	*C*	*M*	*PO*
C	-	no	partly
M	no	-	partly
PO	no	no	-

Figure 2: Trust matrix 2

10.6.6 Requirements

The requirements are derived from the balance properties and the obligations in section 10.6.4 above.

10.6.6.1 Scenarios including the provision of basic services only

From balance property B1 and obligations O1 and O2 follows:
R1 NO *and* SP *need proof that* U *used a certain service in a certain way.*
From balance property B2 and obligations O1 and O2 follows:
R2 U *and* SP *need proof of payment to* SP *and* NO *respectively.*

10.6.6.2 Scenarios including the provision of value added services

From the balance properties B1 and B2 and obligations O1 and O2 follow in addition to the requirements of 10.6.6.1:
R3 VASP *needs proof that* U *used a certain service in a certain way.*
R4 NO *needs proof of payment to* VASP.

10.6.6.3 Scenario including electronic commerce over UMTS

From balance property B1 and obligations O3 and O4 follows:
R5 C *needs proof that* M *made a certain offer and* C *made a corresponding order.*
R6 M *and* PO *need proof that* C *made a certain order and that* M *made the corresponding delivery.*
From balance property B2 and obligation O4 follows in addition:
R7 C *and* PO *need proof of payment.*
These requirements cover all threats listed in 10.6.3 above except T4 and T19 which cannot be countered by technical means.

10.6.7 Requirements on Trusted Third Parties in order to provide secure billing

In the following we go a step further than it is usually done in a first analysis of risks and security requirements. Although we have so far said nothing about the involved security measures we assume that crypto functions will be used. The proofs mentioned in 10.6.6 will be probably ensured by digital signatures. Thus, an asymmetric key management service is needed. Therefore, we assume that Trusted Third Parties (TTP) are involved. More exactly we assume that a TTP network as described in section 9 of this report will be established. The TTP network consists of a number of TTPs with certain trust relations among them which allows any two entities which have trust agreements with any two TTPs in the network to find at least one path of mutual trust between each other.
The TTP network will at least distribute certifications for public keys. This brings up the problem of revocation of certificates and the solution of this issue might involve the problem of time-stamps.
If advice of charge is used as a supplementary service within the billing concept it might be important to ensure confidentiality of the charging data on the air interface by means of symmetric cryptoalgorithms. Thus, in addition, symmetric key management is necessary. But, from the view of secure billing, symmetric key management is not necessarily an issue for the TTP. Once the public keys are distributed, the shared keys can be established by key agreement or key translation without the involvement of the TTP. (This just reflects the view of secure billing and does not involve other requirements like key escrow.)
The minimal requirements for a TTP in order to provide secure billing is support for asymmetric key management and distribution of correct time stamps. Thus, to support secure billing the minimum is the provision of the following *off-line* TTP services,

- generation of private/public keys (In principle, this could be done by the users. But it is supposed that in the near future the majority of the UIM will not be able to generate these key pairs. Still, this option should be possible from the view of secure billing.)
- certification of public keys,

and the following *on-line* TTP services,

- key distribution,
- time-stamping for notarisation,
- revocation support (it is not clear whether this will be a part of the demonstration).

As explained above, symmetric key management might become an issue for the TTP if, for example, advice of charge is used.