ETSI TR 103 502 V1.1.1 (2017-09)

Reconfigurable Radio Systems (RRS);
Applicability of RRS with existing
Radio Access Technologies and core networks;
Security aspects

Reference
DTR/RRS-0314

Keywords
radio, safety, security

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - NAF 742 C
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° 7803/88

Contents

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to the present document may have been declared to ETSI. The information pertaining to these essential IPRs, if any, is publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the ETSI Web server (https://ipr.etsi.org/).

Pursuant to the ETSI IPR Policy, no investigation, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

Foreword

This Technical Report (TR) has been produced by ETSI Technical Committee Reconfigurable Radio Systems (RRS).

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the ETSI Drafting Rules (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Executive summary

The introduction of RRS capability is shown not to inhibit the provision of the security mechanisms of existing radio access technologies. The reason, shown in the present document, is that the security capabilities for common radio technologies (e.g. LTE/SAE, IEEE 802.11™) are present at layers 2 and higher in the OSI protocol stack, whereas the novel features of RRS apply to the means to provision layer 1. It is highlighted however that a Radio Application addresses a complete protocol stack and provision of all layers of the protocol stack in a single software package may require special provisions in the Reconfigurable Equipment to enable full network communication.

1 Scope

The present document shows a mapping of existing Radio Access Technologies (RATs) to the Reconfigurable Radio System (RRS) model in order to identify missing security requirements, in particular identify the boundary of an RRS Radio Application with regard to the security functions present in existing RAT. Recognizing that a RAT is not bound to a single link but may be supported by functions in the network the present document also considers the role of core networks in supporting any triggering of the Reconfigurable Equipment to reconfigure itself using a push mechanism.

2 References

2.1 Normative references

Normative references are not applicable in the present document.

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication ETSI cannot guarantee their long term validity.

The following referenced documents are not necessary for the application of the present document but they assist the user with regard to a particular subject area.

[i.1] ETSI TS 133 401: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; 3GPP System Architecture Evolution (SAE); Security architecture (3GPP TS 33.401)".

[i.2] ISO/IEC 7498-1:1994 "Information technology -- Open Systems Interconnection -- Basic Reference Model: The Basic Model".

[i.3] ETSI TR 102 945: "Reconfigurable Radio Systems (RRS); Definitions and abbreviations".

[i.4] ETSI EN 303 095: "Reconfigurable Radio Systems (RRS); Radio Reconfiguration related Architecture for Mobile Devices".

[i.5] Len Sassaman, Meredith L. Patterson, Sergey Bratus, Michael E. Locasto, Anna Shubina: "Security Applications of Formal Language Theory".

[i.6] Noam Chomsky: "On certain formal properties of grammars", Information and Computation/information and Control, vol. 2, pp. 137-167, 1959.

[i.7] Michael Sipser: "Introduction to the Theory of Computation", Second Edition, International Edition, Thompson Course Technology, 2006.

[i.8] Seymour Ginsburg and Sheila Greibach: "Deterministic context free languages", in Proc. 6th Symp. Switching Circuit Theory and Logical Design, 1965, pp. 203-220.

[i.9] Dan Kaminsky, Meredith L. Patterson and Len Sassaman: "PKI Layer Cake: New Collision Attacks Against the Global X.509 Infrastructure".

[i.10] Travis Goodspeed, Sergey Bratus, Ricky Melgares, Rebecca Shapiro and Ryan Speers: "Packets in Packets: Orson Welles' In-Band Signaling Attacks for Modern Radios", 5th USENIX Workshop on Offensive Technologies, August 2011.

[i.11] Open Mobile Alliance™. OMA-AD-Push-V2-3: "Push Architecture".

NOTE: Available at http://www.openmobilealliance.org/.

[i.12] WAP Forum™ WAP-230-WSP: "Wireless Session Protocol".

NOTE: Available at http://www.openmobilealliance.org/.

[i.13] WAP Forum™: "Wireless Application Protocol".

NOTE: Available at http://www.openmobilealliance.org/.

[i.14] IETF RFC 2616: "Hypertext Transfer Protocol -- HTTP/1.1".

[i.15] IETF RFC 793: "Transmission Control Protocol".

[i.16] IETF RFC 3261: "Session Initiation Protocol".

[i.17] ETSI TS 123 228: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; IP Multimedia Subsystem (IMS); Stage 2 (3GPP TS 23.228)".

[i.18] ETSI TS 122 146: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; Multimedia Broadcast/Multicast Service (MBMS); Stage 1 (3GPP TS 22.146)".

[i.19] Open Mobile Alliance™: "OMA Mobile Broadcast Services".

NOTE: Available at http://www.openmobilealliance.org/.

[i.20] 3GPP2 C.S0070-0: "Broadcast Multicast Services (BCMCS) Codecs and Transport Protocols".

[i.21] ETSI EN 302 304: "Digital Video Broadcasting (DVB); Transmission System for Handheld Terminals (DVB-H)".

[i.22] ETSI TS 123 041: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); Technical realization of Cell Broadcast Service (CBS) (3GPP TS 23.041)".

[i.23] ETSI TS 142 009: "Digital cellular telecommunications system (Phase 2+); Security aspects (3GPP TS 42.009)".

[i.24] ETSI TS 125 302: "Universal Mobile Telecommunications System (UMTS); Services provided by the physical layer (3GPP TS 25.302)".

[i.25] ETSI TS 123 040: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); Technical realization of the Short Message Service (SMS) (3GPP TS 23.040)".

[i.26] ETSI TS 123 204: "Universal Mobile Telecommunications System (UMTS); LTE; Support of Short Message Service (SMS) over generic 3GPP Internet Protocol (IP) access; Stage 2 (3GPP TS 23.204)".

[i.27] ETSI TS 124 011: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Point-to-Point (PP) Short Message Service (SMS) support on mobile radio interface (3GPP TS 24.011)".

[i.28] ETSI TS 144 003: "Digital cellular telecommunications system (Phase 2+) (GSM); Mobile Station - Base Station System (MS - BSS) Interface Channel Structures and Access Capabilities (3GPP TS 44.003)".

[i.29] ETSI TS 143 020: "Digital cellular telecommunications system (Phase 2+) (GSM); Security related network functions (3GPP TS 43.020)".

[i.30] ETSI TS 144 064: "Digital cellular telecommunications system (Phase 2+) (GSM); Mobile Station - Serving GPRS Support Node (MS-SGSN); Logical Link Control (LLC) Layer Specification (3GPP TS 44.064)".

[i.31] ETSI TS 124 008: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Mobile radio interface Layer 3 specification; Core network protocols; Stage 3 (3GPP TS 24.008)".

[i.32] ETSI TS 144 018: "Digital cellular telecommunications system (Phase 2+) (GSM); Mobile radio interface layer 3 specification; GSM/EDGE Radio Resource Control (RRC) protocol (3GPP TS 44.018)".

[i.33] ETSI TS 124 341: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; Support of SMS over IP networks; Stage 3 (3GPP TS 24.341)".

[i.34] GSM Association IR.92: "IMS Profile for Voice and SMS".

[i.35] ETSI TS 131 102: "Universal Mobile Telecommunications System (UMTS); LTE; Characteristics of the Universal Subscriber Identity Module (USIM) application (3GPP TS 31.102)".

[i.36] ETSI TS 131 101: "Universal Mobile Telecommunications System (UMTS); LTE; UICC-terminal interface; Physical and logical characteristics (3GPP TS 31.101)".

[i.37] IEEE 802.11™: "IEEE Standard for Information technology -- Telecommunications and information exchange between systems Local and metropolitan area networks -- Specific requirements -- Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications".

[i.38] ETSI TS 122 016: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); LTE; International Mobile station Equipment Identities (IMEI) (3GPP TS 22.016)".

[i.39] ETSI TS 123 003: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); Numbering, addressing and identification (3GPP TS 23.003)".

[i.40] ETSI TR 103 087: "Reconfigurable Radio Systems (RRS); Security related use cases and threats in Reconfigurable Radio Systems".

[i.41] IEEE 802.11p™: "IEEE Standard for Information technology -- Local and metropolitan area networks -- Specific requirements -- Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications Amendment 6: Wireless Access in Vehicular Environments".

[i.42] ETSI TS 133 102: "Digital cellular telecommunications system (Phase 2+); Universal Mobile Telecommunications System (UMTS); 3G security; Security architecture (3GPP TS 33.102)".

[i.43] GSM Association TS.06 (DG06): "IMEI Allocation and Approval Guidelines, Version 6.0".

NOTE: Available at http://www.gsma.com/newsroom/wp-content/uploads/2012/06/ts0660tacallocationprocessapproved.pdf.

[i.44] IEEE 801.1X™: "IEEE Standard for Local and metropolitan area networks - Port-Based Network Access Control".

3 Definitions and abbreviations

3.1 Definitions

For the purposes of the present document, the terms and definitions given in ETSI TR 102 945 [i.3] apply.

3.2 Abbreviations

For the purposes of the present document, the abbreviations given in ETSI TR 102 945 [i.3] and the following apply:

2G	Second Generation
3G	Third Generation
ASN	Abstract Syntax Notation
BCMCS	Broadcast and Multicast Service
CBCH	Cell Broadcast CHannel
CBS	Cell Broadcast Service
CTCH	Common Traffic Channel
DCCH	Dedicated Control Channel
IMS	IP Multimedia Subsystem
LLC	Logical Link Control
MM	Mobility Management
OMA	Open Mobile Alliance™
P-CSCF	Proxy-Call Service Control Function
P-GW	PDN GateWay
PDN	Packet Data Network
RR	Radio Resource
SACCH	Slow Access Control Channel
SDCCH	Standalone Dedicated Control Channel
SIP	Session Initiation Protocol
SGSN	Service GPRS Support Node
SMS-SC	SMS Service Centre
SQL	Structured Query Language
WAP	Wireless Application Protocol
XMPP	eXtensible Messaging and Presence Protocol

4 OSI stack mapping to RRS

4.1 OSI protocol stack overview

The conventional 7-layer model for Open Systems Interconnection defined in ISO/IEC 7498-1 [i.2] allocates particular functionality to each of the abstract layers. A strict interpretation of the OSI model is that radio technologies apply to layer 1 only, as the nature of the physical media is only apparent at that layer. However, whilst a radio signal is physical the Radio Access Technology is more complex. The physical layer is where the modulation and RF stuff is done, the link layer, layer-2 in the OSI model, is where the logical connectivity is built up and this is carried through to the network layer, layer-3 in the OSI model, where the radio device becomes connected to the wider network. Thus for conventional telecommunications modelling the broadly accepted model has been slimmed down to only 3 layers:

- Layer 1 - physical media, the radio path.
- Layer 2 - the link layer that offers a reliable connection across the radio path between 2 points.
- Layer 3 - the network layer that builds the terminal into a wider network of devices.

The nature of the transmission environment dictates to a very large extent the means by which data can be transmitted and for radio the environment is hostile. The hostility itself takes a number of forms but the corollary of wavelength and transmission range, the impact of noise, the impact of fading, and so on all are addressed across the layers of the protocol stack. In addition the radio resource is highly regulated with regulation defining most of the physical characteristics, e.g. radiated output power, bandwidth, adjacent channel interference restrictions (i.e. how much power can be broadcast in adjacent channels), and for fixed sites also addresses such things as antenna height and antenna gain. Any implementation of RRS has to give assurance the regulation is not abused and the means of enforcing any claims for equipment to the Declaration of Conformity (DoC) has to be added to the security model of the radio device when that device is an RRS capable of being modified on the fly to support one or more Radio Access Technologies (RATs).

In terms of commonly applied security functions the assignments in table 1 are generally applied.

Table 1: OSI mapping of security and transmission functions

n	Name	Basic function for transmission	Commonly applied security functions
4++	Application domain	Everything else	Application security (CIA paradigm)
3	Network layer	Routing, multi-node networking	Authentication, Key management (as part of mobility management in cellular networks)
2	Link Layer	Logical channels creating a reliable link (LLC)	Encryption, Transmission packet integrity, Transmission FEC
1	Physical layer	The radio bits	PHYLAWS, QKD, QE and similar physical layer security

4.2 OSI layers and security mechanisms

4.2.1 Threats and countermeasures

The technical domain of security is often described in terms of the CIA paradigm (Confidentiality Integrity Availability) wherein security capabilities are selected from the CIA paradigm to counter risk to the system from a number of forms of cyber attack. The common model is to consider security in broad terms as determination of the triplet {threat, security-dimension, countermeasure} such that a triple such as {interception, confidentiality, encryption} is formed. The threat in this example being interception which risks the confidentiality of communication, and to which the recommended countermeasure (protection measure) is encryption.

The very broad view is thus, for communications systems, that security functions are there to protect user content from eavesdropping (using encryption) and networks from fraud (authentication and key management services to prevent masquerade and manipulation attacks). Where encryption is applied to content it is applied prior to radio encoding but the key assignment scheme may be optimized for the radio framing structure.

4.2.2 Radio link specificity of countermeasures

Countermeasures should always be designed to fit the context in which they are applied. On the understanding that radio links are inherently unreliable (bit error rates of up to 10 % are common) any security process applied has to be robust in the face of error. For TDM schemes (e.g. TETRA, GSM) the slot and frame numbering may be used to give a time invariant parameter that is used to drive certain cryptographic modes (e.g. in GSM the 22 bit frame number is used as an input with Kc (the traffic encryption key) to the key stream generator (A5/X algorithm) to generate a variable number of key stream bits (limited to 114 bits for single half slot encryption). For cellular systems codes specific to cell sites, colour codes, are also used to give radio path differentiation and in high security systems such as TETRA site specific keys are used as key modifiers to give assurance of cryptographic separation between cells.

NOTE 1: In TETRA and evolving LTE scenarios the same plaintext may be delivered simultaneously to multiple cell sites and means have been defined in TETRA such that an exploit of a single site cannot be transposed to another site.

The means by which media (spectrum) is shared has some impact on the viability of security measures. The TDMA/FDMA schemes are well known in most simple digital cellular radio models (e.g. GSM, DECT, TETRA). In collision awareness based systems such as those used in WLAN (IEEE 802.11™ [i.37]) the model is also relatively simple. For simple FDMA/TDMA and collision avoidance systems at any one point in time at any chosen frequency only one user holds the right to transmission. The more complex modes in CDMA however break that model such that at any point in time at any chosen frequency many users may be using the same resource but the codes used by each user, as they are orthogonal to each other, provide mathematical and thus physical separation. The challenge for FDMA systems is to give assurance of frequency synchronization between Alice and Bob, for TDMA systems to give assurance of time (clock) synchronization between Alice and Bob, and for CDMA systems to give assurance of power synchronization amongst all users of the channel. For CDMA as each code pair provides a level of rejection of any other code pair that is translated to an equivalence of co-channel interference rejection then the received power from radios in a single cell site have to be aligned to be within a few dB of each other where the granularity of power control is sufficient for the code-based signal rejection to operate.

In CDMA systems the codes used by each ME are assumed to be orthogonal to each other but this orthogonality applies at zero phase offset. In a multi-path environment the same encoded signal is received multiple times at small phase offsets and errors may arise through insufficient rejection of the interfering multipath signal. The level of signal rejection is dependent on the level of processing gain, the ratio of the spreading code rate to the signal rate. For a nominal 1 kHz signal spread to 100 kHz the processing gain is nominally 20 dB (in other words an interfering signal of 1 kHz in the received 100 kHz will be spread across the entire 100 kHz giving a 20 dB interference rejection).

The modulation scheme in use has a significant impact on overall performance, particularly in the presence of interference. The normal convention of modulating a symbol where a transition in the modulated signal indicates the symbol transmitted is only accurate in transmission if the transition can be accurately measured. For a software based implementation of modulation there are a number of real time aspects to consider therefore as the processing (i.e. encoding or decoding of the modulated signal) has to respect the real time transitions. Any phase distortion introduced by processing delay across the I/Q channel is visible as a significant reduction in signal to noise.

NOTE 2: In a homodyne baseband receiver the I and Q channels are assumed to be exactly 90° out of phase. An error of even a few tenths of a degree caused either by physical path or processing inequalities introduces noise and may give rise to an increase in inter-symbol-interference. Similar errors may be introduced by DC offset. Frequency and phase correction fields used in the transmission bursts of most digital communication systems will compensate for any I/Q channel imbalance.

4.3 Core elements of the RRS model

The current model of RRS applies only to the physical layer. This then would apply to the modulation and to a small extent the framing synchronization of a RAT. However, as noted above a RAT may be seen as covering layers 1 and 2 as there is a case for stating that the TDM, FDM or CDM access modes and channelization are defined at layer 2. Similarly the overall performance of a radio link is not defined solely by the physical layer.

The RVM enables an RA to choose one among multiple available protection classes for code to be executed on the RVM as well as a protection class for the RF front-end. Depending on the combination of chosen RF & RVM protection classes, the required re-certification process of the software reconfigurable radio platform will be more or less complex. The basic principle is illustrated in figure 1.

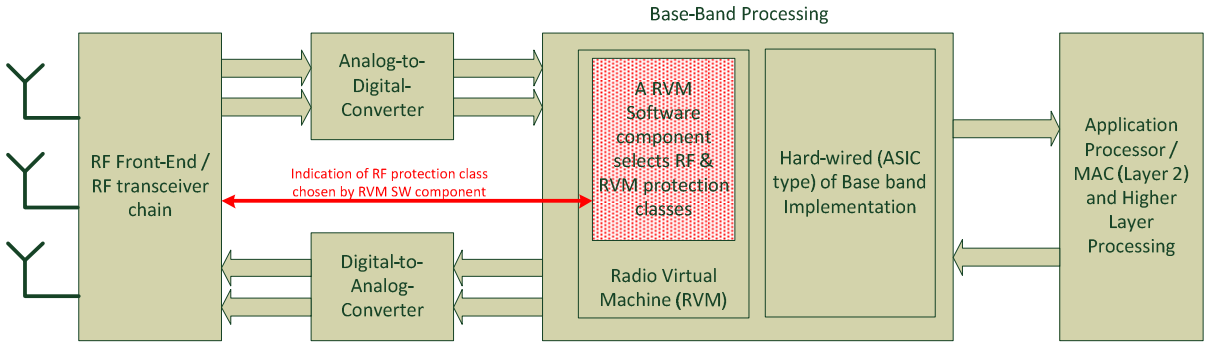

Figure 1: A typical radio equipment architecture comprising an RVM Software Component selecting RF and/or RVM protection class(es) (from ETSI EN 303 095 [i.4])

A typical radio equipment architecture includes an RF Transceiver chain, Analog-to-Digital converters, Digital-to-Analog converters, Base Band Processing, etc. An RVM controls RF Transceiver chain, in particular for selection of an RF Protection Class.

4.4 Applicability of RVM to radio terminal computing

Whilst it is possible to perform general purpose computing using the RVM concept it is not optimized for many of the higher layer protocol and other security capabilities. The state machine based protocols of ISDN-era telephony can be emulated using the RVM although translation of existing code to the RVM model may be problematic. The concern, and security risk, is that existing code, and hardware accelerated firmware, has been mature for many technologies for some years and moving such code to a new platform may introduce performance or functionality uncertainty.

As stated in clause 4.3 the RVM controls the RF Transceiver chain, thus the MAC and higher layer processing is outside the scope of the RVM.

From a security perspective there are many issues and risks from the means used to programme any function. The role of compilers, parsers, and operating systems combine to make absolute claims as to the overall security of executable code difficult. For some elements in the chain strong assertions of functionality can be made. One approach that may be considered if that of Language Theoretic Security (LangSec) described in some detail in annex A, and which is summarized as the application of formal language theory to improve the security of complex systems. LangSec addresses security on interfaces by treating the expected input as a formal language (i.e. as a formal definition of valid input for the interface) and the respective input-handling routines as a recognizer for that language. Under this formalism, the security property of an implementation under its input is dependent on the complexity of the language grammar. For the vast majority of cases the role of the parser to make "correct" interpretations of input is key, thus the goal in an RRS and RVM environment is for the parser to be able to identify valid input and to reject invalid input.

4.5 Notification of Radio App availability

Application domain notification may be enabled by general purpose messaging protocols operating at the application layer of the OSI stack. The overall model is that of an Over The Top (OTT) service and is often vendor specific. As the notification capability will run in the application layer there is no direct impact on the radio elements.

A detailed description of notification technologies is given in annex B.

5 Security provisions in 3GPP SAE and LTE

5.1 Security architecture for 3GPP

Figure 2 gives an overview of the complete security architecture of 3GPP SAE and LTE.

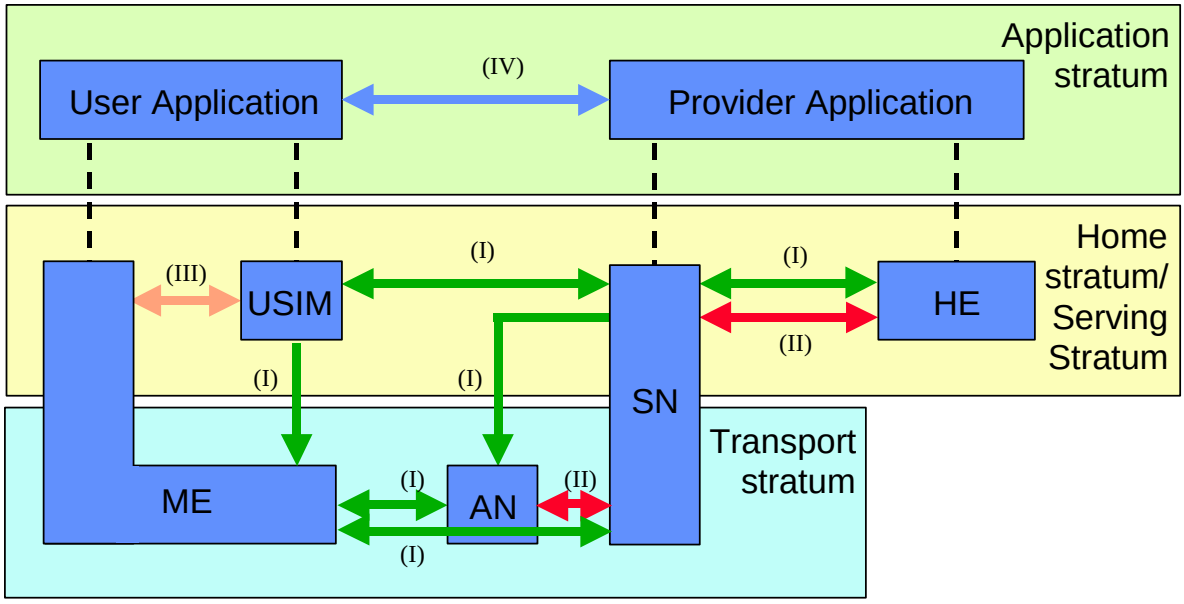

Figure 2: Overview of the security architecture (from ETSI TS 133 401 [i.1])

Five security feature groups are defined. Each of these feature groups meets certain threats and accomplishes certain security objectives:

- **Network access security (I) (NAS):** the set of security features that provide users with secure access to services, and which in particular protect against attacks on the (radio) access link.

- **Network domain security (II) (NDS):** the set of security features that enable nodes to securely exchange signalling data, user data (between AN and SN and within AN), and protect against attacks on the wireline network.
- **User domain security (III) (USD):** the set of security features that secure access to mobile stations.
- **Application domain security (IV) (ADS):** the set of security features that enable applications in the user and in the provider domain to securely exchange messages.
- **Visibility and configurability of security (V) (VCS):** the set of features that enables the user to inform himself whether a security feature is in operation or not and whether the use and provision of services should depend on the security feature.

The elements considered of most relevance to examination in an RRS context from the security model given in ETSI TS 133 401 [i.1] are those representing NAS, NDS, UDS and VCS. It is asserted that the application stratum is sufficiently divorced from the radio layers to not be designed with any consideration given to the radio system itself.

For RRS the VCS element may become critical if the selection of a particular RAT is dependent on the level of communications security it can provide.

5.2 Radio channels in 3G SAE/LTE

The logical radio channels of 3G SAE/LTE are defined at layer 2, the physical radio channels, that is the combination of radio frequency (FDM) and slot number (TDM) are defined at layers 1 and 2 and the TDMA structures should then be encoded outside of the RVM.

5.3 Security functions mapping to radio in 3G SAE/LTE

5.3.1 General overview

The security features of 3G SAE/LTE are defined for application at layers 2 and 3 of the protocol stack. For the present document there is no barrier to RRS (providing radio specific services at layers 1 and 2) to support the communications and security protocols of 3G SAE/LTE.

5.3.2 u-SIM and identity management

5.3.2.1 Overview

The 3G LTE device makes use of 2 key identifiers:

- IMEI (International Mobile Equipment Identifier), optional to provide.
- IMSI (International Mobile Subscriber Identity), mandatory to provide.

The latter is closely linked to the MSISDN used in telephony but not all IMSIs have a related MSISDN. The structure of the identifiers, and their role in protocol, is of interest to RRS for a number of reasons.

5.3.2.2 Provision of subscriber identity

The expectation in LTE (and for 2G and 3G communications systems in general) is that the core subscriber identity and primary secret is delivered on a Subscriber Identity Module (SIM) which is a dedicated smart card. The specification of the uSIM (for UMTS applications) is given in ETSI TS 131 102 [i.35] and provides the interface for authentication and key management. The electrical and other physical connections are defined in ETSI TS 131 101 [i.36]. The SIM acts as the distribution media for the IMSI and manages the security functions of authentication and key management.

When the IMSI is authenticated the uSIM delivers the encryption and integrity keys to the radio platform. The IMSI is assigned by the national numbering authority and assigned to the operator for allocation to devices (via the SIM).

Whilst an external protocol is unable to determine if the IMSI is provisioned on a hardware SIM (either on an UICC (removable) or eUICC (soldered to the UE, supports remote provisioning)) or by some other means the security experts in 3GPP SA3 only support the provision of the IMSI and Ki (secret key used to authenticate the IMSI to the network) via a hardware SIM. The impact for RRS is that any RRS hardware platform used for a terminal requires support of a hardware SIM card reader (where RRS is applied at the cell site no SIM is required).

NOTE: The further development of a "soft" SIM may in due course alter the requirement for a hardware SIM card reader but for the purposes of the present document the assumption is that only a hardware SIM will be allowed for provisioning of the IMSI, Ki and the authentication and key management algorithms.

5.3.2.3 Provision of device identity

For RRS the IMEI structure has be identical to that of any other 3G device and allocated in an identical manner. It is noted that in 3G/LTE there is no means for IMEI recovery over the radio interface and there are no conformance tests for allocation and availability of IMEI, but the management of devices at retail points requires, for most carriers, that the IMEI is recorded prior to sale in order to assist in market controls (fraud prevention, fake handset detection and so forth).

The IMEI is a composite identifier defined as follows:

- IMEI = TAC || Serial number || Check code (15 digits)

Where TAC = RBI || METI (Reporting Body Identifier || Mobile Equipment Type Identifier), and the operator || indicates concatenation. Thus IMEI = RBI || METI || Serial number || Check code, i.e. it contains 3 items of information that may be asserted in some formal process. In all current implementations there is no cryptographic assertion of any of these elements. There is a requirement for the IMEI to be tamper proof and immutable after the ME's final production process.

NOTE: The IMEI and the TAC element of it can be decoded to identify the constituent elements (manufacturer, model and authorising body) using a number of available tools, however the present document does not endorse any of them and leaves it to the reader to investigate for themselves.

The allocation and structure of the IMEI is defined for operators by the GSMA in TS-06 (DG06) [i.43].

In addition there is an extension to IMEI defined that allows for carriage of the software version number:

- IMEISV = TAC || Serial number || SVN (16 digits)

For the IMEI, calculation of the check code is optional but if added has to be encoded using the Luhn Formula as defined in ISO/IEC 7812-1 [i.2] (and in ETSI TS 122 016 [i.38]).

The purpose of the IMEI is described in GSMA TS-06 as follows (the text has been edited for presentation in the present document thus the mandates given in the referenced document have been removed):

- The IMEI uniquely identifies each individual unit of ME.
- As per ETSI TS 123 003 [i.39], ETSI TS 122 016 [i.38], the IMEI has to be provisioned such that it is not possible to be changed after the ME's final production process. This means that it is able to resist tampering, i.e. manipulation and change, by any means (e.g. physical, electrical and software).
- Where repairs necessitate the need to replace the components that contain the IMEI a new IMEI has to defined and used.
- Where a ME has variants that operate in other bands/modes then the ME should be constructed in such a way so that it is not possible to interchange components to permit the IMEI being swapped between the variants.

For the purposes of RRS the final bullet point is crucial as the opportunity presented by RRS is to have a single hardware device present itself as multiple modes or RATs.

The device class, represented as the Mobile Equipment Type Identifier, is attested to by the manufacturer and maintained by the Reporting Body. The METI identifies the forms of RAT assigned to the ME. The ME is the specific instance of Device from the generic model shown in figure 3.

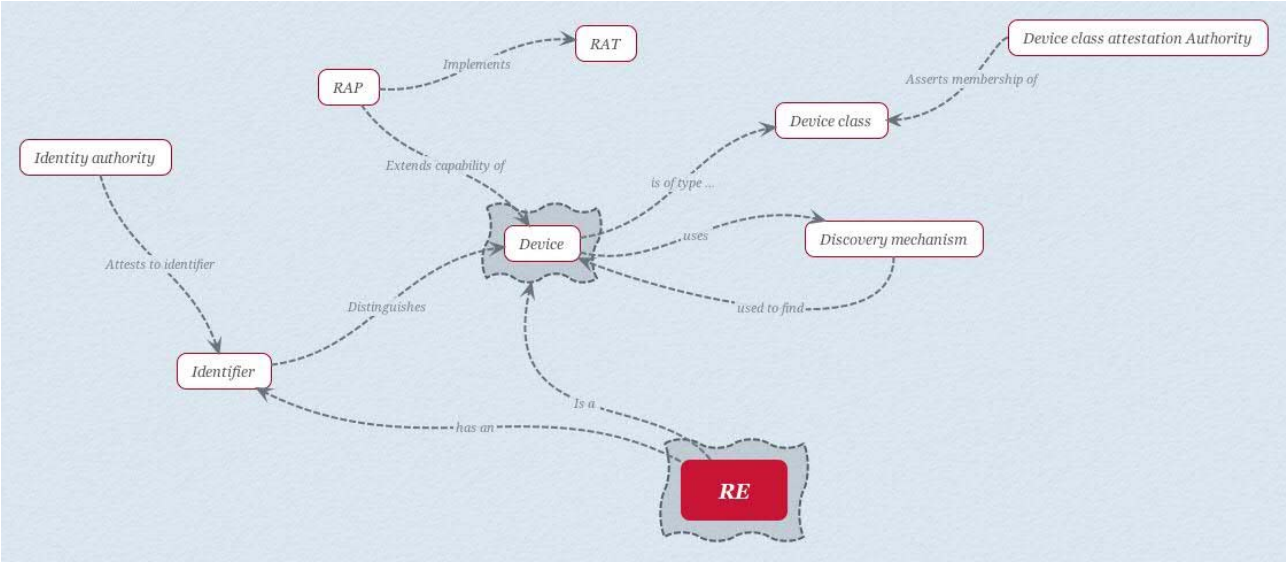

Figure 3: Radio Equipment device identifiers and authorities

The device class attestation authority, having responsibility for the METI, is also expected to be responsible for the assurance of METI compliance statements (that may be asserted in the DoC).

Table 2: Mapping of IMEI to DoC attestations

IMEI element	RRS, or DoC, identity element	Notes
RBI	N/A	There is no direct equivalence to the RBI in the DoC custody chain. As the DoC is an element of CE marking it may be considered that the RBI is somewhat equivalent to the market surveillance authority.
METI	RE Type	The METI is managed by the manufacturer and maintained by the reporting body. This is at slight variance from the definition of RE-Type from ETSI TR 103 087 [i.40] wherein the definition indicates that the RE-Type is a composite of OEM ID and the HW Platform ID in such a manner that the RE Type uniquely identifies the capabilities of the RE.
Serial Number	RE Serial	The definition of RE-Serial from ETSI TR 103 087 [i.40] is given as "serial number uniquely identifying a device within the RE Type". This is close in spirit to the intent of the Serial Number component of the IMEI and should be considered as directly equivalent.

From the perspective of the security model for the Declaration of Conformance described in ETSI TR 103 087 [i.40] there is no simple mapping of the device class and the device class attestation authority.

6 Security provisions in IEEE 802.11™ systems

6.1 System overview

By default IEEE 802.11™ [i.37] operates in clear as part of the overall purpose to define a single medium access control (MAC) for several distinct physical layers (PHY) to allow wireless connectivity for fixed, portable, and moving stations (STAs) within a local area network (LAN). Thus IEEE 802.11™ [i.37] operates at layers 1 and 2 of the ISO interconnection model (see clause 4). The layer 3 provisions of IEEE 802.11™ [i.37] for security credential management are either handed to the provisions of IEEE 802.1x in order to manage security associations, or are managed manually by the end user.

IEEE 802.11™ [i.37] provides several cryptographic algorithms to protect data traffic, including: WEP, TKIP, and CCMP. WEP and TKIP are based on the ARC4 algorithm (the publicly available version of the RC4 streaming cipher), and CCMP is based on the advanced encryption standard (AES). The IEEE 802.11™ [i.37] specification defines means for STAs to select the algorithm(s) to be used for a given association. It is noted though that in many installations the selection of connection mode may limit the selection. For example many home routers are enabled for infrastructure mode only and for CCMP (Wi-Fi Protected Access®) by default.

There have been a number of security issues identified with implementations of the ARC4 algorithm in IEEE 802.11™ [i.37] that make recommendations of using WEP ill advised, which were provisionally addressed in Wi-Fi Protected Access® and WPA2™ and formally addressed in the TKIP protocols. Whilst it is not the purpose of the present document to select security mechanism it is certainly not recommended that any IEEE 802.11™ [i.37] RAT provided in the context of RRS should support WEP.

NOTE: The common term used in consumer circles to refer to IEEE 802.11™ [i.37] as Wi-Fi is misleading as the Wi-Fi Alliance does (WFA) not endorse all modes of IEEE 802.11™ [i.37], and in fact some have been designed for specific application areas. This includes restriction of the IEEE 802.11p™ [i.41] mode to vehicular ITS.

CCMP is an AES based counter mode encryption and integrity check sum protocol. In addition to the encryption mode shown in figure 4 the CCMP mode adds a Cipher Block Chaining Message Authentication Code (CBC-MAC) to each packet.

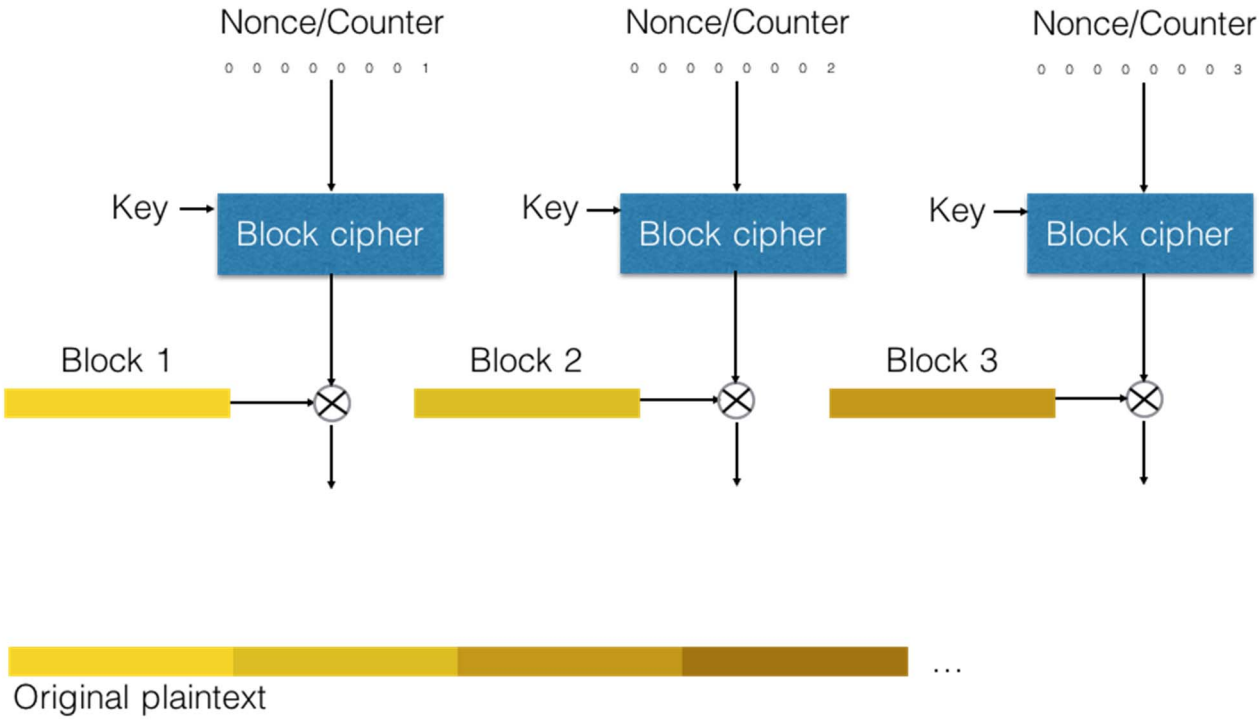

Figure 4: Encryption Counter mode

6.2 IEEE 802.11™ key management systems

There are a number of key management systems available to IEEE 802.11™ [i.37] that are closely aligned to the encryption modes (WEP, TKIP, CCMP). All of the key management facilities in IEEE 802.11™ [i.37] exist at layers 2 and to a smaller extent at layer 3 and for the purposes of the present document have no impact on RRS.

6.3 IEEE 802.1X key management systems

The IEEE 802.1X [i.44] key management protocols (for port based access control) operate at layer 3 and thus have no direct impact on RRS.

7 Physical layer security provisions in RRS

Whilst there are no standardized or openly marketed mechanisms supporting radio based physical layer protection it cannot be stated with certainty if RRS is applicable. However on the understanding that physical layer provisions provide a key (e.g. the PHYLAWS project) and that the actual encryption process uses standard mechanisms there is no barrier to RRS implementing physical layer cryptographic protection.

For PHYLAWS the secrecy code is used in specialization of standard channel encoding (i.e. Reed-Muller, Low-Density Parity-Check (LDPC)) but examples are given for modification of existing LTE and Wi-Fi algorithm modes to use a physical layer derived key as a key modifier.

Annex A:
Language-theoretic security in RRS

A.1 Overview

A.1.1 Introduction

Language-theoretic security, also known as LangSec, is the application of formal language theory to improve the security of complex systems. It addresses security on interfaces by treating the expected input as a formal language (i.e. as a formal definition of valid input for the interface) and the respective input-handling routines as a recognizer for that language. Under this formalism, the security property of an implementation under its input is dependent on the complexity of the language grammar. In particular, for the parser to be able to identify all invalid input, it has to be of computational strength at least equivalent to the complexity of the input language grammar, and there exist grammar types that cannot give secure parser implementations.

The underlying assumption is that, since the security of a system is defined by the operations that can and cannot occur under all possible input, a system whose valid input cannot be determined cannot be made secure. More specifically, the computational power required to parse input increases with the complexity of the input language, until recognizing valid input becomes an undecidable decision problem. Undecidability puts the burden on the implementer to make assumptions so that the input recognition becomes a decidable problem. Such assumptions, in turn, make it impossible to detect all invalid inputs.

The applications of language-theoretic security are of critical importance to the security of computer systems since external interfaces (i.e. input language recognizers) are the primary vessel through which other vulnerabilities in the implementations can be exposed. A corollary is that ad-hoc programming of input handling does not yield good security assurance.

Language-theoretic security does not only consider the input to computer programs, but also their state. That is, it treats input languages both on the syntactic and semantic levels. This also applies to protocols. When the semantic of an input language is ambiguous, independent implementations of the same protocol may have been designed with different assumptions to resolve an ambiguity (or the same implementation may make different assumptions as a producer or as a consumer). This discrepancy introduces vulnerabilities which can be exploited to obtain results that are not normally provided by the protocol (that is, expected to be provided) or, as with input recognition, to put the implementation under an unexpected state.

A.1.2 Weird machine

The term "weird machine" designates a set of states and transitions which, while valid within a system, are not part of its expected behaviour - for the designer or the programmer. From a security standpoint, the weird machine represents a compromise of the system is executes on.

The existence of the weird machine stems from incorrected assumptions about the computational nature of the system and of input handling. Specially crafted input that should have been rejected triggers a transition from a valid state to an unexpected state. This is commonly referred to as an exploit. If one considers the hypothesis that a machine B is hidden inside a machine A, then the exploit allowing a transition from A to B proves the existence of B.

It is important to note that the weird machine comes with all the original functionalities of the system (e.g. the operating system and its libraries) augmented with those of the input that caused the transition to the weird machine. In other words, an adversary can borrow the original functionalities of a system, and program on top of those, by means of a specially crafted input.

A.1.3 Grammar type, computational complexity and decidability

The purpose of formal verification is to prove certain properties of a system, namely:

- safety (nothing bad happens); and
- liveness (something good eventually happens).

If every execution path of the system satisfies a given property, the system is safe (or live) with respect to that property. In the general case, such verification is an undecidable problem. However, when the verification is restricted to a specific subset of the system, such as an interface, or if the whole system is simple enough, it may become decidable depending on the computational complexity of the underlying model.

Language-theoretic security focuses on two specific decision problems, among others, which are:

- the recognition problem (safely accepting and handling input);
- the equivalence problem (identical interpretation of a message between two end-points). See clause A.1.4 for more information.

A decision problem is undecidable when it is known that no single algorithm can be constructed, that always lead to a correct yes-or-no answer - that is to say, while valid input can be recognized, it is not known whether the parser will halt on invalid input, or run indefinitely.

An input language is determined (generated) by its grammar. The Chomsky hierarchy classifies formal grammars according to their expressive power and establishes a correspondence between the grammar and the minimum strength of the computational model necessary to parse the grammar:

- regular languages form the weakest class; they need only a finite state machine and can be parsed with regular expressions;
- unambiguous context-free grammars allow recursively nested data structures; they need a deterministic pushdown automaton (adding a stack to the limited memory of a finite state machine);
- ambiguous context-free grammars need a non-deterministic pushdown automaton to account for ambiguity;
- context-sensitive grammars require a linear-bounded automaton;
- recursively enumerable grammars require a Turing machine.

One observation is that a given automaton can decide an equivalently powerful or less powerful language, but cannot decide a stronger language. This is the reason why regular expressions are often insufficient for input validation.

For the recognition problem, all grammar classes are decidable (they will always halt, in an "accept" or "reject" state), except recursively enumerable languages [i.6], [i.7]. For the latter, the recognizer halts in an "accept" state for all strings in the language, but may either reject or fail to halt on input that is not in the language.

The equivalence problem is decidable only for regular and deterministic (unambiguous) grammars [i.8].

Table A.1 summarizes the decidability of the recognition and equivalence problems relative to the language grammar.

Table A.1: Chomsky hierarchy and computational strength of the parser

Grammar type	Computational model	Recognition Problem	Equivalence problem
Regular	Finite state machine	Decidable	Decidable
context-free (unambiguous)	Deterministic pushdown automata		
context-free (ambiguous)	Non-deterministic pushdown automata		Undecidable
context-sensitive	Linear bounded automata		
recursively enumerable	Turing machine	Undecidable	

In a nutshell, the difficulty of safe input handling is due to the complexity in the design of the input grammar, which introduces problems that can be hard or impossible to solve. While common language that describe structured data - not computation - are not Turing complete, imperative languages with conditional branching and the ability to change an arbitrary amount of memory are. If the recognizer cannot decide on the language, no amount of testing effort can compensate for the weakness of the input handling.

A.1.4 Semantic security and computational equivalence of protocol endpoints

When two components communicate, the semantic of the exchanged message is as important as its syntax. An implementation is usually more than its interfaces, thus the message will continue to affect the internal state of the implementation beyond the input handling routines - that is, after the message has been parsed. When the source and destination implementations do not agree on the meaning of a given message (the semantic), the destination may perform actions that the source did not expect.

This can be formalized by considering the equivalence of automata used in the source encoder and the destination decoder. Consider the encoding function E of the source and the decoding function D of the destination over a language L. The automata are equivalent when for any message M of L, D(E(M)) is equal to M. Subsequently, if one considers that the implementation operates a conditional function C, over the input message, then the composition D(E(M)) • C(D(E(M))) leads to a valid state at the destination (assuming a formal verification of the behaviour of C).

Determining whether two automata implementing a given grammar are equivalent is a decidable problem when the grammar is regular or unambiguous (deterministic) context-free. For any other grammar type, the problem is undecidable. In such case an attacker may take advantage of the discrepancies between the two automata by:

- finding a message M' of L for which D(E(M') is not equal to M'; or
- finding a message M' of L which is not in the subset of messages the source can generate, but is in the subset of messages the destination can operate on (which are both subsets of L).

In such case the composition D(E(M')) • C(D(E(M'))) would lead to an invalid state which is the entry point to a weird machine hidden in the implementation.

The undecidability of the equivalence problem for languages of higher complexity has been leveraged in attacks such as targeting ASN.1 encodings in X.509 certificates [i.9].

NOTE: Ensuring the decidability of the equivalence problem is orthogonal to the protection of messages at the transport level. Automata equivalence for a language L ensures that the result of any message M of L at the destination is predictable. It does not ensure the integrity of the message as intended by the source, nor authenticates the source as a legitimate peer in the communication.

A.1.5 Trustworthiness of a system as a composition of sub-systems

The trustworthiness of a system is the property of the composed system, not only of the individual elements but also of the composition itself. Doing a composition securely is a primary challenge of secure system construction. This challenge is two-fold.

Firstly, as the system is a composition of components, each component has to accept inputs from one or more interfaces. This creates an attack surface that propagates through the system starting from its external interfaces. In other words, a vulnerability that is present within an internal system component may be exploited as a malicious input transitions - and transformed - from one component to the other.

In addition, care has to be taken when composing input languages, that the result indeed expresses the expected input. A well-known example is the use of SQL within an application with user-controlled input, which spans both the application logic and the database server logic.

Secondly, when two components within a system are end-points for a given protocol, their implementation may not interpret exchanged messages in an identical manner. The threat model in that case is that unexpected result of using the protocol can be achieved from a weakness in the definition of the input grammar, resulting in the attacker gaining an advantage from the new state of the system.

NOTE: The trustworthiness of a system is relative to a set of expected properties, of which security is only one element. Other properties such a user privacy and equipment compliance may be at risk if the system enters an invalid state.

The fact that a system is a composition of subsystems has consequences on the effectiveness of security countermeasures that explicitly or implicitly rely on the validation of an input string to detect unwanted behaviour, such as a Network Intrusion Detection System or system call whitelisting. In the former case, the detection system could attempt to emulate a virtual system for which the equivalence problem to the original system can rapidly become undecidable. In the latter case, system calls are only a subset of the possible execution flow of a program, which means that a system call monitor would not be able to validate all the possible state of the program or block the entirety of illegitimate system calls. Similarly, signature-based detection (i.e. the of invalid behaviour can very often only detect a subset of all invalid behaviours.

A.1.6 Core principles

A.1.6.1 Simplicity and decidability

Language-theoretic security validates the intuitive rule that complexity is the enemy of security.

Limiting the input grammar to the simplest form necessary to fulfilling the role of an interface can significantly reduce the insecurity of an implementation, while increasing the complexity of the input language could render the handling of input unsafe, and could render systems composed of protocol end-points to susceptible to differential parse tree attacks. It follows that a requirement of system design should be to stay within the boundaries of decidable problems. Protocol complexity should be part of a risk assessment.

When there is no other choice available than using a language of high complexity (e.g. because of interoperability requirements), agreeing on a limited sub-language of the target language has the advantages of limiting known states and potentially reducing the computational model complexity so that the recognition and equivalence problems are decidable.

It is better to avoid nested grammars when possible, as the resulting composition could lead to increased complexity.

A.1.6.2 Strength of the recognizer

For the language recognizer to effectively identify all valid and invalid input, it should be of equal or superior computational strength compared to the complexity of the language it is designed to recognize.

A.1.6.3 Principle of minimal computation power

Because computation power is exposed on the interface and thus available to an attacker (which could be leveraged for resource exhaustion attacks), it is beneficial that the parser be no more computationally powerful than necessary for the parsing role. Quoting [i.5]:

"For Internet engineers, this principle can be expressed as follows:

- *a parser must not provide more than the minimal computational strength necessary to interpret the protocol it is intended to parse; and*
- *protocols should be designed to require the computationally weakest parser necessary to achieve the intended operation."*

A.1.6.4 Secure composition with parser computational equivalence

Input grammars permitting the equivalence problem to be decidable allow verification of syntactic and semantic equivalence of messages exchanged between two protocol end-points.

A.1.7 Language-theoretic approach as a tool for security auditors and adversaries

The language-theoretic approach allows for the introduction of automated verification tools for input languages. Combined with provably correct parsers generators, these can be leveraged for formal verification of implementation correctness. Since the specification of the input grammar drives the generation of the parser, automated tools allow for quick verification of a new implementation after a change in the input grammar.

In addition, the applications of language-theory for the identification of vulnerabilities can often be more efficient than plain fuzzing. While fuzzing relies on random or crafted input to identify invalid input, and may be limited in the way invalid input can be detected, the language-theoretic approach builds on the formalism of language and automata theory. One particularly potent tool for auditing and protocol analysis is the parse tree differential attack. In this attack, two different implementations of the same protocol are given identical states and input parameters. Any difference between the parse trees representing decoding operations in each of them indicates that different assumptions were made, possibly leading to an exploitable vulnerability at the level of the protocol or the implementation.

A.2 Applicability to Reconfigurable Radio Systems

Several aspects of Reconfigurable Radio Systems can benefit from the application of language-based security:

- Controlled mutability:
  - While RRS allows for mutability of the radio characteristics on the RE, there is a risk of uncontrolled mutability. Active elements of the architecture are potential weird machines that can be triggered by invalid input. This includes the Radio Application, starting from its exposure to radio input on the physical layer. A language-theoretic approach can also be applied at the physical layer in order to avoid vulnerabilities such as those exposed in [i.10].
- Internal interfaces and sandboxing:
  - The implementer of standardized RRS interfaces can apply language-theoretic principles in order to avoid exposure of internal vulnerabilities, when they exist, and better enforce isolation.
- Complement to language-based security:
  - Language-based security applies in the enforcement of security policies. In the context of RRS this also extends to other enforceable properties of the system, such as compliance policies via the machine interpretation of the policies found in the DoC (this may be derived from parsing the DoC and deriving a machine interpretable model of the DoC for policy enforcement), as well as mobility policies. Language-theoretic security complements language-based security by ensuring that the relevant interfaces can be verified with respect to their expected properties.
- Design guidelines for the integration of existing RAT:
  - As the technologies employed in existing RAT can use language grammars resulting in decision problems to be undecidable, language-theoretic principles such as complexity reduction can be used to minimize risks during the implementation phase.
- RRS as a composed system:
  - Building on the language-theoretic approach, other verification methods, and other assumptions (such as the nature of Radio Application and of the RPI), a model can be built to allow formal verification of expected properties of RRS, such as security and compliance.

NOTE: There is an impact on system flexibility in favour of system security of adopting a LangSec approach as its adoption deliberately limits the flexibility by restriction of the vocabulary of the definition language and by also restriction of the state transitions allowed through the inputs.

Annex B:
Review of push mechanisms

B.1 Overview

This annex provides an overview of the push mechanisms available across IP networks, and other technologies such as cellular networks. It is not meant to be a detailed description of each technology, but rather highlights the core properties of each mechanism as well as their security expectations, in the context of their potential use with Reconfigurable Radio Systems.

Such potential uses could be, for example, messages being sent to the Reconfigurable Equipment in order to trigger a connection to the Radio Application Store, or as part of the remote control framework.

Three core technologies are evaluated: generic IP-based mechanisms such as those related to common smartphone platforms, IP-based mechanisms adapted to cellular networks such as OMA Push, and core features provided by 3GPP systems.

B.2 Generic IP-based push mechanism

B.2.1 Introduction

The central assumption of generic IP-based push mechanisms is that they operate on an open network, such as the Internet, where devices are not restricted in their capabilities to access the mechanism.

B.2.2 Services operated by third-parties

As of writing the present annex, several push services integrated with smartphone operating systems are available on the market. These services are aimed at application developers and handle the details of notification routing and delivery to devices. In this case, developers can benefit from a rich and scalable push environment operated by a third-party.

Figure B.1 provides a high-level description, in generic terms, of such service.

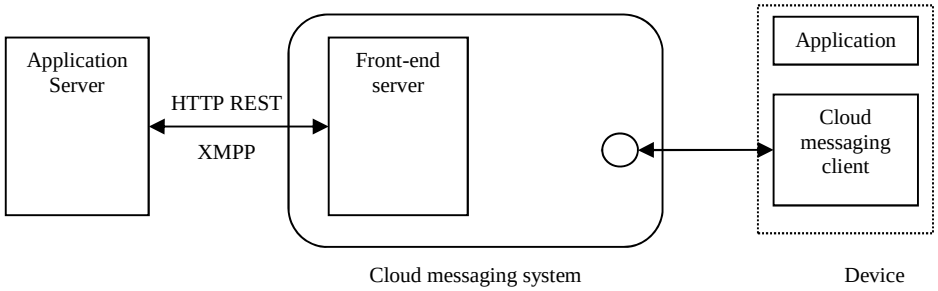

Figure B.1: Overview of a generic IP-based push mechanism as can be found on a smartphone platform

The front-end server is the entry point to users of the push service and exposes the service features via an API (e.g. on top of HTTP, or with XMPP). It accepts messages and other commands from the user's Application Server, which is located within the user domain of control. End-points exist to connect the messaging infrastructure with a dedicated client on the device - usually tied to the device Operating System - which finally transfers the message to the appropriate application.

From an operational perspective, messages are not pushed to a device but to a specific application on the device. Since the push service is shared between many users, they can only address applications which they have developed or for which they have been granted access by another user. Within the domain of a given application, messages can usually be routed according to three targets: a single device, a group of devices, or a topic (following the publish-subscribe model).

Additional services can be provided, for example:

- storage of per-device/per-application messages or configuration data, and management of application status from within the infrastructure;
- analytics services.

B.2.3 Security considerations

Security features exposed to users usually cover the interactions of the push service platform with external elements:

- communication security and authentication of end-points between the Application Server and the Front-End Server, and between the messaging infrastructure and the device or client on the device;
- token-based authentication for users and applications in order to ensure legitimate routing of messages;
- control configuration for supplementary services of the infrastructure (such as access control of message and assets stored within the infrastructure).

While these external properties can be evaluated on a case-by-case basis, it can be more difficult for users to determine the internal behaviour of the service, for example regarding integrity and confidentiality, near real-time delivery, or guaranteed delivery of messages. In such case, additional measures may be necessary as summarized in table B.1.

Table B.1: Security strategies with cloud messaging services

Expected property	Strategy
message integrity	contractual agreement, end-to-end protection of payload
message confidentiality	contractual agreement, end-to-end protection of payload
real-time delivery	contractual agreement
guaranteed delivery	contractual agreement

B.2 Push mechanism adapted to cellular networks

B.2.1 Introduction

This clause introduces the OMA Push [i.11] enabler as an example of push mechanism that has been designed to support cellular networks.

B.2.2 General principles

OMA Push provides a generic push architecture and builds on a common message format and set of delivery mechanisms adapted to IP and non-IP network bearers via specific push protocols - one of the objectives of OMA Push is to make use of efficient delivery mechanisms of cellular data services when they are available. The general architecture is as follow.

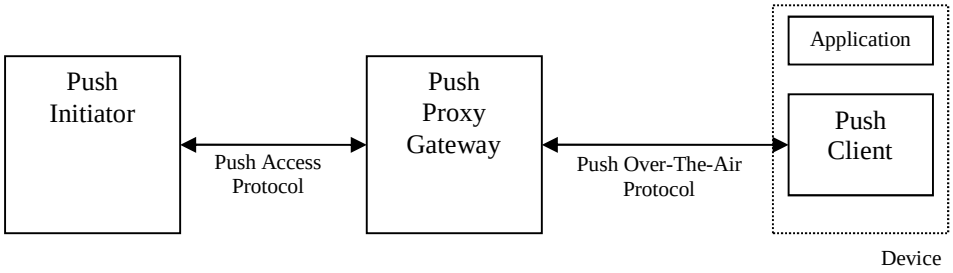

Figure B.2: Overview of the OMA Push architecture

The format of the Push Message allows for transporting arbitrary content from the Push Initiator to the consuming application. The Push Client dispatches the content to the proper application by means of an application identifier present in the Push Message. The metadata in the Push Message allow for other functionalities, such a cache control and retrieval of remote content - that is to say, the content to be pushed to the terminal does not have to be in the Push Message, and instead can be retrieved by the terminal from information contained therein.

B.2.3 Adaption to network bearers

B.2.3.1 Overview of adaption process

Two categories of access are defined: connectionless and connection-oriented delivery. In the latter case, a permanent connection is established between the Push Client and the Push-Proxy Gateway. Delivery of messages can be point-to-point or point-to-multipoint depending on the network bearer.

B.2.3.2 Point-to-point delivery

Three methods are available for point-to-point delivery in OMA Push: the Wireless Session Protocol (WSP), the Hypertext Transfer Protocol (HTTP), and the Session Initialization Protocol (SIP).

WSP [i.12] is part of the Wireless Application Protocol (WAP [i.13]) which aimed at developing an application environment optimized for wireless communication network as well as push primitives - hence OMA Push builds and extends on previously developed push mechanisms, in particular WAP. One connectionless primitive of interest in WSP is the SMS bearer type which is available in all 3GPP networks and has equivalents in many other cellular technologies.

The HTTP [i.14] delivery method relies on an HTTP server on the terminal listening for push requests (embedded in HTTP POST requests) on behalf of the Push Client. Since the terminal may operate on public and private network alike, a method is available for the terminal to initiate the underlying TCP [i.15] connection when the Push-Proxy Gateway cannot do so by itself - by means of Service Initiation Request messages instructing the terminal to perform the connection. The method is operational once the terminal has registered to the Push-Proxy Gateway.

The SIP [i.16] method is intended for operation of the Push Client with a SIP Core, for example the IMS Core [i.17]. In such case the Push Client acts as a SIP client. In connectionless mode, the Push Message is delivered with a SIP MESSAGE, while in connection-oriented delivery a SIP INVITE is sent to the Push Client. In case the terminal is not registered with the SIP Core while a Push Message is pending delivery, such registration can be triggered with a Service Initiation Request message.

B.2.3.3 Point-to-multipoint delivery

Point-to-multipoint delivery allows for simultaneous reaching large groups of Push Client with the same Push Message. It is achieved via broadcast or multicast bearer channels. Optimizations are usually in place so that terminal do not have to continuously listen to the channel. OMA Push leverages several such bearers:

- 3GPP MBMS [i.18].
- OMA BCAST [i.19], which in turn provides a file delivery service over bearers such as 3GPP MBMS [i.18], 3GPP2 BCMCS [i.20] and DVB-H [i.21].
- 3GPP CBS [i.22].

When point-to-multipoint delivery is in use, the Push Client has to be configured to listen on the Push Channel used by the Push-Proxy Gateway. This can be done via the Service Initiation Request message or via out of band means such as OMA Device Management. Because of the unidirectionality of the broadcast/multicast bearer, another bearer is needed for the Push Client to send confirmation messages back to the Push-Proxy Gateway. In case the broadcast/multicast Push Message is meant to trigger download of addition content, access to a bidirectional bearer may also be required.

B.2.4 Security considerations

Push messaging systems designed to cover a variety of delivery configurations often involve complex protocol encapsulation and incomplete reuse of existing, standardized building blocks, due to the need to simultaneously support very different technologies.

This can have various consequences. In terms of service safety, caution is advised as the semantics of a given operation in a protocol may not be conveyed properly once the protocol has been encapsulated. For example, when a Push Client acting as a SIP client answers a SIP MESSAGE containing a Push Message with a "200 OK" error code, this cannot be translated into a confirmation that the Push Client could successfully deliver the Push Content to the final application.

In terms of security, client implementations may have to support several functionalities that are not needed to implement the push service. This results in an increase of the attack surface. In addition, the trust boundaries may not be clear-cut which makes it difficult to evaluate security properties of the complete push messaging system.

The use of broadcast or multicast bearer for delivery of cryptographically protected material implies that receiving terminal be provisioned with common key material (e.g. group keys). This can increase the complexity of the key management infrastructure.

B.3 Security properties of data and notification services in 3GPP networks

B.3.1 Introduction

The Short Message Service (SMS) and the Cell Broadcast Service (CBS) are core methods enabling push delivery in conjunction with the network infrastructure of a Mobile Network Operator, as detailed in clause B.2. In addition, the data service can be leveraged to connect to IP-based push infrastructure operated in foreign networks.

The purpose of this clause is to give an overview of these services and of their capabilities across some of the existing 3GPP variants (namely 2G, 3G and LTE systems). The objective is not to provide a detailed and exhaustive description, but to give enough information for the identification of security properties.

NOTE 1: For simplicity the systems are named after their brand, not their technical designation.

NOTE 2: The security properties are given in a broad sense and do not consider the security strength of the related technical solutions.

B.3.2 Data service

The data service provided on the user plane - i.e. from the terminal up to the SGSN in the case of 2G (GPRS) and 3G and up to the P-GW for LTE, in order to access the Public Data Network - is not integrity protected (i.e. there is no protection against manipulation of the data). It is possible to provide confidentiality protection, but this is not mandatory in deployments.

B.3.3 Cell Broadcast Service

The Cell Broadcast Service [i.22] allows for local delivery (in the cell) of a Short Message to all devices present in the cell. It is primarily intended for public warning systems although other uses have been standardized (e.g. OMA Push).

In a nutshell, the service is controlled by the Cell Broadcast Centre in the core network. Messages are transferred over the BTS-MS interface with the Cell Broadcast Channel (CBCH) for 2G systems, and over the Uu interface with the Common Traffic Channel (CTCH) for 3G systems [i.22]. There is no security protection in both cases [i.23], [i.24].

B.3.4 Short Message Service

The Short Message Service [i.25] is controlled by the SMS Service Centre (SMS-SC) within the core network. Several delivery mechanisms are available: A/Gb mode (2G), A/Gb mode with GPRS, Iu mode (3G), S1 mode (LTE)/SMS over SG, and SMS over the IP Multimedia Subsystem (IMS) [i.26].

In A/Gb mode, the message is distributed over the SDCCH and SACCH channels which are subtypes of the DCCH. The DCCH is subject to confidentiality protection, but not integrity protection. The encryption end-point is the base station [i.27], [i.28], [i.29].

In A/Gb mode with GPRS, the message is distributed from the SGSN to the terminal over the LLC layer [i.27], [i.30]. This layer is subject to confidentiality protection, but not to integrity protection [i.29].

In Iu mode, the procedures are based on the Mobility Management sublayer which is subject to integrity protection [i.31]. The MM-sublayer requires the establishment of an RR connection, which has an optional ciphering procedure [i.32].

In S1 mode, messages are tunnelled through the MME and benefit from the integrity and confidentiality protection on the S1-MME interface.

In the case of the IP Multimedia Subsystem, an Application Server beyond the IMS Core implements the role of an IP-Short-Message-Gateway, which is to route Short Messages between the IMS-enabled terminal and the SMS-SC. Short Messages are embodied as SIP MESSAGE with specific Content-Type headers, for example "application/vnd.3gpp.sms". The P-CSCF is the entry point of the IMS-enabled terminal to send and receive Short Messages [i.33]. Once IP connectivity has been obtained (i.e. via SGSN in 2G/3G of P-GW in LTE), the UE performs a SIP registration with the P-CSCF of its serving IMS which, when successful leads to the establishment of IPsec security associations. When [i.34] is followed, integrity protection is provided, while confidentiality protection is optional.

B.3.5 Security considerations

The level of protection for Short Messages depends heavily on the technology selected by the Mobile Network Operator and may not always satisfy the security requirements of the overall push solution. Acceptable levels of protection can be achieved via contractual agreement with the Mobile Network Operator or via end-to-end protection of the push payload. Table B.2 summarizes the level of protection provided in some of the common settings.

Table B.2: Security services in common deployments

Method	Integrity protection	Confidentiality protection	End-point
CBS in 2G systems	-	-	Base Station
CBS in 3G systems	-	-	Node B
SMS in A/Gb mode (2G)	-	*	SGSN
SMS in A/Gb mode with GPRS	-	*	SGSN
SMS in Iu mode (3G)	+	*	SGSN
SMS in S1 mode (LTE)/SMS over SG	+	+	MME
SMS over the IP Multimedia Subsystem (IMS)	+	*	P-CSCF
NOTE: "-" denotes a functionality that is not provided, "+" denotes a functionality that is provided, "*" denotes a functionality that is optional			

Finally, Short Message authentication is that of the SMS-SC, which can only be inferred from network authentication at the time the terminal registers. The same goes for trust towards the access to the Public Data Network provided by the data service. However, authentication of the home network was not available in 2G systems but has been available starting with 3G systems [i.42]. Authentication of the serving network is available in LTE [i.1].

Annex C: Bibliography

- Sergey Bratus, Michael E. Locasto, Meredith L. Patterson, Len Sassaman and Anna Shubina: "From Buffer Overflows to "Weird Machines" and Theory of Computation"; login: VOL. 36, NO. 6.
- Len Sassaman, Meredith L. Patterson, Sergey Bratus and Anna Shubina: "The Halting Problems of Network Stack Insecurity"; login: VOL. 36, NO. 6.
- Len Sassaman, Meredith L. Patterson, Sergey Bratus: "A Patch for Postel's Robustness Principle", IEEE Security & Privacy, March/April 2012.
- Sergey Bratus, Trey Darley, Michael Locasto, Meredith L. Patterson, Rebecca "bx" Shapiro and Anna Shubina: "Beyond Planted Bugs in "Trusting Trust"; The Input-Processing Frontier", IEEE Security & Privacy, January/February 2014.
- Robert J. Hansen, Meredith L. Patterson, Guns and Butter: "Towards Formal Axioms of Input Validation".

History

Document history		
V1.1.1	September 2017	Publication