Securing USIM-based Mobile Communications from Interoperation of SIM-based Communications

Eric Southern, Abdelkader Ouda, Abdallah Shami
Western University - London - Canada
Electrical and Computer Engineering

Abstract

Mobile networks security is constantly evolving and adapting to meet the needs of users and network operators. It is a requirement that there be interoperation of legacy security frameworks into modern mobile networks. Mobile networks originally had no real security which proved to be a deployment that was attacked constantly and the providers were defrauded of millions of dollars. To address the issues the SIM authentication protocols were developed to secure the resources of the network providers. The original SIM security framework developed in GSM networks had weaknesses brought about by the one way authentication protocol as well as weaknesses in the algorithms used to secure the communication. The evolution of authentication in mobile networks to address the problems in the SIM framework brought about the creation of the USIM protocols used in UMTS, LTE and WiMAX to secure the network from the SIM framework security issues. The integration of those two SIM and USIM frameworks brought forward the major weaknesses first found in the SIM framework. This paper proposes simple and effective solutions to reduce the possible attacks on the USIM protocols due to the above integration. First we propose a subtle modification to the SIM based GSM security protocols as a stand-alone solution, and then a modification to the USIM based UMTS security protocols is proposed as a second solution.

1. Introduction

Wireless communication allows for easy connectivity of devices without the expensive requirements of laying a physical network. One of the main difficulties in deploying wireless networks is the ability to secure information and resources on a medium that by its very nature broadcasts all information. A key aspect of securing wireless communication is the authentication protocol used to allow access to the network. The two major types of wireless networks are the stationary networks generally defined by the IEEE 802.11 standards and the mobile networks defined as 2G, 3G and 4G networks. As security requirements have changed the protocols for authentication have adapted with those changes. Authentication is the process of determining whether someone or something is, in fact, who or what it is declared to be. The traditional method of authentication in computing is the challenge-response mechanism. There is a shared secret between the two parties that is used in an algorithm so that one party poses a question as a challenge and the other party must reply with a correct answer as a response. Both of these network types have faced significant security problems that have needed to be addressed with stronger protocols and more secure cryptographic algorithms. When creating the new more powerful algorithms and protocols the older hardware cannot implement them due to the more strenuous requirements.

The demands on mobile communication and networks have been constantly increasing. Originally the need was simply to have a phone system that could meet most of the requirements of the standard plain old telephone service (POTS) in most homes. The original first generation (1G) systems, such as the advanced mobile phone system (AMPS), were analog cellular networks which met this need without considering the inherent issues that arise due to using a wireless medium as opposed to a wired one. Security was a major issue that was not properly addressed when developing the 1G systems and therefore the phones were susceptible to cloning. This was due to the phones broadcasting their identities without encryption or integrity when phone calls are placed. Attackers could then take this information and apply it to their own phone to then use it to connect to the provider network allowing them to call anywhere without having a legitimate account with the provider. The cloning defrauded many providers of large amounts of money while inappropriately making unauthorized use of their resources. There are many benefits and

requirements of security in mobile wireless communication.

The second generation of mobile communications (2G) strove to solve the phone cloning issue and while meeting the expanding requirements of consumers with GSM/2G networks. Global system for mobile communications (GSM) networks also addressed some of the issues with using a wireless medium when sending information. The new network authenticates the user against the network in a cryptographically secure method to limit the potential of phone cloning security issues as well as ensuring that the network resources are not accessed inappropriately. This made phone cloning a much more difficult proposition for attackers to inappropriately make use of provider networks while allowing providers to be much more certain that their resources were not being fraudulently used by unauthorized devices. The problem with GSM networks was that they did not appropriately protect the user from many other types of attacks such as the false base station attack that would allow an attacker to listen in or modify the communication from the GSM user. The false base station attack and other security issues in GSM networks were attempted to be resolved by providers with the third generation of mobile communication (3G).

3G mobile communications allowed for much better use of the spectrum available allowing much "smarter" devices to be on the network. Even though the cloning issue was mostly resolved with the GSM networks there were other security issues that needed to be addressed in universal mobile telecommunications systems (UMTS) networks. To address these new issues the 3rd generation used mutual-authentication between the mobile device and the provider network. The UMTS networks also have much higher speeds for IP communication to allow for users to make extensive use of the network resources.

The next generation of mobile communication will make even further use of the available spectrum and in-crease the ability of smart devices to do much more robust communication with media and other applications. The authentication in the fourth generation (4G) is still going to be the same authentication protocols as the USIM 3G to make certain that resources are not misappropriated. (4G) long term evolution (LTE) networks will allow wider bandwidths, higher efficiency and a fully IP network for all communication.

GSM networks have by far the largest installed base of users with over 5 billion GSM devices in use around the world [1]. This large market of devices has made it a business requirement of all providers to allow for the legacy GSM system to be integrated into new systems to ensure that these users can use the network resources and be billed appropriately for that usage. The interoperation of legacy systems needs to be executed with the utmost care to ensure that issues in the legacy system do not manifest themselves in the new integrated system. There are many security concerns when integrating legacy systems and the evolution of those systems to handle new requirements. The authentication done in the GSM network was maintained in the new UMTS networks to allow these devices to connect. This integration allows some of the security issues in GSM networks to be exploited in the new network.

This work discusses authentication in mobile wireless networks as well as the needs of those networks to interoperate and the security issues brought about by that integration. The authentication protocols in the new generations of mobile wireless networks are designed to interoperate (not replace) the existing protocols as the infrastructure for the existing system is deployed nationally and is very expensive to replace requiring time, effort and expense. Therefore the integration of the different protocols to allow this interoperation gives the mobile operators the ability to upgrade their networks while still maintaining coverage for their customers.

It worth mentioning that, the integration of the old (flawed) security protocols is not always a right option. For instance, the new authentication protocol described by IEEE 802.11i to protect stationary wireless networks replaced (not interoperate) the legacy Wired Equivalent Privacy (WEP) due to the problems found in the earlier algorithms and protocols. Eric, *at al* [2] compared between the interoperation solution in mobile networks and the replacement solution employed in stationary wireless networks.

In this paper, the interoperation will be described considering security flaws brought about by integrating the old protocols into the new systems. This will include a description the authentication and key agreement (AKA) protocols of the legacy SIM based 2G GSM networks, and the modern USIM based 3G UMTS networks, 4G LTE networks and WiMAX networks. The protocols and methods used for the integration of the legacy systems into the modern AKA systems will discussed. In order to protect the integration in mobile wireless networks against these security flaws in SIM based networks two solutions are proposed.

This paper is organized as follows. Section 2 describes the SIM based authentication mechanism used in GSM networks. The USIM based authentication mechanism used UMTS 3G, LTE 4G, and WiMax 4G networks are described in Section 3. Section 4 describes the mobile user hand-over between different network, and the related security issues. Two solutions are proposed to the problem of integration in

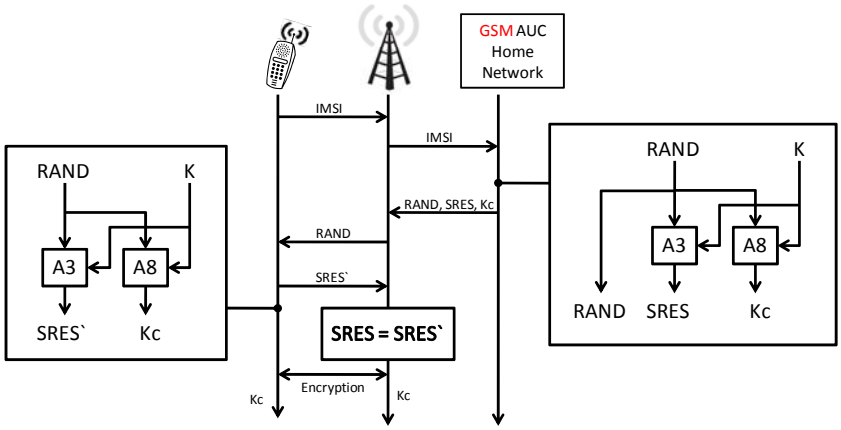


Figure 1. GSM Authentication

mobile wireless networks in Section 5. Section 6 concludes this work.

2. SIM-based Authentication Mechanism

Mobile service providers needed to secure their networks from attack and misappropriation of networking resources. In the attempt to achieve the goals set out in GSM of protecting access to mobile services and to protect any relevant item from being disclosed on the radio path [3]; the GSM security protocols were developed. There are many technical constraints that needed to be addressed when adding security to mobile communication. When authenticating against a mobile wireless network the mobile equipment needs to be able to send from one base station to another without a loss of communication or interruption to an active connection. The requirement to roam without interruption was a major factor in development of mobile networks that would allow a user to be able to authenticate to and use all parts of the network seamlessly. A major difficulty faced by mobile networks is the ability for a user to roam from one network to another network operator which allows mobile network providers to bill foreign users and systems. The authentication protocol deployed to address these problem was the SIM based GSM protocol.

The authentication in GSM is a one-way authentication algorithm to authenticate the mobile device to the service provider network. As shown in Figure 1 the algorithm uses a secret key K that is shared between the GSM home network and the mobile device. The mobile device identifies itself to the network by sending its international mobile subscriber identity (IMSI) to the base station (BS). The BS forwards the IMSI to the home network of the device. Based on the IMSI the home network recognizes the corresponding key K that is used along with a random challenge (RAND) to generate a session key K_c =A8(RAND, K) and the expected response to the challenge SRES=A3(RAND, K), where A8 and A3 are two hashing functions. The home network sends the authentication vector (RAND, SRES, K_c) to the BS who will retain SRES and K_c and sends the RAND to the mobile device as a challenge. Using the shared secret key K along with the received RAND the mobile generates the response SRES' and generates the same session key K_c. The BS responds to the BS with the SRES which the BS then matches against the SRES to verify the identity of the mobile device. This authentication in GSM gave the service providers the ability to address the issue of cell phone cloning by issuing a challenge to the device that would appropriately be responded to with the SRES'. GSM also added encryption using the key K_c to the channel to allow the confidentiality on the information transmitted across the air interface.

Even with all of these new security enhancements to wireless communication, there are many problems with the authentication and security in GSM. The encryption and hashing algorithms were developed in secret design, in violation of Kerckhoff's principle [5]. It says a cryptosystem should be secure even if everything about the system, except the key, is public knowledge, which led to the system being less secure than if they had used known algorithms that had been vetted by cryptographers not involved in the design. In addition, the stream cipher A5 is used for encrypting the communication channels. The adopted A5/1 encryption algorithm in GSM can be broken in real time [6] and the A5/2 algorithm is easily broken in seconds [7] meaning that the intent to keep communication of the customer on the network private is no longer truly provided by the protocol. The GSM framework does allow providers to choose different algorithms for both the hashing and encryption but due to the established base and weaknesses in the protocol this is not entirely feasible for the encryption protocol

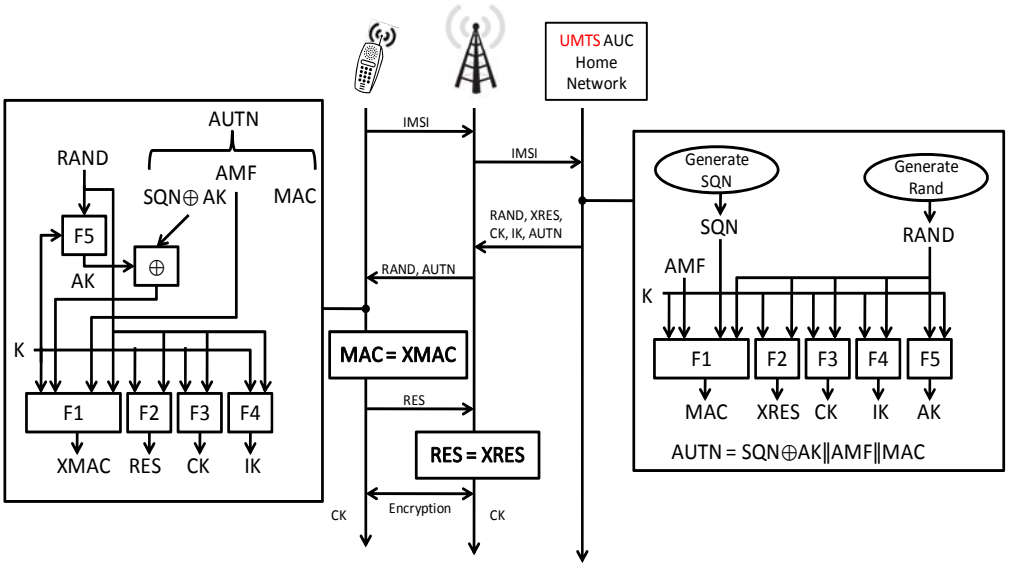


Figure 2. UMTS-AKA Authentication Protocol

(hashing protocols can be set specifically for each device at the discretion of the provider). The XRES and other values are also limited by their length as required in the GSM protocol.

The authentication protocol has many flaws that allow for denial of service, and false base station attacks since the subscriber does not authenticate the network. Note that, GSM uses one-way authentication. A false base station attack is visible due to the mobile device not authenticating the network. The false base station attack is a classic man-in-the-middle attack that generally passes most of the communication from the handset to the tower but will modify some of the transactions to attack the network. These attacks have a method that can retrieve the IMSI of the device and they can have the false tower also force the device to not use encryption for communication which allows the attacker to listen to the conversation and possibly inject information into the channel. Again, the fact that GSM protocol authenticates only the phone and leaves the network unauthenticated allows for these base station attacks to neutralize any increase in the quality of the encryption algorithms since the devices will support the older implemented algorithms and no encryption. The insecurity brought about by the protocol allows these attacks to compromise the confidentiality and integrity of the user communication with the network.

3. USIM-based Authentication Mechanism

3.1. UMTS-AKA Authentication Protocol

UMTS networks have mutual authentication in which the mobile devices is authenticated to the network as well as the network authenticating the phone as shown in Figure 2. This mutual authentication allows the device to discern whether or not the network they are connecting to is a legitimate network. The authentication protocol also makes use of integrity to ensure that the communication is not modified when selecting algorithms for encryption and integrity. The authentication protocol follows many of the same network steps in the GSM protocol with some important changes. The authentication token (AUTN) as well as the integrity key (IK) are sent from the home network. The AUTN token along with the RAND are then sent to the mobile device which processes the RAND with the key to verify the AUTN token by validating the MAC section of the token sent from the network against the XMAC created by using the key, sequence, authentication management field (AMF), and RAND. Note that, AMF is a section of the AUTN token. The mobile equipment also does a validation of the sequence to ensure that it is within the desired range. This verification allows the mobile device to trust the connection to the network.

The algorithms are at the discretion of the providers but generally the Kasumi [8] algorithm is used for both integrity and encryption with an option of no encryption. The UMTS protocol does not allow the system to operate without integrity, which in conjunction with the authentication allows the mobile device and network to have a reasonable expectation that there has been no modification of the communication. This method of authentication with integrity limits many attacks in a purely UMTS network. The Kasumi algorithm is a modified MISTY1 algorithm [8] that was chosen for its suitability for implementation in hardware. The algorithm has some weaknesses but is not susceptible

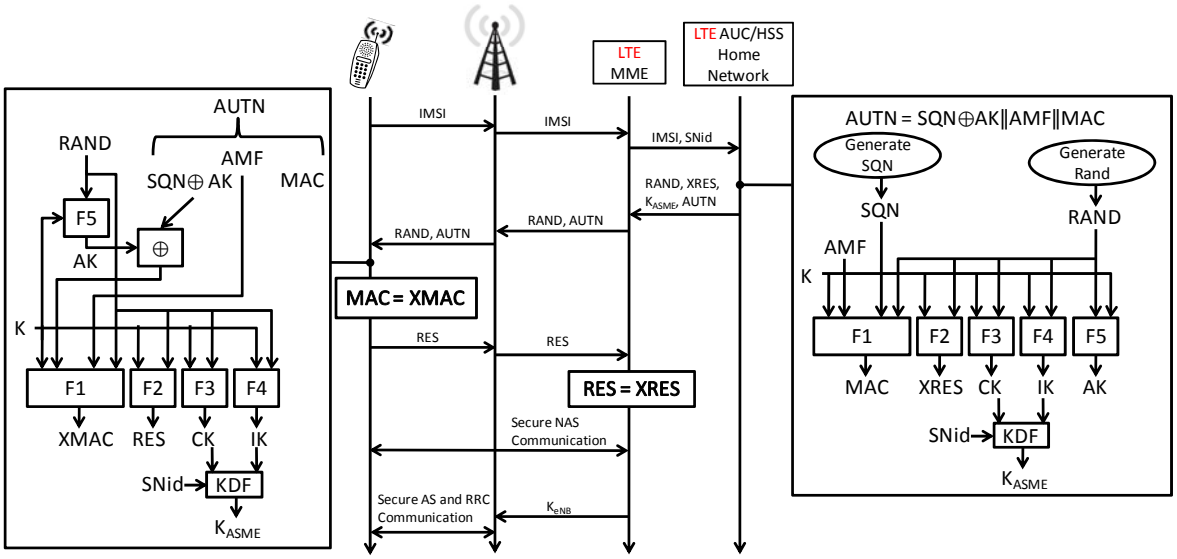


Figure 3. EPS-AKA Authentication

to real-time attacks [9]. Currently the international telecommunication union (ITU) is still developing the standards for 4G mobile communications but the authentication protocols are the same as those of the UMTS network [10].

3.2. EPS AKA, LTE Authentication Protocol

LTE networks were developed to meet the growing mobile data usage of users. The new network moved voice off of a circuit switched network to a packet switched IP based VoIP protocol. There is better utilization of the bandwidth and increased speed and capacity available for providers to meet the constantly growing needs of their users.

LTE networks have expanded the authentication key agreement used in UMTS. The beginning of the protocol is identical with the IMSI request being forwarded by the base station to the authentication center (AuC) as can be seen in Figure 3. The changes begin with the evolved packet system (EPS) authentication vector (AV) which has RAND, AUTN, XRES, and K_{ASME} which is the access security management entity (ASME) instead of CK and IK. The CK and IK values in the USIM along with the serving network's identity are the input into a key derivation function (KDF) to generate K_{ASME}. Then the similarities to the UMTS protocol continue with the user equipment (UE) validating the MAC and then responding with the RES for the network to complete the authentication procedure by comparing it with XRES. The major change in EPS is that the K_{ASME} is used to generate keys in a key hierarchy. Keys are generated for three different traffic types: the non-access stratum (NAS), access stratum (AS) and radio resource control (RRC).

The Key Agreement within the EPS-AKA protocol allows for 6 different keys to be generated as shown in Figure 4 which shows the different equipment that will be used to derive each key, as follows.

- K is a 128-bit secret key stored permanently in USIM and AuC.
- CK and IK are pair of 128-bit keys derived in AuC and USIM during the AKA process.
- K_{ASME} is a 256-bit intermediate key derived in the home subscriber server (HSS) and UE from CK and IK, during the AKA process. K_{ASME} is then forwarded to MME as a part in the EPS AV along with RAND, XRES and AUTN.
- K_{eNB}, K_{NASint}, K_{NASenc}, are 256-bit Intermediate Keys derived in MME and UE as well from K_{ASME} when UE transits to EPS Connection Management ECM state or by UE and target base station eNB using the previous K_{eNB} during eNB handover. K_{eNB} is then forwarded to the eNodeB (eNB). K_{NASint} is an integrity key for protection of NAS data derived in MME and UE. K_{NASenc} is an encryption key for protection of NAS data derived in MME and UE.
- K_{UPenc}, K_{RRCint} and K_{RRCenc} are 256-bit keys derived from K_{eNB} in eNB and UE. K_{UPenc} is an encryption key for protection of user data derived in eNB and UE. K_{RRCint} is an integrity key for protection of user data derived in eNB and UE. K_{RRCenc} is also an encryption key for protection of RRC data derived in eNB and UE.

Therefore, these keys are all based off of the pre-shared key K. They follow the same processes in UMTS as is evident in Figure 3 which are used to generate the different keys and values required for the key agreement and authentication. The difference arises when the keys CK and IK as well as the serving network identity (SNid) are used as input into a key

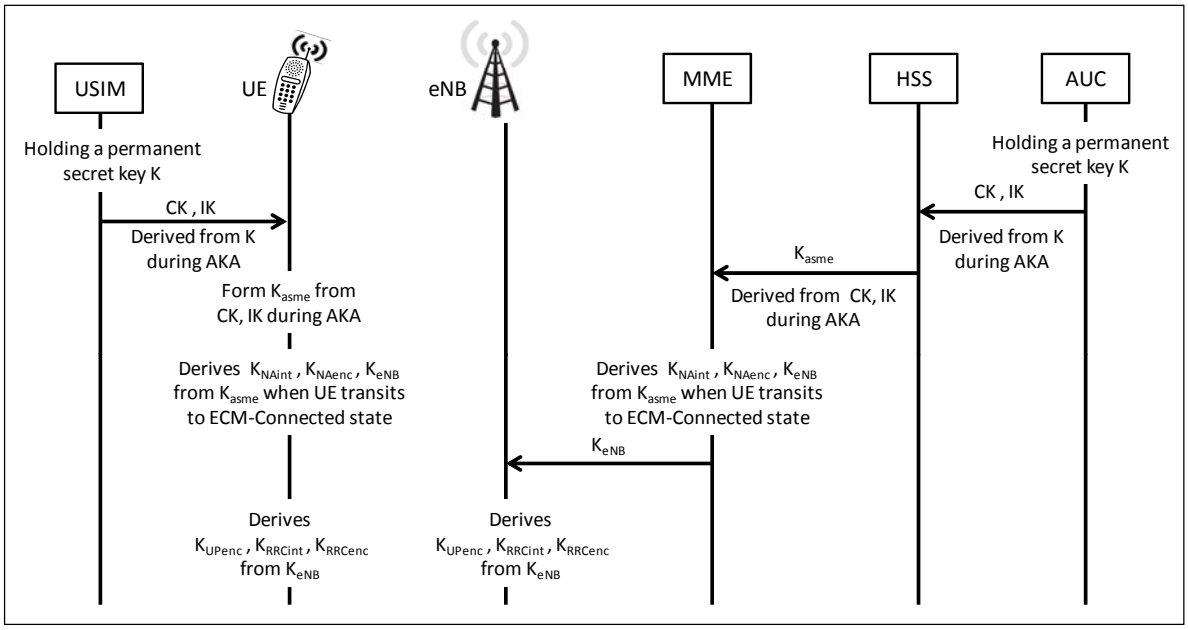


Figure 4. EPS-AKA Key Derivation

derivation function KDF generating K_{ASME} which is then used to generate all the other keys.

The authentication and key agreement in EPS-AKA has identical steps for mutual authentication as UMTS only the key agreement and which devices perform the key generation steps differ.

3.3. EAP-AKA, WiMAX Authentication Protocol

Worldwide Interoperability for Microwave Access (WiMAX) is an IP based, wireless broadband access technology that provides performance similar to 802.11/Wi-Fi networks with the coverage and quality of service (QOS) of cellular networks. In a fixed wireless configuration it can replace the telephone company's copper wire networks, the cable TV's coaxial cable infrastructure while offering Internet service provider (ISP) services. In its mobile variant, WiMAX has the potential to replace cellular networks. It is an IEEE standard designated 802.16-2004 (fixed wireless applications) and 802.16e-2005 (mobile wireless).

To allow connections to WiMAX by current USIM cellular devices an extensible authentication protocol method for UMTS Authentication and Key Agreement or for short EAP-AKA was developed to integrate the UMTS-AKA algorithm into the extensible authentication protocol (EAP) framework as seen in Figure 6. The start of the protocol requires the UE to initiate the connection with the access point (AP) by sending an EAPOL-Start message. The AP will then respond with the EAPOL-request/identity message. The remainder of the protocol is very similar to the UMTS-AKA algorithm but with the elements wrapped in their equivalent EAP message types. Therefore the UE will respond to the EAPOL-request/identity message with an EAPOL-response/identity which contains the IMSI of the USIM. The IMSI will be sent from the AP to the home authentication, authorization and accounting server (HAAA) which will control all future communication with the UE through the AP. The HAAA will then forward the IMSI to the AuC which will then create the authentication vector, identical to the one created in UMTS. The AuC will send the authentication vector of CK, IK, RAND, XRES, and AUTN to the HAAA. The HAAA will then send an EAP-request/AKA-challenge containing the RAND, AUTN, and MAC to the UE. The UE will verify the MAC and respond to the HAAA with the RES in an EAP-response/AKA-challenge message which is then to be validated by the HAAA. The HAAA will then respond with an EAP-success message.

The EAP framework adds some extra overhead to the UMTS-AKA protocol with the addition of the EAP standard messages that complete the requirements of the EAP framework but the overall protocol uses the same messages and mutual authentication requirements.

4. Legacy Integration of SIM with USIM

When the time came for industry to move to UMTS networks the market was already saturated with a large number of GSM devices and network equipment. The integration offered by the protocol allows for the providers to make use of the already embedded

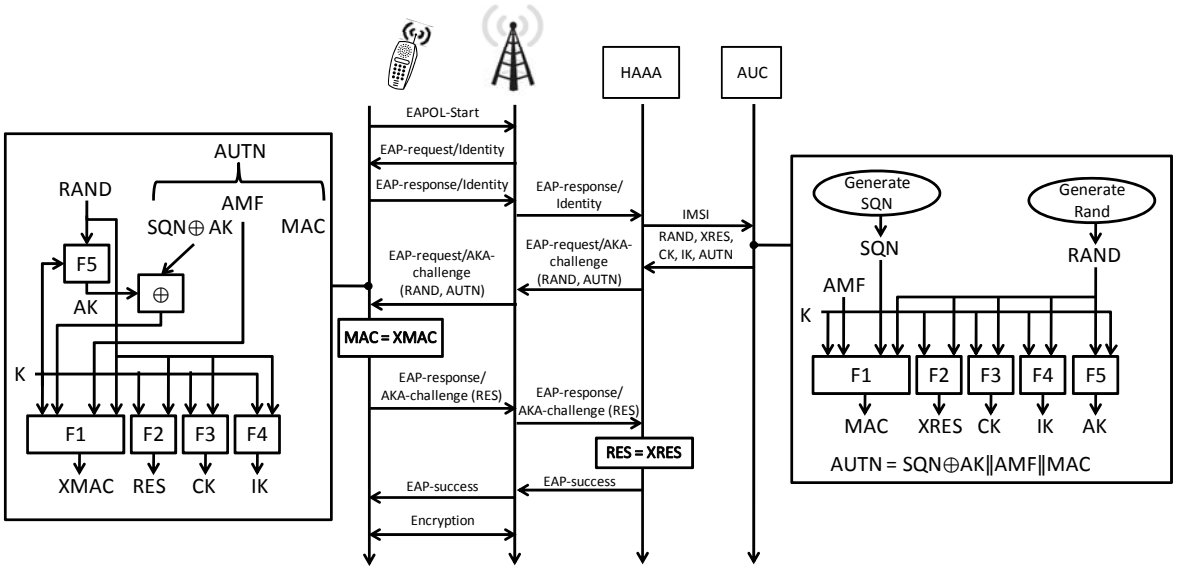


Figure 5. EAP-AKA Authentication

systems. To make the transition cost effective and to make maximum use of the existing user and network hardware, GSM backwards compatibility was built into the UMTS protocols [15]. The interoperation between the two systems allows GSM devices on the UMTS network and allows the network to be slowly upgraded to the new infrastructure. A provider can then support the large number of devices owned by customers as well as have a planned strategy for upgrading their network infrastructure.

To achieve the integration there are some equations that are used to convert the keys from UMTS *CK* and *IK* to GSM K_c and vice versa. Those equations allow the mobile device and network to continue to operate without requiring re-authentication to roam from one network configuration to another. Those equations to create K_c are:

$$K_c = CK_1 \oplus CK_2 \oplus IK_1 \oplus IK_2 \quad (1)$$

$$where, \quad CK = CK_1 \| CK_2 \quad (2)$$

$$and \ IK = IK_1 \| IK_2 \quad (3)$$

To create *CK* and *IK* from K_c when moving from a GSM context to a UMTS context the following equations are used:

$$CK = K_{c1} \| K_{c2} \quad (4)$$

$$IK = K_{c1} \oplus K_{c2} \| K_{c1} \| K_{c1} \oplus K_{c2} \quad (5)$$

$$where, \quad K_c = K_{c1} \| K_{c2} \quad (6)$$

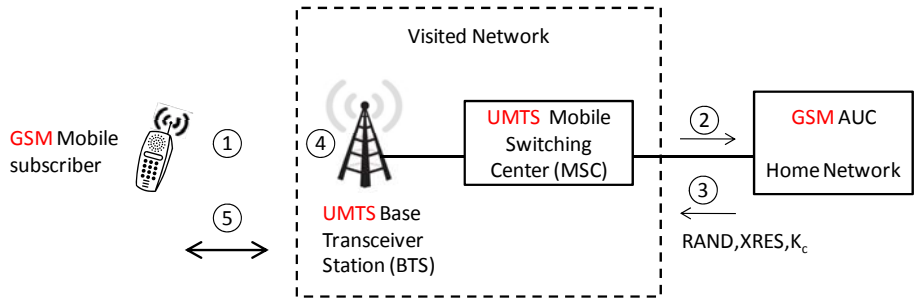
The following sub-section will be exploring 3 different authentication scenarios of GSM and UMTS equipment to show the methods of integrating these two generations of mobile communications.

The 3GPP attempted to address these issues with the security upgrades to the USIM protocol in LTE. They do adequately address protecting the existing USIM keys when moving to the less secure GSM or UMTS network configurations but there are possible issues with security spoofing that may bring the GSM issues forward into the LTE framework. When moving to the less secure UMTS network the proposed specification [16] states that the key K_{ASME} will be used with the KDF to generate a CK' and IK' to be used in the UMTS network. This will protect the LTE framework from an attacker gaining information during the subsequent UMTS or GSM communication and trying to learn information about K_{ASME} to attack the previous LTE communication.

The LTE specification also states that when moving into LTE from UMTS that a check of CK should be done to see if the first 64 and last 64 bits match. If they do it can be assumed that the connection was at one time a GSM connection. These are to be dropped unless there is an ongoing emergency communication occurring. It may be possible to spoof this status of emergency communication as an attacker due to the fact that an attacker could have full control of the communication from the UE. It also doesn't seem entirely practical to refuse the authentication transfer if an active non-emergency conversation is occurring.

4.1. GSM Mobile Device with UMTS Network

When a GSM Mobile device is on a UMTS network as shown in Figure 6, and as per the order of



(1) GSM Mobile subscriber requests a secure connection to UMTS BTS
(2) UMTS MSC requests from the GSM home network the authentication vector (RAND,XRES,K_c).
(3) UMTS MSC receives the GSM authentication vector and forward it to the UMTS BTS
(4) UMTS BTS perform the GSM Authentication protocol with GSM Mobile subscriber
(5) When the authentication process in (4) succeeded, the GSM Mobile and the UMTS BTS can communicate securely applying the UMTS encryption algorithms using the UMTS key CK and the integrity key IK. These keys are generated using the GSM K_c

Figure 6. The GSM Mobile subscriber is authenticated via a UMTS BTS, which is connected to a UMTS MSC

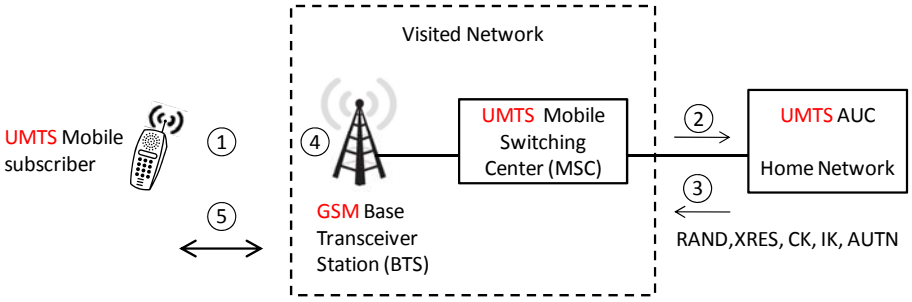
the circled numbers, GSM Mobile subscriber requests a secure connection to UMTS BTS. The UMTS MSC requests from the GSM home network the authentication vector (RAND, XRES, K_c). The UMTS MSC receives and then forwards the authentication vector to the UMTS BTS.

The UMTS BTS then perform the GSM Authentication protocol with GSM Mobile subscriber as described in Section 2 and Figure 1 above. If this authentication process succeeded, the GSM Mobile and the UMTS BTS can communicate securely applying the UMTS encryption algorithms using the UMTS key *CK* and the integrity key *IK*.

Note that, the system will create K_c at the home AuC of the GSM which will then be expanded with Equations 4 and 5 to create *CK* and *IK* in an enhanced GSM mode to increase the security of the communication. The issue brought about by this configuration is that when K_c has already been discovered by an attacker when the phone is operating in a fully GSM context the expanded *CK* and *IK* are easy to discern from the equations and all of UMTS communication can be discovered by an attacker.

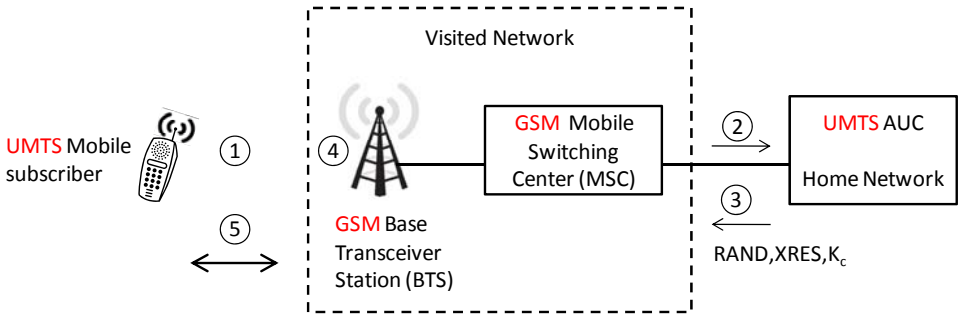
4.2. UMTS Mobile Device with GSM BTS

When connecting to the network it is possible for a UMTS mobile device to connect to a GSM BTS. As shown in Figure 7, and as per the order of the circled numbers, the UMTS Mobile subscriber requests a secure connection to GSM BTS. Accordingly, the UMTS MSC requests from the UMTS home network the authentication vector (RAND, XRES, *CK*, *IK*, AUTN). The UMTS MSC receives the UMTS authentication vector and proceeds to generate a GSM



(1) UMTS Mobile subscriber requests a secure connection to GSM BTS
(2) UMTS MSC requests from the UMTS home network the authentication vector (RAND,XRES, CK, IK, AUTN).
(3) UMTS MSC receives the UMTS authentication vector and proceeds to generate a GSM K_c and forwards K_c to the GSM BTS
(4) GSM BTS performs the GSM Authentication protocol with UMTS Mobile subscriber
(5) When the authentication process in (4) succeeds, the UMTS Mobile and the GSM BTS communicate using the GSM encryption algorithms using the GSM K_c. Which is insecure due to the attacks available against the GSM algorithms.

Figure 7. The UMTS Mobile subscriber is authenticated via a GSM BTS, which is connected to a UMTS MSC



(1) UMTS Mobile subscriber requests a secure connection to GSM BTS
(2) GSM MSC requests from the UMTS home network the authentication vector (RAND,XRES, K_c) which is generated by using the UMTS authentication vector (RAND,XRES, CK, IK, AUTN).
(3) GSM MSC receives the GSM authentication vector and forwards K_c to the GSM BTS
(4) GSM BTS performs the GSM Authentication protocol with UMTS Mobile subscriber
(5) When the authentication process in (4) succeeds, the UMTS Mobile and the GSM BTS communicate using the GSM encryption algorithms using the GSM K_c. Which is insecure due to the attacks available against the GSM algorithms.

Figure 8. The UMTS Mobile subscriber is authenticated via a GSM BTS, which is connected to a GSM MSC

K_c using Equation 1 and then forwards it to the GSM BTS. The GSM BTS performs the GSM authentication protocol with UMTS Mobile subscriber as described in Section 2 and Figure 1 above. If this authentication process succeeded, the UMTS Mobile and the GSM BTS communicate using the GSM encryption algorithms using the GSM K_c.

This type of connection is created either during authentication or during handover to this type of network. The only network device that uses the GSM protocols in this type of connection is the BTS. The MSC, Mobile and AUC are all UMTS devices. The MSC will retain the *CK* and *IK* generated by the UMTS authentication but all encryption between the Mobile and the GSM BTS is done using the K_c created with equation 1. K_c is created by the Mobile and by the UMTS MSC and the GSM BTS is oblivious to this operation. The communication between the Mobile and the BTS can be considered as secure as that of normal GSM communication. When moving to other network configurations the MSC will use the *CK* and *IK* that were originally generated instead of using the K_c generated for the BTS. We know that, the K_c can be compromised during communication with the BTS and will therefore give 64 bits of information relating to the original *CK* and *IK*.

4.3. UMTS Mobile Device with GSM BTS and MSC

Figure 8 shows another scenario when a UMTS mobile device connecting to a GSM network. Following the order of the circled number in the Figure, the UMTS Mobile subscriber requests a secure connection to GSM BTS. Accordingly, the GSM MSC requests from the UMTS home network the authentication vector (RAND, XRES, K_c) where it is generated using the UMTS authentication vector (RAND, XRES, *CK*, *IK*, AUTN). The GSM MSC receives the GSM authentication vector and forwards K_c to the GSM BTS. The GSM BTS then performs the GSM Authentication protocol with UMTS Mobile subscriber as described in Section 2 and Figure 2 above. If this authentication process succeeded the UMTS Mobile and the GSM BTS communicate using the GSM encryption algorithms using the GSM K_c.

In this type of connection authentication or handover occurs when a UMTS authenticated session moves to a GSM network. The GSM MSC and GSM BTS can only handle the K_c for GSM communication. Therefore the UMTS authenticated network transfers K_c derived from equation 1 to the GSM MSC. The new K_c will be used to create any future *CK* and *IK* as well as for all communication between the GSM BTS and the Mobile using equations 5 and 6. This decreases the security of the system beyond the 64 bits of knowledge shown in the previous weakness to a full break of all future communication. All future communication until a new authentication request can be discovered and modified by a false base station. This is the worst case scenario for a UMTS device as it is fully compromised.

5. Proposed Solution to Problem of GSM Integration in UMTS

To solve the issues brought about by integrating the large install-base of the GSM platform and network equipment into the new and more secure UMTS system we have two solutions. We cannot do large modifications to the existing GSM system to protect the communication that will happen when in a GSM context and will therefore assume that when

communication happens in a GSM context that K_c will be compromised and known to attackers. Our focus is on protecting the UMTS communication from attacks through the integration with GSM. First we show a modification to GSM that will allow future communication to be secure when on an UMTS network. Our second proposal is a larger modification to the UMTS protocols to harden the communication in UMTS from attacks due to the GSM integration. It is worth mentioning that, both of the proposals do nothing to increase the security in GSM. GSM is still insecure but we are protecting UMTS from the integration with GSM.

5.1. Proposed Modification to GSM

The change we are proposing to the GSM authentication protocol shown in Figure 9 is simple and yet very effective. As all GSM devices have a hashing algorithm available, such as A3 and A8, and this operation need only happen once when moving from tower to tower the overhead should be minimal. It may be simple to implement this change to existing GSM system hardware. A hashing algorithm is able to keep the source material unknown while creating the same output if given identical input.

This is because it is computationally hard to discover the input if the output is known. Therefore we propose that the encryption in GSM is done with a new key K_h which is a hash of K_c instead of K_c directly, as it is shown in Equation 7.

$$K_h = hash(K_c) \tag{7}$$

This would leave the GSM communication open to all of the previous attacks but when compromised would give the attacker access to K_h instead of K_c. We will now describe how this change protects the communication in each of the previously described scenarios.

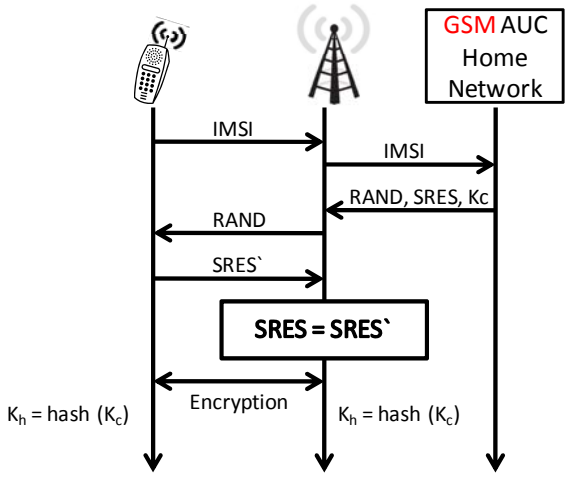


Figure 9. Proposed modification to GSM Protocol

5.1.1. GSM Mobile Device with UMTS Network

Figure 10 shows how GSM authentication takes place with the proposed modification, we see that the air-interface between the mobile subscriber and the BTS is encrypted using shared key K_h. If we assumed an attacker has successfully compromised K_h due to the insecurity of GSM, still the attacker has no access to the value of K_C. This means the values of *CK* and *IK* that are derived from K_c (see Equations 4 and 5) are not compromised. Therefore, in this scenario UMTS security will not compromised and its strength depends on the security of the cryptographic hash function used in Equation 7.

5.1.2. UMTS Mobile Device with GSM BTS

When encrypting the communication again between the mobile and the GSM BTS using the key K_h (see Figure 9), the value of K_c will be shielded by the cryptographic hash function. This hash would keep the attacker far from deriving 64 bits of *CK* and *IK* when the user moves to other networks as the attacker would not be able to discern anything beyond K_h when the system is communicating in this scenario. Again, knowing the value of K_h gives no significant knowledge of K_C and therefore no partial knowledge of *CK* and *IK*.

5.1.3. UMTS Mobile Device with GSM BTS and MSC

Similarly in this scenario, the cryptographic hash function protects K_c from the attacker. This has a much larger implication in this scenario as the *CK* and *IK* that will be used in the future are completely derived from K_c and will be protected from attack due to that the hash function is one-way function. Therefore, the compromised K_h will not give the attacker significant knowledge of K_c and through that will protect all future communication using *CK* and *IK* that are derived directly from K_c.

5.2. Proposed Modification to UMTS

The change to the UMTS protocol is two-fold as it needs to protect information when moving to a GSM network and protect the user when moving back to a UMTS network context. First we recommend that instead of using the equations developed for integration of the legacy GSM protocols we propose that a hash of *CK* and *IK* be used to create the key K_c to be used when communicating in the GSM network. I.e., Equation 1 above will be modified as follows:

$$K_c = CK_1 \oplus CK_2 \oplus IK_1 \oplus IK_2 \tag{8}$$

$$where, \quad hash(CK) = CK_1 \| CK_2 \tag{9}$$

$$and \quad hash(IK) = IK_1 \| IK_2 \tag{10}$$

The advantage to using this equation as opposed to Equation 1 is that the attacker will be unable to find information relating to *CK* and *IK* by knowing the value of K_c. This modification would protect the information sent before moving to the GSM context by securing the values of *CK* and *IK* from creating the value of K_c.

The second change to the protocol is to have the UMTS mobile device and the network do a simple hash of K_c, *K* and a RAND to create a new *CK* and *IK* for use after leaving the GSM context. This would be a simple request/response from the new UMTS network to the UMTS AuC to create the new *CK* and *IK* to be used for communication similar to a location update as can be seen in Figure 10. The small request would be much less overhead than a full re-authentication in UMTS to limit resource utilization on the network. The message sent would be similar to the location update by sending the TMSI along with K_c to the UMTS AuC. The UMTS AuC would then perform a hashing operation as to create a new set of keys for *IK* and *CK* that we will call $K_{CK||IK}$ shown as follows:

$$K_{CK||IK} = hash(K_c \| K \| RAND) \quad (11)$$

$$where, \quad K_{CK||IK} = CK \| IK \quad (12)$$

The AuC will proceed to respond with the new $K_{CK||IK}$ and a RAND to be sent to the mobile device to perform the same operation. This would by necessity have to occur before or immediately after handover to a fully UMTS context. The mobile device and the UMTS network would then be able to communicate securely without considering the fact that the K_c could have been compromised during the GSM communication context. The next sections will describe the impact of this change on the different network scenarios.

5.2.1. GSM Mobile Device with UMTS Network

This context would use the new $K_{CK||IK}$ created in Equation 11 for the keys *CK* and *IK* to be used in the UMTS encrypted communication. This would make the communication secure from any possible attack if the value of K_c had been discovered previously during a fully GSM context. The new values of *CK* and *IK* are not derived with Equation 1 and therefore do not directly come from K_c which makes future communication secure from a compromised GSM context.

5.2.2. UMTS Mobile Device with GSM BTS

The communication in this context would be encrypted using a K_c derived from Equation 8. The communication during this GSM based context would be compromised but communication that occurred before this point would be secure due to the hash in Equation 8 that creates the key K_c and communication after this context would be secure due to the fact that K_c would have been created from a hash and therefore the existing *CK* and *IK* can be used with confidence for future communications as no information on the existing *CK* and *IK* has been discovered.

5.2.3. UMTS Mobile Device with GSM BTS and MSC

In this context, once again the hash in Equation 8 protects *CK* and *IK* from the attacker and therefore all previous communication is secure and no significant knowledge of *CK* and *IK* is available to the attacker. K_c is still available to be compromised by an attacker in this configuration and therefore, when moving to another context from this context we will be creating a new *CK* and *IK* from Equation 11 that will make future communication secure.

6. Conclusion

Wireless network communication requires that user equipment be able to securely connect to the network and maintain integrity of that communication. In stationary networks there is no requirement for user equipment to be able to use all access points and to communicate while roaming between access points. Mobile networks have different requirement and legacy protocols needed to be integrated into new network systems.

To help manage the transition from the legacy SIM

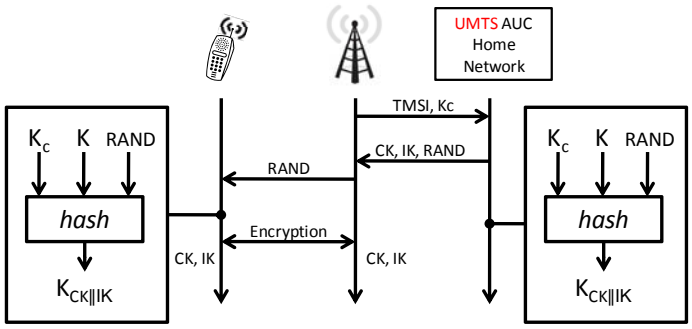


Figure 10. Request/Response to retrieve new CK and IK

based GSM system, protocols were devised to integrate the billions of existing devices into the new USIM based UMTS network. The integration protocols that allow for the integration of those legacy devices also inadvertently brought the insecurity of the GSM system into the new much more secure UMTS system. The GSM key K_c can be compromised and therefore, due to the method of integrating the two systems together which uses simple Equations 1, 4, and 5 to create the keys *CK*, *IK* and K_c used for encryption and integrity, an attacker that has discovered K_c can discern either all or part of *CK* and *IK*. This integration has allowed previous attacks on the GSM system to be effective against attacking the UMTS network negating the positive changes brought about by the mutual authentication in UMTS.

We have proposed two different changes to the protocols in mobile networks to protect against the legacy integration of GSM. One is a very simple change to the GSM protocol to protect K_c by creating K_h a hash of K_c shown in Equation 7 which is to be used when encrypting. This will protect K_c from attackers and therefore, protect the UMTS communication that depends on the keys devised from Equation 1, 4 and 5. The other change we propose is for the UMTS protocol to be modified to remove the equations 2, 5 and 6 used to generate *CK*, *IK* and K_c and replaces those equations with two Equations 8, and 11 which both use a hash function. We also create a simple request/response protocol to generate a new *CK*, *IK* pair generated from Equation 11 to be used in future communication. The changes we have proposed will help resolve the insecurity brought about by the legacy integration of the GSM equipment and protocols into the new UMTS system. The integration that was required due to the large and growing install-base of GSM devices.

Out of the two solutions proposed we recommend the solution of a GSM hash since it changes the protocol that has introduced the problems with a minimal amount of effort. GSM already has cryptographically strong hash functions available for use and should be able to be modified to do the single hash of the K_c value to increase the security of communication. The modification should be easily applied to UMTS devices in their support of the GSM protocols and add the increased security that the change would provide. The other advantage of this modification is that when the GSM protocols are no longer required in the future this change will then be removed as well making it much more self contained than the changes to the UMTS protocol that we propose.

7. References

[1] "GSM World statistics", GSM Association. 2010; http://www.gsma.com. (8 June 2010).

[2] Eric Southern, Abdelkader Ouda and Abdallah Shami, "Wireless Security: Securing Mobile UMTS Communications from interoperation of GSM", submitted to Special Issue on "Security in Wireless Ad Hoc and Sensor Networks with Advanced QoS Provisioning", Wiley Journal on Security and Communication Networks.

[3] GSM 02.09, "Digital cellular telecommunications system (Phase 2+); Security aspects", version 6.1.0, Release 1997.

[4] Scott Fluhrer, Itsik Mantin, and Adi Shamir. "Weaknesses in the key scheduling algorithm of RC4", Lecture Notes in Computer Science, 2259:1–24, 2001.

[5] A. Kerckhoffs, "La cryptographie militaire," Journal des sciences militaires, vol. IX, p. 538, Jan 1883.

[6] A. Biryukov, A. Shamir, and D. Wagner, "Real time cryptanalysis of a5/1 on a pc," in In FSE: Fast Software Encryption. Springer-Verlag, 2000, pp. 1–18.

[7] E. Barkan, E. Biham, and N. Keller, "Instant ciphertext-only cryptanalysis of gsm encrypted communication." Springer-Verlag, 2003, pp. 600–616.

[8] 3GPP, "Security Objectives and Principles," http://www.3gpp.org/ftp/Specs/html-info/33120.htm, 3rd Generation Partnership Project (3GPP), TS 33.120, (Apr. 2001).

[9] O. Dunkelman, N. Keller, and A. Shamir, "A practical-time attack on the a5/3 cryptosystem used in third generation gsm telephony," Cryptology ePrint Archive, Report 2010/013, 2010.

[10] G. Mapp, M. Aiash, A. Lasebae, and R. Phan, "Security models for heterogeneous networking," in Security and Cryptography (SECRYPT), Proceedings of the 2010 International Conference on, july 2010, pp. 1–4.

[11] G. Karsai, F. Massacci, L. Osterweil, and I. Schieferdecker, "Evolving embedded systems," Computer, vol. 43, no. 5, pp. 34 –40, may 2010.

[12] P. Vieira-Marques, S. Robles, J. Cucurull, R. Cruz-Correia, G. Navarro, and R. Marti, "Secure integration of distributed medical data using mobile agents," Intelligent Systems, IEEE, vol. 21, no. 6, pp. 47 –54, dec. 2006.

[13] P. Argyroudis, R. McAdoo, S. Toner, L. Doyle, and D. O'Mahony, "Analysing the security threats against network convergence architectures," in Information Assurance and Security, 2007. IAS 2007. Third International Symposium on, aug. 2007, pp. 241 –246.

[14] P. Trakadas, S. Maniatis, P. Karkazis, T. Zahariadis, H. Leligou, and S. Voliotis, "A novel flexible trust management system for heterogeneous wireless sensor networks," in Autonomous Decentralized Systems, 2009. ISADS '09. International Symposium on, march 2009, pp. 1 –6.

[15] 3GPP TS 33.102, "3G Security; Security architecture", Release 9, 2009.

[16] 3GPP TS 33.401, "3G security; Security architecture, 3GPP System Architecture Evolution (SAE); Security architecture", Release 11, 2011.