Wireless Security: Securing Mobile UMTS Communications from interoperation of GSM

Eric Southern, Abdelkader Ouda, Abdallah Shami

University of Western Ontario
Electrical and Computer Engineering
{esouther, aouda, ashami2}@uwo.ca

Abstract. Wireless communications have revolutionized the way the world communicates. An important process used to secure that communication is authentication. As flaws in the security of a wireless network are discovered new protocols and algorithms are required to meet those security issues. When creating new algorithms and systems it is possible that the existing equipment may not be able to implement the new protocols, which means that integration may be required to transition from an old security protocol to the new more secure protocol. Stationary wireless networks were created without a strong need to integrate protocols and have simply developed slightly more secure protocols to protect old equipment. New protocols in stationary wireless networks are implemented without integration as a requirement. Mobile wireless networks have the requirement of allowing old equipment to use the entire network as it is advantageous to allow new mobile equipment to connect to old networking equipment to increase coverage areas and for old equipment to be able to connect to new towers for roaming and billing. This requirement for mobile networks means that integration is required. There are flaws in this integration of GSM into UMTS networks. Those flaws are analyzed and two practical solutions are proposed.

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 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.

This work discusses authentication in wireless networks as well as the needs of those networks to interoperate and the security issues brought about by that integration. The authentication in 802.11 stationary wireless networks will be described with a focus on Wired Equivalent Privacy (WEP) security flaws and the related solutions offered by Wi-Fi Protected Access (WPA) and WPA2. The aim of this discussion is to give an understanding on how the authentication in the stationary wireless networks has changed due to the problems found in the earlier algorithms and protocols. In contrast with this 802.11 security replacement, the authentication protocols in the new generations of mobile wireless networks are designed to interoperate (not replace) the existing protocols. Therefore the integration of the different protocols to allow this interoperation will be described considering security flaws brought about by integrating the old protocols into the new systems. This will include GSM 2G networks and UMTS 3G networks, 4G networks are not discussed due to the fact that they use the same security protocols and algorithms as the 3G networks. In order to protect the integration in mobile wireless networks against these security flaws two solutions are proposed.

This paper is organized as follows. Section 2 describes the authentication protocols in 802.1. The authentication protocols in GSM 2G networks and UMTS 3G 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 mobile wireless networks in Section 5. Section 6 concludes this work.

2 Authentication in 802.11 Wireless Networks

To understand the security environment in mobile wireless networks it is worthwhile to review the security in stationary networks since both types of networks have undergone a phase of broken security and a migration of equipment from the less secure to more secure environments. Stationary wireless networks allow user equipment to connect to a network without the need of a physical wire. This allows for more user mobility and to create a network quickly and in environments where it is difficult or expensive to deploy physical networks. Generally, there is no need in these types of networks to manage the mobility of the user from one network access point to another as the connection does not need to be maintained if a user roams from one network area to another. The main difference for stationary networks is that the wireless users generally have modern or more powerful equipment that connects to the network and the network operator will generally have more control over all devices on the network. Stationary network providers did not have the same need to make their network allow access to old devices. Another major consideration in the evolution of security in stationary networks is that the equipment manufacturers were in control of the development and migration of the security framework and therefore did not have a strong vested interest in maintaining older hardware and would prefer to sell the new hardware that meets the new standard.

2.1 Wired Equivalent Privacy

The first type of security devised for wireless communication in the 802.11 standard is WEP. The algorithm relies on a shared Key (WEP key) of 40 bits or 104 bits as well as an Initiation Vector (IV) of 24 bits. As can be seen in Figure 1, WEP authentication process starts when a user equipment UE requests to associate with the access point AP, where UE must authenticate itself to the AP. Based on this request, AP sends a challenge nonce R (random number) to the UE, and waits for the response. The UE then encrypt the challenge R using a stream symmetric cipher RC4 as follows.

- The challenge R is first checksummed using CRC32 that is added to R to form the data payload.
- Then the UC creates a 24-bit random initialization vector (IV).
- The IV along with WEP key are used as a seed to generate RC4 key stream K.
- The ciphertext is produced by XORing the key stream K with the data payload.

UC then transmits the ciphertext and the IV to the AP as its response. The AP uses the IV that it received and the shared WEP key to decrypt the data and verify the checksum. If a match is found, the authentication declared successful and the association is formed.

Note that, the cryptosystem used in WEP is a stream symmetric cipher RC4, and the key that encrypts the data is the same key that will be used for decryption to recover the data.

Scott Fluhrer, et al. [1] described in their work titled "Weaknesses in the Key Scheduling Algorithm of RC4", number of weakness in WEP protocol. The flaws were related the way RC4 was implemented. They have mentioned that WEP can be cracked if enough traffic can be intercepted. This is because there are only 16 million possible IV's (24-bit), so after intercepting enough packets, there are sure to be repeats in the IV's. When IVs repeat, the RC4 key stream can be easily discovered and hence a known-plaintext attack can be utilized to recover the plaintext without the need for the WEP key. The end result is that WEP has suffered from key management problems, implementation errors, and overall weakness in the encryption mechanism.

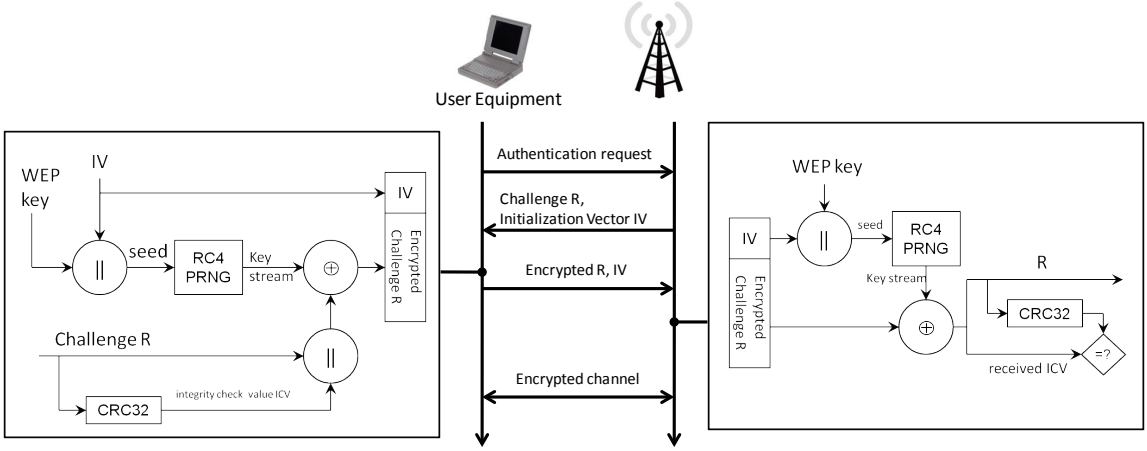


Figure 1. WEP authentication protocol

2.2 Wi-Fi Protected Access (WPA)

The major flaws in WEP made it necessary for the Wi-Fi Alliance to create a stronger protocol to increase the security of wireless networks without replacing the legacy hardware. There was a rush to create a more secure wireless network and therefore WPA was developed as a pre-standard 802.11i protocol that would be able to be loaded as an update to most WEP firmware and would improve the security of existing wireless networks until the 802.11i protocol could be ratified. WPA has the endorsement of the Temporal Key Integrity Protocol (TKIP) and Message Integrity Check (MIC) by the Wi-Fi Alliance. Authentication under WPA is completely different than that in WEP as shown in Figure 2.
The AP sends a random A-nonce to the UE. The UE takes the Pairwise Master Key (PMK), a pre-shared key given to the UE and AP, the received A-nonce, a generated S-nonce, along with AP and UE MAC addresses to compute a Pairwise Transient Key (PTK). This is done by using the Pseudo-Random Functions PRF-512. The PTK is then used to create a Hash-based Message Authentication Code (HMAC) created by the Message-Digest Algorithm (MD5) by giving the Key confirmation key (KCK) which is the first 128 bits of the PTK and the S-nonce as the input into the HMAC-MD5 algorithm. The S-nonce and produced MIC are then sent to the AP. The AP can perform the same PRF-512 done by the user equipment to generate the PTK and then use the PTK to verify the MIC. Once verified the AP will send an encapsulated Group Temporal Key (GTK) and MIC back to the UE for verification. The UE will then respond with an Acknowledgement of successful authentication. The PTK is also used to generate the Key Encryption Key (KEK) and the Temporal Key (TK). The KEK is used to encapsulate the GTK and other handshaking encryption and the TK is used for encrypting the communication over the link. The encryption in TKIP is done using RC4 similar to the encryption in WEP. The methodology used for the encryption of packets in TKIP greatly increases the security compared to WEP as the TK is constantly updated by the larger IV.

2.2 Wi-Fi Protected Access 2 (WPA2)

The Wi-Fi Alliance completed 802.11i as WPA2 to secure communication on wireless networks due to the weaknesses of WEP and WPA. The protocol relies on a shared key called the same Pairwise Master Key (PMK) generated in WPA which is designed to last the entire session and is exposed as little as possible. WPA2 uses the same four-way handshake to authenticate the user equipment (UE) to the access point (AP) and create keys for communication which can be seen in Figure 2. Similar to WPA using TKIP, WPA2 uses counter mode (CTR) with cipher-block chaining message authentication code (CBC-MAC) Protocol (CCMP) to perform many operations including securing the communication channel. There are some differences in the authentication between WPA and WPA2 such as the PRF used to generate the PTK in WPA2 is 384 bits. The MIC in the authentication is SHA-1. The encryption in CCMP uses the advanced encryption standard (AES). There are

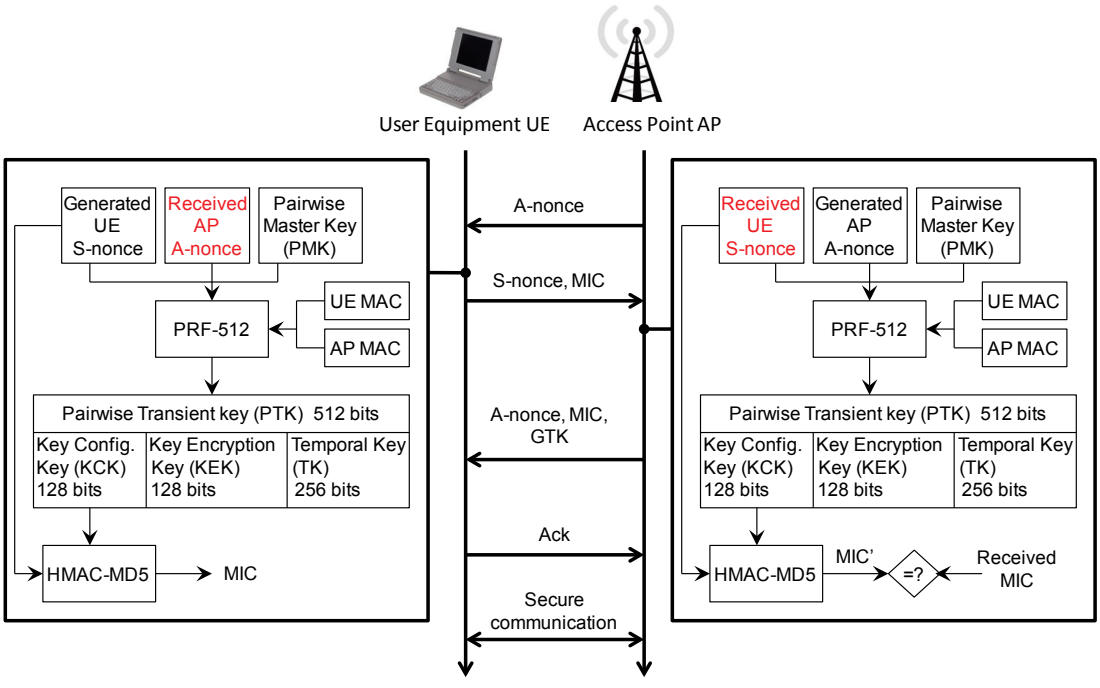


Figure 2. WPA authentication against the access point

major differences in the way the encryption is completed in CCMP compared to TKIP but those differences are not being investigated in this paper as we are focusing on authentication. The change to using the more secure SHA-1 for the MIC instead of MD5 creates a much more secure authentication.

The migration from WEP/WPA to WPA2 could be accomplished relatively quickly due to the fact that most mobile equipment (laptops and other powerful equipment) is upgraded frequently and has very few requirements to run on minimal resources. The migration of the network from WEP/WPA to WPA2 is handled by the network provider which was only limited by each organization mandate and could be accomplished when needed. Overall the cost of the upgrade has involved a massive replacement of equipment on a very large worldwide scale. The capacity of network devices has also grown with the migration from 802.11a to b to g to n, therefore, most providers would have upgraded their networks with the new technology and most users would upgrade their devices at the same time as well to make use of new computing power. The mobile networks have very different considerations when upgrading or integrating protocols. Mobile network operators have agreements with many other operators to allow almost any devices onto their network. To facilitate this requirement the network needs to operate in both 2G and 3G security contexts which we will show in the following section.

3 Authentication in Mobile Wireless Networks

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 forced the development of a network 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. This support limits the control a network provider has over the hardware connecting to their network. These networks also tend to be built out nationally, a very large investment, which needs to be leveraged as long as possible to have connectivity for all users. Some users are also likely to keep a functioning phone for a much longer time than a functioning laptop. GSM phones will operate as a worthwhile and functioning phone for more than a decade to many users that see no reason to upgrade their device.

3.1 GSM Authentication

The authentication in GSM is a one-way authentication algorithm to authenticate the mobile device to the service provider network. As shown in Figure 3 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

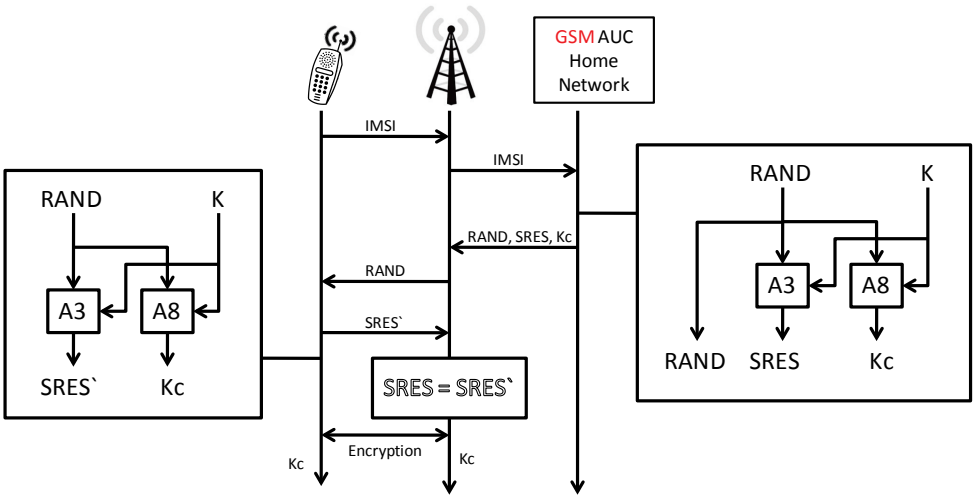


Figure 3. GSM Authentication

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 [2] (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 [3] and the A5/2 algorithm is easily broken in seconds [4] 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 (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.

3.2 UMTS Authentication

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 4. 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 [5] 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 [5] that was chosen for its suitability for implementation in hardware. The algorithm has some weaknesses but is not susceptible to real-time attacks [6]. Currently the ITU (International Telecommunication Union) is still developing the standards for 4G mobile communications but the authentication protocols are the same as those of the UMTS network [7].

4 Legacy Integration of Authentication Protocols

To make use of existing hardware and equipment it may be required to update protocols or create an integration protocol due to the cost or effort required to replace the equipment. Network providers for each type of network need to adapt to the changing security environment. The 802.11 protocols exist side by side on the same equipment, which allowed network providers to stage the upgrading of their networks to the new protocols. The protocols in mobile networks were integrated to allow for maximum use of existing equipment while roaming and to give users and providers as much flexibility as possible when connecting to mobile networks.

4.1 WPA as an Upgrade to Legacy WEP Equipment

WPA was created as a measure to increase the security of WEP equipment. WPA does not address all of the security flaws in WEP but was able to be installed on equipment that supports WEP allowing network administrators to increase the security of their existing equipment. The new WPA protocol can be attacked and broken with only slightly more difficulty than WEP. The overall network would only be as secure as the weakest connection which means that the existing WEP and WPA equipment would need to be segregated from the rest of a secure WPA2 network until the equipment can be upgraded. WPA did give network operators a significant increase in the security of their networks while they could prepare to deploy new WPA2 equipment. Most new 802.11 equipment will support any of the security protocols depending on the needs of the network operator, allowing them to stage their integration of the new WPA2 security protocol to meet the needs of their users. Mobile providers could also stage the upgrade of their networking equipment but will have an extended period of time where they will be hampered by the needs of users to connect using the GSM protocols. The integration also allowed providers to use their existing networks to provide service to the new UMTS devices by having them implement the GSM protocol, which will be shown in the next section.

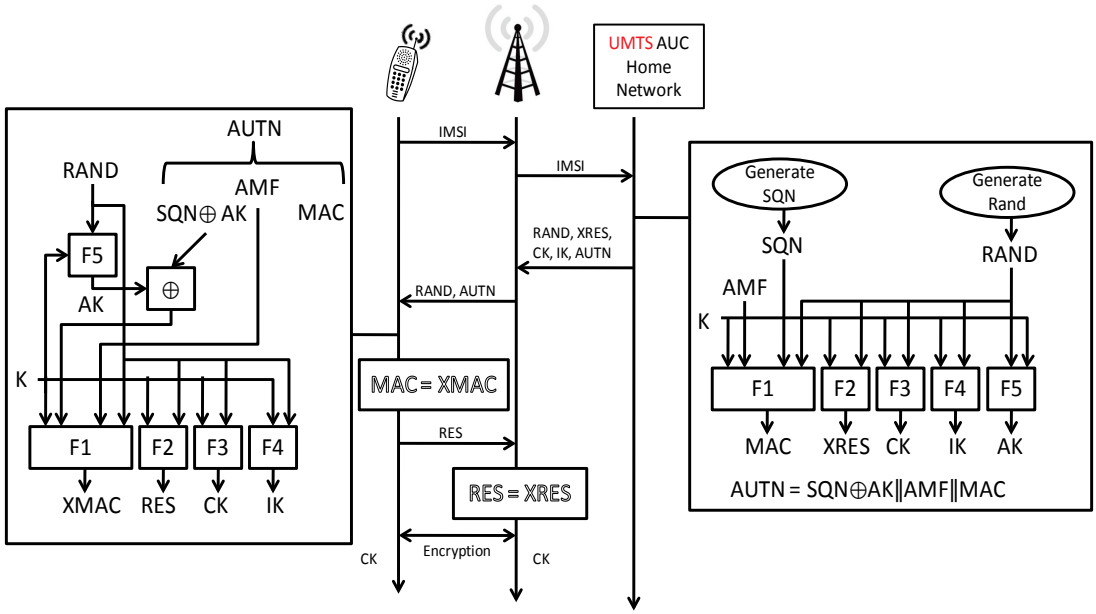


Figure 4. UMTS Authentication

4.2 Legacy Integration of GSM with UMTS

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 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 [12]. 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:

$$K_c = CK_1 \oplus CK_2 \oplus IK_1 \oplus IK_2 \tag{1}$$

$$where, \quad CK = CK_1 \parallel CK_2 \tag{2}$$

$$and \ IK = IK_1 \parallel IK_2 \tag{3}$$

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

$$CK = K_c \parallel K_c \tag{4}$$

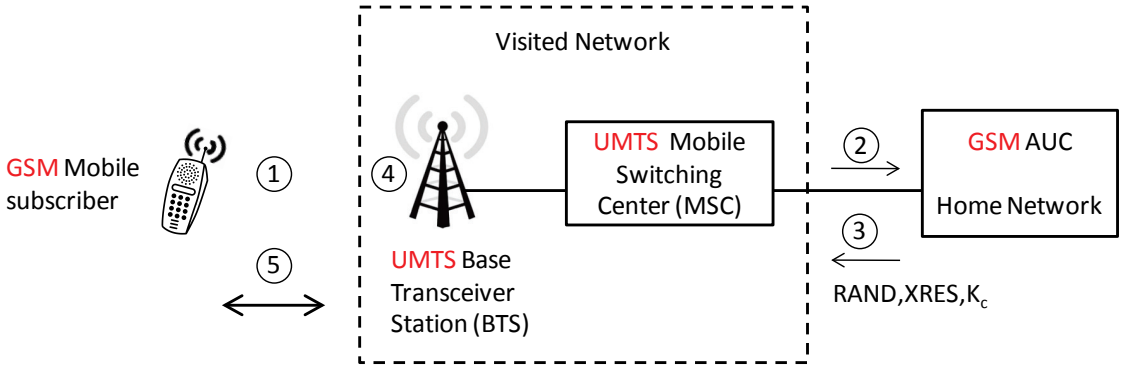
$$IK = K_{c1} \oplus K_{c2} \parallel K_c \parallel K_{c1} \oplus K_{c2} \tag{5}$$

$$where, \quad K_c = K_{c1} \parallel K_{c2} \tag{6}$$

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.

4.2.1 GSM Mobile Device with UMTS Network

When a GSM Mobile device is on a UMTS network as shown in Figure 5, and as per the order of 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 3.1 and Figure 3 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*.



(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

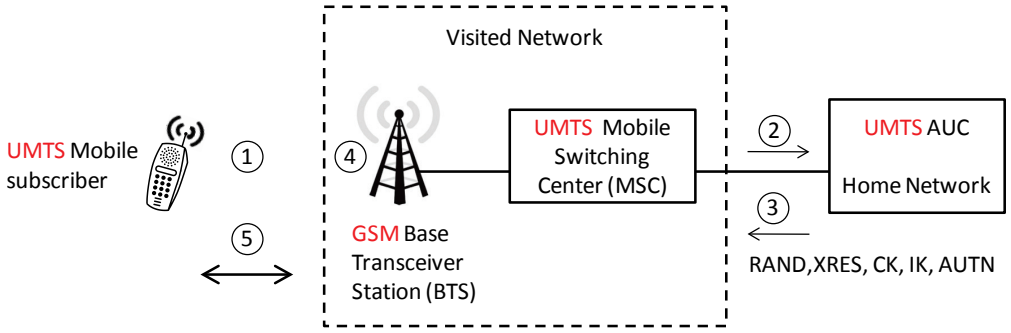
Figure 5. The GSM Mobile subscriber is authenticated via a UMTS BTS, which is connected to a UMTS MSC

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.

4.2.2 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 6, 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 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 3.1 and Figure 3 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*.



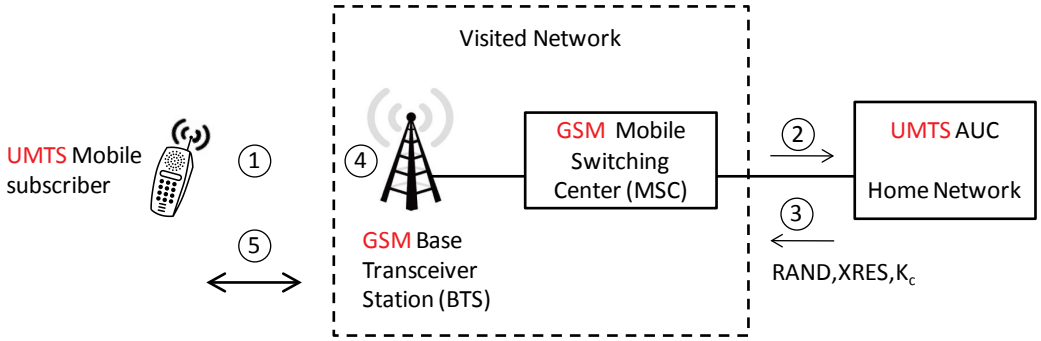
(1) UMTS Mobile subscriber requests a secure connection to GSM BTS
(2) UMTS MSC requests from the UMTS home network the authentication vector (RAND,XRES, CK, IK, AUTN).
(3) UMTS MSC receives the UMTS authentication vector and proceeds to generate a GSM K_c and forwards K_c to the GSM BTS
(4) GSM BTS 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 6. The UMTS Mobile subscriber is authenticated via a GSM BTS, which is connected to a UMTS MSC

4.2.3 UMTS Mobile Device with GSM BTS and MSC

Figure 7 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 3.2 and Figure 4 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



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

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 8 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.

5.1.1 GSM Mobile Device with UMTS Network

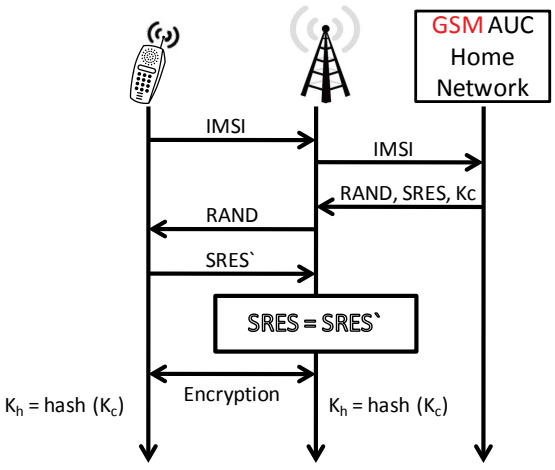


Figure 8. Proposed modification to GSM Protocol

Figure 9 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 7), 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:

$$K_c = CH_1 \oplus CH_2 \oplus IH_1 \oplus IH_2 \tag{8}$$

$$where, \quad hash(CK) = CH_1 \parallel CH_2 \tag{9}$$

$$and \quad hash(IK) = IH_1 \parallel IH_2 \tag{10}$$

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

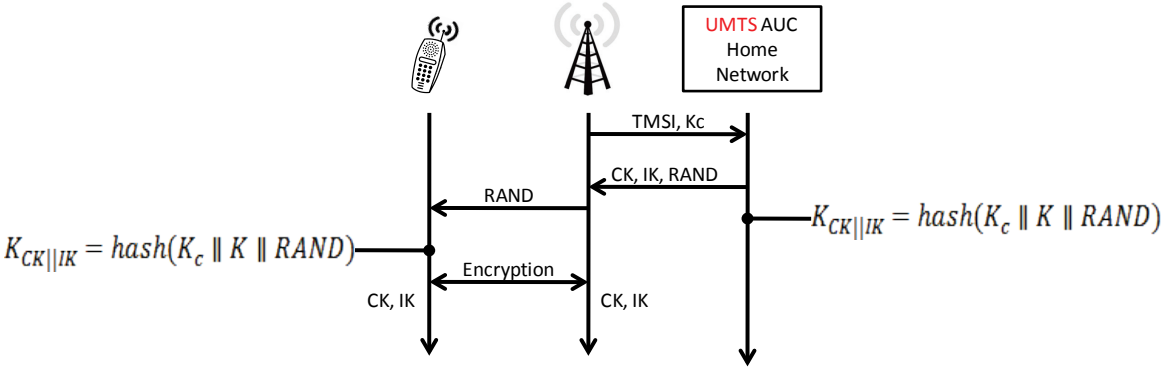


Figure 9. Request/Response to retrieve new CK and IK

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 7. 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:

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

$$where, \quad K_{CK||IK} = CK \parallel IK \tag{12}$$

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 GSM system, protocols were devised to integrate the billions of existing devices into the new 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.

References

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

[2] A. Kerckhoffs, “La cryptographie militaire,” Journal des sciences militaires, vol. IX, p. 538, Jan 1883.

[3] 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.

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

[5] 3GPP, “Security Objectives and Principles,” 3rd Generation Partnership Project (3GPP), TS 33.120, Apr. 2001. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/33120.htm

[6] 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, http://eprint.iacr.org/.

[7] 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.

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

[9] 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, nov.-dec. 2006.

[10] 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.

[11] 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.

[12] 3GPP, “3G security; Security architecture,” 3rd Generation Partnership Project (3GPP), TS 33.102, Jun. 2008. [Online]. Available: http://www.3gpp.org/ftp/Specs/html-info/33102.htm