Signaling-oriented DoS Attacks in UMTS Networks

[1]Georgios Kambourakis, [1]Constantinos Kolias, [1]Stefanos Gritzalis,
[2]Jong Hyuk-Park

[1]Laboratory of Information and Communication Systems Security
Department of Information and Communication Systems Engineering
University of the Aegean, GR-83200 Samos, Greece
{gkamb, kkolias, sgritz}@aegean.gr

[2]Department of Computer Science and Engineering,
Kyungnam University, Korea
parkjonghyuk1@hotmail.com

Abstract. The Universal Mobile Telecommunication Standard (UMTS) is the Third Generation (3G) mobile technology with the widest public acceptance. Although, enhanced in matters of security, comparing to its predecessor i.e., the GSM, it still has vulnerabilities that can lead to security breach. In this paper we investigate the vulnerabilities of the UMTS architecture that can be exploited by a malicious entity to launch Denial of Service (DoS) attacks. We examine the methodologies that an attacker would possibly follow, as well as the possible outcome of such class of attacks. We also give some suggestions that would provide greater tolerance to the system against DoS attacks.

Keywords: UMTS; Denial of Service; Signaling; Security.

1 Introduction

Beyond doubt, handheld devices have changed the modern way of communication and information access. The increasing demand for high quality voice services along with the need for modern pervasive applications has given birth to the Universal Mobile Telecommunication System (UMTS). UMTS is the outcome of a collaborative effort of many international organizations gathered around the 3rd Generation Partnership Project (3GPP) consortium [1]. Today, 3rd Generation (3G) mobile networks based on the UMTS standard are deployed in Europe and USA (3GPP2) with great success [2]. Users of these networks benefit from the higher quality of voice and video calls, higher transfer rates, communication with the internet, and enjoy advance applications and value-added services such as e-commerce, e-banking etc. In the years to come, most people will use their handheld device to make wireless security-sensitive transactions like e-banking, stock trading, and shopping. Therefore, with the introduction of such new applications to the mobile world, security, now more than ever, is a crucial aspect. Nevertheless, the inherited weaknesses of the UMTS that derive mostly from its wireless nature and Second Generation (2/2.5G) networks make it prone to a substantial number of security

threats. That is, even though UMTS is characterized by many security enhancements comparing to its 2G predecessor the GSM it still presents architectural weaknesses that render it vulnerable to several security threats.

The primary target of the designers of UMTS was to maintain maximum compatibility with the 2G systems. Additionally, its designers took into account the constraints in computational power of the various mobile devices, and for that reason they adopted relatively lightweight security techniques, such as symmetric encryption [3]. Until now, the majority of research in UMTS has focused on ways to preserve the privacy and confidentiality of the end users [4-6]. Although privacy and confidentiality are always of top priority in any wireless system, we believe that the availability of the services should not be neglected. Unfortunately, UMTS in its current form makes it easy for Denial of Service attacks (DoS) to be launched.

A DoS attack is the type of attack preformed by a malicious entity in order to render a service unavailable to its intended users. Numerous attack incidents verify the catastrophic potential of this class of attacks [7], and several researchers characterize DoS attacks as the second most dangerous threat after viruses. The methodology and target of a DoS attack may vary spanning from simple DoS to well orchestrated distributed attacks able to paralyze entire network infrastructures. While this type of attacks has its roots on the Internet realm, its philosophy and purpose has derived to the GSM networks and lately to UMTS, since wireless communications offer a new challenging terrain for attackers.

In this paper we particularly focus on signaling –oriented DoS attacks that can be launched against UMTS systems. We point out architectural and protocol vulnerabilities that can be exploited to unleash such attacks and give directions for possible improvements. The remainder of the paper is structured as follows: the next section gives background information regarding UMTS security architecture. Section 3 points out certain UMTS system vulnerabilities and discusses requirements and methodologies that can be exploited by an aggressor to achieve DoS. Section 4 presents our suggestions and gives pointers to future work. The last section draws a conclusion.

2 UMTS security architecture

The UMTS security architecture defines a set of procedures that the user's mobile equipment as well as the network should execute in order to receive increased confidentiality and integrity during their communication. In the heart of the UMTS security architecture lies the user authentication mechanism known as Authentication and Key Agreement (AKA) [8]. This mechanism is somewhat similar to the authentication in GSM. The idea to use public keys in the process of authenticating the users, was abandoned, mainly due to backwards compatibility (with GSM) and for performance considerations. The authentication in UMTS is based on a 128-bit symmetric secret key, namely Ki, which is stored in the user's tamper-resistant Universal Integrated Circuit Card (UICC) and in the corresponding Home Subscriber Server (HSS) of the user's Home Network (HN). The AKA scheme is a combination of the well known challenge response-protocol found in GSM and the authentication

mechanism based on sequence number as defined by the ISO organization [9]. The network entities that take part in the user's authentication procedure are:

- The User's Equipment (UE) and more specifically the USIM application stored in the UICC.
- The Serving GPRS Support Node (SGSN) of the HN or the Serving Network (SN).
- The HSS of the user's HN.

The authentication procedure in UMTS is mutual, which means that both the network is authenticated to the UE and the UE is authenticated to the network. After successful authentication the two ends agree on the use of two additional 128-bit symmetric keys. These keys are derived from the master key Ki and renewed every time the user is authenticated. The procedure typically initiates after the MS attaches to the network and sends its identity. Note, that the user can be identified either by a permanent ID, i.e., the International Mobile Subscriber Identity (IMSI) or, usually, a temporary one known as Temporary Mobile Subscriber Identity (TMSI). During the process, the user's ID is forwarded from the Radio Access Network sub-network to the core network, that is, the SGSN serving that particular area. In any case, the latter entity may send an authentication data request message to the HSS of the user's HN in order to acquire Authentication Vectors (AV) required to authenticate the user. This happens only in cases that no AV for that particular user is available locally in the SGSN. For instance, the user attaches for the first time to this SGSN or the available in the SGSN AVs for that user have been already consumed. Since the HSS possesses the master key (Ki) for each user is capable of creating the corresponding Authentication Vectors (AV). The vectors are sent back to the SGSN in charge by making use of a control message known as *authentication data response*. A vector can be used only once except the case the SGSN does not receive an answer from the MS.

After the SGSN in charge acquires some AVs (they are sent usually in batch), it sends an authentication request to the user. The request contains two parameters: (a) a RAND which is a random number and, (b) the AUTN, i.e., the authentication token. These parameters are transferred in the tamper resistant environment of the UICC/USIM and stored for further processing.

The USIM is also aware of the Ki, and uses it along with the received parameters RAND and AUTN to perform a series of calculations similar to those that took place for the generation of the corresponding AV in the HN's HSS. The outcome of this procedure enables USIM to verify that the AUTN parameter was indeed created by the HSS of the HN and also that it is fresh (i.e., it is not a message replay). In case that the above verifications have a positive outcome the RES (result) parameter is calculated and sent back to the corresponding SGSN by utilizing a user authentication response message. Upon that, the SGSN compares the received RES with the XRES (Expected Response) which is contained in the initial AV. If the two values match then the user is granted access to the network.

Moreover, as already mentioned, two other keys that will be used for confidentiality and data integrity are calculated by the USIM. Using a security mode command the same keys, which are contained in the initial AV, are transmitted by the

SGSN to the corresponding Radio Network Controller (RNC). These keys are known as CK (cipher key) and IK (Integrity Key). Note that while these keys are part of the corresponding AV and thus immediately available to the SGSN, the USIM has to calculate them by itself. An overview of the authentication sequence described above is depicted in Figure 1.

It is to be noted that this section presents only basic information on UMTS security architecture. For a more detailed analysis the reader may refer to [1,8].

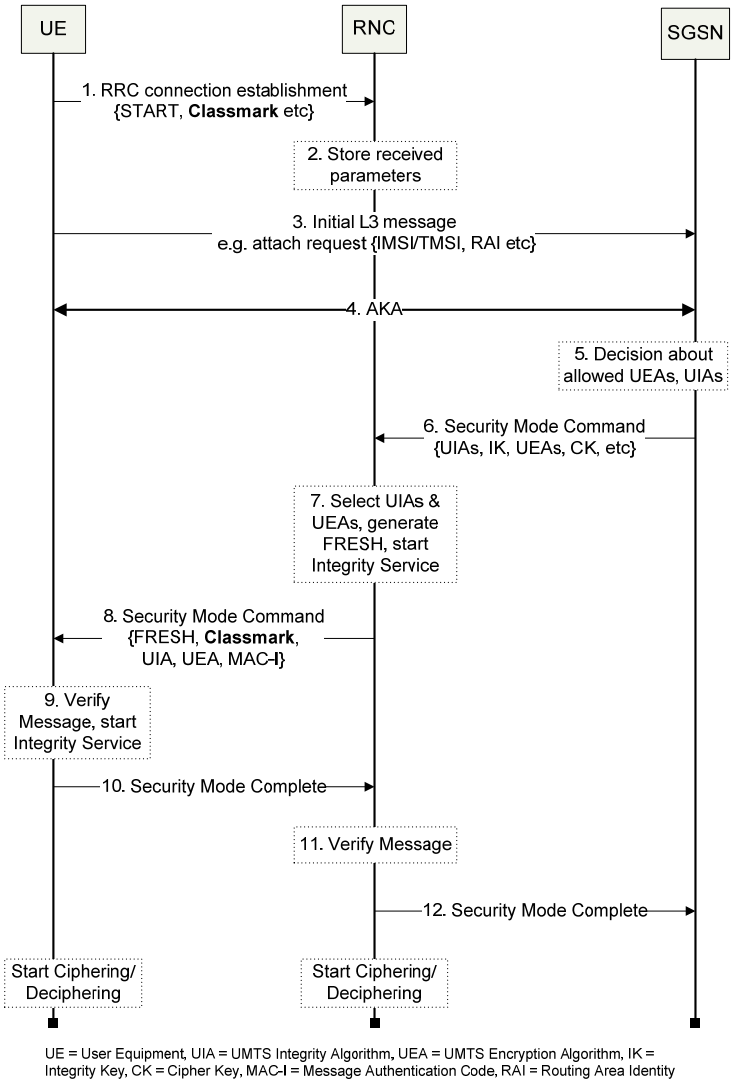

Fig. 1 Start of Security services in UMTS

3 DoS attacks in UMTS

In this section we shall describe some vulnerabilities of the UMTS architecture which can be exploited to launch DoS or more generally DoS-type attacks. Also take into account that this paper considers only signaling-oriented DoS attacks. Typically, an attacker would seek unprotected control messages which would attempt to modify in order to manipulate specific procedures or make them repeat. The expected outcome varies: from lower quality of service (QoS) that a specific user may experience to a massive denial of any underlying service. For example, in [10] the authors identify a critical vulnerability to UMTS architecture and exploit it to perform an HLR flooding. This is achieved by modifying a single message.

In the attacks described below the attacker carries some special equipment, e.g., a false Base Station (BS) and/or a specially modified UE) with the help of which are able to perform as a man-in-the-middle entity. Having such equipment the attacker must be able to intercept a valid UE-to-BS session, analyze traffic, and modify the data of UMTS frames. Also in some cases it is important for the attacker to build a database of valid (intercepted) IMSIs. Research on the field [4,10] proves that this is a relatively straightforward procedure and, in some cases, requires equipment which is easy to obtain or self-fabricate.

A very simple but primitive DoS attack unfolds as follows: An attacker with a false BS equipment moves close to its target victims. All users's mobile terminals will deceived into connecting to the false BS if its signal is stronger than the legitimate BS. After the victim is connected to its fake equipment the attacker would simply drop every packet that is transmitted from and towards the UE. This is usually described as a black hole attack and could be considered as the higher layer equivalent of radio jamming. UMTS security architecture in its current form is not able to counteract these types of attacks [9]. On the other hand, an attacker would rarely adopt such methods to launch DoS attacks because: (a) the attack persists only when the attacker is active, (b) it affects only a small number of users, and (c) it cannot be directed to inflict specific targets (users) only, without affecting others as well. For these reasons it is likely that an attacker would seek more intelligent ways of launching DoS attacks. Hereunder we shall elaborate on more sophisticated attacks.

3.1 Dropping ACK signal

The protection of IMSI is considered a very important issue in UMTS. Therefore, an effort has been made by the designers of system in order for the IMSI to be transmitted and used as seldom as possible. Instead, as already mentioned, temporary identities known as TMSIs are distributed to the users and thereafter are used for all signaling communication. TMSIs are assigned to users, right after the initiation of ciphering. Also new TMSIs are assigned every time a user roams to an area monitored by a different SGSN. Although, a TMSI is transmitted encrypted to the UE the SGSN does not associate the IMSI with the TMSI unless it receives a TMSI Allocation Complete message from the MS. If this message never reaches the intended SGSN then both the associations {IMSI, TMSI_old} and {IMSI, TMSI_new} are considered valid by the SGSN in charge for uplink communication and the UE is free to use any

of them. Contrariwise, for the downlink, the IMSI must be used because the network has no means to know which one of TMSI_new or TMSI_old is valid at the UE side at this particular moment. In this case, the SGSN will instantly instruct the mobile station to delete every available TMSI. In either of the two cases the network may initiate the normal TMSI allocation procedure. Of course, repeated failure of TMSI reallocation may be reported for further maintenance actions by the provider.

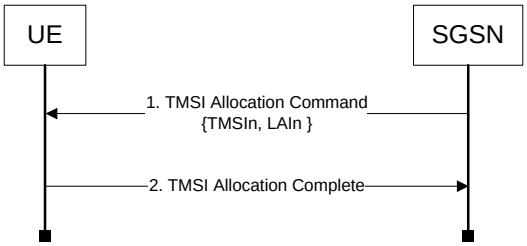

Fig. 2 TMSI allocation procedure

In such an attack, the aggressor might wish to position his equipment to a strategic location, for instance circumferential to a given network cell (where typically new TMSIs are assigned to subscribers entering the cell after a hand-off). Then, he would monitor for *TMSI Allocation Command* messages and then drop any following *TMSI Allocation Complete* message as depicted in Figure 2. This would cause new TMSIs to be created repeatedly, which would be expressed as DoS to all the users entering the particular routing area. Although the creation of a new TMSI is a multi-step procedure it cannot be considered resource demanding. So, extending this attack to become a flooding attack is considered rather difficult. This attack however can be used to expose and collect a large number of IMSIs and then use them to launch more dangerous and persistent attacks like the one described further down in section 3.3.

3.2 Modifying unprotected RRC messages

The Radio Resource Control (RRC) messages are considered vital for the smooth and normal operation of the UMTS system. Therefore, these signaling information messages are protected by integrity mechanisms, i.e., by applying a message authentication function. While this is true for most of the RRC messages exchanged between a mobile station and the corresponding RNC, many messages exist that are not integrity protected and therefore are vulnerable to manipulation. Table 1 presents some of the unprotected RRC messages. This might happen either because these messages are exchanged during the early stage of a connection - where the AKA procedure has not yet completed and thus an IK is not present -, or for reasons of efficiency.

Modifying, dropping or substituting unprotected RRC messages is expected to cause general system instability, or at least commotion, which may lead to lower QoS or more probably DoS for the end user. Theoretically, the ways and possibilities to stress the system with this method are many. Let us consider the following example: an attacker would insert an *RRC Connection Release* message during a valid ongoing

session. By acting the same way, an attacker could substitute a valid *RRC Connection Setup Complete* with a *RRC Connection Reject* message.

Table 1. List of unprotected RRC messages

Handover to UTRAN Complete
Paging Type 1
Push Capacity Request
Physical Shared Channel Allocation
RRC Connection Request
RRC Connection Setup
RRC Connection Setup Complete
RRC Connection Reject
RRC Connection Release
System Information (Broadcast Information)
System Information Change Indication
Transport Format Combination Control (TM DCCH only)

3.3 Modification of the initial security capabilities of MS

This is an extension of an attack already proposed in [10]. It involves the modification of an *RRC Connection Request* message and more specifically the field which defines the UE security capabilities. This message is not integrity protected since the MS and SGSN do not share a common IK yet. This happens because the AKA procedure takes place at a later stage. Any modification of this message will go unnoticed until eventually the AKA procedure completes and the *Security Mode Command* message is sent to the MS. This message includes the user's equipment security capabilities as received from the *RCC Connection Request* message in order to be verified by the UE. In case of mismatch the connection will terminate, but during the process sufficient resources will have been already consumed at both sides.

In such a scenario, the attacker aims to overstress the system by inducing a heavyweight repeating procedure. If successful, the system may become paralyzed being incapable of serving its legitimate users. Moreover, if the attacker has a large database of stolen IMSIs at hand he would be able to cause a much more serious damage compared to that caused by a single IMSI as the authors propose in [10]. By utilizing the proper equipment the attacker could create a very large number of simultaneous connection requests with bogus classmarks, thus causing steps 1 to 9 of Figure 1 to constantly repeat. Obviously, this would stress the system since many heavyweight procedures both bandwidth and computationally intense would a take place at the same time and for a large period of time.

3.4 DoS by modifying periodic authentication messages

Periodic local authentication in UMTS is a procedure meant to provide an additional security mechanism. Potentially, it can provide some sort of integrity protection in the U-plane. According to this procedure the volume of data transmitted during the RRC connection is periodically checked by both the RNC and the UE. The system makes

use of two variables to keep track of the user data transmitted from the mobile station towards the network. The first one namely Count-C_{UE} tracks the volume of user data transmitted by the user equipment, while the other, known as Count-C_{RNC}, stores the volume of user data actually received by the corresponding RNC. The value of these variables is cross-checked at regular intervals upon initiation by the RNC in charge. If a significant inconsistency is found then the RNC may decide to abruptly release the connection assuming that someone is injecting or dropping messages on the network. Assuming that the network provider supports this option, the aforementioned procedure is normally triggered when the value of the Count-C_{RNC} variable reaches a predefined limit.

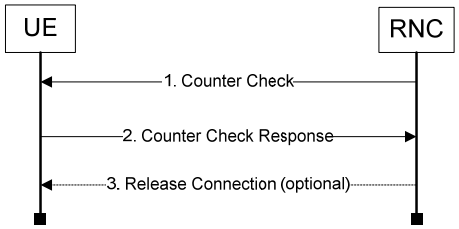

Fig. 3 UMTS Periodic authentication procedure

When this threshold is reached, the RNC sends a Counter Check message which contains the most significant bits of Count-C of each active radio bearer. The user equipment compares the Count-C value(s) received from the RNC with its local value of any matching active radio bearer, computes the difference, if any, and constructs a *Counter Check Response* message containing all differences. If one or more of the values contained in the *Counter Check Response* message is greater than 0 (null) then the RNC may decide to send a *Release Connection* message. Otherwise the procedure completes successfully. The above procedure is depicted in Figure 3.

According to the UMTS specifications [1] all the messages described above are integrity protected. As a result, an attacker is not able to modify any of these messages (for example change the value(s) contained in a *Counter Check Response* message). If so, the system (RNC) will notice that the received message has been somehow tampered. At this point however, the specification does not define any steps that should be made if such a situation occurs. This would be interpreted as releasing the connection or waiting indefinitely for a valid *Counter Check Response* message to arrive. Without doubt, this issue is provider-specific which of course leaves room for possible errors or misconfigurations.

4 Suggestions and Future work

In any case the number of signaling messages that do not afford an integrity service must be limited. Signaling takes place at three different layers, i.e., RRC, Radio Link Control (RLC) and Medium Access Control (MAC). However, RRC layer signaling is the most sensitive one thus its integrity is protected by using the IK. On the other hand, RLC and MAC signaling is protected by means of encryption. Consequently, threats to signaling do exist especially for messages preceding the AKA procedure. In

this context, an integrity mechanism should exist to protect all message exchanges before the IK is in place. As discussed in section 3.2 all RRC messages should also be integrity protected; otherwise the attacker is equipped with the simplest means to launch a simple but effective DoS attack.

Our ongoing and future work concentrates on two issues. First off, find an alternative way to provide an integrity mechanism for protecting the network against flooding attacks. Our intension is not to replace or patch the standard UMTS integrity protection but to provide a simple method to safeguard signaling before AKA execution. In this direction we are examining some variations of the client puzzle scheme [12-15]. This mechanism requires that every client (e.g., a mobile station) would have to solve a small cryptographic puzzle upon requesting services from the network. The basic idea is that the client should commit some of its resources first (do some cryptographic functions that require computational resources) before the server commits its own. The puzzles should be easy for the server to verify so that the server can do this process massively; while at the same time be computationally inefficient for the client for large numbers. By doing so, a potential attacker would be discouraged to massively make new connection requests.

At the same time, we also working on the kind of actions, in terms of protocols, that should be executed when malicious traffic injects into the network, e.g., the received messages systematically do not pass the underlying integrity controls.

5 Conclusions

Several known weaknesses in GSM seem to be now fixed in UMTS, through further study and investigation. Mobile station to network mutual authentication, stronger confidentiality provided in the U-plane, and the protection of signaling messages integrity seem to overhaul certain GSM security gaps towards making mobile communications safer, trustworthy, and thus, more attractive to consumers. Nevertheless, this might not prove adequate against serious attackers since several flaws are documented in the literature. In this paper we introduced some additional flaws that can be relatively easy exploited by attackers to launch dangerous DoS attacks. The inner workings of such an attack capitalize mostly on weaknesses found in signaling to achieve its goals. So, giving the fact that attackers become more and more resourceful there is an urgent need for more effective and carefully designed DoS countermeasures. This will allow the systems to deliver smooth and quality services to their subscribers.

References

1 3GPP Organization, http://www.3gpp.org/, accessed on 13/01/2008.

2 3rd Generation Partnership Project 2, 3GPP2, http://www.3gpp2.org/.

3 Kazumi Algorithm Specification, ETSI TS 135 202 V7.0.0, http://www.etsi.org/website/document /algorithms/ts_135202v070000p.pdf, accessed on 13/01/2008.

4 C. Tang, D.O. Wu, “Mobile Privacy in Wireless Networks-Revisited”, IEEE transactions on the wireless communications, vol. 7, no. 3, pp. 1035-1042, March 2008.
5 U. Meyer, S. Wetzel “A Man-in-the-Middle Attack on UMTS”, WiSe’04, October 1, Philadelphia, Pennsylvania, USA, 2004.
6 Yi-Bing Lin, Ming-Feng Chang, Meng-Ta Hsu, and Lin-Yi Wu, “One-Pass GPRS and IMS Authentication Procedure for UMTS”, IEEE Journal on selected areas in communications, Vol. 23(6), June 2005.
7 Gibson, S., “DRDoS Distributed Reflection Denial of Service”, http://grc.com/dos/drdos.htm, 2002.
8 ETSI TS 133 102 “Security architecture” December 2006.
9 ISO/IEC 9798-4 (1999). Information Technology; Security Techniques; Entity Authentication Part 4: Mechanisms using a cryptographic check function, 1999.
10 M. Khan, A. Ahmed, A.R. Cheema, “Vulnerabilities of UMTS Access Domain Security Architecture”, 9th ACIS International Conference on Software Engineering, Networking, and Parallel/Distributed Computing, Phuket, Thailand, Aug. 2008.
11 3GPP TR 33.900 (1.2.0), “A Guide to 3G Security”, Jan., 2000.
12 W. Feng, E. Kaiser, W. Feng, and A. Luu, “The design and implementation of network puzzles,” in Proc. INFOCOM, 2005.
13 V. Gligor, “Guaranteeing access in spite of service-flooding attacks,” in Proc. Security Protocols Workshop, 2003.
14 X. Wang and M. Reiter, “Defending against denial-of-service attacks with puzzle auctions,” in Proc. IEEE Security and Privacy, pp. 78–92, 2003.
15 B. Waters, A. Juels, J. Halderman, and E. Felten, “New client puzzle outsourcing techniques for DoS resistance,” in Proc. Computer and Communications Security, pp. 246–256, 2004.