Retrofitting Mutual Authentication to GSM using RAND Hijacking

Mohammed Shafiul Alam Khan* and Chris J Mitchell

Information Security Group, Royal Holloway, University of London
Egham, Surrey TW20 0EX, United Kingdom
shafiulalam@gmail.com, me@chrismitchell.net

Abstract. As has been widely discussed, the GSM mobile telephony system only offers unilateral authentication of the mobile phone to the network; this limitation permits a range of attacks. While adding support for mutual authentication would be highly beneficial, changing the way GSM serving networks operate is not practical. This paper proposes a novel modification to the relationship between a Subscriber Identity Module (SIM) and its home network which allows mutual authentication without changing any of the existing mobile infrastructure, including the phones; the only necessary changes are to the authentication centres and the SIMs. This enhancement, which could be deployed piecemeal in a completely transparent way, not only addresses a number of serious vulnerabilities in GSM but is also the first proposal for enhancing GSM authentication that could be deployed without modifying any of the existing network infrastructure.

Keywords: GSM, mutual authentication, SIM application toolkit, RAND

1 Introduction

This paper proposes a way of adding network-to-phone authentication to the GSM mobile phone system, in a way that is completely transparent to the existing network infrastructure. Currently, GSM only supports authentication of the phone to the network, leaving the system open to a wide range of threats (see, for example, [21]). Despite the introduction and deployment of 3G (UMTS) and 4G (LTE) mobile phone systems, which rectify the GSM problem by providing mutual authentication between phone and network, GSM remains of huge practical importance worldwide and is not likely to be replaced for many decades to come. As a result, finding ways of improving the security offered by GSM, without the need for changes to the deployed phones and access networks, is clearly of great practical significance. This observation motivates the work described in this paper.

It is somewhat counterintuitive to propose that authentication of the network to the phone can be achieved without modifying the way in which the existing

* The author is a Commonwealth Scholar, funded by the UK government.

network and phones operate. This apparently paradoxical result is achieved by using a technique we refer to as *RAND hijacking*. This involves using the *RAND* value, which serves as a nonce in the existing unilateral authentication protocol and is sent from the network to the phone, to contain data which enables the recipient SIM to verify its origin and freshness. That is, the *RAND* is hijacked to act as a communications channel between a home network and a SIM.

The remainder of the paper is structured as follows. Key facts about the GSM network, including details of the operation of the GSM authentication and key establishment (AKA) protocol, are given in Sect. 2. This is followed in Sect. 3 by an introduction to the notion of *RAND hijacking*. In Sect. 4, the novel enhanced version of the GSM authentication scheme is described, and Sect. 5 describes how the SIM can use the results of the network authentication to affect UE behaviour. An analysis of the novel system is provided in Sect. 6. The relationship of the proposed scheme to the prior art is discussed in Sect. 7, and the paper concludes in Sect. 8.

2 GSM

2.1 Terminology

We start by providing a brief overview of key terminology for mobile systems. We focus in particular on the GSM network, but much of the description applies in slightly modified form to 3G and 4G networks. A more detailed description of GSM security features can, for example, be found in Pagliusi [22].

A complete mobile phone is referred to as a *user equipment (UE)*, where the term encapsulates not only the *mobile equipment (ME)*, i.e. the phone, but also the *subscriber identity module (SIM)* within it, where the SIM takes the form of a cut-down smart card. The SIM embodies the relationship between the human user and the issuing *home network*, including the *International Mobile Subscriber Identity (IMSI)*, the telephone number of the UE, and other user (subscriber) data, together with a secret key shared with the issuing network which forms the basis for all the air interface security features.

To attach to a mobile network, a UE connects via its radio interface to a radio tower. Several radio towers are controlled by a single *radio network controller (RNC)* which is connected to one *mobile switching center/visitor location register (MSC/VLR)*. The MSC/VLR is responsible for controlling call setup and routing. Each MSC/VLR is also connected to the carrier network's *home location register (HLR)* where corresponding subscriber details can be found. The HLR is associated with an *authentication center (AuC)* that stores cryptographic credentials required for communicating with the SIM; specifically, the AuC shares a unique secret key K_i with each SIM issued by the network to which it belongs. The RNC and the MSC/VLR are part of the *visiting/serving network* whereas the HLR and the AuC are the *home network* component.

2.2 GSM Authentication Protocol

To prevent unauthorised mobile devices gaining access to network service, GSM incorporates an authentication procedure which enables the network to verify that the SIM in a UE is genuine. The authentication procedure operates as follows. Further details can be found in technical specifications GSM 03.20 [10] and GSM 04.08 [12].

1. The UE visits a network, and is initially identified using its IMSI.
2. The visited network identifies the UE's home network from the supplied IMSI, and contacts the home network for authentication information.
3. The home network's AuC generates one or more *authentication triples* (*RAND*, *XRES*, K_c), and sends them to the visited network, where *RAND* is a 128-bit random 'challenge' value, *XRES* is the 64-bit 'expected response', and K_c is a 64-bit short-term session key to be used to encrypt data sent across the air interface between the UE and the network.
4. The visited network sends *RAND* to the UE as an authentication challenge.
5. The ME receives the *RAND*, and passes it to the SIM.
6. The SIM computes $SRES = \mathrm{A3}_{K_i}(RAND)$ and $K_c = \mathrm{A8}_{K_i}(RAND)$, where A3 and A8 are network-specific cryptographic functions; A3 is a MAC function and A8 is a key derivation function. Note that precisely the same computation was performed by the AuC in step 3 to generate *XRES* and K_c.
7. The SIM passes *SRES* and K_c to the ME.
8. The ME keeps the session key K_c for use in data encryption, and forwards *SRES* to the serving network.
9. The serving network compares *SRES* with *XRES*; if they are the same the UE is deemed authenticated, and K_c can now be used for traffic encryption using any of the standardised algorithms (i.e. one of A5/1, A5/2 and A5/3), as selected by the serving network.

2.3 Vulnerabilities

The GSM AKA protocol clearly only provides one-way authentication. As widely documented (see, for example, [21]), this permits a 'false' base station to impersonate a genuine network and interact with a UE. This in turn gives rise to a range of security weaknesses. We are particularly interested in attacks of the following types.

– Because the network always decides whether or not to enable encryption, it is possible for a malicious party to act as an intermediary between a UE and a genuine network, impersonating the network to the UE and using a genuine SIM of its own to talk to the network. All traffic sent via the man-in-the-middle is simply relayed. The false network does not enable encryption on the link to the UE, so the fact that it does not know the encryption key does not matter. If the genuine network chooses to enable encryption, then the man-in-the-middle can communicate with it successfully since it is using

its own SIM for this leg of the communications. As a result, the man-in-the-middle can seamlessly listen to all the voice traffic sent to and from the victim UE, at the cost of paying for the call.

– The fact that the network decides whether or not to enable data encryption also enables the well known Barkan-Biham-Keller attack, [2]. This attack is designed to recover the encryption key K_c, and hence enable unlimited interception of phone calls. The attack takes advantage of three key facts: A5/2 is very weak, the network decides which algorithm to use, and the same key K_c is used with all three encryption algorithms. One possible scenario for the attack is as follows.
  Suppose an eavesdropper intercepts the AKA exchange between the network and a UE (notably including the *RAND*), and also some of the subsequent encrypted voice exchanges involving that UE. Suppose also that the UE is subsequently switched on within the range of a fake network operated by the attacker. The fake network inaugurates the AKA protocol with the UE, and sends the previously intercepted *RAND*, causing the SIM in the UE to generate the same K_c as was used to encrypt the intercepted data. The UE responds with *SRES* (which the fake network ignores) and the fake network now enables encryption using A5/2. The UE will now send data to the network encrypted using A5/2 with the key K_c; because of certain details of the GSM protocol, the plaintext data will contain predictable redundancy. The fake network now takes advantage of the weakness of A5/2 to recover K_c from the combination of the ciphertext and known redundancy in the corresponding plaintext. The key K_c can now be used to decrypt all the previously intercepted data, which may have been encrypted using a strong algorithm such as A5/3.

The lack of mutual authentication has been addressed in 3G and later networks. As a result it is tempting to suggest that trying to fix GSM is no longer of relevance. However, GSM continues to be very widely used worldwide and will continue to be for many years to come; so finding ways of upgrading GSM post-deployment appears to be worthwhile. However, any such solution must work with the existing infrastructure, i.e. the existing serving network systems. We are therefore interested in a solution which only requires SIMs and the home network to be upgraded. Such a solution can be rolled out piecemeal with no impact on the existing global infrastructure, and this is the focus of the remainder of this paper.

2.4 Proactive SIM

Before proceeding we need to briefly review a key piece of GSM technology which enables the SIM to send an instruction to the ME. *Proactive SIM* is a service operating across the SIM-ME interface that provides a mechanism for a SIM to initiate an action to be taken by the ME. It forms part of the *SIM application toolkit (STK)*, which was introduced in the GSM 11.14 technical specification [11]. Communications between an ME and a SIM are command/response based,

and STK provides a set of commands which allow the SIM to interact and operate with any ME which supports them.

The GSM technical specification [13] states that the ME must communicate with the SIM using either the T=0 or T=1 protocol, specified in ISO/IEC 7816-3 [16]. In both cases the ME is always the *master* and thus initiates commands to the SIM; as a result there is no mechanism for the SIM to initiate communications with the ME. This limits the possibility of introducing new SIM features requiring the support of the ME, as the ME needs to know in advance what actions it should take. The proactive SIM service provides a mechanism that allows the SIM to indicate to the ME, using a response to an ME-issued command, that it has some information to send. The SIM achieves this by including a special status byte ('91' followed by the length of the instruction to send) in the response application protocol data unit. The ME is then required to issue the *FETCH* command to find out what the information is [14]. The ME must now execute the SIM-initiated command and return the result in the *TERMINAL RESPONSE* command. To avoid cross-phase compatibility problems, this service is only permitted to be used between a SIM and an ME that support the STK commands. The fact that an ME supports specific STK commands is revealed when it sends the *TERMINAL PROFILE* command during SIM initialisation.

The SIM can make a variety of requests using the proactive SIM service. Examples include: requesting the ME to display SIM-provided text, initiating the establishment of on demand channels, and providing local information from the ME to the SIM. The commands of interest here are *GET CHANNEL STATUS*, which requests the ME to return the current status of all available data channel(s), and *CLOSE CHANNEL*, which requests the ME to close the specified data channel. Both of these STK commands are marked as 'class e', which means that an ME that supports 'class e' STK commands is capable of executing both commands of interest [9]. Although support of STK is optional for an ME, if an ME claims compliance with a specific GSM release then it is mandatory for the ME to support all functions of that release. Since 1998 almost all of the mobile phones produced have been STK enabled, and today every phone on the market supports STK [1].

3 RAND Hijacking

We use the term *RAND hijacking* to refer to the idea of using the *RAND*, sent from the network to the UE during AKA, as a way of conveying information from the AuC to the SIM. That is, instead of generating the *RAND* at random, it is generated to contain certain information; this information is typically sent in encrypted form so that to an eavesdropper it is indistinguishable from a random value.

This idea was apparently first described in a patent due to Dupré [6]. However, the use Dupré makes of the idea is rather different to that proposed here. Later, Vodafone introduced the concept of a *special RAND* [23] in 3GPP TSG document S3-030463. As for Dupré, the purpose of the *special RAND* was com-

pletely different to that proposed here. The other published references to the notion appear in papers [4, 5, 18] that independently propose the use of *RAND hijacking* for improving the privacy properties of GSM, 3G and 4G networks. As far as the authors are aware, no previous authors have proposed the use of this technique for providing mutual authentication in GSM networks.

4 Server-to-SIM Authentication

We now propose a way of using *RAND hijacking* to enable authentication of the network to the SIM. For this to operate the SIM must be programmed to support the scheme, as well as possess certain (modest) additional data, as detailed below. The AuC of the network issuing the 'special' SIM must also store certain additional data items for each such SIM, and must generate its *RAND* values in a special way for such SIMs. No other changes to existing systems are required. It is important to note that the system could be deployed gradually, e.g. by including the additional functionality in all newly issued SIMs, whilst existing SIMs continue to function as at present.

4.1 Prerequisites

In addition to sharing K_i, A3 and A8 (as required for executing the standard GSM AKA protocol), the SIM and AuC must both be equipped with the following information and functions:

- functions $f1$ and $f5$, where $f1$ is a MAC function and $f5$ is a cipher mask generation function, both capable of generating a 64-bit output;
- a secret key K_a to be used with functions $f1$ and $f5$ which should be distinct from K_i — to minimise memory requirements, K_a and K_i could, for example, both be derived from a single SIM-specific master key;
- a 48-bit counter to be used to generate and verify sequence numbers[1].

The functions could be precisely the same as their counterparts used in 3G (UMTS). Indeed, the function names and string lengths have deliberately been made identical to those used in 3G systems to make implementation and migration as simple as possible.

4.2 Protocol Operation

The novel AKA protocol only differs from the 'standard' GSM AKA protocol (as described in section 2.2 above) in steps 3 and 6. Thus, since these steps only involve the AuC and SIM, it should be clear that the scheme is inherently transparent to the serving network and the ME. We describe below how these steps are changed.

[1] As in 3G, an AuC might choose to manage a single counter shared by all user accounts (see, for example, [20]).

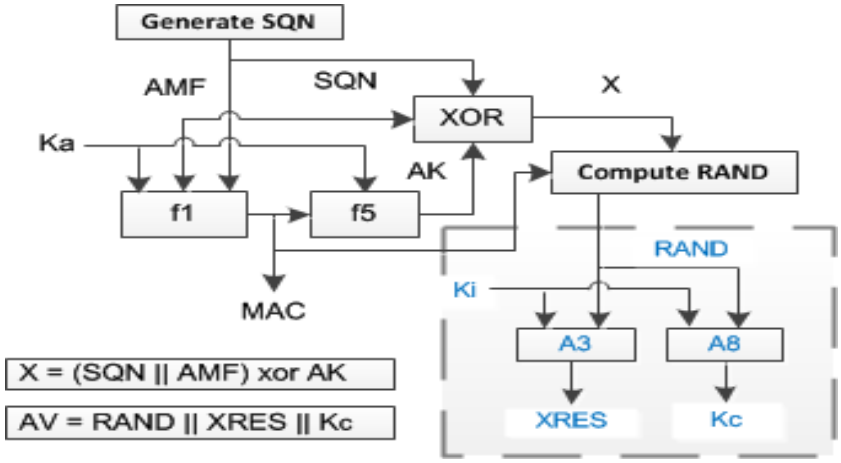

Fig. 1. Modifications at the the AuC

4.3 Revised Steps

Step 3 is changed to the following step 3*. To generate a new authentication triple, the AuC proceeds as follows (see Fig. 1, in which the dotted block represents the usual operation of the AuC).

3.1 The AuC uses its counter value to generate a 48-bit sequence number SQN, which must be greater than any previously generated value for this user account.

3.2 A 16-bit value AMF is also generated, which could be set to all zeros, or could be used for purposes analogous to the AMF value for 3G networks.

3.3 A 64-bit tag value MAC is generated using function $f1$, where

$$MAC = f1_{K_a}(AMF||SQN),$$

and, as throughout, || denotes concatenation of data items.

3.4 A 64-bit encrypting mask AK is generated using function $f5$, where

$$AK = f5_{K_a}(MAC).$$

3.5 The 128-bit $RAND$ is computed as

$$RAND = ((AMF||SQN) \oplus AK)||MAC,$$

where, as throughout, $\oplus$ denotes the bitwise exclusive or operation.

3.6 The $XRES$ and K_c values are computed in the standard way, that is $XRES$ = $\mathrm{A3}_{K_i}(RAND)$ and $K_c = \mathrm{A8}_{K_i}(RAND)$.

Step 6 is changed to step 6*, as follows (see Fig. 2, in which the dotted block represents the usual operation of the SIM).

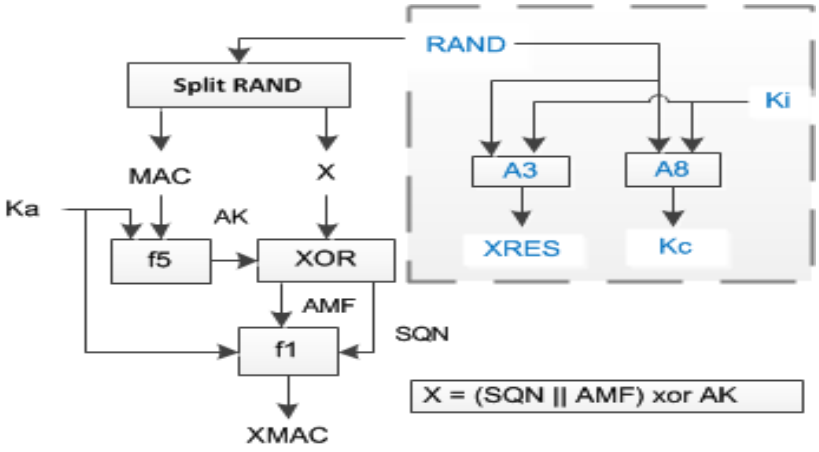

Fig. 2. Modifications at the SIM

6.1 On receipt of the 128-bit $RAND$ value, the SIM first splits it into two 64-bit strings X and $MAC*$, where $X||MAC* = RAND$.

6.2 A 64-bit decrypting mask $AK*$ is generated using function $f5$, where

$$AK* = f5_{K_a}(MAC*).$$

6.3 A 16-bit string $AMF*$ and a 48-bit string $SQN*$ are computed as:

$$AMF*||SQN* = X \oplus AK*.$$

6.4 A 64-bit tag $XMAC$ is computed as:

$$XMAC = f1_{K_a}(AMF*||SQN*).$$

6.5 The recovered sequence number $SQN*$ is compared with the SIM's stored counter value and $XMAC$ is compared with $MAC*$:

- if $SQN*$ is greater than the current counter value **and** $XMAC = MAC*$, then:
  - the network is deemed to be successfully authenticated;
  - the SIM's counter value is updated to equal $SQN*$; and
  - $SRES$ and K_c are computed as specified in the current step 6;
- if either of the above checks fail then:
  - network authentication is deemed to have failed;
  - the SIM's counter value is unchanged; and
  - $SRES$ and K_c are set to random values.

It should be clear that, in step 6*, $AK*$, $AMF*$, $MAC*$ and $SQN*$ should respectively equal the AK, AMF, MAC and SQN values originally computed by the AuC in step 3*.

4.4 Design Rationale

The composition of the $RAND$ value in the above scheme has been made as similar as possible to the 128-bit value $AUTN$ used to provide server-to-UE authentication in the 3G AKA protocol. This is for two main reasons. Firstly, as stated above, by adopting this approach it is hoped that implementation of, and migration to, this new scheme will be made as simple as possible for network operators. Secondly, the 3G AKA protocol is widely trusted to provide authentication, and it is hoped that trust in the novel scheme will be maximised by adopting the same approach.

The only differences between the 3G $AUTN$ and the above construction of $RAND$ are relatively minor, and are as follows.

- In 3G, the AK value is computed as a function of the the $RAND$, whereas here it is necessarily only computed as a function of the last 64 bits of $RAND$. However, these last 64 bits are computed as a function of data which changes for every authentication triple, and hence the AK should still do an effective job of concealing the content it is used to mask.
- In 3G the AK is only 48 bits long, and is only used to encrypt (mask) the SQN. Here we use it to mask the SQN and the AMF, to ensure that a 'new style' $RAND$ is indistinguishable from an 'old style' randomly generated $RAND$ to any party without the key K_a.
- In 3G, the MAC is computed as a function of the $RAND$, SQN and AMF, whereas in the above scheme it is computed only as a function of SQN and AMF, again for obvious reasons. This is the only significant difference from the perspective of authenticating the network to a UE, but we argue below in section 6.2 that this change does not affect the security of the protocol.

The $AUTN$ checking process proposed here and that used in 3G are essentially the same.

One other issue that merits mention is the fact that it is proposed that the SIM outputs random values if authentication fails. It is necessary for the SIM to output values of some kind, since this is part of the existing SIM-ME protocol. That is, placeholder values are required. It is important for reasons discussed below that the SIM should *not* output the correct session key K_c. The only other 'obvious' placeholder values would be to use fixed strings, but the use of random values seems less likely to be obvious if these values are sent across the network (in the case of the $SRES$ value) or used for encryption purposes (for K_c). There may be advantages in not revealing to a casual eavesdropper the fact that authentication has failed.

5 Using the Authentication Results

In the previous section we showed how the SIM can authenticate the network; that is, as a result of step 6*, the SIM will know whether or not the $RAND$ genuinely originates from the AuC and is fresh. However, we did not describe

any way for the ME to know whether authentication has failed or succeeded — indeed, the ME will not understand the concept, as we are assuming it is a 'standard' GSM device.

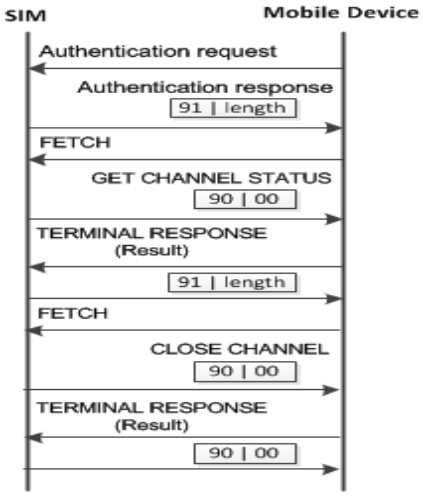

Fig. 3. SIM-ME interactions to drop the established connection

We propose that the proactive SIM feature described in section 2.4 be used to achieve the desired objective. That is, in the event of a network authentication failure, when sending the *SRES* and K_c (in this case random) values back to the ME, the SIM should signal to the phone that it has information to send. When, as a result, the ME sends the *FETCH* command to the SIM, the SIM should respond with the *GET CHANNEL STATUS* command to learn about the established channels in the present connection. Upon receiving the channel information in the *TERMINAL RESPONSE* command, the SIM uses the response status byte in its response to request the ME to send a further *FETCH* command. Once it receives the *FETCH* command, the SIM responds with a *CLOSE CHANNEL* command, specifying the channel information it received from the ME in response to its previous *CHANNEL STATUS* command. The interactions between a SIM and an ME are summarised in Fig. 3. The STK commands issued by the SIM should cause the phone to drop the connection, and (hopefully) prevent any attempted use of the *SRES* or key K_c. The values 90 and 91, shown in Fig. 3, represent the value of the *status byte* sent by the SIM in response to the previous command, where the value 90 means OK, and the value 91 instructs the ME to issue a *FETCH* command to retrieve data from the SIM. The 'length' with the status byte 91 indicates the length of the data in bytes which the SIM wants to send.

6 Analysis

6.1 Deployment Issues

We next consider certain practical issues that may arise when using the scheme proposed in section 4.3.

It seems that at least some GSM networks issue authentication triples in batches (see section 3.3.1.1 of GSM 03.20 [10]), thereby reducing the inter-network communications overhead. Currently, the order in which GSM authentication triples are used does not matter. However, under the scheme described above, triples must be used in ascending order of SQN. This may seem problematic; however, since the requirement to use authentication datasets in the correct order already applies to the corresponding 5-tuples used in 3G, serving networks will almost certainly already be equipped to do this.

In existing GSM networks it is possible, although prohibited by the technical specifications [10], for serving networks to 're-use' authentication triples, i.e. to send the same $RAND$ value to a UE on multiple occasions. This will no longer work with the new scheme, since the SIM will detect re-use of a $RAND$ value. Arguably this is good, since re-use of $RAND$ values is highly insecure: such behaviour would allow the interceptor of a $RAND/SRES$ pair to impersonate a valid UE and perhaps steal service at that UE's expense, an attack that would be particularly effective in networks not enabling encryption.

Finally note that, in order to fully implement the scheme as described in section 4, MEs need to support 'class e' STK commands, although, as discussed above, this proportion seems likely to be very high. It is not clear what proportion of mobile phones in current use support those STK commands.

6.2 Security

We divide our security discussion into three parts: confidentiality and privacy issues, authentication of network to SIM, and authentication of SIM to network.

Confidentiality and Privacy Issues In 'standard' GSM the $RAND$ value is randomly selected, and so does not reveal anything about the identity of the phone to which it is sent. In the scheme proposed in section 4.3, the $RAND$ is a function of a SIM-specific key as well as a potentially SIM-specific SQN value. However, the SQN is sent encrypted, and, assuming the functions $f1$ and $f5$ are well-designed, an interceptor will not be able to distinguish an intercepted $RAND$ computed according to the new scheme from a random value. Thus the scheme does not introduce a new threat to identity confidentiality.

The new scheme does not change the way the data confidentiality key K_c is generated, so the strength of data confidentiality is not affected.

Network-to-SIM Authentication The novel protocol for network-to-SIM authentication bears strong similarities to the corresponding protocol for 3G.

It also conforms to the one-pass unilateral authentication mechanism specified in clause 5.1.1 of 9798-4 [15, 17]. All the protocols in this standard have been formally analysed (and shown to be secure) by Basin, Cremers and Meier [3]. Whilst these arguments do not provide a completely watertight argument for the protocol's security, it is clearly a significant improvement over no authentication at all.

An interesting side observation deriving from the novel scheme is that the 3G and 4G AKA protocols appear to be overly complex. The randomly generated *RAND* value sent from the network to the SIM, which is used to authenticate the response from the SIM to the network, is actually unnecessary, and the *AUTN* value could be used in exactly the same way as the *RAND* is currently. Whilst such a change is not possible in practice, it would have avoided the need for the AuC to generate random values and saved the need to send 16 bytes in the AKA protocol.

It is interesting to speculate why this design redundancy is present. It seems possible that the network-to-SIM authentication was added as a completely separate protocol to complement the GSM-type SIM-to-network authentication mechanism, and no-one thought how the two mechanisms could be combined and simplified (as in the mechanism we propose).

SIM-to-Network Authentication The novel scheme does not affect how the existing SIM-to-network authentication protocol operates, except that a random *RAND* is replaced by one which is a cryptographic function of a sequence number. The new-style *RAND* remains unpredictable to anyone not equipped with the key K_a, and is deterministically guaranteed to be non-repeating (a property that only holds in a probabilistic way for a random *RAND*). To see why the *RAND* is non-repeating, suppose two separate *RAND* values sent to the same USIM incorporate the same MAC values (as necessary if they are to be the same). It follows that the AK values used to mask the *SQNs* embedded in the *RAND* values will also be the same and thus, since the *SQN* values themselves will be different, the two *RAND* values will also differ. That is, it possesses precisely the qualities required by the existing protocol, and hence the security of SIM-to-network authentication is unaffected.

6.3 Impact on Known Attacks

We conclude our analysis of the protocol by considering how it affects possible attacks on GSM networks.

Fake Network Attacks As discussed in section 2.3, if a phone joins a fake GSM serving network, then this fake network can send any *RAND* value it likes as part of the AKA protocol, and the UE will complete the process successfully. If the network does not enable encryption, then communications between the UE and the network will work correctly, which could enable the network to act as an eavesdropping man-in-the-middle by routing calls from the captured UE via a

genuine network. This will no longer be true if the new scheme is implemented, since the SIM will instruct the ME to drop the connection when supplied with a non-genuine *RAND* value.

Of course, it may be possible for a fake network to avoid the AKA protocol altogether, and simply start communication with a newly attached UE. Whether MEs will accept unauthenticated communication is currently not clear to the authors.

Barkan-Biham-Keller Attacks We next consider a particular type of fake network attack, namely the Barkan-Biham-Keller attack outlined in section 2.3. As described there, the attack requires the re-sending of an 'old' *RAND* to a UE. The new scheme will clearly prevent such an attack, i.e. the Barkan-Biham-Keller attack will be prevented, at least in most practical scenarios.

7 Relationship to the Prior Art

This is by no means the first practical proposal for enhancing GSM to incorporate mutual authentication. Indeed, the 3G AKA protocol, discussed widely in this paper, can be regarded as doing exactly that. Although several 3GPP TSG documents [7, 8] proposed the introduction of network authentication into the GSM network, none were adopted, presumably because of cost/feasibility issues. The Ericsson proposal [8] suggested transferring authentication responsibility to the terminal by implementing the core of the UMTS AKA protocol entirely in software, which in turn raised other security threats. Other proposals have been made, including by Kumar et al., [19]. However, all previous proposals are completely impractical in that they would require replacing all the GSM infrastructure. Such a major change to an existing very widely deployed scheme is simply not going to happen.

The most similar proposals to that given here are some of the other schemes using *RAND hijacking*, summarised in section 3. In particular, van den Broek, Verdult and de Ruiter [4] propose a similar structure for a hijacked GSM *RAND*, in their case including a sequence number, a new temporary identity for the SIM, and a MAC, all encrypted in an unspecified way. However, their objective is not to provide authentication of the network to the SIM, but to provide a way to reliably transport new identities from the AuC to the SIM.

8 Concluding Remarks

We have proposed a method for enhancing the GSM AKA protocol to provide authentication of the network to the UE, complementing the UE-to-network authentication already provided. This provides protection against some of the most serious threats to the security of GSM networks. This is achieved in a way which leaves the existing serving network infrastructure unchanged, and also

does not require any changes to existing MEs (mobile phones). That is, unlike previous proposals of this general type, it is practically realisable.

A number of practical questions remain to be answered, including the proportion of MEs supporting 'class e' STK commands, the behaviour of MEs in networks which never perform the AKA protocol, and whether serving networks can be relied upon to use GSM authentication triples in the intended order. Discovering answers to these questions remains as future work.

Acknowledgements

We thank Fabian van den Broek and the anonymous reviewers for their thoughtful feedback and suggestions which have improved the paper.

References

1. SIM Toolkit. `http://www.bladox.cz/devel-docs/gen_stk.html`, (Online) Accessed: 2016-05-31
2. Barkan, E., Biham, E., Keller, N.: Instant ciphertext-only cryptanalysis of GSM encrypted communications. In: Boneh, D. (ed.) Advances in Cryptology — CRYPTO 2003, 23rd Annual International Cryptology Conference, Santa Barbara, California, USA, August 17-21, 2003, Proceedings. Lecture Notes in Computer Science, vol. 2729, pp. 600–616. Springer-Verlag, Berlin (2003)
3. Basin, D.A., Cremers, C.J.F., Meier, S.: Provably repairing the ISO/IEC 9798 standard for entity authentication. In: Degano, P., Guttman, J.D. (eds.) Principles of Security and Trust — First International Conference, POST 2012, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2012, Tallinn, Estonia, March 24 – April 1, 2012, Proceedings. Lecture Notes in Computer Science, vol. 7215, pp. 129–148. Springer-Verkag, Berlin (2012)
4. van den Broek, F., Verdult, R., de Ruiter, J.: Defeating IMSI catchers. In: Ray, I., Li, N., Kruegel, C. (eds.) Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security, Denver, CO, USA, October 12–16, 2015. pp. 340–351. ACM (2015)
5. Choudhury, H., Choudhury, B.R., Saikia, D.K.: Enhancing user identity privacy in LTE. In: IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications (TrustCom), 2012. pp. 949–957. IEEE (2012)
6. Dupré, M.: Process to control a Subscriber Identity Module (SIM) in mobile phone system. US Patent Office (February 2004), US Patent 6,690,930, Filing date–25 May, 1999
7. Ericsson: Enhancements to GSM/UMTS AKA. 3GPP TSG SA WG3 Security, S3-030542, Povoa de Varzim, Portugal (October 6–10 2003)
8. Ericsson: On the introduction and use of UMTS AKA in GSM. 3GPP TSG SA WG3 Security, S3-040534, Acapulco, Mexico (July 6–9 2004)
9. European Telecommunications Standards Institute (ETSI): ETSI TS 101 267 V8.18.0 (2007-06): Technical Specification; Digital cellular telecommunications system (Phase 2+); Specification of the SIM Application Toolkit (SAT) for the Subscriber Identity Module–Mobile Equipment (SIM–ME) interface (3GPP TS 11.14 version 8.18.0 Release 1999)

10. European Telecommunications Standards Institute (ETSI): ETSI-GSM Technical Specification; European digital cellular telecommunication system (phase 1); Security-related network functions (GSM 03.20) (February 1992)
11. European Telecommunications Standards Institute (ETSI): GSM 11.14: Technical Specification; Digital cellular telecommunications system (Phase 2+); Specification of the SIM Application Toolkit for the Subscriber Identity Module–Mobile Equipment (SIM–ME) interface (December 1996)
12. European Telecommunications Standards Institute (ETSI): GSM Technical Specification; Digital cellular telecommunication system (phase 2+); Mobile radio interface layer 3 specification (GSM 04.08) (July 1996)
13. European Telecommunications Standards Institute (ETSI): GSM Technical Specification; Digital cellular telecommunications system (Phase 2+); Specification of the Subscriber Identity Module–Mobile Equipment (SIM–ME) interface; (GSM 11.11) (July 1996)
14. European Telecommunications Standards Institute (ETSI): ETSI TS 102 223 Version 11.1.0; Smart Cards; Card Application Toolkit (CAT) (2012)
15. International Organization for Standardization, Genève, Switzerland: ISO/IEC 9798–4: 1999, Information technology — Security techniques — Entity authentication — Part 4: Mechanisms using a cryptographic check function, 2nd edn. (1999)
16. International Organization for Standardization: ISO/IEC 7816–3; Identification cards—Integrated circuit cards; Part 3: Cards with contacts—Electrical interface and transmission protocols (November 2006)
17. International Organization for Standardization, Genève, Switzerland: ISO/IEC 9798–4: 1999/Cor 1:2009, Technical Corrigendum 1 (2009)
18. Khan, M.S.A., Mitchell, C.J.: Improving air interface user privacy in mobile telephony. In: Chen, L., Matsuo, S. (eds.) Security Standardisation Research — Second International Conference, SSR 2015, Tokyo, Japan, December 15–16, 2015, Proceedings. Lecture Notes in Computer Science, vol. 9497, pp. 165–184. Springer-Verlag, Berlin (2015)
19. Kumar, K.P., Shailaja, G., Kavitha, A., Saxena, A.: Mutual authentication and key agreement for GSM. In: 2006 International Conference on Mobile Business (ICMB 2006), 26–27 June 2006, Copenhagen, Denmark. p. 25. IEEE Computer Society (2006)
20. Mitchell, C.J.: Making serial number based authentication robust against loss of state. ACM Operating Systems Review 34(3), 56–59 (July 2000)
21. Mitchell, C.J.: The security of the GSM air interface protocol. Tech. Rep. RHUL-MA-2001-3, Mathematics Department, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK (August 2001), available at `http://www.ma.rhul.ac.uk/techreports`
22. Pagliusi, P.S.: A contemporary foreword on GSM security. In: Davida, G.I., Frankel, Y., Rees, O. (eds.) Infrastructure Security, International Conference, InfraSec 2002 Bristol, UK, October 1–3, 2002, Proceedings. Lecture Notes in Computer Science, vol. 2437, pp. 129–144. Springer-Verlag, Berlin (2002)
23. Vodafone: Cipher key separation for A/Gb security enhancements. 3GPP TSG SA WG3 Security, S3-030463, San Francisco, USA (July 15–18 2003)