Security in future mobile multimedia networks

Chris J. Mitchell[1] and Liqun Chen[1]

1. Introduction

Current digital mobile networks, e.g. those based on the ETSI GSM standards, provide a robust set of security facilities to protect communications across the air interface. The main GSM security services are confidentiality of user and signalling data (across the air interface), user authentication to a base station, and user identity confidentiality (across the air interface). Because of their universal nature and the extra requirements of high data rate multimedia traffic, standards for future networks will need to support a larger range of security services. Possible new services include: end-to-end data confidentiality and integrity, incontestable charging, and a more robust user identity confidentiality.

There is also much to be gained by standardising management aspects of security provision. In GSM, although the management security requirements are clear, the exact way in which user key information is generated, stored and accessed is left to Network Operators (NOs) and equipment providers to arrange. This can make security service provision costly for all concerned, since every NO may arrange security management differently. In future mobile networks, possibly operating in a rather more deregulated environment than at present, standardised support for security management will be a very important feature. Without such standards, the required co-operation between the likely large numbers of competing NOs and Service Providers (SPs) could become impossibly complex to arrange.

In this paper we examine some of the security provisions in the emerging ETSI UMTS (Universal Mobile Telecommunications System) and ITU FPLMTS (Future Public Land Mobile Telecommunications System) standards for future mobile telecommunications networks. After a brief review of some of the most significant areas for the provision of security services, we focus our attention on the simultaneous provision of identity and location privacy for the mobile user and mutual authentication between mobile user and base station. In doing so we describe research into security for future mobile networks performed by the DTI/EPSRC-funded project 'Third Generation System Security Studies' (3GS3), part of the LINK Personal Communications Programme. The project collaborators were Vodafone Ltd, GPT Ltd. and Royal Holloway, University of London. The authors would like to acknowledge the invaluable support and advice of colleagues in 3GS3, without which this paper could not have been written.

---

[1] Information Security Group, Royal Holloway, Univ. of London, Egham, Surrey TW20 0EX, UK. Fax: +44-1784-439786. Email: {liqun,cjm}@dcs.rhbnc.ac.uk.

2. Third-generation mobile systems

The term *3rd generation* refers to mobile systems which will follow existing digital networks such as GSM, DCS1800 and DECT; such systems are currently being standardised by ITU (FPLMTS) and ETSI (UMTS). They are characterised by the following: multiple operators, multiple environments (residential cordless, mobile, satellite, etc.), multi-vendor and standardised interfaces, use of the WARC-assigned FPLMTS band, higher bit-rates (up to 2Mb/s), and migration from existing systems.

Current GSM systems support security features such as confidentiality of user and signalling data on the air interface, authentication of users, and user identity confidentiality. There are areas where security can be enhanced in 3rd generation systems, partly based on lessons from 2nd generation systems, but mostly deriving from the new characteristics noted above.

In our discussion of security in mobile communications we use a simple model with four roles: Users, NOs, SPs and Intruders, which are defined fully in Clause 3.2 of [1]. Briefly:

- a *user* is an entity authorised to use particular network services,
- a *Network Operator (NO)* is an entity providing network capabilities to support particular services, and which allows users to access the network to use the services,
- a *Service Provider (SP)* is an entity responsible for the provision of particular services, and will typically do so by means of contractual relationships with NOs, and
- an *intruder* is an entity that abuses the network infrastructure or services on the network.

3. Security features for future networks

Initial studies in 3GS3 identified the likely security threats to future mobile networks in the context of role and functional models (also defined by the project; see [1]). Security features necessary to address these threats were identified and classified, including the following.

- *Entity authentication*. Entity authentication between a user and Network Operators and/or Service Providers was studied. A number of mechanisms, based on various cryptographic methods, were examined, classified, and tested (formally and informally). As a result, an entity authentication mechanism was proposed to both UMTS and FPLMTS, and subsequently was incorporated into both sets of draft standards. This mechanism, briefly described in [2], is considered in Section 5.1; a further mechanism is considered in Section 5.2. Problems arising when some of the 'authentication servers' within a system may be unreliable, [3,4], and the effect of the properties of the underlying components of an authentication mechanism on its design, [2], were also considered.

- *Novel techniques for key distribution.* Maurer has shown, [5], how Wyner's 'Wire tap channel' concept, [6], can be used much more widely than originally envisaged. The idea makes use of the universal presence of noise in communications channels to help two users agree a secret key using only 'public' channels. The practicality of this idea was investigated, and new theoretical results were discovered, [7].

- *End-to-end encipherment, and warranted interception facilities.* Multi-media terminals will place demanding bandwidth requirements on the mobile network. These requirements have relatively little direct effect on security feature provision, except that any directly data-related security features, such as the provision of data confidentiality, must be implemented using methods which can handle high-bandwidth data. In practice this means that air interface encryption methods must be able to handle high throughput rates. However, this should not be too difficult since multi-media terminals will not be low cost items, and the provision of processing capabilities to handle high data-rate encipherment should not add significantly to the overall cost of such devices.

  More significant to the design of security features are the likely needs of users of these multi-media services; these needs are potentially very different from those of 'voice' users of existing networks. Of particular importance are likely to be issues such as end-to-end integrity and confidentiality, albeit that existing networks do not support integrity, and only provide encryption for the air interface. Of all the end-to-end security features, end-to-end confidentiality raises most problems. The problems are mainly political rather than technical, and arise from the need of law enforcement agencies for access to certain communications paths, when a warrant exists. Such access is valuable in combating criminal activity, but also needs to be carefully controlled because of the civil liberties issues. This issue has given rise to a public debate on 'key escrow' schemes, starting with the US *Clipper* scheme; see, for example [8]. There is a growing consensus that 'Trusted Third Parties' (TTPs) offer a means of supporting warranted access at the same time as meeting legitimate user needs for confidentiality. 3GS3 has developed a TTP-based scheme for warranted access, offering considerable advantages over some other proposed schemes, [9].

- *Identity and location privacy*. In mobile telecommunications systems, each user must let its SP know where he/she is so that its call route can be maintained by the system. This is achieved by the registration and location update mechanisms which a user employs to tell its current location to its SP via the NO for the current location area. This has the side effect that anyone wanting to track this particular user can do so by monitoring the identity and location messages transmitted during the registration and location update processes.

  Users of public telecommunications networks are likely to regard the possibility of their location being revealed by these mechanisms as an unacceptable breach of personal privacy. Thus, in order

to prevent users' identity and location information being disclosed to unauthorised parties, an Identity and Location Privacy (ILP) mechanism is needed; such a mechanism protects users against tracing of their physical location by illegal means.

Current GSM networks provide a level of user identity confidentiality, but the mechanism used is less appropriate for future networks, not least because of the multi-operator environment likely to prevail. New mechanisms, based on both public key and 'conventional' cryptographic techniques, have been examined, and are the focus of the remainder of this paper.

- *Simultaneous multiple access channel coding and encipherment*. The claim that CDMA, a likely multiple access method for future mobile networks, is inherently secure was considered and rejected. Options for using CDMA sequences for encipherment were also examined, [10].
- *Terminal-related security*. Current networks enable black-listing of stolen terminals, and detection of non-type-approved terminals. The need for such facilities in future was reviewed, given that most mobile terminals are likely to be relatively low-cost. Whether a universal scheme is adopted, or a scheme only applying to valuable (e.g. multimedia) terminals, remains a topic for debate.

A predominant feature on the work of 3GS3 was its commitment to standards contributions, both in ETSI and in ITU. Apart from mechanisms proposed and adopted, much of the draft standards' text on security features classification and analysis is based on 3GS3 contributions.

4. Identity and Location Privacy

The remainder of this paper is concerned with two particularly important security services for future mobile multimedia networks: Identity and Location Privacy (ILP) for the mobile user, and mutual authentication between mobile user and base station. We start by considering in detail the provision of ILP services. Subsequently we consider two mutual authentication mechanisms also providing ILP.

4.1. The GSM approach

In GSM, ILP is achieved by using Temporary Identities (TIs) over the air interface instead of Real Identities (RIs)[2]. The TI is chosen by an NO[3] and is valid only in a given location area. The SP[4] maintains a database of current TI/RI relationships and can therefore determine the real identity of a user, i.e. it can determine the RI from the TI. TIs are changed on each location update and on certain other network-defined occasions.

---

[2] In GSM TIs and RIs are called *Temporary Mobile Subscriber Identities* (TMSIs) and *International Mobile Subscriber Identities* (IMSIs) respectively.

[3] In GSM an NO is a *Base Station Subsystem, Mobile Switching Centre and Visitor Location Register* (BSS/MSC/VLR).

In more detail, the user identifies himself by sending the old TI during each location update process (this occurs prior to authentication, and the TI must therefore be sent unencrypted). The new TI is returned after authentication is complete and a new session key has been generated, and hence the new TI can be, and is, encrypted when sent to the user. This prevents an interceptor from linking an one TI to the next, and blocks tracing of user movements by linking TIs. If a TI is unavailable or invalid, e.g. if during the initial location registration the old NO is unreachable or the old TI is unknown, [11], then a user has to identify itself using its RI. In this event a new TI is allocated and returned encrypted.

4.2. Possible threats to the GSM approach

In this section we consider seven possible threats to the GSM ILP scheme.

T1. **Intercepting communications between user and NO**. An intruder can obtain a RI from the GSM air interface whenever a RI is sent in clear text, i.e. in the following cases: initial location registration, 'old visitor location register unreachable', and 'no old TI available'.

T2. **Impersonating a user**. In a mobile telecommunications environment an intruder may be able to fabricate and/or interfere with a user's messages to an NO. An intruder could modify the user's TI and/or the Location Area Identifier, both of which are sent from user to NO in clear text. This will mean that the NO fails to recognise the user (or is unable to contact the 'old' NO), causing the NO to ask the user to send its RI unencrypted over the Air Interface. Such a procedure could be repeated, enabling an intruder to track a user.

T3. **Impersonating an NO**. In GSM, user authentication is unilateral, i.e. the NO verifies the user's identity, but the user does not verify the NO's identity. Hence an intruder could impersonate an NO and instruct a user to send its RI unencrypted over the Air Interface. As is the case for threat T2, such a procedure could be repeated as often as required, enabling an intruder to track a user.

T4. **Intercepting channels between NOs and SPs**. If an intruder could monitor the channel between NO and SP, it could observe a user's identity and location information, and hence track a user, because each updated location message is sent from an NO to an SP, possibly in clear text.

T5. **Malicious NOs**. It is possible for a malicious NO to track a user because TIs are chosen by NOs, and hence NOs have access to a user's RI.

T6. **Impersonating an SP to an NO**. In GSM the SP verifies the user's identity during the user authentication process, but no mechanisms are provided for the NO and/or the user to verify the SP's identity. In practice where such a threat exists proprietary techniques are used to protect SP/NO communications, and hence (indirectly) protect the user against an intruder impersonating

---

[4] In GSM an SP is an *Authentication Centre and Home Location Register* (AuC/HLR).

an SP. However, if the NO does not authenticate the SP, then an intruder could impersonate an SP to an NO to obtain the user's identity and location information (and thereby track the user).

T7. **Malicious SPs**. A user's physical location could be disclosed to an intruder if an SP abuses the user's identity and location information. However, it is essential that the SP knows the user's identity and location since the user has a contractual/charging relationship with its SP. Hence SPs will need to protect their users against breaches of privacy, and utilise secure access control and audit mechanisms for their user databases.

4.3. Requirements for an ILP mechanism

We now list general requirements for ILP mechanisms, based on our analysis of GSM.

- The user's RI should never be transmitted unprotected across the air interface (hence addressing threats T1, T2 and T3).
- The user's RI should never be transmitted unprotected between network entities (NOs and/or SPs), unless the comms. path is inherently secure (hence addressing threat T4).
- The user's RI should only be given to parties needing it for correct network operation; in the limit this could mean that the user's SP is the only entity knowing the user's RI (addressing threat T5). For service provision, only the user's SP needs to know the user's RI, since when an NO provides service to a user it only needs to know the user's TI and who the user's SP is, so that the NO can subsequently charge the SP for service provided to the user (the SP will also need to keep the TI so that the charge can be matched against the user's RI).
- Third parties should be unable to track users by impersonating an SP to an NO, an NO to a user, a user to an NO, or an NO to an SP (hence addressing threats T2, T3 and T6).

Not all these requirements can always be met in a practical system, although at least the first requirement should always be met (unlike in GSM).

4.4. General approaches for providing ILP

We now discuss two general approaches for providing ILP, which typically occurs in combination with entity authentication. Section 5 contains examples of the two approaches.

The fundamental problem is to meet the first identified requirement, i.e. to avoid transmission of users' RIs on the air interface. Note that the reason why addresses of some kind need to be sent across the air interface is because it is a broadcast medium; NOs need to have a means of distinguishing between users, and users need to have a way of deciding which communications are meant for them.

In the first approach, where symmetric encipherment is used, addresses cannot be enciphered. This is because the NO needs to know which key to use to decipher an address, i.e. the NO needs to read the address *before* deciphering it. Similarly, a user needs to read an address embedded in an enciphered data string before deciding whether it should attempt to decipher it. Of course, these problems disappear if all entities use the same key, but this is very insecure and we do not consider this approach further here. This has led to the use of temporary identities (as in GSM) where the RI is not used as an address, and instead a 'temporary' address (TI) is used to identify a user, and this TI changes at regular intervals. The new TI is chosen by the NO and sent to the user in enciphered form, thus preventing an intruder from linking old TIs to new ones. The problem with the GSM approach is the need to use the user's RI prior to setting up an initial TI; this problem can be avoided by using two levels of TIs, as in the approach of Section 5.1. Apart from this example, another scheme using TIs has been proposed by Mu and Varadharajan, [12], who refer to *subliminal identities* instead of TIs.

In the second approach, where asymmetric encipherment is used, it is possible to encipher addresses, at least on the 'up link', i.e. in communications between mobile users and an NO. This is because users can encipher data sent to the NO using the NO's public encipherment key. Protecting the 'down link' is rather more problematic, and still requires the use of some form of TI. However the 'set up' problems associated with GSM can probably be avoided by using this approach. The only remaining problem is to ensure that a user knows which NO it is sending to (and hence can use the right public key), and possesses reliable copies of public encipherment keys for all NOs it may wish to use. An example of such a scheme is given in section 5.2; another scheme of this type is in section 9.4 of [13].

In addition, Beller et al., [14], give one symmetric based and three asymmetric based authentication protocols for use in mobile systems. Whilst the symmetric based mechanism does not provide ILP services, the asymmetric based protocols provide a level of ILP by encrypting the user identity using the public key of an entity roughly corresponding to our NO, thus ensuring that only the NO knows the user's true identity. Carsen, [15], proposed some enhancements to the protocols in [14], although the ILP mechanisms remain the same. Federrath et al., [16], proposed an ILP scheme for mobile systems which prevents a user's SP from tracking a user's movements, and Jackson, [17], in the same proceedings, proposed a very similar scheme to prevent 'management' from spying on users. In these schemes, a mobile user needs to have know the entire route from himself to his SP (consisting of a number of NOs), all these NOs' public keys, and also has to compute asymmetric encryptions several times (one for each NO in the route) during every location update process. These requirements are probably unrealistic for the real mobile user with limited computational power and memory.

4.5. Legal and operational limitations on ILP

In the discussion of ILP requirements in section 4.3, we ignored the domain management requirements applying to NOs. There are two issues, applying in some domains, affecting the provision of ILP.

- The *Calling Line Identifier (CLI)* requirement necessitates that called entities are provided with the CLI (which typically means the telephone number) of the party calling them.
- The *Warranted Interception* requirement means that law enforcement agencies must be given access to certain calls starting or terminating within their domain, typically when an interception warrant has been issued. In principle this requirement could also be applied to all calls routed through a domain, even if they do not start or terminate within that domain, although this is unlikely (see [8]). For details of evolving European rules see [18].

This means that some NOs may need to know the RIs of users sending and/or receiving calls within their network. However, this does not mean it is essential for the ILP scheme used to always transfer a user's RI to an NO. It may be more appropriate to have RIs routinely transferred from SPs to NOs only when NOs need them for legal and/or operational reasons. Thus one could envisage a situation where some NOs will (by law) not provide service to a user unless the user's SP is prepared to provide the user's RI to the NO, whilst some users may be so concerned about privacy that they refuse to use their mobile telephone in networks where their RI has to be divulged. Hence, if the ILP mechanism can avoid the need for the user's RI to be distributed outside the SP, a whole range of privacy options become possible, giving both users and government agencies the maximum flexibility to manage ILP.

5. Mechanisms for mutual authentication providing ILP

In GSM networks it is theoretically possible for an intruder to masquerade as an NO by imitating a base station, as GSM only provides *unilateral* authentication of a user to an NO. For GSM it is hard to see how the intruder could gain much from doing this; however, in 3rd generation systems it is likely that NOs will have much more over-the-air control of users. For instance, they may be able to disable faulty terminals directly, or write billing data direct to the UIM (User Identity Module, the UMTS equivalent of a Subscriber Identity Module or SIM). Thus *mutual* (two-way) entity authentication is necessary.

In both mechanisms described, the NO is not automatically given the user's RI; if required for legal or operational reasons, the RI can be sent from SP to NO in addition to the specified information. Also in both mechanisms the SP acts as a TTP to help provide authentication and key establishment.

5.1. A mechanism based on symmetric cryptography

5.1.1. Background

This mechanism was previously outlined in [2]. It has the advantage that it establishes a temporary user-NO key, i.e. there is no need for NO-SP communication once a user has registered with an NO. This contrasts with the GSM scheme, which needs regular NO-SP communications to transfer challenge-response pairs. It combines the provision of ILP, entity authentication and session key generation in a single mechanism, and also conforms to the relevant ISO/IEC standard, [19].

The mechanism provides the following security features:

1. Mutual entity authentication between user and NO.
2. User identity confidentiality over the communications path between user and NO.
3. Session key establishment between user and NO for use in providing other security features, e.g. for confidentiality and/or integrity for data passed between user and NO.

The mechanism makes use of the following types of cryptographic key:

- user - SP key: K_{SU}, a secret key known only to a user and its SP, and which remains fixed for long periods of time.
- user - NO key: K_{NU}, a secret key known only to a user, its SP and its 'current' NO. These keys may remain fixed while a user is registered with an NO. Associated with every such key is a Key Offset (*KO*), which is used in conjunction with the user - SP key K_{SU} to generate K_{NU}.
- session key: K_S, a secret key known only to a user and its current NO, i.e. the NO with whom the user is registered. A new session key, for use in data encipherment and/or other security features, is generated as a result of every use of the authentication mechanism.

The mechanism makes use of the following cryptographic algorithms:

- user authentication algorithm: A_U, which takes as input a secret key and data string and outputs a check value *RES*.
- SP authentication algorithm: A_S, which takes as input a secret key and data string and outputs a check value *RES*. This algorithm may be the same as A_U.
- identity hiding algorithm: C_U, which takes as input a secret key and data string and outputs a string *CIPH* used to conceal a user identity.
- session key generation algorithm: A_K, which takes as input a secret key and data string and outputs a session key K_S.
- user - NO key generation algorithm: A_N, which takes as input a secret key and data string and outputs a user - NO secret key K_{NU}. This algorithm may be the same as A_K.

The mechanism makes use of the following types of temporary identifiers:

- temporary user identity for NO: TI_N, an identity used to identify a user to the NO with which they are currently registered. It is known to the user and the current NO.
- temporary user identity for SP: TI_S, an identity is used to identify a user to its SP. It is known to the user and its SP.

There are two versions of the mechanism, depending on whether or not the user is currently registered with the NO; we consider them separately, although they are closely related. In the description, as throughout, $X||Y$ denotes the concatenation of data items X and Y.

5.1.2. Current registrations

We first consider the case where the user is already registered with the NO, so that the user and NO share a valid temporary identity TI_N and secret key K_{NU}. The mechanism for this case consists of three messages exchanged between user and NO (the SP is not involved):

1. **user** $\rightarrow$ **NO**: TI_N, RND_U
2. **NO** $\rightarrow$ **user**: RND_N, $TI'_N \oplus CIPH_N$, RES_N
3. **user** $\rightarrow$ **NO**: RES_U

RND_U and RND_N are random 'challenges' generated by user and NO respectively. RES_U and RES_N are 'challenge responses' generated by user and NO respectively, where $RES_N = A_U(K_{NU}, RND_N||RND_U||TI'_N)$, and $RES_U = A_U(K_{NU}, RND_U||RND_N)$. TI'_N is the 'new' user TI for use with the NO, and will replace the current value TI_N. $CIPH_N$ is a string of bits used to conceal TI'_N whilst in transit between NO and user, where $CIPH_N = C_U(K_{NU}, RND_U)$. The user and NO can compute a session key K_S as $K_S = A_K(K_{NU}, RND_U||RND_N||TI'_N)$.

5.1.3. New registrations

We second consider the case where the user is not registered with the NO, and so user and NO do not share any information. The mechanism for this case consists of five messages exchanged between user, NO, and the user's SP.

1. **user** $\rightarrow$ **NO**: TI_S, RND_U
2. **NO** $\rightarrow$ **SP**: TI_S, RND_U
3. **SP** $\rightarrow$ **NO**: $TI'_S \oplus CIPH_S$, KO, K_{NU}, RES_S
4. **NO** $\rightarrow$ **user**: $TI'_S \oplus CIPH_S$, KO, RES_S, RND_N, $TI'_N \oplus CIPH_N$, RES_N
5. **user** $\rightarrow$ **NO**: RES_U

First note that we assume that a secure channel is available for exchanging messages 2 and 3 between NO and SP. As previously, RND_U and RND_N are random 'challenges' generated by user and NO respectively, and RES_U, RES_N, and RES_S are 'challenge responses' generated by user, NO, and SP respectively. RES_N and RES_U are calculated as in Section 5.1.2, and $RES_S = A_S(K_{SU}, RND_U||KO||\ TI'_S)$. TI'_S is the 'new' user TI for use with the SP, and will replace the current value TI_S. As previously, TI'_N is the 'new' user TI for use with the NO. $CIPH_S$ is a string of bits used to conceal TI'_S whilst in transit between SP and user, where $CIPH_S = C_U(K_{SU}, RND_U)$. $CIPH_N$ (computed as previously) is a bit-string used to conceal the new TI TI'_N whilst in transit between NO and user. On receipt of message 4, the user can compute the NO secret key $K_{NU} = A_N(K_{SU}, KO||NOID)$, where $NOID$ is the NO's identifier; the same calculation is done by the SP on receipt of message 2. As previously, user and NO can compute session key $K_S = A_K(K_{NU}, RND_U||RND_N||TI'_N)$. As a result of the mechanism, user and NO will share a secret key K_{NU} and a TI TI'_N.

5.2. A mechanism based on asymmetric cryptography

5.2.1. Requirements

This mechanism is based on a combination of public key encipherment and symmetric cryptographic techniques. Nonces are used for checking timeliness. The following cryptographic functions are used.

- A public key encipherment function E (which the user and SP must implement). We use $E_{K+}[X]$ to denote public key encipherment of data X using public encipherment key $K+$.
- A cryptographic check function f (which the user, NO and SP must implement). We use $f_K(X)$ to denote the (check-value) output of f given input data X and key K.
- A symmetric encipherment function e (which user, NO and SP must implement). We use $e_K(X)$ to denote the output of e given input data X and key K. This encipherment algorithm must provide integrity and origin authentication (c.f. requirements (a), (b) in Clause 4 of ISO/IEC 11770-2, [20]). If necessary, the encipherment algorithms used by the two pairs: user/SP, and NO/SP, can be distinct; we have assumed that a single algorithm is used to simplify the presentation.

The following keys need to be in place:

- The SP needs to generate a public key/private key pair for the public key encipherment algorithm. The user must have a reliable copy of the SP's public encipherment key, K_{S+}.
- The user and SP must share a secret key K'_{US} for the cryptographic check function f.
- The two entity pairs: user/SP, and NO/SP, both need to share a secret key for the symmetric encipherment algorithm, denoted by K_{US} and K_{NS} respectively.

In addition the user, NO and SP must be able to generate non-repeating nonces, the user must be able to generate temporary identities, and the SP must be able to generate session keys.

5.2.2. The protocol

The following protocol (partly) conforms to Key Establishment Mechanism 9, specified in Clause 6.3 of ISO/IEC 11770-2, [20]. One point at which it significantly diverges from the standard is that the user's identity U is never sent in clear text and is known only to the SP and itself (the standard protocol would require U to be sent in clear text in message **M2**).

M1. **user → NO**: $R_U || S || N || E_{K_{S+}}[U || T_U || f_{K'_{US}}(R_U || U || N || T_U)]$

M2. **NO → SP**: $R_N || R_U || S || N || E_{K_{S+}}[U || T_U || f_{K'_{US}}(R_U || U || N || T_U)]$

M3. **SP → NO**: $e_{K_{NS}}(R_N || K_{UN} || T_U) || e_{K_{US}}(R_U || K_{UN} || N || T_U)$

M4. **NO → user**: $T_U || R'_N || e_{K_{US}}(R_U || K_{UN} || N || T_U) || f_{K_{UN}}(R'_N || R_U || T_U)$

M5. **user → NO**: $f_{K_{UN}}(R_U || R'_N || N)$

The protocol procedure is as follows; if at any point a check fails, then the protocol is aborted.

1. The user generates and stores a nonce R_U, and generates a new temporary identity, T_U. The user then sends the NO an authentication request **M1**, in which it lets the NO know its SP is S. The user's real identity (U) is enciphered using K_{S+} so that only the SP can read it.
2. In **M2** the NO forwards the user's request to the SP, appending (and storing) a nonce R_N.
3. On receipt of **M2** the SP deciphers the enciphered string using its private key. The SP then checks the output of f using its copy of K'_{US}. The SP retrieves the temporary identity T_U, generates a session key K_{UN} for use by user and NO, and distributes them in **M3**. The SP maintains a database of relationships between users and temporary identities.
4. On receipt of **M3**, the NO deciphers (and simultaneously integrity checks) the first part of the message. The NO then checks that the nonce it contains is correct, and also uses the nonce to link the message with the correct 'transaction'. The NO then retrieves the new temporary identity T_U and session key K_{UN}, and uses the latter to generate the check-value in message **M4** which is a function of a second nonce, R'_N, which the NO also stores. Note that, when using broadcast channels, the user's address must be embedded in any message sent to it. Thus message **M4** is prefixed with T_U, to indicate that, if necessary, T_U can be used as the broadcast address for user U without compromising user U's anonymity.
5. On receipt of **M4**, the user deciphers (and integrity checks) the enciphered part. The user checks that the nonce it contains is correct, and retrieves the new session key K_{UN}, which is then used to verify the check-value in the message and to generate message **M5**.

6. On receipt of **M5**, the NO verifies the check-value by recomputing it.

The NO is not given the user's RI, and can only identify a user by the temporary identity T_U supplied by the SP. The NO will use the temporary identity T_U when communicating with SP in order to be recompensed for the cost of providing service to the user.

6. Conclusion

The two protocols have the following advantages over the GSM approach mentioned in section 4.1.

1. User RIs are never transmitted in clear text in the mobile radio path (or, for the 2nd mechanism, in the NO-SP channel).
2. NOs are not given access to a user's RI.
3. Authentication of both NO and SP is implicitly included.

The protocols can prevent threats T1-T6 in section 4.2. Threat T7, i.e. that an SP abuses user identity and location information, can only be prevented by internal management controls imposed by an SP.

A variant of SVO logic, [21], has been used to verify the mechanisms' correctness; in fact logical analysis revealed a subtle flaw in a previous version of the first mechanism which has now been corrected.

The cost of the second mechanism as compared with conventional protocols, for example that presented in section 5.1, is as follows. Each SP must have a public key known to its all users and keep a corresponding private key secret, and each user has to compute $E_{K_{S+}}[U||T_U||f_{K'_{US}}(R_U||U||N||T_U)]$, which has then to be checked by the SP. Note that, for the RSA algorithm, an encryption operation can be made significantly more efficient than a signature operation, since a relatively small public exponent can be chosen. Moreover, transmission of a user-computed signature could also potentially compromise the confidentiality of a user, if the user's public verification key is widely known.

Finally note that both protocols rely on the shared key K_{US} remaining secret long term; other slightly more complex versions of the mechanisms can be devised which do not have this requirement. Also, variants of the 2nd protocol can be devised to deal with various location update requirements, including a 3-message scheme corresponding to the case of section 5.1.2.

References

[1] UK DTI/EPSRC LINK PCP 3GS3 Technical Report 1, *Security features for third generation systems*. Vodafone Ltd., GPT Ltd., Royal Holloway, Univ. of London. Final version, Feb. 1996.

[2] L. Chen, D. Gollmann and C. Mitchell, 'Tailoring authentication protocols to match underlying mechanisms'. In: J. Pieprzyk and J. Seberry (eds.), *Information Security and Privacy*, Springer-Verlag LNCS 1172 (1996) pp. 121-133.

[3] L. Chen, D. Gollmann, and C.J. Mitchell, 'Distributing trust amongst multiple authentication servers'. *Journal of Computer Security* **3** (1994/95) pp. 255-267.

[4] L. Chen, D. Gollmann, and C.J. Mitchell, 'Authentication using minimally trusted servers'. *ACM Operating Systems Review*, submitted (April 1997).

[5] U.M. Maurer, 'Secret key agreement by public discussion from common information'. *IEEE Transactions on Information Theory* **39** (1993) pp. 733-742.

[6] A.D. Wyner, 'The wire-tap channel'. *Bell System Technical Journal* **54** (1975) pp. 1355-1387.

[7] C.J. Mitchell, 'A storage complexity based analogue of Maurer key establishment using public channels'. In: C. Boyd, (ed.), *Cryptography and Coding - Proceedings 5th IMA Conference, Cirencester, December 1995*, Springer-Verlag LNCS 1025 (1995) pp. 84-93.

[8] M.P. Hoyle and C.J. Mitchell, 'On solutions to the key escrow problem'. In: *Proceedings of State of the Art and Evolution of Computer Security and Industrial Cryptography*, Leuven, June 1997 (to appear in Springer-Verlag LNCS).

[9] N. Jefferies, C. Mitchell and M. Walker, 'A proposed architecture for trusted third party services'. In: E. Dawson and J. Golic, (eds.), *Cryptography: Policy and Algorithms*, Springer-Verlag LNCS 1029 (1996) pp. 98-104.

[10] J. Brown, 'Combined multiple access and encryption for CDMA systems'. In: *Proceedings of the 3rd International Symposium on Communication Theory and Applications*, Ambleside, UK, July 1995.

[11] ETSI/PT12 GSM-03.20, *Security related network functions*. European Telecommunications Standards Institute, August 1992.

[12] Y. Mu and V. Varadharajan, 'On the design of security protocols for mobile communications'. In: J. Pieprzyk and J. Seberry (eds.), *Information Security and Privacy*, Springer-Verlag LNCS 1172 (1996) pp. 134-145.

[13] UK DTI/EPSRC LINK PCP 3GS3 Technical Report 2, *Security mechanisms for third generation systems*. Vodafone Ltd., GPT Ltd., Royal Holloway, Univ. of London. Final version, May 1996.

[14] M.J. Beller, L. Chang and Y. Yacobi, 'Privacy and authentication on a portable communications system', *IEEE J. on Selected Areas in Comms.* **11** (1993) pp. 821-829.

[15] U. Carlsen, 'Optimal privacy and authentication on a portable communications system'. *ACM Operating Systems Review* **28 no 3** (July 194) pp. 16-23.

[16] H. Federrath, A. Jerichow and A. Pfitzmann, 'MIXes in mobile communication systems: Location management with privacy'. In: R. Anderson (ed.), *Information Hiding*, Springer-Verlag LNCS 1174 (1996) pp.121-135.

[17] I.W. Jackson, 'Anonymous addresses and confidentiality of location'. In: R. Anderson (ed.), *Information Hiding*, Springer-Verlag LNCS 1174 (1996) pp.115-120.

[18] European Union Council Resolution, *International requirements for the lawful interception of communications*. January 1995.

[19] ISO/IEC 9798-4, *Information technology - Security techniques - Entity authentication - Part 4: Mechanisms using a cryptographic check function*. ISO, 1995.

[20] ISO/IEC 11770-2, *Information technology - Security techniques - Key management - Part 2: Mechanisms using symmetric techniques*. International Organization for Standardization, Genève, Switzerland, 1996.

[21] P. Syverson and P.C. van Oorschot, 'On unifying some cryptographic protocol logics'. In: *Proceedings of the 1994 IEEE Computer Society Symposium on Research in Security and Privacy*. IEEE Computer Society Press (1994) pp. 14-28.