Security in future mobile networks

Chris J. Mitchell
Computer Science Department, Royal Holloway, University of London,
Egham, Surrey TW20 0EX, England.
Fax: +44-1784-443420, Email: cjm@dcs.rhbnc.ac.uk.

1. Introduction

Current digital mobile networks, such as those based on the ETSI GSM standards, provide a robust set of security facilities for protecting communications across the air interface. The main security services supported by GSM are confidentiality (of user and signalling data) for data passed across the air interface, authentication of mobiles to a base station, and user identity confidentiality (across the air interface).

Because of their universal nature, and because of the additional requirements generated by high data rate multimedia traffic, standards for future mobile networks will need to support a larger range of security services. Possible new services include:

- end-to-end data confidentiality,
- end-to-end data integrity, and
- incontestable charging.

Moreover, there is much to be gained by standardising some of the management aspects of security provision. In GSM, although the management requirements for security are clear, the exact means by which user key information is generated, stored and accessed is a matter for Network Operators and equipment providers to arrange. This can make the provision of these services rather costly for all concerned, since security management may well end being differently arranged by every Network Operator.

In future mobile networks, possibly operating in a rather more deregulated environment than is the norm at present, standard support for security management will be a very important feature. Without such standards, the required co-operation between the likely large numbers of competing Network Operators and Service Providers could become impossible complex to arrange.

In this paper we will examine some of the security provisions in the emerging ETSI UMTS and ITU-T FPLMTS standards for future mobile telecommunications networks. In particular, we describe a mutual authentication algorithm (which also allows for the provision of mobile user identity confidentiality) that has been proposed for use in these systems. We then go on to consider some of the problems associated with the provision of end-to-end security services in future mobile systems, which may be of particular relevance to future multi-media applications.

In doing so we will describe on-going research into security features and security management for future mobile networks which is being performed by the DTI/EPSRC-funded project ‘Third Generation System Security Studies’ (3GS3), which is a part of the LINK Personal Communications Programme. The collaborators in this project are Vodafone Ltd, GPT Ltd. and Royal Holloway, University of London.

2. Third-generation mobile systems

The term ‘Third-generation’ refers to the next generation of mobile systems beyond existing digital networks such as GSM, DCS1800 and DECT. Such systems are currently being standardised by the ITU (as FPLMTS) and ETSI (as UMTS). They are characterised by the following:

1. a multiple operator environment,
2. multiple environments (residential cordless, mobile, satellite, etc.),
3. multi-vendor/standardised interfaces,
4. use of the WARC-assigned FPLMTS band,
5. higher bit-rates (up to 2Mb/s),
6. migration from existing systems.

Currently, 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 some areas where security might be enhanced in third-generation systems, partly based on lessons from second-generation systems, but mostly deriving from the additional characteristics of third-generation systems noted above.

3. Security features for future networks

The 3GS3 project has considered the provision of security in future mobile networks. Initial studies have identified the likely security threats to mobile networks, within the context of role and functional models for future mobile networks (also defined by the project). Security features necessary to address these threats have been identified and classified.

As part of the 3GS3 project, a number of security features have been selected for detailed study. In particular we are studying what types of security mechanism can best be used to provide these features in future mobile networks. The selected features were chosen using two main criteria:

1. their importance to future networks, and
2. the difficulty of providing the features (that is we have focused our efforts on those features which we expect to be most difficult to provide).

The main security features on which we have concentrated our efforts during the first two years of the project (and which will continue to be the main focus of our research until the project completes in early 1996) are as follows.

- *Entity authentication*. We have focused our attention in particular on entity authentication between the user and Network Operators and/or Service Providers. We have examined, classified, and tested (formally and informally) a large number of possible mechanisms, based on a variety of cryptographic techniques. As a result of our work, we have proposed an entity authentication mechanism to both (ETSI) UMTS and (ITU-R) FPLMTS which has been incorporated into both sets of draft standards. This mechanism is considered in some detail in Section 4. We have also considered problems arising when some of the ‘authentication servers’ within a system may be unreliable, [1].

- *Novel techniques for key distribution*. We have focused in particular on recent work of Maurer (see, for example, [2]). Maurer has shown how the concept of the ‘Wire tap channel’, introduced by Wyner, [3], can be used in a much wider class of situations than was envisaged by Wyner. The basic idea is to make use of the universal presence of noise in communications channels to enable two users to agree a secret key using only ‘public’ channels. The applicability of this idea to practical networks has been investigated, and some interesting new theoretical results have been discovered.

- *End-to-end encipherment, and warranted interception facilities*. These topics are a current subject for research within 3GS3. An overview of the goals of our continuing work, and its potential relevance to multi-media networks, is given in Section 5.

- *Identity and location privacy*. Current GSM networks provide a level of user identity confidentiality. However, 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. Of particular note is the fact that the authentication mechanism described in Section 4 also provides user identity confidentiality.

- *Simultaneous multiple access channel coding and encipherment*. We have looked in particular at CDMA, which appears to be a likely multiple access method for future mobile networks. The claim that CDMA is inherently secure has been critically examined and rejected. The options for using CDMA sequences as part of an encipherment process have also been examined, and a paper has been prepared on this topic, [4].

- *Terminal-related security*. Current mobile networks contain provisions for the black-listing of stolen terminals, and the detection of not type-approved terminal equipment. The need for such facilities in future networks has been critically examined, particularly as the majority of mobile terminals are likely to be relatively low-cost items. Whether a universal scheme is adopted, or a scheme which only applies to valuable terminals (such as multi-media devices), remains a topic for debate.

One predominant feature in much of the work of the 3GS3 project has been a continuing commitment to standards contributions, both in ETSI and in ITU-R. Apart from mechanisms developed and adopted, much of the text in the draft standards covering the classification and analysis of security features has been based on 3GS3 contributions.

4. A mechanism for mutual authentication

In GSM networks it is theoretically possible for an intruder to masquerade as a network operator by imitating a base station, as GSM only provides *unilateral* authentication of a user to a network operator. In the case of GSM it is difficult to see how the intruder could obtain much benefit from doing this. However, in third-generation systems it is likely that network operators will have considerably 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 (the UMTS equivalent of a SIM). For this reason a *mutual* (two-way) entity authentication mechanism is necessary.

The mechanism we describe here is based on the use of secret key cryptography and provides mutual entity authentication between the user and the network operator.

4.1. Advantages

It sets up (and uses) a temporary key between a user and a network operator. This means that there is no need for communication between a network operator and a service provider once a user has successfully registered with a network operator. This is in contrast with the existing GSM mechanism, which requires regular communications between network operator and service provider to transfer challenge-response pairs.

It also combines the provision of user identity confidentiality, entity authentication and session key generation in a single mechanism.

The mechanism also conforms to the relevant ISO/IEC standard, [5].

4.2. Possible Restrictions on User Identity Confidentiality

Note that, in the mechanism described here, the network operator is not automatically given the user's *IMUI* (International Mobile User Identity). If this is necessary for legal and/or operational reasons it can be included in the third message of the 'new registration' authentication mechanism.

4.3. Security Features Provided by the Mechanism

The mechanism provides the following security features:

1. Mutual entity authentication between the user and the network operator.
2. User identity confidentiality over the communications path between the user and the network operator.
3. Session key establishment between the user and the network operator for use in providing other security features, possibly including confidentiality and/or integrity for data passed between the user and network operator.

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

- *user - service provider key K_{SU}*. These are secret keys known only to a user and their service provider. These secret keys remain fixed for long periods of time.
- *user - network operator key K_{NU}*. These are secret keys known only to a user and their 'current' network operator. These keys may remain fixed while a user is registered with a particular network operator. Associated with every such key is a Key Offset (*KO*), which is used in conjunction with the user - service provider key K_{SU} to generate K_{NU}.
- *session key KS*. These are secret keys also known only to the user and their current network operator (i.e. the network operator with whom they are registered). A new session key is generated as a result of every use of the authentication mechanism. These keys can be used for data encipherment, and/or for the provision of other security features.

The mechanism makes use of the following cryptographic algorithms:

- *user authentication algorithm A_U*. This algorithm takes as input a secret key and a data string and outputs a check value *RES*.
- *service provider authentication algorithm A_S* This algorithm takes as input a secret key and a data string and outputs a check value *RES*. This algorithm may be the same as or distinct from the algorithm A_U.
- *identity hiding algorithm C_U*. This algorithm takes as input a secret key and a data string and outputs a string *CIPH* used to conceal a user identity.
- *session key generation algorithm A_K*. This algorithm takes as input a secret key and a data string and outputs a session key K_S.
- *user - network operator key generation algorithm A_N*. This algorithm takes as input a secret key and a data string and outputs a user - network operator secret key K_{NU}. This algorithm may be the same as or distinct from the algorithm A_K.

The mechanism makes use of the following types of identifiers:

- *International Mobile User Identity IMUI*. This is an identity permanently associated with a user. The IMUI is never passed across the air interface, thus preventing its unauthorised disclosure.

- *network operator identity NOID*.

- *temporary user identity for network operator $TMUI_N$*. This (temporary) identity is used to identify a user to the network operator with which they are currently registered. It is known to the user and to the current FPLMTS network operator.

- *temporary user identity for service provider $TMUI_S$*. This (temporary) identity is used to identify a user to its service provider. It is known to the user and to its service provider.

4.4. The mechanism

There are two versions of the mechanism, depending on whether or not the user is currently registered with the network operator. We consider the two cases separately (although they are closely related).

4.4.1. Current registrations

We first consider the case where the user is already registered with the network operator. This means that the user and the network operator will share a valid temporary identity $TMUI_N$ and secret key K_{NU}. The mechanism for this case consists of three messages exchanged between the user and the network operator. The service provider is not involved.

The three messages are as follows.

1. **user → NO**: $TMUI_N$, RND_U

2. **NO → user**: RND_N, $TMUI'_N \oplus CIPH_N$, RES_N

3. **user → NO**: RES_U

The values RND_U and RND_N are random 'challenges' generated by the user and the network operator respectively.

The values RES_U and RES_N are 'challenge responses' generated by the user and the network operator respectively. RES_N is calculated using the user authentication algorithm A_U with key input K_{NU} and data string input the concatenation of RND_N, RND_U and $TMUI'_N$. RES_U is calculated using the user authentication algorithm A_U with key input K_{NU} and data string input the concatenation of RND_U and RND_N.

$TMUI'_N$ is the 'new' temporary user identity for use with the network operator. This will replace the current temporary identity $TMUI_N$

$CIPH_N$ is a string of bits used to conceal the new temporary identity $TMUI'_N$ whilst it is in transit between the network operator and the user. It is calculated using the identity hiding algorithm C_U with secret key input K_{NU} and data string input RND_U.

The user and the network operator can compute a session key K_S as the output of the session key generation algorithm A_K when given secret key input K_{NU} and data string input the concatenation of RND_U, RND_N and $TMUI'_N$.

4.4.2. New registrations

We second consider the case where the user is not registered with the network operator. This means that the user and the network operator do not share any information. The mechanism for this case consist of five messages exchanged between the user, the network operator, and the service provider of the user.

The five messages are as follows.

1. **user → NO**: $TMUI_S$, RND_U

2. **NO → SP**: $TMUI_S$, RND_U

3. **SP → NO**: $TMUI'_S \oplus CIPH_S$, KO, K_{NU}, RES_S

4. **NO → user**: $TMUI'_S \oplus CIPH_S$, KO, RES_S, RND_N, $TMUI'_N \oplus CIPH_N$, RES_N

5. **user → NO**: RES_U

First note that we assume that a secure channel is available for exchanging messages 2 and 3 between the network operator and service provider.

As previously, the values RND_U and RND_N are random 'challenges' generated by the user and the network operator respectively.

The values RES_U, RES_N, and RES_S are 'challenge responses' generated by the user, network operator, and service provider respectively. RES_N and RES_U are calculated as in the previous case. RES_S is calculated using the service provider authentication algorithm A_S with key input K_{SU} and data string input the concatenation of RND_U, KO and $TMUI'_S$.

$TMUI'_S$ is the 'new' temporary user identity for use with the service provider. This will replace the current temporary identity $TMUI_S$. As previously, $TMUI'_N$ is the 'new' temporary user identity for use with the network operator.

$CIPH_S$ is a string of bits used to conceal the new temporary identity $TMUI'_S$ whilst it is in transit between the service provider and the user. It is calculated using the identity hiding algorithm C_U

with secret key input K_{SU} and data string input RND_U. As previously, $CIPH_N$ is a string of bits used to conceal the new temporary identity $TMUI'_N$ whilst it is in transit between the network operator and the user.

On receipt of message 4, the user can compute the network operator secret key K_{NU} as the output of the network operator key generation algorithm A_N, when given as secret key input K_{SU}, and data string input the key offset *KO* concatenated with the network operator identity *NOID* (this same calculation is done by the service provider on receipt of message 2).

As previously, the user and the network operator can compute a session key K_S as the output of the session key generation algorithm A_K when given secret key input K_{NU} and data string input the concatenation of RND_U, RND_N and $TMUI'_N$.

Note that, as a result of the above mechanism, the user and the network operator will share a secret key K_{NU} and a temporary identity $TMUI'_N$.

4.5. Conclusion

The proposed mechanism provides several security features that will be required for third-generation systems. It is only one building block in a complete security architecture for third-generation systems. Further work is required to establish the efficiency of the mechanism, and to determine how the mechanism can be managed. Current on-going work within the 3GS3 project involves using a variant of the SVO logic, [6], to verify the correctness of this mechanism; in fact, analysis of the protocol using this logic has already revealed a subtle flaw in an earlier version of the protocol which has now been corrected.

5. End-to-end security features

5.1. Multi-media security requirements

Multi-media terminals will clearly place demanding bandwidth requirements on the mobile network. These requirements have relatively little direct effect on the provision of security features, except for the need to ensure that any directly data-related security features, such as the provision of data confidentiality, are implemented using techniques which can handle high-bandwidth data. In practice this will, for example, mean that encipherment techniques used on the air interface must be capable of handling high throughput rates, even when implemented on mobile terminals. However, this should not present too much of a problem since multi-media terminals will not be low cost items, and the provision of appropriate processing capabilities to handle high data-rate encipherment should not add significantly to the overall cost of such devices. Note that, whichever technique is chosen for providing air interface encipherment, it will need to be capable of implementation on the whole range of possible user terminals, since use of more than one encipherment algorithm would cause considerable practial difficulties.

Much more significant to the designers of security features for future mobile networks, are the likely security requirements of the users of these multi-media services, i.e. we need to identify what types of security services these users will need. These needs are potentially very different from those of ‘voice’ users of existing networks. Of particular interest are likely to be issues such as end-to-end integrity and confidentiality protection (by comparison with the existing networks which do not address integrity provision, and only provide encryption for the air interface).

Of course this need to identify the likely security requirements of users applies equally to other users of data transfer facilities in future mobile networks, who will probably also have requirements for end-to-end security features.

5.2. End-to-end encipherment and legal interception

Of all the end-to-end security features, the provision of end-to-end confidentiality is by far the most problematic. The problem is a ‘political’ rather than a technical one, and arises from the need for law enforcement agencies to be given access (when the appropriate legal authorisation exists) to any specified communications path. The need for such access is clearly extremely valuable in combating criminal activity, but also needs to be carefully controlled because of the civil liberties issues which arise.

The US approach (with *Clipper*) has been to define a secret algorithm, for which the government has to supply the keys. Whilst it clearly solves the problem of lawful interception (the government will always know the key!) it also has serious shortcomings.

- Controlling access will be extremely difficult (once a key is divulged to a law enforcement agency, all traffic for that user can be read `for ever’).

- The scheme is clearly of no value to other countries, and an international solution is needed (particularly in a part of the world like Europe with many different national governments and legal frameworks).

Thus there is a need for a more flexible approach which offers the desired trans-national facilities. This is not an easy problem to solve, particularly given there is at present no agreed policy between governments!

There is a growing consensus that ‘Trusted Third Parties’ (TTPs) offer a possible means of providing the desired warranted access, at the same time as meeting legitimate user needs for confidentiality. Any user wishing to make use of end-to-end confidentiality will need to register with a TTP in their country. These TTPs will be ‘licensed’ in some way, and will be required to give up keying information to government agencies when provided with appropriate warrants. A key management system will then need to be devised which will allow a TTP to provide warranted access to any specified user’s incoming and outgoing traffic, without compromising any other user or alerting any other TTPs (or any other national governments). The TTP should also provide its users with keys for communication with all other users (and hence the TTP will be seen as providing a valuable role to its registered users, as well as to the law enforcement agencies).

3GS3 is in the process of developing possible TTP-based solutions to the warranted access problem, which we believe will offer considerable advantages over other proposed schemes.

6. Concluding remarks

Future networks are likely to have much more stringent security requirements than current digital networks. This is to a large extent due to the much wider variety of uses to which communications channels to and from mobile terminals are likely to be put. This trend to wider use will, of course, be accelerated by the likely growth in multi-media terminal capability and availability.

In the 3GS3 project, the mechanisms to meet these likely future needs are being examined (and, where necessary, designed) and solutions to these security problems are being offered to the relevant international standards bodies. This approach is exemplified by the entity authentication scheme which has now been incorporated into both the UMTS and the FPLMTS draft standards

The project is also actively examining ways to solve the warranted intercept problem. Solutions to this difficult and politically sensitive problem are essential if end-to-end encipherment is to become a feature in future networks. The availability of such a service may also have a large bearing on the acceptability of such networks for use for multi-media traffic.

References

[1] L. Chen, D. Gollmann and C.J. Mitchell, ‘Key distribution without individual trusted authentication servers’. To be presented at the *The IEEE Computer Security Foundations Workshop VIII*, Ireland, June 1995.

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

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

[4] J. Brown, ‘Combined multiple access and encryption for CDMA systems’. Submitted to the *3rd International Symposium on Communication Theory and Applications*, July 1995.

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

[6] 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, pages 14-28, 1994.