28

Difficulties in Achieving Security in Mobile Communications

I. Nurkic
Security Analyst
16/501 Wilson St. Chippendale. N.S.W. 2008, Australia, tel. 61-2-99025725, fax. 61-2-9902-5141, medan@ee.su.oz.au

Abstract

This work investigates specific challenges in implementing secure mobile digital communications. The objective was to look at some aspects of the problem which may contribute to an exchange of new ideas. Use of an asymmetric key protocol relationship for key management and subscriber authentication, in combination with a chosen symmetric algorithm to protect mobile sessions, is suggested. Dependence on the security of the chip card (smartcard, Integrated Circuit (IC) card), for the mobile unit, is addressed. Consideration of the emerging security protocols and key management methods for open networks (Internet) is included. Ideas are presented generically for a mobile digital system and some comments on the existing Global System for Mobile Communications (GSM) security implementation are included.

Keywords

Security, mobile, communications, GSM, cryptography, asymmetric, smartcard, key

1 BRIEF SUMMARY OF SECURITY REQUIREMENTS

Three overall network security requirements, confidentiality, data integrity and availability (Ford, 1994), apply to mobile voice communications as well. Specifically for mobile communications, security requirements may be expressed as:

- Privacy of information: Prevent the eavesdropping on the radio path. Although the same requirement may be also applied to the traditional telephony (fixed networks), the nature of mobile communications (availability of signals for reception) makes the requirement for privacy stronger in wireless telephony.
- Privacy of movements of mobile users: Prevention of signal tracking, i.e. "confidentiality of signaling information" (Unknown, 1995).
- User authentication: In order to protect both the business and services of the mobile service provider and its customers, service should be available only to the legitimate

J. L. Encarnação et al. (eds.), *Mobile Communications*

(authenticated) users. In other words, service stealing either by an impersonation or other sophisticated means, should be prevented.

- Subscriber anonymity: Avoid the transfer of genuine user identification information over the network.

If taking into account that there are some security issues related to the charging service, for example how to prevent subscribers repudiating used services, non-repudiation may also be a relevant security requirement in future systems, although it is not specifically addressed in this paper.

2 CRYPTOGRAPHY IN MOBILE COMMUNICATIONS

Security relies on cryptography to achieve data privacy and user authentication. While it is beyond the scope of this paper to present/ analyse cryptographic algorithms, it is important to say that there are some concepts which are common to the cryptographic algorithms currently used in the commercial sector, namely they use a secret parameter - encryption key - to achieve the needed 'unpredictability' of the encryption process. Then, depending on whether an algorithm requires exactly the same encryption key value both to encrypt and decrypt data, or there are two mathematically correlated values used, where one is required for encryption and the other for decryption, algorithms are classified in two families: symmetric and asymmetric algorithms, respectively. The best known algorithms of both classes are publicly documented, some of them standardised either on the international or on a national level, and most of them are licensed.

If looking at some implemented solutions in the mobile environment, 'standard' cryptographic algorithms did not seem applicable in the mobile telephony. There seem to be three reasons for this:

1. Publicly known symmetric key algorithms, like DES (Data Encryption Standard), at the time when considered, could not satisfy mobile environment's specific security (key management) requirements. (Also, some characteristics of older symmetric algorithms - namely key lengths, already were not good enough for security of emerging mobile systems.)
2. Asymmetric key algorithms, like RSA (Rivest-Shamir-Adleman), were not considered for the implementation in the design of first secure mobile systems either (see more in Section 5).
3. Standard GSM security has adopted a specific security design which employs very specific algorithm(s).

Identified specific implementation difficulties, when implementing security for mobile networks, are:

- Difficulties in synchronising cryptographic key material (keys, initial vectors, seeds) between the mobile unit and the centre. There are only few options available in achieving cryptographic key synchronisation requirement:
    - Broadcast a seed value which is an input parameter in the key generation process. Any solution with broadcasting actually relies on the secrecy of either the encryption algorithm (method) used or some unique value (e.g. user authentication code) in the mobile unit.

  - Rely on sets of pre-generated keys which are stored in the mobile unit. However, it would be impossible to store a large number of random encryption keys (without any functional relationship among the keys), presumably required to last for the lifetime of the mobile telephone. It is not acceptable to use permanent session keys because of the risks of exposure over an extended lifetime. Also, it is not acceptable to store non-random keys because of the risk of discovering the key relationship.
  - Broadcast new keys encrypted under the initially set, 'encrypting' (master) keys. For the schemes using symmetric cryptographic algorithms only, this would result in an unacceptable burden of managing (generating, loading into the handset, storing in the service provider's centre) large number of symmetric master keys (similar as if using symmetric algorithms for secure communications in large networks with a key centre, in general).
  - Broadcast new keys using asymmetric keys (see more in Section 5).
- Air-emitted information is accessible to anyone with a receiver without any additional effort required (see more in Section 4).
- Performance factor: In voice communications, due to the speed required to complete all cryptographic operations, the complete family of current asymmetric algorithms is still unusable, at least in achieving data privacy. However, these algorithms may have a practical value in user authentication, if implemented in a more sophisticated chip card technology, with an integrated modulo arithmetic co-processor.
- Commercial factor (cost) up to date was an obstacle in getting more processing power in the chip card. However situation in this area changes rapidly since the cost is actually dictated by the amount of silicon used - therefore for new micro chips it is no longer true that an enhanced chip design would automatically cost more.

3 END-TO-END SECURITY: UTOPIA?

There are two 'credos' in the data communications security framework. Firstly, complete security relies on an implementation of a standard (universal) method over the complete path. Secondly, a security design is as good (secure), as its weakest component.

An end-to-end security design is still a desired aim of the modern telephony. Namely, the traditional voice communications network provide neither data privacy (confidentiality) nor user authentication using sophisticated cryptographic techniques. For that part of the communication path data information (voice and signaling information) in most cases will not be protected in any existing scenario.

Also, the current mobile digital communications (GSM) carriers base the security on the corresponding European Technology Standards Institute (ETSI) recommendations, which protect information only within the radio domain of that carrier. Note also that there exist weaknesses in the GSM security management, i.e. that the security information generated in the Authentication Centre (AUC) and required in the remote locations of HLR (Home Location Register) and VLR (Visitor Location Register), are distributed over the fixed network. Securing of this sensitive information (triplets used for the challenge-response

authentication method) is left to the discretion of the network provider (e.g. link encryptors).

Even further, ETSI recommendations (ETSI GSM 02.09., 1992), (ETSI GSM 03.20., 1992) are broad enough to allow significant differences in the level of security (i.e. quality of implementation) among the implemented systems, which all formally comply to the referenced standards.

For example, the referenced ETSI recommendation allows the use of different authentication and key generation algorithms (corresponding algorithms A3 and A8, respectively, are not specified). It is not clear if the algorithms used by the network providers had to pass some verification process before implementing them. It is believed that, as long as there are no issued guidelines/ standardised requirements on these algorithms, quite simple reversible functions may have been implemented. Preferably, these algorithms should have the features of an one-way hash function. (Due to the explicit compatibility requirements, the third algorithm used in GSM, the encryption algorithm A5, is a specified, although kept confidential, stream cipher.)

Also, a quality of random number generators used is not mandated in ETSI standards. There may exist a large discrepancy in the quality of implemented generators. Some of them may still use ordinary seeds (e.g. date and time) where all samples can easily be collected in a database used for a 'dictionary attack'. Desirably, implemented generators should use more unpredictable sources of random information and more complex functions for deriving a final 'random' value (namely, RAND used as a challenge in the challenge-response scheme used in GSM). Once again, if there are some guidelines available to follow, this rapidly developing industry would follow them.

4 SECRET ENCRYPTION ALGORITHMS: ARE THEY REALLY A NECESSITY?

Today, when achieving communications security in the majority of all industries/ areas involved in electronic communications, beyond the military/ state security sector, public cryptographic algorithms (e.g. DES; RSA) are used. In such public algorithms the encryption method is known and available for an endless process of crypto-analysis therefore their resistance to crypto-analysis may be proven. Security of those systems relies on the secrecy of encryption keys. (A number of security measures is involved but almost all fit into 'key management', i.e. mechanisms and controls, including physically secure, tamper-resistant devices, around cryptographic keys.)

Current mobile security implementations (GSM) are arguably based on the secrecy of encryption algorithm(s). Reasons behind this phenomenon seem to be the following:

1. Air-emitted information is accessible to anyone with a receiver without any additional effort required i.e. it is easy to collect unlimited amounts of 'crypto-material'. As a consequence, it is considered that the risks of attack are much higher than in the conventional communications media.

   (Note that the aims of attack in mobile communications are quite different. Specific nature of the voice communications is such that it is still too complex to change live messages. The prime aim is either a passive attack - to hear personal information - or a special version of user impersonation - to find out subscriber specific information to use mobile services with no charge, etc. However, in generic data communications, an

attacker may be more interested in changing sensitive messages, e.g. value transactions.)

2. Difficulties in synchronising keys between the subscriber's mobile telephone unit and the mobile service centre (as stated in the Section 2 above).
3. Mobile networks are seen as a phenomenon of new technologies, but the concept of wireless communications was already mature in the military domain. Therefore, it seems that security for mobile telephony just followed some data confidentiality principles established in the military scientific circles. In addition, international standards which would dictate a more common implementation approach and which could be based on a public standard algorithm and corresponding key generation process instructions, are not yet available.
4. It seems that the politics (export restrictions on encryption technology) could be responsible for the current restricted access to the security algorithms used in GSM. According to the (Dmargrav, 1995), cellular telephone manufacturers must agree to non-disclosure and obtain special licenses from the British government, since the algorithms were developed in Britain.

The current situation may be overcome with the promises of new technologies, e.g. the IC cards technology improves daily its processing and storage power versus cost. It seems that the advantage of asymmetric key (public key) cryptography, which eliminates the need for key distribution, in combination with a chosen symmetric encryption algorithm, may provide some very good security tools in the area of mobile networks. Also, there is one set of algorithms, so-called zero-knowledge techniques, which may be potentially attractive for future mobile digital systems (Unknown, 1995), but they are not discussed in this paper. The suggested scenario of using asymmetric algorithms is presented in the next section.

5 USAGE OF ASYMMETRIC KEY CRYPTOGRAPHY

Asymmetric key cryptography has three merits very important to resolve issues in network security and, specifically, in mobile communications:

1. Public key, of the asymmetric key pair, of any entity participating in communication may be given to all potential users freely* without any need to hide it and to distribute it via secret channels.
2. Decryption of the original message or of a hash of the message by the sender's own secret key (digital signature), before transfer, represents on its own three guarantees (see also merit no. 3):
    - that the message has not been changed in transport;
    - that it has been originated by the entity which 'signed' the message;
    - that the sender cannot deny that it sent the message (non-repudiation).

---

* The key should be properly certified, i.e. there should not be any suspicion on who stands behind it. Certificates are data formats containing 'signed' public key and some other specified control information like the expiry date, granted by a 'Certification Authority', which is a trusted party whose own public key is too well known and trusted to initiate the chain of trust. Therefore public keys should always be provided together with their certificates to be trusted.

3. There is no need that the entity's secret key ever leave the location where key pair is generated. If it is stored securely (i.e. this location is a tamper resistant entity), this contributes to the security value of the digital signature as an ultimate method to achieve entity authentication and data integrity.

These three features of asymmetric key algorithms can resolve all major key management problems in setting up security for a mobile network system. Actually, there exist all prerequisites to successfully implement security with asymmetric cryptography, within the domain of the mobile digital provider:

- The scheme of the establishment of trust is simplified since the service provider itself may act as a trusted party (Certification Authority) for its customers.
- Each handset contains a piece of tamper-resistant hardware (SIM chip) where both the service provider's public key and the asymmetric key generation software are pre-loaded.

The proposed scenario of setting up security with asymmetric algorithms is the following:

- Mobile service provider's public key is pre-loaded into all mobiles serviced. The centre's secret key is stored securely, probably in a Secure Cryptographic Device (SCD) which has actually generated the centre's key pair and which would perform all other cryptographic services for the central system.
- Subscriber's asymmetric key pair is generated in the mobile's SIM (Subscriber's Identity Module) card. The generated public key is input into the subscriber's database at the time of the customer's registration, probably centrally or as a part of a manually controlled procedure. Corresponding secret key never goes out from an inaccessible IC location.
- At the time of registration, which has to be on-line in this scenario, the public key in the handset has to be certified (signed by the service provider's secret key). This certification function is performed in the SCD's secure environment.
- The resulting User Public Key Certificate has to be loaded into the SIM card. This must be an automatic process, since these certificates are, for required key lengths (more that 512 bits), long unintelligible strings of digits, and it would be both impractical and error prone if they have to be manually entered. (Another, less secure, option is that this step of entering the user's public key into the subscribers database (registration) plays in itself a role of 'public key certification'.)
- The established public key relationship then enables a basis for a safe broadcast of any other keys, which may be organised in various simple or more complex hierarchies, depending on the specific key management design.

In choosing algorithms, their security qualities, speed of cryptographic operations and other factors, are researched. In a commercial environment it is much more attractive then to use a certified public algorithm than a secret algorithm of unknown qualities. This also applies to the choice of data encryption algorithms.

In addition, the quality of key generators is very important. Any fixed pattern in generating subsequent keys should be prohibited.

6 CHIP CARD TECHNOLOGY

Being a separate discipline in modern data security, chip card security is obviously relevant for security of mobile communications. Ideally, mobile devices should implement IC cards which represent the best possible security design in the chip card technology. As indicated in Section 2, cost should not be any longer a reason to employ older chip technology and also not to use chips specialised for calculations used in asymmetric algorithms (i.e. with arithmetic co-processors).

In recent years, a question on how secure are, after all, chip cards, imposes on all industries which use the chip card technology and whose services rely on security of this media. While for some threats to mobile service providers industry, e.g. 'device cloning', security of chip design is irrelevant (since it is much cheaper to replace the original chip), a risk of successful intentional compromise of the complete system by a successful probing into the SIM in the handset, is a viable threat.

To prevent this happening, it may be necessary to mandate a set of minimum security requirements for the Integrated Chips used in mobile communications (hardware, software, initialisation process, etc.).

For example, recent experiences show that the access to the test mode is the most vulnerable feature in chip design. Assuming that an organised crime is ready to invest in equipment (and expertise) required, this vulnerability of the vast majority of all chips manufactured at present, is an open door into the 'hidden secrets' of chip internals.

Assuming that it is possible to access the chip internals by exploring the mentioned weakness in the chip design, another requirement becomes very relevant, namely the internal logical design (access control features of the application and of the underlying operating system).

7 NEW TECHNIQUES - DYNAMIC SETUP OF CRYPTOGRAPHIC KEYS

This Section may also be titled 'What can be learnt from new Internet security developments?'. Actually, only one aspect present in the recent Internet communications security models is relevant to this Section and it is a dynamic setup of cryptographic key relationship required for user authentication and communications privacy.

Current intensive development of security techniques and of underlying key management methods for totally open networks with massive access (Internet), may help to derive ideas on new methods applicable in mobile communications; since if a dynamic key distribution method is viable in the Internet environment, it may be viable for mobile communications as well.

If the new security concepts for emerging Internet applications prove to be adequately secure (for example using a specialised language, like Java, which is designed such that it prevents many harmful activities, like writing into external data space, etc.), features they could support like a dynamic downloading of key generation code, may revolutionise mobile security as well. In particular, such 'dynamic' systems may provide solutions for the third generation systems (beyond GSM) which may require distributed databases.

8 REFERENCES

Ford, W. (1994) Computer Communications Security. *Principles, standard protocols and techniques*. PTR Prentice Hall, New Jersey.
ETSI GSM 02.09., (1992) Recommendation 02.09. *Security Aspects.*
ETSI GSM 03.20., (1992) Recommendation 03.20. *Security-related Network Functions.*
Dmargrav, (1995) GSM Security and Encryption. //www.utw.com/~dmargrav/paper
Unknown author, (1995) Report on security mechanisms in mobile telecommunications system, Trondheim, Norway.

9 BIOGRAPHY

Ilhana has a B.Sc. in Mathematics from University of Sarajevo, Bosnia and Herzegovina and a Masters in Electrical Engineering from the University of Technology in Sydney, Australia. She started to work on communications security issues in eighties in Sarajevo, when she also participated in the CEC COST 11ter security project. Currently she works as a bank's security consultant and for the software company 'Microhit Australia'. She is involved in the work of the corresponding security committees in Standards Australia.