Authentication and Security in Mobile Phones

Greg Rose

QUALCOMM Australia

ggr@qualcomm.com

ABSTRACT

Mobile telephone systems have a checkered reputation regarding security and authentication features after incidents such as “Squidgy” in the U.K. and eavesdropping on Newt Gingrich’s calls in the U.S.. Cellphone fraud accounted for approximately US$750M in lost revenue in North America in 1996. Authentication and security in cellular phones are therefore important issues, and there is existing and ongoing efforts both in the United States and Europe. This paper provides an introduction to these facilities.

Introduction

Security and authentication features were rudimentary in the original analog cellular phones. This led to a strong industry performing cellular phone fraud. In 1996 this industry cost the legitimate service providers some US$750M in revenue in the United States alone, which represented approximately 3% of total industry revenue.[ctia] Authentication and security in cellular phones are therefore important issues, and there is existing and ongoing work both in the United States and Europe.

Second generation phones have attempted to address the authentication problem, and to a lesser extent the security and privacy problems. The reason for this ordering is simply that authentication is linked to revenue, while at the moment customers are generally unwilling to pay for security and privacy. The original cellular phone standard is colloquially known as *AMPS* (Analog Mobile Phone Standard). The second generation analog standard is, not surprisingly, called *NAMPS* (New AMPS). There are also a number of different digital standards. Australia uses *GSM* (Groupe Speciale Mobile), developed in Europe, while the U.S. has two competing standards, *TDMA* and *CDMA* (Time and Code (resp) Division Multiple Access). TDMA is similar in concept to GSM, but very different in detail, while CDMA is radically different being a spread spectrum system.

Note: The term *PCS* (Personal Communication System) is often thought of as a different type of cellular system. Really the only difference between PCS and Cellular communication is the frequencies used. Cellular systems operate in the 800-900 MHz band, while PCS is in the 1.8-1.9GHz band. There are PCS versions of the digital standards, but not for analog.

All of the cryptographic content of the various standards is based on secret key technology. The main reasons for this are:

- the processors in phones, and particularly for GSM as will become evident, are quite limited.

- air bandwidth for call setup argues for short messages, while public key systems tend to send large chunks of data around.

This may change in the next few years, or it may not. Public Key technology based around Elliptic Curves along with faster processors may alleviate these problems.

As far as the authentication and security infrastructure goes, the new standards fall into two groups; GSM stands alone, while the various U.S. standards share quite a lot of infrastructure and architecture. Since the standard which defines the common infrastructure was called "the ANSI accredited TIA IS-41 standard[is41]" these are commonly called IS-41 systems, even though the standard has subsequently been accredited by ANSI. In the following discussion I will use the IS-41 name for anything which has an analog in GSM.

At the time of writing, third generation standards are being developed. The new European standard is called UMTS; the international standard is called IMT-2000, and will probably incorporate both UMTS and future US-developed standards. Where decisions have been made in these efforts, they will be noted below.

Keys

At the top level, both systems have a master key, called the A-Key in IS-41. This is 64 bits (IS-41), or 128 bits (GSM). (In future, all keys in both systems will be 128 bits.) This is used for authentication of the mobile, and for generation of subkeys.

- In GSM, the top level key is used to generate authentication signatures (64 bits) and session keys (64 bits) directly.
- In IS-41, intermediate keys called "Shared Secret Data" are generated. There is an *SSD-A* which is used in authentication signatures, and an *SSD-B* which is used for cryptographic key generation. These are each 64 bits. There are also three session keys generated from SSD-B:
    - The CMEAkey (64 bits)
    - The Voice Privacy Mask (520 bits)
    - The DataKey (32 bits)

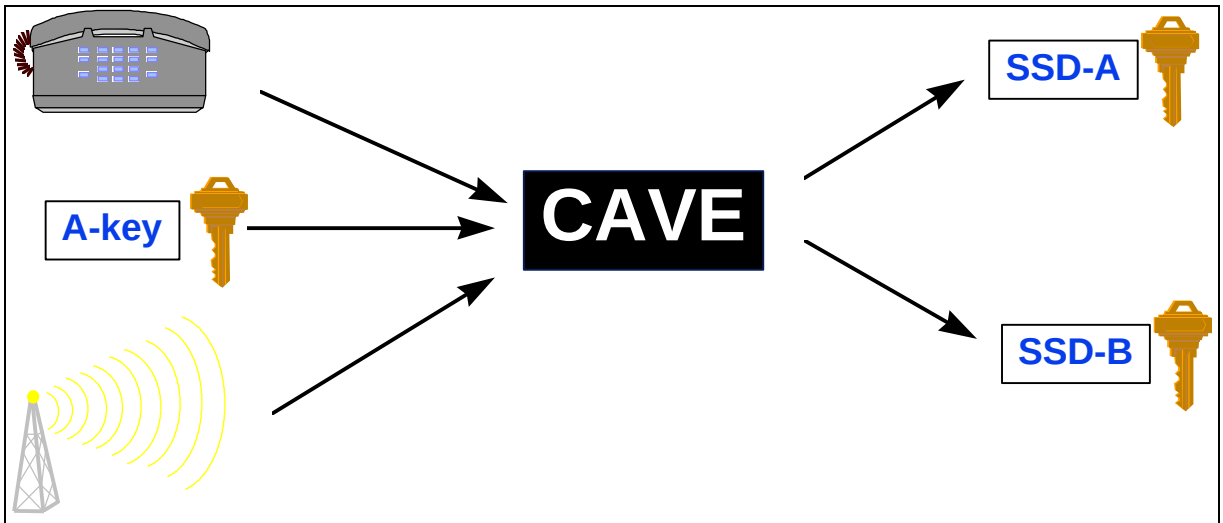

Figure 1 - generation of Shared Secret Data in IS-41.

Details of the algorithms which use these various keys appear below. Figure 1 shows the manner in which the various keys are generated in IS-41. The inputs from the mobile phone are things like the Equipment Serial Number and software revision. The network broadcasts a random number, which is also input to the algorithm.

Authentication Signatures

Before allowing a mobile phone to access the network, the phone must present a response to a challenge, based on the SSD-A (IS-41) or the master key (GSM).

In IS-41 the phone itself contains the secret key, and the algorithm (*CAVE*) which is used to calculate the signature. CAVE is detailed in an export-controlled appendix to the standard, and hence is standardised in all the phones. Figure 2 details the signature calculation process.

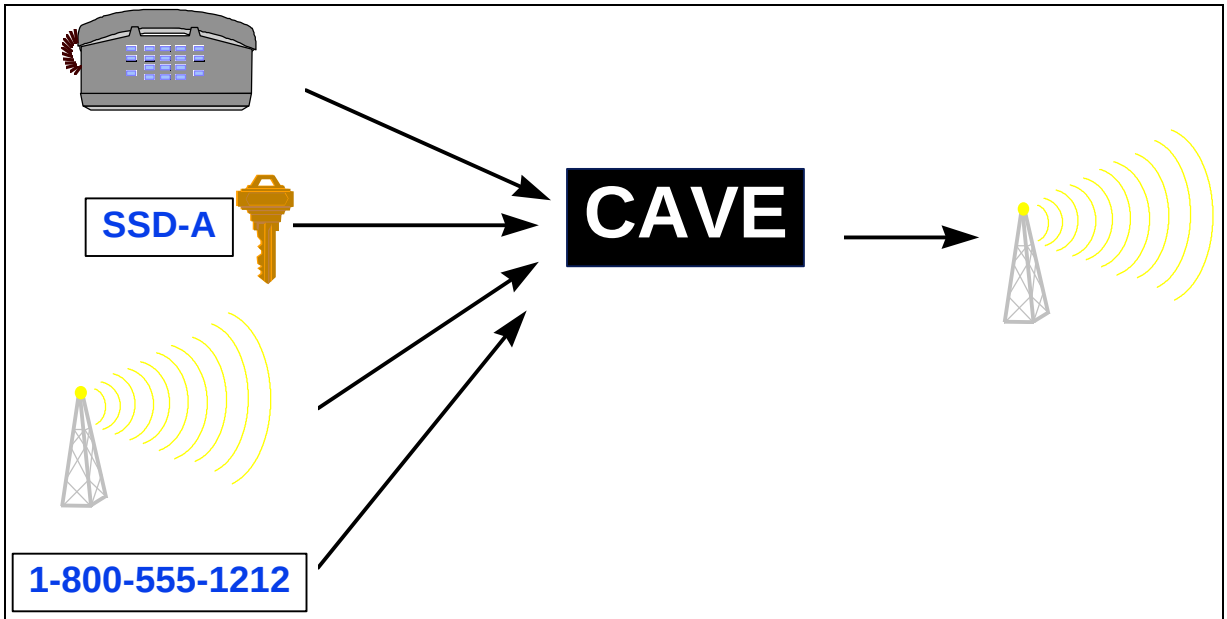

Figure 2 - IS-41 authentication signature calculation.

Note that, for originating a call, the phone number is also input to the algorithm. This is because the random challenge is broadcast and changes regularly, instead of being unique to the origination. This saves message overhead both over the air and within the network. The output from CAVE is truncated to 18 bits, meaning that there is about 1 chance in ¼ million of faking a call by sending a random signature. At these odds, I would invest in the lottery. The shared secret data can be sent to the visited system while roaming, allowing local authentication. If encryption is supported, the various privacy keys are generated soon afterward.

CAVE is a hashing algorithm which works by using a shift register driven walk over the input data and a somewhat random table, and shuffling the inputs. It takes 23 octets of input and produces 16 octets of output.

Future IS-41 systems will replace most of the functionality of CAVE with SHA-1, the US FIPS-180-1 Secure Hash Standard.

In GSM, the picture is quite different, although conceptually similar. The challenge is unique, and is generated within the home system (the system where the phone is registered). The algorithm and the master key are both stored on a smart card called a *SIM* (Subscriber Identity Module). This allows for the possibility that the algorithm may actually vary with different service providers, and indeed this is the case for about 40% of phones. The interface to which the algorithm adheres is called *A3*, and it accepts a 64 bit challenge and produces a 64 bit response, based on the secret key in the SIM. At the same time, an algorithm whose interface is called *A8* calculates the corresponding session key for privacy during the call. The "standard" algorithm performing these functions together is called *COMP128*. This algorithm is held tightly secret by the GSM *MoU* (Memorandum of Understanding Group); only the interface to it is public. Because the algorithm might not even be known at a visited system, the home system has to perform all of the verification and key generation functions. As an optimisation for network traffic, a number of triplets are forwarded upon the first access. These consist of:

1. A challenge to be sent to the mobile station
2. The expected response
3. The session key to be used after authentication succeeds.

Relying on the secrecy of the algorithm is rarely a good move, and indeed COMP128 was disclosed in 1998. Furthermore, the algorithm is weak, allowing disclosure of the A-Key with a few million interactions with the SIM card.

The architecture for UMTS is largely settled. The "triples" will be called the Authentication Vector, and have five elements. There are now two session keys, one for privacy and one for message origin authentication (MAC). The fifth element authenticates the network to the mobile; combined with freshness guarantees, this prevents any form of man-in-the-middle attacks.

Encryption for secrecy in GSM

In GSM the situation for secrecy of voice, signaling data and user data is simple. Once the session has been authenticated, encryption is turned on and everything is protected by the same algorithm, a stream cipher notionally known as *A5*. Actually, there are three different

algorithms, which are negotiated between the phone and the network. A5/1 is based on three shift registers with complicated stepping control, similar to the algorithm mentioned in Schneier[sch96] but not quite. The exact algorithm was reverse engineered early in 1999. A5/2 is a weakened version of A5/1, in which the stepping is controlled by a fourth independent shift register. There is also the option of "no encryption", colloquially called A5/0. The first two of these algorithms are also tightly controlled by the GSM MoU. Unlike the A3/A8 algorithm(s), though, these ones are built into the phone itself, because the SIM doesn't have enough CPU power to calculate the outputs in real time.

Encryption for secrecy in IS-41

In the U.S. standards the situation is much more complicated.

Voice privacy

AMPS, NAMPS – none at all.

TDMA – a fixed voice privacy mask is XORed with the digitised voice. Cryptographically insecure.

CDMA – most of the voice privacy mask is ignored, but the last 42 bits are used as an offset for the output of a linear feedback shift register. Cryptographically this is also not very strong, but this output is used as a spreading code for the spread spectrum transmission. This means that, without knowing the code in advance, it is difficult to even sort out the signal from the background noise.

Signaling data privacy

Data such as numbers dialed, short messages (paging), and DTMF tones are put into data packets and are encrypted using *CMEA* (Cellular Message Encryption Algorithm). This is a variable length block cipher, which works by a table walk using a key-derived somewhat random table, a self-inverse "folding" and the inverse of the first step. This makes the algorithm itself self-inverse, which isn't such a hot idea in retrospect. The cipher is used in ECB (Electronic Code Book) mode, ditto. The packet formats differ for the three standards, but the algorithm is the same. CMEA has been broken[wag97]. A short term drop-in replacement has been proposed and adopted, but is not yet widely deployed.

Data encryption

A stream cipher called ORYX (not an acronym, but upper case anyway) is used for data encryption. It has three Galois configuration linear feedback shift registers, one of which steps normally, and controls the other two; one of these steps according to two different polynomials; the other steps once or twice per byte. The high bytes are passed through a lookup table and combined to form one octet of the output mask. The key length is intentionally very short, for export purposes. ORYX has also been broken.

Future development

There is a long term plan to replace the architecture in IS-41 so that hopefully there won't be so many different algorithms in use. New algorithms in IS-41 will be strong, publicly

scrutinised algorithms. In the short term both CMEA and ORYX are being replaced with a strengthened version of CMEA.

[ctia] Cellular Telecommunications Industry Association, exerpt from confidential report, June 1997.

[is41] TIA IS-41C, Telecommunications Industry Association, 1995.

[sch96] B. Schneier, "*Applied Cryptography*", Second edition, Wiley 1996

[wag97] B. Schneier, J. Kelsey and D. Wagner, "*Cryptanalysis of the Cellular Message Encryption Algorithm*", Proc. Crypto '97.