An Authentication and Security Protocol for Mobile Computing

Yuliang Zheng
Monash University
McMahons Road, Frankston, Melbourne, VIC 3199, Australia
Phone: +61 3 9904 4196, Fax: +61 3 9904 4124
Email: yzheng@fcit.monash.edu.au

Abstract

The main contributions of this paper are: (1) to analyze an authentication and key distribution protocol for mobile computing proposed by Beller, Chang and Yacobi in 1993, and reveal two problems associated with their protocol. (2) to propose a new authentication and key distribution protocol that utilizes a broadcast channel in a mobile network. A particularly interesting feature of the new proposal is that it allows the authentication of a base station by a mobile user to be conducted "at the background", which yields a very compact protocol whose total number of moves of information between a mobile user and a base station is only 1.5 !

Keywords

Authentication, Cryptography, Key Distribution, Mobile Computing, Security

1 SECURITY ISSUES IN WIRELESS NETWORKS

Recent years have seen an explosive growth of interest in wireless (information) networks that support the mobility of users (and terminals). These networks serve as a foundation of future universal, mobile and ubiquitous personal communications systems.

Emerging wireless networks share many common characteristics with traditional wire-line networks such as public switched telephone/data networks, and hence many security issues with wire-line networks also apply to the wireless environment. Nevertheless, the mobility of users, the transmission of signals through open-air and the requirement of low power consumption by a mobile user bring to a wireless network with a large number of features distinctively different from those seen in a wire-line network. Especially, security and privacy becomes more eminent with wireless networks. To this end, we will be primarily concerned with security issues related to or caused by the mobility of users/terminals, open-air transmission of signals and low power supply of a mobile user.

When examining security in a wireless network, a range of issues have to be taken into account. These issues include:

1. identification of a mobile user
2. anonymity of a mobile user (protection of identity)
3. authentication of a base station
4. security of information flowing between a mobile user and a base station
5. prevention of attacks from within a base station
6. hand-over of authentication information
7. the communication cost of establishing a session key between a mobile user and a base station, which is indicated primarily by the total number and length of messages to be exchanged
8. the cost of communications between a mobile user's home domain and a foreign domain where he is currently located, as well as security requirements on the communication links between the two domains
9. the computational complexity of achieving authenticity and security
10. the complexity of computations to be carried out by a mobile user's terminal which is in general much less powerful than a base station

Some issues contradict one another. For instance, to prevent mobile network resources from being abused by a fraudulent user, a network relies on the identification/authentication of a mobile user, which generally requires the user to reveal his or her identity. On the other hand, however, a user who wishes to make anonymous communications may be unwilling to reveal his or her identity. Two recent articles (Brown 1995, Wilkes 1995) survey in details many issues related to security and privacy in mobile networks.

In this extended summary, we assume that the reader is familiar with basic concepts in cryptography, including digital signature, public-key and private-key encryption systems.

2 PREVIOUS PROPOSALS

A concise summary of the authentication and security protocol employed by the global system for mobile telecommunication or GSM (Rahnema 1993) can be found in (Brown 1995). A description of the proposed security and privacy mechanism used in the cellular digital packet data (CDPD) in the US is provided in (Frankel, Herzberg, Karger, Krawczyk, Kunzinger & Yung 1995), where potential threats and attacks to the mechanism, together with possible solutions, are also discussed. Other notable works include (Beller, Chang & Yacobi 1993), (Aziz & Diffie 1994), (Molva, Samfat & Tsudik 1994), and more recently, (Herzberg, Krawczyk & Tsudik 1994), (Asokan 1994) and (Samfat, Molva & Asokan 1995). In the full version of this paper, an outline of each of these protocols, together with a comprehensive comparison of various aspects of these protocols, will be described.

2.1 Beller-Chang-Yacobi Protocol

Among the protocols mentioned above, the one presented in (Beller et al. 1993) deserves special attention, as it represents one of the earliest solutions employing a combination of both private-key and public-key

encryption algorithms. The protocol is based on two computationally infeasible problems: factorization and discrete logarithm. (For this reason, Beller, Chang and Yacobi call their proposal *MSR+DH protocol*.)

As Beller-Chang-Yacobi protocol is partially based on the Diffie-Hellman public key distribution scheme (Diffie & Hellman 1976), it uses a large prime N and a generator α for the multiplicative group $GF(N)^*$. (Note: in their exposition, Beller, Chang and Yacobi also consider a more general case where N can be the product of two large primes.) Both N and α are public. A mobile user m has a pair of public-secret keys (P_m, S_m), where $P_m \equiv \alpha^{S_m} \pmod{N}$. Similarly a base station b too has a pair of public-secret keys (P_b, S_b), where $P_b \equiv \alpha^{S_b} \pmod{N}$.

As in many security solutions, the protocol further requires the existence of a trusted certification authority *ca* who issues a certificate to each mobile user as well as each base station to certify their public keys. (See for instance (Chokhani 1994) for discussions on certification services.) The core part of a certificate issued by the certification authority *ca* to the mobile user m is a digital signature defined by

$$sig_{ca,m} \equiv \sqrt{h(m, P_m)} \pmod{N_{ca}}$$

and similarly, the core part of a certificate issued to the base station b is defined by

$$sig_{ca,b} \equiv \sqrt{h(b, N_b, P_b)} \pmod{N_{ca}}$$

where h is a one-way hash function known to the public, P_m and P_b are the public keys of the mobile user m and the base station b respectively, N_b is the product of two large primes associated with the base station b, and similarly, N_{ca} is the product of two large primes associated with the certification authority *ca*. While N_b and N_{ca} are made public, their prime factors must be kept secret by their respective owners.

Finally a private key encryption algorithm is used in the protocol. One may choose a private key encryption algorithm from a large set of potential candidates, including DES (National Bureau of Standards 1977), IDEA (Lai 1992), RC5 (Lai 1992) and SPEED (Zheng 1996). In the following discussions, we assume that a private key encryption algorithm has been selected, and denote by $Encrypt_K(M)$ the encryption of a message M under a key K, and by $Decrypt_K(C)$ the decryption of a ciphertext C under K.

As indicated below, Beller-Chang-Yacobi protocol consists of five (5) moves (or steps) of information between a mobile user and a base station.

Beller-Chang-Yacobi Protocol

---

1. Mobile User $m \Longrightarrow$ Base Station b
   a message indicating a request for services.
2. Mobile User $m \Longleftarrow$ Base Station b
   Upon receiving the request from m, b sends to m four numbers:
   $$(b, N_b, P_b, sig_{ca,b})$$
   where $P_b \equiv \alpha^{S_b} \pmod{N}$ and $sig_{ca,b} \equiv \sqrt{h(b, N_b, P_b)} \pmod{N_{ca}}$ is the digital signature issued to b by the trusted certification authority *ca*. h, α, N and N_{ca} are all public parameters.
3. Mobile User $m \Longrightarrow$ Base Station b
   Upon receiving $(b, N_b, P_b, sig_{ca,b})$ from b, m checks if
   $$h(b, N_b, P_b) \equiv sig_{ca,b}^2 \pmod{N_{ca}}$$
   The protocol is aborted if the numbers fail to pass the check.

Otherwise, m sends to b two numbers e_2 and e_3 defined by

$$\begin{aligned} e_2 &\equiv x^2 \pmod{N_b} \\ e_3 &= Encrypt_x(m, P_m, sig_{ca,m}) \end{aligned}$$

where x is a number chosen uniformly at random from between 1 and $N_b - 1$, $P_m \equiv \alpha^{S_m} \pmod{N}$ is the public key of m and $sig_{ca,m} \equiv \sqrt{h(m, P_m)} \pmod{N_{ca}}$ is the signature issued to m by the certification authority ca.
Upon receiving (e_2, e_3) from the mobile user m, the base station b extracts x from e_2:

$$x \equiv \sqrt{e_2} \pmod{N_b}$$

by the use of the two secret prime factors of N_b. b then uses x to decrypt e_3:

$$(m, P_m, sig_{ca,m}) = Decrypt_x(e_3)$$

and checks whether

$$h(m, P_m) \equiv sig_{ca,m}^2 \pmod{N_{ca}}$$

The protocol is aborted if m, P_m and $sig_{ca,m}$ fail to pass the test.

4.&5. Mobile User $m \Longleftrightarrow$ Base Station b
Now the mobile user m can calculate $\eta \equiv P_b^{S_m} \pmod{N}$ and $sk = Encrypt_\eta(x)$. Symmetrically the base station b can calculate $\eta \equiv P_m^{S_b} \pmod{N}$ and $sk = Encrypt_\eta(x)$.
To confirm that they have the same session key sk, m and b exchange two known messages which are encrypted under sk. For instance, m can send $Encrypt_{sk}(m)$ to b in exchange of $Encrypt_{sk}(b)$ from b. If the messages are decrypted correctly, sk becomes an authentic session key between m and b.

Note that in practice, a mobile user's computations are all carried out by his personal smart card and/or mobile terminal.

2.2 Problems with Beller-Chang-Yacobi Protocol

After a close examination, we have identified two problems with Beller-Chang-Yacobi protocol. The first problem is related to the inefficiency of the protocol, while the second is concerned with replay-attacks that can be mounted against the protocol.

As the protocol is based on public-key cryptography, an attacker can obtain the public key, as well as its associated digital signature, of a mobile user m, namely $(m, P_m, sig_{ca,m})$. The attacker can then impersonate the mobile user and initiate the protocol with a base station b. The attacker will have no problems in successfully carrying out all the operations involved in the first three (3) moves of the protocol, including passing the test by the base station b in the third move of the protocol. Therefore, the fourth and fifth moves of the protocol must be executed in order for a base station and a genuine mobile user to confirm the consistency of their session keys. Such a 5-move protocol may be inefficient for applications in mobile computing.

The next problem has more serious consequences. Consider an attacker who is malicious towards a mobile user m. The attacker may record communications, including the five moves in Beller-Chang-Yacobi protocol, between the mobile user and a base station b. Some time after m and b complete their communication session, the attacker can initiate a communication with the base station b by replaying

messages previously sent to b by m in Beller-Chang-Yacobi protocol. Clearly the messages will pass all the tests by the base station b, which results in the attacker being successful in impersonating the mobile user m. Now the attacker may be able to transmit, using the name of the mobile user m, an arbitrarily long, but perhaps random and meaningless, message to an arbitrary third user, even though it is computationally infeasible for the attacker to find out the session key sk. This could result in the mobile user m being charged a large amount of money for a communication he never conducted !

3 A NEW PROPOSAL

This section proposes an authentication and key distribution protocol based on a broadcast channel in a mobile network. This protocol is remarkably simple: it consists of only 1.5 moves.

3.1 Certification Authority

As in the case of Beller-Chang-Yacobi protocol, we assume that a mobile network involves a trusted certification authority ca which provides participants of the network, including mobile users and base stations, with public key certification services.

We further assume that the certification authority employs DSS or Digital Signature Standard (National Institute of Standards and Technology 1994). An equally good candidate is a digital signature scheme by Schnorr (Schnorr 1991). The two signature schemes are closely related to each other, and both are based on discrete logarithm over a finite field.

DSS involves three public parameters (p, q, g), where

1. p is a large prime.
2. q is a (large) prime factor of $p-1$.
3. $g \equiv h^{(p-1)/q} \pmod{p}$ with h being an integer satisfying $1 < h < p-1$ and $h^{(p-1)/q} \bmod p > 1$. Note that g is also said to have order $q \bmod p$.

These three parameters are known to all network participants. In addition, DSS requires each user to have a pair of public-secret keys. In particular, the pair of public-secret keys of the certification authority ca is (y_{ca}, x_{ca}), where $y_{ca} \equiv g^{x_{ca}} \pmod{p}$ and x_{ca} is a secret number chosen randomly from $[1, q-1]$.

The pair of public-secret keys of a mobile user m, denoted by (y_m, x_m), and that of a base station b, denoted by (y_b, x_b), are defined in a similar way.

Now the certification authority ca can use DSS to create a certificate for a participant, say a base station b, by digitally signing on a message M using x_{ca}, where M may contain such information as certificate serial number, validity period, the ID of b, the public key of b, the ID of ca, the public key of ca, etc. (See (Chokhani 1994) and (ITU 1993) for a proposed standard format of a certificate.) The digital signature of ca on M is composed of two numbers r and s which are defined as

$$\begin{aligned} r &\equiv (g^k \bmod p) \pmod{q} \\ s &\equiv (h(M) + x_{ca} \cdot r)/k \pmod{q} \end{aligned}$$

where k is a random number chosen from $[1, q-1]$, and h is a one-way hash function. NIST specifies SHS (National Institute of Standards and Technology 1995) as the one-way hash function used in DSS.

Given (M^*, r^*, s^*), one can verify whether (r^*, s^*) is indeed a genuine signature of the certification authority on M^* by the following steps:

1. calculates $v \equiv (g^{h(M^*)/s^*} \cdot y_{ca}^{r^*/s^*} \bmod p) \pmod q$.
2. accepts (r^*, s^*) as a genuine signature of ca on M^* only if $v = r^*$.

3.2 Making Use of a Broadcast Channel

Typically a mobile network uses a broadcast channel to continuously propagate from a base station to mobile users various types of control information such as synchronization parameters, available services, network time data, base station ID etc. The authentication protocol to be proposed in the following uses part of the capacity of the broadcast channel for a base station to propagate to mobile users the certificate associated with its public key. For simplicity, we assume that the certificate takes the form of

$$cert_{ca,b} = (b, y_b, sig_{ca,b})$$

where y_b is the base station's public key, and $sig_{ca,b}$ is the certification authority's digital signature on (b, y_b) created using the DSS scheme. (More information on the format of a certificate can be found in (Chokhani 1994) and (ITU 1993).)

To keep himself abreast of the various types of network information such as synchronization data, types of services, current network time, and the public key and certificate of a base station, a mobile user (through his mobile terminal) continuously monitors the broadcast channel. In doing so, it will be able to check whether $sig_{ca,b}$ is indeed a genuine signature of the certification authority ca on the base station's public information b and y_b, and hence to authenticate the base station "at the background". For this reason, we say that this process contributes 0.5 move to the protocol.

3.3 Key Distribution and Authentication of a Mobile User

As discussed above, a mobile user can authenticate a base station "at the background". In particular, this process can be completed immediately after his mobile terminal is switched on, or he roams into a new cell, without being noticed by the mobile user.

Now we assume that the mobile user m is in the cell covered by a base station b, has successfully authenticated the base station "at the background", and wishes to initiate a communication session. The mobile user m sends two data items c_1 and c_2 to the base station. Here c_1 and c_2 are constructed in the following way:

$$\begin{aligned} c_1 &\equiv g^x \pmod p \\ c_2 &= G(y_b^x \bmod p) \oplus (K, T, cert_{ca,m}, tag) \end{aligned}$$

where $tag = h(K, T, cert_{ca,m}, y_b^{x_m + x} \bmod p)$.

The meanings of other symbols used in c_1 and c_2 are as follows: x is a random number from $[1, p-1]$, $\oplus$ denote bit-wise exclusive-or, K is a random session key, both chosen by the mobile user m, x_m is the secret key, $cert_{ca,m} = (m, y_m, sig_{ca,m})$ is the certificate and y_m is the public key of m, while y_b is the public key of the base station, T is the current network time stamp taken from the base station's broadcast channel, G is a cryptographically strong pseudo-random number generator, and finally h is a

one-way hashing function such as SHS or HAVAL (Zheng, Pieprzyk & Seberry 1993). As is the case for $cert_{ca,b}$, more information may be included in the certificate $cert_{ca,m}$.

Note that the involvement of T, the current network time stamp taken from the base station's broadcast channel, is to ensure the freshness of the message. Also note that the main ideas behind the formation of (c_1, c_2) are from (Zheng & Seberry 1993), where three practical public key cryptosystems have been designed to resist against chosen ciphertext attacks.

Upon receiving (c_1^*, c_2^*) which may differ from (c_1, c_2), the base station calculates $G((c_1^*)^{x_b} \bmod p) \oplus c_2^*$, and splits the result into four parts

$$K^*, T^*, cert_{ca,m}^*, tag^*$$

Here $cert_{ca,m}^*$ consists of $(m^*, y_m^*, sig_{ca,m}^*)$.

The base station b then verifies the certificate $cert_{ca,m}^*$ and also checks the freshness of T^*. It aborts the protocol if either $cert_{ca,m}^*$ is invalid or T^* deviates too far from the current network time. Otherwise, if both $cert_{ca,m}^*$ and T^* are OK, the base station performs the hashing operation

$$d = h(K^*, T^*, cert_{ca,m}^*, (y_m^* \cdot c_1)^{x_b} \bmod p)$$

The base station is convinced of the identity of the mobile user and accepts K^* as a valid common key shared with the mobile user only if $tag^* = d$.

The new proposal is summarized in the following:

A 1.5 Move Protocol

---

1. Mobile User $m \Longleftarrow$ Base Station b
   The base station b broadcasts to mobile users

   $$cert_{ca,b} = (b, y_b, sig_{ca,b})$$

   where y_b is the base station's public key, and $sig_{ca,b}$ is the certification authority's digital signature on (b, y_b).
   The mobile user m monitors the broadcast channel and verifies, "at the background", the authenticity of the certificate and hence of the base station.
2. Mobile User $m \Longrightarrow$ Base Station b
   When the mobile user m wishes to initiate a communication session, he sends to the base station two data items c_1 and c_2 constructed by

   $$\begin{aligned} c_1 &\equiv g^x \pmod p \\ c_2 &= G(y_b^x \bmod p) \oplus (K, T, cert_{ca,m}, tag) \end{aligned}$$

   where $tag = h(K, T, cert_{ca,m}, y_b^{x_m + x} \bmod p)$. The meanings of other symbols used in c_1 and c_2 are: x is a random number in $[1, p-1]$, K is a random session key, both chosen by the mobile user m, x_m is the secret key, $cert_{ca,m} = (m, y_m, sig_{ca,m})$ is the certificate and y_m is the public key of m, while y_b is the public key of the base station, T is the current network time stamp taken from the base station's broadcast channel, G is a cryptographically strong pseudo-random number generator, and finally h is a one-way hashing function.
   Upon receiving (c_1^*, c_2^*) which may differ from (c_1, c_2), the base station b calculates $w = G((c_1^*)^{x_b} \bmod p) \oplus c_2^*$, and then splits w into four parts

   $$K^*, T^*, cert_{ca,m}^*, tag^*$$

where

$$cert^*_{ca,m} = (m^*, y^*_m, sig^*_{ca,m})$$

The base station then verifies the certificate $cert^*_{ca,m}$ and also checks the freshness of T^*. It aborts the protocol if either $cert^*_{ca,m}$ is invalid or T^* deviates too far from the current network time. Otherwise, if both $cert^*_{ca,m}$ and T^* are OK, the base station checks whether

$$tag^* = h(K^*, T^*, cert^*_{ca,m}, (y^*_m \cdot c_1)^{x_b} \bmod p)$$

The base station is convinced of the identity of the mobile user and accepts K^* (which should be identical to K) as a valid common key shared with the mobile user only if the above equation is satisfied.

Once a shared session key is established between the mobile user and the base station, they can use the session key together with a private-key (block or stream) cryptosystem to conduct secure communications.

4 REMARKS

The following observations on the new proposal can be made:

1. Due to the participation of the mobile user's secret key x_m in the formation of (c_1, c_2), the chance for an attacker to make a valid pair (c^*_1, c^*_2) is negligibly small, even if the attacker has the full knowledge of y_m and $cert_{ca,m}$. Hence successful completion of the protocol guarantees that $K^* = K$, namely, the mobile user and the base station have an identical shared key. Consequently, unlike Beller-Chang-Yacobi protocol, there is no need to confirm the consistency of the keys through the exchange of known messages.
2. As (c_1, c_2) employs network time information, replay attacks, such as the one applicable to Beller-Chang-Yacobi protocol, can be effectively thwarted by limiting the valid life span of (c_1, c_2), say, to a fraction of a second.
3. The proposed new protocol consists of 1.5 moves of information: 0.5 move for the authentication of a base station by a mobile user, and a single move for the authentication of the mobile user by the base station and the establishment of a session key.
4. The protocol provides anonymity of the mobile user with respect to an onlooker: as all messages exchanged between a mobile user and a base station, including the identity of the mobile user which is part of his certificate $cert_{ca,m}$, are transported in their encrypted form, an outsider or onlooker cannot figure out which mobile user is communicating with the base station.
5. The protocol prevents the impersonation of a mobile user by a fraudulent base station: the messages sent from a mobile user to a base station contains enough information for the base station to authenticate the mobile user, but not enough for a fraudulent base station to masquerade the mobile user. In fact the only entity who can create a correct pair of (c_1, c_2) is the mobile user who knows his secrete key x_m. The reader is directed to (Zheng & Seberry 1993) for more information on the unforgability of (c_1, c_2).
6. The creation of (c_1, c_2) can be partially pre-computed before the mobile user wishes to start a communication session. This further shortens the time required to establish a session.
7. The protocol can also be applied to networks and distributed computing where broadcast channels present.

Currently we are in the process of conducting a detailed analysis of the proposed new protocol, covering time complexity of the protocol, strategies for pre-computation by a mobile user, roaming, the procedure for a visited base station to contact the "home network" of a mobile user and other issues.

REFERENCES

Asokan, N. (1994), Anonymity in a mobile computing environment, *in* 'Proceedings of 1994 IEEE Workshop on Mobile Computing Systems and Applications'.

Aziz, A. & Diffie, W. (1994), 'Privacy and authentication for wireless local area networks', *IEEE Personal Communications* **1**(1), 25–31.

Beller, M., Chang, L.-F. & Yacobi, Y. (1993), 'Privacy and authentication on a portable communications system', *IEEE Journal on Selected Areas in Communications* **11**(6), 821–829.

Brown, D. (1995), 'Techniques for privacy and authentication in personal communications systems', *IEEE Personal Communications* **2**(4), 6–10.

Chokhani, S. (1994), 'Toward a national public key infrastructure', *IEEE Communications Magazine* pp. 70–74.

Diffie, W. & Hellman, M. (1976), 'New directions in cryptography', *IEEE Transactions on Information Theory* **IT-22**(6), 472–492.

Frankel, Y., Herzberg, A., Karger, P., Krawczyk, H., Kunzinger, C. & Yung, M. (1995), 'Security issues in a CDPD wireless network', *IEEE Personal Communications* **2**(4), 16–27.

Herzberg, A., Krawczyk, H. & Tsudik, G. (1994), On travelling *incognito*, *in* 'Proceedings of 1994 IEEE Workshop on Mobile Computing Systems and Applications'.

ITU (1993), Information technology — open systems interconnection — the directory: Authentication framework, Recommendation X.509, International Telecommunications Union.

Lai, X. (1992), *On the Design and Security of Block Ciphers*, ETH Series in Information Processing, Hartung-Gorre Verlag Konstanz, Zürich.

Molva, R., Samfat, D. & Tsudik, G. (1994), 'Authentication of mobile users', *IEEE Network* .

National Bureau of Standards (1977), Data encryption standard, Federal Information Processing Standards Publication FIPS PUB 46, U.S. Department of Commerce.

National Institute of Standards and Technology (1994), Digital signature standard (DSS), Federal Information Processing Standards Publication FIPS PUB 186, U.S. Department of Commerce.

National Institute of Standards and Technology (1995), Secure hash standard, Federal Information Processing Standards Publication FIPS PUB 180-1, U.S. Department of Commerce.

Rahnema, M. (1993), 'Overview of the GSM system and protocol architecture', *IEEE Communications Magazine* pp. 92–100.

Samfat, D., Molva, R. & Asokan, N. (1995), Untraceability in mobile networks, *in* 'Proceedings of MobiCom'95'.

Schnorr, C. P. (1991), 'Efficient signature generation by smart cards', *Journal of Cryptology* **4**(3), 161–174.

Wilkes, J. (1995), 'Privacy and authentication needs of PCS', *IEEE Personal Communications* **2**(4), 11–15.

Zheng, Y. (1996), 'The SPEED cipher'. Submitted for publication.

Zheng, Y., Pieprzyk, J. & Seberry, J. (1993), HAVAL — a one-way hashing algorithm with varialbe length of output, *in* J. Seberry & Y. Zheng, eds, 'Advances in Cryptology — AUSCRYPT'92', Vol. 718 of *Lecture Notes in Computer Science*, Springer-Verlag, Berlin, New York, Tokyo, pp. 83–104.

Zheng, Y. & Seberry, J. (1993), 'Immunizing public key cryptosystems against chosen ciphertext attacks', *IEEE Journal on Selected Areas in Communications* **11**(5), 715–724.