行政院國家科學委員會補助專題研究計畫成果報告

行動通訊系統中的安全身份驗證協定

A Secure Authentication Protocol in Mobile Communication System

計畫類別：√個別型計畫    □整合型計畫
計畫編號：NSC  89-2213-E-009-004
執行期間：88年8月1日至89年7月31日

計畫主持人：謝續平教授

本成果報告包括以下應繳交之附件：

□赴國外出差或研習心得報告一份
□赴大陸地區出差或研習心得報告一份
□出席國際學術會議心得報告及發表之論文各一份
□國際合作研究計畫國外研究報告書一份

執行單位：國立交通大學資訊工程學系

中華民國 89年 10月 1日
A Secure Authentication Protocol in Mobile Communication System

Abstract

Due to the mobility of new generation mobile computing networks, more and more applications are introduced for commerce, entertainment, personal assistant services, and so on. Security is the one of critical factors affecting the quality of mobile services. In this project, we consider three critical security issues and propose the solutions. First, we discuss the security problems of establishing communication channels under different mobile environment restrictions. For the inter-domain on-line roaming environment, we propose a chain authentication scheme. For the off-line roaming environment, we propose an IC card-based billing scheme for credit card phone services. We also propose a secure message exchange protocol for the “weak connection” environment. In this protocol, every message itself provides authentication, confidentiality, and integrity of message data. Finally, the piracy problem of some reusable service contents, such as mobile codes is considered. We herein design a software authorization and protection model to protect valuable service contents (e.g. mobile codes) from being copied or distributed by unauthorized users.

Keywords: Mobile computing, security, authentication, software authorization, software piracy, mobile code, and statistical database.

1. Authenticating Mobile Users in Inter-domain On-line Roaming Networks

The inter-domain on-line roaming environment is the trend of network infrastructures of future mobile computing services. And authentication for communication entities is the base of security functions at the access phase of mobile services. However, the conventional schemes either do not authenticate all communication entities or suffer from heavy overhead on networks. Therefore, these schemes are impractical for global mobile networks. The scheme proposed herein authenticates all communication entities and guarantees the confidentiality of all data transmissions. In addition, our scheme is suitable for multiple service providers within a geographical area, and reduce the connection overhead and satisfy the security requirements in an enormous and heterogeneous network.

1.1 Chain Authentication

The chain implies that all domains, visited by MS, constitute a roaming path, which is a virtual trusted chain. And the chain originates HLR and ends at the VLRn that he is currently visiting. A trusts B only if A can successfully authenticate B by a pre-defined protocol. Therefore, each entity in the
trusted chain must authenticate the neighbors and trust them. Figure 1. depicts the authentication history for MS. When MS lies in the domain of HLR, the authentication of both MS and HLR is trivial because HLR knows its subscriber MS. If MS roams to VLR1, some authentication procedure must be applied to establish mutual trust. Since VLR1 is strange to MS before registration of MS, VLR1 must query HLR to authenticate MS. Prior to authentication, however, HLR and VLR1 should authenticate each other. Upon establishing authentication, HLR can authenticate MS for VLR1, and relay MS’s security related information to VLR1. VLR1 and MS must then authenticate each other. Afterwards, if they can trust each other, VLR1 can be included in the MS’s trusted chain. These steps are repeatedly executed as MS travels until MS roams to VLRi, the trusted chain being from HLR to VLRi-1; i.e., they all trust MS, and vice versa.

![Figure 1. The chain authentication protocol](image)

Since the proposed protocol uses only VLRi-1 to establish authentication between VLRi and MS, we use VLRo and VLRn to denote the old domain VLRi-1 and the new visited domain VLRi, respectively (Figure 1). If some authentication procedures can be applied to guarantee that (1) VLRo trusts the authentication request claimed by MS, (2) VLRn trusts the response of VLRo, and (3) MS trusts the authentication result issued from VLRo, who is trusted; then, VLRn and MS can trust each other. Therefore, we define a basic rule for the authentication in VLRn using VLRo when MS moves from the trusted VLRo to the new visited domain VLRn. The rule of the chain authentication protocol is as follows:

given

- MS and VLRo trust each other,

if

1) VLRo and VLRn trust each other,
2) VLRn proves to VLRo that MS has arrived

in the new domain, and

3) VLRn proves to MS that VLRo trusts and authorizes VLRn.

then,

MS and VLRn trust each other.

1.2 Advantages of the Chain Authentication Protocol

The chain authentication protocol has the following merits:

- No assistance from HLR
- Low overhead
- Subscriber identity confidentiality
- Communication confidentiality
- Authenticating overall participant communication entities
- Consideration of multiple service providers in a local area
- Domain separation
- Session key confidentiality
- Low cost for preventing replay attacks

2. A Credit Card-based Billing Scheme for Inter-domain Off-line Roaming Environment

Another important mobile network is the inter-domain off-line roaming environment. A well-known and existing example is the credit card-based commercial networks. The credit card is the MS, the issuer (bank) of the card is the home service provider, and each store is a visited service provider. Whenever the user uses the credit card, a new service session is created again and initiated from the access phase. During the service session, the mobile user does/can not roam to another domain, that is, the handover process is unnecessary in this environment.

Although Visa and MasterCard have jointly developed the Secure Electronic Transaction (SET) protocol as a secure payment method for card transactions over open networks, it still has drawbacks and cannot fit in some applications, such as telephone systems. In the project, we propose a credit card-based payment scheme, which can securely authenticate cardholders without exposing their secret information on networks. And the payment scheme supports the capability of non-repudiation. Thus the mobile user can not deny the bills generated by service providers. The scheme also supports anonymity of cardholders, that is, the service provider (VLR) does not know who requires the service and the credit card company does not know what service is demanded by the user.
2.1 The proposed credit card-based billing scheme

The credit card-based billing scheme is shown as Figure 2.

Step 1: Alice inserts her credit card A to the card reader of T1. First, Alice dials her personal identification number (PIN) to enable the card. If PIN is correct, A will get the identity of the telephone company N from T1 and generate the message $X_{11} = \{ID_A, AT, N\}P_C$. Then A will give T1 the following data: the identity of CCAC C, its authentication token AT, and the message $X_{11}$. Finally, Alice dials the callee’s telephone number $T_2$, and T1 sends the message (T1, $T_2$, N, C, AT, $X_{11}$) to N.

Step 2: When the telephone network system N receives the message, it will verify if C is a legal and contracted credit card company that supports the service of credit card phone. If yes, it will keep the data T1, $T_2$, C, and AT in the database itself. Then N generates a timestamp $t_0$ that denotes the startup time for billing. And the message $X_{31} = (N, t_0, AT, X_{11})$, where $X_{11}$ was received from T1, is sent to C.

Step 3: The credit card authentication center C uses its secret key $S_C$ to decrypt the message $X_{11}$. Then C knows the caller’s credit card number $ID_A$, the authentication token AT, and the telephone company N. Since only C and A know the credit card number $ID_A$ and the current token AT, C can distinguish whether the message is new and generated for a phone call by A. If the verification is successful, the message $X_{11}$ is not masqueraded. N hence is the telephone company which is chosen by the caller to deliver the call. After the authentication, C makes sure whether $t_0$ is within the valid interval or not. If $t_0$ is much larger or smaller than the local clock in C, the message may be modified by a hostile or masqueraded by N itself. C thus discards the message and denies the request for the phone call. Otherwise, C generates a virtual phone card $VPC = \{C, AT, N, t_0\}S_C$ for this phone call, and randomly selects a new authentication token $AT_n$ for A. To secretly transmit $AT_n$ to A, C uses A’s public key to generate the message $X_{31} = \{ID_A, AT_n, VPC\}P_A$. Then C sends N the message $X_{32} = \{P_A, VPC, X_{31}\}S_C$. Note that A’s public key $P_A$ is encapsulated in $X_{32}$. That is because N does not know A’s credit card number and its public key.

Step 4: After receiving $X_{32}$, N uses C’s public key to verify the message: $P_C$ is used to decrypt $X_{32}$ to get $P_A$, $VPC$, and $X_{31}$. $P_C$ is used again to decrypt $VPC$ to get four numbers: $C'$, $AT'$, $N'$, and $t'$. If 1) $C'$ and $N'$ are respectively equal to $C$ and $N$, 2) $AT'$ is equal to $AT$ received from A in Step 2, and 3) $t'$ is equal to the timestamp $t_0$ that A itself sent to C in Step 2, $VPC$ is a legal virtual phone card. That is, C has authenticated the caller as a legal credit card user and permitted her to make the credit card phone. Since $VPC$ is trusted, $P_A$ is also trustworthy. Then N generates a new timestamp $t_1$ that indicates the expiration time of the first conversation interval. And N sends T1 the message $X_{41}$ that contains the startup time $t_0$, the expiration time $t_1$, and the message $X_{31}$.

Step 5: When T1 receives $X_{41}$, it forwards the message to the IC card A. A uses its secret key to decrypt the ciphertext $X_{31}$ and adopts the same process mentioned in Step 4 to verify whether $VPC$ is legal. If $VPC$ is legal, A generates a virtual coin $VC_1$ as the first evidence of the following conversation interval, where $VC_1 = \{VPC, t_1\}S_A$. While N receives $VC_1$, it should use $A$’s public key $P_A$, $VPC$, and $t_1$ to verify whether the evidence is legal. If yes, it establishes the communication channel between T1 and T2. The virtual coin should be collected to charge the credit card company the cost of calls in the future.

Step 6: In advance, the telephone network system N needs to negotiate a reconfirmation interval with the credit card authentication center C by a contract. If the conversation between T1 and T2 is longer than the interval time, N should issues a reconfirmation signal to T1 periodically. The
reconfirmation message should contain the expiration time of the next interval \( (t_2 \text{ in Figure } 2) \).

**Step 7:** When \( T_1 \) receives the reconfirmation message, it forwards the message to \( A \). Then \( A \) will generate a new virtual coin \( VC_2 = \{VC_1, t_2\}S_A \) and send it to \( N \) as the next conversation evidence. \( N \) also needs to verify whether the evidence is legal with \( A \)’s public key \( P_A, t_2 \), and the old evidence \( VC_1 \). If yes, the channel between \( T_1 \) and \( T_2 \) is kept for another time interval. And the evidence \( VC_2 \) is saved.

Suppose \( VC_1 = \{VC_{1,i}, t_j\}S_A \) is the last evidence. The telephone company \( N \) sends \( C \) the message \( (N, C, AT, t_0, t_j, VPC, VC_j) \) as a bill for the charge of this call. When \( C \) gets this bill, the token \( AT \) and the startup time \( t_0 \) are used as indices to search the corresponding credit card number \( ID_A \) and its public key \( P_A \). Then \( C \) verifies the virtual coin:

1) Decrypt \( VPC \). Check if the content is equal to \((C, AT, N, t_0) \) or not.
2) Repeatedly decrypt the virtual coin \( VC_j \) for \( j \) times to get the result \((VPC', t_j) \). Check if \( VPC' \) is equal to \( VPC \) or not.
3) Check if the time list \((t_0, t_1, \ldots, t_j) \) computed from the above step is valid (the sequence should be monotonically increased and the interval should be equal to the predefined value).

If it is correct, \( C \) has to pay this call. The charge of this call will appear on the caller Alice’s monthly credit card statements.

### 2.2 Protocol Analysis

Since in telephone systems, we cannot guarantee the caller is in a secure environment. Thus, it is very important to protect callers’ privacy and guarantee that the telephone company can charge the money for the phone calls it served. The proposed scheme can not only support the authentication of callers to guarantee they will pay for these calls via their credit card accounts, but also has four important features:

1) **Confidentiality.** In the protocol, all sensitive data is encrypted with the receiver’s public key, and only the desired receiver can decrypt and share the information. Consequently, we can make sure that the credit card number \( ID_A \) is never disclosed and shared only by the credit card \( A \) and CCAC \( C \). With this shared secret, nobody can forge \( X_{1i} \), and \( C \) can verify if the request is issued by \( A \). Note that the new token \( AT_n \) is also kept secret in Step 3 and 4. If the token is transmitted in the plaintext form, it is possible to be hostilely modified. Both values of \( AT_n \) in \( C \) and \( A \) will be different, so the authentication for the next phone call will fail.

2) **Anonymity.** Since \( A \) and \( C \) do not expose the credit card number \( ID_A \) in messages, the telephone company does not know who make the phone call. In addition, \( T_1 \) and \( T_2 \) are known only by \( N \), the credit card company cannot use the information to trace where its customer made the call and know who is the callee. The personal privacy is protected. Of course, if there exists an argument about the bill, \( C \) and \( N \) can cooperate to disclose the details of the phone calls, such as \( ID_A, T_1, T_2, \) and \( t_0 \).

3) **Efficient reconfirmation.** The proposed protocol adopts a periodic payment scheme to resolve the problem that the caller may repudiate he/she made a call. A phone call consists of many conversation intervals. At the ending of an interval, \( N \) must reconfirm whether the caller will continue the conversation. If yes, the caller must give \( N \) a ‘coin’ to buy the next conversation interval. As the reconfirmation phase needs only two messages shown in Section 4.3, the time spending for the reconfirmation is very short.

4) **Non-repudiation.** Since the virtual phone card \( VPC \) and the corresponding virtual coins \( VC \) are encrypted with the secret keys of \( C \) and \( A \), respectively, the caller (and CCAC) can not repudiate that the call service has been supported by the telephone company.

### 3. A Secure Single-Message Exchange Protocol for Weak Connection Environments

The third type of mobile computing networks is the weak connection environment. In the general strong connection environment, the MS and the service provider can negotiate a session key at the access phase and then adopt the key to encrypt service contents at the service phase. But, in the weak connection environment, the session key scheme is infeasible because that the authentication and the session key generation/exchange must be executed again whenever the connection is
re-established. The overhead spent by the access phase of re-connection will be proportional to the unreliability of radio paths. Herein, we propose a secure message exchange protocol, which is originally designed for supporting the security function of non-session-oriented applications. The proposed scheme needs only one message to authenticate the validity of messages. And, the confidentiality and integrity of service contents carried within the same messages is guaranteed by a message-oriented encryption key. Therefore, the service provider and MS do not need extra messages to establish a secure channel for transmitting the following service contents. The proposed message exchange protocol is especially suitable for the weak connection environment, e.g. wireless communication systems. In addition, it is also suitable for other non-session-oriented application, such as the security management of telecommunication management network (TMN), and Internet management systems.

The message structure of SSMEP looks similar to X.509. SSMEP includes a new verification scheme to verify the freshness of messages, namely, the synchronized-nonce (SN) scheme. The SN scheme is used to replace the timestamp of X.509 (see the next section for the details). The basic message of SSMEP is presented in Figure 3. The message consists of three segments: certificate, authentication, and data segments.

![Figure 3. The secure single-message exchange protocol (SSMEP)](image)

4. A Software Authorization and Protection Model for Mobile Codes

Encrypting service contents only can prevent from eavesdropping or modifying of hostile outsiders. If the content is software, other security threats may occur. For example, after a legal mobile user downloads demanded programs, he can arbitrarily copy and distribute these codes. That is the piracy problem. In the chapter, we propose a new software authorization and protection model. In this model, a software consists of multiple mobile codes, which can be dynamically downloaded from service providers. Some critical codes are executed on trusted proxies and the others are executed on the MS. The execution of a software is conducted by cooperation of the MS and the proxies containing codes. There are two important features in the proposed model.

- Preventing from software piracy: The unauthorized user holding part of mobile codes of the software will not be able to use the software without the help of these proxies.
- Reducing computation load of MSs: The powerful proxies can afford most of computation to enhance the performance of mobile computing services.

4.1 The Proposed Authorization and Protection Model

In mobile code systems, a program (software) is composed of a number of mobile codes. A code can be downloaded dynamically from the remote machine and executed on the local machine, and a job can be processed by the cooperation of these mobile codes. In the section, an authorization and protection model is proposed to enhance the security and protection of mobile codes by delegating some critical execution services to one or more trusted and protected proxies.

The execution of a program can be considered to include three parts: incoming messages, transformation processes, and outgoing messages. A mobile code participates in the transformation process for a message if the code sends or receives the message. If some critical codes are removed from a program, execution of the program cannot proceed.

With the RMI (Remote Method Invocation) technology for Java language that enables cooperating of computers on the network, we proposed a model that protects software with the help of trusted, protected computational proxy servers, instead of tamper-resistant hardware devices installed in the user’s environment. In this model, mobile codes of the software are partitioned into two types, general and privileged codes. The users can acquire only general codes. And the privileged codes are forced to be executed in a protected environment, that is, the trusted computation proxy.

The trusted computational proxy provides computation services for privileged codes, as shown in Figure 4. Only a trusted proxy has the capability to get privileged codes and execute them. The proxy executes the mobile codes on behalf of an
authorized user and returns the result to the user (MS). In this way, an unauthorized user cannot acquire the results of privileged codes, and therefore benefit little from the software. A program may delegate all privileged codes to many proxies, and each proxy executes only a subset of the privileged codes. Thus, a compromised proxy will not leak all privileged codes. In the proposed model, mobile codes to be downloaded are encrypted by code keys, and the code keys for each mobile code are different. These keys are only available to trusted proxies or authorized users. The model consists of six major components:

- **Software Vendor**: The company who developed the software.
- **Certification Authority**: The party who signs and issues the certificate containing the user’s public key.
- **Software Authentication Center**: An accredited organization that authenticates the software developed by software vendors, and signing legitimate parts of the software.
- **Service Provider**: The server who stores mobile codes provided by software vendors. When an MS wants to execute a mobile code, it first connects to the service provider and downloads the code from the server.
- **Trusted Computational Proxy**: The server that provides computational services of privileged codes for users (MS). These proxies need a service provider to support a communication channel to the MS.
- **MS**: The user’s local host for executing the software. For simplification, we also use ‘MS’ to denote the user in the following discussion.

5. Conclusions

In this report, we address the security problems of mobile computing services. First, the features of mobile computing networks are investigated. According to the environment, the possible security threats and the critical security issues are discussed. We also analyse GSM to illustrate the security problems occurring in modern mobile communication systems. Then, we propose our solutions to protect mobile computing services.

5.1 Comments on Protecting Mobile Communications

In the project, we consider three main types of mobile communication environment:
- the inter-domain on-line roaming networks,
- the inter-domain off-line roaming networks,
- the weak connection networks.

We propose three authentication protocols to protect communications between service providers and MSs under these environment restrictions.

The first environment is the trend of current mobile telecommunication systems. The proposed protocol, called chain authentication protocol, can efficiently authenticate all communication entities with the assistance of old visited service provider. In addition, our scheme can be adopted in networks that have multiple service providers within a geographical area. It is thus practical for an enormous and heterogeneous network.

The second environment is similar to the model of credit card-based mobile computing networks. Although the SET standard has been proposed to provide a secure environment for card transactions, the SET is not suitable for all credit card-based applications, such as the credit card-based phone services. In the project, we propose a credit card-based billing scheme for the inter-domain off-line roaming networks. The scheme does not only support the capability of authenticating entities, but also guarantee the anonymity of callers. Furthermore, we propose the concept of virtual phone cards to support a secure, non-repudiated, and efficient payment environment, which makes the communication load much lower than the SET standard.

The third environment is designed for wireless networks. In such environment, the connection has lower bandwidth and higher error rate, and is subject to frequent disconnection. We propose a secure single-message exchange protocol for
efficiently and securely transmitting data. The data and a segment of security-related information are carried within a single message. The information provides the features of authenticating and enciphering the data. The proposed protocol adopts a synchronized nonce scheme, instead of a timestamp, to prevent from replay attacks. Thus, our scheme needs not synchronize the clocks of MSs and service providers.

5.2 Comments on Protecting Contents in Service Providers

In mobile computing networks, not all security threats occur on communications between MSs and service providers. Security mechanisms are also needed to protect contents in service providers. The contents may be mobile service contents (e.g. software programs) and management data (e.g. subscribers’ accounting data). The project herein introduces a security mechanism to protect mobile code-based software.

We propose a software authorization and protection model for the new generation software, mobile code systems. In this model, a software consists of multiple mobile codes, and only independent mobile codes are executed on the MS. The other codes are executed on trusted proxies. Each mobile code produced by the software vendor has a publication license which is issued by a trusted third party. Before using software, the MS must request an execution license for demanded mobile codes. We do not only introduce the system components of the model for issuing the licenses, but also discuss how to partition software based on the considerations of computation/communication load and security. Thus, the proposed model can prevent from software piracy and reduce computation load of MSs. In addition, by verifying the mobile code with its publication license, the MS and proxy can detect whether the code has been unauthorizably modified or not. This feature can reduce the risk of attacks from Trojan Horse or viruses.

6. Reference


