Chain Authentication in Mobile Communication Systems

Chern-Tang Lin, Shiuh-Pyng Shieh
Department of Computer Science and Information Engineering
National Chiao Tung University
Taiwan, ROC 30050
Tel: +886-35-731876
Fax: +886-35-724176
Email: ssp@csie.nctu.edu.tw
URL: http://dsns.csie.nctu.edu.tw

Abstract

Digital mobile telecommunication systems have become a future trend in personal communications services (PCS) networks. To satisfy the demand for high quality services, security functions, including the authentications of communication parties and the confidentiality of communication channels, must be embedded into mobile communication systems. This work presents an authentication scheme to support the security functions. The proposed protocol significantly improves the performance of authentications and ensures the security of mobile communications in a large-scale communication network with multiple service providers.

1. Introduction

As communication and computer technologies develop, users desire that all communication services, including audio, video, image, and data, be available anytime and everywhere to everybody. The Federal Communications Commission (FCC), U. S. A., defines personal communications services as "a family of mobile or portable radio communication services which could provide services to individuals and businesses and could be integrated with a variety of competing networks. The primary focus of PCS is to meet the communications requirements of people on the move." Such an expectation implies that PCS must, at least, possess the features of mobility, digitization, and data variability [1]. Therefore, digital mobile telecommunication systems have become a future trend PCS networks.

Unlike in the conventional computer (or telephone) networks, in a mobile communication system, an end user subscribing in his home domain may request services after or during moving from one domain to another. The service provider of a visited domain, who has no information about the user, should immediately identify the user and provide authorized services, and then inform the service provider of the home domain to accumulate user's accounting data for demanded services. Intuitively, the users' mobility increases the risk of masquerading legal users; the radio channels are also more vulnerable to eavesdroppers. Thus, while a user arrives in a new visited domain, the preparation, called the registration, must, at least, contain the authentication and the generation of the session key to guarantee the security of services against impersonating and eavesdropping by evil intruders. In mobile communication systems, the confidentiality of mobile users' identities is also necessary to protect against tracing users' location by listening to the message exchanges on the radio channel. However, with the restriction of security requirements, the registration scheme must still sustain efficiency to provide excellent services.

Many digital mobile telecommunication systems, e.g. the Global Systems for Mobile Telecommunications (GSM) [2] in several countries, Cellular Digital Packet Data (CDPD) [3] in U.S.A., and the Cellular IS-41 Standard in North American [4], provide simple authentication and ciphering schemes to prevent security threats, e.g. eavesdropping and unauthorized access. In practice, however, those systems authenticate only the mobile users who request services, not other communication parties, e.g. the service providers of visited/home domains. In addition, they guarantee the confidentiality of messages only between mobile users and the service providers of visited domains to prevent against eavesdropping on the radio channel, but not between the service providers of all domains [5]. Recently, IETF/IAB announced the mobile-IP specification that allows mobile computers to move freely between various domains of the Internet [6]. However, it still encounters the same problem: based on the specification, only the authentication between the mobile computer and his home agent is mandatory. To enhance the security of existing mobile networks, Molva et al. proposed an authentication scheme [7] that not

only certified each communication parties, including users and service providers, but also guaranteed the confidentiality of each communication channel.

Unfortunately, each registration request of all above practical systems and schemes, except GSM, must be transmitted back to the home domain to authenticate the mobile user. When the scope of the network is large, communication between the visited and the home domains is expensive and negatively affects the performance of the mobile communication system. Furthermore, if the user roams between different domains, the service provider is left with insufficient time to authenticate. Recently, attempts have been made to reduce the overheads of the mobile user registration, accelerate call connections and reduce network traffic [8,9]. For example, to enhance performance of the registration procedure in the Mobile-IP, the architecture of domains is hierarchically rearranged [10]. However, the authentication of communication parties remains unexplored.

This paper presents an authentication scheme, called *chain authentication*, which does not require assistance from the home domain while authenticating a mobile user. Like in GSM, when a mobile user roams to a new visited domain, the service provider of the old visited domain authenticates the user in the new visited domain. Thus, the proposed scheme combining the adaptive registration procedure (such as the methods mentioned in Ref [8,9]) is more efficient and particularly appropriate for large-scale networks since the old visited domain is generally closer to the new one than the home domain. And, unlike in GSM, the scheme proposed herein authenticates all communication parties and guarantees the confidentiality of all data transmissions. In addition, our scheme is suitable for an area with multiple service providers, and can effectively reduce the connection overhead and satisfy the security requirements in an enormous and heterogeneous network.

Section 2 briefly describes the well-known authentication protocols in modern communication systems, such as GSM, CDPD, IS-41, and Mobile-IP, and analyzes their security weaknesses. Molva's scheme is also described to compare the differences with our scheme. Next, section 3

and 4 present the problems to be resolved, the assumptions for our protocol, and the chain authentication scheme. Section 5 discusses the protocol analysis and compares the present model with traditional protocols. Concluding remarks are finally made.

2. Previously Published Authentication Schemes

We first define a generic environment for mobile communication systems to simplify our discussion. This environment is used herein to describe the existing and the present authentication schemes. The environment is based on the architecture of modern cellular mobile telecommunication systems [11], and uses the similar notations as GSM. The environment comprises three important parties:

- MS --- *mobile station*. It is a portable communication component with limited computation power to provide the security functions. MS also denotes the user of the mobile station throughout the article.
- HLR --- *home location register*. It is a database in a subscriber's home domain that contains the subscriber's/MS's management information, including authorized services and accounting data. Interchangeably, HLR may also denote the service provider of the home domain or the domain itself.
- VLR --- *visitor location register*. It is the database of the service provider in the visited domain, where MS is roaming. This database stores personal and temporary information to manage the visiting MSs. Thus, VLR denotes the service provider of the visited domain or the domain itself. Regarding MS, except the HLR subscribed by MS, other domains are VLRs. If MS roams from an old visited domain to a new one, VLR_n and VLR_o denote the new visited and the old domains, respectively.

Under such an environment, when MS arrives in a new domain VLR_n and seeks a service, VLR_n must identify MS and justify in real time whether MS is authorized to acquire this service. Since

VLR_n lacks this MS's information in advance, it must seek assistance from a third party. There are two candidates of the third party: HLR and VLR_o.

Seeking assistance of HLR is an intuitive solution to authenticate MS because the latter has subscribed in HLR. Many practical systems utilize through this mechanism, including IS-41 [4], CDPD [3], and Mobile-IP [6]. IS-41, Intermin Standard 41, was defined by Electronic Industries Association (EIA) and by Telecommunications Industries Association (TIA) for the mobility management of MSs who roam across cellular telecommunication systems, such as AMPS, IS-95, and so forth [27, 28]. As MS moves from VLR_o to VLR_n, VLR_n forwards the authentication request submitted by MS to MS's HLR (see Fig. 1). The request consists of the authentication result *AUTHR* and other security related information. The request is directly transmitted to HLR via VLR_n. If the verification of *AUTHR* is successful in HLR, HLR responds to VLR_n and provides security related information to VLR_n who establishes the private communication channel between MS and VLR_n. In IS-41, HLR, VLR_n, and the network between them are fully trusted. Thus, this protocol only verifies the validity of MS's request and guarantees the confidentiality between VLR_n and MS.

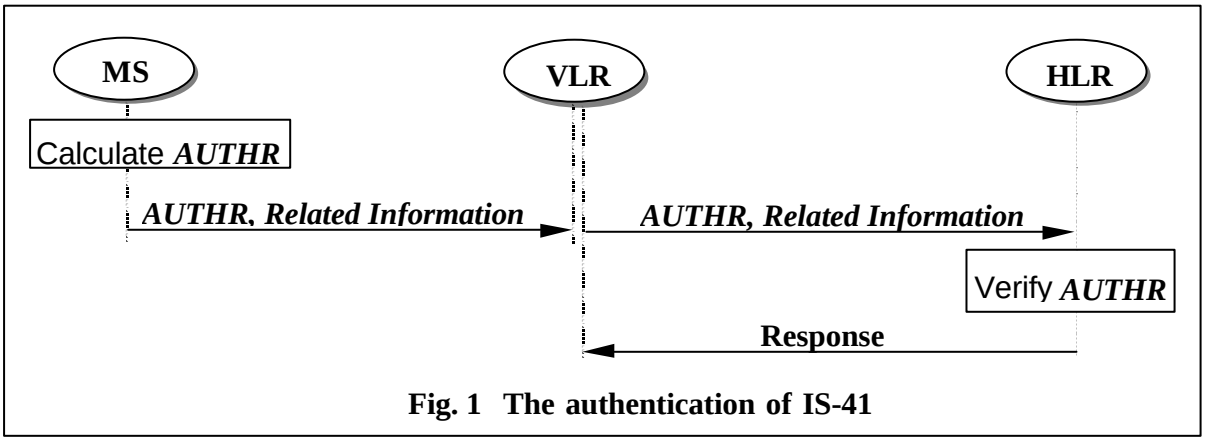

Fig. 1 The authentication of IS-41

The authentication protocol of CDPD resembles IS-41 except that MS and VLR_n must first determine their sharing secret key with the Diffie-Hellman key exchange protocol [12]. HLR, upon approving the authentication request, sends a new credential to MS, in the clear via VLR_n, for the next authentication. Obviously, CDPD assumes that the fixed/wired network is secure.

Like IS-41 and CDPD, Mobile-IP adopts the same registration mechanism: MS-VLRn-HLR-VLRn-MS. MS inserts an authenticator into the registration request, and HLR authenticates the request with the authenticator. (The authenticator is produced with the secret information, called the security association, which is shared between MS and HLR.) MS similarly verifies the registration reply from HLR by the HLR-MS authenticator. The authentications of HLR-VLR and VLR-MS occur if they can share some security associations in advance by other appropriate mechanism(s) [13]. However, establishing this job for each pair of HLR-VLR and VLR-MS is a difficult task. Hence, in the Mobile-IP specification, only the authentication between HLR and MS is mandatory.

The second candidate supporting MS's information to assist VLR_n in authenticating MS is VLR_o, where MS originates from. The authentication protocol proposed in GSM resembles this mechanism. GSM was developed by European Telecommunications Standard Institute (ETSI) [11]; it is the first digital cellular mobile telecommunication system providing security functions for its subscribers to guarantee the confidentiality of communications and avoid frauds. In GSM, each MS carries a unique and permanent subscriber identity, generated by HLR, and a temporary identity (*TID*), generated by VLR where MS is roaming in. Besides, MS and HLR share some security-related-information that is used in the authentication. When MS leaves his home domain and arrives in a new visited domain, VLR must query HLR for security-related-information to authenticate MS. If MS roams to another new domain VLR_n, the latter will demand security-related-information from VLR_o. Fig. 2 depicts the authentication procedure, which is embedded in the registration protocol. Although GSM provides better security functions in modern telecommunication systems, the system still has many weaknesses [5]. For instance, GSM only authenticates MS like IS-41 and CDPD. Furthermore, although VLR_n directly obtains security-related-information from VLR_o rather than from HLR to reduce the traffic of networks, the exchanged messages are by far numerous. The number of messages is large because the security-related-information transmitted from VLR_o to VLR_n should suffice to satisfy sequential

authentications. Besides, VLR_n should inform HLR to update MS's location; the overhead of MS's registration is still large.

Obviously, modern telecommunication systems, such as IS-41, CDPD, Mobile-IP, and GSM, fail to satisfy the requirements of high security assurance in the PCS. Future communication networks will be heterogeneous and integrate multiple service providers to support demanded communication services [25]. Therefore, providing confidentiality and authentication to each party in the communication networks, including subscribers and service providers, is necessary.

To resolve the security problems of mobile telecommunication systems, Molva et al. proposed an authentication protocol [8]. The protocol authenticates all the individuals, i.e., MS, VLR, and HLR, and protects all exchanged messages between the parties as ciphertext. No intruder can therefore impersonate the legal party in the communication network and gain access to unauthorized services, or eavesdrop on secret information in communication channels. Like IS-41, VLR_n in Molva's protocol directly requests HLR to authenticate MS. Fig. 3 presents the details of this procedure. The first message MS sends to VLR_n is *AUTHKur(..)*, which is an authentication token and is encrypted by a location-dependent key K_{ur}. The key K_{ur} is generated by MS and is used only in the authentication phase. Since VLR_n does not yet know K_{ur}, VLR_n

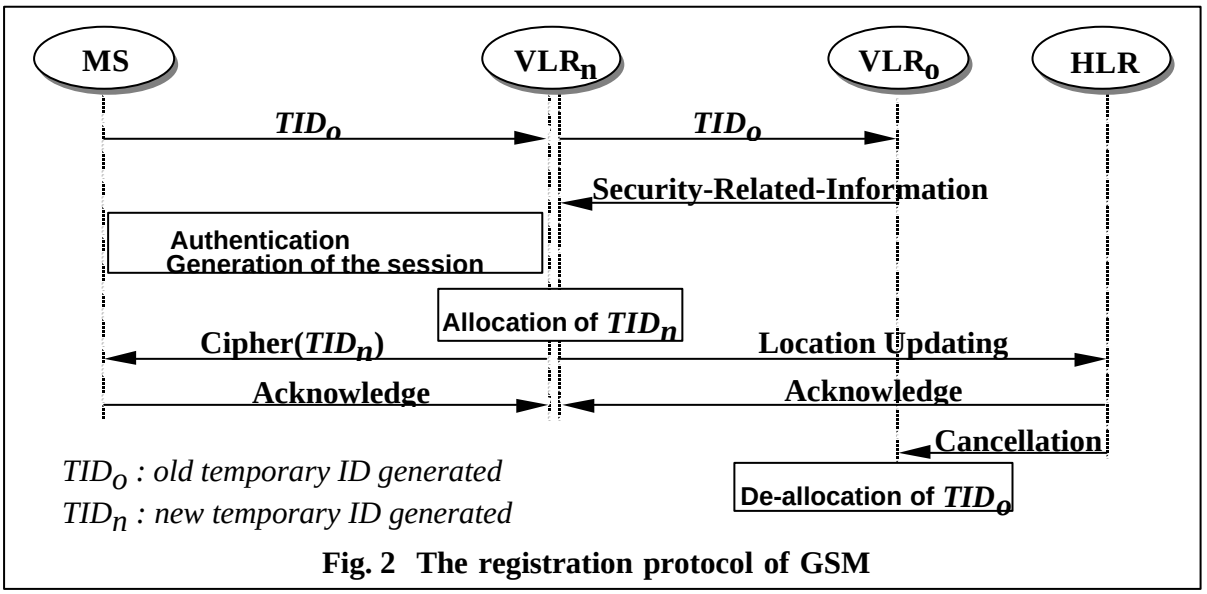

Fig. 2 The registration protocol of GSM

just recomputes a new authentication token with the undecrypted token N_r^2 and its certificate information and sends to HLR. Molva et al. assumed that HLR and VLR_n share a long-term key K_{rh} that is distributed by a secure key distribution procedure involving a mutually trusted third party. With the shared key K_{rh}, HLR authenticates VLR_n and MS by the token and, if successful, sends a ticket to VLR_n that contains the secret key K_{ur}. VLR_n can then verify the authentication token, $AUTH_{Kur}(\ldots)$, received from MS. Upon receiving the correct authentication token, VLR_n sends a ticket containing the session key K_S shared by MS and VLR_n for the subsequent communication to MS.

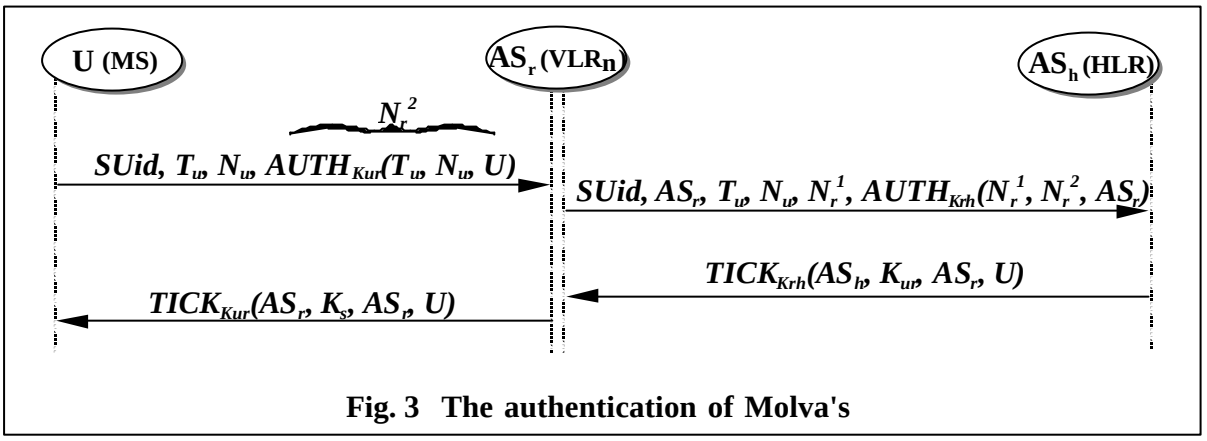

Fig. 3 The authentication of Molva's

Because future PCS will provide services in any city, such extension necessitates long connection between HLR and VLR_n. Thus, Molva's protocol, although superior to the previous four schemes in terms of security, suffers from the communication overhead between VLR_n and HLR. The overhead not only includes the message transmission for the long connection, but also the added complexity of establishing a secure key distributed system to guarantee the security between HLR and the remote VLR. To guarantee smooth services, the time taken for the registration (including the authentication) should be kept to minimum while MS roams into a new domain during service. Therefore, eliminating the traffic between HLR and VLR_n can improve the performance of the registration/authentication.

3. Problems and Assumptions

Many advanced schemes have recently been proposed to reduce the overhead by avoiding the assistance of HLR. Jain et al. proposed a caching strategy, incorporating the original location strategy, to improve the performance of finding the called user's location [8]. Perkins proposed an architecture of hierarchical VLRs for reducing the overhead of registrations in Mobile-IP [10]. In this scheme, the registration request does not need to be transmitted back to HLR, but must cross several domains to a targeted VLR that can authenticate MS. Nevertheless, this scheme also suffers from the disadvantage of only authenticating MS and HLR.

In 1995, Jain et al. proposed a pointer-forwarding scheme to reduce the registration cost [9]. To find the called MS's location and deliver the call to him, the conventional systems resort to record MS's current location at HLR whenever the user moves. However, Jain's registration scheme creates only a pointer from VLR_o to VLR_n when MS moves from VLR to another, instead of immediately updating his location record in HLR. Updating of MS's location in HLR is unnecessary until a call delivery for MS arrives. As Fig. 4 reveals, when MS is called, the new location pointer P is created by traveling from HLR to VLR_4 (dashed lines), and the original pointers are deleted. This scheme significantly benefits the registration when accessing HLR is rather expensive, e.g., for when HLR and VLR_n are distant or when the network topology is quite complicated. However, the authentication protocol was not addressed in their work. The authentication procedure must be completed before the location updating during the registration phase. Jain's registration scheme is incompatible with most conventional authentication protocols because they must connect back HLR to authenticate MSs. Herein, we present an authentication scheme that is especially appropriate for Jain's registration scheme. The overhead of the registration, including the authentication, is therefore significantly lower in our scheme.

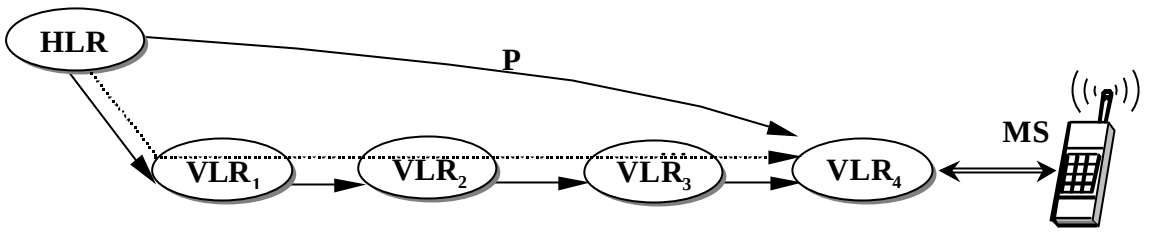

Fig. 4 The pointer-forwarding scheme for registration

In fact, the proposed chain authentication is suitable not only for the pointer-forwarding-based registration, but also for the conventional registrations, which record MS's current location at HLR. In the conventional schemes, the chain authentication will not reduce the network burden for registration but will shorten the response time for authentication. This feature is very useful for communication services, e.g., the voice services, which have to smoothly continue when MSs roam from a domain to another. (That is because the service can be immediately re-started after the interruption for authentication, and then the rest of the registration can be concurrently proceeded with the service.) Because the detail of the registration is out of the scope of this paper, we will not discuss the impact of different location updating schemes combined with the chain authentication.

The following assumptions about the PCS environment are used in this study.

- PCS network is a distrusted communications environment. Users can travel to anywhere in the PCS network that consists of many service providers. All communications are likely to be eavesdropped through wireless or wired channels. The messages can be destroyed to confuse services or be replayed to access unauthorized services. An intruder may try to impersonate every role, including that of MS, VLR, and HLR, to cheat others.
- The network topology is extremely large and complicated. Thus, VLR is distant HLR and the path contains many switches (or routers). Therefore, the communication back to HLR is expensive and may significantly degrade the performance of the authentication. In addition,

the covered area of the administrative domain of a service provider may overlap with other domains, and each service provider may not support all the services. That is, a local area may contain multiple service providers who support different services. Each service provider must continuously broadcast a beacon containing VLR-related information, such as the identity of VLR and supported services. By listening to the beacons, MS can detect the domains he arrives in, and select an appropriate VLR to connect with.

- Inter-domain authentication is available. In a multi-domain environment, all service providers must cooperate to support services for all subscribers under the control of the contract. Thus, authentications of participant service providers and the non-repudiation of their operations are deemed necessary. The public key cryptography will be the best candidate; many security standards and applications use the public key cryptography, e.g., CCITT X. 509 [14], ANSI X9.30, and Internet Privacy Enhance Mail [15-17]. This work assumes that VLR can access the public keys of other VLRs. We do not stipulate that the public key infrastructure established for the certificates and key management (Readers interested in this issue can refer to [18] for the details.). Chokhani [18] reported that the cost of obtaining other public keys relies on the distance between them. Since MSs generally travel from a domain to its neighbor, the distance between VLR_o and VLR_n is typically shorter than the distance between VLR_n and HLR. This assumption is therefore reasonable and practical.
- The administration of HLR/VLR is trusted. That is, the servers of HLR/VLR are protected from malicious intruders or renegade system-operators. This assumption may fail because security-management schemes are still imperfect now. However, schemes, such as access control and auditing, fall beyond the scope of the present paper. Readers interested in those schemes can refer to related articles [19, 20].

- The pointer-forwarding scheme for subscriber registration is applied. We emphasize the design of the authentication protocol. The next section describes the relation between the present protocol and the pointer-forwarding scheme.

4. Chain Authentication

The *chain* implies that, in the proposed protocol, all domains, visited by MS, constitute a virtual trusted chain, which originates HLR and ends at the VLR_n 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. Fig. 5 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 VLR_1, some authentication procedure must be applied to establish mutual trust. Since VLR_1 is strange to MS before registration of MS, VLR_1 must query HLR to authenticate MS. Prior to authentication, however, HLR and VLR_1 should authenticate each other. Upon establishing authentication, HLR can authenticate MS for VLR_1, and relay MS's security related information to VLR_1. VLR_1 and MS must then authenticate each other. Afterwards, if they can trust each other, VLR_1 can be included in the MS's trusted chain. These steps are repeatedly executed as MS travels until MS roams to VLR_i, the trusted chain being from HLR to VLR_{i-1}; i.e., they all trust MS, and vice versa.

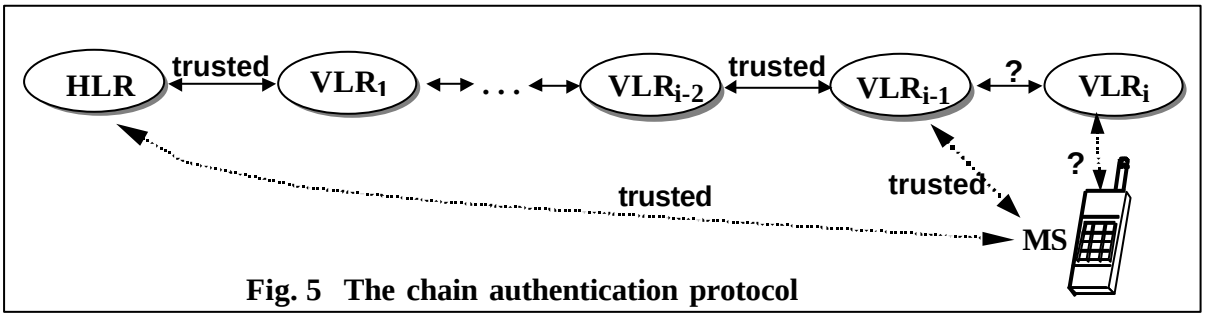

Fig. 5 The chain authentication protocol

Since the proposed protocol uses only VLR_{i-1} to establish authentication between VLR_i and MS, we use VLR_o and VLR_n to denote the old domain VLR_{i-1} and the new visited domain VLR_i,

respectively (Fig. 5). If some authentication procedures can be applied to guarantee that (1) VLR_o trusts the authentication request claimed by MS, (2) VLR_n trusts the response of VLR_o, and (3) MS trusts the authentication result issued from VLR_o, who is trusted; then, VLR_n and MS can trust each other. Therefore, we define a basic rule for the authentication in VLR_n using VLR_o when MS moves from the trusted VLR_o to the new visited domain VLR_n. The rule of the chain authentication protocol is as follows:

given

MS and VLR_o trust each other,

if

1) VLR_o and VLR_n trust each other,

2) VLR_n proves to VLR_o that MS has arrived in the new domain, and

3) VLR_n proves to MS that VLR_o trusts and authorizes VLR_n,

then,

MS and VLR_n trust each other.

The notations used in the chain authentication protocol are defined as follows. Notably, we append a subscript *o* to a notation to denote its relation to VLR_o, and a subscript *n* to show its relation to VLR_n.

- *IMSI* – international mobile subscriber identity. It is a unique and permanent MS identity, generated by HLR when MS subscribes the services. *IMSI* is confidential; only MS, HLR, and trusted VLRs are aware of this information.
- *TMSI* – temporary mobile subscriber identity, generated by VLR whenever MS arrives and completes the registration. To preserve the confidentiality of MS's identity, MS uses *TMSI,* rather than *IMSI,* to identify itself in the local domain. When MS leaves this VLR and registers in another domain, this identity is canceled. It is assumed that *TMSI* contains the information of VLR that generated it. *TMSI* consists of VLR's domain address and a

temporary sequence number to establish its identity. VLR_n can then gain the address of VLR_o while receiving $TMSI_O$ submitted by MS.

- *VID* – the unique and permanent identity (or address) of VLR. By listening to the beacon broadcasted from the service provider, MS can detect his current domain and get the domain-related information, including *VID*. If the local area is covered with the administrative domains of multiple service providers, MS can determine which service provider he wants to connect with by their beacons. VID_n denotes the identity of the new service provider (VLR_n) which MS finally chooses to connect.
- K, K^{-1} – the public/private key pair of VLR with the asymmetric cryptosystem [21]. The public/private key pair is applied to the digital signature and encryption of messages between VLRs/HLR. VLR generates a signature by encrypting data with the private key K^{-1}. Other VLRs can then authenticate the validity of data with the sender's public key K. In addition, sensitive data is encrypted with the receiver's K by the sender and transmitted in the fixed network to prevent from eavesdropping by a third party. Only the public keys of VLRs demanding communications are needed in advance to upgrade the authentication's performance. Since the VLRs are usually neighbours, obtaining and caching these keys should be economical.
- K_C – the session key shared by MS and VLR_o. It is based on the symmetric cryptosystem. However, the algorithm to be chosen relies on system requirements and is ignored in this work. We use K_C to encrypt/decrypt the data transmitted on the radio path to guarantee the confidentiality of communications. To distinguish from K_C, the session key shared by MS and VLR_n is denoted as K_C'.
- *AT* – the authentication token provided by VLR. It is a random number and is only known to MS and to the VLR who generates it. *AT* establishes mutual authentication between MS and VLR in the chain authentication protocol (for details, see section 5(A)).

- N – a nonce generated by VLR_n. VLR_n uses N to verify the freshness of the response message from VLR_o.
- *Seed* – a random number generated by MS, and used by MS and VLR_n to generate the new session key shared only by them.
- $\{Message\}K$ – a message encrypted by an encryption key K. The key may be the session key or, in the asymmetric cryptosystems, the public key K. If the encryption key is the private key K^{-1} of the message sender, this notation denotes the digital signature of *Message*.
- $f(INPUT_1, INPUT_2)$ – the result of an irreversible one-way function f with two inputs, $INPUT_1$ and $INPUT_2$. The output of the function can therefore not be forged without knowing the two inputs. Thus, the one-way function can safely authenticate the sender if only the sender and the receiver share the inputs.

The chain authentication protocol consists of three procedures: the subscription procedure, the initial authentication procedure, and the subsequent authentication procedure. The first procedure is initiated as MS subscribes to a new account of communication services in HLR. The roaming MS in a new domain invokes the second procedure to complete the registration in VLR_n. The last procedure is finally invoked if MS has registered and the authentications are demanded for subsequent services within the same domain.

The Subscription Procedure

For each new MS, HLR provides a unique and permanent identity *IMSI*, a temporary identity *TMSI*, a session key K_C, and an authentication token *AT*. The off-line method directly saves the information in the mobile station or subscriber's smart card, used in GSM, and evades security problems in this phase. MS can then directly announce itself by *TMSI* and communicate with HLR by K_C; no secret information is disclosed. Except *IMSI*, the three parameters *TMSI*, K_C, and

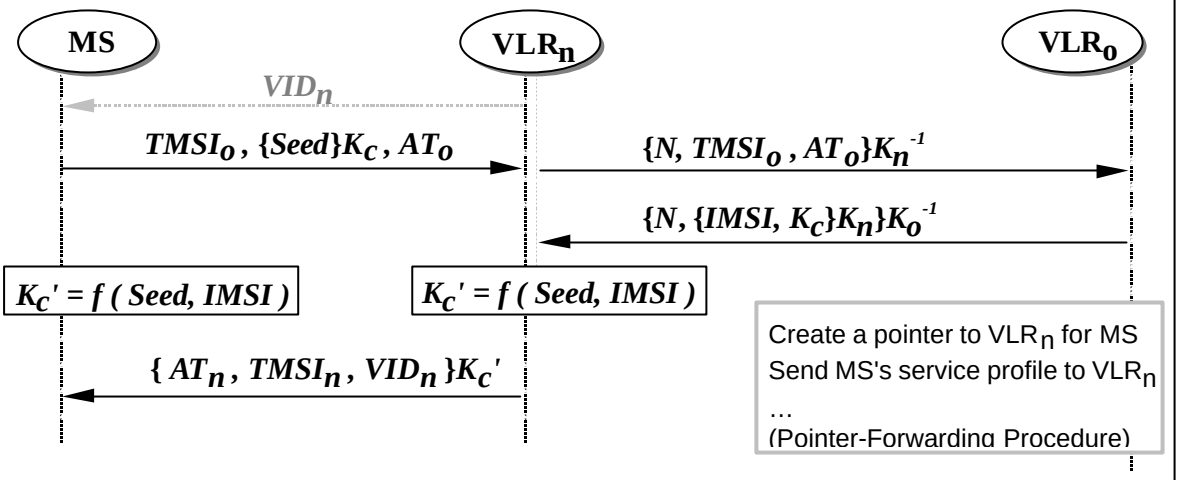

Fig. 6 The basic initial authentication procedure of the chain authentication (for a single service provider in a local area)

AT are used only when MS remains in the home domain. If MS moves and registers elsewhere, the three parameters are rendered illegal and useless.

MS must invoke the initial authentication procedure upon arrival in a new domain to register when VLR_n will generate three new parameters $TMSI_n$, K_c', and AT_n for MS in the newly visited domain. A single domain usually covers a local area and contains only one service provider. Herein, we propose a basic procedure for initial authentications in such environments. However, the future PCS will cover multiple service domains in a local area. That is, two or more service providers concurrently compete for a new MS arriving in their administrative regions (domains). Thus, we propose another enhanced initial authentication procedure for such complicated environments.

The Basic Initial-Authentication-Procedure

Step 0) While detecting a new domain name VID_n in the received beacon, MS should record VID_n and invoke the following procedure.

Step 1) First, MS generates a random number *Seed* and enciphers it with K_c. Then, MS sends its $TMSI_O$, $\{Seed\}K_c$, and AT_O to VLR_n.

Step 2) VLR_n generates a nonce N and ciphers the message (N, $TMSI_O$, AT_O) with its private key K_n^{-1}. Finally, VLR_n sends the ciphered message to VLR_o. VLR_n retains the ciphertext $\{Seed\}K_C$.

Step 3) VLR_o uses the public key K_n of VLR_n to decipher the received message. If the deciphering is successful, VLR_o believes that the message was sent by VLR_n. VLR_o then uses $TMSI_O$ to find out the corresponding *IMSI*, *Kc*, and AT_O in its database. Since AT_O is known only to MS and VLR_o and is used once, if VLR_o receives original AT_O, VLR_o believes that this message is not replayed and that MS is in the domain of VLR_n.

Step 4) VLR_o sends $\{N, \{IMSI, K_C\}K_n\}K_O^{-1}$ to VLR_n. Because $\{IMSI, K_C\}K_n$ is ciphered with K_n, VLR_n alone can disclose this information.

Step 5) VLR_n uses the public key K_O of VLR_o to decipher the received ciphertext and verifies the authenticity of returned nonce N, after which, VLR_n believes that a fresh message arrived from VLR_o. Thus, VLR_n believes that MS has been successfully authenticated by VLR_o. VLR_n can create a new registration record for MS and generate a new temporary identity $TMSI_n$ and an authentication token AT_n for MS. Another important task in this step is to produce the new session key K_C' shared only by MS and VLR_n. VLR_n initially uses the MS-VLR_o session key K_C to decipher the ciphertext $\{Seed\}K_C$ kept at step 2 and obtains the random number *Seed*. The number and MS's *IMSI* are then assigned to a one-way function f to produce the new session key $K_{C'}$, shared by MS and VLR_n. That is, $K_{C'} = f(Seed, IMSI)$.

Step 6) VLR_n sends $\{AT_n, TMSI_n, VID_n\}K_{C'}$ to MS.

Step 7) Before MS receives the message sent at step 6, he uses *Seed* and his *IMSI* to generate $K_{C'}$ using the same one-way function f used by VLR_n. MS then deciphers the message using the new session key and, finally, verifies the varible VID_n with the identity listened from the beacon. The matched identity implies that the authentication request is correctly passed to VLR_o via VLR_n, and VLR_o must trust VLR_n, otherwise, VLR_o

disapproves the request and does not return the previous session key K_C to VLR_n. Without the correct K_C, VLR_n fails to compute K_C' and thus cannot generate the fourth message $\{AT_n, TMSI_n, VID_n\}K_C'$. Therefore, MS can use VID_n to verify the correctness of the fourth message in case that the message may be corrupted.

The basic initial authentication procedure is simple, but vulnerable to the complicated environment containing multiple service providers in a local area. Consider, for example, that MS arrives in a new area containing two service providers, VLR_1 and VLR_2. MS chooses VLR_1 to register and sends the first authentication message to VLR_1. VLR_2 also likely receives (eavesdrops on) this messages and sends the second message to VLR_o. Consequently, arrival of two messages from VLR_1 and VLR_2 confuses VLR_o who fails to distinguish the service provider selected by MS. In the worst case, if VLR_2 can block VLR_1's message, VLR_2 masquerades as VLR_1 to serve MS. However, we improve the basic procedure to help VLR_o distinguish the valid VLR_n. Fig. 7 depicts the enhanced initial authentication procedure. The procedure includes the following steps.

The Enhanced Initial-Authentication-Procedure

Step 0) While detecting a new domain name VID_n in the received beacon, MS should record VID_n and invoke the following procedure.

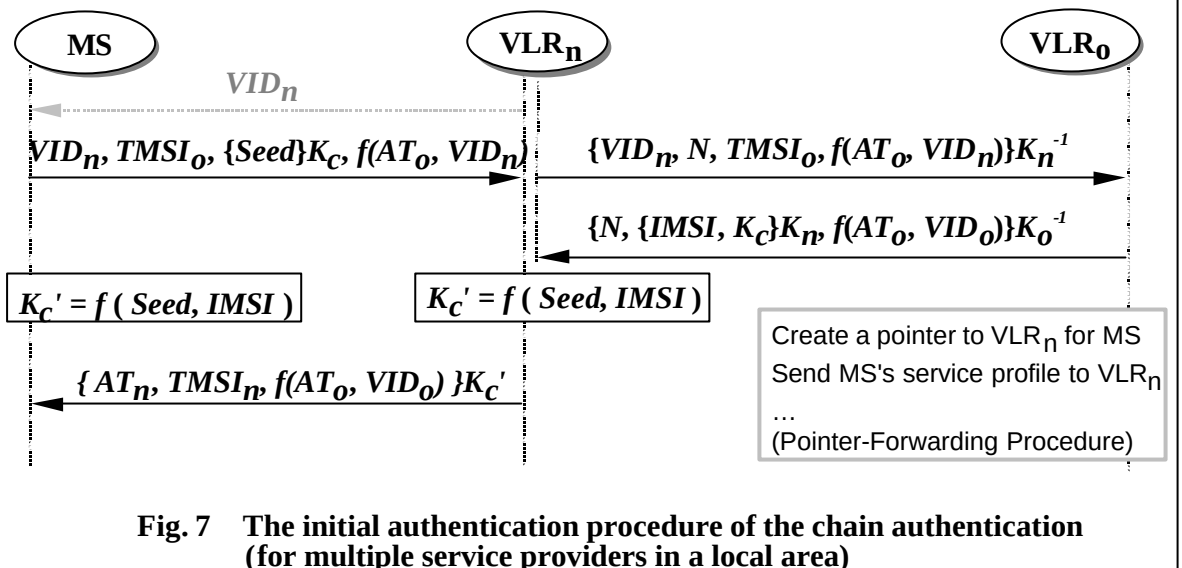

Fig. 7 The initial authentication procedure of the chain authentication (for multiple service providers in a local area)

Step 1) MS initially generates a random number *Seed* and then transmits (VID_n, $TMSI_O$, $\{Seed\}K_C$, $f(AT_O, VID_n)$ to VLR_n, where VID_n is the identity of the service provider chosen by MS. Since AT_O is known only to MS and VLR_O, nobody, even VLR_n, can forge $f(AT_O, VID_n)$.

Step 2) By using VID_n, VLR_n assures itself that the message is authentic. The following steps resemble the basic procedure: VLR_n sends $\{VID_n, N, TMSI_O, f(AT_O, VID_n)\}K_n^{-1}$ to VLR_O. The ciphertext $\{Seed\}K_C$ is retained in VLR_n.

Step 3) VLR_O deciphers the received message and computes $f(AT_O, VID_n)$ using the same one-way function. If the computed result is similar to $f(AT_O, VID_n)$ received from VLR_n, MS's authentication is successful and VLR_O believes that MS chose VLR_n to support services.

Step 4) VLR_O re-computes $f(AT_O, VID_O)$ and sends $\{N, \{IMSI, K_C\}K_n, f(AT_O, VID_O)\}K_O^{-1}$ to VLR_n.

Step 5) VLR_n deciphers the received ciphertext and generates $TMSI_n$, AT_n, and the new session key K_C' as in the basic procedure.

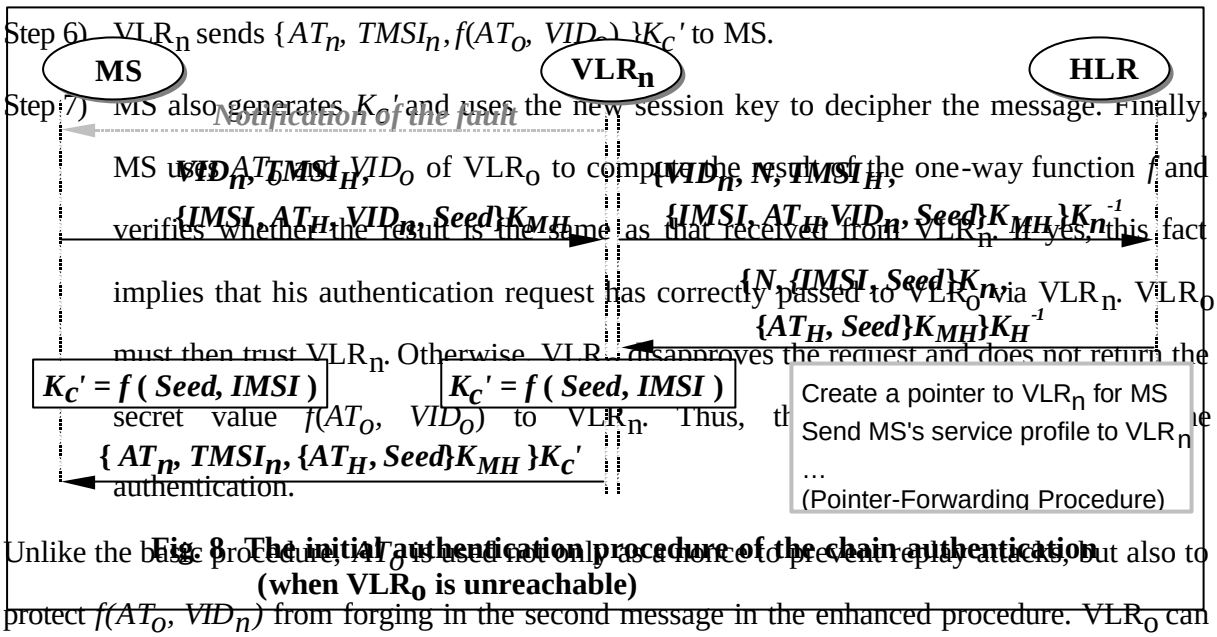

Fig. 8. The initial authentication procedure of the chain authentication (when VLR_O is unreachable)

Step 6) VLR_n sends $\{AT_n, TMSI_n, f(AT_O, VID_n)\}K_C'$ to MS.

Step 7) MS also generates K_C' and uses the new session key to decipher the message. Finally, MS uses AT_n, $TMSI_n$, VID_O of VLR_O to compute the result of the one-way function f and verifies whether the result is the same as that received from VLR_n. If yes, this fact implies that his authentication request has correctly passed to VLR_O via VLR_n. VLR_O must then trust VLR_n. Otherwise, VLR_O disapproves the request and does not return the secret value $f(AT_O, VID_O)$ to VLR_n. Thus, the fault is detected in the authentication.

Unlike the basic procedure, AT_O is used not only as a nonce to prevent replay attacks, but also to protect $f(AT_O, VID_n)$ from forging in the second message in the enhanced procedure. VLR_O can thus recognise the service provider selected by MS. Furthermore, since MS can predict $f(AT_O, VID_O)$, verification of this value allows MS to detect faults in the last message, like the information VID_n used in the basic procedure. (See section 5 for proof)

Notably, the initial authentication procedure can be combined with the registration process which adopts the pointer-forwarding scheme, as mentioned in section 3, to track MS's current location. Details of the combination are not within the scope of this paper. The following example is made. After VLR_O has approved the authentication request and returned the response to VLR_n, it can create a pointer to VLR_n for MS in itself database. At this moment, VLR_O can also transmit other MS's personal information, such as the profile of subscribed services, to VLR_n. If these messages are confidential, VLR_O is entitled to generate a shared secret key and send it to VLR_n by inserting the secret key into $\{IMSI, K_C\}K_n$ mentioned in step 4.

Fault Tolerance for the Initial Authentication Procedure

With our initial authentication procedure, the authentication fails if VLR_O is unreachable, such as when VLR_O crashes, the link between VLR_O and VLR_n is broken, or VLR_n does not know VLR_O. However, the fault probability is low due to the high fault tolerance of telecommunication

systems. In the event of this fault, an additional scheme is necessary to authenticate MS and to continue the registration process. Existing schemes can be used. For instance, by signalling by VLR_n to MS of the fault, both parties adopt Molva's scheme or other conventional authentication protocols with the assistance of HLR to complete the authentication and registration process. Alternatively, our initial authentication procedure can be modified to tolerate the fault with additional secret information. Fig. 8 presents a feasible solution. K_{MH} is a long-term key shared only by MS and HLR. AT_H and $TMSI_H$ denote the old AT and $TMSI$ used, respectively, in the last authentication between MS and HLR. The procedure resembles the previous two authentication procedures, except that MS does not use one-way function to protect data and that *Seed* is disclosed by HLR, rather than by VLR_n itself. Note that K_C, the old session key coresponding to AT_H and $TMSI_H$, in this procedure, will not be used in encipher data because it has been known to the first visited VLR after MS leaving HLR (referring to the previous section). Instead, MS uses AT_H to prove the origin and the freshness of the authentication request because only MS and HLR know AT_H. Thus, VLR_n continues the authentication process with this modified initial authentication procedure while VLR_o is unreachable. Moreover, the additional efforts of MS and HLR only retain K_{MH}, AT_H, and $TMSI_H$.

In addition to the condition that VLR_o is unreachable, this procedure is also suitable to the condition that HLR is nearer than VLR_o, e.g., the new domain in which MS arrives is HLR. If VLR_n always selects the nearest VLR_o/HLR to help authenticate MS, the network burden caused by the authentication will be significantly reduced. However, to achieve this benefit, VLR_n must be intelligent enough to determine the communication cost of contacting VLR_o and HLR.

The Subsequent Authentication Procedure

Some practical systems require re-authentication by MS, who must decide/establish a new session key when seeking an authorized service, after MS registers in the current domain. The design of the subsequent authentication is trivially based on the proposed initial authentication

procedure. The procedure consists of only two messages that can be embedded in the service request or isolated from the request. Herein, we merely describe the contents of authentication messages (Fig. 9). K_C and AT_O denote the session key and the authentication token generated by the previous initial/subsequent authentication procedure, while K_C' and AT_n present the new session key and the authentication token generated presently.

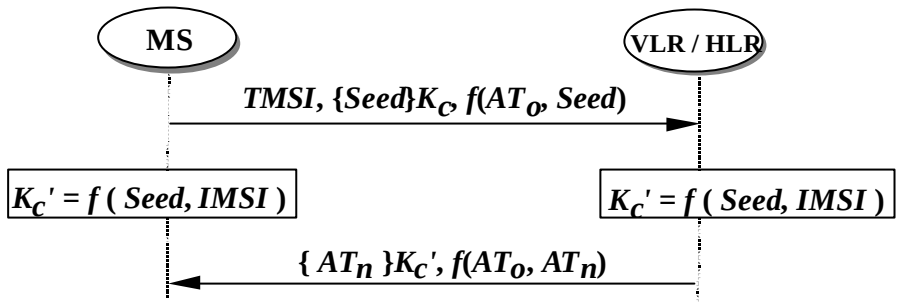

Fig. 9 The subsequent authentication procedure of the chain authentication

Step 1) MS generates a random number *Seed* and enciphers it with K_C shared between MS and the current service provider, i.e., VLR or HLR. Then, MS sends his *TMSI*, the ciphertext $\{Seed\}K_C$, and $f(AT_O, Seed)$ to his service provider.

Step 2) The service provider uses *TMSI* to determine the corresponding *IMSI*, K_C, and AT_O in its database. Next, *Seed* disclosed from the ciphertext with K_C is used to compute $f(AT_O, Seed)$. If the computed result is the same as $f(AT_O, Seed)$ received from MS, the authentication is successful. Although hostile attacks may replace $\{Seed\}K_C$, such a condition is easily detected by verifying the correctness of $f(AT_O, Seed)$; the output of the one-way function cannot be masqueraded.

Step 3) MS and the service provider generate the new session key, that is, $K_C' = f(Seed, IMSI)$, by the key-generation method mentioned in the initial authentication procedure.

Step 4) The service provider generates a new authentication token AT_n and transmits $(\{ AT_n \}K_C', f(AT_O, AT_n))$ to MS. Correct generation of K_C' by the legal service provider

enables MS to decipher this message and obtain the new token AT_n for the subsequent authentication.

If faults occur in the subsequent authentication procedure, MS must only invoke the (basic/ enhanced) initial authentication procedure to restart the services with the present VLR.

5. Protocol Analysis and Comparison

This section analyzes the security and performance of the proposed protocol and compares it with traditional mechanisms. (Only the enhanced initial authentication procedure is considered in this section.)

(A) Security Analysis

Once trust is established between MS and VLR_o, whether MS and VLR_n can trust each other can be determined using three criteria:

- VLR_o and VLR_n trust each other,
- VLR_n proves to VLR_o that MS has arrived in the new domain, and
- VLR_n proves to MS that VLR_o trusts and authorizes VLR_n.

We now explain how our protocol uses four messages to accomplish authentication and will use the BAN-logic to verify why it can correctly authenticate each communication party. (The BAN-logic formally verifies the correctness of authentication protocols. BAN-logic is neither sufficient nor complete; however, it can help verify the correctness of an authentication protocol to some extent. Refer to Ref. [22] for details.)

We first idealize our protocol to the BAN-logic form:

$MS \rightarrow VLR_n$: $\{Seed\}K_c, \langle VID_n \rangle_{AT_o}$ (M1)

$VLR_n \rightarrow VLR_o$: $\{N, TMSI_o, \langle VID_n \rangle_{AT_o}\}K_n^{-1}$ (M2)

$VLR_o \rightarrow VLR_n$: $\{N, \{IMSI, MS \overset{Kc}{\longleftrightarrow} VLR_o\}K_n, \langle VID_o \rangle_{AT_o}\}K_o^{-1}$ (M3)

$VLR_n \rightarrow MS$: $\{AT_n, TMSI_n, \langle VID_o \rangle_{AT_o}\}K_c'$ (M4)

Notably, we represent the one-way function $f(Y, X)$ as $\langle X \rangle_Y$, the form of the shared secret

defined in BAN-logic [22]. $<X>_Y$ means that X is combined with the secret formula Y. The shared secret formula Y is AT_O in our protocol. Since AT_O is shared only by MS and VLR_O, and is used only once, it is difficult to forge or replay $f(AT_O, X)$, where X is VID_N or VID_O. Thus, $f(AT_O, X)$ can prove the origin of the message and guarantee its freshness. That is,

$$VLR_O \textbf{ believes fresh}(<VID_N>_{AT_O}) \text{ and}$$
$$MS \textbf{ believes fresh}(<VID_O>_{AT_O}). \quad \text{(D1)}$$

The following analysis first employs the BAN-logic to describe the deductions obtained upon receipt of each message. The detailed proof is presented in Appendix. We then verify that our protocol meets the three criteria.

After (M1) – VLR_N suspects all information received because it cannot verify the message. Thus, no deduction is derived.

After (M2) – The secret key K_N^{-1} of VLR_N encrypts the message, VLR_O believes that VLR_N is the source of M2. Based on the deduction (D1) mentioned above, we deduce that

$$VLR_O \textbf{ believes } VLR_N \textbf{ believes } (N, TMSI_O, <VID_N>_{AT_O}). \quad \text{(D2)}$$
$$VLR_O \textbf{ believes } (N, TMSI_O, <VID_N>_{AT_O}). \quad \text{(D3)}$$

After (M3) – As for (M2), VLR_N believes M3 is sent by VLR_O, because the nonce N is generated by VLR_N itself, and VLR_N can verify if N is fresh. Thus, we deduce that

$$VLR_N \textbf{ believes } VLR_O \textbf{ believes } (N, \{IMSI, MS \overset{K_C}{\longleftrightarrow} VLR_O\}K_N, <VID_O>_{AT_O}). \quad \text{(D4)}$$
$$VLR_N \textbf{ believes } (N, \{IMSI, MS \overset{K_C}{\longleftrightarrow} VLR_O\}K_N, <VID_O>_{AT_O}). \quad \text{(D5)}$$

By (D5), it follows that VLR_N believes also $\{IMSI, MS \overset{K_C}{\longleftrightarrow} VLR_O\}K_N$. So, we deduce that

$$VLR_N \textbf{ believes } IMSI. \quad \text{(D6)}$$
$$VLR_N \textbf{ believes } MS \overset{K_C}{\longleftrightarrow} VLR_O. \quad \text{(D7)}$$

After (M4) – If VLR_N obtains the correct *Seed* in M1, MS and VLR_N will share the same session key $K_{C'}$. Thus, by (D1), we make the following deductions

$$MS \textbf{ believes } VLR_N \textbf{ believes } (AT_N, TMSI_N, <VID_O>_{AT_O}) \quad \text{(D8)}$$
$$MS \textbf{ believes } (AT_N, TMSI_N, <VID_O>_{AT_O}). \quad \text{(D9)}$$

Since K_c' is generated by MS, $K_c' = f(Seed, IMSI)$, therefore

MS **believes** ($MS \overset{K_c'}{\longleftrightarrow} VLR_n$). (D10)

Using these deductions allow us to demonstrate that our protocol fulfills the three criteria to complete the authentication.

Criterion 1 – VLR_o and VLR_n must trust each other. Deduction (D2) and (D3) prove that message 2 meets the requirement that VLR_o trusts VLR_n. On the other hand, deductions (D4) and (D5) prove that message 3 satisfies the requirement that VLR_n trusts VLR_o. Therefore, criterion 1 is satisfied by messages 2 and 3.

Criterion 2 – VLR_n must prove to VLR_o that MS has arrived in the new domain. This requirement is trivial because our protocol guarantees (D2 and D3) and that only real MS can generate $f(AT_O, VID_n)$. VLR_o therefore believes MS who announces identity by $TMSI_O$ and $f(AT_O, VID_n)$. Besides, since $f(AT_O, VID_n)$ contains the identity of VLR_n and cannot be forged by VLR_n, $f(AT_O, VID_n)$ suggests the location of MS.

Criterion 3 – VLR_n must prove to MS that VLR_o trusts and authorizes VLR_n. MS can check this condition by decrypting message 4 with the new session key K_c' generated by MS. If MS can decrypt it and correctly verify $f(AT_O, IMSI)$, MS accepts the authority of VLR_n granted by VLR_o. Deductions (D9) and (D10) prove that our protocol satisfies this requirement.

Although the chain authentication protocol can fulfill the above criteria, we cannot infer that "VLR_n **believes** K_c'" in the above deductions. Our protocol only guarantees that VLR_n trusts MS (see (D5)) and is able to get the correct session key K_c (see (D7)), but it does not imply that VLR_n obtains the correct *Seed* (see (d15) in Appendix). This inability is because $\{Seed\}K_c$ may be replaced by hostile intruders in message 1. VLR_n cannot confirm the validity of session key K_c' until the former correctly decrypt the data encrypted using K_c' by MS. Fortunately, even in the worst case, when $\{Seed\}K_c$ was replaced, only K_c' of VLRn is inconsistent with MS's and the following communication will fail. Thus, the security of systems will not be compromised.

We guarantee, as explained above, that MS and VLR_n can trust each other. That is, our protocol can authenticate the three communication parties, including MS, VLR_o, and VLR_n.

(B) Performance Evaluation

A simple HLR/VLR network model is illustrated in Figure 10 to show the benefit of the proposed authentication protocol. All HLR/VLRs exchange control signals, such as the registration messages, through middle switches. (In telecommunication systems, the switch is commonly called the Signalling Transfer Point (STP) in the Common Channel Signalling network with a Signalling System No. 7 (SS7) protocol [26].) For simplicity, we assume that the cost of signalling between HLR/VLRs is dependent on the number of switches. Therefore, the geographically contiguous domains have the smallest cost, defined as 1. And we assume that the proposed chain authentication is combined with an intelligent VLR described in the previous section, so that VLR_n will require VLR_o or HLR to authenticate MS based on their costs. To compare the improvement, the traditional authentication scheme with the assistance of HLR, such as IS-41 or Molva's scheme, is used in the performannce evaluation.

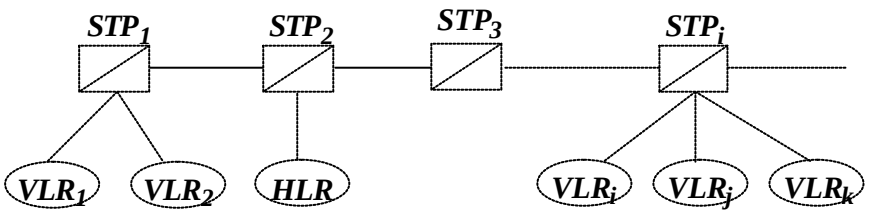

Fig. 10 A network model for signal exchange between HLR/VLRs

Whenever MS roams from an old domain to a new domain, one of three conditions occurs. The mobile user may roam from his HLR to a VLR, from a VLR to his HLR, or from a VLR to another VLR. The following table summarizes the three conditions and the corresponding costs of the first authentication process with different protocols. For variant behavior of mobile users, we define three probabilities, P_1, P_2 and P_3, for these conditions and their sum is equal to 1. And we assume P_1 is equal to P_2 since a user will finally return to his home after a long travel. With the assumption of the cost described in the previous paragraph, we can calculate the costs for

each condition as shown in the following table.

Condition	Old domain	New Domain	Probability	Cost with traditional schemes	Cost with our scheme
I	*HLR*	VLR_J	P_1	$x = y$	x
II	VLR_I	*HLR*	P_2	*0*	*0*
III	VLR_i	VLR_J	P_3	y	Min$[x, y]$

x: the number of switches between the new and old domain

y: the number of switches between the new domain and HLR

Table.

Consequently, the average costs of both schemes are

$$\text{COST}_{\text{traditional}} = x * P_1 + y * P_3 \text{ and}$$
$$\text{COST}_{\text{our}} = x * P_1 + \text{Min}[x, y] * P_3.$$

Intuitively, the cost of our scheme is smaller than the traditional schemes and the difference is

$$\Delta C = \text{COST}_{\text{traditional}} - \text{COST}_{\text{our}} = (y - \text{Min}[x, y]) * P_3.$$

From the equation, we conclude two factors that affect the network burden:

- The mobile user should frequently roam among VLRs.
- The cost of message transmission between the new domain and HLR should be larger than that between the new and old visiting domains.

If we assume the user handset is always power-on while traveling, the authentication and registration process should be immediately invoked when MS arrives in a new domain. That is, the new and old domains are geographically contiguous, and x is equal to 1. Thus, the difference can be simplfied as

$$\Delta C = (y - 1) * P_3.$$

In order to clearly show the reduction of network burden due to the proposed chain authentication protocol, we normalize ΔC as the improvement rate R,

$$R = \frac{\Delta C}{COST_{traditional}} = \frac{(y-1) * P_3}{1 * P_1 + y * P_3}.$$

Figure 11 shows the result under variant behaviors of mobile users. Our scheme significantly reduces the network burden caused by the message exchanges between the new domain and HLR when MS frequently roams out of his home domain. The improvement will be degraded in the

real world since user handsets are not always power-on during their trips. Thus, x may be larger than y, and VLR will require HLR to help authenticate MS. Consequently, in the worst case, our scheme has the same network burden as traditional schemes.

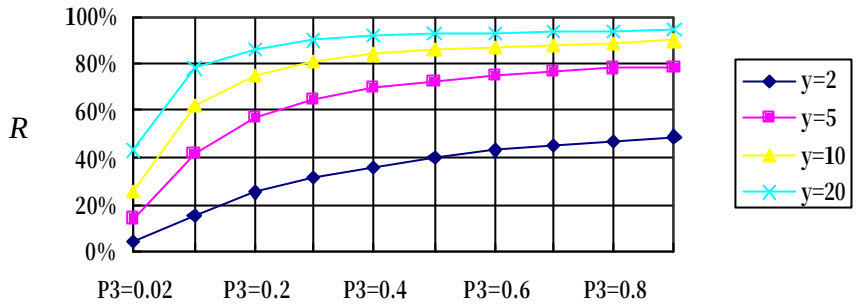

Fig. 11 The performance evaluation for the proposed scheme

(C) The Advantages of the Chain Authentication Protocol

The chain authentication protocol has the following merits:

No assistance from HLR – The distance between VLR_n and VLR_o is generally shorter in a large PCS network than that between VLR_n and HLR. A long connection path causes long propagation delay of messages, reduces the reliability of the communication channel, and makes far more traffic in the network. The proposed protocol merely uses VLR_o to help VLR_n authenticate MS; our scheme authenticates MS rapidly. An efficient authentication scheme is particularly crucial when time is of essence.

Another merit of the mechanism is that the certification between VLR_n and VLR_o is easier and faster than between VLR_n and HLR, because VLR_o lies closer to VLR_n than does HLR. The certification is established by the public key scheme with a hierarchical architecture [18]. Therefore, the cost of the certification relies on the distance between the two communicating parties in the architecture.

Low overhead – Only four messages need be exchanged in the chain authentication protocol. Since VLR_n must contact with VLR_o in the network to query the user information, four messages are the minimum requirement. The proposed scheme does not need the assistance from

HLR, consequently, the chain authentication can significantly reduce the network burden caused by the authentication process when MS tries to register in a new domain.

Furthermore, computation power required is low. Using a one-way function between MS and VLR_n and a symmetric cryptosystem, such as DES [23], guarantees the confidentiality of messages. These two mechanisms demand only simple computation and have been applied in existing mobile telecommunication systems, such as GSM. Although we use the public key cryptosystem between VLR_n and VLR_o, which is more complex than above two mechanisms, VLRs can easily perform the task because they offer better computation power in practical systems.

Subscriber identity confidentiality – If IMSI is directly used to identify MS's messages, MS's moving from a domain to another can be traced by listening to his identity on the radio path. The relation between the transmitted user data and MS is also available. To prevent the privacy of user location and to improve other security features, e.g., user data confidentiality, assurance the confidentiality of subscriber identities in the communication channels is imperative. We recommend that the user be registered as a temporary identity; this mechanism resembles that used by GSM [24]. A temporary subscriber identity *TMSI* used in the chain authentication protocol to identify MS's request of services. *TMSI* is generated by VLR of the domain and submitted to MS in ciphertext when MS arrives in a new domain. Since the temporary identity is changed as the user travels between domains, tracing the user location on the radio path is impossible.

Communication confidentiality – To maintain confidentiality of communications between MS and VLR, a session key is needed to encipher/decipher the data transmitted on the channels. Herein, we only suggest using symmetric cryptosystems because the system demands low computational power of MS. The proposed protocol uses a one-way function with two parameters, a random number *Seed* and the user identity *IMSI* to generate the session key K_C. In

addition, within the initial/subsequent authentication procedure, all exchanged secret information, such as $IMSI$, AT_n, and $TMSI_n$, is transmitted as ciphertext to prevent eavesdropping.

Authenticating overall participant communication parties – The modern telecommunication system only authenticates those subscribers seeking services. Furthermore, the system assumes that the network is trustworthy. In contrast, the proposed protocol authenticates all communication parties participating in the protocol, i.e. MS, VLR_n, and VLR_o. Between VLR_o and VLR_n, the public key cryptosystem are used to authenticate each other. Between MS and VLR_o, the secret information, AT_o and $IMSI$, is used to authenticate each other. Based on the two mutual authentications, MS and VLR_n can authenticate each other, as mentioned earlier in this section.

Consideration of multiple service providers in a local area – Future PCS network will include multiple competing service providers in a local area. The proposed enhanced initial authentication procedure enables VLR_o to distinguish the service provider chosen by MS.

Domain separation – Both the session key K_C and the temporary identity $TMSI$ are local information. They are only valid within the domain that generated them. Thus, all domains are separated by this local and secret information. (If the administration of HLR/VLR betrays, an evil system-operator armed with MS's secret information, i.e. $TMSI$, K_C, and AT, can masquerade as MS in a different domain. Before the masquerade, however, if MS moves and is registered elsewhere, the evil operator (although holding these secrets) is powerless).

Session key confidentiality – Our session key generation relies on a random number and a secret information $IMSI$. MS and VLR_n generate the key, that is unknown to anyone, including HLR and VLR_o. This scheme ensures confidentiality of the new session key and reduces the probability eavesdropping by a third party. (Many practical systems, including GSM, adopt this scheme.)

Low cost for preventing replay attacks – We use an authentication number *AT* rather than the timestamp in the exchanged messages to ensure the freshness. Therefore, the clock synchronization is unnecessary and message replay is difficult.

The following figure shows the comparison between our protocol and other protocols.

Protocol	Assistance of HLR	Messages needed	Authenticated parties	Confidentiality of authentication messages	Clock synchronization
IS-41	Yes	5**	MS	Only between MS and VLR_n	No
CDPD	Yes	6	MS	Only between MS and VLR_n	No
GSM	No*	6***	MS	Only between MS and VLR_n	No
Mobile-IP	Yes	4	only MS-HLR is mandatory	All	Yes
Molva's	Yes	4	MS, VLR_n, HLR	All	Yes
Chain	No	4	MS, VLR_n, VLR_o	All	No

* The registration still requires the assistance of HLR.

** It is the S authentication scheme in IS-41.

*** It does not include the messages of the location updating and acknowledges.

Fig. 12 Comparisons of the protocols

6. Conclusions

To enhance the quality of communication services, users and service providers desire a more secure environment to prevent accessing unauthorized services or disclosing confidential information. Numerous modern mobile telecommunication systems contain simple security functions, such as subscribers' authentication and the confidentiality of the communication on radio paths. These mechanisms need management servers in home domains to authenticate subscribers. However, in a large communication network, the overhead of accessing HLR from the visited domain significantly degrades the system performance. This paper presents a method,

referred to herein as the chain authentication protocol. This protocol contains a series of procedures, including the preparation for subscribing in HLR, the initial authentication for registering in a new domain, and the subsequent authentication for querying a service. In the initial authentication procedure, we exemplify two cases regarding a local area containing a single or multiple service provider(s). Furthermore, we also consider the occurrence of the fault that VLR_0 is unreachable during the initial authentication procedure, and a possible solution is proposed by modifying the original procedure.

Our protocol guarantees the confidentiality of exchanged messages and of the subscriber's identity; furthermore, the protocol uses minimal messages to authenticate all communicating parties (including MS and all participative service providers), does not require the clock synchronization, and, importantly, operates independently of HLR for MS authentication. The protocol can be applied in large communication networks with multiple service providers, such as the global PCS network.

Acknowledgement

This work is supported in part by FarEasTone Telecommunications Co., Ltd. The authors are grateful to Dr. Herman Rao and Dr. Hung-Fa Sun of FarEasTone Telecommunications Co. for their many suggestions that help improve the paper.

Reference

[1] Bennett Z. Kobb, "Personal Wireless," *IEEE SPECTRUM*, Jun. 1993.

[2] M. Mouly, M. B. Pautet, "The GSM System for Mobile Communications," ISBN: 2-9507190-0-7, 1992.

[3] CDPD Consortium, "Cellular Digital Packet Data System Specification," Release 1.0, July 1993.

[4] EIA/TIA, "Cellular Intersystem Operations (Rev. C)," *Technical Report IS-41*, EIA/TIA, 1995.

[5] S. P. Shieh, C. T. Lin, and J. T. Hsueh, "Secure Communication in Global Systems for Mobile Telecommunications," *Proceedings of First Workshop on Mobile Computing*, pp. 136-142, 1995.

[6] C. Perkins, Editor, "IP Mobility Support," *RFC 2002*, Oct. 1996.

[7] R. Molva, D. Samfat, and G. Tsudik, "Authentication of Mobile Users," *IEEE Network*, Mar./Apr. 1994.

[8] R. Jain, Y. B. Lin, C. Lo, and S. Mohan, "A Caching Strategy to Reduce Network Impacts of PCS," *IEEE Journal on Selected Areas in Communications*, Vol. 12, No. 8, Oct. 1994.

[9] R. Jain, Y. B. Lin, and S. Mohan, "A Forwarding Strategy to Reduce Network Impacts of PCS," *IEEE INFOCOM*, 1995.

[10] C. Perkins, "Mobile-IP Local Registration with Hierarchical Foreign Agents," *IETF Internet-Draft*, Feb. 1996.

[11] "GSM 02.09: Security Aspects," European Telecommunications Standards Institute, Jun. 1993.

[12] W. Diffie and M. Hellman, "New Directions in Cryptography," *IEEE Trans. Information Theory*, Nov. 1976.

[13] R. Atkinson, "Security Architecture for the Internet Protocol," *RFC-1825*, Aug. 1995.

[14] "Recommendation X.509 and ISO 9594-8, Information Processing Systems - Open Systems Interconnection - The Directory - Authentication Framework," *CCITT Technical report*, Mar. 1988.

[15] J. Linn, "Privacy Enhancement for Internet ElectronicMail, Parts I: Message Encryption and Authentication Procedures," *RFC-1421*, SRI Network Information Center, Feb. 1993.

[16] S. Kent, "Privacy Enhancement for Internet ElectronicMail, Parts II: Certificate-Based Key Management," *RFC-1422*, SRI Network Information Center, Feb. 1993.

[17] D. Balenson, "Privacy Enhancement for Internet ElectronicMail, Parts III: Algorithms, Modes, and Identifiers," *RFC-1423*, SRI Network Information Center, Feb. 1993.

[18] S. Chokhani, "Toward a National Public Key Infrastructure," *IEEE Communications Magazine*, Sep. 1994.

[19] Ravi S. Sandhu and Edward J. Coyne, "Role-Based Access Control Models," *IEEE Computer*, Feb. 1996.

[20] "Trusted Computer System Evaluation Criteria," *DoD STD-5200.28*, Dec. 1985.

[21] R. L. Rivest, A. Shamir, and L. Adelman, "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems," *Commun. of the ACM*, Vol. 21, No. 2, pp. 120-126, Feb. 1978.

[22] M. Burrows, M. Abadi, and R. Needham, "A Logic of Authentication," *ACM Transaction on Computer Systems*, Vol. 8, No. 1, Feb. 1990.

[23] NBS FIPS PUB 46-1, "Data Encryption Standard," National Bureau of Standards, U.S. Department of Commerce, Jan. 1977.

[24] "GSM 03.20: Security Related Network Functions," *European Telecommunications Standards Institute*, Jun. 1993.

[25] K. Buchanan, R. Fudge, D. McFarlane, T. Phillips, A. Sasaki, and H. Xia, "IMT-2000: Service Provider's Perspective," *IEEE Personal Communications Magazine*, Vol. 4, No. 4, Aug. 1997.

[26] A. R. Modaressi and R. A. Skoog, "Signalling System No. 7," *A tutorial, IEEE Communications Magazine*, pp. 19-35, Jul. 1990.

[27] Y. B. Lin, *Introduction to Mobile Network Management*, Wei-Keg Publishing Co., 1997.

[28] EIA/TIA, "Mobile Station-base Station Compatibility Standard for Dual-Mode Wideband Spread Spectrum Cellular System," *Technical Report TIA/EIA/IS-95-A*, EIA/TIA, 1995.

Appendix

In this proof, we use the same notations and logical postulates as the BAN-logic [15].

- (P1): for the message using shared key, we postulate:

$$\frac{P \text{ believes } P \overset{K}{\longleftrightarrow} Q, \quad P \text{ sees } \{X\}K}{P \text{ believes } Q \text{ said } X}$$

- (P2): for the message using public key, we postulate:

$$\frac{P \text{ believes } \overset{K}{\mapsto} Q, \quad P \text{ sees } \{X\}K\text{-}1}{P \text{ believes } Q \text{ said } X}$$

- (P3): for the message using secret information, we postulate:

$$\frac{P \text{ believes } Q \overset{Y}{\rightleftharpoons} P, \quad P \text{ sees } \langle X \rangle Y}{P \text{ believes } Q \text{ said } X}$$

- (P4): the nonce-verification rule:

$$\frac{P \text{ believes } fresh(X), \quad P \text{ believes } Q \text{ said } X}{P \text{ believes } Q \text{ believes } X}$$

- (P5): the jurisdiction rule:

$$\frac{P \text{ believes } Q \text{ controls } X, \quad P \text{ believes } Q \text{ believes } X}{P \text{ believes } X}$$

- (P6):

$$\frac{P \text{ sees } (X, Y)}{P \text{ sees } X}$$

- (P7):

$$\frac{P \text{ sees } \langle X \rangle Y}{P \text{ sees } X}$$

- (P8):

$$\frac{P \text{ believes } P \overset{K}{\leftrightarrow} Q, \quad P \text{ sees } \{X\}K}{P \text{ sees } X}$$

- (P9):

$$\frac{P \text{ believes } \overset{K}{\mapsto} P, \quad P \text{ sees } \{X\}K}{P \text{ sees } X}$$

- (P10):

$$\frac{P \text{ believes } \overset{K}{\mapsto} Q, \quad P \text{ sees } \{X\}K\text{-}1}{P \text{ sees } X}$$

- (P11):

$$\frac{P \text{ believes fresh}(X)}{P \text{ believes fresh}(X, Y)}$$

As mentioned in section 5, we have the following four idealized messages.

(M1) $MS \rightarrow VLR_n$: $\{Seed\}K_c, \langle VID_n \rangle AT_o$

(M2) $VLR_n \rightarrow VLR_o$: $\{VID_n, N, TMSI_o, \langle VID_n \rangle AT_o\}K_n^{-1}$

(M3) $VLR_o \rightarrow VLR_n$: $\{N, \{IMSI, MS \overset{K_c}{\leftrightarrow} VLR_o\}K_n, \langle VID_o \rangle AT_o\}K_o^{-1}$

(M4) $VLR_n \rightarrow MS$: $\{AT_n, TMSI_n, \langle VID_o \rangle AT_o, MS \overset{K_c'}{\leftrightarrow} VLR_n\}K_c'$

Before the proof, the following assumptions are made.

(A1) VLR_o **believes** $\overset{K_n}{\mapsto} VLR_n$

(A2) VLR_n **believes** $\overset{K_o}{\mapsto} VLR_o$

(A3) VLR_n **believes** $\overset{K_n}{\mapsto} VLR_n$

(A4) VLR_o **believes** $VLR_o \overset{AT_o}{\leftrightarrow} MS$

(A5) MS **believes** $VLR_o \overset{AT_o}{\leftrightarrow} MS$

(A6) MS **believes** $MS \overset{K_c'}{\leftrightarrow} VLR_n$

(A7) VLR_o **believes fresh**(AT_o)

(A8) MS **believes fresh**(AT_o)

(A9) VLR_n **believes fresh**(N)

(A10) VLR_o **believes** VLR_n **controls** $(VID_n, N, TMSI_o, \langle VID_n \rangle AT_o)$

(A11) VLR_n **believes** VLR_o **controls** $(N, \{IMSI, MS \overset{K_c}{\leftrightarrow} VLR_o\}K_n, \langle VID_o \rangle AT_o)$

(A12) VLR_n **believes** VLR_o **controls** $\{IMSI, MS \overset{K_c}{\leftrightarrow} VLR_o\}K_n$

(A13) MS **believes** VLR_n **controls** $(AT_n, TMSI_n, \langle VID_o \rangle AT_o, MS \overset{K_c'}{\leftrightarrow} VLR_n)$

A**)** After message 1

No deduction is derived because MS and VLR_n do not share any key or information.

B) After message 2

1. By (P2), (A1) and (M2) imply
   VLR_o **believes** VLR_n **said** $(VID_n, N, TMSI_o, <VID_n>_{AT_o})$. (d1)

2. By (P11), (A7) implies
   VLR_o **believes fresh**$(<VID_n>_{AT_o})$ and (d2)
   VLR_o **believes fresh**$(VID_n, N, TMSI_o, <VID_n>_{AT_o})$. (d3)

3. By (P4), (d1) and (d2) imply
   VLR_o **believes** VLR_n **believes** $(VID_n, N, TMSI_o, <VID_n>_{AT_o})$. (d4)

4. By (P5), (A10) and (d4) imply
   VLR_o **believes** $(VID_n, N, TMSI_o, <VID_n>_{AT_o})$. (d5)

C) After message 3

1. By (P10), (A2) and (M3) imply
   VLRn **sees** $(N, \{IMSI, MS \overset{K_r}{\longleftrightarrow} VLR_o\}K_n, <VID_o>_{AT_o})$. (d6)

2. By (P2), (A2) and (M3) imply
   VLR_n **believes** VLR_o **said** $(N, \{IMSI, MS \overset{K_r}{\longleftrightarrow} VLR_o\}K_n, <VID_o>_{AT_o})$. (d7)

3. By (P11), (A9) implies
   VLR_n **believes fresh**$(N, \{IMSI, MS \overset{K_r}{\longleftrightarrow} VLR_o\}K_n, <VID_o>_{AT_o})$. (d8)

4. By (P4), (d7) and (d8) imply
   VLR_n **believes** VLR_o **believes** $(N, \{IMSI, MS \overset{K_r}{\longleftrightarrow} VLR_o\}K_n, <VID_o>_{AT_o})$. (d9)

5. By (P5), (A11) and (d9) imply
   VLR_n **believes** $(N, \{IMSI, MS \overset{K_r}{\longleftrightarrow} VLR_o\}K_n, <VID_o>_{AT_o})$. (d10)

6. By (P5), (A12) and (d9) imply
   VLR_n **believes** $\{IMSI, MS \overset{Kc}{\longleftrightarrow} VLR_O\}K_n$. (d11)

7. By (P9), (A3) ,
   VLR_n **sees** $(IMSI, MS \overset{Kc}{\longleftrightarrow} VLR_O)$. (d12)

8. (d11) and (d12) imply
   VLR_n **believes** $(IMSI, MS \overset{Kc}{\longleftrightarrow} VLR_O)$. (d13)

9. By (P8), (d13) and $\{Seed\}K_c$ that received in (M1) imply
   VLR_n **sees** *Seed*. (d14)

D) After message 4

1. By (P1), (A6) and (M4) imply
   MS **believes** VLR_n **said** $(AT_n, TMSI_n, <VID_O>_{ATo}, MS \overset{Kc'}{\longleftrightarrow} VLR_n)$. (d15)

2. By (P11), (A8) implies
   MS **believes fresh**$(<VID_O>_{ATo})$ and (d16)
   MS **believes fresh**$(AT_n, TMSI_n, <VID_O>_{ATo}, MS \overset{Kc'}{\longleftrightarrow} VLR_n)$. (d17)

3. By (P4), (d15) and (d17) imply
   MS **believes** VLR_n **believes** $(AT_n, TMSI_n, <VID_O>_{ATo}, MS \overset{Kc'}{\longleftrightarrow} VLR_n)$. (d18)

4. By (P5), (A13) and (d18) imply
   MS **believes** $(AT_n, TMSI_n, <VID_O>_{ATo}, MS \overset{Kc'}{\longleftrightarrow} VLR_n)$. (d19)