WIRELESS COMMUNICATIONS AND MOBILE COMPUTING
Wirel. Commun. Mob. Comput. 2005; **5**:665–678
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/wcm.335

Integration of authentication and mobility management in third generation and WLAN data networks

Wenye Wang*,†, Wei Liang and Avesh K. Agarwal
Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27613, U.S.A.

Summary

The successful deployment of wireless local area networks (WLAN) for high speed data transmission and cellular systems for wide coverage and global roaming has emerged to be a complementary platform for wireless data communications. In order to fully exploit potentials in 3G/WLAN integration, authentication of roaming users crossing different networks, must be coupled with mobility management, which is a challenging, yet not resolved issue. The focus of this paper is on state-of-art solutions to Wi-Fi and cellular networks based on IP infrastructure. Moreover, we introduce a new authentication architecture for fast authentication during inter-networking handoff and large-scale heterogeneous networks. We show that the new architecture can reduce authentication latency significantly and be adaptive to user mobility and traffic. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS: authentication; mobility management; 3G systems; wireless local area networks (WLANs)

1. Introduction

Wireless networks have evolved into a heterogeneous collection of network infrastructures providing a wide variety of options for user access such as Wi-Fi and cellular networks. Based on Wi-Fi technology, wireless local area networks (WLANs) have demonstrated an exceptional success in recent years because of their high speed data transmission and flexible deployment. Mobile users can easily access the internet through a laptop or a portable digital assistant (PDA) with embedded or removable 802.11 cards, thus experiencing data communications over the IP backbones. However, WLAN systems are not capable of providing mobility and roaming support due to local authorization and registration. This limit is complemented perfectly by established cellular networks on which many users have depended for universal roaming. On the other hand, even with the advances in general packet radio service (GPRS) as 2.5G systems, as well as universal mobile telecommunication systems (UMTS) specified by third generation partnership project (3GPP) and cdma2000 specified by 3GPP2, low transmission rate and expensive data service have made 3G cellular networks not preferable for mobile data applications compared to WLAN systems [1].

Thus, broadband wireless access is promising and achievable with complementary Wi-Fi and cellular technologies because of their strength in different perspectives. The missing ingredient is ubiquitous mobility support that allows users to access two network domains. One of the main challenging issues in the integration or interworking of 3G/WLAN systems is the authentication and mobility management.

*Correspondence to: Wenye Wang, Department of Electrical and Computer Engineering, North Carolina State University, Raleigh, NC 27613, U.S.A.
†E-mail: wwang@eos.ncsu.edu

Authentication requires that a network, either a 3G system or a WLAN authenticate users as they claim who they are. However, a network cannot authorize an unknown user without having users' identity and trustworthy records. Moreover, authentication is always prior to location registration and service delivery, which are two operations in mobility management. Therefore, authentication must be considered in support of universal roaming. In particular, due to the open medium of wireless and the mobility of roaming terminals, along with the use of IP backbone networks, sensitive data information and user identity, even wireless network itself, are vulnerable to attack. Although security issues can occur in both wired and wireless networks, risks in wireless domain are greater than those in wired networks due to weaknesses in wireless protocols [2]. The increasing concerns about security and universal access, therefore, necessitates the integration of authentication and mobility management.

There are many security concerns in 3G/WLAN interworking, such as user identity privacy, data integrity and confidentiality as described in Reference [1]. In this paper, we focus on the issue of authentication of mobile users in support of mobility. Authentication is inherently a security technique, which is designed to protect networks against acceptance of a fraudulent transmission by establishing the validity of a transmission, or an originator. During the authentication process, a user must provide verifiable credentials. When a mobile node requests service from a network other than its home network from which it subscribes the service, it must provide sufficient individual information for authorization and register its location to the home network for subsequent service. This process of authentication and registration plays a very important role in protecting the confidentiality and integrity of wireless networks through denying an unauthorized transmission and preventing intrusions [3–6].

Therefore, authentication has a great impact on network security and mobility management in wireless data networks. First, authentication is aimed at ensuring network resources are used by authorized users, which is a security mechanism to prevent resources from any illegal use or damage. Second, authentication involves negotiation of credentials for secure communications. Credentials are usually the records or identification for attesting the truth of certain stated facts such as users identity. While the major purpose of authentication in wireless networks is to authorize networking access, it also has significant influence on the quality of on-going service because authentication may induce signaling overhead. The delay caused by the authentication may increase packet loss and even reduce throughput. All of these factors will degrade the quality of service (QoS). Therefore, authentication and mobility support cannot be segregated and must be considered together as an integral in 3G/WLAN systems.

The rest of the paper is organized as follows. We introduce a generic authentication architecture and design concerns in Subsection 2.1. Then we present an overview of up-to-date authentication solutions to 3G and WLAN systems in Section 2. A new authentication scheme developed for inter-domain roaming is described in Section 3. Finally, we conclude the paper and discuss future challenges in Section 4.

2. Authentication and Mobility Support in 3G/WLAN Systems

Authentication can be performed during the process of registration in that the system needs to determine the current location of a mobile client for service delivery given available information stored in system databases. Location registration and service delivery are critical to mobility support and QoS provisions. Therefore, authentication and authorization mechanisms must be designed in combination with mobility management, especially for inter-domain roaming. In this section, we describe authentication in UMTS/cdma2000 systems, WLANs, and 3G/WLAN integration.

2.1. System Architecture and Design Issues

Mobility support in wireless networks has been researched extensively in the past decade [7]. Especially for UMTS systems, the mobility management is very similar to 2G systems because core network is separate from radio access network (RAN). The authorization of user identity is implemented through registration, a process for validating users records in centralized databases home location register (HLR) and authentication center (AuC). Registration is usually discussed in the context of mobility management in both cellular and Mobile IP networks [7–9]. Mobile IP is proposed by IETF as a network layer protocol to facilitate macro mobility in WLAN systems, which is not only recommended for WLAN systems, but more importantly, it is considered as a networking protocol by 3GPP2 for 3G cellular

systems [10]. Thus, Mobile IP with authentication, authorization, and accounting (AAA) extensions is developed to provide secure communications [11,12].

Previous efforts on mobility management for seamless roaming have been mainly on the architectural design and vertical handoff between 3G and WLAN systems [13,14]. To be more specific, the integration of authentication and mobility management is focused on authentication architecture and signaling interoperation to reduce handoff latency [15–19]. However, authentication in wireless networks brings about new challenges and design considerations, which go far beyond conventional security solutions for wired networks and mobility management for wireless networks. These design considerations, which are the driving force for developing new solutions rather than using existing mobility support technologies, can be summarized as follows:

- *Architecture*: In distributed or heterogeneous wireless networks, there maybe many mobile users roaming among network domains with different technical specifications, signaling formats, identity authorization credentials, network protocols, and so on. The coverage of each autonomous network varies from tens of meters in WLANs to tens of kilometers in 3G systems, depending on the design and architecture of each network. Inside an autonomous network, there is an authentication server (AS), which is a centralized server for authenticating users within the coverage of the network as shown in Figure 1.

  An access router (AR) can either be a radio network subsystem (RNS) in UMTS or an access point (AP) in WLAN. Each mobile node, the device used by mobile users, has a permanent authentication association with the AS in its home network from which a mobile user subscribes service. Authentication servers trust each other based on security associations (SAs), which is a one-way relationship between a sender and a receiver for security service defined in IP security (IPSec). Authentication architecture, which inter-connects authentication servers, has a great impact on performance of authentication mechanisms and protocols for heterogeneous environments because user credentials will be transmitted through the authentication architecture; thus, it is an important design issue in 3G/WLAN integration.
- *Scalability*: Due to traffic volume as well as service coverage, the number of mobile nodes in each network varies from less than ten in an WLAN to thousands of users in 3G networks. Also, the roaming pattern from one network to another may be different in time periods, for example, during peak hours, many people may drive from home to work place and vice versa, so called 'high-mobility,' whereas lower mobility may be the case after business hours for the same group of people in the same area. Therefore, authentication solutions developed for wireless networks must be scalable enough to adapt to various user densities and roaming patterns.
- *Security service*: In order to protect information secrecy, data integrity, and resource availability for users, security architecture and protocols are designed. Information secrecy means to prevent improper disclosure of information; data integrity is concerned about improper modification of data; and availability is to prevent improper denial of service provided by the system. As an important approach

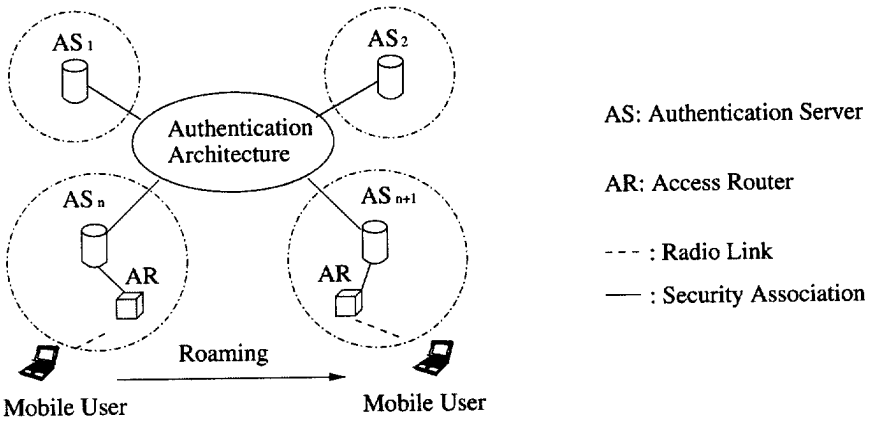


Fig. 1. System model of authentication in wireless networks.

in providing security for network service, authentication is a critical part in maintaining these service. During the authentication process, especially for inter-domain roaming, a mobile node needs to negotiate cryptographic algorithms with an authentication server and obtains keys for subsequent data transmission. In addition, authentication can mitigate the attack of denial of service in which bandwidth is overwhelmed by useless packets sent into the network on purpose to block legitimate network traffic. Complicated authentication protocols can use security associations for each connection segment and enable data encryption throughout the entire session of data delivery.

- *Overhead*: Since authentication is a necessary procedure before actual transmissions over radio channels, it unavoidably introduces overhead into the network. An authentication process includes negotiation of encryption/decryption algorithms, encryption/decryption of messages, transmission of messages, and credential verification. Therefore, the overhead includes signaling, verification, and transmission cost used to exchange credentials between mobile clients, home networks, and authentication servers. Also, encryption/decryption algorithms require strong processing and computation capabilities, which must be considered for power-limited mobile devices. Meanwhile, authentication delay and inefficient bandwidth utilization may be caused by authentication signaling and other processes such as encryption/decryption for high mobility nodes.

2.2. Authentication Architecture and Protocols in 3G Systems

UMTS are envisioned to provide always-on, wide-area connectivity with relatively low data rates to users with high mobility, whereas WLANs offer much higher data rates to users with low mobility over smaller areas. In order to enable wide-area roaming capability for data networks, the security architecture of UMTS has been improved compared to 2G systems. For both data and voice services, an RNS communicates with mobile stations equipped with UMTS subscriber identity module (USIM), which is controlled by a radio network controller (RNC).

Authentication in UMTS is defined as part of Network Access Security in UMTS security architecture [20]. Network access security defines a set of security features that provides users with secure access to 3G service and prevents the attacks on the (radio) access link in particular. They are related to entity authentication in two aspects: user authentication and network authentication. User authentication means that the serving network corroborates the identity of the user; whereas, network authentication concerns how the user corroborates that he/she is connected to a serving network that is authorized by the user's home environment to provide him/her service. The set of security features that enables nodes in the provider domain to securely exchange signaling data and prevents the attacks on wired networks is defined as Network Domain Security. Security features that protect communications between users and mobile stations inside the coverage of service providers are defined as User Domain Security and Application Domain Security, respectively.

To achieve these objectives, it is assumed that the entity authentication should occur at each connection set-up between the user and the network. Two mechanisms have been included for entity authentication: an authentication vector delivered by the user's home environment (HE) to the serving network and a local authentication mechanism. Using an authentication vector between HE and serving network is like the inter-domain authentication procedure since a local authority does not have the credentials of the mobile node and must request the authentication approval from the HE of the mobile node. A local authentication mechanism is like the intra-domain authentication, in which a local authority has enough information to authenticate a mobile node, and it uses integrity key established between a user and its serving network during the previous execution of the authentication and keys establishment procedure. However, when authentication vector is used, the trust relationship between the local authentication entity, a visitor location register (VLR) that may be collocated with serving GPRS service node (SGSN), and an HE is always assumed to exist. The mutual authentication procedure between the USIM and the SGSN/VLR is called UMTS authentication and key agreement (AKA), which has been introduced in a good level of details in Reference [21].

In addition to UMTS, another important standard for 3G systems is cdma2000, which is specified by 3GPP2. The major difference between cdma2000 and UMTS is that cdma2000 systems support AAA for Mobile IP. Thus, for integration with WLANs, it is required that 802.11 gateways support Mobile IP functionalities [18]. The authentication and key

agreement for cdma2000 is very similar to UMTS, except specific algorithms used for random number generation, which is standardized in cdma2000, but not included in UMTS [22]. Also, key generation algorithms are slightly different, including encryption and decryption algorithms, which subsequently affect the authentication vectors for re-authentication and tracking. More details of access security in cdma2000 can be found in Reference [22].

2.3. Mobile IP with AAA Extensions

Standards for terminal mobility over the internet, Mobile IP enables WLAN roaming and promises to enable terminals to move from one subnetwork to another, as packets are being sent, without interrupting this process [23]. Mobile IP is of particular important because it is the basis for the 3G/WLAN integration [1,18].

2.3.1. *Authentication in Mobile IP*

In basic Mobile IP architecture, an authentication extension (AE) is defined for registration messages, which consists of a security parameter index (SPI) and an authenticator calculated by using a keyed-hash function. It is designed to provide entity authentication which protects home agent (HA) and mobile nodes (MNs) against replay attacks, either by timestamps or by nonces, a random number. However, it does not provide data protection between foreign agent (FA) and the MN. The protocol assumes that security associations between FAs and HAs already exist. This assumption requires a huge effort to manage, and it may not be effective for scaling up networks. To strengthen relay protection, Mobile IPv4 challenge/response extensions (MICRE) is developed. This protocol provides replay protection for all messages exchanged with Mobile IP protocol by defining two new types of message extensions: challenge extension for FA advertisement messages and mobile challenge response extension for registration messages [24]. When an MN wants to authenticate itself, it must send an authentication request message with the challenge value received from the FA advertisement. By checking the challenge value to see if it has already been used, the FA can avoid a malicious replay attack from an MN.

The verification of the challenge value depends on the security association between the MN and its HA, while the security association between the FA and the MN may or may not exist. A secure scalable authentication (SSA) is aimed to provide Mobile IP with a strong, scalable authentication mechanism based on public key cryptography [25]. When an MN is moving close to an FA, it receives an advertisement with authentication extension and certificate extension broadcast by the FA. The MN then extracts and validates the certificate with a public key issued by a certificate authority. After the verification, the MN uses the public key of the FA from the certificate to verify the digital signature in the FA authentication extension, which is created using the FA's private key. Then, the MN obtains the secret key of the FA; thus, the communication between the MN and the FA can be protected. The secret key exchanged between the FA and the HA is generated with the same method. This scheme is able to provide security protection between an FA and an MN, and between an FA and HA for large-scale networks.

In order to expedite the authentication process for inter-domain roaming, a rapid authentication (RA) for mobile IP protocol is designed to provide Mobile IP nodes and agents with necessary keys and information needed to establish mobility security associations within a foreign network [26]. By deploying a key distribution center (KDC) in a foreign network, this protocol is able to speed up the subsequent authentications for those mobile nodes that have already registered in a foreign domain. Upon entering a foreign domain, an MN first registers with the KDC, optionally using public key to obtain a secret key used for subsequent authentications. Therefore, authentication latency for intra-domain is decreased, while the MNs are required to differentiate an initial authentication from a subsequent authentication request. In summary, this scheme can reduce latency and computational burden introduced by public-key based, key management protocols in Mobile IP networks.

2.3.2. *AAA framework and protocol*

Authentication process may incur a significant delay due to transmitting credentials, which is critical to MNs that are far away from their home networks. Often time, it is necessary for access routers to keep track of pending requests while the local authority contacts the appropriate external authority. With these requirements, RFC 2977 further provides authentication architectures for Mobile IP networks with AAA extensions [19]. In the proposed basic architecture shown in Figure 2(A), each local AAA server (AAAL) should share security association with a home AAA

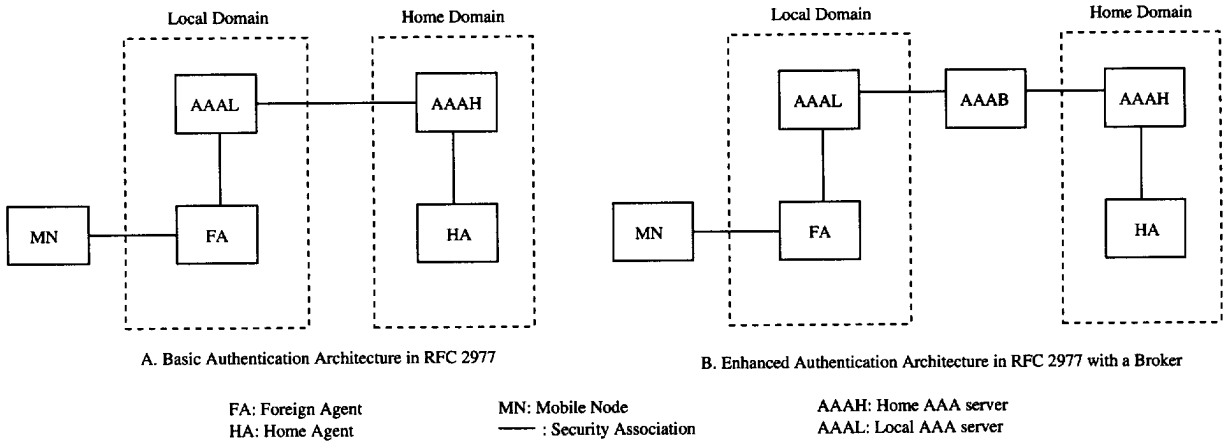


Fig. 2. Basic and enhanced authentication architecture in Mobile IP.

server (AAAH) of a roaming MN in current domain, so that the AAAL can securely transmit MNs' credentials. This configuration, however, may cause a quadratic growth in the number of trust relationships, as the number of AAA authorities (AAAL and AAAH) increases, which is a problem identified by roamops working group [19,27]. Using brokers is a possible solution to the scalability problems associated with requiring direct business/roaming relationships between every pair of administrative domains. In order to provide scalable networks in many service providers and large numbers of private networks, multiple layers of brokers should be supported like the broker model described in Figure 2(B).

Integrity or privacy of information between the home and serving domains may be achieved by either hop-by-hop security associations or end-to-end security associations established with the help of the broker infrastructure. A broker may play the role of a proxy between two administrative domains, which have security associations with the broker, and be able to relay AAA messages back and forth securely [28]. It may also enable roaming in two domains with which it has associations, but domains themselves do not have a direct association for carrying messages. Though this mechanism may reduce latency in the transmission of messages between domains after the broker has completed its involvement, there may be a large amount of overhead messages as a result of additional copies of authorization and accounting to the brokers. There may also be additional latency for the initial access to the network, especially when a new security association needs to be created between AAALs and AAAHs. These delays may become important factors for latency-critical applications such as voice over IP.

DIAMETER protocol is published by IETF as a practical solution for AAA in Mobile IP networks [3,29]. A DIAMETER server is defined as an authority center, which is able to authenticate, authorize, and collect accounting information for Mobile IPv4 service rendered to a mobile node. The DIAMETER base protocol is intended to provide an AAA framework for applications, such as network access or IP mobility, and work in both local AAA and roaming situations. Nowadays, DIAMETER is being deployed as a more flexible successor to the widely-deployed RADIUS protocol for AAA. Security is enhanced between AR and either HA or MNs during AAA and registration process [29]. With these advantages, the DIAMETER protocol is recommended to be the authentication standard in Mobile IP networks.

2.4. Authentication for Interworking 3G/WLAN

2.4.1. *Authentication servers and proxy*

For interworking of 3G/WLAN, the main issue is how to validate a user's credentials that are kept in authentication servers. In [28], a 3G authentication server is proposed as a new functional component in 3G systems, which behaves as a gateway between WLANs and a 3G systems to support interworking. The AAA server will terminates all AAA signaling from WLANs and route to other components in 3G systems in which case it is called AAA proxy. The counterpart in WLANs is WLAN AAA proxy, which routes AAA messages to 3G servers. WLANs are identified based on network address identifier (NAI), which is sent by mobile nodes in their access requests.

Signaling and interfaces for different scenarios can be found in the paper, including control messages between AAA proxy in WLANs and 3G systems.

A similar idea that uses a security gateway (SGW) instead of authentication proxies to integrates IP mobility and security management together is proposed in Reference [30]. IPSec tunnel mode is enabled between the SGW and its MN by which the SGW sets up an SA for each MN in its network. The MN maintains a single SA between itself and the SGW in its home network, whether the MN is in its home subnet or it roams to a foreign subnet. The HA is only responsible for Mobile IP registration and relaying packets to MN's CoA. The MN is protected by the IPSec tunnel between the SGW and MN. While the MN is roaming, there is no window of clear data transmission over the wireless link, and there is no need to re-establish an IPSec tunnel between the SGW and MN. Therefore, this scheme provides a secure communication segment between an roaming MN and its HA without requiring that the foreign network participate in the process.

2.4.2. *AAA and inter-domain roaming*

The most important objective of Mobile IP with AAA extension is to support mobility for inter-domain roaming, especially heterogeneous. When a mobile object moves out of the coverage of its home network, the network address of this mobile object, such as an IP session address, is useless. To efficiently solve these problems, a common architecture for handling inter-system terminal mobility is developed by the authors with Mobile IP authentication architecture as shown in Figure 3 [17]. In this architecture, mobility support is integrated with AAA functions through carefully designed signaling messages. In other words, before an FA confirms the registration of a visiting node, it contacts a foreign AAA server with an access request message. Therefore, the AAA functions are completed along with the registration. Since no separate signaling is needed for authentication and registration, the number of packets exchanged is reduced.

In Reference [18], two architectures for 802.11 and 3G integration are proposed: tightly-coupled and loosely-coupled interworking. In a tightly-coupled architecture, the 802.11 network would emulate functions in 3G; that is the 802.11 hides all details for the 3G networks and appears as either a packet control function (PCF) in cdma200, or as an SGSN to the core network. As a result, two domains would share the same authentication server for billing and accounting. To avoid the use of a common authentication mechanism, based on USIM for 3G or removable user identity module (R-UIM) cards for authentication on WLANs, a loosely-coupled architecture can be used in which two domains can use their individual authentication solutions.

The interworking of authentication requires that a new 802.11 gateway support Mobile IP functionalities and new AAA servers in 3G networks. Consequently 3G networks can collect records of users in 802.11 through AAA service in two different network domains. This would not be a serious problem for interworking cdma2000 and WLANs because cdma2000 supports Mobile IP and AAA, whereas it requires additional specifications for UMTS networks. The main component in 802.11 gateway resides in so called IOWA gateway in which a RADIUS AAA server is used. The server provides authentication

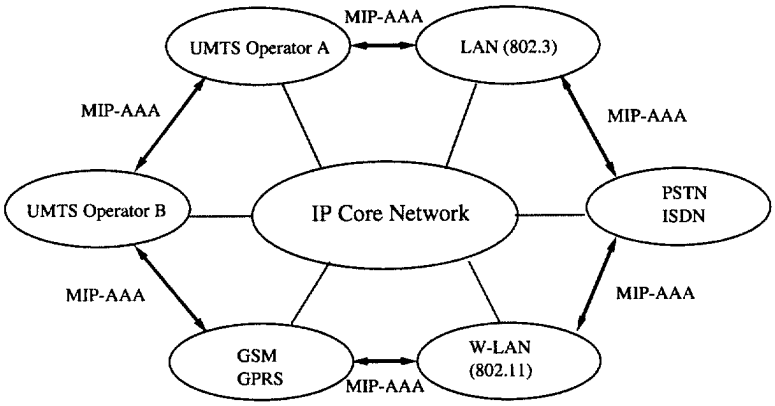


Fig. 3. Mobility support in heterogeneous networks.

service to 802.11 domain and also maintain an authentication agreement with 3G systems.

2.4.3. Authentication and access control

To accomplish secure transmissions in wireless networks, both authentication architecture and protocols are of importance. The architecture determines the platform for implementing various protocols, while the protocols on the architecture are necessary to perform authentication procedure and integrate access control and security mechanism. Therefore, the integrated design of authentication architecture and protocols is a challenging issue, which nevertheless received few attentions before. An access control architecture for wireless IPv6 networks is proposed [31], in which gateway access router connects the access network with the public internet or a private Intranet, for example campus network. Authentication protocol for this architecture uses the standard IPSec encryption header [32] to encrypt the secret credentials of other users or a session key in order to avoid the need for a secure key exchange mechanism and a special protocol extension. The proposed protocol is intended to run over IPv6 with public-key management to maintain scalability and security between mobile users and their home network.

3. A New Authentication Scheme for Wireless IP Networks

As aforementioned, there are not much research on the integration of authentication and mobility management by taking system performance into consideration. One reason is that a massive amount of effort is needed to complete an architecture design as well as a set of protocols. Another reason is that the current existing architectures and protocols are able to provide secure transmission over wireless networks, though not satisfactory. The challenge is how to relate secure transmission with overall system performance in terms of scalability and quality of service since the ultimate goal of wireless networks is to deliver a high quality service to customers. We believe that this issue deserves an in-depth study to improve secure communications in 3G/WLAN integration. In this section, we propose a new authentication architecture on which a dynamic control scheme for efficient authentication can be applied [33]. The new architecture is composed of multiple, licensed authentication centers (LACs) in different wireless network domains. Each domain can be either a WLAN, a wireless personal area network (WPAN) or a 3G system regardless of technical specifications in each network.

3.1. Security Association

The motivation behind our scheme is that existing solutions of authentication architectures cannot satisfy various requirements of security and service in wireless networks because of the use of static security associations (SSA). A security association SA is defined in IPSec, which describes how two or more entities will utilize security services to communicate securely. For example, an IPSec SA defines the encryption algorithm, an authentication algorithm, and the shared session key to be used during the IPSec connection. A static SA (SSA) is defined as an SA that does not change for a long time, for example one month or throughout the period of service enrollment. In addition, current architectures do not support control policies for dynamic authentication management based on user mobility. An SA is a one-way trust relationship between communicators, which affords security service on the traffic with parameters of encryption/decryption method, shared key and lifetime. A flexible SA (FSA), in contrast to SSA, is an SA created on demand to provide a temporary security service for one communication session. An FSA can be established by a four-way handshake protocol in transport layer security (TLS) and changed by adjusting its parameters [34]. In other words, the trust relationship can be removed when a time threshold is reached or the service is terminated.

The advantage of applying FSA is that it can reduce the number of SAs between wireless networks, which has been identified as an important factor of security and manageability in References [27,35]. A huge number of SAs impose great effort to manage the SAs and keep them safe at ACs, which causes manageability problems in wireless networks [27]. Moreover, time to establish an SA is extensive [36], which may cause long authentication delays and further deteriorate the bandwidth efficiency to systems. In addition, the longer an SSA exists, the more vulnerable it becomes to potential attack on it, which is more serious in the open medium of wireless networks. For example, an SA, which is applied at wired equivalent privacy (WEP) in 802.11b and 802.1x between an MN and an AR, can be cracked within 15 min for its constant, small size, encryption key with small size [37].

3.2. Authentication Architecture and Licensed Authentication Center

An LAC is an authority that is responsible for authenticating MNs applying for service in a wireless network. Each wireless network only has one LAC with two functions. One function is to process intra-domain authentication for subscribers in the current network, and the other function is to authenticate inter-domain roaming MNs from other networks. For the intra-domain authentication, because each MN in the home network shares an SSA with the LAC, it can be processed in a short time period by using DIAMETER or RADIUS to authenticate local users SSA [3]. For the inter-domain authentication, the LAC can implement any control scheme to dynamically adjust the lifetime of an SA based on QoS requirements and mobility patterns.

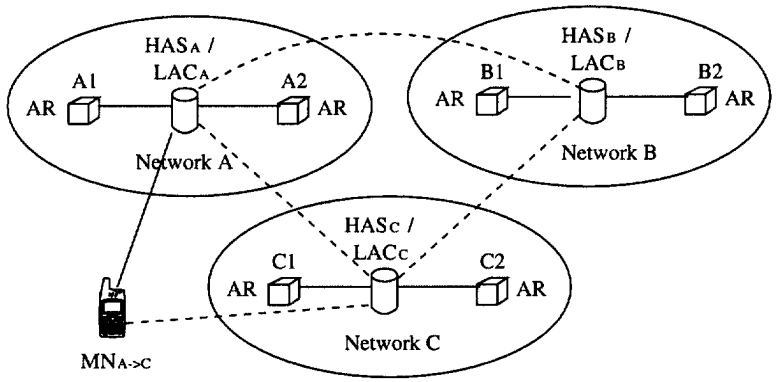
The proposed authentication architecture is shown in Figure 4, in which every AR in a wireless network shares an SSA with the LAC. All of the MNs subscribing to the service in a wireless network are trusted by a HAS in the network, which is also an LAC for visiting MNs. The LACs are connected to each other by FSAs. In this example, there are three wireless networks A, B, and C. In network A, LAC_A is a local authentication authority. A1 and A2 are two ARs in wireless network A. A1 and A2 share SSAs with LAC_A. Network B is another wireless network, in which LAC_B is also a local authentication authority. B1 and B2 are two ARs in wireless network B, which share SSAs with LAC_B. In the third network C, a local authentication authority, LAC_C is responsible for authentications. LAC_C is connected with two ARs, C1 and C2, by SSAs.

We define the average number of SAs between LACs in a period of time as N_N shown as follows, which represents the effectiveness of using FSAs between LACs:

$$\begin{aligned} N_N &= \lim_{T_p \to \infty} \frac{\sum_{i=0}^{M_N-1} \sum_{j=0, j\neq i}^{M_N-1} \int_0^{T_p} n_{ij}(t)\mathrm{d}t}{T_p} \\ &= \sum_{i=0}^{M_N-1} \sum_{j=0, j\neq i}^{M_N-1} t_{ij}A_{ij} \\ &\leq \sum_{i=0}^{M_N-1} \sum_{j=0, j\neq i}^{M_N-1} t_m A_m = t_m A_m M_N(M_N - 1) \end{aligned} \tag{1}$$

where M_N is the number of LACs in our architecture, T_p is an observation time within which we count the number of SAs, i and j are the indexes of the LACs in a wireless network. t_m is the maximal lifetime of FSA between any two wireless networks, A_m is the maximal arrival rate of inter-domain authentication requests. Because the unit of t_m is milliseconds and A_m in a second is also small in reality, the condition $t_m A_m \ll 1$ is satisfied in most cases. If the FSA is alive, the value of $n_{ij}(t)$ is 1 for there only exists one FSA from the LAC i to the LAC j, and the value of $n_{ij}(t)$ is 0, if the FSA does not exist. Then, the actual number of SAs from the LAC i to the LAC j, $n_{ij}(t)$, is shown in Figure 5(a).

The main advantages of the proposed architecture compared to existing solutions can be summarized as follows.



LAC: Licensed Authentication Center
AR: Access Router
MNA->C: A Mobile Node is moving from home network A to network C
- - - - : Flexible Security Association
——— : Static Security Association

Fig. 4. Authentication architecture and licensed authentication centers.

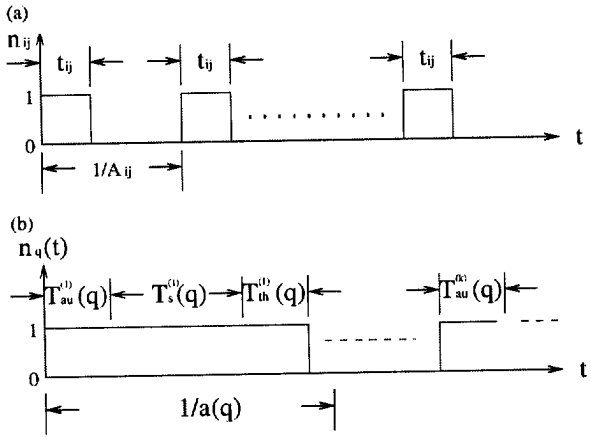


Fig. 5. Flexible security associations. (a) Number of SAs from LAC i to LAC j. (b) Number of SAs between visiting MN q and LAC.

- No gateway routers are required in existing solutions, gateway routers are proposed to deal with inter-working of different systems, which is a reasonable solution. However, we find out in our experimental study, that the entity, which is able to act as a gateway router, has been a major concern. It involves many changes in the protocol stack as well as interfaces, which is not trivial.
- No enforcement of Mobile IP: As opposed to the generic inter-domain authentication in [17,26], Mobile IP and AAA are used to provide authentication among different systems. In our architecture, an on-demand security association will be established based on user mobility and service requirement, rather than going through the core network in cellular systems and home agent in Mobile IP.
- No bottleneck in local authentication servers: In our scheme, the number of security associations depends on authentication requests per user-basis as well as user movement pattern. Security associations can be released from time to time. Therefore, this architecture is adaptive to distributed networking environments.

3.3. Authentication Process

For inter-domain roaming users, the LAC manages two kinds of FSAs to process authentication requests: one between the visiting MN and the LAC and the other between the LAC and HAS of the visiting user. The HAS can also be an LAC for intra-domain authentication. The operation processes of LACs in handling different roaming scenarios are shown in Figure 6 in three cases depending on the existence of security associations between the MN, the LAC, and the HAS of roaming MNs.

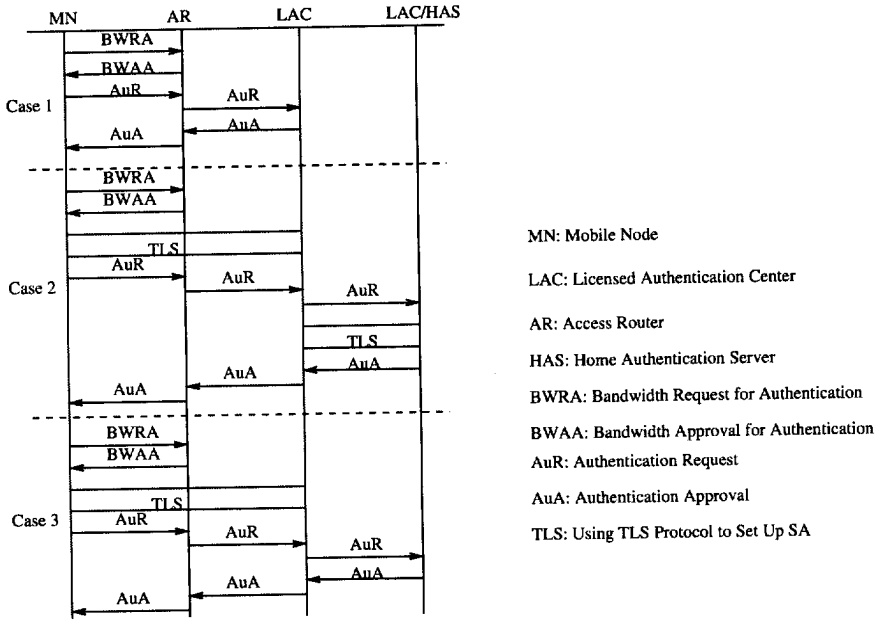


Fig. 6. Authentication for inter-domain roaming.

- *Case 1*: If an MN has previously been registered and authenticated, the SA used between the visiting MN and the local LAC may not be expired. Then, in the subsequent communications, the MN can reuse this SA to authenticate itself in current domain by indicating the SA and sending the encrypted credentials of MN to the LAC. After the LAC decrypts and verifies the secrete data, if the plaintext is identical with the data stored in the LAC, the MN will be authenticated, after which the LAC generates, encrypts, and transfers a session key for the MN and the AR serving for the MN. In this case, the time interval between the end of service time and next authentication request is important because the MN may reuse the previous SA.
- *Case 2*: If an MN is a first-time visitor or its previously used SA with the local LAC has expired, the MN needs to establish an SA with the local LAC first. The LAC must establish an FSA with the HAS of the visiting MN to authenticate the MN. With the establishment of two FSAs, the credentials of the MN are encrypted and sent to the HAS for authentication. If this request is approved and returned to the local LAC, the LAC will generate a session key for the MN and the AR serving for the MN.
- *Case 3*: An MN from a foreign domain does not share an SA with the local LAC, but the local LAC has the SA with the HAS of the MN because of previous authentications. In this case, the MN needs to establish an SA with the local LAC. Then, the credentials of the MN are encrypted and sent to the HAS for authentication. If this request is approved and returned to the local LAC, the LAC then generates a session key for the MN and the AR serving for the MN.

With the designed authentication architecture, intelligent control schemes can be exerted for efficient authentication by managing flexible security associations for inter-domain roaming between 3G/WLAN integrated environments. For example, the objective function can be an optimal threshold time based on handoff delay requirements and mobility patterns. So the new authentication architecture is adaptive and flexible for inter-domain authentication. Many design factors can be considered in control schemes, such as bandwidth efficiency, delay, scalability, and so on. For example, the objective function can be an optimal threshold time based on handoff delay requirements and mobility patterns. We use a simple control scheme with the aim to reduce the average number of SAs and authentication delay in this context because it is a scalability issue as described in Subsection 3.2.

3.4. Location Registration and Service Delivery

The proposed authentication architecture and process can be easily extended to location registration and service delivery for mobility management. In the operation of location registration, mobile users update their current location to their home networks from which they subscribe their services.

- *Location registration*: If a user has already registered with a network, which is the *Case 1* described in the previous section, that means, within the lifetime of FSA, there is no need to register or update with the user's home network again. The latest records of this user's location is the current network ID stored in the user's HLR or home agent, which is completed when the user registered with the new network. If the mobile node does not have an existing SA, or it has not registered with the current network, which are described as *Case 2* and *Case 3* in Figure 6, then locations will be updated during the communications between a local LAC and the LAC or HAS at a mobile user's home network. Therefore, there is no separate process needed for location registration.
- *Service delivery*: When an incoming call arrives to a mobile node, the current network is required to setup the connection for service delivery. In this case, both caller and the callee will be authenticated for their service. From the location registration and authentication process, in fact, the caller is authenticated. Thus, there is no separate process needed to authenticate and authorize caller according to the proposed solution. However, the authentication of callee may still be needed. We assume that service delivery will follow the normal operation in the user's home network, that is, a service request will first be delivered to the user's home network, then the home network forwards this request to the current network serving the callee. The authentication process described in the previous section can be modified by changing the initiator, that is, the LAC will send a request to a mobile user, asking for credentials to confirm that the callee is authorized to take the service. This is very similar to the authentication process for location registration, except for authentication request initiated by the LAC. To avoid repetition the signaling diagram of service delivery is omitted here because it is similar to the process of location registration.

3.5. Numerical Results

We evaluate average authentication latency, bandwidth efficiency, and average number of SAs of the proposed scheme. We define the average authentication latency, T_{av}, as the ratio of the sum of all the authentication latencies of inter-domain authentication requests to the number of these requests. We focus on the inter-domain authentication and compare the results with existing schemes in hierarchical authentication architecture [19], in which hierarchical AAA brokers (AAABs) are trusted by many AAALs to relay credentials. When an AAAB cannot find an SA for two networks, it goes through AAABs in higher levels to find it. With this model, the number of SAs can be reduced. Therefore, the manageability of networks can be maintained. However, the hierarchical architecture of AAABs may take a long time to search upper AAABs during inter-domain authentication, and chaining AAA servers may result in a number of security threats, for example, man-in-middle attack [27].

To improve the bandwidth efficiency during authentication, we develop a dynamic security association control scheme for efficient authentication. The new distributed authentication architecture which utilizes fewer SAs than the hierarchical architecture, and provides great security and independence between different wireless networks as well as scalability.

Two networks are considered in the simulation with two LACs and two ARs in the proposed architecture. For the hierarchical architecture, there is one PAC, two ACs, and two ARs in the networks. Important parameters are defined in Table I. All of the time variables are supposed to be exponentially distributed with mean values. A_{MN} is arrival rate of new inter-domain authentication requests, and T_{int} is a time interval defined as $T_{int} = 1/a - T_{au} - T_s$ according to Figure 5(b). The time variables are obtained from [36,38]. Since the fastest successful crash time of an SA in WEP is around 10 min [37], we use 10 min as the value of $1/\lambda$.

Table I. Simulation Parameters.

Parameters	Values	Parameters	Values
Radio channel BW	3840 (kbps)	$U(f)$	$e^{(k*f)}$
B_s	50 (kbps)	λ (s^{-1})	$\frac{1}{600}$
B_{un}	B_s	k	50
T_s	1 (s)	α	0.1
A_{MN}	0.2 (calls/s)	C	5
MAC access delay time	0.01 (s)	β	0.5
Transmission and propagation time (hierarchical arch)	0.07 (s)	T_V/T_{VH}	0.1 (s)
Transmission and propagation time (proposed arch)	0.05 (s)	T_{int}	10 (s)

In Figure 7, we observe that T_{av} is reduced greatly with the proposed scheme. Compared with T_{av} in hierarchical architecture without the control scheme, the improvement is between 24 and 34% although T_{av} is increasing slowly in the proposed scheme with the increase of A_{AM}. The benefit comes from the control on the establishment of FSAs for visiting MNs. Part of previously authenticated MNs do not need to establish the SA with the LAC during authentication. However, when A_{MN} increases, the number of new visiting MNs is increased. They cannot obtain the benefit from the management of SA between an MN and the LAC in our control scheme. Therefore, T_{av} increases with the increase of A_{MN}.

In Figure 8, the average number of SAs, N_N, is compared between our control scheme on new architecture and the hierarchical architecture without the control scheme. The improvement of N_N with the control scheme is between 70 and 90%. This benefit comes with the small authentication latency and arrival rate of authentication requests between LAC and HAS, which is shown in Equation (1). N_N decreases with the increase of A_{MN} because the increased A_{MN} causes more authentication requests to share an FSA between the LAC and HAS for authentication. Compared to the individual authentication in different time, the share reduces the average number of SAs.

Therefore, the proposed new architecture allows dynamic security association control scheme for

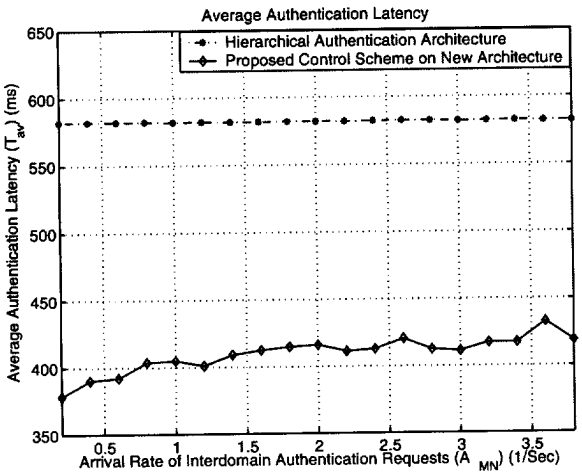


Fig. 7. Authentication latency versus arrival rate of authentication requests.

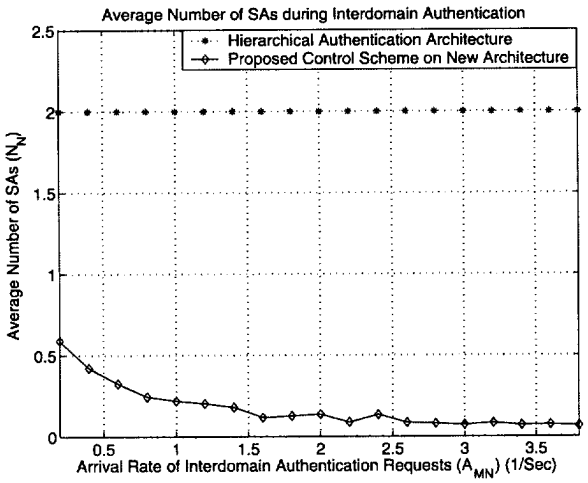


Fig. 8. Average number of SAs versus arrival rate of authentication requests.

efficient authentication for inter-domain roaming, which is applicable to heterogeneous environments such as 3G/WLAN integration. Intelligent management of SAs between different wireless networks, and between visiting mobile devices and local ACs can be enforced and applied based on user mobility and even quality of service. By choosing an optimal threshold time for the SAs based on authentication traffic and residence time in a network domain, our scheme can reduce authentication latency and average number of SAs between different wireless networks, simultaneously.

4. Conclusion

In summary, we first presented a generic authentication architecture and discussed factors that would be considered in new authentication solutions for heterogeneous environments. Then, we provided an overview of state-of-art solutions with respect to authentication issue for 3G/WLAN integration. In addition, we proposed a new architecture for dynamic authentication management which enables control schemes to be applied for efficient authentication and mobility management. Numerical results are provided to demonstrate the effectiveness of the proposed scheme in terms of improving scalability and reducing authentication latency. In the future, wireless services will be based on IP infrastructure for public access; thus there are many opening issues in this field, such as how to integrate QoS and mobility management into authentication protocols because authentication process has a great impact on the QoS and how to design security policy to be adaptive to customized applications.

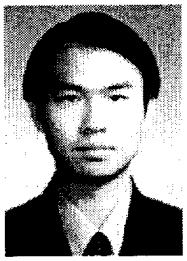
References

1. Køien G, Haslestad T. Security aspects of 3G-WLAN interworking. *IEEE Communications Magazine* 2003; **41**: 82–88.
2. Karygiannis T, Owens L. Wireless Network Security 802.11, Bluetooth and Handheld Devices. *NIST Special Publications* 800–848, November 2002.
3. Calhoun P, Loughney J, Guttman E, Zorn G, Arkko J. Diameter Base Protocol. *draft-ietf-aaa-diameter-17.txt*, December 2002.
4. Jacobs S. Security and authenticaiton in mobile IP. In *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, Vol. 3, September 1999; pp. 1103–1108.
5. Perkins C, Calhoun P. Mobile IPv4 challenge/response extensions. *IETF RFC 3012*, November 2000.
6. Xu M, Upadhyaya S. Secure communication in PCS. In *IEEE Vehicular Technology Conference, 2001*, Vol. 3, 2001; pp. 2193–2197.
7. Akyildiz I, McNair J, Ho J, Uzunalioğlu H, Wang W. Mobility management in next-generation wireless systems. In *Proceedings of the IEEE*, Vol. 87, August 1999; pp. 1347–1384.
8. Chiussi F, Khotimsky D, Krishnan S. Mobility management in third-generation all-IP networks. *IEEE Communications Magazine* 2002; **40**: 124–135.
9. Lin Y-B, Yang S-R. A mobility management strategy for GPRS. *IEEE Transactions on Wireless Communications* 2003; **2**: 1178–1188.
10. Patel G, Dennett S. The 3GPP and 3GPP2 movements toward an All-IP mobile network. *IEEE Wireless Communications* 2000; **7**: 62–64.
11. Johnson DB, Perkins C, Arkko J. Mobility support in IPv6. *IETF Internet Draft, draft-ietf-mobileipv6-17.txt*, May 2003.
12. Calhoun P, Loughney J, Zorn G, Arkko J. Diameter base protocol (request for comments 3588). http://www.ietf.org/rfc/rfc3588.txt, September 2003.
13. Misra A, Das S, McAuley A, Das S. Autoconfiguration, registration, and mobility management for pervasive computing. *IEEE Personal Communications* 2001; **8**: 24–31.
14. Ahmavaara K, Haverinen H, Pichina R. Interworking architecture between 3GPP and WLAN Systems. *IEEE Communications Magazine* 2003; **41**: 74–81.
15. Lin P, Lin Y-B, Feng V, Lai Y-C. Gprs-based wlan authentication and auto-configuration. *Computer Communications* 2004; **27**: 739–742.
16. Dell'Uomo L, Scarrone E. The mobility management and authentication/authorization mechanisms in mobile networks beyond 3G. In *IEEE Personal, Indoor and Mobile Radio Communications 2001 12th IEEE International Symposium*, Vol. 1, 2001; pp. c44–c48.
17. Cappiello M, Floris A, Veltri L. Mobility amongst heterogeneous networks with AAA support. In *IEEE ICC 2002*, Vol. 4, 2002; pp. 2064–2069.
18. Buddhikot M, Chandranmenon G, Han S, Lee Y, Miller S, Salgarelli L. Integration of 802.11 and third-generation wireless data networks. In *IEEE INFOCOM'03*, April 2003.
19. Glass S, Hiller T, Jacobs S, Perkins C. Mobile IP authentication, authorization and accounting requirements. *RFC2977*, October 2000.
20. ETSI. UMTS 3G Security architecture. *3GPP TS 33.102 version 4.5.0 Release 4*, December 2002.

21. Køien G. An introduction to access security in UMTS. *IEEE Wireless Communications* 2004; **11**: 8–18.
22. Rose G, Køie G. Access security in cdma2000, including a comparison with umts access security. *IEEE Wireless Communications* 2004; **11**: 19–25.
23. Perkins C. IP mobility support for IPv4. *Request for Comments (RFC) 3220*, January 2002.
24. Perkins C, Calhoun P. Mobile IP challenge/response extensions. *draft-ietf-mobileip-challenge-09.txt*, February 2000.
25. Jacobs S. Mobile IP public key based authentication. *draft-jacobs-mobileip-pki-auth-02.txt*, March 1999.
26. Sanchez L, Troxel G. Rapid authentication for mobile IP. *draft-ietf-mobileip-ra-00.txt (expired)*, November 1997.
27. Aboba B, Vollbrecht J. Proxy chaining and policy implementation in roaming. *RFC2607*, June 1999.
28. Salkintzis A. Interworking techniques and architectures for wlan/3g integration toward 4g mobile data networks. *IEEE Wireless Communications* 2004; **11**: 50–61.
29. Calhoun P, Johansson T, Perkins CE. Diameter base protocol. *IETF AAA Working Group,draft-ietf-aaa-diameter-mobileip-13.txt*, October 2002.
30. Barton M, Atkins D, Lee J, *et al.* Integration of IP mobility and security for secure wireless communications. In *2002 IEEE International Conference on Communications*, 2002; pp. 1045–1049.
31. Schmid S, Finney J, Wu M, Friday A, Scott A, Shepherd W. An access control architecture for microcellular wireless IPv6 networks. In *26th Annual IEEE Conference on Local Computer Networks (LCN2001)*, 2001; pp. 454–463.
32. Kent S, Atkinson R. Security architecture for the internet protocol. *IETF RFC2401*, November 1998.
33. Liang W, Wang W. A dynamic security association control scheme for efficient authentication in wireless networks. In *IEEE GLOBECOM'04*, November 2004.
34. Aboba B, Simon D. PPP EAP TLS authentication protocol. *RFC2716*, October 1999.
35. F. S. Inc. Secure authentication, access control, and data privacy on wireless LANs. In *White Paper*, January 2002.
36. Gupta V, Gupta S. Experiments in wireless internet security. In *IEEE WCNC 2002*, Vol. 2, March 2002; pp. 17–21.
37. Stubblefield A, Ioannidis J, Rubin A. Using the Fluhrer, Martin, and Shamir attack to break WEB. *AT&T Labs*, August 21, 2001.
38. Hess A, Schafer G. Performance evaluation of AAA/Mobile IP authentication. In http://www-tkn.ee.tu-berlin.de/publications/papers/pgts2002.pdf, 2002.

Authors' Biographies

Wenye Wang (M'98/ACM'99) received her B.S. and M.S. degrees from Beijing University of Posts and Telecommunications, Beijing, China, in 1986 and 1991, respectively. She also received her M.S.E.E. degree and her Ph.D. from Georgia Institute of Technology, Atlanta, GA in 1999 and 2002, respectively. She is now an assistant professor with the Department of Electrical and Computer Engineering, North Carolina State University. Her research interests are in mobile and secure computing, quality-of-service (QoS) sensitive networking protocols in single- and multi-hop networks. She has served on program committees for IEEE INFOCOM, ICC, and ICCCN in 2004. She has been a member of the Association for Computing Machinery since 2002.



Wei Liang (SM'04) received his B.S. degree from the Department of Electrical Engineering, Tsinghua University, Beijing, China, in 1998. He then went to the Institute of Electronics, Chinese Academy of Sciences, Beijing, China, where he received his M.S. degree in 2001. He is currently a Ph.D. candidate at North Carolina State University. Since 2001, he has been a teaching assistant in the Department of Electrical and Computer Engineering, North Carolina State University. He took part in the networking group in 2002. His research interests include mobile and secure computing, authentication in wireless networks, quality-of-service (QoS) in mobile networks, mobility management, modeling and performance analysis of wireless information networks.

Avesh K. Agarwal received his B.E. degree from the Motilal Nehru Regional Engineering College (Now MNNIT), India, in 2000. He is currently working towards his Ph.D. in Computer Science at the North Carolina State University. His current research interests are in the area of simulations and real-time performance evaluation of security mechanisms and routing protocols in wireless LANs and wireless ad hoc networks.