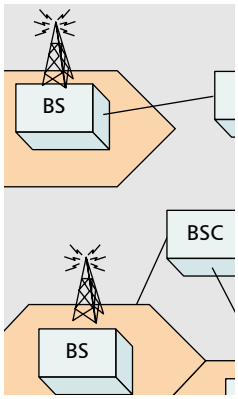
A LIGHTWEIGHT RECONFIGURABLE SECURITY MECHANISM FOR 3G/4G MOBILE DEVICES

JALAL AL-MUHTADI, DENNIS MICKUNAS, AND ROY CAMPBELL,
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN



Existing security schemes in 2G and 3G systems are inadequate, since there is greater demand to provide a more flexible, reconfigurable, and scalable security mechanism that can evolve as fast as mobile hosts are evolving into full-fledged IP-enabled devices.

ABSTRACT

Wireless communications are advancing rapidly. We are currently on the verge of a new revolutionary advancement in wireless data communications: the third generation of mobile telecommunications. 3G promises to converge mobile technology with Internet connectivity. Wireless data, multimedia applications, and integrated services will be among the major driving forces behind 3G. While wireless communications provide great flexibility and mobility, they often come at the expense of security. This is because wireless communications rely on open and public transmission media that raise further security vulnerabilities in addition to the security threats found in regular wired networks. Existing security schemes in 2G and 3G systems are inadequate, since there is a greater demand to provide a more flexible, reconfigurable, and scalable security mechanism that can evolve as fast as mobile hosts are evolving into full-fledged IP-enabled devices. We propose a lightweight, component-based, reconfigurable security mechanism to enhance the security abilities of mobile devices.

INTRODUCTION

Wireless communications are advancing rapidly. We are currently on the verge of a new revolutionary advancement in wireless data communications: the third generation (3G) of mobile telecommunications, and General Packet Radio Service (GPRS), the stepping stone preceding 3G. Wireless data, multimedia applications, and integrated services will be among the major driving forces behind 3G. While wireless communications provide great flexibility and mobility, they often come at the expense of security. This is because wireless communications rely on open and public transmission media that raise further security vulnerabilities in addition to the security threats found in regular wired networks.

To realize the full potential of wireless data while providing secure communications, a more robust and flexible security mechanism is needed in mobile devices, particularly mobile phones, portable communications services (PCS), and 3G devices. This security mechanism has to be lightweight, reconfigurable, and capable of capturing the dynamism and agility of mobile environments. Moreover, our objective is to base the proposed security mechanism on a proven commercially available security infrastructure that is widely used in existing IP networks, thereby granting us flexibility, scalability, and the ability to interoperate securely with existing security mechanisms like Kerberos [1] or SESAME [2]. Such a system should be able to provide confidentiality, authentication, integrity, and nonrepudiation services. Moreover, we aim at providing security services that are flexible enough to secure IP-based applications and meet e-commerce security needs while providing enhanced security and authentication for regular voice communications.

Our approach is to employ a component-based, lightweight, portable security mechanism that we developed based on the SESAME architecture: *Tiny SESAME*. This article discusses Tiny SESAME and its deployment in mobile networks. This article is organized as follows. We briefly describe the evolution of Tiny SESAME and motivate its use in mobile environments. We discuss Tiny SESAME's architecture and describe the use of Tiny SESAME in mobile networks. We briefly mention Tiny SESAME implementation and the testbed we used to experiment with mobile communications.

TINY SESAME

SESAME is a European security architecture for distributed systems. SESAME extends Kerberos by providing additional services. These services include support for asymmetric cryptography,

This research was funded in part by the Motorola Center for Communications at the University of Illinois at Urbana-Champaign and NSF 98-70736.

1070-9916/02/$17.00 © 2002 IEEE

use of special certificates called privileged attribute certificates (PACs) to authenticate principals and identify their privileges, security attributes, and access rights. SESAME also provides authorization services through the Role-Based Access Control (RBAC) model [3]. Moreover, SESAME defines different key management protocols and cryptographic profiles while maintaining backward compatibility with Kerberos. In order to allow application developers to utilize its security services, SESAME adopts the Internet standard *generic security services application programming interface* (GSS-API) [4]. GSS-API is a standard programming interface for generic security services. ECMA standardized SESAME in [5]. Further details about the original SESAME architecture and implementation can be found in [6, 7].

The popularity and widespread of mobile and handheld devices, and the evolution of mobile telecommunication technologies demand a lightweight, component-based, reconfigurable security mechanism that is able to adapt to environments with scarce resources. Nonetheless, this component should be able to reconfigure and evolve as soon as further resources can be spared. The SESAME system with its various protocols and complicated data structures demands high resources in terms of memory and processing power which exceed the typical capacity of mobile devices. *UIUC SESAME* [8] is a portable version of SESAME implemented completely in Java. *Tiny SESAME* is a component-based subset of UIUC SESAME that supports authentication, protocol negotiation, various levels and strengths of encryption, and access control based on the RBAC model. Tiny SESAME's lightweightness is achieved through the employment of a dynamically reconfigurable component-based architecture. This architecture allows Tiny-SESAME-enabled devices to dynamically negotiate the security services, protocols, and cryptographic support needed. Furthermore, it omits the more advanced capabilities of SESAME such as restricted delegation, PV/CV, and intensive auditing. Tiny SESAME employs a Java-based GSS-API that enables applications on the mobile device to interface with the security services.

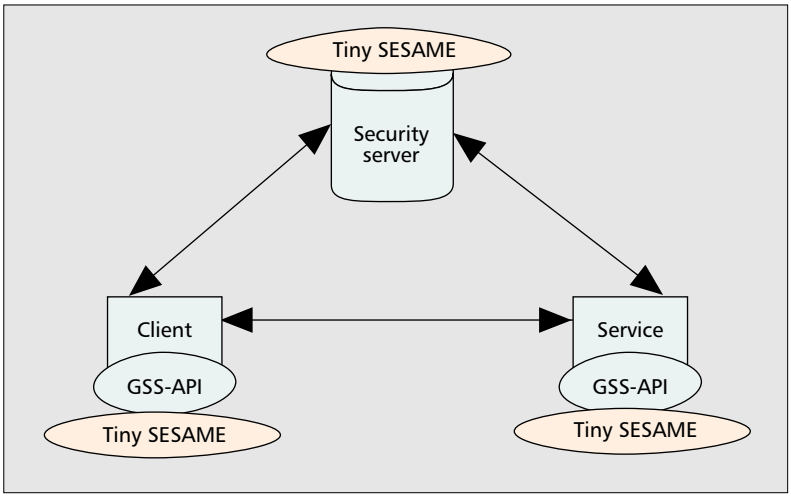


Figure 1. *General overview.*

Tiny Sesame Architecture

In this section a brief description of Tiny SESAME's architecture is in order.

General Overview

A secure environment that utilizes Tiny SESAME consists of at least three major components (Fig. 1): the client application or the initiator who is attempting to securely contact a

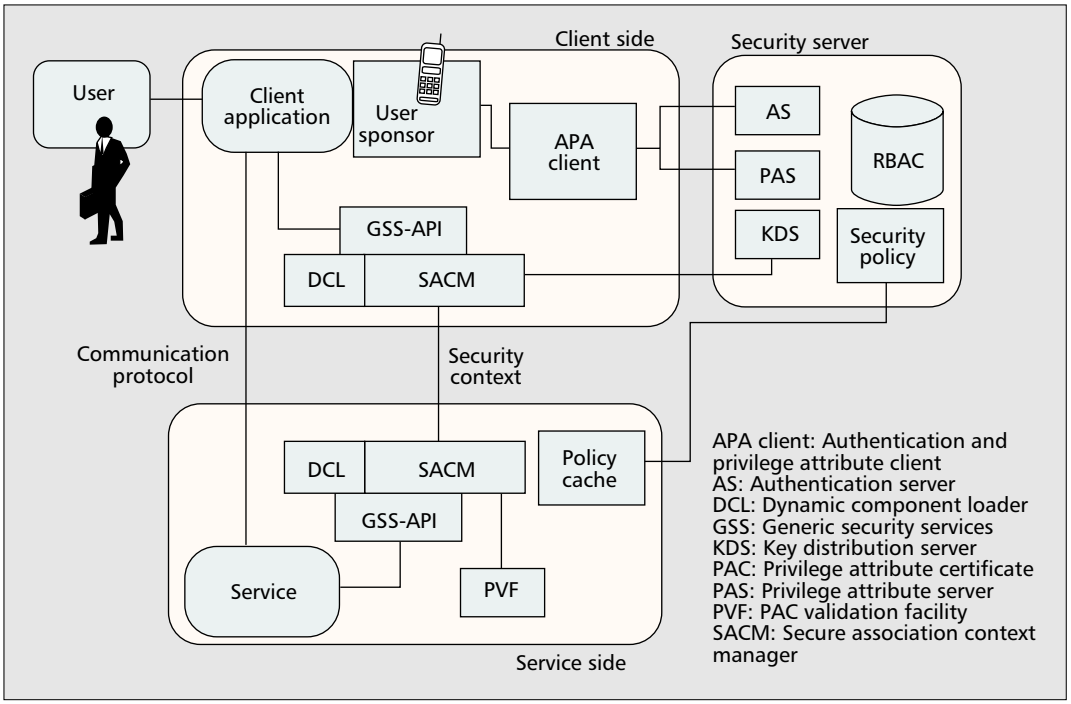


Figure 2. *Tiny SESAME architecture.*

Mobile devices differ greatly in their capabilities, processing powers and security needs. What is needed is a security model that can adapt to the particular capabilities and security requirements of a mobile or an embedded device.

server or get a service, the application server or the service, and the security server that authenticates the users and makes access control decisions. The client and server use GSS-API to authenticate and request security services for the communication channel. A detailed view of the architecture is shown in Fig. 2.

Client Side

The client side is the client application that would like to incorporate security services to access a service or an application server. The client application resides on some *user sponsor*. The user sponsor could be a personal computer, a handheld, a mobile host, or any device capable of running Java. The secure association context nanager (SACM) provides data integrity and confidentiality services for the communication between the client and the service.

The GSS-API on top of SACM provides Java programs with a standard interface to access SACM services. Since SESAME's protocols are rather complex and resource demanding, in Tiny SESAME different protocols, cryptographic profiles, and access control models are implemented as separate software components. These components are loaded dynamically on demand. Once a component is no longer needed it can be unloaded on the fly, allowing only necessary components to be loaded in memory at one time. The dynamic component loader (DCL) is responsible for on-demand loading of required components. A later section describes the currently implemented components in Tiny SESAME. The authentication and privilege attribute client (APA-Client) is responsible for secure connections with the authentication server.

Security Server

The security server includes an authentication server (AS), which provides a single sign-on point for the distributed environment. Tiny SESAME supports two authentication methods. The first is based on Kerberos authentication, which is based on passwords and symmetric keys. The second is based on X.509's strong two-way authentication protocol [9], which uses public key cryptography.

The key distribution server (KDS) manages the cryptographic keys that are used for mutual authentication between the client and remote server. Like its Kerberos counterpart, it relies on symmetric encryption. It is worth noting that in SESAME, if the PAC validation facility (PVF) uses long-term public keys, the KDC can be bypassed. In this scenario, the session keys are generated by the client side's SACM, consequently, the security server is not able to decrypt the communication between the client and the application server.

The privilege attribute server (PAS) provides information about the privileges and security attributes of users and user sponsors. This information is provided to principals in the form of a special data structure referred to as the PAC. The PAC is signed, as a protection against tampering, using the private key of the PAS. Different PACs can be issued to different users and entities to identify them. The PAC can contain the role names assigned to a particular principal. The PAC and the security attributes and roles contained within are used for access control decisions based on a role-based access control (RBAC) model [3]. RBAC is employed for its power of expression and flexibility. RBAC makes it easy to define well structured rules and practices for regulating and managing sensitive data and resources. These rules and regulations form *security policies* that are defined and stored within the security server. An example of a security policy may include restricting access to subscription-based services or reserving certain bandwidth for emergency services, and so on.

Service Side

The service side represents the application servers providing specific services that users try to access remotely. Here the SACM, GSS-API, and DCL are present and function exactly as their counterparts on the client side. The PAC validation facility (PVF) is the component responsible for checking the validity of PACs and detecting any tampering. The policy cache provides caching for security policies relevant to the service. This eliminates the need to consult the security server for every access decision.

Component-Based Architecture

Mobile devices differ greatly in their capabilities, processing powers, and security needs. What is needed is a security model that can adapt to the particular capabilities and security requirements of a mobile or an embedded device. Tiny SESAME achieves this by incorporating a component-based design for the client and service sides. In this design, the different security services, protocols, and cryptographic profiles are implemented as separate components that can be loaded, unloaded, or reconfigured on demand. Thus, the security requirements of a device determine which components, or code modules, are loaded, effectively keeping Tiny SESAME's size manageable.

Tiny SESAME consists of a core component that contains the essential functionality. This includes the Tiny SESAME factory object, the DCL, and the GSS interface without the actual implementation of the security services. The DCL is capable of loading other Tiny SESAME components from the device's built-in memory on demand. The Tiny SESAME package contains several components. Currently, these include several components that implement SACM using different encryption algorithms (SACM-RC5, SACM-DES, and SACM-A5), the SACM with a full-crypto library (includes the most popular encryption algorithms together in one component), the APA client component, and the PVF component. For security reasons, Tiny SESAME's critical components are loaded only from local memory. However, it is possible to extend or update Tiny SESAME's components through special add-on components that are certified and digitally signed by a trusted certificate authority. Upon verifying their integrity and trustworthiness, the components can be

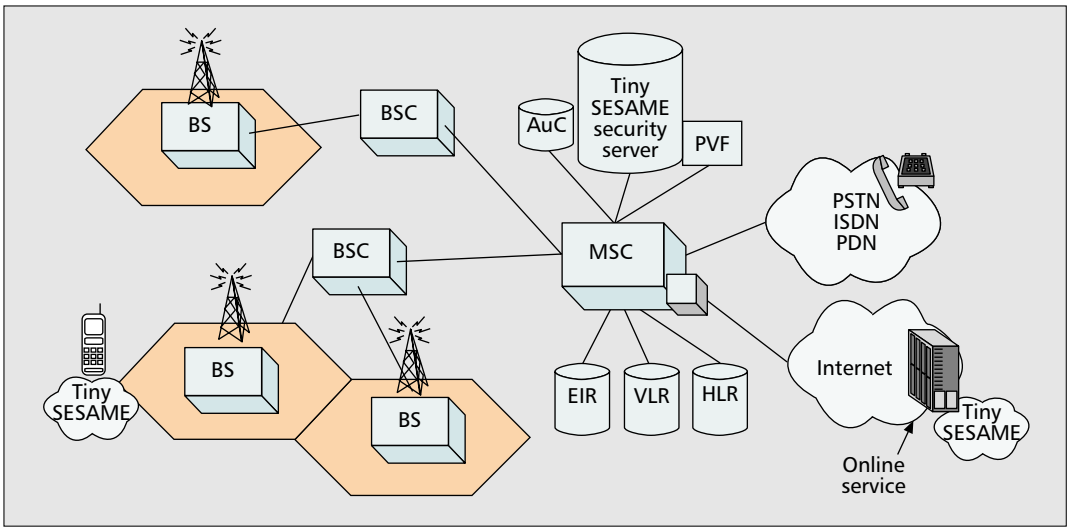


Figure 3. *2G/3G architecture with Tiny SESAME.*

loaded into Tiny SESAME. Additional cryptographic algorithms can be written as certified and digitally signed code modules that can be loaded and used.

Employing Tiny Sesame in 2G/3G

Existing Security

In this section we present a brief description of the existing security in 2G and 3G systems. Further details on 3G security can be found in the 3GPP specification [11]; the details of 2G security are documented in the GSM standard [10].

Initially, the mobile host identification number is transmitted to the base station (BS) unprotected. The receiving BS identifies the mobile host and contacts the home environment of the subscriber. In both 2G and 3G systems, a random challenge (RAND) is generated by the subscriber's home location register (HLR) and AuC. Additionally, a session key (SK) is derived. From this random challenge and the subscriber's individual key (K), which is stored on the 2G SIM or 3G USIM card, the expected response (RESP) is calculated and the session key is derived at the mobile host. This session key is 64 bits in 2G systems and 128 bits in 3G systems. Moreover, in 3G systems security is further enhanced through the use of a 128-bit integrity key (IK) and an authentication token (AUTN) that serves to authenticate the home location to the mobile station, enabling mutual authentication. At the subscriber's HLR, the security information is grouped together to form a *triplet* in the 2G case: $<CHAL_{2G}, RESP_{2G}, SK_{2G}>$ and a *quintuplet* in the 3G case: $<CHAL_{3G}, RESP_{3G}, SK_{3G}$, IK, AUTN>. This information is sent to the requesting visited location register (VLR). In 2G and 3G the authentication of the user is based on his or her knowledge of the individual key (K). Upon receiving the random challenge, the mobile host calculates the RESP and derives the session key. If the mobile station returns the correct response, then the authentication is complete in 2G systems and data transmitted will be encrypted with the session key. In 3G systems, however, additional steps are required, during which the AUTN is sent to the mobile station for mutual authentication. Finally, the 3G mobile host derives the integrity key, and data can be encrypted and validated.

Limitations of Existing Systems

Existing authentication and security in 2G/3G have several limitations. First, the subscriber's ID confidentiality is threatened as a result of the initial sending of the mobile host's identification number unprotected. Second, the key sizes, and encryption and decryption algorithms are fixed. This makes the existing systems inflexible and less secure whenever a security vulnerability is discovered in an existing algorithm, as was the case for GSM's A5/1 algorithm [12]. Having a dynamic security mechanism that is able to negotiate and load new encryption modules with different key lengths on demand allows greater flexibility. In addition to securing mobile communications, the security mechanisms in mobile devices should be able to provide security services for multimedia applications and IP-based services. Moreover, the security mechanism should have the potential to provide more advanced security services that include the creation and enforcement of flexible security policies that are customized for individual services, network administrative domains, organizations, and users.

Using Tiny SESAME for Initial Authentication and Call Setup

In our existing work, Tiny SESAME for mobile devices supports two different authentication protocols. Due to its component-based design, additional authentication protocols can be added later through additional add-on loadable modules. The following subsections will briefly discuss the different approaches.

3G Compatibility Mode — This is the default authentication mechanism. It is backward compatible with the proposed authentication and security aspects for 3G [11]. This involves the formation of the security quintuplet (or triplet in case of a 2G system) exactly as described in

Existing authentication and security in 2G/3G have several limitations. First, the subscriber's ID confidentiality is threatened as a result of the initial sending of the mobile host's identification number unprotected. Second, the key sizes, and encryption and decryption algorithms are fixed. This makes the existing systems inflexible and less secure.

DES 56 encryption (for 32 bytes of data)	82.0 ms
DES 56 decryption (for 32 bytes of data)	82.4 ms
RC5 128 encryption (for 32 bytes of data)	42.05 ms
RC5 128 decryption (for 32 bytes of data)	47.45 ms
RSA 512 encryption (for 32 bytes of data)	236.0 ms
RSA 512 decryption (for 32 bytes of data)	384.0 ms

Table 1. *Encryption/decryption speed.*

3G compatibility authentication	8.04 s
Tiny SESAME authentication through PACs.	10.02 s
Tiny SESAME authentication with higher-subscriber confidentiality (through public key encryption).	16.34 s

Table 2. *Performance of the proposed authentication and call setup schemes.*

the previous section. The mobile host's ID is transmitted unprotected. For this mode, the network parts of a 3G system do not need any modification.

Tiny-SESAME Authentication — In this model, a Tiny SESAME security server should be added to the mobile switching centers (MSCs) (Fig. 3); alternatively, the authentication center (AuC) can be enhanced to provide Tiny SESAME security server's services. The Tiny SESAME security server will host the AS, KDS, PAS, and PVF components (as described in an earlier section). To be able to access the services securely, users are authenticated to the system by the AS component within the security server using an enhanced version of the Kerberos authentication scheme. This scheme will authenticate a user based on his or her knowledge of the subscriber's individual key stored in the SIM or USIM. Initially, the host's ID (unprotected) and an authentication request are sent by the mobile host to the BS. The BS securely passes the request to the security server to authenticate the user. Upon successful authentication, the user sponsor obtains a PAC from the PAS. The PAC contains the user's roles, security attributes, and the basic key. This basic key will be used to derive the session key for securing the communication. The PAC is signed using the private key of the PAS. This PAC can now act as a proof of identity for that user. Furthermore, the security attributes and role names stored in the PAC identify which services that user may access and can be used in creating security policies.

Since a PAC has security attributes that protect it from tampering and has an expiration period, it can be stored within the mobile host. This improves performance and limits the messages exchanged since the mobile host can use the PAC to authenticate itself to different base stations without waiting for security information to come from the HLR of the subscriber's home environment. The security server at the MSC will validate the PAC through its PVF component. If the PAC was issued from another domain, the SESAME interdomain authentication protocol can be employed.

Last but not least, through a slight modification of the protocol, a higher subscriber confidentiality can be achieved. In this case, the asymmetric cryptography capabilities of Tiny SESAME are utilized to encrypt the host's ID with the public key of the nearby MSC rather than sending that information unprotected.

Using Tiny SESAME to Provide Security Services for Multimedia Applications

In addition to adding security to voice and data calls, in mobile devices Tiny SESAME can be used to provide security services for multimedia or IP-based applications. This is essential, particularly in an all-IP 3G network. Since Tiny SESAME is backward compatible with Kerberos, applications on the mobile host can interact securely with "Kerberized" or "Sesamized" applications or services (e.g., the ktelnet application). When the mobile host wishes to use network services, the stored PAC is sent using the SACM to the desired service. The data sent is transferred by any protocol (radio signals and/or IP etc.) to the service. If the service employs SESAME as well, it passes the PAC to the PVF, which verifies, from the integrity protection, that the PAC is genuine, and a secure association is established over which the user can communicate with the remote service. However, if mutual authentication is desired, where a mobile host wants to make sure it is contacting the intended service, the PAC of the target device is sent back before the establishment of the secure association. Once all PACs are validated successfully, the secure association can take place.

Tiny SESAME Implementation

Tiny SESAME was originally implemented in Java. To experiment with security in mobile devices, a version of Tiny SESAME was ported to PersonalJava™ running on Windows CE 2.11 devices.

Our test platform was an HP Jornada 680 running Windows CE 2.11. The Jornada has a Hitachi SH3 processor running at 133 MHz, 16 Mbytes RAM, and a screen resolution of 640 × 240, 256 colors. The Jornada is equipped with wireless Ethernet. We wrote code to simulate the functionality of the VLR, HLR, and AuC, and ran it on fixed workstations on the LAN. Using the Jornada we were able to get some performance results of selected encryption algorithms on the Jornada (Table 1). We assume this would be close to the performance of a 3G device. We tested the different authentication and call setup schemes mentioned in an earlier section and got the preliminary performance measures shown in Table 2.[1]

Future Work

We are using OPNET Modeler to more accurately simulate 2G and 3G security aspects, and measure the impact of employing Tiny SESAME in

[1] *The results are obtained under the assumption that the propagation delay between the mobile unit and the BS is 1 ms, and the propagation delay between the VLR and HLR is 5 ms with a maximum bandwidth of 2 Mb/s (wireless Ethernet).*

different scenarios and compare the results with conventional 2G/3G security. We are also working on porting Tiny SESAME to Java 2 Micro Edition (J2ME), a targeted Java API for writing wireless applications that run on small devices including mobile phones, PDAs, and pagers.

Acknowledgment

This article is based on previously published material from 3Gwireless 2001 organized by Delson Group (http://www.delson.org).

References

[1] MIT's Kerberos Homepage: http://web.mit.edu/kerberos/www/index.html
[2] The SESAME homepage https://www.cosic.esat.kuleuven.ac.be/sesame
[3] R. Sandhu *et al.*, "Role Based Access Control Models," *IEEE Comp.*, vol. 29, no. 2, Feb. 1996.
[4] J. Linn, "Generic Security Service Application Program Interface Version 2," Jan. 1997, RFC 2078.
[5] ECMA-219, "Authentication and Privilege Attribute Application with Related Key Distribution Functions," 2nd ed., Mar. 1996, http://www.ecma.ch
[6] P. Kaijser, T. Parker, and D. Pinkas, "SESAME: The Solution to Security for Open Distributed Systems," *Comp. Commun.*, vol. 17, no. 7, July 1994, pp. 501–18.
[7] P. Kaijser, "A Review of SESAME Development," *Proc. 3rd ACISP Conf.*, 1998, pp. 1–8.
[8] M. Chandak, "UIUC-SESAME: Achieving a Portable Authentication, Access Control, and Delegation Protocol," Masters Thesis at Department of Computer Science, Univ. of IL at Urbana-Champaign, 1999.
[9] ITU-T Rec. X.509 (rev.), "The Directory — Authentication Framework," Geneva, Switzerland, 1993.
[10] ETSI GSM 02.09, "Digital Cellular Telecommunications System (Phase 2+); Security Aspects."
[11] 3GPP Draft Tech.Spec. 33.102, "3G Security Architecture."
[12] A. Biryukov and A. Shamir, "Real-Time Cryptanalysis of the Alleged A5/1 on a PC," presented at Fast Encryption Software Wksp. 2000, Apr. 2000.

Biographies

JALAL AL-MUHTADI (almuhtad@uiuc.edu) is a graduate research assistant and Ph.D. student at the University of Illinois at Urbana-Champaign. His research interests include security in wireless communications and ubiquitous computing. He received his B.Sc. in computer science from King Saud University, Saudi Arabia, and his M.S. in computer science from the University of Illinois at Urbana-Champaign. He has published several papers on security, particularly security in ubiquitous computing environments.

DENNIS MICKUNAS (mickunas@cs.uiuc.edu) holds a B.S. degree in mathematics from the Illinois Institute of Technology, and M.S. and Ph.D. degrees in computer science from Purdue University. He is associate head of the Department of Computer Science at the University of Illinois at Urbana-Champaign. He conducts research on security policy architectures and dynamically adaptable operating systems, and has co-authored more than 40 papers on these subjects.

ROY H. CAMPBELL (rhc@uiuc.edu) holds a B.Sc. degree in mathematics with honors from the University of Sussex, and both M.Sc. and Ph.D. degrees in computing from the University of Newcastle upon Tyne. He is a professor in the Department of Computer Science at the University of Illinois at Urbana-Champaign. He conducts research on dynamically adaptable operating systems, mobile computer security, multimedia, and video, and has co-authored more than 154 papers on these subjects.

In addition to adding security to voice and data calls, in mobile devices Tiny SESAME can be used to provide security services for multimedia or IP-based applications. This is essential, particularly in an all-IP 3G network.