(4) Security Technology for Next-generation Mobile Communication Systems

Christian Schaefer, Hu Wang, Anand Prasad,

Thomas Walter and Peter Schoo

This article presents the research work on security for next-generation mobile communication systems that is carried out in our lab, emphasizing network layer security and application security support.

1. Introduction

The importance of Information Technology (IT) security has massively increased in recent years. The most important driving factors are the growing use of the Internet and web-based technologies. This also applies to Mobile Communications Systems (MCS), where there is a growing need for appropriate security measures. In the future it will become increasingly important for Mobile Network Operators (MNO) to conduct research into security technologies for the following reasons:

1) As acceptance of MCS increases, more users anticipate the potential consequences and effects these systems have in their everyday life. Consequently, users change their expectations of the confidence they have in such systems.
2) A greater awareness of users regarding security increases the importance of security in MNO business models. Security gets more visible as a product quality and a competence topic of an MNO. Therefore, higher security can strengthen its position on the market and supports a prospering business.
3) Today's telecommunications systems, to which MCS belong, are understood as a major component of the public infrastructure. Many other infrastructure components like the energy, finance and health sectors depend on them. Mobile network operators therefore need MCSs that operate securely, correctly and continuously.

Against this background, we are conducting research into next generation mobile communication systems and security in the provision of services and networks, building on the achievements DoCoMo has already made.

The security level can range from "no particular damage is expected, thus no protection is required" to "it can take catastrophic dimensions and business can be stopped, thus the highest level of security has to be adopted". This exemplifies that there exists no unique security solution and that the protection requirements must be identified for each and every IT application or system individually. Determined by the required security level and the remaining risks one is willing to take, security solutions should be integrated into design and development processes. Most often this type of work is thus integration, since implementing security mechanisms can be based on existing and standardized solutions or products.

To qualify the term "security", five standardized and generic security services are distinguished. These are partial elements [1] that are also seen in the standardization specifications for security-ensuring or risk-mitigating solutions.

- Authentication: the process of verifying an identity claimed by or for a system entity;
- Authorization: the permission that is granted to an entity to access a resource;
- Confidentiality: the property that information is not made available or disclosed to unauthorized individuals, entities, or processes;
- Integrity: the accuracy and consistency of the information; and
- Non-repudiation: protection against false denial of involvement in an action or communication.

DoCoMo currently provides both services and network facilities. Here, we are also conducting research into the two fields (projects) of network layer security and application security support.

Network layer security encompasses protocols, interfaces and interchange formats which enable and support the secure transmission of information, basic mechanisms and data formats, and infrastructure components and access protocols which support trusted communications. We are also engaged in a European research project aimed at applying these constituent security technologies to network infrastructures with heteroge-

neous access technologies. In Chapter 2 we discuss the scope of this project and issues for network layer security.

The scope of application security support includes application-specific security solutions, such as secure access to and communication with private Local Area Networks (LANs) or corporate systems, secure co-operation between devices that can form federations, and recommendations for data security and data protection. It also covers the means that can be used by applications in order to protect their data or resources, such as document signature (signing/verifying). This facilitates the construction of secure federations in mobile environments. A project dealing with these issues is described in Chapter 3.

2. Next-generation Network Layer Security

The International Telecommunication Union-Radiocommunication sector (ITU-R) [2] produced a vision document that lays down the required developments for next-generation mobile systems (i.e., Third-Generation (3G) and beyond). This vision document encompasses the idea of "Optimally Connected Anywhere, Anytime". The main characteristics are: i) a new air interface aiming for 50–100 Mbit/s (the characteristics of Fourth-Generation (4G) wireless access) deployable about 2010, and ii) the integration of existing systems to interwork with each other and with the new air interface—this is what defines systems Beyond 3G (B3G).

One of the challenges to be addressed by next-generation mobile communication networks is supporting heterogeneous networks by allowing the core network to connect to existing wireless access and cable access schemes as well as new wireless access schemes. This raises the important issue of arriving at the same level of security (the desired standard) regardless of the access technology or network technology being used.

It is also impossible to ignore the acceptance of users in the realization of next-generation networks. Fundamental factors for gaining and improving user acceptance include providing technology transparent services and security features, respecting the users' privacy, and making the services easy to use. Security is also an important factor relating to the inherent characteristics of wireless access (including the ability to access a network at multiple unspecified points, unlike cable networks), and it is important to study every aspect of security to exploit the capabilities of wireless technology to their full extent.

In next-generation systems there is a trend towards packet-based network technology that uses Internet technology. Vulnerabilities and risks that come when using this technology are well known, although permanently changing. However, this does not mean that packet-based network technology itself is causing the problem. Rather, it is factors such as the widespread use of unmanaged networks and applications and the multitude of user interests not being controllable that raise the risks. In vulnerable packet-based network environments of this sort, it is very difficult to find a solution that strikes the right balance between the quality and cost of security measures.

In this field, we are conducting studies into risk analysis methods, infrastructure protection, and security measures for heterogeneous access networks and multi-player environments. To study security for next-generation mobile communications networks, we are conducting a Scenario-Based Security Study (SBSS) project in which we first need to understand the possible combination of networks and technologies (Section 2.1) and then understand their security issues (Section 2.2).

2.1 Wireless Network Technology Combinations

The possible combinations of various wireless system technologies are expressed by the relationships shown in **Figure 1** [3]. In this figure, a blue arrow shows that technically combination is possible but current market or frequency overlap might make it unrealistic. A green arrow or a line between two wireless network technologies shows that the combination is possible. Further reasoning based on factors like technology, geographic availability etc. led to a set of meaningful combinations of technologies [3].

2.2 Security Concerns and Issues

This section presents a non-exhaustive list of security issues for future-generation mobile communication systems.

1) Trust Management

One of the primary security concerns is trust—i.e., the extent to which a user who relies on a system can be confident that it will behave correctly. This does not differ much from relationships of trust in everyday life.

Similarly in MCS, the creation of trust is a basic step that has an impact on the security of communication and on business. A user should have an appropriate level of trust in the MNO and vice versa. Furthermore, when the user moves to a new network not owned by the MNO he or she is subscribed to then there must be a trust relationship between the two networks.

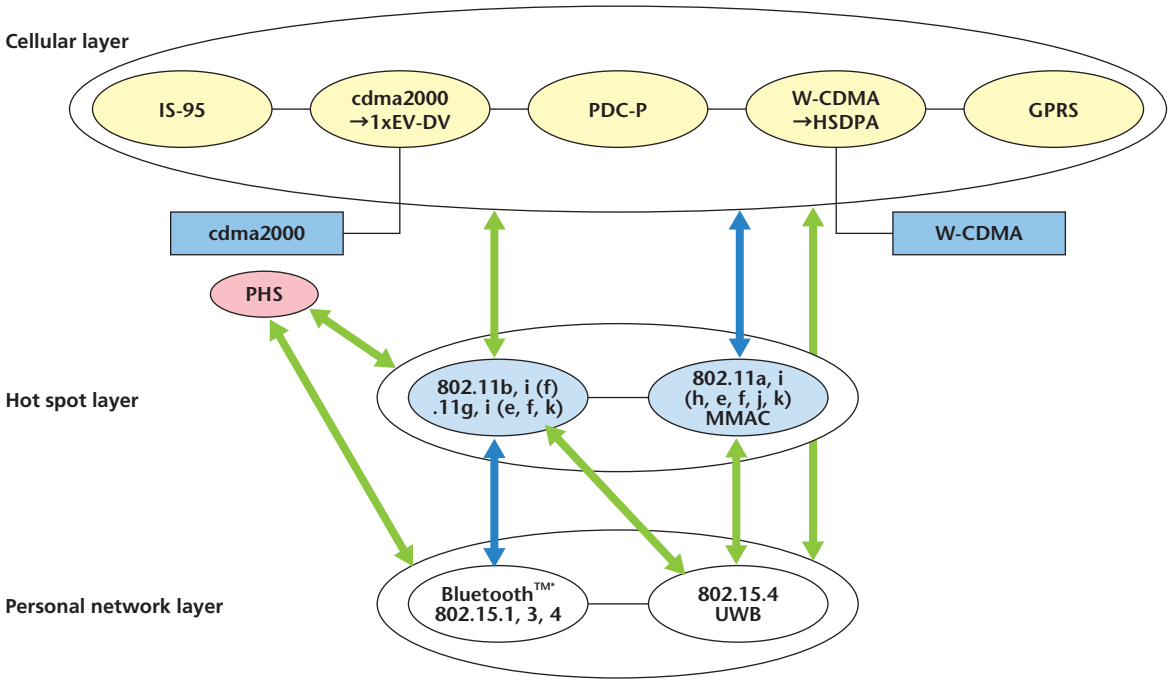

Figure 1 Possible combinations of network technologies

2) Authentication and Key Management

Authentication mechanisms may vary depending on the security technologies and policies used by different networks. This raises the issue of how mobile terminals should adapt to these differences at handovers between domains and do it fast enough such that the service is not affected. Similarly key management mechanisms will vary depending on the network and the system, so it is also important to find a way of performing reliable key management in environments containing mobile terminals with high mobility.

3) Service Level Agreement

In general, when a device moves from one network to another, it is important to maintain the Service Level Agreement (SLA). This process involves two security issues; i) since a SLA also includes an agreement of level of security guaranteed to the user, this must not be affected by the handover, and ii) when negotiating the SLA with the new network, it is important to make sure that there is no possibility for intruders to cause changes in the agreement.

4) Secure Attachment and Detachment

Attaching/detaching securely to/from the network is a very important issue. Normally, when one first connects to a network, important information is exchanged between the network and the mobile terminal. One of the security issues for Internet Protocol (IP) networks is Dynamic Host Configuration Protocol (DHCP). Several security threats may arise from the use of DHCP, like man-in-the-middle attacks through faked DHCP servers, or incorrect configuration information.

5) Accounting and Billing

Another important aspect, also of concern to businesses and service providers, is accounting and billing. If handover happens during an active session of service, especially between different networks, various issues arise such as how to maintain the integrity and consistency of accounting records, or how to implement non-repudiation (e.g., protection against a false denial by the customer that he or she has used the service).

6) Lawful Interception and Anonymity

Lawful interception of the networks operated by MNOs is a legal requirement according to the national laws of several countries. This means that any security solution must also fulfill this requirement when being used for public telecommunications purposes.

Anonymity (protection of user identity) and the protection of location privacy have become standard requirements. Without the network operator's authorization, a user's identity or location must never be disclosed to a third party. It is also essential to make it impossible for outsiders to derive information such as user IDs by recording signaling data after a secure access connection has been established.

* Bluetooth is a registered trademark of Bluetooth SIG, Inc.

7) Network Resources

The IP layer together with layer 2 requires large headers which uses a lot of wireless bandwidth. IP header compression schemes have been proposed to reduce the overhead problems, but current header compression schemes do not offer adequate security. In some cases, it may become possible for attacks such as denial-of-service attacks to exploit the weak header compression. Most protocols used for security and mobility—for example, Internet Protocol Security (IPSec) and Mobile Internet Protocol (Mobile IP)—require a lot of message exchange which uses valuable bandwidth capacity and battery life. In other words, they use up the resources of the network and mobile terminals. In particular, a frequently moving user will cause a very large quantity of messages to be exchanged. The resource demands caused by these protocol overheads must be considered alongside the security requirements.

8) Infrastructure Security

The security of a MNO's infrastructure is also of prime importance and beyond the design of secure protocols. Infrastructure security means security of the network itself, network nodes like routers and servers, and any information available in the network. Attacks on these network nodes can cause the entire network to crash. Secure management of network nodes and timely software upgrades are key issues for infrastructure security.

3. Next-generation Application Security Support

As new applications appear and mobile terminals become increasingly diversified, support for user mobility will become even more important in studies of application security support. Section 3.1 presents an example of application security support for business applications where user mobility is taken into account, and identifies the requirements that need to be met and the issues that need to be dealt with. Section 3.2 introduces the concept of federations and security modules, and introduces a method for building federations in a mobile environment as an original contribution. And Section 3.3 summarizes the issues that mobile network operators will have to deal with in the future.

3.1 Motivation and Application Example

Consider for example a traveling salesperson using a Personal Digital Assistant (PDA) to administer customer contact details and a personal schedule (agenda) (**Figure 2**).

In order to update contacts and agenda information he or she connects to the corporate network via a mobile terminal. In addition, the salesperson gets the latest sales figures out of the corporate sales database and receives the data on his or her mobile terminal. For easy processing, he or she transfers the spreadsheet containing the sales figures to his or her notebook PC or a public terminal.

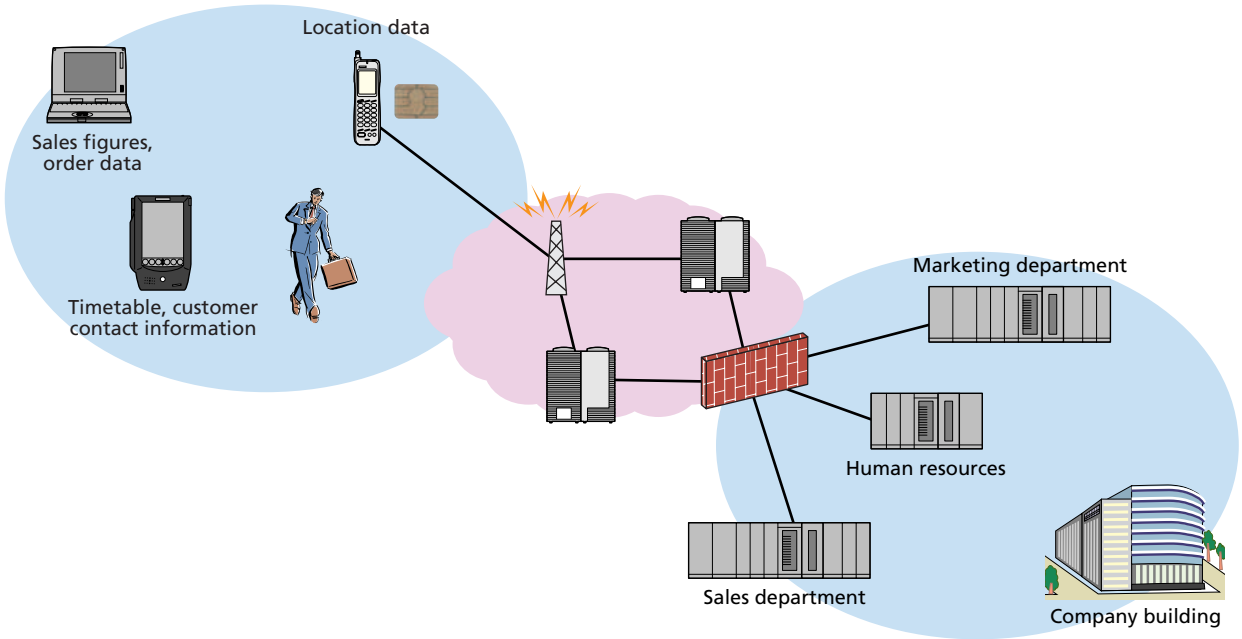

Figure 2 Federation structure for a traveling salesperson

The main security requirements that need to be fulfilled in order to implement the above scenario in a secure environment are as follows:

- Mutual authentication of clients and servers as well as of clients sharing corporate data and resources,
- Authorization of users and applications to use corporate resources
- Data confidentiality and integrity, and non-repudiation of communication events and transactions.

The following section discusses a procedure for setting up a secure communication environment between groups of mobile devices and between mobile devices and corporate networks so that the above security requirements are met; particularly non-repudiation services.

3.2 Security for Federated Devices

The application of security technology to business-to-employee applications (as in the traveling salesperson example above) are being investigated in the Wireless Trust for mobile business (WiTness) project, which is part of the 5th Framework Program of the European Union [4]. Here, secure mobile business applications are developed around two basic concepts: federations and security modules.

1) Federations

The definition of a federation builds upon the concept of devices forming collaborative networks (like the co-operation of mobile phone and notebook in the above example) and the roles of nodes are determined (a priori or ad hoc) on the basis of capability and trust. Every device stores a security certificate containing details such as a private and public key pair, public-key certificates and trust certificates.

Therefore, a federation describes a linking relationship between communicating equipments with different owners. A federation thus extends the concept of a trusted platform beyond a single entity to a distributed environment.

2) Security Module

The security module stores a user's secret data such as private key(s), and performs security-related functions such as digital signing and decryption with the private key. Note that these secret data are generated within the corporate network and stored during initialization on the module. A security module can thus be described as a unique trusted element within the federation. It is implemented in tamper-resistant hardware (hardware designed in a secure manner to prevent illegal modifications), such as a smart card for example.

The security module is accessed whenever a mobile device in a federation requires access to secret data (stored in the module) or needs to perform a secure operation. A federation without a security module is not trusted by the corporate network, and is thus not authorized by the corporate network to perform specific operations or gain access to resources.

3) Bootstrapping a Secure Communications Environment

Referring to the above example, Fig.2, the establishment of a federation requires the following three steps:

step 1. The security module and corporate network mutually authenticate each other.

step 2. The security module authenticates the mobile devices that are to be federated (the notebook PC and PDA in our example).

step 3. The mobile devices are then authenticated with the corporate network.

Step 1 above can be implemented by following established procedures (for instance Secure Socket Layer (SSL) connection set-up). Both entities have pre-installed certificates which are exchanged and validated during connection establishment. The session key used for subsequent communication is encrypted with the security module's public key before it is sent to the security module. Step 2 can also be implemented in more or less the same way, whereby the session key is securely transferred after authentication.

In addition to authenticating federated mobile devices, the security module also generates and transfers a time-limited security token called a temporary certificate that is used by these mobile devices. This procedure allows a mobile device to indicate to the corporate network that it is included in a federation. Step 3 follows the same lines as step 1, and enables the corporate network to exchange a session key with a mobile device. In the above example, this sequence of operations results in the establishment of a secure and trusted federation between the security module, notebook PC and PDA on the user side and the corporate network on the other. This allows the desired business tasks to be performed across secure channels.

3.3 Challenges for Mobile Network Operators in Application Security Support

Mobile communication networks have made it possible for users to access corporate applications from any location, and have promoted new business activities. This benefits MNOs due to the increased amount of traffic generated by mobile business applications, and it also presents service providers with new business opportunities.

1) The Relationship between Security and New Business

As presented here, the smart card built into a mobile phone performs two distinct roles. First, it is a security token used by the MNO to identify the mobile phone. Second, it is a security token used by a company to identify its employees individually, thereby facilitating secure End-to-End (E2E) communication between the company and its employees. Obviously, this requires that the MNO allows third parties (i.e., other companies) to access the smart card so that they can store and manage application-specific security data (keys, certificates, etc.). It is thus likely that close relationships will be established between MNOs and companies.

This approach is also compatible with FirstPass [5], a client authentication service recently established by NTT DoCoMo. Freedom Of Mobile multimedia Access (FOMA) users may acquire a public-key certificate which is downloaded to the FOMA card. In FirstPass, the smart card is still under control of the MNO, but in the WiTness project introduced in Section 3.1, it is also possible to give third parties the right to access smart cards.

In the future, third parties such as other companies may participate in services of this sort to allow customers to be authenticated using their client certificates.

The WiTness project implements an application security support framework that allows MNOs as well as third parties to provide new applications and integration models that may be seen as "complementary" business models because they do not depend on an increasing customer base or increasing data traffic but provide value-added services.

2) The Evolution of Application Security

In recent years, computing and communications environments have changed as shown in **Table 1** according to the needs of users and businesses. Current research into application security support is focused mainly on business applications and can be said to cover all the items listed under "Today" in this table and most of the "Near future" items as well. However, Peer-to-Peer (P2P) applications have not yet received much attention. On the other hand, new challenges are expected to emerge in the "Long-term future" with respect to ubiquitous environments. To mention one example: in the above scenario, security is placed under the control of a single domain (i.e., a single company defines all requirements), and the infrastructure needed to do this has been set up. However, in the long-term future it is thought that security will apply to multiple administrative domains such as device providers, mobile users, and service providers. To establish security and trust in such cases, it is necessary to have a trustworthy third party that holds relationships with all involved partners. A mobile network operator that provides the mobile terminal and understands the needs of users and providers could also act as this trustworthy third party, and it is thought that this position will become even more valuable in the future.

4. Conclusion

This article described two of our current research activities. In the discussion of network layer security in Chapter 2, we listed the combinations of wireless access technologies that should be investigated in the future, and identified the main issues that need to be addressed. It can be said that well designed and secured access, handover and mobility procedures are essential parts of next-generation systems. Further work will be done on

Table 1 Computing and communications paradigms

	Past	Today	Near future	Long-term future
Application	Application-specific	Client-server (chiefly web browsing and web services)	Client-server, web browsing, and web services, P2P	Ubiquitous mobile terminals and communication
Terminal	Desktop system	Desktop systems, Mobile devices (PDAs, mobile phones)	Mobile devices forming federations of mobile terminals (PDAs, mobile phones)	
Network	Intranet	Intranet, VPN	Intranet, VPN, Internet, wired and wireless networks	

the design and evaluation of secure handover procedures, and on expanding the scope of our studies to include issues such as maintaining security amongst different administrative domains. On the other hand, with regard to application security support, we have discussed federations and security module technology which are essential for the implementation of protocols, and we have described a procedure for building federations in mobile environments. In the future, we plan to continue developing this scheme and utilize it in applications such as context-recognition communication.

REFERENCES

[1] IETF Internet RFC 2828-Internet Security Glossary, May 2000.
[2] International Telecommunication Union Radiocommunication sector, http://www.itu.int/ITU-R/.
[3] S. Rohr, B. Kpatcha, C. Eckert, A.R. Prasad, P. Schoo and H. Wang: "Feasible and Meaningful Combinations of Access and Network Technologies for Future Mobile Communications," submitted for publication, WWRF#10 [PS1], Oct. 27–28 2003, New York, USA.
[4] WiTness consortium, Wireless Trust for mobile business, http://www.wireless-trust.org.
[5] N. Nakamura, H. Yamamoto and M. Onogawa: "FirstPass Service Overview," NTT DoCoMo Technical Journal, Vol. 5, No. 3, pp. 4–10, Dec. 2003.

ABBREVIATIONS

cdma2000: Code Division Multiple Access 2000
DHCP: Dynamic Host Configuration Protocol
E2E: End-to-End
FOMA: Freedom Of Mobile multimedia Access
GPRS: General Packet Radio Service
HSDPA: High-Speed Downlink Packet Access
IP: Internet Protocol
IPSec: Internet Protocol Security
IS-95: Interim Standard-95
IT: Information Technology
ITU-R: International Telecommunication Union- Radiocommunkation sector
LAN: Local Area Network
MMAC: Multimedia Mobile Access Communication systems
Mobile IP: Mobile Internet Protocol
P2P: Peer-to-Peer
PDA: Personal Digital Assistant
PDC-P: PDC mobile Packet data communication system
PHS: Personal Handy-phone System
SLA: Service Level Agreement
SSL: Secure Socket Layer
UWB: Ultra Wide Band
VPN: Virtual Private Network (fast packet access on downstream links)
W-CDMA: Wideband Code Division Multiple Access
WiTness: Wireless Trust for mobile business