Study on Secure Mobile Communication based on the Hardware Security Module

Junho Lee, Haeng-Seok Ko, SangHyun Park, Myungwon Seo, and Injung Kim
System R&D Division, The Attached Institute of ETRI
P.O.Box 1, Yuseong-gu, Daejeon-si, KOREA
e-mail: {gladday, hsko, shpark, mwseo, cipher}@ensec.re.kr

***Abstract*— This paper presents a survey of information protection methods for secure mobile communication. Most existing software-based information protection systems have a greater risk for the loss or theft and have difficulty in maintaining. In order to solve these problems, we considered methods of information protection method based on hardware security module that are best adapted to mobile communication environment.**

Keywords - hardware security; secure mobile communication.

I. INTRODUCTION

Over the past four years since Apple's release of the 1st-generation iPhone in 2007, there has been an explosive growth in the demand for smartphones and other mobile devices such as tablet PCs. This surge in the use of mobile devices, however, has not been accompanied by adequate security policies to ensure the safety of communication. As a result, the mobile environment is affected by the same problems that have been plaguing the fixed PC-based Internet environment, such as the spread of malicious codes, hacking and then the resulting leakage of private information [1]. Cell phone tapping and information theft are particularly serious threats to the safety of using mobile devices for government agencies and companies as well as for the general public.

To resolve these security issues, the US Army, for example, is using special smartphones like the L-3 Guardian® Secure Mobile Environment Portable Electronic Device (SME PED) by L3 Communications, in a robust security move [2]. These types of special smartphones, which were designed for secure communications, are however, too onerous for civilian use, whether by businesses or by consumers.

For standard smartphones that are not for encrypted communication, one way of increasing the security of communication is using a separate cipher device to enable cipher communications [3]. Another popularly used method is using a software encryption solution, which hooks IP packets [4]. Both methods require the modification of the hardware or software of the smartphone, which necessitates assistance from the manufacturer. However, smartphone manufacturers are generally unwilling to assist with the process, for reasons pertaining to device stability or costs.

This paper discusses ways of protecting information stored in mobile devices from various forms of cyber threats. Fig. 1 below illustrates the structure of UMTS (the name of the 3G mobile network selected by the 3rd Generation Partnership Project [3GPP]). As it can be noted from this image, the UMTS network consists of two main parts: a core network and UMTS terrestrial radio access network (UTRAN), where the user equipment connects to the UTRAN. Communication within the UMTS network takes place in two separate domains: voice data is transmitted through the circuit-switched (CS) network, while data packets are transmitted in packets through the packet switched (PS) network. In this paper, we will focus on end-to-end methods for protecting user data for secure communication within 3G PS networks. We will begin by examining the characteristics of the environment for end-to-end mobile communication, and then we will proceed to discuss methods of information protection that are best adapted to this environment.

II. PREVIOUS APPROACHES

One of the most widely used methods for enabling secure communication in a smartphone is by connecting a cipher device to the earphone jack [3]. The voice data is received through the microphone of the cipher device and is encrypted by its digital signal processor (DSP) before it is transmitted to the receiving party. The DSP of the cipher device decrypts the received encrypted data also before it is transmitted through the earphone. This method has the advantage in that it does not require the modification of the phone module and can be used for any type of mobile phone. However, this approach is limited to circuit services, and doesn't provide secure communication for data packets. Another major disadvantage is its costliness, due to the DSP, battery, and codecs used for the module.

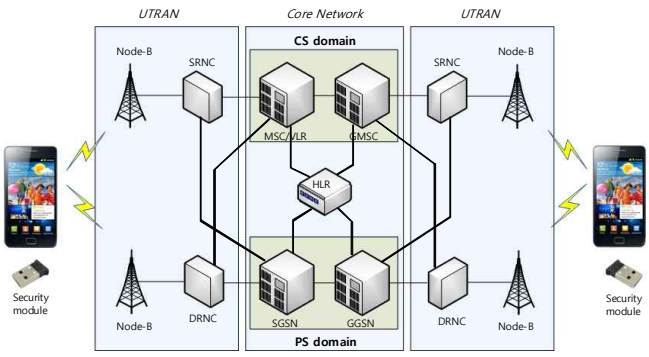

Figure 1. Structure of the UMTS network

 ISBN: 978-1-61208-171-7

Another method consists of encrypting the data by hooking up IP packets by rooting the smartphone. This method is advantageous in that data can be encrypted for all applications used in the smartphone. But the encrypted data also runs the risk of becoming exposed to malicious software attacks or hacking attempts because of the rooting process. More recently, a new method for cipher communication using a secure USIM, instead of a simple USIM (Universal Subscriber Identity Module), has been proposed [4]. This paper proposes a WAP public key infrastructure (WPKI) method, which allows for the remote management of card applications through the USIM chip. Using the WPKI concept, secure communication can be enabled on smartphones by adding a cryptographic algorithm to the security function of the USIM. In order to use the secure USIM, one must be able to hook IP packets. This requires that a number of modifications be done to the hardware as well as the software of a smartphone. However, as has been mentioned earlier, most manufacturers are refusing to assist with introducing such modifications to their phones on the basis of stability risks or costs. Another flaw with this method is that it is vulnerable to tampering attempts.

There are also ways of enabling secure communication through a smartphone application. Although software-based security solutions, involving no separate hardware module, cost less and offer a greater degree of user-friendliness, storing security algorithms and keys inside the device itself makes them run the risk of becoming stolen through hacking. Information that is stolen from within the smartphone can be remotely deleted or controlled, when the phone is lost or stolen, but this does not work when the network is blocked.

III. INFORMATION PROTECTION METHODS OPTIMALLY ATTUNED TO THE MOBILE DEVICE ENVIRONMENT

The existing methods discussed above are not precisely aligned with the specific environment for mobile devices; hence, they are poorly adapted for use with mobile devices. In this section, we will analyze the characteristics that are specific to the mobile device environment and we will propose methods for the protection of information that are best adapted to this environment.

A. Characteristics of the Mobile Device Environment

1. An explosive growth in data traffic, resulting from a sharp surge in the use of smartphones

The widespread use of services like mobile voice over IP (mVoIP) with smartphone and mobile devices in general has caused the sharing of data in overall telecommunications traffic to jump, shrinking the sharing of voice communication commensurately [5]. mVoIP is a technology for converting voice data into packets of IP data and transmitting them using a real-time transport protocol (RTP). This technology makes calls dramatically cheaper and is, for this reason, rapidly replacing traditional voice call services. What this means is that, going forward, secure communication solutions will be needed mostly for data, rather than for voice communication as such.

2. Limited Resources

The biggest advantage of a smartphone is its portability. Users are, therefore, naturally highly sensitive to the issue of battery life. Even the best of smartphone security systems will be shunned if such systems shorten life of the battery and take up an excessive amount of resources, including memory, which results in the slowing down of the device. Therefore, an information protection solution for a smartphone must be designed in a manner that is adapted to the mobile device environment; namely, it must have a very little impact on battery life.

3. Risk of Loss or Theft

There is a greater risk for mobile devices being lost or stolen, as they are portable devices. It is, therefore, important for a security policy to take this risk into consideration. The method, currently used, which consists of deleting the information stored in a smartphone when it is lost or stolen, by remotely controlling the device, has the fatal flaw of ceasing to be effective as soon as the network is blocked. Accordingly, any solutions for protecting information in the event of the loss or theft of a smartphone must be hardware-based, to be more effective.

4. User-friendliness

Since the huge success enjoyed by Apple's iPhone, developed with a focus on user interface (UI), it has now become an accepted fact that user-friendliness is the prime factor to consider when designing a smartphone. An application, no matter how great, will fall out of favor and become irrelevant, if it is not easy to use. The same is true for secure communications applications. A secure communication solution involving complicated procedures or requiring multiple interconnected devices runs the risk of becoming rejected by users. Therefore, security modules for secure communications must not be complex and the device connection must also be as simple as possible.

B. Hardware Security Modules

For secure communication, cryptographic algorithms are generally placed in the device with the secret key kept offline. Password or certificate-based access is the most popularly used method. Using the password or the private key of the certificate, a session key is generated. The encryption takes place through the encryption algorithms hidden inside an ActiveX control or other software. During the encryption, malicious attempts, such as virus attacks or hacking, are monitored by the security software in real time. However, the security software alone is insufficient for detecting all malicious attempts to breach the security of a mobile device. This is because there are security threats other than malicious code. Hardware attacks, which extract data by causing interference in the hardware of the device, should also be contended with. There are also a great variety of hacking methods based on hardware attacks for mobile devices, such as the injection of errors into the device by decomposing it and disabling the internal logic. Currently, there are no countermeasures to attacks of this kind. Also, the cycle for OS (Operating System) upgrade by device manufacturers is quite short nowadays, and is usually only twice a year. Software patches applied at the upgrade are costly as well as

 ISBN: 978-1-61208-171-7

time-consuming. Hardware security modules are, therefore, the alternative to software solutions for resolving issues that cannot be properly resolved by the latter. As has been mentioned earlier in the discussion of the mobile device environment, resources for mobile devices are severely limited, and user-friendliness is paramount. When using a hardware security module, the following are some of the essential considerations:

1. Limited resources: There is no built-in power supply device for the hardware security module, and electric power should be supplied from the mobile device. Therefore, it must be designed in a manner to minimize power consumption and memory usage.

2. User-friendliness: To ensure its user-friendliness, the hardware security module should be made in such a way that it is automatically recognized by the mobile device, as soon as it is connected to the device.

C. *Method of Information Protection*

Fig. 2 shows a block diagram of the secure mobile communications system using a security module. Here, the hardware security module either directly handles the encryption and decryption of communications data or generates key streams needed for encryption and decryption. The security application of the mobile device, meanwhile, ensures the security of communication using the hardware security module. When the mobile device and module are connected, the caller is authenticated using a PIN or other similar methods. The caller must then share the session key with the receiver in order to proceed to secure communication.

There are two different ways of implementing a secure communications system. One is having the security module directly handle the encryption and decryption process. In this case, the most important consideration is the power consumption of the module (in other words, the module's impact on battery life) and the time delay resulting from the process. In order to process the voice data of VoIP in real time, the security module should be able to encrypt and decrypt them at a relatively fast speed. A normal phone conversation becomes difficult with any delay of 30msec or more [6]. Also, as the security module depends on the mobile device for its power supply, having it directly perform the encryption-decryption process then negatively affects battery life.

The other method is having the security module, which generates the key streams that are needed for encryption and decryption. The secure communications application for the mobile device generates a cipher text by doing exclusive-or (XOR) operation with plain text, using the key stream obtained from the security module, and transmitting it to the call receiver. This method has a number of advantages over the first one. It uses less power, and the encryption time delay is minimal; making it ideal for secure VoIP calls. However, any loss or repetition of packets tends to affect all packets, making the call impossible. The capacity of self-synchronization is, therefore, necessary in order to remedy this issue.

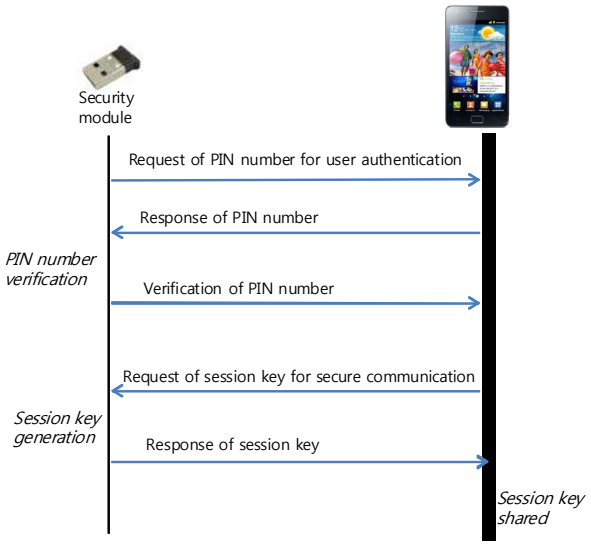

Figure 2. Secure mobile communications system

D. *A Key System and Key-sharing Method Adapted for the Mobile Device Environment*

In this section, we will look into the kind of key system and encryption-decryption method that are best adapted for the mobile device environment. Let us begin by comparing symmetric and asymmetric keys to see which of them are more suitable for mobile devices. For a group of N number of people to use a symmetric key-based encryption system, the number of secret keys shared between them is $N(N\text{-}1)/2$. The corresponding number for an asymmetric key-based encryption system is only N. For example, a group of 1 million people using a symmetric key-based encryption system will need 500 million keys [7]. Therefore, an asymmetric key-based encryption system is more efficient than a symmetric key-based one in terms of the management of keys. On the other hand, the asymmetric key system requires a large amount of computation to calculate private keys from public keys; hence, this system is not well adapted for mobile devices that have limited resources. In this paper, we are primarily concerned with companies or organizations in which the number of secure communication users is limited, rather than with mass secure communication. Therefore, for the purpose of this study a secret key system based on a symmetric key system is a more suitable choice. When the number of users is N, we need a key matrix, which may be expressed as follows:

TABLE I. EXAMPLES OF A N×N KEY MATRIX

	User #1	User #2	⋯	User #N
User #1	K_{11}	K_{12}	⋯	K_{1N}
User #2	K_{12}	K_{22}	⋯	K_{2N}
⋮	⋮	⋮	⋱	⋮
User #N	K_{1N}	K_{2N}	⋯	K_{NN}

 ISBN: 978-1-61208-171-7

User #1 and User #2 can have secure communication using K_{12}, and User #1, by encrypting and decrypting files stored within the mobile device, using K_{11}, can keep them safe from attacks from external sources. Here, the keys must be generated and assigned a priori by the highest level of organization that is reliable in the group. The size of the key matrix, meanwhile, is determined according to the number of users. A company with 1,000 employees, using a 256bit key, would need a storage space of 15MB, for example.

In order to share the session key generated from the secret key, the algorithms must satisfy the following two conditions: first, they should be able to rapidly achieve initial synchronization. In order for users to have a trouble-free session of secure communication, the initial synchronization should occur within the first second from the initiation of a VoIP call. Secondly, the encryption algorithms must provide stability against replay attacks. For mobile devices, methods like IKE by IPSec [8] are preferable to Kerberos [9], which is more resource-intensive.

E. Other Considerations

- Host Mode Support for Mobile Devices

For a mobile device to communicate with a security module through interfaces like a micro USB, it must be able to operate in the host mode for USB. Samsung's Galaxy S2, released in early May, for example, is enabled to operate in host mode. Also given the increasing use of the host mode, smartphones based on Android OS are likely to support it in the near future as well.

- Tamperproof Capacity against Hardware Tampering

As mobile devices can be easily lost or stolen, tamperproof capacity is essential for security solutions. The effectiveness of software-based remote deletion and the control of information have proved limited. The security module discussed in this paper needs to be equipped with features for preventing tampering, which will erase all key information stored in a device, when either abnormal access or an attempt to open the module to steal keys or other secret data is detected.

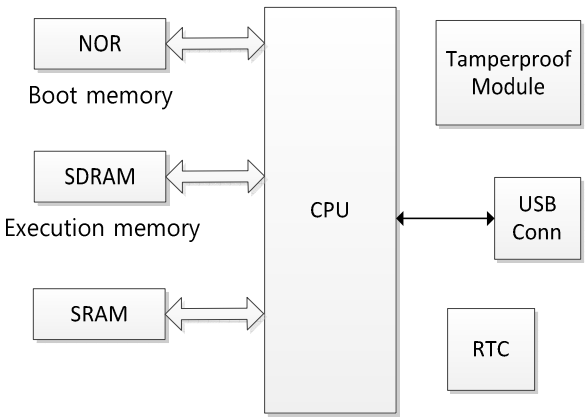

Figure 3. Block diagram of the security module

F. Essential Components of a Hardware Security Module

Based on the various requirements for a security module, discussed above, we can define its essential components as shown in Fig. 3 below. As can be seen in Fig. 3, a hardware security module consists of a processor for the encryption and decryption of data and communication with applications in a mobile device; NOR flash memory for managing boot programs; SRAM memory for storing secret keys; SDRAM memory for imaging the encryption-decryption process; a hardware module for preventing tampering; and a RTC to supply time data to the processor and peripheral devices.

IV. CONCLUSION

This paper has been a discussion of methods for information protection that are the most appropriate for mobile devices. Most existing information protection systems use software-based security solutions, which are poorly prepared for the eventuality of loss or theft of the device, as well as being complicated to use and maintain. The best secure communications solution for mobile devices must, therefore, be hardware-based and be equipped with a tamperproof capacity. Furthermore, given the extremely limited resources of a mobile device, a symmetric key system is more desirable than an asymmetric key system. The mode of sharing session keys should be designed to ensure rapid initial synchronization and stability in the event of a replay attack. Finally, when block encryption algorithms are used, the time delay caused by encryption and decryption must be less than 30msec, as any delay beyond this is not suitable for VoIP calls. If stream encryption algorithms are used, the security module needs to be equipped with a self-synchronization capacity so as to prevent packet losses or repetition from making voice communication impossible.

REFERENCES

[1] Dan Wallach, Smartphone Security: Trends and Predictions, Rice University, Feb. 2011.

[2] http://www.l-3com.com/cs-east/ia/smeped/ie_ia_smeped.shtml

[3] Yeonsu Kim, Joonhee Youn and Hyun Park, "A Voice Encrypted Communication Module for Mobile Communication Terminals", KR 10-2007-0089750, Sep. 2007.

[4] Jae Hyung Joo, Jeong-Jun Suh, and Young Yong Kim, "Secure Remote USIM (Universal Subscriber Identity Module) Card Application Management Protocol for W-CDMA Networks, " *ICCE*, Las Vegas, NV, pp. 101-102, Jan. 7-11, 2006.

[5] Global Mobile Data Traffic Forecast 2010-2015, Technical report, CISCO VNI.

[6] Karie Gonia, Latency and QoS for Voice over IP, Technical report, SANS Institute.

[7] William Stallings, Cryptography and Network Security, 4th Ed., Prentice Hall, Nov. 2005.

[8] Kerberos: The Network Authentication Protocol, http://web.mit.edu/Kerberos/

[9] The Internet Key Exchange (IKE), RFC2409, http://tools.ietf.org/html/rfc2409

 ISBN: 978-1-61208-171-7