

The Value of Hardware-Based Security Solutions and its Architecture for Security Demanding Wireless Services

Rongzhen Yang*, Liang He*, Shoufeng Yang*, Forni Gary*, Fei Liu**,
Jiayue Chang**, and Manxue Guo**

* Intel APAC Security Program Office, Shanghai, China

** R&D, China Mobile Communications Corp, Beijing, China

Abstract - In the paper, the generic vulnerability of security software solution is depicted. For selected wireless services such as Device Management, Digital Rights Management and M-Payment, the vulnerability of software solution and the advantage of hardware security solution are identified and analyzed, finally a hardware based wireless security platform for 'Best Practice' is proposed.

Keywords: Hardware Security, Mobile Device, Wireless Service, Wireless Terminal Security

1. Introduction

Currently, the cellular industry is in the edge of a historic boom for the rich of new wireless services, to step up ARPU and attract the new users. However, some of top hot wireless services, including DRM [1], DM [2], and M-Commerce [3][4][16][17][18], are highly concerned for the security vulnerability of the current software-only security solutions [5]. They are increasingly requesting more security to mitigate the risk from viruses, malicious applications, the theft of contents and privacy [6] [7]. Indeed, pure software DRM solutions have never lasted more than a few weeks without being broken and pure software M-Commerce and DM solutions are not safe from key loggers and other Trojans.

The code of SW security solution and relevant sensitive data are stored and executed in the open platform memory and application processor that are thought of exposure to un-authorized entities which could be made of software or hardware [8] [9]. This is the reason why the current most of commercial operating system and applications are vulnerable to so many security threats [13] [14]. By comparison, a HW implementation allows realizing a well defined security boundary, which can be formally analyzed by well established security criterion [10] [11] [12] [15] [19].

The remaining portions of this paper are arranged as follows:

1. Generic Vulnerability of SW Security Solution
2. The Vulnerability Identification and Analysis of SW Security Solution and the Advantage of HW Security Solution for Selected Wireless Services
3. A Recommendation of HW-based Security Platform for 'Best Practice'

2. Generic Vulnerability of Security SW Solution

The security threats of mobile terminal platform are essentially belonged to the tampering of code and the exposure of sensitive data. The result will be the different kinds of frauds and the denied of some wireless services. So platform should be able to detect tampered code and as well protect sensitive data. However, SW-only security solutions have three generic vulnerabilities:

Firstly, SW-only based code can be tampered. It makes impossible to ensure a SW-only based security solution code to be able to detect the tampering of other SW based code.

Secondly, the storage of sensitive data should be protected by encryption. But it will be impossible for SW-only security solution to protect encryption key. Because under the circumstance of SW-only based security solution., the root key of the encryption key ring has to be stored in plaintext in open memory. Eventually the situation leads to the meaningless to protect the sensitive data by encryption.

Thirdly, the SW-only security solution of cryptographic algorithm is executed in platform's open memory and open application processor. In the processing of encryption, decryption, and the key

generation, the plaintext key can be fetched easily. At the same time, the medium results can be tampered easily too, result in the un-trusted result of execution.

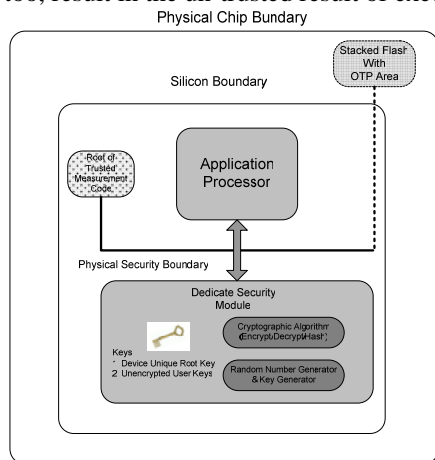


Fig. 1

PHYSICAL SEPARATION OF CRITICAL SECURITY FUNCTIONS

To solve above three vulnerabilities, indeed the dedicated security solution system is required, which is strictly separated from the open application processor and open platform memory to ensure that applications do not affect the dedicated security solution system. The core traits of the dedicated security solution system must be,

- Sealed storage for sensitive data and cryptographic algorithms.
- Atomic operation over crypto services.
- Root of Trust Measurement can not be replaced or tampered.

The notion [5] with the core implementation of HW-based security solution is depicted in Fig.1.

3. The Vulnerability Identification and Analysis of SW Security and the Advantage of HW Security Solution for Selected Wireless Services

As the evidence in previous sections, how to implement security functionality is key to security demanding wireless services. This section continuously analyzes and identifies the vulnerability of SW-based security solution and the advantage of HW-based security solution for DM, DRM and M-Commerce.

3.1 Security Use Case for OMA Device Management for Firmware Updated over the Air (DM FOTA)

The security mechanism of DM relies on the authentication key used to establish the identity of DM server. The distribution server's authentication key (such as the manufacturer's public key) stored in mobile terminal is defined for the authentication and integrity check of distributed code. In a SW-only based solution, the key would be easily tampered in storage and in the processing of downloaded code. The DM client itself, which should authenticate and integrity check downloaded data before installment, can be no difficult tampered without knowing as well. These could represent vulnerability which potentially leads to disasters,

- The server's public key installed on the mobile terminal could be maliciously modified to lead client to trust a rogue DM server.
- The tempered DM client without knowing can bypass the processing of authentication and integrity check of download code and then creates the possible intrusion of mal-ware in firmware.

If using HW-based security solution, when in manufacturing mobile terminal, the hash of manufacture's public key (or DM server's public key) can be stored in ROM or OTP flash; then manufacture's public key can be stored in flash (open memory); the signed hash of DM client image can be stored in flash (open memory); finally DM client is installed in client.

When executing DM FOTA in client, firstly the mobile terminal does authentication and integrity check of public key and DM client,

- The terminal does the hash of the public key stored in open memory then compares the hash with the one stored in ROM or OTP flash. If the comparing result is not the same, public key is tampered.
- The terminal uses the good public key to un-sign the signed hash of original DM client image, hashes DM client image, and then compare. If the compare result is not the same, DM client is tampered.

And then the terminal does the authentication and integrity check of downloaded code,

- The terminal uses the good public key to un-sign the signed hash of the downloaded code, hashes the downloaded code, and then compares. If the compare result is the same, the downloaded code is from right server and the code is not tampered.

In addition, all processes of the authentication and integrity check for manufacturer's public key, DM

client, and downloaded data are executed in the dedicated security module.

3.2 Security Use case for OMA Digital Rights Management (DRM)

The security processing of DRM solutions involves a right object to decrypt encrypted content. To authenticate client (the mobile terminal) and server and to protect the right object that includes the symmetric key to decrypt encrypted content, DRM architecture involves PKI infrastructure that allows the DRM server to use client assigned public key to protect the content encryption key (CEK) and then to use its private key to sign the right object. The purpose of the architecture is to keep the the access privilege of content individualized on a client basis.

Like DM mentioned above, in SW-only security scenario, the client's private key would be exposed and fetched in storage and in the processing of decryption. DRM agent and Media-Player can be tampered and replaced easily without knowing. These could represent vulnerability that potentially leads to breakdowns,

- If user purchased content legally, and the exposed device's private key is fetched, the fetched device's private key can unpack the encrypted CEK. The CEK can decrypt encrypted content. The decrypted content can be distributed widely. Eventual the security model of DRM content can be totally collapsed.
- When DRM agent is tampered without knowing, all keys and content can be easily stolen.
- Media-player (or named as browser in some systems) is used to render the plaintext content and then playback. If it is tampered maliciously without knowing, the plaintext content can be saved by the player then distributed illegally.

Based on the HW-based security mechanism described on Section 4, all keys in client can be securely protected in both of storage and the cryptographic processing. The protected keys and the cryptographic processing in HW base dedicated security module can ensure to detect tampered DRM agent and Media-Player.

3.3 Security Use Case for M-Commerce

A typical M-Commerce involves four parties: the consumer, merchant, M-Payment portal and bank (or other card issuers). Currently, M-Commerce begin hot and attract people's interest. However, customer's concern on security is the top obstacle to inhibit its growth.

The security requirements in M-Payment can be divided into four major categories [4]:

- Authentication: The entities in the transaction must be authenticated mutually;
- Confidentiality: Any sensitive transaction data can't be accessed by any un-authorized parties.
- Integrity Protection: All transaction data during transaction must be protected from being altered.
- Non-Repudiation: All transaction parties can't deny their participated actions.

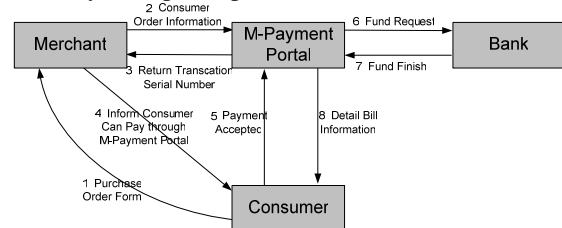


Fig. 2
ONE EXAMPLE M-COMMERCE TRANSACTION PROCESS FLOW

Currently, there are many proprietary M-Commerce implementation models with security mechanism for secure transaction. For all of them, if in the SW-Only security scenario, the secret (the private key or something else) would be exposed and fetched by un-authorized entities, the information used to authenticate the other parties, such as server's public key, can be tampered and replaced easily without knowing, all software components also can be tampered easily without knowing. These vulnerabilities will bring calamitous results:

- The server's public key installed on the mobile terminal could be maliciously modified to lead consumer to trust a forged M-Payment portal, result in the possible fraud and the leakage of consumer's secret information.
- Criminal or criminal organization may use the authorized user's secret to fraud the bank, to steal the money in the account.
- Malicious attacker may forge or tamper the transaction data, bring un-expected loss to consumer, merchant or bank.
- The user's transaction may be intercepted by un-authorized parties, result in the violation of privacy protection.

Based on the Hw-based security platform depicted on section 4, all secret in client can be securely protected in both of storage and the cryptographic processing. The protected keys and the cryptographic processing in HW based dedicated security module can ensure to detect tampered code and data. An all-around protection for M-Commerce in terminal can be ensured, to build up a secure environment for transaction.

4. A Recommendation of a HW-based Security Platform for “Best Practice”

The Root of Trust (ROT) is the basis of wireless trusted platform. It is the component of the platform that can be guaranteed to be trustworthy under all conditions. It constitutes the first element in the transitive trust model. In all cases of power-up, the ROT must be the first thing that executes. In addition, the ROT must be free from being modified or compromised and its execution must be free from interrupted by applications in device.

Fig. 3 depicts a recommendation of a HW-based trusted platform combining with secure SW solution stacks.

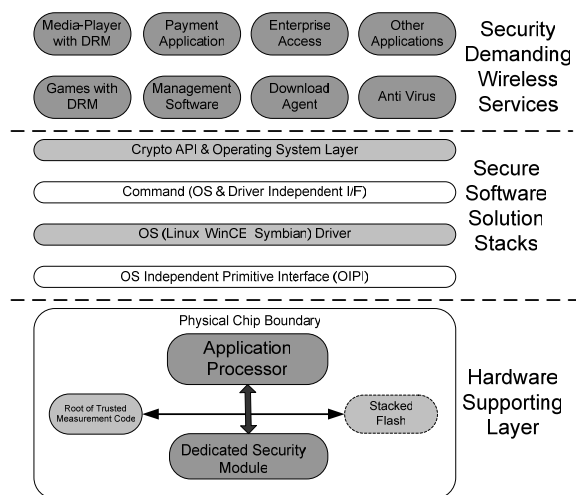


Fig. 3

HW-BASED TRUSTED WIRELESS PLATFORM COMBINING WITH SECURE SW SOLUTION STACKS

In the recommendation of a trusted platform, the ROT consists of dedicated security module and dedicated ROM code for root of trusted measurement. Above the ROT HW based building block, OIPI (OS Independent Primitive Interface) provides a layer which is OS independent to manage the resources, implements authentication and authorization protocols for entities to use Dedicated Security Module.

OS related driver and Crypto APIs provide a standard interface for OS and applications to access the trusted platform resources, without the requirements of specific information about the split of functions between hardware and software.

Based on the platform, the secure-demanding applications, such as DRM, DM, Anti Virus, etc, can setup a clear solution with complete security.

4.1 HW ROT Building Blocks

4.1.1 Trusted Boot ROM

A dedicated security ROM is used to contain the trusted boot code, to prevent the code being modified, and to ensure the code always executed on a power up event firstly, cannot be bypassed.

The trusted boot code initiates the trust boot process that based on the transitive trusted model, to validate the integrity of the platform and boot the device to a “known good” configuration. In the trusted boot process, any platform modifications caused by coding errors or malicious software, such as viruses, will be detected, and the measured value are stored securely that can be used to be presented to some entities to prove the platform trustiness.

In addition, to enable a secure service, the Trusted Boot ROM provides the interface to be invoked by the OS or application at any time after power on, to restart trusted boot process. This allows the internal or external entities to ensure the health of a platform before starting strict security-demanding services.

After the boot process is completed, control is transferred to the OS. Healthy virus scanner, which is validated in the trust boot, may be employed as additional security measure.

4.1.2 Protected Storage

The Trusted Platform solution provides two kinds protected storage to protect sensitive information such as symmetric keys, asymmetric public/private key pairs, passwords, and digital rights management data, etc.

The first one is the hidden physical protected storage that is located in Dedicated Security Module. It is not observable, and hardware tamper-resistance. It provides the space to contain limited secrets in plaintext. Some of them (protected plaintext secrets) are never departed from Dedicated Security Module, such as root key used to support key hierarchy management. Some of them are required for hidden process executed in Dedicated Security Module.

The second one is encryption protected storage in the system storage space that is outside of Dedicated Security Module, such as stacked flash. All sensitive information in storage is protected for both privacy and integrity by using encryption and signature, and the related process are executed inside of Dedicated Security Module. In this way, it is impossible for an attacker to observe the sensitive data or modify it without knowing.

By using the platform specific access control model that may be defined in OS, the hidden physical protected storage and the encryption protected storage are well-knit to provide an integrated security storage protection solution for wireless trusted platform.

4.1.3 Dedicated Security Module

An on-chip dedicated security module with its own processor provides a trusted, non observable execution environment to process confidential data. The module includes a suite of cryptographic engines to support a core set of cryptographic services used to build a higher level of security solutions such as protected storage, security protocols and services, and platform attestation. A key hierarchy structure is also generated in the Module so both asymmetric and symmetric keys can be protected when in use and in store outside the module by encryption.

The objective of the dedicated security module architecture is to provide a trusted execution environment for security processing, in where:

- Cryptographic code and keys, intermediate results of cryptographic operations, the module states, the trusted boot measure result, and other sensitive data are protected from observation and modification by the applications processor.
- Security operations run to completion without interruption and observation by external agents

The need for a trusted execution environment is being driven by the TCG for terminal security platform and Open Mobile Alliance (OMA) for security demanding wireless services such as DRM & DM.

4.2 Secure SW Solution Stack based on HW ROT Building Block

Secure key management, secure protection of intermediate results, atomicity of operations, provability and determinism are unique characteristics of HW based solutions. A security architecture that combines the flexibility of software and the robustness and resistance of HW has the potentiality of providing the best of both worlds. In this case, the HW would be treated by the SW as an opaque entity that could programmatically perform trusted pre-defined operations via an interface to benefit all wireless security demanding services. Major benefits list as below:

- HW based solution protects keys and other sensitive data from being exposed, to ensure a complete security for DM, DRM, M-Commerce, etc.
- SW layer supports a widely used Cryptographic APIs for developers to take advantage of the

strong security features provided by the HW security layer, and reduce the time-to-market.

- HW based Trusted Boot process ensures the platform integrity, to reduce the risk of handset abnormal behaviors.
- Dedicated Security Module has the processor that handles security operations, will improve system performance by reducing application processor workload.

5. Conclusion

In the paper, generic vulnerability of SW-based security solution is depicted. For selected security demanding wireless solutions such as DM, DRM, M-Commerce, the vulnerability of SW security solution and the advantage of HW security solution are identified and analyzed. Finally HW-Based security platform architecture for 'Best Practice' is proposed.

Obviously, software-only security solution just provides obfuscation. This is why software alone is not robust enough to withstand many common attacks. For any type of encryption, it is absolutely critical to protect the keys. But the plaintext keys can not be safely protected in memory or CPU registers so it is perhaps easily reachable by a simple software debugger.

HW-based security solution is very hard to be breakdown. If it is breakdown, the cost paid for relevant attack mechanism will be extremely high. The kind of high cost will significantly reduce or even eliminate the incentive to tamper with device. Therefore HW-based security solution provides the necessary security strength to enable new and pervasive wireless solutions for diversified mobile subscribers. Hence a robust HW-based security is what the end users need to protect their usages of security demanding wireless services.

6. References

- [1]. Open Mobile Alliance, "OMA Digital Rights Management V2.0", Sep.15, 2005, http://www.openmobilealliance.org/release_program/drm_v2_0.html
- [2]. Open Mobile Alliance, "OMA Device Management V1.2", Aug.28, 2005, http://www.openmobilealliance.org/release_program/dm_v1_2.html
- [3]. Mobile Payment Forum, "Risks and Threats Analysis and Security Best Practices for Mobile 2-Way Messaging systems", May 13, 2003, http://www.mobilepaymentforum.org/pdfs/MPF_Security_Best_Practices.pdf
- [4]. Mobile Payment Forum, "Mobile Payment Forum White Paper", Dec, 2002,

- http://www.mobilepaymentforum.org/pdfs/mpf_white_paper.pdf
- [5]. Intel Corp, "Intel(R) Wireless Trusted Platform: Security for Mobile Devices White Paper", 2004, <http://www.intel.com/design/pca/applicationsprocessors/whitepapers/300868.htm>
- [6]. Trusted Computing Group, "Trusted Computing Group Unveils Mobile Phone Security Use Cases", Sep 2005, https://www.trustedcomputinggroup.org/news/press/tcg/2005/Press_Release_for_CTIA.pdf
- [7]. Trusted Computing Group, "Mobile Device Security and Trusted Computing - Next Steps", Sep 2005, https://www.trustedcomputinggroup.org/groups/mobile/MPWG_Primer.pdf
- [8]. Trusted Computing Group, "Mobile Phone Work Group Use Cases", Sep 2005, https://www.trustedcomputinggroup.org/groups/mobile/Final_use_cases_sept_22_2005.pdf
- [9]. Trusted Computing Group, "Mobile Device Security and Use Cases FAQ", Sep 2005, https://www.trustedcomputinggroup.org/groups/mobile/Mobile_Phone_FAQ_Final_092505.pdf
- [10]. Trusted Mobile Platform, "Hardware Architecture Description", Oct 2004, http://www.trusted-mobile.org/TMP_HWAD_rev1_00.pdf
- [11]. Trusted Mobile Platform, "Software Architecture Description", Oct 2004, http://www.trusted-mobile.org/TMP_SWAD_rev1_00.pdf
- [12]. Trusted Mobile Platform, "Protocol Specification", Oct 2004, http://www.trusted-mobile.org/tmp_protocol_spec_rev1_00.pdf
- [13]. Heiko Rossnagel, Tobias Murmann, "How Secure are Current Mobile Operating Systems?" Eighth IFIP TC-6 TC-11 Conference on Communications and Multimedia Security, <http://sec.isi.salford.ac.uk/cms2004/Program/CMS2004final/p2a2.pdf>
- [14]. Arto Kettula, "Security Comparison of Mobile OSes", HUT TML 2000, <http://www.tml.tkk.fi/Opinnot/Tik-110.501/2000/papers/kettula.pdf>
- [15]. Mobey Forum, "Mobile Device Security Element", 2005, <http://www.mobeyforum.org/public/pressreleases/Mobey%20Forum%20Security%20Element%20Analysis%20Summary%202005.pdf>
- [16]. Mobey Forum, "Mobey Forum White Paper on Mobile Financial Services", 2003, http://www.mobeyforum.org/public/material/Mobey%20Forum%20White%20Paper%20on%20Mobile%20Financial%20Services%20v1_14.pdf
- [17]. Mobey Forum, "Preferred Payment Architecture for Local Payments", Sep 2002, <http://www.mobeyforum.org/public/material/Local%20Payments%20Discussion%20Document%201.0.pdf>
- [18]. Mobey Forum, "The Preferred Payment Architecture Technical Documentation", Jun 2000, <http://www.mobeyforum.org/public/material/PPATechnical.pdf>
- [19]. Evgenia Pisko, Kai Rannenber, Heiko Roßnagel,

"Trusted Computing in Mobile Platforms", Sep 2005, DuD • Datenschutz und Datensicherheit <http://www.wiiw.de/publikationen/TrustedComputingInMobilePlatfo1479.pdf>