SWPY003 – January 2002
White Paper

Reducing the Security Threats to 2.5G and 3G Wireless Applications

Sunil Hattangady, Program Manager, OMAP Security
Chris Davis, Wireless Security Architect
Texas Instruments Wireless Terminals Business Group

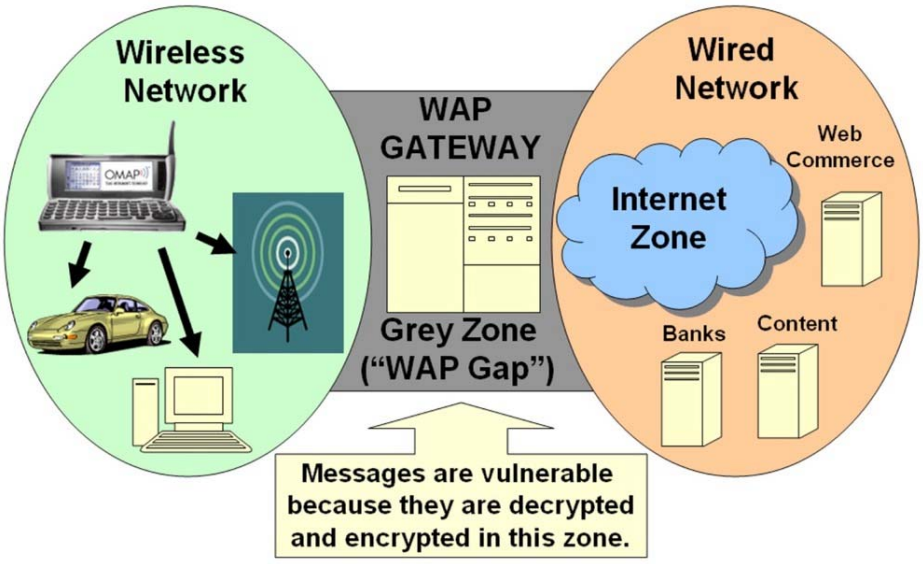
Introduction

The fragile security of 2.5G and 3G wireless applications was abundantly evident in Japan recently when malicious e-mails to wireless handsets unleashed a malevolent piece of code which took control of the communications device and, in some cases, repeatedly called Japan's national emergency number. Other cell phones merely placed several long distance calls without the user's knowledge while others froze up, making it impossible for subscribers to use any of the carrier's services.

Incidents like this one and others involving spamming, denial-of-service, virus attacks, content piracy and malevolent hacking are becoming rampant. There's no doubt about it. The security breaches that have posed a constant threat to desktop computers over the last 10 years are migrating to the world of wireless communications where they will pose a threat to mobile phones, smartphones, personal digital assistants (PDAs), laptop computers and other yet-to-be-invented devices that capitalize on the convenience of wireless communications.

Unfortunately, protecting wireless communications and the applications that use this medium will be more difficult than securing desktop computer applications. Unlike wireless devices, desktop computers have limited and identifiable points of entry, and these entry points can be controlled and safeguarded. But with wireless communications, important and often vital information is often placed on a mobile device that is vulnerable to theft and loss. In addition, this information is frequently transmitted over the unprotected airwaves. Now, some new applications like mobile-commerce (m-commerce) require that this critical information be decrypted by a server somewhere in the communications chain before it is encrypted again and forwarded to its destination. (For example, the 'WAP Gap' in Figure 1 presents such a problem.) Every point in the wireless communications chain where information is decrypted represents vulnerability in the security of the system.

Fig. 1: Wireless Communication Chain



Solving the question of wireless security is not simple because the wireless marketplace is far from monolithic. Clamping severe yet unwarranted security measures on an application would

only frustrate users by slowing down the responsiveness of the application. Not all transactions need the security of a Fort Knox. Purchasing a soda from a vending machine with a cell phone, for example, must be fast and spontaneous. Security measures must match the nature of the application to ensure satisfied users and, at the same time, they should be strong enough to instill a sense of trust that the transaction or the download is not jeopardizing personal information, privacy and content ownership rights.

The users of 2.5G and 3G wireless communications will be quite varied, ranging from socializing teenagers to busy road warrior businessmen, harried housewives and many other demographic groups with diverse expectations and requirements. As a result, the applications, types of wireless terminal devices and usage patterns will vary widely. Successful carriers will seek differentiated service offerings that provide a competitive advantage in each market segment. At the same time, mobile device OEMs will want to simplify and reduce development and deployment costs by settling on a basic terminal device platform that is flexible, adaptable, scalable and powerful enough to thwart security threats over an extended period of time.

As the supplier of the basic architecture that is used in the vast majority of 2G wireless devices, TI has the depth of knowledge and real-world experience that is imperative for understanding how critically important security will be to the success of 2.5G and 3G. TI's enhanced OMAP™ architecture is computationally-robust so that it can carry out the most complex and demanding security protocols and algorithms, and yet it can be scaled and adapted to the simplest of security applications. Coupled with its compatibility with a wide range of third-party security software and hardware, the OMAP platform provides a total security solution for carriers, OEMs and users.

Security Strata

For a number of reasons that have been dictated by the marketplace, the security measures associated with a financial funds transfer between international banks would not be appropriate for a download of a movie preview, for example. A one-size-fits-all mentality would burden low-level applications with unnecessary complexities, hampering their spontaneous use. In contrast, high-level transactions or downloads involving significant monetary value will require the strongest security measures even if this means that the execution of the transaction will take a little longer. A slight delay is a small price to pay for a secure transaction. (See Table 1 for the various user benefits provided by wireless security.)

Table 1: Benefits of Security to End Users
▪ **Encouraging a high-level of trust and data integrity to support a wide-range of mobile real-time financial and content transactions over the Internet and VPNs**
▪ **Ease of integration with standard browsers and non-browsers applications**
▪ **High-performance and strong encryption including on-device, disposable key generation to create a highly secure environment for electronic wallet, VPN and mobile office applications**
▪ **Very fast secure transactions for applications involving high data transmission rates as in content and media distribution (streaming media) and other high-end applications**
▪ **Enhanced user experience through transparency, ease-of-use, extended-battery life and a highly secure environment**

To fit the right security measures with individual 2.5G and 3G applications, carriers must balance the expectations of users with the sort of multimedia experience provided by the application and the financial implications of the transaction or download. This will cause applications to fall into one of several strata, which will make up this emerging marketplace.

Low-Level Security Needs

When important or personal information is not jeopardized or when the value of a transaction is fairly low, the security of an application can be adequately safeguarded with low-level encryption techniques and public key infrastructure (PKI) technology. Examples of these sorts of low-level applications include small mobile-commerce (m-commerce) purchases, point-of-sale transactions like entertainment tickets, and transactions that have a small and fixed liability for both of the parties.

Users at this stratum in the marketplace are impatient. Their purchases will be spur-of-the-moment and impulsive. When the alternative to a small m-commerce purchase with a cell phone is a credit card or cash, the ease-of-use and responsiveness of the m-commerce application is just as critical as its security. Keeping the user's expectations in mind, low-level security measures still must maintain the integrity of the information transmitted and received over the wireless communications channel while ensuring the authenticity and non-repudiation of the transaction.

Mid-Level Security Needs

An e-wallet application that stores someone's personal information such as driver's license number, credit card accounts and passport on a mobile device is a prime example of an application with mid-level security needs. Other examples include an individual's personal financial transactions like bank deposits and withdrawals, or the buying and selling of stocks. Downloading copyrighted content must also be protected at this level.

The processing demands placed on the mobile device will increase from low-level applications because more complex encryption and PKI algorithms, in conjunction with a secure boot loader, digital rights management, filtering and anti-spamming software, will be deployed for mid-level applications. Adding stronger security techniques could come at the expense of responsiveness and the general operations of the client device unless on-processor hardware modules can be included in the mobile device's architecture to accelerate security functions.

High-Level Security Needs

Generally, applications with high-level security needs will start with very strong encryption and PKI algorithms, and increase from there. A dedicated hardware/software security module consisting of hardware-based random number generators, hardware-protected memory where root keys can be stored, secure input/output (I/O) channels, and accelerator modules to improve processing performance will be deployed at this level. This could be accomplished by implementing a security module integrated into the device's processor, by using an add-on security card such as a subscriber identity module (SIM)/wireless identity module (WIM) or a smartcard, or by implementing an integrated on-chip security module. This can be supplemented with other add-on functionality such as biometric sensors as options.

Applications at this end of the security spectrum will include those that involve very large monetary transactions, virtual private network (VPN) access and mobile office applications, and content protection for very valuable software or copyrighted video/audio files.

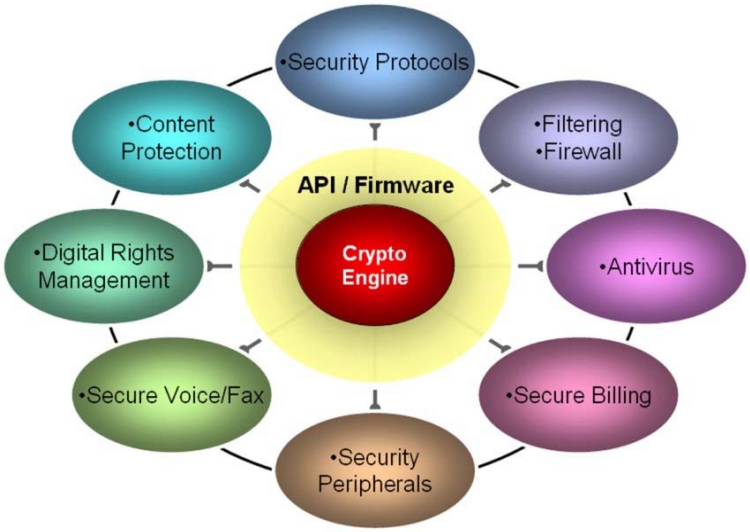
A Security Platform

When it comes to computer and communication security, the one constant is change. Hackers, software pirates and just plain malicious people will continue to ply their trades, improving their skills as technology advances. New security threats in the future will trigger new security techniques and technologies. And wireless carriers will want to deploy these new techniques to protect their subscribers. In fact, a recent study by the Gartner Group's Gartner G2, predicts that service providers who can promote the security of their services will have a competitive advantage in the marketplace. The study found that 86 percent of Americans are very concerned about the security of online financial transaction. (Table 2 -- Benefits of Security to Carriers)

Table. 2: Benefits of Security to Carriers
➢ **Protection from fraudulent theft of services**
➢ **Protection from unauthorized use of mobile devices by someone other than the owner of the device**
➢ **Competitive advantage over insecure services offered by other carriers**
➢ **Protection from denial-of-service attacks**
➢ **Digital rights management protection for copyrighted content, including software and multimedia content such as video and audio files**

As a consequence, service providers and carriers are realizing that the hardware and software architecture of the terminal devices they support must be very scalable so it can support some or all of the components that might make up a complete security solution (See Figure 2).

Fig. 2: Core Security Components



Applications will not require every single security component, but a base platform architecture flexible enough to meet the needs of each stratum in the marketplace will simplify the development and deployment of new user applications and their associated security requirements. This, in turn, will accelerate an application's time-to-market. In addition, the architecture shown in Figure 2 can be adapted to respond to and overcome new threats as they emerge.

At the core of this architecture is a powerful crypto engine surrounded by firmware and an application-programming interface (API) to speed the integration of various security applications and peripherals.

Crypto Engine

The crypto engine for next-generation security will be a combination of hardware and software that is capable of protecting a device's resources from incursion and able to safeguard communications from interception or illicit use. From the carrier's perspective, the crypto engine will be asked to guard against the fraudulent use of the device and the services it provides.

To carry out these tasks, the crypto engine must be computationally robust and equipped with certain hardware-based accelerators tuned to the operations of cryptographic security algorithms. For example, the crypto engine most certainly will incorporate a true hardware-based random number generator (RNG), which forms the foundation of the security of the device. The random bytes generated by the RNG serve as a ‘seed’ to create the ‘secret or private’ keys used for encryption and decryption. The hardware accelerator modules of the crypto engine will empower efficient execution of common cryptographic algorithms (Table 3), including symmetric and asymmetric key functions, as well as hashing techniques for message validation.

Table 3: Common Crypto Algorithms

Type of Algorithm	Algorithm Name
Symmetric	DES, 3DES, RC2, ARC4, AES
Asymmetric	RSA, DSA, DH, NTRU, ECC
Hash	SHA1, MD2, MD5

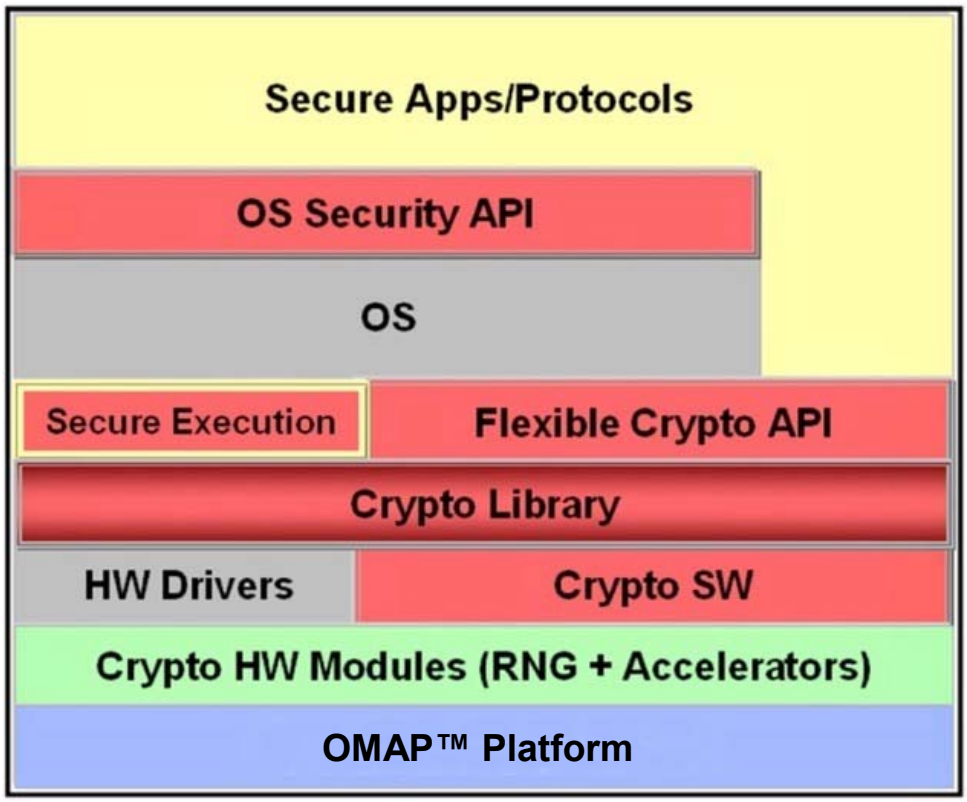
Applications requiring high-level security will demand a secure execution environment to protect the terminal device from incursion. The crypto engine must therefore support a secure mode of operation, where sensitive information, and specifically cryptographic keys, will be protected from access or tampering by un-trusted software.

Security Firmware / API

An important feature of this security architecture is the cryptographic application programming interface (API), that is designed to operate as the interface software component in the system. This API is flexible, supporting a wide range of cryptographic functions and allowing the crypto engine to interface with the higher levels of the system, such as any of the several operating

systems (OSs) currently in use on mobile devices, industry-standard security protocols (SSL, WTLS, IPSec) and interfaces such as Microsoft's CAPI (crypto API) or PKCS (Public Key Cryptography Standards), which perform bulk encryption, key exchanges and hashing, and any add-on security software applications or hardware peripherals. A benefit of this API in the security platform is that once system architecture is implemented, the underlying crypto hardware can be modified or expanded for higher performance without changing any higher-level software (Figure 3).

Fig. 3: Architecture



The firmware layer will include a secure bootloader. This ensures that the system's hardware is not turned over to an OS or other system-level software program that has been maliciously altered as a result of a security breach. When the system is powered up or reset, the bootloader, which is permanently stored in non-volatile memory, initializes the system-level software and brings up the operating system. Unfortunately, software contained in electronically alterable storage like flash memory is a security risk. Viruses embedded in Internet downloads, for example, might modify OS software in flash memory, wreck havoc on the host device and propagate themselves to other mobile devices.

Read-only memory (ROM) is much more secure because it can only be modified by changing the hardware in the device. As a result, the secure bootloader will be stored in the on-processor ROM. As the secure bootloader initializes the system, it will only hand over control to operating or system-level software that has been verified as safe and secure. The secure bootloader can make use of several public key or symmetric key techniques to verify the integrity of the OS software.

Security Applications and Peripherals

The third layer of this security platform architecture includes the industry-standard security protocols that the terminal will need to interoperate with other devices and servers. This layer will also be composed of security applications like anti-virus programs, firewalls, software filters and other software modules which will be dictated by the requirements of the 2.5 and 3G applications running on the mobile device. Some of these security applications will be provided by third parties (Table 4) that have earned a reputation for developing "best-in-breed" security applications.

Table 4: Key Third Party OMAP™ Security Developers

Architectural Providers Level	Key Third Party
Crypto Engine	SafeNet, NTRU, SnapShield, Certicom
Security Protocols	Certicom
Filtering/Firewall	WhiteCell
Anti-Virus	McAfee, WhiteCell
Secure Billing	mSAFE, WhiteCell
Peripherals	AuthenTec, STIP
Secure Voice/Fax	SnapShield, Veratron
Digital Rights Management	Microsoft, 4C, Real, Lockstream, Purplecast, DMOD

Add-on security hardware modules will also be accommodated in this layer of the security architecture. These could include biometric peripherals, such as fingerprint readers or voice scanners, as well as other types of hardware modules, which might accomplish voice encryption and other functions. Some of these hardware modules will be packaged in add-on cards or SIM/WIM cards so that they can be easily integrated into a mobile communications device when needed.

Getting Real

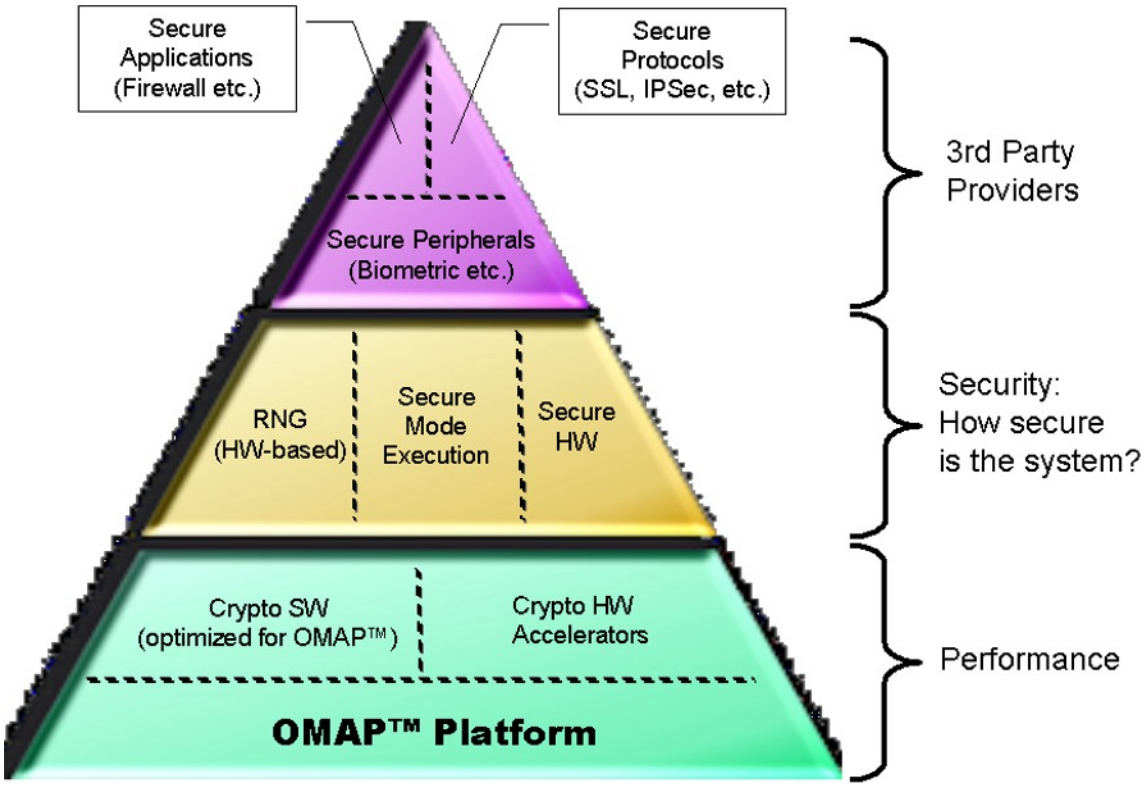
Given the logical structure of this 2.5G and 3G security platform, the OEM of mobile communications devices must determine how best to map this onto a real-world hardware architecture. Fortunately, the abstract security platform maps rather effectively onto the multi-processor architecture that has been deployed in the vast majority of 2G wireless communications devices. This architecture, which was pioneered by TI, includes a digital signal processor (DSP) and a general-purpose processor.

Building on what it has learned from the adoption of its multi-processor architecture by mobile handset OEMs, TI has significantly enhanced the platform and released its next-generation OMAP multimedia platform. Mapping the security requirements of 2.5G and 3G applications onto the OMAP platform demonstrates how effectively a multiprocessor architecture can provide the flexibility and adaptability that will be needed to thwart current and future security threats.

Figure 4 summarizes the value proposition of the OMAP security solution. The foundational element in any wireless security strategy must be computational performance because without

enough processing power to carry it off, latency effects in a wireless security system will leave users frustrated and searching for alternative solutions. At the same time though, handheld devices operate off of battery power and, as a result, they must consume as little power as possible. On both of these counts TI's DSPs and its OMAP architecture fit the bill. Over the last decade, TI's DSPs have consistently established new industry benchmarks for processing power and low power consumption. With a TI DSP as a processing engine, a variety of basic security algorithms including embedded cryptographic software can be executed more efficiently. In addition, high-level security will require dedicated hardware accelerators to ensure responsiveness.

Fig. 4: OMAP™ Security Value Proposition



Based on these sorts of performance-enhancing factors, the next level in this value proposition will include security assurance factors like hardware-based random number generators; secure mode execution and various security peripherals. This level addresses just how secure the system really is. For example, an application requiring a high degree of security might deploy additional hardware-based security measures, like random number generators or hardware-protected memory modules.

The last level in this wireless security value proposition involves those capabilities that are provided by third parties. Many third-party security applications like firewalls, filtering mechanisms and security protocols are currently considered industry-accepted building blocks. In addition, these facilities have demonstrated extensive interoperability with other security measures that are deployed in the market. These third-party capabilities leverage off of the enhanced performance and security features inherent to the OMAP platform, but, because of

their longstanding acceptance in the marketplace, they bring considerable value to the security makeup of wireless systems.

Over and above its value proposition, the OMAP architecture also incorporates extensive modularity for a variety of security algorithms, applications and peripherals. The seamless OMAP platform functions as a common programming environment for a wide range of hardware configurations, each matched to the processing requirements dictated by the user applications and security measures running on a particular mobile device. The portability of third party and proprietary security software, as well as the ease with which hardware-based security peripherals can be integrated into the OMAP platform, will accelerate a mobile device manufacturer's time-to-market. And the flexibility of the OMAP architecture ensures a mobile device can be differentiated on the basis of its security and multimedia capabilities. (Table 5)

Table 5: Benefits of OMAP™ Security to Mobile Device OEMs
▪ Comprehensive solution for transactional-level security and client-level (content protection) security
▪ Portability through seamless integration with major OS platforms to include: Windows® CE, Symbian OS™, Palm OS® and Linux®
▪ Flexibility through low-cost, software-only solutions, or integrated secure hardware solutions
▪ Customizability based on user, device and security level requirements
▪ Compliant with major protocols, including WTLS, SSL and IPSec
▪ Interoperability through functionality with biometric sensors, smartcards and SIM/WIM/SWIM modules
▪ Performance enhancement from high-performance TI DSPs and acceleration with hardware encryption blocks
▪ End-to-end security including true authentication (biometric) at the client

Learning from Experience

Security experts preach that hackers, software vandals, content pirates and other security threats will never be totally eliminated. The tools of the hackers' trade -- the viruses, worms and other assorted collections of malicious code—have a way of morphing and mutating into new forms and shapes. As a result, hacking and other security threats cannot be defeated in the sense that they will never be totally eliminated. But individual security threats can be foiled by innovative and powerful security counter-measures. For mobile wireless communications devices that means identifying the vulnerabilities, adopting a security strategy that takes into account all possible weaknesses, and deploying an architecture that's powerful enough to defeat today's threats yet adaptable enough to meet head-on the unimagined threats of tomorrow.