Mobile Handset Security: Securing Open Devices and Enabling Trust

WHITE PAPER

David Rogers
Industry Relations Manager, OMTP Limited

Abstract

As open mobile terminals become more prevalent across the user-base, developers want to be able to create innovative applications with access to all the features of the handset. Finding a balance which allows developers to cultivate the mobile community, yet which prevents malicious actions by a minority is a difficult yet necessary task. Openness brings new challenges and threats such as Malware. Alongside these new threats is the traditional embedded hacking community who continue to attempt to undermine the underlying hardware and firmware of the device. Establishing underlying trust for users and businesses at the low level of the handset is difficult under these circumstances. The mobile industry, led by the efforts of OMTP is facing these threats in order that users can have a complete mobile experience with the minimum of disruption and so that corporations can use open mobile terminals without the fear of business compromise.

About the Author

David Rogers is the Industry Relations Manager for OMTP, liaising with standards bodies and other members of the mobile phone industry. He is also the Programme Manager for OMTP's Advanced Trusted Environment and Incident Handling tasks. Formerly he led Panasonic Mobile's Product Security initiatives in Europe as well as leading a development team for Customer Engineering. David holds an MSc in Software Engineering from Oxford University and a HND in Mechatronics from the University of Teesside.

About OMTP

OMTP is focussed on improving the mobile customer experience. It works with key mobile operators and vendors to gather, unify and recommend mobile terminal requirements. It is technology-neutral, with its recommendations intended for deployment across the range of technology platforms, operating systems (OS) and middleware layers. Carriers, content providers, middleware vendors, handset manufacturers and users all stand to gain from OMTP's recommendations. More information regarding OMTP can be found at www.omtp.org.

Contents

Why should I care about security?

Security is something that when it works well goes largely unnoticed, but when it goes wrong is spectacularly visible. Security is something I need on my PC because it is connected to the internet and it's a bit wild out there. But do I care as a user whether there is security on my phone?

People care about being able to use their handset in the way in which they want to use it, without fear of malicious interference from anyone else. They want to know that a phone can be fully disabled if they lose it or if it is stolen and that their personal data will be secure. People don't want to have viruses on their phone and they certainly don't want to be surprised by premium rate calls on their bill that they haven't made. Phones are developing and rapidly converging with PC technology and it is important that users have a better, secure user experience in the mobile world than they have had in the PC world.

Security is not just there to protect you from attack. It is also there to protect the businesses and tools that are providing you with a service. It is controversial subject, but quite simply, certain business models are not possible without security. Without adequate security in place, companies wouldn't be able to let you use pay-per-view at home to watch Hollywood films. Equally, the software that you pay good money for would probably not be available if you could just copy it to everyone you know. The music industry has only been able to go fully digital through the use of DRM (Digital Rights Management), opening up huge areas of potential in the future. Today's bank cards and remote access banking methods are only possible because there is security to ensure the bank doesn't lose their money and that you as a customer are protected from having all your money stolen. In this world of fantastic new technologies, there are also significant new threats, from all over the globe – the mobile phone is under attack.

Security issues on mobile devices

Introduction and some history

Security threats against mobile devices come in many shapes and forms. As the complexity of devices increases and mobile operating systems become more open, numerous new attack vectors are presented to a wider audience. The possibilities for attacking handsets should theoretically increase greatly.

Attacks on mobile phones have been occurring for many years; however as traditional device operating systems and software have been of a proprietary design, attacks have been dealt with in-house by individual companies and there has been limited external exposure to the issues. As such, an underworld of hacking against embedded devices has largely gone unnoticed by the security community. This area of hacking is specialised and highly technical, requiring a combination of electronics and software skills. A whole industry has been built-up around working out how to remove SIM locks, change IMEIs and modify other information on mobile devices. Some of this is driven by black market activity in stolen devices – street theft of mobile phones is a major problem in cities across the world and mobile phones are one of the most stolen products during transportation, according to the Transported Asset Protection Association (TAPA). As manufacturers and operators have increased the security on handsets to combat these attacks, so the hacking community has moved to consolidate their skills and increase their funding to hack the latest phones. This is hacking for profit on a major scale. The mobile phone industry is well aware that efforts to implement increased security measures on devices, utilising new technology and techniques should continue if devices are to remain secure once they are on the market.

Eavesdropping is something that is still with us in the mobile world. Due to the encryption on the air interface, it is extremely difficult in most countries to eavesdrop on calls without authorisation. This however, remains a particularly high target for attack for a variety of attackers from intelligence agencies to journalists to suspicious spouses. Attacks against call encryption continue to be developed but as devices have become more open, targeting has shifted to being able to plant something on the device that can either bug a call or enable messages to be viewed remotely. Mobile data theft has already been occurring quite frequently against celebrities in order to cause embarrassment, with Paris Hilton's synced mobile phonebook and

photograph data stolen and private mobile phone pictures of celebrities such as Vida Guerra and Charlotte Church circulating the internet and media.

New technology on mobile phones has brought new breeds of attacks. The introduction of Bluetooth on phones opened up a new attack vector which was exploited in a number of different ways. The anarchistic pleasure of 'Bluejacking' somebody soon spread, with electronic signs and commuters repeatedly falling victim. More sinister Bluetooth attacks came, amongst them, 'Bluesnarfing' - a silent way of extracting the phonebook and other contents of a device. Developments continue on Bluetooth hacking and further attacks should be expected in the future.

The introduction of open OS's such as Symbian S60 ushered in an interesting new area for hackers. It wasn't long before simple viruses appeared that could disable functionality or cause mischief in the handset which subsequently caused unease amongst users.

The source code for the initial viruses was soon being spread around the virus writing community - the source code of 'Cabir' was included in the annual newsletter of the devilishly named 29A group (29A is 666 in hexadecimal). The development continued throughout late 2004 with the aim of creating a mass outbreak. The closest this came was with the creation of 'Commwarrior'. This application was capable of sending itself stealthily via MMS to everyone in the victim's phonebook. This worm did spread, but in a sporadic fashion. Only a small percentage of the mobile market was running the Symbian OS and the user still needed to be fooled into installing the application. Nevertheless, some users found bills for hundreds of pounds on their doorstep and the media spelt impending doom for all mobile phone users.

Despite this doomsday scenario, the growth and spread of mobile malware has to date been somewhat limited. This is partly to do with increased user awareness, the efforts of the industry on protection measures and also to the still relatively fragmented mobile market which is not conducive to the spread of malware.

Although little impact has thus far been felt by the average consumer, the different types of attack and the associated media attention has concerned corporations, particularly about the threat of a virus or worm that could damage their company. Given their experiences of past PC viruses and the havoc they can wreak, the enterprise sector continues to keep a vigilant eye on developments.

Threats and Attacks

The ultimate victims of attack are the user of the device, or a business or the network that the device is connected to. There are many different ways of attacking the handset or the user but the majority of attacks can broadly be categorised into the following:

- **Data Theft** – Taking someone's pictures, messages, phonebook or file data without the permission of the owner
- **Copyright Abuse** – Violating paid for content – e.g. by recording and distributing pay-per-view films
- **Device Theft** – phones are a highly lucrative item and phone theft is a massive problem. Re-enabling stolen phones is a key driver for hacking phones
- **Theft of Service** – stealing someone else's minutes or data or getting free service from the network
- **Denial of Service** – preventing normal operating of a phone or preventing the access to or operation of a network
- **Disruptive / Anarchistic Attacks** – attacks in which the intention is to cause upset, distress and disturbance to the user, network or corporation such as a virus (could include Denial of Service).
- **Interception** – listening into someone's calls or getting access to messages / data during transit
- **Facilitators** – some attacks are deliberately designed to create a staging post for other forms of attack.
- **Fraud** – Getting financial gain by deceptive means

The common factor with most of the attacks is the use or abuse of the mobile phone. By building strength in depth within the phone itself, the different avenues of attack which enable these models of abuse are turned off. On their own these issues may not present a big problem, however the distribution and global nature of mobile phones make this highly lucrative on a grand scale.

Examples of Attack

Attacks come in a variety of forms but mostly with the same general core goals as listed above. The examples listed below show some of these goals as they have been manifested on the phone itself, often in multiple variations. These represent some of the most prevalent and damaging attacks.

Attack: SIM lock removal, IMEI changing and low level data extraction
Attack type: personal data theft, fraud (subsidy), facilitator (mobile phone theft)

This type of attack has been the main area of focus amongst the embedded hacking community. Removing subsidy locks remains a highly lucrative area and is a multi-million pound business around the world. Changing the identity of the device is a key target too, particularly as the industry and police have reacted to mobile phone theft by blocking stolen IMEI numbers. The hackers have to be able to get into the device at a low level, remove or avoid security mechanisms and then be able to repeat that on a mass-scale. Most often this requires software to be written which leverages the attack automatically to make it distributable for a general, relatively unskilled audience.

The resource, time and the complexity of the tools that go into attacking the device at this level mean that once one high security area of the device has been cracked, it is often possible that other opportunities are available for the hacker. Data extraction tools are often a side benefit and are even sold to some law enforcement agencies that use them for forensic data extraction.

Attacks at this level are extremely critical to the whole integrity of the device. Services and applications running above a compromised platform are potentially at risk, through various methods of subsequent attack on applications and data on the device. These hacking solutions are deployed for sale on the internet and are frequently purchased by shops and stalls offering unlocking to the general public.

'Hacking Boxes' - Hardware created by hackers to protect hacking software and sold online to be used for SIM unlocking and IMEI changing.
Top: A Griffin Box from 2004, Bottom Left & Right: Griffin Boxes awaiting shipment

Attack: Attacks on enterprise data and the security of corporate information
Attack type: Data theft

With the increased capabilities of handsets in the area of enterprise applications and email access, companies are increasingly becoming aware of the potential issues that could arise if an employee's handset becomes infected with malware, gets lost or stolen or even that the handset could compromise the company network remotely. Mobile phones are a weak point in a company's compliance to international security and auditing standards such as ISO/IEC 27001 or Sarbanes-Oxley. Corporations are keen to solve this issue, by bringing corporately issued devices under the same umbrella as other technology that connects to the internet or has access to business sensitive information. Without the ability to disable or regulate some features of commercially available handsets, the purchase of them is prohibitive to corporations.

One recent attack on the iPhone allowed remote access through the browser to be able to extract information such as the phonebook. Such attacks could be critical to a business.

Attack: Bluesnarfing, Bluejacking and other Bluetooth attacks
Attack type: Interception and Disruptive / Anarchistic

Between 2003 and 2005, a number of Bluetooth attacks were discovered, mainly due to flaws in the implementation at a low level in the handset. The most infamous of these, 'Bluesnarfing', was demonstrated at the House of Lords in Britain in 2004 by security expert Adam Laurie. The attack exploited a flaw which allowed remote access to controlling functions in some handsets allowing the interception of calls and extraction of data including the phonebook. Software upgrades were not widely available outside service centres so the flaws remained in handsets in the field, leaving them open to attack.

Bluetooth remains a good anonymous transport method over short-ranges and 'Bluejacking' for fun is still prevalent. The Bluetooth Security Expert Group has done an excellent job of implementing a testing and verification process against particular flaws and issues.

Attack: FlexiSpy
Attack type: Interception

One particular application that could be used to intercept calls is called 'FlexiSpy'. This tool provides a number of features which can be used to spy on people. It also tracks location to a certain extent by accessing the cell ID information which pertains to which base station you are connected to. It comes in a number of variant packages, designed to bug rooms, listen into calls and read SMS and email messages. The attack comes when the tool is used for malicious or illegal intent. By its nature, it is not easy to see when the tool is running and so if surreptitiously installed on the device by a third party it is a violation of the user's privacy.

Attack: Commwarrior
Attack type: Theft of Service, Disruptive / Anarchistic

Commwarrior was a worm that spread virally via MMS and also Bluetooth. The most prevalent infection method was that the user would usually receive an MMS from one of the friends with an installation file attached. The social engineering aspects were concerning; the text of the messages would vary from infection to infection, enticing users to install the file. Also, people naturally trust their friends, so despite installation warnings and prompts about untrusted software, a lot of people installed the application (if they had a Symbian phone that it could be installed on). The application was bogus and launched Commwarrior. In most cases, the user was completely unaware that the device was sending MMS's to everybody in their phonebook. The device did not prompt the user that it was sending the messages. The user was also unaware that the application was attempting to spread via Bluetooth. A clever extension of Commwarrior could incorporate the sending of premium rate messages – an attack that was manifested in another piece of mobile phone malware in Russia called RedBrowser.

RedBrowser

The attacks listed above are by no means exhaustive. The principles employed however are broadly similar. It is these underlying factors and drivers that have been and continue to be addressed within the mobile phone industry.

Security developments so far by the mobile phone industry

Phones in general have always been an interesting target, from the early days of 'Phreaking' (attacks on land lines allowing the user to get free calls) to today's complex call interception methods. The design of GSM and 3G networks very specifically had security in mind. The existence of the SIM card is a direct result of this and has proven to be a successful security token, preventing lots of potential fraud, with excellent scope for the future. The success of digital telephony meant an explosion in the amount and types of attack, but the mobile industry has not been at a standstill. A lot of work has been done on different fronts to secure mobile phones. Some key developments and why these help are listed below.

Phone Theft

The industry has responded to the massive increase in mobile phone theft by working together with governments on the issue, helping to create the Mobile Phone Re-programming Act 2002 legislation in the UK and also creating the 'Security Principles Related to Handset Theft' and 'IMEI Weakness and Reporting Process' documents in 2004. This effort, the result of cooperation between the manufacturers, represented by the EICTA Mobile Terminals Group and the operators led by the GSMA Security Group showed that the industry could pull together on security and this spirit has continued in OMTP. The creation of the National Mobile Phone Crime Unit in the UK helped to address the issue of mobile phone theft from a Police perspective. Ultimately, there has been a relative reduction in phone thefts and increased user awareness about the problem.

Trusted Computing

The bit level specifications for Trusted Computing which span both the computing and mobile world have come to fruition; the TCG Mobile Phone Working Group specifications were published in June 2007. These will help to provide more protection to the user, for example for securing data confidentiality and also aid future business models to be executed such as pay-per-view TV and mobile banking.

GlobalPlatform have defined specifications which would enable the provision of secure software platforms, further developing Device Application Security Management. These kinds of implementations would significantly move mobile technology forward with phones being able to securely run transport ticketing, view new release films, and the latest TV series. Users would be able to also download and run high quality games and music on the go from the top publishers.

Prevention of Interception

ETSI & 3GPP have continued to specify the network security architecture for mobile telephony. Their work in the area of air-interface security algorithms and enhanced protection of the SIM card has been extremely successful in the face of determined attack, giving people call privacy and increasing user trust in the integrity of digital mobile networks. It is testament to the work of these individuals that the A5/1 air interface algorithm has stood the test of time much longer than other cryptographic algorithms and it is only now at the start of a phase-out for next generation security for call encryption.

Technological Developments in the Handset

Manufacturers have responded to repeated targeting of SIM locks and IMEI numbers by increasing security in devices and building up significant levels of security expertise in-house. This has provided the ground-work for secure open devices and means that the mobile industry is more equipped for the opening up of mobile phones to 3rd party applications and new types of networks than the PC industry was at the end of the 1980's.

At a lower level, because of manufacturer and mobile network operator demands on increased security, chipset manufacturers have developed new hardware enabling enhanced technological security on the handset, which in turn has increased the drive and knowledge for having high-security features on future phones.

Technology in the mobile world moves extraordinarily quickly. Devices are rapidly converging with other technologies and through this there are new opportunities but also new threats. One thing remains clear, no matter what security is implemented on a device, there will always be somebody willing to invest the time, effort and expenditure necessary to break it open.

What are OMTP doing and how are they doing it?

The Open Mobile Terminal Platform was established in 2004. Its aim is to make things easier and less complicated for the consumer. The main domain areas are security, user experience and device management. The organisation has around 40 members from across the industry including operators, manufacturers and hardware and software vendors. OMTP has worked on a number of security related recommendation documents from both its security and hardware working groups.

OMTP Application Security Framework (ASF)

OMTP's Application Security Framework is designed to create the policy framework under which applications can run in a secure manner on an open mobile device. The OMTP task defragments the way that application access policies were handled previously by the mobile phone industry, providing the consistency that is so vital in this area.

The matrix policy defines a number of levels of access to features of the phone, based on levels of trust. These trust levels are applied by checking the digital signature of the application to verify its source and the integrity. The underlying technology to enable this is a Public Key Infrastructure (PKI) solution - the use of digital certificates allows differentiation between the different signatories and therefore the level of trust. This allows manufacturers and operators to apply the principle of least privilege in the way that the handset deals with applications (in this case, only allowing an application to access the parts of the phone it really needs to run properly). The policy itself is protected by the trusted environment of the device as this is liable to be attacked using embedded hacking methods.

Applications created by the manufacturer of the device are given the highest level of trust and therefore the most access. An unsigned or self-signed application provides the greatest deal of risk to the user and the network which the device is using so therefore has the least amount of access to the device.

Overall, there are five trust levels, from unapproved to manufacturer approved. Some of the security sensitive functions within the matrix include:

- The ability to start applications automatically
- Access to Bluetooth
- Access to location data
- Access to keys pressed by the user
- Access to data stored on the SIM.

The use of appropriate prompting allows the user to make further decisions on what to trust; a manufacturer signed application would have extremely few prompts, but an unapproved application would have more given the relative risk to the user. The user would be able to see, for example: when a request to send a message or access location data was made, he would then be able to decide for himself whether that was the intended, acceptable behaviour of the application.

All of the protective mechanisms defined by the ASF combine to narrow the window in which malicious developers can operate, ultimately removing any incentive for them to attack the device. Their aims of causing damage to the device, theft of data or service and spreading havoc across the mobile world cannot be achieved.

Social Engineering

Having locked out malware writers from technically generating an attack through use of the ASF, they are forced down the route of duping the user into doing something they don't want to. This is called social engineering. There is always the risk of social engineering by malware writers to coerce the user into accepting prompts but this can be mitigated in two forms. The first is that the number of applications certified by the signing schemes should increase, meaning less prompts for the user as these will then be trusted to a certain degree. Secondly, the prompts provided via the ASF should have value – if a message needs to be sent by the device, the user is likely to be charged for that and for an unapproved application, the user would want to know if they would be put at risk. OMTP carried out a usability study on prompting during 2006 the results of which are incorporated into the ASF recommendations.

Good Application goes Bad

It is important to note that this policy and the accompanying process of certification and revocation (as described in the OMTP Signing Schemes

Recommendations) not only provides the framework for protection against intentional malware (as described above) but any application which may unintentionally begin to exhibit malware-like symptoms through a bug in software. Such bugs have the potential to cause damage to the user financially and through a reduction in handset usability, but also have the potential to cause network issues, for example if there were constant connection requests it would effectively be a Denial of Service attack on the network. If there were large numbers of devices with the same problems, it could become an extremely serious, yet unintentional Distributed Denial of Service attack.

Increasing Uptake of Application Signing

Although the ASF works hand-in-hand with the Signing Schemes process, it is clear that not all applications will go through a process to digitally sign them by a trusted third party. Indeed this is the reason why self-signed applications have the least access to security sensitive functions. There still remain issues with how to re-sign and approve applications quickly if there is a requirement to issue a software update. This is one of the key issues in developer uptake of signing their applications. The increase in automated analysis and signing mechanisms, plus lowering the cost of entry to developers should alleviate this problem in the future. ASF is a pioneering piece of work which will bring increased importance to the signing process.

Applying a Corporate Application Security Policy to the Framework

The ASF offers mobile network operators potential beyond the basic framework. A corporate customer buying devices from an operator could, for example request that the operator sets down a trust level for them based on their own requirements. This allows operators to provide corporations with the ability to extend their Information Systems policies to mobile devices owned by the company.

The ASF is an encouraging factor in the raising of development standards. Mobile network operators have a duty of care over their users and it is important that the industry, whilst opening up platforms for development by third parties, still maintains the level of quality and security expected by the end-user.

OMTP Signing Scheme Requirements

Applications can be digitally signed to provide assurance about their origin. This allows users, manufacturers and network operators to identify which applications may harm their devices based on how much trust they put in the signatory.

The Signing Schemes recommendations complement the ASF by providing a set of requirements to organisations that digitally sign applications. Greater trust can be put in applications that are signed by a scheme, knowing that it has been provably tested and that the application is traceable to its source. If an application is later shown to be malware, it can be revoked so that it cannot execute further on devices it has been installed on and so that it cannot be installed on other handsets. The prime aim for OMTP is to reduce the chance of malware getting onto devices in the first place so the signing schemes document only addresses requirements in the area of application security. Other applications may be trusted and functional but have a major security vulnerability exposed at a later stage. It is important that users are protected from this scenario via revocation of the problem application until the issue can be rectified.

The recommendations are designed to help gain further adoption of signing schemes by setting down common, technology agnostic requirements regardless of whether the application goes through any one of a number of schemes from Java Verified™ to Symbian Signed. The recommendations cover the following items necessary to provide support to the Application Security Framework on handsets:

- Key Management and Certification Processing
- Identification and Authentication
- Application Verification
- Legal Assurance
- Revocation

Operators need to protect their users and their own network from attack by malicious applications. OMTP aims to increase security protection by increasing the amount of applications that are signed on devices through further usage of signing schemes. To effectively do this, it is necessary to unify the industry. The provision of universal guidelines on signing schemes to operators, signing schemes and developers alike defragments this area, making it ultimately easier for developers to get their applications on mobile phones and safer for end-users.

OMTP Trusted Environment: Basic and Advanced

The trusted environment provides the basis of security within the handset. From component level it defines the lowest level of security requirements within the device to ensure that the device itself can be trusted. Within the Basic Trusted Environment (TR0), the requirements laid down reflect the underpinnings of trust for a mobile phone: securing the debug ports of the device, the mobile device ID and binding it to a device, SIM lock, secure booting and updating of the device and the basic DRM security requirements.

The Advanced Trusted Environment (TR1) extends the work of TR0 and is also forward looking. To prepare the way for TR1, a comprehensive threat assessment and analysis was produced. The document utilises new techniques and defines an enhanced level of security to that of the Basic Trusted Environment, addressing the pertinent threats to the device posed by embedded hackers. The recommendations for secure storage and trusted execution environments are the basic enablers within TR1 which interact with a number of extended enablers. These enablers introduce requirements which allow high security services such as broadcast and e-commerce to run securely and in a trusted manner. The extended enablers ensure the following:

- that the link between the UICC (the SIM) and the phone is secure
- that the drivers used for items such as the keyboard and the display are secured
- software can be patched in a flexible and secure manner
- critical data on the device is checked regularly to assure its integrity
- high security cryptographic key generation mechanisms are available for new services such as Multimedia Broadcast and IMS

The recommendations take into account the most up-to-date specifications within industry, including the Trusted Computing Group's Mobile Phone Working Group specifications, 3GPP's work on Generic Bootstrapping Architecture and the ETSI SCP work on a secure link between the device and the SIM card.

Without the trusted environment, none of the other security related functions above such as the Application Security Framework could exist securely. It is these solid foundations that will give developers and investors the confidence to create exciting solutions for handsets such as e-commerce and banking and the broadcast of premium TV and film content. The way in which the documents have been created allows real innovation in the way that Trusted

Environments can be implemented, allowing developers to exploit the strengths of the mobile system such as the inherent security of the SIM card.

It is interesting to note that the initial breach to the SIM lock on the Apple iPhone in summer 2007 could have been prevented with elements of the OMTP trusted environment.

OMTP Incident Handling

Mass outbreaks of issues on technology can have major implications. With billions of subscribers connected to mobile networks around the world, the mobile world is no different. Virus outbreaks have been few and far between due to the proprietary nature of most mobile operating systems, however the number of mobile phone models with an open OS is growing fast. Despite technological measures aimed at preventing incidents, it is inevitable that at some point, an incident will happen. It is also prudent to be prepared for the worst case scenario of a mass-outbreak of an extremely damaging virus-type scenario which could cripple mobile networks worldwide. Malware does not respect national boundaries, nor does it respect the boundaries between different network operators. A piece of malware affecting one operator will almost certainly affect another. Operators need to have a platform to share information on incidents and potential incidents as they occur.

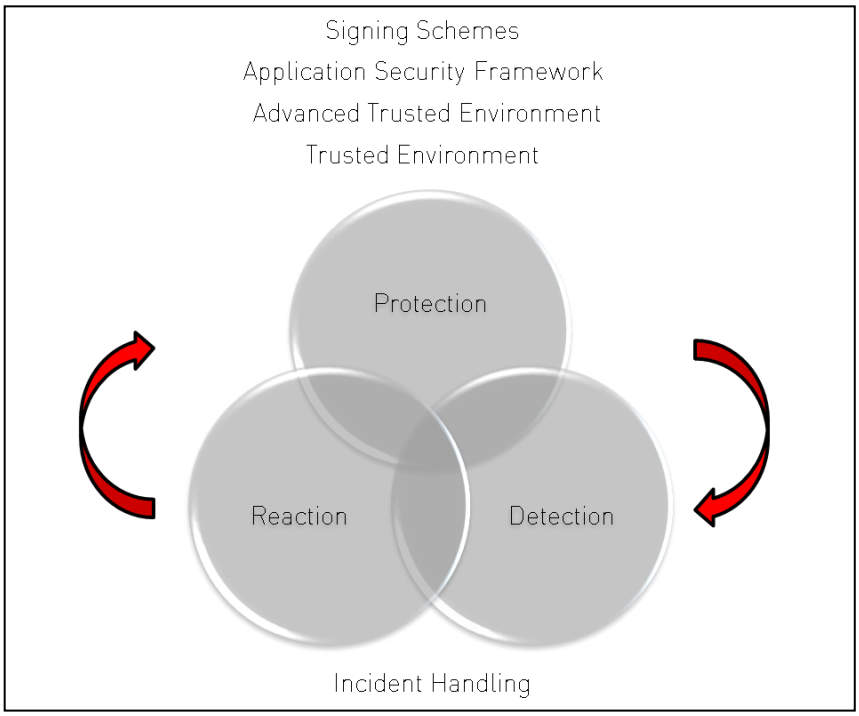
There are many different stakeholders within the mobile industry when it comes to dealing with an incident. The network operators, manufacturers, signing schemes providers, platform vendors, business partners such as anti-virus vendors all have a role in helping to identify and resolve major incidents. OMTP are defining the process that links this together in order for a coordinated and organised cross-industry approach can be taken to resolving major issues. Incident handling procedures in the PC world are well established and run by organisations such as the SANS Institute, as well as part of government sponsored programmes. In the UK, the government created the Warning Advisory and Reporting Point (WARP) which allows UK companies and industry consortia to create their own WARPs, targeted towards their own particular industry. As the mobile industry is slightly different in the way that it is setup and operates, some tailoring of these types of approach is necessary, but the process steps of preparation, identification, containment, eradication, recovery and follow-up remain the same.

Each stakeholder will have their own internal methods of handling incidents and also their own incident handling teams but it is the promotion of these to industry level which needs to be defined as well as defining how information is

collated and shared. It is necessary to define the roles and responsibilities for dealing with potentially multiple incidents, along with communication points and single points of contact (SPOCs). The aim of the industry is to have a centrally handled incident handling process that will be able to deal with problems that occur efficiently, quickly and in a coordinated manner.

Mobile networks now form part of the critical national infrastructure in countries and outage can be particularly harmful. A number of incidents in the computer world during 2007 have shown that a cyber-attack is also part of the war arsenal for many countries and that communication systems will be targeted via the use of Denial of Service attacks, malware and other methods. The mobile world is likely to be part of future military and terrorist strategies in this area.

How the OMTP tasks inter-relate to create a secure device



The handset security lifecycle can be described as the process of Protection, Detection and Reaction as shown in the diagram above. The handset recommendations defined by OMTP fit primarily into the Protection part of the lifecycle as most of OMTP's work is aimed at the design of the product itself.

However, to ensure the most effective security design, one must deal with real world threats and acknowledge that as security expert Bruce Schneier would say, security is a process, not a product. Manufacturers and operators both have their own mechanisms for dealing with issues in the market, but the Incident Handling process helps to complete the circle on behalf of the whole industry. The wealth of knowledge that has been built up can be used to combat the people and organisations bent on breaking the security and targeting the billions of users.

Parts of some of the protective recommendations have their own detection or reaction capabilities, whilst others lay the ground for future activity in this area. The examples below briefly show how the recommendations could work together in real-world situations.

Example 1: Running a Trusted Application From a Third Party Developer
OMTP Recommendations: Application Security Framework, Basic and Advanced Trusted Environment, Signing Schemes

The most effective way in which all the key security recommendations fit together is in the installation of a trusted application from a developer. The developer uses a signing scheme to get his application certified and cleared to access some sensitive features of users' phones. The application runs securely within the Application Security Framework at the Approved trust level, having only access to the features it is supposed to access with the phone prompting the user when necessary to get further permissions. The application and ASF are protected by the underlying hardware security of the device, provided by the Basic and Advanced Trusted Environment. This includes protection to ensure that the certificates used to verify applications have not been replaced, whilst the user can trust that their data is stored and used securely on the phone.

Example 2: Receiving Commwarrior
OMTP Recommendations: Application Security Framework, Incident Handling

The Application Security Framework sets the policy of what applications can access and the necessary user prompts. In this case, the user is warned that the application is untrusted. If the user still goes ahead and installs the application, the user is further protected by the fact that Commwarrior cannot automatically access some very important features, namely the phonebook, the Bluetooth connection and the ability to send MMS's. The user is immediately alerted when the application wants to access these features. This therefore prevents unauthorised sending of data and the application is

limited in its effect as it relies on the ability to send a lot of messages quickly, causing maximum havoc. It also relies being able to send messages by stealth, without the user noticing. The Incident Handling process would be able to deal with any fallout from the problem and also be instrumental in informing the general public if the issue actually did escalate.

Example 3: iPhone SIM Lock Software and Hardware Hacks
OMTP Recommendations: Advanced and Basic Trusted Environment

In August 2007, the first working SIM lock hack was released for the Apple iPhone. This was a good example where the implementation of both the Basic and Advanced Trusted Environment could have prevented this exploit and significantly hampered research and attack efforts. The basics of the attack were that some tools were created that were used as attack facilitators, both on the device and interfacing with it. Extensive probing of the PCB and components was necessary during the research phase and eventual extraction of the firmware of the phone via the debug ports. The next stage was to get the modified software back onto the handset and involved fooling the phone into thinking that the memory was empty and allowing new software to be put on the phone. This also involved a hardware trick which could have been avoided by taking measures to sandwich accessible wire tracks within the circuit board itself. Subsequently, an all software solution has been released, which means that the hack can be replicated and used much more easily by users and shops.

The hack itself could have been prevented by a number of methods defined in the OMTP Basic and Advanced Trusted Environment documents including:

- Locking down the debug ports would have prevented extraction of the firmware and subsequent modification
- Not allowing unsigned software or the replacement of the public keys on the handset would prevent the re-introduction of a modified firmware build
- Secure storage of the SIM lock state itself and the protection of it by a Trusted Execution Environment would have meant that it could not be modified
- The use of a hardware based run-time integrity checking mechanism would have flagged the change to the software and also any subsequent unexpected data change – for example the SIM lock.

Example 4: Installing an Application That Subsequently Goes Bad
OMTP Recommendations: Application Security Framework, Incident Handling, Signing Schemes

There is always the possibility that an application that has been through a signing scheme and is accepted as trusted could subsequently be proven to be malware. There are many reasons as to why this could occur, but the most obvious one is that having shut down the easier routes to spreading malware on mobile devices, a concerted effort will be made to get some well disguised malware through a signing scheme and out into the mass market. It will become a big aim to get something signed that could stealthily work without drawing the attention of the user. The odd text message to a chargeable number here and there, multiplied by millions of users amounts to a very lucrative operation.

If it is discovered that a previously signed and trusted application does go bad, there is always the option of revocation: the signing scheme processes and incident handling process kick-in to ensure that a problem can be nipped in the bud before it becomes an epidemic issue. This mechanism if used appropriately across the industry could become a very effective tool in reducing the incentive to malware authors as the projected profits of such a scheme are vastly reduced.

The examples above give a demonstration of some of the ways in which OMTP is defining the security infrastructure on mobile phones that will underwrite financial risk and engender trust from both a corporate and consumer point of view.

The challenges facing the mobile industry remain tough; good security is a game of continual monitoring and improvement. The mobile industry is showing that it can react to and be proactive about its security measures, given the arsenal at the hands of hackers. Given the different relationship between network and user to the PC world, users feel that they want to be protected by their network operator.

The complete package of security recommendations from OMTP creates a multi-layered security solution providing strength-in-depth to the mobile phone on multiple fronts. Embedded hacking is made extremely difficult through the measures in the trusted environment. Attacks at the application layer are difficult due to the application security policy of the device. Mobile malware authors are forced to attempt getting their malware signed by a signing scheme in the hope that it will not be discovered in testing. If it ever

did get through however, it could be quickly revoked and potentially removed from the actual device.

How does the OMTP's work benefit the end user?

The measures defined in both the Basic and Advanced Trusted Environment push further the technological barrier to entry for embedded hackers on mobile phones. One of the core enablers for mobile phone thieves is the ability to either hack the SIM lock or the IMEI. Large quantities of handsets are shipped out of their country of origin after theft to enable them to work again. The subsequent tools that become available are one factor in enabling devices to function overseas.

A lot of work in this area has been carried out by industry, most notably the 'Security Principles Related to Handset Theft' from the GSMA and EICTA which were 9 principles aimed at securing the IMEI number against hacking. The TR0 and TR1 documents take these measures a lot further, looking at how threats have moved on as hackers have responded to new security measures by manufacturers. The latest measures introduced in TR1 increase the amount of time and effort needed to get into the device and create a viable hacking solution to extremely high levels. All of this creates the grounding for trust in the handset; users are able to place more trust in applications that run on the phone and more easily be able to make decisions about the things they aren't sure about or the use of features which may cost them money. Although the user may not actually know about the existence of the ASF or the trusted environment, they silently benefit from the protection it provides and from the wealth of new services and applications they can use.

How does OMTP's work benefit developers?

Developers going through signing schemes benefit from more access to the functionality of the handset, allowing them to create innovative new applications. Currently, this area is not only extremely fragmented, but access to certain APIs is restricted for security reasons. Operators and manufacturers no longer have to put up the shutters in order to protect themselves. The combination of the ASF and signing schemes process is what allows them to do this, with the signing schemes actively working to prevent malware application developers from getting applications signed. The incident handling process will be there to deal with any issues that do occur.

How does OMTP's work benefit the business community?

The OMTP security recommendations are generated by the world's leading mobile operators, manufacturers and suppliers. This common approach to the global problem of security and trust on devices has created an environment in which companies can be sure that their business models are secured from the ground up. OMTP enables opportunities that previously were not possible due to the associated risk from device compromise. Now is the time that trust can be put in the underlying security of mobile phones.

What does the future hold?

As has been described in this document, security is a continual process. Just as attackers will not stop targeting phones, the mobile industry will not stop in its efforts to enhance security.

Future Technology

The trend of technology is towards further openness on devices. This means that greater functionality will be presented to applications. Naturally, this is a wider surface area to cover in terms of security and fraud management. New types of software execution environments are likely to be developed which will sit on top of the device, perhaps providing a full user-experience to customers. The browser is probably one of the most significant areas that will develop in functionality - currently it is still in relative infancy on mobile phones, bringing future possibilities for attack just as in the PC world now. The increasing number of connections to the internet, combined with fixed rate data charges will result in always-on connections, increasing this exposure.

Access to networks that are out of the scope and control of network operators is another natural evolution of the industry, to suit the users' needs and location. The number of interfaces will increase and the amount of user and corporate data stored on the device will be massive, particularly with the introduction of large-scale flash memory.

The introduction of so-called 'lock and wipe' functionality and further anti-theft measures on devices should help to mitigate issues with mobile phone theft and the theft of corporate and personal data after a device has been lost or stolen.

Future Attacks

In terms of the methods that are used to attack the phone in the future, these will evolve to adapt to the changing technology.

Malware on devices will evolve in a different way to the PC world; the mobile world has the benefit of hind-sight in this respect and this has already been proven by the very fact that mobile phones have not been significantly hit yet. The introduction of firewalls and anti-virus software for mobiles is a benefit to open mobile devices and this kind of technology is now very mature. Viral malware could still be an issue though and more intelligent mechanisms of combining social engineering with longer incubation times could present a real challenge to the mobile phone industry in the future. As technology stabilises device turnover could reduce, increasing the exposure of devices.

Increasing emphasis on a sole asset - the SIM, for high security storage for identity, e-commerce and could see this more heavily targeted. Breakthroughs in technology to attack this kind of hardware are very likely, along with low-level hardware attacks aimed at the core assets of the mobile phone itself.

The mobile phone is a focal point for technology convergence and as such will come under attack from many quarters, for various different reasons and motivations; one technology could be used in a social engineering attack against the other to for example, gain a PIN. The introduction of NFC (Near Field Communication) enabling mobile wallet functionality to phones will be the most likely thing to increase mobile phone theft and security departments will have to deal with all the issues that the banking world currently have to deal with, including things like Phishing and frauds that have only previously affected the network. Single sign-on mechanisms and connectivity to home devices such as heating and security systems are likely to stir more doomsday scenarios in the media. This is not without precedent however; attacks on computer systems have led to shutdowns in the real world such as that which affected a shipping port in the USA.

Targeted thefts of mobiles are unfortunately set to increase as most companies will allow remote access to email and their networks. The difference between is the ease of which access can be got – a mobile device by definition is going to go everywhere the user is, the loss of company data wouldn't be restricted to your user leaving their laptop in a London Tapas bar.

Summary

Security is difficult. Security engineers have to design for the unexpected, against a determined, intelligent and well financed adversary. There are many, many different ways a handset could be attacked and for many different reasons. OMTP helps make the design of security easier. By aggregating some of the industry's leading experts in mobile phone security, their experience has been lent to defining and evaluating the most pertinent and dangerous threats to mobile devices.

Going forward, the public will demand more and more advanced services on open terminals – access to emails, premium games, transferring money, paying for goods, using their phone to get into a concert – the list is nearly endless. Equally, the public will only use these services if they can trust them. To summarise, In order to fulfil these needs securely, the key areas for the mobile industry are addressing are:

- Increasing hardware security: enabling trust for high security services and reducing embedded hacking
- Application security: prevention of the execution and spread of malware and the engendering of trust through certification
- Increasing user and corporate data security
- Developing a unified approach to dealing with security incidents across the industry

OMTP continues to pioneer and develop new security tasks with expert input from across the mobile value chain so that the world's largest mobile network operators can adopt common security recommendations in their device requirement specifications. It's an uncertain world out there, and while no one can predict exactly what will happen, taking the right approach to device security now will help prevent any nasty surprises in the future.

All published OMTP Recommendation Papers, including those referenced here, can be downloaded from www.omtp.org/publications.html. OMTP Advanced Trusted Environment and OMTP Incident Handling recommendations are scheduled to be released in 2008.