A Reference Model for Wireless System Security

Peter G. Cook (HYPRES, Inc. Elmsford, NY, USA pgcook@pgcook.com)

ABSTRACT

Security in a communications system is the attribute that ensures that authorized traffic content is accurately delivered only to intended recipients and that use of system resources is accurately recorded. Introduction of RF links in mobile wireless introduces security considerations beyond those of wireline telephony because interception of the signal cannot be prevented. Implementation of radio links with software defined radio (SDR) technology requires further security measures to preclude introduction of software that can compromise existing security measures.

In this paper we describe a security reference model developed by the SDR Forum Ad Hoc Security Working Group for wireless communication using SDR technology. We also describe a security module architecture developed by HYPRES, Inc. using superconducting microelectronics (SME).

1. INTRODUCTION

Protection needed for content privacy and integrity is only part of the concern associated with SDR wireless implementations, where the full cycle of download, storage, installation, and instantiation (DSII) for software over wireless links must be considered.

DL-SIN [1], prepared by the Download Working Group discusses considerations introduced in the literature, and presents the SDR Forum Reference Model. In this paper we present an overview of that model, and consideration of the architectural implications of security requirements.

At the present there is no single solution to the problems associated with wireless system security. Researchers and system designers working in this area must familiarize themselves with the rich resources available, make the needed trade-offs, and select approaches and components that are best suited to the specific issues at hand

2. REFERENCE MODEL OVERVIEW

To assist in navigating this complex space, the SDR Forum has developed the SDR Security Reference Model shown in Figure 1. This model is derived from study of a large number of relevant works in the area, and it attempts to provide a framework within which specific issues can be placed to put them in perspective with other topics.

The model consists of three levels, representing the wireless link, security threats, and the intervening security provisions for protecting the link from the threats. In each of the levels issues relating to the information source, central network and radio infrastructure, the wireless link, and the remote destination are considered individually.

For simplicity, this version of the model considers a path that is a one-way transfer of information from a central source to a remote location over a wireless link. Most of the security concerns are the same for outbound, inbound, or full duplex operation. Where any differences exist, they can easily be illustrated with variations of the basic model.

Level 1. Communication Channel.

The Communication Channel Level consists of the physical components of the system. The Source includes the entire infrastructure needed to originate information intended for delivery to the Destination, the deployed remote terminal. In the case of a voice call, the source is the origination telephone system, all of the HLR and VLR facilities necessary to locate the recipient and validate account status, and the MSC and BS facilities to originate and maintain the call. For a software module to be downloaded into a mobile, the Source would include the originator of the code, all of the mechanisms for Configuration Management, validation of the remote terminal's right to participate in the network, authorization to proceed with the download, and network facilities to deliver the code, packaged as data, to the Radio Access Network air interface addressed to the Destination.

The central component of the Communication Channel is the Transmission Channel, the wireless link through which the data travels. There are essential differences between wireless links and the Wireline network that was the origin of telephony. One is the mobility provided to the destination terminal. Another is the erratic nature of the link – a higher bit error rate, fading, and loss of signal as the mobile moves out of range. The key difference, from a security viewpoint is that the signal can be intercepted anywhere the signal strength is sufficient, and false transmissions can be made.

The remote terminal is the equipment and software used by the recipient to accept data over the transmission channel and handle it locally.

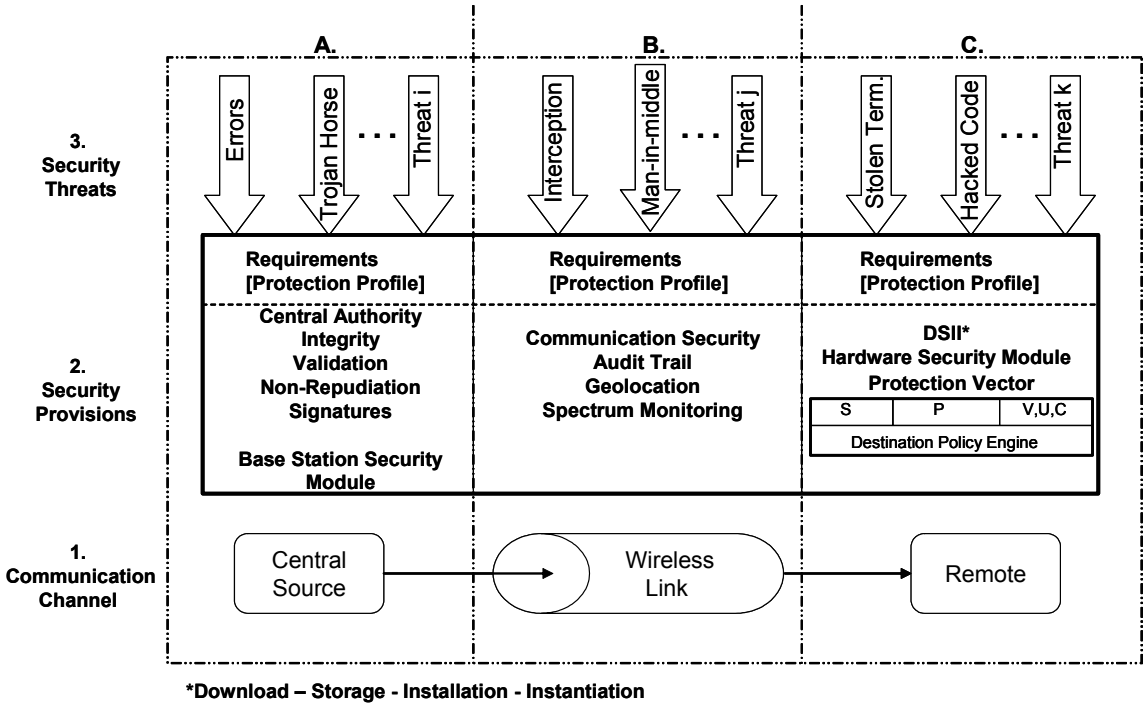

Figure 1. SDR Forum Security Reference Model

Level 2. Security Provisions.

This layer provides an appliqué of protection mechanisms installed in the system to thwart attempts to violate system security. It can include signatures, certificates, and encryption as well as physical security and secrecy to ensure that information entering the channel is valid and to protect it for transmission. After successful download of new software, the Protection Vector (PV) provides a disciplined approach to evaluating the credentials of the package, and making a policy decision to accept or reject the downloaded module after it has arrived at the destination terminal. The evaluation considers the source, reliability of the channel, and local operating considerations in the light of established security and operating policy considerations. This level provides consideration of mechanisms that can protect the bottom level from intrusion.

Protection profiles (PP) provide a structured method to describe security requirements. A protection profile is defined as being "an implementation-independent set of security requirements"[2].

The security provisions indicated in the model are intended to be suggestive rather than exhaustive. The intent is to indicate what provisions are appropriate to which of the columns.

Level 3. Threat.

Any occurrence that detracts from perfect operation of the system is a threat, and is considered in Level 3. Threats generate requirements for security protection. This level, is concerned with sources of possible intrusion, disruption, or interception.

The model provides a tool for consideration of the very complex space of wireless download security. In areas where the threat structure can be identified, systems may benefit from an approach that considers threats first, and then build the Level 2 constructs to mitigate those threats. Other systems will start with a proposed Level 1 structure, build the Level 2 provisions on top, and then test the Level 2 provisions against a threat taxonomy to explore for possible weaknesses.

3. MODEL ELEMENTS

In the following sections we will consider the nine areas of the model using the level number and column letter. We will describe them in some detail, and give examples of the kinds of considerations that are relevant. Our intent is not to provide security solutions, but a tool to facilitate exploration and consideration of this very complex space.

By the nature of the security problem space, preventative measures are a response to threats that have been observed in the past. It is impossible to predict and thwart all future threats, and there are always trade-off between the cost of protection, operating convenience, and the probability of penetration. For purposes of logical sequencing, in this section we will consider Level 3, Threats, before Level 2, Security Provisions.

Level 1.
In the commercial wireless world, Service Providers sell access to a wireless infrastructure. They may own that infrastructure or lease capacity on it. They specify what terminal models will operate in their service, and either sell them directly to customers or through a third party vendor.

Block 1A
The public service telephone network (PSTN) has many years of history providing connectivity between two people wishing to have a conversation. The past two decades have seen two dramatic developments extending PSTN functionality. One is introduction of the internet and widespread availability of facilities to transfer data. The ultimate extension of internet technology is its use to provide telephone service itself, voice over internet protocol (VoIP).

The other major development is use of radio links to remove dependence on a wired local loop. Wireless phones now carry a substantial amount of the world's phone calls.

Block 1A in the diagram includes all of the Wireline facilities as well as the wireless infrastructure, down to the base station antenna. Traffic carried in this network is at low risk because system nodes can be made physically secure and most of the transmission links are implemented with inherently secure fiber technology.

SDR technology introduces software to change the RF characteristics of the base stations that are included in this block. Under some circumstances software may transit the system in the form of data to be installed in remote terminals, a process called software download. It is used for correction of software errors in the initial code, insertion of new applications, addition of functionality, or conversion to a completely different service.

The download process must offer a high degree of assurance that it will not improperly impact the basic performance of the terminal. That assurance, in turn, requires that the source of the software can be trusted. In practice, only the original manufacturer of the terminal should be in a position to change operational software, although third parties will often provide code for applications. Manufacturers must adopt protective means to ensure that radio software can not be impacted by the download, installation, and execution of other types of software.

Block 1B
Block 1B is the wireless link between the infrastructure and the user's mobile terminal. It is inherently insecure, and in the transition from Wireline service to the first generation of cellular phones a great deal of trouble was encountered due to lack of security. Not only were conversations overheard and made public, but substantial revenue was lost from hackers stealing phone numbers over the air.

That experience led to architectural provisions for greatly increased security in the second generation of telephones. In addition provisions for digital protocols, directional sectors, antenna beam steering, and low power operation have improved security in the radio link.

Block 1C
Block 1C is the receiving terminal, usually a small handheld device in the PCS service. An important consideration is the need for a hardware module that cannot be altered after shipment to positively identify the terminal and to carry a secret key.

Handsets typically operate on more than one service or frequency band. At the present that is accomplished by having several different protocol modules in the device, and a switcher than can choose between them. In the future they will be SDRs, and able to be reprogrammed to accommodate new services or operation in a different geographical area.

Many phones now incorporate a user identity module (UIM) that is identified with a specific user. That way the users personal information can be moved to a different unit. In order to deter theft, the UIM can carry a password. In the future biometric devices may provide stronger validation of the user.

Level 3.
The threat space is one of the most complex aspects of wireless security because it is very large and sparsely populated. The threat space is large because an attacker can attack any one of millions of terminals anywhere in the world. Further the attack can take place at any point in the system, so there are billions of possible threat points or places that an attack could conceivable occur. It is sparsely populated because attacks are rarely seen in the normal course of events as the number of calls which comes under attack is a miniscule fraction of the calls completed without incident. The real threat to a particular system is a subset of the potential threat, and varies with the intent and skill of the individuals involved. There are people with a variety of different motivations in contact with the system on a daily basis. Some of them are curious, some careless, and some malicious.

The threats in the model are shown as arrows attacking individual parts of the system, and deterred by the security provisions protecting those components. Individual threats can strike at any part of the system at any time, and they can be new or ones previously seen. There is no way to develop an exhaustive list of threats, because ingenuous individuals will always find new ways to attempt to crack the system.

In addition to malicious threats, disruption of normal operation can be caused inadvertently by users of the wireless system, although good system design should mitigate most such potential problems. In particular, the

system should handle overload gracefully to handle situations such as many drivers in a traffic jam providing a heavy offered load in a small section of freeway.

Block 3A.
Threats in the network may come from personnel who have been given access to facilities as part of their work. Such an individual can be a threat if they violate their trust. System design must identify the source of all actions within the system, and segment allowable procedures by job function. Staff that are careless or under-trained are also be part of the threat space.

All of the security provisions applicable to any computer system on the internet also must be provided here. If an intruder can cause damage by accessing an internal IP address that is a point of vulnerability.

Block 3B
The RF link is a major source of threat. An individual can access a base station masquerading as a legitimate user. Alternately, a fake base station can attempt to attract terminals to log on and attempt to extract information from them.

A more sophisticated attack is the "man-in-the-middle", in which a perpetrator picks up a terminal looking like a legitimate base station. Then the fraudulent station logs on to a real base station, and relays the information. By monitoring the exchange between the terminal and base station the fraudulent station can acquire real log-on procedures.

Block 3C
Threats in this block arise from misuse of a handset. A thief can steal a unit, and attempt to use it to make calls charged to the real user's account. Alternately, an individual can buy a legitimate terminal, and try to revise its internal software to attack the system.

As we have indicated, and as experience with attacks on personal computers has demonstrated, an exhaustive determination of threats is not possible.

Level 2.
This level is an appliqué of security mechanisms to protect information traveling through the Level 1 mechanisms, and is not required for information transfer. The Communication Channel is perfectly capable of transferring information with no specific security provisions. In practice, however, security provisions are required, and a trade-off between cost, efficiency, and protection is needed.

Block 2A.
Block 2A is involved with assurance that the source of the information to be transferred is reliable, trusted, and authorized to undertake the transaction. This may involve adding information to establish that the source is bona fide, and may include ancillary information such as the results of software testing. Certificates may be used to ensure that there has been no perturbation in the information during transmission. Signatures may be used to authenticate that an individual or office is who they say they are. A Central Authority is needed to endorse both parties to the transactions.

Signatures are a mechanism where by some individual or organization indicates its approval of a transaction by signing it. Multiple signatures, representing the concurrence of involved organizations will often be needed. An example would be approval of a software correction patch by the equipment manufacturer, service provider, and network operator. Ultimately, however, trust comes into play, and appropriate steps taken to ensure that the needed level of trust is in place

In this block a security wrapper can be placed around specific information, such as software to be downloaded. The base station security module provides this security, and ensures that protected information put onto the wireless link is accurate. Section 4 describes such a module developed by HYPRES, Inc.

Block 2B.
Block 2B includes whatever encryption, transmission security, or low probability of detection facilities are used to protect the transmitted information. It is not possible to prevent interception on a wireless link, but interpretation can be made very difficult. Security is rarely absolute, but measured in terms of the effort required to access content.

An audit trail of the path transited by the download package en route to the destination can be used to establish that the package was not diverted and manipulated en route. Geolocation and spectrum monitoring can develop information to identify anomalies in terminal behavior from attempts to penetrate the system.

Block 2C.
Block 2C extracts the original information from what ever encryption or other protection mechanism has been used. Then it presented to the terminal user as voice or screen display. If it is downloaded software, it is put into a storage area. Then, after assuring that it has been received correctly and has the proper credentials, it can be introduced into the local software library and utilized.

Included in this block is a process primarily derived from Mobile VCE work described in [3]. That work is recast into the form of a Protection Vector (PV), a series of numeric values for various system security aspects. Those

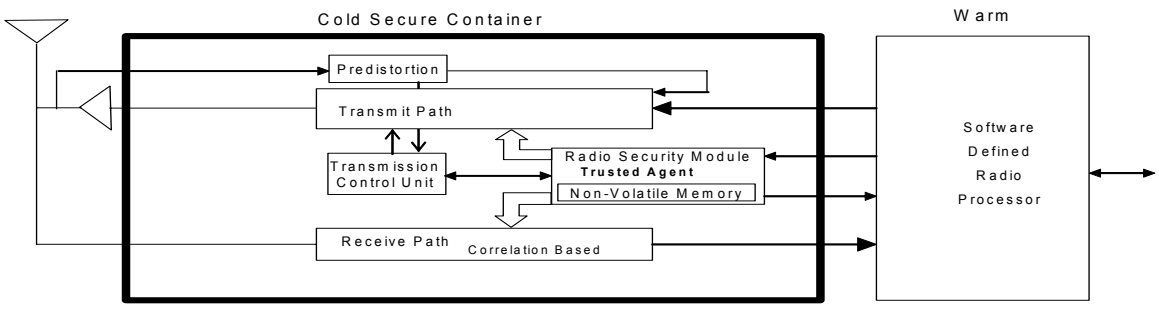

Figure 2. HYPRES Digital Security Module

vector elements are provided to a rule-based Destination Policy Engine that renders a verdict as to whether a given software downloaded should be accepted or not.

The Source (S) parameter is an assessment of the reliability of code that is a candidate for download, and is related to Block 2A. Levels of assurance are a statement by the code originator that it is acceptable, identification of the author(s), development by trusted authors, independent trusted third party test, and formal proof of correctness.

The Path Vulnerability (P) parameter is an evaluation of the intermediate path between the developer and the target terminal. Path protection mechanisms are digital signatures, public key infrastructure, audit trail and path histories, and trusted intermediate repositories.

Three parameters provide local context for policy evaluation and decision making. First is the Inherent Value (V) of the content, with a priority structure of worth such as mission critical, money, executable code, data, and audio or visual material. The second Block 2C parameter is Urgency (U), rated from high to low. Under some circumstances it is appropriate to accept information that is extremely urgent even though some risk is involved because its other parameters were lower than desired. The final parameter is Criticality (C). That is concerned with the impact on system operations if the received information is faulty. Low criticality would be assigned to a correction for a function such as a spelling error, minor feature addition, or improvement in the user interface.

When all of the PV elements are collected together, a Destination Policy Engine uses policy rules to determine whether the code should be installed or not. Much of the philosophy here is similar in concept to the Protection Profile (PP) described in [4].

The Protection Profile is a set of requirements cast in the form of prevention before system design or during system evaluation. It is also involves policy instantiation and execution during run time. The Protection Vector is a go-no go decision after a specific instance of information transfer.

Security Provisions will vary from system to system. The military goes to considerable effort and expense to provide secure tactical radios, but makes extensive use of commercial mobile radios and wireless PCS communications for administrative traffic. Wireless telephony discovered with systems such as AMPS that inadequate security leads to loss of revenue from stolen service and customer dissatisfaction from lack of privacy.

4. HYPRES DIGITAL SECURITY MODULE

HYPRES, Inc. has developed a very high performance RF capability called Digital RF using superconducting micro-electronics (SME). This facility is a candidate for the bse station security module in block 2A.

Digital RF

The following Digital RF capabilities are enabled by the HYPRES technology's ability to process the channel information stream in near-real-time.

Pre distorted transmission.
A signal sampled after the power amplifier is compared with the data input to the exciter. Errors inserted in processing are detected, and reinserted with their phase inverted to eliminate the error in the transmitted signal. This process is dynamically self-correcting.

Correlation-based reception.
The received signal is correlated to a known waveform. Distortions due to multi-path, path delay, fading, etc are detected, and resolved in the digital domain to recreate the original signal

Digital Security Module
An adjunct to the Digital RF front end is provision for a digital security module (DSM). Figure 2 shows the structure of the DSM. It is the HYPRES Digital RF capability with the addition of a security module and communication links to it. The security module with its Trusted Agent has an ability to control the transmitter and receiver of the radio.

Because the superconducting circuits must inherently be maintained in a separate environment to maintain cryogenic temperatures, they can easily be considered as a different architectural domain. That separation is used in the following ways.

Physical Security
The box containing the Digital RF circuits can easily be made as secure and tamperproof as needed with key access, strong enclosure, separate power supply, and RF shielding and filtering.

Trusted Agent (TA)
Software can operate inside the DSM module at warm, HTS, and LTS temperatures. That software can be designed to provide a number of functions as needed to ensure the security of the radio. With the agent in position, compliance with the rule set provided to it is guaranteed.

*Radio Contro*l
Because the DSM is co-resident with the transmission and reception logic in the Digital RF facility, the TA can be given control over those paths, providing the ability to modify channel data or stop it entirely. The default condition is to stop operation, and will be invoked according to the rule set or when power is removed.

10. REFERENCES

[1] DL-SIN, SDRF-02-W-0005-V0.30
[2] DL-SIN Appendix D
[3] E. Gallery,"A Policy-based Framework for the Authorisation of Software Downloads in a Mobile Environment, SDR '03, Paper SY2-002
[4] Common Criteria, http://www.commoncriteriaportal.org /public/files/ccintroduction.pdf.