A Primer Focussing on Handhelds, Wireless Communication and Security

Christopher Lambacher W.F.S. Poehlman

Computer Science Technical Report

April, 2003

Contents

1 Introduction

Digital communication in the past 30 years has grown significantly, both in importance and capability. The Internet has changed how people communicate and do business. It is almost impossible to operate a computer that does not have some connection to another computer.

Handheld devices also have had revolutionary effects on what can be done. Miniaturization has allowed devices with capabilities comparable to 386 computers to be held in the human hand. The major drawback is their isolation from other computers. Recent advances in wireless communications have had significant effects, both on mobile computers, and in how network infrastructure for stationary computers is perceived.

This paper discusses the state of the art in wireless communications standards and attempts at providing security for use with those standards. This is put into context for use with handheld devices by providing a description of current handheld device capabilities and limitations. An attempt is made to analyse the relative advantages and disadvantages of the standards and provide enough detail to make and educated decision on what standards are useful in any particular situation.

2 Wireless Communication Standards

2.1 802.11

2.1.1 Overview and History

The IEEE 802.11 working group oversees a family of standards for wireless LANs consisting of 802.11a, 802.11b, 802.11g and the original 802.11 standard. The protocols are designed to run in the unlicensed 2.4GHz ISM band, and the national information structure(U-NII) bands in the 5GHz-6GHz range. All use a common MAC layer that is similar to Ethernet. 802.11 only operates at 1 and 2 Mbps and runs in the 2.4GHz frequency range. 802.11b operates at up to 11Mbps. 802.11g extends the 802.11b standard to 54Mbps, though is not yet fully standardized. Even so, devices are available to operate with 802.11g which operate on preliminary versions of the standard. 802.11a is not compatible with the other 3 standards as it operates in the 5-6GHz frequency range. It provides transfer speeds from 6Mbps-54Mbps. 802.11b, also known as Wi-Fi is widely used in home and business networks.

After the original 802.11 standardization was completed, work began on the 802.11a standard. It was targeted at business uses where higher transfer rates and higher subscriber densities would be needed. Due to technical problems with the standard, the process was slowed and 802.11b was started and standardized with speeds up to 11Mbps. 802.11b quickly gained popularity with both home and business users. 802.11g was started as a higher speed extension to 802.11b, allowing speeds up to 54Mbps. 802.11a is now standardized and equipment being produced, but adoption is slow due to already widespread use of 802.11b and relative expense of 802.11a. 802.11g products are now available without standardization being complete. Users may choose to use 802.11g instead of 802.11a since 802.11g is backwards compatible with existing networks. New dual band devices are available that support all the standards. This might provide a reasonable upgrade path for companies to consider purchasing 802.11a enabled equipment.

802.11 has become quite popular since it provides a reasonable operating range and speeds that are useful for home and most office use. Industry support for 802.11 products is extensive. Just about every network card vendor provides 802.11 products. Intel's new Centrino mobile platform requires 802.11b wireless support in order for vendors to use the Centrino brand on their products[20]. (See section 5.3.1.) 802.11b is also relatively cheap, and can potentially greatly reduce infrastructure costs that would normally be associated with the establishment of a LAN, such as wiring.

2.1.2 Protocol Information

All of the 802.11 protocols use Carrier Sense Multiple Access with Collision avoidance(CSMA/CA)[14] to share the wireless medium. This simply means that before transmission, the wireless device monitors for activity before sending. If it detects traffic, it waits until no traffic is present. If a collision occurs, then both wireless devices wait a random amount of time before attempting transmission again. The random wait time avoids the problem of attempting to transmit again at the same time.

801.11, 802.11b, and 802.11g all use a similar, backwards compatible physical layer. Each incoming packet is packaged inside a physical layer transmission packet called a PPDU(PLCP Protocol Data Unit)[12]. PLCP protocol is the name given in the standard for the physical layer protocol. Each PPDU contains a fixed length header, which is always transmitted at the lowest bit rate of 1MBps and a variable length data section which contains the MAC information and data payload. When a packet is sent, the transmitter starts by sending alternating 1s and 0s to allow the receiver to synchronize to the incoming signal. The beginning of the frame data is then marked by the frame delimiter, which is the binary sequence: 1111001110100000. The rest of the header follows the frame delimiter. The header indicates the transmission speed of the data section, which it transmits as the rate divided by 100kbps. The size of the packet is given in microseconds that it will take to transmit the full frame, including the header. A 16 bit CRC is also sent in order to test that the header was received correctly.

Channel	Frequency
1	2.412GHz
2	2.417GHz
3	2.422GHz
4	2.427GHz
5	2.432GHz
6	2.437GHz
7	2.442GHz
8	2.447GHz
9	2.452GHz
10	2.457GHz
11	2.462GHz
12	2.467GHz
13	2.472GHz
14	2.484GHz

Table 1: 802.11b frequencies[12]

Once the header data is formatted correctly, the packet is sent to the transmitter. The transmitter uses DSSS(direct sequence spread spectrum)[12] to spread the signal over a 30MHz portion of the 2.4GHz frequency band. This makes the signal more tolerant to interference. 802.11 provides 14 transmission frequencies separated by 5MHz. (See table 1.) This means that consecutive channels cannot be used. A separation of 5 channels is needed to ensure that signals do not interfere with each other. Canada and the US are limited to using channels

1 to 11, which means that a maximum of 3 channels can be used, 1, 6, and 11. The method of spreading the signal varies based on the transmission speed. At low speeds of 1 and 2 Mbps the spreading method is constant, while higher speeds require optimization for each bit sequence.

At 1Mbps, the transmitter shifts the phase of the center transmit frequency to provide the distinction between a 1 and a 0. In order to achieve 2Mbps 4 phase shifts are used, corresponding to the possible values for each pair of bits in the packet. Similar methods are used for 5.5 and 11Mbps.

The resultant signal is amplified and sent to the antenna where it is sent through space to the receiver where it is demodulated, despread and decoded.

802.11a goes through a similar yet still fairly different process. The differences are necessary in order to provide the high speeds for which it was designed. 802.11a uses OFDM(Orthogonal Frequency Division Multiplexing)[11] instead of DSSS. As in 802.11b there is a header, which is always transmitted at the lowest speed, 6Mbps. The signal starts with a field to allow the receiver to synchronize with the incoming OFDM signal. The transmission rate for the data section follows. There is a unique binary code for each of the transmission rates which are 6, 9, 12, 18, 24, 36, 48 and 54Mbps. The length is provided as the number of bytes in the frame and a parity field provides a rudimentary check to see that the data were properly received. The data follow and may be padded with zeros in order to fit the number of bits required for an OFDM transmission symbol.

OFDM in 802.11a separates the signal into 52 separate subcarriers[23]. Four of the subcarriers transmit known sequences and are used to detect problems in the signal. This leaves 48 carriers for data, which are sent in parallel. The subcarrier spacing is 0.3125MHz providing 64 subcarrier slots in the 20MHz frequency space allotted for each channel.

The US allows for 12 channels spaced 20MHz apart. Each channel has one of 3 possible operating powers. Table 2 lists the channels, frequencies and allowed power outputs.

Channel	Frequency	Maximum Power
36	5.180 GHz	40mW
40	5.200 GHz	40mW
44	5.220 GHz	40mW
48	5.240 GHz	40mW
52	5.260 GHz	200mW
56	5.280 GHz	200mW
60	5.300 GHz	200mW
64	5.320 GHz	200mW
149	5.745 GHz	800mW
153	5.765 GHz	800mW
157	5.785 GHz	800mW
161	5.805 GHz	800mW

Table 2: 802.11a Frequencies and Power Ratings[11]

2.1.3 Problems and Limitations

Overall, the problems with 802.11 are minimal and not completely insurmountable. The two main problems are frequency space pollution and multipath interference.

Because 802.11, 802.11b and 802.11g operate in the 2.4GHz frequency space they are subject to interference from other consumer devices including portable phones, garage door openers and microwaves. Bluetooth also operates in the 2.4GHz range. See section 2.3 for more information about 802.11 and Bluetooth interoperability. 802.11a does not suffer from these problems, although any devices operating in the 5GHz range could interfere with its communications. Frequency space restrictions also mean that there is a limited number of usable channels. This could pose a problem if there are several closely situated wireless networks.

Multipath interference is caused by the signal being reflected off objects causing part of the signal to arrive at one time, while more of the signal arrives later. This can confuse the receiver, especially if the delay between the reflections is large[13]. This is less of a problem for 802.11a because it uses ODFM instead of DSSS. In industrial applications where there is a lot of metallic machinery and multipath interference is significant, 802.11a is a better choice. Multipath interference in homes and offices is generally insignificant.

As a final problem, 11Mbps is not sufficient for supporting very many users. Because of this, areas having a high density of users should use 802.11g or 802.11a.

2.2 Bluetooth

2.2.1 Overview and History

Bluetooth was started in 1994 by Ericsson and named after the Harald Blaatand "Bluetooth" II, king of Denmark 940–981AD[8]. A special interest group was formed in 1998 and which had been joined by 1900 companies by June 2002.

Bluetooth is designed to provide short-range wireless communications. It provides a reasonable replacement to IrDA, while extending greatly what can be done by overcoming the IrDA 1m distance limitation and line of sight requirements. Bluetooth allows communications at up to 10m and permits connections with multiple devices without orienting the transmitter in any particular direction.

Bluetooth provides connection rates of up to 780kbps which can be used for 721kbps unidirectionally or 432.6kbps bidirectionally. This does not provide particularly fast speeds, but does allow for faster than serial connections, which makes it possible for Bluetooth to replace serial connections in many instances.

Bluetooth is most useful for small devices such as phones and PDAs. It can allow communication between the devices and other devices such as computers for synchronization. The Bluetooth standard already allows interaction between cell phone and PDAs such as the provision of Internet access to the PDA, or the directive, by PDA, for the phone to dial a given phone number. PDA synchronization and printing are also key possibilities.

Bluetooth will be the basis of the IEEE 802.16 working group's wireless personal area network (WPAN) standard. This may result in a convergence between the two groups.

2.2.2 Protocol Information

Bluetooth devices organize themselves in groups of 2 to 8 called piconets. One device is the master and the other devices are slaves. Devices can be slaves for more than one piconet, but masters of only one. One extension to the standard that is being developed is the ability for devices connected to more than one piconet to act as a bridge between the two piconets. Two or more bridged piconets is called a scatternet(see Figure 1).

Whether or not bridging is available, it is possible to have several piconets operating in the same area. Also, like 802.11b, transmissions are made in the unlicensed

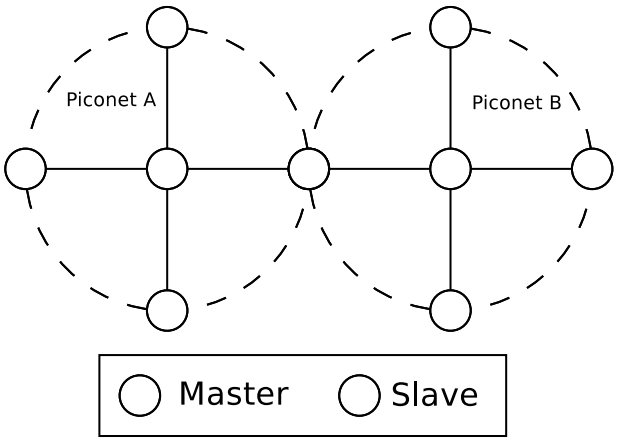

Figure 1: Bluetooth scatternet diagram(based on [8]).

ISM 2.4GHz frequency band. This means that there is interference from many other devices. To help alleviate problems with this, Bluetooth devices in a piconet are synchronized to a frequency hopping pattern which is unique to each piconet. Using this pattern, devices change frequencies 1600 times per second. Between hops is a time slot for the transmission of data. A particular packet may require a transmission duration that spans several hops. In this case the frequency is held for the duration of the transfer.

Bluetooth defines channels for transmission of both voice and data. The standard allows for one asynchronous data link and up to 3 synchronous voice links. Voice transmission uses reserved time slots to guarantee reasonable transmission, but never retransmits dropped packets.

The transmission control is done by the link manager which handles connection setup and security(see section §3.6).

Above the link manager, software protocols are defined. L2CAP is the lowest software layer of the protocol and provides access to the link layer. Protocols not defined by Bluetooth, including TCP/IP and vCard/vCalendar operate on top of this. Bluetooth defined protocols that run on top of L2CAP include RFCOMM, a simple serial data transfer link, Telephony Control Protocol Specification(TCS) which provides voice and data call control.

Also included in the Bluetooth standard is the service

discovery protocol(SDP). As the name implies it can be used to find out what services are available on remote devices. In an environment where services can come in and out of radio range randomly, it is important to have a method of discovering such information.

2.2.3 Limitations and Problems

As in 802.11, frequency space pollution is a potential problem. Although Bluetooth uses frequency hopping to avoid problems with this, eventually to many devices using 2.4GHz frequency range could cause problems that frequency hopping will not be able to avoid.

The 780kbps transfer rate limit may prove to be too slow in the near future. It is sufficient for current applications and applications requiring more speed can use 802.11. Similarly range could be perceived as an issue, again 802.11 might be a better choice if range becomes a problem with some applications of Bluetooth.

2.3 Bluetooth and 802.11 interoperability

Because Bluetooth and 802.11b and 802.11g use the same 2.4GHz frequency space for communication, they interfere with each other. Several papers have been written on the performance of 802.11b WLANs and Bluetooth devices in the presence of each other[3, 5, 9] and techniques for making them peacefully coexist[4, 10]. Current research indicates that this problem is not insurmountable. The upcoming 802.16 WPAN standard based on Bluetooth is intended to address current issues with Bluetooth, including the interoperability problem.

2.4 3G Wireless Telephone

Wireless telephone systems have been in place since 1946[1]. It was not until 1984 when Advanced Mobil Phone System(AMPS), an implementation of the cellular system proposed by AT&T in 1968, began to be deployed that mobile phone systems really took off. In 1991, US digital cellular was released and in 1994 IS-95 code division multiple access(CDMA) was introduced. Recently third generation(3G) wireless systems based on CDMA technologies, have been released.

3G is the latest version of wireless cellular telephone communications. 3G is intended to provide better data services and enhance existing voice communications. 3G provides data rates of up to 2Mbps, a reasonably useful rate for users requiring Internet connectivity in areas covered by the 3G network. This may however be too little too late as far as data laptop connectivity is concerned as that speed is easily outstripped by 802.11. With the number of 802.11 hotspots growing rapidly, the need for 3G communications in urban areas is diminishing. 3G does however have the advantage of consistency of service provider and provider dedication to expansion of coverage area.

The use of 3G in conjunction with Bluetooth and handheld devices does hold promise. Recent handheld devices such as Palm use Bluetooth to access the Internet through the data connection provided by the phone. This is used for low bandwidth applications such as email and instant messaging. Such applications and digital applications on the phone itself provide interesting services for users who already carry a mobile phone.

3G's main focus is still voice traffic, and it and its predecessors are essentially the only options for mobile telephone systems.

2.5 Other Technologies

Other technologies not mentioned in this paper, such as HomeRF, an 802.11 competitor, are either not in widespread use or severely limited in their capabilities so as to make them not generally useful.

3 Wireless Security

Wireless security primarily consists of two parts: network authentication and data integrity assurance. Network authentication ensures that only authorized devices are using the network and may also ensure that only an authorized subscriber is using the network through the authorized device. Once subscribers are permitted access to the network they would like verification that their communications can only be viewed by the intended recipient. This is termed data integrity assurance. The broadcast nature of wireless communications means that anyone can capture the data traffic. Because of this, the content of the traffic should be obscured. This is accomplished through

various forms of encryption. Most security protocols provide mechanisms for both of these aspects of security.

There are several places where security can be inserted into the system. Each of them has implications for usability and what actually gets secured. Points for security are referred to by the level in the OSI networking model in which they are introduced. Current solutions occur at the data link layer, the network layer, the application layer, or above the application layer.

The data link layer, level 2, is one level up from the physical medium, level 1. It is concerned with such aspects as framing, data flow control and error control. This is the level of Ethernet and the MAC. The network level, level 3, provides the ability to route datagrams between network nodes. IP operates at this level. The application level is level 7. This is where application protocols such as HTTP, FTP, and SMTP operate. Above the application protocol level applications such as mail programs, and web browsers exist. The distinction between security at the application level and above the application level may be difficult to see. Application level security is built into the protocol as in the case of SSL. Security above the application level may pertain to the device itself such as password access to use the device at all, or may be an application data mechanism such as encrypted email, where it is some portion of the application's data that is encrypted rather than the protocol itself.

For each of these layers, the security provides different levels of coverage. Security at any level only provides protection against attacks at that level and below. For instance, if data link layer security is used, then the security only protects that network segment. As soon as an IP packet is routed onto another network segment, the security no longer applies to the communication. Application protocol level support however ensures that all communication between a client and server computer is safe. Layers above application level support, such as in an encrypted email, ensure that only the intended recipient can view the content.

3.1 Cryptography

Cryptography is commonly used in network authentication and data integrity assurance. The reason for this is that it allows the communication to be obscured while allowing the intended recipient to view the data stream. The basic algorithm takes the message to be encrypted and a key to transform the message to a different space. The idea being that it is easy to calculate the transformation when the key is known, but difficult in any reasonable amount of time when the key is unknown[15]. There are three forms of cryptography in use: symmetric key, public key and hash functions.

Symmetric key systems use the same key at either end to do both the transformations, encryption and decryption. Public key systems use two keys: one key that is known publicly and one key that is private. When a message is encrypted with the public key, it can only be decrypted with the private key. When the message is encrypted with the private key it can only be decrypted with the public key. Hash functions are one way mappings. The idea is that for every input, a hash function creates a unique output that cannot be decrypted in any way.

Public key systems provide a very flexible framework for providing authentication and encryption. With public key systems and an appropriate trust model which provides a method of obtaining public keys, anyone can send a message intended for a particular recipient and be assured that only the intended recipient will be able to view the message. On the other hand, a sender can encrypt a message with their private key, and anyone who decrypts it can be assured that it truly came from the sender indicated. It is common to establish communications using public key cryptography and use a negotiated symmetric key, which is significantly less computationally intensive, for the rest of the session.

Hash functions provide a reasonable way to ensure that a message was not altered en route. While this does not prevent someone from eavesdropping, it does guarantee the validity of the information received.

3.2 Network Authentication Methods

Common methods for network authentication include password checking, public key systems and biometric detection. In each case, a secret that is supposed to only be known to the authenticator and the authenticatee is checked to determine identity. The difference is how the exchange is carried out. Provided that no one can acquire the secret, any solution based on this method should be effective.

Challenge response systems built on public key cryptography are very popular because of the difficulty involved in spoofing. For instance, if a simple password authentication system is used then some measure must be taken to be sure that the security cannot be circumvented and the password obtained at a level between password collection and comparison. With private key systems, the authenticating agent asks the client to encrypt some random piece of data using the private key. The authenticator then attempts to decrypt the returned encrypted data and compares the result with the initial challenge data. As long as the private key is not compromised, this will provide very good security. In this example, note that the randomness of the challenge is important since using the same challenge word repeatedly allows an attacker to know the response to the challenge without computing the cryptographic transformation. It may also allow the attacker to guess the private key.

Some systems require two challenge items, such as a private key and a password, in order to enhance security. In this case there is some protection against a device being used by the wrong person. For instance, the machine being used may be authenticated with public key cryptography, while a particular user authenticated by password, the password being protected by cryptographic methods.

Current wireless communication systems provide some form of network authentication. 2G and 3G wireless telephone systems, 802.11b and Bluetooth all provide shared key systems to authenticate users. These provide varying levels of success.

The other common security systems used at other network levels such as SSL and IPSec also provide key based authentication, but may also allow other methods such as password exchange instead of or in addition to the cryptographic methods.

Note that systems that do not use cryptographic methods to do the authentication may still use them to secure the transmission during the authentication process.

3.3 Data Integrity Methods

The only way to properly ensure data integrity is to use encryption. The most common method for providing this is a public key session setup, followed by symmetric key encryption. Public key cryptography allows the assurance that only the intended recipient can view communications. The computations for public key encryptions are excessive for use for the full session. Once the session is established, the parties can agree on a symmetric key which is generated randomly after session setup. If a session lasts a significant amount of time, the symmetric key should be changed. When this occurs, the two parties exchange the new key under the encryption of the last key. At all times, the weaker symmetric key is protected by short usage time and exchange while encrypted.

Many flaws in encryption can be tied to key leaks that occur because the key is somehow exposed either during session setup or through extended use with the same key.

Both 802.11b and Bluetooth provide built in data link layer encryption. 802.11b's encryption is called WEP(Wired Equivalent Protection), see section §3.5. Virtual private networking(VPN) solutions use network level encryption to extend a private network out to other isolated computers or join two networks together through the Internet. While there are many competing standards for doing this, IPSec is an attempt at providing common ground for vendor interoperation. Due to problems in WEP, many 802.11b networks use IPSec to protect communication. Most application level encryption uses some version of SSL.

3.4 3G Wireless Telephone Systems

Wireless telephone systems are unique in the systems evaluated in this document because they are fully administered by the network owners. This means that the network owners provide the appliances and the network and say who can and cannot use the network. The security of the system is built around this premise. Blanchard[2] provides a description of 3G security and its differences from second generation digital wireless telephone system security.

The security in 3G is centered around the information provided in the Subscriber Identity Module(SIM), which is a removable module that is inserted into the device and administered by the home environment operator. Because the SIM is removable, it necessarily provides security independent of the mobile device in which it is inserted.

The SIM contains a user identity and an authentication key. The key is used in a challenge response authentication mechanism which means that the key is not directly required by the serving network and is never transmitted

over the air or exposed across the interface between the SIM and the mobile device. The challenge response occurs between the SIM module and the home environment operator's authentication center. This allows individual home environments to claim responsibility for authenticating their own devices when the client is connected to another network.

An unfortunate drawback of the old 2G system is that the challenge response authentication is one directional. The client is given no assurance that they have connected to a valid serving network. This was addressed as a limitation in the 3G standard by adding the 3GPP authentication and key agreement. This essentially allows the client to also do a challenge response authentication of the serving network, which is also handled by the home environment and the SIM.

Data integrity in 3G is used not only to protect the confidentiality of the user, but also ensures that the validity of the authentication made at the start of the session is maintained. It is not, however a mandatory feature due to restrictions on encryption in some countries. In these cases connection integrity protection is added to the signaling messages instead. This however does not keep someone from eavesdropping on conversations when encryption is not used. When encryption is used to protect data integrity, a symmetric synchronous stream cipher based on the Kasumi algorithm is used.

3.5 802.11b WEP

WEP(Wired Equivalent Protection) is the 802.11b attempt to provide security that is at least as good as having a wired connection. The goal is simply to keep eavesdroppers from being able to read the contents of packet. These are not very lofty security goals and would have been sufficient to make security problems no worse than those on a wired connection. Unfortunately several oversights were made about how 802.11b was to be used. WEP uses the RC4 stream cipher[6]. While well know and commonly used, the nature of wireless communications makes its use in 802.11b questionable. Because the key never changes, it is possible, after the collection of 5-10 million encrypted packets to compute the key used for the encryption. 5 million packets seems like a lot, but in a production corporate environment where there is much traffic 5 million packets can be acquired relatively quickly. Several freely available programs, including AirSnort[7], can do the work of collecting packets and computing the key for the user. Fluhrer, Mantin and Shamir[6] discuss the issue fully.

Because of the security problems in 802.11b, network administrators are using encryption at other network layers to protect their networks and users. Many 802.11b equipment vendors are offering proprietary alternatives to WEP in an attempt to help fill the gap in security. Future versions of 802.11b are expected to address this issue.

3.6 Bluetooth

Bluetooth provides device authentication and data integrity assurance through its security architecture. An overview of the architecture is provided by Träskbäck[22]. The security model provides authorization for services through its authentication method. Authentication is provided in the form of a shared key challenge-response mechanism. The establishment of a key is done in a process called pairing. Pairing is done by entering a common PIN for both devices. The key exchange can also be done wirelessly if desired, preferably in a secure area where there is some assurance that no one can eavesdrop.

Once a device is authenticated, an encrypted link can be established if desired. The encryption is provided by stream cipher E0 and only affects the packet payload.

Träskbäck[22] notes that the current security has some limitations, the biggest being that only devices are authenticated, not users. This means that potentially a device can be stolen and used with another already paired device. Nguyen et. all[16] suggest the use of an extension to the current system that also authenticates the remote user with a shared password that is exchanged after the device has been authenticated and encrypted communications established.

3.7 IPSec and VPN

IPSec is the Internet Engineering Task Force's network layer security protocol. IPSec essentially provides VPN(Virtual Private Networking) and encryption so that secure networks can be connected together through insecure networks such as the Internet.

One extension of this is to have individual computers connect into a central corporate network through the Internet. Many network administrators are using IPSec as a method of adding security to 802.11b networks.

3.8 SSL and TLS

SSL(Secure Sockets Layer), which in later versions became TLS(Transport Layer Security) was developed as a method of encrypting HTTP(Hyper Text Transport Protocol) traffic for web applications such as banking. Many other application protocols now use SSL and TLS to provide security, especially for protection of password exchange in protocols such as POP(Post Office Protocol) and IMAP(Internet Message Access Protocol).

4 Handheld Complications

4.1 Characteristics of Handheld Devices

The power of handheld devices is in their portability. Common handheld devices include Palm computers and smaller tablet PCs. Their intended portability makes them prime candidates for wireless communication technology. Unfortunately, they also have strict constraints on size, weight and heat dissipation. Also, because they do not have a wired power source they must be battery operated, while at the same time the battery size has to be limited to fulfill size and weight requirements making already limited power supplies even more limited.

Size, power, and heat dissipation, constraints limit the amount of processing power, and memory available for use. The more processing power, the more power required for the processor and the more heat that will have to be dissipated. Memory takes up a large amount of space, and many economical forms of memory use a large amount of power to store data, and even more to retrieve them.

Finally, because of their small size, handheld devices are easily compromised through theft or loss.

4.2 Implications on Wireless and Security

Easy compromise means that steps must be taken to ensure that if the device is used by someone other than the intended user, access to the device is limited or eliminated all together. Palm pilot like devices, whose focus is on ease of use, have for a long time omitted features related to this issue. Third party programs were necessary to provide even the smallest of automated security measures. While Palm devices did provide the ability to hide or mask data stored on the device, the security setting needed to be applied manually after viewing of secure data. This left the device open to viewing if the user forgot to resecure the device. As of version 4.0 of the Palm OS[17], Palm instituted the automatic remasking or hiding of such data at power off. This still leaves most of the security concerns for the device open to the individual applications running rather than any kind of all encompassing security policy. Third party applications such as TealLock[21] provide system wide security by requiring password authentication for access to the device. While this can turn off serial, Bluetooth and infrared communication, it cannot stop someone with physical access from attempting to attach a device to read data directly from the memory. Also, applications that require administration by someone other than the end user will need to provide their own security.

The power usage of wireless communications is directly proportional to the transmission range. This means that a tradeoff needs to be made between battery life and the range of communications. 802.11 implementations are inherently going to use up the battery faster than Bluetooth which has a comparably much smaller range.

Encryption is very computationally intensive. Many algorithms used on desktops and servers for providing encryption are noticeably slow on handhelds. Ravi et all[18] state that it takes “a PalmIIIx 3.4 minutes to perform 512-bit RSA key generation, 7 seconds perform digital signature generation and can perform (single) DES encryption at only 13kbps, even if the CPU is completely dedicated to security processing.” This poses two problems: one, it is not likely that the CPU will be able to be completely dedicated to security, and two, the power requirements of both the wireless transmission and the 100% CPU use would drain the battery far to fast to be useful for any length of time. Ravi et al[18] call this the “Wireless Security Processing Gap”.

5 Possible Solutions to Handheld Problems

5.1 Algorithm Optimization

There are three ways to look at optimizing cryptography algorithms for handheld devices. One, look at what features can be removed altogether from protocols. Two, find ways to reduce the number of instructions performed while providing the same features. Three, define new cryptography systems that are less computationally intensive yet provide similar levels of security.

Ravi et al[18] point out, that many of the wired security protocols that would be tempting to use on handheld devices are far too complicated for the needs of handheld devices. It is possible however to select a useful subset of protocol features to reduce the amount of processing that must be performed by the device.

Selecting optimal software implementations of the algorithms used can also significantly reduce the amount of processing required. On processors with such limited processing capabilities, this can make a significant difference.

Many cryptography protocols "use RSA-based public key cryptography for authentication." The RSA algorithm is based on the NP-hard problem, integer factorization. This requires the key(the modulus) to be a large integer in the range of 1024-2028 bits to provide proper security. Because of this long key, encryption is computationally intensive. Newer cryptography systems provide similarly high levels of security, with lower computing and memory requirements. Ravi et al[18] recommend the use of Elliptic Curve Cryptosystems(ECC).

5.2 Data Preprocessing

Limited size has implications on handheld devices that go beyond cryptography calculations. It also means that in some cases data being sent to the device may exceed its needs. For instance, web pages formated for desktop clients are not suitable for handhelds as too much information is present, graphics are to big and may contain information that, due to display limitations, is unnecessary. Preprocessing of this data can help to limit the amount of computation the handheld device must perform to make the information usable and will reduce the amount of data that needs to be transmitted. Reducing data transmission can help alleviate the so called wireless security processing gap.

5.3 New Chipsets

By working on new chip designs, there are two ways to help the problem. The first method is to attempt to significantly reduce the power requirements of existing more powerful CPUs. The second is to add extra processing ability to existing low power chip designs.

5.3.1 Intel Centrino Mobile Platform

Intel recently released its latest effort at the reducing the power requirements of an existing more powerful CPU with its new Pentium-M processor. The Pentium-M is part of what it is marketing as the Centrino Mobile Processing Platform.

Anand Tech[20] takes a close look at the technology found in Intel's newest mobile processing platform, dubbed Centrino. Released in the first quarter of 2003, Centrino is comprised of three components required for Centrino branding by vendors. The first component is a new lower power processor, the Pentium-M. The second component is the new low power DDR memory chipset, the 855PM. The third component is Intel's PRO/Wireless 2100 802.11b card.

Intel started with their Pentium III processor when developing the Pentium-M processor, however the final result is far from that beginning. The methods used in developing this mobile processor are significantly different from the methods used for previous chips. In the past, existing desktop processors were taken and scaled down to fit into the requirements for mobile computing. Newer CPUs like the Pentium 4 consume too much power to be successfully modified in this fashion.

The Pentium-M design team initially looked at the Pentium 4 as a starting point, however they quickly decided that a 20 stage pipeline was far too long for the power levels they wanted to reach. Instead they started with a Pentium III and put it through heavy redesign. The first thing to notice is that the Pentium-M has a longer pipeline than the Pentium III to allow for higher clock speeds. Normally this would lead to a lower IPC(Instructions Per Cycle), however Intel managed to increase the IPC on the

Pentium-M over the Pentium III. This means the Pentium-M can perform more instructions at lower clock speeds than the Pentium III or the Pentium 4. This directly translates into power savings as the CPU needs to be less active for similar processing.

One big problem with a longer pipeline is that if branch prediction is incorrect, then there is a greater performance penalty due to more instructions being lost on a misprediction. Pentium 4 processors help alleviate this problem with a trace cache, which allows instructions to start later in the pipeline in the event of a misprediction. A trace cache however contains many gates and consumes much power. The Pentium-M design team instead decided to improve branch prediction and not use a trace cache. They designed a branch predictor that resulted in about 20% fewer mispredictions than that of a Pentium III[20, p. 3]. This meant less power use than a trace cache, without the performance cost of not having anything to compensate for the cost of mispredictions due to the longer pipeline.

The Pentium-M has a 1MB L2 cache. The cache is highly optimized for power savings. Not only is the gate layout optimized for low power consumption, but the 8 way set associative cache conserves more power by subdividing each of the 8 ways into quadrants. Only the quadrant that is necessary is activated at any time. This lowers the power consumption. The cache layout adds a latency hit, which is made up for by the gains of having 1MB of L2 cache.

The Pentium-M will also reduce clock speed to preserve battery. As can be seen from Table 3 the the drop is significant for the higher speed processors. This idea is similar to the SpeedStep technology used in Pentium-4M processors. The major difference here is that the Pentium-M processor starts with portions of the CPU turned off, and only when they are needed are they turned on. This will result in some performance latency, however it will also have a significant effect on the power use.

The second part of the Centrino Mobile Platform is the new DDR memory and AGP 4x chipset, the 855PM. Intel applied similar techniques to the design of this chipset, in the end making it consume only half the power of Intel's equivalent desktop DDR and AGP chipset, the 845[20, p. 9]. Intel is also offering the 855GM which adds an integrated power optimized version of the 845G graphics core to the 855PM.

The final part, the Intel PRO/Wireless 2100 802.11b card, does not add any power saving features but shows that 802.11b is growing to be a strong force in mobile communications.

The Pentium-M will help blur the distinction between small tablet PCs and Palm type devices with its new power and size saving features.

5.3.2 MOSES

In *System design methodologies for a wireless security processing platform*[19] Ravi et all describe their efforts with the MOSES project which attempts to bridge the "wireless security processing gap"[18] by taking an existing extremely low power processor and adding extra instructions geared at increasing cryptography capabilities.

Their focus was not on the creation of hardware implementations of cryptography, but rather identifying common operations that are used in cryptography that are computationally intensive and can be implemented as extra instructions. Ravi et all[19] decided to take the Xtensa T1040 processor from Tensilica, Inc which provides a 32-bit RISC like CPU core, but allows a great amount of customization, including the addition custom instructions. This means that the developers can decide on the extra operations that they need the processor to perform without the need to design a whole CPU.

The method used was to first collect cryptographic algorithms, then analyse them to determine the most efficient alternatives of the various implementations of the same algorithm. They were then reimplemented to take advantage of library functions for complex operations. Then an analysis was done of the algorithms in use to determine the most useful extra processor instructions based what portions of the library were used most. The library functions that were added as instructions were then converted to use the new instructions.

Using comparing the hardware accelerated implementations to optimized software only implementations Ravi et all[19, p. 200] found that the MOSES platform provided a speed up of 2.18X speedup for small transactions where public key algorithm computations dominate. In large transactions where private-key algorithm computations start to dominate the overall computation, a 3.05X speedup was observed.

This is a significant increase in performance and indicates that new hardware and software combinations can

	Frequency		Voltage		
Processor	Performance	Battery Mode	Performance	Battery Mode	Power Use
Pentium-M 1.60	1.60GHz	600MHz	1.48V	0.96V	24.5W
Pentium-M 1.50	1.50GHz	600MHz	1.48V	0.96V	24.5W
Pentium-M 1.40	1.40GHz	600MHz	1.48V	0.96V	24.5W
Pentium-M 1.30	1.30GHz	600MHz	1.39V	0.96V	22.0W
Low Voltage Pentium-M 1.10	1.10GHz	600MHz	1.18V	0.96V	12.0W
Ultra Low Voltage Pentium-M 900	900MHz	600MHz	1.00V	0.96V	7.0W

Table 3: Pentium-M Operating Frequencies and Voltages[20, p. 8]

help overcome processor performance problems that hinder the use of cryptography in handheld devices.

6 Conclusion

Wireless and security are complex topics. At the present time, there is no definitive combination to provide maximum flexibility and security. In complex cases hybrid systems may be necessary, though this is complicated by interference generated in using them at the same time. As the protocols evolve, better coexistence should alleviate this problem.

The number of readily available programs that exploit security weaknesses show that the default security provided by any of the systems discussed is not sufficient for sensitive data.

The tight power and processing constraints of handheld devices further complicate matters, limiting how long connections can be maintained and how well they can be secured using encryption. This is unfortunate because handheld and other portable devices stand to gain the most from the availability of wireless communication systems.

It is clear that there are many hurdles to overcome in making wireless communications and handheld devices secure, but none of the current problems look like they will be issues for more than a couple of years. Ongoing research is pushing the envelope for processing requirements of encryption, increasing processing power and reducing power consumption. In all the future of handheld wireless communications and its applications looks very interesting.

References

[1] C.T. Abdallah and R. Jordan. Wireless communications and networking: an overview. *IEEE Antenna's and Propagation Magazine*, 44(1):185–193, February 2002.

[2] C. W. Blanchard. Wireless security. *BT Technology Journal*, 19(3):67–75, 2001.

[3] C.F. Chiasserini and R.R. Rao. Performance of IEEE 802.11 WLANs in a bluetooth environment. In *Wireless Communications and Networking Conference*, volume 1, pages 94–99. IEEE, 2000.

[4] C.F. Chiasserini and R.R. Rao. Coexistence mechanisms for interference mitigation between IEEE 802.11 WLANs and bluetooth. In *Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings.*, volume 2, pages 590–598. IEEE, 2002.

[5] M. Fainberg and D. Goodman. Analysis of the interference between IEEE 802.11b and bluetooth systems. In *Vehicular Technology Conference, 2001. VTC 2001 Fall. IEEE VTS 54th*, volume 2, pages 967–971, 2001.

[6] Scott Fluhrer, Itsik Mantin, and Adi Shamir. Weaknesses in the key scheduling algorithm of RC4. In *8th Annual Workshop on Selected Areas in Cryptography*, volume 2259 of *Lecture Notes in Computer Science*, pages 1–24, Toronto, Canado, August 2001. Springer-Verlag, Berlin Germany.

[7] The Shmoo Group. Airsnort homepage. `http://airsnort.shmoo.com/`, 2003.

[8] N. Gunasekaran, S.R. Redd, and K.V.S.S.S.S. Sairam. Bluetooth in wireless communication. *IEEE Communications Magazine*, 40(6):90–96, June 2002.

[9] I. Howitt. Bluetooth performance in the presence of 802.11b WLAN. *Vehicular Technology, IEEE Transactions on*, pages 1640–1651, 2002.

[10] Kyunghun Jang, Yongsuk Kim, and Bin Zhen. The analysis of coexistence mechanisms of bluetooth. In *Vehicular Technology Conference, 2002. VTC Spring 2002. IEEE 55th*, volume 1, pages 419–423, 2002.

[11] Jim Geier. 802.11a physical layer revealed. *802.11 Planet*, 14 March 2003. http://www.802.11-planet.com/tutorials/article.php/2109881.

[12] Jim Geier. 802.11b physical layer revealed. *802.11 Planet*, 10 March 2003. http://www.802.11-planet.com/tutorials/article.php/2107261.

[13] Jim Geier. Multipath a potential wlan problem. *802.11 Planet*, 10 March 2003. http://www.802.11-planet.com/tutorials/article.php/1121691.

[14] Matt Kirk. 802.11 term definition. http://searchnetworking.techtarget.com/sDefinition/0,,sid7_gci341007,00.html, 2003.

[15] Katherine Ladniak. Engineering and society 4x03 inquiry: Cryptography in telecommunications. *http://www.iprimus.ca/~hither/inquiry.pdf*, 2002.

[16] L. Nguyen, R. Safavi-Naini, W. Susilo, and T. Wysocki. Secure authorization, access control and data integrity in bluetooth. In *Networks, 2002. ICON 2002. 10th IEEE International Conference on*, pages 428–433, 2002.

[17] Palm, Inc. New features - Palm OS upgrade with mobile connectivity. http://www.palm.com/software/palmosupgrade_4-1/details.html, 2003.

[18] Nachiketh Potlapally, Anand Raghunathan, and Srivaths Ravi. Securing wireless data: system architecture challenges. In *Proceedings of the 15th international symposium on System Synthesis*, pages 195–200. ACM Press, 2002.

[19] Nachiketh Potlapally, Anand Raghunathan, Srivaths Ravi, and Murugan Sankaradass. System design methodologies for a wireless security processing platform. In *Proceedings of the 39th conference on Design automation*, pages 777–782. ACM Press, 2002.

[20] Anand Lal Shimpi. Intel's Centrino CPU (Pentium-M): Revolutionizing the mobile world. *Anand Tech*, 12 March 2003. http://www.anandtech.com/mobile/showdoc.html?i=1800.

[21] TealPoint Software. Teallock handheld security. http://www.tealpoint.com/softlock.htm, 2003.

[22] Marjaana Träskbäck. Security of bluetooth: An overview of bluetooth security. http://www.cs.hut.fi/Opinnot/Tik-86.174/Bluetooth_Security.pdf, 2000.

[23] VOCAL Technologies, Ltd. 802.11a modulation. http://www.vocal.com/data_sheets/80211a_mod.html, 2002.