RFID Systems and Security and Privacy Implications

Sanjay E. Sarma, Stephen A. Weis, and Daniel W. Engels

Auto-ID Center
Massachusetts Institute of Technology
Cambridge, MA 02139
www.autoidcenter.org

Abstract. The Auto-ID Center is developing low-cost radio frequency identification (RFID) based systems with the initial application as next generation bar-codes. We describe RFID technology, summarize our approach and our research, and most importantly, describe the research opportunities in RFID for experts in cryptography and information security. The common theme in low-cost RFID systems is that computation resources are very limited, and all aspects of the RFID system are connected to each other. Understanding these connections and the resulting design trade-offs is an important prerequisite to effectively answering the challenges of security and privacy in low-cost RFID systems.

1 Introduction

Automatic Identification (Auto-ID) systems have become commonplace in access control and security applications, in industries requiring the tracking of products through the supply chain or manufacturing process, and in industries requiring the identification of products at the point of sale or point of service. Perhaps the most widely recognized Auto-ID system is the bar code system developed during the early 1970's. More recently, Radio-Frequency Identification (RFID) systems have begun to find greater use in automatic identification applications. RFID systems consist of Radio Frequency (RF) tags, or transponders, and RF tag readers, or transceivers. The transponders themselves typically consist of integrated circuits connected to an antenna [8]. The use of silicon-based microchips enables a wide range of functionality to be integrated into the transponder. Typical functionality ranges from large read/write memories to integrated temperature sensors to encryp! tion and access control functionality. The transceivers query the transponders for information stored on them. This information can range from static identification numbers to user written data to sensory data.

The potential applications for RFID systems are numerous. Consider, for example, supply chain management applications and the use of EAN-UCC bar codes. Today, over 5 billion bar codes are scanned daily world-wide [6]. Yet, most bar codes are scanned only once during the lifetime of the item, namely at the check out. RFID systems, if strategically deployed, are a single platform on which a number of supply chain management applications can be simultaneously

B.S. Kaliski Jr. et al. (Eds.): CHES 2002, LNCS 2523, pp. 454–469, 2003.
© Springer-Verlag Berlin Heidelberg 2003

implemented, benefiting all parties involved in a commercial transaction: the manufacturers, the retailers, the users, and even regulatory bodies (such as the Food and Drug Administration (FDA) in the United States). Automated item level inventory identification using RFID systems will revolutionize supply chain management by enabling applications such as automated real-time inventory monitoring (at the shelf and in the warehouse), automated quality control, and automatic check-out.

The significant benefits that an inexpensive, open standards-based RFID system can provide are widely understood and acknowledged. At the same time, typical low-cost transponders are priced in the range of US$0.50-US$1.00, and RFID systems lack widely accepted and implemented standards for communication and functionality, thereby limiting their practical usefulness and keeping their system costs too high for many applications. In order to achieve significant item-level penetration within most supply chain applications, transponders will need to be priced well under US$0.10, and preferably under US$0.05. These cost targets cannot be achieved without a system-level approach that encompasses every aspect of the RFID technology, from IC design to RF protocols, from reader design to back-end data systems, and from IC manufacturing to antenna manufacturing. The challenge has been to develop a complete open standards-based system that enables the design and manufacture of lo! w-cost RFID systems.

The Auto-ID Center, an industry sponsored research center with laboratories at Massachusetts Institute of Technology, USA, Cambridge University, UK, and the University of Adelaide, AU, has designed, developed, and deployed within a large-scale field trial an open standards-based system that enables the unique identification of and retrieval of information on ubiquitously tagged items. The Center, in conjunction with its sponsors, has also undertaken projects to design and manufacture open standard low-cost RFID transceivers and transponders capable of little more than communicating a unique identifier stored within them. Low-cost transponders enable the tagging and unique identification of virtually all man-made items.

The commercial availability of low-cost, Auto-ID Center standards-based RFID systems by mid-2003 has poised these systems to be one of the earliest and perhaps most explosive opportunities in ubiquitous computing. As these systems leave the industrial applications and enter our daily lives, privacy and security related issues will play an increasingly important role in their use and ubiquity. The pupose of this paper is to explain the technology, the challenges, and the opportunities ubiquitous RFID systems present to the security and privacy communities.

2 A Brief Introduction to RFID Systems

2.1 Basic System Components

All RFID systems are comprised of three main components:

– the RFID tag, or *transponder,* which is located on the object to be identified and is the data carrier in the RFID system,

- the RFID reader, or *transceiver,* which may be able to both read data from and write data to a transponder, and
- the *data processing subsystem* which utilizes the data obtained from the transceiver in some useful manner.

Typical *transponders* (*trans*mitters/re*sponders*) consist of a microchip that stores data and a coupling element, such as a coiled antenna, used to communicate via radio frequency communication. Transponders may be either active or passive. Active transponders have an on-tag power supply (such as a battery) and actively send an RF signal for communication while passive transponders obtain all of their power from the interrogation signal of the transceiver and either reflect or load modulate the transceiver's signal for communication. Most transponders, both passive and active, communicate only when they are interrogated by a transceiver.

Typical *transceivers* (*trans*mitter/re*ceivers*), or RFID readers, consist of a radio frequency module, a control unit, and a coupling element to interrogate electronic tags via radio frequency communication. In addition, many transceivers are fitted with an interface that enables them to communicate their received data to a data processing subsystem, e.g., a database running on a personal computer. The use of radio frequencies for communication with transponders allows RFID readers to read passive RFID tags at small to medium distances and active RFID tags at small to large distances even when the tags are located in a hostile environment and are obscured from view.

The basic components of an RFID system combine in essentially the same manner for all applications and variations of RFID systems. All objects to be identified are physically tagged with transponders. The type of tag used and the data stored on the tag varies from application to application.

Transceivers are strategically placed to interrogate tags where their data is required. For example, an RFID-based access control system locates its readers at the entry points to the secure area. A sports timing system, meanwhile, locates its readers at both the starting line and the finish line of the event. The readers continuously emit an interrogation signal. The interrogation signal forms an interrogation zone within which the tags may be read. The actual size of the interrogation zone is a function of the transceiver and transponder characteristics. In general, the greater the interrogation signal power and the higher the interrogation signal frequency, the larger the interrogation zone. Sending power to the transponders via the reader-to-tag communication signal is the bottleneck in achieving large read range with passive tags. Active tags do not suffer from this drawback; thus, they typically have larger communication ranges than an otherwise equivalent passive tag.

The transceivers and transponders simply provide the mechanism for obtaining data (and storing data in the case of writable tags) associated with physical objects.

Passive RFID systems are the most promising to provide low-cost ubiquitous tagging capability with adequate performance for most supply chain management applications. These low-cost RFID systems are, of necessity, very resource

limited, and the extreme cost pressures make the design of RFID systems a highly coupled problem with sensitive trade-offs. Unlike other computation systems where it is possible to abstract functionality and think modularly, almost every aspect of an RFID system affects every other aspect. We present a brief overview of the critical components of RFID technology and summarize some of these trade-offs in passive RFID design.

2.2 Transceiver-Transponder Coupling and Communication

Passive RFID tags obtain their operating power by harvesting energy from the electromagnetic field of the reader's communication signal. The limited resources of a passive tag require it to both harvest its energy and communicate with a reader within a narrow frequency band as permitted by regulatory agencies. We denote the center of this frequency band by f, and we refer to RFID systems operating at frequency f with the understanding that this is the center frequency of the band within which it operates.

Passive tags typically obtain their power from the communication signal either through inductive coupling or far field energy harvesting. Inductive coupling uses the magnetic field generated by the communication signal to induce a current in its coupling element (usually a coiled antenna and a capacitor). The current induced in the coupling element charges the on-tag capacitor that provides the operating voltage, and power, for the tag. In this way, inductively coupled systems behave much like loosely coupled transformers. Consequently, inductive coupling works only in the near-field of the communication signal. The near field for a frequency f extends up to $1/(2\pi f)$ meters from the signal source.

For a given tag, the operating voltage obtained at a distance d from the reader is directly proportional to the flux density at that distance. The magnetic field emitted by the reader antenna decreases in power proportional to $1/d^3$ in the near field. Therefore, it can be shown that for a circularly coiled antenna the flux density is maximized at a distance d (in meters) when $R \cong \sqrt{2} \cdot d$, where R is the radius of the reader's antenna coil. Thus, by increasing R the communication range of the reader may be increased, and the optimum reader antenna radius R is 1.414 times the demanded read range d.

Far field energy harvesting uses the energy from the interrogation signal's far field signal to power the tag. The far field begins where the near field ends, at a distance of $1/(2\pi f)$ from the emitting antenna. The signal incident upon the tag antenna induces a voltage at the input terminals of the tag. This voltage is detected by the RF front-end circuitry of the tag and is used to charge a capacitor that provides the operating voltage for the tag.

There is a fundamental limitation on the power detected a distance d away from a reader antenna. In a lossless medium, the power transmitted by the reader decreases as a function of the inverse square of the distance from the reader antenna in the far field.

A reader communicates with and powers a passive tag using the same signal. The fact that the same signal is used to transmit power and communicate data creates some challenging trade-offs. First, any modulation of the signal causes a

reduction in power to the tag. Second, modulating information onto an otherwise spectrally pure sinusoid spreads the signal in the frequency domain. This spread, referred to as a "side band," along with the maximum power transmitted at any frequency, is regulated by local government bodies in most parts of the world. These regulations limit the rate of information that can be sent from the reader to the tag. RFID systems usually operate in free bands known as Industrial-Scientific-Medical (ISM) bands, where the emitted power levels and the side band limits tend to be especially stringent.

The signaling from the tag to the reader in passive RFID systems is not achieved by active transmission. Since passive tags do not actively transmit a signal, they do not have a regulated limit on the rate of information that can be sent from the passive tag to the reader. In the near field, tag to reader communication is achieved via *load modulation.* Load modulation is achieved by modulating the impedance of the tag as seen by the reader. In the far field, tag to reader communication is achieved via *backscatter.* Backscatter is achieved by modulating the radar cross-section of the tag antenna. Comprehensive reviews of the operation of tags and readers are available in [8] and [17].

The powering of and communication with passive tags with the same communication signal places restrictions on the functionality and transactions the tags are capable of. First, there is very little power available to the digital portion of the integrated circuit on the tag. This limits the functionality of the tag. Second, the length of transactions with the tag is limited to the time for which the tag is expected to be powered and within communication range. Governmental regulations can further limit communication timings. In the US in the 915 MHz ISM band, regulations require that, under certain operating conditions, the communication frequency change every 400 ms. Since every change in frequency may cause loss of communication with a tag, transponders must not be assumed to communicate effectively for longer than 400 ms. Finally, it is important to minimize state information required in passive tags. In many practical situations, power supplied to the tag may be errat! ic, and any long-term reliance on state in the tag may lead to errors in the operation of a communication protocol.

2.3 Data Coding

The data, consisting of ones and zeroes, communicated between tags and readers must be sent in a reliable manner. There are two critical steps to reliable communication, the encoding of the data and the transmission of the encoded data, that is, the modulation of the communication signal. The combination of coding and modulation schemes determines the bandwidth, integrity, and tag power consumption.

The coding and modulation used in RFID communications is limited by the power and modulation/demodulation capabilities of the tags. Another limiting factor is the bandwidth occupied by the signal. Readers are capable of transmitting at high power but are limited to narrow communication bands by communications regulations; therefore, the encoding used from reader to tag usually

needs to occupy a low bandwidth. Passive tags, however, do not actively transmit a signal; therefore, the encoding used for tag to reader communication can occupy a high bandwidth.

There are two broad categories of codes used in RFID: level codes and transition codes. Level codes represent the bit with their voltage level. Transition codes capture the bit as a change in level. Level codes, such as Non-Return-to-Zero (NRZ) and Return-to-Zero (RZ), tend to be history independent; however, they are not very robust. Transition codes can be history dependent, and they can be robust. Figure 1 illustrates several codes.

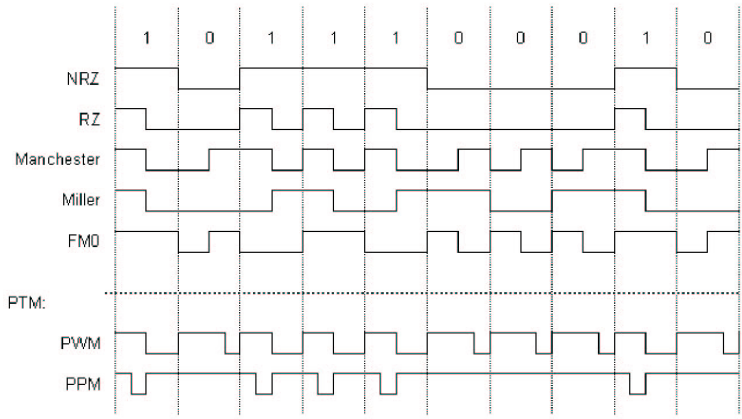

Fig. 1. Examples of several coding schemes.

The simplest code is Pulse Pause Modulation (PPM) in which the length between pulses is used to convey the bit. PPM codes provide low bit rates but occupy only a small bandwidth and are very easy to implement. In addition, these encodings can be adapted easily to ensure uninterrupted power supply since the signal does not change for long periods of time.

The Manchester code is a higher bandwidth transition code that represents a 1 as a negative transition at the half period and a 0 as a positive transition at a half period. The Manchester Code provides for efficient communication since the bit rate is equal to the bandwidth of the communication.

In RFID, the coding technique must be selected with three considerations in mind: 1) the code must maintain power to the tag as much as possible, 2) the code must not consume too much bandwidth, and 3) the code must permit the detection of collisions. The collision detection ability of a code is discussed further in Section 2.5. Depending on the bandwidth available, most RFID systems use PPM or PWM to communicate from reader to tag and Manchester or NRZ to communicate from tag to reader.

2.4 Modulation

The data coding scheme determines how the data is represented in a continuous stream of bits. How that stream of bits is communicated between the tag and the reader is determined by the modulation scheme. For convenience, RF communications typically modulate a high frequency carrier signal to transmit the baseband code. The three classes of digital modulation are Amplitude Shift Keying (ASK), Frequency Shift Keying (FSK) and Phase Shift Keying (PSK). The choice of modulation is based on power consumption, reliability requirements, and bandwidth requirements. All three forms of modulation may be used in the return signal although ASK is most common in load modulation at 13.56 MHz, and PSK is most common in backscatter modulation.

A problem unique to RFID systems is the vast difference in power between the signal outgoing from the reader and that returning to the reader as reflected from the tag. In some situations, this difference may be in the range of 80-90 dB [8], and the return signal may be impossible to detect. To avoid this problem, the return signal is sometimes modulated onto a sub-carrier, which is then modulated on to the carrier. For example, in the ISO 15693 standard for RFID, a sub-carrier of 13.56/32 (= 423.75 KHz) is used.

2.5 Tag Anti-collision

When multiple tags respond simultaneously to a reader's signal, their communication signals can interfere with one another. This interference is referred to as a collision and typically results in a failed transmission. In order for a reader to communicate with multiple tags, a method for collision free tag communication must be employed. These methods are referred to as anti-collision methods. An anti-collision method must be employed if an application will typically have more than one tag communicating with a reader at the same time.

Anti-collision methods, or algorithms, in tags have similarities to anti-collision algorithms in networking. Unlike standard networking however, RFID tags pose a number of problems that arise from the very limited resources that they are provided with. First, they can afford only limited computation power. Second, state information, such as what portion of the tags identifier has already been read, may be unreliable. Third, collisions may be difficult to detect due to widely varying signal strengths from the tags. Finally, as in most wireless networks, transponders cannot be assumed to be able to hear one another.

A common classification of anti-collision algorithms, either *probabilistic* or *deterministic,* is based upon how the tags respond during the anti-collision algorithm. In probabilistic algorithms, the tags respond at randomly generated times. There are several variations of probabilistic protocols depending on the amount of control the reader has over the tags. Many probabilistic algorithms are based on the Aloha scheme in networking [3]. The times at which readers can respond can be slotted or continuous. The ISO 15693 protocol, for example, supports a slotted Aloha mode of anti-collision.

Deterministic schemes are those in which the reader sorts through tags based on their unique identification number. The simplest deterministic scheme is the

binary tree-walking scheme, in which the reader traverses the tree of all possible identification numbers. At each node in the tree, the reader checks for responses. Only tags whose identifier is a child of the checked node respond. The lack of a response implies that the sub-tree is empty. The presence of a response gives the reader an indication as to where to search next.

The performance metrics that are traded-off by these algorithms and their variants include: 1) the speed at which tags can be read, 2) the outgoing bandwidth of the reader signal, 3) the bandwidth of the return signal, 4) the amount of state that can be reliably stored on the tag, 5) the tolerance of the algorithm to different types of noise in the field, 6) the cost of the tag, 7) the cost of the reader, 8) the ability to tolerate tags which enter and leave the field during the inventory-taking process, 9) the desire to count tags exactly as opposed to sampling them, and finally, 10) the range at which tags can be read.

The impact of regulated reader-to-tag bandwidth on the anti-collision protocol can be severe. In the US, for example, two common operating frequencies for RFID systems are the 13.56 MHz and the 915 MHz ISM bands. The regulations on the 13.56 MHz band offer significantly less bandwidth in the communication from the reader to the tag than do the regulations on the 915 MHz band. For this reason, Aloha-based anti-collision algorithms are more common in systems that operate in the 13.56 Mhz band and deterministic anti-collision algorithms are more common in the 915 Mhz band.

In practice, most RFID anti-collision algorithms tend to be an amalgam of probabilistic and deterministic concepts. Almost all require a unique ID to sort through the tags. This in itself has implications on privacy, as we will discuss later. The interplay between the anti-collision algorithm, the identifier, and the bandwidth available has an impact on all transactions between the reader and the tag. Approaches to security and privacy must therefore be geared to these very subtle trade-offs. Protocols to secure the tag at 13.56 Mhz, for example, must use far less signaling from reader-to-tag than at 915 Mhz. Either way, when there are several tags in the field, it is best to leverage the anti-collision algorithms as much as possible for efficiency.

2.6 Reader Anti-collision

RFID systems have traditionally been used in sparse applications where the readers tend to be far apart. In the applications we have explored, particularly those in supply chain management, the density of readers will often be very high, creating a new class of problems related to reader interference. We first reported the *Reader Collision Problem* in [7]. The solution to a reader collision problem allocates frequencies over time to a set of readers. The solution may be obtained in either a distributed or centrally controlled manner.

Reader collision problems have some similarities to frequency assignment problems in mobile telephone systems. However, the approaches that work in mobile telephones do not translate to RFID systems due to the limited functionality in RFID tags. The inability of the transponders to aid in the communication process means that they are unable to discriminate between two readers

communicating with them simultaneously. As a result, two readers that may communicate with the same tag must communicate at different times.

In a cooperative, trusted environment, reader collisions can be handled in a fairly seamless way. However, complications may arise in the execution of commands that change the state of the tag. If the reader executing a series of state changing actions is interrupted by another reader, it may be forced to relinquish control over the tag. The new reader that acquires the tag may further change the state of the tag without the cooperation of the first reader. Transactions between readers and tags must therefore be brief and atomic.

2.7 Frequencies and Regulations

The operation of RFID systems worldwide is regulated by local governmental bodies which control the electromagnetic spectrum in a region. Most RFID systems operate in so-called Industrial-Scientific-Medical (ISM) bands. These bands are freely available for use by low-power, short-range systems. The ISM bands are designated by the International Telecommunications Union (ITU) [11]. A comprehensive summary of the standards is available in [17]. The most commonly used ISM frequencies for RFID are 13.56 MHz and 902-928 MHz (in the US only). In addition, the low frequency band 9kHz-135 kHz is available for unlicensed use in most regions, and the 868MHz-870MHz band is available for use by nonspecific short-range devices in Europe. Each band has its own radiation power and bandwidth regulations.

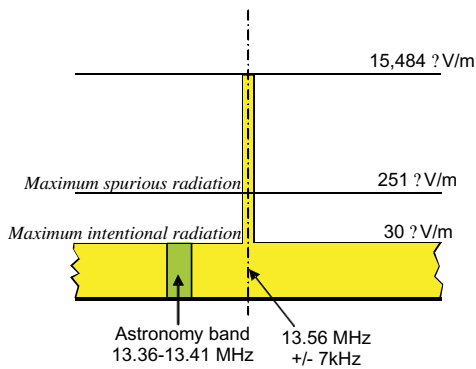

Fig. 2. The 13.56 MHz ISM band US power regulations.

Each frequency band brings its own challenges and advantages in terms of operation. The 13.56 MHz band shown in Figure 2 offers a great deal of asymmetry between the forward and reverse communication. Since readers must power the tags in passive RFID systems, the reader-to-tag communication must be at maximum power for maximum communication range. This limits the bandwidth

in reader-to-tag communication to a total of 14kHz. However, there is a great deal more bandwidth available for the low-power return communication.

The 915 MHz ISM band in the US, in contrast, allows multiple reader-to-tag communication options. The option that enables the longest communication range, the most commonly used mode in RFID systems, requires the reader to change its communication frequency every 0.4 seconds. The reader may 'hop' between 50 or more channels, each with a maximum bandwidth of 250 kHz. Frequency hopping imposes several limitations on RFID systems. The most severe of these limitations is that a tag cannot be assumed to be in continuous communication across a frequency hop. This means that transactions with 915 MHz RFID systems in the US should be limited to within the 0.4 second intervals allocated to any single frequency sub-band. Constraints of this type also point to the need for limited length, atomic transactions in RFID systems, a requirement which must be respected in the design of security and privacy systems.

3 The EPC System: A Minimalist Approach

At the Auto-ID Center, we have developed and implemented a system that enables all physical objects to be connected in real-time to the Internet by affixing an RFID tag to the object [14]. The scale of the system (essentially a several quadrillion node network), combined with the trade-offs in RFID design, created an intriguing design challenge. We utilized a minimalist strategy for the RFID tags (the most replicated component of the system) to enable extremely low-cost RFID systems. The result is a system that minimizes the functionality on the tag by moving that functionality to the 'network.'

The four key components of this system are the Electronic Product Code (EPC), the Object Name Service (ONS), the Savant, and the RFID transponders.

The EPC. The Electronic Product Code (EPC) is an identification scheme designed to enable the unique identification of all physical objects. This is the only data required to be stored on a tag, for once the unique identity of an object is established, information about that object can be obtained from the network. As such, the EPC acts like a pointer to this information.

The ONS. The Object Name Service (ONS) is a directory service that maps the EPC to an IP (Internet Protocol) address where information about the associated object can be written and/or accessed. The ONS is based entirely on the Domain Name Service (DNS), the directory service used on the Internet today to map a domain name (e.g., www.mit.edu) to an IP address (e.g., 18.181.0.31). At the IP address pointed to by the ONS, data about the particular object is stored in XML [20] format, and can be accessed by standard methods like HTTP and SOAP [19].

ONS reduces the burden on the transponders, and provides several advantages simultaneously. First, it reduces the memory and power requirements on the tag. Second, by transferring much of the data communication to a much higher-bandwidth back-end network, it saves precious wireless bandwidth. Third,

it makes the system more robust: while it is difficult to store and recover information from a failed tag, it is possible to back up databases inexpensively. Finally, this approach significantly reduces the footprint of the tag's microchip, reducing the cost of the transponder. (The cost of the microchip is proportional to its area [15].)

Savant. The Savant system is a hierarchical control and data management building block that can be used to provide automated control functionality and manage the large volumes of data generated by the RFID readers. A Savant enables the distributed creation of a reader network by acting as a gateway to the next higher level in the Savant hierarchy, effectively isolating the reader sub-network. The use of Savants enables distributed security by providing convenient points for network isolation.

The Savant network further reduces the burden on the tags while providing several advantages. First, it reduces the memory and power requirements on the tags by transferring the computationally intensive functionality to a powered system. Second, it makes the system more robust: any single point of failure has local effects. Third, it enables the entire system to be scalable as systems and reader sub-networks may be added seemlessly. Finally, the Savant network significantly reduces the footprint of the tag's microchip, reducing the cost of the transponder.

RFID Transponders. RFID transponders are the most numerous and cost sensitive of our system components. We have designed RFID protocols for both 13.56 MHz and 915 MHz, both with the aim of having minimum cost identification tags with acceptable performance for supply chain applications. Both transponders are designed to store a unique identifier, an EPC, and have that identifier retrieved as part of the anti-collision algorithm. The 915 MHz, UHF, transponder utilizes a directed tree search anti-collision algorithm, while the 13.56 MHz, HF, transponder utilizes a slotted Aloha-based anti-collision algorithm. Both transponders also implement a password protected `Self Destruct` command, that enables the owner of the tag to electrically and permanently destroy the tag.

The implementation cost of securing the `Self Destruct` command was weighed against the benefits of that security. It was determined that a secret key must be used to execute the `Self Destruct` command; therefore, requiring the destruction of a single tag at a time. The secret key is meant only to discourage the unauthorized destruction of tags. In a pervasive reader environment, unauthorized `Self Destruct` commands can be detected by the readers, enabling a proper reaction to the issuance of these commands.

We have taken a building-block approach to RFID transponder design in that these minimal functionality tags form the foundation of the functionality that will be found in higher-functionality tags. These higher functionality tags may be used in applications that can afford the additional cost of the transponder and require the transponder to implement the functionality.

4 RFID Security Benefits and Threats

Universally deploying RFID tags offers many potential security benefits, yet may expose new privacy threats. Otherwise intrusive or cumbersome security practices, such as airline passenger and baggage tracking, can be made practical by using RFID systems. Authentication systems already take advantage of RFID technology, for example car key-less entry systems. Embedding RFID tags as seals of authenticity in documents, designer products, and currency may discourage forgery. While RFID tags improve certain security properties in these applications, they may exacerbate privacy threats or pose new security risks.

RFID systems are different from other means of identification because RF communication is non-contact and non-line-of-sight, whereas other means of identification are either contact-based or require line-of-sight. In other words, it is more difficult for the owner of the RF tag to physically impede communication with the tag. The promiscuity of RF tags is not unique; magnetic stripe cards, for example, are promiscuous, but we assume that the owner of the card takes the physical responsibility of preventing unauthorized users from physically accessing the card. Of course, the propagation characteristics of electromagnetic fields do limit the range from which passive RFID cards can be read. In fact, most tags operating at 13.56 MHz cannot be read from more than a meter away, and 915 MHz tags are difficult to read through most materials. Yet, as the information stored on the tag becomes more and more valuable, it is necessary to think through some of the security and privacy! related issues in RFID. We present such a discussion in this section, ending with a proposed approach.

4.1 Previous Work

The contactless interface and constrained computational resources of RFID devices present a unique set of characteristics most closely related to smart cards. Many relevant lessons may be gleaned from the wealth of smart card and tamper resistant hardware research. [1] discusses a range of smart card protocols and analyzes cost and security trade-offs. Many RFID tags will operate in hostile environments and may be subject to intense physical attacks. Analysis of smart cards operation in hostile environments is presented in [9], while [18] provides an excellent overview of many physical attacks and countermeasures. Several specific lower cost physical attacks are detailed in [2] and are part of ongoing research at the University of Cambridge's TAMPER Lab [16]. Many results pertaining to implementation of cryptographic primitives on smart cards apply to RFIDs. Cautionary information regarding implementation of AES i! n smart cards appears in [5]. Being contactless and passively powered may make RFID devices especially susceptible to fault induction or power analysis attacks. Both [4] and [12] highlight many of these issues in cryptographic devices.

4.2 Security Goals

It is useful to state clear security goals when discussing security properties of various RFID designs. Tags must not compromise the *privacy* of their holders. Information should not be leaked to unauthorized readers, nor should it be

possible to build long-term tracking associations between tags and holders. To prevent tracking, holders should be able to detect and disable any tags they carry. Publicly available tag output should be randomized or easily modifiable to avoid long-term associations between tags and holders. Private tag contents must be protected by access control and, if interrogation channels are assumed insecure, encryption.

Both tags and readers should *trust* each other. Spoofing either party should be difficult. Besides providing an access control mechanism, mutual authentication between tags and readers also provides a measure of trust. Session hijacking and replay attacks are also concerns. Fault induction or power interruption should not compromise protocols or open windows to hijack attempts. Both tags and readers should be resistant to replay or man-in-the-middle attacks.

4.3 Low-Cost RFID Issues

With these security goals in mind, consider the security properties of passive factory-programmed, read-only tags. Each tag contains a unique identifier such as an EPC. While no more "promiscuous" than an optical bar code, automated monitoring of RF tags is possible. This basic design clearly violates the privacy goal since tracking tag holders and reading tag contents are possible if the tag is properly presented to a reader's interrogation field. Neither tags nor readers are authenticated; therefore, no notion of trust exists either.

To address these deficiencies, suppose we adopt a policy of erasing unique serial numbers at the point of sale. Consumer held tags would still contain product code information, but not unique identification numbers. Unfortunately, tracking is still possible by associating "constellations" of particular tag types with holder identities. For example, a unique penchant for RFID-tagged Gucci shoes, Rolex watches and Cohiba cigars may betray your anonymity. Furthermore, this design still offers no trust mechanism.

Providing the stated security goals requires implementing access control and authentication. Public key cryptography offers a solution. A particular (type of) reader's public key and a unique private key may be embedded into each tag. During interrogation, tags and readers may mutually authenticate each other with these keys using well understood protocols. To prevent eavesdropping within the interrogation zone, tags may encrypt their contents using a random nonce to prevent tracking. Unfortunately, supporting strong public key cryptography is beyond the resources of low cost (US$0.05-0.10) tags, although solutions do exist for more expensive tags [13] .

Symmetric message authentication requires each tag to share a unique key with a reader or for a key to be shared among a batch of tags. To support a unique key per tag, a complex key management overhead is necessary. If keys are to be shared, tags must be resilient to physical attacks described in [18]; otherwise, compromising a single tag effective compromises an entire a batch. Implementing secure memory on a low cost tag with a logic gate count in the hundreds is a daunting task, especially in light of the difficulty in securing memory on relatively resource abundant smart cards. Even supporting strong symmetric encryption is a challenge in the short term.

4.4 Some Approaches to RFID Protection

Accepting short-term limitations on low-cost tag resources, we discuss a simple RFID security scheme based on a one-way hash function. In practice, a hardware optimized cryptographic hash function would suffice, assuming it may be implemented with significantly fewer resources than symmetric encryption. In this design, each hash-enabled tag contains a portion of memory reserved for a "meta-ID" and operates in either an unlocked or locked state. While unlocked, the full functionality and memory of the tag are available to anyone in the interrogation zone.

To lock a tag, the owner computes a hash value of a random key and sends it to the tag as a lock value, i.e. *lock=hash(key)*. In turn, the tag stores the lock value in the meta-ID memory location and enters the locked state. While locked, a tag responds to all queries with the current meta-ID value and restricts all other functionality. To unlock a tag, the owner sends the original key value to the tag. The tag then hashes this value and compares it to the lock stored under the meta-ID. If the values match, the tag unlocks itself.

Each tag always responds to queries in some form and thus always reveals its existence. Tags will be equipped with a physical self-destruct mechanism and will only be unlocked during communication with an authorized reader. In the event of power loss or transmission interruption, tags will return to a default locked state. A trusted channel may be established for management functions, such as key management, tag disabling or even tag writes, by requiring physical contact between a control device and a tag. Requiring physical contact for critical functionality helps defend against wireless sabotage or denial of service attacks.

The hash-based lock mechanism satisfies most of our privacy concerns. Access control to tag contents is restricted to key holders. Individuals may both locate and disable tags they may be carrying since tags always respond to queries. Long-term associations can be avoided since locked tags only respond with the correct meta-ID. One caveat is that stale meta-ID values may be used to build tracking associations over time. This necessitates periodically refreshing meta-ID values by unlocking and re-locking tags.

Although authenticating readers and providing a trusted channel satisfies some of our trust requirements, this design does sacrifice several security properties to save costs; specifically tag authentication. Tag MAC functionality would allow tags to authenticate themselves, but is beyond current low-cost tag resources. Lacking authentication exposes tags to man-in-the-middle attacks since an attacker can query tags for meta-IDs, rebroadcast those values to a legitimate reader, and later unlock the tags with the reader's response keys. Many key-less car entry systems currently possess the same vulnerability. Regardless, attackers without access to an authorized reader cannot access tag contents outside physical channels.

4.5 Future Research Directions

While this candidate design partially satisfies some desired security properties, more secure implementations require several developments. One key line of research is the further development and implementation of low cost cryptographic

primitives. These include hash functions, random number generators and both symmetric and public key cryptographic functions. Low cost hardware implementations must minimize circuit area and power consumption without adversely affecting computation time. RFID security may benefit from both improvements to existing systems and from new designs. More expensive RFID devices already offer symmetric encryption and public key algorithms such as NTRU [10,13]. Adaptation of these algorithms for the low-cost (US$0.05-0.10), passive RFID devices should be a reality in a matter of years.

Protocols utilizing these cryptographic primitives must be resilient to power interruption and fault induction. Compared to smart cards, RFID tags possess more vulnerabilities to these types of attacks. Protocols must account for disruption of wireless channels or communication hijack attempts. Tags themselves must gracefully recover from power loss or communication interruption without compromising security.

Continually improving technology will steadily blur the line between RFID devices, smart cards and ubiquitous computers. Research benefiting the security of RFID devices will help pave the way for a universal, secure ubiquitous computing system. Developments related to RFID tags and other embedded systems may all contribute to the creation of a robust and secure infrastructure offering many exciting potential applications.

5 Conclusions

This article is a summary of a research effort underway by three universities, more than 60 companies, and more than 50 researchers world-wide.

The effort has been fueled by the potential economic impact of inexpensive, ubiquitous item identification in the supply chain. The roadmap towards cheap tags has been laid out, but like any research effort, uncertainty is a part of the challenge. Several technology alternatives will need to be tested for each component of the system before the optimal one is determined. Even after the first cheap tags have been manufactured, scaling production to the volumes needed to meet expected demand will be a challenge. It may be years before the supply meets the enormous demand that a technology of this type is projected to generate. However, it is these very volumes that make it necessary for the technology to be carefully thought out to save every fraction of a cent in the cost of a tag and to ensure the security and privacy of its future users.

Acknowledgments. The authors wish to thank Ron Rivest and Peter Cole for their continued support of this work.

References

1. M. Abadi, M. Burrows, C. Kaufman, and B. W. Lampson. Authentication and delegation with smart-cards, In *Theoretical Aspects of Computer Software*, pages 326–345, 1991.

2. R. Anderson and M. Kuhn. Low cost attacks on tamper resistant devices. In *IWSP: International Workshop on Security Protocols, LNCS*, 1997.
3. B. Bing. *Broadband Wireless Access,* Boston, Kluwer Academic Publishers, 2000.
4. D. Boneh, R.A. DeMillo, and R.J. Lipton. On the importance of checking cryptographic protocols for faults. In *EUROCRYPT'97*, volume 1233, pages 37–51. Lecture Notes in Computer Science, Advances in Cryptology, 1997.
5. S. Chari, C. Jutla, J.R. Rao, and P. Rohatgi. A cautionary note regarding evaluation of AES candidates on smart-cards. In *Second Advanced Encryption Standard (AES) Candidate Conference*, Rome, Italy, 1999.
6. EAN International and the Uniform Code Council, Note to Editors, http://www.ean-int.org/index800.html
7. D. Engels. The Reader Collision Problem. Technical Report. MIT-AUTOID-WH-007, 2001. http://www.autoidcenter.org/research/MIT-AUTOID-WH-007.pdf.
8. K. Finkenzeller. *RFID Handbook,* John Wiley & Sons. 1999.
9. H. Gobioff, S. Smith, J.D. Tygar, and B. Yee. Smart cards in hostile environments. In *2nd USENIX Workshop on Elec. Commerce*, 1996.
10. J. Hoffstein, J. Pipher, and J.H. Silverman. NTRU: A ring-based public key cryptosystem. *Lecture Notes in Computer Science*, volume 1423, 1998.
11. International Telecommunications Union. Radio Regulations, Vol. 1, 1998.
12. B.S. Kaliski Jr. and M.J.B. Robshaw. Comments on some new attacks on cryptographic devices. RSA Laboratories' Bulletin No. 5, July 14, 1997. Available from http://www.rsasecurity.com/rsalabs/bulletins/.
13. NTRU. GenuID. http://www.ntru.com/products/genuid.htm.
14. S. Sarma, K. Ashton, D. Brock. The Networked Physical World, Technical Report MIT-AUTOID -WH-001, 1999. http://www.autoidcenter.org/research/MIT-AUTOID-WH-001.pdf.
15. S. Sarma. Towards the 5 cent Tag, Technical Report MIT-AUTOID -WH-006, 2001. http://www.autoidcenter.org/research/MIT-AUTOID-WH-006.pdf.
16. TAMPER Lab. University of Cambridge Tamper and Monitoring Protection Engineering Research Lab, http://www.cl.cam.ac.uk/Research/Security/tamper.
17. T. Scharfeld. An Analysis of the Fundamental Constraints on Low Cost Passive Radio-Frequency Indentification System Design. MS Thesis, Department of Mechanical Engineering, Massachusetts Institue of Technology, Cambridge, MA 02139, 2001.
18. S.H. Weigart. Physical security devices for computer subsystems: A survey of attacks and defences. *CHES 2000*, *Lecture Notes in Computer Science*, volume 1965, pages 302–317, 2000.
19. World Wide Web Consortium. http://www.w3c.org/SOAP/
20. World Wide Web Consortium. http://www.w3c.org/XML/