Security and Privacy in Two RFID Deployments, With New Methods For Private Authentication and RFID Pseudonyms

by David Alexander Molnar

Research Project

Submitted to the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley, in partial satisfaction of the requirements for the degree of **Master of Science, Plan II**.

Approval for the Report and Comprehensive Examination:

Committee:

Professor David Wagner
Research Advisor

(Date)

* * * * * * *

Professor Michael Franklin
Second Reader

(Date)

Contents

List of Figures

Abstract

We study security and privacy in deployments of Radio Frequency Identification (RFID) technology and propose novel mechanisms for improving RFID privacy. In the first part of the thesis, we consider two real deployments of RFID technology: library books and electronic passports. For each deployment, we set out security and privacy issues. Then we analyze existing RFID technology in the context of these issues. We relate these issues to concrete technical problems, such as the problem of *private authentication*: how can Alice and Bob determine that they share a secret key without an eavesdropper learning their identities?

The second part of the thesis describes new techniques for solving these problems. We describe a symmetric-key private authentication protocol which requires work logarithmic in the number of RFID tags in a system, while all previous solutions required linear work. Then we discuss using a trusted third party called an "infomediary" to enforce a privacy policy and a way to realize the infomediary by "recoding" RFID tags. We move beyond recoding with a method for tags to generate a new one-time pseudonym on each reading. Our pseudonym scheme requires work logarithmic in the number of tags for an infomediary to learn the real tag ID from a pseudonym, while all previous schemes required linear work. A drawback is that our scheme loses some, but not all, privacy if individual tags are compromised; we show that the result is a tradeoff between privacy and reader efficiency. Our scheme also supports *delegation* to third parties of the ability to learn tag IDs for a limited number of reads. We show that delegation enables the transfer of an RFID-tagged item between two mutually distrustful parties. Finally, we close with open problems and future directions.

Acknowledgments

I have many people to thank.

David Wagner is a model advisor, collaborator, and mentor. Without his encouragement and help, I probably would not be at Berkeley and possibly not even in graduate school. Certainly this work would not exist. There is much more I could say, but I will just leave it at a resounding "Thank you, David!"

Michael Franklin graciously agreed to act as the reader for this report. I thank him for his time, especially considering that this has turned out to be longer than a typical Master's report. I am also grateful to him for spending a Berkeley Database Group meeting on applications of the pseudonym protocol described in this report.

I have been fortunate to work with wonderful people during the projects reported here. I am indebted to them for their insight, good humor, and patience. In addition to David Wagner, I have been lucky to work with Nathaniel Good, John Han, Ari Juels, Elizabeth Miles, Deirdre Mulligan, Laura Quilter, Andrea Soppera, Ross Stapleton-Gray, and Jennifer Urban. A more detailed account of our collaborations appears in the Introduction. Thank you!

I also thank Simson Garfinkel and Beth Rosenberg for organizing the Addison/Wesley book on *RFID Privacy, Security, and Applications* and for including our chapter. Editing a large multi-author book into a coherent whole is a difficult job, and getting it out in time to remain relevant is even more difficult. They have pulled it off. I am honored to have been a part of it.

For the work on RFID in library books, I thank the following people for their feedback and advice: Alicia Abramson, Rebekah E. Anderson, Oleg Boyarsky, Justin Chen, Karen Coyle, Karen Duffy, Elena Engel, Gerard Garzon, Craig Gentry, Jackie Griffin, Steve Halliday, John Han, Craig K. Harmon, Susan Hildreth, Kris Hildrum, Eric Ipsen, Jayanth Kumar Kannan, Doug Karp, Judith Krug, Elizabeth Miles, Dan Moniz, Julie Odofin, Matt Piotrowski, Zulfikar Ramzan, Pam Samuelson, Karen Saunders, Rupert Scammell, David Schultz, Paul Simon, Al Skinner, Laura Smart, Ross Stapleton-Gray, Lee Tien, Peter Warfield, and Hoeteck Wee. I also thank the anonymous reviewers of ACM CCS 2004 for helpful comments. The work was supported by DARPA NEST contract F33615-01-C-1895.

For the work on RFID pseudonyms, I thank Deepti Agrawal, Gildas Avoine, Michael Backes, Trevor Burbridge, Etienne Dysli, Arnaud Jacquet, Pascal Junod, Vivekanand Korgaonkar, Philippe Oechslin, and the anonymous reviewers of

CHES 2005 and SAC 2005 for helpful comments. I also thank the members of the Berkeley Center for Emerging Networked Trustworthy Systems, the cryptography group of EPFL, and members of IBM Research Zurich for feedback on early presentations of this work. Together with my co-authors, I gratefully acknowledge support from NSF grant CCR-0093337, the Sloan Research Fellowship, and British Telecom.

For our work on electronic passports, I thank Neville Pattinson for helpful discussions and for giving us access to his white paper. I thank Rei Onishi, Seth Schoen, and Lee Tien for helpful discussions on e-passports and their policy ramifications. I am also indebted to Bart Jacobs, Tamas Visegrady, and researchers at IBM Zurich for discussions regarding European electronic passports. Christian Cachin, Jan Camenisch, Susan Hohenberger, Guenter Karjoth, and Anna Lysyanskaya served as gracious hosts during the visit to IBM Zurich during which I discussed the work. The anonymous reviews from SecureComm 2005 helped us improve our presentation, as did comments from Markus Kuhn, Avishai Wool, and the other attendees. The research was supported by NSF grants CCR-0093337 and CCR-0325311 and by a generous donation from British Telecom.

Satish Rao showed me that it is possible to teach algorithms to 130 people and have a lot of fun doing it; I thank him and my fellow TAs, Shyam Lakshmin and Eric Kuo, for making my first semester a wonderful introduction to Berkeley. The early stages of this work would not have been possible without the time afforded by an Intel Open Collaboration Research Fellowship. More recently, I am indebted to the National Science Foundation for the support provided by an NSF Graduate Research Fellowship.

I have been lucky to have an amazing group of colleagues and friends at Berkeley, too many to list all of them here. With the security group, the Berkeley Theory Group, and the Network Embedded Sensor Technologies Group, among others, there is never a boring moment. Alex Fabrikant and Alex Simma, my roommates, have been everything one could ask for and more.

I would not be half the researcher I am now without the experiences that happened prior to attending Berkeley. I am grateful to M.O. Rabin for all manner of research and commercial experiences, and for allowing me to see the way he works. Dennis Shasha showed me what a quick and omniheuristic mind can do, and his enthusiasm is infectious. No one could ask for more or better. Stuart Shieber helped me navigate through my undergraduate career and taught an excellent computational linguistics class.

Stuart Schecter introduced me to the economic approach I discuss in the Conclusions chapter. Rachel Greenstadt and Tony Vila collaborated on the previous analysis of web site privacy as a lemons market, which directly inspired the approach to economics of RFID killing. I thank jbash, Nikita Borisov, and several anonymous livejournal commenters for further comments on the idea.

The Free Haven Project gave me my first serious exposure to research in privacy-enhancing technologies. I am grateful to Roger Dingledine, Michael J. Freedman, David Hopwood, Nick Mathewson, Michael Mitzenmacher, Ron Rivest, Joseph Sokol-Margolis, and other Free Haven members and affiliates.

Finally, I thank Mom, Dad, and Susan. Without you I would never have made it this far.

Chapter 1

Introduction

In the near future, millions of people will interact directly with a Radio Frequency Identification (RFID) device. In RFID, a "tag" is applied to an item or a container. The tag can then be interrogated via radio waves by an RFID reader to return a small amount of information, such as a serial number. The technology promises benefits for supply chain management, item tracking, anti-counterfeiting, and other areas.

At the same time, RFID technology raises security and privacy issues. For example, applying RFID tags to individual items raises the possibility that the movement of these items can be tracked, or that individuals can be scanned to learn what items they carry. For another example, anti-counterfeiting uses of RFID rely on the tag being bound tightly to the item it authenticates. We explore these issues in the context of two real deployments, derive problems to solve, and give new techniques for solving these problems.

1.1 Motivation

Our motivation is that unless we study RFID architectures that can provide privacy and security now, we will find ourselves stuck with literally millions of legacy tags that provide no support for privacy and security. Our experience with the Internet teaches us that systems designed without adversaries in mind fall victim to ever more clever exploits.

Now is a good time to raise issues regarding security and privacy, because the use of RFID technology is growing. The most widely known RFID technology is the supply chain RFID tags used by Wal-Mart and the US Department of Defense. In 2003, Wal-Mart announced that its top 100 suppliers must implement RFID tags for tracking pallet shipments to warehouses by January 2005. The Department of Defense, in contrast, has used RFID tags for logistics management since the late 1990s, after bad experiences with logistics during operation Desert Storm. While both organizations currently use RFID for pallet and case level tagging primarily, the push in both Wal-Mart and the Department of De-

fense is towards "item-level tagging," in which each individual item is assigned its own RFID tag.

The vision of RFID tags for supply chain management started at the MIT Auto-ID Center. Originally started in 1999, the Center turned over its operations in 2003 to a commercial consortium, EPCglobal. The stated goal of EPCglobal is a tag aimed at the supply chain that costs 5 US cents; cheap enough to apply to nearly every item. The consortium also specifies an "electronic product code" for uniquely identifying every instance of every item ever tagged. Early work on privacy issues in this context was carried out by Weis, Sarma, Rivest, and Engels, who define the problem and set out solutions for RFID tags in the supply chain context [84].

Since the seminal work of Weis et al. [84], there has been a surge of interest in RFID privacy. Part of this interest has stemmed from greater public awareness of RFID technology. For example, Wal-Mart, Benetton, and Tesco Supermarkets have all been widely reported as conducting trials of RFID tags on individual items. The Metro Future Store in Germany, for another example, paired item-level tagging with an RFID shopper loyalty card. These reports raised significant privacy concerns and furthered interest in designing more secure and more private RFID architectures. In some cases, these reports led to boycotts of the organizations involved or demonstrations against the use of RFID technology. The key issues here are that RFID tags can be *read silently at a distance* and may *carry sensitive information*. In most cases, detailed information about the deployment was not available, leaving the public to assume the worst.

Despite such resistance, momentum is growing behind deployments of RFID technology. Many of these deployments are in a wider variety of contexts than the supply chain. Hitatchi, for example, has introduced a "mu-chip" RFID for bank notes, concert tickets, and item tracking. As we will see later in the thesis, library books and electronic passports are two other real world applications of the technology.

A further motivation of our work is to expand the scope of RFID privacy and security work beyond the supply chain. While the supply chain is undoubtedly one of the largest applications for RFID by tag volume, today's supply chain tags are typically applied to containers, not individual items. In contrast, other applications of RFID make use of item-level tagging today or in the near future, raising more serious privacy concerns. The technology assumptions in these deployments are also different, raising the possibility of different methods that may be too computationally expensive or otherwise ill-suited to the supply chain setting. For example, the most high-end devices, such as those used in electronic passports, have computational power comparable to a smart card, including the ability to perform public-key cryptography. Despite this computational power, there remain privacy and security problems in electronic passport deployments that require novel thinking and novel solutions. Simply having cryptography is not enough to automatically ensure privacy.

1.2 Learning From Deployments

Privacy is notoriously difficult to pin down. Before we can build better RFID architectures, however, we need to formulate clear technical problems. We must also show that solving these problems will make concrete improvements in the privacy and security offered by RFID technology.

To find these problems, we look at two real deployments of RFID. Our method is to evolve the "right" problems by looking at the security and privacy issues in deployments, then attempting to generalize. The advantage of looking at real deployments is that we can gain a lot of insight into the real impact of RFID technology and some insight into possible "unintended consequences" of implementation decisions.

The first deployment we consider is the application of RFID to libraries. In this deployment, RFID tags are used to manage inventory and check-out of library materials, such as books, videotapes, or DVDs. In Figure 1.2 we show a picture of a hand-held RFID reader scanning library books at a real RFID deployment, the Santa Clara City Library. Over 130 libraries in North America alone have implemented RFID tags for tracking items, and more are on the way.

Library reading habits are widely considered private information. We explain this in the context of the United States this by situating libraries in a broader framework of "information goods" and showing how information goods are protected by norms and law. Because an RFID tag is applied to each individual book, the ability to read RFID tags remotely raises privacy concerns. In Chapter 3 we investigate current library architectures and set out what information is and is not leaked by these architectures. In particular, we focus on *tracking* attacks, in which a book's movements are tracked, and *hotlisting*, in which an adversary can learn if an individual carries a specific sensitive book.

The second deployment we consider is that of electronic passports, or "e-passports." These promise to reduce passport fraud by tying a passport holder to a secure biometric embedded in an RFID on the passport. An early adopter in this area is Malaysia, which began issuing such passports in 1998 and now has over five million in circulation. We refer to Figure 1.2 for a description of how Malaysian e-passports are used in immigration at Kuala Lumpur airport and elsewhere. More recently, the United States, Australia, and Belgium, among other nations, have moved to adopt e-passports of their own. In Chapter 4 we outline the security and privacy issues in e-passports. We show that the use of biometrics means that the data contained on an e-passport is more sensitive than might be expected.

From these deployments, we extract two key problems for RFID architectures. First, we consider *private authentication*, or how an RFID tag and a reader can determine that they share a secret without leaking their identities. Traditional cryptographic methods are not sufficient in the RFID context, as most such methods assume that the identity of the parties is known. As a result, adopting these methods to the RFID setting often means that a reader must try all possible shared secrets before finding the "right" secret. This leads to the reader doing work linear in the number of possible tags, which is impractical as

Figure 1.1: Close-up of a library RFID reader and a picture of the reader in use. Pictures courtesy Santa Clara City library.

Figure 1.2: Instructions for using a Malaysian biometric e-passport. Over five million such e-passports have been issued to Malaysian citizens.

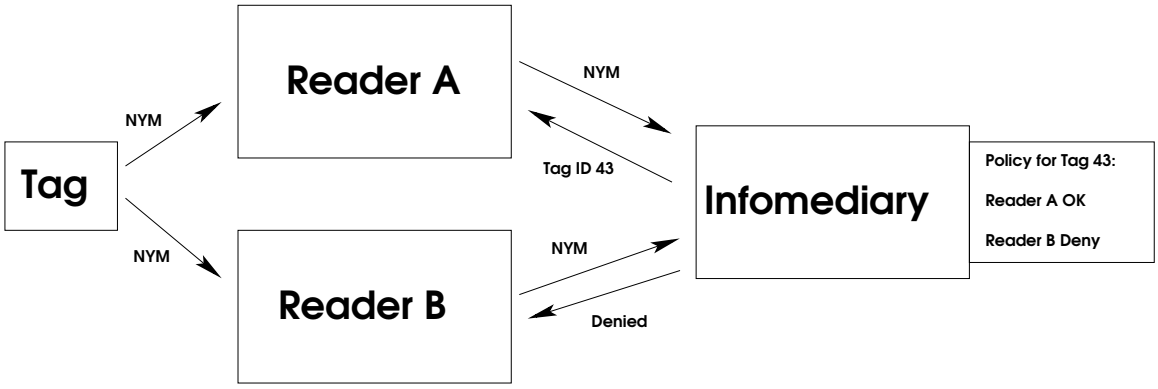

Figure 1.3: An infomediary interacts with two RFID readers, Reader A and Reader B. Both readers read the same Tag and receive a pseudonym, then query the Infomediary. The Infomediary checks the privacy policy for that Tag. As a result, the Infomediary tells Reader A the tag's ID and denies access to Reader B.

there may be thousands or millions of tags in a deployment. Nevertheless, we show that private authentication enables RFID security and privacy, because it prevents unauthorized readers from reading or modifying the state of an RFID tag.

Second, we discuss building an RFID *infomediary*, which is a trusted third party that controls access to tag data. We show how rewritable tags enable a basic kind of infomediary, but still fall victim to tracking and hotlisting attacks. To fix this problem, we turn to *RFID pseudonyms*, as introduced by Ohkubo, Kinoshita, and Suzuki [69]. Unlike the case of private authentication, an RFID pseudonym is simply a specially encrypted version of the RFID tag's ID. Each time the RFID tag is read, it responds with a different pseudonym. With the right secret key, the pseudonym can be decrypted to yield the correct ID. Without the key, however, no two sightings of the same tag can be linked together, because the tag returns different encrypted pseudonyms each time. Pseudonyms are appropriate for applications where the reader simply wants to know the ID of a tag and does not need to send any commands to the tag. Just as for private authentication, however, applying traditional cryptographic approaches to this problem results in linear work for the RFID reader.

We also introduce an extension to RFID pseudonyms called *delegation*. In delegation, we give a reader the ability to decrypt a limited number of pseudonyms. This means that the reader can identify an RFID tag for a limited time. After this time the tag's pseudonyms become unlinkable again. Delegation allows us to perform "revocation" without requiring the RFID tag to be explicitly notified. We also show how delegation enables *ownership transfer* of an RFID-tagged item between two mutually distrusting parties. The problem of ownership transfer is particularly important in supply chain applications of RFID, where many different entities may handle an object before it reaches its

final destination.

Another key motivation for delegation is that it limits the amount of trust that must be placed in any single RFID reader. For example, if a deployment has one million tags and ten thousand readers, with existing methods offline operation of a reader requires secrets sufficient to read any tag at any time. Therefore, the compromise of even a single reader may be catastrophic. With delegation, each reader can be given only the secrets it needs for a period of time. If a reader is compromised, the damage is limited to those secrets alone.

After setting out the problems, in Chapter 5 and Chapter 6 we give novel solutions to these problems for RFID tags that can compute pseudo-random functions. Recent advances in circuit design for AES by Feldhofer et al. suggest this is practical for a wide range of RFID devices [25]; we note that e-passport RFID tags *already* support 3DES and tags exist in the inventory space that can support 3DES (albeit at extra cost compared to non-cryptographically enhanced tags). A main feature of our solutions is that they require work only logarithmic in the number of RFID tags for the reader, in contrast with previous solutions that require linear work. We close by pointing out open problems and future directions for research in RFID privacy and security.

Beyond its intrinsic interest, RFID is an area that showcases the interaction between traditional computer security and physical security. Studying the interplay between the two in the RFID setting gives us insight into designing against threats spanning the two worlds. As computation becomes more embedded in the physical world, we expect this insight will apply to new and different areas.

1.3 Contributions and Statement on Joint Work

The original work in this thesis is joint work. I am indebted to my collaborators for their insight, good humor, and patience. The contributions of the thesis and specific collaborations are as follows:

- In Chapter 3 we analyze vulnerabilities of RFID deployments in libraries. We show that all current RFID tags used for libraries expose library patrons to tracking and hotlisting attacks, and we suggest concrete improvements. This work is joint with David Wagner and appears at the 2004 ACM Conference on Computer and Communications Security [60].

- Chapter 3 also includes a discussion of legal, normative, and policy questions concerning privacy in "information goods," of which library books are one example. This work is joint with Nathan Good, Jon Han, Elizabeth Miles, Deirdre Mulligan, Laura Quilter, Jennifer Urban, and David Wagner. A summary of the work appears as a short paper in the 2004 ACM Workshop on Privacy in the Electronic Society [32].

- In Chapter 4 we analyze security and privacy issues in "e-passports," which are passports that carry RFID devices. In particular, we consider the recent standard for e-passports published by the International Civil Aviation

Organization (ICAO) and deployments of ICAO e-passports by the United States and other countries. We show that the original deployment choices of the United States in rolling out ICAO e-passports put both the security and the privacy of U.S. e-passport holders at risk. Since the initial publication of our work, U.S. policy has changed to improve e-passport privacy protections; we briefly summarize the new policy. This work is joint with Ari Juels and David Wagner. It appears at the IEEE SecureComm 2005 Conference. Our work also appears on the eprint.iacr.org preprint archive [43] and as part of a formal comment to the U.S. Department of State filed by the Electronic Frontier Foundation.

- In Chapter 5 we give a symmetric key scheme for private authentication with $O(\log n)$ work, where n is the number of RFID tags in a deployment. This is the first symmetric key private authentication scheme with logarithmic work, answering an open question of Weis et al. [84]. This work is joint with David Wagner and appears in our publication at the 2004 ACM Conference on Computer and Communications Security [60].

- In Chapter 6, as part of formulating the problem of RFID pseudonyms, we discuss a method for preserving privacy by using RFID infomediaries. We show how an infomediary can enforce a privacy policy for reading tag data, given the ability to "recode," or re-write, an RFID tag. This work is joint with Ross Stapleton-Gray and David Wagner. Our work also appears in the book *RFID Applications, Security, and Privacy*, edited by Simson Garfinkel and Beth Rosenberg [59].

- In Chapter 6 we give a scheme for RFID pseudonyms that requires only $O(\log n)$ work for a trusted authority to map a pseudonym to the real tag identity. This is the first RFID pseudonym protocol with logarithmic work for the reader. The scheme also supports delegating the ability to map pseudonyms to real identities for a limited time only. We show how this delegation ability allows for ownership transfer of RFID tags between mutually distrusting parties. This work is joint with Andrea Soppera and David Wagner and appears at Selected Areas in Cryptography 2005 [58].

Chapter 2

RFID Technology

The term Radio Frequency Identification (RFID) has come to stand for a family of technologies that communicate data wirelessly from a small chip called a "tag" to a reading device. Many RFID devices, including all the devices considered in this thesis, are *passive*, that is, they carry no on-board source of power and are powered only through energy provided by the reader's signal. Beyond this, RFID technologies may have different characteristics depending on the intended application. Within an application, often several different types of RFID are available; as a result, it is hard to make sweeping statements about RFID technology as a whole. We now survey some of the different RFID types and applications they serve.

Perhaps the most well-known RFID application is the use of RFID tags to improve efficiency of the supply chain, as used by Wal-Mart, the U.S. Department of Defense, and others. Supply chain tags are designed to be as simple and cheap as possible, with minimal additional features beyond holding a single identifier. For example, the only privacy feature in the tags specified by the industry body EPCglobal is a special "kill" command that renders the tag permanently inoperative. The target cost for supply chain tags is US$0.05 in high volume, although today's tags cost closer to US$0.20. Supply chain tags typically operate at a frequency of 915 MHz and have an intended read range of three to five meters.

Two major types of RFID are used in the supply chain. First, the International Organization for Standardization (ISO) has published a standard, ISO 18000-6. Second, EPCglobal, a consortium of several RFID manufacturers together with the Uniform Code Council, publishes specifications for RFID tags. Recently, EPCglobal completed its "Gen II" standard; at this writing, compliant RFID tags are starting to appear from manufacturers such as Alien and Matrics.

We will consider two other applications besides the supply chain: libraries and electronic passports. Both of these applications use RFID tags that operate at a frequency of 13.56 MHz. We go into more detail about the type of RFID used in these applications in the following chapters. A table summarizing some

RFID Type	Frequency	Application	Range
EPC	915 MHz	Supply Chain	3m
EPC	13.56 MHz	Supply Chain	.5m
ISO 18000-6	915 MHz	Supply Chain	3m
ISO 18000-3	13.56 MHz	Inventory	.5m
ISO 15693	13.56 MHz	Inventory	.5m
ISO 14443	13.56 MHz	Smart Card	10cm

Figure 2.1: Representative RFID tag specifications, with the frequency of operation, intended application, and intended read range.

representative RFID types is found in Figure 2.

We write *intended* read range to mean the ranges achievable with vendor-standard readers. An adversary willing to build its own readers may achieve longer read ranges, especially if it is willing to violate applicable laws regulating radio devices. In fact, there are at least four different ranges of interest in the RFID setting.

- The *powering range*, the range at which an RFID device can be provided with enough energy to function properly, given a specified amount of energy available to the adversary's terminal.
- The *direct communication range*, the range at which an adversary can communicate with an already powered RFID device.
- The *reader-to-tag eavesdropping range*, the range at which an adversary can passively overhear communication from an RFID reader to a tag.
- The *tag-to-reader eavesdropping range*, the range at which an adversary can passively overhear communication from an RFID tag to a reader.

Different RFID technologies will have different values for each of these ranges. Therefore, each range should be considered in the context of a specific RFID technology. It is not clear, for example, that a long range for one technology implies a long range is achievable for a different technology.

At this writing, relatively little hard data is available for these ranges, but some information is known. Tests by the U.S. National Institute of Standards and Technology were widely reported as showing a 30 foot eavesdropping range on ISO 14443 devices [89], but technical data is not yet available. Flexilis Security demonstrated a 69 foot direct read range for EPC tags, but a technical writeup of the demonstration is not yet available [81].

Kfir and Wool use simulations to estimate the powering range for ISO 14443 13.56 MHz devices at roughly 50cm. They present a design for a device that powers the RFID to enable communication by a different antenna located further away. They also show how man in the middle attacks can take advantage of a long-range communications method to make it appear that an RFID tag is present to a reader, when in fact the tag is arbitraily far away [47]. Hancke

reports on a practical implementation of the relay portion of their attack [35]. Recently, Hancke, and independently Kirschenbaum and Wool reported on practical implementations of skimming for ISO 14443 devices; both report skimming ranges of roughly 25cm [36, 49]. Hancke also reports a 4 meter eavesdropping range, but speculates that this range may improve with more sophisticated signal processing techniques.

Chapter 3

Library RFID

3.1 Overview of Library RFID

Many libraries are starting to tag every item in their collections with radio frequency identification (RFID) tags, raising patron privacy concerns. Several libraries, such as the Santa Clara City Library in California, the University of Nevada, Las Vegas library, and the Eugene, Oregon public library have already tagged many of the books, tapes, CDs, or other items in their collections. In an item-level tagging regime, the ability to track tags raises the possibility of surveillance of library patrons and their reading habits. We investigate privacy risks in libraries' use of RFID technology and methods for minimizing such risks.

Most supply chain applications focus on tagging cases or pallets holding merchandise. A key question has been the feasibility, security, and privacy of *item-level tagging*, in which each individual item is given its own RFID tag. Many have raised concerns over the privacy implications of item-level tagging. Still, item-level RFID tagging is often considered to be 5 or more years in the future for retail RFID applications, due to the cost of tags, reader infrastructure, and uncertainty about near term applications. In contrast, library RFID applications require item-level tagging, because RFIDs are used to manage each

Tag Type	Example Library	Example Vendors
Checkpoint WORM	Santa Clara City	Checkpoint
Checkpoint writeable	None	Checkpoint
TAGSYS C220-FOLIO	U. Delaware	VTLS, TechLogic
ISO 15693/18000-3 MODE 1	National U. Singapore	3M, Bibliotheca, Libramation
ISO 18000-3 MODE 2	Not yet available	Coming soon
EPC Class 1 13.56MHz	Not for library	WalMart
EPC Class 0 915MHz	Not for library	WalMart
EPC Class 1 915MHz	Not for library	WalMart

Figure 3.1: Summary of the RFID technologies used in libraries.


Figure 3.2: On the left, a Checkpoint library RFID tag. On the right, an exit gate.

item in a library collection. Thus, library RFID applications may be the first major deployment of item-level tagging. This provides an opportunity to study the privacy implications of item-level RFID tagging in a concrete, real-world setting.

3.2 Library RFID Issues

3.2.1 Why Libraries Want RFID

RFID technology promises several important benefits for libraries. The major motivating factors may change from library to library, but several common themes have emerged.

First, RFID may reduce the incidence of repetitive stress injuries (RSI). A study at the San Francisco Public Library found that circulation desk work involves several motions likely to cause injury [52]. More importantly, library employees may face permanent disability from RSI injuries; at least one Berkeley Public library employee has been forced to retire permanently due to a RSI-related disability. While formal studies on the RSI benefit of RFID are not yet available, vendors claim significant reductions in the number of motions required for checkout.

Second, RFID promises to streamline mechanisms for patron *self-check*, allowing patrons to check out items without the help of library staff. Self-check machines that work with library magnetic strip security systems can only check out one book at a time. An RFID-based approach can, in theory, read and check out a stack of books placed on a self-check machine without the patron needing to handle each book individually. Several vendors also suggest the use of RFID-enabled patron cards, which offer the promise of completely hands-free checkout.

Third, librarians hope to make inventory management easier by using RFID tags. Hand-held RFID readers promise the ability to sweep a shelf once and obtain a list of all books on the shelf. Ease of inventory was one of the major considerations cited by the Vatican library [30]; because the library does not allow persons (except for the Pope) to check out materials , item checkout is not a concern. The Singapore national public library credits their RFID system

with reducing inventory time from a week to hours [88].

Finally, RFID acts as an enabler for automatic sorting on book check-in. Sorting systems, such as those from TechLogic or FlashScan, can read a bar code from the RFID and look up the shelf location from the bibliographic database. The book can then be sorted into a cart directed at the appropriate location.

The exact extent of benefits from RFID in the library setting is still being worked out. We stress that we do not attempt to evaluate all aspects of library RFID, only the privacy and security risks. At the same time, we believe it is important to understand the reasons why librarians may want this technology.

3.2.2 Information Goods in the United States

We can better understand the privacy issues in library materials by situating them in a broader context of "information goods." We use the term "information goods" to refer specifically to books, music, and film. Individuals have strong expectations of personal privacy in their choice of information goods that are reinforced in social norms, public policy, and law. We will consider these expectations as expressed in the norms and law surrounding information goods in the United States. In particular, we look at connections between privacy, the First Amendment, and information goods. These connections show that RFID privacy issues are of special interest in the library context.

Individuals' expectations of privacy when buying or borrowing books, music, and film stem from traditional ways to access those media with relative anonymity. Currently, individuals can purchase each of these goods with cash. In this case few means remain beyond the point of sale to discover the buyer's identity or to monitor what use the buyer makes of the work. Without identifying themselves, people can browse information on the Internet or in a library without checking materials out. Although library borrowing requires identification and registration, libraries have historically been staunch defenders of patron privacy, providing elaborate policy mechanisms to ensure records are kept secret from third parties when at all possible.

Traditionally, libraries have championed First Amendment rights to free speech and freedom of inquiry, viewing themselves as defenders of due process in the face of threats to free and anonymous inquiry. In an Interpretation of the Library Bill of Rights, the American Library Association instructs that "[i]n a library (physical or virtual), the right to privacy is the right to open inquiry without having the subject of one's interest examined or scrutinized by others." To this end, "[r]egardless of the technology used, everyone who collects or accesses personally identifiable information in any format has a legal and ethical obligation to protect confidentiality." In addition to this broad policy statement, libraries' privacy policies typically implement Fair Information Practices-they hold patrons' information for the shortest time possible, keep minimal patron records, and restrict access to patron borrowing records, even where not required by law to do so.

Established public policy reinforces these normative customs of relatively anonymous or confidential access to information. A patchwork of existing law

protects the unique privacy interests in information goods from a number of would-be intrusions in a range of settings. While the privacy protections surrounding information goods are neither complete nor uniform, taken as a whole they reflect a core policy principle: that our democratic society guarantees the right to freely speak and listen without the potential chilling effect of personal identification with the subject at hand.

The Constitution

The Constitution protects individual rights of free and private inquiry against government intrusion in the First Amendment's prohibition of any law that abrogates freedom of speech and the Fourth Amendment's limits on government surveillance. The Supreme Court has pronounced that the First Amendment protects the right to inquire freely as the logical corollary to freedom of speech: "The right of freedom of speech and press includes not only the right to utter or to print, but the right to distribute, the right to receive, the right to read . . . and freedom of inquiry. [78]"

Constitutional interests in open, surveillance-free use of information works limits the Government's power to discover the nature of its citizens' intellectual consumption. The Supreme Court provided an example of this boundary in *United States v. Rumely*, holding that Congress could not compel a wholesaler of politically controversial books to disclose sales records at a congressional hearing. The Constitution also limits the extent to which the Government can require citizens to disclose their choices in information access. In *Denver Area Educ. Telecommunications Consortium v. FCC*, the Supreme Court struck down a statutory provision requiring subscribers of indecent cable television programming to first register in order to receive those programs. The Court found that the requirement abridged the broadcaster's speech rights and represented an unconstitutional restriction on individuals' right to view privately. Further, the Court struck down a statute requiring individuals to identify themselves in order to receive controversial material, recognizing the burden such rules place on accessing information.

Protection of book sales records received keen public attention during the Clinton presidency in the Kramer Books-Monica Lewinsky matter. In 1998, Kramer sued to stop subpoenas from Independent Counsel Kenneth Starr for Monica Lewinksi's book purchase records. The store's owner stated that it is their company policy to "not turn over any information about [their] customers' purchases." Kramer was successful in blocking Starr's subpoenas. Many organizations, including the Association of American Publishers, the American Library Association, the Publishers Marketing Association, and the Recording Industry Association of America, lauded the action and announced formal support for bookstore defense of consumer privacy as a matter of policy.

Legislation

Congress and state legislatures have created a patchwork of industry-specific statutes that shield records of individual inquiry from disclosure to public and private parties alike. These laws are generally based on Fair Information Practices and limit the collection, retention, and disclosure of data. These laws are further evidence of the importance placed in the United States on privacy in information goods.

Several federal laws protect data collection and use relating to information goods. These statutory protections, while still patchwork and incomplete, are also typically stricter than for other goods. For example, at the federal level, the Cable Television Privacy Act of 1984 protects cable television subscribers from unfair data collection and use, and the Video Privacy Protection Act protects the video rental records from release without a court order.

More directly related to our purposes, similar laws protect library check-out and circulation information from release with without a court order in 48 states. The American Library Association notes "these laws mirror the express policy of the American Library Association. Eleven state constitutions guarantee a right of privacy or bar unreasonable intrusions into citizens' privacy. Forty-eight states protect the confidentiality of library users' records by law, and the attorneys general in the remaining two states have issued opinions recognizing the privacy of users' library records. [7]"

As an example of such laws, California state law provides:

> All registration and circulation records of any library which is in whole or in part supported by public funds shall remain confidential and shall not be disclosed to any person, local agency, or state agency except as follows: (a) By a person acting within the scope of his or her duties within the administration of the library. (b) By a person authorized, in writing, by the individual to whom the records pertain, to inspect the records. (c) By order of the appropriate superior court. As used in this section, the term "registration records" includes any information which a library requires a patron to provide in order to become eligible to borrow books and other materials, and the term "circulation records" includes any information which identifies the patrons borrowing particular books and other material [65].

The wording of laws in Alabama, Illinois, and New York is similar [64, 66, 67].

The recent USA PATRIOT Act modified the process by which a court order can be obtained for disclosure of library records. In particular, under some conditions, such an order can be issued as a National Security Letter; in this case, the details of the order are sealed and libraries are forbidden from revealing that they have been served with such an order. The move towards such secret disclosure of patron data has been near-universally resisted within the library community. We note, however, that even this weakening of the protection afforded library patrons still requires a hearing before an order is released.

Risks of Using RFID

Whatever the applicable law, the policy goal of protecting private inquiry may become more difficult as RFID is implemented. In the pre-RFID world, individuals can pay in cash leaving no records and can hide the fact of the purchase to limit third party knowledge of their reading habits. Moreover, before widespread retail and library use of RFID, providers of information goods, from wholesalers to retailers to renters and lenders, have control over their own records, and are often bound legally to demand due process of law before disclosing private records. Data holders can examine subpoenas for authenticity and cause, and challenge them in court before disclosing private information. In the RFID-enabled world, however, anyone with an RFID reader can potentially discover individuals' informational preferences without their permission. When information goods can be "interrogated" over the radio, revealing the goods' identity (or other information) to the immediate surroundings, no providers, librarians, the individual, sellers of goods, nor the law stand between people and those who seek to know what information they consume. In the next section, we make this observation more precise by describing current library RFID architectures and the exact information leaked by library RFID.

Also unanswered is the question of what will constitute intentional interception of radio transmissions or unlawful access to information stored on RFID tags for purposes of the Wiretap Act as amended by Electronic Communications Privacy Act (ECPA). Violation of these laws requires a reasonable expectation of privacy on the part of the speaker, and such expectation may not be reasonable when an individual broadcasts information by radio frequency [77]. Indeed, from 1986 to 1994 the law specifically exempted the radio portion of cordless phone conversations of phone conversations from protection because such transmissions were so easily intercepted. Though a subsequent amendment deleted the exception, courts have said that "broadcasting communications into the air by radio waves is more analogous to carrying on an oral communication in a loud voice or with a megaphone than it is to the privacy afforded by a wire." To realize its purpose, ECPA may require further amendment or interpretation by courts that extends its protections to the radio transmissions of RFID.

Finally, no library privacy law we are aware of specifically requires libraries to protect data stored on RFID tags from eavesdropping. In contrast, such laws explicitly prohibit libraries from turning over patron records without a court order. What is the legal status of a library that adopts an RFID system which leaks information to someone with an RFID reader? Does it depend on which information, specifically, is leaked by the system? Does it depend on the read range of the RFID technology in use? In our discussions with librarians in California, several have suggested that they do not believe California law compels them to protect data on an RFID tag. While bills relating to the use and deployment of RFID tags have been introduced in California, Utah, and other states, we are not aware of any that directly address the question of legal responsibilities of libraries with respect to data stored on RFID tags.

3.2.3 Current Library RFID Architectures

Once a library selects an RFID system, it is unlikely that anything short of catastrophe could motivate a library to spend the money and labor required to physically upgrade the tags. Currently, tags cost in the neighborhood of US$0.75 (exact prices are confidential and may vary widely) [15], while readers and other equipment may cost multiple thousands of dollars.

Libraries make use of a *bibliographic database* to track circulation information about items in a collection. Each book, upon being acquired by the library, is assigned a unique number, usually called a *bar code*. There is no fixed relation between author, title, and bar code. In today's library RFID deployments, tags are programmed with at least the bar code. In addition, some vendors suggest placing extra information on the tag, such as shelf location, last checked out date, author, and title [50].

Check-out occurs at either a circulation desk or a special "self-check" machine that allows patrons to check out their own books. In both cases, the RFID tag is read and the association between ID number and book looked up in the bibliographic database, and the status of the book is changed to "checked out" in the bibliographic database. Later, when the book is checked in, the tag is read again and the bibliographic database updated.

The RFID tag also acts as a security device. Special RFID *exit sensors* are placed at the exit of a library, just as most libraries today have exit sensors for magnetic strip anti-theft devices. When a patron exits, the sensors scan for books that have not been checked out.

Depending on the vendor, the security check is achieved in at least one of two ways. One method, used by 3M, VTLS, and Libramation among others, stores the status of the book on the tag; a specific bit, often called a "security bit," reveals whether the book is checked in or checked out. It is important to note that the security bit does not necessarily affect whether the tag can be read. The security bit must be correctly set at every check-in and check-out, or else false alarms may be triggered. A second method does not store the circulation status on the tag. Instead, the readers query the bibliographic database for the circulation status of the book as it passes through the exit sensors; this introduces issues of latency due to query time.

While this does not require writing the tag on each check-in and check-out, it introduces extra latency as the database is queried. To reduce latency, queries may be answered by a cache on the same LAN as the exit sensor. If a cache is used, then methods must be used to maintain cache consistency with the bibliographic database. In addition, such a cache creates a separate location with potentially sensitive data, i.e. the list of bar codes currently checked out of the library.

Many vendors also offer hand-held RFID readers. These readers are intended for use in finding lost books and taking inventory. Typically these consist of a Windows CE or PalmOS device attached to an RFID reader. The range of hand-held readers is on the order of 4 to 6 inches. Hand-held readers can be programmed for a variety of tasks, including mass inventory of books and

searching for a specific target RFID.

Privacy concerns in today's deployments have focused on the bibliographic database and short range of RFID readers. Without the bibliographic database, an adversary cannot directly map a bar code number to the title and author of a book, and so cannot immediately learn the reading habits of people scanned. Some library RFID proponents have argued that an adversary without the database and with only short-range readers poses little to no risk. In the next section, we show this is not the case.

3.2.4 Attacks on Current Architectures

Threat Model

In what follows, unless otherwise specified, we assume the adversary does not have access to the bibliographic database. We do assume that the adversary has access to an RFID reader and where indicated has the power to perform passive eavesdropping or even active attacks. Our attacks are summarized in Figure 3.3.

Detecting Tag Presence

Current RFID tags do not prevent an unauthorized reader from detecting a tag's presence. Detecting a new library RFID tag means someone or something moved a book into detection range, typically signalling the presence of a human being. Detecting human presence enables applications such as alarm systems, advertisements that respond when someone comes near, or real-time tracking of specific tags. The ability to detect a human presence might, in some cases, be considered an infringement on that person's privacy.

Static Tag Data and No Access Control

Referring to Figure 3.3, we see that none of today's library RFID tags employ read passwords or other read access control.[1] Because the identifier on the RFID tag never changes throughout its lifetime, the ability to read the tag at will creates several privacy risks.

First, the adversary may determine which library owns the book and infer the origin of the person carrying the book. In particular, bar codes for libraries with the Innovative bibliographic database have well-known, geographically unique prefixes. Vendors may also place library IDs on tags to prevent tags from one library from triggering readers at another. Learning origin data can be a privacy problem. For example, police at a roadblock may scan for patrons from specific city libraries in predominantly minority areas and search them more carefully; this would raise issues of racial profiling.

[1] Proprietary tag formats may raise the cost of building unauthorized readers, but such minor barriers will inevitably be defeated. As always, security through obscurity is not a good defense.

Second, any static identifier can be used both to *track* and *hotlist* books. In book tracking, the adversary tracks a book by correlating multiple observations of the book's RFID tag. The adversary may not necessarily know the title and author of the book unless the bibliographic database is available, but the static identifier can still be used to track the book's movements. Combined with video surveillance or other mechanisms, this may allow an adversary to link different people reading the same book. In this way, an adversary can begin profiling individuals' associations and make inferences about a particular individual's views, e.g. "this person checked out the same books as a known terrorist" or "mainly younger people have been seen with this book, so this person is young-thinking."

In hotlisting, the adversary has a "hotlist" of books in advance that it wishes to recognize. To determine the bar codes associated with these books, the adversary might visit the library to read tags present on these books. Later, when the adversary reads an RFID tag, it can determine whether that tag corresponds to a book on the hotlist. With current architectures, hotlisting is possible: each book has a single static identifier, and this identifier never changes over the book's lifetime.

Hotlisting is problematic because it allows an adversary to gather information about an individual's reading habits without a court order. For example, readers could be set up at security checkpoints in an airport, and individuals with hotlisted books set aside for special screening. For another example, readers could be set up at the entrance to stores and used to tailor patron experience or target marketing; these readers would look almost identical to the anti-theft gates used today.

Hotlisting is not a theoretical attack. We recall FBI warnings regarding almanacs as an indicator of terrorist activity [22]. We have also heard anecdotal reports from librarians that they refuse requests by law enforcement to track specific titles, and there are troubling historical precedents surrounding law enforcement and libraries. In the 1970s, the FBI Library Awareness Program routinely monitored the reading habits of "suspicious persons"; this was stopped only after public outcry and the passage of library privacy laws in many jurisdictions. Under the USA PATRIOT act, however, patron records may be accessed by order of the Foreign Intelligence Surveillance Court, or via a National Security Letter, as well as by a regular court order[23].

We have experimentally verified that tags can be read without access control at two library deployments of RFID. One library is the César Chávez branch of the Oakland Public Library, which uses ISO 15693 tags; the other is the University of Nevada, Las Vegas library, which uses Texas Instruments Tag-It! tags. We used a TAGSYS Medio S002 short-range reader for our experiments. We saw both deployments use static identifiers that enable tracking and hotlisting.

Collision-Avoidance IDs

Even if RFID tags were upgraded to control access to bar codes using read passwords or some other form of access control, many tags can still be identified uniquely by their radio behavior. In particular, many tags use a globally unique and static *collision ID* as part of their collision-avoidance protocol. This typically will allow unauthorized readers to determine the tag's identity merely through its collision-avoidance behavior. We give some concrete examples of this issue.

- In ISO 18000-3 MODE 1 tags, the current draft of the standard specifies that each tag will have a globally unique, 64-bit "MFR Tag ID." Further, tags are mandated to support an "Inventory" command that returns the MFR Tag ID as part of the response; no access control is in place for this command. Thus, an attacker with a reader could learn the tag's identity simply by asking for it.

  This ID is also used for the collision-avoidance protocol of MODE 1, which introduces a second way that the tag's identity can leak. The MODE 1 collision-avoidance protocol operates in two modes: slotted or non-slotted. In non-slotted mode, the reader broadcasts a message with a variable-length *mask*. All tags with least significant bits matching the mask respond, while others remain silent. To learn a tag's ID, an adversary need only make two mask queries per bit and see to which one the tag responds. By extending the mask by one bit each time, the adversary can learn a tag's collision ID in 64 queries. Because in the MODE 1 collision-avoidance protocol this ID is the same as the MFR Tag ID, this allows unique identification of the tag. In the slotted version of the MODE 1 protocol, time is divided into 16 slots based on the most significant bits of the ID, and the process is similar.

  EPC Class 1 13.56 MHz tags use their EPC identifier directly in a similar collision-avoidance protocol [18].

- ISO 18000-3 MODE 2 also specifies a 64-bit manufacturer ID. The ID is not used directly for collision avoidance. The collision avoidance protocol requires the generation of random numbers, however, and the standard specifies the use of "at least a 32-bit feedback shift register or equivalent." While it is not explicitly specified, we expect that each tag will have a globally unique seed in practice. In particular, we note that 32 bits of the 64 bit manufacturer ID are defined to be a globally unique "specific identifier"; it would be natural to use this specific identifier to seed a PRNG.

  If a 32-bit LFSR is used, then tags can be uniquely identified. Specifically, if as few as 64 outputs of the LFSR are observed in the collision-avoidance protocol, the entire state of the LFSR can be reconstructed using the Berlekamp-Massey algorithm and run backwards to obtain the

unique seed. In general, if a weak PRNG is used with the ISO 18000-3 MODE 2 protocol, tags can be identified.

- In EPC 915 MHz tags, there are three different modes for "singulation" or collision avoidance, one of which uses the globally unique Electronic Product Code (EPC) ID. The choice of modes is controlled by the reader. An adversarial reader can simply ask the tag to use its EPC ID; because there is no authentication of this command, the tag will obey.

As a consequence, any library system using one of these tags will be vulnerable to tracking and hotlisting of books and patrons. The collision-avoidance behavior is hard-coded at such a low layer of the tag that, no matter what higher layers do, privacy will be unachievable. This is unfortunate, because it means that much of today's RFID hardware is simply incompatible with privacy for library patrons. It is also dangerous, as vendors and libraries may implement privacy-enhancing methods that focus on tag data and then be unaware that tags are not in fact private.

Write Locks, Race Conditions, and Security Bit Denial of Service

In deployments with rewritable tags, some method must be used to prevent adversaries from writing to the tag. Otherwise, an adversary can commit acts of vandalism such as erasing tag data, switching two books' RFID data, or changing the security status of tags with "security bits." Unfortunately, vandalism is a real threat to libraries, especially from people who feel certain books should not be available; it would be naive to expect such people to ignore RFID-based vandalism for long.

Unfortunately, several current specifications have write protection architectures that are problematic in the library application. The EPC 13.56 MHz tag specification, as well as ISO 18000-3 MODE 1, include a "write" and a "lock" command, but no "unlock" command. In addition, write commands are not protected by password; this is consistent with a supply chain application that writes a unique serial number to a tag, then never needs to re-write the number. While the lock command is only an optional part of the ISO 18000-3 MODE 1 standard, it is supported by many tags, including the Phillips ICode tags purchased by the National University of Singapore to supplement its 3M library system [24]. In ISO 18000-3 MODE 2, locking is also irrevocable, but protected by a 48-bit password.

Once locked, a page of memory cannot be unlocked by any reader. A page containing a security bit needs to be unlocked when a book is checked in or out, or else the status of the bit can not be changed. An adversary can change the security bit to "not checked out" and then lock that page of memory. The resulting tag is then unusable, as the memory cannot be unlocked; physical replacement of the tag is required before the book can be checked out. We refer to irrevocable locking of the security bit as a *security bit denial of service*.

In addition to the issues with implementing security bits, there is a privacy concern as well. If there exists unlocked memory on the tag, an adversary can

write its own globally unique identifier and track tags based on this ID; the RF-DUMP software by Grunwald makes this a one-click operation [33]. This attack could bypass other mechanisms intended to prevent tracking or hotlisting of tags, such as rewriting tag IDs. Therefore, care should be taken to always lock all unused memory on writeable library RFID tags.

In our experiments with ISO 15693 tags in a real library deployment, we experimentally verified that none of the tag data blocks were locked. We also verified that tag blocks could be locked irrevocably on these tags, enabling security bit denial of service. We have since learned that the Phillips ICode tags used in this library have an extra security bit and so can be used safely in practice; this security bit, however, is not part of the ISO 15693 standard, and so readers must be specifically compatible with the Phillips extension. This makes such tags less desirable for a library that would like to avoid being locked in to a specific supplier.

TAGSYS C220 tags avoid security bit denial of service by having a special area of memory dedicated to the security bit built into the tag, separate from regular data storage. Checkpoint tags, in contrast, do not implement security bits, but rely on a database of checked-out books.

An alternative RFID architecture might implement separate "unlock," "write," and "lock" commands, either on a per tag or per data page basis. Such an architecture is suggested by Weis et al. in the context of "hash locks" [84]. Weis et al. note that session hijacking is possible in such an architecture. In such a system, it is also possible for an active adversary to bypass the write lock mechanism by racing a legitimate reader. After waiting for the legitimate reader to unlock the tag, the adversary can then send write commands which will be accepted by the tag.

In practice, tags may be left unlocked by accident if a tag is prematurely removed from a reader's field of control before the tag can be re-locked. We have anecdotal evidence that this occurs in self-check stations when patrons place a large stack of books on the machine, but remove them before all can be locked. In this case, the tag is vulnerable to malicious writes of all unlocked data.

In addition, several tag types support command sequences that force a tag to restart collision avoidance protocols. If a unlock-write-lock architecture is overlaid on these tags, special care must be taken that tags transition to the "locked" state on receipt of any such commands.

Tag Password Management

The ISO 18000 standard and EPC specifications only allow for static passwords sent in the clear from reader to tag. As noted, current deployments do not seem to use read passwords, but write passwords are employed. There are two natural approaches to password management: (1) use a single password per site; or, (2) endow each tag with its own unique password.

If a single password is used for all tags, then a compromise of any tag compromises the entire system. In deployments that use writable security bits, the write password is used on every self-checkout; in systems with read passwords,

exit sensors must use the read password every time a book leaves the library. In either case, passwords are available to a passive eavesdropper. Consequently, eavesdropping on a single communication reveals the password used by every tag in the system, a serious security failure. Once learned by a single adversary, a password can be posted on the Internet. Then, anyone with a reader can mount the attacks we have discussed.

If different passwords per tag are used, then some mechanism is required to allow the reader to determine which password should be used for which tag. Unfortunately, most obvious mechanisms for doing so, such as having a tag send an index into a table of shared secrets to the reader, provide tags with static, globally unique IDs. These globally unique IDs allow tracking and hotlisting of tags, which would defeat the entire purpose of read access control. Thus, privacy appears incompatible with prudent password management. We will return to this question in Chapter 5 and demonstrate a scheme that can reconcile these two demands.

Stealing Books With Aluminum Foil

Because detecting the RFID tag on exit is the primary security mechanism, blocking the tag signal allows an adversary to steal the book undetected. A blocked tag will pass by the exit sensors in a library without triggering any alarm. While a blocker tag could be used for this purpose, it would be easier and cheaper to use materials such as aluminum foil or mylar, which can absorb or diffuse an RFID signal [72]. As Boss notes, in a library with RFID, carrying common aluminum foil becomes evidence of intent to steal books [15]. We are certainly not the first to notice this issue in library RFID, and the severity of this risk is limited—tag security is primarily intended to keep honest people honest, not as a foolproof theft-prevention mechanism—but we note it for completeness.

No Forward Privacy

Current architectures for library RFID do not have *forward privacy*. If the adversary collects a database of tag readings and later obtains the bibliographic database, then all the title and author information of those readings is revealed. The adversary then learns everything about the reading habits of the people observed. The database could be revealed via a search warrant, but also by network intrusion, computer misconfiguration, throwing out backup tapes accidentally, or the work of an insider.

3.3 Improving Library RFID

3.3.1 Today's Tags

Unfortunately, as we have shown, many types of current tags can be uniquely identified by their collision-avoidance behavior. This identification is independent of any read access control on the tag data. Consequently, it appears to be

Tag Type	Read PW	Write PW	DoS	Priv. C.A.	Priv. Auth.
Checkpoint WORM	No	n/a	n/a	Unknown	No
Checkpoint writeable	No	Yes	n/a	Unknown	No
TAGSYS C220 FOLIO	No	Yes (32 bits)	Unknown	Unknown	No
ISO 15693/18000-3 MODE 1	No	No (Lock)	Yes	No	No
ISO 18000-3 MODE 2	Yes (48 bits)	Yes (48 bits)	Yes*	No*	No

Figure 3.3: Summary of attacks. The fourth column indicates whether the tag type is vulnerable to security bit denial of service; the fifth and sixth columns show whether the tag supports private collision-avoidance and private authentication protocols. Note that all but the ISO 18000-3 MODE 2 tag lack access control and hence are vulnerable to straightforward hotlisting and tracking attacks. ISO 18000-3 MODE 2 tags leak their identity through the collision-avoidance protocol (unless a crypto-strength PRNG is used), and are vulnerable to security bit DoS attacks if the password is known.

impossible to build privacy-preserving architectures for library RFID on many of today's tags.

3.3.2 Tags With Private Collision Avoidance

If we have a tag with private collision avoidance, then we have a hope for achieving a private library RFID architecture.

Random Transaction IDs on Rewritable Tags

Our first proposal is similar to the Anonymous ID scheme proposed by Ohkubo et al. [48]; we adapt it to the library setting. On each check-out, the reader picks a new random number r, reads the tag data D, and stores the pair (r, D) in a backend database. The RFID reader then erases D from the tag and writes r. On check-in, the library reader reads r, looks up the corresponding D, and writes D back to the tag. While tracking a book is still possible with this scheme, hotlisting is not. This scheme also offers a measure of forward privacy if the database securely deletes r after the book is checked in. Special care must be taken that the book's identifier has been written correctly, as RFIDs have difficulty writing at a distance. For example, the process may involve reading back and validating the new ID at check-in and check-out. In chapter 6 we discuss how to avoid the tracking risk by using RFID tags that support "RFID psedonyms."

Private Authentication

If only readers authorized by the library could read RFID tags, many of our attacks would fail. Therefore, a natural approach is to have some kind of password or other authentication protocol between library RFID tags and readers. Good security practice dictates that each tag have a distinct secret key, raising the issue of how a reader knows which secret to use when presented with a new tag. Trying each secret in turn will take too much communication to be

feasible. At the same time, most straightforward ways for accomplishing this goal provide unique identifiers for the tag, which defeats the purpose of read access control in the library RFID setting. This is the symmetric-key *private authentication* problem: how can two parties that share a secret authenticate each other without revealing their identities to an adversary? We will discuss this problem in more depth in chapter 5.

3.4 Related Work

In the retail RFID space, the EPCGlobal suite of RFID specifications mandates that tags support an irrevocable "kill" command. In the library setting, however, tags must be re-used to check in loaned items. Irrevocably killing a tag is not an option.

Juels, Rivest, and Szydlo propose a device called a "blocker tag" [44]. The blocker tag exploits the tree-walking collision-avoidance protocol of 915 MHz EPC tags to "block" readers attempting to read tags of a consumer. Because of bandwidth constraints, the 13.56 MHz tags used in library settings do not use tree-walking, so their scheme is not applicable; a new scheme would have to be designed. In addition, a blocker tag would enable stealing library books because it would prevent exit gates from scanning tags on books leaving the library.

Several activist groups have raised the issue of patron privacy for library RFID. The Electronic Frontier Foundation wrote a letter to the San Francisco Public Library raising several important policy questions surrounding library RFID [80]. A general "RFID Bill of Rights" was proposed by Garfinkel [28]; it proposes a right to notice that RFIDs are in use and a right to RFID alternatives.

Some vendors also have literature addressing the issue of library RFID and patron privacy. The 3M "eTattler" newsletter claims that the proprietary nature of 3M RFID tags and the low read range make privacy less of a concern [1]. The VTLS white paper on patron privacy cites low read range and also mentions that "encryption" can be used to protect tag data [17]. While library RFID read ranges may be low, they are still enough to provide for reading in doorways or other close spaces from vendor standard readers; adversaries willing to break the law and build more powerful readers may achieve greater range. Past experience also teaches us that it is dangerous to rely solely on security through obscurity and proprietary protocols.

Finally, the Berkeley Public Library has put together a series of "best practices" for library RFID [51]. These practices include limiting the data on the tag to a bar code only and prohibiting patrons from searching the bibliographic database by bar code. We have shown that privacy risks still exist even when data is limited to a bar code and the adversary does not have access to the bibliographic database, although in light of our results, the Berkeley practices seem to be the best possible with today's tags.

3.5 Summing Up Library RFID

Current library RFID tags do not prevent unauthorized reading of tag data. Therefore, information such as title, author, shelf location, patron information, or last checkin/checkout time should in no circumstance be stored on library RFID tags.

At the same time, both *tracking* and *hotlisting* are possible whenever a static identifier is used. Therefore, if a static identifier is in place on the RFID tag, it is imperative to prevent unauthorized tag reads. We stress that static identifiers may include collision IDs that are not protected by access control mechanisms intended to protect tag data. To avoid tracking tags by collision ID, some mechanism for private collision avoidance must be used.

Would these library RFID security and privacy problems automatically go away if tags advanced to the point where hash functions and symmetric encryption on tags became feasible? Our results on identification via collision avoidance, private authentication, and write locks show the answer is no. Careful design of the entire system is required to support privacy-enabled RFID applications.

What is more, libraries want RFID now. Over 130 libraries in North America alone have installed RFID technology, and more are considering it. Waiting for next generation tags that support cryptography may not be acceptable, especially at increased cost. Tag vendors, in addition, may be unwilling to introduce special modifications for what is a comparatively small market. Unfortunately, such changes will require time, effort, and money, and no current library RFID system supports them. There will be a substantial cost for privacy and security in the library RFID setting.

Is the cost of privacy and security "worth it?" Put another way, should a library refuse to buy RFID until systems are available that resist these attacks? We cannot dictate answers to this question. What we have done, instead, is provide the means for libraries and their communities to make an informed decision.

Chapter 4

Electronic Passports

4.1 Introduction

Major initiatives by the United States and other governments aim to fuse Radio Frequency Identification (RFID) and biometric technologies in a new generation of identity cards. Together, RFID and biometric technologies promise to reduce fraud, ease identity checks, and enhance security. At the same time, these technologies raise new risks. We explore the privacy and security implications of this worldwide experiment with a new type of authentication platform, with particular attention to its deployment in passports.

As part of its US-VISIT program, the United States government has mandated adoption by October 2006 of biometrically-enabled passports by the twenty-seven nations in its Visa-Waiver Program (VWP), among them Japan, most of the nations of Western Europe, and a handful of others [1]. By the end of 2006, all passports produced in the U.S. will carry biometric information. These passports are based on guidelines issued by the International Civil Aviation Organization (ICAO), a body run by the United Nations with a mandate for setting international passport standards [37]. The ICAO guidelines, detailed in ICAO Document 9303, call for incorporation of RFID chips, microchips capable of storing data and transmitting it in a wireless manner, into passports. Such chips will be present in initial deployments of biometrically enabled United States passports, and in the biometrically enabled passports of other nations as well. Next-generation passports, sometimes called *e-passports*, will be a prominent and widespread form of identification within a couple of years.

The ICAO standard specifies face recognition as the globally interoperable biometric for identity verification in travel documents. Thus e-passports will contain digitized photographic images of the faces of their bearers. The standard additionally specifies fingerprints and iris data as optional biometrics. The US-

[1] The deadline for adoption was originally October 2005, but was pushed back in response to concerns from other nations over delays in procurement and concerns from U.S. citizens over privacy. At this writing in January 2006, however, e-passport trials are beginning in several U.S. airports, including San Francisco International Airport.

VISIT program in fact requires visitors to provide two fingerprint images in addition to a headshot. The ICAO standard also envisions that e-passports will someday include a write capability for storage of information like digital visas.

Interestingly, one nation has already deployed e-passports in a project predating the ICAO standard. Since 1998, Malaysian passports have included a chip containing an image of a thumbprint of the passport holder; a second generation of e-passports rolled out in 2003 that contains extracted fingerprint information. When flying through Kuala Lumpur International Airport, a Malaysian citizen passes through an automated gate that reads the thumbprint from the chip and compares it to the thumb pressed on a scanner. Today, over 5,000,000 first generation and 125,000 second generation Malaysian e-passports are in circulation.

While e-passports are important in their own right, they also merit scrutiny as the harbinger of a wave of a fusion of RFID and biometrics in identity documents. Another next-generation ID card slated for deployment in the near future in the United States, for example, is the Personal Identity Verification (PIV) card. These cards will serve as ID badges and access cards for employees and contractors of the federal government in the United States. A standard for government ID cards (FIPS 201) is seeing rapid development by the National Institute of Standards and Technology (NIST). We expect PIV cards will include the same blend of technical mechanisms as e-passports: a combination of RFID and biometrics. The biometric of choice for PIV cards, however, will probably be fingerprint recognition. At the time of writing, the U.S. House of Representatives recently passed a bill called the Real ID Act; this seems a likely impetus for states to issue drivers' licenses containing biometrics, and probably RFID tags as well [56].

The goal of the ICAO and PIV projects is the same: strong authentication through documents that unequivocally identify their bearers. Data integrity and physical integrity are vital to the security of ID cards as authenticators. For authorities to establish the identity of John Doe with certainty, for example, Doe's passport must carry a photograph of irrefutable pedigree, with a guarantee that no substitution or tampering has taken place. Without this guarantee, passports can be forged, enabling unauthorized persons to enter a country.

Strong authentication requires more than resistance to tampering. *Data confidentiality*, i.e. secrecy of data stored on ID cards, is also critical. Protecting biometric and biographical data is essential to the value and integrity of an authentication system. In particular, data secrecy affords an important form of protection against forgery and spoofing attacks. Therefore protecting e-passport data against unauthorized access is a crucial part of the security of the entire system.

Confidentiality protection for stored data is important for other reasons as well. Both RFID and biometrics are highly privacy-sensitive technologies. Sensitive data, such as birthdate or nationality, are carried on passports. The privacy, physical safety, and psychological comfort of the users of next-generation passports and ID cards will depend on the quality of data-protection mechanisms and supporting architecture.

We identify security and privacy threats to e-passports generally, then evaluate emerging e-passport deployments with respect to these threats. We primarily analyze the ICAO standard and the specific deployment choices of early adopter nations. Where appropriate, we also discuss the Malaysian e-passport. Here is a summary of the major points we touch on:

1. **Clandestine scanning**: It is well known that RFID tags are subject to clandestine scanning. Baseline ICAO guidelines do not require authenticated or encrypted communications between passports and readers. Consequently, an unprotected e-passport chip is subject to short-range clandestine scanning (up to a few feet), with attendant leakage of sensitive personal information, including date of birth and place of birth.

2. **Clandestine tracking:** The standard for e-passport RFID chips (ISO 14443) stipulates the emission (without authentication) of a chip ID on protocol initiation. If this ID is different for every passport, it could enable tracking the movements of the passport holder by unauthorized parties. Tracking is possible even if the data on the chip cannot be read. We also show that the ICAO Active Authentication feature enables tracking when used with RSA or Rabin-Williams signatures.

3. **Skimming and cloning:** Baseline ICAO regulations require digital signatures on e-passport data. In principle, such signatures allow the reader to verify that the data came from the correct passport-issuing authority. [2] The digital signatures used in the baseline ICAO standard do not, however, bind the data to a particular passport or chip, so they offer no defense against passport cloning.

4. **Eavesdropping**: "Faraday cages" are an oft-discussed countermeasure to clandestine RFID scanning. In an e-passport, a Faraday cage would take the form of metallic material in the cover or holder that prevents the penetration of RFID signals. Passports equipped with Faraday cages would be subject to scanning only when expressly opened by their holders, and would seem on first blush to allay most privacy concerns.

   Faraday cages, however, do not prevent eavesdropping on legitimate passport-to-reader communications, like those taking place in airports. Eavesdropping is particularly problematic for three reasons.

   - *Function creep:* As envisioned in the ICAO guidelines, e-passports will likely see use not just in airports, but in new areas like e-commerce; thus eavesdropping will be possible in a variety of circumstances.

[2] Digital signatures and indeed, e-passports and secure ID cards in general, do not solve the problem of validating *enrollment*. Depending on how new users are validated, it may be possible to obtain an authentic ID by presenting inauthentic credentials or through circumventing issuing guidelines. Indeed, the 9/11 hijackers had perfectly authentic drivers' licenses. Digital signatures would merely have confirmed their validity. We do not treat the issue of enrollment here, but we note that it is pivotal in any ID system.

- *Feasibility:* Unlike clandestine scanning, eavesdropping may be feasible at a longer distance, given that eavesdropping is a passive operation [89].
- *Detection difficulty:* As it is purely passive and does not involve powered signal emission, eavesdropping is difficult to detect (unlike clandestine scanning).

5. **Biometric data-leakage**: Among other data, e-passports will include biometric images. In accordance with the ICAO standard, these will initially be digitized headshots, while thumbprints are used for the Malaysian e-passport. These images would not need to be secret to support authentication if the physical environment were strictly controlled. Existing and proposed deployments of e-passports, however, will facilitate automation, and therefore a weakening of human oversight. This makes secrecy of biometric data important.

6. **Cryptographic weaknesses:** ICAO guidelines include an optional mechanism called "Basic Access Control" for authenticating and encrypting passport-to-reader communications. The idea is that a reader initially makes optical contact with a passport, and scans the issue date, date of birth, and passport number to derive a cryptographic key K with two functions:

   - The key K allows the passport to establish that it is talking to a legitimate reader before releasing RFID tag information.
   - The key K is used to encrypt all data transmitted between the passport and the reader.[3]

   Once a reader knows the key K, however, there is no mechanism for revoking access. A passport holder traveling to a foreign country gives that country's Customs agents the ability to scan his or her passport in perpetuity. Further, we find that the cryptography relied upon by the ICAO standard itself has some minor flaws.

Related Work

Existing media stories, e.g., [74], have recognized the first three. The other issues, more technical in nature, have seen less exposition. Pattinson's whitepaper outlines the privacy problems with e-passports that may be readable by anyone and argues, as we do, for Basic Access Control [70]. Pattinson also points out the need for a direct link between optically scanned card data and secret keys embedded in an e-passport. He does not, however, consider the issue of biometric data leakage or the cryptographic issues we address. Karger

[3]The need for optical scanning of passports seems to negate the benefits of wireless communication conferred by RFID. Our supposition is that ICAO guidelines favor RFID chips over contact chips because wireless data transmission causes less wear and tear than physical contact.

and Kc report on work performed at IBM on e-passport privacy independently and in parallel with our analysis which agrees with our assessment of e-passport vulnerabilities [46].

Organization

In section 4.4 we give a detailed discussion of the data contained in e-passports deployments and the risks posed by data exposure. We focus on the ICAO standard and the choices of specific countries in implementing the standard, and also briefly describe the Malaysian program as an illustration of likely deployment features. We consider the cryptographic security measures of the ICAO standard in section 4.5, illuminating some potential weaknesses. In section 4.6, we sketch a few countermeasures to the security weaknesses we highlight. We discuss security issues likely to arise in future e-passport and ID-card systems in section 4.7. We conclude in section 4.8 with summary recommendations for improved e-passport deployment and with pointers to ID projects with similar underpinnings.

4.2 Terms: "Contactless Smart Cards" vs. "RFID"

Two different terms are used in the e-passport context. The first, "contactless smart card," emphasizes the fact that the devices used in e-passports can perform substantial computation, up to and including public-key cryptography. The other term used is "RFID," emphasizing the radio interface used in today's e-passports. Both, in this context, refer to the same technology: smart-card class processors communicating wirelessly with the ISO 14443 standard. We will continue to use the term RFID for consistency with the rest of the thesis, but we note the alternative term for completeness.

4.3 Biometrics in Brief

Biometric authentication is the verification of human identity through measurement of biological characteristics. It is the main mechanism by which human beings authenticate one another. When you recognize a friend by her voice or face, you are performing biometric authentication. Computers are able to perform very much the same process with increasing efficacy, and biometric authentication is gaining currency as a means for people to authenticate themselves to computing systems. We use the term *biometrics* in this paper to refer to human-to-computer authentication.

The range of practical biometrics for computing systems is different than for hu man-to-human authentication. Popular computer-oriented biometrics, for instance, include fingerprints, face recognition, and irises; these are the three biometrics favored for e-passport deployments.

Face recognition involves photographic imaging of the face; it is essentially the automated analog of the ordinary human process of face recognition. Fin-

gerprint recognition likewise relies on imaging and an automated process very loosely analogous to the fingerprint matching used in criminal investigations (but often based on a different class of fingerprint features). Fingerprint scanners can take on optical or silicon-sensor forms. Iris recognition also involves imaging. The iris is the colored annular portion of the eye around the pupil. Someone with "blue eyes," for instance, has blue irises. (The iris is not to be confused with the retina, an internal physiological structure.) Iris scanning in biometric systems takes place via non-invasive scanning with a high-precision camera. The device that captures user data in a biometric system is often called a *sensor*.

The process of biometric authentication is roughly similar in most systems. An authenticated user enrolls by presenting an initial, high-quality biometric image to the sensor. The system stores information extracted during enrollment in a data structure known as a *template*. The template serves as the reference for later authentication of the user. It may consist of an explicit image of the biometric, e.g, a fingerprint image, or of some derived information, such as the relative locations of special points in the fingerprint. To prove her identity during an authentication session, the user again presents the biometric to a sen sor. The verifying entity compares the freshly presented biometric information with that contained in the template for the user in a process generally called *matching*. The template and authentication image are deemed to match success fully only if they are sufficiently similar according to a predetermined—and often complicated and vendor-specific—metric.

While conceptually simple, the process of biometric authentication abounds with privacy and security complications. A key issue is *biometric authenticity*: How does the verifying entity know that the image presented for authentication is fresh and comes from a human being rather than a prosthetic or a digital image? The manufacturers of biometric sensors try to design them to re sist spoofing via prosthetics; the designers of biometric systems employ data se curity techniques to authenticate that the origin of biometric information is a trusted sensor. As we shall explain, however, the *secrecy* of templates is ultimately quite important and yet insufficiently assured in the baseline ICAO standard.

4.4 E-passport Threats

4.4.1 Data Leakage Threats

Without protective measures, e-passports are vulnerable to "skimming," meaning surreptitious reading of their contents. Even a short read range is enough for some threats. For example, a 3-foot read range makes it possible to install RFID readers in doorways; tags can then be read from anyone passing through the doorway. Such readers could be set up as part of security checkpoints at airports, sporting events, or concerts. Alternatively, clandestine readers could be placed in shops or entrances to buildings. Such readers might look much

like the anti-theft gates already used in thousands of retail stores. A network of such readers would enable fine-grained surveillance of e-passports.

Skimming is problematic because e-passports contain sensitive data. The ICAO standard for e-passports mandates that the RFID chip contain the passport holder's name, date of birth, and passport number. Actual deployments will include further biometric information, including at a minimum a photograph. Optional data items include such data as nationality, profession, and place of birth. First generation Malaysian e-passports contain an image of the passport holder's thumbprint as the biometric instead of a photograph. Second generation ICAO e-passports may also store a thumbprint template, as well as a small amount of writable memory for storing recent travel locations.

The RFID protocols executed by an e-passport may also leak information. For example, consider the ISO 14443 collision avoidance protocol, used by ICAO and Malaysian second generation passports. This protocol uses a special UID value to avoid link-layer collisions. If the UID value is fixed and different for each e-passport, then it acts as a static identifier for tracking the movement of e-passports. A static identifier also enables *hotlisting*. In hotlisting, the adversary builds a database matching identifiers to persons of interest. Later, when the identifier is seen again, the adversary knows the person without needing to directly access the e-passport contents. For example, a video camera plus an RFID reader might allow an adversary to link a face with a UID. Then subsequent detections of that UID can be linked with the face, even if no video camera is present.

Leakage of e-passport data thus presents two problems with consequences that extend beyond the e-passport system itself:

Identity Theft: A photograph, name, and birthday give a head start to a criminal seeking to commit identity theft. With the addition of a social security number, the criminal has most of the ingredients necessary to build a new identity or create a fake document.

Tracking and Hotlisting: Any static identifier allows for tracking g the movements of an RFID device. By itself, the movements of an individual may not be that interesting. When combined with other information, however, it can yield insight into a particular person's movements. Further, this information only becomes more useful over time, as additional information is aggregated.

Hotlisting is potentially more dangerous than simple tracking, because it explicitly allows targeting specific individuals. One unpleasant prospect is an "RFID-enabled bomb", a device that is keyed to explode at a particular individual's RFID reading [34]. In the case of e-passports, this might be keyed on the collision avoidance UID. Of course, one can detonate bombs remotely without the help of RFID, but RFID paves the way for unattended triggering and more comprehensive targeting.

4.4.2 The Biometric Threat

Leakage of the biometric data on an e-passport poses its own special risks: compromise of security both for the e-passport deployment itself, and potentially

for external biometric systems as well.

While designated as optional, biometric information will play a central role in e-passport systems. A facial image—a digitized headshot—is designated the "global interchange feature," meaning that it will serve as the international standard for biometric authentication. Indeed, ICAO guidelines describe it as the mandatory minimum for global interoperability [38]. Optional fields exist for iris and fingerprint data, which may be used at the issuing nation's discretion. We note that the US-VISIT program requires fingerprint biometrics from visitors; these fingerprints could be stored in the appropriate fields on an ICAO e-passport.

Advocates of biometric authentication systems sometimes suggest that secrecy is not important to the integrity of such systems. The fact that an image of John Doe's fingerprints is made public, for instance, does not preclude verification of Doe's identity: Comparison of the public image with the prints on her hands should still in principle establish her identity. This is all the more true when such comparison takes place in a secure environment like an airport, where physical spoofing might seem difficult to achieve.

At first glance, secrecy would seem particularly superfluous in the US-VISIT initiative and first deployments of ICAO passports. The globally interoperable biometric, as mentioned above, is face recognition. Thus the biometric image stored in passports will be headshots, which is in some sense public information to begin with.

Data secrecy in biometric systems, however, is a subtle issue. Two trends erode security in the face of public disclosure of biometric data:

1. *Offline vs. Online Trials:* Possession of the biometric template and the algorithm used to perform identification allows the adversary to mount an attack on the system in the safety of his or her home. For example, the adversary may create a prosthetic fingerprint or face and check it multiple times against the template until a match is found. Then, the adversary has a high degree of confidence that the prosthetic will pass the real inspection. In contrast, without the template, the adversary must work from an independently gathered sample of the biometric or directly with the system to be fooled. This is similar to the distinction made in password-based systems between online and offline attack.

2. *Automation:* Because biometric authentication is an automated process, it leads naturally to the relaxation of human oversight, and even to self-service application. This is already the case with e-passports. At Kuala Lumpur International Airport, Malaysian citizens present their e-passports to an "AutoGate" and authenticate themselves via a fingerprint scanner, without any direct human contact. If the fingerprint matches the e-passport data, the gate opens and the e-passport holder continues to his or her flight [45]. Australia plans to introduce similar "SmartGate" technology with face recognition in conjunction with its e-passport deployment. These deployments are instructive, because they tell us what

airport procedures might look like in a world where e-passports are ubiquitous.

The pressures of passenger convenience and airport staff costs are likely to reinforce this trend towards unattended use of biometrics. The result will be diminished human oversight of passenger authentication and greater opportunities for spoofing of biometric authentication systems.

3. *Spillover:* As biometrics serve to authenticate users in multiple contexts, compromise of data in one system will threaten the integrity of other, unrelated ones. In particular, biometric authentication is gaining in popularity as a tool for local authentication to computing devices and remote authentication to networks. For example, Microsoft is initiating support for optical fingerprint scanning devices in 2005 [62]. Even if the secrecy of John Doe's fingerprint image is relatively unimportant at a supervised immigration station in an airport, it may be of critical importance to the security of his home PC or corporate network if they also rely on biometrics for authentication, as an attacker able to simulate Doe's finger in these settings may do so in the absence of human oversight. (An unclassified State Department whitepaper recognizes the need to protect the privacy of iris and fingerprint data, but does not explain why [79].)

   Also, multiple enrollments of the same biometric can cause subtle security problems, even if none of the biometric data is "compromised." Recently, Barral, Coron, and Naccache proposed a technique for "externalized fingerprint matching" [11] [90], also a research prototype from GemPlus under the name BioEasy. The goal is to enable storing a fingerprint template on a low-cost chip, without requiring the overhead of traditional cryptography. In their scheme, a chip stores a fingerprint template $f(D)$ of a fingerprint D together with a set of randomly chosen fingerprint minutae r. When queried, the chip returns $t := f(D) \cup r$ and challenges the reader to determine which minutae belong to $f(D)$ and which belong to r. The authors argue that even if an adversary queries the chip remotely and learns t, recovering the template $f(D)$ without access to the fingerprint D is difficult because of the additional minutae r.

   If the same user enrolls in two different organizations A and B with the same finger, however, these organizations will give the user cards with $t_A = f(D) \cup r_A$ and $t_B = f(D) \cup r_B$ (we assume that the template algorithm can tolerate some fuzziness in the fingerprint reading and obtain the same or very similar f(D)). If the adversary scans the user, then it will learn both t_A and t_B. Then the adversary can compute $t_A \cap t_B = f(D) \cup (r_A \cap r_B)$. If r_A and r_B were chosen independently, we expect their intersection to be small, so the adversary can gain an advantage at determining the fingerprint template not envisioned in the original design of the system. This vulnerability illustrates the issues that could arise when fingerprints are used both for e-passports and for other forms of identification. The designers, in a patent application on the technology,

Type	Feature Name	Purpose
Mandatory	Passive Authentication	Prevent data modification
	Biometric: Photo	Identify passport holder
Optional	Active Authentication	Anti-cloning
	Basic Access Control	Data confidentiality
	Biometric: Fingerprint	Identify passport holder

Figure 4.1: Summary of ICAO security features.

suggest to use a second finger as the source of false minutae; this avoids the attack we have described but demonstrates the need for careful design in a world with e-passports [12].

These risks apply even to passport photos. While John Doe's face is a feature of public record, his passport photo is not. Passport photos have two special properties:

1. *Image quality:* Doe's passport photo is likely to be of a higher quality than the image of Doe's face that an attacker can obtain in casual circumstances. Passport photos are taken under rigorously stipulated conditions. One example is particularly illuminating with respect to these conditions: To comply with the technical requirements of facial recognition, applicants for U.K. passports may not smile for their photos [13].

2. *Disclosure may enable forgery:* Passport photos are the target authenticator: they are the reference point for an attacker aiming to spoof a facial recognition system. Forgery of a face in a biometric authentication systems may seem implausible, but Adler shows that holding up a photo is sufficient to spoof some face-recognition systems [3]. As noted above, knowledge of the passport photo allows an adversary to mount an offline attack on the biometric system.

Going further, iris scans and fingerprints are secondary biometrics specified in the ICAO document, and fingerprints are the primary biometric for Malaysian e-passports. In unattended settings, spoofing these biometrics is also possible given enough preparation time. For example, Matsumoto showed how several fingerprint recognition systems could be fooled when presented with gelatin "fingers" inscribed with ridges created from pictures of fingerprints [55].

4.5 Cryptography in E-passports

4.5.1 The ICAO Specification

As we have explained, the ICAO guidelines specify a large range of mandatory and optional data elements. To ensure the authenticity and privacy of this data, the guidelines include an array of cryptographic measures, discussed next.

The ICAO standard specifies one *mandatory* cryptographic feature for e-passports [37, 38]:

Passive authentication: The data stored on a e-passport will be signed by the issuing nation [38]. Permitted signature algorithms include RSA, DSA and ECDSA. As noted in the ICAO guidelines, passive authentication demonstrates only that the data is authentic. It does *not* prove that the container for the data, namely the e-passport, is authentic.

The ICAO guidelines additionally specify two *optional* cryptographic features for improved security in e-passports.

Basic Access Control and Secure Messaging: To ensure that tag data can be read only by authorized RFID readers, Basic Access Control stores a pair of secret cryptographic keys $(K_{\text{ENC}}, K_{\text{MAC}})$ in the passport chip. When a reader attempts to scan the passport, it engages in a challenge-response protocol that proves knowledge of the pair of keys and derives a session key. If authentication is successful, the passport releases its data contents; otherwise, the reader is deemed unauthorized and the passport refuses read access. The keys K_{ENC} and K_{MAC} derive from optically scannable data printed on the passport, namely:

- The passport number, typically a nine-character value;
- The date of birth of the bearer;
- The date of expiration of the passport; and,
- Three check digits, one for each of the three preceding values.

E-passports use the ISO 11770-2 Key Establishment Mechanism 6:

Reader		Tag
	$\xrightarrow{\text{Get challenge}}$	
		$r_T \in_R \{0,1\}^{64}$
	$\xleftarrow{r_T}$	
$r_R, k_R \in_R \{0,1\}^{64}$		
$S_R := r_R \|\| r_T \|\| k_R$		
$C_R := E_{K_{\text{ENC}}}(S_R)$		
$M_R := \text{MAC}_{K_{\text{MAC}}}(E_R)$		
	$\xrightarrow{C_R \|\| M_R}$	
		$k_T \in_R \{0,1\}^{64}$
		$S_T := r_T \|\| r_R \|\| k_T$
		$C_T := E_{K_{\text{ENC}}}(S_T)$
		$M_T := \text{MAC}_{K_{\text{MAC}}}(E_T)$
	$\xleftarrow{C_T \|\| M_T}$	

Here E is two-key triple-DES in CBC mode with an all-0 IV, and M is the ANSI "retail MAC" [39]. In this protocol, the Tag first checks the message authentication code (MAC) M_R and then decrypts the value C_R. The Tag then checks that the r_T in the decrypted value matches the r_T which it previously sent. If either check fails,the Tag aborts.

Similarly, when the Reader receives the messages C_T and M_T, it first checks the MAC M_T and then decrypts C_T. The Reader then checks that the correct r_R appears in the decryption of C_T. If either check fails, the Reader aborts. Otherwise, the Reader and Tag proceed to derive a shared session key from the "key seed" $k_R \oplus k_T$, by using the key derivation mechanism in Section E.1 of the ICAO PKI report [38].

The intent of Basic Access Control is clearly spelled out in the ICAO report: the Basic Access Control keys, and hence the ability to read the passport contents, should be available *only* when a passport holder intends to show his or her passport. Unfortunately, the scheme falls short of this goal in two ways.

First, the entropy of the keys is too small. The ICAO PKI Technical Report warns that the entropy of the key is at most 56 bits. The ICAO report acknowledges that some of these bits may be guessable in some circumstances. We believe that the key length is in fact slightly shorter for a general population. We estimate that the birth date yields about 14 bits of entropy and the expiration date, which has a 10-year maximum period, yields roughly 11 bits of

Country	RFID Type	Deployment	Security	Biometric
Malaysia Gen1	non-standard	1998	Passive Authentication + Unknown	Fingerprint
Malaysia Gen2	14443	2003	Passive Authentication + Unknown	Fingerprint
Belgium	14443	2004	Unknown	Photo
U.S.	14443	2006	Passive, Active Authentication, BAC	Photo
Australia	14443	2005	Unknown	Photo
Netherlands	14443	2005	Unknown	Photo

Figure 4.2: Current and near-future e-passport deployments. The Belgium, U.S., Australia, and Netherlands deployments follow the ICAO standard, while Malaysia's deployment predates the standard. The chart shows the type of RFID technology, estimated time of first deployment, security features employed, and type of biometric used. "Unknown" indicates a lack of reliable public information. "BAC" stands for Basic Access Control.

entropy. The remaining entropy depends on the passport number scheme of the issuing nation. For concreteness, we discuss the passport number scheme of the United States [4].

United States passports issued since 1981 have 9-digit passport numbers. The first two digits encode one of fifteen passport issuing offices, such as "10" for Boston or "03" for Los Angeles. The remaining seven digits are assigned arbitrarily. Probably some two-digit leading codes are more likely than others, as some offices presumably issue more passports than others, but we will conservatively ignore this effect. Given fifteen passport issuing agencies currently in the United States, U.S. passport numbers have at most $\lg(15 \times 10^7) \approx 27$ bits of entropy. This means Basic Access Control keys have a total of about 52 bits of entropy. Other nations may have more or less entropy in passport number assignment; for example, Riscure estimates that Dutch e-passport keys have only 33 bits of entropy [73].

Furthermore, the passport number is not typically considered a secret. Entities such as cruise ships, travel agents, airlines, and many others will see the number and may include it on paper documents.

Second, a single fixed key is used for the lifetime of the e-passport. As a consequence, it is impossible to revoke a reader's access to the e-passport once it has been read. If a passport holder visits a foreign nation, he or she must give that nation's border control the key for Basic Access Control. Because the key never changes, this enables that nation to read the e-passport in perpetuity. This capability may be misused in the future, or databases of keys may be inadvertently compromised.

Despite its shortcomings, Basic Access Control is much better than no encryption at all. Still, the United States originally elected not to include Basic Access Control in its e-passport deployment. Concern over e-passport privacy eventually resulted in the U.S. State Department delaying the rollout of e-passports to allow for Basic Access Control to be included.

"Active Authentication": The ICAO spec urges use of another, optional

security feature called "Active Authentication." While Basic Access Control is a confidentiality feature, Active Authentication is an anti-cloning feature. It does not prevent unauthorized parties from reading e-passport contents. Active Authentication relies on public-key cryptography. It works by having the e-passport prove possession of a private key. The corresponding public key is stored as part of the signed data on the passport. The ICAO guidelines are somewhat ambiguous, but appear to specify an integer factorization based signature such as RSA or Rabin-Williams. To authenticate, the passport receives an 8-byte challenge from the reader. It digitally signs this value using its private key, and returns the result. The reader can verify the correctness of the response against the public key for the passport. The ICAO guidelines specify use of the ISO/IEC 7816 Internal Authenticate mechanism, with ISO 9796-2 Signature Scheme 1 padding for the underlying signature:

Reader		Tag
$r_R \in_R \{0,1\}^{64}$		
	$\xrightarrow{r_R}$	
		$M_1 \in_R \{0,1\}^{64}$
		$X := M_1 \| r_R$
	$\xleftarrow{\mathrm{Sig}_{SK}(X)}$	

Here $\mathrm{Sig}_{SK}(X)$ is an RSA or Rabin-Williams signature with 9796-2 padding signed with the secret key SK of the e-passport. Notice that X contains both a random nonce generated by the Tag and a challenge from the reader; we speculate that this may be intended to counteract padding attacks such as those of Coron, Naccache, and Stern [19]. The 9796-2 padding itself makes use of a hash function, which may be SHA-1 or another hash function; the ICAO standard does not restrict the choice of hash. The signature can then be verified with the public key supposedly associated with the passport. If the signature verifies, the Reader gains some confidence that the passport presented is the contained which is supposed to hold the presented biometric data. The U.S. Concept of Operations for e-passports further specifies in Section C.2.7.2.2 a security policy that e-passport chips must support, namely that data cannot be overwritten on the chip after personalization [20]. Signing the chip's public key is a statement that the chip with the corresponding secret key is trusted to implement the security policy.

The public key used for Active Authentication must be tied to the specific e-passport and biometric data presented. Otherwise a man-in-the-middle attack is possible in which one passport is presented, but a different passport is used as an oracle to answer Active Authentication queries. The ICAO specification recognizes this threat, and as a result mandates that Active Authentication occur in conjunction with an optical scan by the reader of the machine-readable zone of the e-passport. As a result, every reader capable of Active Authentication and compliant with the ICAO specification also has the hardware capability

necessary for Basic Access Control. Deployments which neglect this part of the specification open themselves to a risk of cloned e-passports.

Active Authentication also raises subtle issues concerning its interaction with Basic Access Control and privacy. The certificate required for verifying Active Authentication also contains enough information to derive a key for Basic Access Control; as a result the certificate must be kept secret. In addition, when Active Authentication is used with RSA or Rabin-Williams signatures, responses with different moduli, and hence from different e-passports, can be distinguished. As a result, Active Authentication enables tracking and hotlisting attacks even if Basic Access Control is in use. We recommend that Active Authentication be carried out only over a secure session after Basic Access Control has been employed and session keys derived. Because Active Authentication requires an optical scan of the e-passport, just as Basic Access Control does, we do not believe this presents more of an operational burden than the existing specification.

4.5.2 Cryptographic Measures in Planned Deployments

At this point, more information is publicly available for the United States deployment of ICAO e-passports than any other of which we are aware. An unclassified State Department memo obtained by the ACLU describes elements of the U.S. PKI architecture as envisioned in 2003 [79]. A Federal Register notice dated 18 February 2005 provides a number of details on U.S. e-passport plans [68]. Appendix D of the State Department Concept of Operations document specifies that readers should support Active Authentication, leaving open the possibility of its future deployment in U.S. and foreign e-passports [20]. The original Federal Register notice, however, stated that U.S. passports would not implement Basic Access Control. The Federal notice offered three reasons for the decision not to implement Basic Access Control: (1) The data stored in the chip are identical to those printed in the passport; (2) Encrypted data would slow entry processing time[4]; and (3) Encryption would impose more difficult technical coordination requirements among nations implementing the e-passport system. Further, this notice intimated that e-passports will carry Faraday cages and that e-passport readers will be shielded to prevent eavesdropping.

Our analysis suggests this reasoning was flawed. Active Authentication requires an optical scan of a passport to provide the claimed anti-cloning benefit. This is why the ICAO spec mandates readers supporting Active Authentication be able to optically scan e-passports; this optical scan capability is also sufficient for Basic Access Control. Reason (3) is also flawed: because all the data required to derive keys for Basic Access Control is present on the data page of the e-passport, no coordination among nations is required. Coordination among vendors is required for interoperability of e-passports and readers, but such coordination is already required for e-passports without Basic Access Control. Finally, as we have argued, Faraday cages are not sufficient to protect

[4]Presumably this refers to the requirement for optical scanning in association with Basic Access Control.

against unauthorized eavesdropping, and so they do not rule out the attacks on security and privacy we have outlined.

In fact, our analysis shows that the original deployment choices of the United States put e-passport holders at risk for tracking, hotlisting, and biometric leakage. The lack of Basic Access Control means that any ISO 14443 compliant reader can easily read data from an e-passport, leading directly to these attacks. We are also concerned that a push towards automatic remote reading of e-passports may lead the U.S. to neglect optical scanning of e-passports, thereby weakening the anti-cloning protections of Active Authentication.

Since the original publication of our work in April 2005, however, the U.S. State Department has reversed itself and indicated that U.S. e-passports will in fact employ Basic Access Control. Frank Moss, the State Department official in charge of e-passport implementation, stated that the change was due to a realization the e-passports could be read at much further distances than previously thought. In addition, initial request for comments on U.S. e-passport policy gathered over 2400 responses (including a draft of this work); 98.5 percent of these responses were against deploying e-passports, most citing privacy concerns.

As it pre-dates the ICAO standard, the Malaysian identity card/passport is not compliant with that standard. Published information suggests that it employs digital signatures ("passive authentication") [21]. There appears to be no reliable public information on other security mechanisms, although the US patent filed on the technology suggests a "proprietary and secret" encryption algorithm is used for mutual authentication between e-passport and reader [87]. Belgium began issuing e-passports to citizens in November 2004, Australia, and the Netherlands expect large-scale issuing by the end of 2005, while the United States has delayed until 2006. For the ICAO e-passport deployments, the specific choices of each country as to which security features to include or not include makes a major difference in the level of security and privacy protections available. We summarize the known deployments, both current and impending shortly, in Figure 4.2.

Other nations may or may not meet the United States mandate for deployment in the next few years. Indeed, the reason that the United States favored a minimal set of security features appears to stem from problems with basic operation and compatibility in the emerging international infrastructure [86].

4.6 Strengthening Today's E-passports

4.6.1 Faraday Cages

One of the simplest measures for preventing unauthorized reading of e-passports is to add RF blocking material to the cover of an e-passport. Materials such as aluminum fiber are opaque to radio waves and could be used to create a Faraday cage, which prevents reading the RFID device inside the e-passport. Before such a passport could be read, therefore, it would have to be physically

opened.

The ICAO considered Faraday cages for e-passports, as shown in a discussion of "physical measures" in Section 2.4 of [38]. Because Faraday cages do not prevent eavesdropping on legitimate conversations between readers and tags, however, Faraday cages were deprecated in favor of Basic Access Control.

While a Faraday cage does not prevent an eavesdropper from snooping on a legitimate reading, it is a simple and effective method for reducing the opportunity for unauthorized reading of the passport at times when the holder does not expect it. Recently, the U.S. State Department indicated that U.S. e-passports may include metallized covers, following discussion of privacy risks by the ACLU and other groups.

The research community has proposed a number of tools for protecting RFID privacy, including "Blocker Tags" [44] and "Antenna Energy Analysis" [26]. While either of these mechanisms would be helpful, in the special context of e-passports they would be no more practical or protective than a Faraday cage, given that passive eavesdropping during legitimate read sessions is likely to constitute perhaps the major vulnerability to data leakage.

4.6.2 Larger Secrets for Basic Access Control

As we have discussed, the long-term keys for Basic Access Control have roughly 52 bits of entropy, which is too low to resist a brute-force attack. A simple countermeasure here would be to add a 128-bit secret, unique to each passport, to the key derivation algorithm. The secret would be printed, together with other passport information, on the passport. Such a secret could take the form of a larger passport ID number or a separate field on an e-passport. To aid mechanical reading, the secret might be represented as a two-dimensional bar code or written in an OCR font to the Machine Readable Zone (MRZ) of each passport.

4.6.3 Private Collision Avoidance

Even if a larger passport secret is used as part of key derivation, the collision avoidance protocol in ISO 14443 uses a UID as part of its collision avoidance protocol. Care must be taken that the UID is different on each reading and that UIDs are unlinkable across sessions. One simple countermeasure is to pick a new random identifier on every tag read. In general, e-passports and other IDs should use *private collision avoidance* protocols. Avoine analyzes several existing protocols and proposes methods for converting them into private protocols [9].

4.6.4 Beyond Optically Readable Keys

The ICAO Basic Access Control mechanism takes advantage of the fact that passports carry optically readable information as well as biometric data. In the passport context, the ICAO approach neatly ties together physical presence

and the ability to read biometric data. In general, however, we cannot count on this kind of tight coupling for next-generation ID cards. Furthermore, the use of a static, optically readable key leads to readers that must be trusted in perpetuity when all that is desired is to allow a single passport read. Therefore an important problem is to create a keying mechanism that limits a reader's power to reuse secret keys and a matching authorization infrastructure for e-passport readers.

Before we can move beyond optically readable keys, a key management problem reveals itself. Which key should an authorized party use to authenticate with a e-passport? The e-passport dare not reveal its identity to an untrusted reader, but at the same time the reader does not know which key to use.

An earlier version of our analysis suggested using the JFKr authenticated Diffie-Hellman key agreement protocol of Aiello et al. for this problem [5]. We also highlighted reader revocation as an open issue in e-passports. We have since learned that the German government has proposed a Diffie-Hellman based protocol for "Extended Access Control" in the ICAO specification [16].

Reader revocation in the German proposal is accomplished by time-expiring certificates issued to readers combined with a time-stamping service run by each nation. On each interaction with a legitimate reader, the reader provides the passport with the most recent known timestamp from that passport's nation. While this raises a denial-of-service risk if a nation ever signs a timestamp far in the future, it fits with the constraints imposed by a mostly-offline reader architecture. In particular, border control readers in southeastern Europe may be offline for weeks or months at a time [83].

4.7 Future Issues in E-passports

4.7.1 Visas and Writable E-passports

Once basic e-passports become accepted, there will be a push for e-passports that support visas and other endorsements. (We note that the presently proposed approach to changes in basic passport data is issuance of a new passport [68]; this may eventually become unworkable.) Because different RFID tags on the same passport can interfere with each other, it may not be feasible to include a new RFID tag with each visa stamp. Instead, we would like to keep the visa information on the same chip as the standard passport data. These features require writing new data to an e-passport after issuance.

A simple first attempt at visas for e-passports might specify an area of append-only memory that is reserved for visas. Each visa would name an e-passport explicitly, then be signed by an issuing government authority just as e-passport credentials are signed. An e-passport might even implement "sanity checks" to ensure that a visa is properly signed and names the correct e-passport before committing it to the visa memory area.

In some cases, however, a passport holder may not want border control to know that she has traveled to a particular location. For example, most Arab

countries will refuse entry to holders of passports which bear Israeli visas. As another example, someone entering the United States via Canada may wish to conceal a recent visit to a nation believed to be harboring terrorists. The first example is widely considered a legitimate reason to suppress visas on a passport; in fact, visitors to Israel from the United States may request special removable visa passport pages for exactly this reason. The second motivation may be considered less legitimate, and preventing this scenario may become a goal of future visa-enabled e-passports.

4.7.2 Function Creep

The proliferation of identification standards and devices is certain to engender unforeseen and unintended applications that will affect the value and integrity of the authentication process. For example, passports might come to serve as authenticators for consumer payments or as mass transit passes. Indeed, the ICAO standard briefly discusses the idea that e-passports might one day support digital commerce.

Function creep has the potential to undermine data protection features, as it will spread bearer data more widely across divergent systems. Moreover, function creep may lead to consumer demands for greater convenience, leading to the erosion of protective measures like optical-scanning-based access control and Faraday-cage use. Passport holders may wish to pass through turnstiles, for instance, without having to pause to have their documents optically scanned.

Web cookies are an instructive example of function creep. Originally introduced to overcome the stateless nature of the HTTP protocol, it was quickly discovered that they could be used to track a user's browsing habits. Today, web sites such as doubleclick.com use cookies extensively to gather information about customers.

4.8 Summing Up E-passports

We have identified principles for secure biometric identity cards and analyzed these principles in the context of the ICAO e-passport standard, current ICAO deployments, and Malaysian e-passports. We can draw several conclusions:

- The secrecy requirements for biometric data imply that unauthorized reading of e-passport data is a security risk as well as a privacy risk. The risk will only grow with the push towards unsupervised use of biometric authentication.

- At a minimum, a Faraday Cage and Basic Access Control should be used in ICAO deployments to prevent unauthorized remote reading of e-passports.

Today's e-passport deployments are just the first wave of next-generation identification devices. E-passports may provide valuable experience in how to build more secure and more private identification platforms in the years to come.

Chapter 5

Private Authentication

5.1 Problem Statement

We have seen that many of the privacy issues in the library and the e-passport setting arise because tags can be read by unauthorized readers. Therefore we would like to restrict access to a tag to only authorized readers. At the same time, adversaries who eavesdrop on a conversation or attempt to communicate with the tag should be unable to determine the identity of the tag. This is the symmetric-key *private authentication* problem: how can two parties that share a secret authenticate each other without revealing their identities to an adversary?

In the private authentication setting, there are n distinct RFID tags and one RFID reader. We assume that the reader has a database of tag keys TK_i; each key is a shared secret between the tag and reader. We consider an adversary that may interact with tags of its choice and eavesdrop on conversations between a legitimate tag and reader. We say a scheme for mutual authentication is *private* if an adversary is unable to distinguish two different tags with different secret keys. We say a scheme is secure if an adversary cannot fool a tag or reader into accepting when the adversary does not in fact know the tag's secret key.

A key performance metric for private authentication is how the amount of work performed by the reader scales with the number of tags in the system. The issue with private authentication is that the RFID reader does not know with which tag it is communicating. In the case that every tag has a different secret key, the "naive" approach is for the reader to try each key in turn. With this approach, the reader's work scales linearly with the number of tags n. This will not scale to large RFID deployments, which may have millions of tags. Our main contribution is a technique for mutual authentication that requires work only logarithmic in the number of tags.

Other performance metrics for private authentication include the amount of computation required by reader and tag. The number of bits which must be communicated is also important, because it affects how many tags may be read

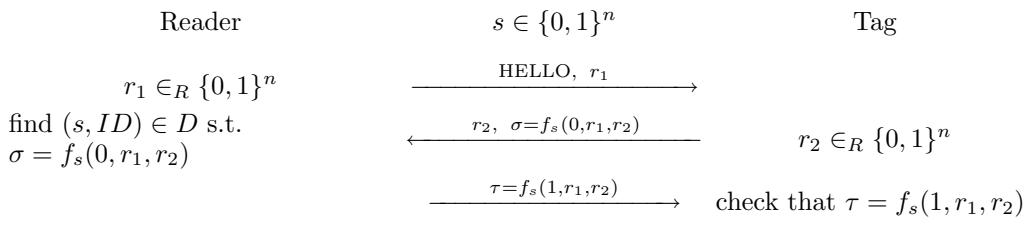

Figure 5.1: Our basic PRF-based private authentication protocol.

in a certain amount of time.

Another key metric for private authentication is how privacy degrades under *tag compromise*. In a tag compromise attack, the adversary learns the secret keys of some subset of tags in the deployment, then attempts to distinguish between tags it has not yet compromised. At one extreme, a system that shares the same secret key for all tags has catastrophic privacy degradation under tag compromise: if one tag is compromised, the entire system is compromised. At the other extreme, in a system that has different keys for each different tag, compromising one tag does not aid the adversary in tracking other tags. The problem here is that a system with different keys for each tag, as noted above, leads to work for the reader that is linear in the number of tags. Our technique shares key material between tags in a way that allows us to make a tradeoff between privacy degradation under tag compromise and reader work.

5.2 Solution: Private Authentication Schemes

In the following, we refer to a private RFID authentication scheme by a triple of probabilistic polynomial time algorithms (G, R, T) (for Generator, Reader, and Tag). Let k be a security parameter. The key generator $G(1^k)$ is a randomized algorithm that outputs a shared secret key K. Then the algorithms $R(K)$ and $T(K)$ interact to perform authentication.

5.2.1 A Basic PRF Private Authentication Scheme

We propose a scheme for mutual authentication of tag and reader with privacy for the tag. Our scheme, shown in Figure 5.1, uses a shared secret s and a pseudo-random function (PRF) to protect the messages communicated between tag and reader. The result is a private authentication scheme with reader workload linear in the number of tags. We refer to this basic PRF scheme as $(G_{\texttt{basic}}, R_{\texttt{basic}}, T_{\texttt{basic}})$.

5.2.2 Tree-Based Private Authentication

Next we discuss how to provide scalable private authentication. We build a new tree-based protocol with reader work $O(\log n)$, $O(\log n)$ rounds of interaction, and $O(\log n)$ tag storage, where n denotes the number of tags. Our scheme, $(G_{\texttt{tree}}, R_{\texttt{tree}}, T_{\texttt{tree}})$, assumes the existence of a subprotocol (G_1, R_1, T_1) that provides private authentication with constant rounds, constant tag storage, and reader work linear in the number of tags.

We consider the n tags as leaves in a balanced binary tree, then associate each node in the tree with a secret. Each secret is generated uniformly and independently. The reader is assumed to know all secrets. Each tag stores the $\lceil \lg n \rceil$ secrets corresponding to the path from the root to the tag. We can think of these secrets as defining a function $H : \{0,1\}^{\leq d} \to K$, where inputs are node identifiers and outputs are secret keys output by G_1.

The reader, when it wishes to authenticate itself to a tag, starts at the root and uses R_1 to check whether the tag uses the "left" secret or the "right" secret. If the reader and the tag successfully authenticate using one of these two secrets, the reader and tag continue to the next level of the tree, doing a depth-first search over the tree of secrets. If the reader passes all secrets in a path, the tag accepts the reader. Otherwise, the tag rejects the reader.

This tree-based scheme requires $\lceil \lg n \rceil$ invocations of R_1 and T_1 with 2 secrets. Therefore the total scheme requires $O(\log n)$ rounds of communication, $O(\log n)$ work for the reader, and $O(\log n)$ storage at the tag. Pseudocode is shown in Figure 5.2.

For example, we can use the basic PRF scheme shown above as our (G_1, R_1, T_1). In this case, we can re-use the nonces r_1 and r_2 at each level of the tree, so long as the nonces are long enough to prevent collisions. For example, we might set r_1 and r_2 to 128 bits each.

For simplicity of exposition, we described the scheme in terms of a binary tree, but nothing restricts the tree-based scheme to binary trees. Larger branching factors reduce the number of rounds of interaction and improve resistance against compromise tags at the cost of somewhat increased reader work.

The main issue with our scheme is the number of rounds of communication.Gentry and Ramzan have pointed out that some underlying protocols may allow performing all levels of the tree in parallel [29]. Such an optimization would yield a protocol with $O(1)$ rounds of interaction and messages with $O(\log n)$ length.

5.2.3 A Two-Phase Tree Scheme

As just described, the tree scheme uses a single fixed security parameter k for all instances of R_1 and T_1, which therefore requires communication cost at least k for each of the $\lceil \log n \rceil$ rounds, or $O(k \log n)$ communication. We now describe how we can create a tree scheme with communication $O(k + \log n)$ by splitting into two phases.

In the first phase, we run the tree scheme using R_1 and T_1 generated with

Algorithm $\mathrm{G}_{\mathtt{tree}}()$:
1. Let $d = \log n$. Let $H : \{0,1\}^{\leq d} \to K$ be a random function,
i.e., pick $H(s) \in_R K$ uniformly at random for each bitstring s of length at most N.
2. Assign each tag an identifier i that is d bits long.
3. For each tag, parse i in binary as $b_1, \ldots, b_d$.
4. Let $TK_i \leftarrow (v_1, \ldots, v_\ell)$, where $v_i = H(b_1, \ldots, b_d)$
5. Let RK be the values of H defined in Steps 2 and 3.

Algorithm $\mathrm{R}_{\mathtt{tree}}, \mathrm{R}_{\mathtt{tree}}(RK, TK)$:
1. Return $\mathrm{DFS}(r, 1, \epsilon)$, where ϵ denotes the empty bitstring.

Algorithm $\mathrm{DFS}(TK, i, s)$:
1. Set $ids := \emptyset$.
2. Parse TK as $(v_1, \ldots, v_\ell)$
3. If running $(R_1(H(s\,0)), T_1(v_i))$ returns true then
4. If $i \geq \ell$ then return $s\,0$
5. else set $ids := ids \cup \mathrm{DFS}(i+1, s\,0)$.
6. If running $(R_1(H(s\,1)), T_1(v_i))$ returns true then
7. If $i \geq \ell$ then return $s\,1$
8. else set $ids := ids \cup \mathrm{DFS}(i+1, s\,1)$.
9. Return *ids*.

Figure 5.2: Unoptimized tree-based private authentication protocol.

a smaller security parameter (that may depend on the level of the tree) to identify the tag. We can affect the probability of accidental mis-identification by trading off the branching factor and the phase-1 security parameter. In the second phase, once the tag is identified, the reader and tag can execute R_1 and T_1 using k as the security parameter.

For a concrete example, consider the basic PRF scheme, $n = 2^{20}$ tags, and a two-level tree with branching factor $2^{10} = 1024$. We give a tag three 64-bit secret keys: two for phase 1 and the final key for phase 2. In both levels, we truncate the PRF output to 10 bits. We then expect to need only one iteration of the first and one of the second level, for a total expected $2 \cdot 2^{10} = 2^{11}$ PRF evaluations for the reader and 4 PRF evaluations for the tag in phase 1, plus 2 each for phase 2. Communication cost is then $10 + 10 + 64 = 84$ bits of PRF output, plus 128 bits for the random nonces, for a total of 212 bits of communication. To fool a tag into accepting, the adversary must pass both phase 1 and phase 2. Ramzan notes that any authentication scheme with n possible tags requires $\Omega(\log n)$ communication cost, because writing a tag identifier requires $\Omega(\log n)$ bits, so we see our two-phase tree scheme is asymptotically optimal [71].

5.2.4 Privacy Under Tag Compromise

Tags in our scheme share parts of their keying material. Therefore, compromising a tag gives the adversary some advantage in identifying other, not yet compromised tags. The branching factor of the tree of secrets allows us to trade off privacy degradation under tag compromise with efficiency for the RFID reader. For example, if the branching factor is set to n, the number of tags, then we recover the "try all keys" scheme that has maximum resistance to tag compromise, but minimum efficiency. On the other hand, because each tag has at least one key unique to that tag, an adversary that compromises tags cannot impersonate any tags not so compromised. This makes our scheme qualitatively different from the approach of giving each tag the same secret key - in that case, a single compromised tag loses privacy and security for the whole system.

Analyzing the exact privacy-efficiency tradeoff for our scheme is outside the scope of this thesis, although we do discuss several other tree configurations in the next chapter. Going further, Avoine, Dysli, and Oechslin quantify the effect on privacy under tag compromise at different branching factors by measuring the adversary's advantage at distinguishing tags [8]. Nohara et al. measure the effect of compromising a single tag using an entropy-based metric for anonymity [63].

Chapter 6

RFID Pseudonyms

6.1 Problem Statement

This chapter discusses RFID *pseudonym protocols*. In a pseudonym protocol, an RFID tag does not emit its "real" identifier when queried by a reader. Instead, the tag returns a special *pseudonym*. Our goal in a pseudonym protocol is to ensure that pseudonyms can be mapped to the real identifier only by parties we trust. Then, by ensuring that pseudonyms cannot be linked to each other or the real identifier, we obtain privacy for the movements of RFID tagged items.

Pseudonyms vs. Access Control. Providing RFID pseudonyms is not the same as access control for RFID tag data. In access control, the object is to prevent an adversary from reading or writing tag data without a secret key. With RFID pseudonyms, the adversary is allowed to read the tag's current pseudonym, but must be unable to determine when it has read the same tag twice.

Access control requires mutual authentication between RFID tag and reader; the tag must know it receives commands from a legitimate reader, and the reader must know it is sending information to the correct tag. Mutual authentication is overkill for many RFID applications, because in most cases we simply want to know the tag's identity. Writing to the tag or other commands from the reader are not necessary. In a common use case for RFID, a large number of items pass by a reader in a short amount of time. Even if performing mutual authentication with a single item is fast, authenticating a reader to dozens of tags at once may be challenging. In addition, mutual authentication cannot be carried out with a single message from tag to reader, while an RFID pseudonym protocol provides exactly this.

Furthermore, a pseudonym protocol works with legacy RFID readers that have not been engineered with pseudonyms in mind. To such a reader, a pseudonym appears to be a legitimate tag ID. The reader simply passes the pseudonym to a back-end database as it would with a standard ID. The database in turn can ask the infomediary to map the pseudonym to the correct tag ID.

Privacy Control. In many settings, we may wish to have a single party manage access to many tags. Thus, we assume the presence of a central trusted entity, which we call the Trusted Center(TC). Given any pseudonym from such a tag, the TC can determine the identity of the tag using a database of secret information..

For example, if a cryptographic pseudonym protocol is used, the tag is loaded with a secret key generated for it by the TC. The TC keeps a database listing for each tag with the secret key that was provided to that tag and any data that is to be associated with that tag (such as its identity or access policy). Upon receipt of a pseudonym, the TC can use its database of secrets to map the pseudonym to the tag's correct ID.

The Trusted Center acts as a trusted third party that manages the privacy policy associated to tags. While the RFID tag manufacturer could act as a Trusted Center in practice, this is not required. RFID tags could be shipped without any secrets written into them. Then, when the tag is first deployed, the Trusted Center can write the relevant secrets to the tag. We can construct tags that may only be written in such a way once, and then do not permit overwriting or reading of secrets thereafter. For example, a library deploying RFID could enroll a tag and write secrets to it when the tag is applied to a library book; the library would then act as the Trusted Center. For another example, the Infomediary we describe in the Introduction is another possible Trusted Center.

Controlled Delegation. In the future, a RFID infrastructure might consist of thousands or even millions of RFID readers deployed across the planet, and we need a way for legitimate readers to be allowed to read the tag. In a naive implementation, a TC for the tag would give a copy of the tag's secret key to each reader that is authorized to read the tag. This form of delegation is too coarse-grained, because the reader then permanently receives the ability to identify this tag for all time. We may not wish to place this much trust in every RFID reader that ever encounters the tag; the challenge is to provide controlled delegation, where a reader's ability to read a tag can be limited to a particular time period.

If readers are online, one simple approach is to have the reader simply act as a dumb relay, passing on the pseudonym from the tag to Trusted Center and letting the TC reply with the identity of the tag. In such a scheme, the TC can indeed authenticate the reader and check the privacy policy of the tag before responding to this reader's request. If a reader Alice wishes to determine a tag's ID, she must ask the TC. The TC can then decide whether Alice is allowed to see this information based on the tag privacy policy stored in the database. However, one limitation of this approach is that it requires a costly interaction between the reader and TC every time a tag is read. Because today's readers may repeatedly broadcast queries to all tags within range at a rate of 50 times per second or so, the burden on the TC and the database may be very high: if there are 10 tags within range, we require 500 round-trip interactions per second with the TC, multiplied times the number of readers.

We instead focus on the problem of *offline delegation*, in which readers need not be connected to the Trusted Center during a tag reading. Offline delegation is helpful for cases where readers have intermittent or low-bandwidth connectivity. When a reader first sees a tag it is unable to recognize, the reader can send the pseudonym it received to the TC. If this reader is authorized this tag, the TC can return not only the tag's identity but also a secret that allows reading the tag for a limited time (say, for 1000 queries). Because tags typically repeatedly query their environment many times a second, this allows any arbitrary number of subsequent queries to be disambiguated locally at the reader, without requiring further interaction with the TC (until the query limit is exceeded). Thus, a scheme that supports delegation can still be used with online readers. Further, the ability to exploit the locality in tag sightings can be used to greatly improve performance of readers.

One could even ask for *recursive delegation*. With recursive delegation, once we have delegated to Alice limited-access to the tag, she can further re-delegate to other readers. Alice can delegate to Bob the power to query this tag, and Bob can further delegate to Carol, and so on. Moreover, the rights delegated can be limited arbitrarily at each step. For instance, if Alice receives a secret that lets her identify the tag for the next 100 queries, she can compute a secret for Bob that will let him read the tag for the next 40 queries, a secret for Bill that lets Bill read the tag for the 30 queries after that, and so on. To the best of our knowledge, no previous work has addressed delegation in the context of RFID tags, let alone recursive delegation.

Ownership Transfer. A related problem to delegation is that of *ownership transfer*, when Alice gives an RFID-tagged item to Bob. After the transfer of ownership, Bob should be able to read the item but Alice should not. Pseudonyms allow us to cleanly deal with ownership transfer from Alice to Bob. If Alice has not been delegated the ability to disambiguate pseudonyms, no further work is needed: the TC simply denies Alice's requests to disambiguate pseudonyms after Bob registers his ownership of the item. If Alice has been delegated such ability, then the pseudonym protocol must somehow support ownership transfer.

Performance Metrics. A major technical challenge in the design of RFID pseudonym systems is how to make them scalable to a large number of tags. Consider a TC with a database of n tags that receives a pseudonym to be disambiguated. Naively, one might check, for each of the n tags known to the TC, whether this pseudonym could have been generated by that tag. This naive strategy, unfortunately, requires $O(n)$ work each time a RFID tag is read, which may not be practical for an RFID system with $n = 10^6$ tags. Therefore, reader work is a key performance metric, just as in the case of private authentication.

In addition, the amount of computation required for a tag is an important metric, because RFID tags may have few gate available for security. The amount of communication is also a key performance metric. Again, this is similar to the case of private authentication.

6.1.1 Threat Model

Our parties are a Trusted Center, a Reader, and many different Tags. We now outline the security goals and threat model for RFID pseudonyms.

First, we want *privacy* for RFID tag readings: without specific permission by the Trusted Party, a reader cannot determine the tag identity from the pseudonym or otherwise link different readings of the same tag. Ideally, privacy should hold even when the adversary is allowed to ask for delegated access to tags of its choice. In particular, an adversary should be unable to map a tag's pseudonym to the tag's ID unless it has been specifically delegated access to the tree leaf currently used by that tag.

Second, we want *replay-only security against impersonation attack*. In an impersonation attack, an adversary wishes to pretend it is a legitimate RFID tag without knowing that tag's secrets. Because a pseudonym protocol uses only one message from tag to reader, it necessarily falls victim to a replay attack in which an adversary records a tag's pseudonym and replays it later to an RFID reader. We want a protocol where replay is the "worst" an adversary can do: without the secret keys of a tag, an adversary cannot generate valid tag pseudonyms it has not yet seen. We believe this limited replay-only security is tolerable, as duplicate readings of the same pseudonym can be detected and handled by a back-end database correlating RFID information.

In our threat model, the adversary is allowed to eavesdrop on all conversations between a legitimate reader and a tag. The adversary is allowed to query the Trusted Center with pseudonyms and learn whether the pseudonym is correctly formed, and to interact with tags of its choice.

6.2 Solutions

6.2.1 Solution: Recoding

Writable RFID tags allow us to *recode* an RFID tag, or rewrite the data it carries. We can use this feature to recode an RFID tag with the name of a Trusted Center and a random identifier. Then a reader can, given the tag reading, query the TC and ask for the tag's "real" ID. The TC can then apply the privacy policy of the tag to decide whether to honor the request. Recoding is a simple way to implement RFID pseudonyms: each random identifier serves as a new pseudonym.

Recoding requires rewriteable tags, but the ability to rewrite a tag must be protected. Otherwise, RFID tag "vandalism" becomes possible, as a vandal can change the data on an RFID tag to make an item appear to be something it is not, or simply erase the tag entirely. Vandalism might be performed to deny service to legitimate users, or there might be some financial motive involved. While RFID tag vandalism has not yet been reported, we suspect it is only a matter of time.

With respect to financial motives, scams have already appeared that switch optical bar code labels. For example, Home Depot suffered nearly half a million

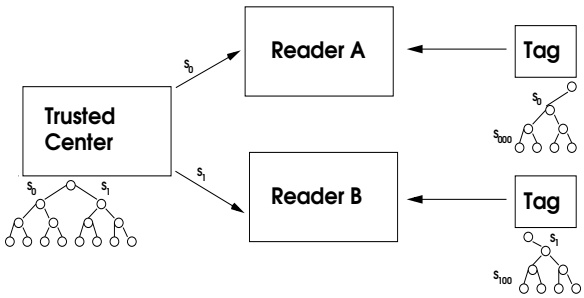

Figure 6.1: The Trusted Center delegates access to two different Readers.

dollars in losses from a group of thieves that created bar code labels for low-cost items, pasted them on top of high-cost items' labels, bought the items at a discount, and then returned the item for the full price. In the RFID setting, we could expect to see a quick "cloning" of other items found in the same store, in which a thief would read a code off a cheap (but similar) product, then overwrite the tag of a more expensive product. As noted in chapter 3, many of today's RFID tags employ a "write then lock" architecture, in which the tag data can be written an unlimited number of times and then irrevocably locked. After locking, the data on the tag cannot be modified or erased. Unfortunately, this irrevocable lock does not work for recoding, because the data on the RFID tag must be modified. Instead, some kind of write password will need to be employed. Therefore, support for infomediaries via recoding requires managing RFID tag passwords. The main benefit of RFID recoding, however, is that it can be implemented using cheap tags that have no cryptographic capabilities, such as today's EPC Gen 1 tags.

6.2.2 Solution: Scalable, Delegatable Pseudonyms

Recoding RFID tags suffers from at least one other drawback: between recodings, the tag contains the same static identifier. Therefore a tag can be tracked and hotlisted between recodings. In a cryptographic pseudonym protocol, in contrast, as introduced by Ohkubo, Suzuki, and Kinoshita [69], a tag emits a different pseudonym each time it is read. A Trusted Center (TC) can disambiguate the pseudonym and reveal a tag identifier using a secret key shared between tag and TC. An adversary who lacks the secret key cannot link such sightings. Possession of the secret key "controls" the ability to link sightings of the same tag. Unlike recoding, the tag changes its own pseudonym on each reading. The result is improved privacy, because the tag may no longer be tracked and hotlisted between interactions with a trusted reader. On the other hand, the RFID tag must be able to compute a pseudo-random function. While this is already available for e-passport devices, it is less clear whether it is reasonable for supply chain tags. We will see, however, that given a pseudo-random function, we can obtain all our desired features for RFID pseudonyms.

Furthermore, Ohkubo et al.'s scheme does not support delegation and has a high workload for the RFID reader. We address these problems with a novel cryptographic pseudonym scheme for RFID tags. Our scheme supports offline and recursive delegation: the TC can compute a time-limited secret that only confers the ability to disambiguate a limited number of tag pseudonyms. In particular, the TC computes a secret that allow to recognize the next q pseudonyms from this tag, where q is arbitrary and can be specified by the privacy policy. This secret can be communicated to Alice, the reader, through any channel, and thereafter the reader does not need to interact with the TC in any way. In Figure 6.1 we show a diagram of how delegation works in our scheme with different RFID readers and the Trusted Center. To the best of our knowledge, no previous RFID schemes support delegation.

Our scheme supports ownership transfer. If Alice has been delegated linking ability, we have two methods for ensuring Alice can no longer link a tag after it is passed to Bob. First, a method we call *soft killing,* and second a method for securely incrementing a tag's leaf counter. We describe both methods in more detail in Section 6.2.6. Previous work on ownership transfer focused on a "recoding" technique with writeable RFID tags, in which a tag is overwritten with a new identifier that does not change between recodings. Therefore the RFID tag is still vulnerable to tracking and hotlisting until it is recoded [59]. Recoding also introduces the problem of managing secure access to the recoding operation. Our scheme addresses both these problems.

Ohkubo et al.'s scheme requires the reader to perform work linear in the number of tags to map from a tag's pseudonym to its ID [69]. In contrast, for our protocol, the TC needs only do $O(\log n)$ work to disambiguate a pseudonym. The logarithmic complexity does not apply to readers who have been delegated access to a subset of tags: a reader can disambiguate each pseudonym in $O(D)$ time, where D is the number of tags delegated to the reader. In practice we expect D will be small compared to the total number of tags; for example, D might be the number of tags in a single shipment of goods. Fortunately, since there is a great deal of locality in tag-reader interactions, most readers will only be associated with a small number of tags, so we expect this performance level to be more than adequate in practice.

6.2.3 Protocol Overview

Privacy in RFID must consider all layers of the device. In particular, devices should have private collision avoidance, i.e. the radio behavior should not leak a unique identifier for a tag or allow linkage of tag sightings. Physical differences due to manufacture might also allow an adversary to link different reads of the same tag. Avoine discusses several ways in which existing RFID tags leak such information and ways to fix them [9]. Our work assumes that these problems have been solved.

The main idea of our scheme is to store a "tree of secrets" on the RFID tag. Our solution requires a pseudo-random function and a non-volatile counter on the RFID tag. Given recent results on AES implementation for RFID by

Scheme	T_{Reader}	S_{Reader}	T_{TC}	S_{TC}	# Msg	Comm	Delegation?
OSK [69]	$O(n)$	$O(n)$	NA	NA	1	$O(1)$	No
AO [10]	$O(n^{\frac{2}{3}})$	$O(n^{\frac{2}{3}})$	NA	NA	1	$O(1)$	No
MW [60]	$O(\log n)$	$O(1)$	NA	NA	$O(\log n)$	$O(\log n)$	No
Basic Scheme	$O(D)$	$O(D)$	$O(\log n)$	$O(2^{d_1})$	1	$O(\log n)$	Yes
Optimized Scheme	$O(D)$	$O(D)$	$O(\log n)$	$O(1)$	1	$O(\log n)$	Yes

Figure 6.2: Comparison to previous RFID privacy schemes. Here T_{TC} and S_{TC} stand for the time and storage requirements of the Trusted Center, with the Reader requirements marked similarly. n is the total number of tags in the system, d_1 is the depth of the Trusted Center's tree, and D is the number of tags delegated to a particular reader. In practice, we expect $D \ll n$. The Optimized Scheme uses a PRF to generate the TC's tree of secrets and truncates the tag outputs, as described in Section 6.3.

Feldhofer et al., this appears reasonable for a large class of tags [25]. Each tag keeps a counter, which is incremented on each read. The counter stores the index of the next leave of the tree to use. The path from root to leaf, combined with a random nonce, determines the tag's response to an RFID reader. After each response, the tag "updates" itself and its key material to the next leaf in the tree. Some key material is shared between the Trusted Center and the tag alone. This key material allows the TC to determine the tag's identity with logarithmic work. The other key material may be given by the Trusted Center to an RFID reader. Because the tag evolves its key with each step, this delegated key material will "expire" after a certain number of tag reads.

For simplicity, we will describe our protocol as if a random number generator exists on the RFID tag. In some RFID technologies, this may not be realistic; therefore we show later how to replace this with an increasing counter. We will also limit the description to a binary tree of secrets, but in practice we will want to pick a tree with a high branching factor to make a tradeoff between reader work and tag communication.

6.2.4 Notations and Background

We use the following notation.

- In the following description we use a pseudo-random functions (PRF) that uses key k from a key space K on input M of length n-bits and output n-bits. $F : K \times \{0,1\}^n \rightarrow \{0,1\}^n$. We write $F_k(M)$.
- A pseudo-random generator (PRG) on input M of length k-bits is defined as: $G : \{0,1\}^k \rightarrow \{0,1\}^k \times \{0,1\}^k$. We write $G_{\{0,1\}}(M)$. By $G_0(M)$ we denote the first k bits output G on input M. By $G_1(M)$ we denote the next k bits output G on input M.

- Let $\{0,1\}^{\leq n}$ denote the set of bitstrings of length at most n. If $s \in \{0,1\}^*$ is a bitstring, let $s_{1..i}$ denote the first i bits of s, and let len(s) denote the length of s (in bits).
- We also view $s_1 2^{n-1} + \cdots + s_{n-1}2 + s_n$. Each bitstring $s \in \{0,1\}^{\leq d}$ identifies a node in the tree; $s = 0$ and $s = 1$ are its left and right children, respectively.
- If $f : S' \to T$ is a function and $S \subseteq S'$, let $f|_S : S \to T$ denote the function f restricted to S. When given a function $h : \{0,1\}^{\leq d_1} \to K$ defined on $\{0,1\}^{\leq d_1}$, we extend it to a function defined on all of $\{0,1\}^*$ as needed by defining $h(sb) = G_b(h(s))$ for every $s \in \{0,1\}^{>d_1}, b \in \{0,1\}$.

We define a rooted full binary tree of depth d with k-bit string stored in the nodes and edges labeled 0 or 1. The tree stores random k-bit strings in all nodes $\leq d_1$. In the nodes of succeeding levels it stores k-bit string computed by applying G as follows. If a k-bit string is stored in an internal node v, then $G_0(v)$ is stored in $v's$ left son and $G_1(v)$ is stored in $v's$ right son.

If $s \in \{0,1\}^d$ is a bitstring representing the position of a node v at the leaf. Let $s_{1..d-1}$ denote the position of $v's$ parent. The ancestor path from leaf to the root is defined by the nodes in position: $(s_{1..d-1}), (s_{1..d-2}), ..., (s_{1..1})$ and the function $h(s_{1..i})$ represents the k-bit string value of the node in position $s_{1..i}$.

6.2.5 Our Protocol

In our RFID protocol scheme, we assume a central Trusted Center that can authenticate and authorize readers. Each Tag has a unique ID, which we would like to keep secret from a Reader unless the Reader's request meets a privacy policy associated with the Tag. The Reader interacts with a Tag and learns a one-time pseudonym p. Then the Reader asks the Trusted Center to identify the Tag.

We first describe the basic "tree of secrets" which is used to generate the one-time pseudonym, including a description of the setup phase. We then describe the process through which a tag responds to the reader. Next we describe the mapping from pseudonym to tag identity, focusing on the problem of delegation. Later we will show how our protocol enables a secure transfer of ownership without need to rekey the tag. The state and algorithms for each party are shown in Figures 6.4, 6.5, and 6.6.

Tree of Secrets. To ensure our privacy goal the pseudonym needs to be updated whenever a tag response is generated. Our protocol is based around a tree of secrets of depth $d = d_1 + d_2$ as shown in Figures 3. Each node in the tree represents a cryptographic secret of length k-bit.

The first d_1 levels of the tree contain node secrets that are chosen independently of each other. The Trusted Center maintains the tree and generates these secrets at system initialization time using the algorithm TC.GenTC. The TC associates each tag with one node of the tree at depth d_1 and the following

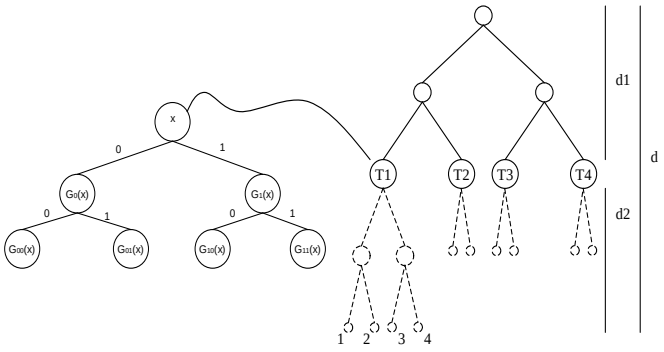

Figure 6.3: An example tree of secrets for four tags in our RFID pseudonym scheme. The nodes drawn with solid lines correspond to secrets shared only between the tags T1,...,T4 and the Trusted Center. Each of these secrets is drawn uniformly at random and independently of each other. The dashed line nodes are secrets in *delegation trees*, where child nodes are derived by the GGM construction of applying a pseudo-random generator to the parent. On each read, a tag updates its state to use the next leaf in the delegation tree for its next pseudonym. To delegate limited-time access to a Tag, the Trusted Center can give out subtrees of the delegation tree; for example, the immediate parent of 1 and 2 allows learning T1's identity in time periods 1 and 2, but not in time periods 3 and 4.

property will always hold: each tag knows all the keys from its node at depth d_1 up to the root node, but not other nodes in the tree. Secrets above d_1 in the tree are shared only between a Tag and the Trusted Center; a Reader will not have access to these secrets. Formally, we model the secret generation as a random function H kept by the Trusted Center and generated during TC.GenTC. All provisioning is done by the TC, which also ensures no tags are given the same secrets at level d_1. This algorithm is shown in Figure 6.5 (see TC.EnrollTag). The TC at enrollment time also records each tag's real identity ID, which may be an arbitrary string.

The next d_2 levels of the tree contain node secrets that are derived using a GGM tree construction [31]: each node is labeled with a secret, and the secrets for its children are derived by applying a PRF. Knowing a secret in the tree allows computation of the secrets for every descendant, i.e. the subtree rooted at that node, but nothing else. From [31], if we denote a secret x stored in a node a depth d_1 then $G_0(v)$ is stored in $v's$ left son and $G_1(v)$ is stored in $v's$ right son. Let $s = s_1 2^{n-1} + \cdots + s_{d-1}2 + s_d$ be a binary string. The value of a node at depth D is $G_{s_d}(G_{s_{d-1}}(\ldots(G_{s_{d_1}}(x))))$. These secrets are shared between a Tag and the Trusted Center and can be shared with a Reader during

Tag State:
c, a counter in $\{0,1\}^d$. Initialized to 0.
h, where $h = H|_S$ for some set $S \subseteq \{0,1\}^{\leq d_1}$.

Algorithm TAG.RESPOND():
1. Pick $r \in_R \{0,1\}^k$ uniformly at random.
2. Set $p := (F_{h(c_{1..1})}(r), F_{h(c_{1..2})}(r), \ldots, F_{h(c_{1..d})}(r))$.
3. Set $c := c + 1$.
4. Return (r, p).

Figure 6.4: Algorithms and state for the RFID tag.

the delegation process.

Tag Responds. Having access to subtrees of secrets is important for a Reader, because these subtrees allow the Reader to map the Tag's pseudonym (r, p) to an ID without needing the Trusted Center. Each Tag T keeps a counter $T.c$. A Tag responds to a query from the Reader by generating a random number r and sending a pseudonym

$$(r, p) = (r, F_{h(c_{1..1})}(r), F_{h(c_{1..2})}(r), ..., F_{h(c_{1..d})}(r))$$

where the $h(c_{1..i})$ values represent secrets along the path in the tree of secrets from the root to the Tag's current leaf $T.c$. The Tag then increments the counter c. In practice, the counter value might be 64 bits.

Pseudocode for computing the response is shown in Tag.RESPOND. We can think of each leaf value c as corresponding to a new pseudonym of the tag. Below we discuss how the Trusted Center and the Reader can use their trees of secrets to map the pseudonym (r, p) to the Tag's ID. Notice that because the counter c increments, the Tag will use a different path of secrets, and therefore a different pseudonym, for every reader response: this is what ensures that the Reader's subtree of secrets will "expire" after a certain number of tag reads. The complexity of Tag.RESPOND depends on the overall depth of the tree, however, not directly to the size of the counter. By varying the branching factor and depth of the tree, we can trade off between the complexity of Tag.RESPOND and the complexity for the reader; we return to this in more depth in Section 9.

Mapping and Delegation. To map a pseudonym p to the Tag's identity, the TC starts at the root of the tree of secrets. Then the TC performs a depth-first search over the tree, looking for the path in the tree that matches the response p. At each node s, the TC can check whether the left child $s\,0$ or the right child $s\,1$ matches entry p_i in the response by checking whether $F_{s0}(r) = p_i$ or $F_{s1}(r) = p_i$, respectively. Pseudocode is shown in Figure 6.5 (see TC.IDENTIFYTAG). Then the TC can map from the identity of the tag's current node to the tag's real identity ID. Based on ID, the identity of the Reader, and a privacy policy, the TC can then decide whether to reveal ID to the Reader. This provides a

mechanism for enforcing a privacy policy regarding which readers are allowed to learn which Tag IDs.

With this approach, the TC must be online for every tag read, which may incur too much overhead for the TC. Our protocol also allows for "offline delegation" the TC to delegate access to a certain interval of pseudonyms to the Reader. This can be thought of as allowing the Reader to perform the mapping itself from a pseudonym (r, p) to the Tag's identity ID, but only if the Tag's counter value is in a prescribed interval $[L, R]$ (for some $1 \leq L \leq R \leq 2^d$).

Recall that each leaf of the tree corresponds to a different pseudonym for a tag. To delegate access to leaves in an interval $[L, R]$, the Trusted Center first determines the set S of all x_i such that the following two conditions hold. First, for all $x \in [L, R]$ there exists $s \in S$ such that s is a prefix of x. Second, for all $s \in S$, for all $t \in \{0,1\}^d$, if s is a prefix of t, then $t \in [L, R]$. It can be shown that S contains at most d_2 elements.

The Trusted Center then sends $H|_S$ to the Reader along with the Tag's identity. Pseudocode is shown in TC.DELEGATE. In terms of our tree, the set S corresponds to the minimal set of nodes that covers exactly the interval $[L, R]$. Now, when the Reader sees the Tag's pseudonym (r, p), the Reader no longer needs to communicate with the Trusted Center. Instead, the Reader computes $F_{h(s)}(r)$ for all $s \in S$, which it can do because it has access to $H|_S$. If the Reader finds a match between the tag response and an s value, then it has learned the Tag's identity. Pseudocode for the Reader's computation is shown in Figure 6.6 (see Reader.IDENTIFYTAG). After the Tag updates itself past the leaf R, however, the Reader can no longer map the Tag's pseudonym (r, p) back to the Tag's identifier ID. This is because the counter TAG.c will have updated past the subtree of secrets known to the Reader. At that point, the Reader must re-apply to the TC for more access.

During the depth-first search, the TC determines which node at level d_1 is currently in use by the Tag. This requires $2d_1$ evaluations of a PRF. Because each tag has at least one node at level d_1 of the tree and none of these values are shared between tags, this requires only $O(\log N)$ evaluations of the PRF. If the TC further wishes to learn the exact counter value used by the Tag, this requires another $2d_2$ evaluations of a PRF.

The Reader, by contrast, must check every value in its delegated subset S to see if it finds a match with an entry of the Tag's response. This takes time $O(D)$, where $D = |S|$.

6.2.6 Ownership Transfer

Ownership transfer in RFID is the following problem: Alice gives an RFID tag to Bob. How do we prevent Alice from later reading the RFID tag? This problem is crucial for limiting the trust required in readers which may need to read tags at some point in the tag's lifetime.

In the case that Alice has not been delegated access to the RFID tag, ownership transfer in our model is simple. The Trusted Center is notified of the transfer and updates a privacy policy associated with the tag. Afterwards, Al-

ice requests access to the tag's ID. The Trusted Center then checks the privacy policy, sees Alice no longer owns the item, and denies access. In case Alice has been already been delegated access to the tag, we introduce two methods for ownership transfer.

Soft Killing. In the first method, *soft killing*, Bob queries the Trusted Center and learns how many leaves were delegated to Alice. Suppose this number is k. Bob then reads the tag $k+1$ times. The tag will then have updated past Alice's access, so she will no longer be able to disambiguate the tag's pseudonyms. Notice that even if Bob knows how many leaves were delegated to Alice, he still cannot distinguish a tag delegated to Alice from any other tag without Alice's help; this is because the tag will emit a new, pseudorandom, pseudonym on each read. Therefore knowing the number of leaves delegated to Alice does not hurt the privacy of our protocol.

The benefit of soft killing is that it does not require shared secrets between the tag and reader. The downside is that soft killing requires many tag reads. Soft killing also opens up the possibility for a denial of service attack if an adversary reads the tag many times; Alice can recover from this by simply asking the Trusted Center to delegate more access.

Increasing The Tag Counter. In the second method, we allow Bob to increase the counter on a tag from c to c'. Bob does so by sending the Tag a random seed r, after which Bob and the Tag can perform mutual authentication and establish a secure channel with the shared secret $F_{h(c)}(r)$. Bob then sends c' to the tag. We require that $c' > c$, so Bob can only increase the tag's counter, not decrease it. Alternatively, Bob can send the Tag a similar message identifying a subtree; the tag then updates itself to the least leaf in that subtree. By doing so, Bob can "leapfrog" the tag over Alice's delegated leaves and be sure that Alice can no longer read the tag. Increasing the counter requires only one read, but also requires that Bob share a secret with the tag. Notice that the Trusted Center need not be involved at all in the transaction in this case.

6.3 Optimizations

We now present some optimizations for our pseudonym protocol. We note that some of these optimizations also apply to the private authentication protocol described in the previous chapter.

Reducing TC Storage. In our protocol as described, the Trusted Center must generate and store 2^{d_1+1} independent random values. We can reduce this storage to $O(1)$ by instead having the Trusted Center use a PRF with a secret that is never revealed to any other party. This PRF evaluated at a nodeID yields the secret for the node.

From PRFs to Weak PRFs. Throughout we have assumed the use of pseudorandom functions for generating tag responses. Because of the structure of our protocol, however, a *weak PRF* would suffice instead [61]. A weak PRF is a

keyed function whose output is indistinguishable from random assuming that the input was chosen uniformly at random. Unlike a PRG, the input is public and known to an adversary; only a fixed key is secret. Unlike a standard PRF, the adversary is not allowed arbitrary access to the function. A weak PRF is sufficient for our protocol if a hardware random number generator is used; in this case the nonce value r is drawn uniformly at random from the appropriate set of bitstrings. It remains an open question whether weak PRFs are more efficient to construct than PRFs in practice on RFID devices.

Trading Randomness for a Counter. In some RFID technologies, it may be difficult to generate random numbers. We can replace the random number generator with a counter in such situations. We refer to Figure 6.3, which shows algorithms NextNonce.Initialize and NextNonce.GetNextNonce that use a counter *ctr* and a pseudo-random function F_k to generate the next nonce for an RFID pseudonym, where k is a uniformly chosen random secret. We stress that the key k for this pseudo-random function is not shared with a Reader at any time. For each pseudonym, we increment the counter c and return $F_k(c)$ as the next nonce. If we use this optimization, however, then we can no longer use a weak PRF, as the inputs to F_k are not chosen uniformly at random.

Truncating PRF Values. Instead of sending full PRF values in a tag response, it is more efficient to send truncated versions. This reduces communication overhead at the cost of following false paths during the depth-first search. To avoid misidentification of tags, we recommend truncating only at the internal nodes and sending the full-length PRF output at the leaves. If internal nodes are truncated to a bits, the tag's response becomes (r, p) where $p := (F_{h(c_{1..1})}(r) \bmod 2^a, ..., F_{h(c_{1..d-1})}(r) \bmod 2^a, F_{h(c_{1..d})}(r))$. With full-length values at the leaves, the probability of misidentification is negligible.

When PRF responses are truncated, identifying a tag requires searching through the tree, and this search might follow false paths that do not correspond to the true tag identity. If the branching factor is exactly 2^a, it is possible to show that the search process is a birth-death process and that the expected complexity of the search is $O(2^a \times \lg N) = O(2^a \times d)$.

Branching Factor and Concrete Examples. Truncation greatly reduces communication overhead while only slightly impacting the complexity of tag identification. For instance, with a binary tree of depth $d = 40$, we might truncate PRF values to 1 bit at internal nodes and use a 64-bit PRF output at the leaves. With these parameters, the response p will be 103 bits long, while the search complexity remains minimal.

In practice, we would use trees with branching factors much larger than 2. A larger branching factor reduces the depth of the tree, thus reducing tag storage and computation, at the cost of more computation for the Trusted Center and reader. For example, consider an RFID system with $N = 2^{20}$ tags, each of which will be read at most 2^{20} times. We construct a four-layer tree of secrets with branching factor $1024 = 2^{10}$ at all levels. Each tag stores two 64-bit secrets s_1, s_2, with the second secret being the root of a GGM tree that covers the final

two tree levels. Each pseudonym requires two PRF invocations to compute s_3, s_4 and four PRF invocations to compute the response. Total tag storage is $2 \cdot 64 = 128$ bits and total tag computation is 6 applications of the PRF. If we truncate the tag's responses to 10 bits at internal nodes and 64 bits at the leaf, and use a 64-bit r, the tag's total communication is $64+30+64 = 158$ bits. The work for the reader, on the other hand, is only $6 \cdot 2^{10}$ applications of the PRF. We show concrete parameters for this and some other examples in Figure 6.8.

6.4 Application Scenarios

Shipping. We note that UPS has recently begun experimenting with RFID [54]. In a shipping scenario, Alice wishes to send a package to Bob using a shipping company Charlie. Charlie needs to read the RFID tag to expedite tracking of Alice's package. After Charlie delivers the item to Bob, however, Bob wants to prevent Charlie from reading the tag in the future.

For example, Bob might be a retail store that has a shipping contract with Charlie; if Charlie can walk into the store and discover that he ships eighty per cent of Bob's items, Charlie might use this information against Bob at the next contract negotiation. Therefore, we would like to use our mechanisms for ownership transfer to ensure Charlie has only limited-time access to the RFID tag.

Here Alice can act as the Trusted Center in our protocol. Together with the shipping information, Alice delegates access to Charlie for enough pseudonyms to allow for tracking the RFID tag during shipping. Should Charlie exhaust his access during shipping, he can contact Alice for more access. Once the item reaches its destination, Bob increments the counter on the Tag, thereby gaining confidence that Charlie can no longer determine the Tag's identity from its pseudonyms.

Warehouses. In this application, RFID tags are applied to cases or pallets in a warehouse owned by Alice. Here we are concerned with competitive intelligence: by hiding an RFID reader near a warehouse or by controlling RFID readers at a point of shipment, Eve might learn something about Alice's business practices [76].

With our protocol, Alice sets up her own Trusted Center off-site and several readers inside the warehouse. Alice then delegates access to the readers based on her knowledge of which items should arrive at the warehouse. The reader, given this access, can quickly determine a tag's identity from the pseudonym emitted by the tag. Eve, in contrast, cannot even link two different sightings of the same rag. Furthermore, even if Eve steals one of Alice's readers, the exposure is limited to the access delegated to that reader.

Supply Chain. In the supply chain, a tag may travel from a manufacturer to a distributor, then to a retail outlet and finally to an end consumer. At each step, RFID readers need to read the tag for a limited time, but should not be able to read the tag after it has passed to the next point in the chain. Our protocol

supports this by allowing the Trusted Center to delegate access to the readers for each of the entities in turn. At point of sale, we can perform soft killing of the tag and therefore guarantee that no previous party can determine the tag's identity.

Pharmaceuticals. The U.S Food and Drug Administration is pushing for the widespread use of radio frequency identification to track the distribution of prescription drugs. RFID should start being used at the pallet level throughout the pharmaceutical supply chain within the next three years. Jump Start, the first RFID trial in the pharmaceutical industry, aims to track products from manufacturers to the distributors [85]. CVS/Pharmacy Corp has led the use of RFID, introducing item-level tracking [53].

Our mechanisms for ownership transfer to different readers are applicable here. After a transfer, the new owner can enforce tag privacy by soft-killing or by increasing the tag counter. Then, if a previous owner or other adversary tries to read the tag, it will be unable to collect any information about the drugs from the resulting pseudonym.

Libraries. As in the warehouse case, the library itself can act as the Trusted Center and delegate access to its own readers. Then if an adversary reads a library book's RFID tag without authorization, the adversary will be unable to determine the book's identity. If patrons wish to read the tag themselves, the library can delegate access to their reader and revoke access once the book is returned.

6.5 Related Work

Weis et al. provide "hash lock" protocols for private mutual authentication [84]. As we have discussed, mutual authentication is not needed in scenarios when only tag identification is required, and it incurs significant performance costs. Their schemes also require readers to perform linear work in the number of total tags and do not support controlled delegation to offline readers.

In the last chapter, we introduced a scheme to reduce reader work in private mutual authentication from linear to logarithmic in the number of tags. In contrast to the pseudonym scheme found in this chapter, this scheme requires at least 3 and possibly as many as $O(\log N)$ rounds of communication between tag and reader, while we achieve one message from tag to reader. Further, our private mutual authentication scheme does not support delegation, nor does it work with legacy readers.

Ohkubo et al. introduce a scheme for RFID pseudonyms [69]. Recovering the tag identity requires work linear in the number of possible tags, while we achieve logarithmic work. The authors propose storing the expected next output of each RFID tag as an optimization, but this cannot be kept up to date unless the trusted authority is online for every tag read. Avoine and Oechslin propose a time-space tradeoff technique that improves the complexity of the Ohkubo et

al. protocol to $O(N^{\frac{2}{3}})$ time with a table of size $O(N^{\frac{2}{3}})$, but their protocol does not support delegation as ours does [10].

Juels gives a scheme for one-use RFID pseudonyms [41]. Unlike our protocol, Juels's scheme does not require a PRF; a simple XOR is enough. Juels also discusses methods for rotating and changing pseudonyms to support ownership transfer. His protocol, however, only allows a tag to emit a limited number of pseudonyms before it must be refreshed through interaction with a trusted reader; Juels outlines an extension to the protocol using a PRG which removes this restriction, but this method requires tight synchronization between reader and tag. Further, his protocol does not work with legacy readers, and it does not support delegation as ours does.

Avoine and Oechslin introduce a method of precomputation for the pseudonym scheme of Ohkubo et al [10]. The benefit of their approach is that no privacy is lost under tag compromise for non-compromised tags, while the cost to identify a tag given the precomputation is modest. Avoine, Dysli, and Oechslin analyze the time requirements for the reader in more detail [8].

Unfortunately, Juels notes that the precomputation scheme itself can imperil privacy: the precomputation requires setting an upper bound on the number of times a single tag is queried. Once any single tag passes this bound, an adversary can distinguish that tag from others by observing that the tag is no longer identified by the system [42]. While Avoine et al. observe that the precomputation approach provides similar performance to our scheme, their analysis sets the upper bound on tag queries to 128. As a result, their approach must be re-evaluated for RFID deployments that may read tags often; for example, some RFID technologies may read a tag as much as 50 times per second.

TC State:
$H : \{0,1\}^{\leq d_1} \to K$, a function.

Algorithm TC.GenTC():
1. Let $H : \{0,1\}^{\leq d_1} \to K$ be a random function,
   i.e., pick $H(s) \in_R K$ uniformly at random
   for each bitstring s of length at most d_1.

Algorithm TC.EnrollTag(ID):
1. Find the smallest integer $t \in \{0,1\}^{d_1}$ that hasn't been assigned to any other tag.
   Assign t to this tag.
2. Set $S := \{t_{1..j} : 0 \leq j \leq d_1\}$.
3. Return $(t\ 0^{d_2}, H|_S)$ as the state for this tag.

Algorithm TC.Delegate(L, R):
1. Let S be the minimal set such that
2. 1) $\forall x \in [L,R], \exists s \in S$ such that s is a prefix of x
3. 2) $\forall s \in S, \forall t \in \{0,1\}^d$, if s is a prefix of t then $t \in [L,R]$.
4. Return $H|_S$.

Algorithm TC.IdentifyTag(r, p):
1. Return DFS($r, p, 1, \epsilon$), where ϵ denotes the empty bitstring.

Algorithm DFS(TK, i, s):
1. Set $ids := \emptyset$.
2. Parse TK as $(v_1, \ldots, v_\ell)$
3. If $F_{H(s\,0)}(r) = p_i$ then
4. If $i \geq \ell$ then return $s\,0$
5. else set $ids := ids \cup \text{DFS}(i+1, s\,0)$.
6. If $F_{H(s\,1)}(r) = p_i$ then
7. If $i \geq \ell$ then return $s\,1$
8. else set $ids := ids \cup \text{DFS}(i+1, s\,1)$.
9. Return ids.

Figure 6.5: Algorithms and state for the Trusted Center.

Reader State:
$h : S \to K$, for some $S \subseteq \{0,1\}^{\geq d_1}$, with S initialized to $\emptyset$.

Algorithm Reader.IdentifyTag($r, p = (p_1, .., p_d)$):
1. For each $s \in S$, do:
2. If $F_{h(s)}(r) = p_{\text{len}(s)}$, then return s.
3. return $\perp$.

Figure 6.6: Algorithms and state for the Reader.

NextNonce.Initialize()
1. Initialize `ctr` to 0
2. Pick secret key $s \in_R K$
NextNonce.GetNextNonce()
1. Return $F_s(\texttt{ctr++})$.

Figure 6.7: Generating nonces with a PRF and a counter.

Number of Tags	Tag Storage	Communication	Tag Computation	Reader Computation
2^{20}	128 bits	158 bits	6	$6 \cdot 2^{10}$
2^{30}	192 bits	168 bits	7	$7 \cdot 2^{10}$
2^{40}	256 bits	178 bits	8	$8 \cdot 2^{10}$

Figure 6.8: Concrete resource use of our scheme for some example parameters. We use a branching factor of 2^{10} in all cases, use a 64-bit r value with truncation, and we assume tags will be read at most 2^{20} times. Tag and reader computation are both measured in expected number of PRF evaluations.

Chapter 7

Conclusions

7.1 Open Problems

7.1.1 Forward Privacy in Log-Work Pseudonyms

The RFID pseudonym protocol of Ohkubo, Suzuki, and Kinoshita [69], and the extension due to Avoine and Oeschlin [10] provide a *forward privacy* property not enjoyed by our tree schemes: if the tag is compromised at a time t, an adversary is unable to link readings of the same tag seen at times $t' < t$. The straightforward attempt to obtain this property in our schemes is to have secrets update by being hashed after each time period.

Unfortunately, this leads to reader work linear in the number of possible time periods, as the reader does not know at which time period the tag is. Releasing the time period "in the clear" addresses this performance issue, but it also allows an adversary to distinguish different tags. Is there a pseudonym or private authentication scheme with work logarithmic in both time periods and number of tags?

7.1.2 Resistance to Tag Compromise

The pseudonym scheme of Ohkubo et al. provides privacy for non-compromised tags if a tag is compromised [69]. In contrast, privacy of other tags degrades in our tree schemes if a tag is compromised; this is because tags share some keying material. This sharing of key material appears essential to our approach for obtaining logarithmic work for the reader. Is there a scheme for private authentication or RFID pseudonyms with logarithmic work but in which tags are keyed independently of each other and so lose no privacy under tag compromise? If not, what is the minimum loss of privacy under tag compromise for a given reader efficiency?

7.2 Future Directions

Unlike the previous section, this section discusses broader research directions. Here the problems are less clearly defined, and the scope of the work required is broader.

7.2.1 Working With RFID Limitations

Both the deployments considered in this thesis use 13.56 MHz RFID technology, which has less stringent limitations than the 915 MHz tags used in the supply chain setting. For example, the cheap tags used by Wal-Mart are unlikely to support a PRF in the near future, and so unlikely to make use of the techniques we have shown. In discussions with supply chain RFID manufacturers, however, we learned two further surprising limitations of 915 MHz supply chain tags.

1. Writing long-term state during a read, even a counter, is difficult. Not enough power is transmitted to an RFID tag at long range to make writing state practical.

2. Generating randomness is hard.

Taken seriously, these limitations mean we should prefer stateless, deterministic RFID security protocols. Unfortunately, this rules out standard notions of cryptographic security, such as indistinguishability of encryptions. How real are these limitations? Can we work around them with weaker, but still useful definitions of security? Juels has taken a step in this direction with an adversarial model capturing "refreshes" of RFID tags by legitimate readers, but more work remains.

7.2.2 Database Integration

An RFID reader and tag is only part of the story. Behind every RFID reader will be a massive database responsible for collecting and managing the observed RFID data[1]. How should this data be managed? In particular, if RFID pseudonyms are used, when should the pseudonyms be mapped back to the real IDs of the tags? Here there are several tradeoffs between efficiency and security. For a concrete example, we could use our delegation property to push subtrees of secrets towards the edge of the network to improve performance. If hardware on the edge is compromised, the number of secrets exposed is limited. On the other hand, if the edge hardware has few secrets, it must communicate more frequently with the Trusted Center.

A database architecture like the HiFi system of Franklin et al. [27] offers a way to manage such tradeoffs. In HiFi, *virtual device drivers,* or VICEs, are responsible for cleaning and post-processing sensor data. One duty of such a VICE might be deciding how to move subtrees of secrets and when to perform

[1]This section results from discussions with the Berkeley HiFi Group led by Michael Franklin, in particular Shawn Jeffery. We are grateful for the helpful discussions.

the mapping from pseudonym to real ID. For more details on VICEs in the HiFi architecture, we refer to Jeffery et al. [40].

7.2.3 Economics of RFID Privacy and Security

Researchers such as Anderson have suggested that economics matters as much or more than technical features in deployments of security technology [6]. We suggest two areas where economic tools may help us analyze RFID privacy and security. First, the incentives for adopting RFID security and privacy features. Second, the effects of "asymmetric information" on RFID privacy.

Adoption

Several RFID privacy techniques require an investment in infrastructure RFID readers to provide privacy. For example, RFID killing requires a reader to be present at point of sale to perform killing. The economic problem here is that the benefit of killing is to the end customer, but the cost of buying the reader for killing falls on the retailer. As a result, the retailer may rationally decide not to buy the RFID reader, therefore weakening the privacy protections supposedly offered by killing. Similar issues arise with recoding.

For example, consider a retailer, such as a family-owned convenience store, that decides that the cost of an RFID reader is not worth the benefits in terms of store inventory. We call such a retailer a "sub-threshold" retailer. If live RFID tags are delivered to a sub-threshold retailer, these tags will "leak" into the population because the sub-threshold retailer lacks the means to kill them. Because end users will not have RFID readers in the near term, such leaked tags may not be detected or killed for a long time, if ever. These leaked tags may raise privacy issues not only for the original purchaser, but also for anyone that later receives the item.

Our RFID pseudonym scheme and private authentication scheme avoid the problem of sub-threshold retailers, only to run into another problem. Specifically, while these schemes do not need infrastructure readers, they do require RFID tags that support a PRF. While RFID tags with a PRF are available in some applications, this raises the cost of the tags. The extra capital cost compared to a "dumb" tag is easy to quantify, while the economic benefit of using our schemes is harder to show on the bottom line. Who will absorb this cost?

We note that RFID technology as a whole, independent of any security or privacy features, still faces significant questions regarding the economics of its adoption. While Wal-Mart and the U.S. Department of Defense have had success in mandating suppliers to adopt RFID, other organizations have had trouble pushing the technology. We suspect some of these questions may become less important as the cost of the technology drops, but others, such as the uneven split between parties regarding costs and benefits, will continue for the foreseeable future.

The RFID Lemons Market

Economics has been used as a tool to analyze privacy behavior in general, and with respect to web privacy issues in particular. Acquisti gives an overview of work in this area [2]. We now give an example of such techniques used to analyze the "kill command" proposed for RFID privacy. Our analysis closely follows the lemons market for web privacy described by Vila, Greenstadt, and Molnar [82].

Consider the following "RFID killing game." In this game, a Customer buys RFID-tagged items from a Merchant. The RFID technology we consider is that of the EPCglobal specifications, in which only a "kill" command is available. We will assume that the Customer does not possess an RFID reader and so cannot check for herself that an RFID tag has actually been killed. We further assume that the Customer has no post point-of-sale uses for the RFID tag and so would always prefer that the RFID tag be killed at point of sale.

The Customer has two options: either to Buy an item or Don't Buy an item. The Merchant also has two options: either to Respect the customer's privacy by killing the RFID tag or to Defect and fail to kill the tag. Now we need to set up the incentives for the two parties. Let us say the payoff Customer receives if it decides to Buy is 1 and the RFID tag is killed, while it receives -1 if the RFID tag is not killed. This models the negative impact to the Customer of having an RFID tagged item. Let the payoff the Merchant obtains for the item be 2 if the RFID tag is not killed, but only 1 if the RFID tag is killed. This models the cost of buying an RFID reader and changing the buying process, plus possibly a benefit to the Merchant from an unkilled tag. If the Customer chooses Don't Buy, neither party receives any benefit. The payoff matrix is then the following:

$$\begin{pmatrix} & \textbf{Respect} & \textbf{Defect} \\ \textbf{Buy} & 1,1 & -1,2 \\ \textbf{Don't Buy} & 0,0 & 0,0 \end{pmatrix}$$

Given this payoff matrix, is the Merchant's optimal strategy to Respect, to Defect, or some mix of the two? What about the Customer's optimal strategy? It turns out that the RFID killing game is an example of Akerlof's "market for lemons," in which the problem of *asymmetric information* leads to an undesirable outcome. Asymmetric information here means that while the Merchant may know whether it will Respect or Defect, the Customer has no way to verify which decision the Merchant has made. The Customer does know, however, that the Merchant has an incentive to Defect, as then the Merchant earns a payoff of 2, which is greater than the payoff 1 the Merchant earns when it Respects. As a result, the equilibrium point for the killing game is to make no trade at all. This pessimistic result does not mean the kill command is useless. It does mean that the kill command alone is not sufficient to provide RFID privacy. Some additional mechanism is needed.

We can go further and model the effect of such a mechanism. Spence introduced the notion of a *signal* in a market, which is a device with a high cost for Defecting sites but a low cost for Respecting merchants [75]. The Customer

can then prefer a Merchant that sends the signal over one which does not. The key question here is finding a real-world signal that exhibits the needed cost difference. Molnar, Soppera, and Wagner discuss an architecture for a "trusted RFID reader [57]." Their design leverages the remote attestation primitive proposed by the Trusted Computing Group to allow anyone to check that the reader follows a particular privacy policy. This check functions as a signal in the lemons market: if the hardware is tamper-resistant, then Respecting sites need do nothing, while Defecting sites who wish to send the signal must pay the cost to break the tamper-resistance of the hardware. Estimating the cost to break tamper-resistant hardware, however, becomes crucial, because this cost determines how useful such hardware may be as a signal[2].

For another approach, Vila, Greenstadt, and Molnar extend the lemons market with a "cost of testing." Here, the cost of testing can be thought of as a cost for the Customer to check that the RFID tag has in fact been killed [82]. Practically, represents the cost of buying an RFID reader and using the reader to check an item after sale. They show that unless the cost of testing is zero, the game tends to a mixed equilibrium in which some merchants Respect and others Defect. They also showed how the cost of testing in the web site privacy setting can be manipulated by unscrupulous web sites that employ purposely obfuscated privacy policies. In the RFID setting, this might correspond to the percentage of Merchants that install RFID readers.

7.2.4 Sensor Network Applications

Both our private authentication and pseudonym schemes can be used by wireless sensor nodes, such as the Berkeley Mica2 and Telos platforms. In particular, our pseudonym scheme can be used to identify a secret key used by a sensor node to encrypt its packets. Our delegation functionality then enables giving third parties limited time access to sensor readings. Are there applications of sensor networks that need these capabilities? Can we use recursive delegation to limit the impact of a base station compromise in a sensor network? Carrying through this direction requires analyzing concrete examples of sensitive sensor network applications, just as we have done for RFID deployments.

7.3 Concluding Remarks

7.3.1 Research and Politics

In both the library and e-passport settings, the decision to use RFID has become heavily politicized. For example, the announcement by the San Francisco Public Library that it would issue a Request for Proposal for an RFID tagging system triggered protests by the Electronic Frontier Foundation and citizens' advocacy groups. Recently, similar protests began in Berkeley following the adoption

[2]We thank Nikita Borisov for raising the question of the exact cost to break tamper-resistant hardware.

by the Berkeley Public Library of an RFID system. Electronic passports have received attention from a broad range of civil liberties groups, including the American Civil Liberties Union, the Electronic Frontier Foundation, and the Electronic Privacy Information Center, all of whom have written white papers or public comments opposing the use of RFID in passports.

In both cases, the issues involved ranged beyond privacy and security of RFID. For example, there is the simple issue of cost in an RFID system: for a library this can run into the tens of thousands of dollars, while the return on investment and benefits of the technology are still being established. In the case of Berkeley Public Library, the cost issue is exacerbated by a budget crisis, which threatened the first staff layoffs in decades. For another example, health concerns over the long-term effect of exposure to radio fields used for RFID played a major role in the debate at San Francisco Public Library, but it is outside our expertise and the scope of our work.

We were also surprised at how the current debate over RFID in San Francisco Public Library tied into long-running disputes between the library staff and citizens' advocacy groups. As an illustration, a leader of one such group had been a self-described library "gadfly" for over seven years, well before the advent of library RFID. Nevertheless, the concerns over RFID privacy made our research directly relevant to these political debates.

Two major issues arose for us because of this political connection. The first is *political engagement* - how do we, as researchers, contribute to the decision-making process surrounding an RFID deployment? For example, a surprise of the work on library RFID was the amount of time spent meeting with library staff, library activists, and in library commission meetings. Finding the right balance between generating new research and disseminating research results has been a challenge. Another challenge in this area has been describing the results of our research to members of the general public, library staff, and decision-makers, few of whom have a background in computer science or computer security.

The second issue is *credibility and partiality*. Do we appear to be unduly favoring one side or the other of the debate? The danger here is that by appearing partial, we may compromise our credibility and hence our opportunity to make an impact. We have been inspired here by Blaze's account of his experiences in the key escrow and export control debates of the 1990s, and his advice to avoid partiality by focusing on questions of technical possibility and impossibility [14].

As we noted in the Introduction, other deployments of RFID have led to political debate, boycotts, and demonstrations. Several states, including California, have also introduced bills to regulate one aspect or another of RFID technology. Therefore, we expect research into RFID security and privacy will continue to have direct political relevance. Researchers working in the area should be prepared for these debates.

7.3.2 The Road To Impact

Beyond the issues of politics and RFID, there is a larger question: how can research into RFID security and privacy make an impact on RFID practice? At the end of the day, this is a question of supply and demand. Before a particular feature will be used, customers of RFID need to demand it, and manufacturers of RFID need to supply it.

Our research is relevant here because we have given an analysis of RFID privacy threats and proposed new features for RFID privacy. Before these features can be used, however, a large amount of work remains to take them from paper to practice. It is not clear how much of this work can be done by academic researchers and how much can be done by commercial entities, such as RFID manufacturers.

7.3.3 The Bottom Line

Several large-scale deployments of RFID devices already exist, and more are on the way. Unless we address the security and privacy issues in RFID now, we will find ourselves locked into vulnerable RFID technologies. We have made concrete progress on technical issues in building better RFID architectures, but more work remains in the technical, political, business, and social realms.

Bibliography

[1] 3M. eTattler newsletter, January 2004. http://cms.3m.com/cms/US/en/0-257/kkruuFX/viewimage.jhtml.

[2] Alessandro Acquisti. The economics of privacy. resource web site., 2005. http://www.heinz.cmu.edu/~acquisti/economics-privacy.htm.

[3] Andy Adler. Sample images can be independently restored from face recognition templates, June 2003. http://www.site.uottawa.ca/~adler/publications/2003/adler-2003-fr-templates.pdf.

[4] U.S. Social Security Administration. Passports as evidence, 2005. http://policy.ssa.gov/poms.nsf/lnx/0302640050?OpenDocument&Click=.

[5] William Aiello, Steven M. Bellovin, Matt Blaze, Ran Canetti, John Ioannidis, Angelos D. Keromytis, and Omer Reingold. Just fast keying: Key agreement in a hostile internet. *ACM Trans. Inf. Syst. Secur.*, 7(2):242–273, 2004.

[6] Ross Anderson. Why information security is hard : An economic perspective. In *Applied Computer Security Applications Conference*, 2001. http://www.ftp.cl.cam.ac.uk/ftp/users/rja14/econ.pdf.

[7] The American Library Association. State privacy laws regarding library records, 2005. http://www.ala.org/Template.cfm?Section=stateifcinaction&Template=/ContentManagement/ContentDisplay.cfm&ContentID=14773.

[8] G. Avoine, E. Dysli, and P. Oechslin. Reducing time complexity in RFID systems. In *Selected Areas in Cryptography (SAC)*, 2005.

[9] Gildas Avoine. RFID privacy: A multilayer problem. In *Financial Cryptography*, 2005.

[10] Gildas Avoine and Philippe Oechslin. A scalable protocol for RFID pseudonyms. In *IEEE PerSec*, 2004.

[11] Claude Barral, Jean-Sébastien Coron, and David Naccache. Externalized fingerprint matching. In *ICBA*, pages 309–315, 2004.

[12] Claude Barral, Jean-Sébastien Coron, David Naccache, and Cedric Cardonnel. Biometric identification method and device adapted to verification on chip cards, patent US2005011946. http://v3.espacenet.com/textdoc?DB=EPODOC&IDX=US2005011946.

[13] BBC. Grins banned from passport pics, 2004. http://news.bbc.co.uk/2/hi/uk_news/politics/3541444.stm.

[14] Matt Blaze. Loaning your soul to the devil: Influencing policy without selling out, 2001. Usenix Security 2001 Invited Talk.

[15] Richard Boss. Library RFID technology. *Library Technology Reports*, Nov/Dec 2003.

[16] German BSI. Alternative extended authentication, 2005.

[17] Vinod Chachra and Daniel McPherson. Personal privacy and use of RFID technology in libraries, October 2003. http://www.vtls.com/documents/privacy.pdf.

[18] EPCGlobal Consortium. EPC ISM Band 13.56MHz Class 1 candidate recommendation, 2004. http://www.epcglobalinc.org/standards_technology/Secure/v1.0/HF-Class1.pdf.

[19] J.S. Coron, D. Naccache, and J. Stern. On the security of RSA padding. In *Advances in Cryptology (CRYPTO)*, 1999.

[20] U.S. State Department. Abstract of the concept of operations for integration of contactless chip in the US passport, 2004. http://www.statewatch.org/news/2004/jul/us-biometric-passport-original.pdf.

[21] Digicert. Digicert PKI toolkit — dcTools specification sheet, 2005. http://www.digicert.com.my/toolkits.htm.

[22] FBI Counterterrorism Division. FBI intelligence memo no. 102, December 2002. http://cryptome.quintessenz.org/mirror/fbi-almanacs.htm.

[23] Charles Doyle. Libraries and the USA PATRIOT Act. Congressional Research Service Report., 2003.

[24] Phillips Electronics. ICode SLI data sheet, 2004. http://www.semiconductors.philips.com/acrobat/other/identification/sl2ics20-fact-sheet.pdf.

[25] Martin Feldhofer, Sandra Dominikus, and Johannes Wolkerstorfer. Strong authentication for RFID systems using the AES algorithm. In Marc Joye and Jean-Jacques Quisquater, editors, *Workshop on Cryptographic Hardware and Embedded Systems (CHES)*, volume 3156 of *Lecture Notes in Computer Science*, pages 357–370, Boston, Massachusetts, USA, August 2004. IACR, Springer-Verlag.

[26] Kenneth Fishkin and Sumit Roy. Enhancing RFID privacy through antenna energy analysis. In *MIT RFID Privacy Workshop*, 2003. `http://www.rfidprivacy.org/papers/fishkin.pdf`.

[27] M. Franklin, S. Jeffery, S. Krishnamurthy, F. Reiss, S. Rizvi, E. Wu, O. Cooper, A. Edakkunni, and W. Hong. Design considerations for high fan-in systems: The HiFi approach. In *CIDR Conference*, 2005.

[28] Simson Garfinkel. Adapting fair information practices to low cost RFID systems. In *Privacy in Ubiquitous Computing Workshop*, 2002. `http://www.simson.net/clips/academic/2000_Ubicomp_RFID.pdf`.

[29] Craig Gentry and Zulfikar Ramzan. Personal communication, 2004.

[30] Carol Glatz. Vatican library begins using computer chips to identify volumes. catholic news. march 29., 2004. `http://www.catholicnews.com/data/stories/cns/20040329.htm`.

[31] Oded Goldreich, Shafi Goldwasser, and Silvio Micali. How to construct random functions. *J. ACM*, 33(4):792–807, 1986.

[32] Nathan Good, John Han, Elizabeth Miles, David Molnar, Deirdre Mulligan, Laura Quilter, Jennifer M. Urban, and David Wagner. Radio frequency id and privacy with information goods. In *Workshop on Privacy in the Electronic Society - WPES*, 2004. `http://www.cs.berkeley.edu/~daw/papers/rfid-wpes04.pdf`.

[33] Lukas Grunwald. RF-DUMP, 2004. `http://www.rfdump.org`.

[34] Tom Halfhill. Is RFID paranoia rational?, 2005. `www.maximumpc.com/reprints/reprint_2005-01-14a.html`.

[35] Gerhard Hancke. A practical relay attack on ISO 14443 proximity cards, 2005. `http://www.cl.cam.ac.uk/~gh275/relay.pdf`.

[36] Gerhard Hancke. Practical attacks on proximity identification systems (short paper). In *IEEE Symposium on Security and Privacy*, 2006.

[37] ICAO. Document 9303, machine readable travel documents, October 2004.

[38] International Civil Aviation Organization (ICAO). PKI for machine readable travel documents offering ICC read-only access, version 1.1, October 2004.

[39] ISO. ISO/IEC 9797-1 algorithm 3, 1999.

[40] Shawn R. Jeffery, Gustavo Alonso, Michael J. Franklin, Wei Hong, and Jennifer Widom. Virtual devices: An extensible architecture for bridging the physical-digital divide, 2005. Under submission.

[41] Ari Juels. Minimalist cryptography for RFID tags. security in communications networks (SCN). c. blundo, ed., 2004. http://www.rsasecurity.com/rsalabs/staff/bios/ajuels/publications/minimalist/index.html.

[42] Ari Juels. RFID security and privacy: A research survey. *IEEE Journal on Selected Areas in Communications (J-SAC)*, 2006. http://www.rsasecurity.com/rsalabs/node.asp?id=2937.

[43] Ari Juels, David Molnar, and David Wagner. Security and privacy issues in e-passports. Cryptology ePrint Archive, Report 2005/095, 2005. http://eprint.iacr.org/.

[44] Ari Juels, Ronald L. Rivest, and Michael Szydlo. The blocker tag: selective blocking of RFID tags for consumer privacy. In *Proceedings of the 10th ACM Conference on Computer and Communication security (CCS)*, pages 103–111. ACM Press, 2003.

[45] Dato' Mohd Jamal Kamdi. The Malaysian electronic passport, 2004. Presentation to ICAO, http://www.icao.int/icao/en/atb/fal/fal12/Presentations/Malaysia.ppt.

[46] Gaurav S. Kc and Paul A. Karger. Preventing security and privacy attacks on machine readable travel documents (MRTDs). Cryptology ePrint Archive, Report 2005/404, 2005. http://eprint.iacr.org/.

[47] Ziv Kfir and Avishai Wool. Picking virtual pockets using relay attacks on contactless smartcard systems. Cryptology ePrint Archive, Report 2005/052 and IEEE SecureComm 2005, 2005.

[48] Shingo Kinosita, Fumitaka Hoshino, Tomoyuki Komuro, Akiko Fujimura, and Miyako Ohkubo. Non-identifiable anonymous-ID scheme for RFID privacy protection, 2003. In Japanese. See English description as part of http://www.autoidlabs.com/whitepapers/KEI-AUTOID-WH004.pdf.

[49] Ilan Kirschenbaum and Avishai Wool. How to build a low-cost, extended-range RFID skimmer. In *Usenix Security*, 2006.

[50] Libramation. Overview of library RFID products, 2004. http://www.libramation.com/prod_radio.html.

[51] Berkeley Public Library. Best practices for library RFID, 2004. http://berkeleypubliclibrary.org/BESTPRAC.pdf .

[52] San Francisco Public Library. Checkout time-motion study, 2004. Presented to SFPL Library Commission, April 4.

[53] Elena Malykhina. RFID tests are positive for CVS and pharmaceuticals. Information Week, http://informationweek.com/story/showArticle.jhtml?articleID=48800464.

[54] David L. Margulius. UPS pilots an RFID rollout. InfoWorld, http://www.infoworld.com/article/04/04/09/15FErfidcase_1.html.

[55] Tsutomu Matsumoto. Gummy and conductive silicone rubber fingers. In *ASIACRYPT*, 2002.

[56] D. McCullugh. House backs major shift to electronic IDs. *CNET News*, 10 February 2005. http://news.zdnet.com/2100-9595_22-5571898.html.

[57] David Molnar, Andrea Soppera, and David Wagner. RFID privacy through trusted computing. In *Workshop on Privacy in the Electronic Society (WPES)*, 2005. http://www.cs.berkeley.edu/~daw/papers/wpes05-camera.pdf.

[58] David Molnar, Andrea Soppera, and David Wagner. A scalable, delegatable pseudonym protocol enabling ownership transfer of RFID tags. In *Selected Areas in Cryptography (SAC)*, 2005. http://eprint.iacr.org/2005/315/.

[59] David Molnar, Ross Stapleton-Gray, and David Wagner. Killing, recoding, and beyond, 2005. Chapter in RFID Applications, Security, and Privacy, Simson Garfinkel and Beth Rosenberg eds., July 2005, Addison/Wesley.

[60] David Molnar and David Wagner. Security and privacy in library RFID: Issues, practices, and architectures. In *ACM Cryptography and Communications Security (CCS)*, 2004.

[61] Moni Naor and Omer Reingold. Synthesizers and their application to the parallel construction of pseudo-random functions. *J. of Computer and Systems Sciences*, 58(2):336–375, April 1999.

[62] Will Ness. Microsoft optical desktop comes with fingerprint reader. cooltechzone.com, January 2005. http://www.cooltechzone.com/index.php?option=content&task=view&id=891&Itemid=0.

[63] Yasunobu Nohara, Sozo Inoue, Kensuke Baba, and Hiroto Yasuura. Quantitative evaluation of unlinkable ID matching schemes. In *Workshop on Privacy in the Electronic Society (WPES)*, 2005.

[64] The State of Alabama. Code of Alabama section 41-8-10, 2004.

[65] The State of California. Cal. gov. code section 6267, 2004.

[66] The State of Illinois. 75 ILCS 70/1, 2004.

[67] The State of New York. NY CLS CPLR Section 4509, 2004.

[68] Department of State. Department of State, 22 CFR Part 51, Public Notice 4993, RIN 1400-AB93, Electronic Passport. *Federal Register*, 70(33), 18 February 2005. Action: Proposed Rule. Available at `http://a257.g.akamaitech.net/7/257/2422/01jan20051800/edocket.access.gpo.gov/2005/05-3080.htm`.

[69] Miyako Ohkubo, Koutarou Suzuki, and Shingo Kinoshita. Cryptographic approach to a privacy friendly tag. In *RFID Privacy Workshop, MIT*, 2003.

[70] Neville Pattinson. Securing and enhancing the privacy of the e-passport with contactless electronic chips, 2004. International Association of Privacy Professionals, `https://www.privacyassociation.org/index.php?option=com_content&task=view&id=272&Itemid=125`.

[71] Zulfikar Ramzan. Personal communication, 2004.

[72] Mark Reynolds. Physics of RFID. In *MIT RFID Privacy Workshop*, 2003. `http://www.rfidprivacy.org/papers/physicsofrfid.ppt`.

[73] Riscure Security. Privacy issues with new digital passport, July 2005. `http://www.riscure.com/news/passport.html`.

[74] R. Singel. No encryption for e-passports. *Wired News*, 24 February 2005. `http://www.wired.com/news/privacy/0,1848,66686,00.html?tw=wn_tophead_1`.

[75] Michael Spence. Job market signalling. The Quarterly Journal of Economics, pages 355-75, 1973.

[76] Ross Stapleton-Gray. Would Macy's scan Gimbels? Competitive intelligence and RFID, 2003. `http://www.stapleton-gray.com/papers/ci-20031027.PDF`.

[77] The United States. Wiretap act as amended by electronic communications privacy act, 18 uscs section 2510(2), 2000.

[78] Justice Douglas Supreme Court of the United States. Griswold v. Connecticut, 381 US 479, 482.

[79] U.S. State Department "Architecture Team". IC embedded passport PKI requirements, 20 October 2003. Available at http://www.aclu.org/passports/PKIRequirements.pdf.

[80] Lee Tien. Privacy risks of radio frequency identification "tagging" of library books, October 2003. `http://www.eff.org/Privacy/Surveillance/RFID/20031002_sfpl_comments.php`.

[81] Phillip Torrone. Make magazine blog: DEFCON RFID world record attempt, 2005. http://www.makezine.com/blog/archive/2005/07/_defcon_rfid_wo.html.

[82] T. Vila, R. Greenstadt, and D. Molnar. Why we can't be bothered to read privacy policies: Models of privacy economics as a lemons market. In *3rd Workshop on Economics and Information Security*, 2003.

[83] Tamas Visegrady. Personal communication, 2005. tvi@zurich.ibm.com.

[84] Stephen A. Weis, Sanjay E. Sarma, Ronald L. Rivest, and Daniel W. Engels. Security and Privacy Aspects of Low-Cost Radio Frequency Identification Systems. In *Security in Pervasive Computing*, volume 2802 of *Lecture Notes in Computer Science*, pages 201–212, 2004.

[85] Rick Whiting. Drug makers 'jumpstart' tagging of bottles, July 2004. Information Week, http://www.informationweek.com/story/showArticle.jhtml?articleID=25600213&tid=5978.

[86] Staff writer. New-look passports. *Economist*, 17 February 2005. http://economist.com/science/displayStory.cfm?story_id=3666171.

[87] Chas Hock Eng Yap and Foong Mei Chua. U.S. patent 6,111,506 method of making an improved security identification document including contactless communication insert unit, 2000.

[88] Jimmy Yap. Are we ready for radio? April 5. MIS Asia, 2004. http://www.smh.com.au/articles/2004/04/05/1081017086098.html.

[89] Junko Yoshida. Tests reveal e-passport security flaw, August 2004. *EE Times*.

[90] David Zhang and Anil K. Jain, editors. *Biometric Authentication, First International Conference, ICBA 2004, Hong Kong, China, July 15-17, 2004, Proceedings*, volume 3072 of *Lecture Notes in Computer Science*. Springer, 2004.