Light-Weight Authentication Schemes

With Applications To RFID Systems

by

Behzad Malek

Thesis submitted to the

Faculty of Graduate and Postdoctoral Studies

In partial fulfillment of the requirements

For the Ph.D. degree in

Electrical and Computer Engineering

School of Information Technology and Engineering

Faculty of Engineering

University of Ottawa, Ontario, Canada

Abstract

The first line of defence against wireless attacks in Radio Frequency Identification (RFID) systems is authentication of tags and readers. RFID tags are very constrained in terms of power, memory and size of circuit. Therefore, RFID tags are not capable of performing sophisticated cryptographic operations. In this dissertation, we have designed light-weight authentication schemes to securely identify the RFID tags to readers and vice versa. The authentication schemes require simple binary operations and can be readily implemented in resource-constrained Radio Frequency Identification (RFID) tags. We provide a formal proof of security based on the difficulty of solving the Syndrome Decoding (SD) problem. Authentication verifies the unique identity of an RFID tag making it possible to track a tag across multiple readers. We further protect the identity of RFID tags by a light-weight privacy protecting identification scheme based on the difficulty of the Learning Parity with Noise (LPN) complexity assumption. To protect RFID tags authentication against the relay attacks, we have designed a resistance scheme in the analog realm that does not have the practicality issues of existing solutions. Our scheme is based on the chaos-suppression theory and it is robust to inconsistencies, such as noise and parameters mismatch. Furthermore, our solutions are based on asymmetric-key algorithms that better facilitate the distribution of cryptographic keys in large systems. We have provided a secure broadcast encryption protocol to efficiently distribute cryptographic keys throughout the system with minimal communication overheads. The security of the proposed protocol is formally proven in the adaptive adversary model, which simulates the attacker in the real world.

Acknowledgements

This dissertation would not have been possible without the guidance and the help of several individuals who in one way or another contributed and extended their valuable assistance in the completion of this study.

First and foremost, my utmost gratitude goes to Professor Ali Miri at the School of Information Technology and Engineering, University of Ottawa. His encouragement and assistance at the most difficult times have always helped me get through life and overcome the obstacles in the completion of this research work. He has been a great friend and a wise mentor. I appreciate all his contributions of time, ideas and funding to make my Ph.D. experience productive and enjoyable.

My colleagues and staff in the Computational Lab in Coding and Cryptography (CLiCC) have been a great source of motivation and exchange of ideas. I am especially thankful for the friendship and support of original CLiCC group members: Saeed Samet, Wilson Poon, Ilker Onat and Augusto Lima. My time at the University of Ottawa was made enjoyable in large part due to the many friends that became a part of my life. I am grateful for the time spent with my best friends and roommates, Assal Abdossalehi and Shahrooz Asadi, and for the great memories we shared together.

Lastly, I would like to thank my family for all their love and encouragement, especially my parents, Mahasti and Hossein, who relentlessly supported me in my greatest endeavors. I am sincerely thankful to my sisters, Mitra and Sogol, and brother, Babak, for their moral support and wisdom throughout my life. Thank you all.

Contents

List of Figures

List of Tables

List of Acronyms

Admin Administrator 104

AES Advanced Encryption Standard 20

BDHE Bilinear Diffie-Hellman Exponent 109

BE Broadcast Encryption 104

BLA Backward Link Authentication 9

CCA Chosen Ciphertext Attack 107

DLP Discrete Logarithm Problem 109

Chapter 1

Introduction

Advancements in technology have enabled mass production of cheap, miniaturized RFID tags that have become rampant in every application ranging from animal/cargo tracking to store item's labeling and payment systems. The main task of RFID tags is wireless identification. That is, an RFID reader (scanner) should be able to find an RFID tag in its reading range and recognize its identity. Every time an RFID tag is being scanned, it responds with an identification code. Depending on the identity that the RFID tag presents to the reader, it will receive a specific service. The service varies greatly from one application to another. It can simply be matching the identity of the tag to a price in a retail store or granting access to a secured door in a building. This basic task puts many security and privacy challenges in front of designers and researchers, in order to build a robust RFID system.

1.1 Identification in RFID Systems

RFID systems communicate via electromagnetic waves and are categorized as radio systems. All radio systems operate in a narrow band to avoid signal interferences with other radio systems. Therefore, available frequencies and transmit power in to every radio system, including RFID systems, are heavily regulated. These regulations and restrictions directly affect an RFID system in its reading range, power, operating frequency and memory.

In general, there are three types of tags: *passive*, *semi-passive* and *active*. Passive tags have no battery on their circuit, and they rely solely on the reader to provide the power for the tag to operate. Passive tags are low-cost and have limited computational and storage resources. Active and semi-passive tags use internal batteries to power their circuits. Semi-active tags can perform more sophisticated operations that passive tags and are more expensive than passive tags. An active tag also uses its battery to broadcast radio waves to a reader, whereas the power to broadcast in semi-passive tags is supplied by the reader. Active tags operate in higher frequencies and have longer reading range than passive and semi-passive tags.

RFID systems operate in different frequency ranges from 135 KHz in the Low Frequency (LF) band to 13.6 MHz range in the High Frequency (HF) band and up to 928 MHz in the Ultra High Frequency (UHF) band [38]. The operating frequency directly impacts the transmission rate and power of the RFID system. Various frequencies impose different restrictions in terms of the speed and amount of data can be transmitted in an RFID system.

All types of RFID tags – passive, semi-passive and active – have some non-volatile memory

to store the identification code and information data. The size of the available memory ultimately determines the price of the RFID tag. Memories vary in capacity from just one byte, in the so-called 'pigeon tag', to 64 kilobytes in microwave tags. The amount of memory available on low-cost RFID tags, such as EPC-Gen 2 tags [46], is very limited (usually less than one kilobytes).

In large scale applications, such as payment cards, passports, cargo tracking, where millions of RFID tags should be deployed, the price of the system is of paramount importance. Currently, the most popular RFID tags in the industry are EPC Gen 2 tags [46]. They are passive, HF tags that have $1K$ bits of memory and are regulated by ISO 14443 standard [39, 40]. Throughout this dissertation, we focus on solutions that can add security and privacy to EPC Gen 2 tags.

1.2 Wireless Attacks in RFID Systems

The RFID system is comprised of multiple components, including the tag, reader and server. Each of these components can be targeted for an attack. The identification data are exchanged between the tag and the reader over a wireless, public link. The identification data transmitted over the wireless link has to be secured against classical wireless attacks including: eavesdropping, replay and relay attacks and tampering attacks.

Eavesdropping: This is the simplest type of attacks, where the attacker is passive, and it only listens to the communication data being exchanged between the tag and the reader. Since, the communication takes place over a wireless channel, it is trivial for the attacker to

capture messages being transmitted in the air. Moreover, it is possible to covertly trace an RFID tag across multiple readers violating privacy of an RFID tag's bearer. Clearly, confidentiality of sensitive data communicated over a public channel has to be protected against eavesdropping attacks.

Replay attack: The attacker in the replay attack launches an eavesdropping attack first attempts first to listen to the data being exchanged between a legitimate tag and the reader. Then, it tries to repeat the data exchange, in order to impersonate as the legitimate tag or the reader. In other words, the attacker tries to use the communication messages exchanged in previous rounds to compromise the new authentication round. To secure RFID tags against the replay attacks, the authentication mechanism has to be randomized, that is every new authentication session has to be different than any previous sessions.

Relay attack: In a slightly different attack than the replay attack, the attacker only intercepts the communication data for a current authentication round and redirects them to legitimate tags or readers. In other words, the tag and the reader remotely authenticate one another, without knowing that the attacker is communicating with the reader via the signals received from a remote tag. If the attacker has relayed the signals promptly, it can covertly surpass the authentication process. It is very challenging to detect a relay attack and to deter it, as the RFID tags immediately respond to any query signal and their response signals can be easily captured in the air. There exist non-cryptographic solutions that safeguard

authentication schemes and protect against the relay attacks in RFID systems.

Tampering: In the tampering attack, the attacker is active. Not only does the attacker eavesdrop on the data exchanged between the tag and the reader, but it also tries to change the data in a meaningful way. A proper change in the communicated data may yield access to the attacker and change the intended result of authentication. For example, the attacker may succeed to change the price of an item from $200 to $20, which is recorded in an RFID tag mounted on the item. Use of cryptographic algorithms can help alleviate tampering with confidential data in RFID tags. However, RFID tags very limited in size, memory, and computation power. Therefore, most classical cryptographic algorithms will not fit into low cost RFID tags, and new light-weight algorithms should be designed.

1.3 Security and Privacy Requirements

As discussed earlier, the main responsibility of RFID tags is reliable and secure identification. The tag has to be able to prove its identity to a reader with a certain level of confidence and trust. Inversely, the reader should be able to authenticate itself to a tag and be trusted by the tag. However, sharing the unique identity of an RFID tag across multiple domains of use creates a privacy concern.

The security of a system is as strong as its weakest link. Since, the RFID system is comprised of many parts including the wireless link, tag, reader and server (computer support system), attacks can be launched at various parts. Therefore, the security and privacy

solution must protect the RFID system at various parts. Due to restriction on the tags' price and size in an RFID system, the main challenge is to design light-weight schemes to meet the security and privacy requirements in passive RFID tags.

1.3.1 Authentication

The very basic goal in the RFID system is identification. The system has to be able to recognize entities carrying RFID tags with them. Depending on the identity of the tag, which is usually represented by a number, the system should take appropriate actions; for example, in a shopping store, the identity of a tag would represent the price, or in a library it would represent a book. Simple identification usually suffices for many industry applications, for example in the library applications. However, in many other applications, the tag or the reader should be able to present a proof of identity (authentication) to the other party.

Authentication is the first line of defence against wireless attacks in an RFID system. Authentication adds trust to the identification procedure and validates an identity to a verifier. In an authentication scheme, the entity being authenticated is referred to as the prover and entity checking the identity is referred to as the verifier. Authentication schemes for RFID systems can be categorized in two main categories based on the prover and verifier: Authenticating the RFID tags to the reader is referred to as Forward Link Authentication (FLA), the prover is the RFID tag and the verifier is the RFID reader. Similarly, authenticating the reader (prover) to the RFID tag (verifier) is referred to as Backward Link Authentication (BLA).

The performance of authentication schemes depends mainly on cryptographic primitives used in the protocol. Various cryptographic primitives have different requirements in terms of the size of secret keys, the mechanism to share keys, the space of randomness, speed and power consumptions, which have to be considered in design of a light-weight authentication scheme for passive RFID tags. Cryptographic solutions are either based on a *symmetric-key* algorithm or an *asymmetric-key* algorithm. In the symmetric-key algorithm, the authentication protocol requires the prover and verifier to share exactly the same cryptographic key. On the other hand, the authentication protocol based on an asymmetric-key algorithm has two sets of keys: a *public key*, which is publicly known to everyone in the system and a corresponding *private key*, which is different from the public key and is known by one owner only (the prover in the authentication protocol). Every two parties in a public key authentication scheme hold a public/private key pair that enables them to establish a secure communication channel. In an asymmetric (public key) cryptosystem, the encryption key is publicly known to everyone, where as the decryption key cannot be (easily) derived from the public encryption key, and it is only known to one entity. In other words, it is computationally impossible to find the decryption algorithm, given only the encryption algorithm in a public key cryptosystem. Some additional secret information, usually referred to as *trapdoor*, must be present to find the decryption key.

In general, authentication schemes based on symmetric-key algorithms are faster and more efficient in hardware implementation than authentication schemes based on asymmetric-key algorithms. Nevertheless, symmetric-key algorithms require the prover and verifier to share

the same secret key over a secure channel that might not exist in practice. Cryptographic keys are updated frequently to remove the compromised keys and to maintain a high security level in the system. In a symmetric cryptosystem, this requires that new secret keys to be communicated regularly with all the key holders over a secure channel. Building such a channel might not be feasible or might be costly after the initial setup. RFID authentication schemes based on symmetric cryptography are useful for closed systems with a few participants. In a closed system, each participant (tags and readers) can get its key and key updates via a secure channel. An example of a closed environment is the car immobilizer, where the symmetric authentication keys can be shared in advance at the manufacturer. Therefore, the manufacturer has the role of the central authority and is responsible for distributing the keys.

In asymmetric authentication systems, the prover and verifier do not need to share the same key. In fact, the verifier and everyone else in the system only know the prover's public key without any knowledge of its private key. Therefore, asymmetric authentication schemes are suitable for large, distributed systems, where it is difficult to establish a secure communication channel to share secret keys. In RFID systems, the challenge is to reduce the complexity of asymmetric authentication systems, in order to achieve an efficient scheme that can be fit into low-cost, passive RFID tags.

1.3.2 Privacy Protecting Identification

All types of RFID tags (passive and active) have non-volatile memory to store data used in identification or authentication protocols. In every identification protocol, an RFID tag is scanned unnoticeably, and it immediately responds with the same identification code. In authentication protocols, the identity of an RFID tag is uniquely verified, and the same tag can be covertly traced across multiple domains of use. Tracing a tag across multiple readers creates great privacy concerns, as it can reveal sensitive information, such as the number of visits to a store, location of the tag's bearer and its shopping preferences. Therefore, the goal is to design a light-weight scheme that identifies (authenticate) an RFID tag to its authorized readers, while tags' identity is kept hidden from readers and cannot be traced over multiple readers. In other words, readers can individually identify (authenticate) an RFID tag without being able to track over another reader. The proposed scheme should still be implementable with little complexity in RFID tags.

1.3.3 Relay Attacks

As discussed earlier, an RFID authentication protocol verifies the unique identity of a prover to its verifier in the RFID system. Nevertheless there are attacks, such as the *relay attacks*, that can circumvent the authentication process without having to break the underlying symmetric or asymmetric cryptosystem. In the relay attack, the attacker simply relays the authentication messages between a legitimate tag and reader, and it does not involve in the authentication process. The tag and the reader remotely authenticate one another without

knowing that the attacker is instead communicating with the reader via the signals received from a remote tag. If the attacker has relayed the signals promptly, it can covertly surpass the authentication process. It is very challenging to detect a relay attack and to deter it, as the RFID tags immediately respond to any query signal and their response signals can be easily captured in the air. Therefore, not only should proper light-weight authentication schemes be designed for RFID tags, but also there has to be a mechanism in place to thwart the relay attacks. Otherwise, the proposed authentication solutions can be easily circumvented in a relay attack.

1.3.4 Key Distribution and Management

In an RFID system, secure distribution and management of cryptographic keys used in the authentication/privacy-protection schemes is a sensitive task, as the entire trust in the system depends on it. In symmetric-key systems, the total number of secret keys used increases quadratically with the total number of parties in the system, since every two parties (the prover and the verifier) must share the same key. In a large system, key management of a symmetric cryptosystem quickly becomes overwhelming, as a central authority has to be devised to track all the keys being shared between various parties. In order to distribute the symmetric keys, one needs to have access to a secure communication channel with every part involved (tag or reader). Establishing a secure communication channel across different authority domains imposes practical problems.

An public-key cryptosystem greatly simplifies the management of cryptographic keys in

the system and allows for securely sharing secret keys over public (non-secure) channels. A Public Key Cryptography (PKC) can expand arbitrarily to any number of tags and readers. A public-key cryptosystem solves the problem of sharing secret keys between two parties only. However in a typical RFID system, a reader is interacting with many tags (probably around thousands) and a tag is usually in contact with a few readers (probably less than five). Efficiently extending the secret-sharing beyond two parties requires non-trivial solutions. *Broadcast encryption* refers to a set of cryptographic tools that use encryption to create a ciphertext that can be decrypted by many parties (i.e. RFID readers) that do not share the same private/public keys. Design of broadcast encryption schemes that securely and efficiently distribute cryptographic keys among more than two entities (RFID readers) is an active research area in cryptography.

1.4 Contributions

We have made contributions to four areas listed earlier in this chapter. The major contributions of this dissertation are listed as follows:

Light-weight Authentication of Tags and Readers: In design of authentication schemes for an RFID system, new restrictions arise mostly from limitations in hardware, and the price of the RFID tags ultimately determines the complexity of algorithms that can be used. We design a light-weight BLA scheme to authenticate readers to the RFID tags in Chapter 3. Complex computational operations are replaced on the RFID tags by simple binary operations on short vectors, making the protocol suitable for most basic tags. We

have submitted our results for publication to the IEEE International Conference on RFID-Technology and Applications (RFID-TA 2011) on September 2011. In Chapter 4, we also design a light-weight (FLA) scheme to authenticate the RFID tags to readers. The proposed scheme is a *zero-knowledge* protocol, where the tags can be easily authenticated by readers without revealing the tag's secret to the readers. Our contributions have been published partly in the Personal, Indoor Mobile Communications (PIMRC) conference on October 2010 and partly in the International Journal of Network Security (IJNS) on November 2010.

Light-weight Privacy Protecting Identification: Robust authentication and private identification of RFID tags appear contradictory. On one hand, the reader has to uniquely identify an RFID tag with a high level of confidence and has to be able to authenticate a tag. On the other hand, RFID tags can be easily traced over multiple readers, as they respond with the same identity. This poses a privacy risk to the bearer of the tag, as tracking can be done without the bearer's consent. In Chapter 5, we have proposed a light-weight privacy protecting identification scheme that can be readily implemented in existing RFID tags. The complexity of adding privacy to the tags is shown to be very little. We show that our scheme is more efficient than other existing privacy protecting schemes. We also provide a formal proof of security for the newly designed scheme. Our results have been accepted for presentation in the Springer Foundations and Practice of Security (FPS 2011) conference on May 2011 and partly submitted to the Journal of Computer Security on June 2010.

Relay Attacks Resistance Scheme: As mentioned before, the relay attack is a very simple, yet effective attack against most authentication/identification schemes. Almost all

countermeasures against the relay attacks rely on the existence of a Ultra Wide Band (UWB) channel and precise timing of messages exchanged in the RFID system. A UWB channel provides a higher resolution in measuring frequency and time. These assumption are mostly difficult to realize in practice, where the size and complexity of RFID tags are minimal. We have taken a new approach and proposed a solution in Chapter 6 based on the chaos-suppression theory in which the analog RFID signals can only be unmasked if signals are emitted from a legitimate reader. In our approach, we apply *chaos suppression* instead of *chaos synchronization* to design a secure communication channel, and our chaos-masking scheme is not reliant on the existence of a UWB and not sensitive to timing discrepancies.

Secure Key Management: In a distributed system, such as the RFID system, cryptographic keys should be distributed to many different parties throughout the system via a secure channel. This is often a challenging task, as many parties (servers, readers and tags) are remotely positioned and they are only accessible via public (insecure) channels. Moreover, parties join and leave the system dynamically. Managing and distributing the keys throughout the system requires a solution that can efficiently establish a secure broadcast channel. The solution has to be adapted to any arbitrary set of readers over different authority domains. In Chapter 7, we have shown how to manage and distribute cryptographic keys to a set of remote users (RFID readers) and propose the first adaptively secure Broadcast Encryption scheme with short ciphertexts. Our results have been accepted (with revision) for publication in the International Journal of Network Security (IJNS) on September 2010.

In summary, the main contributions of this dissertation have been accepted or submitted

for publication in the following venues:

- B. Malek and A. Miri, *Backward Link Authentication for RFIDs*, In the IEEE International Conference on RFID-Technology and Applications (RFID-TA 2011), Barcelona, Spain, September 2011 (submitted).

- B. Malek and A. Miri, *Private Identification of RFID Tags*, In the Springer Foundations and Practice of Security (FPS 2011) conference, Paris, France, May 2011.

- B. Malek and A. Miri, *Forward Link Authentication for RFIDs*, In the IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2010), Istanbul, Turkey, September 2010.

- B. Malek and A. Miri, *Securing Harari's Authentication Scheme*, International Journal of Network Security (IJNS), November 2010.

- B. Malek and A. Miri, *Adaptively Secure Broadcast Encryption with Short Ciphertexts*, International Journal of Network Security (IJNS), September 2010.

- B. Malek and A. Miri, *Privacy Protecting Identification for RFIDs*, IOS Press, Journal of Computer Security, June 2010 (submitted).

- B. Malek and A. Miri, *Authentication Mechanisms for RFID Systems*, Book Chapter, IGI Global, December 2009 (in press).

- B. Malek and A. Miri, *Chaos-Masking for Securing RFIDs against Relay Attacks* (in preparation for submission).

1.5 Organization

In Chapter 2, we provide an overview of existing security solutions for RFID systems. Our light-weight BLA scheme to authenticate RFID readers is given in Chapter 3. In Chapter 4, We develop an FLA scheme to authenticate RFID tags. In Chapter 5, we design a light-weight asymmetric identification scheme that preserves the privacy of RFID tags throughout the system. Chapter 6 includes our results in chaos-masking scheme to counter the relay attack. Key management is addressed in Chapter 7, where we provide the details of our adaptively secure Broadcast Encryption scheme with short ciphertexts. Finally in Chapter 8, we provide conclusions and outlines of the future work.

Chapter 2

Literature Survey

Various cryptographic primitives have different requirements in terms of size of secret keys, key-sharing methods, speed and power consumption. In this chapter, we provide an overview of the state of the art in security solutions for RFID systems. We review a few of well-known, related symmetric and asymmetric schemes that are specifically designed for RFID authentication and private identification in Sections 2.1 and 2.2. The non-cryptographic solutions against relay attacks are given in Section 2.3. In Section 2.4, we list some of the existing related works that securely distribute cryptographic keys in a system.

2.1 Symmetric Authentication Schemes

Many authentication schemes proposed for RFID systems are based on a classical symmetric algorithm or cryptographic hash function [35, 36, 41, 80, 90]. For example, Molnar and Wagner [80] proposed a randomized hash-based scheme to authenticate RFID tags in

a challenge-response protocol. In their scheme, the reader first challenges the tag with a random number r_1. Then, the tag generates another random number r_2. Then, the tag hides its identity ID by sending $\text{ID} \oplus f_k(r_1, r_2)$ to the reader, where f_k is a hash function parameterized with the symmetric key k. The reader and tag should exchange the random numbers r_1 and r_2 before the tag sends $\text{ID} \oplus f_k(r_1, r_2)$. Although most RFID tags are capable of executing cryptographic hash functions, current commercial RFID tags do not implement hash functions inside RFID tags due to higher production costs.

Symmetric-key encryption is more favorable than hash-based authentication, as symmetric-key cryptography supports additional security measures, such as message integrity and confidentiality. In [35], a simple challenge-response authentication scheme based on Advanced Encryption Standard (AES) is proposed. In [35], the reader first sends a random number C_r to the tag. Then, the tag picks a random number as the challenge C_t and includes it in the response message $\mathcal{E}_K(C_t, C_r)$ encrypted using AES encryption function. Using the shared symmetric key K, the reader has to decrypt $\mathcal{E}_K(C_t, C_r)$ to retrieve C_t. The reader authenticates itself by sending $\mathcal{E}_K(C_r, C_t)$ back to the tag. Note that changing the position of C_t and C_r in $\mathcal{E}_K(\cdot, \cdot)$ implies a modification on C_t and C_r combined together. This is usually a simple modification, such as a concatenation or addition of C_t and C_r by a fixed number. The tag uses K to decrypt $\mathcal{E}_K(C_r, C_t)$ and to check the modification of C_t and C_r performed by the reader. If the modification matches with the tag's modification of C_t and C_r, it authenticates the reader, otherwise rejects it. This scheme is considered highly secure and well suited for hardware implementations. Nevertheless, standard cryptographic

algorithms, such as AES, require a great amount of resources (power, memory and space) on the RFID tags. For very sensitive applications, expensive RFID tags equipped with standard symmetric cryptographic algorithms might be justifiable. In routine industrial applications, such as cargo-tracking, where a minimum level of security should be present, expensive tags might not be a viable solution. It should be noted that RFID systems are widespread, mainly because they can cheaply replace barcodes and can provide additional functionalities.

Code-based cryptography is another very efficient and very high speed tool to design light-weight authentication schemes, as compared to algebraic cryptography. Therefore, code-based cryptographic systems are desirable choice for small devices. The first popular code-based system is HB+ by Juels and Weis [60]. HB+ is a modification of the authentication protocol by Hopper and Blum (HB) that was originally designed to detect human users. HB+ is implementable in very small, low-cost RFID tags [46]; it only requires simple binary operations, such as XOR ($\oplus$) and dot-product ($\cdot$) and has little memory requirements (merely 800 bits of memory) [60]. Although HB+ is implementable in EPC-Gen 2 tags [46], it has a high false rejection rate – for 80 bits of security, the false rejection rate is estimated at 44% [44, 45]. In HB+, The communication overheads increase linearly with the security parameter. In other words, lowering the communication overheads (transmission rate) lowers the security of the scheme.

Gilbert et al. [44] improved HB+ protocol and proposed HB# by expanding the size of secret keys by increasing the transmission rate. Both HB+ and HB# (with some modifications) are widely well-known and specifically proposed for implementation in EPC-Gen 2

tags [46], and they can be used as a good reference [44, 60]. We will extensively refer to HB+ and HB# protocols in the following chapters to comparatively analyze the performance of our proposed schemes. Therefore, it is necessary to provide an overview of HB#, which is proved to be more secure than HB+, in this section.

2.1.1 HB# Authentication Scheme

The HB# protocol is shown in Algorithm 1. In the HB# protocol, two $k_X \times m$ and $k_Y \times m$ matrices, denoted by X and Y respectively, are used. These matrices serve as the symmetric key shared between the tag and the reader. The tag uses a random binary source of noise. An m-bit vector v is added as a noise in which the probability of a bit to be 1 is η.

Algorithm 1 HB# authentication protocol for RFIDs

1.	$T:$	pick v randomly s.t. $v \in \{0,1 \mid \text{prob}('1') = \eta\}^m$
	$T:$	pick randomly $b \in \mathbb{F}_{2^{k_Y}}$
	$T:$	$b \longrightarrow R$
	$R:$	pick randomly $a \in \mathbb{F}_{2^{k_X}}$
5.	$R:$	$a \longrightarrow T$
	$T:$	compute $z = aX \oplus bY \oplus v$
	$T:$	$z \longrightarrow R$
	$R:$	check $w(z \oplus aX \oplus bY) \leq t$

The authentication process is performed in one round as follows: The tag sends a random m-bit vector b to the reader. Then, the reader challenges the tag by sending a random m-bit vector a back to the tag. The tag responds to the challenge by calculating a noisy vector z as follows:

$$z = aX \oplus bY \oplus v, \tag{2.1}$$

where $\oplus$ denotes XOR of two binary vectors. For verification, the reader has to only count the number of difference positions, denoted by e, between the received z and a self-generated vector $aX \oplus bY$. In other words, e equals the weight of $z \oplus aX \oplus bY$. If $e \leq t$ for some threshold t, the tag is authenticated. It is recommended in [44] to set $t = um$, for some $u \in [\eta, \frac{1}{2}]$. Note that HB+ and HB# are shown to be insecure against the man-in-the-middle attack [45], which alters the bits communicated between the tag and the reader.

There are other proposals in the literature [28, 86, 87, 88, 89] promising efficient security for RFID systems, but most of them have been proved to have some security weaknesses making them vulnerable to passive and active attacks [1, 2, 21, 69, 70]. The challenge in design of new cryptosystems for RFIDs is to provide a proof of security and confidence in the proposed scheme.

2.2 Asymmetric Authentication Schemes

There are various asymmetric cryptographic algorithms found in the literature. They vary greatly in terms of size, security, power consumption and communication overhead. Among

many asymmetric cryptosystems, one can refer to RSA [79], Rabin [68], Elliptic Curve Cryptography (ECC) [11, 15, 53] and NTRU [56, 57]. The RSA and Rabin cryptosystems are prevalent in software and over the internet. However, ECC is well suited for small devices, as it offers shorter keys for the same level of security than RSA, and it yields savings in the memory [47]. Asymmetric cryptosystems can be used for a variety of applications, such as data integrity, mutual authentication, confidentiality and non-repudiation. Asymmetric-key (public key) cryptosystems are superior to symmetric-key cryptosystems in key management, and they are highly recommended for large-scale, distributed systems. Nevertheless, they demand more resources and often require complex mathematical operations with large numbers. It has been shown that it is possible to implement some of the asymmetric cryptosystems, such as ECC and NTRU, on RFID tags [6, 37, 55, 57, 62, 63, 65]. However, the price of tags will increase drastically to accommodate a standard asymmetric algorithm in the RFID tag.

There exist other asymmetric schemes that are code-based and therefore efficient in hardware implementation [31, 42]. Gaborit and Girault [42] propose an asymmetric authentication scheme based on error-correcting codes. Their scheme is an adaptation of Stern authentication scheme [95] based on error-correcting codes. By using double circulant matrices, they manage to reduce the storage requirement from $\mathcal{O}(N^2)$ to $\mathcal{O}(N)$, where N is the size of the underlying code. Nevertheless, Gaborit and Girault's scheme requires implementation of a cryptographic hash function in the RFID tag [42]. In their authentication scheme (type B), the transmission rate is $42,336$ bits and the work-factor is $2^{20.7}$ binary operations, which are overwhelming for low-cost RFID tags.

Recently, it has been shown that the size of parameters in code-based cryptography can be drastically reduced, while preserving a high level of security [7]. Berger et al. [7] reduce the size of public keys in the code-based cryptosystem. This reduction opens new possibilities to revisit the application of asymmetric code-based cryptography in RFID systems. Here, we revisit one of the classical, code-based authentication schemes by Harari [54] that has inspired many other code-based authentication schemes [31, 23, 25, 95], including our FLA schemes given in Chapter 4. Harari's authentication scheme is unique as it is a *zero-knowledge system*. In a zero-knowledge authentication scheme, the verifier can readily verify that the prover has the correct credentials to pass the authentication, but it does not learn anything about the prover's credentials.

2.2.1 Harari's Authentication Scheme

Harari's scheme [54] is given as follows: a trusted center (such as the administrator) chooses a random binary code of length N and dimension k. It is recommended that $N \geq 2000$ and $k \geq 1000$. The generator of the code is denoted by a $k \times N$-matrix G (or equivalently by its parity-check $(N-k) \times N$-matrix H). Let μ be a small odd quantity chosen at random in the interval $[50, 100]$. Let s be a random *codeword* of weight μ. Note that since s is a codeword, then $Hs^T = \mathbf{0}$. The vector s is the secret key of user A (prover) that will be used in the authentication rounds. The trusted center publicizes the generator matrix G (or equivalently the parity-check matrix H) as well as μ. Note that s is a codeword of weight μ and is kept private as A's secret key. The following protocol would allow A to authenticate

herself to B (verifier):

- A sets $l \in [100, 200]$, as the number of random vectors used in the challenge and chooses l random binary vectors $r_i \in \mathbb{F}_{2^N}$ for $i = 1, \cdots, l$. Note that r_i's are kept private from B and they have disjoint supports, i.e. for any $i \neq j$, we have $w(r_i \oplus r_j) = w(r_i) + w(r_j)$, where $w(\cdot)$ denotes the Hamming weight function of a binary vector. We denote the weight of r_i by w_i. Then A computes the syndromes $t_i = Hr_i^T$, where r_i^T is the transpose of vector r_i.

- B receives the set $\{(w_i, t_i) | i = 1, \cdots, l\}$ from A and chooses an odd weight binary vector e of length l, where $\frac{l}{3} \leq w(e) \leq \frac{2l}{3}$. B returns e to A.

- A randomly computes a permutation π of $\{1, \cdots, l\}$ and sends it back to B.

- A and B both compute $t = \pi(e)$, where $\pi(\cdot)$ denotes applying π permutation on a binary vector.

- B randomly chooses b as the challenge selection-vector, such that $b \in \{t, \bar{t}\}$, where $\bar{t}$ is the binary complement of t. Each bit of b is denoted by b_i for $i = 1, \cdots, l$.

- A computes the masking, random vector r, such that $r = \sum_{i|b_i=1} r_i$ and returns the response vector y, such that $y = r \oplus s$ to B. Note that only r_i's are used in calculating r for which $b_i = 1$.

- B checks three conditions to verify A's identity

1. First condition on weights:

$$w(y) \neq \sum_{i|b_i=1} w_i$$

2. Condition on syndromes:

$$Hy^T = \sum_{i|b_i=1} t_i$$

3. Second condition on weights:

$$-\mu + \sum_{i|b_i=1} w_i \leq w(y) \leq \mu + \sum_{i|b_i=1} w_i$$

Harari's security analysis only provides a quantitative proof of security based on the known attacks on the ciphertexts only. Later Véron [98] showed that Harari's initial scheme is insecure. Later in Chapter 4, we remove the security flaw from Harari's scheme and propose a modified secure scheme. Moreover, Harai's scheme inspired us to propose a FLA scheme that is suitable for RFID systems. The details of our FLA scheme are given in Chapter 4.

Securing RFID systems is not limited to cryptographic authentication schemes. There exist other analogue attacks, such as relay attacks, which compromise any RFID authentication scheme. There exist other non-cryptographic approaches both to securely authenticate an RFID tag or to protect against the relay attacks.

2.3 Non-cryptographic Solutions

Non-cryptographic solutions heavily rely on the physical characteristics of the RFID system and the communication channel. This is a fairly new topic of study in RFID systems,

and only two main categories of systems had been explored before our novel approach in Chapter 6. The first category uses the intrinsic characteristics of the physical RFID tag. The second category uses the limitation in the communication channel (mostly time of signal propagation) to identify authentic RFID tags. This approach is usually referred to as the *distance bounding* technique. We briefly review both of these approaches in this section.

2.3.1 Physical Unclonable Function

The first category of identification mechanisms uses physical characteristics of RFID tag's circuitry, which are difficult to model precisely. These characteristics are represented as a Physical Unclonable Function (PUF). It has been shown that reliable silicon PUFs can be produced in which the PUF is known to the manufacturer only [12, 72, 75]. A PUF computes its output by exploiting the inherent variability of wire delays and gate delays in its circuitry [12]. These delays depend highly on some unpredictable factors, including but not limited to manufacturing variations, quantum mechanical fluctuations, thermal gradients, electromagnetic effects, parasitics, and noise [12]. A good PUF is therefore hard to be modeled precisely even when an identical hardware is used, and they can replace a cryptographic hash function. Bolotnyy and Robins [12] propose a PUF-based signature protocol for RFID tags that verifies the integrity of a messages. Although PUFs are highly desirable in small, resource constrained RFID systems, they have to deal with some practicality issues. In practice, the output of a PUF is only probabilistically consistent with its expected value, and it can be greatly distorted, depending on the noise in the system. This is still an ongoing research

topic in the literature.

2.3.2 Distance Bounding Techniques

Physical distance can serve as a measure of trust and privacy. In some applications, users can be granted access based on their perceived location. In low range RFID systems, it can be stated that the main task is proximity identification, as low range RFID tags can be identified only within a few meters (less than 3 meters) form a reader [39, 40]. As mentioned before, RFID systems are vulnerable to the relay attacks. In the relay attack, the attacker can remotely impersonate a tag and pretend that the tag is in proximity of the reader. The attacker relays the communication between the reader and a remote (authentic) tag without altering the communication messages or attempting to crypt-analyze the authentication mechanism.

Distance-bounding is referred to a suit of mechanisms that use the physical characteristics of the communication system to enforce an upper bound on the maximum distance an RFID tag can be placed from a reader. Measuring the physical distance has been used extensively as a countermeasure to thwart the relay attack [51, 64, 81, 93]. Some physical characteristics of the communication channel have been used in distance-bounding methods, such as signal's strength, angle of arrival and time of fly. The received signal's strength is inversely proportional to the distance of a transmitter. Therefore, it is possible to estimate the distance of the transmitter on an RFID chip from a reader by measuring the power of the received's signal [50]. This approach, however, is not effective against a relay attack,

as the attacker can easily increase the transmission power to fake an authentic transmitter in proximity. Angle of an arrived signal can also be used to determine the location of a transmitter [50]. This approach fails against sophisticated attackers that can reflect signals from various directions. Time of fly measures the time elapsed for a message exchange from a reader to a tag and then back to the reader. By knowing the propagation speed in the communication channel, one could easily find the distance of the reader from the tag. This approach has received great attention in the security community [49, 51, 64, 81, 82, 91], as it provides some level of security against relay attacks.

A major contribution in distance-bounding protocols for RFID systems is by Hancke and Kuhn [51, 81]. In their protocol, a random challenge is sent to the tag as a series of bits. The reader measures the time between sending each bit of the challenge and receiving the corresponding response. Measuring the time of fly between sent and received bits, the reader can estimate the distance of the transmitter in this protocol. The exchange of challenge and response bits is performed over a fast UWB channel that provides a fine resolution for measuring the distance. Munilla et al. [81, 82] propose a distance bounding protocol without using a UWB channel, but they still rely on precise timing of the communication signals.

The additional delay introduced by the attacker in the relay attack can be kept very small [49, 52], making it quite feasible to bypass the timing requirements set in ISO standards [39, 40]. A small tolerance of a few micro seconds in measuring time of fly in distance-bounding protocols will result in hundreds of meters difference in the measured distance. Novel protective solutions must be designed to protect RFID systems against the relay at-

tacks.

2.4 Key Management for Large Systems

Cryptographic security and privacy solutions for an RFID system require management and distribution of cryptographic keys between tags, readers and servers in the system. It is necessary to change and update the symmetric/asymmetric keys on many readers and tags in the system or to distribute new keys to servers for authenticating the readers. If the number of tags and readers are large in an RFID system, establishing a secure communication channel among them becomes challenging. Unlike classical PKC, where only two parties are involved, we need to establish a secure channel between many parties in an RFID system. The sender should be able to create a ciphertext (broadcast message) that can be decrypted by many parties. There are existing solutions proposed for other applications, such as video conferencing, Digital Rights Management systems, secure IP multi-casting, or group-key multi-casting, that can be used for RFID systems as well.

An attempt to efficiently scale group key sharing from two (as in the PKC) to three entities is found in [59]. The protocol then is extended in [5] to large groups consisting of n members. It requires $\mathcal{O}(\log_3 n)$ communication rounds for n members. Choi et al. [30] propose a constant-round protocol using bilinear pairings – that is defined later in Chapter 7. We also make extensive use of bilinear parings in our key management protocol in Chapter 7. A good collection of protocols based on bilinear pairings is gathered in [26, 76]. The most important requirements found in the literature [13, 14, 16, 18, 24, 26, 76] for key management

protocols are communication efficiency and collusion resistance.

Communication Efficiency: Regarding the communication bandwidth as a limited natural resource, design of a protocol that trades off computations or storage complexity for minimal communication overheads is desirable. Therefore, we would prefer solutions in which the communication overheads to securely share secret keys among a dynamic group of entities are very small. Ideally, the size of the broadcast message (ciphertext) should be $\mathcal{O}(1)$, that is it is fixed regardless of the number of entities involved.

Collusion Resistance: A broadcast message (ciphertext) can only be decrypted by members intended initially and included in the recipients' set. Existing members in the broadcast group that are not included in the recipient's set are not able to access the key encrypted in the broadcast message. In other words, the excluded members should not be able to cooperate together to obtain decryption of a (broadcast) ciphertext or private keys of other members in the broadcast group.

Most collusion resistant protocols impose great communication overheads [34, 48, 83] depending on the size of entities involved in the protocol. That is the ciphertext in these schemes grows (usually) linearly with the number of (privileged) entities in the group. Nevertheless, Boneh et al. [16] have proposed a system that has short ciphertexts, i.e. the size of the broadcast message is fixed and does not change with the size of the broadcast group. Their collusion resistant broadcast encryption is designed in a *static* security model; the static model is a weak security model, as it does not simulate the attacker in the real world. It provides security against certain attacks not all the attacks from the adversary in practice.

In the static model, the adversary has to announce in advance the set of entities S' (members in the system) that it will attack. Thus in the static model, it is known in advance which entities the adversary will attack. Therefore, the private keys of other entities (non-attacked entities) in the system will not be revealed to the adversary – which is not the case in a real application.

Gentry and Waters [43] proposed a broadcast encryption with short ciphertexts in the *semi-static* security model, where the adversary similarly must commit to a set S' of indices before the setup phase, but can query the private keys of an arbitrary subset of entities in S'. Note that the adversary still cannot query the private keys of other entities in the system. Gentry and Waters state that *"a semi-static adversary ... is stronger than a static adversary, in that its choice of which subset of S' to attack can be adaptive"* [43].

In the fully *adaptive* security model, no initial commitment is required from the adversary. The entities that the adversary can attack are not known in advance, as in the real world. Security in the adaptive model only allows the adversary to see the public parameters of the system and to access the private keys of any arbitrary member. Nevertheless, the non-attacked members remain secure in the adaptive security model. In Chapter 7, we propose a broadcast encryption scheme in the adaptive model that has short ciphertexts, and it is suitable for management and distribution of cryptographic keys in an RFID system.

In this chapter, we have reviewed some of the existing solutions that address security and privacy problems in the RFID system, and we have discussed their shortcomings and deficiencies. We have given a brief overview of existing symmetric and asymmetric authentication

schemes proposed for RFID systems, and reviewed HB# and Harari's scheme in greater details, as they will be referenced extensively in Chapters 3, 4 and 5. We have highlighted shortcomings of existing non-cryptographic countermeasures against the relay attacks. Later in Chapter 6, we will show how our chaos-masking scheme can secure RFID systems against relay attacks. Finally, the requirements for a secure key distribution scheme and some of the related work have been listed in this chapter. In Chapter 7, we will provide our key distribution system that is superior to the existing schemes.

Chapter 3

Backward-Link Authentication

RFID tags usually carry sensitive data, such as user's location, purchase orders, shopping preferences, dates of birth or other personal information. It's clear that one would need to protect these data and limit access to the tag's memory. The readers that can access the tag's memory have to be authenticated before it is authorized to read/write the memory's content. Therefore, RFID tags must be equipped with a mechanism to authenticate the readers before they are granted access to the tag's data. Most of existing authentication/encryption protocols are based on cryptographic primitives that are often very complex for low-cost RFID tags, such as basic EPC-Gen 2 tags [46] that have very limited resources for security. It is estimated that EPC Gen 2 tags can only accommodate 500-5000 gates, whereas a simple encryption function, such as AES, requires around $20,000$-$30,000$ gates [66]. These limitations demand design of new authentication schemes that are secure and light-weight, suitable for low-cost RFID tags.

Code-based cryptosystems are five to 10 times faster [10] than other number theoretic cryptosystems, such as ECC [65]. Therefore, we focus on designing authentication scheme based on error-correcting codes. In this chapter, we show how to modify a code-based cryptosystem to fit it in low-cost RFID tags. The required background to understand the main scheme is provided in Section 3.1. In Section 3.2, we use the same cryptosystem both for authenticating the reader and providing confidentiality to messages exchanged with the reader. We discuss the security of the proposed BLA scheme in Section 3.3. Finally in Section 3.4, the performance of our protocol is compared to other schemes that are suitable for RFID systems.

3.1 Coding Theory Background

Code-based cryptography shows great promises for simple design and efficient implementation in hardware, as it mostly requires matrix and vector operations over binary fields. Error-correcting codes that can be used in cryptography must possess known complexity problems that can be used to design code-based cryptographic systems and to provide a certain level of security in those system. In this section, we have provided the common definitions in coding and complexity assumptions that are used in this dissertation.

3.1.1 Linear Error-correcting Codes

In this section, we provide a brief introduction on error-correcting codes and code-based cryptography, sufficient to understand the proposed schemes in this dissertation. Further

details can be found in [61].

A linear binary code of length N is a vector subspace of $\mathbb{F}_{2^N}$. In a binary linear code, sum of every two codewords modulo 2 is also a codeword. A linear binary code $\mathcal{C}$ is denoted by $[N, k, d]$, where there are 2^k binary codewords of length N and a minimum *Hamming distances* of d between every two codewords. The Hamming distance of two codewords $a = (a_1, \cdots, a_N)$ and $b = (b_1, \cdots, b_N)$ is the number of positions, where the codewords are different.

When working with binary codes defined over $\mathbb{F}_2$, the *Hamming weight* w of a codeword is the number of non-zero positions in the codeword. The $k \times N$ matrix G is called a generator matrix of the code if k linearly independent codewords form the rows of G. A vector v of length k is mapped (encoded) into a codeword c of length N using the generator matrix via $c = vG$. The *parity-check matrix* is a $(N - k) \times N$ matrix, denoted by H, that satisfies $HG^T = \mathbf{0}$, where G^T is the transpose of matrix G. The parity-check matrix can be applied to check whether a codeword belongs to a code $\mathcal{C}$, since for every codeword $c \in C$, we have: $Hc^T = HG^T v^T = \mathbf{0}$.

It can be shown that there always exists a polynomial-time algorithm to obtain G from H and vice-versa [61]. The parity-check matrix is mainly used in error correcting, as one can figure out if a codeword is erroneous. The *syndrome* of a vector y is defined as yH^T. If y is comprised of an error vector e and a codeword c added together, the syndrome of y is zero if and only if e is zero, as shown in Eq. (3.1).

$$Hy^T = Hc^T \oplus He^T = \mathbf{0} \oplus He^T = He^T. \tag{3.1}$$

Eq. (3.1) gives a method to solve $(N-k)$ equations with N unknown variables of the error vector e, in order to correct the unknown error. It is known that for a code $[N,k,d]$, error vectors of weight $\lfloor \frac{d-1}{2} \rfloor$ or less can be corrected in the codewords. All $[N,k,d]$ codes satisfy the *singleton bound*, i.e. $d \leq N-k+1$.

3.1.2 Complexity Assumptions in Coding

Linear codes used in cryptography must satisfy the following complexity properties. The first complexity assumption is based on the Syndrome Decoding (SD) problem [78].

Definition 1 (Syndrome Decoding Problem)**.** *Let H be a parity-check matrix of a linear binary code $[N,k,d]$. Suppose i is the syndrome of a vector e, that is $i = He^T$, where e^T is the transpose of vector e. Suppose p is a given integer in the space of all possible weights of vector e. The question is if one can easily (in polynomial time) find a vector e' of length N, such that $He'^T = i$ and $w(e') < p$, where $w(\cdot)$ denotes the Hamming weight of a binary vector.*

The SD problem is solved if one can find a pair (m,e), such that $y = mG \oplus e$, where $w(e) = p$. It has been shown in [8] that decoding an arbitrary linear code is an NP-Hard problem. There is also a decision-variant of the SD problem, which is given as follows:

Definition 2 (Syndrome Decoding Decision Problem)**.** *Let G be a generator matrix of a binary code $[N,k,d]$, y a binary vector of length N and p an integer. The question is if there exists a vector e of length N and weight p, such that $y \oplus e \in \mathcal{C}$.*

The SD Decision problem is solved if one can find a pair (m, e), such that $y = mG \oplus e$, where $w(e) = p$. It has been shown that one can go from the parity-check matrix in the SD to the generator matrix in the SD decision problem (and vice-versa) in polynomial time [78].

A related complexity assumption is the difficulty of the Learning Parity with Noise (LPN) problem. This assumption has been the basis in the security proof of HB+ and HB# and also the BLA scheme proposed in this chapter.

Definition 3 (Learning Parity with Noise Problem)**.** *Let A be a random $(q \times k)$-binary matrix, let x be a random k-bit binary vector, $\eta \in (0, 1/2)$ be a noise parameter, and v be a random q-bit binary vector, such that $w(v) \leq \eta q$, where $w(\cdot)$ denotes the weight of a binary vector. Given A, η and $z = xA \oplus v$, find a k-bit binary vector y, such that $w(yA \oplus z) \leq \eta q$.*

LPN is an NP-Hard problem, that is any efficient (polynomial time) solver has only *negligible* probability of success in solving the problem. A function is negligible if it approaches zero faster than the inverse of any polynomial.

The security of many coding-based cryptographic protocols, such as McEliece [78] and Niederreiter cryptosystem [84], are based on the NP-Hardness of the SD problem stated above. In this thesis we will introduce modifications to McEliece cryptosystem that would allow us to embed it inside RFID tags. Thus, it is necessary to first provide a brief introduction to McEliece cryptosystem.

3.1.3 McEliece Cryptosystem

The McEliece cryptosystem [78] is a public-key cryptosystem based on the SD decision problem. Let $[N, k, d]$ be a linear code with length N, dimension k and minimum distance d. Let G be the generator matrix of an efficient code for which an efficient decoding algorithm exists. G is a k-by-N matrix. Take a random invertible $k \times k$-matrix S and a random permutation $N \times N$-matrix P. Compute the public-key matrix F as given in Eq. (3.2).

$$F = SGP \tag{3.2}$$

The matrix F and $[N, k, d]$ are publicly known, while the efficient decoding algorithm for G, matrix S and matrix P are kept private. It is easy to see that using F, anyone can encrypt a plaintext, denoted by $m \in \mathbb{F}_{2^k}$. To encrypt a message vector m, pick a random error vector $e \in \mathbb{F}_{2^N}$, such that its weight is equal to $\frac{d-1}{2}$. Then compute $y = mF + e$ and output $y \in \mathbb{F}_{2^N}$.

With the knowledge of the decoding algorithm for G and the matrices S and P, the ciphertext y is decrypted as follows: First calculate $yP^{-1} = mSG + eP^{-1}$ and then decode the result. It should be noted that since P^{-1} is a permutation matrix, eP^{-1} has the same weight as e. Decode yP^{-1} to obtain mS and remove S applying S^{-1}.

A more secure, randomized version of McEliece is proposed recently by Nojima et al. [85]. Randomized McEliece [85] is a probabilistic cryptosystem in which the encryption algorithm encrypts $[r|m]$ instead of m itself. Let's suppose that k_1 is the length of the random string r, and k_2 is the length of the message m. k_1 and k_2 are chosen such that $k = k_1 + k_2$ and

$k_1 = bk$, where $b < 1$ is a positive rational number, e.g. $b = \frac{9}{10}$. The decryption algorithm is almost the same as in the original McEliece, except that it outputs only the last k_2 bits of the decrypted string. For more details, we refer to [85]. Using the randomized McEliece cryptosystem, we design an identification protocol, where a tag can present its identity to a reader without other readers finding any information about the tag's identity.

3.2 Backward-Link Authentication

The main problem in code-based cryptography is that large parameters have to be used to provide a sufficient level of security in practice. The size of encoding matrix (F) that can be used in (randomized) McEliece cryptosystem is very large and requires huge memory on RFID tags. The recommended parameters for a secure McEliece cryptosystem were originally $[N, k, d] = [1024, 524, 50]$ [9, 78], but to secure against new attacks, this was later increased to $[N, k, d] = [2960, 2288, 56]$ requiring around $6.5M$ bits in the memory [9]. Recently, Berger et al. [7] showed that by using quasi-cyclic alterant codes, the McEliece parameters can be compressed to $[N, k, d] = [450, 225, 56]$. The new, compressed parameters require around $100K$ bits of memory, which is still not practical for most low-cost RFID tags, e.g. EPC-Gen 2 tags [46, 89]. We modify the McEliece cryptosystem in such a way that the RFID tag can efficiently store the public-key matrix and perform the encryption operations.

3.2.1 Modified McEliece Cryptosystem

Our adaption is based on the fact that an m-by-n Toeplitz matrix can be stored by saving only $m+n-1$ elements of the Toeplitz matrix, whereas storing a totally random m-by-n matrix, one would need all the elements (mn elements). In the original McEliece cryptosystem, the matrix S is selected as a purely random, invertible matrix in computing the public-key matrix F. Here, we modify the calculation of public-key matrix in such a way that it will be turned into a large Toeplitz matrix. This is a similar technique that has been used to reduce the overheads in HB# protocol [44]. However, the security of HB# protocol is proven based on the LPN assumption on purely random matrices, and this security proof cannot be expanded to Toeplitz matrices in HB# protocol [44].

Let $[N, k, d]$ be a linear code. We set G as a k-by-N matrix as the generator matrix of an efficient code, e.g. Goppa Code [78]. As before, P is a random N-by-N permutation matrix. However, the invertible k-by-k matrix S is calculated as follows: Take a random N-by-N, Toeplitz matrix T and calculate the N-by-k matrix X, such that $XGP = T$. Using the Gaussian elimination, one could easily find such a matrix. Notice that the matrix S in the original McEliece system [78] is replaced by a k-by-k sub-matrix of X. As in the original McEliece system, S has to be an invertible matrix. It should be clear that the public-key matrix F is a k-by-N matrix, which is formed from k rows of T. The same rows are selected as the rows in X that has formed S. In other words, F is a Toeplitz matrix (missing a few rows), and it can therefore be stored very efficiently in the memory.

3.2.2 Main BLA Protocol

Using the modified McEliece cryptosystem that was described in the previous section, we propose the following BLA protocol: Let us denote the tag by T and the reader by R. Each reader R has a modified McEliece public-key F defined over a given linear code $[N, k, d]$. The tag T and reader R are engaged in the BLA protocol, as shown in Algorithm 2.

Algorithm 2 Backward Link Authentication scheme for RFID tags

1. T : picks e randomly s.t. $w(e) = \lfloor \frac{d-1}{2} \rfloor$

   T : selects a random $r \in \mathbb{F}_{2^k}$

   T : computes $y = rF \oplus e$

   T : $y \longrightarrow R$

5. R : decrypts y to recover r

   R : $r \longrightarrow T$

   T : checks correctness of r, o.w. quit

At the setup, when the tag is initialized, a trusted center (programmer or manufacturer) stores F into the RFID tag's memory. The public-key matrix F can be stored very efficiently in the memory, as the tag only needs to store the first row and the first column of the matrix requiring $k + N - 1$ bits. When the tag T is queried, the reader is identified to the tag. The tag then searches its memory to find the corresponding public-key F of the reader R. If T does not recognize R or does not find the corresponding public-key in its memory, it quits

the protocol. Otherwise, it proceeds by picking a random error vector $e \in \mathbb{F}_{2^N}$, such that its weight is equal to $\lfloor \frac{d-1}{2} \rfloor$. Then, it has to take a new random vector $r \in \mathbb{F}_{2^k}$ and compute $y = rF \oplus e$. Computing y is very simple. To calculate y, T only needs to calculate rF, where F is a Toeplitz matrix, and to perform a binary addition with e. Calculating rF requires only shift and addition of rows in F corresponding to non-zero bits in r. Let's denote the i-th bit in the random vector r by r_i and the i-th row in F by f_i. Then, rF is calculated as sum of f_i-s for which $r_i = 1$. Since, F is a Toeplitz matrix, f_i-s are easily computed from shifting the first row and column, and the tag does not need to store f_i-s separately.

Note that y is a valid McEliece ciphertext, which can be decrypted if R has knowledge of the corresponding McEliece decryption algorithm corresponding to F. The reader R decrypts y to retrieve r and sends it back to T as the response. The tag T receives r from the reader and checks the response against the stored value in the memory. If they are not equal, T refuses R and authentication fails. Otherwise, the authentication succeeds and the tag accepts the reader.

The proposed authentication protocol is a simple challenge-response based on McEliece cryptosystem. In any RFID system, usually a reader is interacting with many tags, whereas each tag is only communicating with a few (maybe less than five) readers. Therefore, the tag will need to store a few F in the memory. The only parameters the reader has to store are the decoding algorithm of the McEliece cryptosystem, matrix S and matrix P, regardless of the number of tags in the system.

Not only the tag should be able to authenticate readers, but it also has to protect its sensitive content. Since data from the RFID tag are transmitted wirelessly to the reader, it is easy for an eavesdropper to capture the data in the air. This is specially the case with passive tags, as a passive tag immediately transmits its content (identity), as soon as its energized by a reader. We address confidentiality of tag's content in the rest of this section.

3.2.3 Data Confidentiality

Confidentiality is about restricting access to sensitive data only to authorized entities. Sensitive data in security applications should be hidden from any unauthorized entity. The proposed BLA authenticates the reader to RFID tags via a challenge-response protocol. As shown in Algorithm 3, the same protocol can be easily adapted to add data confidentiality to RFID tag as well as authenticating readers. Let's suppose that a confidential message $m \in \mathbb{F}_{2^k}$ has to be transmitted to the reader wirelessly. The message can include the tag's identity if privacy of the tag is also required in the system.

The protocol is given in Algorithm 3. The tag simply picks a random session key $s \in \mathbb{F}_{2^k}$ and computes $c = m \oplus s$. The session key s is encrypted as $z = sF \oplus e$, and z is sent to the reader. The reader uses its private key to decrypt z and retrieve s. Then, it simply XORs s with c to recover m.

We have reused the proposed BLA scheme to add confidentiality to the data exchanged with the reader in the air. The reader or the tag does not need to store an encryption key

Algorithm 3 Confidential exchange of data for RFID tags

1. T : picks e randomly s.t. $w(e) = \frac{d-1}{2}$

   T : selects a random $s \in \mathbb{F}_{2^k}$

   T : computes $z = sF \oplus e$

   T : computes $c = m \oplus s$

5. T : $\{c, z\} \longrightarrow R$

   R : decrypts z to recover s

   R : computes $m = c \oplus s$

locally. The tag shares a secret (session) key with the reader every time a new message is to be transmitted. It should also be noted that the proposed scheme in Algorithm 3 inherently protects the privacy of tags, as a new random number (i.e. s) is XORed with the tag's identity (represented in m) in every session. This implies that even if the same tag is queried, a new random number will be transmitted and therefore the message (i.e. c) that the adversary captures in the air will be different. Also note that using a public-key cryptosystem removes the need for the tag to store session keys locally. In every new round, a new session key is generated at random and then it is shared with reader through the protocol shown in Algorithm 3.

3.3 Security Discussions

The proposed BLA scheme should provide a verifiable, secure means of authentication for low-cost RFID tags. Our scheme is based on a modification of McEliece cryptosystem given in Section 3.2. We have to show that the modification does not compromise the security of McEliece cryptosystem.

The modification in McEliece cryptosystem is applied to the matrix S, we change it from a purely random k-by-k invertible matrix to a calculated k-by-k invertible, matrix S. The invertible matrix S is derived from $XGP = T$ after eliminating (last) $N - k$ rows from the N-by-k matrix X. The randomness of the McEliece F public-key matrix is reduced from the space of all possible random matrices, i.e. 2^{N^2}, to the space of all possible Toeplitz matrices, i.e. 2^{2N-1}. Nevertheless, the space of possible Toeplitz matrices is still very large for a large N (typically $N \approx 450$).

There have been various attacks suggested to cryptanalyze McEliece cryptosystem, including: brute-force key search attacks, exploiting specific structures of Goppa codes, generalized decoding and syndrome decoding attacks [7, 9, 19, 20]. None of these attacks have any substantial advantage in breaking the modified McEliece cryptosystem, since the proposed modification does not change the structure of Goppa codes. The randomness of T will result in the randomness of S, but over a smaller space. Note that G and P are purely random over their original space with the same distribution as in the original McEliece cryptosystem.

It should be noted that changing the structure of the McEliece public key F from a random-looking matrix to a Toeplitz matrix could possibly reduce the security of the au-

thentication scheme. The challenge message is generated as $y = rF \oplus e$ in Algorithm 2. The adversary wins if he can successfully guess r from y and $w(e)$. It should be clear that this problem is an instance of LPN problem. Therefore for a large matrix F, it is computationally infeasible to guess r by knowing only F, y and $w(e)$.

In the next section, we will show in greater details that our proposed authentication protocol will require a few resources to be implemented on low-cost RFID tags.

3.4 Resource Requirements of BLA Scheme

Let's use the McEliece parameters $[N, k, d] = [450, 225, 56]$ proposed by Berger et al. [7] to achieve a greater improvement by reducing the size of public key in the McEliece cryptosystem. With the given McEliece parameters, we have $r \in \mathbb{F}_{2^{225}}$ and $e \in \mathbb{F}_{2^{450}}$, where $w(e) = 27$. Therefore, the RFID tag only needs to store $k + N - 1$ bits, i.e. 674 bits, in its memory to store F. The computations required in the tag are very trivial. The tag has to calculate rF by approximately $4,000$ XORs (i.e. $\mathcal{O}(N \log N)$) and adding e to the resultant by 450 XORs. Eventually, the tag will have to verify the reader's response with r that can easily be done with 225 XORs. The number of XOR operations in the tag is estimated at $5,000$ XORs, requiring around $5K$ circuit gates and $1,349$ bits of memory, which can be very well implemented in low cost RFID tags [89].

We have compared in Table 3.1 the resources required in our scheme with popular symmetric and asymmetric authentication schemes specifically proposed for low-cost RFID tags. The comparison in Table 3.1 is based on the suggested practical parameters for HB+ [60],

HB# [44] based on using Toeplitz matrices and Gaborit (type B) [42].

Table 3.1: Resource requirements in the BLA scheme compared with similar code-based authentication schemes

Scheme	Memory (bits)	Message (bits)	XORs	Key Sharing
HB+ [60]	800	8,040	16,000	Symmetric
HB# [44]	1,472	953	194,481	Symmetric
Gaborit (Type B) [42]	1,041	42,336	1,690,584	Asymmetric
BLA Scheme	**1,349**	**450**	**5,000**	**Asymmetric**

As it can be seen from Table 3.1, our scheme is an asymmetric authentication scheme and it outperforms the symmetric-key schemes of HB family in terms of communication and computation overheads. Also, our scheme has better a communication rate and computation complexity compared to Gaborit (type B) [42]. Note that Gaborit's authentication scheme also requires implementation of a cryptographic hash function in the RFID tag [42], which further increases the complexity of the tag and its price. As seen in Table 3.1, our proposed BLA scheme requires greater memory than HB+ and Gaborit (type B), which is still less than HB#. Nevertheless, memory sizes in RFID tags range from just 1 byte (as in pigeon tags) to 64K bytes (microwave tags with SRAM) [38]. Therefore, it is quite feasible to implement the proposed BLA scheme in low-cost RFID tags.

The amount of memory required is directly proportional to the recommended size of public key matrix in McEliece cryptosystem. The total memory required for the RFID tag to perform the operations in Algorithm 2 is estimated at $2(k+N)-1$ bits, where k and N are the dimensions of the public key matrix in McEliece cryptosystem. The recommended size of parameters for McEliece cryptosystem offer a very high level of security. The recommended parameters ($k = 225, N = 450$) yield 80 bits of security [7]. This means that the complexity of the best decoding attack on McEliece cryptosystem with the compressed parameters is 2^{80} operations. This is considered to be computationally infeasible to break within existing computers. The proposed BLA scheme can be easily fit in EPC-Gen 2 tags if we set $k = 170$ and $N = 340$ requiring $1,019$ bits of memory.

Chapter 4

Forward-Link Authentication

In this chapter, we design a zero-knowledge protocol to authenticate RFID tags to readers, i.e. an FLA scheme. Recent improvements in reducing the size of code-based cryptosystems has made classical zero-knowledge protocols, such as Harari's scheme, more efficient than other code-based schemes [22, 95]. However, Harari's initial design is shown to be insecure [98]. First in Section 4.1, we secure Harari's scheme and provide a formal proof of security for the modified scheme. Then in Section 4.2, we show that the improved Harari's scheme can be used as an FLA scheme in sophisticated RFID tags with great computational and memory resources. Later in this chapter, we present a light-weight (zero-knowledge) FLA scheme. Our proposed FLA scheme is given in Section 4.3. We show in Section 4.4 that the proposed scheme is secured based on the difficulty of the SD problem. In Section 4.5, the required resources of the proposed FLA scheme are analyzed, and we show that the proposed scheme is suitable for low-cost RFID tags.

4.1 Securing Harari's Scheme

We remove the security flaw in the Harari's scheme to make it secure against Véron's attack [98]. We achieve this without increasing the transmission rate or computational overheads of the Harari's original scheme given in Chapter 2, Section 2.2. The modified Harari's scheme is given as follows.

The trusted center publicizes the random parity check matrix H, as before. The notations used in our adaptation are the same as in Harari's original scheme with a few changes; the private key of A is chosen as a *random vector* $s \in \mathbb{F}_{2^N}$ of weight μ instead of a codeword. Since, s is a random vector instead of codeword, it has a non-zero syndrome denoted by d, where $d = Hs^T$. The trusted center publishes $\{d, \mu\}$ as A's public parameters. The following identification protocol would allow A to authenticate herself to B:

- A chooses l random binary vectors, $r_i \in \mathbb{F}_{2^N}$ for $i = 1, \cdots, l$, where the supports of the vectors are disjoint, i.e. for any $i \neq j$, we have $w(r_i \oplus r_j) = w(r_i) + w(r_j)$. We denote the Hamming weight of r_i by w_i, that is $w_i = w(r_i)$. Then A computes the syndromes $t_i = Hr_i^T$, where r_i^T is the transpose of the vector r_i.

- B receives the set $\{(w_i, t_i) | i = 1, \cdots, l\}$ and chooses an odd weight binary vector e of length l, where $\frac{l}{3} \leq w(e) \leq \frac{2l}{3}$. B sends e to A. Each bit of e is denoted by e_i for $i = 1, \cdots, l$.

- A computes the random masking vector r, as $r = \bigoplus_{i|e_i=1} r_i$ and returns $y = r \oplus s$ to B.

- B checks three conditions on y to verify A's identity

  1. First condition on weights:

$$w(y) \neq \sum_{i|e_i=1} w_i$$

  2. Condition on syndromes:

$$Hy^T = d \oplus \bigoplus_{i|e_i=1} t_i$$

  3. Second condition on weights:

$$-\mu + \sum_{i|e_i=1} w_i \leq w(y) \leq \mu + \sum_{i|e_i=1} w_i$$

Note that in the above conditions, d is added for verification, when checking the condition on syndromes. With this modification, an attacker can no longer send arbitrary vectors to the victim as described in the Véron's attack. The adversary might be able to find a random vector x of weight μ, instead of s, to satisfy the conditions on weights, but it would fail on the condition on syndromes. It is shown in the next section that addition of d in the condition on syndromes prevents the adversary from sending an arbitrary vector x to surpass the authentication process. In the rest of this section, we prove that the proposed modification in the Harari's scheme makes it as secure as solving the SD problem.

4.1.1 Security of Modified Harari's Scheme

Let's suppose that there is an algorithm $\mathcal{A}$ that generates vectors to surpass the checking conditions in the proposed scheme. Also, there is a simulator algorithm $\mathcal{B}$ that tries to find

a vector of syndrome d^* and weight less than p, that is $\mathcal{B}$ tries to solve an instance of the SD problem.

The algorithm $\mathcal{B}$ challenges $\mathcal{A}$ for authentication in the proposed scheme once with d and another time with d', such that $d \oplus d' = d^*$. It sets μ and μ' as the weights of d and d', respectively, such that $\mu + \mu' \leq p/2$.

Algorithm $\mathcal{A}$ sends $\{(w_i, t_i)|i = 1, \cdots, l\}$ in the first challenge and $\{(w'_i, t'_i)|i = 1, \cdots, l'\}$ in the second challenge. Algorithm $\mathcal{B}$ sets e and e' such that $\bigoplus_{i|e_i=1} t_i = \bigoplus_{i|e'_i=1} t'_i$ and also $|\sum_{i|e'_i=1} w_i - \sum_{i|e'_i=1} w'_i| \leq p/2$. Note that since the vectors r_i and r'_i have disjoint supports, we should have: $\sum_{i|e_i=1} w_i \leq N$ and also $\sum_{i=1|e'_i=1} w'_i \leq N$. Therefore, the two series of w_i and w'_i are bounded by N, and it is possible to find two close enough series that satisfy:

$$|\sum_{i|e_i=1} w_i - \sum_{i|e'_i=1} w'_i| \leq p/2$$

The algorithm $\mathcal{A}$ by using e and e', assigns $r = \bigoplus_{i|e_i=1} r_i$ and $r' = \bigoplus_{i|e'_i=1} r'_i$, and then returns $y = r \oplus s$ and $y' = s' \oplus r'$, respectively. It is straightforward to show that the following the inequalities are true:

$$|w(y \oplus y')| - |w(r) - w(r')| \leq |w(y) - w(r)| + |w(y') - w(r')| \leq \mu + \mu'$$

Since we have $\mu + \mu' \leq p/2$ and the algorithm $\mathcal{B}$ has set the weights such that $|w(r) - w(r')| \leq p/2$, we can then conclude:

$$|w(y \oplus y')| \leq p$$

On the other hand, the condition on the syndromes returns:

$$\begin{aligned} Hy^T \oplus Hy'^T &= H(y^T \oplus y'^T) \\ &= (d \oplus d') \oplus (\bigoplus_{i|e_i=1} t_i \oplus \bigoplus_{i|e'_i=1} t'_i) \\ &= (d \oplus d') = d^* \end{aligned}$$

That is $y \oplus y'$ is a binary vector of syndrome d^* and weight less than p. Thus, the vector $y \oplus y'$ is a solution to the given instance of the SD problem.

4.2 Requirements of Modified Harari's Scheme

In this section, we analyze the amount of resources required to implement the modified Harari's scheme in RFID tags. Using the recommendations on the size of the parity check matrix in [7], we use a 225×450 parity-check matrix and pick $l \in [50, 100]$, as suggested in [54]. In Table 4.1, we have compared the performance of the proposed scheme to results of Véron [98], pages 266-267, and the newest implementation [22] of Stern's scheme [95]. Note that the prover's computational overhead in Table 4.1 is approximated by $\mathcal{O}(N^2)$, where N is the size of the codewords.

As it can be seen in Table 4.1, the modified Harari's scheme requires less resources than other similar zero-knowledge protocols. It has a significantly better transmission rate than the newest implementation [22] of Stern's scheme [95] with a comparable matrix size and

Table 4.1: Requirements comparison of zero-knowledge authentication schemes

Scheme	$H_{k\times N}$	Memory (bits)	Prover's Work (XOR)	Message (bits)
Harari[54]	1000×2000	1,002,000	$2^{28.2}$	153,900
Véron[98]	1000×2000	1,002,000	$2^{32.2}$	3,078,000
Stern [22]	256×512	1,718	2^{18}	40,000
Modified Harari	225×450	6,750	$2^{17.7}$	17,100

computational overheads. Despite the fact that Stern's implementation [22] has lower memory overheads, it requires implementation of hash functions that will further increase the complexity of the circuitry on an RFID tag.

In the modified Harari's protocol, we have removed the security flaw from Hararis scheme and provided a formal proof of security based on the hardness of the SD problem. We have also shown that our proposed scheme outperforms other code-based zero-knowledge schemes in terms of communication and computation overheads. The proposed protocol is based on an asymmetric-key algorithm that simplifies the key management. However, the memory requirements are overwhelming for low-cost RFID tags, e.g. EPC-Gen 2 tags [46]. In the next section, we proposed a light-weight FLA scheme that can directly be fit into low-cost RFID tags.

4.3 Code-based Authentication of RFID Tags

We propose a scheme based on error-correcting codes in which complex matrix operations are moved from the resource-constrained RFID tags to the (usually more powerful) reader, without any sacrifice in the security of the scheme. Our proposed FLA scheme is given as follows.

Let's denote the tag by T and the reader by R. In the proposed protocol, a trusted center chooses a random binary code $\mathcal{C}$ of length N and dimension k that has the generator matrix G and the corresponding parity-check matrix H. Let $s \in \mathbb{F}_{2^N}$ be a random vector of weight μ and $c \in \mathcal{C}$ a random codeword. Vectors s and c are given to the tag T as its secret key. The trusted center publishes $d\,(\neq 0)$, $w(c)$, H and μ as the tag's public parameters, where $d = Hs^T$ and $w(c) > \mu$. As shown in Algorithm 4, the reader R first has to retrieve the tag's certificate that includes $\{d, w(c), H, \mu\}$. Then, it initiates the authentication process shown in Algorithm 4.

In order to authenticate an RFID tag T to the reader R, the reader R picks a random number $t \in [N/3, 2N/3]$ and sends it to T. Then, T chooses a random binary vector $r \in \mathbb{F}_{2^N}$ and computes $y_0 = s \oplus r$ and $y_1 = r \oplus c$. Then, it sends $\{y_0, y_1\}$ back to the reader. The reader R calculates $H(y_0 \oplus y_1)^T$ and then verifies the tag's identity by using the tag's public parameters. It checks the following conditions:

1. Syndrome Condition: The reader first checks if $d = H(y_0 \oplus y_1)^T$. The correctness of the first condition can be easily verified; since $c \in \mathcal{C}$, we have $H(c)^T = 0$ and we get:

$$H(y_0 \oplus y_1)^T = H(r \oplus s \oplus r \oplus c)^T = H(s \oplus c)^T = H(s)^T = d$$

2. First Weight Condition: The reader verifies the condition on the weight of $y_0 = s \oplus r$ by checking $t - \mu \leq w(y_0) \leq t + \mu$. This can be easily verified, since μ is small and $w(r) = t$.

3. Second Weight Condition: The reader verifies the condition on the weight of y_1 by checking $t - w(c) \leq w(y_1) \leq t + w(c)$. This can be easily verified, since $y_1 = r \oplus c$.

4. Third Weight Condition: The reader verifies the condition on the weights of y_0 and y_1 by checking $w(c) - \mu \leq w(y_0 \oplus y_1) \leq w(c) + \mu$. This is also correct, since $y_0 \oplus y_1 = s \oplus c$.

In the following section, we investigate the security of the proposed scheme and show that the proposed FLA scheme provides a secure means to authenticate RFID tags.

4.4 Security Analysis

In this section, we show that the security of the proposed authentication scheme is based on the difficulty of the SD problem. The adversary, without any knowledge of the tag's secret key $\{s, c\}$, has to solve SD problem to retrieve the tag's secret key and to surpass the authentication mechanism. Let's assume, the adversary's knowledge is restricted to the public parameters of the tag $\{d, w(c), H, \mu\}$ and the messages communicated in the authentication protocol.

Algorithm 4 FLA scheme for an RFID system

1. $R:$ pick a random number $t \in [N/3, 2N/3]$

   $R:$ send $t \longrightarrow T$

   $T:$ pick a random $r \in \mathbb{F}_{2^N}$, s.t. $w(r) = t$

   $T:$ compute $y_0 = s \oplus r$, $y_1 = r \oplus c$

5. $T:$ send $\{y_0, y_1\} \longrightarrow R$

   $R:$ compute $H(y_0 \oplus y_1)^T$

   $R:$ check if

   1. $d = H(y_0 \oplus y_1)^T$,
   2. $t - \mu \leq w(y_0) \leq t + \mu$,
   3. $t - w(c) \leq w(y_1) \leq t + w(c)$,
   4. $w(c) - \mu \leq w(y_0 \oplus y_1) \leq w(c) + \mu$.

Guessing s: Guessing s is equivalent to obtaining a vector of length N of a given syndrome d and weight μ, which is an instance of NP-Hard problem. Total number of vectors of weight μ is $\binom{N}{\mu}$, but since $d \neq 0$, s is not a codeword. The probability of randomly finding a correct s of weight μ is $2^{-(N-k)}\binom{N}{\mu}$, which is very small for large values of N, regardless of μ. As an example, for $N = 450$, $k = 225$ and $\mu = 25$, the probability of correctly guessing s is less than 2^{-89}.

Guessing c: It is straightforward for the adversary to find $s \oplus c$ from the messages exchanged in the proposed protocol. Correctly guessing a codeword of weight $w(c)$ is an instance of SD problem, which is considered NP-Hard for large values of N and a random H. Therefore, the probability of successfully guessing c for large values of N is also negligible.

Guessing r: If a correct r is found, the adversary can easily find s and c from y_0 and y_1, respectively. However, the total number of random vector r of weight t is $\binom{N}{t}$, which is an extremely large number for large values of N and $t \in [N/3, 2N/3]$ as set in Algorithm 4. Therefore, the probability of successfully guessing r is also negligible.

Deriving s from y_0, y_1-s: Let's suppose that the adversary has access to m number of messages communicated in the authentication process. We denote all messages that have been communicated by $A = \{t, y_0, y_1\}$ for m rounds. Every (y_0, y_1) pair yields a fixed $(s \oplus c)$ and one new equation $(s \oplus r)$ for every new r. Therefore, the adversary only has $m + 1$ equations

in A with $m+2$ unknowns. Thus, the adversary's attempt to find s from A is futile.

Replay Attack: An adversary who has intercepted the messages communicated in previous authentication rounds can succeed in a new authentication round only if the same t is used in the challenge. Note that since $t \in [N/3, 2N/3]$, it is limited to $N/3$ different choices. For very large values of N, the probability of selecting the same t is slim, but selecting large values of N increases the size of the public parameters and the transmission rate in the protocol. Repeating the protocol many times (n times) is one solution to decrease the adversary's probability of success in the replay attack. As an example, if we have set $N = 450$ and repeated $n = 10$ times, the probability of choosing the same t is less than $2^{-n\log(N/3)} = 2^{-70}$.

Man-in-the-middle Attack: In the man-in-the-middle attack, the adversary alters the signals communicated between the RFID tag and reader, in order to obtain the secrets or surpass the authentication round. Formally, we prove in Theorem 2 that the proposed authentication scheme is secure against the man-in-the-middle attack based on difficulty of the SD problem.

Theorem 1. *If there is a man-in-the-middle attack with non-negligible advantage (ϵ) in breaking the proposed authentication protocol, one can solve the SD problem with the same probability (ϵ).*

Proof. Let's suppose that the adversary has access to a polynomial-time algorithm $\mathcal{A}$ that surpasses the proposed authentication protocol by probability ϵ. The algorithm $\mathcal{A}$ generates

valid (y_0, y_1) tuples for any random number $t \in [N/3, 2N/3]$. If such an algorithm $\mathcal{A}$ exists, we show that it can be used to solve the SD problem with probability ϵ.

As a part of the proof, we build a simulator algorithm $\mathcal{B}$ that receives an instance of SD problem; that is $\mathcal{B}$ is given a random parity-matrix H and is challenged to find a vector e, such that $w(e) < p$ and $He^T = i$. $\mathcal{B}$, on the other hand, prepares an appropriate challenge to algorithm $\mathcal{A}$ and uses its response to solve the SD problem at hand: $\mathcal{B}$ simply picks a random μ and $w(c)$ such that $\mu + w(c) < p$ and $w(c) > \mu$. It then sets $d = i$ and publishes $\{d, w(c), H, \mu\}$ as the public parameters in the proposed authentication scheme. Then, $\mathcal{B}$ challenges $\mathcal{A}$ by sending a random number $t \in [N/3, 2N/3]$. The algorithm $\mathcal{A}$ then returns a (y_0, y_1) tuple that passes the syndrome and weight conditions of the authentication process. Therefore, we must have $d = H(y_0 \oplus y_1)^T$ and $w(y_0 \oplus y_1) \leq w(c) + \mu < p$. Therefore, algorithm $\mathcal{B}$ can return $e = y_0 \oplus y_1$ as a solution to the given instance of the SD problem.

In other words, $\mathcal{B}$'s advantage in solving the SD problem is exactly $\mathcal{A}$'s advantage in surpassing the authentication in the proposed scheme. Since, SD problem is NP-Hard for large values of N with a negligible probability of success, no $\mathcal{A}$ exists that has a non-negligible advantage in breaking the proposed authentication scheme. □

The proposed FLA scheme is a light-weight protocol, which is formally proved to be secure against the man-in-the-middle attack. In the proposed authentication scheme, the complex matrix operations are moved from the RFID tags to the reader, which usually has more computational power. We further discuss the performance of the proposed FLA scheme in the following section.

4.5 Resource Requirements of FLA Scheme

In the proposed FLA protocol, we do not require the RFID tag to store the entire public-key H, and the RFID tag carries the pre-computed values of the codeword c. If we set $[N, k, d] = [450, 225, 56]$, storing each pair of $\{s, c\}$ requires 900 bits of memory. The only computation needed at the RFID tag is calculation of y_0 and y_1. Computing $y_0 = s \oplus r$ and $y_1 = r \oplus c$ can be done very fast by a simple XORing of 450-bit vectors. Table 4.2 summarizes the resources required in the RFID tag in comparison with other efficient authentication schemes for recommended for RFID systems.

Table 4.2: Performance of the FLA scheme in comparison with similar authentication schemes

Scheme	Memory (bits)	Message (bits)	XORs	Key Sharing
HB+ [60]	800	8,040	16,000	Symmetric
HB# [44]	1,472	953	194,481	Symmetric
Gaborit (type B) [42]	1,041	42,336	1,690,584	Asymmetric
FLA Scheme	900	900	1,800	Asymmetric

From the reader to the tag, only one vector t of $\log_2(N/3)$ bits needs to be transmitted. From the tag to the reader, the messages are limited to two 450 bit-long vectors. The computational operations by the tag are only limited to calculating $s \oplus r$ and $r \oplus c$ with a total of 900 XORs. The computation complexity for the reader, on the other hand, is heavier

than the tag, and it is in order of $\mathcal{O}(N^2) \approx 203,000$ in each authentication round. The total memory required for the tag to participate in Algorithm 4 is $2N$, where N is the size of the codewords used in McEliece cryptosystem. The underlying McEliece cryptosystem with the given parameters offers 80 bits of security [7]. As it can be seen from Table 4.2 that our scheme is more suitable for EPC-Gen 2 tags than any other authentication schemes for RFID tags. The proposed FLA scheme requires only 900 bits of memory, which can be easily fit into EPC-Gen 2 tags.

Using an asymmetric FLA scheme allows the system to expand according to any number of tags and readers. The only parameters the reader needs to store are the public-key of the tags that can be downloaded from the server at the time of authentication if they do not already exist at the reader. The public-keys can be discarded after every authentication saving the reader's memory. This is extremely important in practice, as a reader might be authenticating a huge number of tags.

The proposed authentication scheme in Chapter 4 is secure against eavesdropping attacks. If there are concerns over the replay attacks, one could repeat the same scheme many times. Repetition of the proposed FLA scheme will increase its security against the replay attacks. With the recommended parameters for a secure McEliece cryptosystem, the adversary's probability of success would be less than 2^{-70} if the proposed FLA scheme is repeated $n = 10$ times.

So far, we have designed light-weight authentication schemes that are suitable for low-cost RFID tags and can either prove the identity of the tag to the reader or verify the reader to

the tag. Privacy is another important factor in RFID systems that we will handle in the next chapter.

Chapter 5

Privacy Protecting Identification

RFID tags are very small in size and usually transparent to their bearer. They are wirelessly activated and scanned without being noticed. It is clear that openly sharing the tag's identity poses privacy concerns. Some sensitive data, such as location or frequency of a visit to a store, even if they are protected via encryption, can be uncovered simply by tracing a tag.

In this chapter, we propose a solution that protects the tag's privacy. Our solutions identifies an RFID tag only to authorized readers, while the tag is not even traceable across the authorized readers. The goal is to design a light-weight scheme that identifies the RFID tags only to their authorized readers, while tags' identity is hidden from unknown readers. With further improvements on McEliece cryptosystem and a new arrangement, we are able to design a lightweight identification protocol that preserves privacy of RFID tags. In Section 5.1, we present our private identification scheme that is suitable for low-cost RFID tags. The security analysis is given in Section 5.2, which is then followed by the resource requirements

are given in Section 5.3.

5.1 Private Protecting Scheme

Our Privacy-Protecting Identification (PPI) scheme for an RFID system is given as follows: Let's denote the tag by T and the reader by R. Each reader R has a public-key F in the randomized McEliece cryptosystem with a given linear code $[N, k, d]$. The tag's identity is uniquely mapped to an element $a_{id} \in \mathbb{F}_{2^{k_2}}$, where k_2 is the size of the message block in the randomized McEliece cryptosystem. Note that in the original randomized McEliece [85], authors propose $k_2 < k_1$, in order to maximize the probabilistic effect and to provide pseudo-randomness.

Tag's identity is represented by a vector a_{id}, such that $a_{id} \in \mathbb{F}_{2^{k_2}}$. The tag's identity is concatenated with a random binary vector $r \in \mathbb{F}_{2^{k_1}}$ to form $[r|a_{id}]$. Recall that in the randomized McEliece cryptosystem [85], F can be denoted as $F^T = [F_1^T|F_2^T]$, where F_1 and F_2 are $k_1 \times N$ and $k_2 \times N$ sub-matrices of F, respectively. Thus, we can divide the encryption of $[r|a_{id}]$ in two parts as shown in Eq. (5.1).

$$y = c \oplus e = [r|a_{id}]F \oplus e = (rF_1 \oplus e) \oplus a_{id}F_2 \tag{5.1}$$

The first part, i.e. $a_{id}F_2$, is fixed and carries the tag's identity. The second part, i.e. $(rF_1 \oplus e)$, is randomized and changes every time to guarantee the probabilistic security in McEliece cryptosystem. At the setup when the tag is initialized, the trusted center calculates rF_1 and $a_{id}F_2$ for the RFID tag and stores them, as well as r in the tag's memory.

The tag's identity a_{id} is a fixed element and F is known for every authorized reader, so the trusted center can pre-compute the set $\{rF_1, a_{id}F_2, r\}$ and store it in the tag's memory at the setup. It should be noted that the tag does not need to store a large matrix F, but it stores only three small vectors of the set $\{rF_1, a_{id}F_2, r\}$. Let $h(\cdot)$ be a (hash) function that returns k_1 bits of its inputs. Note that $h(\cdot)$ can have a very simple design and does not need to be cryptographically secure. For instance, $h(\cdot)$ can be implemented by simply returning the first k_1 bits of its input. The identity of the tag is then securely transmitted to an authorized reader R via the PPI protocol shown in Algorithm 5.

When the tag T is queried, the reader first sends its identification to the tag. The tag then searches its memory to find the corresponding pre-computed values $rF_1, a_{id}F_2$ and r corresponding to the reader R. If T does not recognize R or does not find the corresponding values in its memory, it quits the protocol. If T does not quit in the previous step, it proceeds by picking a random error vector $e \in \mathbb{F}_{2^N}$, such that its weight is equal to $\lfloor \frac{d-1}{2} \rfloor$. It then computes $y = rF_1 \oplus a_{id}F_2 \oplus e$ and sends it to R. Computing y is very simple, as T only needs to retrieve the values of rF_1 and $a_{id}F_2$ from the memory and XOR them together along with e. Note that y is a valid McEliece ciphertext, which can be decrypted by R if it has the corresponding McEliece private key.

The reader R, on the other side, decrypts y to retrieve r, a_{id} and e. The reader picks new random vectors $r' \in \mathbb{F}_{2^{k_1}}$ and $t \in \mathbb{F}_{2^N}$. Using t, R generates a circular matrix denoted by A_r. It then prepares the response set $S = \{d_0, d_1, d_2, d_3\}$, where $d_0 = (r' \oplus h(e))$, $d_1 = (r'F_1 \oplus eA_r)$,

Algorithm 5 Private identification of RFID tags

1.	$T:$	find $\{rF_1, a_{id}F_2, r\}$ for R, o.w. quit
	$T:$	pick e randomly s.t. $w(e) = \lfloor \frac{d-1}{2} \rfloor$
	$T:$	compute $y = rF_1 \oplus a_{id}F_2 \oplus e$
	$T:$	$y \longrightarrow R$
5.	$R:$	decrypt y and get rF_1, e, a_{id}
	$R:$	pick a random $r' \in \mathbb{F}_{2^{k_1}}$
	$R:$	return $\{(r' \oplus h(e)), (r'F_1 \oplus eA_r),$
		$(rF_1 \oplus t), (r \oplus h(e))\}$
	$R:$	$S = \{d_0, d_1, d_2, d_3\} \longrightarrow T$
10.	$T:$	check if $(d_3 \oplus h(e)) = r$, o.w. quit
	$T:$	compute $d_2 \oplus rF_1$ and calculate eA_t
	$T:$	calculate $d_1 \oplus eA_t$
	$T:$	replace rF_1 by $d_1 \oplus eA_t$
	$T:$	replace r by $d_0 \oplus h(e)$ and quit

$d_2 = (rF_1 \oplus t)$ and $d_3 = (r \oplus h(e))$. The tag receives S from the reader and expands it to $S = \{d_0, d_1, d_2, d_3\}$. Then, T proceeds as follows: if its stored value of r from memory is not equal to $(d_3 \oplus h(e))$, it quits the protocol notifying that the reader R has failed to decrypt y correctly. Otherwise, it generates a circular matrix A_t from $d_2 \oplus rF_1$ and calculates eA_t.

Calculating eA_t can be done very efficiently, since A_t is a circulant matrix and e has only a few non-zero digits in it. Finally, T replaces rF_1 and r in the memory by $d_1 \oplus eA_r$ and $d_0 \oplus h(e)$, respectively. The next identification round is executed similarly with new values of $r'F_1$ and r'.

It is easy to check the correctness of the proposed protocol. The first condition that T checks is to verify $(d_3 \oplus h(e)) = r$, where $d_3 = r \oplus h(e)$. The tag then finds the random vector t by computing $d_2 \oplus rF_1 = (rF_1 \oplus t) \oplus rF_1 = t$. If a correct t is recovered, then $A_t = A_r$ and therefore eA_r at the reader is equal to eA_t at the tag. The operation $d_1 \oplus eA_t = (r'F_1 \oplus eA_r) \oplus eA_t$ returns $r'F_1$, which replaces rF_1 in the memory. The corresponding random vector r' corresponding to $r'F_1$ is retrieved from $(d_0 \oplus h(e)) = (r' \oplus h(e)) \oplus h(e) = r'$. The proposed identification scheme achieves several goals that are listed below:

Privacy: The proposed method protects privacy of the tags, since the tag uses a different F, so different rF_1 and $a_{id}F_2$ values, to communicate with various (known) readers. Furthermore, each rF_1 is updated with a new $r'F_1$ every time the tag is queried by the same reader. In other words, when the tag is communicating with the same reader many times, the messages exchanged between the tag and the reader are randomized every time and new parameters (except for $a_{id}F_2$) are generated that are not distinguishable to unauthorized

readers. For an adversary who is eavesdropping on readers, identifying a specific tag has two hurdles: (1) associating two different messages S and S' sent to the reader R to the same tag T^* and (2) matching the messages S_i and S_j $(i \neq j)$ sent to two different readers R_i and R_j, respectively to the same tag T^*. Later in Section 5.2, we show that both of these tasks are computationally impossible.

Protected Identities: Encrypting the identity of an RFID tag protects it from being queried by unauthorized readers. In other words, the tag always responds with an encrypted identity to the queries, and the requester has to have the decryption key (corresponding to F used in Algorithm 5) to be able to uncover the identity. Without a proper decrypting algorithm, the requester cannot retrieve the tag's identity. This scheme can be readily extended to encrypt tag's content as well as its identity; one would simply encrypt the entire data on the tag with a secret key before storing the data on to the tag. The secret key can be a variation of a_{id} that is obtained if the reader can retrieve a_{id} in the proposed identification method.

Efficiency: In the original McEliece cryptosystem, the size of public-keys is often large, but the proposed PPI method requires only pre-computed values of $a_{id}F$, not the entire public-key F to be stored in the tag. Even if the compressed version of McEliece $[N, k, d] = [450, 225, 56]$ found in [7] was not used, storing each $\{rF_1, a_{id}F_2, r\}$ in their original size $[N, k, d] = [2960, 2288, 56]$, as suggested in [9], would require merely $7.8K$ bits. Computing $y = rF_1 \oplus a_{id}F_2 \oplus e$ can be done very fast by XORing rF_1 and $a_{id}F_2$ values from the memory with e, which mostly contains $'0'$ except at 27 random positions (for $d = 56$). The most

complex operation at the tag is computation of eA_t. Note that A_t is a circulant matrix derived derived from vector t. Therefore, computing eA_t can be easily done by shifting t and XORing it to itself 27 times (for $d = 56$). The rest of the PPI protocol is performed with simple XOR operations.

Scalability: Using a PKC allows the system to expand practically to accommodate any number of tags and readers. In any RFID system, usually a reader is in interaction with a huge number (probably around thousands) of tags, where as each tag is only communicating with a few (maybe less than five) readers. Unlike other symmetric-key protocols proposed for RFID systems, the parameters that reader has to store in our protocol does not increase with the number of tags. On the other hand, the number of parameters for the RFID tag increase linearly with the number of readers. As we will see later in Section 5.3, the size of parameters to be stored on the RFID tag is not an issue even for low-cost tags, as the number of readers is usually kept small in most practical applications.

5.2 Security Analysis

We have shown that the RFID tags will require little memory along with capabilities for simple binary operations to perform the protocol. This provides a security and privacy enhancement for inexpensive RFID tags. We investigate the security of the proposed scheme in the rest of this section.

As mentioned before, the security and privacy of the proposed identification scheme comes from two assumptions: (1) it is difficult to distinguish a particular tag among many messages

sent to one reader R, and (2) it is also computationally difficult to track one RFID tag by messages it sends to many readers. It can be shown that if any of the above assumptions is not valid, one could build a protocol to solve the SD problem. We further analyze the security of the scheme in Algorithm 5 under these assumptions.

Assumption 1: Let's suppose that a tag, T, has communicated with the reader R and the tag's identity a_{id} is known. The adversary is given $a_{id}, F_1, F_2, a_{id}F_2$, where F_1 and F_2 are sub-matrices of a randomized McEliece public-key F. The adversary's challenge is to answer by observing the messages exchanged in the protocol, i.e. $y = rF_1 \oplus a_{id}F_2 \oplus e$ and $S = \{(r' \oplus h(e)), (r'F_1 \oplus eA_r), (rF_1 \oplus t), (r \oplus h(e))\}$, if the same T is present.

Informally, the adversary would need knowledge of t, in order to find eA_r and to recover $r'F_1$. Without knowing $r'F_1$, it is shown in [85] that the randomized McEliece cryptosystem is probabilistic and a_{id} cannot be distinguished from a random identity. Clearly without knowing t and eA_r, rF_1 and $r'F_1$ are completely masked in $(rF_1 \oplus t)$ and $(r'F_1 \oplus eA_r)$, respectively. Because t is an arbitrary binary vector, and it is therefore not disjoint from eA_r. The adversary could easily find $t \oplus e$ from $a_{id}F_2, y$ and $(rF_1 \oplus t)$, but this will fail to return A_r.

If the adversary attempts to extract either r, r' or e from $(r \oplus h(e))$ and $(r' \oplus h(e))$, he will not succeed, as there are three unknowns and only two equations. In every round of identification, new random vectors, i.e. r' and e, are selected, making it impossible to extract any of the unknowns from other identification sessions. Next, we formally show that if the above assumption is not valid, one could build a protocol to solve the LPN problem. We

need the following lemma before giving the full security proof.

Lemma 1. *Let $a, b \xleftarrow{R} \mathbb{F}_{2^N}$, that is a and b are binary vectors selected uniformly at random from $\mathbb{F}_{2^N}$. A circulant $N \times N$-matrix B is derived from b. The vector a, b and c are kept secret, but $(a \oplus b)$ and $(c \oplus aB)$ are disclosed. The vector $(c \oplus aB)$ is pseudo-random.*

To prove $(c \oplus aB)$ is pseudo-random, we show that if $(c \oplus aB)$ is *distinguishable* from a random vector, then the protocol that distinguishes $(c \oplus aB)$ with non-negligible advantage can also solve the LPN problem.

Proof: Let's assume that there is an algorithm that $\mathcal{A}$ that has a non-negligible advantage ϵ in distinguishing $(c \oplus aB)$ from a random vector. Suppose that there is a simulator algorithm $\mathcal{B}$ that receives a random instance of the LPN problem. That is the challenger takes a random looking $(q \times k)$-binary matrix A from $\mathcal{B}$ and returns $z = sA \oplus e$, where e is a random binary vector with probability distribution of the noise parameter $\eta \in (0, 1/2)$.

$\mathcal{B}$ chooses a random matrix T, and sends the matrices $(I \oplus T)$ and T to the challenger. The challenger returns accordingly $z_1 = s(I \oplus T) \oplus e_1$ and $z_2 = sT \oplus e_2$. Then, $\mathcal{B}$ calculates: $s(I \oplus T) \oplus e_1 \oplus (sT \oplus e_2)T^{-1} = e_1 \oplus e_2 T^{-1}$ and $s(I \oplus T) \oplus e_1 \oplus sT \oplus e_2 = sI \oplus e_1 \oplus e_2 = s \oplus e_1 \oplus e_2$.

In the next round, $\mathcal{B}$ forms a circulant matrix from $e_1 \oplus e_2 T^{-1}$, denoted by $E_1 \oplus E_2'$, and sends it to the challenger as the random matrix. The challenger, then, returns $s(E_1 \oplus E_2') \oplus e_3$. The simulator $\mathcal{B}$ is now ready to solve the LPN problem by using an algorithm $\mathcal{A}$ that solves Lemma 1. It simply sets $a \oplus b = s \oplus (e_1 \oplus e_2)$ and $c \oplus aB = s(E_1 \oplus E_2') \oplus e_3$ and sends them as the problem instance to $\mathcal{A}$. This instance corresponds to a problem parameterized as follows: $a = s$, $b = e_1 \oplus e_2$ and $c = sE_2' \oplus sE_2 \oplus e_3$.

This is a genuine association, since for $a \oplus b = s \oplus (e_1 \oplus e_2)$, we have to have $c \oplus aB = sE_2' \oplus sE_2 \oplus e_3 \oplus s(E_1 \oplus E_2) = s(E_1 \oplus E_2') \oplus e_3$ as given. It is easy to see that this instance has the same probability distribution as in the proposed PPI protocol. If $\mathcal{A}$ has any non-negligible advantage in solving Lemma 1, it can be used to find s and to solve the LPN problem.

Lemma 1 proves that the adversary's advantage in solving the lemma cannot be more than his advantage in solving LPN problem. The security of the proposed protocol is provided by the following theorem.

Security of PPI: Let's suppose that the tag, T, has communicated with a reader R and the tag's identity a_{id^*} is known. The adversary's challenge can be formulated as follows: the adversary is given $a_{id^*}, F_1, F_2, a_{id^*}F_2$, where F_1 and F_2 are sub-matrices of a randomized McEliece public-key F. Let's flip a fair coin b and set $y_b = r^*F_1 \oplus a_{id^*}F_2 \oplus e^*$ if $b = 0$, otherwise set $y_b = rF_1 \oplus a_{id}F_2 \oplus e$, where $id \neq id^*$ and r and e are appropriate random vectors. Send y_b to the adversary without revealing b. His challenge is to guess b with non-negligible probability from y_b, S^*, S, where $S^* = \{r' \oplus h(e^*), (r'F_1 \oplus e^*A_r), (r^*F_1 \oplus t), r^* \oplus h(e^*)\}$ and $S = \{r'' \oplus h(e), (r''F_1 \oplus eA_r''), (rF_1 \oplus t''), r \oplus h(e)\}$. It should be clear that S^* represents the set of messages communicated during identification of a_{id^*}, and S contains the messages communicated during identification of a_{id}. The adversary wins if he picks a_{id} (a_{id}'') when $b = 0$ ($b = 1$).

Theorem 2. *The adversary has a negligible advantage in correctly guessing b if the LPN problem is NP-Hard.*

Proof: The adversary has received a set of messages including $(r'F_1 \oplus e^*A_r), (r^*F_1 \oplus t)$ and $(r^*F_1 \oplus a_{id^*}F_2 \oplus e^*)$. We have assumed that the adversary knows $a_{id^*}F_2$, thus it can obtain $(r^*F_1 \oplus e^*)$. To prove that the adversary has any a negligible advantage in guessing b in the proposed PPI protocol, we prove that any noticeable advantage could solve Lemma 1. Simply calculate $(r^*F_1 \oplus t) \oplus (r^*F_1 \oplus e^*) = (e^* \oplus t)$. It is easy to see that $(e^* \oplus t)$ and $(r'F_1 \oplus e^*A_r)$ form a valid instance in Lemma 1, where $a = e^*$, $b = t$ and $c = r'F_1$. Thus, the adversary's advantage in guessing b is the same as his advantage in solving Lemma 1 and consequently solving the LPN problem.

Assumption 2: The adversary might try to distinguish a specific tag T by the messages communicated over various readers. When communicating with two authorized readers R_i and R_j $(i \neq j)$, the adversary's challenge is to find out if the same tag T is present. We show that if the adversary cannot successfully perform a cryptanalysis of McEliece ciphertexts, he cannot single out the tag T from other tags.

The messages sent from T to R_i and R_j are calculated using public-key F_i and F_j, respectively. Let's make the assumption that reader R_j is compromised and it reveals a_{id} as well as r_j. Then, the adversary can successfully track the tag T over R_i only if he can either compromise R_i or match messages communicated with R_i to T. Note that it is proven in [85] that knowledge of since $r_j \neq r_i$ McEliece cryptosystem is randomized and $r_iF_i \oplus e_i$ is indistinguishable from $r_jF_j \oplus e_j$ even if $F_j = F_i$. Therefore, the adversary cannot identify T, as it still needs to obtain r_i (or r_iF_i) to distinguish T.

Under these assumptions, we conclude that the proposed PPI scheme is secure. The

proposed scheme allows authorized readers to query and update tags' memory, while tracking the tag across multiple readers is (computationally) impossible. In the next section, the performance of the proposed PPI scheme is given in further details.

5.3 Resource Requirements of PPI Scheme

In the randomized McEliece cryptosystem, a random vector $r \in \mathbb{F}_{2^{k_1}}$ is concatenated with the identity vector $a_{id} \in \mathbb{F}_{2^{k_2}}$. Therefore, the size of a_{id} is less than the dimension of underlying codes ($k_2 < k$). According to the EPC standards [40], the identity vectors a_{id} are at least 96 bits in length (enough to store a Universal Product Code (UPC)). In McEliece cryptosystem $[N, k, d] = [450, 225, 56]$, we suggest to set $k_2 = 110$ leaving enough bits to represent a_{id} and $k_1 = 115$, which complies with the condition $k_2 < k_1$. Therefore, we will have $a_{id} \in \mathbb{F}_{2^{110}}$ and $r \in \mathbb{F}_{2^{115}}$. Our scheme can be directly realized in the low-cost RFID tags, even EPC-Gen 2 tags [46]. Table 5.1 summarizes the resources used in the proposed PPI scheme on the RFID tag and reader, and provides a comparison with HB+ and HB# protocols. The comparison is based on the suggested practical parameters for HB+ [60] and HB# [44].

The RFID tag needs to store two binary vectors, i.e. rF_1 and $a_{id}F_2$, of 450 bits each, and one random vector r of 115 bits in its memory, totalling $1,015$ bits which is less than $1K$-bit limit in the EPC-Gen 2 tags [46]. The scheme is only one round and the communication messages are limited to one vector $y \in \mathbb{F}_{2^{450}}$ from the tag to the reader. From the reader to the tag, the size of S is $1,130$ bits, as it is comprised of four binary vectors $d_1, d_2 \in \mathbb{F}_{2^{450}}$ and $d_0, d_3 \in \mathbb{F}_{2^{115}}$. Computations required on the tag are limited to checking $r = d_3 \oplus h(e)$

Table 5.1: Resource requirements of the proposed PPI scheme in comparison with HB+, HB#

Scheme	Side	Memory (bits)	Message (bits)	XORs	Key Sharing
HB+	Reader	800	8,000	8,000	Symmetric
	Tag	800	8,040	16,000	
HB#	Reader	1,472	80	194,481	Symmetric
	Tag	1,472	953	194,481	
PPI	Reader	101,250	1,130	202,500	**Asymmetric**
	Tag	**1,015**	**450**	**5,300**	

by 230 XORs, computing $t = d_2 \oplus rF_1$ by 450 XORs, calculating eA_t by around $4,000$ XORs (approximately $\mathcal{O}(N \log N)$) and finally evaluating $d_1 \oplus eA_t$ and $d_0 \oplus h(e)$ with 450 and 115 XORs, respectively. The number of XOR operations in the tag is estimated at $5,300$ XORs. The computation complexity for the reader is larger than the tag, and it is in the order of $202,500$ XORs.

In this chapter, we have presented a light-weight PPI scheme that keeps the identity of RFID tags hidden from eavesdroppers. The resources required to implement the proposed PPI scheme can be realized within the limited resources available on EPC-Gen 2 tag[46]. In the proposed PPI scheme, authorized readers can query and update the tag's memory, while tracking the tag is impossible even across authorized readers. In the next chapter, we will provide a novel solution that further protects the RFID tags against the relay attacks.

Chapter 6

Chaos-Masking against Relay Attacks

Securing RFID systems against the relay attack – a fairly simple yet effective attack against RFID tags – has been an obstacle for any authentication/identification scheme. In the relay attack, the adversary spoofs the identity of an authentic tag to acquire access to restricted resources. This is accomplished by relaying and presenting the credentials of the authentic tag in the authentication process. The relay attack is simple to launch, as the adversary is not required to break the authentication process and it only has to act as a proxy. The adversary transfers the challenge signals from an authorized reader to the tag without altering any bit of the signals. The relay attack can be launched against any wireless system, but the damage is more severe in RFID systems. This is mainly because the minimalist design of tags does not leave the capacity for a sophisticated authentication mechanism and the success rate of the relay attack will be significant.

Let's imagine that in a vicinity payment system, users can make a quick payment simply

by holding an RFID-enabled payment card close to a reader. The RFID cards are activated remotely. The payer's identification number is read from the RFID card he/she is holding. The payer is authenticated – often using the secret key stored in the RFID card – and billed if authentication is successful. Using the relay attack, an attacker could surreptitiously pass the bills to another bystander with an RFID card. The attacker succeeds by relaying the authentication messages between the reader and the victim's card. Successful cases of the relay attack has been reported in SpeedPass RFID cards [58], where a thief could get free gas by passing the charges to another SpeedPass card holder nearby.

In this chapter, a masking mechanism is proposed that can distinguish a legitimate reader from a proxy by checking for the watermark signal. The watermark signal is weak signal sent at a given range of frequencies and amplitudes to remove the chaotic signal from the RFID tag's signals. We provide an introduction to chaos and chaotic systems in Appendix A. In summary, the essential properties of a chaotic system are sensitivity to initial conditions and unpredictability. These are very good properties to be used in design of cryptographic systems. It has also been shown that two chaotic systems can be synchronized together [94], so it is possible to build secure communication system that transmit chaotic signals. Many chaotic systems have been developed in chaotic secure communication and chaotic spread spectrum communication [17, 27, 32, 71, 102], some of which have been broken mostly due to generalized chaos synchronization techniques [33, 99] or adaptive synchronization techniques [97, 96]. Instead of chaos synchronization, our proposed chaos-masking scheme is based on *chaos suppression*, which is described in full details in Section 6.1. The watermark signal is

in the form of chaos suppression signal that is known by legitimate readers only and removes the randomness from the RFID tag's signals. This is the first countermeasure against the relay attack based on chaos suppression. We describe the proposed method in details in Section 6.2. Discussions on the security of the proposed chaos-masking scheme are given in Section 6.3.

6.1 Chaos Suppression

A very interesting phenomenon in the study of chaotic systems is suppression of chaos by an external force. Chaos suppression is about using a small force, denoted by $f(t)$, to remove unwanted effects of chaos [29]. Choe et al. [29] show that applying a weak external force can harmonize the chaotic oscillators. The external force is a small perturbation on the chaotic system. It can be a periodic signal as $f(t) = k\cos(\omega t)$ or even a random signal as $f(t) = \epsilon(t)$. The notion $\epsilon(t)$ describes a bandwidth-limited noise with the power spectral density, which is given in Eq. (6.1).

$$P_\epsilon(\omega) = \begin{cases} p(\omega), & \omega_1 < |\omega| < \omega_2; \\ 0, & \text{otherwise.} \end{cases} \tag{6.1}$$

It is important to note that the chaos suppressing force is weak, i.e. k and $p(\omega)$ are small compared to the output power of the chaotic system. Clearly, a strong periodic force may dominate the system dynamics and turn it periodic, but it also dominates the overall system. It is interesting to find a small modification of the system that results removes. In order for

any nonlinear dynamic system to operate in the chaotic region, parameters of the system has to be set properly. Therefore, if a key parameter r is perturbed as $r(1 + f(t))$, it is possible to control the behavior of the dynamic system.

In the absence of a parametric driving force $f(t)$, the Lorenz equations demonstrate different dynamical regions on variation of system parameters. The behavior of Lorenz system changes drastically by changing the parameter r. Thus, the chaotic behavior of the Lorenz system can be diminished by controlling r. In chaos suppression, we apply a small external force $f(t)$ to oscillate the parameter r around its nominal value. In other words, parameter r in Eq. (A.1) is replaced by $r(1 + f(t))$, where $|f(t)| \ll 1$. Let's assume that the external force is a small periodic signal given as $f(t) = k\cos(\omega t)$. Note that changes to r due to $f(t)$ are small and only around r's nominal value.

Let's define the *characteristic frequency* of Lorenz system, denoted by ω_0, as the mean-time derivative of the phase of Lorenz system. Thus, ω_0 is given as in Eq. (6.2).

$$\omega_0 = \lim_{T\to\infty} \frac{2\pi N(T)}{T}, \tag{6.2}$$

where $N(T)$ is the number of turns performed in T. Assuming $\omega_0 \ll \omega$, Choe et al. [29] calculate an equivalent parameter (r_{eff} as given in Eq. (6.3)) that defines the slowly varying dynamics of Lorenz in the presence of an external force. Similar to the original unforced Lorenz system, the chaotic region of the slowly varying envelope functions is mostly affected by r_{eff}.

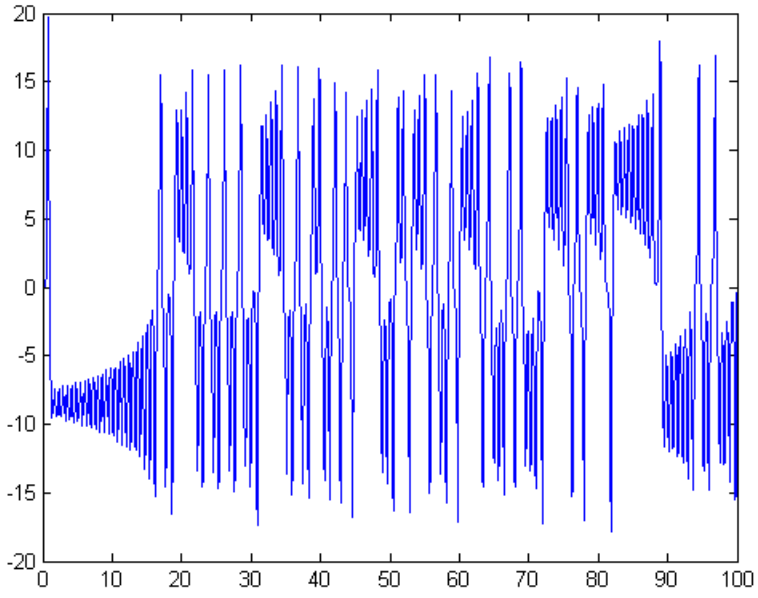
$$r_{\text{eff}} = r(1 - rK_\omega), \quad \text{where } K_\omega = \frac{\sigma k^2}{2\omega^2} \tag{6.3}$$

The unforced Lorenz system, which is parameterized with Saltzman parameters, i.e. $(\sigma, b, r) = (10, 8/3, 28)$, behaves chaotically and operates in Region 1. We would like to suppress chaos and move the system from chaos (Region 1) to stability at origin (Region 3). In order to have the averaged system operate in Region 3, we have to set $r_{\text{eff}} \leq 1$ and to bound parameters of $f(t)$, i.e. k and ω.

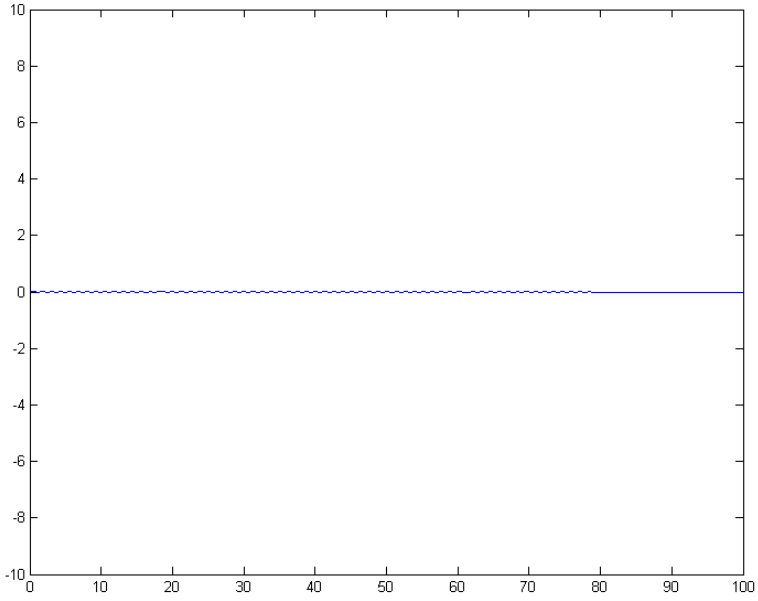
$$K_\omega \geq \sqrt{\frac{(r-1)}{r^2}} \tag{6.4}$$

If k and ω are chosen such that K_ω satisfies the condition in Eq. (6.4), the averaged system converges to the stability point at the origin, and chaos is suppressed. This is numerically simulated and shown in Figure 6.1 for Lorenz system with Saltzman parameters. The mean frequency of the unforced oscillations of the Lorenz system is found to be $\omega_0 = 8.24$ in Eq. (6.2) for the above parameters [29]. Set the frequency of the rapid periodic force to $\omega = 70$ ($\gg w_0 = 8.24$), and $k = 6.5$ to satisfy Eq. 6.4.

It can be seen from Figure 6.1 that chaos is diminished, and the state variables of Lorenz system converge to the origin at $(X, Y, Z) = (0, 0, 0)$. We have only shown the state variable x converges to 0 in Figure 6.1, but it is easy to check that state variables y and z also converge to the origin. This experiment shows that chaos can be removed by a small perturbation if the external perturbation force is positioned at a correct amplitude and frequency. This is the basis of our proposed chaos-masking scheme, which is explained in the following section.



(a) State variable x of Lorenz $(\sigma, b, r) = (10, 8/3, 28)$ is in chaotic region before a small perturbation is applied to r



(b) Chaos at state variable x is suppressed when a small perturbation signal $f(t) = 6.5\cos(70t)$ is applied to r

Figure 6.1: Suppression of chaotic behavior in Lorenz by a small periodic signal

6.2 Chaos-Masking Scheme

To increase the reading range in the relay attack, the attacker needs to place a proxy interfacing the authentic tag, as shown in Figure 6.2. The proxy intercepts the signals from the tag and sends them over to a mole interfacing the authorized reader. The mole masquerades the authentic tag's identity by reflecting its responses.

Over a long range, the attacker needs to establish a fast, reliable communication link between its proxy and mole. At the proxy, the attacker recovers the bits transmitted by the tag and sends them to its mole. Therefore, the proxy has to be able to demodulate and then decode the signals sent from the tag. From the proxy to the mole, the attacker has the same challenges facing classical communication systems; he needs to encode/modulate the signals at the transmitter and then demodulate/decode at the receiver. At the mole, the signal again has to be encoded/modulated according to the communication standards of the actual RFID tag/reader. Thus, the very first challenge in the relay attack is to recover bits transmitted by the RFID tag at the proxy. If the proxy cannot recover the transmitted bits, the relay attack fails. We design an RFID system in which chaotic system that confiscates the signals sent from an RFID tag. Only a proper watermark signal in the analogue realm can remove the chaos. The main idea is to place a watermark signal on the analogue signals, such that the adversary's processing of the signals loses the watermark.

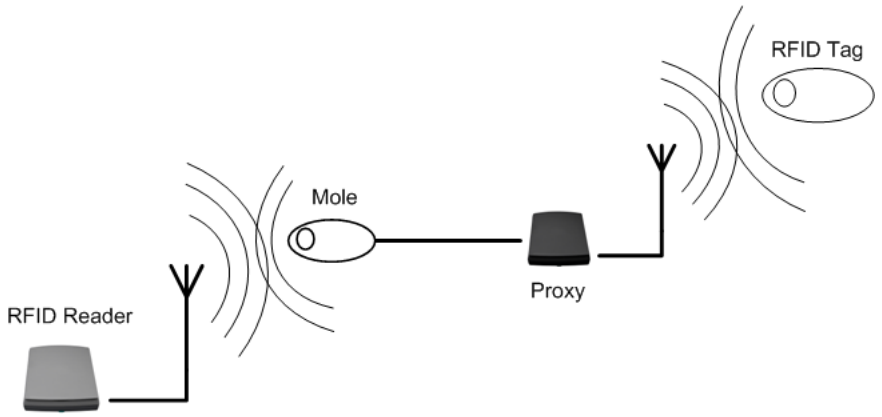


Figure 6.2: Proxy and mole are used to increase the range in the relay attack

6.2.1 Building Blocks

We continue to use the Lorenz system that is given in Appendix A. It should be noted that Lorenz system is not the only available chaotic system, and other chaotic systems might be used considering the implementation size, complexity and power consumption of the system with regards to the resources available at the tag. We design the watermark signal as a small perturbation to remove chaos in the Lorenz system. The general structure of tag-reader system that communicates using chaotic watermark signals is shown in Figure 6.3.

The chaotic system is initialized with random inputs. Truly random inputs can be readily derived from unique physical characteristics of the tag's circuitry, such as manufacturing variations, quantum mechanical fluctuations, thermal gradients, electromagnetic effects, and parasitics [12]. The RFID tag confiscates its data signal with the chaotic signal through a *mixer*. The mixer is a nonlinear function, denoted by $g(x, u)$, which combines the state variable x of Lorenz system with the data signal u. Our initial design of g is proposed

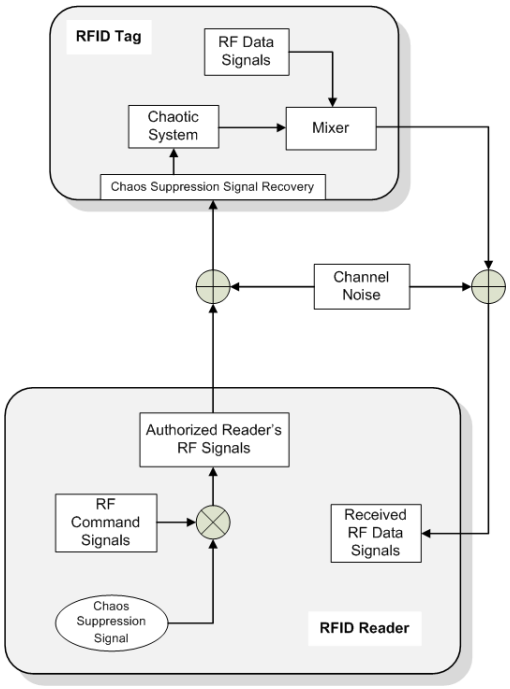


Figure 6.3: Tag-reader communication blocks combined with the chaos-masking system

as a simple Voltage Controlled Oscillator (VCO) that is controlled by the state variable x of Lorenz system. The output of the oscillator is a fixed amplitude signal with a varying frequency, depending on the input voltage. The quiescent frequency[1] f_q of the VCO is set at the frequency $100/(2\pi)$. The output of the mixer modulates the data signals of the tag. This is shown in Figure 6.4. Clearly, proper modulation is achieved only when the VCO is operating at the quiescent frequency.

Let's assume that the data signal is in the form of $u = a\cos(\omega_u t)$ and is amplitude-modulated by the carrier signal $c = \sin(2\pi f t)$. The sensitivity of the VCO should be high to make sure that the output of the mixer, in presence of chaos, rapidly changes. If we represent

[1]Quiescent frequency is the frequency of the VCO's output, when the input voltage is zero

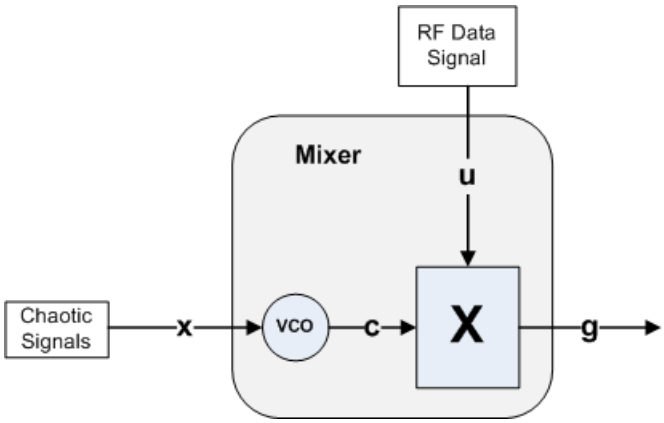


Figure 6.4: Circuit diagram of the chaos mixer $g(x, u)$

the behavior of the VCO by function $h(\cdot)$, the mixer g can be formulated as in Eq. (6.5). The impact of g on the security of the proposed scheme is discussed in Section 6.3.

$$g(x, u) = u \sin(2\pi h(x)t), \tag{6.5}$$

In the absence of chaos suppressing signal ($x \neq 0$), g returns a noise like signal with a varying frequency. If the watermark is a chaos suppressing signal, chaos at the tag is suppressed, i.e. $x \to 0$, and g returns the properly modulated data signal, i.e. $u \sin(2\pi f_q t)$. Therefore, only authentic readers capable of generating a chaos suppressing watermark could diminish chaos and communicate with the tags properly.

The circuitry of the proposed scheme is very simple both for RFID reader and tag. For the reader, it only requires generating and adding watermarking signals to its signals. The watermarking signal is a chaos suppressing signal in form of small periodic force. It is shown that any periodic signal that satisfies Eq. (6.4) can suppress chaos in the Lorenz system. Thus, the reader can change watermarking signals randomly and pick different signals, as

long as they satisfy Eq. (6.4). Moreover, the watermark recovery circuit at the tag, as shown in Figure 6.3, has a small footprint. By using a simple bandpass filter, the tag can easily retrieve the watermark signal. The chaotic system and the mixer can be implemented in RFID tags with a few integrators, adders and multipliers.

6.2.2 Simulation Results

In this section, we use MATLAB to numerically simulate the proposed chaos-masking scheme. The chaotic system is Lorenz initialized with Saltzman parameters $(\sigma, b, r) = (10, 8/3, 28)$. We inspect time and frequency responses of signals of various components in the proposed scheme throughout the simulation.

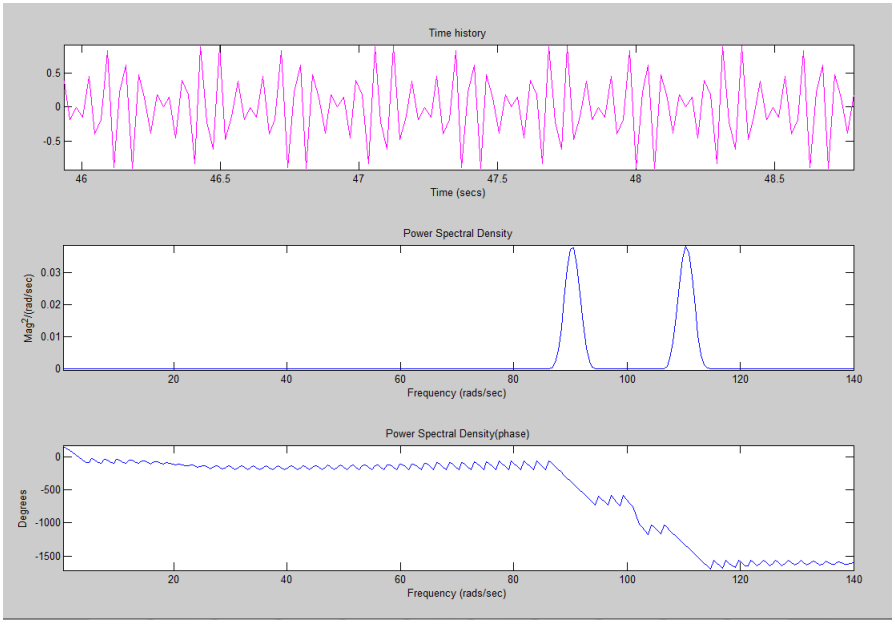


Figure 6.5: Modulated data signal that is output from the RFID tag

First, let's assume that the chaotic watermarking signal does not exist with the command

signal sent to the RFID tag. For simplicity in simulation, we have ignored the channel noise. Let's assume that the data signal $u(t)$ is a periodic wave with $\omega_u = 10rad/sec$ and $a = 1$, i.e. $u(t) = \cos(10t)$. This selection is only for demonstration purposes, and ω_u has to be readjusted to the standard underlying frequency. Let's set the frequency of the carrier and therefore the quiescent frequency f_q of the VCO at $\frac{100}{2\pi}$ Hz, where $\pi = 3.14$. The data signal is then modulated and transmitted as shown in Figure 6.5.

The mixer is defined as in Eq. (6.5). The output signal of the mixer $g(x, u)$ is shown in Figure 6.6, when no chaos suppression is introduced to the tag. In other words, $g(x, u)$ is the scrambled data signal that RFID transmits in the absence of a proper watermark. We have assumed that the watermark signal would be transmitted over a white Gaussian noise channel with SNR=10dB.

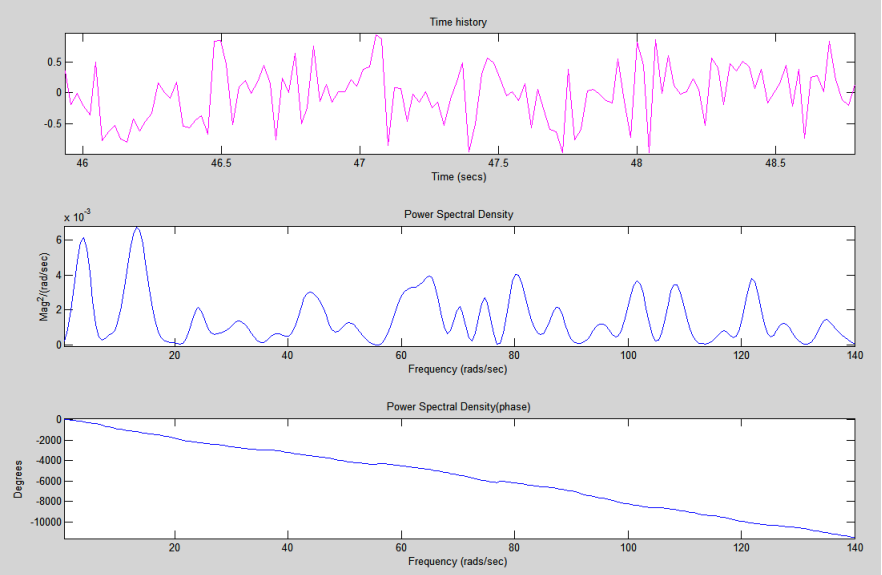


Figure 6.6: Output signal of the chaos mixer $g(x, u)$ without chaos suppression in presence of AWGN (SNR=10dB)

Let's apply the chaos suppressing watermark to the RFID tag. The data signal is as before $u(t) = \cos(10t)$. The watermark signal is generated as $f(t) = k\cos(\omega t)$, where $k = 6.5$ and $\omega = 70$. The watermark signal is sent over the same AWGN channel (SNR=10dB) and then applied to the parameter r in the Lorenz system. The output of the mixer $g(x, u)$ is depicted in Figure 6.7, where it is shown that chaos is suppressed completely and the output of the mixer, when a proper chaos-suppressing watermark is present, equals the original RF data signal in Figure 6.5.

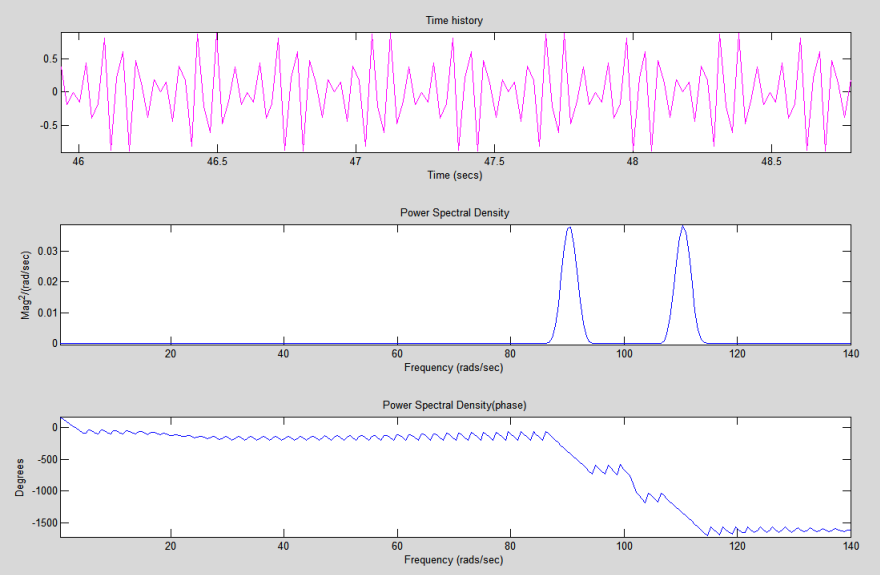


Figure 6.7: Output signal of the chaos mixer $g(x, u)$ with a proper chaos suppressing signal, $f(t) = 6.5\cos(70t)$, in presence of AWGN (SNR=10dB)

Let's now assume that the attacker attempts to generate an estimate of chaos suppressing watermark in an attempt to suppress chaos and remove the mask from data signals. The adversary needs to guess the amplitude and the frequency of a proper watermark signal that satisfy Eq. (6.4), where the Lorenz system parameters are hidden from the attacker. So, the

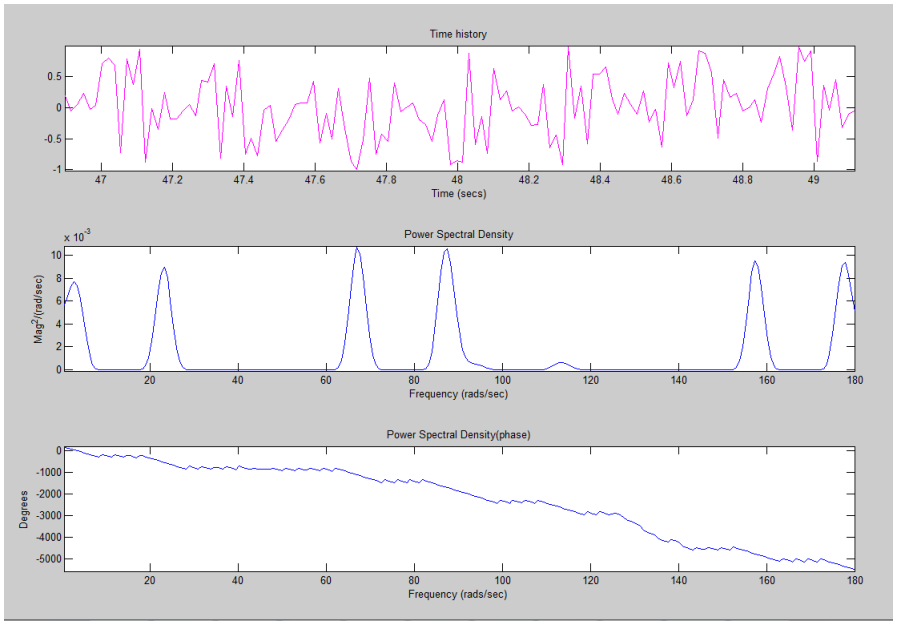
attacker will have to guesstimate the watermark signal.

First, we assume that the adversary's watermark signal $f'(t)$ is correctly estimating the amplitude of the watermark, i.e. $f'(t) = k\cos(\omega' t)$. In other words, the adversary's watermark signal is skewed in frequency ($\omega' \neq \omega$). The effect of a watermark signal, which is skewed 30% in frequency from the proper frequency is shown in Figure 6.8. The adversary might pick a watermark signal at such a high frequency that Eq. (6.4) is not satisfied. As shown in Figure 6.8a, chaos is not suppressed and the data signal is not recoverable.

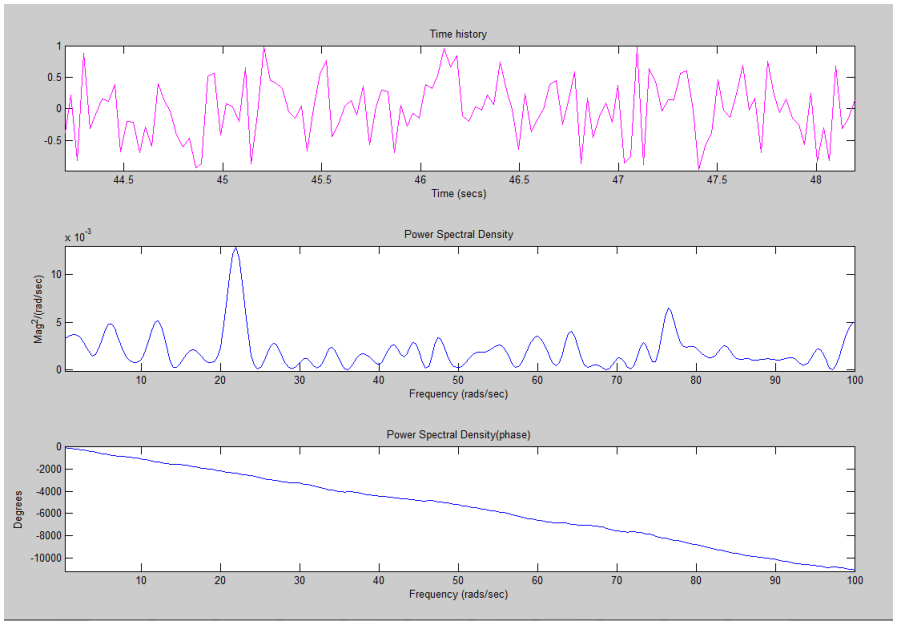
In Figure 6.8b, the watermark signal does indeed satisfy the suppression condition given in Eq. (6.4). However, the frequency $\omega' = 50$ is not big enough compared to the characteristic frequency $\omega_0 = 8.24$. Therefore, Eq. (6.4) is not a valid condition anymore. The effect of an external force with a low frequency ($\omega' = 50$ rad/sec) on the chaotic system is shown in Figure 6.9. We can see from Figure 6.9 that the state variable x does not converge to the origin, although the chaos suppression condition in Eq. (6.4) is satisfied – for a better comprehension, contrast x in Figure 6.1 to x in Figure 6.9.

The adversary's watermark signal can also be deflected in amplitude, i.e. $f'(t) = k'\cos(\omega t)$. In Figure 6.10, the output of the mixer is shown for watermark signals, which deviate in amplitude (k') from the proper k that satisfies Eq. (6.4).

It can be seen from Figure 6.10a that the suppression condition in Eq. (6.4) is not satisfied. Thus, chaos is not suppressed in Figure 6.10a and the output signal of the mixer is combined with chaos. When the amplitude of external force k' is much greater than k, as shown in Figure 6.10b, it is not considered a small perturbation. A large external force rules



(a) $f'(t) = 6.5\cos(90t)$



(b) $f'(t) = 6.5\cos(50t)$

Figure 6.8: Output of the chaos mixer $g(x, u)$ for frequency-skewed watermarks

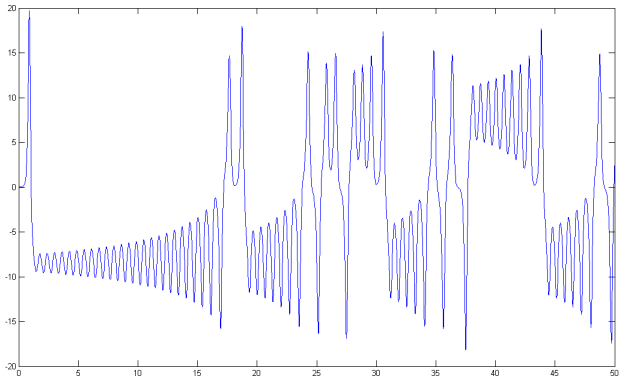
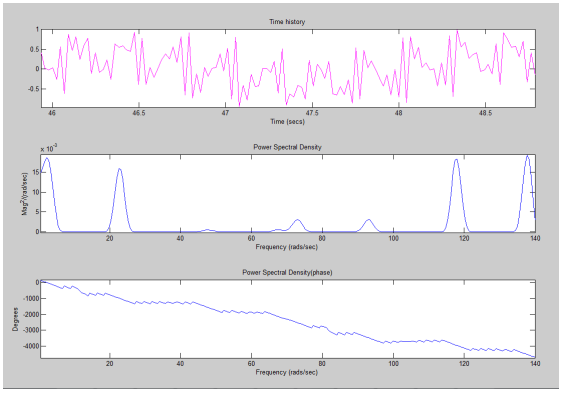


Figure 6.9: State variable x in chaotic Lorenz when a low frequency perturbation is applied

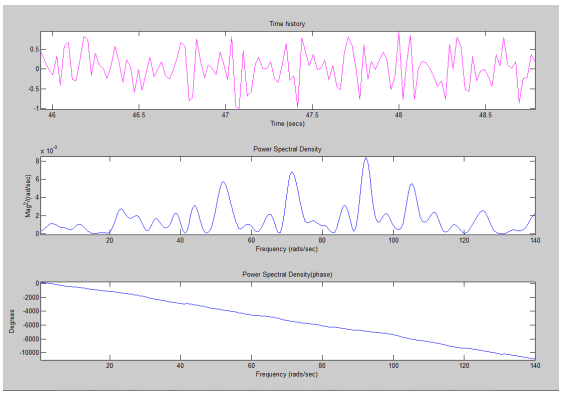
over the chaos suppression and drives the Lorenz system away from its stability point at the origin. Therefore, the output of the mixer has masked the data signal. It should be noted that for larger values of k, the watermark still satisfies Eq. (6.4), and as seen in Figure 6.10c, the data signal can be correctly demodulated.

6.3 Security Discussions

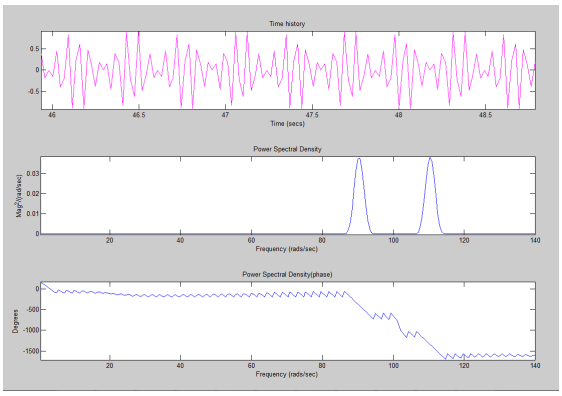
The chaotically secure schemes are essentially different from the classical cryptography, and it is difficult to investigate their security within the classical cryptanalysis methods. A prime difference is that the proposed scheme is designed in the analogue realm, where no classical cryptanalysis exists. Thus, it is necessary to test the security of the chaos-masking system with new cryptanalysis methods. Therefore, the adversary will need to use other techniques to break the system. We have identified three possible methods to break the proposed scheme; methods based on chaotic synchronization, analyzing the mixer and searching for



(a) $f'(t) = 4.5\cos(70t)$



(b) $f'(t) = 10\cos(70t)$



(c) $f'(t) = 8.5\cos(70t)$

Figure 6.10: Output of the chaos mixer $g(x, u)$ for amplitude-skewed watermarks

the watermark signal.

6.3.1 Chaotic Synchronization Attacks

The essential characteristic of chaos is sensitivity to initial conditions. This means that any trajectory that starts in a point infinitely close to a chaotic trajectory will diverge exponentially from it. Since, infinite precision in measuring the initial conditions is impossible, chaotic systems are unpredictable. However, it is well known in chaos theory that one can find an approximation system that synchronizes with the chaotic system [94]. The approximation system generates signals that closely follow the chaotic system.

Most secure chaotic systems have been broken due to chaotic synchronization methods. In the chaotic synchronization method, the adversary usually succeeds in generating a signal that follows the chaos very closely, even without knowing the initial conditions of the chaotic system that generated the chaos. Using the synchronized signal, the adversary removes (or reverses) the chaos to obtain the data signal. Many developments have been made in this field, such as generalized chaos synchronization [99], synchronization errors [77], parameter-adaptive synchronization [33], chaotic synchronization based on random optimization [92], adaptive synchronization techniques [97, 96] and dynamic minimization algorithm [73]. All of these chaotic systems are based on chaotic synchronization between a drive and response system. The synchronization signal is transmitted from the drive system to the response system via a direct link (usually in air). The link is necessary for the drive and response system to synchronize together and establish a reliable communication channel. The synchronization

signal often unveils one (or many) of the state variables of the underlying chaotic system to the response system. This synchronization signal is mostly used to break the underlying chaotic communication system.

The proposed scheme is based on chaotic suppression rather than chaotic synchronization theory. Thus, no state variable or synchronization link exist that can directly lead to the system's compromise due to synchronization attacks. The adversary has to invert g to recover $h(x)$ and reverse it for synchronization. The output of the mixer is the only visible signal with information about the chaotic system, that is x in $u\sin(2\pi h(x)t)$. It is easy to see that the output of the mixer (g) hides the states variables of the underlying chaotic system. For any $g = u\sin(2\pi h(x)t)$, there exists infinitely many possible $h(x)$ and x-s. We are not aware of any (adaptive) synchronization method that synch directly with $h(x)$ by using the outputs of g. Thus, we can conjecture that none of the state variables of the underlying chaotic system is observable, and no link is available to the adversary. Without having access to any observable state of the chaotic system or knowing its parameters, no synchronization can be found and the proposed scheme is secure against synchronization-based attacks.

6.3.2 Mixer's Security

The characteristics of the underlying chaotic system are confiscated by applying a nonlinear function on state variables ($h(x)$) and using a periodic function in the mixer $u\sin(2\pi h(x)t)$. Therefore, characteristic-based attacks that aim at finding an approximation of the chaotic system using its intrinsical characteristics, are no longer effective. This includes: short-time

zero-crossing rate of chaotic switching [99, 101] and the spectrogram property of chaotic orbit [67, 100]. Various mixers with a sophisticated design can be used to hide the intrinsic characteristics of the chaotic system. The proposed mixer, as given in Eq. (6.5), is a simple function that can be easily implemented in small RFID tags.

6.3.3 Space of Possible Watermarks

Assuming that the mixer completely hides the state variables and prevents characteristic-based attacks, the adversary's attack would be limited to searching for a proper watermark. There are two sets of parameters in the proposed scheme that help narrowing the search for a proper watermark: the parameters of the underlying chaotic system and the parameters of the chaos suppression signal (watermark). To successfully regenerate a watermark signal, the adversary needs to find parameters of a proper watermark (k and ω) that satisfy Eq. (6.4). However, Eq. (6.4) depends on the parameters of the underlying chaotic system. Eq. (6.4) is rewritten in Eq. (6.6) to express the suppression condition in terms of Lorenz system parameters.

$$\frac{k}{\omega} \geq \sqrt{\frac{2(r-1)}{\sigma r^2}} \tag{6.6}$$

The adversary's exhaustive search for a proper watermark has to return a range of $\frac{k}{\omega}$ that satisfies Eq. (6.6). In order to estimate the key space, we need to find applicable chaotic parameters of Lorenz system. Lorenz system has been studied extensively and its chaotic regions are widely known to researchers [3, 4, 29]. The selection of parameters for

the underlying chaotic system is limited to a range in which the system behaves chaotically.

The chaotic parameters of Lorenz system have been studied partially for fixed values of σ and b (usually $\sigma = 10, b = 8/3$) and a range of r. For large values of r, different situations appear; for the Saltzman classical value $r = 28$ there is a globally-attracting chaotic attractor. For larger values of r, a large chaotic region appears up to $r \simeq 146$, then a regular region up to $r \simeq 166$, then another chaotic region up to $r \simeq 214$ and then a regular region [4]. Barrio and Serrano [4] prove by both numerical and theoretical analysis that the chaotic range is bounded on $\sigma - b$ plane for a fixed $r > 0$. The chaotic parameters of Lorenz system is contained in the sphere given in Eq. (6.7). The chaotic region will be inside this sphere.

$$\Omega = \{(x, y, z) | x^2 + y^2 + (z - \sigma - r)^2 \leq R^2\},\ R^2 = \begin{cases} \frac{(\sigma+r)^2 b^2}{4(b-1)}, & \sigma \geq 1, b \geq 2; \\ (\sigma + r)^2, & \sigma > \frac{b}{2}, b < 2; \\ \frac{(\sigma+r)^2 b^2}{4\sigma(b-\sigma)}, & \sigma < 1, b \geq 2\sigma. \end{cases} \tag{6.7}$$

We can observe in Eq. (6.7) that σ and b in the chaotic region are covering a portion of the $b - \sigma$ plane for a fixed r. Therefore, chaotic parameters are not independent from each other. It should also be noted that Eq. (6.6) is independent of b. Nevertheless, parameter b will affect the characteristic frequency (ω_0) of Lorenz system and will vary the suppressing frequencies (ω). For a fixed b, we pick the ranges for $0 < \sigma < 200$ and $24.74 < r < 500$. Let's assume that the RFID tag only allows watermarks for which the k/ω ratio can be at maximum 30% more than the passing condition in Eq. (6.6). This will decrease the adversary's chance of success.

By knowing the system parameters, the attacker can chose a proper k and ω. The range of k and ω is also bounded due to practical reasons; the amplitude or frequency of the watermark cannot be arbitrarily chosen. The range of ω is set by the bandwidth of the bandpass filter at the RFID tag. Also, k is bounded by the maximum power of the watermark signal. To be able to estimate the size of the key space, we should find the sensitivity of chaos suppression in Lorenz system to parameters that still satisfy Eq. (6.6). To have a better chance of success, the adversary would increase k and lower ω as much as possible. As seen in Figure 6.7 and Figure 6.10c, both $k = 6.5$ and $k = 8.5$ suppress chaos at a given frequency $\omega = 70$ rad/sec. It is important to hide ω from the adversary, as it is the key in finding a correct k' satisfying Eq. (6.6). It is also possible to narrow the search for ω by finding the bandwidth of the bandpass filter and trying only frequencies within the bandwidth of the filter. Not knowing the suppression frequencies, the adversary is deterred by difficulties in reverse engineering the tag's hardware. Practicality of components used in the proposed has been tested in an implementation by Choe et al. [29]. This allows for greater acceptance and integration in industry.

A key component in our chaos-masking scheme is mixing chaotic signals with data signals at the RFID tag. In our initial design, we have suggested to use a non-linear mixer as in a VCO, which can be readily implemented in an RFID tag. In future work, we intend to investigate the effects of various mixers on performance of the proposed scheme. Our design has been in the analogue realm at the RFID tag. However, it is possible to design systems that implement a digital watermark recovery method inside RFID tags. Our scheme has been

based on Lorenz chaotic system, since it has been widely studied in theory and practice. We have used Lorenz system only in Region 1 (chaos) and Region 3 (stability at origin). It is interesting to see how Lorenz system can be used in Region 2 (stability at two points) to stop the relay attack. Not only Lorenz system can be used in the proposed scheme, but also any other chaotic system might be used in the proposed scheme as well. A comparative study of various chaotic systems and their security when applied in the proposed scheme will be necessary.

Chapter 7

Key Management & Secure Broadcast Encryption

In every symmetric or asymmetric authentication system, cryptographic keys must be distributed throughout the system via a secure channel. Establishing a secure channels between all the (remote) entities in a distributed system, such as RFID system, is a great challenge. Firstly, the cryptographic keys must be only delivered to their intended recipient or group of recipients. Secondly, distribution of keys should not overload the network, as new keys and updates will be frequently broadcast through the network.

Let's assume that a secret message, such as the price of an item, needs to be communicated with a few selected RFID tags at certain domains. For example, the prices or any other settings have to be changed on Readers 1, 2, 4 and 5 that are logistically grouped into Readers Group A and B, as shown in Figure 7.1. Note that these readers might have different

roles or be located at different domains. The server in Figure 7.1 should be able to send a short broadcast message that can only be accessed by the privileged readers. The broadcast message has to be encrypted to be secure and be short to avoid flooding the network. Readers have different public/private keys pairs, and the broadcast message needs to be generated using every (privileged) reader's key. Any colluding subset of readers in the network must not be able to access the broadcast message if they are excluded from the set of privileged recipients.

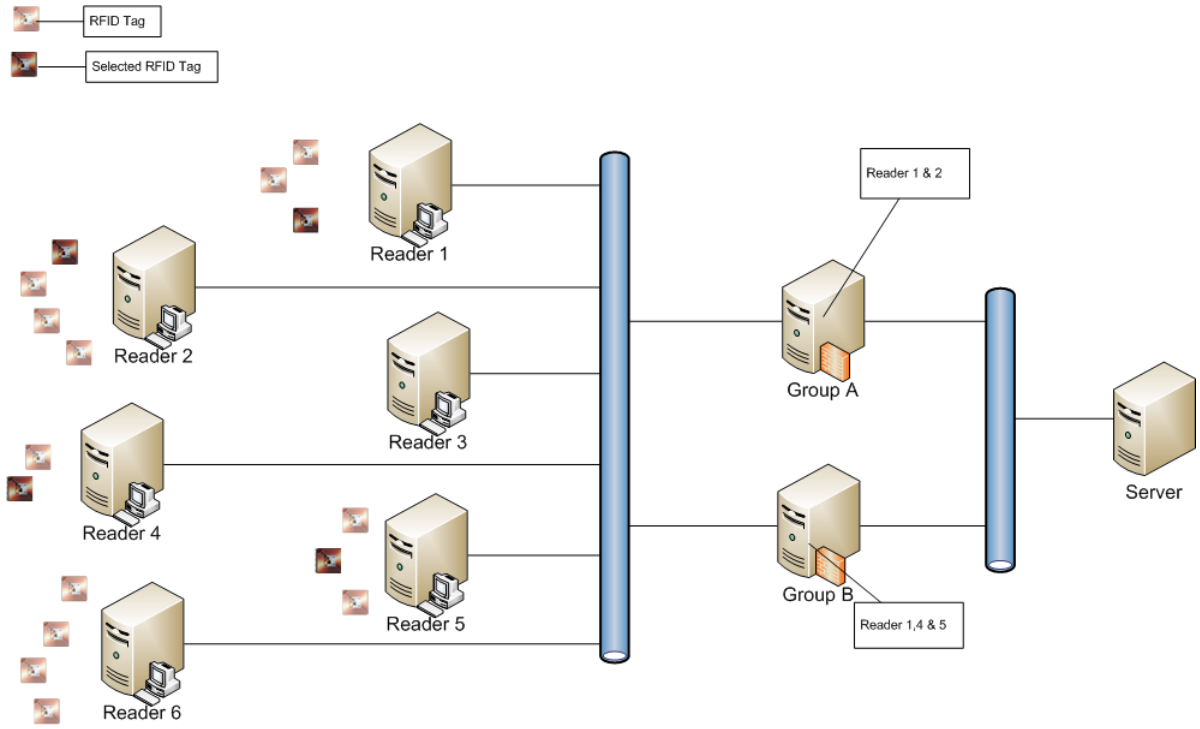


Figure 7.1: Sharing secrets among a dynamic set of RFID readers in a broadcast group

Readers can arbitrarily select any subset of members of a broadcast group for sharing a cryptographic key. Members leave and join the group depending on the credentials

they receive from the group manager at any time. We refer to the group manager as Administrator (Admin) who is responsible for managing the group and distributing keys to group members. The Admin can also be viewed as a central authority who manages the entire group membership and is responsible for sharing the public keys with members of the broadcast group.

In this chapter, we build the first broadcast encryption scheme that is secure in a fully adaptive model and has short ciphertexts. After the preliminaries in Section 7.1, the main Broadcast Encryption (BE) protocol is given in Section 7.2. Security of the proposed protocol in its underlying attack model is formally proved in Section 7.3. The security model is based on a known intractable problem and provides a very strong model that closely simulates the adversary in the real world. Our model does not require implementation of hash functions, i.e. the proof is completed without using a random oracle. In Section 7.4, the complexity of the proposed broadcast encryption scheme is discussed.

7.1 Preliminaries

In this section, we begin by giving a formal definition of a broadcast encryption system. Then, we present the adaptive security model in which our broadcast encryption system is secured. Later in this section, we introduce the cryptographic primitives used as the security basis of our scheme.

7.1.1 Broadcast Encryption Systems

A public-key broadcast encryption system is defined in the following framework: The BE scheme is comprised of four algorithms:***Setup(***λ, n***)***, ***KeyGen(***i, SK***)***, ***Encrypt(***S, PK***)*** and ***Decrypt(***$S, i, D_i,$ Hdr$, PK$***)***.

Setup(λ, n***):*** Takes as input the number of receivers (n) and the security parameter λ of a broadcast recipient group. It outputs a public/secret key pair $\langle PK, SK\rangle$ belonging to the `Admin`. Note that SK is called a secret key, as the security of the given broadcast encryption system depends on it.

KeyGen(i, SK***):*** Takes an input an index $i \in \{1, \cdots, n\}$ and the secret key SK. It outputs a private key d_i for the i-th member (identity). We will see later that this private key is used for decryption in the ***Decrypt()*** algorithm.

Encrypt(S, PK***):*** Takes as input a subset $S \subseteq [1, n]$ and a public key PK. If the size of the subset ($|S|$) satisfies $|S| \leq l$, it outputs a pair $\langle \text{Hdr}, K\rangle$, where Hdr is called the header and $K \in \mathcal{K}$ is a message encryption key. We will show later that K is used as the encryption key and Hdr contains data for intended recipients to find the encryption key. The broadcast to members in S consists of $\langle S, \text{Hdr}\rangle$.

Decrypt($S, i, D_i, \text{Hdr}, PK$***):*** Takes as input a subset $S \subseteq [1, n]$, an index $i \in \{1, \cdots, n\}$,

private key D_i corresponding to i, a header Hdr for the given S and the public key PK. If $|S| \leq l$ and $i \in S$, then the algorithm outputs the message encryption key $K \in \mathcal{K}$.

7.1.2 Security Model

The security of our protocol is defined in the chosen ciphertext security against an *adaptive* adversary using the following game between the *challenger* and an attack algorithm $\mathcal{A}$. The challenger builds a BE scheme and the attack algorithm tries to compromise it. The game is organized as follows: Both $\mathcal{A}$ and the challenger are given n and λ in the beginning. The adversary is adaptive; that is it does not need to commit to a subset of members before seeing the public parameters PK. We improve the security model of Gentry and Waters' [43] by adding the decryption query round in which the adversary, in addition to adaptively obtaining the private keys of the attack set, can send decryption queries to the challenger for the challenge set. The security model presented in this chapter represents a stronger model, as it captures a wider range of attacks. It is therefore closer to the adversary in real world as compared to others in the literature. Our model is defined as follows:

Setup: The challenger runs ***Setup***(λ, n) to obtain a public key PK, which is then revealed to the adversary.

Key Query Phase: Algorithm $\mathcal{A}$ adaptively issues private key (D_i) queries for any set of indices $S' \subset [1, n]$.

Challenge: The challenge set is specified as $S^* = [1, n] \setminus S'$. Note that for all private keys (D_i) of member i queried in the **Key Query Phase**, we have $i \notin S^*$. The challenger then runs ***Encrypt***(S^*, PK) and outputs $\langle \mathrm{Hdr}^*, K \rangle$. The challenger secretly picks a random $Z \xleftarrow{R} \in \mathcal{K}$. It then sets $b \xleftarrow{R} \in \{0, 1\}$ and returns $\langle \mathrm{Hdr}^*, K^* \rangle$ to the adversary, where $K^* \leftarrow K$ if $b = 0$, otherwise $K^* \leftarrow Z$.

Decryption Query Phase: The adversary issues adaptively decryption queries $q_1, \cdots, q_D$, where a decryption query consists of the triple (i, S, Hdr) for any $S \subset [1, n]$ including $S \subset S^*$. The only constraint is that $\mathrm{Hdr} \neq \mathrm{Hdr}^*$. The challenger responds with ***Decrypt***$(S, i, D_i, \mathrm{Hdr}, PK)$.

Guess: The adversary uses algorithm $\mathcal{A}$ to output its guess $b' \in \{0, 1\}$ for b and wins the game if $b' = b$.

We refer to the game described above as the adaptive Chosen Ciphertext Attack (CCA). Using an algorithm $\mathcal{A}$ to break the broadcast encryption system (BE) with parameters (λ, n), i.e. to guess the correct value of b, the adversary's advantage is defined as follows:

$$Adv_{\mathcal{A},\mathrm{BE}}(\lambda, n) = |Pr[b' = b] - \frac{1}{2}|,$$

where b' is the algorithm $\mathcal{A}$'s guess of b. If the adversary's guess is correct, the attack algorithm $\mathcal{A}$ has succeeded in compromising the BE scheme.

Definition 4. *A broadcast encryption system* BE *is adaptively* $(\mathrm{negl}(\lambda), n, q_D)$ *CCA secure*

if for all polynomial-time algorithms $\mathcal{A}$ *that make a total of* q_D *decryption queries, we have* $Adv_{\mathcal{A},\mathrm{BE}}(\lambda, n) = \mathrm{negl}(\lambda)$. *The adversary has a negligible advantage if* $\mathrm{negl}(\lambda)$ *can be made smaller than* $\frac{1}{\mathrm{poly}(\lambda)}$ *for any arbitrary polynomial* $\mathrm{poly}()$.

We make extensive use of *bilinear maps* at the core of our proposed BE scheme, but first we need to define it.

7.1.3 Bilinear Maps

Let $\mathbb{G}$ and $\mathbb{G}_T$ be groups of order p, and let g be a generator of $\mathbb{G}$. A bilinear map is an efficiently computable function from $\mathbb{G} \times \mathbb{G}$ onto $\mathbb{G}_T$, such that it has the following properties:

1. *Bilinearity*: For all $g, g', h, h' \in \mathbb{G}$,

$$e: \mathbb{G} \times \mathbb{G} \to \mathbb{G}_T,$$

$$e(gg', h) = e(g, h)e(g', h),$$

$$e(g, hh') = e(g, h)e(g, h')$$

   Note that $e(\cdot, \cdot)$ is symmetric, that is $e(g^a, g^b) = e(g^b, g^a) = e(g, g)^{ab} \quad \forall a, b$.

2. *Non-degeneracy*: If $e(g, h) = 1$ for all $h \in \mathbb{G}$, then $g = I$ (identity).

Weil pairing and Tate pairing are two implementations of an efficient bilinear map over elliptic curve groups useful for cryptography. For a more detailed discussion on bilinear maps and pairings, we refer the reader to [11]. Bilinear maps for cryptography has to have certain complexities to be used in cryptographic algorithms. This is explained further in the following section.

7.1.4 Complexity Assumptions

The security of our schemes is based on a complexity assumption that has been used extensively in cryptographic algorithms [13, 14, 16, 43]. Complexity assumptions found in the literature have slightly different settings, but they are all related to the difficulty of solving Discrete Logarithm Problem (DLP) over large algebraic group. Our main construction, which is given later in Section 7.2, is based on a narrower variant of the DLP assumption, referred to as the Bilinear Diffie-Hellman Exponent (BDHE)-Sum assumption. This is the same complexity assumption that has been used in Gentry and Waters' scheme [43]. We have simplified the definition to relate directly to our security proof.

Definition 5 (BDHE-Sum Assumption (for n):)**.** *As usual, let $\mathbb{G}$ and $\mathbb{G}_T$ be groups of order p with bilinear map $e : \mathbb{G} \times \mathbb{G} \rightarrow \mathbb{G}_T$, g a generator for $\mathbb{G}$ and $\alpha, s \xrightarrow{R} \mathbb{Z}^*{}_p$. Set $S = [-2n, 2n]$. Given $\{y_i = g^{\alpha^i} : i \in S\}$, compute $e(g, g)^{\alpha^{4n+1}}$, without knowing α.*

There is also a decision-variant of the BDHE-Sum assumption, which is stated as follows:

Definition 6. *Let $\widehat{y}_{g,\alpha,n} = \{y_i = g^{\alpha^i} \ \forall i \in S\}$. An algorithm $\mathcal{B}$ that outputs $b \in \{0, 1\}$ has advantage ϵ in solving the decision BDHE(-Sum) for n in $\mathbb{G}$ if*

$$Pr\left[\mathcal{B}(g, \widehat{y}_{g,\alpha,n}, e(g, g)^{\alpha^{4n+1}}) = 0\right] - Pr\left[\mathcal{B}(g, \widehat{y}_{g,\alpha,n}, Z) = 0\right] \geq \epsilon,$$

where the probability is over the random choice of the generator $g \in \mathbb{G}$, the random choice of $\alpha \in \mathbb{Z}^*{}_p$, the random choice of $Z \in \mathbb{G}_T$, and the random bits consumed by $\mathcal{B}$. We refer to the distribution on the left as Pr_{BDHE} and the distribution on the right as R_{BDHE}.

We say that the (decision) (ϵ, n)-BDHE-Sum assumption holds in $\mathbb{G}$ if no polynomial-time algorithm has significant advantage greater than ϵ in solving the (decision) BDHE-Sum problem for a given n in $\mathbb{G}_T$.

7.2 Adaptively Secure Broadcast Encryption

We refer to BE scheme in the adaptive model as BE_A. Let's denote the maximum number of members in the multicast group by n. Our BE_A scheme is given in the following:

Setup(λ, n***)*** Run $\langle \mathbb{G}, \mathbb{G}_T, e \xleftarrow{R} \mathit{GroupGen}(\lambda, n)\rangle$. Set $\alpha \xleftarrow{R} {\mathbb{Z}^*}_p$, the generator $g \in \mathbb{G}$, identity values $x_1, \cdots, x_n \xleftarrow{R} \mathbb{G}^n$ and two secret values $\gamma \xleftarrow{R} {\mathbb{Z}^*}_p$. Set PK to include a description of $\mathbb{G}, \mathbb{G}_T, e$, $\{x_1, \cdots, x_n\}$, $\{g^{\alpha^i}, \forall i \in [0, 2n]\}$ and $e(g, g)^{\alpha^{2n+1}}$ as the session key. The group's secret SK is set as $\langle \gamma, \alpha \rangle$, which is known by Admin only. Output $\langle PK, SK \rangle$.

KeyGen(i, SK***)*** Pick $r_i \xleftarrow{R} {\mathbb{Z}^*}_p$ and for all $j \in [0, n]$, pick randomly $B_j \xleftarrow{R} {\mathbb{Z}^*}_p$. Release to member i the following set of private keys $D_i \leftarrow \{r_i, d_{i,j}, T_{i,j}\}$, where:

$$d_{i,j} = g^{\gamma \alpha^n B_j \frac{\alpha^2 - r_i}{\alpha - x_i}} \ \forall\ j \in [1, n] \text{ and } j \neq i$$

$$T_{i,j} = g^{\frac{\alpha^j}{\gamma B_i}} \ \forall\ j \in [0, n]$$

We emphasis that r_i and $d_{i,j}$ values are used for decryption and $T_{i,j}$ values are used to create the broadcast encryption message.

Encrypt(S, i, D_i, PK***)*** The set S includes the index of members for which the message

will be sent, as well as the index of the encrypting member i. Pick $t \xleftarrow{R} \mathbb{Z}^*_p$ and set Hdr $\leftarrow \langle C_1, C_2 \rangle$, where $C_1 \leftarrow g^t$ and

$$C_2 \leftarrow g^{t(\gamma B_i)^{-1}\alpha^{n-|S|}\prod_{j\in S}(\alpha - x_j)},$$

where i is the sender's index i. Let's denote $p(\alpha) = \alpha^{2n-|S|}\prod_{j\in S}(\alpha - x_j)$. It should be clear that $p(\alpha)$ is a polynomial of degree n, and therefore $g^{(\gamma B_i)^{-1}p(\alpha)}$ can be readily calculated from $T_{i,j}$-s and x_j-s. The session key (K) is set as follows:

$$K \leftarrow e(g,g)^{t\alpha^{2n+1}}.$$

Output $\langle \text{Hdr}, K \rangle$.

Decrypt($S, i, D_i, \text{Hdr}, PK$ ***)*** If $i \in S$, find the sender's index (j) and then expand Hdr to $\langle C_1, C_2 \rangle$ and output

$$K \leftarrow e(C_1, g^{p_i(\alpha)} \cdot g^{r_i h_i(\alpha)})e(C_2, d_{i,j}).$$

where $p_i(\alpha) = \alpha^{2n+1} - \frac{\alpha^{n+2}p(\alpha)}{\alpha - x_i}$ and $h_i(\alpha) = \alpha^n \frac{p(\alpha)}{\alpha - x_i}$. Note that for $i \in S$, p_i is a polynomial of degree $2n$ and therefore $g^{p_i(\alpha)}$ can be easily calculated from g^{α^j}-s and x_j-s, when $i \in S$. Similarly, $r_i h_i(\alpha)$ is polynomial of degree $2n-1$ that can be easily calculated from g^{α^j}-s, x_j-s, and r_i.

Correctness: Let's check that decryption recovers the correct value of K. Recall that the secret key of a member is set as $d_{i,j} = g^{\gamma\alpha^n B_j \frac{\alpha^2 - r_i}{\alpha - x_i}}$ for the sender's index j. Then, we

have the following proceedings:

$$\begin{aligned} e(C_1, g^{p_i(\alpha)} \cdot g^{r_i h_i(\alpha)}) e(C_2, d_{i,j}) &= e(g^t, g^{p_i(\alpha)+r_i h_i(\alpha)}) \\ &\times e(g^{t(\gamma B_j)^{-1} p(\alpha)}, g^{\gamma \alpha^n B_j \frac{\alpha^2 - r_i}{\alpha - x_i}}) \\ &= e(g,g)^{t(p_i(\alpha)+r_i h_i(\alpha))} e(g,g)^{t\alpha^n p(\alpha) \frac{\alpha^2 - r_i}{\alpha - x_i}} \\ &= e(g,g)^{t\alpha^n (\alpha^{n+1} - \frac{\alpha^2 p(\alpha)}{\alpha - x_i} + r_i \frac{p(\alpha)}{\alpha - x_i} + p(\alpha) \frac{\alpha^2 - r_i}{\alpha - x_i})} \\ &= e(g,g)^{t\alpha^{2n+1}} \end{aligned}$$

as required.

Authentication: Let SymEnc and SymDec be symmetric encryption and decryption, respectively. Let M be a random verification message to be broadcast to the set S, and let $C_M \xleftarrow{R} \text{SymEnc}(K, M)$ be the randomized encryption of M under the session key K, which is broadcast to the set S. The broadcast to members in S consists of $\langle S, \text{Hdr}, M, C_M \rangle$. The privileged receiver, a member in the set S, can easily verify the sender of the broadcast message as follows:

First, member i (if $i \in S$) retrieves the session key K from Hdr by using the decryption key ($d_{i,j}$) corresponding to the sender (member j) of the message. Then, member i checks if $M = \text{SymDec}(K, C_M)$. If it passes, it verifies the sender, otherwise, it refuses the authentication.

7.3 Security Analysis

In this section, we prove the fully security of the proposed BE_A scheme in the adaptive model.

Theorem 3. *Let $\mathbb{G}_T$ be a bilinear group of prime order p. For any positive integers $n, 2n$ (s.t. $n < 2n < p$) our n-broadcast encryption system is $(\text{negl}(\lambda), 2n)$ adaptively secure assuming the decision $(\text{negl}(\lambda), 2n)$-BDHE-Sum assumption holds in $\mathbb{G}_T$.*

Proof. As usual, we start by the assumption that there is an algorithm $\mathcal{A}$ with advantage $\epsilon = \text{negl}(\lambda)$ in attacking the proposed BE_A scheme. If this is true, we prove that $\mathcal{A}$ can be used to solve the decision n-BDHE-Sum in $\mathbb{G}$. We build a simulation machine $\mathcal{B}$ that receives an instance of the decision n-BDHE-Sum problem. This is comprised of $Z \in \mathbb{G}$ and the set $\{g^{a^{-2n}}, \cdots, g^{a^{2n}}\}$.

No Commit: It has to be emphasized that the adversary's algorithm $\mathcal{A}$ does not commit to a predetermined set of indices S^* to attack, before seeing the public parameters of the scheme. Without loss of generality, we assume $|S^*| = 2$. This implies that the adversary can attack and retrieve the private keys of all members, except two members that will be used in the challenge round. One non-attacked (non-compromised) member is used to generate a broadcast message (Hdr^*) only for the other non-attacked member. Otherwise, it is obvious that the adversary will be able to recover the session key, as it already has the private key of other members.

Setup: $\mathcal{B}$ disguises the parameters of the challenge problem into parameters of the proposed BE_A scheme. $\mathcal{B}$ puts $\alpha = a$ and using the challenge instance, it sets the public parameters as: $g^{\alpha^i} = g^{a^i}$ for $i \in [0, 2n]$. For the public identity's of members (x_i), $\mathcal{B}$ picks

$x_i \overset{R}{\leftarrow} \mathbb{Z}^*{}_p$ and publishes PK as $\mathbb{G}, \mathbb{G}_T, e$, $\{x_1, \cdots, x_n\}$, and $\{g^{\alpha^i}, \forall i \in [0, 2n]\}$. Then, $\mathcal{B}$ picks a random $y_0 \overset{R}{\leftarrow} \mathbb{Z}^*{}_p$ and sets $\gamma = y_0 a^{-2n}$. The session key, as before, is the following $K = e(g,g)^{\alpha^{2n+1}} = e(g,g)^{a^{2n+1}}$. The secret key SK includes the set $\{\alpha = a, \gamma = y_0 a^{-2n}\}$.

Private Keys Query: Algorithm $\mathcal{A}$ queries private keys ($d_{i,j}$) for any arbitrary subset S' of $[1, n]$, where $\max(|S'|) = n - 2$. We make the restriction that $\mathcal{A}$ has to query all private keys at once for members it will attack. Let's denote the set of non-attacked members by S^*. Thus, we have $S^* \cup S' = [1, n]$. We have assumed that $|S^*| = 2$, so the notation $j \oplus 1$ refers to the other index in S^* than j in our notes. Having known the set of attacked members (S') and the set of non-attacked members (S^*), $\mathcal{B}$ picks a random $b_j \overset{R}{\leftarrow} \mathbb{Z}^*{}_p$ and sets $B_j = b_j a^n$ for $j \in S'$, but it sets $B_j = b_j a^{n-1}(a - x_{j\oplus 1})$ and $B_{j\oplus 1} = b_{j\oplus 1} a^{n-1}(a - x_j)$ for $j \in S^*$.

For $i \in S'$, $\mathcal{B}$ responds to the query for member i's private keys as follows: it returns $D_i \leftarrow \langle r_i, d_{i,j}, T_{i,j} \rangle$, recall that in the real protocol, we have to return r_i as a random value, $d_{i,j} = g^{\gamma \alpha^n B_j \frac{\alpha^2 - r_i}{\alpha - x_i}}$ and $T_{i,j} = g^{\frac{\alpha^j}{\gamma B_i}}$. Therefore for all $i \in S'$ and for $j \in S'$, $\mathcal{B}$ sets $r_i = x_i^2$ and returns $d_{i,j} = g^{y_0 b_j (a + x_i)}$, but for $j \in S^*$, it has to return $d_{i,j} = g^{y_0 b_j a^{-1}(a - x_{j\oplus 1})(a + x_i)}$ and $d_{i,j\oplus 1} = g^{y_0 b_j a^{-1}(a - x_j)(a + x_i)}$ to the adversary. Finally, $T_{i,j} = g^{(y_0 b_i)^{-1} a^{n+j}}$ for all $i \in S'$ and for $j \in [0, n]$. Note that all these parameters can be readily calculated from the BDHE-Sum instance and the random parameters of $\mathcal{B}$. It is easy to check that the private keys are matched with the parameters in the real protocol, and they are valid. The set of indices in the non-attacked set S^*, which have not been queried, will be used in the challenge phase.

Challenge: In the challenge phase, $\mathcal{B}$ creates a broadcast encryption message for $i \in S^*$. It sets $S \subset S^*$ and generates $C_1^* = g^t$ and $C_2^* = g^{t(\gamma b_i)^{-1} p(\alpha)}$, where $i \in S^*$. Let's suppose that the broadcast message is generated by member i, where $i \in S^*$ and therefore $S = \{i \oplus 1\}$. The challenge is then calculated as follows: pick a random $t_0 \xleftarrow{R} \mathbb{Z}^*{}_p$, and set $t = t_0 a^{2n}$. Then, calculate the broadcast $\text{Hdr}^* \leftarrow \langle C_1^*, C_2^* \rangle$, where $C_1^* = g^{t_0 a^{2n}}$ and $C_2^* = g^{t_0 a^{2n} (y_0 a^{-2n} B_i)^{-1} a^{n-1}(a - x_{i\oplus 1})}$, which yields $g^{t_0 (y_0 b_i)^{-1}}$ for $B_i = b_i a^{n-1}(a - x_{i\oplus 1})$. Note that both C_1^* and C_2^* can be directly calculated from the BDHE-Sum instance and the random parameters of $\mathcal{B}$. Therefore, $\text{Hdr}^* = \{C_1^*, C_2^*\}$, where $C_1^* = (g^{a^{2n}})^{t_0}$ and $C_2^* = (g)^{t_0 (y_0 b_i)^{-1}}$, is a valid ciphertext for indices in $S \subset S^*$. The corresponding session key would then be $K = e(g,g)^{t_0 a^{2n} a^{2n+1}} = e(g,g)^{t_0 a^{4n+1}}$. $\mathcal{B}$ outputs Hdr^* and $K^* = Z^{t_0}$, where Z is the challenge from the BDHE-Sum instance, as the new challenge to $\mathcal{A}$.

Decryption Query: We further allow $\mathcal{A}$ to use the set of private keys it received to generate a broadcast message for any $i \in [1, n]$ and even for $i \in S^*$. $\mathcal{B}$ is able to derive the private keys $d_{i,j}$ for $i \in S^*$, in the same way as in the **Private Keys Query** phase, except $T_{i,j}$ values for $i \in S^*$. Nevertheless, this does not stop $\mathcal{B}$ from returning correct decryptions, since only r_i and $d_{i,j}$ are need for decrypting and $T_{i,j}$ values are used to create the broadcast encryption. By setting $r_i = x_i^2$ and $d_{i,j} = g^{y_0 b_j (a + x_i)}$ for any $i \in S^*$ and $j \in S'$, $\mathcal{B}$ is able to respond correctly to the decryption queries as in the real application.

Guess: The algorithm $\mathcal{A}$ outputs its guess $b' \in \{0, 1\}$ and wins the game if $b' = b$. $\mathcal{B}$ sends

b' to the challenger in the proposed BE_A scheme to solve the BDHE-Sum instance. From $\mathcal{A}$'s perspective, $\mathcal{B}$'s simulation has almost the same distribution as the adaptive security model defined earlier in Section 7.1. The public and private keys are appropriately distributed, since x_i-s and therefore r_i-s are uniformly random. When $b = 0$ in the adaptive game, $\langle \text{Hdr}^*, K^* \rangle$ is generated according to the same distribution as in the real application with a valid session key $K^* = e(g,g)^{ta^{2n+1}}$, where $t = t_0 a^{2n}$. Thus, the challenge is a valid ciphertext under the randomness of t_0. From $\mathcal{B}$'s simulation, when $b = 0$, we can easily find the solution to BDHE-Sum problem, by outputting $Z = K^{*1/t_0} = e(g,g)^{a^{4n+1}}$.

When $b = 1$ in the adaptive game, $\langle \text{Hdr}^*, K^* \rangle$ is generated with K^* being replaced by a random key. Since $K^* \xleftarrow{R} \mathbb{G}_T$ is a uniformly random element of $\mathbb{G}_T$, this distribution is identical to that of $\mathcal{B}$'s simulation where Hdr^* is a valid ciphertext. From this, we see that $\mathcal{B}$'s advantage in deciding n-BDHE-Sum problem is precisely $\mathcal{A}$'s advantage against the BE_A scheme. □

7.4 Complexity Analysis

In this section, we analyze the overheads of the proposed scheme over previously known schemes. A fully adaptive BE_A with short ciphertexts is achieved. The design is aimed at constructing a fully secure BE scheme with $\mathcal{O}(1)$ communication overheads (Hdr), regardless of the size of the broadcast group. This is achieved without using the Random Oracle Model (ROM) and hash functions. In comparison with Gentry and Waters' BE scheme and its variants [43], the security of the proposed scheme is proved in an attack model that is

stronger than Gentry and Waters' BE scheme. In our security model, we allow the adversary to query the private keys of all members under attack and also to receive decryption of broadcast messages intended for all members. Gentry and Waters' semi-static BE scheme requires $\mathcal{O}(1)$ private keys and $\mathcal{O}(n)$ public keys to be stored by each member in the broadcast group. They extend the security from semi-static to fully adaptive by increasing the size of broadcast message to $\mathcal{O}(\sqrt{n})$ without using random oracles (hash functions). In our scheme, the increase in security has led to an increase in the size of private keys – that is each member in our scheme has to store $\mathcal{O}(n)$ private keys. Next, we show that the basic group operations to add or remove members from the broadcast group can be performed without any overhead in the proposed BE_A scheme.

7.4.1 Group Operations

In the proposed BE_A scheme also, removing from the group membership do not affect existing members. Excluding a member simply means not including the index of excluded member in calculate the ciphertexts (Hdr). If a member is permanently removed from the group, only the identity parameter (x_k) of the excluded member is removed and no further changes to private keys of members are required. Keys of members remain the same as the group membership changes without compromising security of the BE_A protocol. It should be added that member removal is performed at no extra communication or computation cost to group members.

Removal: membership removal is inherent in the BE_A scheme. Excluding a member is

achieved by not including the index of the excluded member in S. This will incur no extra communication or computation overhead on group members to remove a member.

Addition: adding a member is authorized by Admin. If the group's maximum capacity, set by n, is not reached, any new member i' can be added to the group. The Admin simply generates a new set of private keys $\{d_{i',j}, T_{i',j}\}$ for the new member and publishes its identity $(x_{i'})$ to the group. Unlike the semi-static scheme of Boneh et al. [16] that did not require further key update at the existing members, we have to send the new decryption key $d_{i,i'}$ for the existing member i to be able to communicate with new member i'.

As said earlier, the maximum size of the group is limited by the order of bilinear underlying group. Moreover, adding new members requires the Admin to broadcast new decryption keys to all current members of the group, where as removing any member does not require any change to keys of existing members. Our proposed scheme is the only scheme that provides authentication of the sender without any increase in the size of the broadcast encryption message.

A broadcast encryption scheme based on cryptographic pairings is proposed in this chapter. The scheme is the first adaptively secure broadcast encryption with short ciphertexts that does not use the random oracle model. The security model of the proposed broadcast encryption scheme is a strong model simulating the adversary in the real world as closely as possible. In our model, the adversary can receive the private keys of any subset of members in the broadcast group as well as decryption of previous broadcast messages. It has also been shown that the communication and computation overheads needed for the protocol to

actively exclude or include memberships are very minimal, i.e. with $\mathcal{O}(1)$ communication and $\mathcal{O}(n)$ computations, where n is the size of the broadcast group. The amount of storage required for each member is trivial when compared to other protocols. Members can join or leave the group, while the security keys of other members will not be affected by the changes in the group. The maximum number of members that can join the group is limited by the underlying algebraic group structure. The maximum size of the broadcast group is bounded by the size of the underlying bilinear group. This implies that the size of the underlying pairing group increases linearly with the maximum size of the broadcast group.

Chapter 8

Conclusions and Future Work

Providing a fully secure authentication scheme for RFID systems is only possible if the authentication process is protected from known existing attacks. One of the main tasks in RFID systems is secure identification over a wireless channel. In this dissertation, we have addressed various challenges in designing light-weight schemes to authenticate RFID tags and readers. We have given the specific requirements and challenges in RFID systems and proposed a solution for every problem that arise in practice, including light-weight authentication, privacy protection, security against relay attacks and key management.

8.1 Light-weight, Asymmetric Authentication

The solutions provided to each problem directly affects the complexity of the RFID tag's circuit, transmission rate, management of the keys and ultimately the price of the system. A review of the state-of-the-art solutions for every problem in RFID systems has been pro-

vided in Chapter 2. We have shown the shortcomings of existing solutions and provided the requirements that should be satisfied. In Chapters 3 and 4, we have designed light-weight, asymmetric authentication schemes based on error-correcting codes. Asymmetric authentication algorithms better handle key management, and code-based systems are efficient in hardware implementation, suitable for RFID systems. In Chapter 3, we have shown that the given modification on McEliece cryptosystem will enable us to authenticate the readers. We have also shown that the same system can be reused to add confidentiality via encryption to the tag's content and protect it from tampering attacks.

The proposed BLA scheme is a randomized protocol, such that it can protect the identity of RFID tags from eavesdroppers and replay-attackers. Tampering attacks are also prevented by authenticating the readers. This will the tags to grant access only to authorized readers to change the content of the tag and access its sensitive data.

In the proposed BLA scheme, the tag is able to authenticate one pre-authorized reader if the public-key of new readers are not stored on the tag's memory in advance. To authenticate new readers, the RFID tag can download the public-key of a new reader and verify its validity. Thereafter, the tag can easily authenticate the new reader using the same BLA scheme.

In Chapter 4, we have presented a zero-knowledge authentication scheme that allows the readers to authenticate RFID tags in an FLA protocol without being able to learn anything about the tag's secret key. The security of the proposed scheme is rooted in a well-known complexity assumption in the coding theory, that is SD problem. We provided a formally proof of security based on the difficulty of solving the SD problem. The authentication

scheme is an asymmetric protocol that can be readily scaled to a system of many tags. The complexity of operations required to perform the authentication is very minimal and suitable for EPC Gen 2 tags.

8.2 Light-weight Privacy Protection

Tag's authentication often verifies the unique identity of an RFID tag. Since passive (low-cost) RFID tags are wirelessly activated and immediately respond to authentication queries, there exist privacy threats in RFID systems. In Chapter 5, we have provided a light-weight PPI scheme, which is suitable for very low-cost RFID tags, such as EPC Gen 2 rags[46]. A light-weight identification mechanism is proposed for an RFID systems in which the privacy of tags is protected against unknown readers while the tags can be correctly identified to trusted readers. The proposed PPI scheme is formally proven to be secure based on the difficulty of LPN problem. In the proposed PPI scheme, we have moved the complex computational operations from the RFID tags to readers, enabling them to quickly authenticate (trusted) readers. Authentic readers can identify an RFID tag, but unknown readers will not be able to trace the tag across many domains of use.

Using an asymmetric cryptosystem in our construction has enabled the proposed PPI scheme to expand to any number of tags and readers. In particular, the size of parameters remain constant for the readers regardless of the number of RFID tags that are usually very large (around thousands) in practice. In the proposed PPI scheme, the identity of the tags are revealed only to the authorized readers. If authorized readers maliciously collude together,

they can track the RFID tags in the system. By randomizing the identity of the RFID tags for each authorized reader, we could further protect the privacy of RFID tags even against colluding authorized readers.

8.3 Protection against Relay Attacks

We have shown in Chapter 6 that most existing identification/authentication schemes are prone to relay attacks. The adversary can simply impersonates an authentic tag by relaying the challenge-response messages between the tag and the reader, gaining access to restricted resources. Existing countermeasures for the relay attacks are mostly difficult to implement in practice. They mostly require access to an UWB channel or relay on precise timing of exchanged signals. Both of these approaches create great restrictions in practice. Therefore, we have proposed a new approach that does not have these restrictions.

In Chapter 6, we have proposed a masking scheme based on chaos suppression theory that hides the analog signals of an RFID tag from unauthorized readers. The proposed approach is the first solution that is based on chaos suppression. Unlike distance bounding solutions, our approach does not require access to a fast channel or precise timing mechanism. In the proposed chaos-masking scheme, the chaos suppression signal is sent as a secret key to remove the mask (scrambling signals) from the RFID tag's data signals. If a correct suppression signal is not presented, the tag continues to scramble its data signals before they are output.

The proposed masking scheme is based on chaotic suppression rather than chaotic syn-

chronization. Therefore, it has a moderate sensitivity to noise in the chaotic systems and parameters mismatch. In addition to practicality, this will also improve the overall security of the chaotic system. In other chaotic communication systems based on chaos-synchronization, synchronization signals must be easily observable in the system and to the receiver. This also leads to the breach of security, as the adversary can easily eavesdrop on the communication channel and receive the synchronization signals.

The proposed scheme have overcome these problems by using a bandlimited suppression signal without having to modify the underlying communication standards. Our work is based on Lorenz chaotic system, which can be easily implemented in EPC Gen 2 tags with three mixers and three integrators. It is possible to reduce the number of analog mixers and integrators by using other chaotic systems that have a lower complexity. It is an interesting area for further research to compare the security and complexity of the proposed chaos-masking system under various chaotic systems.

8.4 Efficient Key Management

Sharing secret keys and securely communicating them in the RFID system is the first requirement to complete any of the proposed in authentication schemes, and any authentication or encryption algorithm needs a cryptographic key. Managing and distributing the keys in the RFID system is a very challenging, administrative problem in reality. The cryptographic keys should be updated frequently to guarantee a high level of security, as new entities (readers, tags or servers) constantly join or leave the system. Therefore, there has to be a

mechanism to distribute the cryptographic keys in the system efficiently and securely. The scheme should not overload the system by updates and should safely deliver the keys to their intended recipients.

We have utilized the broadcast encryption scheme for remote distribution of cryptographic keys or any other secret data to the readers. Broadcast encryption allows us to securely communicate cryptographic keys concurrently to many readers over a public, insecure channel. Designing a secure and efficient broadcast encryption scheme has always been a great challenge. The readers in a broadcast group should not be able to collude together to gain unrestricted access. Moreover, we should keep the communication overheads minimal in the system to avoid network overflow or delay that could lead to security problems.

In Chapter 7, we have proposed the first broadcast encryption scheme that is provably secure in the most realistic attack model, adaptive adversary security model. The proposed scheme has short ciphertexts, and it is suitable for large networks with many readers. The size of the broadcast messages is fixed, i.e. $\mathcal{O}(1)$, and does not increase with the number of recipients in the broadcast group. The computation and storage requirements for the distributor (usually the administrator) increase linearly with the number of RFID readers in the broadcast group. The RFID system evolves over time and grows in size. Nevertheless, RFID readers and servers do not have the resource limitation of RFID tags and can easily accommodate the increase.

Bibliography

[1] Mihály Bárász, Balázs Boros, Péter Ligeti, Krisztina Lója, and Dániel Nagy. *Breaking LMAP*. In *Conference on RFID Security (RFIDSec'07)*, 2007.

[2] Mihály Bárász, Balázs Boros, Péter Ligeti, Krisztina Lója, and Dániel A. Nagy. *Passive Attack Against the M2AP Mutual Authentication Protocol for RFID Tags*. In *EURASIP International Workshop on RFID Technology*, 2007.

[3] Roberto Barrio and Sergio Serrano. *A three-parametric study of the Lorenz model*. *Physica D*, 229:43–51, 2007.

[4] Roberto Barrio and Sergio Serrano. *Bounds for the chaotic region in the Lorenz model*. *Physica D: Nonlinear Phenomena*, 238(16):1615–1624, 2009.

[5] Rana Barua, Ratna Dutta, and Palash Sarkar. *Extending Jouxs protocol to multi party key agreement*. In *Advances in Cryptology: INDOCRYPT'03*, Lecture Notes in Computer Science, vol. 2904, pages 205–217. Springer-Verlag, 2003.

[6] Lejla Batina, Nele Mentens, Kazuo Sakiyama, Bart Preneel, and Ingrid Verbauwhede. *Low-Cost Elliptic Curve Cryptography for Wireless Sensor Networks*. In *Workshop*

on Security and Privacy in Ad-Hoc and Sensor Networks: ESAS'06, Lecture Notes in Computer Science, vol. 4357, pages 6–17. Springer-Verlag, 2006.

[7] Thierry P. Berger, Pierre-Louis Cayrel, Philippe Gaborit, and Ayoub Otmani. *Reducing Key Length of the McEliece Cryptosystem.* In *Advances in Cryptology (AFRICACRYPT'09)*, Lecture Notes in Computer Science, vol. 5580, pages 77–97. Springer-Verlag, 2009.

[8] Elwyn R. Berlekamp, Robert McEliece, and Henk C. A. Van TilBorg. *On the inherent intractability of certain coding problems. IEEE Transactions on Information Theory*, 24(3):384–386, 1978.

[9] Daniel J. Bernstein, Tanja Lange, and Christiane Peters. *Attacking and Defending the McEliece Cryptosystem.* In *2nd International Workshop on Post-Quantum Cryptography (PQCRYPTO'08)*, volume 5299, pages 31–46. Springer-Verlag, 2008.

[10] Bhaskar Biswas and Nicolas Sendrier. *McEliece cryptosystem implementation: theory and practice.* In *Second International Workshop in Post-Quantum Cryptography: PQCrypto'08*, Lecture Notes in Computer Science, vol. 5299, pages 47–62, 2008.

[11] Ian F. Blake, Gadiel Seroussi, and Nigel P. Smart, editors. *Advances in Elliptic Curve Cryptography.* Cambridge University Press, 2005.

[12] Leonid Bolotnyy and Gabriel Robins. *Physically Unclonable Function-Based Security and Privacy in RFID Systems.* In *Fifth IEEE International Conference on Pervasive*

Computing and Communications (PERCOMP'07), pages 211–220. IEEE Computer Society, 2007.

[13] Dan Boneh and Xavier Boyen. *Efficient selective-ID identity based encryption without random oracles*. In Christian Cachin and Jan Camenisch, editors, *Advances in Cryptology: EUROCRYPT'04*, Lecture Notes in Computer Science, vol. 3027, pages 223–238. Springer, 2004.

[14] Dan Boneh, Xavier Boyen, and Eu-Jin Goh. *Hierarchical Identity Based Encryption with Constant Size Ciphertext*. In *Advances in Cryptology:EUROCRYPT'05*, Lecture Notes in Computer Science, vol. 3494, pages 440–456. Springer-Verlag, 2005. Available at `http://www.cs.stanford.edu/~xb/eurocrypt05a/`.

[15] Dan Boneh and Matthew Franklin. *Identity-Based Encryption from the Weil Pairing*. *SIAM Journal of Computing*, 32(3):586–615, 2003.

[16] Dan Boneh, Craig Gentry, and Brent Waters. *Collusion Resistant Broadast Encryption with short ciphrertexts and private keys*. In *Advances in Cryptology (CRYPTO'05)*, Lecture Notes in Computer Science, vol. 3621, pages 258–275. Springer-Verlag, 2005.

[17] Samuel Bowong, F.M. Moukam Kakmeni, and M. Siewe Siewe. *Secure communication via parameter modulation in a class of chaotic systems. Communications in Nonlinear Science and Numerical Simulation*, 12(3):397–410, 2007.

[18] Ran Canetti and Benny Pinkas. *A taxonomy of multicast security issues.* Available Online, November 1998. IBM Research, the Weizmann Institute.

[19] A. Canteaut and F. Chabaud. *A new algorithm for finding minimum-weight words in a linear code: Application to McEliece's cryptosystem and to narrow-sense BCH codes of length 511. IEEE Transactions on Information Theory*, 44(1):367–378, 1998.

[20] A. Canteaut and N. Sendrier. *Cryptanalysis of the original McEliece cryptosystem.* In *Advances in Cryptology: ASIACRYPT'98*, Lecture Notes in Computer Science, vol. 1514, pages 187–199. Springer-Verlag, 1998.

[21] Julio Cesar Hernandez Castro, Juan M. Estevez-Tapiador, Pedro Peris-Lopez, and Jean-Jacques Quisquater. *Cryptanalysis of the SASI Ultralightweight RFID Authentication Protocol with Modular Rotations.* Technical report, CoRR Labs, 2008.

[22] Pierre-Louis Cayrel, Philippe Gaborit, and Emmanuel Prouff. *Secure implementation of the Stern authentication and signature schemes for low-resource devices.* In *Eighth Smart Card Research and Advanced Application Conference: CARDIS'08*, Lecture Notes in Computer Science, vol. 5189, pages 91–205. Springer-Verlag, 2008.

[23] Florent Chabaud and Jacques Stern. *The cryptographic security of the syndrome decoding problem for rank distance codes.* In *Advances in Cryptology: ASIACRYPT'96*, Lecture Notes in Computer Science, vol. 1163, pages 368–381, 1996.

[24] Yacine Challal and Hamida Seba. *Group Key Management Protocols: A Novel Taxonomy. International Journal of Information Theory*, 2(1):105–118, 2005.

[25] Kefei Chen. *A new identification algorithm.* In *Cryptography Policy and Algorithms*, Lecture Notes in Computer Science, vol. 1029, pages 244–249. Springer-Verlag, 1996.

[26] Liqun Chen, Michael Cheng, and Nigel P. Smart. *Identity-based Key Agreement Protocols From Pairings. International Journal of Information Security*, 6(4):213–241, 2007.

[27] Shao cheng Qu, Mei jing Gong, and Xiao yan Wang. Parameter modulation secure communication based on active control. In *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pages 167–170, 2008.

[28] Hung-Yu Chien. *SASI: A New Ultralightweight RFID Authentication Protocol Providing Strong Authentication and Strong Integrity. IEEE Transactions on Dependable and Secure Computing*, 4(4):337–340, 2007.

[29] Chol-Ung Choe, Klaus Höhne, Hartmut Benner, and Yuri S. Kivshar. *Chaos suppression in the parametrically driven Lorenz system. The American Physical Society, Physical Review*, 72(3):no. 036206, 2005.

[30] Kyu Young Choi, Jung Yeon Hwang, and Dong Hoon Lee. *Efficient ID-based Group Key Agreement with Bilinear Maps.* In *Public Key Cryptography*, Lecture Notes in Computer Science, vol. 2947, pages 130–144. Springer-Verlag, 2004.

[31] Nicolas T. Courtois. *Efficient Zero-Knowledge Authentication Based on a Linear Algebra Problem MinRank*. In *Advances in Cryptology (ASIACRYPT'01)*, Lecture Notes in Computer Science, vol. 2248, pages 402–421, 2001.

[32] Kevin M. Cuomo, Alan V. Oppenheim, and Steven H. Strogatz. *Synchronization of Lorenz-B ased Chaotic Circuits with Applications to Communications. IEEE Transcation on Circuits and System-11: Analog and Digital Signal Processing*, 40(10):624–633, 1993.

[33] A. d'Anjou, C. Sarasola, F.J. Torrealdea, R. Orduna, and M. Grana. *Parameter-adaptive identical synchronization disclosing Lorenz chaotic masking. Physics Review E*, 63:no. 046213–1–5, 2001.

[34] Yevgeniy Dodis and Nelly Fazio. *Public Key Broadcast Encryption for Stateless Receivers*. In *Digital Rights Management: DRM'02*, Lecture Notes in Computer Science, vol. 2696, pages 61–80. Springer-Verlag, 2002.

[35] Martin Feldhofer, Sandra Dominikus, and Johannes Wolkerstorfer. *Strong authentication for RFID systems using the AES algorithm*. In *Workshop on Cryptographic Hardware and Embedded Systems (CHES'04)*, Lecture Notes in Computer Science, vol. 3156, pages 357–370. Springer-Verlag, 2004.

[36] Martin Feldhofer and Christian Rechberger. *A Case Against Currently Used Hash Functions in RFID Protocols*. In *OTM 2006 Workshops*, Lecture Notes in Computer Science, vol. 4277, pages 372–381. Springer-Verlag, 2006.

[37] Martin Feldhofer and Johannes Wolkerstorfer. *Strong Crypto for RFID Tags – A Comparison of Low-Power Hardware Implementations.* In *IEEE International Symposium on Circuits and Systems (ISCAS'07)*, pages 1839–1842. IEEE, May 2007.

[38] Klaus Finkenzeller. *RFID Handbook: Fundamentals and Applications in Contactless Smart Cards and Identification.* Wiley, 2003.

[39] International Organization for Standardization. *ISO/IEC 15693: Identification cards – Contactless integrated circuit cards – Vicinity cards*, 2006.

[40] International Organization for Standardization. *ISO/IEC 14443: Identification cards – Contactless integrated circuit cards – Proximity cards*, 2008.

[41] International Organization for Standardization. ISO/IEC 9798-2: Information Technology Security techniques Entity Authentication Mechanisms. *Part 2: Entity authentication using symmetric techniques.* ISO/IEC, 1993.

[42] Philippe Gaborit and Marc Girault. *Lightweight code-based authentication and signature.* In *IEEE International Symposium on Information Theory (ISIT'07)*, pages 191–195, 2007.

[43] Craig Gentry and Brent Waters. *Adaptive Security in Broadcast Encryption Systems (with Short Ciphertexts).* In *Advances in Cryptology: EUROCRYPT'09*, Lecture Notes in Computer Science, vol. 5479, pages 171–188. Springer-Verlag, 2009.

[44] Henri Gilbert, Matthew J.B. Robshaw, and Yannick Seurin. *HB#: increasing the security and efficiency of HB+.* In *Advances in Cryptology (Eurocrypt'08)*, Lecture Notes in Computer Science, vol. 4965, pages 361–378. Springer, 2008.

[45] Henri Gilbert, Matthew J.B. Robshaw, and Herv Sibert. *An Active Attack Against HB+: A Provably Secure Lightweight Authentication Protocol. IEE Electronics Letters*, 41(21):1169–1170, 2005.

[46] EPC Global. *Class-1 Generation-2 UHF air interface protocol standard.* Accessd Online, July 2010. `http://www.epcglobalinc.org/standards/`.

[47] Nils Gura, Arun Patel, Arvinderpal Wander, Hans Eberle, and Sheueling Chang Shantz. *Comparing Elliptic Curve Cryptography and RSA on 8-bit CPUs.* In *Workshop on Cryptographic Hardware and Embedded Systems (CHES'04)*, Lecture Notes in Computer Science, vol. 3156, pages 119–132. Springer-Verlag, August 2004.

[48] Dani Halevy and Adi Shamir. *The LSD Broadcast Encryption Scheme.* In *Advances in Cryptology (CRYPTO'02)*, Lecture Notes in Computer Science, vol. 2442, pages 145–161. Springer-Verlag, 2002.

[49] Gerhard P. Hancke. *Practical Attacks on Proximity Identification Systems.* In *IEEE Symposium on Security and Privacy*, volume 328-333, 2006.

[50] Gerhard P. Hancke and Saar Drimer. *Security in RFID and Sensor Networks*, chapter 8: Secure Proximity Identification for RFID, pages 170–194. CRC Press, 2008.

[51] Gerhard P. Hancke and Markus G. Kuhn. *An RFID distance bounding protocol*. In *Security and Privacy for Emerging Areas in Communications Networks (SecureComm'05)*, pages 67–73. IEEE, 2005.

[52] Gerhard P. Hancke and Markus G. Kuhn. *Attacks on time-of-flight distance bounding channels*. In *First ACM conference on Wireless network security: WiSec'08*, pages 194–202. ACM, 2008.

[53] Darrel Hankerson, Alfred J. Menezes, and Scott Vanstone. *Guide to Elliptic Curve Cryptography*. Series: Springer Professional Computing. Springer-Verlag, 2004.

[54] Sami Harari. *A new authentication algorithm*. In *3rd International Colloquium on Coding Theory and Applications*, Lecture Notes in Computer Science, vol. 388, pages 91–105. Springer-Verlag, 1989.

[55] Jeffrey Hoffstein, Nick Howgrave-graham, Jill Pipher, Joseph H. Silverman, and William Whyte. *NTRUSign: Digital signatures using the NTRU lattice*. In *Topics in Cryptology – The Cryptographers' Track at the RSA Conferenc*, Lecture Notes in Computer Science, vol. 2612, pages 122–140. Springer-Verlag, 2003.

[56] Jeffrey Hoffstein, Jill Pipher, and Joseph H. Silverman. *NTRU: A Ring-Based Public Key Cryptosystem*. In *Algorithmic Number Theory: ANTS III*, Lecture Notes in Computer Science, vol. 1423, pages 267–288. Springer-Verlag, 1998.

[57] Jeffrey Hoffstein, Joseph H. Silverman, and William Whyte. NTRU report 012. *Estimated breaking times for NTRU lattices*. Technical Report 12, NTRU Cryptosystems, Inc., June 2003.

[58] RFID Journal. *Attacks on a Cryptographic RFID Device*. Accessed Online, Feburary 2005. `http://www.rfidjournal.com/article/view/1415/1/82`.

[59] Antoine Joux. *A One Round Protocol for Tripartite Diffie-Hellman*. In *the 4th International Symposium on Algorithmic Number Theory*, Lecture Notes in Computer Science, vol. 1838, pages 385–394. Springer-Verlag, 2000.

[60] Ari Juels and Stephen A. Weis. *Authenticating pervasive devices with human protocols*. In *Advances in Cryptology (Crypto'05)*, Lecture Notes in Computer Science, vol. 3126, pages 198–293. Springer, 2005.

[61] G. Kabatiansky, E. Krouk, and S. Semenov. *Error Correcting Coding and Security for Data Networks*. John Wiley and Sons, 2005.

[62] Jens-Peter Kaps. *Cryptography for Ultra-Low Power Devices*. PhD thesis, ECE Department of Worcester Polytehcnic Institute, 2006.

[63] Maurice Keller and William Marnane. *Low Power Elliptic Curve Cryptography*. In Nadine Azémard and Lars J. Svensson, editors, *Integrated Circuit and System Design. 17th International Workshop on Power and Timing Modeling, Optimization and Simu-*

lation (PATMOS 2007), Lecture Notes in Computer Science, vol. 4644, pages 310–319, Gothenburg, Sweden, September 2007. Springer-Verlag.

[64] Chong Hee Kim and Gildas Avoine. *RFID Distance Bounding Protocol with Mixed Challenges to Prevent Relay Attacks.* In *International Conference on Cryptology and Network Security: CANS'09*, Lecture Notes in Computer Science, pages 119–133, Berlin, Heidelberg, 2009. Springer-Verlag.

[65] SungJin Kim, YoungSoo Kim, and SeokCheon Park. *RFID Security Protocol by Lightweight ECC Algorithm.* In *Sixth International Conference on Advanced Language Processing and Web Information Technology (ALPIT'07)*, pages 323–328. IEEE Computer Society, 2007.

[66] RSA Laboratories. *Technical Characteristics of RFID.* Online, August 2010. `http://www.rsa.com/rsalabs/node.asp?id=2121`.

[67] Min Lei, Guang Meng, and Zhengjin Feng. *Security analysis of chaotic communication systems based on VolterraWienerKorenberg model. Chaos, Solitons & Fractals*, 28(1):264–270, 2006.

[68] Arjen K. Lenstra and Eric R. Verheul. *Selecting cryptographic key sizes. Journal of Cryptology: The Journal of the International Association for Cryptologic Research*, 14(4):255–293, 2001.

[69] Tieyan Li and Robert Deng. *Vulnerability analysis of EMAP - an efficient RFID mutual authentication protocol*. In *International Conference on Availability, Reliability and Security*, pages 238–245, 2007.

[70] Tieyan Li and Guilin Wang. *Security analysis of two ultra-lightweight RFID authentication protocols*. In *IFIP International Information Security Conference (IFIP-SEC'07)*, pages 14–16. Springer, 2007.

[71] Chuandong Lia, Xiaofeng Liao, and Kwok wo Wong. *Chaotic lag synchronization of coupled time-delayed systems and its applications in secure communication. Physica D: Nonlinear Phenomena*, 194:187–202, 2004.

[72] D. Lim, J. Lee, B. Gassend, G. Suh, M. Dijk, and S. Devadas. *Extracting secret keys from integrated circuits. IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 13(10):1200–1205, 2005.

[73] Ying Liu, Yu Mao, Wallace K. S. Tang, and Ljupco Kocarev. *Cryptanalysis of chaotic communication schemes by dynamical minimization algorithm. International Journal of Bifurcation and Chaos*, 19(7):2429–2437, 2009.

[74] E.N. Lorenz. *Deterministic nonperiodic flow. Atmospheric Science*, 20:130–141, 1963.

[75] Mehrdad Majzoobi, Farinaz Koushanfar, and Miodrag Potkonjak. *Techniques for design and implementation of secure reconfigurable PUFs. ACM Transactions on Reconfigurable Technology and Systems (TRETS)*, 5(2):1–33, 2009.

[76] Mark Manulis. *Security-Focused Survey on Group Key Exchange Protocols*. Technical Report November, HGI Network and Data Security Group, 2006. Available at `http://www.manulis.eu/papers/TR0603-GKEPS.pdf`.

[77] A. Maybhate and R. E. Amritkar. *Use of synchronization and adaptive control in parameter estimation from a time series. Physics Review E*, 59:284–293, 1999.

[78] Robert J. McEliece. *A public-key cryptosystem based on algebraic coding theory: DNS Progress Report*. Technical report, Jet Propulsion Laboratory, 1978.

[79] Alfred J. Menezes, Paul C. van Oorschot, and Scott A. Vanstone. *Handbook of Applied Cryptography*. CRC Press, 1996.

[80] David Molnar and David Wagner. *Privacy and security in library RFID: Issues, practices, and architectures*. In *Conference on Computer and Communications Security (CCS'04)*, pages 210–219. ACM, ACM Press, 2004.

[81] Jorge Munilla and Alberto Peinado. *Distance bounding protocols for RFID enhanced by using void-challenges and analysis in noisy channels. Wireless Communication and Mobile Computing*, 8(9):1227–1232, November 2008.

[82] Jorge Munilla and Alberto Peinado. *Enhanced Low-cost RFID Protocol to Detect Relay Attacks*. In *Wireless Communications and Mobile Computing*, volume 10, pages 361–371, 2010.

[83] Dalit Naor, Moni Naor, and Jeffrey B. Lotspiech. *Revocation and Tracing Schemes for Stateless Receivers*. In *Advances in Cryptology (CRYPTO'01)*, Lecture Notes in Computer Science, vol. 2139, pages 41–62. Springer-Verlag, 2001.

[84] Harald Niederreiter. *Knapsack-type cryptosystems and algebraic coding theory. Problems in Control and Information Theory*, 15:159–166, 1986.

[85] Ryo Nojima, Hideki Imai, Kazukuni Kobara, and Kirill Morozov. textitSemantic security for the McEliece cryptosystem without random oracles. *Designs, Codes and Cryptography*, 49(1):289–305, 2008.

[86] P. Peris-Lopez, J. C. Hernandez-Castro, J. M. Estevez-Tapiador, and A. Ribagorda. *M2AP: A minimalist mutual-authentication protocol for low-cost RFID tags*. In *UIC'06*, Lecture Notes in Computer Science, vol. 4159, pages 912–923. Springer-Verlag, 2006.

[87] Pedro Peris-Lopez, Julio César Hernández Castro, Juan M. Estévez-Tapiador, and Arturo Ribagorda. *EMAP: An efficient mutual authentication protocol for low-cost RFID tags*. In *OTM Workshops On the Move to Meaningful Internet Systems*, Lecture Notes in Computer Science, vol. 4277, pages 352–361. Springer-Verlag, 2006.

[88] Pedro Peris-Lopez, Julio César Hernández Castro, Juan M. Estévez-Tapiador, and Arturo Ribagorda. *LMAP: A real lightweight mutual authentication protocol for low-cost RFID tags*. In *Workshop on RFID and Lightweight Cryptography*, 2006.

[89] Pedro Peris-Lopez, Julio César Hernández Castro, Juan M. Estévez-Tapiador, and Arturo Ribagorda. *Advances in Ultralightweight Cryptography for Low-Cost RFID Tags: Gossamer Protocol.* In *International Workshop Information Security Applications: WISA'08*, Lecture Notes in Computer Science, vol. 5379, pages 56–68. Springer, 2009.

[90] Norbert Pramstaller, Stefan Mangard, Sandra Dominikus, and Johannes Wolkerstorfer. *Efficient AES Implementations on ASICs and FPGAs.* In *Advanced Encryption Standard (AES'05)*, Lecture Notes in Computer Science, vol. 3373, pages 98–112. Springer, 2005.

[91] Jason Reid, Juan M. Gonzalez Nieto, Tee Tang, and Bouchra Senadji. *Detecting relay attacks with timing-based protocols.* In *2nd ACM Symposium on Information, Computer and Communications Security: ASIACCS'07*, 2007.

[92] H. Sakaguchi. *Parameter evaluation from time sequences using chaos synchronization. Physics Review E*, 65:no. 027201–1–4, 2002.

[93] Dave Singelée and Bart Preneel. *Distance Bounding in Noisy Environments.* In *European Workshop on Security and Privacy in Ad-Hoc and Sensor Networks: ESAS'07*, Lecture Notes in Computer Science, vol. 4572, pages 101–115. Springer-Verlag, 2007.

[94] Peter Stavroulakis, editor. *Chaos applications in telecommunications.* CRC Press, 2006.

[95] Jacques Stern. *A new identification scheme based on syndrome decoding.* In *Advances in Cryptology (CRYPTO'93)*, Lecture Notes in Computer Science, vol. 773, pages 13–21. Springer-Verlag, 1993.

[96] Fang Tang. *An adaptive synchronization strategy based on active control for demodulating message hidden in chaotic signals. Chaos, Solitons & Fractals*, 37(4):1090–1096, 2008.

[97] Fang Tang and Ling Wang. *An adaptive active control for the modified Chua's circuit. Physics Letter A*, 346(5):342–346, 2005.

[98] Pascal Véron. *Cryptanalysis of Harari's Identification Scheme.* In *Proceedings of the 5th IMA Conference on Cryptography and Coding*, Lecture Notes in Computer Science, vol. 1025, pages 264–269. Springer-Verlag, 1995.

[99] T. Yang, L. B. Yang, and C. M. Yang. *Break chaotic switching using generalized synchronization: examples. IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 45(10):1062–1067, 1998.

[100] T. Yang, L.B. Yang, and C.M. Yang. *Breaking chaotic secure communication using a spectrogram. Physics Letter A*, 247:105–111, 1998.

[101] Tao Yang. *Recovery of digital signal from chaotic switching. International Journal of Circuit Theory and Applications*, 23(6):611–615, 1995.

[102] Tao Yang. *A Survey of Chaotic Secure Communication Systems. International Journal Computantional Cognition*, 2:81–130, 2004.

Appendix A

Introduction to Chaos Theory

Chaos is a very universal and robust phenomenon in many nonlinear systems. Chaos is formally defined in Chapter 1, Page 2 [94] as *"an aperiodic long-term behavior in a deterministic system that exhibits sensitive dependence to initial conditions"*. In other words, the chaotic trajectories that start in a point infinitely close to another chaotic trajectory will diverge exponentially from it. Since infinite precision in measuring initial conditions is impossible, chaotic systems are unpredictable and aperiodic in nature. Three main components of chaos are described as follows:

1. *Aperiodic long-term* behavior means that the system's trajectory in phase space does not settle down to any fixed points (steady state), periodic orbits, or quasi-periodic solutions as time tends to infinity. This part of the definition differentiates aperiodicity due to chaotic dynamics from the transient aperiodicity of, for example, a periodically oscillating system that has been momentarily perturbed.

2. *Deterministic* systems can have no stochastic (meaning probabilistic) parameters. It is a common misconception that chaotic systems are noisy systems driven by random processes. The irregular behavior of chaotic systems arises from intrinsic nonlinearities rather than noise.

3. *Sensitive dependence on initial conditions* requires that trajectories originating from very nearly identical initial conditions will diverge exponentially quickly. The meaning of this will be made clear in the following discussion.

Among many chaotic systems, we use Lorenz system as a chaotic system. Lorenz system is a classical low-dimension chaotic system [74], which has been studied extensively in the literature and its chaotic regions are very well known. Lorenz system is defined in Eq. (A.1) by its state space variables x, y and z.

$$\begin{aligned} \dot{x} &= \sigma(y - x) \\ \dot{y} &= rx - y - xz \\ \dot{z} &= xy - bz \end{aligned} \tag{A.1}$$

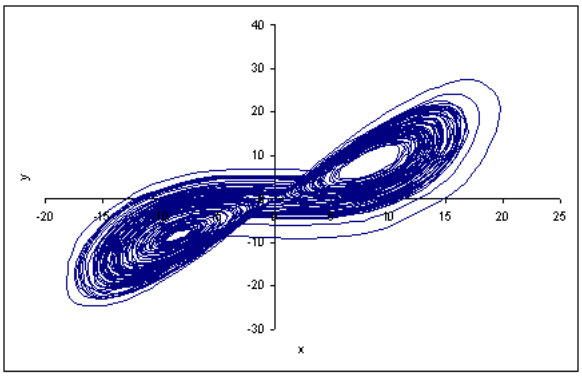
where the dots stand for the derivatives in time of the state space variable. Note that σ, r, and b are system parameters in Lorenz system. For the following values, i.e. $\sigma = 10, b = 8/3$, and $r > 24.74$, it is known that no stable fixed points exist in the system and chaotic dynamics occur [74]. A limited range of $24.74 < r < 214$ is a chaotic region and is referred to as Region 1. It is shown that the chaotic regions in Lorenz system depend on three system parameters, but they are bounded on the (σ, b) plane for a fixed r [4]. If we fix $\sigma = 10$ and $b = 8/3$, a large

chaotic region appears up to $24.74 < r < 146$, then a regular region up to $146 < r < 166$, then another chaotic region up to $166 < r < 214$ and then a regular region. Region 2 is for $1 < r \leq 24.74$, where two fixed points at $(x, y, z) = (\pm\sqrt{b(r-1)}, \pm\sqrt{b(r-1)}, r-1)$ appear. Finally, Region 3 is for $r < 1$, where there exists only one stable fixed point at the origin $(x, y, z) = (0, 0, 0)$.

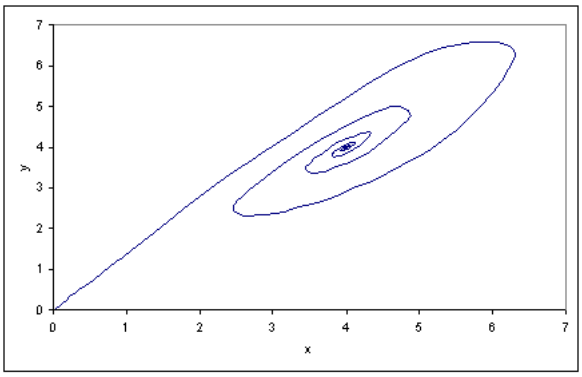
In Figure A.1, we have shown the trajectories of Lorenz system in the x-y plane for different regions, all initialized with $(x_0, y_0, z_0) = (0.002, 0.001, 0.001)$. As it can be seen from Figure A.1a, x and y values never repeat themselves, and they show new details in every orbit in the chaotic region. Figure A.1b shows that x and y values converge and finally oscillate around two fixed points at $(x_1, y_1, z_1) = (4, 4, 6)$ and $(x_2, y_2, z_2) = (-4, -4, 6)$. Finally, Region 3 is shown in Figure A.1c, where x and y values converge to zero.

Dependence on initial conditions can be clearly seen in Figure A.2, where the behavior of Lorenz system initialized with two very close initial conditions is depicted. Lorenz system defined in Eq. (A.1) is initialized once with $(x_0, y_0, z_0) = (0.0020, 0.001, 0.001)$ and again with $(x_0, y_0, z_0) = (0.0021, 0.001, 0.001)$. The x state of Lorenz system is depicted in Figure A.2a and in Figure A.2b respectively based on the initial conditions. It can be easily seen that as time evolves the outputs of the same system quickly diverges if initialized on a slightly different initial condition.

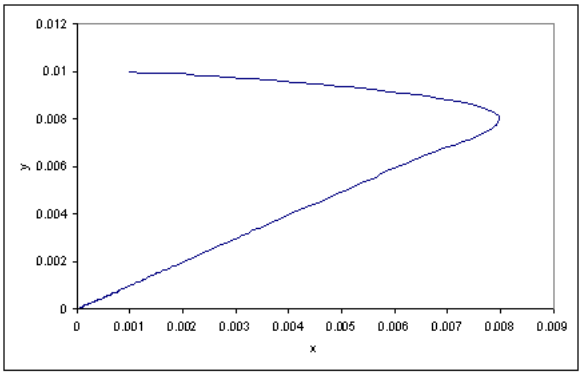
Chaotic effects described here are very desirable in security systems. They can be used to mask an analog signal and make it indistinguishable from a noise signal. Although attractive, there are other factors that we need consider to apply chaotic systems in RFIDs. A set of



(a) Region 1: Lorenz $(\sigma, b, r) = (10, 8/3, 28)$ in chaotic region

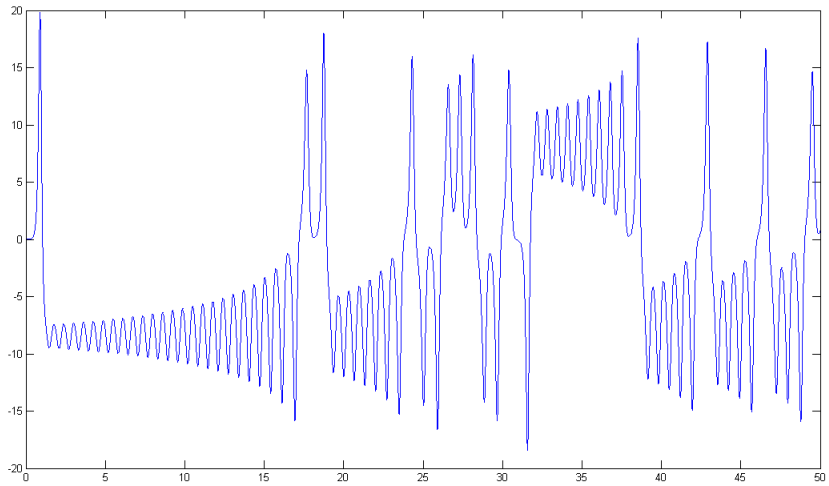


(b) Region 2: Lorenz $(\sigma, b, r) = (10, 8/3, 7)$ with two fixed points

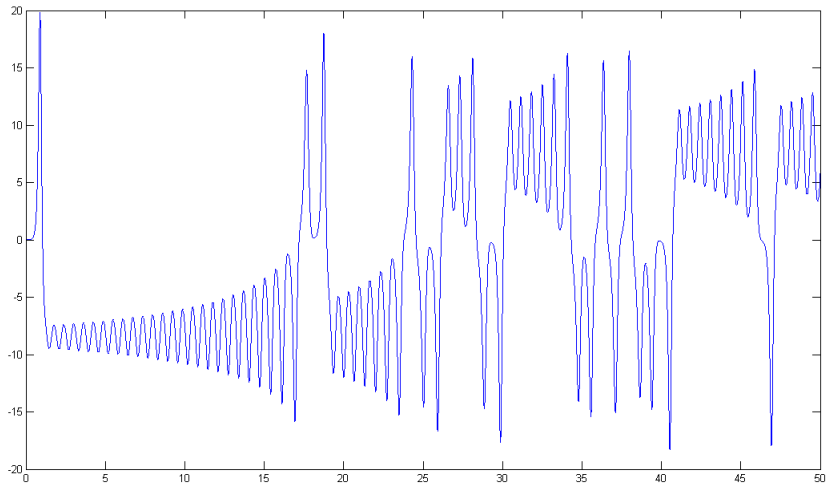


(c) Region 3: Lorenz $(\sigma, b, r) = (10, 8/3, 0.90)$ with one fixed point as the origin

Figure A.1: Various regions in Lorenz system



(a) Lorenz $(\sigma, b, r) = (10, 8/3, 28)$ initialized with $(x_0, y_0, z_0) = (0.0020, 0.001, 0.001)$



(b) Lorenz $(\sigma, b, r) = (10, 8/3, 28)$ initialized with $(x_0, y_0, z_0) = (0.0021, 0.001, 0.001)$

Figure A.2: Sensitivity to the initial conditions in Lorenz system

requirements that we consider in designing a relay-attack proof system for RFIDs is listed as follows:

- **Chaotic Behavior:** we re-emphasize that the system has to be aperiodic and sensitive to initial conditions. Chaotic properties are beneficial in security.
- **One Way:** by observing the chaotic outputs of a system, it has to be difficult to find the system parameters. This property provides that the chaotic system parameters can be chosen as security parameters of the system.
- **Large Key Space:** the range of parameters for which the system behaves chaotically should be very large. In other words, there must be exponentially many parameters that can cause a chaotic behavior. Depending on system parameters, chaos suppression happens at different values of coupling.
- **Easy Setup:** not only should it be straightforward to find a set of parameters for chaotic behavior, but also the parameters for which chaos is suppressed or synchronized should be widely available. With the knowledge of system and its parameters, one should be able to easily set the values of a suppressing/coupling force that quenches the chaos.
- **Simple Design:** RFID tags are bounded by size and cost. Therefore, additional circuitry to implement a chaotic system into an RFID tag should be small.
- **Energy Efficiency:** the main constraint in RFID tags is power. Passive tags receive

their power from the reader and they do not require any battery on board. Thus, passive tags are smaller and cheaper than active tags that come with a batter.

The first two requirements concern about the strength of the chaotic system to be used in a security application. On one hand, we would like the chaotic system to be difficult to predict. On the other hand, it should be straightforward to find the chaotic regions of the system. The last two requirements deal with implementation complexities of a chaotic system into an RFID tag.