Gröbner Basis Attacks on Lightweight RFID Authentication Protocols

Daewan Han*

Abstract—Since security and privacy problems in RFID systems have attracted much attention, numerous RFID authentication protocols have been suggested. One of the various design approaches is to use light-weight logics such as bitwise Boolean operations and addition modulo 2^m between m-bits words. Because these operations can be implemented in a small chip area, that is the major requirement in RFID protocols, a series of protocols have been suggested conforming to this approach. In this paper, we present new attacks on these lightweight RFID authentication protocols by using the Gröbner basis. Our attacks are superior to previous ones for the following reasons: since we do not use the specific characteristics of target protocols, they are generally applicable to various ones. Furthermore, they are so powerful that we can recover almost all secret information of the protocols. For concrete examples, we show that almost all secret variables of six RFID protocols, LMAP, M^{2}AP, EMAP, SASI, Lo et al.'s protocol, and Lee et al.'s protocol, can be recovered within a few seconds on a single PC.

Keywords—RFID, Authentication Protocol, Algebraic Attack, Gröbner Basis

1. INTRODUCTION

Recently, RFID systems have attracted much attention due to their wide applications including inventory control, logistics, supply chain management, etc.

Since communications in RFID systems are executed on public RF channels and any reader can get the tag's information easily, the systems without proper protection are vulnerable to eavesdropping, tracking, tag forgery, and cloning. These security and privacy problems prohibit more rapid and widespread deployments of RFID. To address these problems, numerous RFID authentication protocols have been suggested in various ways. One of them is to design authentication protocols by using only lightweight logics, such as bitwise Boolean operations and addition modulo 2^m, between m-bits words. Because these operations can be implemented in a small chip area, which is the major requirement in RFID protocols, a series of protocols have been suggested that conform to this design approach. The vulnerabilities of those protocols have also been presented. However, they were so specific to each target protocols that they were not applicable the other protocols. Furthermore, in most cases their concern was not to recover secure information but to point out the simple security flaws of the protocols.

In this paper, we present new attacks on these lightweight RFID authentication protocols using the Gröbner basis. Our attacks are superior to previous ones for the following reasons: since

Manuscript received June 16, 2011; accepted August 29, 2011.
Corresponding Author: Daewan Han
* The Attached Institute of ETRI, Daejeon, Korea (dwh@ensec.re.kr)

Table 1. Summary of our results

Protocol names	Eavesdropping (sessions)	Average time (sec.)	Recovery rates (%)	Guess (trial)
LMAP	3	26.5	100	
M^2AP-1	2	7.9	100	
M^2AP-2	5	3.8	94.9	
EMAP	3	4.7	99.9	
SASI	5	151.8	100	96^2
Lo *et al.*	1	0.4	100	96

we do not use the specific characteristics of target protocols, our attacks are generally applicable to various ones. Furthermore, the attacks are so powerful that we can recover almost all of the secret information of the protocols. For concrete examples, we show that almost all secret variables of the six RFID protocols, LMAP, M^2AP, EMAP, SASI, Lo *et al.*'s protocol(LSY), and Lee *et al.*'s protocol(LHYC), can be recovered within a few seconds. In the attack we need to solve systems of multivariate non-linear equations over GF(2). All the equations are induced from the mathematical relations between secret variables and public ones that are obtained by passively eavesdropping on a few consecutive communication sessions. For solving the equation systems we used the Gröbner basis method, which is an emerging technique in recent cryptanalysis.

The summary of our results are given in Table 1. In the table below, the second column denotes the number of eavesdropping sessions needed for the attacks and the third one means the average time consumed in recovering the secret variables by a Gröbner basis algorithm. The fourth column denotes the rates of the number of recovered variables over all the secrets and the final column means the number of guess needed for the attack.

Our results remind us that designing secure and lightweight RFID authentication protocols is not easy work. Using only simple arithmetic operations should especially be carefully considered.

The rest of this paper is organized as follows: in Section 2 we introduce the preliminaries of this paper, which are RFID security, algebraic attacks, the Gröbner basis, and the notations used in the paper. In Section 3 we briefly introduce the RFID authentication protocols of our concern and the overview of our attack. We describe the detailed cryptanalysis on six protocols and their simulation results in Section 4. Finally we conclude the paper in Section 5.

2. PRELIMINARIES

2.1 Survey on RFID and its Security

A RFID system is an automated identification technology in which a small transponder (tag), attached to a real world object, receives and responds to radio-frequency queries from a transceiver (reader). It has been already widely used in daily life for access control, various payment systems, electronic identification cards, and so on. However, the most notable and revolutionary application of RFID system in the near future will be the replacement of the current barcode

system for supply chain management, inventory control, and anti-counterfeiting. The attractiveness of the RFID over the barcode is twofold. First, unlike a barcode scanner, an RFID reader does not require line-of-sight or physical contact to scan an RFID tag. This feature reduces the cumbersome need for manual intervention in the scanning process. Secondly, an RFID tag assigns a unique serial number to an individual item, while a barcode typically specifies the type of product it is printed on. The unique identifier associated with an object can serve as a pointer to a database entry containing the detailed history of the object. Thanks to these features of automated scanning and unique identification, the RFID promises fine-grained tracking of inventory on an unprecedented scale.

Although RFID systems have many benefits as described above, they have some problems with security and privacy, which prohibit the more rapid and widespread deployment of them. Since the communication between a tag and a reader is executed on public RF channels and any reader can get the tag's information easily, the systems are vulnerable to eavesdropping, tag forgery, cloning, tracking, and so on. We refer to [28] for more details on these issues. To address these problems, numerous physical protections and logical protocols have been suggested [3, 18]. Though physical protections such as the blocker tag [19] might be efficient solutions for some applications, they have limits for general and broad uses; hence they have not attracted much attention. On the contrary, logical protocols for RFID security have been intensively studied in various directions. In the early days several protocols using cryptographic hash functions and block ciphers were suggested [15, 28]. However, this approach was proved to be inadequate for RFID security, since currently used hash functions and block ciphers are too expensive to be operated in low-cost RFID tags [16]. The second approach was to design protocols by using hard mathematical instances such as LPN [20]. However, their securities are currently uncertain and still evolving. Another approach was to design protocols using only lightweight logics such as Boolean operations and integer modular addition. These protocols are our concerns in this paper, so we will introduce them minutely in Section 3.1.

2.2 Algebraic Attacks and the Gröbner Basis

Any cryptographic primitive can be modeled by a system of multivariate equations over a finite field. The basic principle of algebraic attacks is to construct and solve the equation systems whose solutions have a correspondence to secret information of the original cryptosystem. The attacks have been successfully applied to some multivariate public key cryptosystems [14] and stream ciphers [10, 11]. The feasibility of them against block ciphers is one of the current research areas in cryptography [5, 8, 9]. The core of the attack is how to solve the equation systems. Indeed, the problem of solving multivariate polynomial equations is known to be NP-complete. However, the systems derived from cryptosystems are mostly sparse and their solutions lie in GF(2), in which cases they can be solved somewhat easily. So far many algorithmic techniques and tools are suggested to resolve it. Among them, the most popular and effective one is to use the Gröbner basis of polynomial ideals.

A Gröbner basis is a high-level mathematical concept with many applications in commutative algebra, algebraic geometry, and computational algebra [1]. For its broadness, a complete introduction of the concept is beyond this paper, and we don't need a deep understanding of the notion to understand this paper. Here, we briefly introduce its relevance to cryptography. As we mentioned above, the main relevance of Gröbner basis to cryptography is the utility in solving

polynomial equation systems. If we have the following equation system over a field F

$$f_1(x_1, \cdots, x_n) = 0, \cdots, f_m(x_1, \cdots, x_n) = 0,$$

then we can find its solution set by computing the ideal $I =< f_1, \cdots, f_m >$ and the associated variety v(I). By the way, the Gröbner basis of I can provide implicit solutions to the equation system over the algebraic closure of F.

There are some reasons for the Gröbner basis to be more effective in solving equation systems arising in cryptography. Firstly, the solution of cryptographic equation systems mostly lie in a small finite field GF(q) and in most cases in GF(2). So, we can restrict the solution to the base field GF(q) by adding the field relations $x_i^q - x_i = 0$ to the original equation system.

This causes the basis to be found more quickly, since Gröbner basis algorithms work more efficiently when there are more equations. Secondly, cryptographic equation systems usually have a unique solution. In those cases we can effectively try the so-called guess-and-determine-attack. Usually Gröbner basis algorithms tend to be terminated more quickly when an equation system has no solution. Suppose an equation system has a unique solution. We can make a modified equation system by guessing some parts of the solution. If we guess inadequately, then the modified system will have no solution. Thus, Gröbner basis algorithms will output a result quickly, which enables us to know that we guessed badly.

There are well-known algorithms for computing the Gröbner basis of a given polynomial ideal [12,13], and practical S/W packages implementing them. In this paper we used the package PolyBoRi [4].

2.3 Notations

Let X,Y be *m*-bits word, where *m* may be any non-negative integers, but mostly 96 and 128 in this paper. In the rest of this paper, we will use the following notations.

- x_i: *i*-th lsb of *X*, that is, $X = x_{m-1} \cdots x_0$.
- X $\oplus$ Y: bitwise XOR of X and Y.
- X $\wedge$ Y: bitwise AND of X and Y.
- X $\vee$ Y: bitwise OR of X and Y.
- X $\boxplus$ Y: addition of X and Y modulo 2^m.
- X $\boxminus$ Y: subtraction of X and Y modulo 2^m.
- $x + y$: addition of *x* and *y* over GF(2).
- $x \cdot y$ (xy): multiplication of *x* and *y* over GF(2).
- X $\gg l$: right *l*-bits shift of X.
- X $\ggg l$: right *l*-bits rotation of X.
- X $\lll l$: left *l*-bits rotation of X.
- $Rot(X)$: *n*-bits left rotation of X, where *n* is the hamming weight of X.
- $S_1(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} x_i$.
- $S_2(x_1, x_2, \cdots, x_n) = \sum_{1 \le i < j \le n} x_i x_j$.

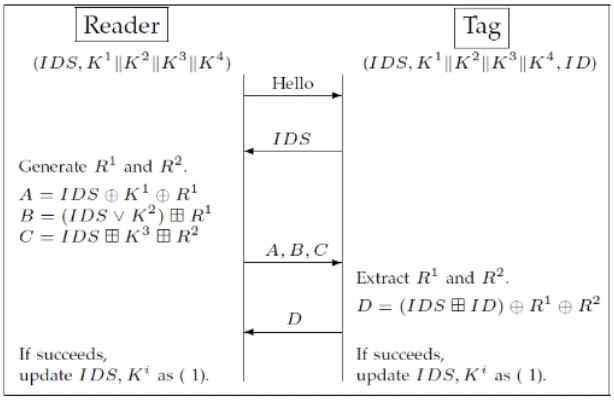

Fig. 1. LMAP Protocol in a simplified form

3. GRÖBNER BASIS ATTACKS ON RFID AUTHENTICATION PROTOCOLS

3.1 RFID Authentication Protocols Based on Lightweight Operations

Since RFID tags are required to be low cost, they can only store hundreds of bits and have 5K-10K logic gates, and only 0.25K-3K gates can be devoted to security tasks [28]. Thus, standard cryptographic algorithms are too heavy to be implemented on such low-gated RFID tags. However, Boolean operations such as XOR, AND, OR, and modular addition are light enough to be implemented on them and multiple combinations of these operations seem to be a good candidates for one-way functions. These merits induce people to use those operations as building blocks for designing RFID authentication protocols. Of course, they are followed by various cryptographic scrutiny [2, 6, 17, 26]. However, since the merits of using only lightweight operations still exist and the defects seem to be easily remedied, modified protocols still have been suggested.

In this paper we show that these protocols are very vulnerable to a new attack by using the Gröbner basis. Concrete targets of our attack are LMAP [23], M^2AP [24], EMAP [25], SASI [27], Lo *et al.*'s protocol (LSY) [22], and Lee *et al.*'s protocol (LHYC) [21]. They have the standard challenge-response mutual authentication structures. So, here we describe only LMAP in detail and the other ones are simply introduced in the following subsections.

LMAP is based on the use of an index-pseudonym (IDS), which is the index of a table where all the information about a tag is stored. Each tag is associated to key K, which is divided into four parts each of 96 bits ($K = K^1 \| K^2 \| K^3 \| K^4$). In each tag, the 96-bits static identification number (ID) with the IDS and the key are stored, and the IDS and K are updated at the end of each session. The mutual authentication procedure is as follows (See also Fig. 1):

The reader sends a hello message to the tag and the tag answers by sending its current IDS. By means of IDS, only an authorized reader is able to access the tag secret key K, which is necessary to carry out the next authentication stage. Then, the reader generates two random numbers R^1 and R^2. With R^1, R^2, K^1, K^2 and K^3 it generates the messages A, B, and C and then transmits them to the tag. Then the tag authenticates the reader and obtains R^1 and R^2 from A, B, and C. The random numbers R^1 and R^2 are used in updating the index-pseudonym and the key. Once these verifications are performed, the tag generates the answer message D to authenticate

and transmit its static identifier in a secure form. The messages A, B, C, and D are calculated as in Fig. 1. After the reader and the tag have been authenticated mutually, the index-pseudonym and the key updating stage are carried out as the following:

$$\begin{aligned} \mathrm{IDS}' &= (\mathrm{IDS} \boxplus (\mathrm{R}^2 \boxplus \mathrm{K}^4)) \oplus \mathrm{ID}, \\ \mathrm{K}^{1\prime} &= \mathrm{K}^1 \oplus \mathrm{R}^2 \oplus (\mathrm{K}^3 \boxplus \mathrm{ID}), \\ \mathrm{K}^{2\prime} &= \mathrm{K}^2 \oplus \mathrm{R}^2 \oplus (\mathrm{K}^4 \boxplus \mathrm{ID}), \\ \mathrm{K}^{3\prime} &= (\mathrm{K}^3 \oplus \mathrm{R}^1) \boxplus (\mathrm{K}^1 \oplus \mathrm{ID}), \\ \mathrm{K}^{4\prime} &= (\mathrm{K}^4 \oplus \mathrm{R}^1) \boxplus (\mathrm{K}^2 \oplus \mathrm{ID}). \end{aligned} \tag{1}$$

3.2 Overview of Our Attacks

The object of our attacks is to find all secret variables of the protocols. It is passive, that is, we only need eavesdropping for a few consecutive sessions for the attack. The number of eavesdropping sessions depends on the number of equations needed for the systems to be over-defined. We need to guess some bits of secret variables in the attack of SASI, LSY, and LHYC, but they are also mounted within practical time.

The general procedure for the attack is as follows:

1) Figure out the relations between secret and public variables obtained from a session of the protocol.
2) Convert the relations to quadratic equations over GF(2).
3) Estimate the number of consecutive sessions for the system to be over-defined.
4) Construct a system of equations over GF(2) from all the relations of the consecutive sessions.
5) Compute a Gröbner basis of the ideal generated by the polynomials in the equation system.
6) If Gröbner basis algorithms are terminated abnormally or they cannot produce sufficient solutions, increase the number of sessions and repeat 4) and 5), until we get all secret variables from the solution.

We used the S/W package PolyBori implemented at the Sage distribution of mathematical software [27] for a Gröbner basis algorithm. The simulations are run on a PC with 2.53GHZ Dual CPU and 3GB memory.

3.3 Converting Bitwise Operations to Algebraic Equations over GF(2)

In Gröbner basis attacks we need to convert relations which are given by bitwise operations and addition modulo 2^m to algebraic equations over GF(2). The following theorem helps us to perform it automatically.

Theorem 1. Let X, Y and Z be m-bits words. Then, the following four relations are converted to the corresponding system of quadratic equations in x_i, y_i and z_i over GF(2).

1) $\mathrm{Z} = \mathrm{X} \oplus \mathrm{Y} \iff \mathrm{S}_1(x_i, y_i, z_i) = 0$ for $i = 0, \cdots, m-1$.
2) $\mathrm{Z} = \mathrm{X} \wedge \mathrm{Y} \iff z_i + \mathrm{S}_2(x_i, y_i) = 0$ for $i = 0, \cdots, m-1$.
3) $\mathrm{Z} = \mathrm{X} \vee \mathrm{Y} \iff \mathrm{S}_1(x_i, y_i, z_i) + \mathrm{S}_2(x_i, y_i) = 0$ for $i = 0, \cdots, m-1$.

4) $\mathrm{Z} = \mathrm{X} \boxplus \mathrm{Y} \iff \mathrm{S}_1(x_0, y_0, z_0) = 0,$
$\mathrm{S}_1(x_1, y_1, z_1) + \mathrm{S}_2(x_0, y_0) = 0,$
$\mathrm{S}_1(x_i, y_i, z_i, x_{i-1}, y_{i-1}) + \mathrm{S}_2(x_{i-1}, y_{i-1}, z_{i-1}) = 0$ for $i = 2, \cdots, m-1$.

Proof. 1), 2), and 3) are easily verified by the definition of each operations and the truth table of $\oplus, \wedge$ and $\vee$.

For 4), we can describe $\mathrm{X} \boxplus \mathrm{Y}$ by using carry bits $c_1, \cdots, c_{m-1}$ as follows.

$$z_0 = x_0 + y_0, \tag{2}$$
$$c_1 = x_0 y_0, \tag{3}$$
$$z_i = x_i + y_i + c_i \text{ for } i \geq 1, \tag{4}$$
$$c_{i+1} = x_i y_i + c_i(x_i + y_i) \text{ for } i \geq 1. \tag{5}$$

The first equation of 4) is equal to (2). The second equation is obtained by cancelling c_1 in (3) and (4) for i=1. Using (4) and (5) for $i \geq 2$, we obtain

$$\begin{aligned} c_i &= x_{i-1}y_{i-1} + c_{i-1}(x_{i-1} + y_{i-1}) \\ &= x_{i-1}y_{i-1} + (z_{i-1} + x_{i-1} + y_{i-1})(x_{i-1} + y_{i-1}) \\ &= x_{i-1}y_{i-1} + z_{i-1}(x_{i-1} + y_{i-1}) + (x_{i-1} + y_{i-1}). \end{aligned}$$

Replacing c_i again by $(x_i + y_i + z_i)$ we obtain the final equation. ■

Remark 1. RFID authentication protocols in this paper consist of more complex operations such as $\mathrm{W} = (\mathrm{X} \boxplus \mathrm{Y}) \oplus \mathrm{Z}$ or $\mathrm{W} = (\mathrm{X} \oplus \mathrm{Y}) \boxplus \mathrm{Z}$. However, these equations also can be converted to quadratic equations without additional variables. For example, $\mathrm{W} = (\mathrm{X} \boxplus \mathrm{Y}) \oplus \mathrm{Z}$ is equivalent to the following system of quadratic equations over GF(2).

$$\mathrm{S}_1(x_0, y_0, z_0, w_0) = 0,$$
$$\mathrm{S}_1(x_1, y_1, z_1, w_1) + \mathrm{S}_2(x_0, y_0) = 0,$$
$$\mathrm{S}_1(x_i, y_i, z_i, w_i, x_{i-1}, y_{i-1}) + \mathrm{S}_2(x_{i-1}, y_{i-1}, z_{i-1}) + w_{i-1}(x_{i-1} + y_{i-1}) = 0 \text{ for } i \geq 2.$$

4. CRYPTANALYSIS OF RFID PROTOCOLS

In this section, we describe the concrete attacks against six RFID authentication protocols and present the simulation results. The simulations were carried out in at least 10 random instances for all cases. Since registration, identifying, authentication, and variable updating procedures of all protocols are similar to the LMAP introduced in Section 3.1. We describe each protocol briefly, only by presenting the variables and relations between them.

4.1 Cryptanalysis of LMAP

4.1.1 LMAP

- Secret variables
  · Shared secrets: K^1, K^2, K^3, K^4, ID
  · Random numbers: R^1, R^2

- Mutual authentication
  · Tag → Reader: IDS
  · Reader → Tag: A ‖ B ‖ C
  · Tag → Reader: D

$$\begin{aligned} A &= IDS \oplus K^1 \oplus R^1, \\ B &= (IDS \vee K^2) \boxplus R^1, \quad (6) \\ C &= IDS \boxplus K^3 \boxplus R^2, \quad (7) \\ D &= (IDS \boxplus ID) \oplus R^1 \oplus R^2. \end{aligned}$$

- Updating variables

$$\begin{aligned} IDS' &= \left(IDS \boxplus (R^2 \boxplus K^4)\right) \oplus ID, \\ K^{1\prime} &= K^1 \oplus R^2 \oplus (K^3 \boxplus ID), \\ K^{2\prime} &= K^2 \oplus R^2 \oplus (K^4 \boxplus ID), \\ K^{3\prime} &= (K^3 \oplus R^1) \boxplus (K^1 \oplus ID), \\ K^{4\prime} &= (K^4 \oplus R^1) \boxplus (K^2 \oplus ID). \end{aligned}$$

4.1.2 Cryptanalysis

Except for (6) and (7), the above relations can be converted to quadratic equations over GF(2) by directly applying the rules in Theorem 1 and Remark 1. Since we know IDS, we can fix each bit of $IDS \vee K^2$ to 1 or k_i^2 according to ids_i. Thus, (6) is converted to quadratic equations in k_i^2 and r_i^1. For (7), it is equivalent to $C \boxminus IDS = K^3 \boxplus R^2$. Since we know the left term of this relation, it can be regarded as a modular addition of two unknowns, so we can also apply Theorem 1 directly. Therefore, we can convert all relations to a system of quadratic equations over GF(2) without any additional variables.

If we eavesdrop on one session of an LMAP, we can obtain 4 relations w.r.t. 7 secret variables. Thus, the converted system of equations cannot be over-defined. However, if we eavesdrop the next session, we can get another 9 relations (5 in updating variables and 4 in the authentication procedure of the second session) w.r.t. 13 unknown variables. Note that ID is not changed in the second session. The converted equation system from these relations may be over-defined. However, this is not true in practice. So, we increase the number of eavesdropping sessions until the converted system becomes to be over-defined. The simulation result in Table 2 shows that we can find all secrets from eavesdropping on 3 sessions.

In Table 2 the first column denotes the number of eavesdropping sessions. The next two columns denotes the number of variables and equations in converted quadratic equation systems over GF(2). The fourth and fifth columns denote the average and maximum time consumed in finding a Gröbner basis. The last column denotes the ratio of the recovered ones over the all

Table 2. Simulation results for LMAP

Sessions	Variables	Equations	Avg. (sec)	Max. (sec)	Rate (%)
2	1,248	1,248	124.7	307.1	93.9
3	1,824	2,112	26.5	27.7	100

secret variables. In conclusion, we can find all the secret variables of LMAP within about 30 seconds by eavesdropping on 3 consecutive sessions.

The basic principle of the attacks against the other five protocols in this paper is same to that of LMAP. Thus, from the next subsection we describe only the notable features of each protocol used in constructing equation systems and simulation results.

4.2 Cryptanalysis of M^{2}AP

4.2.1 M^{2}AP

- Secret variables
  · Shared secrets: K^1, K^2, K^3, K^4, ID
  · Random numbers: R^1, R^2

- Mutual authentication
  · Tag → Reader: IDS
  · Reader → Tag: $A \| B \| C$
  · Tag → Reader: $D \| E$

$$A = IDS \oplus K^1 \oplus R^1, \quad (8)$$
$$B = (IDS \wedge K^2) \vee R^1, \quad (9)$$
$$C = IDS \boxplus K^3 \boxplus R^2, \quad (10)$$
$$D = (IDS \vee K^4) \wedge R^2, \quad (11)$$
$$E = (IDS \boxplus ID) \oplus R^1. \quad (12)$$

- Updating variables

$$IDS' = (IDS \boxplus (R^2 \boxplus R^1)) \oplus ID, \quad (13)$$
$$K^{1\prime} = K^1 \oplus R^2 \oplus (K^3 \boxplus ID), \quad (14)$$
$$K^{2\prime} = K^2 \oplus R^2 \oplus (K^4 \boxplus ID), \quad (15)$$
$$K^{3\prime} = (K^3 \oplus R^1) \boxplus (K^1 \oplus ID), \quad (16)$$
$$K^{4\prime} = (K^4 \oplus R^1) \boxplus (K^2 \oplus ID). \quad (17)$$

4.2.2 Cryptanalysis

We divide the attack into two phases. In the first phase, we recover all secrets, except K^2 and K^4. Next, we recover most bits of K^2 and K^4.

Firstly, suppose we eavesdropped on consecutive two sessions of M^{2}AP. Then, we obtain 9 relations ((8), (10), (12), (13), (14), (16) in the first session and (8), (10), (12) in the second session) w.r.t. 9 unknown variables K^1, K^3, $K^{1\prime}$, $K^{3\prime}$, R^1, R^2, $R^{1\prime}$, $R^{2\prime}$, and ID. If we convert these relations to an equation system over GF(2) and solve it, we can recover all variables. Table 3 shows the simulation results. The meaning of each column of the table is same to that of Table 2.

Table 3. Simulation results for M^{2}AP in the first phase

Sessions	Variables	Equations	Avg. (sec)	Max. (sec)	Rate (%)
2	864	864	7.9	8.2	100

Table 4. Simulation results for M^2AP in the second phase

Sessions	Variables	Equations	Avg. (sec)	Max. (sec)	Rate (%)
4	768	1,344	6.9	13.8	89.3
5	960	1,728	3.8	5.1	94.9
6	1,152	2,112	5.1	6.2	96.5

In the second phase, we may assume that all variables except for K^2 and K^4 are revealed in all sessions from the first phase attack. Thus, the relations of our concern are (9), (11), (15), and (17). We constructed equation systems from those relations and tried to solve them. The simulation results are given in Table 4. Most of the systems that were obtained from less than three sessions were not solved. In the case of three sessions, about a half of the systems were solved, but the recovering rates are less than 80%. From four sessions we could get solutions in almost all instances. These results accord with the well-known properties of Gröbner basis algorithms that work better with an increasing number of equations.

4.3 Cryptanalysis of EMAP

4.3.1 EMAP

- Secret variables
  · Shared secrets: K^1, K^2, K^3, K^4, ID
  · Random numbers: R^1, R^2

- Mutual authentication
  · Tag → Reader: IDS
  · Reader → Tag: $A \| B \| C$
  · Tag → Reader: $D \| E$

$$\begin{aligned} A &= IDS \oplus K^1 \oplus R^1, \\ B &= (IDS \vee K^2) \oplus R^1, \\ C &= IDS \oplus K^3 \oplus R^2, \\ D &= (IDS \wedge K^4) \oplus R^2, \\ E &= (IDS \wedge R^1 \vee R^2) \oplus ID \oplus K^1 \oplus K^2 \oplus K^3 \oplus K^4. \end{aligned} \quad (18)$$

- Updating variables

$$\begin{aligned} IDS' &= IDS \oplus R^2 \oplus K^1, \\ K^{1\prime} &= K^1 \oplus R^2 \oplus (ID(95:48) \| F(K^4) \| F(K^3)), \\ K^{2\prime} &= K^2 \oplus R^2 \oplus (F(K^1) \| F(K^4) \| ID(47:0)), \\ K^{3\prime} &= K^3 \oplus R^1 \oplus (ID(95:48) \| F(K^4) \| F(K^2)), \\ K^{4\prime} &= K^4 \oplus R^1 \oplus (F(K^3) \| F(K^1) \| ID(47:0)), \end{aligned}$$

where X(*i*:*j*) means (*i*-*j*+1)-bits string $x_i \| x_{i-1} \| \cdots \| x_j$ and F(X) means 24-bits string X(95:72)⊕ X(71:48) ⊕X(47:24) ⊕X(23:0).

4.3.2 Cryptanalysis

Since IDS is known and X(*i*:*j*), F(X) are linear, all relations of EMAP except for (18) are converted to linear equations over GF(2). Thus, the equation systems of EMAP are much simpler than those of previous protocols. This is the reason why EMAP is more easily broken than LMAP and M^{2}AP, which is checked from the simulation results given in Table 5.

Table 5. Simulation results for EMAP

Sessions	Variables	Equations	Avg (sec)	Max. (sec)	Rate (%)
2	1,248	1,440	1.69	1.72	99.54
3	1,824	2,400	4.69	4.87	99.97

4.4 Cryptanalysis of SASI

4.4.1 SASI

- Secret variables
  · Shared secrets: K^1, K^2, ID
  · Random numbers: R^1, R^2

- Mutual authentication
  · Tag → Reader: IDS
  · Reader → Tag: $A \| B \| C$
  · Tag → Reader: D

$$A = IDS \oplus K^1 \oplus R^1,$$
$$B = (IDS \vee K^2) \boxplus R^2,$$
$$X^1 = (K^1 \oplus R^2) \lll K^1, \quad (19)$$
$$X^2 = (K^2 \oplus R^1) \lll K^2, \quad (20)$$
$$C = (K^1 \oplus X^2) \boxplus (X^1 \oplus K^2),$$
$$D = (X^2 \boxplus ID) \oplus ((K^1 \oplus K^2) \vee X^1), \quad (21)$$

- Updating variables

$$IDS' = (IDS \boxplus ID) \oplus (R^2 \oplus X^1),$$
$$K^{1\prime} = X^1,$$
$$K^{2\prime} = X^2.$$

4.4.2 Cryptanalysis

Since we do not know the rotation bits of (19) and (20), in the attack of SASI we should try the guess-then-determine-attack.

Firstly, we assume that the rotation bits of (19) and (20) are known exactly. Then, we can construct an equation system of SASI like the previous subsections. Since (21) is so complicated, we divide it into two relations adding an additional variable as follows:

$$T = (K^1 \oplus K^2) \vee X^1,$$
$$D = (X^2 \boxplus ID) \oplus T.$$

Table 6. Simulation results for SASI

Sessions	Variables	Equations	Avg. (sec)	Max. (sec)	Rate (%)
3	1,728	2,208	246.4	629.1	100
4	2,208	2,976	61.5	80.1	100
5	2,688	3,744	151.8	184.0	100

With only one session, the converted equation system seems to be over-defined. But, the simulation showed that we need more than 3 sessions. In fact, with less than three sessions, PolyBori is terminated abnormally. With three sessions, in 7 over 10 instances we could find all variables. We could recover all secrets in all instances with four sessions. Simulation results are given in Table 6.

Now we should mention the complexity and method of guessing rotation bits in (19) and (20). In each relation the number of different ways of rotation is at most 96, because the bit-length of all variables is 96. Thus, the number of guesses needed for the attack is not more than 96^2, which is also practical in a PC. If equation systems have no solution, PolyBori informs us that the systems are inconsistent. In those cases it tends to finish faster than when systems have solutions as described early in Section 2.2. If we incorrectly guess the rotation bits of SASI, the possibility that equation systems are consistent will be extremely low, and PolyBori will finish earlier than the average time in Table 6. From the simulation, we verified these phenomena. For example, in the case of 4 sessions PolyBori finished within 34 seconds on average for the wrongly guessed instances, while it returned the exact solutions after 61 seconds for the consistent ones.

4.5 Cryptanalysis of LSY

4.5.1 LSY

- Secret variables
  · Shared secrets: K, ID
  · Random numbers: R^2, R^3, R^4, R^5

- Mutual authentication and updating variables
  · Reader → Tag: R^1
  · Tag → Reader: $GID \parallel A^1 \parallel B^1$
  · Reader → Tag: $A^2 \parallel B^2$
  · Tag → Reader: $A^3 \parallel B^3$
  · Reader → Tag: Select GID' and K', and sends $A^4 \parallel B^4 \parallel C^4 \parallel D^4 \parallel E^4$
  · Tag updates GID, K to GID', K' from B^4, C^4.

$$A^1 = K \oplus R^2, \tag{22}$$
$$B^1 = (K \vee ((R^1 \boxplus K) \gg 1) \oplus (R^2 \ggg ((R^1 \boxplus K) \gg 1)), \tag{23}$$
$$A^2 = (R^2 \wedge K) \oplus R^3, \tag{24}$$
$$B^2 = (((R^2 \wedge K) \vee R^3) \lll (R^2 \gg 1)) \oplus R^2,$$
$$A^3 = (R^3 \vee K) \oplus R^4,$$
$$B^3 = (((R^3 \vee K) \wedge R^4) \ggg (R^3 \gg 1)) \oplus ID,$$
$$A^4 = (R^4 \wedge K) \oplus R^5,$$

$$B^4 = (((R^4 \wedge K) \vee R^5) \lll (R^4 \gg 1)) \oplus GID',$$
$$C^4 = (R^4 \wedge K) \oplus K',$$
$$D^4 = (GID' \vee R^5) \oplus (R^4 \wedge K),$$
$$E^4 = (R^4 \wedge K \wedge R^5) \oplus K'.$$

4.5.2 Cryptanalysis

Firstly, we consider only two equations (22) and (23). Then, the secret variables are only K and R^2, and the converted system of equations is over-defined in most cases. So, we can get K and R^2 by solving the system. The simulation results are given in Table 7.

Next, R^3 can be recovered from (24), and the remaining secret variables R^4, ID, R^5, and K' can be also recovered from the next relations sequentially.

Like SASI, we should guess the amount of rotated bits in (23). The plausibility of a guess-then-determine strategy is the same as SASI, and the complexity of guess is at most 96.

Table 7. Simulation results for LSY

Sessions	Variables	Equations	Avg. (sec)	Max. (sec)	Rate (%)
1	384	480	0.38	0.41	100

4.6 Cryptanalysis of LHYC

4.6.1 LHYC

- Secret variables (128 bits)
  · Shared secrets: K
  · Random numbers: R

- Mutual authentication
  · Tag → Reader: IDT
  · Reader → Tag: A ‖ B
  · Tag → Reader: C

$$A = K \oplus R, \tag{25}$$
$$B = Rot(K) \oplus Rot(R), \tag{26}$$
$$C = (K \vee Rot(R)) \oplus (Rot(K) \wedge R).$$

- Updating IDT and secret keys

$$IDT' = K \oplus Rot(R),$$
$$K' = Rot(K) \oplus R.$$

4.6.2 Cryptanalysis

Like SASI and LSY, we need to guess the hamming weights of K and R. If they are correctly guessed, the equation system induced by (25) and (26) in only one session is enough to get K and R and the system is even linear. Thus, we do not need Gröbner basis algorithms for solving

the system and we omit the simulation. The complexity for guessing is at most 128^2.

5. CONCLUSION

In this paper we showed that six RFID authentication protocols based on bitwise operations and an addition of modulo 2^m are very vulnerable to an algebraic attack with the Gröbner basis. We could recover almost all of the secret variables of those protocols within a few seconds by our attack.

Our result shows that RFID authentication protocols should be more carefully designed, if we want to use simple arithmetic operations. Another contribution of this paper is that we presented the evidence that Gröbner basis is a powerful toolkit for cryptanalysis.

REFERENCES

[1] W.W. Adams and P. Loustaunau, "An Introduction to Gröbner Bases," *Graduate Studies in Mathematics*, Vol.3, AMS, 1994.

[2] B. Alomair, L. Lazos, and R. Poovendran, *"Passive Attacks on a Class of Authentication Protocols for RFID," Proceedings of ICISC 2007*, LNCS 4817, Springer-Verlag, 2007, pp.102-115.

[3] G. Avoine, Cryptography in Radio Frequency Identification and Fair Exchange Protocols [dissertation]. Lausanne, Switzerland: EPFL; 2005.

[4] M. Brickenstein, A. Dreyer, PolyBoRi: *"A Framework for Gröbner Basis Computations with Boolean Polynomials," Electronic Proceedings of the MEGA 2007 - Efficient Methods in Algebraic Geometry*, Strobl, Austria, 2007.

[5] J. Buchmann, A. Pyshkin, and R-P Weinmann, *"Block Ciphers Sensitive to Gröbner Basis Attacks," Proceedings of CT-RSA 2006*, LNCS 3860, Springer-Verlag, 2006, pp.313-331.

[6] T. Cao, L. Bertino, and H. Lei, "Security Analysis of the SASI Protocol," *IEEE Transactions on Dependable and Secure Computing*, Vol.6, No.1, 2009, pp.73-77.

[7] H.-Y. Chien, "SASI: A New Ultralightweight RFID Authentication Protocol Providing Strong Authentication ans Strong Integrity," *IEEE Transactions on Dependable and Secure Computing*, Vol.4, No.4, 2007, pp.337-340.

[8] C. Cid, S. Murphy, and M. Robshaw, *Algebraic Aspects of the Advanced Encrpytion Standard*, Springer-Verlag, 2006.

[9] N. Courtois, and J. Pieprzyk, *"Cryptanalysis of Block Ciphers with Over-defined System of Equations," Proceedings of Asiacrypt 2002*, LNCS 2501, Springer-Verlag, 2002, pp.267-287.

[10] N. Courtois, *"Fast Algebraic Attacks on Stream Ciphers with Linear Feedback," Proceedings of Crypto 2003*, LNCS 2729, Springer-Verlag, 2003, pp.176-194.

[11] N. Courtois and W. Meier, *"Algebraic Attacks on Stream Ciphers with Linear Feedback," Proceedings of Eurocrypt 2003*, LNCS 2656, Springer-Verlag, 2003, pp.345-359.

[12] J.-C. Faugère, "A New Efficient Algorithm for computing Gröbner bases (F4)," *Journal of Pure and Applied Algebra*, Vol.139, 1999, pp.61-88.

[13] J.-C. Faugère, *"A New Efficient Algorithm for computing Gröbner bases without Reduction to Zero (F5)," Proceedings of ISSAC 2002*, pp.75-83.

[14] J.-C. Faugère, and A. Joux, *"Algebraic Cryptanalysis of Hidden Field Equation (HFE) Cryptosystems using Gröbner bases," Proceedings of Crypto 2003*, LNCS 2729, Springer-Verlag, 2003, pp.44-60.

[15] M. Feldhofer, S. Dominikus, and J. Wolkerstorfer, *"Strong Authentication for RFID Systems Using AES Algorithm," Proceedings of CHES 2004*, LNCS 3156, Springer-Verlag, 2004, pp.357-370.

[16] M. Feldhofer and C. Rechberger, *"A Case against Currently Used Hash Functions in RFID Protocols," Proceedings of RFIDSec 2006*.

[17] C. Hung-Yu and H. Chen-Wei, "Security of ultra-lightweight RFID authentication protocols and its improvements," *ACM SIGOPS Operating Systems Review*, Vol.41, No.4, 2007, pp.83-86.

[18] A. Juels, "RFID Security and Privacy: A Research Survey," *IEEE Journal on Selected Areas in Communications*, Vol.24, No.2, 2006, pp.381-394.

[19] A. Juels, R. Rivest and M. Szydlo, *"The Blocker tag: Selective Blocking of RFID Tags for Consumer Privacy," Proceedings of CCS 2003*, ACM Press, 2003, pp.103-111.

[20] A. Juels and S. A. Weis, *"Authenticating Pervasive Devices with Human Protocols," Proceedings of Crypto'05*, LNCS 3621, Springer-Verlag, 2005, pp.293-308.

[21] Y.-C. Lee, Y.-C. Hsieh, P.-S. You, T.-C. Chen, *"A New Ultralightweight RFID Protocol with Mutual Authentication," Proceedings of WASE 2009*, Vol.2 of ICIE, 2009, pp.58-61.

[22] N.-W. Lo, H.-S. Shie, K.-H. Yeh, *"A Design of RFID Mutual Authentication Protocol Using Lightweight Bitwise Operations," Proceedings of JWIS 2008*.

[23] P. Peris-Lopez, J.C. Hernandez-Castro, J.M. Estevez-Tapiador, and A. Ribagorda, *"LMAP: A Real Lightweight Mutual Authentication Protocol for Low-cost RFID tags" Proceedings of UIC 2006*, LNCS 4159, Springer-Verlag, 2006, pp.912-923.

[24] P. Peris-Lopez, J.C. Hernandez-Castro, J.M. Estevez-Tapiador, and A. Ribagorda, *"M^{2}AP: A Minimalist Mutual-Authentication Protocol for Low-cost RFID tags" Proceedings of UIC 2006*, LNCS 4159, Springer-Verlag, 2006, pp.912-923.

[25] P. Peris-Lopez, J.C. Hernandez-Castro, J.M. Estevez-Tapiador, and A. Ribagorda, *"EMAP: An Efficient Mutual Authentication Protocol for Low-cost RFID tags," Proceedings of IS 2006*, LNCS 4277, Springer-Verlag, 2006, pp.352-361.

[26] R. C.-W. Phan, "Cryptanalysis of a New Ultralightweight RFID Authentication Protocol-SASI," *IEEE Transactions on Dependable and Secure Computing*, Vol.6, No.4, 2009, pp.316-320,.

[27] *Sage distribution of mathematical software*, http://www.sagemath.org

[28] S. A. Weis, Security and Privacy in Radio-Frequency Identification Devices [dissertation]. Massachusetts: Massachusetts Institute of Technology (MIT); 2003.

Daewan Han

He received his B.S., M.S., and Ph.D. degrees in Mathematics from Seoul National University in 1995, 1997, and 2007, respectively. He has been a senior researcher at the Attached Institute of ETRI since 2001. His research interests include Symmetric and Asymmetric Encryption, the Cryptanalysis of Cryptographic Algorithms and Protocols.