Algebraic Cryptanalysis and RFID Authentication

C. Cid[1], L. Ferreira[2], G. Procter[1], and M.J.B. Robshaw[3]

[1] Information Security Group, Royal Holloway University of London
Egham, TW20 OEX, UK
[2] Applied Cryptography Group, Orange Labs
38–40 rue de General Leclerc, 92794 Issy les Moulineaux, France
[3] Impinj, 701 N. 34th Street, Suite 300,
Seattle, WA 98103, USA

Abstract The standardization group ISO/IEC SC31/WG7 is working on a set of *cryptographic suites* to provide security to wireless devices including UHF RFID tags. These cryptographic suites are presented as independent parts to a single standard ISO/IEC 29167. Within this multi-part standard 29167-15 is based around very simple operations and intended to provide tag, interrogator, and mutual authentication. Here we show that these proposals can be fully compromised using algebraic cryptanalytic techniques; the entire key can be recovered after eavesdropping on just four authentications.

1 Introduction

It is perhaps a sign of commercial maturity that standardization on cryptography for low-cost UHF RFID tags has begun. With such standards to hand tag, interrogator, and mutual authentication, along with secure tag-interrogator communications, can all be considered and securely implemented in the future. Increasingly referred to as RAIN RFID after the foundation of the <u>RA</u>dio <u>I</u>dentificatio<u>N</u> (RAIN) *Industry Alliance* [20], the technology will be widely deployed and is likely to become an integral part of the *Internet of Things (IoT)*. However, such tags pose a major challenge when deploying cryptography since they are very limited in terms of the space available in silicon and the power available for on-tag computation. By comparison HF RFID tags that we find in public transport ticketing and NFC applications are positively luxurious. It is within this context that ISO/IEC 29167-15 has been proposed; to provide security services to resource constrained devices.

One goal of this paper is to highlight the poor—in fact non-existent—security provided by ISO/IEC 29167-15. While the scheme can be compromised in many ways, this paper uses algebraic cryptanalysis to provide an elegant and efficient attack. But an arguably more important goal is to provide a warning of how even technically poor proposals can advance far into the standardization process. Given that there already exist many sound (and standardized) cryptographic designs for HF and UHF RFID tags, we hope our analysis will deter

standardization bodies and product developers from adopting schemes that have had little technical scrutiny.

1.1 The Standardization Landscape for UHF RFID

It may not be immediately apparent why yet another group within ISO/IEC is standardizing cryptographic mechanisms. To see why this important work is underway we need to understand the role of other standards in the field.

The commands that can be sent to a (standardized) UHF RFID tag are defined in two documents that have been published[1] by EPCglobal, part of GS1. The dominant standard covering all current large-scale deployments is known as Gen2v1 [8] and the final update to this standard was published in 2008. In 2013, however, the Gen2v2 standard was published [9] and as the version number implies, this extends the functionality of Gen2v1. The most significant and far-reaching additions are optional over-the-air commands that allow the deployment of security functionality.

Gen2v2 defines the over-the-air commands for UHF RFID but this is all it does. For instance, a command AUTHENTICATE is defined and this can be used to develop a solution for tag, interrogator, or mutual authentication. However all the security commands in Gen2v2, including AUTHENTICATE, have been deliberately designed to be both flexible and crypto-agnostic; they are completely independent of any specific cryptographic technology. By way of illustration, the format of the AUTHENTICATE command is given below, with the field descriptions, lengths, and possible values given by the three rows of the table. The `handle` and `CRC-16` are part of the communication protocol while `SenRep` and `IncRepLen` are application options. The most important fields for our purposes are marked $\star$ and their values are not defined by Gen2v2. The `CSI` field identifies the cryptographic algorithm/protocol while the `Length`/`Message` fields identify the cryptographic payload being carried by the command.

		command	RFU	SenRep	IncRepLen	CSI	Length	Message	RN	CRC
length		8	2	1	1	8	12	*variable*	16	16
value		d5$_x$	00$_b$	0$_b$/1$_b$	0$_b$/1$_b$	$\star$	$\star$	$\star$	handle	CRC-16

For the cryptographic technology itself we would likely turn to the usual sources. NIST standardizes cryptographic technologies such as the *Advanced Encryption Standard (AES)* [16]. Other cryptographic technologies have been standardized within ISO/IEC SC27 and some, such as PRESENT [2,14] and CRYPTOGPS [10,15], are explicitly targeted at constrained environments.

However we can see there is an implementation gap between the over-the-air commands and the cryptographic primitives. For example, the AUTHENTICATE command says nothing about how to achieve tag authentication using, say, a *challenge-response* authentication protocol. It doesn't even say what algorithms might be supported on the tag or interrogator. Similarly, the AES standard

---

[1] The Gen2v1 specifications are also standardized within ISO/IEC 18000-63 with Gen2v2 standardization underway as a revision.

(FIPS-197 [16]) doesn't tell us how to use the AES block cipher to perform tag authentication; instead FIPS-197 tells us how a 128-bit output is derived from a 128-bit input and a key. It is the goal of the work in ISO/IEC SC31/WG7, therefore, to provide a mapping between the cryptographic primitive and the generic over-the-air command; that is, to fill in the information marked $\star$ in the command above. This mapping is referred to as a *cryptographic suite* and the ISO/IEC 29167 standard consists of several parts, each describing a cryptographic suite and a solution. If one wishes to perform tag authentication using AES-128 then ISO/IEC 29167-10 is the cryptographic suite of interest. For tag authentication with PRESENT-80, Grain-128a or CRYPTOGPS, parts 29167-11, 29167-13, and 29167-17 are, respectively, the ones to use.

Many cryptographic suites in ISO/IEC 29167 are built on trusted or standardized primitives. Some, however, are built around new and immature proposals. Despite very negative comments during the development of 29167-15, the ISO/IEC voting structure is such that even a technically poor proposal can advance far through the standardization process. The current status of 29167-15 is unclear, though it may be moved to a *Technical Specification*. Technical specifications are sometimes used when there are irreconcilable differences of opinion and they provide the opportunity for public comment. After three years the work in the technical specification is then either abandoned or re-introduced to the standards process. This paper can be viewed, therefore, as input to this process and provides compelling support for the view that future work on this standard should be resisted.

2 Early Versions

One reason for the longevity of ISO/IEC 29167-15 is that patches have been applied at several stages during the process. All variants propose mechanisms for tag authentication, interrogator authentication, and mutual authentication. All variants are simple and built around the supposed difficulty of analyzing the combination of bitwise exclusive-or and integer addition, though first proposals were even simpler; see Table 1.

There are many problems with the proposal in Table 1 but the most pressing is that there is no security. The sole use of a single operator (in this case bitwise exclusive-or) gives a differential attack. By eavesdropping an attacker recovers RI, RT, A, and B. The adversary can then make a fake tag that fools a legitimate reader *without* knowing the secret key K. The attack is outlined in Table 2 where variables in a subsequent run of the protocol are denoted *. To confirm that the fake tag is always accepted as genuine we observe that

$$
\begin{aligned}
\text{REVERSE}(\texttt{T}^*) &= \text{REVERSE}(\texttt{SK}^* \oplus \texttt{B} \oplus \text{REVERSE}(\texttt{X}) \oplus \Delta_{SK}) \\
&= \text{REVERSE}(\texttt{SK} \oplus \texttt{B}) \oplus \texttt{X} = \text{REVERSE}(\texttt{SK} \oplus (\texttt{SK} \oplus \texttt{RCI})) \oplus \texttt{X} \\
&= \texttt{CI} \oplus \texttt{X} = \texttt{CI} \oplus \Delta_A \oplus \Delta_{SK} \\
&= (\texttt{SK} \oplus \texttt{A}) \oplus (\texttt{A} \oplus \texttt{A}^*) \oplus (\texttt{SK} \oplus \texttt{SK}^*) = \texttt{A}^* \oplus \texttt{SK}^* = \texttt{CI}^*
\end{aligned}
$$

A tag can be cloned after eavesdropping one legitimate authentication.

```
┌────────────────────────────────────────────────────────────────┐
│ Interrogator (secret key K)              Tag (secret key K)     │
├────────────────────────────────────────────────────────────────┤
│                              RI                                 │
│   Choose random RI       ─────────→                            │
│                              RT                                 │
│                          ←─────────   Choose RT               │
│   SK = K ⊕ RI ⊕ RT                    SK = K ⊕ RI ⊕ RT         │
│   Choose CI                                                     │
│                              A                                  │
│   A = SK ⊕ CI            ─────────→                            │
│                                       CI = SK ⊕ A             │
│                                       RCI = REVERSE(CI)        │
│                              B                                  │
│                          ←─────────   B = SK ⊕ RCI           │
│   T = SK ⊕ B                                                   │
│   REVERSE(T) =? CI                                             │
└────────────────────────────────────────────────────────────────┘
```

Table 1. The first version of the tag authentication scheme where all variables are 64 bits long.

```
┌────────────────────────────────────────────────────────────────┐
│ Interrogator (secret key K)                  Fake Tag          │
├────────────────────────────────────────────────────────────────┤
│                                RI*                              │
│   Choose random RI*        ─────────→                          │
│                                RT*                              │
│                            ←─────────   Choose RT*            │
│   SK* = K ⊕ RI* ⊕ RT*                   SK* is unknown        │
│                                         Δ_I = RI ⊕ RI*         │
│                                         Δ_T = RT ⊕ RT*         │
│                                         Save Δ_SK = Δ_I ⊕ Δ_T  │
│   Choose CI*                                                    │
│                                A*                               │
│   A* = SK* ⊕ CI*           ─────────→                         │
│                                         A ⊕ A* = Δ_A          │
│                                         X = Δ_A ⊕ Δ_SK         │
│                                B*                               │
│                            ←─────────   B* = B ⊕ REVERSE(X) ⊕ Δ_SK │
│   T* = SK* ⊕ B*                                                │
│   REVERSE(T*) =? CI*                                           │
└────────────────────────────────────────────────────────────────┘
```

Table 2. Fooling a legitimate reader during tag authentication. The attacker has eavesdropped on one run of the protocol in Table 1. The (changing) parameters in this second run are indicated using *.

Interrogator (secret key K)	**Tag** (secret key K)

$$\text{Choose random } R$$

$$S = R \oplus K \quad \xrightarrow{\quad S \quad} \quad R = S \oplus K$$

$$A = R + 0x55 \cdots 55$$

$$B = T \oplus K \quad \xleftarrow{\quad T \quad} \quad T = A \oplus K$$

$$B \stackrel{?}{=} R + 0x55 \cdots 55$$

Table 3. A second tag authentication scheme. The shared secret key K and all intermediate values are 64 bits long, $\oplus$ denotes bitwise exclusive-or, and $+$ denotes integer addition modulo 2^{64}.

After this inauspicious start a second proposal is illustrated in Table 3. Interrogator authentication is provided by reversing the roles of the two participants while mutual authentication is derived by interleaving two sessions that establish tag and interrogator authentication.

The weaknesses are immediately apparent and, as before, there are too many problems to list. It is sufficient to show that we can recover the key in a passive attack with high reliability. Indeed, suppose an attacker intercepts S and T from a legitimate authentication session. We then have that

$$S \oplus T = (R + 0x55 \cdots 55) \oplus R.$$

The least significant bit of $S \oplus T$ is always set to 1 and further analysis of $S \oplus T$ is easy to make. For instance, the 2^{32} values of R for which $R \wedge 0x55 \cdots 55 = 0x00 \cdots 00$ will give $S \oplus T = 0x55 \cdots 55$ and other observations based on $S \oplus T$ can be used to recover R and, via S, the shared secret key K.

For an alternative approach we simplify the notation by setting $X = S \oplus T$ and $C = 0x55 \cdots 55$. This means that $X = R \oplus (R + C)$ and considering this equation bit-by-bit gives, for bit position j with $j \geq 0$,

$$X_j = R_j \oplus ((R_j + C_j + c_{j-1}) \bmod 2) = C_j \oplus c_{j-1}$$

where c_{j-1} denotes the carry given at bit $j-1$ generated within the integer addition $R + C$. Setting $c_{-1} = 0$, the carry bit c_j for $j \geq 0$ is computed as:

$$c_j = \text{MAJ}(R_j, \ C_j, \ c_{j-1}) = (R_j \wedge C_j) \oplus (R_j \wedge c_{j-1}) \oplus (C_j \wedge c_{j-1})$$
$$= (R_j \wedge (C_j \oplus c_{j-1})) \oplus (C_j \wedge c_{j-1})$$

where MAJ denotes the *majority function*. Hence, for $j \geq 0$,

$$X_{j+1} = C_{j+1} \oplus (R_j \wedge (C_j \oplus c_{j-1})) \oplus (C_j \wedge c_{j-1})$$
$$= C_{j+1} \oplus (R_j \wedge X_j) \oplus (C_j \wedge (X_j \oplus C_j))$$
$$= C_{j+1} \oplus (R_j \wedge X_j) \oplus (C_j \wedge (X_j \oplus 1))$$

This means that, for $j \geq 0$, we have $\mathtt{R}_j \wedge \mathtt{X}_j = \mathtt{X}_{j+1} \oplus \mathtt{C}_{j+1} \oplus (\mathtt{C}_j \wedge (\mathtt{X}_j \oplus 1))$, which we write, setting $\mathtt{V}_j = \mathtt{R}_j \wedge \mathtt{X}_j$, as

$$\mathtt{V}_j = \mathtt{X}_{j+1} \oplus \mathtt{C}_{j+1} \oplus (\mathtt{C}_j \wedge (\mathtt{X}_j \oplus 1)) \text{ for } j \geq 0.$$

Looking at the expression for $\mathtt{V}_j$ we see that it consists entirely of arguments from $\mathtt{X}$, which is available to an eavesdropper, and $\mathtt{C}$ which is fixed and known. Thus if $\mathtt{X}_j = 1$ for bit j, which we expect half the time, then we can compute $\mathtt{R}_j$ directly and the corresponding bit of the shared secret $\mathtt{K}$ is given by

$$\mathtt{K}_j = \mathtt{R}_j \oplus \mathtt{X}_j.$$

Each bit $\mathtt{K}_j$, for $0 \leq j \leq 62$ can be recovered and we expect to be able to recover all but the most significant bit of $\mathtt{K}_j$ with two intercepted authentications. The work-effort is negligible. Note that this gives us two possible values for the full 64-bit shared secret $\mathtt{K}$. However these two keys are *equivalent*, that is they behave identically in the authentication protocol, and so they can both be used to impersonate a tag.

3 More Advanced Versions

Those not involved in ISO/IEC standardization may be somewhat mystified by what is happening with ISO/IEC 29167-15. Early versions were clearly weak and offered little promise. Yet voting was such that the scheme moved forward towards standardization anyway. Once we arrive at a *committee draft (CD)* the document should, in theory, be technically mature since each subsequent stage of the process, namely *draft international standard (DIS)* and *final draft international standard (FDIS)* provide little opportunity for technical modification before publication. Yet as we will show in this section, even advanced versions of these schemes were far from being technically mature and were, in fact, completely insecure.

3.1 Applying Algebraic Cryptanalysis

To repair earlier weaknesses a patched version was briefly proposed; see Table 4.

Conventional observations. Again, we can immediately see that the least significant bit of $\mathtt{S} \oplus \mathtt{T}$ is always set to 0. It is easy to find other faults and by setting $\mathtt{C} = \mathtt{0x55} \cdots \mathtt{55}$ we have $\mathtt{S} = (\mathtt{RI} + \mathtt{C}) \oplus \mathtt{K}$ and $\mathtt{T} = (\mathtt{K} + \mathtt{C}) \oplus \mathtt{RI}$ so

$$\mathtt{S} \oplus \mathtt{K} = ((\mathtt{K} + \mathtt{C}) \oplus \mathtt{T}) + \mathtt{C}. \tag{1}$$

We can use Equation 1 as a distinguisher to check if a possible value for the key $\mathtt{K}$ is a correct candidate. This can be done in several ways, but we illustrate a byte-by-byte approach, first considering the least significant byte of $\mathtt{K}$. Suppose k, a, b are the least significant bytes of $\mathtt{K}$, $\mathtt{S}$, and $\mathtt{T}$ respectively. Then any x

Interrogator (secret key `K`)	**Tag** (secret key `K`)
Choose `RI`	
$S = (RI + C) \oplus K \xrightarrow{\ S\ } RI = (S \oplus K) - C$	
$T \oplus RI \stackrel{?}{=} K + C \xleftarrow{\ T\ } T = (K + C) \oplus RI$	

Table 4. Another patched version of the tag authentication scheme. All variables are 64 bits long and $C = \texttt{0x55} \cdots \texttt{55}$.

satisfying $a \oplus x = ((x + \texttt{0x55}) \oplus b) + \texttt{0x55}$ is a good candidate for k. After eight to sixteen runs, one value should be predicted with close to 100% reliability and the least significant byte of the key is recovered. In *parallel* we can process other bytes of `K` in the same manner. There is a slight complication due to the possibility of a carry from one byte to another in the integer addition; at the same time the most significant bit of each byte might also need some attention. However, closer analysis when using particular values `S` and `T` can be used to avoid significant carry propagation. This allows us to filter incorrect values and the few key candidates that remain can be tested against the tag to find the right one. Passively eavesdropping on eight to sixteen authentication runs appears to be sufficient to recover each byte of the key `K` with good reliability. The work effort is negligible since all bytes can be treated in parallel.

Algebraic cryptanalysis. Algebraic attacks are powerful techniques that have been successfully applied against several stream ciphers (*e.g.* [5,6]) and considered against block ciphers and other cryptographic primitives [3,4]. In algebraic attacks, one writes the entire cryptographic operation as a large set of multivariate polynomial equations and then uses equation-solving techniques to recover the value of some of the unknown variables (*e.g.* the encryption key). The scheme in ISO/IEC 29167-15 uses a set of very simple operations and algebraic cryptanalysis proves to be very effective. We describe our attack below.

Let n be the size of the set of all variables in the protocol; in our case we have $n = 64$. We will assume that an attacker can eavesdrop on m runs of a uni-directional (tag or interrogator) authentication protocol. Since the mutual authentication protocol consists of two interleaved runs of a uni-directional protocol, observing m runs of the mutual authentication protocol will give identical results to $2m$ runs of the uni-directional protocol.

Without loss of generality we have implemented the attack on the protocol for uni-directional (tag or interrogator) authentication. Denote the value of `S` on the tth run of the protocol by S^t, similarly for `T` and `RI`. We use variables p_i, s_{t_i}, r_{t_i}, so_{t_i}, a_{t_i}, and b_{t_i} for $0 \leq i \leq 63$ and $0 \leq t < m$ assuming that all strings are written using big-endian convention with the most significant bit on the left, *i.e* $K = p_{n-1} \dots p_1 p_0$.

We represent the computation of S^t using the equations

$$s_{t\_i} = p_i + r_{t\_i} + a_{t\_i} + c_i$$

for $0 \leq i < n$, $0 \leq t < m$ and where a_{t_i} is the carry bit during the modular addition of RI^t and C. This gives, for $1 \leq i \leq n$,

$$a_{t\_0} = 0$$
$$a_{t\_i} = \mathrm{MAJ}\big(r_{t\_(i-1)}, c_{i-1}, a_{t\_(i-1)}\big)$$
$$= r_{t\_(i-1)} * c_{i-1} + r_{t\_(i-1)} * a_{t\_(i-1)} + a_{t\_(i-1)} * c_{i-1}$$

Similarly the computation of T^t can be represented, for $1 \leq i < n$ and $0 \leq t < m$, as:

$$b_{t\_0} = 0$$
$$so_{t\_i} = p_i + r_{t\_i} + b_{t\_i} + c_i$$
$$b_{t\_i} = p_{i-1} * c_{i-1} + p_{i-1} * b_{t\_(i-1)} + b_{t\_(i-1)} * c_{i-1}$$

where the b variables denote the carry bits. Finally, we define equations to represent the observed values of S^t and T^t and the defined value of C; $s_{t\_i} = S^t_i$, $so_{t\_i} = T^t_i$, and $c_i = C_i$. Our system is presented as polynomials over the finite field $\mathbb{F}_2$, i.e. all variables and coefficients take values in $\{0, 1\}$. We may therefore include the equations of the form $x^2 = x$ for every variable x.

The complete set of equations can be summarized as follows:

$$
\begin{cases}
s_{t\_i} = p_i + r_{t\_i} + a_{t\_i} + c_i & 0 \leq i < n & 0 \leq t < m \\
a_{t\_(i+1)} = r_{t\_i} * c_i + r_{t\_i} * a_{t\_i} + a_{t\_i} * c_i & 0 \leq i < n-1 & 0 \leq t < m \\
a_{t\_0} = 0 & 0 \leq t < m \\
so_{t\_i} = p_i + r_{t\_i} + b_{t\_i} + c_i & 0 \leq i < n & 0 \leq t < m \\
b_{t\_(i+1)} = p_i * c_i + p_i * b_{t\_i} + b_{t\_i} * c_i & 0 \leq i < n-1 & 0 \leq t < m \\
b_{t\_0} = 0 & 0 \leq t < m \\
s_{t\_i} = S^t_i & 0 \leq i < n & 0 \leq t < m \\
so_{t\_i} = T^t_i & 0 \leq i < n & 0 \leq t < m \\
c_i = C_i & 0 \leq i < n \\
x^2 = x & \text{for all } x
\end{cases}
$$

This system of polynomial equations includes $5nm + 2n$ variables and $11nm + 3n$ equations of degree at most two. Of course this system can be greatly simplified, by substituting the variables that have a fixed value (e.g. c_i, a_{t_0}, b_{t_0}), as well as the ones observed in the protocol runs (s_{t_i} and so_{t_i}). This reduces the number of variables to $(3n - 2)m + n$, and the number of equations to $(7n - 4)m + n$. For the parameter values of relevance to ISO/IEC 29167-15, the entire system will consist therefore of $444m + 64$ equations in $190m + 64$ variables, which can be constructed for the very small values of m that are required to recover the key. We use Gröbner bases algorithms to solve this system [4].
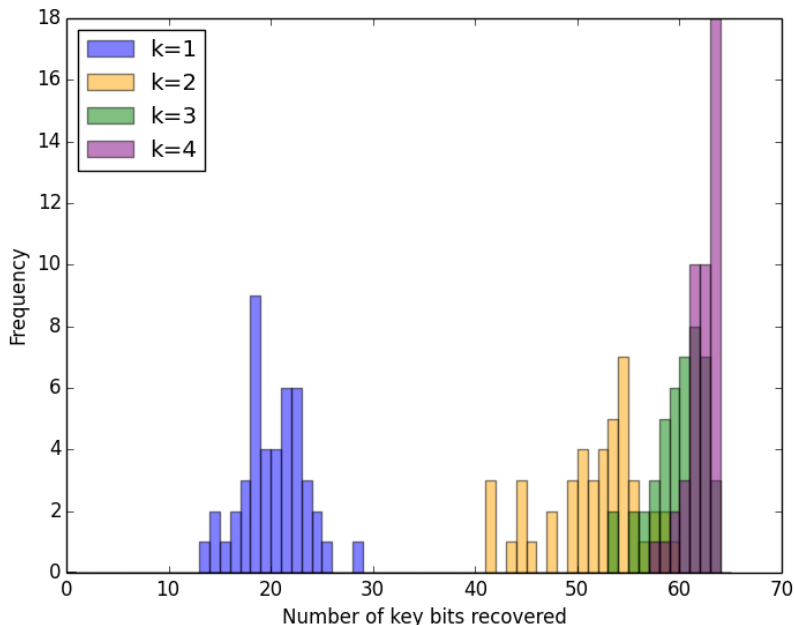
Figure 1. Results from experiments on the scheme of Section 3.1. The number of protocol runs is given by k in this figure and each set of experiments was repeated 45 independent times.

Results. The average number of key bits recovered and the average time required to solve the system of equations are given in the following table.

Number of protocol runs	1	2	3	4
Average number of key bits recovered	19.7	51.1	59.4	61.7
Average run time (s)	8	83	113	255

The attack was implemented on *SageMathCloud* [21] and timed using Python's `timeit` function; any set-up time is assumed to be pre-computed or amortized over many protocol runs. Figure 1 provides a visualization of the number of key bits recovered after observing m runs of the uni-directional authentication protocol and performing 45 sets of experiments. The values $(\mathtt{K}, \mathtt{RI}^1)$ were chosen randomly for each trial when $m = 1$. As the number of protocol runs was increased $\mathtt{K}$ remains unchanged but fresh random choices were used for $\mathtt{RI}^2$, $\mathtt{RI}^3$, and $\mathtt{RI}^4$ as would be expected in a real-life implementation.

Our experiments suggest that after witnessing four uni-directional runs of the protocol – or just two mutual authentication runs – the attacker would be able to recover 62 out of 64 bits of the secret key in around 84% of the time. While the entirety of the key can often be recovered, we conjecture that the

"missing" bits that occur from time-to-time are neutral bits; the values of these bits cannot be determined by that particular instance of the equation system. This is a feature to many cryptanalytic techniques and is often exhibited in the most significant bits of operations such as integer addition. While it might be an interesting exercise to provide an exact explanation of this phenomenon, it is not relevant to the essential message of our cryptanalysis.

3.2   The Most Advanced Version

The fourth iteration to be described in this paper is the mutual authentication protocol outlined in Table 5. As in previous versions tag and interrogator authentication are derived from the relevant halves of the full authentication protocol. Several changes have been made to this latest proposal to complicate the task of the cryptanalyst. The use of the (unknown) bitwise rotation seems to prevent the attacker from aligning bits in the challenge and response in a trivial way. For example a single bit change in the challenge will change the Hamming Weight of the candidate `RI` derived by the tag; this would result in different rotation amounts being used for the computation of the final response. Despite these complications it is straightforward to compromise the scheme. Our initial analysis suggested that even with the rotation operation, the scheme can be compromised using conventional cryptanalysis after intercepting around 32 uni-directional authentication runs. However, as demonstrated in the previous section, it is more elegant and efficient to use algebraic cryptanalysis against the scheme. This technique allows us again to recover the shared secret key `K` after eavesdropping on as few as four authentication runs.

Algebraic cryptanalysis. As before, we need to set up a system of multivariate polynomial equations. The system used to describe this latest scheme is similar to that of Section 3.1. However it is helpful to introduce some additional variables to take account of the rotation: $\mathtt{m_{t\_i}}$ corresponds to $\mathtt{K}'_i$ in the tth protocol run and $\mathtt{n_{t\_i}}$ corresponds to $\mathtt{RI}'_i$ in the tth protocol run. In truth these variables have been introduced to improve the exposition of the attack. It would be straightforward to work without them if there were significant advantage in doing so.

In this variant of the scheme we need to take account of the unknown rotation amount. The simplest way to do this is to guess the rotation amount, and to solve each set of equations that arise for each guess. To include this within the equation system we introduce an array rot_guess where rot_guess[t] is a guess for the Hamming weight of $\mathtt{RI}^t$.

The complete set of equations is summarized overleaf and the system includes $7nm + 2n$ variables and $15nm + 3n$ equations of degree at most two. Again, this system can be greatly simplified by substituting fixed/known value variables, as well as redundant ones, to a system with $(3n-2)m+n$ variables, and $(7n-4)m+n$ equations. For the parameter values of relevance to ISO/IEC 29167-15, the entire system consists of $444m + 64$ equations in $190m + 64$ variables, which can be constructed for the small values of m that are required to recover the key.

Interrogator (secret key K)	**Tag** (secret key K)

Choose RI

$S = (RI + C) \oplus K \quad \xrightarrow{\quad S \quad} \quad RI = (S \oplus K) - C$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad w_i = \text{HW}(RI)$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad K' = K \lll w_i$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad RI' = RI \lll w_i$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad T = (K' + C) \oplus RI'$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ Choose RT

$w_i = \text{HW}(RI) \quad \xleftarrow{\quad T, U \quad} \quad U = (RT + C) \oplus K$

$K' = K \lll w_i$

$RI' = RI \lll w_i$

$(T \oplus RI') \overset{?}{=} (K' + C)$

tag authenticated

$RT = (U \oplus K) - C$

$w_t = \text{HW}(RT)$

$K'' = K \lll w_t$

$RT'' = RT \lll w_t$

$V = (K'' + C) \oplus RT'' \quad \xrightarrow{\quad V \quad} \quad w_t = \text{HW}(RT)$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad K'' = K \lll w_t$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad RT'' = RT \lll w_t$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad (V \oplus RT'') \overset{?}{=} (K'' + C)$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ interrogator authenticated

Table 5. The last version of the mutual authentication scheme. The *Hamming weight* of A is denoted HW(A) while $A \lll w$ (*resp.* $A \ggg w$) denotes a left (*resp.* right) bitwise rotation of A by w bits. The constant C takes the value $0x55 \cdots 55$.

$$
\begin{cases}
s_{t\_i} = p_i + r_{t\_i} + a_{t\_i} + c_i & 0 \leq i < n & 0 \leq t < m \\
a_{t\_(i+1)} = r_{t\_i} * c_i + r_{t\_i} * a_{t\_i} + a_{t\_i} * c_i & 0 \leq i < n-1 & 0 \leq t < m \\
a_{t\_0} = 0 & 0 \leq t < m \\
m_{t\_i} = p_{(i+\text{rot\_guess}[t]\%n)} & 0 \leq i < n & 0 \leq t < m \\
n_{t\_i} = r_{t\_(i+\text{rot\_guess}[t]\%n)} & 0 \leq i < n & 0 \leq t < m \\
so_{t\_i} = m_{t\_i} + n_{t\_i} + b_{t\_i} + c_i & 0 \leq i < n & 0 \leq t < m \\
b_{t\_(i+1)} = m_{t\_i} * c_i + m_{t\_i} * b_{t\_i} + b_{t\_i} * c_i & 0 \leq i < n-1 & 0 \leq t < m \\
b_{t\_0} = 0 & 0 \leq t < m \\
s_{t\_i} = S_i^t & 0 \leq i < n & 0 \leq t < m \\
so_{t\_i} = T_i^t & 0 \leq i < n & 0 \leq t < m \\
c_i = C_i & 0 \leq i < n \\
x^2 = x & \forall x
\end{cases}
$$

Guessing the rotation amount. The equation-system depends on rot_guess, a guess for the Hamming weight of `RI`. We expect that a correct guess for the Hamming weight of `RI` will yield a system of equations that is easily solved to reveal many key bits. The values `RI` are random and so assuming that they are generated uniformly the random variable $\text{HW}(\texttt{RI})$ will be distributed according to a binomial distribution with parameters $(64, \frac{1}{2})$. It is therefore straightforward to compute the probability that a randomly chosen `RI` has a particular Hamming weight. Of particular relevance to our attack is the fact that a substantial fraction of all possible `RI` have a Hamming weight lying within a small range.

x	32	33	34	35	36
$\Pr_{\texttt{RI}}[\text{HW}(\texttt{RI}) = x]$	0.10	0.10	0.09	0.08	0.06

This means that the probability $\text{HW}(\texttt{RI})$ lies in the interval $[32 - \delta, 32 + \delta]$ is:

δ	0	1	2	3	4
$\Pr_{\texttt{RI}}[\text{HW}(\texttt{RI}) \in [32 - \delta, 32 + \delta]]$	0.10	0.29	0.46	0.62	0.74

Now assume an attacker who eavesdrops on one uni-directional authentication session. He could simply run the equation solving algorithm three times with the guesses of 31, 32, and 33 for the rotation amount. With a probability close to 30% one of these guesses would be correct. Alternatively, he could elect to further try the values 28, 29, 30, 34, 35, and 36 which would require nine runs of the equation solving algorithm. The probability that the eavesdropped session is covered by one of these nine guesses is close to 74%. This has been confirmed by experiments.

It turns out that the polynomial system solving algorithm is a good method to verify whether the guessed rotation amount is correct. Empirically it seems that selecting the wrong rotation makes the system inconsistent, *i.e.* there will be no valid solution and this is quickly detected by the Gröbner bases algorithm. Of course it cannot be ruled out that cases exist when an incorrect guess of rotation results in a system for which solutions exist (corresponding to an incorrect key). However this does not appear to be common; in over 20 experiments we never recovered false solutions for any $\delta \leq 5$. Of course, even if they did occur, false alarms could easily be filtered out using further intercepted authentication attempts or even a forgery attempt.

Results. The average number of key bits recovered and the average time required to solve the system of equations are given in the following table.

Number of protocol runs	1	2	3	4
Average number of key bits recovered	21.3	51.4	60.6	63.2
Average run time (s)	10	53	87	193

Again, the attack was implemented on *SageMathCloud* [21] and timed using Python's `timeit` function; any set-up time is assumed to be pre-computed or amortized over many protocol runs. Figure 2 provides a visualization of the
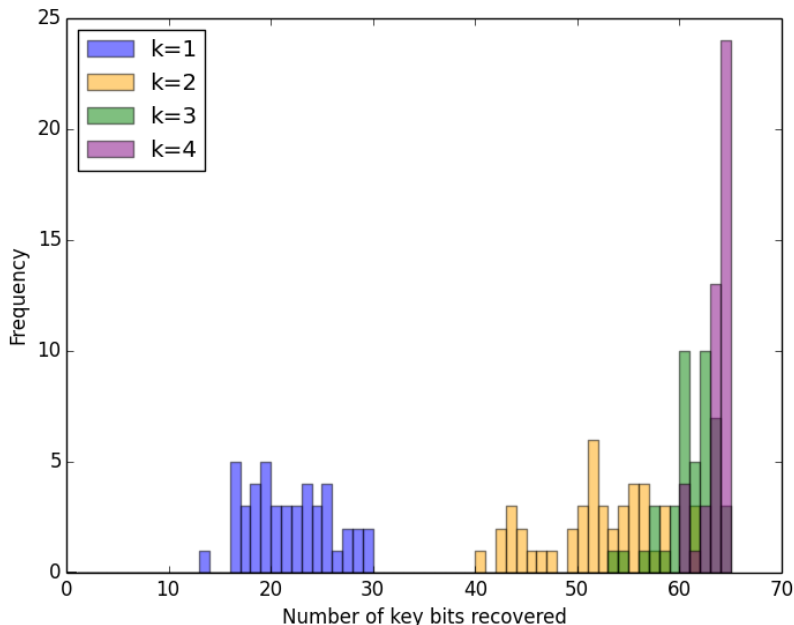
Figure 2. Results from experiments on the scheme of Section 3.2. The number of protocol runs is given by k in this figure and 45 random instances were generated.

number of key bits recovered after observing m runs of the uni-directional authentication protocol over all 45 sets of experiments. The values $(\texttt{K}, \texttt{RI}^1)$ were chosen randomly for each trial when $m = 1$. As the number of protocol runs was increased $\texttt{K}$ remains unchanged but fresh random choices were used for $\texttt{RI}^2$, $\texttt{RI}^3$, and $\texttt{RI}^4$ as would be expected in a real-life implementation. These results assume a correct guess for $\mathrm{HW}(\texttt{RI}^j)$, as discussed in the previous section.

Witnessing four uni-directional runs, or two mutual authentication runs, of the scheme in Section 3.2 and guessing the rotation amount correctly gives a probability of approximately 85% for recovering all but two bits of the key. However different attack strategies are possible.

Note that the attack using a single observed run is much faster than the attack with four observed runs, although it would recover a smaller proportion of the secret key bits. Moreover, as discussed above, when attempting to solve the resulting system of equations it is straightforward to recognize when an incorrect rotation has been guessed. Thus an efficient way to perform the full attack is as follows: we try the attack on four individual single authentications with different guesses to $\mathrm{HW}(\texttt{RI}^t)$. Once we have identified the correct rotation for four runs of the protocol we apply the four-run attack that simultaneously uses the information from all four authentications to recover almost all bits of the key

with high probability. For example, using this method, one could mount the full attack to recover all except two bits of the key with probability of approximately 69%, taking around 30 minutes to run on *SageMathCloud*.

4 Results and Discussion

Over the course of this paper we have seen several incremental versions of a tag, reader, and mutual authentication scheme. We have shown that none offer any substantive security. The results are summarized here, with most of the results in this paper being confirmed by implementations.

Version	Type of Attack	Net Result	# Authentications
Table 1	Passive	Tag cloning	1
Table 3	Passive	Key recovery	2
Table 4	Passive	Key recovery	8–16
Table 4	Passive (Algebraic)	Key recovery	4
Table 5	Passive (Algebraic)	Key recovery	4

We have concentrated on uni-directional authentication but mutual authentication consists of two inter-leaved versions of a tag and interrogator authentication. This means that all attacks will also apply to mutual authentication, but often with less effort since twice as much information is leaked during each protocol run. For all but the first variant the long term K key can be recovered.

For completeness we note that there have also been a proposal for a method to provide a secure channel, but first versions of the encryption method were wholly insecure. It might also be noted that there is some ambiguity at times as to whether protocol variables should be considered 64-bit values or 65-bit values. However this is likely to be a case of poor notation since (i) 65-bit values make little mathematical sense and (ii) they carry a significant implementation penalty. Even if we set these objections aside, the discussion quickly becomes academic since our attacks apply in a similar fashion either way.

We note that there will be many other opportunities for the attacker; we happily recognize that our application of algebraic techniques has been straightforward. A little more analysis might yield more efficient attacks, both algebraic and non-algebraic. However, in our view the point is already made and there is little to be gained by adding other incremental attacks to the mix. Those interested at looking at things more will find other work in a similar vein by other authors, *e.g.* [11], as well as other simple proposals [18,19]. Our code is available via `www.isg.rhul.ac.uk/~ccid/publications/iso-iec-29167-15.htm`.

5 Conclusion

The schemes described in this document are intrinsically weak due to the simple operations used by the tag and the interrogator. Future versions, if based on identical principles, are unlikely to provide additional security (see Appendix A).

In fairness, it should be noted that the combination of exclusive-or, integer addition, and bitwise rotation *can* be a good basis for the design of a secure primitive. So-called ARX designs are rightly popular and they featured prominently in the NIST SHA-3 initiative [17]. However these are typically multi-round algorithms while the most advanced variant of the schemes under analysis has probably just achieved a "single round" of computational complexity (if one can make the analogy).

The project editors for 29167-15 motivate their use of the simplest operations by stating that this will result in a low-area solution. However this is misguided. The bulk of the area for an implementation comes from the cryptographic state which is governed by the size of the variables. So even though ISO/IEC 29167-15 uses simple operations it doesn't lead to a dramatic implementation advantage. More importantly, there are already very good cryptographic solutions for RAIN RFID tags that provide good security; the AES is one option while PRESENT [14,13] or Grain128a [1] provide different implementation trade-offs.

One goal of the paper was to demonstrate that the simple authentication schemes being considered for standardization are flawed; this helps to emphasize the contributions cryptographers can make to a variety of ISO/IEC initiatives [7]. But a second, more important, goal was to stress that cryptography for RFID does not need to be bad cryptography. The state of the art is such that well-studied standardized schemes are available and these can be deployed in even the most demanding environments.

Acknowledgements

References

1. M. Ågren, M. Hell, T. Johansson, and W. Meier. Grain-128a: A New Version of Grain-128 with Optional Authentication, International Journal of Wireless and Mobile Computing, **5**, (1), pp. 48–59, Inderscience, 2011.
2. A. Bogdanov, L.R. Knudsen, G. Leander, C. Paar, A. Poschmann, M.J.B. Robshaw, Y. Seurin, and C. Vikkelsoe. PRESENT: An Ultra-Lightweight Block Cipher. In P. Paillier and I. Verbauwhede, editors, *Proceedings of CHES 2007*, Lecture Notes in Computer Science, vol. 4727, 450–466, Springer, 2007.
3. C. Cid, S. Murphy, and M.J.B. Robshaw. Algebraic Aspects of the Advanced Encryption Standard. Springer, 2006.
4. C. Cid and R.P. Weinmann. Block ciphers: algebraic cryptanalysis and Groebner bases. In *Groebner bases, coding, and cryptography*. pp. 307–327. Springer, 2009.
5. N. Courtois. Cryptanalysis of Sfinks. *Proceedings of ICISC 2005*, LNCS 3935, pp. 261–269. Springer, 2005.
6. N. Courtois and W. Meier. Algebraic Attacks on Stream Ciphers with Linear Feedback. *Proceedings of EUROCRYPT 2003*, LNCS 2656, pp. 345–359. Springer, 2003.

7. J. P. Degabriele, V. Fehr, M. Fischlin, T. Gagliardoni, F. Günther, G. Azzurra Marson, A. Mittelbach, and K. G. Paterson. Unpicking PLAID - A Cryptographic Analysis of an ISO-standards-track Authentication Protocol. Cryptology ePrint Archive, Report 2014/728, 2014, Available via `http://eprint.iacr.org/`.

8. EPCglobal. EPC Radio Frequency Identity Protocols, Generation 2 UHF RFID. Specification for RFID Air Interface Protocol for Communications at 860 MHz 960 MHz Version 1.2.0. Available via `www.gs1.org/gsmp/kc/epcglobal/uhfc1g2`.

9. EPCglobal. EPC Radio Frequency Identity Protocols, Generation 2 UHF RFID. Specification for RFID Air Interface Protocol for Communications at 860 MHz 960 MHz Version 2.0.0. Available via `www.gs1.org/gsmp/kc/epcglobal/uhfc1g2`.

10. M. Girault, G. Poupard and J. Stern. On the Fly Authentication and Signature Schemes Based on Groups of Unknown Order. *Journal of Cryptology*, vol. 19, number 4, 463–488, 2006.

11. D. Han. Gröbner Basis Attacks on Lightweight RFID Authentication Protocols. *Journal of Information Processing Systems*, Vol. 7, number 4, 691–706, 2011

12. J. D. Hunter, Matplotlib: A 2D graphics environment. Computing In Science & Engineering **9** (3) pp.90–95, IEEE 2007.

13. ISO/IEC 29167-11:2014 – Information technology – Automatic identification and data capture techniques – Part 11: Crypto suite PRESENT-80 security services for air interface communications.

14. ISO/IEC 29192-2:2011 – Information technology – Security techniques – Lightweight cryptography – Part 2: Block ciphers.

15. ISO/IEC 29192-4:2013 – Information technology – Security techniques – Lightweight Cryptography – Part 4: Asymmetric Techniques.

16. National Institute of Standards and Technology. FIPS 197: Advanced Encryption Standard, November 2001.

17. National Institute of Standards and Technology. SHA-3 competition. Available via `csrc.nist.gov/groups/ST/hash/sha-3/index.html`.

18. 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 H. Jin, L.T. Yang, and J. Tsai, editors, *Proceedings of UIC 2006*, Lecture Notes in Computer Science, vol. 4159, 912–923, Springer, 2006.

19. 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. In Z. Tari, editor, *On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops*, Lecture Notes in Computer Science, vol. 4277, 352–361, Springer, 2006.

20. RAIN RFID. Available via `www.rainrfid.org`.

21. W. A. Stein *et al.*, Sage Mathematics Software (Version 6.3), The Sage Development Team, 2014, `www.sagemath.org`.

Appendix A: Obvious Variants are Not Secure

It is our belief that the entire rationale for ISO/IEC 29167-15 is misguided. With the hope of discouraging further patches we pro-actively anticipate some modifications that might be made with the hope of increasing security. We show that none of the obvious variants provide a significantly increased level of security.

One variant would be to increase the size of all parameters, using a secret key of size $n = 64 + n'$ bits. The intention would be to increase the size of the equations systems since algebraic cryptanalytic schemes do not scale well. However this is particularly ineffective for the scheme of Table 4 since one can simply discard the n' high bits of the transmitted values and run precisely the same attack, recovering many of the lower 64 bits. We could then guess the value of the carry bit a_{j_64} which would allow the remaining bits to be attacked independently. We view this kind of attack as "slicing" the problem and we will return to this below.

We can apply a similar technique to remove any advantage from increased parameters in the scheme of Table 5. The rotation $r = \text{HW}(\text{RI})$ will, with high probability, lie in an interval $(\frac{n}{2} - \delta, \frac{n}{2} + \delta)$ for small δ. This means we can consider a subset of equations consisting of $S_0^j \ldots S_{s-1}^j$ and $T_r^j \ldots T_{r+s-1}^j$ (and the corresponding p_i, r_{j_i}, a_{j_i}, b_{j_i}, *etc.*) for some value of the rotation r. This attack requires us to guess the values of the rotations and, additionally, we must guess b_{j_r}. But this does not significantly affect the complexity of the attack. (In this case we do not need to guess a_{j_0}, however had we considered a subset of equations not including S_0^j we would have needed to guess another carry bit.) From two runs of the uni-directional authentication the attacker guesses the value of four bits and, for the correct rotation amount, the number of key bits recovered when using a single 16-bit or 24-bit slice in experiments is shown in Figure 3.

One strategy that may improve the attack would be to first attack a small number of bits (via the slicing attack) for several guessed values of $\text{HW}(\text{RI}^j)$. Then we could carry out a full attack against multiple protocol runs (simultaneously) once the correct rotation values have been identified. This way one could recover almost all the key bits without having to run the full attack many times.

An alternative modification to the protocol might be to change the value of the constant C. We implemented an attack assuming that two runs of the uni-directional authentication protocol were observed. We implemented a 16-bit slice attack and repeated the whole set of experiments for 256 different values of C. The 256 C values were built up as a single byte pattern repeated eight times. The number of key bits that was recovered in our attacks for different C is illustrated in Figure 4. Zero is the only value for C that leaked no key bits. However $00 \cdots 00$ is a particularly bad choice for C since this makes $S = T$ and forgeries are trivial. Every other choice of C leads to at least four key bits out of 16 being recovered. This suggests that changing the value of C is unlikely to improve the security of the proposed protocol.
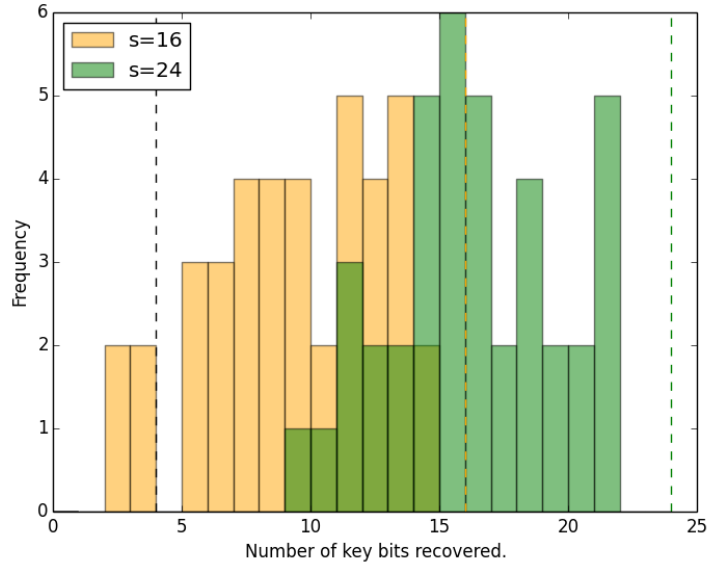
Figure 3. Implementing the "slicing" attack against the scheme of Table 5 with slices of size s. Dashed lines (green and yellow) highlight the size of the slice used, which is an upper bound on the number of key bits that can be recovered within that slice. In both cases two runs of the uni-directional authentication scheme were used and the value of the four bits were guessed (highlighted by the black, dashed line).
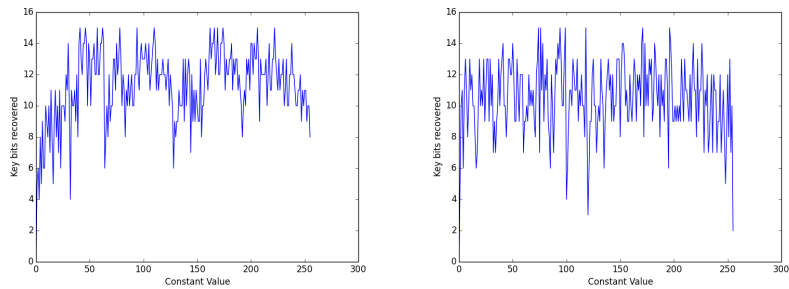

Figure 4. A set of variants of the scheme in Table 5 were tested where the constant C is varied to take 256 different values (see text). Experiments were run over two instances of the uni-directional authentication scheme and the left and right charts give the results for two randomly generated keys.