Public-Key Cryptography for RFID-Tags

L. Batina¹, J. Guajardo², T. Kerins², N. Mentens¹, P. Tuyls², and I. Verbauwhede¹

¹ Katholieke Universiteit Leuven, ESAT/COSIC, Belgium

² Philips Research Laboratories, Eindhoven, The Netherlands

Abstract

RFID-tags are a new generation of bar-codes with added functionality. An emerging application is the use of RFID-tags for anti-counterfeiting by embedding them into a product. Public-key cryptography (PKC) offers an attractive solution to the counterfeiting problem but whether a public-key cryptosystem can be implemented on an RFID tag or not remains unclear. In this paper, we investigate which PKC-based identification protocols are useful for these anti-counterfeiting applications. We also discuss the feasibility of identification protocols based on Elliptic Curve Cryptography (ECC) and show that it is feasible on RFID tags. Finally, we compare different implementation options and explore the cost that side-channel attack countermeasures would have on such implementations.

1 Introduction

In recent years, the growth of counterfeit goods has experienced a rather steep increase. This increase translates into a large source of losses for manufacturers. For example, it has been estimated that the world market for counterfeit goods was worth between 350 and 385 billion USD in 2001 and it was expected to surpass the 500 billion USD per year mark by 2004 [13]. Notice that the above numbers only point to the economical consequences of counterfeit products. However, in the particular case of the pharmaceutical industry, counterfeit products have a direct (negative) impact on the health and life of thousands of people worldwide. It is clear that new technologies need to be put in place to thwart the counterfeiting threat. RFID has been identified as one of these technologies as shown for example by legislation introduced in the US mandating use of RFID technology as anti-counterfeiting technology for at-risk pharmaceuticals for all medicines in the supply chain by the end of 2010 [17]. However, the use of RFID as an anti-counterfeiting technology is at present rather primitive.

The whole security relies on the premise that an RFID tag is harder to copy than a bar code. However, it has already been demonstrated several times that simple RFID tags can be cloned. For example, Bono et al. [6] have shown how an RFID transponder device manufactured by Texas Instruments and used in many car keys can be successfully cloned with off-the-shelf equipment and minimal RF expertise (see also [30] for another example). Thus, sound technological solutions for the counterfeiting problem need to be developed. By sound, we mean solutions based on cryptography, fundamental physical properties of materials that make them unclonable or a combination of both. Notice that the anti-counterfeiting problem can also be rephrased as an authentication problem. In other words, how can a reader tell that a certain RFID tag is really the one that it intended to talk to? In this setting, RFID-tags contain some secret reference information that is used to check their authenticity. In order to avoid counterfeiting, RFID-tags have to be unclonable. First, this implies that it should be hard to make a physical clone. Secondly, this also means that retrieving the secret reference information by attacking the protocols that are carried out between the reader and a tag (proving its authenticity) should be infeasible. Protection against physical unclonability is provided by using physical countermeasures such as Physical Unclonable Functions [28] and protection against active or passive attacks on the protocols is provided by cryptographic techniques such as digital signatures and secure identification protocols. In short, RFID-based identification is an example of an emerging technology which requires authentication as a cryptographic service. This property can be achieved by symmetric as well as asymmetric primitives.

* A preliminary version of this paper appears in the Workshop for RFID Security 2006 which had no formal proceedings.
Previous work considered only symmetric-key algorithms e.g. AES [8]. It is still not clear whether Public-Key (PK) algorithms can be implemented in constrained devices, such as RFID tags, and still comply with the area, performance, and power requirements typical of these applications. Recently, a few papers [28, 31] discussed feasibility of ECC based PKC on RFID-tags. Here, we extend that line of work and discuss implementations aspects of even stronger PK-based protocols in more detail. In particular, the protocols investigated in [28] were only secure against passive attacks. Thus, in this paper, we investigate the efficiency of protocols (Okamoto-identification protocol) that are also secure against active and concurrent attacks. It is shown that only a small price for much additional security has to be paid. In addition, we present ECC-based implementation of the above mentioned protocols and compared different implementations methods available.

The remainder of the paper is organized as follows. Section 2 provides an overview of related work. In Sect. 3 we state our assumptions and review the protocol of [28] describing only the PUF-based protocol for the off-line authentication case. The Okamoto identification protocol and its hardware implementation for off-line verification are described in Sect. 4. Finally, our results are presented in Sect. 5.

2 Related Work

Protocols for cheap authentication have been presented [16, 15]. However, they focus on the on-line situation in which the reader shares a secret with the tag being authenticated. In addition, they do not take physical cloning into account. In [28], RFID-tags that withstand general cloning attacks (including physical ones) are introduced. Based on an Integrated PUF (I-PUF) [9, 29] a PUF-Certificate-Identity Based identification scheme was introduced. This scheme allows for off-line authentication. In [28] the implementation of the Schnorr Identification scheme was investigated for this purpose. This protocol is only secure against passive attacks but it is very efficient. There have not been many attempts at hardware implementations of PKC on RFID tags or other low-power application platforms e.g. sensor nodes. Gaubatz et al. [11] showed that RSA is not a feasible solution while NtruEncrypt can be implemented in about 3000 gates. More recent work of Wolkerstorfer [31] is the first to claim possible to have low-power and compact implementation of ECC that meets the constraints imposed by the EPC standard. However, our solution is smaller as the off-line authentication in our case does not require full ECDSA signature generation to be executed on the RFID tag. This allowed for further area optimizations.

3 Off-line Authentication

We distinguish between on-line and off-line authentication. Off-line authentication is the most attractive one from a practical point of view but also the most challenging one, as costs grow much more in this case. The particular case of on-line authentication was considered in [28] and it does not make use of PK cryptography, thus we do not discuss it any further.

3.1 Assumptions

We consider RFID-tags embedded in a product or its package for detection and prevention of product counterfeiting. The tag is manufactured and embedded into the product by a legitimate authority which is assumed to be trusted. We consider an active attacker that knows the position of the tag in the product or its package, so she can remove the tag from the package to investigate it. We also assume that the attacker can (passively) eavesdrop on the channel between a reader and the tag, or can install a fake reader that communicates with the tag (active attack). Finally, we assume that the attacker can physically attack the tag; i.e. she can try to read out its memory. The goal of the attacker is to produce a fake RFID-tag containing reference information such that it can only be distinguished from a real tag with small probability . Clearly, by embedding such a fake tag into a fake product, the fake product is identified as an authentic one.

3.2 Authentication Protocol

In [28] a PUF-Certificate-Identity based Identification scheme was proposed. For the sake of completeness we describe it briefly here but refer to [28] for the details. Given the following algorithms and definitions:

- a tag with identity $I$ and a PUF,
- a standard identification scheme $SI = (K_g, P, V)$, where $K_g$ denotes the key generation algorithm, and $P, V$ denote the interactive protocols run by the prover and verifier respectively, and
- a secure signature scheme $SS = (SK_g, Sign, V_f)$, with $SK_g$ denoting the key generation algorithm, $Sign$ denoting the signing algorithm and $V_f$ the verification algorithm run by a verifier.

a PUF-Certificate-Identity based Identification scheme $(MK_g, UK_g, \hat{P}, \hat{V})$ can be constructed as follows.

During enrollment the issuer uses $SK_g$ as the master-key generation algorithm $MK_g$ for the secure signature scheme. The result is the issuer’s master-key pair $(mpk, msk)$. The algorithm $UK_g$ creates for each tag a public-secret key pair...
(pk, sk) using the algorithm \( K_g \) for the SI-scheme. The issuer runs a protocol with the tag to determine the PUF’s challenge \( c \) and helper data \( w \) such that the PUF response \( x(c) \) maps onto the secret key \( sk \). The helper data \( w \) are written into the ROM (EEPROM) memory of the tag. Finally, the issuer, using his master secret-key msk to sign, creates the following certificate that is also stored in the ROM of the tag \( \text{Cert} \leftarrow (pk, \text{Sign}(msk, pk||ID)) \).

During authentication the algorithms \( \mathcal{P} \) and \( \mathcal{V} \) are run as follows. The tag (in the role of the prover) sends the certificate \( \text{Cert} \) to the reader. If \( \text{Cert} \) is valid, the tag and the reader run the SI-protocol. If the tag passes this protocol too, the reader decides that the tag is authentic and otherwise not. Note that in order to run this last step, the tag has to challenge its PUF and use the helper data to obtain the secret key \( sk \) from the measured response \( y(c) \).

The security of the scheme depends on three factors: (i) the security of the PUF as a secure storage of the secret key, (ii) the security of the identification scheme used, and (iii) the security of the signature scheme used. It was shown in [28] that if the PUF is unclonable and a good Fuzzy Extractor is used for key extraction, the PUF provides a secure way of storing secret keys. The security of the scheme against impersonation attacks depends on the security of the identification scheme used against those attacks. Therefore, it is of crucial importance to understand which trade-off is being made between efficiency and security.

4 Okamoto’s ID Protocol Based on ECDLP

In [28], Schnorr’s identification protocol [26] is used as the SI in the Cert-IBI. Furthermore, it is shown that the elliptic curve version of Schnorr’s identification protocol can be efficiently implemented. Schnorr’s protocol is, however, only resistant against passive attacks under the discrete logarithm assumption. Another protocol that is also resistant against active and concurrent attack under the discrete logarithm assumption is Okamoto’s identification protocol [25]. We investigate therefore the efficiency of the implementation of this protocol here in detail. Notice that we are considering Okamoto’s identification protocol as it provides security against active adversaries and it is based on the hardness of the DL problem. Other protocols found in the literature include Beth’s identification protocol [5] and the XDL-IBI scheme in [3]. Beth’s protocol only requires one point multiplication but it remains an open problem to prove its security against active adversaries. The XDL-IBI scheme also requires only one point multiplication but is only secure against passive adversaries and concurrent attacks (under a modified assumption). Thus, it seems that by analyzing both Schnorr’s and Okamoto’s we cover the efficiency of all available ID protocols based on the hardness of the DL problem. Protocols based on the hardness of integer factorization also exist (see [3, 23] for a thorough classification of ID-based protocols) but their performance and scalability properties are worse, in general, than those based on ECDLP as the comparisons in [10] show.

4.1 Elliptic Curve Cryptography

The main operation in any ECC-based primitive is the scalar multiplication. The point scalar multiplication is achieved by repeated point addition and doubling. We can use the basic double-and-add algorithm in both cases. In the case of Schnorr’s identification protocol, we can also use the Montgomery ladder method [22] and benefit from the Lopez-Dahab projective coordinates [19].

4.1.1 Point Addition and Doubling

The point addition/doubling depend on the type of projective coordinate used. Table 1 summarizes the number of operations required for known projective coordinates in terms of multiplications, squarings, and additions. The number of operations is assuming general values for \( Z \). However, this value varies based on the values of the curve coefficients \( a, b \) are considered.\(^1\)

\( ^1 \)The EC is defined as \( y^2 + xy = x^3 + ax^2 + b \), with \( a, b \in \mathbb{F}_{2^n} \).

**Figure 1. Okamoto’s identification protocol.**

**Table 1. Operation Counts for EC point addition and doubling**

<table>
<thead>
<tr>
<th>Coordinate System</th>
<th>Addition</th>
<th>Doubling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jacobian projective ((X/Z)^2, Y/Z, Z) (a \neq 0)</td>
<td>14</td>
<td>5</td>
</tr>
<tr>
<td>Jacobian projective ((X/Z)^2, Y/Z, Z) (a = 0)</td>
<td>14</td>
<td>4</td>
</tr>
<tr>
<td>Lopez-Dahab ((X/Z, Y/Z)^2, Z) (b = 1)</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Lopez-Dahab ((X/Z, Y/Z)^2, Z) (b \neq 1)</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>Modified Lopez-Dahab ((X/Z, Y/Z)^2, Z) (b = 1)</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Modified Lopez-Dahab ((X/Z, Y/Z)^2, Z) (b \neq 1)</td>
<td>6</td>
<td>1</td>
</tr>
</tbody>
</table>
tively cheaply in hardware. Notice that our emphasis is on minimizing area rather than performance. Addition of two elements \(C = A + B \in \mathbb{F}_{2^n}\) is performed via an \(n\)-bitwise logical XOR operation. The simplest multiplier is based on Horner’s scheme for multiplication. In particular, to compute the product \(C = A \cdot B \in \mathbb{F}_{2^n} \cong \mathbb{F}_2[x]/f(x)\), and \(A = \sum_{i=0}^{n-1} a_i x^i, B = \sum_{j=0}^{n-1} b_j x^j, f = x^n + \sum_{i=0}^s f_i x^i, s < n\), we process one bit of \(B\) at the time as follows

\[
C = \sum_{j=0}^{n-1} \sum_{i=0}^{n-1} a_i b_j x^{i+j} \mod f = A \sum_{j=0}^{n-1} b_j x^j \mod f
\]

The multiplication process then requires \(n\) iterations. Digit serial multiplication [27] is a generalization of this in which several coefficients of \(B\) are processed in parallel. Thus, we trade off area for performance. Algorithm 1 describes how to perform Most Significant Digit multiplication according to [12] which reduces the area of the multiplier with respect to [27] and [2] by not requiring either an additional reduction circuit as in [27] or additional \(n + 2D - 1\) MUXes as in [2]. The cost is a longer critical path, which at the frequencies of operation that we are considering does not affect the overall performance. Squaring \(c = a^2 \in \mathbb{F}_{2^n}\) is a special case of multiplication. It is well known that \(a^2 = \sum_{i=0}^{n-1} a_i x^{2i}\) which can then be reduced modulo \(f\) to a field element in \(\mathbb{F}_{2^n}\).

### 4.2 ECC processor

Our Elliptic Curve Processor (ECP) for RFID has the following operational blocks: a Control Unit (CU), an Arithmetic Unit (ALU), and Memory (RAM and ROM). The ECC parameters and the constants are stored in ROM. On the other hand, RAM contains all input/output and intermediate variables and it therefore communicates with both, the ROM and the ALU. The CU controls scalar multiplication and point operations. In addition, the controller commands the ALU which performs field multiplication, addition and squaring. The bits of the scalar multiplier \(k = \sum_{i=0}^{n-1} k_i 2^i\), \(k_i \{0, 1\}, n = \lceil \log_2 k \rceil\), are evaluated from MSB to LSB.

The CU consists of a number of simple state machines and a counter and its area cost is small.

### 5 Results and Discussion

In this section, we provide results for the latency and the area complexities of both Schnorr’s and Okamoto’s protocols. As we are interested in implementations of identification protocols (e.g. Schnorr, Okamoto) the operation required is one point multiplication in the case of Schnorr’s protocol or multiple-point multiplication in the case of Okamoto’s scheme.

#### 5.1 Implementation of Okamoto’s Scheme

In [28], the feasibility of the ECC version of Schnorr’s identification protocol in an RFID system was investigated and area and latency estimates were provided. Here, we provide detailed numbers and we also investigate the feasibility of the Okamoto’s scheme as it provides security against active adversaries which Schnorr’s scheme does not.

Table 2 summarizes the number of operations that the different addition formulae imply. Here we assume that for the Jacobian coordinates on average we perform \(n\) doublings and \(n/2\) addition operations and that for the Lopez-Dahab coordinates we use the Montgomery ladder with \(n\) iterations². Notice that the modified formulae presented in [28] provide simple side-channel attack resistance if implemented without the use of a dedicated squarer. In this case, the formula of [28] are almost 3 times as slow as the standard formulae from [19] which does not provide side-channel resistance.

<table>
<thead>
<tr>
<th>Operation</th>
<th>(L)</th>
<th>(C)</th>
<th>(S)</th>
<th>Total Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\mathbb{F}_{2^n}) addition</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>(\mathbb{F}_{2^n}) squaring</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>(\mathbb{F}_{2^n}) multiplication</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>(d + 3)</td>
</tr>
</tbody>
</table>

²The actual number of doublings is \(n - 1\) for the binary method, but we chose \(n\) in the interest of space.
computing at each of $\lceil \frac{n}{w} \rceil$ steps $w$ doublings and 1 addition from a list of $2^w$ pre-calculated values of the form $iP + jQ$. As the width of the window $w$ is variable, this allows to trade-off area for speed. We chose the smallest window i.e. $w = 1$. In this way, the memory requirements are minimized as only 3 points have to be stored: $P, Q, P + Q$. The expected running time of the algorithm for $w = 1$ is $\frac{3}{w^2}$ point additions and $(n - 1)$ point doublings.

We have implemented the Schnorr scheme in VHDL and obtained area and timing values for a 0.25µm CMOS library. We have used these values to estimate the performance of Okamoto’s identification protocol using Shamir’s Trick and Jacobian coordinates. Table 3 summarizes the results. We notice that the amount of logic required to support Okamoto’s protocol is not significantly larger than that corresponding to the implementation of Schnorr’s. However, the required RAM to implement Okamoto’s identification protocol is more than twice the required RAM required for Schnorr’s. In practice this means an increase in area anywhere from 20% to 50% depending on the chosen RAM implementation (i.e. whether a RAM cell is implemented as a register requiring at least 6 equivalent gates worth of area or as dedicated embedded RAM requiring somewhere between 1.5 and 2 equivalent gates [14]). In terms of latency, Okamoto’s identification protocol is almost twice as slow as Schnorr’s over elliptic curves due to the fact that the coordinate representation introduced in [19] is only applicable to the Montgomery Ladder method of exponentiation. In addition, simultaneous double exponentiation is naturally about 25% slower than the regular binary method for exponentiation. With respect to the most compact solution, as required due to low gate-count and low-power requirements, implementing curve-based protocols with shorter bit-lengths appears to be an attractive option. For example, in the case of ECC one could use 130-bit long parameters. This solution would still maintain a suitable level of security [18], especially for low-cost RFIDs, and the gate complexity would scale-down accordingly resulting in more attractive solutions from the area and performance points of view. We conclude by noticing that the performance of the simultaneous point multiplication (as well as the binary method) can be easily improved by using Non-Adjacent Form representation for the multiplier. Such methods in the binary case would for example reduce the number of multiplications from a half on average to a third, providing significant performance improvements (see for example [21]).

5.2 A Word Regarding Power

At the present moment, we do not have an actual chip and we lack explicit power measurements for our simulations. Nevertheless, we believe that attaining the power values required for RFID applications using our design is possible. In fact, our processor architecture is very similar to the architecture presented in [31]. One particular characteristic of both designs is the usage of an Arithmetic Logic Unit (ALU) with a full-precision data path. The differences are on the details of our implementation: field size, choice of finite field arithmetic methodology (Montgomery vs. dedicated trinomial or pentanomial circuits), support for multiple fields versus support for a single field, hashing versus no hashing required in our implementation, etc. In general, our design is aimed at making our implementation as specific as possible to our particular application. This methodology leads to significant complexity reduction in the area requirements and thus also to power savings. Thus, since [31] was able to attain the power requirements of an RFID system, we are confident that our design, being smaller and simpler, will also attain the required power figures at the same or lower operating frequencies.

5.3 Generation of Randomness

Finally notice that a source of randomness is needed on the tag. In our case, this can be done by applying a random challenge to the PUF e.g. in a range out of its specification. The random challenge can be generated by the reader. For a construction of a random number generator based on a PUF, we refer to [24].

5.4 On Performance and Other Implementations

Recently, McLoone and Robshaw [20] explore the hardware cost of the GPS scheme modified for ultra-low cost devices, such as RFID tags. The authors achieve costs equivalent to those of symmetric-key encryption as described in the work of Feldhofer et al. [8], in other words in the order of 1700 equivalent gates for their largest implementation and requiring not more than 900 cycles, which at 100 KHz translates into 9 msec and under 3.4 µA of current. However, it is not possible to compare fairly the work of this paper with that of [20]. First, the implementation of [20] is based on the idea of coupons. In other words, the RFID tag stores a number of challenge-response pairs that have been

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>D=1</td>
<td>3032</td>
<td>3961</td>
<td>2596</td>
<td>0.90</td>
<td>1.00</td>
<td>3694</td>
</tr>
<tr>
<td>D=2</td>
<td>1758</td>
<td>2311</td>
<td>996</td>
<td>1.00</td>
<td>1.10</td>
<td>2608</td>
</tr>
</tbody>
</table>

Table 3. Implementation results @ 175 kHz and assuming a dedicated squarer circuit.
precomputed and only performs the last part of the 3-pass authentication protocol which in general is much cheaper to implement and requires an addition and a single modular multiplication. This implies that the RFID tag can only authenticate itself a limited number of times. On the other hand, our solution can participate in an unlimited number of authentications. Second, the hardware requirements of [20] do not include the storage cost of the coupons in their estimates, whereas our implementation does include this cost. Obviously, which solution is best will depend on the requirements of the application. If the application only requires that the RFID tag be operational for just a few authentications, then the solution described in [20] will be adequate, otherwise our flexible elliptic curve solution would be the one of choice.

6 Concluding Remarks

In this paper we discussed the feasibility of public key based secure identification protocols for RFID-tags. As an example we investigated the implementation of Okamoto’s identification protocol in detail. It was shown that it is just slightly more expensive than Schnorr’s identification protocol. Finally, we noticed that the performance of Okamoto’s protocol can be further improved using the techniques presented in [1] and recently improved in [4]. Such improvements will be considered in future work.

Acknowledgments

We would like to thank the anonymous referees for their helpful comments. This work is also partially funded by FWO G.0450.04 and K.U.Leuven (OT).

References


[33] The storage costs in [20] will not be as high as in our implementation as they only require read-only memory which is cheaper than RAM.