SCIS 2007 The 2007 Symposium on Cryptography and Information Security
Sasebo, Japan, Jan. 23-26, 2007
The Institute of Electronics, Information and Communication Engineers

Securing HB$^+$ against GRS Man-in-the-Middle Attack

Dang Nguyen Duc * Kwangjo Kim *

Abstract— In Crypto'05, Juels and Weis proposed an efficient and provably secure authentication protocol for RFID devices, namely HB$^+$. The protocol is adapted from a human authentication protocol called HB which was proposed earlier by Hopper and Blum. Although HB$^+$ is more secure than HB in order to be secure in a RFID environment, HB$^+$ still suffers from an inherent weakness of HB. That is HB+ is not shown to be provably secure against a stronger yet practical type of attack, *e.g.,* man-in-the-middle attack. This problem was quickly demonstrated by Gilbert *et al.* They presented a man-in-the-middle-attack with linear complexity which can discover a secret information shared by a RFID tag and a RFID reader. Till then, an efficient variant of HB$^+$ which is secure against active adversaries remains an open question. In this paper, our goal is to solve this open question. We propose an augmented version of HB$^+$ and show that the new protocol is secure against man-in-the-middle attacks. Comparing to HB$^+$, our improved protocol requires only one more secret and minimal additional computation at tag and reader's side. Therefore, HB* is still usable for RFID devices.

Keywords: human authentication, LPN problem, RFID security

1 Introduction

Research on lightweight cryptographic protocol has attracted significant attention in the cryptologic community. A lightweight cryptographic protocol can be informally defined as an extremely efficient one yet obtain a reasonable level of security comparing to the conventional protocols. The main motivation behind this trend is the blossom of various kinds of pervasive devices as we enter the so-called ubiquitous computing era. Pervasive devices like mobile phones, personal assistance devices (PDA), sensors, smart cards and RFID tags, *etc.* share a common characteristic that its computational ability is very limited, sometimes even extremely basic as in the case of passive RFID tags. As a result, it is inappropriate to use most of conventional security protocols, which have been designed for fully functional computers, in these devices.

RFID security is one of the hottest subjects in the cryptologic community in recent years. RFID system is a promising technology to replace with Barcode-based recognition system and provides much more powerful applications. By tagging each and every object with a unique identification which can be read by RFID readers using radio communication, people can virtually identify and keep track of everything. And this potential results in limitless applications, most notably automated supply chain management, smart home appliances, library management, *etc.* However, besides its prospective usefulness, RFID technology also brings a long security threat to personal and business sectors. The security concern is two-fold: Fake RFID tags results in impersonation and counterfeiting products; The availability of unique identification results in the disclosure of personal belongings, preferences and movements. Many protocols for RFID devices have been proposed to address the above security issues [16, 17, 18, 19, 20, 21]. Among these protocols, HB$^+$ protocol by Juels and Weis is considered to be the most interesting one because their protocol is very efficient to implement on extremely low-cost hardware and bases its security on a well-studied hard problem called *Learning Parity in the Presence of Noise* (LPN for short). The LPN problem is relatively new to the cryptologic applications but better known in the machine learning area and has been shown to be NP-hard. The origin of HB$^+$ can be traced back to the work of Hopper and Blum's Asiacrypt'01 paper [14]. Hoppper and Blum presented two provably secure human authentication protocols [14], one of which depends on the hardness of the LPN problem (and usually referred to as HB protocol). Because HB protocol can be carried out by a human, it is conceivable that HB is also suitable for computationally limited devices. Note that, in case of human authentication, a person authenticates to a machine and we can assume that the machine is trusted. However, it is different in RFID environment because RFID tags and readers communicate in an automated manner so neither tags nor readers need to be trustful. As a consequence, Juels and Weis designed HB$^+$ from HB in a way that a malicious reader has little chance of violating security of the protocol, *e.g.,* extracting secret information stored in a tag. HB$^+$ protocol is shown to be as secure as HB, this in turn means that breaking HB$^+$ is as hard as solving LPN problem.

Unfortunately, HB$^+$ is only provably secure against active adversaries (also known as secure in *detection*

* International Research Center for Information Security (IRIS), Information and Communication University (ICU), 119 Munjiro, Yuseong-gu, Daejeon, 305-732 Republic of Korea, e-mail: {nguyenduc, kkj}@icu.ac.kr

model). Resistant against more advanced attacks like man-in-the-middle attack was not achieved in [21]. This drawback was quickly shown by Gilbert *et al* (GRS attack for short) in [22]. By presenting a man-in-the-middle attack with linear complexity, they proved that tag's secret can be recovered with high probability. Until now, there is no published work to defend HB$^+$ against this attack.

Our contribution. In this paper, we present an augmented version of HB$^+$ protocol to thwart the man-in-the-middle attack like GRS attack. Our proposed protocol introduces reasonable computational and communication overhead comparing to HB$^+$ protocol. Furthermore, we prove the security of our protocol the stronger security model called *prevention model*.

Organization of this paper. In Section II, we briefly review the previous work. In Section III, we describe our proposed protocol followed by security analysis in Section IV. Finally, we conclude with the final remarks and future work in Section V.

2 The Previous Works

2.1 HB Human Authentication Protocol

The HB protocol involves the computation of binary inner product of two k-bit numbers. The operation is defined as follows: given two k-bit number $a = (a_0 a_1 ... a_{k-1})_2$ and $x = (x_0 x_1 ... x_{k-1})_2$, the binary inner product of a and x, denoted as $a \cdot x$ is computed as follows: $a \cdot x = (a_0 \wedge x_0) \oplus (a_1 \wedge x_1) \oplus \cdots \oplus (a_{k-1} \wedge x_{k-1})$.

This binary inner product operation can be carried out relatively easy by a human as well as by low-cost devices (like RFID tag). It is easy to show that binary inner product operation follows distributive law: $(a_1 \oplus a_2) \cdot b = (a_1 \cdot b) \oplus (a_2 \cdot b)$.

Goldreich and Levin proved an interesting unpredictability result of the binary inner-product operation in [3]. Let $g(a, x) = (f(a), x)$ where f is a one-way function, then it holds that $b(a, x) = a \cdot x$ is a hard-core predicate of g. In other words, given $g(a, x)$, it is hard to predict the output of $b(a, x)$ with probability significantly greater than $\frac{1}{2}$. This result serves as a basis for a construction of a secure pseudo-random generator (see **Proposition 3.12** of [28]). An alternative interpretation of the above result was also given in [28]. It states that, given only a, $b(a, x)$ appears as a random bit, otherwise, one can efficiently compute x. Those positive results about binary inner-product did not generate any further cryptographic applications until the work of Hopper and Blum appeared in 2001 [14].

In the HB protocol, the human (denoted as $\mathcal{H}$, also called the prover) and a machine (denoted as $\mathcal{C}$, also called the verifier) share a secret x of k-bit long. The protocol consists of several executions of a basic challenge-response protocol which is described in Fig. 1.

The basic protocol starts with $\mathcal{C}$ sending a k-bit random challenge a to $\mathcal{H}$. $\mathcal{H}$ computes a 1-bit response z as the binary inner product between a and x. Before sending z to $\mathcal{C}$, $\mathcal{H}$ decides whether to flip the value of z depending on the probability $\eta \in (0, \frac{1}{2})$. The probability η is fixed and z is flipped independently for every protocol round. We can say that, the noise bit v is drawn from Ber_η which is a Bernoulli distribution with an expected value η. Sending a noisy response z will prevent an eavesdropper who captures k instances of the basic protocol, *i.e.,* k different pairs (a, z), from recovering the secret x through Gaussian elimination (k bits of x are k unknowns and each pair of (a, z) constitutes an equation with k unknowns). $\mathcal{C}$ verifies $\mathcal{H}$ by counting the number of correct responses in r rounds of the basic protocol. Due to the effect of η, the genuine human should send roughly ηr false responses. Therefore, $\mathcal{C}$ accepts $\mathcal{H}$ only if it receives less than ηr (threshold value) false responses. Hopper and Blum showed that any naive adversary attempting to play random guess of the response z has success probability at most $e^{-c_0 r}$ where c_0 is a constant depending on η and greater than $\frac{2}{3}$. Unfortunately, HB protocol is not secure against active adversaries since an attacker can retransmit the same challenge a for one protocol session then he can learn a noise-free value of $a \cdot x$, *i.e.*, one valid equation with k unknowns. By applying such attack with k linear independent a, the attacker can recover x using Gaussian elimination.

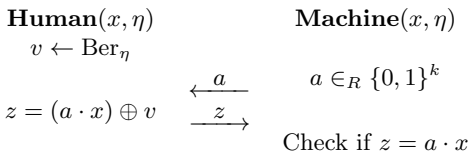

Figure 1: Basic protocol of HB protocol

2.2 Learning Parity with Noise Problem

It is straightforward that HB protocol is secure only if an eavesdropper observing messages exchanged between $\mathcal{H}$ and $\mathcal{C}$ has a negligible chance of impersonating $\mathcal{T}$. More specifically, an eavesdropper $\mathcal{A}$ obtains r pairs of (a, z) and tries to deduce a k-bit number x' such that using x' to carry out HB protocol, $\mathcal{A}$ would get accepted by $\mathcal{C}$. Let $\mathbf{M}$ be a $r \times k$ binary matrix such that each row of $\mathbf{M}$ is a k-bit challenge a sent by $\mathcal{C}$. Let us view x' as a column vector of dimension k and r responses z observed by $\mathcal{A}$ as vector $\mathbf{z}$. Then $(\mathbf{M} \cdot x') \oplus \mathbf{z} = \mathbf{v}$ where $\mathbf{v}$ is a column vector of dimension r. It is easy to see that each '1' bit in $\mathbf{v}$ corresponds to one incorrect response checked by $\mathcal{C}$. In order to be accepted by $\mathcal{C}$, the Hamming weight of $\mathbf{v}$, denoted as $\mid \mathbf{v} \mid$ has to be less than or equal to ηr. The problem of finding such x' is called *Learning Parity in the Presence of Noise* problem (LPN). However, as noted by Katz and Shin in [24], finding x' is essentially equivalent to finding x itself. We now formally define LPN problem as follows.

Definition (**LPN Problem**). Let $\mathbf{M}$ be a random $r \times k$ binary matrix. Let $\eta \in (0, \frac{1}{2})$ be a noise factor and $\mathbf{v} = (v_0, v_1, \cdots, v_{r-1})^T$ be a noise vector of r dimensions whose each member is generated independently according to noise factor η, i.e., $\mathbf{Pr}(v_i = 1) = \eta$.

Choose a random k-bit secret x and compute vector $\mathbf{z} = (\mathbf{M} \cdot x) \oplus \mathbf{v}$. Given only $\mathbf{M}$, $\mathbf{z}$ and η, compute x.

The LPN problem has been extensively studied in several research works including [8, 10, 12, 14]. Those results show that LPN problem is very likely an intractable problem. To solve LPN problem as defined above, the best known algorithm by Blum *et al.* has sub-exponential complexity of $2^{O(\frac{k}{\log k})}$. Hopper and Blum even conjectured that there is no polynomial algorithm to solve LPN problem with randomly chosen instance $(\mathbf{M}, \mathbf{z}, \eta)$. The latest hardness result of LPN problem is due to Regev [23], and Katz and Shin [24]. They showed that if LPN problem is hard, a k-bit string $(a, (a \cdot x) \oplus v)$ is indistinguishable from a true random k-bit string. In fact, Katz and Shin used this result to give more elegant security proofs of HB protocol family than ones provided by Juels and Weis. We are going to use their technique in the analysis of our proposed protocol.

2.3 HB$^+$ Authentication Protocol

In the HB$^+$ protocol, a RFID tag (denoted as $\mathcal{T}$) plays a role as a human and a RFID reader (denoted as $\mathcal{R}$) plays a role as a machine. Comparing to HB protocol, $\mathcal{T}$ and $\mathcal{R}$ share an additional k-bit secret y. To prevent a malicious reader from extracting the secrets stored in tag's memory, $\mathcal{T}$ first selects a random k-bit blinding factor and sends it to $\mathcal{R}$. This blinding factor can effectively eliminate the threat of losing tag's secret to malicious readers. The detail of HB$^+$ protocol is given in Fig. 2.

Juels and Weis have proved that breaking security of HB$^+$ protocol can be reduced to that of HB protocol [21]. Subsequently, they induced a similar reduction from HB$^+$ attack to solving LPN problem. All of proofs by Juels and Weis assume that the protocols are carried out in sequential manner. Fortunately, in [24], Katz and Shin further assures the security of HB and HB$^+$ protocols by proving that they remain secure under parallel execution of the corresponding basic protocols. This result also implies performance advantage as concurrently executing HB and HB$^+$ basic protocols without sacrificing security makes it possible to run HB and HB$^+$ in 2 and 3 rounds, respectively.

2.4 Man-in-the-middle Attack on HB$^+$

In [22], Gilbert *et al.* presented a very efficient man-in-the-middle attack which could allow an attacker to discover the secret x and y. The attack requires an attacker to intercept the challenge a sent by $\mathcal{R}$ and replace it with $a' = a \oplus \delta$. $\mathcal{T}$ then innocently computes the response z using a'. We have: $z = (a' \cdot x) \oplus (b \cdot y) \oplus v = ((a \oplus \delta) \cdot x) \oplus (b \cdot y) \oplus v = (\delta \cdot x) \oplus (a \cdot x) \oplus (b \cdot y) \oplus v$.

The attacker uses the same δ for r different challenges in one session of the protocol. And if $\mathcal{R}$ accepts $\mathcal{T}$, with high probability, $\delta \cdot x = 0$ since δ does not change the value of the correct response $z = (a \cdot x) \oplus (b \cdot y) \oplus v$. Otherwise, it is likely that $\delta \cdot x = 1$. By collecting k linear independent δ, the attacker can discover x using Gaussian elimination.

Tag(x, y, η)		**Reader**(x, y, η)
$v \leftarrow \text{Ber}_\eta$		
$b \in_R \{0,1\}^k$	$\xrightarrow{b}$	
	$\xleftarrow{a}$	$a \in_R \{0,1\}^k$
$z = (a \cdot x) \oplus (b \cdot y) \oplus v$	$\xrightarrow{z}$	
		Verify $z = (a \cdot x) \oplus (b \cdot y)$

Figure 2: Basic protocol of HB$^+$ protocol

3 Our Proposed Protocol

Key Idea. We now present our variant of HB$^+$ protocol (denoted as HB*) which can resist against man-in-the middle attacks including GRS attack. We observe that in the HB$^+$ protocol, the response z is always computed by associating the secrets x and y with the challenge a and the blinding factor b, respectively. This partly helps the GRS attack because an attacker knows that his modified challenge a' will be counted with respect to x. Note that, in terms of security, there is no distinction between the roles of x and y. Therefore, we think that it is possible to eliminate GRS attack by randomly swapping the roles of x and y when computing the response z. Furthermore, both the tag and the reader should fairly involve in such process so that if either party acts maliciously, security of the protocol will not be compromised. However, we do not want to use special cryptographic primitives like block cipher to achieve our goal. The reason is that it is desireable to preserve the efficiency of HB$^+$ and base security of HB* solely on the LPN problem.

Construction.. In the new protocol, there are 4 k-bit secrets, x, y, r and t shared by the tag and the reader. The new secrets r and t will be used to securely communicate 2 bits between the tag and the reader. The key idea is to embed a bit γ into a pair (a, w) where a is a random k-bit number and $w = (a \cdot s) \oplus \gamma$. If γ is generated at a fixed probability, then a collection of (a, w) form an instance of the LPN problem. Under the assumption that the LPN problem is computationally hard, the pair (a, w) appears as a random $(k+1)$-bit string [24]. Therefore, γ can be securely communicated via (a, w). A detail description of the basic protocol of HB* is given in Fig. 3.

In the basic authentication protocol of HB*, the role of x and y are swapped according to γ and γ' sent by the tag and the reader, respectively. Unlike the HB$^+$ protocol, $\mathcal{T}$ is accepted after r rounds of the basic authentication protocols only if all of r responses are correctly verified. This is because we no longer need to apply noise to the response z as the change in how z is computed for each basic protocol round already does the job. This property is called *perfect completeness* and is another advantage of HB* comparing to HB$^+$ and HB. In case of HB$^+$, even though a genuine RFID tag follows the protocol properly, there is still a chance it is not accepted by the RFID reader. This is clearly not desirable in practical applications. Another big difference between HB* and HB$^+$ is that the basic authentication protocol of HB* is only a 2-round protocol.

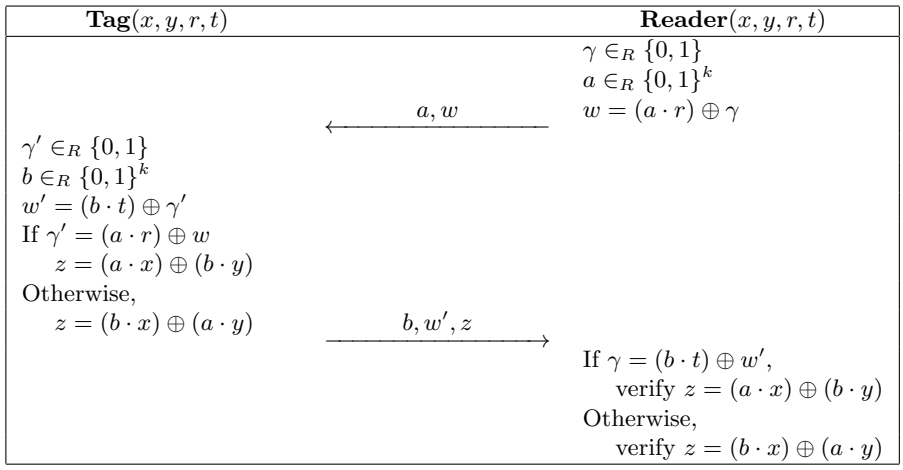

Figure 3: Basic authentication protocol of HB*

In fact, the basic authentication protocol of HB$^+$ can also be 2-round by allowing the tag to send the blinding factor b together with the response z. However, the security proof provided by Juels and Weis requires that the blinding factor b must be sent before the challenge a. Note that, in HB*, the two noise values γ and γ' are chosen at random. Therefore, (a, w) and (b, w') form two instances of the LPN problem with the noise factor $\frac{1}{2}$. As the LPN problem is most intractable with true random noise, HB* is immune against new algorithms to solve instances of the LPN problem with small noise factor. We can also consider other variants of the LPN problem which are possibly harder than the one defined in **Section** 2.2. For example, it is possible to use the noise factor as a secret or a variable value in HB*.

4 Concluding remark

In this paper, we have presented HB* protocol, an augmented version of HB$^+$ protocol which can prevent the man-in-the-middle attack described in [22]. Our protocol can be seen as a combination of two instances of the HB and HB$^+$ protocols. Comparing with the HB$^+$ protocol, our protocol requires one more additional secret, two more binary inner product computation and one more bit to transfer by the reader. Therefore, HB* can still be useful for tightly resource-constrained devices like RFID tags and sensor nodes.

There are several interesting open questions related to HB protocol family, The first problem noted in [24] is to tighten the security reduction from the HB protocol family to the LPN problem. In addition, removing the limitations of imperfect completeness in the HB family would be greatly useful. Finally, we could see that LPN is a fairly well understood hard problem and further applications of LPN problem, especially in designing lightweight cryptographic primitives, should be explored.

References

[1] E. R. Berlekamp, R. J. McEliece and H. C. A Van Tilborg, "On the Inherent Intractability of Certain Coding Problems", *IEEE Transactions on Information Theory*, Vol. 24, pp. 384–386, 1978.

[2] Amos Fiat and Adi Shamir, "How to Prove Yourself: Practical Solutions to Identification and Signature Problems", *Proceedings of CRYPTO'86*, A. M. Odlyzko (Ed.), Springer-Verlag, LNCS 263, pp. 186-194, 1987.

[3] Oded Goldreich and L.A. Levin, "Hard-core Predicates for Any One-Way Function", *21st ACM Symposium on Theory of Computation*, pages 25–32, 1989.

[4] Tsutomu Matsumoto and Hideki Imai, "Human Identification Through Insecure Channel", *Proceedings of EUROCRYPT'91*, D. W. Davies (Ed.), Springer-Verlag, LNCS 547, pp. 409-421, 1991.

[5] Chih-Hung Wang, Tzonelih Hwang and Jiun-Jang Tsai, "On the Matsumoto and Imai's Human Identification Scheme", *Proceedings of EUROCRYPT'95*, L. C. Guillou and J. J. Quisquater (Ed.), Springer-Verlag, LNCS 921, pp. 382-392, 1995.

[6] Oded Goldreich, N. Nisan and A. Wigderson, "On Yao's XOR-lemma", Available at `http://eccc.uni-trier.de/eccc-reports/1995/TR95-050/`.

[7] Tsutomu Matsumoto, "Human-Computer Cryptography: An Attempt", *Proceedings of the Third ACM Conference on Computer and Communications Security*, C. Neuman (Ed.), ACM Press, pp. 68–75, 1996.

[8] Johan Hastad, "Some Optimal Inapproximability Results", *Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing*, ACM Press, pp. 1–10, May, 1997.

[9] Moni Naor and Benny Pinkas, "Visual Authentication and Identification", *Proceedings of CRYPTO'97*, Jr. B. S. Kaliski (Ed.), Springer-Verlag, LNCS 1294, pp. 322–336, 1997.

[10] Michael Kearns, "Efficient noise-tolerant learning from statistical queries", *Journal of ACM* Volume 45, Issue 6, ACM Press, pp. 983–1006, November, 1998.

[11] Xiang-Yang Li and Shang-hua Teng, "Practical Human-Machine Identification over Insecure Channels", *Journal of Combinatorial Optimization*, Volume 3, Kluwer Academic Publishers, pp. 347–361, 1999.

[12] Avir Blum, Adam Kalai and Hal Wasserman, "Noise-tolerant Learning, the Parity Problem, and the Statistical Query Model", *Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing*, ACM Press, pp. 435–440, 2000.

[13] Nicholas Hopper and Manuel Blum, "A Secure Human-Computer Authentication Scheme", Technical Report CMU-CS-00-139, Carnegie Mellon University, May, 2000.

[14] Nicholas Hopper and Manuel Blum, "A Secure Human-Computer Authentication Scheme", *Proceedings of ASIACRYPT'01*, Bart Preneel (Ed.), Springer-Verlag, LNCS 2248, pp. 149-153, 2001.

[15] Stephen Weis, "Security and Privacy in Radio Frequency Identification Devices", Master Thesis, Available at `http://theory.lcs.mit.edu/\~sweis/masters.pdf`, May 2003.

[16] Miyako Ohkubo, Koutarou Suzuki, and Shingo Kinoshita, "Efficient Hash-Chain Based RFID Privacy Protection Scheme", *Proceedings of International Conference on Ubiquitous Computing, Workshop Privacy*, September 2004.

[17] Gildas Avoine and Philippe Oechslin, "A Scalable and Provably Secure Hash-Based RFID Protocol", *Proceedings of Workshop on Pervasive Computing and Communications Security - PerSec'05*, March 2005.

[18] Gildas Avoine, Etienne Dysli, and Philippe Oechslin, "Reducing Time Complexity in RFID System", *Proceedings of Selected Areas in Cryptography (SAC)'05*, Bart Preneel and Stafford Tavares (Ed.), Springer-Verlad, LNCS 3897, pp. 291–306, 2005.

[19] D. Molnar, A. Soppera and D. Wagner, "A Scalable, Delegatable Pseudonym Protocol Enabling Ownership Transfer of a RFID Tag", *Proceedings of Selected Areas in Cryptography (SAC)'05*, Bart Preneel and Stafford Tavares (Ed.), Springer-Verlad, LNCS 3897, pp. 276–290, 2005.

[20] Tassos Dimitriou, "A Lightweight RFID Protocol to Protect against Traceability and Cloning Attacks", *Proceedings of SecureComm'05*, September 2005.

[21] Ari Juels and Stephen Weis, "Authenticating Pervasive Devices with Human Protocols", *Proceedings of CRYPTO'05*, Victor Shoup (Ed.), Springer-Verlag, LNCS 3261, pp. 293-308, 2005.

[22] Henri Gilbert, Matthew Robshaw and Hervé Silbert, "An Active Attack Against HB^+ - A Provably Secure Lightweight Authentication Protocol", Available at `eprint.iacr.org/2005/237.pdf`.

[23] Oded Regev, "On Lattices, Learning with Errors, Random Linear Codes, and Cryptography", *Proceedings of 37th ACM Symposium on Theory of Computing*, ACM, pp. 84–93, 2005.

[24] Jonathan Katz and Ji Sun Shin, "Parrallel and Concurrent Security of the HB and HB^+ Protocols", Available at `http://eprint.iacr.org/2005/461.pdf`.

[25] Ari Juels, "Strenthening EPC Tag against Cloning", *ACM Workshop on Wireless Security (WiSe)*, M. Jakobsson and R. Poovendran (Ed.), pp.67-76. 2005.

[26] Ari Juels, "RFID Security and Privacy: A Research Survey", *To Appear in the Proceedings of IEEE JSAC'06*.

[27] Ilan Kirschenbaum and Avishai Wool, "How to Build a Low-Cost, Extended-Range RFID Skimmer", Available at `http://eprint.iacr.org/2006/054`.

[28] Oded Goldreich, "Modern Cryptography, Probabilistic Proofs and Pseudorandomness", ISBN 3-540-64766-x, Springer-Verlag, Algorithms and Combinatorics, Vol 17, 1998.