Revisiting Lightweight Authentication Protocols Based on Hard Learning Problems

Panagiotis Rizomiliotis
Dep. of Information and Communication Systems Engineering
University of the Aegean
Karlovassi, Samos, GR 83200, Greece
prizomil@aegean.gr

Stefanos Gritzalis
Dep. of Information and Communication Systems Engineering
University of the Aegean
Karlovassi, Samos, GR 83200, Greece
sgritz@aegean.gr

ABSTRACT

At the 2011 Eurocrypt, Kiltz et al., in their best paper price awarded paper, proposed an ultra-lightweight authentication protocol, called $AUTH$. This new protocol is supported by a delegated security proof, against passive and active attacks, based on the conjectured hardness of the Learning Parity with Noise (LPN) problem. However, $AUTH$ has two shortcomings. The security proof does not include man-in-the-middle (MIM) attacks and the communication complexity is high. The weakness against MIM attacks was recently verified as a very efficient key recovery MIM attack was introduced with only linear complexity with respect to the length of the secret key. Regarding the communication overhead, Kiltz et al. proposed a modified version of AUTH where the communication complexity is reduced at the expense of higher storage complexity. This modified protocol was shown to be at least as secure as AUTH.

In this paper, we revisit the security of $AUTH$ and we show, somehow surprisingly, that its communication efficient version is secure against the powerful MIM attacks. This issue was left as an open problem by Kiltz et al. We provide a security proof that is based on the hardness of the LPN problem to support our security analysis.

Categories and Subject Descriptors

E.3 [**Data Encryption**]: Miscellaneous; H.4 [**Information Systems Applications**]: Communications Applications

General Terms

Security

Keywords

RFID authentication protocols, provable security, LPN

WiSec'13, April 17-19, 2013, Budapest, Hungary.

1. INTRODUCTION

The last few years, the design of ultra-lightweight authentication protocols has gained a lot of attention. Motivated mainly by the restrictions that the Radio Frequency Identification (RFID) technology imposes on the available resources for security, several protocols have been proposed [2]. Among them, the most promising family of authentication protocols is the family of HB-like protocols that are based on the so-called *Learning Parity with Noise (LPN)* problem.

The LPN problem is an average-case version of the following problem: given a set of noisy binary equations, find a solution that maximally satisfies the equations. In the worst case version LPN is related to the well studied decoding of a random linear code problem that has been proved to be NP-hard by Berlekamp et al. in [3]. Apart from the authentication protocols, several other cryptographic applications, like encryption schemes ([12]), Message Authentication Codes ([19]), string commitment schemes and zero-knowledge proofs ([16]), have been recently introduced based on the LPN problem.

In [15], Juels and Weis proposed HB^+, a symmetric key authentication scheme, inspired by HB ([14]), the work of Hopper and Blum for the secure identification of human beings. The HB^+ has a very simple circuit representation, as it performs only a few dot-product and bit exclusive-or computations. However, the most interesting feature of the protocol is the elegant proof that supports its security analysis. Specifically, in [15], a concrete reduction of the LPN problem to the security of the HB^+ protocol in two attack models was shown. In the first model the attacker is passive and can only eavesdrop the communication between the prover (tag) and the verifier (reader), while in the second model she is active and she can also send queries to the prover. The original proof was further improved in [**?**], [17].

This security proof does not consider more powerful adversaries that can manipulate messages exchanged between the prover and the verifier. Thus, shortly after the introduction of HB^+, a simple key recovery man-in-the-middle (MIM) attack was proposed ([10]). Motivated by this MIM attack, several variants of HB^+ have been introduced ([5], [6], [7], [8], [11], [21], [22], [26], [28]). However, most of these schemes have been shown to be weak against a MIM attacker.

In this short paper, we will revisit one of these proposals. At the 2011 Eurocrypt, Kiltz et al., in their best paper price awarded paper, proposed an ultra-lightweight authentication protocol, called $AUTH$. This new protocol is supported

by a delegated security proof based on the conjectured hardness of the LPN problem against passive and active attacks. To be more precise, they build on a modified version of the LPN problem, the so-called subset LPN problem. $AUTH$ has two shortcomings. Firstly, the security proof does not include MIM attacks and this weakness against MIM attacks was recently verified as a very efficient key recovery MIM attack was introduced with only linear complexity with respect to the length of the secret key [29]. Secondly, $AUTH$ has rather high communication complexity. To cope with the communication overhead, Kiltz et al. proposed a modified version of $AUTH$ in which the communication complexity is reduced at the expense of higher storage complexity. The authors used a technique adapted by Gilbert et al. to enhance the security of HB^+ in [11]. The size of the exchanged messages between the tag and the reader is reduced, while the shared secret key is increased from a vector to a matrix. We will call this protocol $AUTH^{\#}$. $AUTH^{\#}$ was shown to be at least as secure as AUTH. However, the evaluation of the resistance of $AUTH^{\#}$ against MIM attacks was left as an open problem.

In this paper, we revisit the security of $AUTH$, the ultralightweight cryptographic protocol for RFID authentication, and we show, somehow surprisingly, that its communication complexity efficient version, $AUTH^{\#}$, is much more secure. More precisely, we show that this version of $AUTH$ can provably resist against powerful MIM attacks. Our security proof is based on the hardness of the LPN.

1.1 Outline

The paper is organized as follows. In Section 2, we establish the necessary background on the LPN problem, while in Section 3, we present the $AUTH$ and the $AUTH^{\#}$ authentication protocols. In Section 4, we provide a proof of the security of $AUTH^{\#}$ against MIM attacks. Finally, conclusions and topics for further research can be found in Section 5.

2. BACKGROUND

2.1 Notation

We try to apply, as possible, the established notation. We use normal, bold and capital bold letters, x, $\boldsymbol{x}$ and $\boldsymbol{M}$ to denote single elements, vectors and matrices, respectively. The Hamming weight $\mathrm{wt}(\boldsymbol{x})$ of a vector $\boldsymbol{x} = [x(0), x(1), \cdots, x(n-1)]$ is the number of nonzero elements and $\boldsymbol{M}^T$ is the transpose of a matrix $\boldsymbol{M}$. Also, $\mathbf{0}_m$ denotes the all zero vector of length m and for real numbers $\eta, \psi \in \mathbb{R}$, $]\eta, \psi[= \{x \in \mathbb{R} | \eta < x < \psi\}$. Let $\boldsymbol{a}$ and $\boldsymbol{b}$ be two binary vectors with length l. We use $\boldsymbol{a}_{\downarrow \boldsymbol{b}}$ to denote the subvector of $\boldsymbol{a}$ obtained by deleting all bits of $\boldsymbol{a}$ where $\boldsymbol{b}$ equals 0 (for instance for $\boldsymbol{a} = 10101000$ and $\boldsymbol{b} = 00011010$ we have $\boldsymbol{a}_{\downarrow \boldsymbol{b}} = 010$) and $\boldsymbol{M}_{\downarrow \boldsymbol{b}}$ to denote the submatrix of $\boldsymbol{M}$ obtained by deleting all rows of $\boldsymbol{M}$ where $\boldsymbol{b}$ equals 0. The matrix $\boldsymbol{M}_{\downarrow \boldsymbol{b}}$ can be written as $\boldsymbol{V}(\boldsymbol{b}) \cdot \boldsymbol{M}$, where $\boldsymbol{V}(\boldsymbol{b})$ is a $\mathrm{wt}(\boldsymbol{b}) \times l$ matrix where each row has only one non-zero element.

We use $x \xleftarrow{\$} X$ to denote the assignment to x of a value sampled from the uniform distribution on the finite set X. We use Ber_η to denote the Bernoulli distribution with parameter η, meaning that a bit $\nu \in Ber_\eta$, then $Pr[\nu = 1] = \eta$ and $Pr[\nu = 0] = 1 - \eta$. A vector $\boldsymbol{\nu}$ randomly chosen among all the vectors of length m, such that $\nu(i) \in Ber_\eta$ and $\eta \in (0, 1/2)$, for $0 \leq i \leq m-1$, is denoted as $\boldsymbol{\nu} \xleftarrow{\$} Ber(m, \eta)$, while we use $\boldsymbol{b} \xleftarrow{\$} \{0,1\}^k$ to denote a random binary vector $\boldsymbol{b}$ with length k.

An algorithm D is probabilistic polynomial time if D uses some randomness of its logic and for any input the computation of the algorithm terminates in a number of steps that are a polynomial function in the length of the input. Finally, we denote an arbitrary polynomial function of x by $poly(x)$ and by $f(x) = negl(x)$ a function f that is negligible as a function of x, i.e. it vanishes faster than the inverse of any polynomial in x.

2.2 Learning Parity with Noise

The last few years, the *Learning Parity with Noise (LPN)* problem has gained a lot of attention. It appears in two versions, the decisional and the computational one. In [18], it was shown that the two versions are equivalent and depending on the application the most adequate is used. In this paper we use the computational version.

More precisely, for a secret vector $\boldsymbol{x} \in \{0,1\}^l$, we define $\Lambda_{\eta,l}(\boldsymbol{x})$ the distribution over $\{0,1\}^{l+1}$ where a sample is given by

$$(\boldsymbol{r}, \boldsymbol{r}^T \cdot \boldsymbol{x} \oplus \nu)$$

where $\boldsymbol{r} \in \{0,1\}^l$ and $\nu \in Ber_\eta$. We use $\Omega_{\eta,l}(\boldsymbol{x})$ to denote the oracle that outputs samples from the distribution $\Lambda_{\eta,l}(\boldsymbol{x})$. Let U_l denote the uniform distribution over $\{0,1\}^l$. For any $\boldsymbol{x}$, $\Lambda_{\frac{1}{2},l}(\boldsymbol{x})$ is the same distribution as U_l. The decisional version of the LPN problem is defined as follows.

DEFINITION 1. *The decisional $LPN_{\eta,l}$ problem is (t, q, ϵ)-hard if for any distinguisher D running in time t and making q oracle queries, it holds that,*

$$|Pr\left[\boldsymbol{x} \xleftarrow{\$} \{0,1\}^l : D^{\Omega_{\eta,l}(\boldsymbol{x})}(1^l) = 1\right] - Pr\left[D^{U_{l+1}}(1^l) = 1\right]| \leq \epsilon.$$

The above description corresponds to the average case LPN problem. In machine learning theory, this problem was introduced by Angluin and Laird [1]. Kearns [20] proved that the class of noisy parity concepts is not learnable within the statistical query model. The worst case version is strongly related to the decoding problem of random linear codes, which is $\mathcal{NP}$-complete [3] and hard to approximate within a factor of 2 [13].

For the average case several studies have been proposed for solving the LPN problem for a constant noise parameter η (for instance see [14], [17], [27]). The most popular algorithm for solving the LPN problem is the BKW algorithm, proposed by Blum, Kalai and Wasserman in [4]. The BKW algorithm was further improved, initially, by Fossorier et al. in [9], and most recently by Levieil and Fouque in [23].

2.3 Subspace and subset Learning Parity with Noise Problems

Several problems have been proposed that are based on the hardness of the LPN problem. In [25], the subspace LWE problem was introduced. The subspace LPN problem is the subspace LWE over a field of size $q = 2$.

Let $\boldsymbol{A}$ be a $l \times l$ binary matrix and $\boldsymbol{b} \in \{0,1\}^l$. We define

the distribution,

$$\Gamma_{\eta,l,d}(\boldsymbol{x},\boldsymbol{A},\boldsymbol{b}) = \begin{cases} \perp, \text{ if } \text{rank}(\boldsymbol{A}) < d \\ \Lambda_{\eta,l}(\boldsymbol{A}\boldsymbol{x} \oplus \boldsymbol{b}), \text{ otherwise} \end{cases}$$

and let $\Gamma_{\eta,l,d}(\boldsymbol{x},\cdot,\cdot)$ denote the oracle which on input $\boldsymbol{A}$ and $\boldsymbol{b}$ outputs a sample $\Gamma_{\eta,l,d}(\boldsymbol{x},\boldsymbol{A},\boldsymbol{b})$.

DEFINITION 2. *Let $l, d \in \mathbb{Z}$ where $d \leq l$. The decisional $SLPN_{\eta,l,d}$ problem is (t,q,ϵ)-hard if for every distinguisher D running in time t and making q queries,*

$$|Pr[\boldsymbol{x} \xleftarrow{\$} \{0,1\}^l : D^{\Gamma_{\eta,l,d}(\boldsymbol{x},\cdot,\cdot)} = 1] - Pr[D^{U_{l+1}(\cdot,\cdot)} = 1]| \leq \epsilon,$$

where $U_{l+1}(\cdot,\cdot)$ on input $\boldsymbol{A},\boldsymbol{b}$ outputs a sample of U_{l+1} if $\text{rank}(\boldsymbol{A}) \geq d$ *and $\perp$ otherwise.*

PROPOSITION 1. *[25] For any $l, d, g \in \mathbb{Z}$ where $d+g \leq l$, if the decisional $LPN_{\eta,d}$ problem is (t,q,ϵ)-hard then the decisional $SLPN_{\eta,l,d}$ problem is (t',q,ϵ')-hard where,*

$$t' = t - poly(l,q)$$
$$\epsilon' = \epsilon + 2q/2^{g+1}.$$

The subset LPN problem ($SLPN^*$) is a weaker version of the $SLPN_{\tau,l,d}$ problem where subsets of the secret $\boldsymbol{x}$ are used. Let $\boldsymbol{v} \in \{0,1\}^l$ and $diag(\boldsymbol{v})$ is the zero matrix with $\boldsymbol{v}$ in the diagonal. We define the distribution,

$$\begin{aligned}\Gamma^*_{\eta,l,d}(\boldsymbol{x},\boldsymbol{v}) =& \Gamma_{\eta,l,d}(\boldsymbol{x}, diag(\boldsymbol{v}), \boldsymbol{0}_l) \\ =& \begin{cases} \perp, \text{ if } \text{rank}(\text{wt}\,\boldsymbol{v}) < d \\ \Lambda_{\eta,l}(\boldsymbol{x}\boldsymbol{v}), \text{ otherwise} \end{cases}\end{aligned}$$

From the $\Gamma^*_{\eta,l,d}(\boldsymbol{x},\boldsymbol{v})$ distribution the subset LPN problem is defined as follows.

DEFINITION 3. *Let $l, d \in \mathbb{Z}$ where $d \leq l$. The decisional $SLPN^*_{\eta,l,d}$ problem is (t,q,ϵ)-hard if for every distinguisher D running in time t and making q queries,*

$$|Pr[\boldsymbol{x} \xleftarrow{\$} \{0,1\}^l : D^{\Gamma^*_{\eta,l,d}(\boldsymbol{x},\cdot)} = 1] - Pr[D^{U_{k+1}(\cdot)} = 1]| \leq \epsilon,$$

where $U_{l+1}(\cdot)$ on input $\boldsymbol{v}$, outputs a sample of U_{l+1}, if $\text{wt}(\boldsymbol{v})$, *and $\perp$ otherwise.*

The security of the $AUTH$ protocol is based on the hardness of the subset LPN problem.

2.4 Definition of security models

We consider three types of attacks: passive, active, and man-in-the-middle attacks.

A passive attacker eavesdrops the communication between a legitimate prover (tag) and the verifier (reader) and then she tries to convince the verifier. An active attacker is more powerful, as she can interrogate a prover for a polynomial number of times and then she interacts with the verifier trying to receive an accept message.

In the man-in-the-middle (MIM) attacks, the attacker can interact with both the prover and the verifier and learn the verifier's decision; *accept* or *reject*. This being the strongest security notion for authentication protocols. It is divided into two phases. In the first phase, the attacker modifies the messages exchanged between the prover and the verifier for q invocations of the protocol, while in the second phase the attacker impersonates the prover. Most of the attacks against HB^+ and its variants are MIM ones.

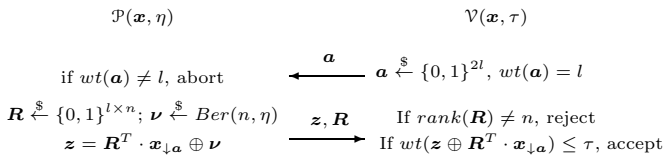

Figure 1: The *AUTH* protocol.

3. THE AUTH AND AUTH# AUTHENTICATION PROTOCOLS

The $AUTH$ protocol is a symmetric key authentication protocol supported by a security proof under the hardness of the subspace LPN problem ([19]). After some initialization phase, the prover $\mathcal{P}$ (the tag) and the verifier $\mathcal{V}$ (the reader) share a secret key $\boldsymbol{x}$ with length $2l$. The basic steps of the protocol go as follows (Fig. 1):

1. The verifier generates a random bit-string $\boldsymbol{a}$ with length $2l$ and sends it to tag $\mathcal{T}$. The Hamming weight of the $\boldsymbol{a}$ must be l.

2. The prover verifies that $\text{wt}(\boldsymbol{a}) = l$ and generates a full rank $l \times n$ random binary matrix $\boldsymbol{R}$ and a bit-string $\boldsymbol{\nu} \in Ber(n,\eta)$. Then, it computes $\boldsymbol{z} = \boldsymbol{R}^T \cdot \boldsymbol{x}_{\downarrow \boldsymbol{a}} \oplus \boldsymbol{\nu}$ and sends to the verifier both $\boldsymbol{z}$ and $\boldsymbol{R}$. If $\text{wt}(\boldsymbol{a}) \neq l$, it aborts the execution of the protocol.

3. The verifier first verifies that the matrix $\boldsymbol{R}$ has rank n and then it accepts if $\text{wt}(\boldsymbol{z} \oplus \boldsymbol{R}^T \cdot \boldsymbol{x}_{\downarrow \boldsymbol{a}}) \leq \tau$, where $n\eta \leq \tau \leq \frac{n}{2}$. If the rank is not correct or the condition is not satisfied, the verifier rejects.

The main disadvantage of AUTH is its extensive communication complexity. In order to reduce this large communication overhead, a trade off between the communication complexity and the key-size was proposed. Actually, they used an idea introduced by Gilbert et al. ([11]) to enhance the security of HB^+. The modified version of the AUTH protocol appears in Fig. 2. We call this modified version $AUTH^\#$. $AUTH^\#$ minimizes the communication complexity, since, instead of sending the $l \times n$ binary matrix $\boldsymbol{R}$, the tag has to send just a l-bit vector $\boldsymbol{r}$. On the other hand, the secret key shared between the verifier and prover increases significantly and a $2l \times n$ matrix $\boldsymbol{X}$ must be stored. The basic steps of the protocol go as follows:

1. The verifier $\mathcal{V}$ generates a random bit-string $\boldsymbol{a}$ with length $2l$, $\text{wt}(\boldsymbol{a}) = l$ and sends it to the prover.

2. The prover verifies that $\text{wt}(\boldsymbol{a}) = l$ and generates a random binary vector $\boldsymbol{r}$ with length l and a bit-string $\boldsymbol{\nu} \in Ber(n,\eta)$. Then, it computes $\boldsymbol{z} = \boldsymbol{r}^T \cdot \boldsymbol{X}_{\downarrow \boldsymbol{a}} \oplus \boldsymbol{\nu}$ and sends to the verifier both $\boldsymbol{z}$ and $\boldsymbol{r}$. If $\boldsymbol{a} \neq l$, it aborts the execution of the protocol.

3. The verifier first verifies that $\text{wt}(\boldsymbol{r}) \neq 0$, otherwise aborts the execution. Then, it accepts if $\text{wt}(\boldsymbol{z} \oplus \boldsymbol{r}^T \cdot \boldsymbol{X}_{\downarrow \boldsymbol{a}}) \leq \tau$, where $n\eta \leq \tau \leq \frac{n}{2}$. Otherwise, the verifier rejects.

In [19], it was proved that AUTH is secure against passive and active attackers given the intractability of the subspace LPN problem. However, recently the very efficient key recovery attack was proposed against AUTH. The attack has

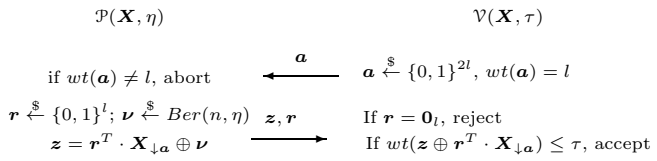

Figure 2: The low communication complexity version of *AUTH* protocol.

linear complexity with respect to the length of the secret key. In [19], it was also shown that the communication efficient variant, $AUTH^{\#}$, was secure against passive and active attacks. The proof is a trivial application of the methodology followed in [11]. However, it is still an an open problem the evaluation of its resistance against MIM attacks. Next, we show that even when the attacker is able to change some of the responses of the prover, then protocol is secure.

Typically, the false rejection rate P_{FR} of the protocol; i.e. the probability to reject a legitimate tag, equals the probability $\mathrm{wt}(\boldsymbol{\nu}) > \tau$ and it is given by

$$P_{FR} = \sum_{i=\tau+1}^{n} \left(\begin{array}{c} n \\ i \end{array} \right) \eta^i (1-\eta)^{n-i}.$$

Finally, the false acceptance rate P_{FA}; i.e. the probability to accept a randomly selected response $\boldsymbol{z}$, can be computed as follows:

$$P_{FA} = \sum_{i=0}^{\tau} \left(\begin{array}{c} n \\ i \end{array} \right) 2^{-n};$$

i.e. it is equal to the number of binary vectors with length n and Hamming weight at most τ.

4. ON THE SECURITY OF AUTH$^{\#}$

4.1 Definition of security models

We use $\mathcal{V}_{\boldsymbol{X},\tau}$ to denote the algorithm that it is run by the verifier and $\mathcal{P}_{\boldsymbol{X},\eta}$ the one run by a legitimate prover. We define two models of security, the $ACT-model$ and the $MIM-model$. In each of the models the adversary runs in two stages. In the first stage she has some interaction with the prover and/or the verifier and in the second she interacts only with the verifier and wins if the verifier returns *accept*. In the $ACT-model$ the active attacker interacts only with an honest prover for a polynomial number of times.

DEFINITION 4. *(ACT-model). In the $ACT-model$ the attack is carried in two phases:*

- **Phase 1.** *The adversary interacts q times with the honest prover.*
- **Phase 2.** *The adversary interacts with the verifier trying to impersonate the prover*

In the $MIM-model$ the attack is carried in two phases and the adversary can manipulate all messages exchanged between the tag and the reader.

DEFINITION 5. *(MIM-model). In the $MIM-model$ the attack is carried in two phases:*

- **Phase 1.** *The adversary interferes for q executions of the protocol. On each execution, the adversary can eavesdrop on all messages exchanged between the honest prover and the honest verifier, including the verifier's decision. In addition, she can modify all these messages with the restriction that all the modifications must have been decided before each execution has started.*
- **Phase 2.** *The adversary interacts with the verifier trying to impersonate the prover.*

In the $MIM-model$, it is assumed that the attacker cannot decide on the alterations of the exchanged messages during the execution of the protocol. This is the class of the most practical MIM attacks, in which the attacker cannot perform computations on the fly during the execution. This class includes all the MIM attacks that have been proposed so far against LPN-based authentication protocols ([10], [24]).

We define the advantage of an adversary $\mathcal{A}$ against $AUTH^{\#}$ protocol in the $ACT-model$ and the $MIM-model$ as the overhead success probability over the false acceptance probability P_{FA} in impersonating the tag:

$$Adv_{\mathcal{A}}^{ACT}(l,n,\eta,\tau,q) = Pr[\boldsymbol{X} \xleftarrow{\$} \{0,1\}^{(2l,n)}, \mathcal{A}^{\mathcal{P}_{\boldsymbol{X},\eta}}(1^k) : \langle \mathcal{A}, \mathcal{V}_{\boldsymbol{X},\tau} \rangle = ACC] - P_{FA}.$$

and

$$Adv_{\mathcal{A}}^{MIM}(l,n,\eta,\tau,q) = Pr[\boldsymbol{X} \xleftarrow{\$} \{0,1\}^{(2l,n)}, \mathcal{A}^{\mathcal{P}_{\boldsymbol{X},\eta},\mathcal{V}_{\boldsymbol{X},\tau}}(1^k) : \langle \mathcal{A}, \mathcal{V}_{\boldsymbol{X},\tau} \rangle = ACC] - P_{FA}.$$

Proof overview. Mainly we adapt the proof of Theorem 2 in [11]. More precisely, we reduce the security in the MIM-model to the security in the ACT-model. The security in the ACT-model has been already proved in [19]. We will show that if there is an attacker $\mathcal{A}^{\#}$ that can efficiently mount a MIM attack with advantage at least δ against $AUTH^{\#}$, then there is an attacker $\mathcal{A}$ that can mount an active attack. Recall that in the MIM-model, the adversary can modify all the messages exchanged between the reader and the tag. The proof goes as follows.

During the first phase $\mathcal{A}$ has to simulate the tag and the reader for $q^{\#}$ times. As $\mathcal{A}$ has access to an honest tag that it can query freely, there is no difficulty in simulating an honest tag to $\mathcal{A}^{\#}$. The main challenge comes with the task of simulating the honest reader. The strategy that we follow for the reader is easy; the reader accepts the tag only when $\mathcal{A}^{\#}$ does not modify any of the messages.

From the point of view of $\mathcal{A}^{\#}$, the tag is perfectly simulated by $\mathcal{A}$. So the success of the attack depends only on the correct simulation of the reader for $q^{\#}$ executions and the success probability of $\mathcal{A}^{\#}$, i.e. $P_{FA} + \delta$. If p_r is the probability of false simulating the reader (for a single execution), then the overall probability of the attack is given by $(1 - q^{\#} \cdot p_r)(P_{FA} + \delta)$.

LEMMA 1. *[11] Let $\boldsymbol{X}$ be a random $l \times m$ binary matrix and let d be an integer, $1 \leq d \leq \frac{m}{2}$. Then, the probability*

$$p(d) = Pr\left[\min_{\boldsymbol{a} \in \mathbb{F}_2^l, \boldsymbol{a} \neq \mathbf{0}_l} (\mathrm{wt}(\boldsymbol{a} \cdot \boldsymbol{X})) \leq d \right],$$

is upper bounded by

$$p(d) \leq 2^{-(1-\frac{l}{m}-H(\frac{d}{m}))},$$

where $H(s) = s \cdot \log_2(\frac{1}{s}) - (1-s) \cdot \log_2(\frac{1}{1-s})$ is the entropy function.

THEOREM 1. *If there is an adversary $\mathcal{A}^{\#}$ that can attack the $AUTH^{\#}$ protocol with parameters (l, n, η, τ) in the MIM-model by modifying $q^{\#}$ protocol executions between the prover and the verifier, with running time $T^{\#}$ and achieving advantage at least $\delta^{\#}$, then, there is an adversary $\mathcal{A}$ that can attack the $AUTH^{\#}$ protocol in the ACT-model with the same parameters by interrogating an honest tag $q^{\#}$ times, with running time at most $T^{\#}$ and with advantage at least $\delta \geq \delta^{\#} - (P_{FA} + \delta^{\#})q^{\#}p_r$, where p_r is a negligible function and P_{FA} is the false acceptance probability.*

PROOF. In the ACT-model, the attacker $\mathcal{A}$ can interrogate a prover. We will show how $\mathcal{A}$ can attack $AUTH^{\#}$ protocol in the ACT-model using the algorithm that the adversary $\mathcal{A}^{\#}$ executes.

During the MIM attack, $\mathcal{A}^{\#}$ is modifying the exchanged messages, and, while, the adversary $\mathcal{A}$ has access to a prover, she has to simulate the behaviour of the verifier. More precisely, her strategy goes as follows.

1. $\mathcal{A}$, simulating the verifier, produces a random bit-string $\boldsymbol{a}$ with length $2l$ and Hamming weight l, and sends it to $\mathcal{A}^{\#}$.
2. $\mathcal{A}^{\#}$ sends $\hat{\boldsymbol{a}} = \boldsymbol{a} \oplus \bar{\boldsymbol{a}}$ to $\mathcal{A}$.
3. $\mathcal{A}$ based on $\hat{\boldsymbol{a}}$ interrogates the prover and sends the produced random binary vector $\boldsymbol{r}$ and the bit-string $\boldsymbol{z}$ to $\mathcal{A}^{\#}$.
4. $\mathcal{A}^{\#}$ produces a new pair $(\hat{\boldsymbol{r}} = \boldsymbol{r} \oplus \bar{\boldsymbol{r}}, \hat{\boldsymbol{z}} = \boldsymbol{z} \oplus \bar{\boldsymbol{z}})$ and sends it to $\mathcal{A}$.
5. $\mathcal{A}$ simulates the verifier as follows. If the triplet $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{r}}, \bar{\boldsymbol{z}})$ is all-zero, the simulated verifier; i.e. $\mathcal{A}$, answers "accept". Otherwise, it rejects.

The previous steps are repeated $q^{\#}$ times. Then, the adversary $\mathcal{A}$ impersonates the prover to a verifier in the ACT-attack, by using the second phase of $\mathcal{A}^{\#}$.

The overall probability p^A of the attack that $\mathcal{A}$ mounts is given by

$$p^A = p_{auth} \cdot (P_{FA} + \delta) \tag{1}$$

where p_{auth} is the probability of successfully simulating a verifier's behaviour and it depends on the ability of the adversary to simulate the last step; i.e. the acceptance or rejection decision.

Next, we compute p^A. In order for the attack to be successful, the adversary $\mathcal{A}$ must be able to simulate the reader's behavior for $q^{\#}$ consecutive executions of the protocol. Let p_r be the probability to fail in a single execution. Then,

$$p_{auth} = (1 - q^{\#} \cdot p_r). \tag{2}$$

The probability of false rejecting, when the triplet $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}})$ is all zero, i.e. when $\mathcal{A}^{\#}$ does not modify any of the messages, is P_{FR}. That is, $p_r \geq P_{FR}$.

When $(\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}) \neq (\mathbf{0}_{2l}, \mathbf{0}_n, \mathbf{0}_l)$, the probability p_r of false simulating the reader is also defined by the probability that the condition $\mathrm{wt}(\hat{\boldsymbol{z}} \oplus \hat{\boldsymbol{r}}^T \cdot \boldsymbol{X}_{\downarrow \boldsymbol{a}}) \leq \tau$, where $n\eta \leq \tau \leq \frac{n}{2}$, is satisfied. We use $FAIL$ to indicate this event.

The sum $\hat{\boldsymbol{z}} \oplus \hat{\boldsymbol{r}}^T \cdot \boldsymbol{X}_{\downarrow \boldsymbol{a}}$ can be written as

$$\begin{aligned}
&\bar{\boldsymbol{z}} \oplus \boldsymbol{r}^{\boldsymbol{T}} \cdot \boldsymbol{X}_{\downarrow \hat{\boldsymbol{a}}} \oplus \boldsymbol{\nu} \oplus \hat{\boldsymbol{r}}^T \cdot \boldsymbol{X}_{\downarrow \boldsymbol{a}} = \\
&\bar{\boldsymbol{z}} \oplus \boldsymbol{r}^{\boldsymbol{T}} \cdot \boldsymbol{V}(\bar{\boldsymbol{a}} \oplus \boldsymbol{a}) \cdot \boldsymbol{X} \oplus \boldsymbol{\nu} \oplus (\bar{\boldsymbol{r}}^T \oplus \boldsymbol{r}^T) \cdot \boldsymbol{V}(\boldsymbol{a}) \cdot \boldsymbol{X} = \\
&\bar{\boldsymbol{z}} \oplus (\boldsymbol{r}^{\boldsymbol{T}} \cdot (\boldsymbol{V}(\boldsymbol{a}) \oplus \boldsymbol{V}(\bar{\boldsymbol{a}} \oplus \boldsymbol{a})) \oplus \bar{\boldsymbol{r}}^T \cdot \boldsymbol{V}(\boldsymbol{a})) \cdot \boldsymbol{X} \oplus \boldsymbol{\nu}.
\end{aligned}$$

Let $\boldsymbol{y}_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}} = \bar{\boldsymbol{z}} \oplus (\boldsymbol{r}^{\boldsymbol{T}} \cdot \boldsymbol{D}\boldsymbol{V}_{\boldsymbol{a}}(\bar{\boldsymbol{a}}) \oplus \bar{\boldsymbol{r}}^T \cdot \boldsymbol{V}(\boldsymbol{a})) \cdot \boldsymbol{X}$ and let $\beta_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}}$ be the Hamming weight of $\boldsymbol{y}_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}}$. Then, $n - \beta_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}}$ bits of $\boldsymbol{y}_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}} \oplus \boldsymbol{\nu}$ follow a Bernoulli distribution of parameter η and the rest $\beta_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}}$ bits follow a Bernoulli distribution of parameter $1-\eta$. That is, the Hamming weight $\mathrm{wt}(\boldsymbol{y}_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}} \oplus \boldsymbol{\nu})$ follows a binomial distribution of expected value $\mu = (n - \beta_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}})\eta + \beta_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}}(1 - \eta)$ and variance $\sigma^2 = n\eta(1 - \eta)$.

Since, the expected value is a function of $\beta_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}}$ we can easily verify that for $\beta_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}} \geq 1 + \lfloor \frac{\tau - \eta n}{1 - 2\eta} \rfloor$, it holds that $\mu > \tau$. For any $\beta_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}} \geq 1 + \lfloor \frac{\tau - \eta n}{1 - 2\eta} \rfloor$, any $\boldsymbol{X}$ and $\boldsymbol{\nu}$, it holds that

$$\begin{aligned}
&P_r[FAIL] = \\
&Pr_{\boldsymbol{\nu}}[FAIL | dmin(\boldsymbol{X}) > \beta_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}}] Pr_{\boldsymbol{X}}[dmin(\boldsymbol{X}) > \beta_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}}] + \\
&Pr_{\boldsymbol{\nu}}[FAIL | dmin(\boldsymbol{X}) \leq \beta_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}}] Pr_{\boldsymbol{X}}[dmin(\boldsymbol{X}) \leq \beta_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}}]
\end{aligned}$$

where $dmin(\boldsymbol{X}) = \min_{\boldsymbol{a} \in \mathbb{F}_2^l, \boldsymbol{a} \neq \mathbf{0}_l}(\mathrm{wt}(\boldsymbol{a} \cdot \boldsymbol{X}))$.

When, $\mu > \tau$; i.e. $\beta_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}} \geq 1 + \lfloor \frac{\tau - \eta n}{1 - 2\eta} \rfloor$ from the Chernoff bound we have that $\mathrm{wt}(\boldsymbol{y}_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}} \oplus \boldsymbol{\nu}) < \tau$ with probability less than $e^{-\frac{(\mu - \tau)^2}{2\mu}}$ and the simulation fails. From the above observation and from Lemma 1, we have that

$$\begin{aligned}
P_r[FAIL] &\leq Pr_{\boldsymbol{\nu}}[FAIL | dmin(\boldsymbol{X}) > \beta_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}}] \\
&\quad + Pr_{\boldsymbol{X}}[dmin(\boldsymbol{X}) \leq \beta_{\bar{\boldsymbol{a}}, \bar{\boldsymbol{z}}, \bar{\boldsymbol{r}}}] \\
&\leq e^{-\frac{(\mu - \tau)^2}{2\mu}} + 2^{-n + 2l + nH(\frac{1 + \lfloor \frac{\tau - \eta n}{1 - 2\eta} \rfloor}{n})}.
\end{aligned}$$

Similarly to [11], in order to ascertain that the first term is negligible, we define $\hat{d}$ the least integer such that $\mu(\hat{(d)} > (1 + c)\tau$ for some $c > 0$ and for all $d \geq \hat{d}$, $e^{-\frac{(\mu - \tau)^2}{2\mu}} \leq e^{-\frac{(c\tau)^2}{2(c+1)}}$. Also, for practical values of the parameters the exponent of the second term is negative, while the P_{FR} is negligible. Thus, from (1) and (2), the overall probability of the attack is lower bounded by

$$\begin{aligned}
&(1 - q^{\#} \cdot (e^{-\frac{(\mu - \tau)^2}{2\mu}} + 2^{-n + 2l + nH(\frac{1 + \lfloor \frac{\tau - \eta n}{1 - 2\eta} \rfloor}{n})})) \\
&\quad \cdot (P_{FA} + \delta) < p^A.
\end{aligned}$$

□

From Theorem 1, any efficient attacker achieving a noticeable advantage $\delta^{\#}$ against the $AUTH^{\#}$ protocol in the MIM-model can be turned into an efficient attacker against the same protocol in the ACT-model. However, from [19], this contradicts the hardness assumption of the subspace LPN problem.

5. CONCLUSIONS

The design of lightweight authentication protocols is a challenging task. One of the most recent proposals, $AUTH$, was introduced in 2011 by Kiltz et al., in their Eurocrypt best paper price awarded paper. One of the main advantages of $AUTH$ is the elegant security proof, against passive and active attacks, based on the conjectured hardness of the LPN problem that supports its security analysis. However, due to its high communication complexity, Kiltz et al. presented a variant of $AUTH$ with significant smaller communication overhead, but with higher storage complexity. It was also proved that this variant was at least as secure as $AUTH$.

In this paper, we have revisited the security of $AUTH$ and have shown that its variant is much more secure. More

precisely, we showed that it can resist powerful MIM attacks and we provided a security proof based on the hardness of the LPN problem to support our security analysis. However, it remains an interesting open problem the designing of a variant of $AUTH$ that has both small storage and communication complexity.

6. ACKNOWLEDGEMENTS

This research is performed in the framework of the INTERREG III Poseidon project, which is funded by the European Union (80%) and National Funds of Greece and Cyprus (20%).

7. REFERENCES

[1] D. Angluin and P. Laird. Learning from Noisy Examples. *Machine Learning*, vol. 2(4), 1987, pp. 343–370.

[2] G. Avoine. *RFID Security and Privacy Lounge*. The list of papers is available at http://www.avoine.net/rfid/download/bib/bibliography-rfid.pdf.

[3] E. R. Berlekamp, R. J. McEliece, V. Tilborg. On the Inherent Intractability of Certain Coding Problem. *IEEE Transactions on Information Theory*,vol. 24, 1978, pp. 384-386.

[4] A. Blum, A. Kalai, and H. Wasserman. Noise-Tolerant Learning, the Parity Problem, and the Statistical Query Model. Journal of the ACM, vol. 4, 2003, pp. 506-519.

[5] J. Bringer, H. Chabanne, EH. Dottax. HB^{++}: a Lightweight Authentication Protocol Secure against Some Attacks. In Proc. of the IEEE Int. Conference on Pervasive Sevices, Workshop - SecPerU, 2006.

[6] J. Bringer, H. Chabanne. *Trusted-HB*: A Low-Cost Version of HB Secure Against Man-in-the-Middle AttackHB^{++}. *IEEE Transactions on Information Theory*, vol. 54, 2008, pp. 4339-4342.

[7] C. Bosley, K. Haralambiev, A. Nicolosi. HB^N: An HB-like protocol secure against man-in-the-middle attacks. Cryptology ePrint Archive, Report 2011/350 (2011), http://eprint.iacr.org.

[8] D.N. Duc and K. Kim. Securing HB^+ against GRS Man-in-the-Middle Attack. In Proc. of the Symp. on Cryptography and Information Security, 2007.

[9] M.P.C. Fossorier, M.J. Mihaljevic, H. Imai, Y. Cui, and K. Matsuura. A Novel Algorithm for Solving the LPN Problem and its Apllication to Security Evalucation of the HB Protocol for RFID Authentication. Cryptology ePrint Archive, Report 2006/197, http://eprint.iacr.org, 2006.

[10] H. Gilbert, M. Robshaw, and Y. Silbert. An Active Attack against HB^+-a Provable Secure Lightweighted Authentication Protocol. Cryptology ePrint Archive, Report 2005/237, http://eprint.iacr.org, 2005.

[11] H. Gilbert, M. Robshaw, and Y. Silbert. $HB^{\#}$: Increasing the Security and Efficiency of HB^+. In Proc. of Eurocrypt, Springer LNCS, vol. 4965, 2008, pp. 361–378.

[12] H. Gilbert, M. Robshaw, and Y.Seurin. How to Encrypt with the LPN Problem. In Proc. of ICALP '08, LNCS 5126, 2008, pp. 679-690.

[13] J. Hastad. Some Optimal Inapproximability Results. *J. ACM*, vol. 48 (4), 2001, pp. 798-859.

[14] N.J. Hopper, and M., Blum. Secure Human Identification Protocols. In Proc. of Asiacrypt, Springer LNCS, vol. 2248, 2001, pp. 52–66.

[15] A, Juels, and S.A. Weis. Authenticating Pervasive Devices with Human Protocols. In Proc. of Crypto, Springer LNCS, vol. 3126, 2005, pp. 293–308.

[16] A. Jain, S. Krenn, K. Pietrzak and Aris Tentes. Commitments and Efficient Zero-Knowledge Proofs from Hard Learning Problems. In Proc. of Asiacrypt, Springer LNCS, vol. 7658, 2012, pp. 663–680.

[17] J. Katz, and A. Smith. Analyzing the HB and HB^+ Protocols in the Large Error Case. Cryptology ePrint Archive, Report 2006/326, http://eprint.iacr.org/, 2006.

[18] J. Katz, and J. Shin. Parallel and Concurrent Security of the HB and HB^+ Protocols. *Journal of Cryptology*, vol. 23, 2010, pp. 402–421.

[19] E. Kiltz, K. Pietrzak, D. Cash, A. Jain, and D. Venturi. Efficient Authentication from Hard Learning Problems. In Proc. of Eurocrypt, Springer LNCS, vol. 6632, 2011, pp. 7–26.

[20] M. Kearns. Efficient noise-tolerant learning from statistical queries. In Proc. of the 25th ACM Symposium on Theory of Computing, 1993, pp. 392–401.

[21] X. Leng, K. Mayes, and K. Markantonakis. *HP-MP*$^+$: An Improvement on the *HB-MP* Protocol. In Proc. of the IEEE Int. Conference on RFID 2008, IEEE Press, 2008, pp. 118–124.

[22] J. Munilla, and A. Peinado. *HP-MP*: A Further Step in the *HB*-family of Lightweight authentication protocols. *Computer Networks*, Elsevier, vol. 51, 2007, pp. 2262–2267.

[23] E. Levieil, and P.A. Fouque. An improved LPN Algorithm. In Proc. of SCN, Springer LNCS 4116, 2006, pp. 348–359.

[24] K. Ouafi, R. Overbeck, V. Vaudenay. On the Security of $HB^{\#}$ against a Man-in-the-Middle Attack. In Proc. of Asiacrypt, Springer LNCS, vol. 5350, 2008, pp. 108–124.

[25] K. Pietrzak. Subspace LWE. 2010. Manuscript available at http://homepages.cwi.nl/pietrzak/publications/SLWE.pdf.

[26] S. Piramuthu. *HB* and Related Lightweight Authentication Protocols for Secure RFID Tag/Reader Authentication. In Proc. of CollECTeR Europe Conference, Basel, Switzerland, 2006.

[27] O. Regev. On Lattices, Learning with Errors, Random Linear Codes, and Cryptography. In Proc. of STOC, ACM, 2005, pp. 84–93.

[28] P. Rizomiliotis. *HB-MAC*: Improving the Random - $HB^{\#}$ Authentication Protocol. In Proc. of the 6th International Conference on Trust, Privacy and Security in Digital Business (TrustBus), Springer, LNCS 5695, 2009, pp. 159–168.

[29] P. Rizomiliotis and S. Gritzalis. On the security of AUTH, a provably secure authentication protocol based on the subspace LPN problem. Accepted for publication in the Int. J. of Inform. Security, 2012.