Secure and Efficient HB-CM Entity Authentication Protocol

Zhijun Li[1], Guang Gong[1], and Zhiguang Qin[2]

[1] University of Waterloo, Waterloo, ON, Canada
leezj@engmail.uwaterloo.ca ggong@calliope.uwaterloo.ca
[2] University of Electronic Science & Technology of China, Chengdu, China
qinzg@uestc.edu.cn

Abstract. The simple, computationally efficient LPN-based HB-like entity authentication protocols have attracted a great deal of attention in the past few years due to the broad application prospect in low-cost pervasive devices. At present, the most efficient protocol is HB$^{\#}$, which is proven to resist the GRS attack under the conjecture that it is secure in the DET-model. In this paper, we introduce an innovative HB-CM^{-} protocol, which significantly reduces the storage requirement while maintaining the same level of communication cost. We develop the concept of equivalence class, and present HB-CM^{-} reductionist proof that overcomes an inherent limitation in the HB$^{\#}$ security proof. In fact, HB$^{\#}$ is only provably resistant to partial instances of GRS attack, while we prove that HB-CM^{-} can prevent the full GRS attack except one trivial case. In addition, we propose a new noise mode for all HB-like protocols in order to thwart the latest OOV man-in-the-middle attack, which can effectively compromise all current HB-like protocols with the basic Bernoulli nose mode. The HB-CM^{-} protocol along with the proposed noise mode constitutes our final protocol: HB-CM.

1 Introduction

Nowadays, there are more and more low-cost pervasive devices employed. Radio frequency identification (RFID) tags are the typical representative. RFID is a technology for automated identification of physical entities using radio frequency transmissions. Typically, RFID systems consist of simple, low-cost tags that are attached to physical objects and powerful readers that queue data from tags. Billions of tags have been deployed; tens of billions are on the way, making RFID tags the most pervasive microchips in the history. The production cost of those pervasive devices partly determines the success of the application systems. In general, RFID tag's price must be below ten cents to be considered affordable for most RFID applications [1]. Consequently, the low cost demand for pervasive devices causes them to be very resource limited.

Security and privacy play important roles in the prevalence of RFID systems. Secure and efficient entity authentication protocols are natural approaches to address the counterfeiting problem, which imposes a serious threat to those low-cost pervasive computing devices. These devices, which lack the computation,

storage, energy, and communication capacities necessary for most cryptographic authentication schemes, call for lightweight authentication approaches.

The HB-like authentication protocols [2, 3] have gained much attention in this field. Their solid security foundation, a well-studied Learning Parity with Noise (LPN) hard problem, and the lightweight computation requirement, imposing only bitwise operations on authentication participants, makes them very attractive for entity authentication in the resource-constrained devices. Among previously proposed protocols, the HB$^{\#}$ protocol [3] is the most practical one because of its security against the the GRS man-in-the-middle attack [4], the relatively small memory requirement, and the extremely reduced communication cost.

Our Contributions. In this paper, we first propose an innovative HB-like entity authentication protocol: HB-CM^{-}, using the technique of circulant matrix with a special property. Compared to HB$^{\#}$, the HB-CM^{-} protocol significantly reduces the memory requirement without degrading other performances. Moreover, our security proof of HB-CM^{-} overcomes an inherent limitation for the HB$^{\#}$ reductionist proof in [3]. In fact, HB$^{\#}$ is only provably resistant to partial instances of GRS attack, while we prove the HB-CM^{-} protocol security against the full GRS attack except one trivial case.

In addition, we introduce a new noise mode for the HB-like protocols to address the security challenge brought about by a general man-in-the-middle attack (OOV attack), which was discovered by Ouafi, Overbeck, and Vaudenay [5]. All HB-like protocols can benefit from the proposed noise mode in terms of resistance to the OOV attack and zero false rejection, though the underlying hard problem is changed, and needs more research. Adopting the proposed noise mode, the HB-CM^{-} protocol turns into our final protocol: HB-CM. We also give the practical parameter recommendation for the HB-CM protocol.

Notation. The following notation is used throughout this paper.

- $\boldsymbol{a} \circ \boldsymbol{x}$: the binary inner-product of two vectors (or matrices) $\boldsymbol{a}$ and $\boldsymbol{x}$.
- $\boldsymbol{a} \oplus \boldsymbol{x}$: the bitwise exclusive-or (xor) operation of two vectors (or matrices) $\boldsymbol{a}$ and $\boldsymbol{x}$.
- $\boldsymbol{a}||\boldsymbol{b}$: the concatenation of two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$.
- $\mathbb{A}||\mathbb{B}$: the concatenation of two matrices $\mathbb{A}$ and $\mathbb{B}$ with a same number of rows; each row in the resulting matrix is the concatenation of two corresponding rows in $\mathbb{A}$ and $\mathbb{B}$.
- Hwt($\boldsymbol{x}$): the Hamming weight of the binary vector $\boldsymbol{x}$.
- Rtt($\boldsymbol{\theta}, i$): the left rotation operation on vector $\boldsymbol{\theta}$, that is, Rtt$(\boldsymbol{\theta}, i) = (\theta_i, \theta_{i+1}, \cdots, \theta_{k\text{-}1}, \theta_0, \theta_1, \cdots, \theta_{i\text{-}1})$, where $\boldsymbol{\theta} = (\theta_0, \theta_1, \cdots, \theta_{k\text{-}1}), i \in \{0, 1, \cdots, k-1\}$.
- $\mathbf{S}^k$: the set of all k-bit vectors except $\mathbf{0}^{(k)}$ (all k bits are 0) and $\mathbf{1}^{(k)}$ (all k bits are 1).

The rest of this paper is organized as follows. We review the LPN problem and previous HB-like protocols in Section 2. Then we present the HB-CM^{-}

protocol in Section 3. The security definitions, proofs, and arguments of HB-CM$^-$ are given in Section 4. Afterwards, we propose the new noise mode for all HB-like protocols to resist the OOV attack, and depict the final HB-CM protocol in Section 5. The practical parameter setting and performance comparison are stated in Section 6. Section 7 concludes our work.

2 Previous Work

2.1 The LPN Problem

Suppose the tag shares a secret k-bit vector $\boldsymbol{x}$ with the reader for authentication. First the reader randomly generates a sequence of binary vectors $\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n$ and sends those challenges to the tag, which responds with $z_i = \boldsymbol{x} \circ \boldsymbol{a}_i$ accordingly. The reader accepts the tag's authentication if and only if $\boldsymbol{x} \circ \boldsymbol{a}_i = z_i$. Unfortunately, after observing $O(k)$ linearly-independent challenge-response pairs of $(\boldsymbol{a}_i, z_i)$, an adversary can readily recover the secret $\boldsymbol{x}$ using the Gaussian elimination.

In the presence of noise, however, where each bit z_i is independently xored by a noise bit taking '1' with probability $\eta \in (0, \frac{1}{2})$, determining $\boldsymbol{x}$ becomes much more difficult. This problem is known as Learning Parity with Noise, or the *LPN Problem*. Formally, it is defined as follows:

Definition 1 (LPN Problem). *View k as a security parameter. Let $\boldsymbol{x}$ be a secret random k-bit column vector, $\mathbb{A}$ be a random $(q \times k)$-binary matrix, $\eta \in (0, \frac{1}{2})$ be a noise level, and $\boldsymbol{v}$ be a random q-bit vector such that* $\mathrm{Hwt}(\boldsymbol{v}) \leq \eta q$. *Given $\mathbb{A}$, η, and $\boldsymbol{z} = (\mathbb{A} \circ \boldsymbol{x}) \oplus \boldsymbol{v}$, find a k-bit column vector $\boldsymbol{y}$ such that* $\mathrm{Hwt}((\mathbb{A} \circ \boldsymbol{y}) \oplus \boldsymbol{z}) \leq \eta q$.

The LPN problem has long been studied as the following equivalent problems: syndrome decoding problem [6, 7] and minimal disagreement parity problem [8]. It has been proven that The problem is NP-Hard [7], and it is still hard to just find a vector satisfying more than half of the challenge-response pairs [9]. Furthermore, Regev [10] introducing a natural extension of the LPN problem, referring to as the Learning With Error (LWE) problem, by replacing $GF(2)$ in the LPN problem with $GF(p)$, where p is a prime number. Regev [10] proved the reduction from worse-case lattice problems, such as Shortest-Vector Problem (SVP), to the LWE problem. Similar to other NP-Hard problems for application in the cryptography, the security of LPN-based authentication protocols depends on the hardness of average case of the LPN problem, while the NP-Hard allegation only guarantees the intractability in the worse case. The computational complexity of algorithms for solving random instances in the LPN-based protocols has been addressed in [11–14].

2.2 HB-Family Authentication Protocols

Hopper and Blum [11] first presented an LPN-based straightforward authentication protocol (HB protocol) on the application area of human identification.

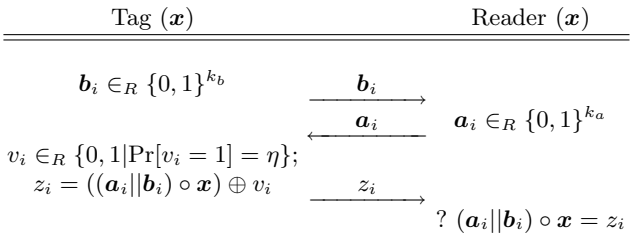

Fig. 1. The ith round of the HB$^+$ authentication protocol, where $\boldsymbol{x}$ is a k-bit column vector, $k_a + k_b = k$, $\eta \in (0, \frac{1}{2})$, $\boldsymbol{b}_i$ is a *blinding vector*, $\boldsymbol{a}_i$ is a *challenge vector*, and $\boldsymbol{a}_i||\boldsymbol{b}_i$ is named a *joint-challenge vector*

The HB protocol is provably secure against passive eavesdroppers [2, 15] under the assumption of the LPN problem's intractability. However, an active attacker can easily overcome the noise by repeating identical challenges (*JW attack*) and then recover the authentication key [2]. Focusing on the lightweight authentication for the RFID systems, Juels and Weis [2] proposed the HB$^+$ authentication protocol, which prevents the aforementioned active attack by adding blinding vectors. One authentication procedure in HB$^+$ consists of n rounds of interactions between the tag and the reader. One single round of HB$^+$ is outlined in Fig. 1 [1]. After n rounds, the reader accepts the tag's authentication if the number of unmatched challenge-response pairs does not exceed a threshold τ.

Juels and Weis proved that the HB$^+$ protocol is secure under the DET-model [2]. The security proof of HB$^+$ in [2] requires the sequential execution (that is, one by one) of n rounds. Katz and Shin [15] gave an elegant security proof of HB$^+$ protocol in the case of parallel and concurrent executions. Katz and Smith [16] further extended these theoretical results to a larger range of noise levels $\frac{1}{4} \leq \eta < \frac{1}{2}$ whereas the Katz-Shin proof [15] is valid only for $\eta < \frac{1}{4}$. Despite that the HB$^+$ protocol has those security proofs in the relatively restrictive DET-model, Gilbert, Robshaw, and Sibert [4] discovered that there exists a simple man-in-the-middle (MIM) attack (GRS attack) which completely compromises the HB$^+$ protocol.

GRS Attack. Let $\boldsymbol{x}^{(a)}||\boldsymbol{x}^{(b)} = \boldsymbol{x}$, where vector $\boldsymbol{x}^{(a)}$ is of k_a bits and vector $\boldsymbol{x}^{(b)}$ is of k_b bits. The original GRS attack [4] is launched as follows. In second pass of every round of one HB$^+$ authentication procedure, an adversary intercepts the challenge $\boldsymbol{a}_i$, and replaces it with $\boldsymbol{a}_i \oplus \boldsymbol{\delta}^{(a)}$, where $\boldsymbol{\delta}^{(a)}$ is a constant vector for one authentication procedure. Acceptance or rejection of this manipulated authentication procedure will reveal to the adversary the result of $\boldsymbol{\delta}^{(a)} \circ \boldsymbol{x}^{(a)}$, that is, one bit of $\boldsymbol{x}^{(a)}$. He simply repeats k_a times of manipulating authentication procedures with linearly independent $\boldsymbol{\delta}^{(a)}$'s, and then fully recovers $\boldsymbol{x}^{(a)}$. After that, the adversary is able to impersonate a valid tag by setting $\boldsymbol{b} = \boldsymbol{0}^{(k_b)}$; or the

[1] In the original HB$^+$ proposal [2], there are two secret vectors rather than one. In essence, it is equivalent to the description presented here.

adversary can recover $\boldsymbol{x}^{(b)}$ by acting as a tag to interact with a genuine reader, using a constant blinding vector $\boldsymbol{b'}$ in one authentication procedure, responding challenge $\boldsymbol{a}_i$ with $\boldsymbol{a}_i \oplus \boldsymbol{x}^{(a)}$, and learning the result of $\boldsymbol{b'} \oplus \boldsymbol{x}^{(b)}$ according to acceptance or rejection.

Apparently, the GRS manipulation strategy can be applied to blinding vectors to recover $\boldsymbol{x}^{(b)}$. Then the adversary can launch the original JW attack to retrieve $\boldsymbol{x}^{(a)}$, totally breaking the protocol. Therefore, in a full GRS attack model, the adversary should be allowed to manipulate both blinding vectors and challenge vectors, replacing the joint-challenge vector $\boldsymbol{a}_i||\boldsymbol{b}_i$ with $(\boldsymbol{a}_i||\boldsymbol{b}_i) \oplus \boldsymbol{\delta}$. We refer to $\boldsymbol{\delta}$ as an *interference vector*.

Even after a series of HB$^+$ enhancement protocols, such as HB^{++} [17], HB* [18], HB-MP [19], modification of HB^{++} [20] and HB-MP$^+$ [21] have been proposed, Gilbert, Robshaw, and Sibert [22] demonstrated that those variants still could be attacked in the linear time while increasing the computational complexity and/or reduced the practicality. The PUF-HB protocol [23] and the Trust-HB protocol [24] make use of a physically unclonable circuit and a lightweight hash function family respectively, intending to thwart the GRS attack. However, the introduction of such ingredients into HB$^+$ might not fully meet the motivation of designing lightweight simple-bitwise-operation-based authentication protocols. Moreover, Frumkin and Shamir [25] recently broke the security of Trust-HB in realistic scenarios.

Gilbert, Robshaw, and Seurin [3] presented the Random-HB$^\#$ protocol and the HB$^\#$ protocol, which are resistant to the GRS attack. In contrast to the secret vector used in HB$^+$, Random-HB$^\#$ employs a secret matrix. Instead of blinding-challenge matrices for one HB$^+$ authentication, the Random-HB$^\#$ protocol only needs blinding-challenge vectors, exceedingly reducing the communication cost. However, the $(k \times n)$ secret matrix in Random-HB$^\#$ imposes too high storage burden to be practical in realistic systems. In order to overcome the drawback, they proposed to replace the random matrix with a Toeplitz matrix, which becomes the HB$^\#$ protocol.

Revised HB$^\#$. The original HB$^\#$ protocol [3] requires two independent Toeplitz matrices, one $(k_a \times n)$, the other $(k_b \times n)$; thus the total memory cost is $k+2n-2$ bits, where $k = k_a + k_b$. We propose a revised HB$^\#$ protocol, which is depicted Fig. 2, by changing to one $(k \times n)$ Toeplitz matrix. The original HB$^\#$ security proof in [3] still holds for the revised HB$^\#$. Since it reduces the memory cost to $k + n - 1$ bits while all other performances stay unchanged compared to the original HB$^\#$, the revised HB$^\#$ protocol should be the final version of HB$^\#$.

In the DET-model, Random-HB$^\#$ is provably secure, while HB$^\#$ is conjured to be secure [3]. On the other hand, Gilbert, Robshaw, and Seurin [3] use a GRS-MIM-model, in which the adversary is only allowed to manipulate the challenges from the reader to the tag, to prove that Random-HB$^\#$ and HB$^\#$ [3] resist the GRS attack. However, the GRS-MIM-model does not completely

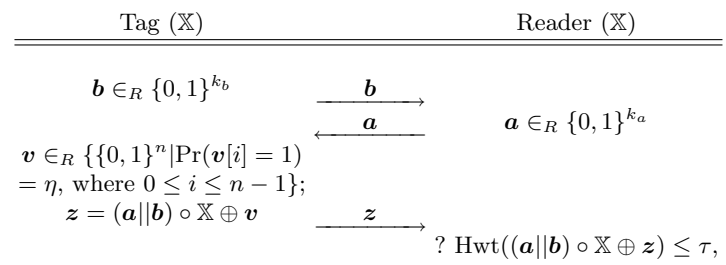

Fig. 2. Revised HB$^{\#}$ authentication protocol, where $\mathbb{X}$ is a $(k \times n)$ Toeplitz matrix, $k_a + k_b = k$, noise level $\eta \in (0, \frac{1}{2})$, pass-threshold $\tau \in (\eta n, \frac{n}{2})$

simulate the full GRS attack. From the practical point of view, it is unreasonable to restrict the GRS attacker's ability to only being able to manipulate challenge vectors. Such a limitation merely results from the fact that the security proofs [3] only hold on that condition for practical parameters. Even though there is no realistic reason to doubt if Random-HB$^{\#}$ and HB$^{\#}$ are vulnerable to the full GRS attack, it is still more desirable that a protocol can provably resist the full GRS attack. On the other hand, the key storage cost for the HB$^{\#}$ protocol is still a little higher than ordinary RFID tags' capacity. For example, even for the revised HB$^{\#}$ protocol, a typical authentication key incorporates above one thousand bits. In order to overcome these two drawbacks of HB$^{\#}$, we propose an innovative HB-CM^{-} protocol, which is provably resistance to the full GRS attack except one trivial case, and consumes almost lesser key storage, without degrading other protocol performances.

3 HB-CM^{-} Protocol

The HB-CM^{-} protocol makes use of circulant matrices with a special property to encode joint-challenges. An $(n \times k)$ circulant matrix $\mathbb{C} = |\mathbb{C}[i,j]|_{n \times k}$ is a special matrix for which $\mathbb{C}[i,j] = \mathbb{C}[(i+1) \mod n, (j+1) \mod k]$, for $i = 0, \cdots, n-1$; $j = 0, \cdots, k-1$. Consequently, an $(n \times k)$ circulant matrix $\mathbb{C}$ can be stored in $\max(n,k)$ bits. A circulant matrix is employed in the MixColumns step of the Advanced Encryption Standard (AES) by [26]. In the HB-CM^{-} protocol, the parameter set is $(d, k = k_a + k_b, \eta, n, \tau)$, where d is the *security level*, k is the *key length*, η is the *noise level*, n is the *interaction expansion*, and τ is the *pass-threshold*. For security matters, we require the key length k to be a *prime* number satisfying that 2 is a primitive element of finite field $GF(k)$ (i.e. a generator of the multiplicative group of the field), and $n \le k-1$. For an k-bit vector $\boldsymbol{r}$, we denote by $\mathbb{C}_{\boldsymbol{r}}$ the $(n \times k)$ circulant matrix whose top row is represented by $\boldsymbol{r}$.

The tag and the reader pre-share a secret k-bit column vector $\boldsymbol{x}$ for authentication. Similar to HB$^{\#}$, the authentication interaction in HB-CM^{-} is only one round of three passes. First the tag chooses a random k_b-bit blinding vector $\boldsymbol{b}$ and sends it to the reader. The reader then randomly selects a k_a-bit challenge vector $\boldsymbol{a}$ and sends it to the tag. Accordingly, both sides generate an $(n \times k)$

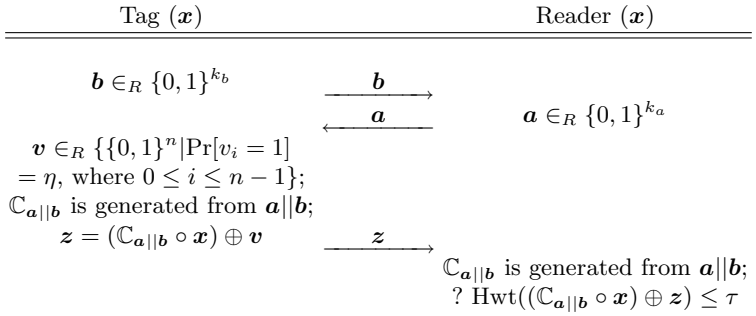

Fig. 3. HB-CM$^-$ authentication protocol, where $\boldsymbol{x}$ is a k-bit column vector, k is a prime number satisfying that 2 is a primitive element of $GF(k)$, $k_a + k_b = k$, n is the interaction expansion and $n \le k-1$, noise level $\eta \in (0, \frac{1}{2})$, pass-threshold $\tau \in (\eta n, \frac{n}{2})$, and $\mathbb{C}_{\boldsymbol{a}||\boldsymbol{b}}$ is an $(n \times k)$ circulant matrix

circulant matrix $\mathbb{C}_{\boldsymbol{a}||\boldsymbol{b}}$ from the joint-challenge vector $\boldsymbol{a}||\boldsymbol{b}$. The tag selects an n-bit random noise column vector $\boldsymbol{v}$, in which every bit independently follows the Bernoulli distribution with parameter η, and computes a response vector $\boldsymbol{z} = (\mathbb{C}_{\boldsymbol{a}||\boldsymbol{b}} \circ \boldsymbol{x}) \oplus \boldsymbol{v}$. Upon receiving $\boldsymbol{z}$, the reader checks if $\mathrm{Hwt}((\mathbb{C}_{\boldsymbol{a}||\boldsymbol{b}} \circ \boldsymbol{x}) \oplus \boldsymbol{z})$ is less than or equal to pass-threshold τ for the tag's authentication. The HB-CM$^-$ protocol is illustrated in Fig. 3.

Error Rates. Since the LPN-based protocols, including the HB-CM$^-$ protocol, use the Bernoulli noise mode, there exist two types of authentication errors. A *false negative*, that is, the authentication of a legitimate tag being rejected, takes place when the number of incorrect responses exceeds the pass-threshold τ. By contrast, a *false positive* is defined that the number of unmatched responses out of random bits is less than the pass-threshold τ. In other words, we assume that an illegitimate tag only can response with random bits. The false negative rate P_{FN} and the false positive rate P_{FP} are determined [14, 3] by $P_{\mathrm{FN}} = \sum_{i=\tau+1}^{n} \binom{n}{i} \eta^i (1-\eta)^{n-i}$ and $P_{\mathrm{FP}} = \sum_{i=0}^{\tau} \binom{n}{i} 2^{-n}$.

4 Security Results

4.1 Equivalence Class

We define the concept of equivalence class and prove several useful lemmas.

Definition 2 (Equivalence Class). *For two vectors in $\mathbf{S}^k$, say $\boldsymbol{a}$ and $\boldsymbol{b}$, if $\exists i \in \{0, \cdots, k-1\}$ such that $\boldsymbol{b} = \mathrm{Rtt}(\boldsymbol{a}, i)$, then we define that $\boldsymbol{a}$ and $\boldsymbol{b}$ are cyclically shift equivalent and they are in an equivalence class.*

An equivalence class can be represented by any one of its members.

Lemma 1. *If k is a prime number, then there are $\frac{2^k-2}{k}$ disjoint equivalence classes in $\mathbf{S}^k$. Each equivalence class contains k elements.*

Proof. An equivalence class in $\mathbf{S}^k$ has at most k elements. And any two different equivalence classes are disjoint—they do not share any common elements. Since $\mathbf{0}^{(k)}$ and $\mathbf{1}^{(k)}$ are not the elements in $\mathbf{S}^k$, every equivalence class contains at least two elements. Suppose there is an equivalence class $\boldsymbol{\Theta}$ that has less than k elements. It means that there exists at least one element $\boldsymbol{\theta'}$ satisfying Rtt($\boldsymbol{\theta'}$, i) = $\boldsymbol{\theta'}$ where $1 < i < k$ (i cannot be 1; otherwise the equivalence class only has one element). Due to the characteristic of equivalence class, the relation Rtt($\boldsymbol{\theta}$, i) = $\boldsymbol{\theta}$ holds for every element $\boldsymbol{\theta}$ in $\boldsymbol{\Theta}$. Consequently, i should be a factor of k. However, it contradicts the fact that k is prime, since k only has two factors 1 and k while $1 < i < k$. Therefore, every equivalence class of $\mathbf{S}^k$ has exact k elements, and there are $\frac{2^k-2}{k}$ disjoint equivalence classes in $\mathbf{S}^k$. □

A proof of Lemma 1 also can be found in [27]. For completeness, we present the proof above.

Lemma 2. *If k is prime and 2 is a primitive element of finite field $GF(k)$, then the polynomial $g(x) = x^{k-1} + x^{k-2} + \cdots + x + 1$ is irreducible over $GF(2)$.*

This lemma is proved in [28]. If and only if $2^i \mod k \neq 1$, $\forall 1 \leq i \leq k-2$, 2 is a primitive element of finite field $GF(k)$.

Lemma 3. *If k is a prime number satisfying that 2 is a primitive element of $GF(k)$, then any $k-1$ elements in every equivalence class of $\mathbf{S}^k$ are linearly independent.*

Proof. Let $\boldsymbol{\theta} = (\theta_0, \theta_1, \cdots, \theta_{k\text{-}1}) \in \mathbf{S}^k$, we define a $((k-1) \times k)$ matrix

$$\mathbb{D}_{\boldsymbol{\theta}} = \begin{bmatrix} \theta_0 & \theta_1 \cdots & \theta_{k-1} \\ \theta_1 & \theta_2 \cdots & \theta_0 \\ \vdots & \vdots \ddots & \vdots \\ \theta_{k-1} & \theta_0 \cdots & \theta_{k-2} \end{bmatrix}.$$

We may view $\mathbb{D}_{\boldsymbol{\theta}}$ as linear feedback shift register sequences of characteristic polynomial $x^k + 1$ over finite field $GF(2)$, according to [27]. Note that $x^k + 1 = (x+1)(x^{k-1}+x^{k-2}+\cdots+x+1)$ over $GF(2)$. Let $g(x) = x^{k-1}+x^{k-2}+\cdots+x+1$. Since k is prime, according to Lemma 1, $\boldsymbol{\theta}$ has period k. Thus, we only need to consider the following two cases.

Case 1: $\theta_0 + \theta_1 + \cdots + \theta_{k-1} = 0$

In this case, the sequence $(\theta_0, \cdots, \theta_{k-1})$ is generated by $g(x)$. Based on Lemma 2, $g(x)$ is irreducible over GF(2) if 2 is a primitive element of finite field $GF(k)$. Since the degree of $g(x)$ is equal to $k-1$, then any $k-1$ vectors in $\mathbb{D}_{\boldsymbol{\theta}}$ are linearly independent.

Case 2: $\theta_0 + \theta_1 + \cdots + \theta_{k-1} = 1$

In this case, the sequence $(\theta_0, \cdots, \theta_{k-1})$ is not generated by $g(x)$ but by polynomial $x^k + 1$. Since $x^k + 1$ has degree k, then all k vectors in $\mathbb{D}_{\boldsymbol{\theta}}$ are linearly independent.

In summary, if k is a prime number and 2 is a primitive element of $GF(k)$, then any $k-1$ elements in every equivalence class of $\mathbf{S}^k$ are linearly independent. □

Lemma 3 constitutes the technical core of this paper. In other words, Lemma 3 proves that all row vectors in an $(n \times k)$ circulant matrix $\mathbb{C}_{\boldsymbol{a}}$ are linearly independent, where $\boldsymbol{a}$ is a k-bit vector, $n < k$.

4.2 Security Definitions

To formally define security models, similar to [2, 3], we denote a tag-reader HB-CM$^-$ authentication system by a pair of probabilistic functions $(\mathcal{T}_{\boldsymbol{x},\eta,n}, \mathcal{R}_{\boldsymbol{x},n,\tau})$, namely a tag function $\mathcal{T}_{\boldsymbol{x},\eta,n}$ and a reader function $\mathcal{R}_{\boldsymbol{x},n,\tau}$. Two models are defined to discuss the protocol's security.

Definition 3 (DET-Model). *In the DET-model, which is identical to the detection-based-model used in [2, 15, 16, 3], the DET attack is carried out in two phases:*

- Phase 1*: Adversary $\mathcal{A}$ interacts q times with the honest tag $\mathcal{T}_{\boldsymbol{x},\eta,n}$. On the ith invocation, $\mathcal{T}_{\boldsymbol{x},\eta,n}$ internally generates a random blinding vector $\boldsymbol{b}_i$, takes a challenge vector $\boldsymbol{a}_i$ from $\mathcal{A}$ as input, and outputs $\boldsymbol{z}_i = (\mathbb{C}_{\boldsymbol{a}_i||\boldsymbol{b}_i} \circ \boldsymbol{x}) \oplus \boldsymbol{v}_i$ to $\mathcal{A}$. This simulates an active adversary querying the tag.*
- Phase 2*: Adversary $\mathcal{A}$ interacts with the reader $\mathcal{R}_{\boldsymbol{x},n,\tau}$, intending to impersonate the tag.*

Definition 4 (GRS-Model). *In the GRS-model, the GRS attack is carried out in two phases:*

- Phase 1*: Adversary $\mathcal{A}$ manipulates any blinding vectors from the tag $\mathcal{T}_{\boldsymbol{x},\eta,n}$ to the reader $\mathcal{R}_{\boldsymbol{x},n,\tau}$ and any challenge vectors from $\mathcal{R}_{\boldsymbol{x},n,\tau}$ to $\mathcal{T}_{\boldsymbol{x},\eta,n}$ for q executions. on the ith invocation, $\mathcal{T}_{\boldsymbol{x},\eta,n}$ internally generates a random blinding vector $\boldsymbol{b}_i$, and sends it to adversary $\mathcal{A}$. Then $\mathcal{R}_{\boldsymbol{x},n,\tau}$ receives a modified blinding vector $\boldsymbol{b}'_i$ from $\mathcal{A}$, generates a random challenge vector $\boldsymbol{a}_i$, and gives it to $\mathcal{A}$. $\mathcal{T}_{\boldsymbol{x},\eta,n}$ takes a modified challenge vector $\boldsymbol{a}'_i$ from $\mathcal{A}$, and sends $\boldsymbol{z}_i = (\mathbb{C}_{\boldsymbol{a}'_i||\boldsymbol{b}_i} \circ \boldsymbol{x}) \oplus \boldsymbol{v}_i$ to $\mathcal{R}_{\boldsymbol{x},n,\tau}$, which then checks if $\mathrm{Hwt}((\mathbb{C}_{\boldsymbol{a}_i||\boldsymbol{b}'_i} \circ \boldsymbol{x}) \oplus \boldsymbol{z}_i) \leq \tau$. If it holds, $\mathcal{R}_{\boldsymbol{x},n,\tau}$ outputs "ACCEPT" to $\mathcal{A}$; otherwise, it outputs "REJECT". This simulates a full GRS attacker. We define the interference vector $\boldsymbol{\delta}_i$ as $(\boldsymbol{b}_i \oplus \boldsymbol{b}'_i)||(\boldsymbol{a}_i \oplus \boldsymbol{a}'_i)$. To simplify the security proof for HB-CM$^-$, we rule out a trivial case of $\boldsymbol{\delta}_i = \mathbf{1}^{(k)}$.*
- Phase 2*: Adversary $\mathcal{A}$ interacts with the reader $\mathcal{R}_{\boldsymbol{x},n,\tau}$, intending to impersonate the tag.*

In the GRS-MIM-model defined in [3], the adversary is only permitted to change the challenges from the reader to the tag. So the GRS-model includes the GRS-MIM-model, and actually simulates the full GRS attack except the trivial case $\boldsymbol{\delta}_i = \mathbf{1}^{(k)}$.

In the GRS-model, since

$$(\mathbb{C}_{\boldsymbol{a}_i||\boldsymbol{b}'_i} \circ \boldsymbol{x}) \oplus \boldsymbol{z}'_i = ((\mathbb{C}_{\boldsymbol{a}_i||\boldsymbol{b}'_i} \oplus \mathbb{C}_{\boldsymbol{a}'_i||\boldsymbol{b}_i}) \circ \boldsymbol{x}) \oplus \boldsymbol{v}_i = (\mathbb{C}_{\boldsymbol{\delta}_i} \circ \boldsymbol{x}) \oplus \boldsymbol{v}_i \ ,$$

the authentication result is equivalently decided by

$$\mathrm{Hwt}((\mathbb{C}_{\boldsymbol{\delta}_i} \circ \boldsymbol{x}) \oplus \boldsymbol{v}_i) \leq \tau \ . \tag{1}$$

By replying a random vector in Phase 2 of both models, the probability that an adversary impersonating the tag will success is the false positive rate P_{FP}. This is the best soundness error we can achieve for the HB-CM$^-$ protocol. We define the advantage of an adversary $\mathcal{A}$ against HB-CM$^-$ in the DET-model and in the GRS-model as its overall success probability over P_{FP} in impersonating the tag:

$$\mathrm{Adv}_{\mathcal{A}}^{\mathrm{DET}}(k,\eta,n,\tau) \stackrel{\mathrm{def}}{=} \Pr[\boldsymbol{x} \stackrel{\$}{\leftarrow} \mathbf{S}^k, \mathcal{A}^{\mathcal{T}_{\boldsymbol{x},\eta,n}}(1^k) : \langle \mathcal{A}, \mathcal{R}_{\boldsymbol{x},n,\tau} \rangle = \mathrm{ACC}] - P_{\mathrm{FP}} \ ;$$

$$\mathrm{Adv}_{\mathcal{A}}^{\mathrm{GRS}}(k,\eta,n,\tau) \stackrel{\mathrm{def}}{=} \Pr[\boldsymbol{x} \stackrel{\$}{\leftarrow} \mathbf{S}^k, \mathcal{A}^{\mathcal{T}_{\boldsymbol{x},\eta,n},\mathcal{R}_{\boldsymbol{x},n,\tau}}(1^k) : \langle \mathcal{A}, \mathcal{R}_{\boldsymbol{x},n,\tau} \rangle = \mathrm{ACC}] - P_{\mathrm{FP}} \ .$$

If an adversary only achieves a negligible advantage against an HB-like protocol in a model, we assert that the protocol is secure in this model.

4.3 HB-CM$^-$ Security in the GRS-Model

Theorem 1. *If there exists an adversary $\mathcal{A}$ attacking the HB-CM$^-$ protocol in the GRS-model, modifying at most q executions of the protocol between an honest tag and an honest reader, running in time t, and achieving $\mathrm{Adv}_{\mathcal{A}}^{\mathrm{GRS}}(k,n,\eta,u) \geq \delta$. Then there exists an adversary $\mathcal{A}'$ attacking the HB-CM$^-$ protocol in the DET-model, interacting at most q oracle queries, running in time $O(t)$, and achieving $\mathrm{Adv}_{\mathcal{A}'}^{\mathrm{DET}}(k,n,\eta,u) \geq \delta - q\epsilon(P_{\mathrm{FP}} + \delta)$ for some negligible function ϵ. Hence, assuming HB-CM$^-$ is secure in the DET-model, HB-CM$^-$ is provably secure in the GRS-model.*

Proof. In Phase 1, $\mathcal{A}'$ can readily simulate the honest tag to $\mathcal{A}$ since $\mathcal{A}'$ has access to $\mathcal{T}_{\boldsymbol{x},\eta,n}$. The main challenge lies on how to simulate the reader $\mathcal{R}_{\boldsymbol{x},n,\tau}$. Similar to the proof method for the Random-HB$^{\#}$ protocol [3], $\mathcal{A}'$ launches Phase 1 of adversary $\mathcal{A}$, and simulates the tag and the reader for q times as follows:

1. $\mathcal{A}'$ obtains a blinding vector $\boldsymbol{b}_i$ from the tag $\mathcal{T}_{\boldsymbol{x},\eta,n}$, and sends $\boldsymbol{b}_i$ as the blinding vector of the simulated tag; $\mathcal{A}$ modifies it into $\boldsymbol{b}_i'$, and sends $\boldsymbol{b}_i'$ to the simulated reader.
2. $\mathcal{A}'$ sends a random vector $\boldsymbol{a}_i$ as the challenge of the simulated reader. $\mathcal{A}$ modifies it into $\boldsymbol{a}_i'$; $\mathcal{A}'$ forwards $\boldsymbol{a}_i'$ to the real tag.
3. The real tag responds with $\boldsymbol{z}_i = (\mathbb{C}_{\boldsymbol{a}_i' || \boldsymbol{b}_i} \circ \boldsymbol{x}) \oplus \boldsymbol{v}_i$ to $\mathcal{A}'$, which uses it as the answer of the simulated tag to the simulated reader.
4. Recall $\boldsymbol{\delta}_i = (\boldsymbol{b}_i \oplus \boldsymbol{b}_i') || (\boldsymbol{a}_i \oplus \boldsymbol{a}_i')$. If $\boldsymbol{\delta}_i = \mathbf{0}^{(k)}$, $\mathcal{A}'$ outputs "ACCEPT" to $\mathcal{A}$ as the authentication result of the simulated reader; otherwise, it outputs "REJECT".

After Phase 1, $\mathcal{A}'$ launches Phase 2 of $\mathcal{A}$. Since Phase 2 in the DET-model is identical to that in the GRS-model, $\mathcal{A}'$ just replicates $\mathcal{A}$'s behavior with the real reader, perfectly simulating the tag $\mathcal{T}_{\boldsymbol{x},\eta,n}$. Therefore, if $\mathcal{A}$ achieves $\mathrm{Adv}_{\mathcal{A}}^{\mathrm{GRS}}(k,n,\eta,u) \geq \delta$, then the probability of $\mathcal{A}'$ successfully impersonating a valid tag is the same as the success probability of $\mathcal{A}$, i.e. $P_{\mathrm{FP}} + \delta$, on the condition that the reader is correctly simulated by $\mathcal{A}'$ in Phase 1.

Now we need to calculate the probability of $\mathcal{A}'$ successfully simulating the reader for $\mathcal{A}$ in Phase 1. Consider one execution of the protocol in Phase 1, based on (1). When $\boldsymbol{\delta}_i = \mathbf{0}^{(k)}$, $\mathcal{A}'$ fails at simulating the reader with a probability equal to the false negative rate P_{FN}. For the case of $\boldsymbol{\delta}_i \neq \mathbf{0}^{(k)}$, since we suppose $\boldsymbol{\delta}_i \neq \mathbf{1}^{(k)}$, thus $\boldsymbol{\delta}_i \in \mathbf{S}_k$. According to Lemma 3, all the row vectors in the $(n \times k)$ circular matrix $\mathbb{C}_{\boldsymbol{\delta}_i}$ are linearly independent. Let $\boldsymbol{d}_i = \mathbb{C}_{\boldsymbol{\delta}_i} \circ \boldsymbol{x}$ denote the error vector added by $\mathcal{A}$. Following the same argument in Theorem 2 of [15], $\boldsymbol{d}_i$ is uniformly distributed over $\{0,1\}^n$, as the row vectors of $\mathbb{C}_{\boldsymbol{\delta}_i}$ are linearly independent. Since the error vector $\boldsymbol{d}_i$ is added before the Bernoulli noises are added by the tag, $\boldsymbol{v}_i$ is independent of $\boldsymbol{d}_i$. Thus the resulting error vector $\boldsymbol{d}_i \oplus \boldsymbol{v}_i$ follows the uniform distribution over $\{0,1\}^n$. As a result, the probability of $\mathcal{A}'$ wrongly outputting "REJECT" is exactly the same as the false positive rate P_{FP}. Overall, $\mathcal{A}'$ fails at simulating the reader in one execution at most with probability $\epsilon = \max(P_{\mathrm{FN}}, P_{\mathrm{FP}})$. The probability of $\mathcal{A}'$ correctly simulating the reader in Phase 1 would be not less than $1-q\epsilon$, and adversary $\mathcal{A}'$ can impersonate a valid tag with success probability not less than $(P_{\mathrm{FP}} + \delta)(1 - q\epsilon)$. Therefore, $\mathcal{A}'$ can achieve advantage

$$\mathrm{Adv}^{\mathrm{DET}}_{\mathcal{A}'}(k, \eta, n, \tau) \geq (P_{\mathrm{FP}} + \delta)(1 - q\epsilon) - P_{\mathrm{FP}} = \delta - q\epsilon(P_{\mathrm{FP}} + \delta) \ .$$

If δ is non-negligible, then $q\epsilon(P_{\mathrm{FP}}+\delta) \leq \frac{\delta}{2}$ for k big enough, and $\mathrm{Adv}^{\mathrm{DET}}_{\mathcal{A}'}(k, \eta, n, \tau) \geq \frac{\delta}{2}$ is non-negligible. Thus if HB-CM$^-$ is secure in the DET-model, HB-CM$^-$ is secure in the GRS-model. In other words, if HB-CM$^-$ is vulnerable to the GRS attack, then it must be vulnerable to the DET attack. □

4.4 HB-CM$^-$ Security in the DET-Model

Similar to HB$^{\#}$, we cannot present a strict reduction from the LPN problem to HB-CM$^-$ security in the DET-model currently. Instead, we conjecture that the HB-CM$^-$ protocol is secure in the DET-model.

Claim 1 *In the DET-model, the HB-CM$^-$ protocol is as secure as the parallel HB$^+$ protocol. Moreover, the HB-CM$^-$ protocol with key length k achieves the same security level d as the HB$^+$ protocol with key length $k + d$, if $n \geq d$.*

Justification. Let's recall the parallel HB$^+$ protocol, which is provably secure in the DET-model [2, 15]. The tag first generates a random $(n \times k_b)$ blinding matrix $\mathbb{B}$ and sends it to the reader; then the reader selects an $(n \times k_a)$ challenge matrix $\mathbb{A}$ at random. After receiving $\mathbb{A}$, the tag computes and sends the n-bit response vector $\boldsymbol{z} = ((\mathbb{A}||\mathbb{B}) \circ \boldsymbol{x}) \oplus \boldsymbol{v}$. Let $\boldsymbol{x}^{(a)}||\boldsymbol{x}^{(b)} = \boldsymbol{x}$, where vector $\boldsymbol{x}^{(a)}$ is of k_a bits and vector $\boldsymbol{x}^{(b)}$ is of k_b bits. Then the response vector is equivalently computed by

$$\boldsymbol{z} = (\mathbb{A} \circ \boldsymbol{x}^{(a)}) \oplus (\mathbb{B} \circ \boldsymbol{x}^{(b)}) \oplus \boldsymbol{v} \ . \tag{2}$$

As for the HB-CM$^-$ protocol, we define $\widehat{\boldsymbol{a}} = \boldsymbol{a}||\mathbf{0}^{(k_b)}$, and $\widehat{\boldsymbol{b}} = \mathbf{0}^{(k_a)}||\boldsymbol{b}$. Then the response vector is equivalently computed by

$$\boldsymbol{z} = (\mathbb{C}_{\widehat{\boldsymbol{a}}} \circ \boldsymbol{x}) \oplus (\mathbb{C}_{\widehat{\boldsymbol{b}}} \circ \boldsymbol{x}) \oplus \boldsymbol{v} \ . \tag{3}$$

In Theorem 2 of [15], to prove the parallel HB$^+$ protocol's security in the DET-model, it requires 2^{n-k_a} to be negligible such that the row vectors in the random challenge matrix $\mathbb{A}$ are linearly independent. Interestingly, the security proof of Random-HB$^\#$ in the GRS-MIM-model (Theorem 2 of [3]) demands 2^{k_a-n} to be negligible. In contrast, the HB-CM$^-$ protocol does not encounter such limitations. The row vectors in circular matrix $\mathbb{C}_{\widehat{\boldsymbol{a}}}$ are linearly independent with the overwhelming probability of $1-(\frac{1}{2})^{k_a}$, because $\widehat{\boldsymbol{a}} \in \mathbf{S}_k$ if $\boldsymbol{a} \neq \mathbf{0}^{(k_a)}$. Similarly, the row vectors in circular matrix $\mathbb{C}_{\widehat{\boldsymbol{b}}}$ are linearly independent with the probability of $1-(\frac{1}{2})^{k_b}$.

The HB-CM$^-$ and parallel HB$^+$ protocols are very alike except that matrix $(\mathbb{A}||\mathbb{B})$ in parallel HB$^+$ is random and $\mathbb{C}_{\boldsymbol{a}||\boldsymbol{b}}$ in HB-CM$^-$ is a circulant matrix. Since the major operation in the authentication is inner product, the linear independence of the row vectors in $\mathbb{C}_{\boldsymbol{a}||\boldsymbol{b}}$ actually enhances the protocol's security.

Let's compare (2) for HB$^+$ with (3) for HB-CM$^-$ side by side for security level. In Phase 1, the adversary can freely choose the challenges. When setting $\mathbb{A} = \mathbf{0}^{(n\times k_a)}$ in (2) against HB$^+$, the adversary can get an LPN instance $(\mathbb{B}, \boldsymbol{z} = (\mathbb{B} \circ \boldsymbol{x}^{(b)}) \oplus \boldsymbol{v})$. Therefore, as concluded in [14], the hardness of HB$^+$ against a DET adversary only relies on the k_b-bit LPN instances. In contrast, even if choosing $\boldsymbol{a} = \mathbf{0}^{(k_a)}$ in (3) against HB-CM$^-$, the adversary obtains an LPN instance $(\mathbb{C}_{\widehat{\boldsymbol{b}}}, \boldsymbol{z} = ((\mathbb{C}_{\widehat{\boldsymbol{b}}} \circ \boldsymbol{x}) \oplus \boldsymbol{v})$. Consequently, the hardness of HB-CM$^-$ depends on the k-bit LPN instances as long as $n \geq k_a$, no matter what challenges the adversary chooses. Similarly, as for Phase 2, in which the adversary can choose arbitrary blinding vectors, k_a in HB$^+$ has to be at least d in order to guarantee d-bit security [14], while the adversary against HB-CM$^-$ is confronting $\mathbb{C}_{\widehat{\boldsymbol{a}}} \circ \boldsymbol{x}$, which actually provides $(k_a + \min(k_b, d))$-bit security.

Therefore, HB-CM$^-$ is at least as secure as the parallel HB$^+$ protocol in the DET-model. In addition, if $n \geq d$ (which is always true in practice), the HB-CM$^-$ protocol with key length k achieves the same security level d as the HB$^+$ protocol with key length $k + d$. □

Provable Security Paradox. The LPN problem forms the security foundation for the HB-family authentication protocols. In order to formally prove an LPN-based protocol's security in a certain model, one has to provide the complete reduction from the LPN problem to the protocol. The reduction procedures from sequential HB to sequential HB$^+$ in [2] and from MHB-puzzle to Random-HB$^\#$ in [3], as part of the full security proofs in the DET-model, are elegant and impressive; they successfully create security equivalence between two primitives. However, to eventually reduce those protocols to the LPN problem, some interesting point emerges as we compare their proof mechanisms to the GRS attack. The method used in [2] is essentially a concrete version of Blum et al.'s asymptotic reduction strategy from [29]; and the key technical lemma used for security proofs of parallel HB and HB$^+$ in [15, 16] is essentially in Section 4 of [10]. The principal technique behind all of those is same, described as follows. Given a sequence of LPN instances $(\boldsymbol{a}_i, z_i = (\boldsymbol{a}_i \circ \boldsymbol{x}) \oplus v_i)$, the adversary, which is required to solve the LPN problem, changes to $(\widehat{\boldsymbol{a}}_i = \boldsymbol{a}_i \oplus \boldsymbol{\theta}, z_i)$ by a random

$\boldsymbol{\theta}$, and feeds $(\widehat{\boldsymbol{a}_i}, z_i)$ to a protocol attacker. If the protocol attacker achieves non-negligible advantage against the protocol using the modified LPN instances, then the adversary can learn $\boldsymbol{\theta} \circ \boldsymbol{x} = 0$; otherwise, $\boldsymbol{\theta} \circ \boldsymbol{x} = 1$. After repeating many times, the adversary can fully recover $\boldsymbol{x}$, solving the LPN problem. Ironically, this technique is a variant of the GRS attack—the security proof procedures [2, 15, 16] that can finally reduce those protocols (HB, Sequential HB$^+$, Parallel HB$^+$) to the LPN problem rely on the fact that the protocols are vulnerable to the GRS attack.

It reminds us of the provable security paradox in the first public-key system with reductionist security. In 1979, Rabin [30] introduced an encryption function that could be proved to be invertible only by someone who could solve the integer factorization problem. However, as Rivest (see [31]) pointed out, the very feature that provides such a security proof would lead to total break if the Rabin encryption system was confronted with the chosen-ciphertext attacker [32].

As for the Random-HB$^\#$ protocol, the security proof [3] in the DET-model is based on the fact that the security proof in [2] implies, as noticed by Katz and Shin [15], that the one-round HB protocol in the passive model is a $(1-\frac{1}{2})$-hard weakly verifiable puzzle (HB puzzle). The authors [3] indirectly reduce the security of MHB-puzzle to the LPN problem by a security induction between MHB-puzzle to the perfect n-fold repetition of the HB-puzzle. However, if we review the MHB-puzzle directly, since it can resist the GRS-attack, the original reduction characteristic from the LPN problem in the HB-puzzle has vanished in the MHB-puzzle. Nevertheless, the argument above does not undermine the outstanding security reduction of [2, 15, 3]. As the security proofs of most *practically employed* public-key systems count on a strong, implausible random oracle model, we do not have such luxury in the LPN-based lightweight authentication protocols. It would be very interesting to design an LPN-based authentication protocol that can be provably reduced to the LPN problem while thwarting all man-in-the-middle attacks.

5 New Noise Mode Preventing OOV Attack

5.1 OOV Attack

At this time, all HB-like protocols (HB$^+$, Random-HB$^\#$, HB$^\#$, and HB-CM$^-$) can be generalized into the following prototype protocol. The tag and the reader hold a secret $(k \times n_2)$ matrix $\mathbb{X}$ for authentication. During the ith protocol procedure, the two sides work together to contribute to a joint-challenge $(n_1 \times k)$ matrix $\mathbb{M}_i$. Let $n = n_1 n_2$. The tag outputs the response matrix $\mathbb{Z}_i = (\mathbb{M}_i \circ \mathbb{X}) \oplus \mathbb{V}_i$, where $\mathbb{V}_i$ is an $(n_1 \times n_2)$ noise matrix, of which each element independently follows the Bernoulli distribution with parameter η, i.e. Bernoulli noise mode. The reader accepts the authentication providing the value of $\mathrm{Hwt}((\mathbb{M}_i \circ \mathbb{X}) \oplus \mathbb{Z}_i)$ does not exceed a threshold τ.

At AsiaCrypt 2008, Ouafi, Overbeck and Vaudenay [5] presented a general man-in-the-middle attack (OOV attack) against all current HB-like protocols, including the HB-CM$^-$ protocol, with the Bernoulli noise mode.

The basic OOV attack is conducted as follows. The attacker first eavesdrops one successful execution of the protocol, obtaining a pair $(\widehat{\mathbb{M}}, \widehat{\mathbb{Z}})$ satisfying $\widehat{\mathbb{Z}} = (\widehat{\mathbb{M}} \circ \mathbb{X}) \oplus \widehat{\mathbb{V}}$ and $\mathrm{Hwt}(\widehat{\mathbb{V}}) \leq \tau$. Then the attacker manipulates many executions of the protocol by xoring the $(\mathbb{M}_i, \mathbb{Z}_i)$ with $(\widehat{\mathbb{M}}, \widehat{\mathbb{Z}})$; thus the authentication result is actually decided by whether $\mathrm{Hwt}(\mathbb{V}_i \oplus \widehat{\mathbb{V}}) \leq \tau$. Based on the overall success probability, the attacker can calculate the exact value of $\mathrm{Hwt}(\widehat{\mathbb{V}})$. After that, the attacker changes $\widehat{\mathbb{Z}}$ by one bit to $\widehat{\mathbb{Z}}'$, uses $(\widehat{\mathbb{M}}, \widehat{\mathbb{Z}}')$ to interrupt with many executions of the protocol, and get the result of $\mathrm{Hwt}(\widehat{\mathbb{V}}')$, where $\widehat{\mathbb{V}}' = (\widehat{\mathbb{M}} \circ \mathbb{X}) \oplus \widehat{\mathbb{Z}}'$. By comparing the values of $\mathrm{Hwt}(\widehat{\mathbb{V}})$ and $\mathrm{Hwt}(\widehat{\mathbb{V}}')$, the attacker can figure out one bit of noise matrix $\widehat{\mathbb{V}}$. Repeating this process, the attacker can obtain the error-free value of $\widehat{\mathbb{M}} \circ \mathbb{X}$. The attacker collects enough equations that he can completely recover the secret $\mathbb{X}$, breaking the protocol.

In addition, Ouafi, Overbeck and Vaudenay [5] demonstrated another simple man-in-the-middle attack (OOV2 attack) against a protocol with a special noise mode. Since an HB-like protocol with the Bernoulli noise mode would incur a certain false negative probability, a natural method to overcome that drawback is to demand the tag to generate a noise $\mathbb{V}_i$ of bounded Hamming weight, that is $\mathrm{Hwt}(\mathbb{V}_i) \leq \tau$, as discussed in [15, 3]. We refer to it as the upper-bounded Binomial noise mode. For one iteration $(\mathbb{M}_i, \mathbb{Z}_i)$ of the prototype protocol with that noise mode, an OOV2 attacker manipulates the response such that the reader receives $\mathbb{Z}_i \oplus \mathbb{D}_i$ rather than $\mathbb{Z}_i$, where $\mathbb{D}_i$ is a random matrix of Hamming weight 2. Let $w_i = \mathrm{Hwt}((\mathbb{M}_i \circ \mathbb{X}) \oplus \mathbb{Z}_i)$ denote the Hamming weight of the noise added by the tag. If and only if $w = \tau - 1$ or τ and the attacker flipped two non-erroneous bits[2] in the response, the reader rejects the authentication. In other words, in the case of a rejection, the attackers learn two bits about $\mathbb{X}$. Assume that the rejection takes place with probability p, the attacker can recover the $(k \times n)$-bit matrix $\mathbb{X}$ with $\frac{kn}{2p}$ trials.[3]

5.2 New Noise Mode

We propose a new noise mode for all HB-like protocols to thwart the OOV attack. A protocol with this noise mode will naturally be false-negative-free, while robust to the OOV2 attack. Let $t = \lfloor \eta n \rfloor$ in the prototype protocol. The tag generates a noise $\mathbb{V}_i$ such that $\mathrm{Hwt}(\mathbb{V}_i)$ is equal to t or $t + 1$. The reader accepts an instance $(\mathbb{M}_i, \mathbb{Z}_i)$ if and only if $\mathrm{Hwt}((\mathbb{M}_i \circ \mathbb{X}) \oplus \mathbb{Z}_i) = t$ or $t + 1$.

Let's review the OOV attack strategy in the context of the proposed noise mode. Algorithm 3 of [5], an important optimization for the OOV basic attack, no longer takes effect. Thus, the attacker only can launch the basic attack. If an OOV attacker uses $(\mathbb{M}', \mathbb{Z}')$ satisfying $\mathrm{Hwt}((\mathbb{M}' \circ \mathbb{X}) \oplus \mathbb{Z}') = w$ to interfere

[2] The two non-erroneous bits come from the only two non-zero elements in $\mathbb{D}_i$.

[3] The equation is accidently typed as $\frac{kn2}{p}$ in [5].

with the authentication of the prototype protocol with the new noise mode, the success probability will be

$$p(w) = \frac{\binom{w}{\lfloor \frac{w}{2} \rfloor}\binom{n-w}{t-\lfloor \frac{w}{2} \rfloor} + \binom{w}{\lceil \frac{w}{2} \rceil}\binom{n-w}{t+1-\lceil \frac{w}{2} \rceil}}{\binom{n}{t} + \binom{n}{t+1}} = \frac{\binom{w}{\lfloor \frac{w}{2} \rfloor}\left(\binom{n-w}{t-\lfloor \frac{w}{2} \rfloor} + \binom{n-w}{t+1-\lceil \frac{w}{2} \rceil}\right)}{\binom{n+1}{t+1}} \quad . \tag{4}$$

Through an observation of one iteration of the protocol, the attacker obtains a pair $(\widehat{\mathbb{M}}, \widehat{\mathbb{Z}})$ such that $\mathrm{Hwt}((\widehat{\mathbb{M}} \circ \mathbb{X}) \oplus \widehat{\mathbb{Z}}) = w$, where $w = t$ or $t+1$. As a result, for $w = [0, t]$, the probability of the attacker correctly generating such a pair $(\mathbb{M}', \mathbb{Z}')$ is not greater than $\frac{\binom{t}{t-w}}{\binom{n}{t-w}}$. Therefore, overall, the success probability of an OOV manipulation authentication

$$P_{\mathrm{OOV}} \leq \max_{0 \leq w \leq t} \left(\frac{\binom{t}{w}\binom{w}{\lfloor \frac{w}{2} \rfloor}\left(\binom{n-w}{t-\lfloor \frac{w}{2} \rfloor} + \binom{n-w}{t+1-\lceil \frac{w}{2} \rceil}\right)}{\binom{n}{t-w}\binom{n+1}{t+1}} \right) \quad . \tag{5}$$

For specific protocols, we can carefully choose practical parameters (n, t) such that P_{OOV} is negligible[4]. Therefore, the OOV attack cannot succeed under polynomial bounds. As for the OOV2 attack, in which the attacks flips k bits of $\mathbb{Z}_i$ for one iteration of protocol, the proposed noise mode prevents the attacker from learning a firm equation $(\mathbb{M}_i \circ \mathbb{X})[j] = 1$ or 0 from the authentication result. Thus the OOV2 attack does not work either.

After changing the Bernoulli noise mode to the new noise mode, the reduction procedures for all HB-like protocols, including Theorem 1 for HB-CM$^-$, still hold. Even better, the inherent zero false negative bridge a security proof gap in many reduction proofs, such as Theorem 2 of [3], which require the false negative rate P_{FN} to be negligible, while a non-negligible $P_{\mathrm{FN}} = 2^{-40}$ is recommended for practical concerns in [3]. However the foundational problem behind an HB-like protocol with the new noise mode is no longer the original LPN problem.

Some may suggest a noise mode in which the Hamming weight of noise vectors is a constant integer t. However, the LPN variant problem on this noise mode is not hard at all. In fact, it would be as easy as the case of noise-free. From each instance $(\mathbb{A}_i, \boldsymbol{z}_i = (\mathbb{A}_i \circ \boldsymbol{x}) \oplus \boldsymbol{v}_i)$ where $\mathrm{Hwt}(\boldsymbol{v}_i) = t$, an attacker, as pointed out in [5], can learn

$$\bigoplus_{j=0}^{n-1} (\mathbb{A}_i \circ \boldsymbol{x})[j] = \bigoplus_{j=0}^{n-1} \boldsymbol{z}_i[j] \oplus \begin{cases} 1 & \text{if } t \text{ is odd} \\ 0 & \text{if } t \text{ is even} \end{cases} \quad .$$

After gathering $O(k)$ instances, the attacker can completely recover $\boldsymbol{x}$.

5.3 HB-CM Protocol

Our final HB-CM protocol, which is essentially the HB-CM$^-$ protocol with the new noise mode, is depicted in Fig. 4.

[4] That is, $P_{\mathrm{OOV}} \leq 2^{-d}$.

Tag ($\boldsymbol{x}$)		Reader ($\boldsymbol{x}$)
$\boldsymbol{b} \in_R \{0,1\}^{k_b}$	$\xrightarrow{\boldsymbol{b}}$	
	$\xleftarrow{\boldsymbol{a}}$	$\boldsymbol{a} \in_R \{0,1\}^{k_a}$
$\boldsymbol{v} \in_R \{\{0,1\}^n \mid \mathrm{Hwt}(\boldsymbol{v}) \in [t, t+1]\}$; $\mathbb{C}_{\boldsymbol{a}\|\|\boldsymbol{b}}$ is generated from $\boldsymbol{a}\|\|\boldsymbol{b}$; $\boldsymbol{z} = (\mathbb{C}_{\boldsymbol{a}\|\|\boldsymbol{b}} \circ \boldsymbol{x}) \oplus \boldsymbol{v}$	$\xrightarrow{\boldsymbol{z}}$	
		$\mathbb{C}_{\boldsymbol{a}\|\|\boldsymbol{b}}$ is generated from $\boldsymbol{a}\|\|\boldsymbol{b}$; ? $\mathrm{Hwt}((\mathbb{C}_{\boldsymbol{a}\|\|\boldsymbol{b}} \circ \boldsymbol{x}) \oplus \boldsymbol{z}) \in [t, t+1]$

Fig. 4. HB-CM authentication protocol, where $\boldsymbol{x}$ is a k-bit binary vector, k is a prime number satisfying that 2 is a primitive element of $GF(k)$, $k_a + k_b = k$, n is the interaction expansion and $n \leq k-1$, noise level $\eta \in (0, \frac{1}{2})$, $t = \lfloor \eta n \rfloor$, and $\mathbb{C}_{\boldsymbol{a}\|\|\boldsymbol{b}}$ is an $(n \times k)$ circulant matrix

The false positive rate in the HB-CM protocol is computed by

$$P'_{\mathrm{FP}} = \sum_{i=t}^{t+1} \binom{n}{i} 2^{-n} = \binom{n+1}{t+1} 2^{-n} \ .$$

The fundamental problem behind the HB-CM protocol is defined as follows.

Definition 5 (LPN-CM Problem). *Let k is a prime number such that 2 is a primitive element of $GF(k)$ and $n < k$. Let $\boldsymbol{x}$ be a random k-bit vector, $t \in (0, \frac{n}{2} - 1)$ be a integer noise degree. For $i \in [0, q-1]$ (q is polynomial in k), let $\boldsymbol{a}_i$ be a random k-bit vector, and $\boldsymbol{v}_i$ be a random n-bit vector such that $\mathrm{Hwt}(\boldsymbol{v}_i) = t$ or $t+1$. Given q pairs $(\boldsymbol{a}_i, \boldsymbol{z}_i = (\mathbb{C}_{\boldsymbol{a}_i} \circ \boldsymbol{x}) \oplus \boldsymbol{v}_i)$, recover the secret vector $\boldsymbol{x}$.*

Let's compare the LPN-CM problem with the LPN problem. The requirement of k be prime in LPN-CM is trivial according to the prime number theorem, which describes the asymptotic distribution of the prime numbers. Even though there is no deterministic number theory result about the distribution of a special class of prime number k satisfying 2 is a primitive element of $GF(k)$, we can conjecture that this kind of integers are mass-distributed and this requirement would not degrade the hardness of LPN-CM compared to LPN. We list all prime numbers between 512 and 1024 satisfying that property: 523, 541, 547, 557, 563, 587, 613, 619, 653, 659, 661, 677, 701, 709, 757, 773, 787, 797, 821, 827, 829, 853, 859, 877, 883, 907, 941, 947, and 1019. Due to the remarkable feature of linear independence of row vectors in the circular matrix, it seems reasonable to assume that the circular-matrix-encoded challenge matrices in LPN-CM are as secure as random matrices in LPN. The main concern left is the new noise mode. For the time being, we presume that the LPN-CM and LPN problems, with the identical key length k and noise level η, provide the equivalent security level.

The security of the HB-CM protocol depends on the following problem.

Definition 6 (HB-CM Problem). *An adversary against the HB-CM protocol tries to achieve non-negligible advantage in the DET-model.*

Assume hardness of the HB-CM problem, the HB-CM protocol resists all known man-in-the middle attacks. How to reduce the HB-CM problem to the LPN-CM problem remains open.

6 Protocol Parameters and Performance

The two parameters k (key length) and η (noise level) dominate the security level d of LPN-based authentication protocols. The desirable value of k is a reflection of the running time of an algorithm solving LPN instances[12–14]. At present, the best algorithm is the LF algorithm, presented by Levieil and Fouque [14]. Based on the assumption of security equivalence between LPN-CM and LPN, we directly adopt the parameter settings recommended in [14] for the HB-CM protocol with an additional concern. Since we exclude the case of $\boldsymbol{\delta}_i = \mathbf{1}^{(k)}$ in the GRS-model, and the interference vector $\mathbf{1}^{(k)}$ in the GRS attack does reveal the parity of $\boldsymbol{x}$, key length k in the HB-CM protocol should be increased by 1. On the other hand, in order to resist the OOV attack, we should choose parameters $(n, t = \lfloor \eta n \rfloor)$ satisfying $P_{\mathrm{OOV}} \leq 2^{-d}$. Therefore, to achieve 80-bit security, we recommend the parameters $k = 859, n = 858, t = 108$ for the HB-CM protocol. Then $P_{\mathrm{OOV}} \leq 2^{-82}, P'_{\mathrm{FP}} \leq 2^{-386}$. Furthermore, we choose $k_a = 256, k_b = 603$.

As for HB$^{\#}$ with the proposed noise mode, to resist the OOV attack, we should use the same n and t. In comparison to HB-CM, to achieve the same security level, it seems that HB$^{\#}$ requires $k_a^{\#} = 80, k_b^{\#} = 858$ according to Claim 1. However, the selection of $k = 859$ in the HB-CM protocol is mainly determined by the limitation of $n < k$. Therefore, we can choose lesser $k_b^{\#}$, which is decided by the LF algorithm [14]. We select $k_b^{\#} = 640$ to guarantee 80-bit LPN instances hardness for $\delta = 0.125$. Consequently, the memory consumptions for HB$^{\#}$ and revised HB$^{\#}$ are $k_a^{\#}+k_b^{\#}+2n-2 = 2434$ bits and $k_a^{\#}+k_b^{\#}+n-1 = 1577$ bits respectively.

To further reduce the storage consumption, we may loosen the requirement to $P_{\mathrm{OOV}} \leq 2^{-40}$, which, of course, needs additional security management against OOV attack, that an authentication key should be revoked if the failed authentication number exceeds 2^{40}. Therefore, we can choose $n = 522, t = 130, k = 523, k_a = 120, k_b = 403$ for HB-CM, and choose $n = 522, t = 130, k_a^{\#} = 80, k_b^{\#} = 522$ for HB$^{\#}$. The performance comparison is summarized in Table 1.

7 Conclusion

In summary, this paper presents the innovative HB-CM^{-} entity authentication protocol, which surpasses the HB$^{\#}$ protocol in the terms of storage cost and provable security against the full GRS attack. We develop the concept of equivalence class, and prove the linearly independence of row vectors in a special class of circulant matrices, which forms the security and efficiency foundation for the

Table 1. Performance comparison between HB-CM and HB$^{\#}$

Protocol	Storage (bits)		Transmission (Bits)	
	$P_{\mathrm{OOV}} \leq 2^{-80}$	$P_{\mathrm{OOV}} \leq 2^{-40}$	$P_{\mathrm{OOV}} \leq 2^{-80}$	$P_{\mathrm{OOV}} \leq 2^{-40}$
HB-CM	859	523	1717	1045
HB$^{\#}$	2434	1644	1578	1124
Revised HB$^{\#}$	1577	1123	1578	1124

HB-CM^{-} protocol. In addition, we introduce a new noise mode, which helps all HB-like protocols resist the OOV attack. We define several problems regarding the security of the final HB-CM protocol, which is the HB-CM^{-} protocol with the new noise mode, to stimulate further research.

References

1. Peris-Lopez, P., Hernandez-Castro, J., Estevez-Tapiador, J., Ribagorda, A.: M^{2}AP: A Minimalist Mutual-Authentication Protocol for Low-Cost RFID Tags. In: Ubiquitous Intelligence and Computing. LNCS 4159 (2006) 912–923
2. Juels, A., Weis, S.A.: Authenticating Pervasive Devices with Human Protocols. In: Advances in Cryptology CRYPTO 2005, Updated version available at: http://www.rsasecurity.com/rsalabs/staff/bios/ajuels/publications/pdfs/lpn.pdf. LNCS 3621 (2005) 293–308
3. Gilbert, H., Robshaw, M.J., Seurin, Y.: HB$^{\#}$: Increasing the Security and Efficiency of HB^{+}. In: Advances in Cryptology EUROCRYPTO 2008, Full version available at: Cryptology ePrint Archive: Report 2008/028. (2008)
4. Gilbert, H., Robshaw, M., Sibert, H.: An Active Attack Against HB^{+} - A Provably Secure Lightweight Authentication Protocol. Technical report, Cryptology ePrint Archive: Report 2005/237 (2005)
5. Ouafi, K., Overbeck, R., Vaudenay, S.: On the Security of HB$^{\#}$ against a Man-in-the-Middle Attack. In: Advances in Cryptology - ASIACRYPT 2008. LNCS 5350 (2008) 108–124
6. MacWilliams, F., Sloane, N.: The Theory of Error-Correcting Codes. North-Holland Amsterdam (1977)
7. Berlekamp, E.R., McEliece, R.J., van Tilborg, H.C.: On the inherent intractability of certain coding problems. IEEE Transactions on Information Theory **24**(3) (1978) 384–386
8. Crawford, J.M., Kearns, M.J.: The Minimal Disagreement Parity Problem as a Hard Satisfiability Problem. Computational Intelligence Research Laboratory and AT&T Bell Labs, Available at http://www.cs.cornell.edu/selman/docs/crawford-parity.pdf (1995)
9. Håstad, J.: Some optimal inapproximability results. In: Proceedings of the twenty-ninth annual ACM symposium on Theory of computing, El Paso, Texas, United States (1997)
10. Regev, O.: On lattices, learning with errors, random linear codes, and cryptography. In: Proceedings of the thirty-seventh annual ACM symposium on Theory of computing, Baltimore, MD, USA, ACM (2005)
11. Hopper, N., Blum, M.: Secure Human Identification Protocols. In: Advances in Cryptology - ASIACRYPT 2001. LNCS 2248 (2001) 52–66

12. Blum, A., Kalai, A., Wasserman, H.: Noise-tolerant learning, the parity problem, and the statistical query model. Journal of the ACM (JACM) **50**(4) (2003) 506–519
13. Fossorier, M., Mihaljevi, M., Imai, H., Cui, Y., Matsuura, K.: An Algorithm for Solving the LPN Problem and Its Application to Security Evaluation of the HB Protocols for RFID Authentication. In: Progress in Cryptology - INDOCRYPT 2006. LNCS 4329 (2006) 48–62
14. Levieil, E., Fouque, P.A.: An Improved LPN Algorithm. In: Security and Cryptography for Networks. LNCS 4116 (2006) 348–359
15. Katz, J., Shin, J.: Parallel and Concurrent Security of the HB and HB+ Protocols. In: Advances in Cryptology - EUROCRYPT 2006. LNCS 4004 (2006) 73–87
16. Katz, J., Smith, A.: Analyzing the HB and HB$^+$ Protocols in the "Large Error" Case. Technical report, Cryptology ePrint Archive, Report 2006/326 (2006)
17. Bringer, J., Chabanne, H., Dottax, E.: HB^{++}: a Lightweight Authentication Protocol Secure against Some Attacks. In Chabanne, H., ed.: Second International Workshop on Security, Privacy and Trust in Pervasive and Ubiquitous Computing (SecPerU 2006). (2006) 28–33
18. Duc, D.N., Kim, K.: Securing HB$^+$ Against GRS Man-in-the-Middle Attack. In: Proceedings of Symposium on Cryptography and Information Security (SCIS 2007), Sasebo, Japan (2007)
19. Munilla, J., Peinado, A.: HB-MP: A further step in the HB-family of lightweight authentication protocols. Computer Networks **51**(9) (2007) 2262–2267
20. Piramuthu, S.: HB and Related Lightweight Authentication Protocols for Secure RFID Tag/Reader Authentication. In: CollECTeR Europe Conference. (2006)
21. Leng, X., Mayes, K., Markantonakis, K.: HB-MP+ Protocol: An Improvement on the HB-MP Protocol. In: IEEE International Conference on RFID. (2008) 118–124
22. Gilbert, H., Robshaw, M.J., Seurin, Y.: Good Variants of HB$^+$ are Hard to Find. In: Financial Crypt 2008. (2008)
23. Hammouri, G., Sunar, B.: PUF-HB: A Tamper-Resilient HB Based Authentication Protocol. In: Applied Cryptography and Network Security. LNCS 5037 (2008) 346–365
24. Bringer, J., Chabanne, H.: Trusted-HB: A Low-Cost Version of HB+ Secure Against Man-in-the-Middle Attacks. IEEE Transactions on Information Theory **54**(9) (2008) 4339–4342
25. Frumkin, D., Shamir, A.: Un-Trusted-HB: Security Vulnerabilities of Trusted-HB. In: The 5th Workshop on RFID Security (RFIDSec 09). (2009)
26. Daemen, J., Rijmen, V.: The Design of Rijndael: AES - The Advanced Encryption Standard. Springer-Verlag (2002)
27. Golomb, S.W., Gong, G.: Signal Design for Good Correlation: For Wireless Communication, Cryptography, and Radar. Cambridge University Press (2004)
28. Lidl, R., Niederreiter, H.: Finite Fields (Encyclopedia of Mathematics and its Applications). Volume 20. Addison-Wesley, (Revised version, Cambridge University Press, 1997.) (1983)
29. Blum, A., Furst, M., Kearns, M., Lipton, R.: Cryptographic Primitives Based on Hard Learning Problems. In: Advances in Cryptology - CRYPTO' 93. LNCS 773 (1994) 278–291
30. Rabin, M.C.: Digitalized signatures and public-key functions as intractable as factorization. Technical report, Technical Report LCS/TR-212, MIT Lab. for Computer Science (1979)
31. Williams, H.C.: A modification of the RSA public-key encryption procedure. IEEE Transactions on Information Theory **26**(6) (1980) 726–729

32. Koblitz, N., Menezes, A.J.: Another Look at "Provable Security". Journal of Cryptology **20**(1) (2007) 3–37