A Novel Probabilistic Passive Attack on the Protocols HB and HB$^+$

José Carrijo* Rafael Tonicelli† Hideki Imai‡ Anderson C. A. Nascimento§

November 11, 2008

Abstract

We present a very simple probabilistic, passive attack against the protocols HB and HB$^+$. Our attack presents some interesting features: it requires less captured transcripts of protocol executions when compared to previous results; It makes possible to trade the amount of required transcripts for computational complexity; the value of noise used in the protocols HB and HB$^+$ need not be known.

1 Introduction

Authentication protocols specially designed for devices with low computational power are an active area of research, see, for instance, Matsumoto and Imai [1], Wang et al. [2], Naor and Pinkas [4], Hopper and Blum [5], Juels and Weis [6]. Among the many proposed schemes, the protocols HB/HB$^+$ have received special attention as they seem to be practical and their security was formally reduced to a well known computation problem: the Learning Parity with Noise - LPN [8].

These protocols are appropriate to be implemented in RFID tags (Radio Frequency Identification) or other devices with low power consumption and computational power. These tags may be used to check medicine bottles, library books, driver's licenses, and so on.

Attacks have been proposed against HB/HB$^+$. Some are active [9], that is they assume an adversary capable of inserting her own messages in the protocol while trying to impersonate a legal party. Other attacks are passive [8], assuming just adversaries that merely listen to protocol executions, capture transcripts and try to recover the secret information available to the legal parties. Clearly, a successful passive attack has devastating effects, as it leaves no trace whatsoever that the original protocol has been broken. The efficiency of passive attacks is usually measured in the literature by two parameters: the computational complexity of the attack; and the amount of protocol transcript executions (also denoted amount of captured protocol rounds when no confusion

*Department of Electrical Engineering, University of Brasilia. Campus Universitario Darcy Ribeiro,Brasilia, CEP: 70910-900, Brazil, Email:carrijo@redes.unb.br

†Department of Electrical Engineering, University of Brasilia. Campus Universitario Darcy Ribeiro,Brasilia, CEP: 70910-900, Brazil, Email:tonicelli@redes.unb.br

‡Faculty of Science and Engineering, Chuo University 1-13-27 Kasuga, Bunkyo-ku, Tokyo, 112-8551, Japan, & Research Center for Information Security (RCIS), National Institute of Advanced Industrial Science and Technology (AIST), Room 1102, Akihabara Daibiru, 1-18-13 Sotokanda, Chiyoda-ku, Tokyo, 101-0021 Japan E-mail:h-imai@aist.go.jp

§Department of Electrical Engineering, University of Brasilia. Campus Universitario Darcy Ribeiro,Brasilia, CEP: 70910-900, Brazil. E-mail: andclay@ene.unb.br

arrises) which is required for running the attack. All the passive protocols presented in the literature are deterministic, i.e. they are guaranteed to produce a valid output. It is an interesting question to study whether relaxing the attacks to be probabilistic gives us any kind of advantage.

Our Result: We propose the first probabilistic passive attack against HB/HB$^+$. Our attack needs much less captured transcripts than previous results while keeping a reasonable computational complexity.

Related Work: The most important attacks against HB/HB$^+$ are based on the BKW algorithm [8]. The problem is attacked by using Gaussian elimination method in samples of sequences with η-percent error. However, the computational complexity of BKW and the amount of captured transcripts of protocol executions required for running the attack are of the same order (exponential). This fact turns cryptanalysis infeasible in situations where those transcripts are not easily obtained. Our method does not have this limitation. It requires much less captured transcripts than BKW for a comparable computational complexity. Moreover, it makes possible to trade captured transcript for computational complexity. [12] proposes an attack against the protocol HB which has superior computational performance when compared to BKW. However, [12] still requires an amount of captured transcripts of the same order as its computational complexity. Additionally, it is not clear how to extend the attack proposed in [12] to the protocol HB$^+$. Our attack is trivially extendable to deal with HB$^+$. Finally, we note that an active attack (man-in-the-middle) was proposed against HB$^+$ in [9]. Our attack is passive.

After the completion and submission of this work we became aware of similar results independently discovered by Golebiewski et. al. in [11].

Organization of the Paper: In Section 2 we review the protocol HB/HB$^+$. In Section 3 we present our attack. In Section 5 we present our results and comparisons with the literature. Finally, we present our conclusions in Section 5.

2 The Protocols HB and HB$^+$

Consider the so called authentication problem. We have two parties, Alice and Bob, connected by means of an insecure channel. We assume that Alice and Bob shared a piece of secret information, a key. In an authentication protocol, Alice and Bob exchange messages over the insecure channel so that, at the end of the protocol, they are sure they are talking to each other or not. Informally, the protocol is said to be secure if no malicious party can impersonate Alice or Bob. This problem becomes particularly difficult when one of the parties has low computational power (smart cards, RFIDs, etc.).

The protocol HB, Hopper and Blum [5], was proposed as a way to obtain authentication for devices with low computational power. We now briefly describe how it works. Assume Alice and Bob pre-shared a k-bits long key $\mathbf{x}$.

Protocol HB

1. For $i = 1$ to r
   (a) Alice chooses a random k-bit string $\mathbf{a}_i \in \{0,1\}^k$, and sends it to Bob.
   (b) Bob computes $z_i = \mathbf{a}_i \odot \mathbf{x} + \nu_i$ where $\odot$ is the inner product, ν_i is a bit equal to 1 with probability $\eta \in (0, 1/2)$ and the sum is modulo 2. Bob sends z_i to Alice.

(c) Alice computes $z_i^* = \mathbf{a}_i \odot \mathbf{x}$ and compares it with $\mathbf{z}_i$.

2. Alice accepts the authentication as valid if $z_i^* \neq z_i$ in less than ηr rounds.

This protocol is clearly insecure if an attacker repeatedly queries Alice using the same string $\mathbf{a}$ for every round. To prevent such kind of attacks, the protocol HB$^+$ was proposed in [6]. We assume Alice and Bob pre-shared two k-bits long keys $\mathbf{x}$ and $\mathbf{y}$. r is a security parameter.

Protocol HB$^+$

1. For $i = 1$ to r

   (a) Alice chooses a random k-bit string $\mathbf{a}_i \in \{0,1\}^k$, and sends $\mathbf{a}_i$ to Bob.

   (b) Bob chooses a random k-bit string $\mathbf{b}_i \in \{0,1\}^k$, and sends $\mathbf{b}_i$ to Alice.

   (c) Bob computes $z_i = \mathbf{a}_i \odot \mathbf{x} + \mathbf{b}_i \odot \mathbf{y} + \nu_i$, where $\odot$ is the inner product ν_i is a binary random variable which is equal to 1 with probability η and the sum is modulo 2. Bob sends z_i back to Alice.

   (d) Alice computes $z_i^* = \mathbf{a}_i \odot \mathbf{x} + \mathbf{b}_i \odot \mathbf{y}$ and compares it to z_i.

2. Alice accepts the authentication as valid if $z_i^* \neq z_i$ in less than ηr rounds.

3 Proposed Method

3.1 Description

We propose a probabilistic passive attack against the HB protocol. Denote by $\mathbf{A}$ the $m \times k$ matrix $[\mathbf{a}_i]_{i=1}^m$ where the i-th row vector equals $\mathbf{a}_i$. This matrix represents the transcript of several executions of the protocol.

Denote by ν the m-dimensional column vector with each entry equal to ν_i and similarly for $\mathbf{x}$. Given $\mathbf{A}$, define $\mathbf{z}$ as $\mathbf{z} = \mathbf{A}\mathbf{x} + \nu$. We can also assume that the hamming weight of ν will be no larger than $0.40m$ (otherwise the authentication procedure becomes too unreliable) and that $m > k$. Let C be a subset of $\{1, \ldots, m\}$ with cardinality $|C|$. Denote the i-th element of C by $C(i)$ and the matrix $[\mathbf{a}_{C(i)}]_{i=1}^{|C|}$ by $\mathbf{A}_C$. Denote the i-th element of the column vector $\mathbf{x}$ by $x(i)$ and the column vector $\{x(C(1)), \ldots, x(C(|C|))\}$ by $\mathbf{x}_C$

Algorithm Inputs: (A,z)

1. Randomly select a subset C with cardinality $n = k + \gamma$ (γ being an integer suitably chosen).

2. Compute, by gaussian elimination, $\mathbf{x}_C$ so that $\mathbf{z}_C = \mathbf{A}_C\mathbf{x}_C$. If this solution does not exist or if there are many solutions, go back to the previous step.

3. Check if the hamming weight of $\mathbf{A}\mathbf{x}_C$ is less than $0.40n$. If it is the case, halt and output $\mathbf{x}_C$ as the desired solution (key). Otherwise go back to the first step.

3.1.1 Extension to HB^+

Our attack can be trivially extended to HB^+ with the overall effect of doubling the key size. Denote by $\mathbf{x}||\mathbf{y}$ the concatenation of two k-dimensional column vectors $\mathbf{x}$ and $\mathbf{y}$ so that $\mathbf{x}||\mathbf{y} = (x_1, \cdots, x_k, y_1, \cdots, y_k)$. Denote by $\mathbf{A}||\mathbf{B}$ the concatenation of the $m \times k$ matrices $\mathbf{A}$ and $\mathbf{B}$ so that $\mathbf{A}||\mathbf{B} = [\mathbf{a}_i||\mathbf{b}_i]_{i=1}^m$. Put $\mathbf{z}$ as $\mathbf{z} = [\mathbf{A}||\mathbf{B}][\mathbf{x}||\mathbf{y}] + \nu$ for previously defined ν.

Algorithm Inputs: (A,B,z)

1. Randomly select a subset C with cardinality $n = 2k + \gamma$ (γ being an integer suitably chosen).
2. Compute, by gaussian elimination, $[\mathbf{x}||\mathbf{y}]_C$ so that $\mathbf{z}_C = [\mathbf{A}||\mathbf{B}]_C[\mathbf{x}||\mathbf{y}]_C$. If this solution does not exist or if there are many solutions, go back to the previous step.
3. Check if the hamming weight of $[\mathbf{x}||\mathbf{y}]_C$ is less than $0.40n$. If it is the case, halt and output $[\mathbf{x}||\mathbf{y}]_C$ as the desired solution (key). Otherwise go back to the first step.

3.2 Complexity of the Attack

After capturing m transcripts of the protocol, the passive adversary creates a linear system $\mathbf{z}_C = \mathbf{A}_C\mathbf{x}_C$ by randomly choosing n challenge-response pairs $\{\mathbf{a}_i, z_i\}$, each pair corresponding to a linear equation of the form $z_i = \mathbf{a}_i \odot \mathbf{x}_C$.

There are m equations to be chosen, where, on average, $(1-\eta)m$ equations are correct and ηm equations are incorrect. The probability p of breaking the cryptosystem is the probability of choosing n linearly independent and correct equations, what is denoted by:

$$p = \frac{\begin{pmatrix} (1-\eta)m \\ n \end{pmatrix} \times \begin{pmatrix} \eta m \\ 0 \end{pmatrix}}{\begin{pmatrix} m \\ n \end{pmatrix}} = \frac{\begin{pmatrix} (1-\eta)m \\ n \end{pmatrix}}{\begin{pmatrix} m \\ n \end{pmatrix}} \tag{1}$$

Thus, on average, it will be necessary $1/p$ linear system resolutions to recover the shared key $\mathbf{x}$. Consequently, the expected complexity of the attack is dominated by:

$$\frac{1}{p} = \frac{\begin{pmatrix} m \\ n \end{pmatrix}}{\begin{pmatrix} (1-\eta)m \\ n \end{pmatrix}} = \frac{m(m-1)\ldots(m-n+1)}{\eta m(\eta m-1)\ldots(\eta m-n+1)} \tag{2}$$

As illustrated in equation 2, the computational effort depends basically on the parameters $\{\eta, m, n\}$, where (in the case of HB) $n = k + \gamma$ and $\gamma \geq 0$. This effort increases when η increases, or when the number of transcripts m is close to the length k of the shared key. Furthermore, under similar conditions and with the same parameters, different executions of the cryptanalytic algorithm may present different execution times due to its probabilistic nature.

4 Results and Comparisons

We compare the computational complexity and amount of required captured protocol transcripts of our attack to that of BKW algorithm and the algorithm presented in [12], here on denoted FMICM. We remark again that BKW and FMICM are deterministic algorithms while ours is a probabilistic one, thus our complexities here presented are *expected* values. Computational complexities are presented *per bit of information* as in [6] and [12]. For details about the performance evaluation of BKW and FMICM we refer to [8] and [12] respectively.

Tables 1,2,3 and 4 show a comparative analysis between expected computational complexities and amount of required protocol transcripts (captured protocol rounds) of our attack and those of BKW and FMICM for noise probability $\eta = 0.25, 0.20, 0.15, 0.10$.

Before specifically commenting about the numbers we present in this section, we would like to stress that in the case of attacks against authentication protocols for smart cards, RFIDs and other devices with narrow communication bandwidth, reducing the amount of required captured protocol transcripts might be more important than reducing the computational complexity (run time). Such a device usually communicates at not so high rates (typically no larger than 1 Mbps). Therefore, knowing that in a single round of the HB protocol we have to communicate about k bits (where k is usually 256 bits), if a certain attack needs to capture 2^{80} rounds of the HB protocol an adversary might have to wait about $2^{88}/10^6$ seconds (which is more than the estimated age of the universe) in order to be able to start doing computations to discover the secret key. As shown in the tables 1,2,3 and 4, our attack dramatically reduces the amount of captured protocol rounds necessary for obtaining the secret key, thus making attacks against the HB protocol much closer to being practical.

In details, we note that for $\eta = 0.20, 0.15, 0.10$ our attack presents better expected computational complexity and required number of transcripts of protocol executions than BKW. For $\eta = 0.25$ and key length larger than 160 bits, BKW presents better computational complexity.

For $\eta = 0.15$ and $\eta = 0.10$ our method possess better computational complexities and requires less amount of captured transcripts than FMICM. For $\eta = 0.20$ our method still requires much less capture of protocol transcripts than FMICM. However, the computational complexities of both methods are about the same (with a slight advantage to FMICM). For $\eta = 0.25$ our method is less efficient in terms of computational complexity when compared to FMICM, but it still requires much less captured protocol rounds.

It is also interesting to remark that for values of noise $\eta = 0,15$ and $\eta = 0,10$ the performance of FMICM significantly degrades when compared to BKW and our method.

5 Conclusion

We presented a novel probabilistic and passive attack against the protocols HB and HB$^+$.

Compared to the BKW and FMICM attacks, the proposed cryptanalytic method presents some key advantages:

- it does not require any pre-processing;
- it does not require any previous knowledge about the noise probability;

	Run Time			Amount of Captured Protocol Rounds		
Key Length	**BKW**	**FMICM**	**Our Method**	**BKW**	**FMICM**	**Our Method**
32	2^{23}	2^{8}	2^{9}	2^{23}	2^{8}	2^{10}
64	2^{35}	2^{16}	2^{21}	2^{35}	2^{16}	2^{12}
96	2^{45}	2^{25}	2^{34}	2^{45}	2^{24}	2^{13}
128	2^{54}	2^{34}	2^{47}	2^{54}	2^{34}	2^{13}
160	2^{62}	2^{43}	2^{60}	2^{62}	2^{42}	2^{14}
192	2^{70}	2^{52}	2^{73}	2^{70}	2^{51}	2^{14}
224	2^{78}	2^{60}	2^{87}	2^{78}	2^{60}	2^{14}
256	2^{86}	2^{69}	2^{99}	2^{86}	2^{69}	2^{14}
288	2^{94}	2^{81}	2^{113}	2^{94}	2^{81}	2^{14}

Table 1: Comparison of the expected computational effort and the amount of captured protocol rounds required to perform the attacks BKW, FMICM and the new cryptanalytic method for $\eta = 0.25$.

	Run Time			Amount of Captured Protocol Rounds		
Key Length	**BKW**	**FMICM**	**Our Method**	**BKW**	**FMICM**	**Our Method**
32	2^{22}	2^{8}	2^{6}	2^{22}	2^{9}	2^{10}
64	2^{33}	2^{16}	2^{15}	2^{33}	2^{10}	2^{12}
96	2^{42}	2^{25}	2^{25}	2^{42}	2^{24}	2^{12}
128	2^{50}	2^{35}	2^{34}	2^{50}	2^{19}	2^{13}
160	2^{58}	2^{44}	2^{45}	2^{58}	2^{42}	2^{13}
192	2^{66}	2^{52}	2^{58}	2^{66}	2^{50}	2^{13}
224	2^{74}	2^{62}	2^{65}	2^{74}	2^{60}	2^{15}
256	2^{82}	2^{71}	2^{75}	2^{82}	2^{65}	2^{15}
288	2^{89}	2^{83}	2^{86}	2^{89}	2^{83}	2^{15}

Table 2: Comparison of the expected computational effort and the amount of captured protocol rounds required to perform the attacks BKW, FMICM and the new cryptanalytic method for $\eta = 0.20$.

	Run Time			Amount of Captured Protocol Rounds		
Key Length	**BKW**	**FMICM**	**Our Method**	**BKW**	**FMICM**	**Our Method**
32	2^{21}	2^{8}	2^{3}	2^{21}	2^{8}	2^{10}
64	2^{31}	2^{19}	2^{9}	2^{31}	2^{19}	2^{13}
96	2^{39}	2^{26}	2^{17}	2^{39}	2^{24}	2^{13}
128	2^{47}	2^{35}	2^{23}	2^{47}	2^{33}	2^{13}
160	2^{55}	2^{43}	2^{31}	2^{55}	2^{43}	2^{13}
192	2^{63}	2^{52}	2^{39}	2^{63}	2^{51}	2^{13}
224	2^{69}	2^{62}	2^{47}	2^{69}	2^{62}	2^{14}
256	2^{76}	2^{71}	2^{56}	2^{76}	2^{71}	2^{14}
288	2^{82}	2^{83}	2^{61}	2^{82}	2^{83}	2^{14}

Table 3: Comparison of the expected computational effort and the amount of captured protocol rounds required to perform the attacks BKW, FMICM and the new cryptanalytic method for $\eta = 0.15$.

	Run Time			Amount of Captured Protocol Rounds		
Key Length	**BKW**	**FMICM**	**Our Method**	**BKW**	**FMICM**	**Our Method**
32	2^{20}	2^{8}	2^{1}	2^{20}	2^{8}	2^{10}
64	2^{28}	2^{17}	2^{4}	2^{28}	2^{16}	2^{10}
96	2^{36}	2^{26}	2^{9}	2^{36}	2^{24}	2^{11}
128	2^{44}	2^{35}	2^{13}	2^{44}	2^{33}	2^{13}
160	2^{50}	2^{44}	2^{18}	2^{50}	2^{43}	2^{13}
192	2^{57}	2^{54}	2^{24}	2^{57}	2^{51}	2^{13}
224	2^{63}	2^{62}	2^{27}	2^{63}	2^{62}	2^{13}
256	2^{70}	2^{71}	2^{31}	2^{70}	2^{71}	2^{14}
288	2^{76}	2^{85}	2^{36}	2^{76}	2^{85}	2^{14}

Table 4: Comparison of the expected computational effort and the amount of captured protocol rounds required to perform the attacks BKW, FMICM and the new cryptanalytic method for $\eta = 0.10$.

– the amount of transcripts needed to break the cryptosystem is dramatically reduced, making it more feasible to be implemented.

We finally remark that our attack allows one to trade computational effort versus amount of required protocol transcript executions. One can, in principle, reduce the computational effort to break the protocols HB/HB$^+$ by increasing the amount of available transcripts. It is also possible to break the protocols with fewer transcripts by increasing the computational effort. The same is not true for the BKW and FMICM algorithms, where, for any case, the amount of transcripts required to break the protocols increases exponentially with the length of the key.

Acknowledgements: We would like to thank an anonymous referee for comments that greatly improved the quality of the presentation of our results. This work was supported by SAGEM ORGA do Brasil.

References

[1] T. Matsumoto, H. Imai. Human identification through insecure channel. In Davies, D.W., ed.: Advances in Cryptology - EUROCRYPT 91. Volume 547 of Lecture Notes in Computer Science, Springer-Verlag (1991) 409-421.

[2] C. H. Wang, T. Hwang, J. J. Tsai. On the Matsumoto and Imai's Human Identification Scheme, In L.C. Guillou, J. J. Quisquater, eds.: Advances in Cryptology - EUROCRYPT 95. Volume 921 of Lecture Notes in Computer Science., Springer-Verlag (1995) 382-392.

[3] T. Matsumoto. Human-computer cryptography: An attempt. In C. Neuman, ed.: 3rd ACM Conference and Communications Security, New Delhi, India, ACM Press (1996) 68-75.

[4] M. Naor, B. Pinkas. Visual authentication and identification. In B. S. Kaliski Jr., ed.: Advances in Cryptology - CRYPTO '97. Volume 1294 of Lecture Notes in Cimputer Science., Springer-Verlag (1997) 322-336.

[5] N. J. Hopper and M. Blum. Secure Human Identification Protocols. In C. Boyd, editor, Advances in Cryptology - Asiancrypt '01, volume 2248 of Lecture Notes in Computer Science, pages 52-66. Springer-Verlag, 2001.

[6] A. Juels and S. A. Weis. Authenticating pervasive devices with Human Protocols. In Shoup, editor, Advances in Cryptology - Crypto 05, Lecture Notes in Computer Science. Springer-Verlag, to appear 2005.

[7] J. Katz and J. S. Shin. Parallel and Concurrent Security of the HB and HB$^+$ Protocols. Advances in Cryptology (EUROCRYPT 2006).

[8] A. Blum, A. Kalai, H. Wasserman. Noise-Tolerant Learning, the Parity Problem, and the Statistical Query Model. Journal of the ACM 50, 4 (July 2003), 506 - 519.

[9] H. Gilbert, M. Robshaw, and H. Silbert. An Active Attack against HB$^+$ - a Provably Secure Lightweight Authentication Protocol. IEE Electronic Letters 41, 21, pgs 1169–1170, 2005

[10] S. Weis, R. Rivest and A. Smith. New Foundations for efficient Authentication, Commutative Cryptography, and Private Disjointness Testing. MASSACHUSETTS INSTITUTE OF TECHNOLOGY - MIT, 2006

[11] Z. Golebiewski, K. Majcher, F. Zagorski, and M. Zawada. Practical Attacks on HB and HB+ Protocols, eprint available at http://eprint.iacr.org/2008/241.pdf

[12] M. Fossorier, M. Mihaljevi, H. Imai, Y. Cuiz, K. Matsuura. A Novel Algorithm for Solving the LPN Problem and its Application to Security Evaluation of the HB Protocol for RFID Authentication. Lecture Notes in Computer Science, vol. 4329, pp. 48-62, Dec. 2006.