Reid et al.'s Distance Bounding Protocol and Mafia Fraud Attacks over Noisy Channels

A. Mitrokotsa, C. Dimitrakakis, P. Peris-Lopez, J. C. Hernandez-Castro

***Abstract*—Distance bounding protocols are an effective countermeasure against relay attacks including distance fraud, mafia fraud and terrorist fraud attacks. Reid *et al.* proposed the first symmetric key distance bounding protocol against mafia and terrorist fraud attacks [1]. However, [2] claims that this is only achieved with a $(7/8)^n$ probability of success for mafia fraud, rather than the theoretical value of $(3/4)^n$ (for n rounds) achieved by distance bounding protocols without a final signature. We prove that the mafia fraud attack success using the Reid *et al.* protocol is bounded by $(3/4)^n$ and reduces as noise increases. The proof can be of further interest as it is the first – to the best of our knowledge – detailed analysis of the effects of communication errors on the security of a distance bounding protocol.**

***Index Terms*—Contactless smart cards, RFID, distance bounding protocols, relay attacks, mafia fraud attacks**

I. Introduction

A number of secure and efficient authentication protocols for RF transponders such as contactless smart cards and RFID tags have been proposed recently. Most assume proximity between readers and transponders due to limited radio channel range. However, an intruder located between the tag T (prover) and the reader R (verifier), can trick the latter into thinking that T is in close proximity. This attack can be divided into three subcategories:

a) Distance fraud: The attacker is a fraudulent tag $\overline{T}$. The attack involves $\overline{T}$ convincing the legitimate reader R of being nearer to the legitimate tag T than it really is.

b) Mafia fraud [3]: The attacker A is a pair $A = \{\overline{T}, \overline{R}\}$: $\overline{T}$ is a fraudulent tag interacting with the legitimate reader R and $\overline{R}$ is a fraudulent reader interacting with the legitimate tag T. Using $\overline{R}$, $\overline{T}$ convinces R that the latter communicates with the legitimate tag T while in reality it communicates with the attacker A. This is achieved without the disclosure to A of the private key shared between T and R.

c) Terrorist Fraud: The attacker $A = \{T, \overline{T}\}$ is a pair of two colluding parties: a legitimate tag T and a terrorist tag $\overline{T}$. The attack enables $\overline{T}$ to convince the legimate reader R of an assertion related to the private key of T. In this attack although the legitimate tag T is dishonest and cooperates with the terrorist tag $\overline{T}$, the secret key shared between the legitimate tag and reader is not revealed to the terrorist tag $\overline{T}$.

Distance bounding protocols were introduced in [4] to hinder distance fraud and mafia fraud attacks, by measuring the round trip delays during a rapid challenge-response exchange of n bits to infer an upper bound on the distance between the verifier and the prover. Subsequently [5] proposed a distance bounding protocol offering protection against mafia fraud only. Later, Reid *et al.* proposed a new protocol [1] with the objectives of (a) being resistant to both mafia and terrorist fraud, and (b) suitable for low-cost RFID tags. This work can be considered a reference point in the design of distance bounding protocols for constrained RF tags. Indeed, key ideas of [1] are used in recent proposals as in [6].

Contribution: We analyse the security of Reid *et al.*'s protocol (henceforth RP) against mafia fraud attack under noisy conditions. Due to power constraints and the wireless medium, RFID systems are particularly susceptible to noise, but its effect on the attacker has not been studied previously. In addition, we clarify RP's security under noise-free conditions. More specifically, [2] claims that the probability of success for a mafia fraud attack is bounded by $(7/8)^n$. However, this claim is based on an incorrect assumption about the Key Derivation Function (KDF) used in the protocol: that if the adversary can control $3/4$ bits of the input to the KDF, then he can guess the output of the KDF more easily. However, the KDF is indistinguishable from a uniform distribution unless all bits are known [1]. Nevertheless, [2] is commonly cited as evidence for the low security of RP. In this paper, we prove that the attack success probability is upper bounded by $(3/4)^n$ in noise-free conditions and refine the results of [5] by showing that it decreases polynomially as noise increases.

Notation: We consider sequences $x = x_1, \dots, x_n$ with all x_i in some alphabet $\mathcal{X}$ and $x \in \mathcal{X}^n$. We write $\mathcal{X}^* \triangleq \bigcup_{n=0}^{\infty} \mathcal{X}^n$ for the set of all sequences. The concatenation of x with some $y \in \mathcal{X}^m$ is written as $x|y \in \mathcal{X}^{m+n}$. If $x, y \in \mathcal{X}^n$ then $x \oplus y \in \mathcal{X}^n$, where $\oplus$ is an appropriate operator (XOR for $\mathcal{X} = \{0, 1\}$). $\mathbb{P}(A)$ denotes the probability of event A, while $\triangleq$ implies a definition.

II. Reid et al.'s Protocol

In RP [1], the reader R and tag T, whose identifiers are $\mathrm{ID}_R, \mathrm{ID}_T \in \mathcal{X}^*$ respectively, share a common secret $x \in \mathcal{X}^n$. The messages exchanged are:

1) $R \to T$: The reader chooses a random number $y_B \in \mathcal{X}^m$ and transmits it and its identity ID_R to the tag.
2) $T \to R$: The tag chooses a random number $y_A \in \mathcal{X}^m$ and transmits it and its identity ID_T to the reader.

A. Mitrokotsa and P. Peris-Lopez are with the ICT Group, Technical University of Delft ({a.mitrokotsa, p.perislopez}@tudelft.nl).

C. Dimitrakakis is with the Informatics Institute, University of Amsterdam (christos.dimitrakakis@gmail.com).

J. C. Hernandez-Castro is with the School of Computing, University of Portsmouth (Julio.Hernandez-Castro@port.ac.uk).

This work was supported by the Netherlands Organization for Scientific Research (NWO) under the RUBICON grant "Intrusion Detection in Ubiquitous Computing Technologies" and the ICIS project, supported by the Dutch Ministry of Economic Affairs, grant nr: BSIK03024.

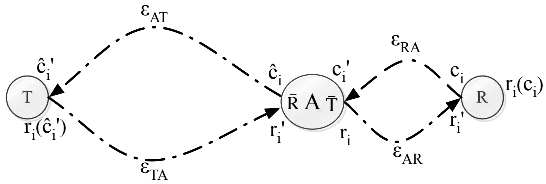

Fig. 1. Mafia Fraud attack over noisy channels

3) R and T: Both entities now use a key derivation function $f_K : \mathcal{X}^n \times \mathcal{X}^* \rightarrow \mathcal{X}^n$ to derive a symmetric encryption key $k = f_K(x, \text{ID}_T|\text{ID}_R|y_A|y_B)$. This key is used to split the secret key into two shares, k and $d = k \oplus x$. f_K is a pseudo-random function, so when x is a secret of high enough entropy, $f_K(x, \cdot)$ is indistinguishable from a uniform distribution.
4) T and R start the rapid bit exchange. The following steps are repeated for n rounds. At each round i:
   a) $R \rightarrow T$: The reader chooses a random bit c_i, transmits it to T and starts a clock.
   b) $T \rightarrow R$: Upon receiving c_i, the tag replies $r_i(c_i)$, with $r_i(c_i) = d_i$ if $c_i = 0$, and $r_i(c_i) = k_i$ if $c_i = 1$.
   c) R: After the reception of r_i, the reader stops the clock and stores the delay time Δt_i and checks r_i. If r_i is incorrect, an extra error message is sent to T. The delay time is highly dependent on the distance[1] between the tag and the reader.

III. Mafia Fraud Attacks

In this section we prove bounds on the security of RP against mafia fraud attacks. It is important to note that these bounds are generally applicable to the class of distance bounding protocols that do not use signed messages (i.e. $sign(c_1|r_1|...|c_n|r_n)$) at the end of the rapid bit exchange phase. In addition, we provide first security results on rapid bit exchange protocols under noisy conditions.

Theorem 1: In RP, the probability that a mafia fraud attack can occur is bounded by $(\frac{3}{4})^n$, when transmission errors due to the noise in the forward (reader-to-tag) and backward (tag-to-reader) channels are zero.

Proof (Sketch): An adversary could transmit an anticipated challenge c_i' to the tag before the reader sends its challenge c_i. Half of the time, $c_i' = c_i$, so the adversary can correctly reply $r_i(c_i)$ to the reader. Otherwise, the adversary can guess randomly, again being correct half of the time. So, the adversary has $3/4$ probability of answering correctly overall. Assuming that the success probability at each round is independent of previous successes, the total probability of success for an adversary is $(3/4)^n$ for n rounds. ■

We only show a sketch proof, as the theorem is a direct corollary of Theorem 2, which also holds when errors can appear due to the noise in the channel (see Fig.1). We now assume that the communication between entities in $\{R, T, A\}$ is not noise free: Whenever a symbol $x \in \mathcal{X}$ is sent from Y to Z, the symbol x' that Z receives may differ from x due to noise. This is modeled as a probability of erroneous transmission from Y to Z, $\varepsilon_{YZ} \triangleq \mathbb{P}(x' \neq x)$, for all $Y, Z \in \{T, R, A\}$.

Now consider that an attacker A performs a mafia fraud attack against the communication between a legitimate RFID reader R and a genuine RFID tag T. We use $c_i \in \mathcal{X}$ to denote the challenge sent during the i^{th} round of the rapid bit exchange and $r_i(c_i)$ for the correct response. Before the rapid bit exchange starts, the attacker A sends a sequence of n challenge guesses $\{\hat{c}_i\}_{i=1}^n$ to the legitimate tag T, which receives $\hat{c}_i'$, with error probability $\varepsilon_{AT} \triangleq \mathbb{P}(\hat{c}_i' \neq \hat{c}_i)$. Then the legitimate tag T calculates the appropriate response $r_i(\hat{c}_i')$ and sends it back to the attacker A, who receives r_i'. If $c_i' = \hat{c}_i$ the attacker sends $\hat{r}_i = r_i'$, the reply received by the tag; otherwise, he selects $\hat{r}_i$ uniformly from $\mathcal{X}$. The reader R sees $\hat{r}_i'$, and $\varepsilon_{AR} \triangleq \mathbb{P}(\hat{r}_i' \neq \hat{r}_i)$.

If the response $r_i(c_i)$, calculated by the legitimate reader R, equals $\hat{r}_i'$, then the adversary is *successful*. We denote this event by $s_i \triangleq \mathbb{I}\{\hat{r}_i' = r_i(c_i)\}$, where $\mathbb{I}$ is an indicator function such that $\mathbb{I}\{A\} = 1$ if A is true and 0 otherwise. The attack is *completely successful* when s_i is 1 for all i.

Theorem 2: Using RP for n rounds, with alphabet $\mathcal{X}$ and channel noise ε, the probability of success of a mafia fraud attack is:

$$\mathbb{P}(s_1, \ldots, s_n) = [A\frac{1}{k} + \frac{1}{k}(1 - \frac{1}{k})]^n$$

where $A \triangleq \frac{(F+1)F}{2} + \frac{(1-F)^2}{k-1}$, $F \triangleq \frac{\varepsilon}{(k-1)} \frac{1-z^2}{1+z^2} + z^2$ and $z \triangleq \frac{k(1-\varepsilon)-1}{k-1}$, $k \triangleq |\mathcal{X}|$. In addition, we assume that $\forall c_i, c_i' \in \mathcal{X}$ such that $c_i \neq c_i'$, $\theta \triangleq \mathbb{P}(r(c_i) = r(c_i')) = 1/k$.

Proof: The probability of correctly guessing the i-th challenge is:

$$\mathbb{P}(s_i) = \mathbb{P}(s_i|\hat{c}_i = c_i')\mathbb{P}(\hat{c}_i = c_i') + \mathbb{P}(s_i|\hat{c}_i \neq c_i')\mathbb{P}(\hat{c}_i \neq c_i').$$

$\forall g, \hat{c}_i \in \mathcal{X}$ and $|\mathcal{X}| = k$, $\mathbb{P}(\hat{c}_i = g) = \frac{1}{k}$, so $\mathbb{P}(c_i \neq c_i') = 1 - \mathbb{P}(\hat{c}_i = c_i') = 1 - \frac{1}{k}$. Combining the above, we obtain:

$$\mathbb{P}(s_i) = \mathbb{P}(s_i|\hat{c}_i = c_i')\frac{1}{k} + \mathbb{P}(s_i|\hat{c}_i \neq c_i')(1 - \frac{1}{k}) \qquad (1)$$

It also holds that $\mathbb{P}(s_i|\hat{c}_i \neq c_i') = \mathbb{P}(\hat{r}_i' = r_i(c_i)|\hat{c}_i \neq c_i')$. In addition, whenever $c_i' \neq \hat{c}_i$, the attacker sends a random response $\hat{r}_i$. In that case, the success probability is:

$$\mathbb{P}(s_i|\hat{c}_i \neq c_i') = \mathbb{P}(\hat{r}_i' = r_i(c_i)|\hat{c}_i \neq c_i') = \frac{1}{k}.$$

Consequently, equation (1) can be written as:

$$\mathbb{P}(s_i) = \mathbb{P}(s_i|\hat{c}_i = c_i')\frac{1}{k} + \frac{1}{k}(1 - \frac{1}{k}) \qquad (2)$$

For the case when $c_i' = \hat{c}_i$, we define the following quantities:

$$A \triangleq \mathbb{P}(s_i|\hat{c}_i = c_i') = \mathbb{P}(\hat{r}_i' = r_i(c_i)|\hat{c}_i = c_i')$$
$$B \triangleq \mathbb{P}(\hat{r}_i' = r_i(c_i)|\hat{c}_i = c_i', \hat{r}_i' = r_i(\hat{c}_i'))$$
$$C \triangleq \mathbb{P}(\hat{r}_i' = r_i(c_i)|\hat{c}_i = c_i', \hat{r}_i' \neq r_i(\hat{c}_i'))$$
$$D \triangleq \mathbb{P}(\hat{r}_i' = r_i(\hat{c}_i')|\hat{c}_i = c_i')$$

[1]Assuming that the information can not travel faster that the speed of light c, the distance between R and T is upper bounded by $c \cdot \Delta t_{max}/2$, where Δt_{max} is the maximum delay time between sending out the bit c_i and receiving the bit r_i back.

It is easy to see that:

$$A = \mathbb{P}(s_i|\hat{c}_i = c'_i) = BD + C(1-D) \quad (3)$$

In order to further simplify the derivation, we shall also define the following: $x_0 \triangleq c_i$, $x_1 \triangleq c'_i = \hat{c}_i$, $x_2 \triangleq \hat{c}'_i$, $\varepsilon_1 \triangleq \varepsilon_{RA}$, $\varepsilon_2 \triangleq \varepsilon_{AT}$.

It can now be easily seen that x_0, x_1, x_2 form a three stage Markov chain, which satisfies the assumptions of Lemma 1 (see Appendix) for $n = 2$. Applying the lemma, we obtain:

$$\mathbb{P}(x_n = x_0) = \mathbb{P}(x_2 = x_0) = \mathbb{P}(\hat{c}'_i = c_i) = \sum_{l=1}^{2} \frac{\varepsilon_l}{k-1} \prod_{j=l+1}^{2} \frac{k(1-\varepsilon_j)-1}{k-1} + \prod_{j=1}^{2} \frac{k(1-\varepsilon_j)-1}{k-1} = F \quad (4)$$

Similarly, we may set: $y_0 \triangleq r_i(\hat{c}'_i)$, $y_1 \triangleq \hat{r}_i = r_i{'}$, $y_2 \triangleq \hat{r}'_i$, ${\varepsilon_1}' \triangleq \varepsilon_{TA}$, ${\varepsilon_2}' \triangleq \varepsilon_{AR}$ to obtain:

$$\mathbb{P}(y_n = y_0) = \mathbb{P}(y_2 = y_0) = \mathbb{P}(\hat{r}'_i = r_i(\hat{c}'_i)|\hat{c}_i = c'_i) = \sum_{l=1}^{2} \frac{\varepsilon'_l}{k-1} \prod_{j=l+1}^{2} \frac{k(1-{\varepsilon_j}')-1}{k-1} + \prod_{j=1}^{2} \frac{k(1-{\varepsilon_j}')-1}{k-1} = D \quad (5)$$

After some calculations, we can simplify the expressions for B, C to $B = F + \theta - \theta F$ and $C = \frac{1-F}{k-1}$, since:

$$\mathbb{P}(\hat{r}'_i = r_i(c_i)|\hat{c}_i = c'_i, \hat{r}'_i \neq r_i(\hat{c}'_i), \hat{c}'_i \neq c_i) = \frac{1}{k-1}.$$

If we assume $\varepsilon_{AT} = \varepsilon_{TA} = \varepsilon_{RA} = \varepsilon_{AR} = \varepsilon$, Corollary 1 applies and equations (4) and (5) can be condensed to:

$$F = D = \frac{\varepsilon}{(k-1)} \frac{1-z^n}{1+z^n} + z^n, \quad (6)$$

where $z = \frac{k(1-\varepsilon)-1}{k-1}$ and $n = 2$.

Finally, by equation (3) and the theorem's assumption that $\theta = 1/k$ and substituting B and C we get:

$$A = \frac{F^2(k-1)+F}{k} + \frac{(1-F)^2}{k-1} \quad (7)$$

Using (6), (7) and (3) we obtain the final result, where the probability of a successful attack only depends on the noise ε, the alphabet size k and the number of rounds n. ■

For $\varepsilon = 0$, equation (6) becomes $\mathbb{P}(s_i) = \frac{2k-1}{k^2}$. Assuming $k = 2$, we obtain $\mathbb{P}(s_i) = \frac{3}{4}$ and via independence of consecutive successes, a total success probability of $(3/4)^n$ over n rounds in the rapid bit exchange. Thus, we recover Theorem 1 and the original result of [1].

The success probability for increasing ε is shown in Fig. 2 for $k \in \{2, 4, 6, 8\}$. One may also see that a successful attack becomes less likely with increased alphabet size, or noise. A larger alphabet may result in either a longer transmission time (which is undesirable) or larger error probabilities, depending on the encoding. On the other hand, increased noise reduces the probability of successful authentication of a legitimate tag (see [5], Sec. 3.2). We may choose n, k, and a tolerance threshold [1], [5] to trade off transmission times with guarantees for false accept or reject rates. This is possible if a bound on the error is known, and the costs of transmission and false acceptance or rejection are well defined. However, we do not examine this issue here.

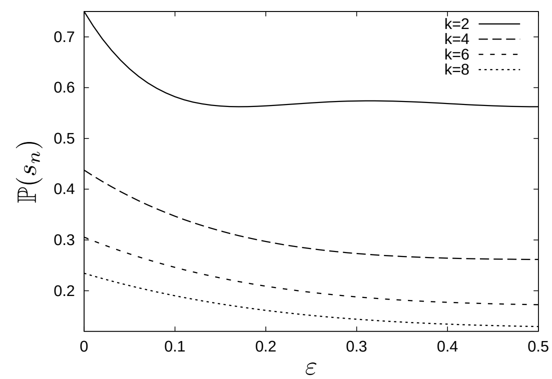

Fig. 2. Probability of Mafia Fraud attack vs. Noise for alphabet size k

IV. Conclusions

We have proved that Reid *et al.*'s protocol is secure against mafia fraud attacks. The probability that an intruder can trick the verifier into thinking that the prover is in close proximity is bounded by $(3/4)^n$ and reduces as noise increases. The result can be extended to the use of a threshold for tolerating a small number of errors [5] by plugging the expression for $\mathbb{P}(s_i)$ in the binomial cumulative distribution function. The security of this protocol can be further increased to $(1/2)^n$ by the inclusion of a signed message of the $2n$ bits sent in the rapid-bit exchange phase [4], [6]. However, such a signed message must be sent by normal communication with error correction [7], which increases authentication time.

References

[1] J. Reid, J. M. Gonzalez Nieto, T. Tang, and B. Senadji, "Detecting relay attacks with timing-based protocols," in *Proc. ASIACCS'07*. ACM, 2007, pp. 204–213.

[2] S. Piramuthu, "Protocols for RFID tag/reader authentication," *Decision Support Systems*, vol. 43, no. 3, pp. 897–914, 2007.

[3] Y. Desmedt, "Major security problems with the "unforgeable" (Feige)-Fiat-Shamir proofs of identity and how to overcome them," in *Proc. SecuriCom '88*, Mar. 1988, pp. 147–159.

[4] S. Brands and D. Chaum, "Distance-bounding protocols," in *Proc. EUROCRYPT'93*, ser. LNCS, vol. 765, 1994, pp. 344–359.

[5] G. Hancke and M. Kuhn, "An RFID distance bounding protocol," in *Proc. SECURECOMM'05*, Sept. 2005, pp. 67–73.

[6] C. H. Kim, G. Avoine, F. Koeune, F.-X. Standaert, and O. Pereira, "The Swiss-knife RFID distance bounding protocol," in *Proceedings of ICISC '08*, ser. LNCS. Springer-Verlag, Dec. 2008.

[7] C. H. Kim and G. Avoine. (2009) RFID distance bounding protocol with mixed challenges to prevent relay attacks. Cryptology ePrint Archive, Report 2009/310. [Online]. Available: http://eprint.iacr.org/2009/310.pdf

V. Appendix

Lemma 1: Assume a Markov chain $x_0, x_1, \ldots, x_n$ with $x_i \in \mathcal{X}$ and $|\mathcal{X}| = k$. The chain has the property that, for some $\{\epsilon_i\}_{i=1}^n$ with $\epsilon_i \in [0,1]$: $\mathbb{P}(x_i \neq x_{i-1}) = \epsilon_i$, for $i = 1, \ldots, n$. In addition, $\mathbb{P}(x_i = x | x_i \neq x_{i-1}) = \frac{1}{k-1}$, $\forall x \neq x_{i-1}$. Then, $\mathbb{P}(x_n = c|x_0 = c)$ equals:

$$\sum_{l=1}^{n} \frac{\epsilon_l}{k-1} \prod_{j=l+1}^{n} \frac{k(1-\epsilon_j)-1}{k-1} + \prod_{j=1}^{n} \frac{k(1-\epsilon_j)-1}{k-1}.$$

Corollary 1: If $\epsilon_i = \epsilon$ for all i, then for any $n \geq 1$:

$$\mathbb{P}(x_n = c|x_0 = c) = \frac{\epsilon}{k-1} \frac{1-z^n}{1-z} + z^n, \quad z = \frac{k(1-\epsilon)-1}{k-1}.$$