Provably Secure Distance-Bounding: an Analysis of Prominent Protocols

Marc Fischlin Cristina Onete

Darmstadt University of Technology & CASED, Germany
www.cryptoplexity.de

Abstract. Distance-bounding protocols aim to prevent man-in-the-middle attacks by measuring response times. At ISC 2011, Dürholz et al. [8] formalized the four attacks such protocols typically address: (1) mafia attacks, where the adversary must impersonate to a verifier in the presence of an honest prover (2) terrorist attacks, where the adversary gets limited offline support from the prover for the impersonation, (3) distance fraud attacks, where provers claim to be closer to verifiers than they really are, and (4) impersonation security, where adversaries try to impersonate the provers in lazy protocol phases. Dürholz et al. [8] also showed a thorough and formal security analysis of (an enhanced version of) the mafia fraud resistant construction of Kim and Avoine [12].

In this paper we give the exact level of security against each of the above-standing attacks for some other distance-bounding protocols according to [8], namely the constructions by Brands and Chaum [4], Hancke and Kuhn [11], Avoine and Tchamkerten [3], Reid et al. [14], and the Swiss-knife protocol due to Kim et al. [13]. Our conclusions are as follows: we (1) show that the Swiss Knife protocol attains the same optimal mafia fraud resistance as Brands and Chaum, but with less expensive primitives. The other constructions have worse mafia fraud resistance levels ([3] also has a good mafia fraud resistance, but requires exponential storage). The same optimal mafia fraud bound is also achieved by the enhanced construction in [8], at the expense of lower distance fraud resistance. Furthermore, we (2) prove that the allegedly terrorist-fraud resistant constructions in [14] and [13] are, in fact, *not* terrorist fraud resistant according to the notion of Dürholz et al. Whereas [13] does not attain even the intuitive notion of terrorist fraud resistance in the literature, the protocol in [14] does attain it. We (3) discuss different approaches to define terrorist fraud resistance, and discuss the strength of the model in [8], showing that though it is strong, it may still be more appropriate than the weaker intuition. Finally, we (4) give an overview of tools used in distance-bounding protocols at large.

Keywords. distance-bounding, location privacy

1 Introduction

Distance-bounding protocols were designed by Brands and Chaum [4] in 1993 to address man-in-the-middle (MITM) relay attacks, also called mafia fraud byDesmedt [7]. Essentially, distance-bounding enhances regular authentication such that the verifier accepts if the prover authenticates *and* it can prove it is within a pre-set distance of the verifier. In practice this is done by mounting a clock on the verifier. The protocol is then divided in lazy phases (where the clock is not used) and time-critical phases (where the verifier measures the time elapsed between sending a challenge and receiving a response). As relay MITM attacks cause processing delays, pure relaying of messages is thus detected by the clock.

Recently, Dürholz et al. [8] formalized the four most significant threats against distance-bounding protocols, particularly for Radio Frequency Identification (RFID) systems. In this setting, provers are the so-called RFID tags, while the verifiers are represented by RFID readers.

The four attacks described in [8] are:

MAFIA FRAUD. The adversary impersonates to the reader while communicating with the genuine tag. Here the clock prevents pure relaying between reader and tag.

TERRORIST FRAUD. The tag helps the adversary authenticate by disclosing useful information in offline phases. However, the tag should not reveal trivial information like the secret key.

DISTANCE FRAUD. The (malicious) tag claims to be closer to the reader than it actually is.

IMPERSONATION RESISTANCE. The adversary impersonates the honest tag during lazy phases only.

Recently, Cremers et al. [6] introduced a new attack, called distance hijacking. This attack, however, involves two provers, one honest and one malicious. As we confine ourselves to the single-prover-single-verifier case, this attack falls outside the scope of this paper.

1.1 RFID Distance-Bounding Protocols

In their distance-bounding framework, Dürholz et al. [8] also quantified and proved the security properties of an enhancement of the well-known protocol due to Kim et al. [12]. On the one hand, such formal analysis can prevent security breaches like those outlined by Abyneh [1], who showed attacks against two allegedly secure schemes. On the other hand, analyzing security properties in a common, formal framework enables easier comparison between protocols.

In this paper, we continue the assessment work of [8] by quantifying the security of several outstanding RFID distance-bounding protocols. Indeed, such protocols abound in the literature. The best known mafia fraud resistance (for N_c time-critical rounds) is about $\frac{1}{2}$ per time-critical round, thus $\left(\frac{1}{2}\right)^{N_c}$ in total, and it was first achieved by the protocol due to Brands and Chaum [4]. However, [4] uses computationally-expensive signature schemes and cannot be used for resource-constrained devices such as RFID tags. The same resistance is achieved in a less expensive way by the Swiss-knife protocol [13]. A very efficient and well-known protocol, due to Hancke and Kuhn [11], achieves only a resistance of $\left(\frac{3}{4}\right)^{N_c}$, since it is vulnerable to the so-called Go-Early strategy (see the proof in section 3.2), where the adversary queries the tag in advance to learn about a half of the correct responses for the time critical rounds, and then guesses the correct answer for the other responses. A better mafia fraud resistance than [11] is attained by Avoine and Tchamkerten [3], where the tree structure for the responses limits the man-in-the-middle's Go-Early success. However, the tree in [3] requires exponential storage (in N_c).

Fewer protocols in the literature address terrorist fraud resistance. An early construction due to Reid et al. [14] claims to achieve terrorist fraud resistance by using a pseudo-random function (PRF) and a symmetric encryption scheme. An adversary that has both the output of the PRF and of the symmetric encryption can recover the secret key. Another protocol aiming to achieve terrorist fraud resistance is the Swiss-knife protocol due to Kim et al. [13]. Interestingly, in this scheme the prover can help the adversary by handing over only half of the responses, and letting the adversary guess the other responses. This gives the adversary an advantage in sessions when the prover helps; however, once the prover stops helping, the probability drops again.

1.2 Our Contributions

In this paper we analyze the distance-bounding properties of the following protocols: the Brands and Chaum protocol [4], the Hancke-Kuhn protocol [11], the Avoine-Tchamkerten [3], and the reputedly terrorist-fraud resistant constructions due to Reid et al. [14] and Kim et al. [13]. For each protocol, we (1) show concrete security bounds for the four attacks described above, thus (2) disproving claims of terrorist fraud resistance of the well-known protocols in [14] and [13]. We also (3) discuss the consequences of this contradiction, balancing the strength of the model against the notion we intuitively wish to achieve. Finally, we (4) outline a few strategies for distance-bounding constructions.

One of our main contributions is (2): we show that the two allegedly terrorist fraud resistant constructions in [13] and [14] are *not*, in fact, terrorist fraud resistant. In particular, for Reid et al.'s construction, a malicious prover could forward the adversary a part (a half) of the correct responses and let the adversary guess the other responses. Thus, the adversary authenticates with probability of about $(\frac{3}{4})$ per time-critical round, whereas the simulator (i.e. the adversary without the prover's support) may only authenticate with probability $(\frac{1}{2})$ per round. This attack can furthermore be scaled down so that it even holds if we introduce a threshold value for the simulator's success probability (this, however, reduces the adversary's advantage).

In the Swiss-knife protocol, the tag can even hand over its secret key, as long as it does not forward the secret identifier (the simulator has then no way of guessing the identifier, and cannot authenticate).

We discuss at length the consequences of our results, with particular emphasis on the terrorist fraud resistant notion captured by [8], which we assess in view of the intuition behind terrorist fraud resistance. Our attack against the construction of Reid et al. raises some questions, as this protocol does attain some intuitive, though weak, notion of terrorist fraud resistance, which excludes attacks where partial, indirect secret key information is forwarded. We argue that, whereas the notion in [8] is very strong, the weaker, intuitive notion is too weak. Furthermore, no other concrete formalization of this intuitive notion is known, as the definition in [2] is very informal. It remains an open question whether the intuition of terrorist fraud resistance can be captured better by a relaxation of the definition in [8]. We also consider it paramount to design constructions which do, in fact, attain the strong notion considered by [8].

Our analysis results are summarized in table 1. In this table we show only the loose security bounds for the four attacks described above (ignoring the small terms). The precise quantification can be found in section 3.

	Mafia	**Terror**	**Distance**	**Impersonation**
[4] [1]	$\left(\frac{1}{2}\right)^{N_c}$	×	×	$\left(\frac{1}{2}\right)^{N_c}$
[11]	$\left(\frac{3}{4}\right)^{N_c}$	×	$\left(\frac{3}{4}\right)^{N_c}$	×
[3] [2]	$\frac{1}{2}(N_c+2)\left(\frac{1}{2}\right)^{N_c}$	×	$\left(\frac{3}{4}\right)^{N_c}$	$\left(\frac{1}{2}\right)^{\lvert V\rvert}$
[14]	$\left(\frac{3}{4}\right)^{N_c}$	×	$\left(\frac{3}{4}\right)^{N_c}$	×
[13]	$\left(\frac{1}{2}\right)^{N_c-T}$	×	$\left(\frac{3}{4}\right)^{N_c-T}$	$\left(\frac{1}{2}\right)^{\lvert V\rvert}$
[8]	$\left(\frac{1}{2}\right)^{N_c}$	×	$\left(\frac{7}{8}\right)^{N_c}$	$\left(\frac{1}{2}\right)^{\lvert V\rvert}$

Figure 1: Distance Bounding at a glance. [1]This protocol uses expensive primitives. [2]This protocol requires exponential storage requirements. We denote by N_c the number of time-critical rounds, by T a tolerance level for faults, and by $|V|$ is the bit length of an authentication string V sent by the verifier.

2 Preliminaries

2.1 Model Overview

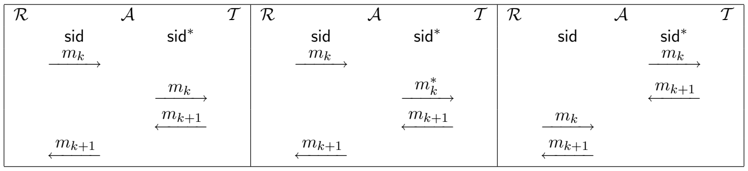
The security model in [8] considers single-verifier-single-prover distance-bounding, focusing on the RFID distance-bounding.Here, the single prover $\mathcal{T}$ is an RFID tag and the verifier $\mathcal{R}$ is an RFID reader, sharing a secret key *sk* generated by a key generation algorithm Kg. The reader uses a clock to measure the time it takes for a reader to send a challenge and receive a response.

Dürholz et al. consider round-based distance-bounding protocols, where rounds are called *time-critical* if the clock is used and *lazy* otherwise. We briefly review here the main possible attacks and refer to [8] for more details.

Mafia Fraud. Mafia fraud resistance is a man-in-the-middle (MITM) attack, where pure relay is prevented by the reader's clock. Informally, the MITM consists of two adversaries: a leech, which impersonates the reader to an honest tag, and a ghost, which impersonates the tag to an honest reader. The adversary aims to authenticate to the reader; however, the clock detects processing delays in the MITM resulting from pure relay. Both the reader and the honest tag are unaware of the MITM attack.

Dürholz et al. [8] formalize mafia fraud by introducing an abstract clock denoted clock, which keeps track of the messages sent in several protocol executions called sessions. There sessions can be: reader-tag (the adversary observes the exchange, but does not insert messages), reader-adversary (the adversary impersonates the tag to the reader), and adversary-tag (the adversary impersonates the reader to the tag). Relaying is considered round-wise (actually, [8] considers phases consisting of several rounds, in case the reader measures time over many rounds). A phase is called tainted if the adversary *purely* relays communication between a reader-adversary and an adversary-tag session. Here, *pure* relay refers to an adversary receiving a message in a session sid and *then* relaying the exact, same message in a session with identifier sid'. Then the adversary does the same, relaying the adversary back again between sid' and sid for all subsequent rounds pertaining to the tainted phase. If the adversary changes any of the challenges or responses it receives in sid before it forwards them in sid', this is not pure relay. Moreover, if the adversary queries sid' with some message m *before* receiving the same m from sid, this is not relay.

We refer to [8] for a precise definition of tainted phases and the motivation, and show the following illustration, also from [8]:



tainted
(pure relay)

untainted
(distinct messages $m_k^* \neq m_k$)

untainted
(distinct scheduling)

Figure 2: Tainted and Untainted Time-Critical Phases.

Terrorist Fraud. A terrorist adversary must also authenticate to the reader; however, the dishonest tag now helps the adversary win, without forwarding information that allows the adversary to authenticate later.In [8], the dishonest tag only interacts with the adversary in lazy phases. A time-critical phase of a reader-adversary session is tainted if during this phase the adversary queries the tag.

A crucial part of the definition for terrorist fraud resistance is defining how much the malicious tag can help the adversary. In [8], terrorist fraud resistance is defined in terms of a simulator: a scheme achieves terrorist fraud resistance if an adversary receiving offline help from a tag is as likely to win as a simulator having access to the information passed from the tag to the adversary. For the sake of fairness, the simulator has as many impersonation attempts as the adversary. Note that the simulator only begins its attack once the adversary successfully authenticates. If the information passed from the tag to the adversary is session-specific, then this data is useless to the simulator. If, on the other hand, the data contains the secret key, then the simulator can recover it and also succeed.

Formally, for every successful terrorist fraud adversary $\mathcal{A}$, there must exist a simulator $\mathcal{S}$ whose success probability is at least as large as the adversary's.

Distance Fraud. In distance fraud attacks, the adversary is the malicious tag itself, who is at a larger distance than allowed from the reader, but aims to convince the reader to the contrary.However, as the reader's clock measures time accurately, the adversary can only beat by anticipating the reader's challenges and respond in advance. In [8], the adversary must commit in advance to the responses for each time-critical phase. If the adversary does not commit to the response of one phase, this phase is called tainted. For distance fraud, reader-tag and adversary-tag sessions are no longer relevant since the adversary is the tag.

Impersonation Resistance. Finally, impersonation resistance refers to lazy-phase tag authentication. The idea, introduced by [3], is that even without the time-critical phases, the prover is still authenticated. Replays, however, are still allowed at this stage. We refer to the definition in [8] for the formalization of this notion.

2.2 Parameters

In [8], apart from the threshold $t_{\max}$, which upper-bounds the roundtrip transmission time, and the number of time-critical rounds N_c, several other parameters are considered. Concretely, [8] allows for max. $T_{\max}$ delayed transmissions (i.e. phases with delayed responses) and for max. $E_{\max}$ erroneous transmissions (phases with wrong responses). Note that though most existing constructions do not provide for erroneous/delayed communication, fault tollerance is essential in resource constrained environments such as RFID.

When specifying the adversary's characteristics we use the parameter t for the running time (including the time of honest parties in the protocol, when the adversary has to wait for the responses) and $q_{\mathcal{R}}$, $q_{\mathcal{T}}$, and q_{OBS} for the number of executions in which the adversary communicates with the honest reader, the honest tag, resp. passively observes a communication between reader and tag. To relate the distance bounding security notions to the security of the underlying cryptographic primitives, we often transform a successful distance-bounding adversary $\mathcal{A}$ into an adversary (or more) $\mathcal{A}'$ against the primitive(s). We use standard notions for these primitives, letting $\mathbf{Adv}_{\mathcal{S}}^{\text{Exp}}(\mathcal{A}')$ denote the maximal probability of an adversary $\mathcal{A}'$ breaking a cryptographic scheme $\mathcal{S}$ in some experiment Exp. For example, $\mathbf{Adv}_{\mathit{Sign}}^{\text{Unf}}(\mathcal{A}')$ denotes the probability of forging signatures, $\mathbf{Adv}_{\mathsf{PRF}}^{\text{d}}(\mathcal{A}')$ denotes the advantage of distinguishing a pseudorandom function PRF from random, and $\mathbf{Adv}_{\mathcal{E}}^{\text{IND-CPA}}(\mathcal{A}')$ is the advantage of breaking the encryption scheme in an IND-CPA attack. The parameters of $\mathcal{A}'$ in these experiments will be specified in terms of the parameters of $\mathcal{A}$.

3 Protocol Assessment

In this section, we analyze the properties the following distance-bounding protocols: Brands and Chaum [4], Hancke and Kuhn [11], Avoine and Tchamkerten [3], Reid et al. [14], and the Swiss-Knife protocol [13]. We concretely show that provable terrorist fraud resistance is harder to attain than intuition indicates.

Also, most earlier constructions of distance-bounding protocols do not allow for fault tolerance. In other words, the assumption is that the prover's behavior is always constant and that both transmissions and transmission times are reliable and constant. Thus the assumption is that any challenge sent by the verifier always arrives at the prover (and takes a constant time), any response sent by the prover is also always received by the verifier (and the time of flight is constant), and the prover always has the same processing delay. This is not always the case, particularly not for resource constrained devices like RFID, where transmissions are not always reliable and transmission times and processing delays may vary. In their framework, Dürholz et al. [8] introduce fault tolerance parameters as outlined in Section 2.2.

3.1 Brands and Chaum

In the Brands and Chaum construction [4], the reader and tag first exchange random bits in N_c time-critical phases, then finally signs the concatenation of these bits under the shared secret key *sk* — see Fig. 3. Though in this case the lazy phases do not pre-date the time-critical phases, the concept of impersonation resistance is still applicable. We only require that the adversary can forward a fresh signature in a winning reader-adversary session, in the sense that the adversary has not seen it before. As the lazy phase has a single message, the only possibly relayed message here is the signature.

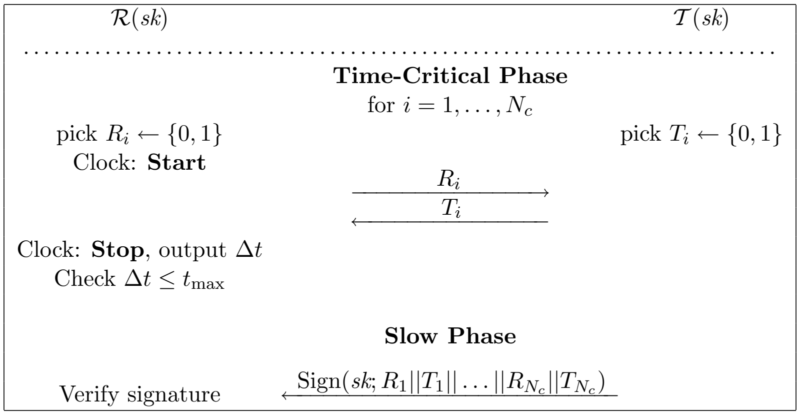

$\mathcal{R}(sk)$
$\mathcal{T}(sk)$
Time-Critical Phase
for $i = 1, \ldots, N_c$
pick $R_i \leftarrow \{0,1\}$
pick $T_i \leftarrow \{0,1\}$
Clock: Start
R_i
T_i
Clock: Stop, output Δt
Check $\Delta t \leq t_{\max}$
Slow Phase
Verify signature
$\mathrm{Sign}(sk; R_1||T_1||\ldots||R_{N_c}||T_{N_c})$

Figure 3: The Brands and Chaum protocol.

Intuitively, the time-critical phases must ensure here that the reader-to-tag distance is no greater than the one associated with $t_{\max}$. Finally, the signature in the lazy phase must ensure that the bit exchange was done by a legitimate prover. This ensures both impersonation resistance and a measure of mafia fraud resistance. Lazy phases are susceptible to pure relay (which otherwise taints time-critical rounds), thus the signature yields neither terrorist, nor distance fraud resistance. Unlike most other distance-bounding protocols, however, the tag's responses are fully random, thus the difficulty in mafia fraud attacks is not to answer the time-critical rounds, but rather to make generate the correct signature at the end.We formally state the properties of this protocol below, without considering any further fault tolerance. If the parameters $E_{\max}$ and $T_{\max}$ are considered, upper-bounding the number of erroneous transmissions and respectively the number of transmissions exceeding $t_{\max}$, the security bound for mafia fraud resistance drop by a factor $2^{T_{\max}}$. Note, however, that should *any* of the transmissions go wrong, the signature will not longer verify: for scenarios where transmissions are not reliable, the tag should append the values $R_1, \ldots, R_{N_c}, T_1, \ldots, T_{N_c}$ to the signature.

Theorem 3.1 (Brands-Chaum Properties) *Let* Sign = $(\mathsf{SKg}, \mathsf{SSign}, \mathsf{SVf})$ *be the signature scheme used above. The Brands and Chaum protocol* $\mathcal{ID}$ *above has the following properties:*

- *It is not resistant to terrorist, nor to distance fraud.*
- *For any* $(t, q_{\mathcal{R}}, q_{\mathcal{T}}, q_{\mathrm{OBS}})$*-impersonation adversary* $\mathcal{A}$ *against the Brands Chaum protocol, there exists an adversary* $\mathcal{A}'$ *against the unforgeability of* Sign *that runs in time t and requests at most* $q_{\mathcal{T}} + q_{\mathrm{OBS}}$ *signatures, and that then outputs a valid forgery with an advantage* $\boldsymbol{Adv}^{Unf}_{\mathrm{Sign}}(\mathcal{A}')$ *such that:*

$$\boldsymbol{Adv}_{\mathcal{ID}}^{imp}(\mathcal{A}) \leq \boldsymbol{Adv}_{\mathrm{Sign}}^{Unf}(\mathcal{A}') + \binom{q_{\mathcal{T}} + q_{\mathrm{OBS}}}{2} \cdot 2^{-N_c}.$$

- *For any $(t, q_{\mathcal{R}}, q_{\mathcal{T}}, q_{\mathrm{OBS}})$-Mafia-fraud adversary $\mathcal{A}$ against the Brands Chaum protocol, there exists an adversary $\mathcal{A}'$ against the unforgeability of* Sign *that runs in time t and requests at most $q_{\mathcal{T}} + q_{\mathrm{OBS}}$ signatures, and that then outputs a valid forgery with an advantage $\boldsymbol{Adv}_{\mathrm{Sign}}^{Unf}(\mathcal{A}')$ such that:*

$$\boldsymbol{Adv}_{\mathcal{ID}}^{mafia}(\mathcal{A}) \leq \boldsymbol{Adv}_{\mathrm{Sign}}^{Unf}(\mathcal{A}') + \binom{q_{\mathcal{T}} + q_{\mathrm{OBS}}}{2} \cdot 2^{-N_c} + q_{\mathcal{R}} \cdot 2^{-N_c}.$$

Proof. The second statement follows trivially. Consider a reader-adversary session sid where the adversary successfully impersonates to the reader. Firstly, there is at most a probability of $\binom{q_{\mathcal{T}}+q_{\mathrm{OBS}}}{2} \cdot 2^{-|N_c|}$ that the adversary has observed a legitimate signature in a reader-tag session. Else, we build an adversary $\mathcal{A}'$ against the unforgeability of the signature scheme as follows: $\mathcal{A}'$ does not query the signing oracle of *Sign*, instead running $\mathcal{A}$ and storing the value of $\mathsf{SSign}(sk; R_1||T_1|| \dots ||R_{N_c}||T_{N_c})$. If the impersonation adversary $\mathcal{A}$ has an advantage $\mathbf{Adv}_{\mathcal{ID}}^{\mathrm{imp}}(\mathcal{A})$, then the unforgeability adversary $\mathcal{A}'$ has a probability at least as large to succeed in its attempt.

The first statement also follows easily: the adversary $\mathcal{A}$ against distance fraud simply commits to randomly chosen values of T_i, then computes the (correct) signature with its own secret key sk, and succeeds with probability 1. For terrorist fraud, the adversary trivially generates random time-critical responses, then forwards these values and the values of R_i to the dishonest prover, and receives the signature. The success probability of this adversary is 1. However, the simulator's success probability is upper bounded by the probability that it can forge a signature for a fresh session. For an unforgeable signature scheme, the simulator's success probability is negligible, and thus the scheme is not terrorist fraud resistant.

We now prove the last statement. Consider the adversary $\mathcal{A}$ mounting a Mafia fraud attack. Consider the reader-adversary session sid where $\mathcal{A}$ successfully authenticates to the reader. We look at previous reader-tag sessions and adversary-tag sessions. There exists a probability of $\binom{q_{\mathcal{T}}+q_{\mathrm{OBS}}}{2}$ that the values R_i in sid are all identical to the fast phase values R_i^* in one of these sessions. In this case, the adversary can simply forward the observed/received values T_i^* to the reader in session sid; at the end, the adversary forwards the observed/received value of the signature. We now assume that the reader-session sid is the only session the adversary has access to. The adversary can start a concurrent adversary-tag session sid^*. We call a round successful if the reader input R_i and the response T_i in round i of sid are the same as the adversary input R_i^* and the response T_i^* in sid^*, *without* pure relay, i.e. the adversary must R_i^* in sid^* before receiving R_i in sid.

We first show that except with negligible probability, the adversary cannot win unless he all time-critical rounds are successful. Indeed if one round is not successful (i.e. R_i^* in sid is different from R_i in sid^*) and yet $\mathcal{A}$ authenticates to the reader in sid, we can build an adversary against the unforgeability of the signature scheme. This adversary's forgery is the message m containing the concatenation of all the queries and responses in session sid and the signature that $\mathcal{A}$ forwards in sid. The signature must be correct – else $\mathcal{A}$ cannot authenticate – and the message is fresh, as at least in round i the value of $R_i||T_i$ is fresh. Thus the probability that $\mathcal{A}$ wins without being successful in each round is $\mathbf{Adv}_{Sign}^{\mathrm{Unf}}(\mathcal{A}')$. To this we add the probability that $\mathcal{A}$ guesses the correct signature in each of its $q_{\mathcal{R}}$ attempts.

Finally, the probability that all the N_c rounds are successful is 2^{-N_c} ($\mathcal{A}$ essentially guesses the values of R_i in advance), thus yielding the stated bound. □

3.2 Hancke and Kuhn

One weakness of the Brands and Chaum construction is the use of the digital signature, which cannot be easily implemented on RFID. Also, [4] offers no distance fraud resistance, as the tag's responses are not challenge specific and can be sent in advance. The protocol due to Hancke and Kuhn uses a pseudorandom function (PRF) instead of the signature scheme, implemented as HMAC. It also gains distance-fraud resistance at the cost of a lower mafia fraud resistance. The protocol consists of a lazy phase, where the parties exchange nonces and pre-compute an HMAC value, which is divided in a left and a right half, and N_c time-critical phases where the reader forwards a random bit and the tag responds with a bit either from the left or right half of the PRF output. This is also depicted in Fig. 4.

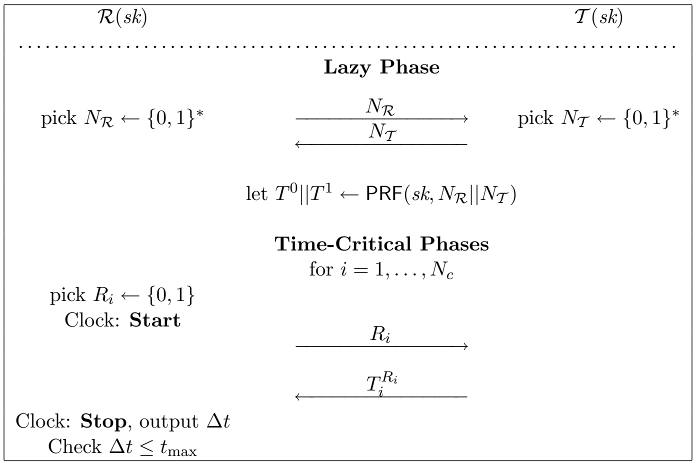

$\mathcal{R}(sk)$
$\mathcal{T}(sk)$
Lazy Phase
pick $N_{\mathcal{R}} \leftarrow \{0,1\}^*$
$N_{\mathcal{R}}$
$N_{\mathcal{T}}$
pick $N_{\mathcal{T}} \leftarrow \{0,1\}^*$
let $T^0||T^1 \leftarrow \mathsf{PRF}(sk, N_{\mathcal{R}}||N_{\mathcal{T}})$
Time-Critical Phases
for $i = 1, \ldots, N_c$
pick $R_i \leftarrow \{0,1\}$
Clock: Start
R_i
$T_i^{R_i}$
Clock: Stop, output Δt
Check $\Delta t \leq t_{\max}$

Figure 4: The Hancke and Kuhn protocol.

The PRF output is $2N_c$ bits, and the left and right halves of this output have equal length. As R_i is chosen randomly, a dishonest tag can't send the responses early, as it cannot predict the reader's challenge bit. On the other hand, the PRF output is unpredictable, thus an adversary without knowledge of *sk* cannot compute it. Though this protocol is more easily implementable on RFID, the lack of reader authentication leads to a decreased Mafia fraud resistance; however, the random challenges provide distance fraud resistance. As in [4], there is no impersonation resistance. We formally state these properties below.

Theorem 3.2 (Hancke Kuhn Properties) *Let $\mathcal{ID}$ be the distance-bounding identification scheme in fig. 4 with parameters $(t_{\max}, N_c)$. This scheme has the following properties:*

- *It is not impersonation resistant, nor resistant to terrorist fraud.*
- *For any $(t, q_{\mathcal{R}}, q_{\mathcal{T}}, q_{\text{OBS}})$-distance-fraud adversary $\mathcal{A}$ against the scheme there exists a (t', q')-distinguisher $\mathcal{A}'$ against* PRF *(where $\mathcal{A}'$ runs in time $t' = t + O(n)$ and makes at most $q' = q_{\mathcal{R}} + q_{\mathcal{T}} + q_{\text{OBS}}$ queries) such that*

$$\boldsymbol{Adv}^{dist}_{\mathcal{ID}}(\mathcal{A}) \leq \boldsymbol{Adv}^{d}_{\mathsf{PRF}}(\mathcal{A}') + \left(\tfrac{3}{4}\right)^{N_c}.$$

- *For any $(t, q_{\mathcal{R}}, q_{\mathcal{T}}, q_{\text{OBS}})$-Mafia-fraud adversary $\mathcal{A}$ against the scheme there exists a (t', q')-distinguisher $\mathcal{A}'$ against* PRF *(where $\mathcal{A}'$ runs in time $t' = t + O(n)$ and makes at most $q' = q_{\mathcal{R}} + q_{\mathcal{T}} + q_{\text{OBS}}$ queries) such that*

$$\boldsymbol{Adv}^{mafia}_{\mathcal{ID}}(\mathcal{A}) \quad \leq \quad q_{\mathcal{R}} \cdot \left(\tfrac{3}{4}\right)^{N_c} + \binom{q_{\mathcal{R}} + q_{\text{OBS}}}{2} \cdot 2^{-|N_{\mathcal{R}}|} + \boldsymbol{Adv}^{d}_{\mathsf{PRF}}(\mathcal{A}') + \binom{q_{\mathcal{T}}}{2} \cdot 2^{-|N_{\mathcal{T}}|}.$$

Proof. The first statement follows easily: an impersonation adversary which simply generates a random nonce during the single lazy phase of this protocol wins with probability 1 as there is no authentication during this lazy phase. For terrorist fraud resistance, a malicious tag equips the adversary during the lazy phase of session sid with the output $T^0||T^1$ of the PRF for nonces $N_{\mathcal{R}}$ and $N_{\mathcal{T}}$. This adversary succeeds with probability 1, as it can answer the queries of the honest reader in sid as though it were the legitimate tag. Now consider a simulator attempting to authenticate with the data obtained from the adversary. However, the simulator's session is fresh and so is the reader's nonce. Thus either the PRF output is different, or we can find a collision in the PRF.

For the second statement, consider an adversary $\mathcal{A}$ against distance resistance. In each time-critical round, this adversary must commit to a value T_i^* before it has received the challenge R_i from the reader. We can assume a truly random distribution of 0 and 1 bits in the PRF output $T^0||T^1$ (otherwise we have a distinguisher $\mathcal{A}'$ that can distinguish the PRF output from a truly random number). For each round $i = 1, \ldots, N_c$, if $T_i^0 = T_i^1$, the tag commits to the correct value with probability 1; if they are unequal, the tag commits to the correct value with a probability of at most $\frac{1}{2}$, giving a total success probability of $\frac{3}{4}$ per round, giving the above statement.

The proof of the last statement consists of the following high-level steps:

1. Show that one can safely replace the PRF runs of honest parties by picking independent random strings $T^0||T^1$ for each new nonce pair $(N_{\mathcal{R}}, N_{\mathcal{T}})$.

2. Show that nonce pairs are (almost) unique, except for possibly one adversary-tag session sid^* having the same nonce pair as a reader-adversary session sid (here the adversary relays the nonces between sessions).

3. Bound the probability that the adversary passes the time-critical phases for at most one adversary-tag interaction.

For the first step we claim that replacing the PRF-values by random (but consistent) values can at most decrease the adversary's success probability by the distinguishing advantage for PRF. This can be seen easily by construction adversary $\mathcal{A}'$ against PRF via black-box simulation of $\mathcal{A}$, each time applying the random or pseudorandom oracle to nonce pairs on behalf of the honest parties. Finally, $\mathcal{A}'$ checks if $\mathcal{A}$ succeeds in some reader-adversary session and outputs 1 if this happens. The distinguishing advantage of $\mathcal{A}'$ then corresponds to the decrease of the success probability of $\mathcal{A}$ when switching to random values $T^0||T^1$.

Next consider the adversary $\mathcal{A}$ mounting a mafia fraud attack and all the pairs of nonces appearing in the attack. Assume that there exist two sessions (between adversary and tag or reader, or between both honest parties) with the same pair $(N_{\mathcal{R}}, N_{\mathcal{T}})$. Then we claim that this can only be a reader-adversary session and an adversary-tag session, except with probability

$$\binom{q_{\mathcal{R}} + q_{\mathrm{OBS}}}{2} \cdot 2^{-|N_{\mathcal{R}}|} + \binom{q_{\mathcal{T}}}{2} \cdot 2^{-|N_{\mathcal{T}}|}.$$

This holds as for each two executions for the reader resp. tag the nonce of this party is picked at random. If three identical nonce pairs appear in some executions then two of them are either in the at most $q_{\mathcal{R}} + q_{\mathrm{OBS}}$ executions with the reader, or in the $q_{\mathcal{T}}$ executions with the tag. Such collisions occur with the above probability.

Declare $\mathcal{A}$ to lose if a collision appears, decreasing its success probability by this negligible term, but allowing us to consider collision-free executions. In particular, except for the matching session, all values $T^0||T^1$ in the attack are independent.

Now consider a reader-adversary session sid in which $\mathcal{A}$ successfully impersonates to $\mathcal{R}$. By assumption the same nonce pair appears in at most one other adversary-tag session. If there exists a (unique) matching adversary-tag session sid^* then we claim that this session taints sid with high probability (if there is no such session, we have the case below, where the adversary does not take advantage of a matching session). Since for this protocol $T_{\max} = 1$, this invalidates session sid. Suppose, to the contrary, that the matching session sid^* taints no time-critical phase in sid.

Consider an untainted time-critical phase of sid where $\mathcal{R}$ sends $R_i = b$ and expects T_i^b. The adversary has thus successfully passed the first $i-1$ time-critical phases and can choose to do one of the following in the i-th phase:

The Go-Early Strategy. In session sid^* the adversary has sent some bit R_i^* to $\mathcal{T}$ before having received R_i (i.e., $\mathsf{clock}(\mathsf{sid}, i+2) > \mathsf{clock}(\mathsf{sid}^*, i+2)$ in the notion of [8]). Then, since R_i is random and independent of all other data, the probability of $R_i^* \neq R_i$ is $\frac{1}{2}$, in which case $\mathcal{A}$ does not receive T_i^* in sid^* and can only guess the value T_i in sid. If $R_i = R_i^*$, however, the adversary returns the correct reply T_i^b with probability 1.

The Go-Late Strategy. In session sid the adversary replies to R_i with some T_i before receiving $(T_i^b)^*$ in session sid^* (i.e., $\mathsf{clock}(\mathsf{sid}, i+3) < \mathsf{clock}(\mathsf{sid}^*, i+3)$ in the notion of [8]). Now $\mathcal{A}$ succeeds only with probability $\frac{1}{2}$ for this phase.

The Modify-it Strategy. The adversary schedules the message such that it receives R_i in sid, sends some R_i^* in sid^*, receives $(T_i^b)^*$ in sid^* and forwards some T_i in sid. Hence, the scheduling corresponds to a pure relay attack, but $R_i \neq R_i^*$ or $T_i \neq (T_i^b)^*$. If R_i^* is wrong then $(T_i^b)^*$ is actually never sent by $\mathcal{T}$ in sid^* and the adversary can thus only guess T_i with probability $\frac{1}{2}$; if $R_i = R_i^*$ then $T_i \neq (T_i^b)^*$ makes the reader reject.

The Taint-it Strategy. The adversary taints this phase of sid through sid^*. This is equivalent here to losing in sid.

The Taint-it strategy may be ignored, as it disables sid. The Go-Late and Modify-it strategies both succeed with probability at most $\frac{1}{2}$. The Go-Early Strategy succeeds with probability $\frac{3}{4}$. As all rounds are independent, and taking into account the $q_{\mathcal{R}}$ trials, this gives the claimed bound. □

3.3 Avoine and Tchamkerten

As noted in the previous section, the Hancke and Kuhn construction has interdependent challenges, allowing the adversary to run a complete number of time-critical exchanges with the tag before attempting to authenticate to the reader. The Avoine and Tchamkerten protocol tries to correct this flaw by providing some reader authentication. The main idea is to store the secret key in one (or more) binary tree(s); the challenges are inter-related, with the responses forming paths from the root to the leaves.

Consider now an adversary impersonating a reader in an adversary-tag session in a Go-Early strategy as in the proof of 3.2. For [11], an adversary can choose challenges R_i^* for each round $i = 1, \ldots, N_c$ and receive responses T_i^* from the tag, having a 50% probability to have queried the honest tag with the correct $R_i^* = R_i$; for these queries, the adversary will know the correct response $T_i = T_i^*$. On the other hand, once an adversary makes an incorrect guess for some R_i^*, none of the future responses will be the correct ones.

This is also depicted in figure 5.

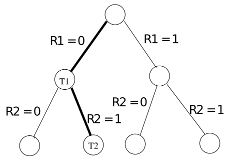
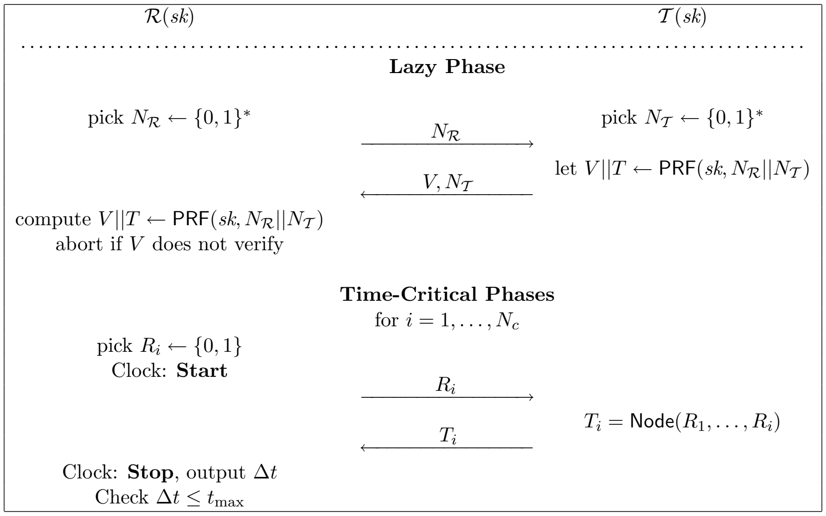


Figure 5: The Avoine-Tchamkerten protocol where the reader sends challenges $R_1 = 0$, $R_2 = 1$.

Note that this protocol requires at least $2^{N_c+1} - 2$ potential response bits; at the expense of security, one could consider more than just one tree. We have, however, lazy phase authentication, leading to a total PRF output of $m + 2^{N_c+1} - 2$ bits for acceptably large m. The first m bits of the PRF output are denoted M; the last $2^{N_c+1} - 2$ bits are denoted T and stored in a tree as follows: from the top downwards and from left to right, each node is labelled with an output bit. The edge between each node and its left child is labelled 0 and the edge to the right child is labelled 1. We denote by $\mathsf{Node}(R_1, \dots, R_i)$ the label of the node that has the path $R_1, \dots, R_i$ from the root. With this notation, the protocol runs as shown in fig. 6. Note in particular that the responses T_i do not correspond directly to the bits of T, but are rather chosen from amongst the bits of T according to the path $R_1, \dots, R_i$.



$\mathcal{R}(sk)$ | $\mathcal{T}(sk)$

Lazy Phase

pick $N_{\mathcal{R}} \leftarrow \{0,1\}^*$ | pick $N_{\mathcal{T}} \leftarrow \{0,1\}^*$

$\xrightarrow{N_{\mathcal{R}}}$

let $V||T \leftarrow \mathsf{PRF}(sk, N_{\mathcal{R}}||N_{\mathcal{T}})$

$\xleftarrow{V, N_{\mathcal{T}}}$

compute $V||T \leftarrow \mathsf{PRF}(sk, N_{\mathcal{R}}||N_{\mathcal{T}})$
abort if V does not verify

Time-Critical Phases
for $i = 1, \dots, N_c$

pick $R_i \leftarrow \{0,1\}$
Clock: **Start**

$\xrightarrow{R_i}$

$T_i = \mathsf{Node}(R_1, \dots, R_i)$

$\xleftarrow{T_i}$

Clock: **Stop**, output Δt
Check $\Delta t \leq t_{\max}$

Figure 6: The Avoine and Tchamkerten protocol.

Intuitively, the m-bit value V offers impersonation resistance. While the inter-dependency between the challenges of the reader increases the protocol's Mafia fraud resistance, this is still not optimal, and it requires a great deal of storage. We formalize the concrete security of this construction below.

Theorem 3.3 (Avoine Tchamkerten Properties) *Let $\mathcal{ID}$ be the distance-bounding identification scheme in fig. 6 with parameters $(t_{\max}, N_c)$. This scheme has the following properties:*

- *It is not resistant to terrorist fraud.*
- *For any $(t, q_{\mathcal{R}}, q_{\mathcal{T}}, q_{\text{OBS}})$-distance-fraud adversary $\mathcal{A}$ against the scheme there exists a (t', q')-distinguisher $\mathcal{A}'$ against PRF (where $t' = t + O(n)$ and $q' = q_{\mathcal{R}} + q_{\mathcal{T}} + q_{\text{OBS}}$) such that*
$$\boldsymbol{Adv}^{dist}_{\mathcal{ID}}(\mathcal{A}) \leq \boldsymbol{Adv}^{d}_{\mathsf{PRF}}(\mathcal{A}') + \left(\tfrac{3}{4}\right)^{N_c}.$$
- *For any $(t, q_{\mathcal{R}}, q_{\mathcal{T}}, q_{\text{OBS}})$-Mafia-fraud adversary $\mathcal{A}$ against the scheme there exists a (t', q')-distinguisher $\mathcal{A}'$ against PRF (where $t' = t + O(n)$ and $q' = q_{\mathcal{R}} + q_{\mathcal{T}} + q_{\text{OBS}}$) such that*
$$\boldsymbol{Adv}^{mafia}_{\mathcal{ID}}(\mathcal{A}) \quad \leq \quad \tfrac{1}{2} q_{\mathcal{R}}[N_c + 2] \cdot 2^{-N_c} + \boldsymbol{Adv}^{d}_{\mathsf{PRF}}(\mathcal{A}') + \binom{q_{\mathcal{R}} + q_{\text{OBS}}}{2} \cdot 2^{-|N_{\mathcal{R}}|} + \binom{q_{\mathcal{T}}}{2} \cdot 2^{-|N_{\mathcal{T}}|}.$$
- *For any $(t, q_{\mathcal{R}}, q_{\mathcal{T}}, q_{\text{OBS}})$-impersonation adversary $\mathcal{A}$ against the scheme there exists a (t', q')-distinguisher $\mathcal{A}'$ against PRF (where $t' = t + O(n)$ and $q' = q_{\mathcal{R}} + q_{\mathcal{T}} + q_{\text{OBS}}$) such that*
$$\boldsymbol{Adv}^{imp}_{\mathcal{ID}}(\mathcal{A}) \quad \leq \quad q_{\mathcal{R}} \cdot 2^{-|V|} + \binom{q_{\mathcal{R}} + q_{\text{OBS}}}{2} \cdot 2^{-|N_{\mathcal{R}}|} + \boldsymbol{Adv}^{d}_{\mathsf{PRF}}(\mathcal{A}') + \binom{q_{\mathcal{T}}}{2} \cdot 2^{-|N_{\mathcal{T}}|}.$$

Proof. The proofs of statements 1 and 3 are identical to those of the Hancke Kuhn protocol.For mafia fraud resistance, we change the Hancke Kuhn proof as follows: for each round, we denote by pass_j the event that an adversary guesses the correct R_i in a Go-Early attack. Again the Go-Early strategy is the most effective, with a success probability given by the iterative expression below:

$$\text{Prob}\left[\bigwedge\nolimits_{j=i}^{N_c}\mathsf{pass}_j \;\middle|\; \bigwedge\nolimits_{j=1}^{i-1}\mathsf{pass}_j\right] \leq \tfrac{1}{2}\cdot\tfrac{1}{2}^{N_c-i+1} + \frac{1}{2}\cdot\text{Prob}\left[\bigwedge\nolimits_{j=i+1}^{N_c}\mathsf{pass}_j \;\middle|\; \bigwedge\nolimits_{j=1}^{i}\mathsf{pass}_j\right].$$

After summing up and iterating, we have the bound above.

Finally, impersonation security follows analogously to that of Dürholz et al. [8]: the success probability is given by the advantage of the distinguisher $\mathcal{A}'$ and the probability that the adversary knows the output of the PRF for the successful reader-adversary session sid (which only happens if it has seen, resp. generated the authentication value V in a reader-tag, resp. adversary-tag session).

□

3.4 Reid et al.

The construction in [3] has better Mafia fraud resistance (but greater data storage) than [11]. The scheme in [3] is additionally impersonation resistant. However, neither scheme is terrorist fraud resistant. We now analyse the protocol due to Reid et al. [14], which adds a symmetric encryption scheme to the Hancke and Kuhn construction (the authors suggest a one-time-pad xor operation). Thus, this protocol inherits the lack of impersonation resistance of [11], as well as the lower Mafia fraud resistance.

We first describe this protocol. In their paper [14] use a so-called key derivation function denoted KDF which can be viewed as a PRF for the sake of simplicity. We thus denote it PRF as for the construction due to Hancke and Kuhn. Furthermore, Reid et al. consider a symmetric IND-CPA encryption scheme denoted $\mathcal{E}$. To this scheme they associate a symmetric ephemeral secret key eph and a long term secret key *sk*. The notation $\mathcal{E}_{\mathsf{eph}}(sk)$ denotes the encryption under eph of the plaintext *sk*. We denote the corresponding decryption process by $\mathcal{D}$. Furthermore, [14] associates to the tag and reader some public identities $\mathcal{ID}_A, \mathcal{ID}_B$.

The main idea here is that both reader and tag compute a symmetric ephemeral key eph as the output of PRF, and then they use eph to encrypt the long-term secret key *sk* with $\mathcal{E}$. For each time-critical round, the reader challenges the tag with a random bit, and the tag responds with either a bit of the encrypted secret key or a bit of eph. Intuitively, terrorist fraud resistance results from the fact that the adversary requires either the long term secret *sk*(which can be later be used by the simulator to authenticate) or both eph and the encryption of *sk* (which can be used by the simulator to compute $\mathcal{D}_{\mathsf{eph}}(sk)$, thus also obtaining the secret *sk*). The scheme is depicted in Figure 7.

Before stating the security properties we note that the informal definition of a successful terrorist fraud attack is that the adversary can authenticate if aided by a malicious tag in a particular session, but cannot authenticate *without* such additional aid. However, this informal definition is deceptive: many protocols claiming to be terrorist fraud resistant are actually resistant to the very restrictive requirement that if the adversary authenticates when aided by the malicious tag, then this facilitates, however little, future authentication sessions. By tying the knowledge (or accurate guessing) of both responses to the secret key, protocols can attain this form of terrorist fraud resistance. In other words, if the prover forwards even a single bit of one of the responses, this is considered trivial help, as it gives the adversary an increased success probability in future rounds.

However, the definition in [8] allows such partial attacks as valid terrorist fraud attacks, concretely requiring that the prover's help does not give the adversary an *equal* advantage for future sessions, i.e. it only excludes attacks where the prover forwards the adversary any fragment of the *secret key itself.* We discuss the merits of both definitions in Section 4.

In the following, we show that this protocol does *not* attain terrorist fraud resistance in the sense of Dürholz et al.

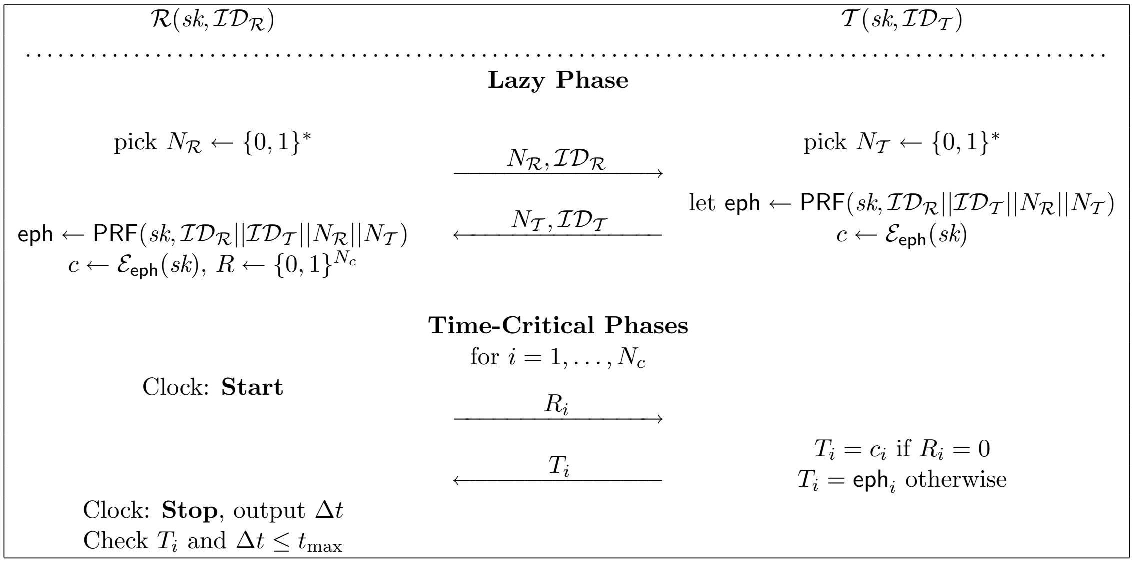

$\mathcal{R}(sk, \mathcal{ID}_{\mathcal{R}})$
$\mathcal{T}(sk, \mathcal{ID}_{\mathcal{T}})$
Lazy Phase
pick $N_{\mathcal{R}} \leftarrow \{0,1\}^*$
pick $N_{\mathcal{T}} \leftarrow \{0,1\}^*$
$\xrightarrow{N_{\mathcal{R}}, \mathcal{ID}_{\mathcal{R}}}$
let $\mathsf{eph} \leftarrow \mathsf{PRF}(sk, \mathcal{ID}_{\mathcal{R}}||\mathcal{ID}_{\mathcal{T}}||N_{\mathcal{R}}||N_{\mathcal{T}})$
$c \leftarrow \mathcal{E}_{\mathsf{eph}}(sk)$
$\xleftarrow{N_{\mathcal{T}}, \mathcal{ID}_{\mathcal{T}}}$
$\mathsf{eph} \leftarrow \mathsf{PRF}(sk, \mathcal{ID}_{\mathcal{R}}||\mathcal{ID}_{\mathcal{T}}||N_{\mathcal{R}}||N_{\mathcal{T}})$
$c \leftarrow \mathcal{E}_{\mathsf{eph}}(sk)$, $R \leftarrow \{0,1\}^{N_c}$
Time-Critical Phases
for $i = 1, \ldots, N_c$
Clock: **Start**
$\xrightarrow{R_i}$
$T_i = c_i$ if $R_i = 0$
$T_i = \mathsf{eph}_i$ otherwise
$\xleftarrow{T_i}$
Clock: **Stop**, output Δt
Check T_i and $\Delta t \leq t_{\max}$

Figure 7: The Reid et al. protocol.

Theorem 3.4 (Reid et al. Properties) *Let $\mathcal{ID}$ be the distance-bounding identification scheme in fig. 7 with parameters $(t_{\max}, N_c)$. This scheme has the following properties:*

- *It is neither impersonation, nor terrorist fraud resistant (assuming the pseudorandomness of* PRF*).*
- *For any $(t, q_{\mathcal{R}}, q_{\mathcal{T}}, q_{\text{OBS}})$-distance-fraud adversary $\mathcal{A}$ against the scheme there exists a (t', q')-distinguisher $\mathcal{A}'$ against* PRF *(where $t' = t + O(n)$ and $q' = q_{\mathcal{R}} + q_{\mathcal{T}} + q_{\text{OBS}}$) or a (t'', q'')-distinguisher $\mathcal{A}''$ against the IND-CPA of $\mathcal{E}$ (where $t'' = t + O(n)$ and $q'' = q_{\mathcal{R}} + q_{\mathcal{T}} + q_{\text{OBS}}$ is the number of requested ciphertexts) such that:*

$$\boldsymbol{Adv}_{\mathcal{ID}}^{imp}(\mathcal{A}) \leq q_{\mathcal{R}} \cdot \left(\tfrac{3}{4}\right)^{N_c} + \boldsymbol{Adv}_{\mathsf{PRF}}^{d}(\mathcal{A}') + \boldsymbol{Adv}_{\mathcal{E}}^{\textit{IND-CPA}}(\mathcal{A}'').$$

- *For any $(t, q_{\mathcal{R}}, q_{\mathcal{T}}, q_{\text{OBS}})$-Mafia-fraud adversary $\mathcal{A}$ against the scheme there exists a (t', q')-distinguisher $\mathcal{A}'$ against* PRF *(where $t' = t + O(n)$ and $q' = q_{\mathcal{R}} + q_{\mathcal{T}} + q_{\text{OBS}}$) or a (t'', q'')-distinguisher $\mathcal{A}''$ against the IND-CPA of $\mathcal{E}$ (where $t'' = t + O(n)$ and $q'' = q_{\mathcal{R}} + q_{\mathcal{T}} + q_{\text{OBS}}$) such that*

$$\boldsymbol{Adv}_{\mathcal{ID}}^{mafia}(\mathcal{A}) \quad \leq \quad \left(\tfrac{3}{4}\right)^{N_c} + \boldsymbol{Adv}_{\mathsf{PRF}}^{d}(\mathcal{A}') + \boldsymbol{Adv}_{\mathcal{E}}^{\textit{IND-CPA}}(\mathcal{A}'') + \binom{q_{\mathcal{R}} + q_{\text{OBS}}}{2} \cdot 2^{-|N_{\mathcal{R}}|} + \binom{q_{\mathcal{T}}}{2} \cdot 2^{-|N_{\mathcal{T}}|}$$

Proof. The proofs of statements 2 and 3 are analogous to the proofs for the protocol due to Hancke and Kuhn [11] and so we do not show them here. The same holds for the proof of impersonation resistance.

In the following we show an adversary $\mathcal{A}$ against terrorist fraud resistance for which there exists no simulator such that $\mathbf{Adv}_{\mathcal{ID}}^{\text{terror}}(\mathcal{A}, \mathcal{S}, \mathcal{T}) \leq 0$. This would therefore show that the scheme is not terrorist fraud resistant. The idea is for the malicious tag $\mathcal{T}'$ to give information that facilitates the adversary's attack, without revealing any essential information about future impersonation attempts. Indeed, let $\mathcal{A}$ receive the value eph from $\mathcal{T}'$ in each of its $q_{\mathcal{R}}$ impersonation attempts. In the subsequent time-critical phases, if the reader sends challenge $R_i = 1$, $\mathcal{A}$ sends $T_i = \mathsf{eph}_i$; else, the adversary guesses T_i. This adversary's probability to win is thus $\frac{3}{4}^{N_c} + \mathbf{Adv}_{\mathcal{E}}^{\text{IND-CPA}}(\mathcal{A}'')$ for the adversary $\mathcal{A}''$ whose advantage to win against the IND-CPA of $\mathcal{E}$ is the largest.

Now consider a simulator $\mathcal{S}$. The simulator has no access to the tag $\mathcal{T}$, but it may run $\mathcal{A}$ internally. However, under the assumption of the pseudorandomness of PRF, there is only a negligible probability that $\mathcal{A}$ knows eph for any of the $q_{\mathcal{R}}$ impersonation sessions where the simulator attempts to authenticate to the reader. Thus, the simulator's probability of winning is $\frac{1}{2}^{N_c} + \mathbf{Adv}_{\mathcal{E}}^{\text{IND-CPA}}(\mathcal{A}'')$. Thus, the adversary has an advantage over any simulator $\mathcal{S}$.

□

3.5 The Swiss-Knife RFID Distance Bounding Protocol

The Swiss-Knife protocol due to Kim et al. [13] aims to achieve privacy as well as mafia, terrorist, and distance fraud resistance. Very notably, the lazy phase of this protocol is divided into two parts: the first precedes the time-critical phase, the second follows it. In the first part, the reader and tag exchange nonces and the tag computes a pseudo-random function (PRF) on input a system constant and a tag-chosen nonce $N_{\mathcal{T}}$. The output of this function, a, is then XORed with the long-term secret key *sk*. In the time-critical rounds, the tag responds with either a or with $a \oplus$ *sk*, depending on the reader's challenge. Note that if this protocol is run in an RFID scenario, the size of *sk*, which equals the number of time-critical rounds, is restricted by the tag's capacity to sustain time-critical rounds. Therefore, the keys are short. Finally, after the time-critical rounds, during the second lazy phase, the tag authenticates by computing the PRF on all the received challenges, its identity, and both the reader and the tag's nonces. The reader may then also authenticate by computing the PRF on input the tag's nonce. This second lazy phase is essential in preventing the recovery of the secret key during a mafia fraud attack. In fact, the fact that the tag computes the second PRF value brings the mafia fraud resistance to loosely $(\frac{1}{2})$ per round.

In the Swiss-Knife scenario, each tag is associated with an identity ID which is stored by the reader in the same database that stores the secret key *sk* of the tag. In order to achieve anonymity, this identity is never sent, and the reader needs to search the database exhaustively to find it. This protocol also has some fault tolerance, i.e. the reader counts a total number of errors consisting of: (1) the number of faulty challenges R_i that the tag receives; (2) the number of faulty responses T_i that the reader receives; and (3) the number of rounds in which the tag's response exceeds the time threshold $t_{\max}$. The protocol is depicted in figure 8. Note that Kim et al. also present a more efficient version, but whereas this second scheme is computationally more efficient than the simplified one, the security properties are comparable. In figure 8, the value const is a system constant.

Theorem 3.5 (Swiss-Knife Properties) *Let $\mathcal{ID}$ be the distance-bounding identification scheme in fig. 7 with parameters $(t_{\max}, N_c)$. This scheme has the following properties:*

- *It is not terrorist fraud resistant (assuming the pseudorandomness of* PRF*).*
- *For any $(t, q_{\mathcal{R}}, q_{\mathcal{T}}, q_{\text{OBS}})$-distance-fraud adversary $\mathcal{A}$ against the scheme there exists a (t', q')-distinguisher $\mathcal{A}'$ against* PRF *(where $t' = t + O(n)$ and $q' = q_{\mathcal{R}} + q_{\mathcal{T}} + q_{\text{OBS}}$) such that:*

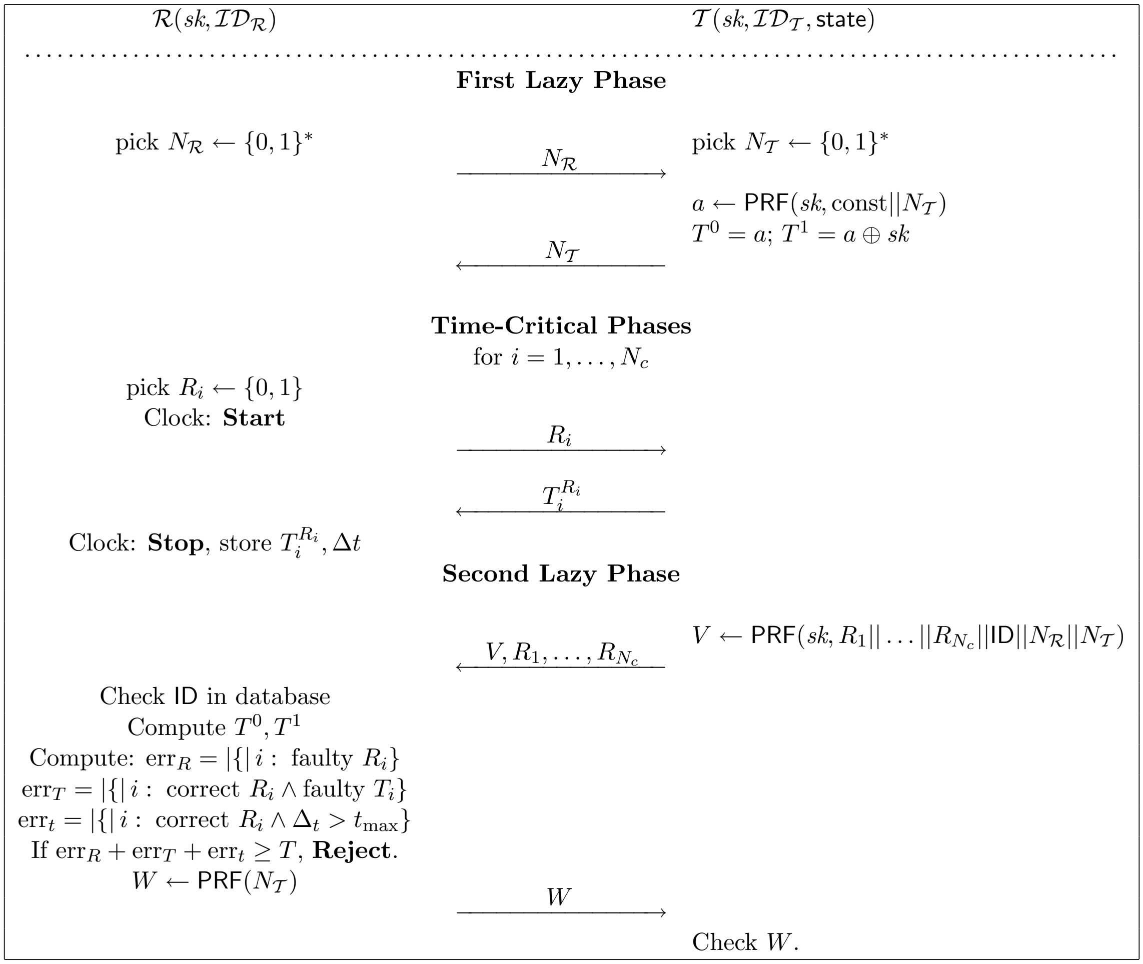
$$\boldsymbol{Adv}_{\mathcal{ID}}^{imp}(\mathcal{A}) \leq q_{\mathcal{R}} \cdot \left(\tfrac{3}{4}\right)^{N_c - T} + \boldsymbol{Adv}_{\mathsf{PRF}}^{d}(\mathcal{A}').$$

- *For any $(t, q_{\mathcal{R}}, q_{\mathcal{T}}, q_{\text{OBS}})$-Mafia-fraud adversary $\mathcal{A}$ against the scheme there exists a (t', q')-distinguisher $\mathcal{A}'$ against* PRF *(where $t' = t + O(n)$ and $q' = q_{\mathcal{R}} + q_{\mathcal{T}} + q_{\text{OBS}}$) such that*

$$\boldsymbol{Adv}_{\mathcal{ID}}^{mafia}(\mathcal{A}) \quad \leq \quad \left(\tfrac{1}{2}\right)^{N_c - T} + \binom{q_{\mathcal{R}} + q_{\text{OBS}}}{2} \cdot 2^{-(|N_{\mathcal{R}}| + \lceil \frac{N_c}{2} \rceil - T)} + \boldsymbol{Adv}_{\mathsf{PRF}}^{d}(\mathcal{A}') + \binom{q_{\mathcal{T}}}{2} \cdot 2^{-(|N_{\mathcal{T}}| + \lceil \frac{N_c}{2} \rceil - T)}.$$

- *For any $(t, q_{\mathcal{R}}, q_{\mathcal{T}}, q_{\text{OBS}})$-impersonation adversary $\mathcal{A}$ against the scheme there exists a (t', q')-distinguisher $\mathcal{A}'$ against* PRF *(where $t' = t + O(n)$ and $q' = q_{\mathcal{R}} + q_{\mathcal{T}} + q_{\text{OBS}}$) such that*

$$\boldsymbol{Adv}_{\mathcal{ID}}^{imp}(\mathcal{A}) \quad \leq \quad q_{\mathcal{R}} \cdot 2^{-|V|} + \binom{q_{\mathcal{R}} + q_{\text{OBS}}}{2} \cdot 2^{-(|N_{\mathcal{R}}| + N_c - T)} + \boldsymbol{Adv}_{\mathsf{PRF}}^{d}(\mathcal{A}') + \binom{q_{\mathcal{T}}}{2} \cdot 2^{-(|N_{\mathcal{T}}| + N_c - T)}.$$



$\mathcal{R}(sk, \mathcal{ID}_{\mathcal{R}})$
$\mathcal{T}(sk, \mathcal{ID}_{\mathcal{T}}, \mathsf{state})$
First Lazy Phase
pick $N_{\mathcal{R}} \leftarrow \{0,1\}^*$
$N_{\mathcal{R}}$
pick $N_{\mathcal{T}} \leftarrow \{0,1\}^*$
$a \leftarrow \mathsf{PRF}(sk, \text{const}||N_{\mathcal{T}})$
$T^0 = a$; $T^1 = a \oplus sk$
$N_{\mathcal{T}}$
Time-Critical Phases
for $i = 1, \ldots, N_c$
pick $R_i \leftarrow \{0,1\}$
Clock: Start
R_i
$T_i^{R_i}$
Clock: Stop, store $T_i^{R_i}, \Delta t$
Second Lazy Phase
$V \leftarrow \mathsf{PRF}(sk, R_1|| \ldots ||R_{N_c}||\mathsf{ID}||N_{\mathcal{R}}||N_{\mathcal{T}})$
$V, R_1, \ldots, R_{N_c}$
Check ID in database
Compute T^0, T^1
Compute: $\text{err}_R = |\{| i: \text{ faulty } R_i\}$
$\text{err}_T = |\{| i: \text{ correct } R_i \wedge \text{faulty } T_i\}$
$\text{err}_t = |\{| i: \text{ correct } R_i \wedge \Delta_t > t_{\max}\}$
If $\text{err}_R + \text{err}_T + \text{err}_t \geq T$, Reject.
$W \leftarrow \mathsf{PRF}(N_{\mathcal{T}})$
W
Check W.

Figure 8: The Swiss-Knife Protocol.

Proof. The proof of statement 1 is trivial: the malicious tag can even send the adversary the secret key *sk* and the output V for the second lazy phase. However, the simulator cannot guess the value of ID except with probability $2^{|\mathsf{ID}|}$.

For the second statement, the proof works as for the previous protocols, except for the fault tolerance level T.

For the third statement, the Go-Early strategy now becomes the only possible strategy: however, if the adversary fails to guess any of $N_c - T$ values R_i, the PRF output in the second lazy phase generated by the tag will not be accurate. We also need to account, even for sessions with matching $N_{\mathcal{R}}$ and $N_{\mathcal{T}}$, for about half the challenges in time-critical rounds.

This last argument also accounts for the impersonation resistance expression in the fourth session. □

4 The Case for Terrorist Fraud Resistance

In this paper, we prove that two constructions which claim to achieve terrorist fraud resistance are in fact *not* terrorist fraud resistant in the framework due to Dürholz et al. Additionally, it appears that intuitive countermeasures against terrorist fraud resistance do not work, as *partial* help from the dishonest prover gives the adversary some advantage over the simulator. Our results may be viewed from two separate points of view. It can be argued, on the one hand, that the model due to Dürholz et al. is too strong, and does not accurately capture the notion of terrorist fraud resistance. On the other hand, our results may be viewed as proof that terrorist fraud resistance is in fact a very powerful attack, which is difficult to counteract in practice. We present and assess both points of view in the considerations below.

Model strength. As noted in section 3.4, the notion achieved by [14] is very weak in the sense that it excludes even prover information that significantly aids adversary authentication while disclosing a relatively insignificant *part* of the secret key. We note that previous definitions, such as the one in Avoine et al.'s framework [2], are ambiguous regarding this point. In fact, Avoine et al. requires, literally, that the prover's help gives the adversary no advantage in future attempts. It is unclear, however, what "further" means in this context: does it refer to the success probability of the adversary *after* the prover helped it, compared to the adversary's success *before* the prover helped it, or rather to the notion captured by [8], i.e. the success probability of the adversary *after* the prover helped it compared to *while* the prover helped it?

If we take the former, weaker notion, the protocol due to Reid et al. [14] is intuitively terrorist fraud resistant. However, we point out that no formal definition in the literature covers the weaker definition of terrorist fraud resistance presented informally above. Thus, it is difficult to say how secure these protocols are, nor how they compare in terms of adversarial advantage.

A further question is which definition best captures the intuition behind terrorist fraud attacks. A strong degree of terrorist fraud resistance is always more desirable, thus from this point of view the definition due to Dürholz et al. sets the standard for protocol design. On the other hand, this definition seems hard to achieve, as it enables attacks where some indirect information about the key is forwarded to the adversary (as in sections 3.4 and 3.5).

The intuition of terrorist fraud resistance is that the malicious prover is willing to assist the adversary in its authentication attempt, but wants to control his access. Thus, the adversary should not be able to authenticate without the prover's help. We note, however, that the adversary always has some (usually negligible in the number of time-critical rounds) probability of authenticating without the prover's help: this is equivalent to the probability that he guesses the correct replies or, equivalently, that he guesses the secret key.

How far does the model in [8] cover this intuitive notion? Dürholz et al. quantify the adversary's success probability in the presence of the malicious prover, and then the simulator's probability (where the simulator does not have access to the prover, only somewhat to the adversary in the state when communicating with the tag). The scheme is considered terrorist fraud resistant if the simulator's probability of success equals (or is greater than) the adversary's probability of success. In other words, an attack is successful if the prover's help enables the adversary to succeed in one session with some probability, but this probability diminishes in future sessions, when the prover is no longer available. In this scenario the prover has the guarantee that the adversary will only be able to authenticate (afterwards) with less probability. This definition also seems too strong, in the sense that Dürholz et al. accept an attack where the prover authenticates with probability 75% (3 out of 4 times), but the simulator can only authentication with probability 50% (1 out of 2 times). This contradicts the spirit of terrorist fraud resistance as it is understood in the literature.

A middle way would be to define a so-called tolerance level for the simulator, i.e. accept attacks as long as the simulator's success probability does not exceed this tolerance level. Note, however, that the attack presented in section 3.4 can be tweaked so that the adversary still has an advantage over the simulator, whereas the simulator succeeds with a probability within the tolerance level (instead of giving half the response, the prover would forward only a number of bits of this response, thus easing the adversary's job).

It is our opinion that the notion described by Dürholz et al., though strong, does capture the intuition of terrorist fraud resistance better than the weaker definition which these protocol seem to attain. A common approach in security is to be conservative and to ask for strong(er) security, rather than to label insecure protocols as secure.

Constructive aspects. A second perspective in which to view our result is a constructive one, i.e. if we consider that the model by Dürholz et al. captures the correct notion of terrorist fraud resistance, then clearly achieving this definition requires a stronger construction. One might argue that the strong requirement posed by the model in [8] would lead to inefficient constructions. We argue, however, that the notion of terrorist fraud resistance, is, in its own right, a very strong notion: here, the (dishonest) prover *helps* the adversary authenticate. The challenge is thus to ensure that *any* information leaked to the adversary automatically will carry over to the simulator.

We also note that there is a clear separation between distance-bounding realizations for RFID and for more powerful devices. Indeed, terrorist fraud resistance might be more easily achieved if it is possible to use, say, public key cryptography. In this sense, we could wonder how realistic a threat terrorist fraud attacks are on RFID systems and whether it is worth addressing them directly in protocol design. With RFID tags used in the pharmaceutical industry, in general logistics, and in public transport [5, 9], it seems quite likely that terrorist fraud attacks are quite likely in practice in these settings. In fact, RFID systems are also used in airport security in many German airports: impersonation MITM attacks have already been mounted on these systems by the Chaos Computer Club (CCC) [10]. Though these attacks were not real-time relay attacks, the incentive to mount mafia and terrorist fraud attacks on RFID authentication protocols is rather high. It remains an open question whether RFID systems can be efficiently protected against terrorist fraud in practice, however. The results in this paper show that terrorist fraud resistance is not trivial to achieve, and that achieving it may be inefficient for RFID devices. As terrorist fraud resistance is, however, both a very strong, and a very desirable goal, the authors of this paper interpret their results as an incentive to construct protocols that *are*, in fact, terrorist fraud resistant in the notion of Dürholz et al.

5 Conclusions

We finally recapture the tools used by various constructions to attain distance-bounding security goals, particularly for RFID scenarios, and look again especially at the hardness of achieving terrorist fraud resistance.

Clearly, the Mafia-fraud resistance of a protocol depends on the success rate of the Go-Early strategy: in particular, the adversary should not be able to use information leaked in advance from the prover. The adversary's success rate is high for the Hancke and Kuhn protocol [11] (which is vulnerable to Go-Early attacks). To address this, the protocol due to Kim and Avoine [12], whose security is assessed in [8], is to give some form of reader authentication during the time-critical rounds: thus, the honest tag has a higher chance of detecting an adversary leeching its responses. Creating a dependency between the responses as in [3] is also a solution, but this protocol involves exponential storage. Finally, the solution due to Kim et al. [13] seems quite elegant: by authenticating the challenges after they are sent, the tag makes sure that even if the adversary leeches some challenges, it can only receive the correct authentication response if it has guessed *all* the challenges in advance. This latter solution, however, involves a second PRF computation.

The partial reader authentication used in [12] diminishes distance-fraud resistance. This follows because the reader authentication introduces a degree of predictability in the tag's responses. Thus the malicious tag in the distance fraud scenario is able to prepare for some of the responses in advance, and can guess the other responses.

As outlined in the previous section, it remains an open question to find effective means to achieving terrorist-fraud resistance. Creating an interdependency between the secret key and the tag responses has thus far been the advertised strategy towards attaining terrorist fraud resistance. However, the attack we show in this paper against Reid et al.'s protocol [14] can be used against *any* scheme. It also remains an open question whether terrorist fraud resistance is, in fact, attainable in RFID distance-bounding scenarios.

References

[1] Abyneh, M.R.S.: Security analysis of two distance-bounding protocols. In: Proceedings of RFIDSec 2011. Lecture Notes in Computer Science, Springer (2011)

[2] Avoine, G., Bingol, M.A., Karda, S., Lauradoux, C., Martin, B.: A formal framework for cryptanalyzing rfid distance bounding protocols. `http://eprint.iacr.org/2009/543.pdf` (2009)

[3] Avoine, G., Tchamkerten, A.: An efficient distance bounding rfid authentication protocol: Balancing false-acceptance rate and memory requirement. In: Information Security. Lecture Notes in Computer Science, vol. 5735, pp. 250–261. Springer-Verlag (2009)

[4] Brands, S., Chaum, D.: Distance-bounding protocols. In: Advances in Cryptology — Eurocrypt'93. pp. 344–359. Lecture Notes in Computer Science, Springer-Verlag (1993)

[5] Cole, P.H., Ranasinghe, D.C.: Networked RFID Systems and Lightweight Cryptography. Springer-Verlag (2008)

[6] Cremers, C., Rasmussen, K.B., Čapkun, S.: Distance hijacking attacks on distance bounding protocols. Cryptology ePrint Archive, Report 2011/129 (2011), ePRINTURL

[7] Desmedt, Y.: Major security problems with the 'unforgeable' (feige)-fiat-shamir proofs of identity and how to overcome them. In: SecuriCom. pp. 15–17. SEDEP Paris, France (1988)

[8] Dürholz, U., Fischlin, M., Kasper, M., Onete, C.: A formal approach to distance bounding RFID protocols. In: Proceedings of the 14th Information Security Conference ISC 2011. pp. 47–62. Lecture Notes in Computer Science, Springer-Verlag (2011)

[9] Editors, Zhang, Y., Kitsos, P.: Security in RFID and Sensor Networks. CRC Press (2009)

[10] H-Security, T.: Chip-based ID cards pose security risk at airports. `http://www.h-online.com/security/news/item/Chip-based-ID-cards-pose-security-risk-at-airports-905662.html` (2010)

[11] Hancke, G.P., Kuhn, M.G.: An rfid distance bounding protocol. In: SECURECOMM. pp. 67–73. ACM Press (2005)

[12] Kim, C.H., Avoine, G.: Rfid distance bounding protocol with mixed challenges to prevent relay attacks. In: Proceedings of the 8th International Conference on Cryptology and Networks Security (CANS 2009). Lecture Notes in Computer Science, vol. 5888, pp. 119–131. Springer-Verlag (2009)

[13] Kim, C.H., Avoine, G., Koeune, F., Standaert, F.X., Pereira, O.: The swiss-knife RFID distance bounding protocol. In: Proceedings of the 14th Information Security Conference ISC 2011. pp. 98–115. Lecture Notes in Computer Science, Springer-Verlag (2009)

[14] Reid, J., Nieto, J.M.G., Tang, T., Senadji, B.: Detecting relay attacks with timing-based protocols. In: ASIACCS. pp. 204–213. ACM Press (2007)