Terrorism in Distance Bounding: Modeling Terrorist-Fraud Resistance

Marc Fischlin and Cristina Onete

CASED & Technische Universität Darmstadt
www.cryptoplexity.de

Abstract. In distance-bounding protocols, verifiers use a clock to measure the time elapsed in challenge-response rounds, thus upper-bounding their distance to the prover. This should prevent man-in-the-middle (MITM) relay attacks. Distance-bounding protocols may aim to prevent several attacks, amongst which terrorist fraud, where a dishonest prover helps the adversary to authenticate, but without passing data that allows the adversary to later authenticate on its own. Two definitions of terrorist-fraud resistance exist: a very strong notion due to Dürholz et al. [6] (which we call SimTF security), and a weaker, fuzzier notion due to Avoine et al. [1]. Recent work [7] indicates that the classical countermeasures to terrorist fraud, though intuitively sound, do *not* grant SimTF security. Two questions are posed in [7]: (1) Is SimTF security achievable? and (2) Can we find a definition of terrorist-fraud resistance which both captures the intuition behind it *and* enables efficient constructions?

We answer both questions affirmatively. For (1) we show the first provably SimTF secure distance-bounding scheme in the literature, though superior terrorist-fraud resistance comes here at the cost of security. For (2) we provide a game-based definition for terrorist-fraud resistance (called GameTF security) that captures the intuition suggested in [1], is formalized in the style of [6], and is strong enough for practical applications. We also prove that the SimTF-insecure [7] Swiss-Knife protocol *is* GameTF-secure. We argue that high-risk scenarios require a stronger security level, closer to SimTF security. Our SimTF secure scheme is also strSimTF secure.

1 Introduction

Authentication protocols, run between a *prover* and a *verifier*, allow the verifier to either accept the prover as legitimate or reject it if it is illegitimate. Authentication is used in e.g. public transport, Passive Keyless and Start (PKES) systems, and personal identification. Secure authentication schemes must prevent impersonation attacks, i.e. the verifier must always reject illegitimate provers. However, security models in authentication do not usually capture man-in-the-middle (MITM) relay attacks, where an adversary authenticates by just forwarding data between the prover and verifier. Such attacks, called mafia fraud [4], have been implemented in various application scenarios like Bluetooth [15,9], smart- and RFID cards [5,10,13], e-Passports [12], e-voting [16], and PKES [8].

M. Jacobson et al. (Eds.): ACNS 2013, LNCS 7954, pp. 414–431, 2013.

Introduced in [3], distance bounding detects mafia fraud, or rather, the delay caused by relays in the MITM adversary. Here the verifier uses a clock to upper-bound its (communication) distance to the prover, by measuring the time elapsed between sending a challenge and receiving the response. If the roundtrip time is at most equal to a threshold $t_{\max}$, the response is *in time*, presumably sent by a prover in the verifier's *proximity*. Thus, $t_{\max}$ denotes a maximum trusted distance to the verifier, which can be a few millimeters, some centimeters, or more. Time measurements are usually round-based; most protocols consist of rounds (or phases [6]), which are either *lazy* (slow) —if the clock is *not* used— or *time-critical* (fast) —if the clock measures time-of-flight. The digital-analog system in [17] ensures that distance-bounding protocols can be implemented in practice, detecting pure relays for up to 41 cm. Many distance-bounding protocols are designed for resource-constrained devices, e.g. RFID tags.

In this paper we focus on one of the four main goals of distance-bounding protocols, namely *terrorist fraud* resistance. Terrorist fraud is an attack where the MITM adversary is helped by a dishonest prover to authenticate (but this help should *not* allow the adversary to authenticate later). For example, simply passing the secret key is prohibited, but revealing some secret information which can be used in a single execution is admissible. Two previous frameworks [1,6] define this attack differently. This controversy is unfortunately not unique in the area of distance bounding, where, though the intuition behind the security model has been known for decades, the formalization of it is still debatable. No previous definition of terrorist-fraud resistance seems quite "right", being either too weak or too strong, depending on the (limitations of the) adversary's power. Essentially, there are two main model features which limit the adversary's power: its interaction with the prover (should it be just in slow, or also in fast phases?), and the restriction on the prover do to help. Both existing frameworks [1,6] allow the prover and adversary to interact only in lazy phases (however, we argue that restriction is unnecessary and artificial). Furthermore, while [1] greatly restrict the prover and dismiss most attacks (provers may only forward data that leaves the secret key statistically hidden [2]), the model of [6] allows the prover to forward almost any data, thus excluding very few attacks (the prover can even send bits of the key if the adversary can use them more in the session where the prover helps than in later sessions). We argue that, while the former model allows very efficient constructions, it is too weak in the sense that it might not prevent real attacks. Yet, the latter notion is too strong in the sense that it is not attained by schemes employing classical (and intuitively effective) countermeasures to terrorist fraud.[1].

[1] Concretely, the attack in [7] is aimed at the protocols of Reid et al. [18] and the Swiss-Knife protocol [14]. In fact, slightly modified versions of these protocols are used, since the circular dependency between the secret key and the time-critical responses in the original schemes makes it hard to prove mafia and impersonation resistance. In [7] a single instance of the secret key *sk* is replaced in each protocol by another key sk^*. Yet, the terrorist attack in [7] works against the original, as well as the modified schemes.

Contributions. In this paper we answer the following questions, posed by [7]:

1. Can the definition of Dürholz et al. actually be achieved?
2. Can we "rightly" define terrorist-fraud resistance, such that we capture the intuition *and* enable efficient constructions?

We mainly focus on (2), but we also answer question (1) affirmatively. We prove that the challenging notion of [6] (which we call SimTF security, because it uses a simulation-based definition) is achievable. Yet, in order to attain SimTF security, our protocol (the first SimTF secure scheme in the literature) becomes more vulnerable to other attacks. This may indicate that SimTF security cannot be achieved *efficiently*. Our scheme modifies the Swiss-Knife protocol [14], introducing a "back door" for the simulator, which can authenticate either by learning the long-term secret (from the adversary's state) or by luck (the verifier accepts an incorrect authentication string with some probability). Our scheme inherits the mafia and distance-fraud resistance of the Swiss-Knife protocol, which many protocols lack [7], but due to the "back door" for proving SimTF with decreased security levels.

In answer to (2) we propose a sufficiently strong, game-based notion of terrorist-fraud resistance, called GameTF-security. We start from the intuition of [1], but formalize it as in [6], striving towards a unified security framework. A protocol is GameTF-secure if any adversary authenticating with the prover's help can authenticate unaided with better-than-mafia-fraud probability. This notion also captures the intuition of terrorist-fraud resistance: it requires that the information gained from the prover during the terrorist attack (which constitutes the terrorist adversary's state) will not lead, once the prover stops helping, to an authentication probability higher than for a mafia adversary. Note that the mafia-fraud success probability is a natural lower bound for the unaided adversary, since, once the prover stops helping, the adversary finds itself exactly in the MITM mafia scenario, with only its state to give it any advantage. This notion captures the exact intuition behind terrorist fraud and indeed, we can prove that the SimTF insecure, modified Swiss-Knife protocol [7], *is* GameTF secure (as intuition indicates it should be).

Our GameTF notion is strong enough for, e.g., public transport ticketing mechanisms. Yet, terrorist fraud affects high-security applications like e-Passports and e-voting much more (see discussion in Section 5); thus stronger definitions are needed. We propose a natural extension of SimTF-security, where adversaries also access the prover online, during the authentication attempt (excepting relay scheduling, of course). Our *strong simulation-based terrorist-fraud model* (strSimTF) is stronger than SimTF security, but also achievable: in fact, our SimTF-secure scheme is also strSimTF-secure.

For completeness, we also give a full security diagram featuring our notions and SimTF security. Interestingly, our strSimTF and GameTF models are independent of each other; however, a scheme that is strSimTF-secure *and* mafia-fraud resistant is also GameTF-secure. We also show that, though our GameTF definition resembles the notion in [1], it does *not* imply mafia-fraud resistance (as [1] argues). The full diagram appears in Fig. 3.

2 Preliminaries

We first review the terminology of [6], particularly terrorist fraud (SimTF) resistance. The setting we consider is that of a single prover $\mathcal{T}$ and a single verifier $\mathcal{R}$, sharing a secret key *sk* generated by an algorithm Kg.[2] In the RFID setting, the provers are RFID *tags* and the verifier is a *reader*; this is the terminology used in [6]. The reader has a clock and stores *sk* in an internal database. The interaction between $\mathcal{T}$ and $\mathcal{R}$, i.e. the protocol, is run in phases, which are either *time-critical* (if $\mathcal{R}$ measures roundtrip times, matching them against a threshold $t_{\max}$), or *lazy* (if the clock is not used). The following *timing parameters* are considered: the number N_c of time-critical phases; the threshold roundtrip time $t_{\max}$; the number $T_{\max}$ of time-critical phases that may exceed $t_{\max}$; and the number $E_{\max}$ of time-critical phases with erroneous responses[3].

In [6], $\mathcal{T}$ and $\mathcal{R}$ interact in *sessions*, indexed by session id's sid and associated with transcripts containing all the exchanged messages in sid. For mafia and terrorist fraud, sessions are run between 2 out of these 3 parties: the tag $\mathcal{T}$, the reader $\mathcal{R}$, and a MITM adversary $\mathcal{A}$. In *reader-tag* sessions, $\mathcal{A}$ observes honest prover-verifier interaction. In *adversary-tag sessions*, $\mathcal{A}$ interacts with the honest $\mathcal{T}$, impersonating a reader. In *reader-adversary* sessions, $\mathcal{A}$ impersonates the prover to $\mathcal{R}$. In reader-tag sessions, $\mathcal{A}$ may not interfere with the protocol run; to run a MITM attack, $\mathcal{A}$ opens parallel reader-adversary and adversary-tag sessions. We quantify the adversary in terms of its runtime t and the number of sessions it runs, i.e. q_{OBS} reader-tag, $q_{\mathcal{R}}$ reader-adversary, and $q_{\mathcal{T}}$ adversary-tag sessions. The advantage ϵ of $\mathcal{A}$ is its success probability (see below).

As in [6], we denote messages i to j exchanged in session sid by $\Pi_{\mathsf{sid}}[i \dots j]$, while $\Pi_{\mathsf{sid}}[1 \dots]$ denotes *all* the messages exchanged in sid. An abstract, *universal* clock variable clock (distinct from the reader's *local* clock) keeps track of the order in which messages are sent. The integer $\mathsf{clock}(\mathsf{sid}, k)$ is assigned to the k-th protocol message, which is delivered in session sid to an honest party. This party's reply is associated with $\mathsf{clock}(\mathsf{sid}, k+1) = \mathsf{clock}(\mathsf{sid}, k) + 1$ (i.e. clock is augmented by 1). If the adversary opens two parallel sessions, then $\mathsf{clock}(\mathsf{sid}, k) < \mathsf{clock}(\mathsf{sid}^*, k)$ if $\mathcal{A}$ sends the k-th message in session sid^* after the k-th message in session sid.

Mafia fraud. In [6], each attack is defined by restricting the adversary's interactions to a number of allowed *tainted* phases. In mafia fraud, a phase is *tainted* if pure relaying takes place (in reality this is detected by the clock). The adversary can taint at most $T_{\max}$ rounds, thus accounting for expected transmission delays; in practice, $T_{\max}$ should be very low. More formally [6]:

[2] Though distance bounding is usually run in a symmetric setting, our results extend to public-key settings too.

[3] The values $T_{\max}$ and $E_{\max}$ are not classical parameters in distance bounding, but were introduced in [6] to account for unreliable time-critical transmissions. Also note that Dürholz et al. use a misnomer (also often found in the literature) in talking about *identification* rather than *authentication* schemes: indeed, the protocols output an accept/reject bit, not an identity.

Definition 1 (Tainted Time-Critical Phase, [6]). *A time-critical phase* $\Pi_{\mathsf{sid}}[k \dots k+2\ell-1] = (m_k, \dots, m_{k+2\ell-1})$ *for* $k, \ell \geq 1$ *of a reader-adversary session* sid*, with the k-th message being received by the adversary, is* tainted by the phase $\Pi_{\mathsf{sid}^*}[k \dots k+2\ell-1] = (m^*_k, \dots, m^*_{k+2\ell-1})$ of an adversary-tag session sid^* *if for all* $i = 0, 1, \dots, \ell-1$ *we have:*

$$\begin{aligned} (m_k, \dots, m_{k+2\ell-1}) &= (m^*_k, \dots, m^*_{k+2\ell-1}), \\ \mathit{clock}(\mathsf{sid}, k+2i) &< \mathit{clock}(\mathsf{sid}^*, k+2i), \\ \textit{and} \qquad \mathit{clock}(\mathsf{sid}, k+2i+1) &> \mathit{clock}(\mathsf{sid}^*, k+2i+1). \end{aligned}$$

Insight: pure relay. The definition excludes *only* pure relay: exact messages sent in the same order between sessions; thus an adversary who receives from $\mathcal{R}$ some input challenge bit b is allowed to flip this bit and relay it to the prover, then relaying the response. In practice, this method can be used against protocols where the computation for one input bit (say $b = 1$) is easier than for the other; in this case, $\mathcal{A}$ can fool the clock by using the faster computation. Since communication is usually very fast in distance bounding, computation delays are very significant. Dürholz et al. restrict mafia adversaries only *minimally*: they assume that the reader's clock only detects same-message relays between parties.

Definition 2 (Mafia Fraud Resistance). *For a distance-bounding authentication scheme* $\mathcal{ID}$ *with parameters* $(t_{\max}, T_{\max}, E_{\max}, N_c)$*, a* $(t, q_{\mathcal{R}}, q_{\mathcal{T}}, q_{\mathrm{OBS}})$*-mafia-fraud adversary* $\mathcal{A}$ wins *against* $\mathcal{ID}$ *if the verifier accepts in a reader-adversary session* sid *such that any adversary-tag session* sid^* *taints at most* $T_{\max}$ *time-critical phases of* sid. *Let* $\boldsymbol{Adv}^{\mathit{mafia}}_{\mathcal{ID}}(\mathcal{A})$ *denote the probability that* $\mathcal{A}$ *wins.*

We say $\mathcal{ID}$ is *mafia-fraud resistant* if any efficient mafia-fraud adversary has at most a negligible advantage to win.

The SimTF notion. In the terrorist fraud resistance notion in [6] (here called SimTF-security), the adversary may not interact with the prover during time-critical phases *at all.* This is reflected in the definition below, which states that if $\mathcal{A}$ and the malicious $\mathcal{T}'$ interact, the phase is tainted.

Definition 3 (Tainted Time-Critical Phase (SimTF)). *A time-critical phase* $\Pi_{\mathsf{sid}}[k \dots k+2\ell-1] = (m_k, \dots, m_{k+2\ell-1})$ *for* $k, \ell \geq 1$ *of a reader-adversary session* sid*, with the k-th message being received by* $\mathcal{A}$*, is* tainted *if there exists a session* sid' *between* $\mathcal{A}$ *and* $\mathcal{T}'$ *such that, for some* i*,*

$$\mathit{clock}(\mathsf{sid}, k) < \mathit{clock}(\mathsf{sid}', i) < \mathit{clock}(\mathsf{sid}, k+2\ell-1).$$

SimTF security is defined in terms of a simulator: once an adversary $\mathcal{A}$ authenticates in a reader-adversary session, its transcripts and randomness (i.e. the view $\mathsf{view}_{\mathcal{A}}$ of $\mathcal{A}$) are passed to a simulator $\mathcal{S}$ which must authenticate, by only using $\mathsf{view}_{\mathcal{A}}$, with at least as much probability. Thus, if the adversary requests (a part of) the secret key, this information is passed on to the simulator.

Definition 4 (SimTF security, [6]). *Let $\mathcal{ID}$ be an authentication scheme for parameters $(t_{\max}, T_{\max}, E_{\max}, N_c)$. Let $\mathcal{A}$ be a $(t, q_{\mathcal{R}}, q'_{\mathcal{T}})$-*SimTF *adversary, $\mathcal{S}$ be an algorithm with runtime $t_{\mathcal{S}}$, and $\mathcal{T}'$ be an algorithm with runtime t'. Let*

$$\boldsymbol{Adv}_{\mathcal{ID}}^{terror}(\mathcal{A}, \mathcal{S}, \mathcal{T}') = p_{\mathcal{A}} - p_{\mathcal{S}}$$

where $p_{\mathcal{A}}$ is the probability that $\mathcal{R}$ accepts in one of the $q_{\mathcal{R}}$ reader-adversary sessions sid *such that at most $T_{\max}$ time-critical phases of* sid *are tainted, and $p_{\mathcal{S}}$ is the probability that, given $\mathsf{view}_{\mathcal{A}}$, $\mathcal{S}$ authenticates to $\mathcal{R}$ in one of $q_{\mathcal{R}}$ subsequent executions.*

Insight: SimTF. In [1], the active adversary succeeds if: it authenticates with the prover's aid; and it authenticates (at all) without it. In fact, the prover's secret must be information-theoretically hidden. This model excludes nearly *any* information-exchange with the adversary, even if the data does not directly help authentication. As most attacks are ruled out, this definition is rather weak.

By contrast, SimTF security focuses on exactly how much the prover's information helps the simulator. Excluded are only attacks where prover data, contained in $\mathcal{A}$'s state, is directly used by $\mathcal{S}$. Thus, even if the simulator's authentication probability is significant, but not as large as the adversary's, the attack is valid. This definition is very broad, enabling syntactic attacks like the one in [7] against the scheme of Reid et al.

3 Flavors of Terrorist Fraud

In this section, we introduce two possible definitions of terrorist-fraud resistance. The first (called GameTF security, see Section 3.1) is a game-based definition capturing the intuition behind a basic terrorist-fraud attack in a manner compatible with the model of Dürholz et al. [6]. This notion is sufficient for many practical applications, e.g. logistics or ticketing in public transport. Our second notion (strSimTF security, see Section 3.2) extends, in a natural way, the simulation-based SimTF definition in [6]; this definition is extremely strong, and should be used only in high-risk applications like e-Passports or e-voting. In what follows we briefly explain our motivations for introducing the two notions, referring to previous models of terrorist-fraud resistance, and sketching our own approach towards defining terrorist-fraud attacks.

We discuss mainly two modeling aspects in defining terrorist fraud: (1) the adversary-prover interaction; and (2) the restriction on how much a prover can help. Both terrorist-fraud models in the literature [1,6] seem to agree on how to handle (1), but fundamentally disagree on how to define (2). We first discuss point (2). In this matter, Avoine et al. [1] demand that the prover's aid gives the adversary "no further advantage" to authenticate, requiring statistically-hiding properties for the prover's secret key. As discussed in Section 1 this restricts the prover very much, and thus attacks where partial key-related information is given are ruled out. By contrast, SimTF security [6] only rules out attacks where the information received from the prover can be used as effectively during the prover-aided session *and* later. We agree with the intuition of Avoine et al. that the adversary should have no "further advantage", but note that the behavior of the adversary after the prover has stopped helping it is that of a mafia fraud attacker who also retains some state information, i.e. what the prover has forwarded it before. Thus, our GameTF notion considers a pair of adversaries: a first, terrorist adversary (aided by the prover); and a second, mafia-fraud adversary sharing state with the first adversary.

We also re-consider the traditional adversary-prover interaction restriction to lazy phases (point (1) above), which seems to assume that time-critical interactions would be detected by the verifier's clock. We disagree: the clock can *only* detect queries to the prover if the messages have a *relay scheduling*, i.e. a MITM adversary receives input from the reader, then sends input to the tag; upon receiving output from the tag, it sends some output (the same, or different) to the reader. This is not the same as *pure relay* as defined for mafia fraud, see Definition 1, since in pure relay, the input and output messages must be the same. Thus, we *may* allow such adversary-prover interactions. We discuss also why we *should* allow them; and why we cannot allow the adversary *even more* freedom, e.g. by using Definition 1.

Why we should allow it. Consider a distance-bounding protocol where the dishonest $\mathcal{T}'$ and $\mathcal{R}$ share, at the end of the lazy phases, pseudo-random strings T^0, T^1, such that $T^0 \oplus T^1 = sk^*$, where sk^* is a secret key (a part of sk or an independent key). This is how terrorist fraud resistance is usually achieved.

Now assume that $\mathcal{R}$ generates challenges as follows: it first draws a random c_1 for the first round, then runs a PRF (with key sk) on input c_1 to generate a string s with $|s| \geq N_c$. Then $\mathcal{R}$ sets challenges $c_2, \dots c_{N_c}$ for the other rounds bitwise to the bits of s. That is, c_2 is set to the most significant bit of s, c_3, to the following bit, etc. Such a protocol does not exist in the literature; nevertheless, our model *should* rule out such dependency of challenges.

In each time-critical phase of the protocol, $\mathcal{R}$ sends a challenge bit c_i and expects a bit from T^{c_i} (i.e. either T^0 or T^1). At the end of the lazy phases, $\mathcal{T}'$ has computed responses T^0 and T^1. When the terrorist adversary $\mathcal{A}$ receives challenge c_1 from $\mathcal{R}$, it sends a random bit r to $\mathcal{R}$, then forwards c_1 to $\mathcal{T}'$. Now $\mathcal{T}'$ computes $c_2, \dots c_{N_c}$ and sends the appropriate responses to $\mathcal{A}$ (without revealing any information about sk^*). The adversary wins with probability $\frac{1}{2}$ (the probability that $r = T^{c_1}$).

Why this is all we can do. Mafia fraud adversaries may use relay scheduling if at least one relayed message is not the exact one $\mathcal{A}$ received from the honest party. We cannot allow this for terrorist fraud, since the dishonest prover may adapt its response in order to bypass our definition. For instance, instead of sending the correct response r for each round, it just sends $1 \oplus r$, that is, the flipped bit. Then $\mathcal{A}$ just flips the bit back and sends it to the verifier.

Consequently, we redefine tainted phases as follows:

Definition 5 (Tainted Time-Critical Phase ($\mathsf{strSimTF}$)). *A time-critical phase* $\Pi_{\mathsf{sid}}[k \dots k+2\ell-1] = (m_k, \dots, m_{k+2\ell-1})$ *for* $k, \ell \geq 1$ *of a reader-adversary session* sid*, with the k-th message being received by the adversary, is* tainted *if there exists an adversary-tag session* sid^* *and messages* $(m^*_k, \dots, m^*_{k+2\ell-1})$ *such that for all* $i = 0, 1, \dots, \ell-1$ *we have:*

$$\begin{aligned} & \mathit{clock}(\mathsf{sid}, k+2i) < \mathit{clock}(\mathsf{sid}^*, k+2i), \\ \mathit{and} \qquad & \mathit{clock}(\mathsf{sid}, k+2i+1) > \mathit{clock}(\mathsf{sid}^*, k+2i+1). \end{aligned}$$

3.1 GameTF Security

Our game-based terrorist fraud resistance GameTF follows the intuition of [1]. The key difference between this and SimTF security is that GameTF security rules out attacks if the attacker gains *any* advantage to authenticate later (even if this advantage is *smaller* than the adversary's success probability). Thus, we match the unaided adversary's success against a MITM attack (mafia fraud).

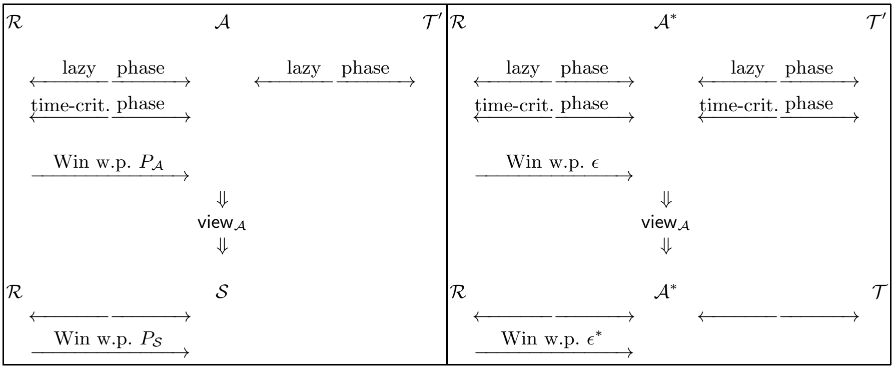
We consider a simulator-free two-step game, with two adversaries $\mathcal{A}$ and $\mathcal{A}^*$ sharing view $\mathsf{view}_{\mathcal{A}}$, as defined in the SimTF security model.[4] Now $\mathcal{A}$ can interact with the dishonest $\mathcal{T}'$ during lazy and time-critical phases as described above (we use the notion of tainted phases in Definition 5). The second adversary $\mathcal{A}^*$ (sharing state, or view, with $\mathcal{A}$) runs a mafia fraud interaction with $\mathcal{R}$ in the presence of the prover (who is this time honest). Thus, $\mathcal{A}^*$ models the adversary *after* the prover stops helping: $\mathcal{A}^*$ must authenticate in a MITM attack, using $\mathsf{view}_{\mathcal{A}}$. In SimTF security, the simulator is passive and just uses $\mathsf{view}_{\mathcal{A}}$ to authenticate; however, in GameTF, $\mathcal{A}^*$ runs an active mafia-fraud interaction *and* uses $\mathsf{view}_{\mathcal{A}}$. We say that $\mathcal{A}$ *is helpful* to $\mathcal{A}^*$ if $\mathcal{A}^*$ authenticates with better than mafia-fraud success probability (i.e. $\mathsf{view}_{\mathcal{A}}$ shouldn't help $\mathcal{A}^*$ at all).

We sketch the differences between SimTF and GameTF security in Fig. 1. Also note that in SimTF security, $\mathcal{A}$ queries $\mathcal{T}'$ in at most $T_{\max}$ time-critical phases (tainting them). However, the GameTF adversary $\mathcal{A}$ may query $\mathcal{T}'$ in *each* time-critical phase, tainting it only if it uses relay scheduling.

Of $\mathcal{A}$ and $\mathcal{A}^*$, the former is the terrorist adversary. Its attack is *invalid* if there exists $\mathcal{A}^*$ such that $\mathcal{A}$ is helpful to $\mathcal{A}^*$, i.e. we rule out attacks where $\mathcal{A}$ learns information useful for later authentication. Schemes are GameTF secure if every terrorist adversary $\mathcal{A}$ either (i) wins with negligible probability; or (ii) there exists an adversary $\mathcal{A}^*$ to which $\mathcal{A}$ is helpful. Let $\mathcal{A}$ run in time t, using q_{OBS} reader-tag, $q_{\mathcal{R}}$ resp. reader-adversary, and $q_{\mathcal{T}'}$ adversary-tag sessions —the latter subject to Definition 5; its success probability is denoted ϵ.

When $\mathcal{A}$ stops, it forwards $\mathsf{view}_{\mathcal{A}}$ to $\mathcal{A}^*$. Then $\mathcal{A}^*$ runs a mafia-fraud interaction with $\mathcal{T}$ (we omit the apostrophe as $\mathcal{T}$ is now honest). W.l.o.g., let $\mathcal{A}^*$ run in time $t^* \leq 3t$ ($\mathcal{A}^*$ runs $\mathcal{A}$ at most twice internally, with the same queries as $\mathcal{A}$), and let $\mathcal{A}^*$ run at most q_{OBS} reader-tag, $q_{\mathcal{R}}$ reader-adversary, and $q_{\mathcal{R}}$ adversary-tag sessions (since $\mathcal{A}$'s queries to $\mathcal{T}'$ deviate from protocol, we give $\mathcal{A}^*$ one adversary-tag session for each reader-adversary session). Let $\mathcal{A}^*$ win w.p. ϵ^*. We now define *helpful* terrorist adversaries and GameTF security.

[4] Note that any other state information is computable from $\mathsf{view}_{\mathcal{A}}$, for higher runtimes.



SimTF Security:
$P_\mathcal{S} \geq P_\mathcal{A}$

GameTF Security:
ϵ non-negl. or $\epsilon^* > \mathbf{Adv}_{\mathcal{ID}}^{\text{mafia}}$

Fig. 1. Simulation and game-based security models

Definition 6. *For an authentication scheme $\mathcal{ID}$ with parameters $(t_{\max}, T_{\max}, E_{\max}, N_c)$, let $\mathcal{A}$ be a $(t, q_{\text{OBS}}, q_\mathcal{R}, q_{\mathcal{T}'})$ adversary running a* strSimTF *interaction with $\mathcal{R}$ and $\mathcal{T}'$, and let* st = *view*$_\mathcal{A}$ *denote its state. We say that $\mathcal{A}$ is* helpful *to an adversary $\mathcal{A}^*$ with input* st*, runtime at most $3t$, running at most $q_{\text{OBS}}, q_\mathcal{R}$, and $q_\mathcal{T} = q_\mathcal{R}$ sessions in a mafia-fraud interaction with $\mathcal{R}$ and $\mathcal{T}$, and winning with probability ϵ^* (taken over* view$_\mathcal{A}$ *and the coins of $\mathcal{A}^*$) if:*

$$\epsilon^* > \boldsymbol{Adv}_{\mathcal{ID}}^{mafia},$$

where $\boldsymbol{Adv}_{\mathcal{ID}}^{mafia}$ denotes the mafia fraud resistance of $\mathcal{ID}$ for a $(t, q_{\text{OBS}}, q_\mathcal{R}, q_\mathcal{T})$-mafia adversary.

Definition 7 (GameTF Security). *A distance-bounding authentication scheme $\mathcal{ID}$ with parameters $(t_{\max}, T_{\max}, E_{\max}, N_c)$ is $(t, q_{\text{OBS}}, q_\mathcal{R}, q_{\mathcal{T}'}, \epsilon)$-*GameTF secure *if for all $(t, q_{\text{OBS}}, q_\mathcal{R}, q_{\mathcal{T}'})$ adversaries $\mathcal{A}$ running a* strSimTF *interaction, one of the following statements hold:*

- *The probability that $\mathcal{A}$ wins is upper bounded by ϵ;*
- *There exists an adversary $\mathcal{A}^*$ such that $\mathcal{A}$ is helpful to $\mathcal{A}^*$ as defined above.*

A scheme $\mathcal{ID}$ is GameTF *secure* if it is $(t, q_{\text{OBS}}, q_\mathcal{R}, q_{\mathcal{T}'}, \epsilon)$-GameTF secure for negligible ϵ.

The Swiss-Knife protocol. This section concerns the Swiss-Knife protocol of [14], modified as in [7], which we depict in Fig. 2. We use the modified version since it is mafia-fraud resistant, noting that both the original and the modified versions are SimTF-insecure [7]. Despite the attack of [6], however, the scheme prevents known terrorist attacks. In fact, we can prove its GameTF-security, confirming intuition; in particular, the syntactic attack in [7] is ruled out because the prover's help gives the adversary a significant advantage. For our GameTF proof, we use the scheme's mafia fraud resistance. In the protocol, PRF denotes a pseudorandom function, $\mathcal{ID}_\mathcal{R}$ and $\mathcal{ID}_\mathcal{T}$ are reader and tag identifiers, and const is a publicly known constant. The difference to the original scheme is the use of an independent key sk^* instead of re-using sk.

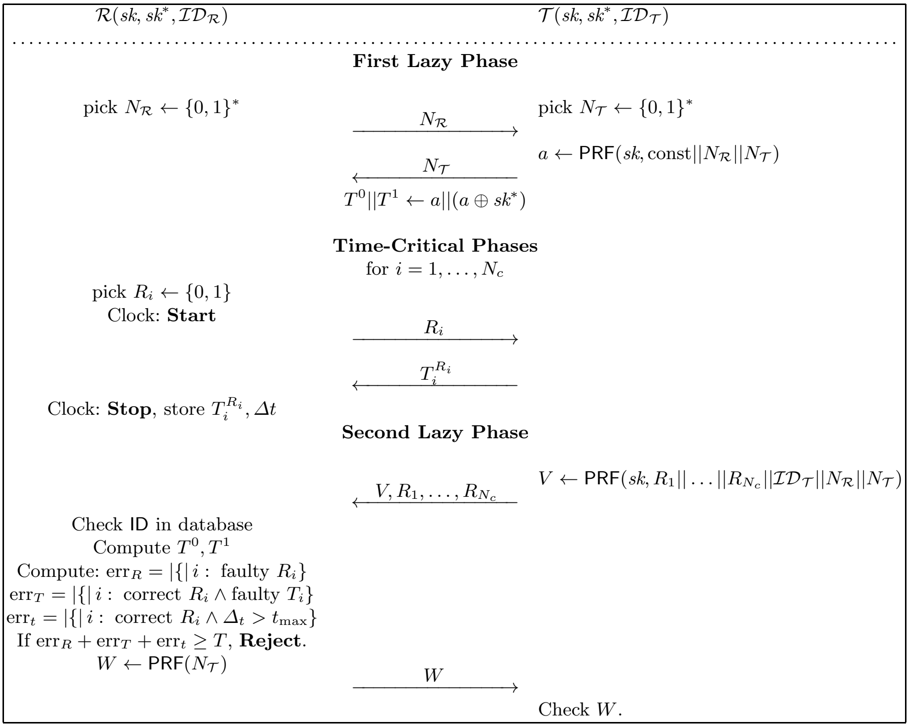

$\mathcal{R}(sk, sk^*, \mathcal{ID}_\mathcal{R})$
$\mathcal{T}(sk, sk^*, \mathcal{ID}_\mathcal{T})$
First Lazy Phase
pick $N_\mathcal{R} \leftarrow \{0,1\}^*$
$N_\mathcal{R}$
pick $N_\mathcal{T} \leftarrow \{0,1\}^*$
$a \leftarrow \mathsf{PRF}(sk, \text{const}||N_\mathcal{R}||N_\mathcal{T})$
$N_\mathcal{T}$
$T^0||T^1 \leftarrow a||(a \oplus sk^*)$
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
$V \leftarrow \mathsf{PRF}(sk, R_1||\ldots||R_{N_c}||\mathcal{ID}_\mathcal{T}||N_\mathcal{R}||N_\mathcal{T})$
$V, R_1, \ldots, R_{N_c}$
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.

Fig. 2. The Modified Swiss-Knife protocol of [7]

Proposition 1 (GameTF **Security).** *Let $\mathcal{ID}$ be the protocol in Fig. 2 with parameters $(t_{\max}, N_c)$. This scheme is $(t, q_{\text{OBS}}, q_\mathcal{R}, q_{\mathcal{T}'}, \epsilon)$-*GameTF *secure, for $\epsilon \geq$ $\boldsymbol{Adv}_{\mathcal{ID}}^{mafia}$.*

Proof. Assume towards contradiction that the scheme is *not* $(t, q_{\text{OBS}}, q_\mathcal{R}, q_{\mathcal{T}'}, \epsilon)$-GameTF resistant. Then there exists a $(t, q_{\text{OBS}}, q_\mathcal{R}, q_{\mathcal{T}'})$ adversary $\mathcal{A}$ such that: (i) $\mathcal{A}$ wins with probability $\epsilon > \mathbf{Adv}_{\mathcal{ID}}^{\text{mafia}}$; *and* (ii) for all $(3t, q_{\text{OBS}}, q_\mathcal{R}, q_\mathcal{R})$-adversaries $\mathcal{A}^*$, initialized with $\mathsf{view}_\mathcal{A}$, running a mafia fraud interaction with $\mathcal{R}$ and $\mathcal{T}$, the success probability ϵ^* of $\mathcal{A}^*$ is such that $\epsilon^* \geq \mathbf{Adv}_{\mathcal{ID}}^{\text{mafia}}$.

We construct, for each $\mathcal{A}$ as in (i) and (ii), an $\mathcal{A}^*$ with input $\mathsf{view}_\mathcal{A}$, winning in the attack above with probability $\epsilon^* \geq \epsilon$. Thus, if $\mathcal{A}$ wins w.p. $\epsilon > \mathbf{Adv}_{\mathcal{ID}}^{\text{mafia}}$ (as in (i)), our $\mathcal{A}^*$ follows the specifications of Definition 6 and wins w.p. $\epsilon^* = \epsilon > \mathbf{Adv}_{\mathcal{ID}}^{\text{mafia}}$ (contradicting point (ii)). Thus, an adversary $\mathcal{A}$ for which points (i) and (ii) both hold does not exist.

We describe $\mathcal{A}^*$. For each session $\mathcal{A}$ runs with $\mathcal{R}$, $\mathcal{A}^*$ runs parallel sessions with $\mathcal{R}$ and resp. $\mathcal{T}$, relaying the lazy phase and running time-critical phases as follows. In the verifier-adversary session sid, $\mathcal{A}^*$ runs $\mathcal{A}$ internally, branching out in two executions, so that: if $\mathcal{A}$ taints a phase, so does $\mathcal{A}^*$ (both succeed w.p. 1 and have 1 less phase to taint); if $\mathcal{A}$ refuses to respond to challenge $\alpha_i =: \alpha$, then $\mathcal{A}^*$ uses a Go-Early strategy (see Proposition 3), querying $\mathcal{T}$ with challenge $\bar{\alpha} = \alpha \oplus 1$ (both $\mathcal{A}$ and $\mathcal{A}^*$ know the same response), and $\mathcal{A}^*$ guesses the response if queried with challenge α in session sid: this gives $\mathcal{A}$ and $\mathcal{A}^*$ equal probability to win; finally, if $\mathcal{A}$ forwards responses r_0 (for a 0 challenge) and r_1 (for a 1 challenge) for this round, $\mathcal{A}^*$ uses the Go-Early strategy, challenging $\mathcal{T}$ with $\alpha \in \{0, 1\}$, and receiving R_i^α. Then $\mathcal{A}^*$ sets $R_i^{\bar{\alpha}} = R_i^\alpha \oplus r_0 \oplus r_1$; given challenge $c \in \{0, 1\}$ in sid, $\mathcal{A}^*$ responds with R_i^c. There are four cases:

- Both values r_0 and r_1 are correct. Then both $\mathcal{A}$ and $\mathcal{A}^*$ win w.p. 1.
- Both r_0 and r_1 are incorrect. Now $\mathcal{A}$ loses the phase and $\mathcal{A}^*$ wins w.p. 1.
- Either r_0 or r_1 is incorrect. Now $\mathcal{A}$ wins the round w.p. $\frac{1}{2}$. As $\mathcal{A}^*$ runs the Go-Early strategy for challenge $\alpha \in \{0, 1\}$, it knows the correct R_i^α, but the wrong $R_i^{\bar{\alpha}}$ (as $r_0 \oplus r_1$ is incorrect), and wins the phase w.p. $\frac{1}{2}$. If they answer wrongly, both adversaries subtract 1 from $E_{\max}$.

Thus, $\mathcal{A}^*$ wins with at least as high probability as $\mathcal{A}$ in each time-critical phase. Thus, $\mathcal{A}$'s success probability ϵ equals that of $\mathcal{A}^*$, i.e. ϵ^*. Furthermore, the parameters of $\mathcal{A}^*$ are as required. Now if there exists an adversary $\mathcal{A}$ with $\epsilon > \mathbf{Adv}_{\mathcal{ID}}^{\text{mafia}}$, then $\mathcal{A}^*$ succeeds with probability $\epsilon^* > \mathbf{Adv}_{\mathcal{ID}}^{\text{mafia}}$. Thus, $\mathcal{A}$ is helpful to $\mathcal{A}^*$, contradicting our assumption. Since the scheme is mafia fraud resistant, it is also GameTF secure. □

3.2 strSimTF Security

Terrorist fraud is a very strong attack. If the incentive is high (e.g. breaking e-Passport security), then dishonest provers may be willing to forward *some* secret information to ensure the adversary's success. However, SimTF security (while strong) restricts the adversary unnecessarily by not allowing it to query the prover in time-critical rounds.

We obtain our strSimTF notion by simply switching the tainted-phase definition from Definition 3 to Definition 5, and then use Definition 4. The strSimTF adversary is stronger: we show in Theorem 1 that there exist SimTF-secure schemes that are strSimTF-insecure. We also show in Section 4 that strSimTF security is achievable; this is a non-trivial statement, since the recent results of [7] cast a doubt whether *any* existing protocol is provably SimTF-secure (they are thus also strSimTF-insecure). Our construction relies on the Swiss-Knife protocol, but we introduce a back door for the simulator to authenticate.

3.3 Relating the Notions

Our full security diagram in Fig. 3 fully relates the notions. Due to space reasons, we only sketch the proofs in the Appendix.

Theorem 1 (Relations between notions). SimTF*,* strSimTF*,* GameTF *security, and mafia-fraud resistance are related as in Fig. 3. Arrows between notions indicate that security against one notion implies security against the other.*

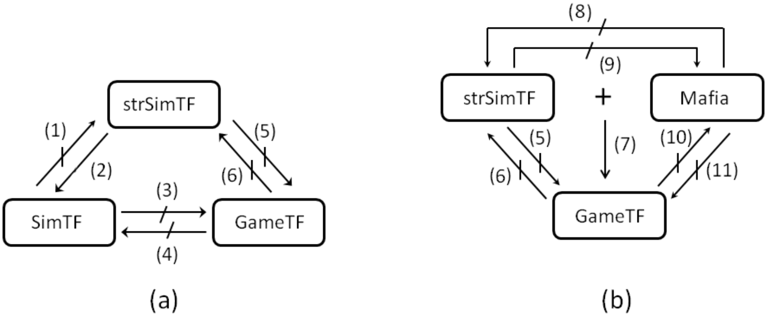


Fig. 3. Full security diagram. The "+" sign beside (7) indicates property composition.

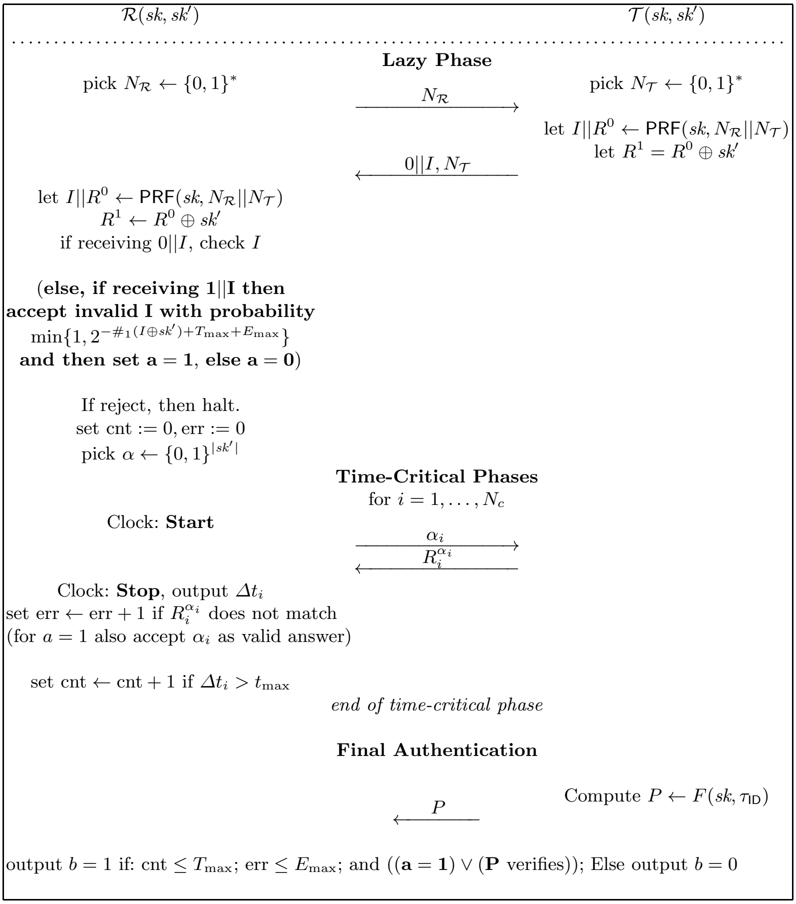
4 Terrorist Fraud Resistant Construction

4.1 The Protocol

Our SimTF- and strSimTF-secure protocol relies on the (modified) Swiss-Knife protocol in Fig. 2, which thwarts MITM attacks by a second authentication phase. We make the following changes: (1) we add a bit to the authentication string, now denoted $0||I$ in Fig. 4 (an honest prover always sends $0||I$, but by sending $1||I$, a dishonest prover or an adversary may switch the flag a for $\mathcal{R}$, see the following point); (2) we add a flag a for $\mathcal{R}$ denoting whether the protocol runs normally (more or less as in the Swiss-Knife protocol) or exceptionally, such that during the time-critical phases, the verifier just expects $\mathcal{T}$ to echo the challenges (also see below). In our proof the simulator will try to make $\mathcal{R}$ run the protocol exceptionally, thus bypassing authentication.

Now, if the prover's first protocol response is a string of the form $1||I, N_{\mathcal{T}}$, $\mathcal{R}$ accepts $1||I$ as valid lazy authentication (continuing the protocol) with probability $\min\{1, 2^{-\#_1(I \oplus sk') + T_{\max} + E_{\max}}\}$; in this case the flag a is set to 1. We denote by $\#_1(I \oplus sk')$ the Hamming distance between I and sk'; thus, if the first bit is a 1, the rest of the string I should be close to sk' (an adversary can't just receive an honest $0||I$ and flip the first bit). The probability is tailored to fit the SimTF definition, where the simulator recovers some bits of sk' from a successful adversary; the bound also accounts for $\mathcal{A}$'s tainted and erroneous-response rounds (see the SimTF proof). The flag a and our second authentication method (using $1||I$ responses, with I close to sk') are artifices enabling us to prove SimTF and resp. strSimTF security. Once the flag is flipped, *any* party in $\mathcal{R}$'s proximity can authenticate, since the reader expects $\mathcal{T}$ to just echo the time-critical challenges. However, mafia fraud attackers cannot make use of this, as honest provers never send $1||sk'$ (but rather a string $0||I$, where I is output by PRF) —for mafia and impersonation security we only lose a term $q_{\mathcal{R}} \cdot 2^{-(2-\log_2 3)N_c+T_{\max}+E_{\max}}$, accounting for the probability of guessing a close-enough authentication string.

For the second lazy authentication phase, $\mathcal{T}$ runs a different PRF than before (namely, F) on the session transcript, denoted τ_{ID}. In Simulator mode (i.e. if $a = 1$), the string P is not checked. See the full protocol in Fig. 4.



$\mathcal{R}(sk, sk')$
$\mathcal{T}(sk, sk')$
Lazy Phase
pick $N_{\mathcal{R}} \leftarrow \{0,1\}^*$
pick $N_{\mathcal{T}} \leftarrow \{0,1\}^*$
$\xrightarrow{N_{\mathcal{R}}}$
let $I||R^0 \leftarrow \mathsf{PRF}(sk, N_{\mathcal{R}}||N_{\mathcal{T}})$
let $R^1 = R^0 \oplus sk'$
$\xleftarrow{0||I, N_{\mathcal{T}}}$
let $I||R^0 \leftarrow \mathsf{PRF}(sk, N_{\mathcal{R}}||N_{\mathcal{T}})$
$R^1 \leftarrow R^0 \oplus sk'$
if receiving $0||I$, check I
(**else, if receiving 1||I then accept invalid I with probability** $\min\{1, 2^{-\#_1(I \oplus sk')+T_{\max}+E_{\max}}\}$ **and then set a = 1, else a = 0**)
If reject, then halt.
set $\mathrm{cnt} := 0, \mathrm{err} := 0$
pick $\alpha \leftarrow \{0,1\}^{|sk'|}$
Time-Critical Phases
for $i = 1, \ldots, N_c$
Clock: **Start**
$\xrightarrow{\alpha_i}$
$\xleftarrow{R_i^{\alpha_i}}$
Clock: **Stop**, output Δt_i
set $\mathrm{err} \leftarrow \mathrm{err} + 1$ if $R_i^{\alpha_i}$ does not match
(for $a = 1$ also accept α_i as valid answer)
set $\mathrm{cnt} \leftarrow \mathrm{cnt} + 1$ if $\Delta t_i > t_{\max}$
end of time-critical phase
Final Authentication
Compute $P \leftarrow F(sk, \tau_{\mathsf{ID}})$
$\xleftarrow{P}$
output $b = 1$ if: $\mathrm{cnt} \leq T_{\max}$; $\mathrm{err} \leq E_{\max}$; and $((\mathbf{a = 1}) \vee (\mathbf{P}$ verifies$))$; Else output $b = 0$

Fig. 4. SimTF secure distance-bounding protocol

4.2 Security

Here we prove our scheme SimTF- and strSimTF-secure, under the assumption that reader-adversary sessions are executed sequentially. Note that not *every* SimTF-secure scheme is also strSimTF-secure, see also Section 3.3. We also state the scheme's full distance-bounding properties, but omit the proofs for space reasons.

Theorem 2 (SimTF Security). *Let* ID *be the distance-bounding authentication scheme in Fig. 4 with parameters* $(t_{\max}, T_{\max}, E_{\max}, N_c)$. *For any* $(t, q_{\mathcal{R}}, q_{\mathcal{T}'})$-SimTF *adversary* $\mathcal{A}$ *against the scheme, mounting a sequential attack, there exists a* $t_{\mathcal{S}}$*-simulator* $\mathcal{S}$ *with* $t_{\mathcal{S}} = 2t + O(nq_{\mathcal{R}})$ *such that we have*

$$\boldsymbol{Adv}_{\mathsf{ID}}^{\mathsf{SimTF}}(\mathcal{A}, \mathcal{S}, \mathcal{T}) \leq 0.$$

Proof. We describe the simulator $\mathcal{S}$. Given $\mathsf{view}_{\mathcal{A}}$, including $\mathcal{A}$'s randomness, $\mathcal{S}$ internally runs $\mathcal{A}$ stepwise with $\mathsf{view}_{\mathcal{A}}$, repeating the same strategy for each of its $q_{\mathcal{R}}$ sessions sid(as many sessions as $\mathcal{A}$). Namely, $\mathcal{S}$ checks if $\mathcal{A}$ sends $1||I$ and succeeds; if so, $\mathcal{S}$ sets $sk'' = I$ for sid. Else, if $\mathcal{A}$ uses $0||I$, the simulator constructs sk'' as follows: each time $\mathcal{A}$ expects α_i in the next time-critical phase, $\mathcal{S}$ branches into two executions, once sending $\alpha_i^0 = 0$ and the other time $\alpha_i^1 = 1$ to $\mathcal{A}$. It waits for $\mathcal{A}$ to answer in both branches, or query $\mathcal{T}'$ (tainting a branch). As we consider sequential executions only, there are no other options. If $\mathcal{A}$ taints or refuses one query, $\mathcal{S}$ picks sk_i'' at random; else it sets $sk_i'' = R_i^0 \oplus R_i^1$. The simulator returns to its main execution and resumes the simulation with the correct α_i. When $\mathcal{A}$ stops, $\mathcal{S}$ has predictions sk_i'' for each bit of sk_i'. If $\mathcal{A}$ succeeds in some sid with $0||C$, then there are four cases for each guessed bit sk_i'':

- The adversary taints the phase or refuses to answer both challenges. Then $\mathcal{S}$'s guessing strategy is good: by comparing the term $\#_1(I \oplus sk') - T_{\max} - E_{\max}$ (i.e. the number of bits $\mathcal{S}$ needs to predict) to the number of phases $\mathcal{A}$ needs to pass, we see that $\mathcal{S}$ gets a "wild card" for each of the at most $T_{\max}$ tainted phases. If $\mathcal{A}$ taints the phase in both branches, it succeeds for one round; however $\mathcal{S}$ then "gains" 1.5 bits by deducting one wild card off $T_{\max}$ and guessing a bit of sk' with probability $\frac{1}{2}$. Thus $\mathcal{S}$ has an advantage over $\mathcal{A}$. If, however, $\mathcal{A}$ taints exactly one branch *and* always responds correctly in the other (it always wins the round), then $\mathcal{S}$ gets half a bit from sk_i' correctly (for the untainted branch, which occurs w.p. $\frac{1}{2}$), and another half a bit from the tainted branch ($\mathcal{A}$ cannot taint another round later). On average $\mathcal{S}$ gets thus as many bits as is $\mathcal{A}$'s success probability.
- If $\mathcal{A}$ returns correct R_i^0, R_i^1, then $sk_i'' = sk_i'$, $\mathcal{A}$ wins the round, and $\mathcal{S}$ gains a bit.
- Analogously, $sk_i'' = sk_i'$ if both replies are incorrect ($\mathcal{A}$ fails here).
- If exactly one of R_i^0 and R_i^1 is correct, then sk_i'' is certainly incorrect. But then $\mathcal{A}$ too fails the phase with probability $\frac{1}{2}$. The reasoning from the first case for $T_{\max}$ applies to $E_{\max}$.

Accounting for at most $T_{\max} + E_{\max}$ tainted and erroneous phases, $\mathcal{A}$ authenticates with probability at most $2^{-\#_1(sk'' \oplus sk') + T_{\max} + E_{\max}}$. By using sk'', $\mathcal{S}$ also authenticates with the same probability. Also, if $\mathcal{S}$ reuses $sk'' = I$ for adversary executions with $1||I$, it succeeds with the same probability as $\mathcal{A}$. □

Proposition 2. *Let* $\mathcal{ID}$ *be the protocol in Fig. 4 with parameters* $(t_{\max}, T_{\max}, E_{\max}, N_c)$. *For any* $(t, q_{\mathcal{R}}, q_{\mathcal{T}'})$-strSimTF *adversary* $\mathcal{A}$ *against* $\mathcal{ID}$*, mounting a* *sequential attack, there exists a $t_{\mathcal{S}}$-simulator $\mathcal{S}$ with $t_{\mathcal{S}} = 2t + O(nq_{\mathcal{R}})$ such that for any $\mathcal{T}'$ running in time $t_{\mathcal{T}'}$*

$$\boldsymbol{Adv}^{terror}_{\mathcal{ID}}(\mathcal{A}, \mathcal{S}, \mathcal{T}) \leq 0.$$

Proof. We extend our SimTF proof to account for time-critical queries to $\mathcal{T}'$, also for sequential executions. We change $\mathcal{S}$ as follows: if $\mathcal{A}$ does *not* interact with $\mathcal{T}'$ during time-critical phases, the simulator is the same. If $\mathcal{A}$ *does* query $\mathcal{T}'$, for each time-critical phase where $\mathcal{A}$ interacts with $\mathcal{T}'$, the simulator branches the execution for both challenges. If $\mathcal{A}$ refuses to forward one response or taints the phase (with relay scheduling), $\mathcal{S}$ guesses the bit in sk'' as before.

The old proof still stands; indeed, if the phase is *not* tainted by relaying, then either $\mathcal{A}$ queries $\mathcal{T}'$ *before* challenge α_i is sent, or $\mathcal{T}'$ responds *after* $\mathcal{A}$ has replied to $\mathcal{R}$ in this phase. In the former case, $\mathcal{T}'$ does *not* know the true challenge, as in the SimTF scenario. In the latter case, the prover's response does not help $\mathcal{A}$, as the responses are pseudorandom and independent of each other, though it may help the simulator instead (since $\mathsf{view}_{\mathcal{A}}$ contains the correct response). □

Proposition 3 (Mafia Fraud Resistance). *Let ID be the scheme in Figure 4 with parameters $(t_{\max}, T_{\max}, E_{\max}, N_c)$. For any $(t, q_{\mathcal{R}}, q_{\mathcal{T}}, q_{\mathrm{OBS}})$-mafia-fraud adversary $\mathcal{A}$ against the scheme there exist: a (t', q')-distinguisher $\mathcal{A}'$ against PRF, a (t'', q'')-distinguisher $\mathcal{A}''$ against F, and a (t''', q''')-distinguisher $\mathcal{A}'''$ (where $t', t'', t''' = t + O(n)$ and $q', q'', q''' = q_{\mathcal{R}} + q_{\mathcal{T}} + q_{\mathrm{OBS}}$) such that:*

$$\begin{aligned}\boldsymbol{Adv}^{mafia}_{\mathsf{ID}}(\mathcal{A}) \leq q_{\mathcal{R}} \left(\tfrac{1}{2}\right)^{N_c-(T_{\max}+E_{\max})} &+ \binom{q_{\mathcal{R}} + q_{\mathrm{OBS}}}{2} \cdot 2^{-(|N_{\mathcal{R}}|+\lceil \frac{N_c}{2} \rceil - T_{\max} - E_{\max})} \\ &+ \binom{q_{\mathcal{T}} + q_{\mathrm{OBS}}}{2} \cdot 2^{-(|N_{\mathcal{T}}|+\lceil \frac{N_c}{2} \rceil - T_{\max} - E_{\max})} + \boldsymbol{Adv}^{d}_{\mathsf{PRF}}(\mathcal{A}') \\ &+ \boldsymbol{Adv}^{d}_{F}(\mathcal{A}'') + 2\boldsymbol{Adv}^{d(\mathcal{D},U)}_{\mathsf{Kg}}(\mathcal{A}''') + q_{\mathcal{R}} \cdot 2^{-(2-\log_2 3)N_c + T_{\max} + E_{\max}}.\end{aligned}$$

Proposition 4 (Distance Fraud Resistance). *Let ID be the scheme in Figure 4 with parameters $(t_{\max}, T_{\max}, E_{\max}, N_c)$. Assume also that Kg is run by either the reader or a trusted third party (*not *the tag), such that it generates keys sk, sk' by drawing them uniformly at random from a distribution $\mathcal{D}$ computationally indistinguishable from the uniform random distribution. For any $(t, q_{\mathcal{R}}, q_{\mathcal{T}}, q_{\mathrm{OBS}})$-distance-fraud adversary $\mathcal{A}$ against ID it holds that,*

$$\boldsymbol{Adv}^{dist}_{\mathsf{ID}}(\mathcal{A}) \leq q_{\mathcal{R}} \cdot \left(\frac{3}{4}\right)^{N_c - T_{\max} - E_{\max}} + \boldsymbol{Adv}^{d(\mathcal{D},U)}_{\mathsf{Kg}}(\mathcal{A}').$$

Proposition 5 (Impersonation Security). *Let ID be the scheme in Figure 4 with parameters $(t_{\max}, T_{\max}, E_{\max}, N_c)$. For any $(t, q_{\mathcal{R}}, q_{\mathcal{T}}, q_{\mathrm{OBS}})$-impersonation adversary $\mathcal{A}$ against ID there exist a (t', q')-distinguisher $\mathcal{A}'$, resp. a (t'', q'')-distinguisher $\mathcal{A}''$ against PRF and resp. F (with $t', t'' = t + O(n)$ and $q', q'' = q_{\mathcal{R}} + q_{\mathcal{T}} + q_{\mathrm{OBS}}$) such that*

$$\begin{aligned}\boldsymbol{Adv}^{imp}_{\mathcal{ID}(\mathcal{A})} \leq q_{\mathcal{R}} \cdot 2^{-|I|} + q_{\mathcal{R}} \cdot 2^{-(2-\log_2 3)N_c + T_{\max} + E_{\max}} + q_{\mathcal{R}} \cdot \boldsymbol{Adv}^{d}_{\mathsf{PRF}}(\mathcal{A}') &+ \\ q_{\mathcal{R}} \cdot \boldsymbol{Adv}^{d}_{F}(\mathcal{A}') + \left(\binom{q_{\mathcal{R}} + q_{\mathrm{OBS}}}{2} \cdot 2^{-|N_{\mathcal{R}}|} + \binom{q_{\mathcal{T}}}{2} \cdot 2^{-|N_{\mathcal{T}}|}\right) \cdot 2^{-N_c}&.\end{aligned}$$

5 Which Model to Use

The abundance of terrorist-fraud resistance definitions in the literature proves that, though this topic is crucial to distance-bounding authentication, no clear solution has been found for it. Even our present work does not give one, but rather two definitions of terrorist-fraud resistance, and proves that, though many existent schemes in the literature fail to achieve one notion (strSimTF security), they do attain the other. Which definition is better? That is a question which cannot be answered in an unequivocal way.

Simulation-based models, like SimTF and strSimTF security, formalize terrorist-fraud resistance in a very strong way, allowing the prover to help the adversary as long as the gained help cannot be used by a simulator given the adversary's view only. This is the case for the SimTF notion of [6], which we extend to better capture the attack. These strong notions *should* be used in high-risk applications, like e-voting or e-Passports, where the strongest possible security is desirable. Indeed both SimTF and strSimTF security *can* be achieved, e.g. by our scheme.

However, simulation-based security is too strong for resource-constrained devices, as it does *not* enable efficient protocols. In such scenarios, our game-based GameTF model is more appropriate, capturing the intuition of terrorist fraud resistance, but enabling more efficient schemes e.g. [14].

References

1. Avoine, G., Bingol, M.A., Karda, S., Lauradoux, C., Martin, B.: A formal framework for analyzing RFID distance bounding protocols. Journal of Computer Security - Special Issue on RFID System Security (2010)
2. Avoine, G., Lauradoux, C., Martin, B.: How secret-sharing can defeat terrorist fraud. In: Proceedings of the Fourth ACM Conference on Wireless Network Security, WISEC 2011, pp. 145–156. ACM Press (2011)
3. Brands, S., Chaum, D.: Distance-bounding protocols. In: Helleseth, T. (ed.) EUROCRYPT 1993. LNCS, vol. 765, pp. 344–359. Springer, Heidelberg (1994)
4. 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)
5. Drimer, S., Murdoch, S.J.: Keep your enemies close: distance bounding against smartcard relay attacks. In: Proc. of the 16th USENIX Security Symposium on USENIX Security Symposium, article no. 7. ACM Press (2007)
6. Dürholz, U., Fischlin, M., Kasper, M., Onete, C.: A formal approach to distance-bounding RFID protocols. In: Lai, X., Zhou, J., Li, H. (eds.) ISC 2011. LNCS, vol. 7001, pp. 47–62. Springer, Heidelberg (2011)
7. Fischlin, M., Onete, C.: Provably secure distance-bounding: an analysis of prominent protocols. Accepted at the 6th Conference on Security and Privacy in Wireless and Mobile Networks ACM WISec 2013, Proceedings will follow (2013), http://eprint.iacr.org/2012/128.pdf
8. Francillon, A., Danev, B., Čapkun, S.: Relay Attacks on Passive Keyless Entry and Start Systems in Modern Cars (2010), http://eprint.iacr.org/2010/332

9. Haataja, K., Toivanen, P.: Two practical man-in-the-middle attacks on bluetooth secure simple pairing and countermeasures. Transactions on Wireless Communications 9(1), 384–392 (2010)
10. Hancke, G.P.: A practical relay attack on ISO 14443 proximity cards (2005), http://www.cl.cam.ac.uk/gh275/relay.pdf
11. Hancke, G.P., Kuhn, M.G.: An RFID distance bounding protocol. In: SECURECOMM, pp. 67–73. ACM Press (2005)
12. Hlaváč, M., Tomáč, R.: A Note on the Relay Attacks on e-Passports (2007), http://eprint.iacr.org/2007/244.pdf
13. Kfir, Z., Wool, A.: Picking virtual pockets using relay attacks on contactless smartcard systems. In: Conference on Security and Privacy for Emergency Areas in Communication Networks – SecureComm 2005, pp. 47–58. IEEE (2005)
14. Kim, C.H., Avoine, G., Koeune, F., Standaert, F.-X., Pereira, O.: The Swiss-Knife RFID distance bounding protocol. In: Lee, P.J., Cheon, J.H. (eds.) ICISC 2008. LNCS, vol. 5461, pp. 98–115. Springer, Heidelberg (2009)
15. Levi, A., Çetintaş, E., Aydos, M., Koç, Ç.K., Çağlayan, M.U.: Relay attacks on bluetooth authentication and solutions. In: Aykanat, C., Dayar, T., Körpeoğlu, İ. (eds.) ISCIS 2004. LNCS, vol. 3280, pp. 278–288. Springer, Heidelberg (2004)
16. Oren, Y., Wool, A.: Relay attacks on RFID-based electronic voting systems. Cryptology ePrint Archive, Report 2009/442 (2009), http://eprint.iacr.org/2009/422.pdf
17. Ranganathan, A., Tippenhauer, N.O., Škorić, B., Singelée, D., Čapkun, S.: Design and Implementation of a Terrorist Fraud Resilient Distance Bounding System. In: Foresti, S., Yung, M., Martinelli, F. (eds.) ESORICS 2012. LNCS, vol. 7459, pp. 415–432. Springer, Heidelberg (2012)
18. 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)

A Full Proofs of Security Diagram

Proof (sketch). For the proofs of (8) and (9) we use the strategy in [6], reusing their counterexample to prove (8). The counterexample for (9) is the scheme in (10). Finally (1) follows trivially from the strSimTF definition. The proofs are out of order, as we group similar proofs together.

Our separation for (2) relies on our scheme in Fig. 4, modified to run $2N_c$ time-critical rounds such that $\mathcal{R}$ reveals the even-indexed challenges in advance, sending them masked with pseudorandom bits during odd-indexed rounds. However, in odd-indexed rounds, the prover must just echo the challenge.The modified scheme preserves the properties of the original one. Though distance fraud adversaries can predict even-indexed challenges, they must guess the odd-indexed ones. Mafia fraud adversaries trivially echo odd-indexed responses, but learn nothing about the encrypted even-indexed challenges. Finally, the SimTF security proof still stands since the odd-indexed rounds are trivial for both $\mathcal{A}$ and $\mathcal{S}$. However, a strSimTF adversary echoes odd-indexed challenges, using its time-critical interactions to forward the encrypted challenges in advance (thus receiving also the even-indexed challenges). Since no key information is leaked, the simulator cannot authenticate.

We prove (7) similarly to Proposition 1: let $\mathcal{ID}$ be a mafia-fraud and $\mathsf{strSimTF}$-secure scheme, and assume it is *not* GameTF resistant. Assume that there exists a $(t, q_{\text{OBS}}, q_{\mathcal{R}}, q_{\mathcal{T}'})$-adversary $\mathcal{A}$ interacting in a $\mathsf{strSimTF}$ way such that: (i) $\mathcal{A}$ wins with non-negligible probability ϵ; (ii) all $(3t, q_{\text{OBS}}, q_{\mathcal{R}}, q_{\mathcal{R}})$ adversaries $\mathcal{A}^*$ using $\mathsf{view}_{\mathcal{A}}$ in a mafia fraud interaction wins w.p. at most $\mathbf{Adv}_{\mathcal{ID}}^{\text{mafia}}$. By $\mathsf{strSimTF}$ security, for adversary $\mathcal{A}$ there exists a simulator $\mathcal{S}$, which, given $\mathsf{view}_{\mathcal{A}}$, wins with probability $p_{\mathcal{S}} \geq \epsilon$. Now $\mathcal{A}^*$ run $\mathcal{S}$ as a black box on $\mathsf{view}_{\mathcal{A}}$, and wins w.p. $p_{\mathcal{S}} \geq \epsilon$. Following point (ii), $\mathcal{A}^*$ must win w.p. at most $\mathbf{Adv}_{\mathcal{ID}}^{\text{mafia}}$; thus $\epsilon \leq p_{\mathcal{S}} \leq \mathbf{Adv}_{\mathcal{ID}}^{\text{mafia}}$. Then $\mathbf{Adv}_{\mathcal{ID}}^{\text{mafia}}$ is non-negligible, contradicting the assumption that $\mathcal{ID}$ is mafia-fraud resistant.

For (8) we use the Hancke-Kuhn protocol [11] except that R^0, R^1 are computed as: $R^0||R^1 \leftarrow \mathsf{PRF}(sk, N_{\mathcal{R}}||N_{\mathcal{T}})$. The mafia fraud resistance of this scheme can be found in [7]; however, a GameTF adversary can query $\mathcal{T}'$ for $R^0||R^1$ in some session sid, giving no help for future authentication. Similarly, Mafia $\nrightarrow$ $\mathsf{strSimTF}$.

For (11) we use a trick from [6], changing the protocol in Fig. 2 to allow an adversary to change a flag that makes $\mathcal{R}$ run in a special mode, expecting the conjugated response values, rather than the originals. Now a mafia adversary passes the challenges to $\mathcal{T}$, but flips the responses. However, a GameTF-adversary cannot use this trick, as relay scheduling taints the phase (even if the bits are flipped). We use the same trick for (5). In $\mathsf{strSimTF}$ security, $\mathcal{S}$ must win with the same probability as $\mathcal{A}$. The helpfulness of GameTF adversaries depends though on mafia fraud resistance. If mafia fraud adversaries authenticate easily, any adversary is unhelpful, even one for which there exists a simulator as in $\mathsf{strSimTF}$ security. We modify the scheme in Figure 4 as in (11), thus making $\mathbf{Adv}_i^{\text{mafia}}d = 1$, for the (still) $\mathsf{strSimTF}$-secure scheme. However, the protocol is GameTF insecure: an adversary $\mathcal{A}$ receiving sk' from $\mathcal{T}'$: (i) wins with probability 1; (ii) all adversaries $\mathcal{A}^*$ with input $\mathsf{view}_{\mathcal{A}}$, win w.p. at most $1 = \mathbf{Adv}_{\mathcal{ID}}^{\text{mafia}}$. The same strategy proves (3), by replacing $\mathsf{strSimTF}$ with SimTF security, and (4) follows from (6) and (2).

For (6), we change the scheme in Fig. 2 to allow a dishonest prover to generate and send the adversary a particular cheating lazy-phase response, making $\mathcal{R}$ run a special mode, where the challenges are predictable, but only by a prover. This breaks $\mathsf{strSimTF}$ security, enabling the prover to help the adversary and then forward the correct challenges; however, a simulator is unable to learn the responses, even if it knows the challenges. By contrast, if a GameTF adversary uses the cheat, it is helpful to an adversary who can then use the Go-Early strategy to learn the correct responses. □