On RFID Privacy with Mutual Authentication and Tag Corruption

Frederik Armknecht[1], Ahmad-Reza Sadeghi[2],
Ivan Visconti[3], and Christian Wachsmann[2]

[1] University of Mannheim, Germany
armknecht@math.uni-mannheim.de
[2] Horst Görtz Institute for IT-Security (HGI), Ruhr-University Bochum, Germany
{ahmad.sadeghi,christian.wachsmann}@trust.rub.de
[3] Dipartimento di Informatica ed Applicazioni, University of Salerno, Italy
visconti@dia.unisa.it

Abstract. RFID systems have become increasingly popular and are already used in many real-life applications. Although very useful, RFIDs also introduce privacy risks since they carry identifying information that can be traced. Hence, several RFID privacy models have been proposed. However, they are often incomparable and in part do not reflect the capabilities of real-world adversaries. Recently, Paise and Vaudenay presented a general RFID security and privacy model that abstracts and unifies most previous approaches. This model defines mutual authentication (between the RFID tag and reader) and several privacy notions that capture adversaries with different tag corruption behavior and capabilities.

In this paper, we revisit the model proposed by Paise and Vaudenay and investigate some subtle issues such as tag corruption aspects. We show that in their formal definitions tag corruption discloses the temporary memory of tags and leads to the impossibility of achieving both mutual authentication and any reasonable notion of RFID privacy in their model. Moreover, we show that the strongest privacy notion (narrow-strong privacy) cannot be achieved simultaneously with reader authentication even if the adversary is not capable of corrupting a tag during the protocol execution.

Although our results are shown on the privacy definition by Paise and Vaudenay, they give insight to the difficulties of setting up a mature security and privacy model for RFID systems that aims at fulfilling the sophisticated requirements of real-life applications.

Keywords: RFID, Security Model, Privacy, Mutual Authentication.

1 Introduction

Radio Frequency Identification (RFID) enables RFID *readers* to perform fully automatic wireless identification of objects that are labeled with RFID *tags*, and is widely deployed to many applications (e.g., access control [1,2], electronic tickets [2,3], e-passports [4]). As pointed out in previous publications (see,

J. Zhou and M. Yung (Eds.): ACNS 2010, LNCS 6123, pp. 493–510, 2010.

e.g., [5,6,7]), this prevalence of RFID technology introduces various risks, in particular concerning the privacy of its users and holders. The most deterrent privacy risk concerns the tracking of users, which allows the creation and misuse of detailed user profiles. Thus, it is desired that an RFID system provides *anonymity* (confidentiality of the tag identity) as well as *untraceability* (unlinkability of the communication of a tag), even in case the state (e.g., the secret) of a tag has been disclosed.

The design of a secure privacy-preserving RFID scheme requires a careful analysis in an appropriate formal model. There is a large body of literature on security and privacy models for RFID (see, e.g., [8,9,10,11,12,13]). Existing solutions often do not consider important aspects like adversaries with access to auxiliary information (e.g., on whether the identification of a tag was successful), or the privacy of corrupted tags whose state has been disclosed. In particular, tag corruption is usually considered to happen only *before* and *after*, but *not during* a protocol execution. However, in practice there are a variety of side-channel attacks (see., e.g., [14,15,16]) that extract the state of a powered tag based on the observation of (e.g., the power consumption of) the tag *while* it is executing a protocol with the reader. Since RFID tags are usually cost-effective devices without expensive tamper-proof mechanisms [1,2], tag corruption is an important aspect to be covered by the underlying (formal) model. Though in literature, tag corruption during protocol execution is rarely considered. To the best of our knowledge, the security and privacy model in [10] is the only one that considers corruption of tags during protocol executions and proposes a protocol in this model. However, this model does not consider issues like the privacy of tags *after* they have been corrupted and privacy against adversaries with access to auxiliary information. Moreover, [10] only provides an informal security analysis of the proposed protocol. Recently, tag corruption during protocol executions has been informally discussed in [13]. However, the formal RFID security and privacy model proposed in [13] assumes that such attacks cannot occur. Moreover, [13] indicates informally (without giving formal arguments and proofs) that tag corruption during protocol execution may have an impact on the formal definitions of [9] and [11,12], which is basis for many subsequent works (see, e.g., [19,20,21,22,23,24,25,26]). The first papers addressing tag corruption during protocol execution in the model of [11] are [24,25], where it is shown that privacy can be achieved under the assumption that tag corruption during protocol execution can be detected by the tag.

In this paper, we focus on the security and privacy model by Paise and Vaudenay [12] (that is based on [11]), which we call the *PV-Model* (Paise-Vaudenay Model) in the following. The PV-Model is one of the most comprehensive RFID security and privacy models up to date since it captures many aspects of real world RFID systems and aims at abstracting most previous works in a single concise framework. It defines mutual authentication between RFID tags and readers and several privacy notions that correspond to adversaries with different tag corruption abilities. However, as we show in this paper, the PV-Model suffers

from subtle deficiencies and weaknesses that are mainly caused by tag corruption aspects: in the PV-Model, each tag maintains a state that can be divided into a persistent and a temporary part.[1] The *persistent state* subsumes all information that must be available to the tag in more than one interaction with the reader (e.g., the authentication secret of the tag) and can be updated during the interaction with the reader. The *temporary state* consists of all ephemeral information that is discarded by the tag after each interaction with the reader (e.g., the randomness used by the tag). The PV-Model shows that if the adversary can obtain *both* the persistent *and* the temporary tag state by tag corruption, then it is impossible to achieve any notion of privacy that allows tag corruption in their model. They address this issue by assuming that each tag erases its temporary state each time it gets out of the reading range of the adversary. However, this assumption leaves open the possibility to corrupt a tag *while* it is in the reading range of the adversary, i.e., *before* its temporary state is erased. In particular, the PV-Model allows the adversary to corrupt a tag *while* it is executing the authentication protocol with the reader.

Contribution. In this paper, we point out subtle weaknesses and deficiencies in the PV-Model. First, we show that the assumption of erasing temporary tag states whenever a tag gets out of the reading range of the adversary made by the PV-Model is not strong enough. We prove that, even under this assumption, it is *impossible* to achieve reader authentication and simultaneously *any* notion of privacy that allows tag corruption. This implies that the PV-Model cannot provide privacy along with mutual authentication without relying on tamper-proof hardware, which is unrealistic for low-cost RFID tags. Consequently, we show attacks on the privacy goals of two of the three schemes presented in [12].

Our second contribution is to show that even under the strong assumption that the temporary tag state is not subject to corruption attacks, some privacy notions still remain impossible in the PV-Model. This implies that the third protocol of [12] does not achieve its claimed privacy goal. On the good side, we prove that some privacy notions can be met under the assumption that temporary tag states are not disclosed by tag corruption.

Although our results are shown on the privacy definition by Paise and Vaudenay, we believe that our work is helpful for developing a mature security and privacy model for RFID systems that fulfills the sophisticated requirements of real-life applications.

2 RFID System and Requirement Analysis

System model. An RFID system consists of at least an operator $\mathcal{I}$, a reader $\mathcal{R}$ and a tag $\mathcal{T}$. $\mathcal{I}$ is the entity that enrolls and maintains the RFID system. Hence, $\mathcal{I}$ initializes $\mathcal{T}$ and $\mathcal{R}$ before they are deployed in the system. $\mathcal{T}$ and $\mathcal{R}$ are called *legitimate* if they have been initialized by $\mathcal{I}$. In many applications $\mathcal{T}$

[1] During a protocol execution, tags could store some temporary information that allows them to verify the response of the reader.

is a hardware token with constrained computing and memory capabilities that is equipped with a radio interface [1,2]. All information (e.g., secrets and data) stored on $\mathcal{T}$ is denoted as the *state* of $\mathcal{T}$. Usually $\mathcal{T}$ is attached to some object or carried by a user of the RFID system. $\mathcal{R}$ is a stationary or mobile computing device that interacts with $\mathcal{T}$ when $\mathcal{T}$ gets into the reading range of $\mathcal{R}$. The main purpose of this interaction usually is the authentication of $\mathcal{T}$ to $\mathcal{R}$. Depending on the use case, $\mathcal{R}$ may also authenticate to $\mathcal{T}$ and/or obtain additional information like the identity of $\mathcal{T}$. $\mathcal{R}$ can have a sporadic or permanent online connection to some backend system $\mathcal{D}$, which typically is a database maintaining detailed information on all tags in the system. $\mathcal{D}$ is initialized and maintained by $\mathcal{I}$ and can be read and updated by $\mathcal{R}$.

Trust and adversary model. The operator $\mathcal{I}$ maintains the RFID system, and thus is considered to behave correctly. However, $\mathcal{I}$ may be curious since he may collect user information (see, e.g., [5,6]). Since $\mathcal{T}$ and $\mathcal{R}$ communicate over a radio link, any entity can eavesdrop and manipulate this communication, even from outside the nominal reading range of $\mathcal{R}$ and $\mathcal{T}$ [27]. Thus, the adversary $\mathcal{A}$ can be every (potentially unknown) entity. Besides the communication between $\mathcal{T}$ and $\mathcal{R}$, $\mathcal{A}$ can also obtain useful auxiliary information (e.g., by visual observation) on whether $\mathcal{R}$ accepted $\mathcal{T}$ as a legitimate tag [9,12]. Most commercial RFID tags are cost-efficient devices without expensive protection mechanisms against physical tampering [1,2]. Hence, $\mathcal{A}$ can physically attack (*corrupt*) $\mathcal{T}$ and obtain its state. In practice, RFID readers are embedded devices that can be integrated into mobile devices (e.g., mobile phones or PDAs) or computers. The resulting complexity exposes them to sophisticated hard- and software attacks (e.g., viruses and Trojans). This problem aggravates for mobile readers that can easily be lost or stolen. Hence, $\mathcal{A}$ can get full control over $\mathcal{R}$ [28,29,30].

Security and privacy objectives. The most deterrent privacy risk concerns the *tracking* of tag users, which allows the creation and misuse of detailed user profiles in an RFID system [6]. For instance, detailed movement profiles can leak sensitive information on the personal habits and interests of the tag user. The major security threats are to create illegitimate (*forge*) tags that are accepted by honest readers, to simulate (*impersonate*) or to copy (*clone*) legitimate tags, and to permanently prevent users from using the RFID system (*denial-of-service*) [10]. Thus, an RFID system should provide *anonymity* (confidentiality of the tag identity) as well as *untraceability* (unlinkability of the communication of a tag) even if the state of a tag has been disclosed. The main security objective is to ensure that only legitimate tags are accepted by honest readers (*tag authentication*). Most use cases additionally require $\mathcal{R}$ to determine the authentic tag identity (*tag identification*). Moreover, there are several applications (e.g., electronic tickets) where reader authentication is a fundamental security property. However, there are also use cases (e.g., electronic product labels) that do not require reader authentication.

3 Notation

For a finite set S, $|S|$ denotes the size of S whereas for an integer n the term $|n|$ means the bit-length of n. The term $s \in_R S$ means the assignment of a uniformly chosen element of S to s. Let A be a probabilistic algorithm. Then $y \leftarrow \mathsf{A}(x)$ means that on input x, algorithm A assigns its output to variable y. The term $[\mathsf{A}(x)]$ denotes the set of all possible outputs of A on input x. $\mathsf{A}_K(x)$ means that the output of A depends on x and some additional parameter K (e.g., a secret key). The term $\mathsf{Prot}[\mathsf{A}: x_\mathsf{A};\ \mathsf{B}: x_\mathsf{B};\ *: x_{pub}] \rightarrow [\mathsf{A}: y_\mathsf{A};\ \mathsf{B}: y_\mathsf{B}]$ denotes an interactive protocol Prot between two algorithms A and B. Hereby, A (resp. B) gets a private input x_A (resp. x_B) and a public input x_{pub}. While A (resp. B) is operating, it can interact with B (resp. A). After the protocol terminates, A (resp. B) returns y_A (resp. y_B). Let E be some event (e.g., the result of a security experiment), then $\Pr[E]$ denotes the probability that E occurs. Probability $\epsilon(l)$ is called *negligible* if for all polynomials f it holds that $\epsilon(l) \leq 1/f(l)$ for all sufficiently large l. Probability $1-\epsilon(l)$ is called *overwhelming* if $\epsilon(l)$ is negligible.

4 The PV-Model

In this section, we recall Paise and Vaudenay's model (PV-Model) [12], which is one of the most comprehensive RFID security and privacy models up to date.

4.1 System Model

The PV-Model considers RFID systems that consist of a single operator $\mathcal{I}$, a single reader $\mathcal{R}$ and a polynomial number of tags $\mathcal{T}$. $\mathcal{R}$ is assumed to be capable of performing public-key cryptography and of handling multiple instances of the mutual authentication protocol with different tags in parallel. Each tag $\mathcal{T}$ is a passive device, i.e., it does not have its own power supply but is powered by the electromagnetic field of $\mathcal{R}$. Hence, $\mathcal{T}$ cannot initiate communication, has a narrow communication range (i.e., a few centimeters to meters) and erases its temporary state (i.e., all session-specific information and randomness) after it gets out of the reading range of $\mathcal{R}$. Each $\mathcal{T}$ is assumed to be capable of performing basic cryptographic functions like random number generation, hashing and symmetric-key encryption. The authors of [12] also use public-key encryption, although it exceeds the capabilities of most currently available RFID tags.

The operator $\mathcal{I}$ sets up the reader $\mathcal{R}$ and all tags $\mathcal{T}$. Hence, there are two setup algorithms to generate the system parameters (e.g., keys) of $\mathcal{R}$ and $\mathcal{T}$. Mutual authentication is covered by a third protocol between $\mathcal{T}$ and $\mathcal{R}$.

Definition 1 (RFID System [12]). *An RFID system is a tuple of probabilistic polynomial time (p.p.t.) algorithms* $(\mathcal{R}, \mathcal{T}, \mathsf{SetupReader}, \mathsf{SetupTag}, \mathsf{Ident})$ *that are defined as follows:*

$\mathsf{SetupReader}(1^l) \rightarrow (sk_\mathcal{R}, pk_\mathcal{R}, \mathtt{DB})$ *On input of a security parameter l, this algorithm creates the public system parameters $pk_\mathcal{R}$ that are known to all entities.*

Moreover, it creates the secret system parameters $sk_{\mathcal{R}}$ and a database DB *that can only be accessed by* $\mathcal{R}$.

$\mathsf{SetupTag}_{pk_{\mathcal{R}}}(\mathtt{ID}) \rightarrow (K, S)$ *uses* $pk_{\mathcal{R}}$ *to generate a tag secret* K *and tag state* S*, initializes* $\mathcal{T}_{\mathtt{ID}}$ *with* S*, and stores* $(\mathtt{ID}, K)$ *in* DB.

$\mathsf{Ident}[\mathcal{T}_{\mathtt{ID}} : S;\ \mathcal{R} : sk_{\mathcal{R}}, \mathtt{DB};\ * : pk_{\mathcal{R}}] \rightarrow [\mathcal{T}_{\mathtt{ID}} : out_{\mathcal{T}_{\mathtt{ID}}};\ \mathcal{R} : out_{\mathcal{R}}]$ *is an interactive protocol between* $\mathcal{T}_{\mathtt{ID}}$ *and* $\mathcal{R}$*.* $\mathcal{T}_{\mathtt{ID}}$ *takes as input its current state* S *while* $\mathcal{R}$ *has input* $sk_{\mathcal{R}}$ *and* DB*. The common input to all parties is* $pk_{\mathcal{R}}$*. After the protocol terminates,* $\mathcal{R}$ *returns either the identity* ID *of* $\mathcal{T}_{\mathtt{ID}}$ *or* $\perp$ *to indicate that* $\mathcal{T}_{\mathtt{ID}}$ *is not a legitimate tag.* $\mathcal{T}_{\mathtt{ID}}$ *returns either* ok *to indicate that* $\mathcal{R}$ *is legitimate or* $\perp$ *otherwise.*

Correctness describes the honest behavior of legitimate tags and the reader $\mathcal{R}$.

Definition 2 (Correctness [12]). *An RFID system (Definition 1) is correct if* $\forall\ l,\ \forall\ (sk_{\mathcal{R}}, pk_{\mathcal{R}}, \mathtt{DB}) \in [\mathsf{SetupReader}(1^l)]$*, and* $\forall\ (K, S) \in [\mathsf{SetupTag}_{pk_{\mathcal{R}}}(\mathtt{ID})]$ $\Pr\left[\mathsf{Ident}[\mathcal{T}_{\mathtt{ID}} : S;\ \mathcal{R} : sk_{\mathcal{R}}, \mathtt{DB};\ * : pk_{\mathcal{R}}] \rightarrow [\mathcal{T}_{\mathtt{ID}} : \mathtt{ok};\ \mathcal{R} : \mathtt{ID}]\right]$ *is overwhelming.*

4.2 Trust and Adversary Model

The PV-Model assumes the issuer $\mathcal{I}$, the backend database $\mathcal{D}$ and the readers to be trusted, whereas a tag $\mathcal{T}$ can be compromised. All readers and $\mathcal{D}$ are subsumed to *one single* reader entity $\mathcal{R}$ that cannot be corrupted. The PV-Model defines privacy and security as experiments, where an adversary $\mathcal{A}$ interacts with a set of oracles that model the capabilities of $\mathcal{A}$. These oracles are:

$\mathsf{CreateTag}^b(\mathtt{ID})$ Allows $\mathcal{A}$ to set up a tag $\mathcal{T}_{\mathtt{ID}}$ with identifier ID by internally calling $\mathsf{SetupTag}_{pk_{\mathcal{R}}}(\mathtt{ID})$ to create (K, S) for $\mathcal{T}_{\mathtt{ID}}$. If input $b = 1$, then $(\mathtt{ID}, K)$ is added to DB. If $b = 0$, then $(\mathtt{ID}, K)$ is *not* added to DB.

$\mathsf{Draw}(\delta) \rightarrow (vtag_1, b_1, \ldots, vtag_n, b_n)$ Initially, $\mathcal{A}$ cannot interact with any tag but must query Draw to get access to a set of tags that has been chosen according to a given probability distribution δ. $\mathcal{A}$ knows the tags he can interact with by temporary tag identifiers $vtag_1, \ldots, vtag_n$. Draw manages a secret table Γ that links each temporary tag identifier $vtag_i$ to the corresponding real tag identifier $\mathtt{ID}_i$ (i.e., $\Gamma[vtag_i] = \mathtt{ID}_i$). Moreover, Draw provides $\mathcal{A}$ with information on whether the tags are legitimate ($b_i = 1$) or not ($b_i = 0$).

$\mathsf{Free}(vtag)$ Makes $vtag$ inaccessible to $\mathcal{A}$ such that $\mathcal{A}$ can no longer interact with $vtag$ until it is made accessible again (under a new temporary identifier $vtag'$) by another Draw query.

$\mathsf{Launch}(\,) \rightarrow \pi$ Makes $\mathcal{R}$ to start a new instance π of the Ident protocol.

$\mathsf{SendReader}(m, \pi) \rightarrow m'$ Sends a message m to instance π of the Ident protocol that is running on $\mathcal{R}$. $\mathcal{R}$ interprets m as a protocol message of instance π of the Ident protocol and responds with a message m'.

$\mathsf{SendTag}(m, vtag) \rightarrow m'$ Sends a message m to the tag $\mathcal{T}_{\mathtt{ID}}$ that is known as $vtag$ to $\mathcal{A}$. $\mathcal{T}_{\mathtt{ID}}$ interprets m as a protocol message of the Ident protocol and responds with a message m'.

$\mathsf{Result}(\pi)$ Returns 1 if instance π of the Ident protocol has been completed and the tag that participated in instance π has been accepted by $\mathcal{R}$. Otherwise Result returns 0.

$\mathsf{Corrupt}(vtag) \rightarrow S$ Returns the current state S (i.e., all information stored in the memory) of the tag $\mathcal{T}_{\mathtt{ID}}$ that is known as *vtag* to $\mathcal{A}$.

The PV-Model distinguishes eight adversary classes, which differ in (i) their ability to corrupt tags and (ii) the availability of auxiliary information (i.e., the ability to access the Corrupt and Result oracle, respectively).

Definition 3 (Adversary Classes [12]). *An adversary is a p.p.t. algorithm that has arbitrary access to all oracles described in Section 4.2. Weak adversaries cannot access the* Corrupt *oracle. Forward adversaries can no longer query any other oracle than* Corrupt *after they made the first* Corrupt *query. Destructive adversaries cannot query any oracle for vtag again after they made a* $\mathsf{Corrupt}(vtag)$ *query. Strong adversaries have no restrictions on the use of the* Corrupt *oracle. Narrow adversaries cannot access the* Result *oracle.*

Tag corruption aspects. Depending on the concrete scenario one could have that the temporary tag state is disclosed under tag corruption. In general, any concrete scenario will range between the following two extremes: (i) corruption discloses the full temporary tag state or (ii) corruption does not disclose any information on the temporary tag state. In Sections 5 and 6, we will prove that in both cases some privacy notions are impossible to achieve in the PV-Model. Thus, *independently* of any possible interpretation of tag corruption, impossibility results exist that revisit the claims of [12].

4.3 Security Definition

The security definition of the PV-Model focuses on attacks where the adversary aims to impersonate or forge a legitimate tag $\mathcal{T}$ or the reader $\mathcal{R}$. It does *not* capture availability and security against cloning.

Tag authentication. The definition of tag authentication is based on a security experiment $\mathbf{Exp}^{\mathcal{T}\text{-aut}}_{\mathcal{A}_{\text{sec}}}$, where a strong adversary $\mathcal{A}_{\text{sec}}$ (Definition 3) must make the reader $\mathcal{R}$ to identify some tag $\mathcal{T}_{\mathtt{ID}}$ in some instance π of the Ident protocol. To exclude trivial attacks (e.g., relay attacks), $\mathcal{A}_{\text{sec}}$ is not allowed to simply forward all the messages from $\mathcal{T}_{\mathtt{ID}}$ to $\mathcal{R}$ in instance π nor to corrupt $\mathcal{T}_{\mathtt{ID}}$. This means that at least some of the protocol messages that made $\mathcal{R}$ to return ID must have been computed by $\mathcal{A}_{\text{sec}}$ without knowing the secrets of $\mathcal{T}_{\mathtt{ID}}$. With $\mathbf{Exp}^{\mathcal{T}\text{-aut}}_{\mathcal{A}_{\text{sec}}} = 1$ we denote the case where $\mathcal{A}_{\text{sec}}$ wins the security experiment.

Definition 4 (Tag Authentication [12]). *An RFID system (Definition 1) achieves tag authentication if for every strong adversary* $\mathcal{A}_{\text{sec}}$ *(Definition 3)* $\Pr[\mathbf{Exp}^{\mathcal{T}\text{-aut}}_{\mathcal{A}_{\text{sec}}} = 1]$ *is negligible.*

Note that tag authentication is a critical property and hence must be preserved even against strong adversaries.

Reader authentication. The definition of reader authentication is based on a security experiment $\mathbf{Exp}^{\mathcal{R}\text{-aut}}_{\mathcal{A}_{\text{sec}}}$, where a strong adversary $\mathcal{A}_{\text{sec}}$ (Definition 3) must successfully impersonate the reader $\mathcal{R}$ to a legitimate tag $\mathcal{T}_{\texttt{ID}}$. Also here, to exclude trivial attacks, $\mathcal{A}_{\text{sec}}$ must achieve this without simply forwarding the protocol messages from $\mathcal{R}$ to $\mathcal{T}_{\texttt{ID}}$. With $\mathbf{Exp}^{\mathcal{R}\text{-aut}}_{\mathcal{A}_{\text{sec}}} = 1$ we denote the case where $\mathcal{A}_{\text{sec}}$ wins the security experiment.

Definition 5 (Reader Authentication [12]). *An RFID system (Definition 1) achieves reader authentication if for every strong adversary $\mathcal{A}_{\text{sec}}$ (Definition 3)* $\Pr[\mathbf{Exp}^{\mathcal{R}\text{-aut}}_{\mathcal{A}_{\text{sec}}} = 1]$ *is negligible.*

4.4 Privacy Definition

The privacy definition of the PV-Model is very flexible and, dependent on the adversary class (see Definition 3), it covers different notions of privacy. It captures anonymity and unlinkability and focuses on the privacy leakage of the communication of tags with the reader. It is based on the existence of a simulator $\mathcal{B}$, called *blinder*, that can simulate every tag $\mathcal{T}$ and the reader $\mathcal{R}$ without knowing their secrets such that an adversary $\mathcal{A}_{\text{prv}}$ cannot distinguish whether it is interacting with the real or the simulated RFID system.

The privacy definition can be formalized by the following experiment $\mathbf{Exp}^{\text{prv-}b}_{\mathcal{A}_{\text{prv}}}$: Let $\mathcal{A}_{\text{prv}}$ be an adversary according to Definition 3, l be a given security parameter and $b \in_R \{0,1\}$. In the first phase of the experiment, $\mathcal{R}$ is initialized with $(sk_{\mathcal{R}}, pk_{\mathcal{R}}, \texttt{DB}) \leftarrow \mathsf{SetupReader}(1^l)$. The public key $pk_{\mathcal{R}}$ is given to $\mathcal{A}_{\text{prv}}$ and $\mathcal{B}$. Now, $\mathcal{A}_{\text{prv}}$ is allowed to arbitrarily interact with all oracles defined in Section 4.2. Hereby, $\mathcal{A}_{\text{prv}}$ is subject to the restrictions of its corresponding adversary class (see Definition 3). If $b = 1$, all queries to the Launch, SendReader, SendTag and Result oracles are redirected to and answered by $\mathcal{B}$. Hereby, $\mathcal{B}$ can observe all queries $\mathcal{A}_{\text{prv}}$ makes to all other oracles that are not simulated by $\mathcal{B}$ and the corresponding responses ("$\mathcal{B}$ sees what $\mathcal{A}_{\text{prv}}$ sees"). After a polynomial number of oracle queries, the second phase of the experiment starts, where $\mathcal{A}_{\text{prv}}$ can no longer interact with the oracles but is given the secret table Γ of the Draw oracle. Finally, $\mathcal{A}_{\text{prv}}$ returns a bit b', which we denote with $\mathbf{Exp}^{\text{prv-}b}_{\mathcal{A}_{\text{prv}}} = b'$.

Definition 6 (Privacy [11]). *Let C be an adversary class according to Definition 3. An RFID system (Definition 1) is C-private if for every adversary $\mathcal{A}_{\text{prv}}$ of C there exists a p.p.t. algorithm $\mathcal{B}$ (blinder) such that the advantage*

$$\mathbf{Adv}^{\text{prv}}_{\mathcal{A}_{\text{prv}}} = \left| \Pr\left[\mathbf{Exp}^{\text{prv-0}}_{\mathcal{A}_{\text{prv}}} = 1\right] - \Pr\left[\mathbf{Exp}^{\text{prv-1}}_{\mathcal{A}_{\text{prv}}} = 1\right] \right|$$

of $\mathcal{A}_{\text{prv}}$ is negligible. $\mathcal{B}$ simulates the Launch, SendReader, SendTag *and* Result *oracles to $\mathcal{A}_{\text{prv}}$ without having access to $sk_{\mathcal{R}}$ and* DB. *Hereby, all oracle queries $\mathcal{A}_{\text{prv}}$ makes and their corresponding responses are also sent to $\mathcal{B}$.*

Figure 1 summarizes all privacy notions defined in the PV-Model and shows the relations among them. It has been shown that strong privacy is impossible while the technical feasibility of destructive privacy currently is an open problem [11].

$$\begin{array}{ccccccc} \text{Strong} & \Rightarrow & \text{Destructive} & \Rightarrow & \text{Forward} & \Rightarrow & \text{Weak} \\ \Downarrow & & \Downarrow & & \Downarrow & & \Downarrow \\ \text{Narrow-Strong} & \Rightarrow & \text{Narrow-Destructive} & \Rightarrow & \text{Narrow-Forward} & \Rightarrow & \text{Narrow-Weak} \end{array}$$

Fig. 1. Privacy notions defined in the PV-Model and their relations

5 Corruption with Temporary State Disclosure

We now point out a subtle conceptual weakness of the PV-Model and revisit two of the claims given in [12] about their protocols, that do not achieve the claimed privacy properties. We first illustrate our adversarial strategy by showing how tag corruption (as defined in the PV-Model) can be exploited to attack one of the protocols proposed in [12]. Then we generalize our attack to the class of narrow-forward private protocols, which finally leads to our first impossibility result: in the PV-Model it is *impossible* to achieve any notion of privacy simultaneously with reader authentication (under temporary state disclosure) except for the weak and narrow-weak privacy notions.

We stress that this impossibility result is due to the fact that, according to the formal definitions of the PV-Model, the adversary can obtain the full state (including the temporary memory) of a tag by corrupting the tag *while* it is executing a protocol with the reader. In face of side-channel attacks (see, e.g., [15,16]), such attacks can be feasible in practice (in particular against low-cost RFID tags) and hence, must be formally considered. Although [12] informally discusses an issue related to tag corruption during protocol execution, we show that such attacks are *not* adequately captured by the formal definitions of the PV-Model. Hence, the only achievable privacy notions are those where the adversary is not allowed to corrupt tags at all. Since in practice tag corruption is realistic, this implies that using the PV-Model is not helpful when reader authentication and a reasonable notion of privacy are needed.

5.1 Illustrative Example

We first show our attack on protocol 3 of [12], which is claimed to provide reader authentication while being narrow-strong private.[2]

Protocol description. Let $l, k, \alpha, \beta \in \mathbb{N}$ be given security parameters. $\mathcal{R}$ is initialized with the credentials database $\mathtt{DB}$ and the secret key $sk_{\mathcal{R}}$ of a CCA-secure public-key encryption scheme, while $\mathcal{T}_{\mathtt{ID}}$ is initialized with the corresponding public key $pk_{\mathcal{R}}$, a given tag identifier $\mathtt{ID}$ and $K \in_R \{0,1\}^k$. The Ident protocol is illustrated in Figure 2 and works as follows: Upon receipt of a, $\mathcal{T}_{\mathtt{ID}}$ stores b in its temporary memory and sends c to $\mathcal{R}$.[3] If $\mathcal{T}_{\mathtt{ID}}$ is legitimate, $\mathcal{R}$ can identify $\mathcal{T}_{\mathtt{ID}}$

[2] In the full version of this paper [31], we show that the same attack can be launched on the second protocol of [12], which is claimed to be narrow-destructive private.

[3] Note that $\mathcal{T}_{\mathtt{ID}}$ interprets the protocol messages sent by $\mathcal{R}$ based on the value of b. If b is empty (i.e., has been erased), then $\mathcal{T}_{\mathtt{ID}}$ considers the message sent by $\mathcal{R}$ as the first, and as the third protocol message otherwise.

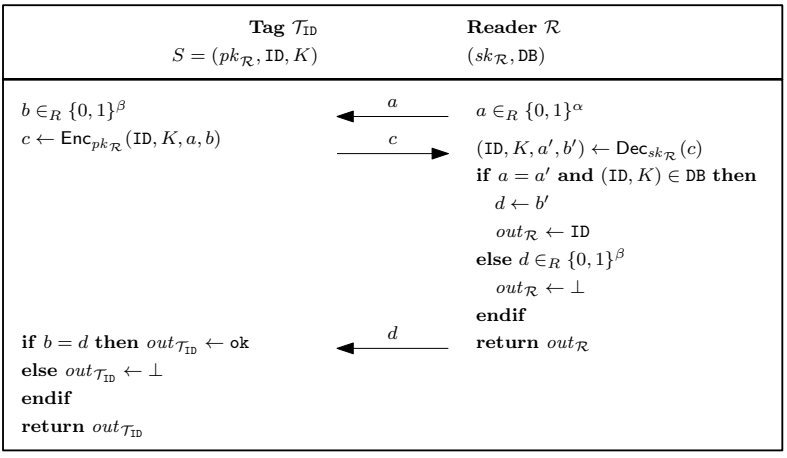

Fig. 2. Protocol 3 of [12]

by verifying that $a = a'$ and checking if $\mathtt{DB}$ contains$(\mathtt{ID}, K)$. If this is the case, $\mathcal{R}$ sends $d \leftarrow b'$ to $\mathcal{T}_{\mathtt{ID}}$ and outputs $\mathtt{ID}$. Otherwise, $\mathcal{R}$ sends a random d and returns $\perp$. $\mathcal{T}_{\mathtt{ID}}$ checks if $b = d$ and returns $\mathtt{ok}$ if this is the case and $\perp$ otherwise.

Attacking the protocol. Theorem 4 of [12] claims that the protocol shown in Figure 2 provides mutual authentication (Definitions 4 and 5) and narrow-strong privacy (Definition 6). Note that the proof of reader authentication given in [12] (which we believe to be correct) requires the encryption scheme to be CCA-secure and $2^{-\beta}$ to be negligible. In the following we show that the protocol in Figure 2 is not narrow-strong private.

Theorem 1. *The RFID protocol shown in Figure 2 does not achieve narrow-strong privacy (Definition 6), reader authentication (Definition 5) and correctness (Definition 2) at the same time.*

The full proof of Theorem 1 can be found in the full version of this paper [31].

Proof (Theorem 1, Sketch). The idea of the proof is to construct a narrow-strong adversary $\mathcal{A}_{\text{prv}}$ that violates Definition 6 by corrupting a legitimate tag $\mathcal{T}_{\mathtt{ID}}$ during the Ident protocol. More precisely, $\mathcal{A}_{\text{prv}}$ corrupts $\mathcal{T}_{\mathtt{ID}}$ right *before* it receives d, which authenticates $\mathcal{R}$ to $\mathcal{T}_{\mathtt{ID}}$. This allows $\mathcal{A}_{\text{prv}}$ to obtain the complete state $S = (pk_{\mathcal{R}}, \mathtt{ID}, K, b)$ of $\mathcal{T}_{\mathtt{ID}}$, including its temporary state b. Hence, $\mathcal{A}_{\text{prv}}$ can perform the computation $\mathcal{T}_{\mathtt{ID}}$ would have done on receipt of d (i.e., $\mathcal{A}_{\text{prv}}$ can check whether $b = d$). In case $\mathcal{A}_{\text{prv}}$ interacted with the real oracles, then (due to correctness) with overwhelming probability this computation must result in acceptance of $\mathcal{R}$ (i.e., it must hold that $b = d$). However, if $\mathcal{A}_{\text{prv}}$ interacted with the blinder $\mathcal{B}$, then the computation done by $\mathcal{A}_{\text{prv}}$ leads to rejection of $\mathcal{R}$ (i.e., it holds that $b \neq d$) with overwhelming probability. This is due to reader authentication (Definition 5). Hence, $\mathcal{A}_{\text{prv}}$ can distinguish the real oracles from

the simulation by $\mathcal{B}$ with non-negligible advantage, which violates narrow-strong privacy (Definition 6). □

5.2 Impossibility of Narrow-Forward-Privacy

Now we generalize the attack shown in Section 5.1. To prove our first impossibility result, we need the following lemma, which we will prove further below.

Lemma 1. *If for every narrow-forward adversary $\mathcal{A}_{\text{prv}}$ there is a blinder $\mathcal{B}$ such that $\mathbf{Adv}^{\text{prv}}_{\mathcal{A}_{\text{prv}}}$ is negligible (Definition 6), then $\mathcal{B}$ can be used to construct an adversary $\mathcal{A}^{\mathcal{B}}_{\text{sec}}$ such that $\Pr[\mathbf{Exp}^{\mathcal{R}\text{-aut}}_{\mathcal{A}^{\mathcal{B}}_{\text{sec}}} = 1]$ is non-negligible (Definition 5).*

Based on this lemma, we set up the following theorem, which we need later to prove our main impossibility result:

Theorem 2. *There is no RFID system (Definition 1) that achieves both reader authentication (Definition 5) and narrow-forward privacy (Definition 6).*

Proof (Theorem 2). Let $\mathcal{A}_{\text{prv}}$ be a narrow-forward adversary (Definition 3). Definition 6 requires the existence of a blinder $\mathcal{B}$ such that $\mathcal{A}_{\text{prv}}$ cannot distinguish between $\mathcal{B}$ and the real oracles. From Lemma 1 it follows that $\mathcal{B}$ can be used to impersonate $\mathcal{R}$ to any legitimate tag $\mathcal{T}$ with non-negligible probability. Hence, the existence of $\mathcal{B}$ contradicts reader authentication (Definition 5). □

Proof (Lemma 1). First, we show how to construct $\mathcal{A}^{\mathcal{B}}_{\text{sec}}$ from $\mathcal{B}$. Second, we prove that $\mathcal{A}^{\mathcal{B}}_{\text{sec}}$ violates reader authentication (Definition 5) if $\mathbf{Adv}^{\text{prv}}_{\mathcal{A}_{\text{prv}}}$ is negligible for every narrow-forward $\mathcal{A}_{\text{prv}}$ (Definition 3).

Let $q_{\mathcal{R}} \in \mathbb{N}$ with $q_{\mathcal{R}} > 0$ be the (expected) number of SendReader queries as specified by the Ident protocol and let $S^{\mathcal{R}}_i$ be the state of $\mathcal{R}$ after processing the i-th SendReader query. The initial reader state $S^{\mathcal{R}}_0$ includes the public key $pk_{\mathcal{R}}$ and the secret key $sk_{\mathcal{R}}$ of $\mathcal{R}$ as well as a pointer to the credentials database DB. Note that during the processing of a SendReader query, $\mathcal{R}$ can update DB. $\mathcal{R}$ can be considered as a tuple of algorithms $(\mathcal{R}^{(1)}_\pi, \ldots, \mathcal{R}^{(q_{\mathcal{R}})}_\pi)$, where $\mathcal{R}^{(i)}_\pi$ represents the computation done by $\mathcal{R}$ when processing the i-th SendReader query in instance π of the Ident protocol. More formally: $(S^{\mathcal{R}}_1, m_1) \leftarrow \mathcal{R}^{(0)}_\pi(S^{\mathcal{R}}_0)$ and $(S^{\mathcal{R}}_{i+1}, m_{2i+1}) \leftarrow \mathcal{R}^{(i)}_\pi(S^{\mathcal{R}}_i, m_{2i})$ for $1 \leq i \leq q_{\mathcal{R}}$. Since tags are passive devices that cannot initiate communication $\mathcal{R}$ must send the first protocol message. Thus, $\mathcal{R}$ generates all protocol messages with odd indices whereas the tag $\mathcal{T}$ generates all messages with even indices. In case the Ident protocol specifies that $\mathcal{T}$ sends the last protocol message, then $m_{2q_{\mathcal{R}}+1}$ is the empty string. Let $q_{\mathcal{T}} \in \mathbb{N}$ with $q_{\mathcal{T}} > 0$ be the (expected) number of SendTag queries as specified by the Ident protocol and let $S^{\mathcal{T}}_i$ be the state of $\mathcal{T}$ after processing the i-th SendTag query. $\mathcal{T}$ can be represented as a tuple of algorithms $(\mathcal{T}^{(1)}, \ldots, \mathcal{T}^{(q_{\mathcal{T}})})$ where $\mathcal{T}^{(i)}$ means the computation done by $\mathcal{T}$ when processing the i-th SendTag query in an instance of the Ident protocol that involves $\mathcal{T}$. More formally: $(S^{\mathcal{T}}_{i+1}, m_{2i}) \leftarrow \mathcal{T}^{(i)}(S^{\mathcal{T}}_i, m_{2i-1})$ for $1 \leq i \leq q_{\mathcal{T}}$. Note that $m_{2q_{\mathcal{T}}}$ is the empty string if Ident specifies that $\mathcal{R}$ must send the last protocol message.

Alg. 1. Adversary $\mathcal{A}_{\text{sec}}^{\mathcal{B}}$ violating reader authentication

1: $\mathsf{CreateTag}(\mathtt{ID})$
2: $vtag \leftarrow \mathsf{Draw}(\Pr[\mathtt{ID}] = 1)$
3: $\pi \leftarrow \mathsf{Launch}(\,)$ ▷ simulated by $\mathcal{B}$
4: $m_1 \leftarrow \mathsf{SendReader}(-, \pi)$ ▷ simulated by $\mathcal{B}$
5: $i \leftarrow 1$
6: **while** $i < q_{\mathcal{R}}$ **do**
7: **if** $i \leq q_{\mathcal{T}}$ **then** $m_{2i} \leftarrow \mathsf{SendTag}(m_{2i-1}, vtag)$ ▷ simulated by $\mathcal{B}$
8: **end if**
9: $m_{2i+1} \leftarrow \mathsf{SendReader}(m_{2i}, \pi)$ ▷ simulated by $\mathcal{B}$
10: $i \leftarrow i + 1$
11: **end while**
12: $out_{\mathcal{T}_{\mathtt{ID}}} \leftarrow \mathsf{SendTag}(m_{2q_{\mathcal{R}}-1}, vtag)$ ▷ computed by $\mathcal{T}_{\mathtt{ID}}$

The idea of $\mathcal{A}_{\text{sec}}^{\mathcal{B}}$ is to internally use $\mathcal{B}$ as a black-box to simulate the final response of $\mathcal{R}$ that makes a legitimate tag $\mathcal{T}_{\mathtt{ID}}$ to accept $\mathcal{A}_{\text{sec}}^{\mathcal{B}}$ as $\mathcal{R}$. $\mathcal{A}_{\text{sec}}^{\mathcal{B}}$ is defined in Algorithm 1 and works as follows: First, $\mathcal{A}_{\text{sec}}^{\mathcal{B}}$ creates a legitimate tag $\mathcal{T}_{\mathtt{ID}}$ (step 1) and makes it accessible (step 2). Both steps are also shown to $\mathcal{B}$, which expects to observe all oracle queries. Then, $\mathcal{A}_{\text{sec}}^{\mathcal{B}}$ makes $\mathcal{B}$ to start a new instance π of the Ident protocol with $\mathcal{T}_{\mathtt{ID}}$ (step 3) and obtains the first protocol message m_1 generated by $\mathcal{B}$ (step 4). Now, $\mathcal{A}_{\text{sec}}^{\mathcal{B}}$ internally runs $\mathcal{B}$ that simulates $vtag$ and $\mathcal{R}$ until $\mathcal{B}$ returns the final reader message $m_{2q_{\mathcal{R}}-1}$ (steps 5–11). Finally, $\mathcal{A}_{\text{sec}}^{\mathcal{B}}$ sends $m_{2q_{\mathcal{R}}-1}$ to the real tag $\mathcal{T}_{\mathtt{ID}}$ (step 12). $\mathcal{A}_{\text{sec}}^{\mathcal{B}}$ succeeds if $\mathcal{T}_{\mathtt{ID}}$ accepts $m_{2q_{\mathcal{R}}-1}$ and returns $out_{\mathcal{T}_{\mathtt{ID}}} = \mathtt{ok}$, which means that $\mathcal{T}_{\mathtt{ID}}$ accepts $\mathcal{B}$ as $\mathcal{R}$. More formally, this means that:

$$\Pr\left[\mathbf{Exp}_{\mathcal{A}_{\text{sec}}^{\mathcal{B}}}^{\mathcal{R}\text{-aut}} = 1\right] = \Pr\left[\mathsf{Ident}\left[\mathcal{T}_{\mathtt{ID}} : S_0^{\mathcal{T}_{\mathtt{ID}}};\ \mathcal{A}_{\text{sec}}^{\mathcal{B}} : -;\ * : pk_{\mathcal{R}}\right] \rightarrow \left[\mathcal{T}_{\mathtt{ID}} : \mathtt{ok};\ \mathcal{A}_{\text{sec}}^{\mathcal{B}} : \cdot\right]\right] \quad (1)$$

We stress that this indeed is a valid attack w.r.t. Definition 5 since $\mathcal{A}_{\text{sec}}$ neither corrupts $\mathcal{T}_{\mathtt{ID}}$ nor just forwards the protocol messages between $\mathcal{R}$ and $\mathcal{T}_{\mathtt{ID}}$.

Next, we show that narrow-forward privacy (Definition 6) ensures that Eq. 1 is non-negligible. Therefore, we assume by contradiction that Eq. 1 is negligible, which implies that with overwhelming probability $p_{\perp}$ message $m_{2q_{\mathcal{R}}-1}$ generated by $\mathcal{B}$ makes $\mathcal{T}_{\mathtt{ID}}$ to return $out_{\mathcal{T}_{\mathtt{ID}}} = \perp$. In the following, we show that if $p_{\perp}$ is non-negligible, then there is a narrow-forward adversary $\mathcal{A}_{\text{prv}}$ that has non-negligible advantage $\mathbf{Adv}_{\mathcal{A}_{\text{prv}}}^{\text{prv}}$, which contradicts narrow-forward privacy (Definition 6). $\mathcal{A}_{\text{prv}}$ is defined in Algorithm 2 and works as follows: First, $\mathcal{A}_{\text{prv}}$ creates a legitimate tag $\mathcal{T}_{\mathtt{ID}}$ (step 1) and makes it accessible (step 2). Then, $\mathcal{A}_{\text{prv}}$ makes $\mathcal{R}$ to start a new instance π of the Ident protocol with $\mathcal{T}_{\mathtt{ID}}$ (step 3) and obtains the first protocol message m_1 from $\mathcal{R}$ (step 4). Now, $\mathcal{A}_{\text{prv}}$ eavesdrops on the execution of the Ident protocol up to to the point *after* $\mathcal{R}$ has sent its last protocol message $m_{2q_{\mathcal{R}}-1}$ (steps 5–11) and corrupts $\mathcal{T}_{\mathtt{ID}}$ just *before* $\mathcal{T}_{\mathtt{ID}}$ received $m_{2q_{\mathcal{R}}-1}$ (step 12). Next, $\mathcal{A}_{\text{prv}}$ performs the computation $\mathcal{T}_{\mathtt{ID}}$ would have done on receipt of $m_{2q_{\mathcal{R}}-1}$ (step 13). If this computation results in $out_{\mathcal{T}_{\mathtt{ID}}} = \mathtt{ok}$, $\mathcal{A}_{\text{prv}}$ returns 0 to indicate that he interacted with the real oracles (step 14). Otherwise, $\mathcal{A}_{\text{prv}}$ indicates the presence of $\mathcal{B}$ by returning 1 (step 15). Note that $\mathcal{A}_{\text{prv}}$ indeed

Alg. 2. Narrow-forward adversary $\mathcal{A}_{\text{prv}}$

1: $\mathsf{CreateTag}(\mathtt{ID})$
2: $vtag \leftarrow \mathsf{Draw}(\Pr[\mathtt{ID}] = 1)$
3: $\pi \leftarrow \mathsf{Launch}(\,)$
4: $m_1 \leftarrow \mathsf{SendReader}(-, \pi)$
5: $i \leftarrow 1$
6: **while** $i < q_{\mathcal{R}}$ **do**
7: **if** $i \leq q_{\mathcal{T}}$ **then** $m_{2i} \leftarrow \mathsf{SendTag}(m_{2i-1}, vtag)$
8: **end if**
9: $m_{2i+1} \leftarrow \mathsf{SendReader}(m_{2i}, \pi)$
10: $i \leftarrow i + 1$
11: **end while**
12: $S_{q_{\mathcal{R}}}^{\mathcal{T}_{\mathtt{ID}}} \leftarrow \mathsf{Corrupt}(vtag)$
13: $out_{\mathcal{T}_{\mathtt{ID}}} \leftarrow {\mathcal{T}_{\mathtt{ID}}}^{(q_{\mathcal{R}})}(S_{q_{\mathcal{R}}}^{\mathcal{T}_{\mathtt{ID}}}, m_{2q_{\mathcal{R}}-1})$
14: **if** $out_{\mathcal{T}_{\mathtt{ID}}} = \mathtt{ok}$ **then return** 0
15: **else return** 1
16: **end if**

is a narrow-forward adversary (Definition 3) since $\mathcal{A}_{\text{prv}}$ never queries Result and none of the oracles defined in Section 4.2 after corrupting $\mathcal{T}_{\mathtt{ID}}$.

Next, we determine $\mathbf{Adv}^{\text{prv}}_{\mathcal{A}_{\text{prv}}}$. Therefore, we first consider the case where $\mathcal{A}_{\text{prv}}$ interacts with the real oracles. Since $\mathcal{T}_{\mathtt{ID}}$ is legitimate, it follows from correctness (Definition 2) that $out_{\mathcal{T}_{\mathtt{ID}}} = \mathtt{ok}$ with overwhelming probability $p_{\mathtt{ok}}$. Hence,

$$\Pr\left[\mathbf{Exp}^{\text{prv-0}}_{\mathcal{A}_{\text{prv}}} = 1\right] = 1 - p_{\mathtt{ok}} \tag{2}$$

is negligible. Now, consider the case where $\mathcal{A}_{\text{prv}}$ interacts with $\mathcal{B}$. Note that by the contradicting hypothesis, $\mathcal{B}$ generates a protocol message $m_{2q_{\mathcal{R}}-1}$ that makes $\mathcal{T}_{\mathtt{ID}}$ to return $out_{\mathcal{T}_{\mathtt{ID}}} = \perp$ with overwhelming probability $p_{\perp}$. Thus, we have

$$\Pr\left[\mathbf{Exp}^{\text{prv-1}}_{\mathcal{A}_{\text{prv}}} = 1\right] = p_{\perp}\,. \tag{3}$$

From Eq. 2 and Eq. 3 it follows that $\mathbf{Adv}^{\text{prv}}_{\mathcal{A}_{\text{prv}}} = \left|1 - p_{\mathtt{ok}} - p_{\perp}\right|$. Note that both $p_{\mathtt{ok}}$ (due to correctness) and $p_{\perp}$ (by assumption) are overwhelming. Hence, $\mathbf{Adv}^{\text{prv}}_{\mathcal{A}_{\text{prv}}}$ is non-negligible, which contradicts narrow-forward privacy (Definition 6). In turn, this means that narrow-forward privacy ensures that Eq. 1 is non-negligible, which finishes the proof. □

Since the impossibility of narrow-forward privacy (see Theorem 2), implies the impossibility of all stronger privacy notions (see Figure 1), we have the following corollary, which corresponds to the first main claim of this paper.

Corollary 1. *In the PV-Model, there is no RFID system (Definition 1) that achieves both reader authentication (Definition 5) and any privacy notion that is different from weak and narrow-weak privacy (Definition 6) under temporary state disclosure.*

6 Corruption without Temporary State Disclosure

Our first impossibility result shows that the PV-Model requires further assumptions to evaluate the privacy properties of RFID systems where tag corruption is of concern. A natural question therefore is, whether one can achieve mutual authentication along with some form of privacy, if the temporary tag state is *not* disclosed. Hence, in this section we consider the case where corruption *only* reveals the persistent tag state but *no* information on the temporary tag state.

The attack and the impossibility result shown in Section 5 critically use the fact that in the PV-Model an adversary $\mathcal{A}_{\text{prv}}$ can learn the temporary state of a tag during the Ident protocol. This allows $\mathcal{A}_{\text{prv}}$ to verify the response of $\mathcal{R}$ (that may have been simulated by $\mathcal{B}$) and hence, due to reader authentication (Definition 5), $\mathcal{A}_{\text{prv}}$ can distinguish with non-negligible advantage between the real oracles and $\mathcal{B}$. However, if an adversary cannot obtain temporary tag states, he cannot perform this verification. Hence, the impossibility result we proved in Section 5 does not necessarily hold if the temporary state is safe to corruption.

6.1 Privacy under Corruption without Temporary State Disclosure

To show that it is possible to achieve a notion of privacy in the PV-Model that captures adversaries who can corrupt tags, we show that the protocol depicted in Figure 2 achieves narrow-forward privacy if corruption *only* reveals the persistent tag state but *no* information on the temporary tag state. Note that the attack presented in Section 5.1 cannot be applied since we now consider the case that the adversary cannot obtain the temporary tag state b.

Theorem 3. *The RFID protocol shown in Figure 2 achieves narrow-forward privacy (Definition 6) if the underlying encryption scheme is CCA-secure and* Corrupt *does not reveal the temporary tag state b.*

The full proof of Theorem 3 is given in the full version of this paper [31].

Proof (Theorem 3, Sketch). We construct a blinder $\mathcal{B}$ and show that $\mathbf{Adv}^{\text{prv}}_{\mathcal{A}_{\text{prv}}}$ is negligible for any narrow-forward adversary $\mathcal{A}_{\text{prv}}$, as required by Definition 6. More precisely, we prove that if there is an $\mathcal{A}_{\text{prv}}$ such that $\mathbf{Adv}^{\text{prv}}_{\mathcal{A}_{\text{prv}}}$ is non-negligible, then $\mathcal{A}_{\text{prv}}$ can be used as a black-box to construct an adversary $\mathcal{A}_{\text{cca}}$ that violates the CCA-security of the underlying encryption scheme. □

6.2 Impossibility of Narrow-Strong Privacy

We now point out a second, conceptually different weakness of the claimed narrow-strong private protocol of [12] (which is depicted in Figure 2). More precisely, we show an attack on this protocol that does not require the adversary to obtain temporary tag states. Moreover, we generalize this attack to prove our second impossibility result: in the PV-Model, it is *impossible* to achieve narrow-strong privacy along with reader authentication. This means that even in case the adversary cannot obtain temporary tag states, the most challenging privacy notion defined in [12] (narrow-strong privacy) remains unachievable.[4] We stress

[4] The impossibility of strong privacy has been shown in [11].

Alg. 3. Narrow-strong adversary $\mathcal{A}_{\text{prv}}$ on the protocol in Figure 2

1: $\mathsf{CreateTag}(\mathtt{ID})$
2: $vtag \leftarrow \mathsf{Draw}(\Pr[\mathtt{ID}] = 1)$
3: $(pk_{\mathcal{R}}, \mathtt{ID}, K) \leftarrow \mathsf{Corrupt}(vtag)$
4: $\mathsf{Free}(vtag)$
5: $\pi \leftarrow \mathsf{Launch}(\,)$
6: $a \leftarrow \mathsf{SendReader}(-, \pi)$
7: $b \in_R \{0,1\}^{\beta}$
8: $c \leftarrow \mathsf{Enc}_{pk_{\mathcal{R}}}(\mathtt{ID}, K, a, b)$
9: $d \leftarrow \mathsf{SendReader}(c, \pi)$
10: **if** $b = d$ **then return** 0
11: **else return** 1
12: **end if**

that introducing even stronger hardware assumptions to further restrict the ability of the adversary to corrupt tags would deviate from the capabilities of real tags. Indeed, most RFID tags are low-cost devices that usually are not equipped with mechanisms that ensure tamper-evidence or tamper-resistance.

Theorem 4. *The RFID protocol shown in Figure 2 does not achieve narrow-strong privacy (Definition 6), reader authentication (Definition 5) and correctness (Definition 2) at the same time.*

Note that the proof of reader authentication given in [12] (which we believe to be correct) requires the underlying public-key encryption scheme to be CCA-secure.

Proof (Theorem 4, Sketch). The idea is that a narrow-*strong* adversary $\mathcal{A}_{\text{prv}}$ can detect the presence of a blinder $\mathcal{B}$ by simulating a corrupt legitimate tag $\mathcal{T}_{\mathtt{ID}}$ to $\mathcal{R}$. In contrast to $\mathcal{R}$, $\mathcal{B}$ does not know the secret decryption key $sk_{\mathcal{R}}$ and thus will fail to generate a correct response d with overwhelming probability. This allows $\mathcal{A}_{\text{prv}}$ to distinguish between $\mathcal{B}$ and the real oracles. Note that in the following attack $\mathcal{A}_{\text{prv}}$ corrupts $\mathcal{T}_{\mathtt{ID}}$ *before* executing Ident with $\mathcal{R}$ and hence, the attack is *independent* of the temporary state of $\mathcal{T}_{\mathtt{ID}}$. $\mathcal{A}_{\text{prv}}$ is defined in Algorithm 3: First, $\mathcal{A}_{\text{prv}}$ creates a legitimate tag $\mathcal{T}_{\mathtt{ID}}$ (step 1), makes it accessible (step 2), corrupts it (step 3), and makes it inaccessible again (step 4). Then, $\mathcal{A}_{\text{prv}}$ makes $\mathcal{R}$ to start Ident with $\mathcal{T}_{\mathtt{ID}}$ (step 5) and obtains a from $\mathcal{R}$ (step 6). Now, $\mathcal{A}_{\text{prv}}$ simulates $\mathcal{T}_{\mathtt{ID}}$ to compute c (step 7–8) and sends c to $\mathcal{R}$ to obtain d (step 9). Next, $\mathcal{A}_{\text{prv}}$ performs the computation $\mathcal{T}_{\mathtt{ID}}$ would have done on receipt of d (step 10), i.e., $\mathcal{A}_{\text{prv}}$ checks if $b = d$. Finally, $\mathcal{A}_{\text{prv}}$ returns 0 to indicate that it interacted with the real oracles (step 10), or 1 to indicate the presence of $\mathcal{B}$ (step 11).

Since $\mathcal{T}_{\mathtt{ID}}$ is legitimate, due to correctness (Definition 2), $\mathsf{SendReader}$ must respond with $d = b$. The real $\mathsf{SendReader}$ oracle can compute b with probability 1 by decrypting c with $sk_{\mathcal{R}}$. Since, due to reader authentication (Definition 5), the public-key encryption scheme is assumed to be CCA-secure and $\mathcal{B}$ does not know $sk_{\mathcal{R}}$, $\mathcal{B}$ can at most guess b with negligible probability. Hence, $\mathcal{A}_{\text{prv}}$ has non-negligible advantage of distinguishing between $\mathcal{B}$ and the real oracles, which violates narrow-strong privacy (Definition 6). □

Now, we generalize this attack to our second impossibility result:

Theorem 5. *In the PV-Model there is no RFID system (Definition 1) that fulfills both reader authentication (Definition 5) and narrow-strong privacy (Definition 6).*

The full proof of Theorem 5 can be found in the full version of this paper [31].

Proof (Theorem 5, Sketch). We show that if there is a blinder $\mathcal{B}$ such that advantage $\mathbf{Adv}^{\text{prv}}_{\mathcal{A}_{\text{prv}}}$ is negligible for every narrow-strong adversary $\mathcal{A}_{\text{prv}}$ (as required by Definition 6), then we can use $\mathcal{B}$ as a black-box to construct an adversary $\mathcal{A}_{\text{sec}}$, who violates reader authentication (Definition 5). Note that $\mathcal{A}_{\text{prv}}$ can interact with tags after corrupting them since $\mathcal{A}_{\text{prv}}$ is a narrow-*strong* adversary. Hence, $\mathcal{A}_{\text{prv}}$ can corrupt a tag $\mathcal{T}$ after its creation and simulate it to $\mathcal{R}$ (that might be simulated by $\mathcal{B}$). This allows $\mathcal{A}_{\text{prv}}$ to verify the authentication messages of $\mathcal{R}$. Hence, $\mathcal{A}_{\text{prv}}$ can detect $\mathcal{B}$ since, due to reader authentication (Definition 5), $\mathcal{B}$ should not be able to successfully authenticate to $\mathcal{T}$ as $\mathcal{R}$. □

7 Conclusion

In this work we proved impossibility results that show that the RFID model proposed by Paise and Vaudenay [12] cannot guarantee the most interesting privacy notions and simultaneously reader authentication (which is the goal of the model). Nevertheless, we pointed out that, by restricting the tag corruption ability of the adversary, at least some, although weak, privacy notions can be achieved.

Acknowledgments. We thank Paolo D'Arco and Alessandra Scafuro for several useful discussions about RFID privacy notions. This work has been supported in part by the European Commission through the FP7 programme under contract 216646 ECRYPT II, 238811 UNIQUE, and 215270 FRONTS, in part by the Ateneo Italo-Tedesco under Program Vigoni and by the MIUR Project PRIN 2008 "PEPPER: Privacy E Protezione di dati PERsonali" (prot. 2008SY2PH4).

References

1. Atmel Corporation: Innovative IDIC solutions (2007), `http://www.atmel.com/dyn/resources/prod_documents/doc4602.pdf`
2. NXP Semiconductors: MIFARE smartcard ICs (September 2008), `http://www.mifare.net/products/smartcardics/`
3. Sadeghi, A.R., Visconti, I., Wachsmann, C.: User privacy in transport systems based on RFID e-tickets. In: International Workshop on Privacy in Location-Based Applications, PiLBA (2008)
4. I.C.A. Organization: Machine Readable Travel Documents, Doc 9303, Part 1 Machine Readable Passports, 5th Edn. (2003)

5. Weis, S.A., Sarma, S.E., Rivest, R.L., Engels, D.W.: Security and privacy aspects of low-cost radio frequency identification systems. In: Hutter, D., Müller, G., Stephan, W., Ullmann, M. (eds.) Security in Pervasive Computing. LNCS, vol. 2802, pp. 50–59. Springer, Heidelberg (2004)
6. Juels, A.: RFID security and privacy: A research survey. Journal of Selected Areas in Communication 24(2), 381–395 (2006)
7. Sadeghi, A.R., Visconti, I., Wachsmann, C.: Location privacy in RFID applications. In: Bettini, C., Jajodia, S., Samarati, P., Wang, X.S. (eds.) Privacy in Location-Based Applications. LNCS, vol. 5599, pp. 127–150. Springer, Heidelberg (2009)
8. Avoine, G.: Adversarial model for radio frequency identification. ePrint, Report 2005/049 (2005)
9. Juels, A., Weis, S.A.: Defining strong privacy for RFID. In: Proceedings of the Fifth IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOMW '07), pp. 342–347. ACM Press, New York (2007)
10. Burmester, M., van Le, T., de Medeiros, B.: Universally composable and forward-secure RFID authentication and authenticated key exchange. In: Proc. of ASI-ACCS, pp. 242–252. ACM Press, New York (2007)
11. Vaudenay, S.: On privacy models for RFID. In: Kurosawa, K. (ed.) ASIACRYPT 2007. LNCS, vol. 4833, pp. 68–87. Springer, Heidelberg (2007)
12. Paise, R.I., Vaudenay, S.: Mutual authentication in RFID: Security and privacy. In: Proc. of ASIACCS, pp. 292–299. ACM Press, New York (2008)
13. Deng, R.H., Li, Y., Yao, A.C., Yung, M., Zhao, Y.: A new framework for RFID privacy. ePrint, Report 2010/059 (2010)
14. Mangard, S., Oswald, E., Popp, T.: Power Analysis Attacks Revealing the Secrets of Smart Cards. Springer, Heidelberg (2007)
15. Hutter, M., Schmidt, J.M., Plos, T.: RFID and its vulnerability to faults. In: Oswald, E., Rohatgi, P. (eds.) CHES 2008. LNCS, vol. 5154, pp. 363–379. Springer, Heidelberg (2008)
16. Kasper, T., Oswald, D., Paar, C.: New methods for cost-effective side-channel attacks on cryptographic RFIDs. In: Workshop on RFID Security, RFIDSec (2009)
17. D'Arco, P., Scafuro, A., Visconti, I.: Semi-destructive privacy in DoS-enabled RFID systems. In: Workshop on RFID Security, RFIDSec (2009)
18. D'Arco, P., Scafuro, A., Visconti, I.: Revisiting DoS Attacks and Privacy in RFI-DEnabled Networks. In: Dolev, S. (ed.) ALGOSENSORS 2009. LNCS, vol. 5804, pp. 76–87. Springer, Heidelberg (2009)
19. Ng, C.Y., Susilo, W., Mu, Y., Safavi-Naini, R.: RFID privacy models revisited. In: Jajodia, S., Lopez, J. (eds.) ESORICS 2008. LNCS, vol. 5283, pp. 251–266. Springer, Heidelberg (2008)
20. Ng, C.Y., Susilo, W., Mu, Y., Safavi-Naini, R.: New privacy results on synchronized RFID authentication protocols against tag tracing. In: Backes, M., Ning, P. (eds.) ESORICS 2009. LNCS, vol. 5789, pp. 321–336. Springer, Heidelberg (2009)
21. Bringer, J., Chabanne, H., Icart, T.: Efficient zero-knowledge identification schemes which respect privacy. In: Proceedings of ASIACCS '09, pp. 195–205. ACM Press, New York (2009)
22. Sadeghi, A.R., Visconti, I., Wachsmann, C.: Efficient RFID security and privacy with anonymizers. In: Workshop on RFID Security, RFIDSec (2009)
23. Sadeghi, A.R., Visconti, I., Wachsmann, C.: Anonymizer-enabled security and privacy for RFID. In: Miyaji, A., Echizen, I., Okamoto, T. (eds.) CANS 2009. LNCS, vol. 5888, pp. 134–153. Springer, Heidelberg (2009)

24. Goldreich, O., Micali, S., Wigderson, A.: How to Play any Mental Game—A Completeness Theorem for Protocols with Honest Majority. In: Proc. of ACMSTOC, pp. 218–229 (1987)
25. Goldwasser, S., Micali, S., Rackoff, C.: The Knowledge Complexity of Interactive Proof-Systems. SIAM J. on Computing 18(6), 186–208 (1989)
26. Sadeghi, A.R., Visconti, I., Wachsmann, C.: PUF-enhanced RFID security and privacy. In: Workshop on Secure Component and System Identification (SECSI) (2010)
27. Kirschenbaum, I., Wool, A.: How to build a low-cost, extended-range RFID skimmer. ePrint, Report 2006/054 (2006)
28. Avoine, G., Lauradoux, C., Martin, T.: When compromised readers meet RFID. In: Workshop on RFID Security (RFIDSec) (2009)
29. Garcia, F.D., van Rossum, P.: Modeling privacy for off-line RFID systems. In: Workshop on RFID Security (RFIDSec) (2009)
30. Nithyanand, R., Tsudik, G., Uzun, E.: Readers behaving badly: Reader revocation in PKI-based RFID systems. Cryptology ePrint Archive, Report 2009/465 (2009)
31. Sadeghi, A.R., Visconti, I., Wachsmann, C.: On rfid privacy with mutual authentication and tag corruption — Extended Version. ePrint (2010)