Computational Soundness of Non-Malleable Commitments

David Galindo\textsuperscript{1*}, Flavio D. Garcia\textsuperscript{2}, and Peter van Rossum\textsuperscript{2}

\textsuperscript{1} Departament of Computer Science, University of Malaga, Spain
dgalindo@lcc.uma.es
\textsuperscript{2} Institute for Computing and Information Sciences, Radboud University Nijmegen, The Netherlands.
{flaviog,petervr}@cs.ru.nl

Abstract. This paper aims to find a proper security notion for commitment schemes to give a sound computational interpretation of symbolic commitments. We introduce an indistinguishability based security definition of commitment schemes that is equivalent to non-malleability with respect to commitment. Then, we give a construction using tag-based encryption and one-time signatures that is provably secure assuming the existence of trapdoor permutations. Finally, we apply this new machinery to give a sound interpretation of symbolic commitments in the Dolev-Yao model while considering active adversaries.

1 Introduction

Over the last few decades, two mainstream approaches have been developed for the analysis of security protocols. On the one hand, the cryptographic approach considers an arbitrary computationally-bound adversary that interacts with honest participants and tries to break a security goal. This model is satisfactory as it deals with every efficient attacker. On the other hand, the symbolic or Dolev-Yao approach idealizes the security properties of the cryptographic primitives, which are axiomatized in a logic. Moreover, the capabilities of the adversary are also specified by a set of inference rules. This approach is appealing because there are automated techniques for the verification of some security properties.

Abadi and Rogaway in \cite{AR02} pioneered the idea of relating these two models and showed that, under appropriate assumptions on the underlying cryptographic primitives, a simple language of encrypted expressions is sound with respect to the computational model in the case of passive adversaries.

Such a relation maps symbolic messages $m$ to distributions over bitstrings $[m]$. This map then should relate messages that are observationally equivalent in the symbolic world to indistinguishable distributions over bitstrings. Such a map allows one to use formal methods, possibly even automated, to reason about security properties of protocols and have those reasonings be valid also in the standard computational model.

\* partially funded by the Spanish Ministry of Science and Education through the projects ARES (CSD2007-00004) and CRISIS (TIN2006-09242)
Several extensions to the original Abadi-Rogaway logic [AR02] have been proposed in the literature. These extensions deal with public key encryption [MW04, Her05], key cycles [ABHS05], partial information leakage [ABS05], active instead of passive adversaries [MW04, JLM05], and more realistic security notions [AW05]. Other extensions add new primitives to the logic such as bilinear pairings [Maz07], modular exponentiation [BLMW07] and hash functions [CKKW06, GvR06]. There are also frameworks dealing with generic equational theories [BCK05, ABW06, KM07]. So far there is no work in the literature, that we are aware of, that relates these two approaches for commitment schemes.

Commitment schemes are fundamental cryptographic primitives and are used in protocols like zero-knowledge proofs [GMW91], contract signing [EGL85], and can be used for bidding protocols. A commitment consists of two phases: the commitment phase where the principals commit to a message without revealing any information; and the opening phase where the principals reveal the message and it is possible to verify that this message corresponds to the value committed to during the commitment phase. After the commitment phase it should be infeasible to open the commitment to a different value than the one committed. This property is called binding. In the context of bidding protocols, non-malleability is also a desirable property. This means that an adversary cannot modify an intercepted commitment, say into a commitment to a slightly higher bid.

Our contribution. The first objective of this paper is to find sufficient security assumptions to give a sound computational interpretation of commitments schemes in the Dolev-Yao model, under active adversaries. Pursuing that objective we propose a new indistinguishability-based security definition for commitment schemes in the presence of adaptive adversaries. Then we give a novel generic construction for a non-malleable commitment scheme based on one-way trapdoor permutations. This construction is secure with respect to our new definition and has some additional properties such as being non-interactive, perfectly binding and reusable, which makes it of independent interest. This new definition allows us to prove soundness of the Dolev-Yao model extended with commitments, following the directions of Micciancio and Warinschi [MW04].

Overview. Section 3 introduces basic notation and definitions from the literature. Section 4 elaborates on different definitions of non-malleability for commitment schemes and discusses the relations among them. In Section 5 we propose a new commitment scheme and we give a security proof. Section 2 describes symbolic protocol executions, its computational counterparts and the map between them and also states the soundness result. Finally in Section 7 there are some concluding remarks.

2 Symbolic Protocols

We are going to apply this theory to give sound computational interpretation to symbolic commitments. Recall from the introduction that the symbolic approach to protocol verification deals with symbolic or algebraic messages and idealized cryptographic primitives. In this setting the adversary is unbounded in running
time and has full control over the communication media but is completely incapable
of breaking the underlying cryptographic primitives.

We now describe the message space and the closure operator. These messages
are used to formally describe cryptographic protocols. The closure represents
the knowledge that can be extracted from a message, and it is used to define what
valid algebraic protocol runs are. Intuitively a protocol run is valid if every message sent
by a principal can be deduced from its knowledge except maybe for some fresh
randomness. Much of this is standard (see, e.g., [AR02, MW04, MP05, GvR06]),
except that we model commitments and decommitments as well as encryption.

**Definition 2.1.** Let **Nonce** be an infinite set of nonce symbols, **Const** a finite set
of constant symbols, **Key** an infinite set of key symbols, and **Random** an infinite set
of randomness labels. Nonces are denoted by $n, n', \ldots$, constants by $c, c', \ldots$, keys
by $k, k', \ldots$, and randomness labels by $r, r', \ldots$. Using these building blocks,
messages are constructed using symbolic encryption, commitments, decommitments,
and pairing operations:

$$
\text{Msg} \ni m := c | n | \{m\}^r_k | \text{com}^r(m) | \text{dec}^r(m) | \langle m, m \rangle.
$$

A message of the form $\{m\}^r_k$ is called an encryption and the set of all such messages
is denoted by \text{Enc}. Similarly, messages of the form $\text{com}^r(m)$ are called commitments
and the set of all these messages is denoted by \text{Com}. The messages of the form
$\text{dec}^r(m)$ are called decommitments and the set of all these messages is denoted
by \text{Dec}. In a protocol run $\text{dec}^r(m)$ is a valid decommitment of $\text{com}^r(m')$ only if
$m = m'$ and $r = r'$. We say that elements in $\text{Const} \cup \text{Nonce} \cup \text{Key}$ are primitive
and we denote this set by \text{Prim}. For a public key $k$ we denote its associated private
key as $k^{-1}$.

The closure of a set $U$ of messages is the set of all messages that can be con-
structed from $U$ using tupling, detupling, commitment, decommitment, and e-
cryption and decryption. It represents the information an adversary could deduce
knowing $U$. Note that, due to secrecy of the commitment scheme, knowing $\text{com}^r(m)$
does not provide an adversary with any information about $m$.

**Definition 2.2** (Closure). Let $U$ be a set of messages. The closure of $U$, denoted
by $\overline{U}$, is the smallest set of messages satisfying: 1. $\text{Const} \subseteq \overline{U}$; 2. $U \subseteq \overline{U}$;
3. $m, m' \in \overline{U} \implies \langle m, m' \rangle \in \overline{U}$; 4. $m \in \overline{U} \land k \in \overline{U} \implies \{m\}^r_k \in \overline{U}$;
5. $\{m\}^r_k \in \overline{U} \land k^{-1} \in \overline{U} \implies m \in \overline{U}$; 6. $m \in \overline{U} \implies \text{com}^r(m), \text{dec}^r(m) \in \overline{U}$; 7. $\text{dec}^r(m) \in \overline{U} \implies m \in \overline{U}$; 8. $\langle m, m' \rangle \in \overline{U} \implies m, m' \in \overline{U}$.

Next we need to find the right security notions to give sound computational
interpretation to symbolic encryption and commitments.

3 Computational Setup

This section introduces syntax and security definitions for different cryptographic
primitives. Much of this is standard, we refer the reader to [GM84, RS92] and [NY90]
for a thorough explanation. Some of this primitives will be used to interpret alge-
braic operations and some of them are used as building blocks for our construction
of Section 5.
3.1 Commitment Schemes

Definition 3.1. A commitment scheme is a triple \( \Omega = (\text{TTP, Snd, Rcv}) \) of probabilistic polynomial-time algorithms. \( \text{TTP} \), the trusted third party, takes as input the security parameter \( 1^n \) and produces a common reference string \( \sigma \). We require that \( |\sigma| \geq p(\eta) \) for some non-constant polynomial \( p \). \( \text{Snd} \), the sender, takes as input \( \sigma \) and a message \( m \) and produces a commitment \( \text{com} \) to this message and a corresponding decommitment \( \text{dec} \). \( \text{Rcv} \), the receiver, takes as input \( \sigma \), \( \text{com} \), and \( \text{dec} \) and produces a message or \( \perp \).

<table>
<thead>
<tr>
<th>Meaningfulness( _\Omega )(A):</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma \leftarrow \text{TTP}(1^n) )</td>
</tr>
<tr>
<td>( m \leftarrow A(\sigma) )</td>
</tr>
<tr>
<td>( (\text{com, dec}) \leftarrow \text{Snd}(\sigma, m) )</td>
</tr>
<tr>
<td>( m_1 \leftarrow \text{Rcv}(\sigma, \text{com, dec}) )</td>
</tr>
<tr>
<td>return ( m \neq m_1 )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Secrecy( _{\text{TTP.Snd}}(A_1, A_2): )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma \leftarrow \text{TTP}(1^n) )</td>
</tr>
<tr>
<td>( m_0, m_1, s \leftarrow A_1(\sigma) )</td>
</tr>
<tr>
<td>( b \leftarrow {0,1} )</td>
</tr>
<tr>
<td>( (\text{com, dec}) \leftarrow \text{Snd}(\sigma, m_0) )</td>
</tr>
<tr>
<td>( b' \leftarrow A_2(s, \text{com}) )</td>
</tr>
<tr>
<td>return ( b = b' )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Binding( _{\text{TTP.Rcv}}(A): )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sigma \leftarrow \text{TTP}(1^n) )</td>
</tr>
<tr>
<td>( (\text{com, dec}, \text{dec}_2) \leftarrow A(\sigma) )</td>
</tr>
<tr>
<td>( m_1 \leftarrow \text{Rcv}(\sigma, \text{com, dec}_1) )</td>
</tr>
<tr>
<td>( m_2 \leftarrow \text{Rcv}(\sigma, \text{com, dec}_2) )</td>
</tr>
<tr>
<td>return ( m_1 \neq \perp \neq m_2 ) ( \land m_1 \neq m_2 )</td>
</tr>
</tbody>
</table>

The following three conditions must hold.

1. For all probabilistic polynomial-time algorithms \( A \), the probability \( \Pr[\text{Meaningfulness}_\Omega(A)] \) is a negligible function of \( \eta \).
2. For all probabilistic polynomial-time algorithms \( (A_1, A_2) \), the advantage \( |\Pr[\text{Secrecy}_{\text{TTP.Snd}}(A_1, A_2)] - 1/2| \) is a negligible function of \( \eta \).
3. For all probabilistic polynomial-time algorithms \( A \), the probability \( \Pr[\text{Binding}_{\text{TTP.Rcv}}(A)] \) is a negligible function of \( \eta \).

Definition 3.2. A commitment scheme is said to be perfectly binding if for all unbounded algorithms \( A \), the probability \( \Pr[\text{Binding}_{\text{TTP.Rcv}}(A)] \) is zero.

Definition 3.3. A commitment scheme is said to be perfectly hiding if for all unbounded algorithms \( (A_0, A_1) \), \( |\Pr[\text{Secrecy}_{\text{TTP.Snd}}(A_1, A_2)] - 1/2| \) is zero.

3.2 Encryption Schemes

Definition 3.4. An encryption scheme is a triple \( \Pi = (\mathcal{K}, \mathcal{E}, \mathcal{D}) \) of probabilistic polynomial-time algorithms. \( \mathcal{K} \) takes as input the security parameter \( 1^n \) and produces a key pair \((pk, sk)\) where \( pk \) is the public encryption key and \( sk \) is the private decryption key. \( \mathcal{E} \) takes as input a public key \( pk \) and a plaintext \( m \) and outputs a ciphertext. \( \mathcal{D} \) takes as input a private key \( sk \) and a ciphertext and outputs a plaintext or \( \perp \). It is required that \( \Pr[(pk, sk) \leftarrow \mathcal{K}(1^n); c \leftarrow \mathcal{E}(pk, m); m' \leftarrow \mathcal{D}(sk, c) : m = m'] = 1 \).

<table>
<thead>
<tr>
<th>IND-CCA ( _\Pi )(( A_0, A_1 )):</th>
</tr>
</thead>
<tbody>
<tr>
<td>( (pk, sk) \leftarrow \mathcal{K}(1^n) )</td>
</tr>
<tr>
<td>( m_0, m_1, s \leftarrow A_0^D(pk) )</td>
</tr>
<tr>
<td>( b \leftarrow {0,1} )</td>
</tr>
<tr>
<td>( c \leftarrow \mathcal{E}(pk, m_0) )</td>
</tr>
<tr>
<td>( b' \leftarrow A_1^E(s, c) )</td>
</tr>
<tr>
<td>return ( b = b' )</td>
</tr>
</tbody>
</table>
**Definition 3.5.** An encryption scheme $H = (K, E, D)$ is said to be IND-CCA secure if for all probabilistic polynomial-time adversaries $A = (A_0, A_1)$ the advantage of $A$, defined as $|P[\text{IND-CCA}_H(A_0, A_1)] - 1/2|$, is a negligible function of $\eta$. This adversary has access to a decryption oracle $D$ that on input $c'$ outputs $D(sk, c')$ with the only restriction that $c \neq c'$.

### 3.3 One-time Signatures

**Definition 3.6.** A signature scheme is a triple $(Gen, Sign, Vrfy)$ of probabilistic polynomial-time algorithms. $Gen$ takes as input the security parameter $1^\eta$ and produces a key pair $(vk, sk)$ where $vk$ is the signature verification key and $sk$ is the secret signing key. $Sign$ takes as input $sk$ and a message $m$ and produces a signature $s$ of $m$. $Vrfy$ takes as input $vk$, a message $m$ and a signature $s$ and outputs whether or not $s$ is a valid signature of $m$.

\[
\begin{align*}
\text{OTS}_\Sigma(A_0, A_1): & \\
(vk, sk) & \leftarrow Gen(1^\eta) \\
m, s & \leftarrow A_0(vk, 1^\eta) \\
\sigma & \leftarrow Sign(sk, m) \\
m', \sigma' & \leftarrow A_1(s, \sigma) \\
\text{return} & \sigma \neq \sigma' \land \text{Vrfy}(vk, (m', \sigma'))
\end{align*}
\]

**Definition 3.7.** A signature scheme $\Sigma = (Gen, Sign, Vrfy)$ is a strong, one-time signature scheme if the success probability of any probabilistic polynomial-time adversary $(A_0, A_1)$ in the game $\text{OTS}_\Sigma(A_0, A_1)$ is negligible in the security parameter $\eta$.

### 3.4 Tag-based Encryption

**Definition 3.8.** A tag-based encryption scheme (TBE) handling tags of length $\ell$ (where $\ell$ is a polynomially-bounded function) is a triple of probabilistic polynomial-time algorithms $(KeyGen, Enc, Dec)$. $KeyGen$ takes a security parameter $1^\eta$ and returns a public key $pk$ and secret key $sk$. The public key $pk$ includes the security parameter $1^\eta$ and $\ell(\eta)$; as well as the description of sets $\mathcal{M}, R, C$, which denote the set of messages, randomness and ciphertexts respectively. These descriptions might depend on the public key $pk$. $Enc$ takes as inputs $pk$, a tag $t \in \{0, 1\}^{\ell}$ and $m \in \mathcal{M}$. It returns a ciphertext $c \in \mathcal{C}$. $Dec$ takes as inputs the secret key $sk$, a tag $t$ and $c \in \mathcal{C}$, and returns $m \in \mathcal{M}$ or $\bot$ when $c$ is not a legitimate ciphertext. For the sake of consistency, these algorithms must satisfy $Dec(sk, t, c) = m$ for all $t \in \{0, 1\}^{\ell}$, $m \in \mathcal{M}$, where $c = Enc(pk, t, m)$.

**Definition 3.9.** Let $E = (KeyGen, Enc, Dec)$ be a TBE scheme. We say $E$ is IND-TBE-CCA secure if for any 3-tuple of PPT oracle algorithms $(A_0, A_1, A_2)$ and any polynomially-bounded function $\ell$ the advantage in the following game is negligible in the security parameter $1^\eta$:

$A_0(1^\eta, \ell(\eta))$ outputs a target tag $t$. $KeyGen(1^\eta)$ outputs $(pk, sk)$ and the adversary is given $pk$. Then the adversary $A_1$ may ask polynomially-many queries to a decryption oracle $D(t', c') = Dec(sk, t', c')$ for pairs tag-ciphertext $(t', c')$ of its choice, with the restriction $t \neq t'$. At some point, $A_1$ outputs two equal length
messages \(m_0, m_1\). A bit \(b \leftarrow \{0, 1\}\) is chosen at random and the adversary is given a challenge ciphertext \(c \leftarrow \text{Enc}(pk, t, m_b)\). \(A_2\) may continue asking the decryption oracle for pairs tag-ciphertext \((t', c')\) of its choice, with the restriction \(t \neq t'\). Finally, \(A_2\) outputs a guess \(b'\).

**IND-TBE-CCA**

\[
\begin{align*}
\text{IND-TBE-CCA}_\varepsilon(A_0, A_1, A_2): \\
& t, s_1 \leftarrow A_0(1^n, \ell(\eta)) \\
& (pk, sk) \leftarrow \text{KeyGen}(1^n) \\
& m_0, m_1, s_2 \leftarrow A_1^D(s_1, pk) \\
& b \leftarrow \{0, 1\} \\
& c \leftarrow \text{Enc}(pk, t, m_b) \\
& b' \leftarrow A_2^D(s_2, c) \\
& \text{return } b = b'
\end{align*}
\]

We define the advantage of \(A\) as \(|\mathbb{P}[\text{IND-TBE-CCA}(A)] - 1/2|\).

### 3.5 Interpretation

Suppose we have an encryption scheme \(\Pi\), a commitment scheme \(\Omega\) and a function that maps symbolic constants to constant bitstrings. Then we can define a mapping \([\cdot]\) from algebraic messages \(m \in \text{Msg}\) to distributions over bitstrings \([m] \in \text{Str}\).

This interpretation maps nonces to random bitstrings of length \(\eta\); encryptions are interpreted by running the encryption algorithm \(E\) and for interpreting commitments and decommitments we use the commit algorithm \(Snd\).

In order to achieve sound interpretation we will explore the security requirements on these cryptographic primitives. For the case of encryption it is satisfactory to use any IND-CCA encryption scheme as shown in [MW04]. For the case of commitments, using standard security definitions is not straightforward as they are not strong enough nor indistinguishability based. To achieve sound interpretation of the idealized Dolev-Yao model, throughout the next section we elaborate on a convenient security definition for commitment schemes.

### 4 Definitions of Non-malleability

As noticed by Fischlin and Fischlin [FF00], there are two different versions of non-malleability for commitment schemes, namely: NM with respect to opening (NMO) and NM with respect to commitment (NMC). NMC was the version originally proposed by Dolev, Dwork and Naor in [DDN91]. It means that given a commitment to a message \(m\), the adversary is unable to build a different commitment to \(m'\), with \(m\) related to \(m'\). This version of non-malleability is appropriate while considering perfectly binding commitments and only makes sense for schemes that are not perfectly hiding.

The other version NMO, seemingly weaker, means that an adversary that is first given a commitment to \(m\) and on a second stage its decommitment, is unable to find a different commitment-decommitment pair that decommits to a message \(m'\) related to \(m\). This notion was studied by Di Crescenzo, Ishai and Ostrovsky [CIO98] and later by Di Crescenzo, Katz, Ostrovsky and Smith [CKOS01]. Intuitively a commitment scheme is non-malleable if the adversary can do no better than a
simulator which has no information at all about the message that was committed to. Next we recall their definition.

\[
\text{NMO}_Ω(A_1, A_2, D, R):
\begin{align*}
σ &\leftarrow \text{TTP}(1^n) \\
m_1 &\leftarrow D \\
\text{com}_1, \text{dec}_1 &\leftarrow \text{Snd}(σ, m_1) \\
\text{com}_2 &\leftarrow A_1(σ, \text{com}_1) \\
\text{dec}_2 &\leftarrow A_2(σ, \text{com}_1, \text{com}_2, \text{dec}_1) \\
m_2 &\leftarrow \text{Rcv}(σ, \text{com}_2, \text{dec}_2) \\
\text{return } \text{com}_1 \neq \text{com}_2 \land R(m_1, m_2)
\end{align*}
\]

**Definition 4.1 (Non-malleability [CIO98, CKOS01])**. Let \(Ω = (\text{TTP}, \text{Snd}, \text{Rcv})\) be a commitment scheme. \(Ω\) is called non-malleable if for all PPT adversaries \((A_1, A_2)\) there is a PPT simulator \(S\) such that for all distributions \(D\) and all relations \(R\),

\[
\mathbb{P}[\text{NMO}_Ω(A_1, A_2, D, R)] - \mathbb{P}[\text{SIM}(S, D, R)]
\]

is a negligible function of \(η\).

**Remark 4.2.** To prevent that the adversary trivially wins, by refusing to decommit, the following restriction over the relation \(R\) is imposed: for all messages \(m\), we have \(R(m, ⊥) = 0\).

### 4.1 NMC-CCA: Non-malleability Against Chosen Commitment Attacks

The previous definition deals with non-malleability with respect to opening. For the relation between symbolic and computational cryptography we need the stronger notion of non-malleability with respect to commitment. Intuitively, this is because in the algebraic setting \(\text{com}'(m')\) cannot be deduced from \(\text{com}'(m)\), with \(m'\) somehow related to \(m\). Therefore we adapt the NMO definition to non-malleability with respect to commitment and we strengthen it by incorporating active adaptive security, allowing the adversary to mount chosen commitment attacks (CCA in short). Specifically, we empower the adversary with access to a decommitment oracle \(D\). To do so, from now on, we restrict our attention to non-interactive, perfectly binding trapdoor commitment schemes. The oracle \(D\) has access to the trapdoor information. It takes as argument a commitment \(c\) with the restriction that \(c\) is not equal to the challenge commitment \(\text{com}_1\). Then if the commitment \(c\) has been correctly generated, the oracle returns a decommitment \(d\) which opens \(c\), and otherwise it outputs ⊥.

\[
\text{NMC-CCA}_Ω(A_0, A_1, R):
\begin{align*}
σ &\leftarrow \text{TTP}(1^n) \\
D, s_1 &\leftarrow A_0^D(σ) \\
m_1 &\leftarrow D(σ) \\
\text{com}_1, \text{dec}_1 &\leftarrow \text{Snd}(σ, m_1) \\
\text{com}_2, s_r &\leftarrow A_1^D(s_1, \text{com}_1) \\
\text{dec}_2 &\leftarrow D(\text{com}_2) \\
m_2 &\leftarrow \text{Rcv}(σ, \text{com}_2, \text{dec}_2) \\
\text{return } \text{com}_1 \neq \text{com}_2 \land R(s_r, m_1, m_2)
\end{align*}
\]

\[
\text{SIM-CCA}_Ω(S_0, S_1, R):
\begin{align*}
σ &\leftarrow \text{TTP}(1^n) \\
D, s_1 &\leftarrow S_0(σ) \\
m_1 &\leftarrow D(σ) \\
\text{com}_2, s_r &\leftarrow S_1(s_1) \\
\text{dec}_2 &\leftarrow D(\text{com}_2) \\
m_2 &\leftarrow \text{Rcv}(σ, \text{com}_2, \text{dec}_2) \\
\text{return } R(s_r, m_1, m_2)
\end{align*}
\]
**Definition 4.3 (NMC-CCA).** Let $\Omega = (TTP, Snd, Rcv)$ be a commitment scheme. $\Omega$ is called **NMC-CCA secure** if for all PPT adversaries $(A_0, A_1)$ there is a PPT simulator $(S_0, S_1)$ such that for all relations $R$ (with the same restriction as in 4.2),

$$P[\text{NMC-CCA}_\Omega(A_0, A_1, R)] - P[\text{SIM-CCA}_{TTP}(S_0, S_1, R)]$$

is a negligible function of $\eta$.

### 4.2 An Indistinguishability Based Definition

Next we introduce an equivalent formulation of NMC-CCA that is more convenient to prove soundness of the Dolev-Yao model with respect to commitment schemes.

**Definition 4.4 (IND-COM-CCA).** Let $\Omega = (TTP, Snd, Rcv)$ be a commitment scheme. $\Omega$ is said to be **IND-COM-CCA secure** if for all PPT adversaries $(A_0, A_1)$

$$P[\text{IND-COM-CCA}_1(A_0, A_1)] = 1 - P[\text{IND-COM-CCA}_0(A_0, A_1)]$$

is a negligible function of $\eta$.

Next we show that NMC-CCA and IND-COM-CCA are equivalent. We discuss it briefly as it is basically the proof that NM-CCA and IND-CCA are equivalent, adapted to commitment schemes.

**Theorem 4.5.** Let $\Omega = (TTP, Snd, Rcv)$ be a commitment scheme. Then $\Omega$ is **IND-COM-CCA secure** if and only if $\Omega$ is **NMC-CCA secure**.

**Proof.** (IND-COM-CCA $\Leftarrow$ NMC-CCA) Let $(B_0, B_1)$ be an adversary for IND-COM-CCA. Then we build the following adversary $(A_0, A_1)$ against NMC-CCA.

**Algorithm** $A_0^D(\sigma)$:

$m_0, m_1, s_1 \leftarrow A_0^P(\sigma)$

$\text{return } m_0, m_1$

**Algorithm** $A_1^D((\sigma, m_0, m_1), c_1)$:

$b \leftarrow B_1^D(s_1, c_1)$

$c_2 \leftarrow Snd(\sigma, m_b)$

$\text{return } c_2, \epsilon$

where $U$ is the uniform distribution. Now take the relation $R(s_r, m_1, m_2)$ as $m_1$ equal to $m_2$. It should be clear, after unfolding $(A_0, A_1)$ in the **NMC-CCA** game, that this adversary has the same advantage that $(B_0, B_1)$ has against IND-COM-CCA.

(IND-COM-CCA $\Rightarrow$ NMC-CCA) Let $(A_0, A_1)$ be an adversary for NMC-CCA. Then we build the following adversary $(B_0, B_1)$ against IND-COM-CCA.

**Algorithm** $B_0^D(\sigma)$:

$m_0, m_1 \leftarrow D(\sigma)$

$\text{return } m_0, m_1$

**Algorithm** $B_1^D((\sigma, m_0, m_1, s_1), c_1)$:

$c_2 \leftarrow A_1^P(s_1, c_1)$

$m \leftarrow D(\epsilon)$

if $m = m_1$ then return 1
else return 0
Again, just by unfolding these adversaries in the IND-COM-CCA game, it is easy to verify that they have the same advantage that \((A_0, A_1)\) has against NMC-CCA. □

It remains to show that such a security notion for a commitment scheme is achievable. In the next section we give a practical construction that achieves IND-COM-CCA security.

5 The Construction

We now propose a new construction for IND-COM-CCA that is computationally hiding, perfectly binding, reusable, non-interactive, non-malleable under adaptive adversaries, and provably secure under the assumption that trapdoor permutations exist.

Next we outline the idea of our construction. As pointed out by Di Crescenzo, Katz, Ostrovsky and Smith [CKOS01], an IND-CCA secure public key encryption scheme can be converted into a perfectly binding non-malleable commitment scheme. Let \(\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})\) be an indistinguishable against adaptive chosen-ciphertext attacks secure public key encryption scheme. The idea is to commit to a message \(m\) by encrypting it using random coins \(r\); commitment is set to be the ciphertext \(c = \text{Enc}(pk, m; r)\); de-commitment is set to be the pair \((m, r)\); finally the opening algorithm takes \((c, m, r)\) and checks whether \(c = \text{Enc}(pk, m; r)\).

When trying to directly use this construction to instantiate an IND-COM-CCA commitment scheme one might not be able to simulate the de-commitment oracle. The reason is that given a ciphertext/commitment \(c\), one recovers the purported embedded message \(m\) by using the decryption algorithm, but not necessarily the randomness \(r\). One way to break through this situation is to include in the commitment a second ciphertext \(c' = \text{Enc}(pk', r; r')\) encrypting the randomness \(r\) used in the first ciphertext \(c = \text{Enc}(pk, m; r)\). This is the key idea of our construction. We additionally use one-time signatures and this together with tag-based encryption schemes ensure the de-commitment oracle does not leak vital information.

Let \(\Pi = (\text{KeyGen}, \text{Enc}, \text{Dec})\) be a tag based encryption scheme and let \(\Sigma = (\text{Gen}, \text{Sign}, \text{Vrfy})\) be a signature scheme. Define \((\text{TTP}, \text{Snd}, \text{Rcv})\) as follows:

- \(\text{TTP}\) runs \(\text{KeyGen}(1^n)\) twice to obtain \((pk_1, sk_1)\) and \((pk_2, sk_2)\). The common reference string includes \(pk_1, pk_2\).
- To commit to a message \(m\), the sender \(\text{Snd}\) computes and outputs the commitment \(C = (vk, c_1, c_2, s)\) where \(c_1 = \text{Enc}(pk_1, vk, m; r_1)\), \(c_2 = \text{Enc}(pk_2, vk, r_1; r_2)\), with \(r_1, r_2 \leftarrow \mathcal{R}\), \((vk, sk) \leftarrow \text{Gen}(1^n)\) and \(s \leftarrow \text{Sign}(sk, (c_1, c_2))\). The decommitment is set to be \((m, r_1)\).
- To de-commit ciphertext \(C = (vk, c_1, c_2, s)\) using \((m, r_1)\), the receiver \(\text{Rcv}\) first checks if the signature on \((c_1, c_2)\) is correct, and afterwards whether or not \(c_1 = \text{Enc}(pk_1, vk, m; r_1)\).

We assume \(\mathcal{R} = \mathcal{M}\).

**Theorem 5.1.** Assume that \((\text{KeyGen}, \text{Enc}, \text{Dec})\) is an IND-TBE-CCA secure tag based encryption scheme and that \((\text{Gen}, \text{Sign}, \text{Vrfy})\) is a one-time strongly unforgeable signature scheme. Then \((\text{TTP}, \text{Snd}, \text{Rcv})\) is an IND-COM-CCA secure commitment scheme.
Proof. We transform an adversary $A$ against the IND-COM-CCA security of the commitment scheme into adversaries against the TBE and the OTS. Next we will describe a sequence of games following the methodology advocated in [Sho04, BR06]. Let $X_i$ be the event that $A$ learns the challenge bit $b$ in the $i$-th game.

**Game 0.** This is the unmodified IND-COM-CCA game. Trivially, $|\mathbb{P}[X_0] - 1/2|$ equals the advantage of $A$ against IND-COM-CCA.

**Game 1.** In this game we disallow decryption queries $C = (vk, c_1, c_2, s)$ s.t. $vk = vk'$ where $(vk', c_1', c_2', s')$ is the challenge commitment. Then, we get that $|\mathbb{P}[X_1] - \mathbb{P}[X_0]|$ is less or equal than the advantage any PPT algorithm has in breaking the one-time strong unforgeability security of the OTS.

**Game 2.** Still decryption queries with $vk = vk'$ are forbidden. In this game we use the IND-CCA security of the second instance of the TBE scheme. The components $c_1'$ and $c_2'$ of the challenge ciphertext are changed to $c_1' = \text{Enc}(pk_1', vk', m_1'; r_1)$, and $c_2' = \text{Enc}(pk_2', vk', r')$ where $r', r_1 \leftarrow \mathcal{R}$. Now, we have $|\mathbb{P}[X_2] - \mathbb{P}[X_1]|$ is less or equal than the advantage any PPT algorithm has in breaking the selective IND-CCA security of the TBE.

Finally it is shown that $|\mathbb{P}[X_2] - 1/2|$ is bounded by the advantage any PPT algorithm has in breaking the selective IND-CCA security of the first instance of the TBE.

Putting everything together, we get that $|\mathbb{P}[X_0] - 1/2|$ is bounded by the advantages in breaking the OTS scheme plus twice the advantage in breaking the selective IND-CCA of the TBE scheme. Next we describe the concrete adversaries.

**Game 0 ≈ Game 1.** Assume that there is an adversary $(A_0, A_1)$ that is able to distinguish the environments of Game 0 and 1. Then we build an adversary $(B_0, B_1)$ against the one-time strong unforgeability of the signature scheme.

![Algorithm B_0(\eta, vk):](image)

![Algorithm B_1(s_1|vk|c_1|c_2|sk_1|sk_2, s) = |b'| ≈ A_1^P(s_1, (vk, c_1, c_2, s))]. Calls to the de-commitment oracle $D(vk', c_1', c_2', s')$ are simulated by firstly verifying the signature $\text{Vrfy}(vk', (c_1', c_2'), s')$. If the verification succeeds then the oracle returns the pair $(\text{Dec}(sk_1, vk', c_1'), \text{Dec}(sk_2, vk', c_2'))$, and otherwise it outputs $\perp$. If the adversary eventually performs a query $D(vk', c_1', c_2', s')$ with $vk' = vk$ then the execution of the adversary is aborted and $B$ outputs $((c_1', c_2'), s')$, thus breaking the one-time strong unforgeability of the signature scheme.
Game 1 \approx Game 2. Assume that there is an adversary \((A_0, A_1)\) that is able to distinguish the environments of Game 1 and 2. Then we build an adversary \((B_0, B_1, B_2)\) against the IND-CCA security of the second TBE. Take \(B_0(1^\eta, \ell(\eta)) = [(vk, sk) \leftarrow \text{Gen}(1^\eta); \text{return } vk, (vk||sk)]\) and \(B_1(s_1, pk_2) = [r', r_1 \leftarrow \mathcal{R}; \text{return } r', r_1, (s_1||r'||r_1||pk_2)]\) and

\[
\text{Algorithm } B_2^{\mathcal{O}_{sk_2}}((vk||sk||r'||r_1||pk_2), c_2) : \\
pk_1, sk_1 \leftarrow \text{KeyGen}(1^\eta) \\
m_0, m_1, s_1 \leftarrow A_0^B(pk_1, pk_2) \\
b \leftarrow \{0, 1\} \\
c_1 \leftarrow \text{Enc}(pk_1, vk, m_1; r_1) \\
s \leftarrow \text{Sign}(sk, (c_1, c_2)) \\
b' \leftarrow A_1^B(s_1, (vk, c_1, c_2, s)) \\
\text{if } b = b' \text{ then return 1} \\
\text{else return 0}
\]

Calls to the decommitment oracle \(\mathcal{D}(vk, c_1, c_2, s)\) are simulated by firstly verifying the signature \(\text{Vrfy}(vk, (c_1, c_2), s)\). If the verification succeeds then the oracle returns \((\text{Dec}(sk_1, vk, c_1), \mathcal{O}_{sk_2}(c_2))\) and otherwise it outputs \(\perp\).

Finally we show that \(|\mathbb{P}[X_2] - 1/2|\) is bounded by the advantage any PPT algorithm has in breaking the selective IND-CCA security of the first instance of the TBE. Assume that there is an adversary \((A_0, A_1)\) for Game 2. Then we build an adversary \((B_0, B_1, B_2)\) against the IND-CCA security of the first TBE. Take \(B_0(1^\eta, \ell(\eta)) = [(vk, sk) \leftarrow \text{Gen}(1^\eta); \text{return } vk, (vk||sk)]\) and

\[
\text{Algorithm } B_1^{\mathcal{O}_{sk_1}}((vk||sk), pk_1) : \\
pk_2, sk_2 \leftarrow \text{KeyGen}(1^\eta) \\
m_0, m_1, s_1 \leftarrow A_0^B(pk_1, pk_2) \\
\text{return } (m_0, m_1), (vk||sk||pk_1||s_1||pk_2||sk_2)
\]

\[
\text{Algorithm } B_2^{\mathcal{O}_{sk_1}}((vk||sk||pk_1||s_1||pk_2||sk_2), c_1) : \\
r' \leftarrow \mathcal{R} \\
c_2 \leftarrow \text{Enc}(pk_2, vk, r') \\
s \leftarrow \text{Sign}(sk, (c_1, c_2)) \\
b' \leftarrow A_1^B(s_1, (vk, c_1, c_2, s)) \\
\text{return } b'
\]

Calls to the decommitment oracle \(\mathcal{D}(vk, c_1, c_2, s)\) are simulated by firstly verifying the signature \(\text{Vrfy}(vk, (c_1, c_2), s)\). If the verification succeeds then the oracle returns \((\mathcal{O}_{sk_1}(c_1), \text{Dec}(sk_2, vk, c_2))\) and otherwise it outputs \(\perp\).

6 Protocol Execution and State Traces

We now prove that it is possible to port proofs in the symbolic framework to the computational one. First, for the sake of self-containment we describe the adversarial model and the execution environment following the directions of Micciancio and Warinschi [MW04]. We refer the reader to this paper for a thorough explanation.
The message space and the closure operator were defined in Section 2. Messages are used to formally describe cryptographic protocols. The closure represents the knowledge that can be extracted from a message, and is used to define what valid algebraic protocol runs are. Intuitively a protocol run is valid if every message sent by a principal can be deduced from its knowledge except maybe for some fresh randomness. In this setting an adversary is in control of the communication media and is able to interact with honest participants. Consider then an adversary that has access to an oracle that will play the role of the honest participants. This adversary can start new sessions of the protocol and send messages to a principal of a given session and get the respective answer back. Formally, the adversary \( A \) can perform one of the following queries to the execution oracle \( O \).

1. \( \text{newsession}([I_1 \ldots I_n]) \) that takes a list of user identities \( I_i \) and returns a new session identifier \( s \).
2. \( \text{send}(s, I, m) \) that delivers the message \( m \) to the principal \( I \) of session \( s \). Then \( O \) updates \( I \)'s state and returns the answer to the adversary.

In case that the adversary performs a query that is not according to the protocol, for the specific state of the receiver, the oracle aborts the execution of this session.

In a formal protocol, the messages exchanged are algebraic expressions from the message algebra. A formal adversary \( A^f \) will interact with the formal oracle \( O^f \) in a symbolic protocol run.

On the other hand, a computational adversary \( A^c \) is a probabilistic polynomial-time Turing machine that operates on bitstrings. For a fixed value of the security parameter there is a set of primitive bitstrings for constants and nonces denoted by \( \text{Prim}_\eta \). The set of bitstrings \( \text{Msg}_\eta \) is build from \( \text{Prim}_\eta \) by tupling, encryptions, commitments and decommitments. There is a set \( \text{Sid} \) of session identifiers; a set \( \text{Uid} \) of user identities and a set \( \text{Vars} \) of variables in the abstract protocol description.

Let \( F : \text{Sid} \times \text{Uid} \rightarrow (\text{Vars} \rightarrow \text{Msg}, \mathbb{N}) \) be the state maintained by the formal oracle \( O^f \). On input \((s, I)\) it returns the state of principal \( I \) in session \( s \) together with his instruction pointer. The instruction pointer indicates on which step of the abstract protocol this principal is. Similarly, \( C : \text{Sid} \times \text{Uid} \rightarrow (\text{Vars} \rightarrow \text{Msg}_\eta, \mathbb{N}) \) is the state maintained by the computational oracle \( O^c \). Assume without loss of generality that all the sessions are created at the beginning. Then, a formal adversary \( A_f \) is just a sequence of \( \text{send}(s, I, m) \) queries. We say that a formal adversary \( A_f \) is a valid Dolev-Yao adversary (\( A_f \in \text{DY} \)) if each message he sends to the oracle is in the closure of his initial knowledge plus the answers he gets from the oracle \( O^f \). A protocol execution, thus, is the sequence of states \( F_0, F_1, \ldots \) of the formal oracle \( O^f \) and is denoted by \( \text{trace}(A_f, O^f) \). After fixing the randomness of the adversary and that of the oracle environment to \( \tau_A \) and \( \tau_O \), we can similarly define a computational execution \( \text{trace}(A_c((\tau_A), O^c(\tau_O))) \) as the sequence of states \( C_0, C_1, \ldots \) of the computational oracle \( O^c \).

**Definition 6.1.** We say that \( [\cdot] : \text{Prim} \rightarrow \text{Prim}_\eta \) is an interpretation function if it is injective and structure preserving (i.e., maps formal nonces to nonce bitstrings, formal commitments to commitments and so on).
Definition 6.2. Let $F = F_0, F_1, \ldots$ be a formal execution trace and let $C = C_0, C_1, \ldots$ be a concrete execution trace. We say that $F \preceq C$ if there exists an interpretation function $[\cdot]$ such that $[F_0] = C_0, [F_1] = C_1, \ldots$.

The following theorem shows that a computational adversary has no more power than an algebraic adversary.

Theorem 6.3. Let $(\text{TTP}, \text{Snd}, \text{Rcv})$ be an IND-COM-CCA secure commitment scheme and let $(K, E, D)$ be an IND-CCA secure encryption scheme. For any computational adversary $A_c$, the probability

$$P[\exists A_f \in DY : \text{trace}(A_f, O_f) \preceq \text{trace}(A_c(\tau_A), O^c(\tau_O))]$$

is overwhelming. Here the probability is taken over the random choices $\tau_A$ of the adversary and $\tau_O$ of the oracle.

Proof. First fix the randomness $\tau_A$ and $\tau_O$. Running the computational adversary $A_c$, it produces a sequence of queries/answers to/from the computational oracle. Because we know all the trapdoor information that the oracle generates and because the adversary has to send properly typed messages, we can de-construct any message sent into primitive terms. Choosing new algebraic terms for each distinct primitive bitstring encountered we build a sequence of algebraic queries which constitutes an algebraic adversary $A_f$. Note that for different random choices of $\tau_A$ and $\tau_O$ we get the same $A_f$ (up to renaming) with overwhelming probability.

It remains to show that the adversary we just built is Dolev-Yao. Suppose that it is not. Then $A_f$ must, at some point, send a query that contains a non-adversarial nonce $n^*$ that is not in the closure of the messages he received before. If this nonce occurs inside an encryption (with an unknown key) then one can build an adversary breaking the IND-CCA security of the encryption scheme [MW04]. Assume then that it occurs inside a commitment. We now build an adversary that breaks the IND-COM-CCA security of the commitment scheme.

This adversary simulates the environment to $A_c$ using the de-commit oracle when necessary except for the query that contains $n^*$. There it generates two interpretations $(n_0, n_1)$ for $n^*$ and gives them as challenge plaintext for the IND-COM-CCA game. The challenger gives back a commitment to $n_b$ where $b$ is the challenge bit. This commitment to $n_b$ is used to answer the oracle queries. At the moment $A_c$ outputs the interpretation of $n^*$ we can check whether it is $n_0$ or $n_1$.

A formal security notion is a predicate $P_f$ on formal traces. A protocol $\Pi \models_f P_f$ if for all adversaries $A_f \in DY$ holds that $\text{trace}(A_f, O_f) \in P_f$. Similarly, a computational security notion is a predicate $P_c$ on computational traces. A protocol $\Pi \models_c P_c$ if for all probabilistic polynomial-time adversaries $A_c$ holds that $\text{trace}(A_f, O_f) \in P_c$ with overwhelming probability (taken over the random choices of the adversary and the ones of the oracle environment). The proof of the following theorem follows as in [MW04].

Theorem 6.4. Let $(\text{TTP}, \text{Snd}, \text{Rcv})$ be a IND-COM-CCA secure commitment scheme and let $(K, E, D)$ be an IND-CCA secure encryption scheme. Let $P_f$ and $P_c$ be respectively formal and computational security notions such that for all formal traces.
ft and all computational traces ct it holds that \((ft \in P_f \land ft \preceq ct) \implies ct \in P_c\). Then

\[ \Pi \models_f P_f \implies \Pi \models_c P_c. \]

7 Conclusions

We presented two equivalent security notions for commitment schemes: a simulation based definition and a indistinguishability based one. We then gave a concrete scheme satisfying this security notion. This construction is of interest on itself as it is generic and has some interesting features like being reusable, perfectly binding and secure against adaptive chosen-commitment attacks. We then applied this new machinery to give sound interpretation of symbolic commitments while considering active adversaries.

References


