Relations between Semantic Security and Anonymity in Identity Based Encryption

Javier Herranz[1], Fabien Laguillaumie[2], and Carla Ràfols[1]

[1] Dept. Matemàtica Aplicada IV, Universitat Politècnica de Catalunya,
C. Jordi Girona 1-3, Mòdul C3, 08034, Barcelona, Spain.
e-mail: {jherranz,crafols}@ma4.upc.edu
[2] GREYC - Université de Caen Basse-Normandie,
Boulevard du Maréchal Juin, BP 5186, 14032 Caen Cedex, France.
e-mail: fabien.laguillaumie@unicaen.fr

Abstract. Semantic security and anonymity are the two main properties that an identity-based encryption scheme can satisfy. Such properties can be defined in either an adaptive or a selective scenario, which differ on the moment where the attacker chooses the identity/ies that are the target of the attack. There are well-known separations between selective and adaptive semantic security on the one hand, and between selective and adaptive anonymity on the other hand.
In this paper we investigate the relations between these selective and adaptive notions, for identity-based encryption schemes enjoying at the same time some security and anonymity properties. On the negative side, we prove that there is a separation between selective and adaptive anonymity even for schemes which enjoy adaptive semantic security. On the positive side, we prove that selective semantic security and adaptive anonymity imply adaptive semantic security.

Keywords. Identity-based encryption, selective/adaptive adversaries, semantic security, anonymity.

1 Introduction

Identity-based cryptography was introduced by Shamir in 1984 [Sha85] to simplify the key management in the public key setting. In this scenario, any sender A can encrypt a message to any receiver B using only a string corresponding to B's identity. Although the case of identity-based signatures was already solved in this seminal paper, the first encryption schemes appeared only in 2001 in independent works by Boneh and Franklin [BF03] and Cocks [Coc01] based on entirely different approaches.

The main security properties that can be required of identity-based encryption (we will sometimes use IBE, for short) are semantic security and anonymity. These properties are defined in either a *selective* or an *adaptive* scenario: in the first one, the attacker chooses its target identity/ies at the beginning of the game, while in the latter, it chooses its target identity/ies after some adaptive queries to an extraction oracle which gives secret keys corresponding to the requested identities.

Identity-based encryption with semantic security *and* anonymity is not only interesting as a cryptographic primitive, but also because it can be used to design other primitives such as public key encryption with keyword search, as proved in [BDOP04,A+08]. The first anonymous IBE scheme is indeed Boneh and Franklin's [BF03], although that was not explicitly stated, but its main drawback is the fact that security proofs are carried out in the random oracle model. The scheme in [AG09] is also fully (or adaptively) anonymous under the quadratic residuosity assumption (in particular, it does not employ bilinear pairings), but again in the random oracle model. There exist IBE schemes which are semantically secure in the standard model (see for instance those from [BB04a,BB04b,Wat05,Nac05,CS05]), but achieving anonymity at the same time seems considerably harder. The first identity-based schemes enjoying anonymity in the standard model are those in [BW06] and [Gen06]. The first fully anonymous *hierarchical* identity-based encryption scheme was provided in [DIP10] as a modification of

the scheme from [LW10]. These schemes are mainly based on bilinear pairings. Recently, some constructions of (hierarchical) identity-based encryption schemes in a lattice setting have been proposed [CHKP10,ABB10a,ABB10b], achieving selective or adaptive security in the standard model.

Intuition and Contributions. There exist generic conversions from a selectively secure/anonymous IBE scheme to an adaptively secure/anonymous IBE scheme, either in the random oracle model or when the size of the space of identities is small [BB04a]. However, in general there is a separation between the two models. For example, Galindo proved [Gal06] a separation result regarding semantic security: any IBE scheme which has selective semantic security can be transformed into another scheme which also has selective semantic security, but does not even enjoy one-wayness against adaptive attacks. The idea of this transformation is to choose a special identity id^* in the setup phase, and add the secret key for id^* in the public parameters.

Similar separation results can be easily proven for the case of anonymity. However, note that the transformation by Galindo leads to a quite artificial IBE scheme, which in particular is not anonymous against adaptive attacks, because ciphertexts addressed to id^* can be easily told apart from the rest of ciphertexts. This observation motivates this work. We want to investigate the relation between selective and adaptive semantic security (respectively, anonymity) for IBE schemes which are not so artificial, for example because they also enjoy some anonymity (respectively, semantic security) property. It is interesting to note that the existing identity-based encryption schemes in the literature which enjoy both semantic security and anonymity have either both properties proved in the selective setting [BW06,BW07,Duc10,ABB10a] or both properties proved in the adaptive setting [Gen06,CKRS09,DIP10].

We provide both negative and positive results, which are summarized in Table 3 of Section 3. On the negative side, we prove that an IBE scheme which is at the same time semantically secure and anonymous in front of selective attacks is not necessarily semantically secure nor anonymous in front of adaptive attacks. Then, we prove that there is a separation between selective anonymity and adaptive anonymity even for IBE schemes which are fully (i.e. adaptively) semantically secure. On the positive side, we prove that the symmetric situation is different: for IBE schemes which are fully (i.e. adaptively) anonymous, the notions of selective and adaptive semantic security are equivalent.

We describe and prove our results in the scenario of chosen-plaintext attackers who cannot make decryption queries for ciphertexts of their choice, but our results extend to a chosen-ciphertext attack scenario. Finally, the same results are valid for hierarchical identity-based encryption, as well.

Organization. Section 2 contains the syntactic definition of an identity-based encryption scheme, as well as the semantic security and anonymity definitions for both selective and adaptive adversaries. The main results about the relations among these security notions are proven in Section 3. In Section 4 we discuss some potential applications of our results, in particular in the scenario of hierarchical identity-based encryption. Finally, we conclude in Section 5.

2 Identity-based Encryption and its Security

In this section we give some basic definitions related to an identity-based encryption scheme, including several different security notions regarding indistinguishability and anonymity.

2.1 Syntactic Definition

Let k be a positive integer and let $\mathcal{ID} = \mathcal{ID}(k)$ be a set of possible identities. An *identity-based encryption (IBE)* scheme Π handling identities in $\mathcal{ID}$ is a tuple of probabilistic polynomial time algorithms $(\mathsf{Setup}, \mathsf{Extract}, \mathsf{Encrypt}, \mathsf{Decrypt})$ defined as follows.

- Setup takes a security parameter 1^k as input and produces the system parameters params and a master key msk.

– $\mathsf{Extract}$ takes a security parameter 1^k, the system parameters params, the master key msk and an identity $\mathsf{id} \in \mathcal{ID}$ as inputs. It outputs the secret key $\mathsf{sk}_{\mathsf{id}}$ corresponding to the identity id.
– $\mathsf{Encrypt}$ takes a security parameter 1^k, the system parameters params, an identity $\mathsf{id} \in \mathcal{ID}$ and a message $m \in \{0,1\}^*$ as inputs and outputs a ciphertext c.
– $\mathsf{Decrypt}$ takes a security parameter 1^k, the system parameters params, a secret key $\mathsf{sk}_{\mathsf{id}}$ and a ciphertext c as inputs, and outputs a message m.

These algorithms have to satisfy the *correctness* property: for all $k \in \mathbb{N}$, $\mathsf{id} \in \mathcal{ID}$ and $m \in \{0,1\}^*$,

$$\Pr\Big[(\mathsf{params}, \mathsf{msk}) \xleftarrow{\$} \mathsf{Setup}(1^k), \mathsf{sk}_{\mathsf{id}} \xleftarrow{\$} \mathsf{Extract}(1^k, \mathsf{params}, \mathsf{msk}, \mathsf{id}),$$
$$c \xleftarrow{\$} \mathsf{Encrypt}(1^k, \mathsf{params}, \mathsf{id}, m) : \mathsf{Decrypt}(1^k, \mathsf{params}, \mathsf{sk}_{\mathsf{id}}, c) = m\Big] = 1\ .$$

2.2 Security Requirements

For simplicity, we will focus on security against adversaries which do not make decryption queries. However, the results in this paper apply also to the scenario of chosen-ciphertext attacks.

Semantic Security. Boneh and Franklin define in [BF03] the concept of chosen plaintext security (also known as semantic security, or indistinguishability) for identity-based encryption under a *chosen* identity attack.

Definition 1 (IND-CPA). *Let $\Pi = (\mathsf{Setup}, \mathsf{Extract}, \mathsf{Encrypt}, \mathsf{Decrypt})$ be an identity-based encryption scheme. Let $k \in \mathbb{N}$ and let $\mathcal{ID} = \mathcal{ID}(k)$ be a set of identities. Let $\mathcal{A} = (\mathcal{A}_f, \mathcal{A}_g)$ be an adversary that runs in two stages with access to an extraction oracle $\mathcal{O}_{\mathsf{Extract}}(\cdot)$. We consider the following random experiments:*

$$\begin{array}{l}\underline{\boldsymbol{Exp}^{\mathsf{ind\text{-}cpa}}_{\Pi,\mathcal{A}}(k, \mathcal{ID})} \\ \quad (\mathsf{params}, \mathsf{msk}) \leftarrow \Pi.\mathsf{Setup}(1^k) \\ \quad (m_0, m_1, \mathsf{id}_{ch}, st) \leftarrow \mathcal{A}_f^{\mathcal{O}_{\mathsf{Extract}}(\cdot)}(1^k, \mathsf{params}) \\ \quad b \xleftarrow{\$} \{0,1\} \\ \quad c \leftarrow \Pi.\mathsf{Encrypt}(1^k, \mathsf{params}, \mathsf{id}_{ch}, m_b) \\ \quad b' \leftarrow \mathcal{A}_g^{\mathcal{O}_{\mathsf{Extract}}(\cdot)}(1^k, c, st) \\ \quad \textit{Return } (b' = b)\end{array}$$

During the two stages, $\mathcal{A}_f$ and $\mathcal{A}_g$ run under the restriction that they do not query their extraction oracle on id_{ch}. The advantage *of $\mathcal{A}$ is defined as*

$$\mathsf{Adv}^{\mathsf{ind\text{-}cpa}}_{\Pi,\mathcal{A}}(k, \mathcal{ID}) = \left|\Pr\left[\boldsymbol{Exp}^{\mathsf{ind\text{-}cpa}}_{\Pi,\mathcal{A}}(k, \mathcal{ID}) = 1\right] - \frac{1}{2}\right|.$$

The scheme Π is said to be indistinguishable under a chosen plaintext attack *if the function $\mathsf{Adv}^{\mathsf{ind\text{-}cpa}}_{\Pi,\mathcal{A}}$ is negligible for any adversary $\mathcal{A}$ whose time complexity is polynomial in k.*

The notion of $\mathsf{IND\text{-}CPA}$ security for identity-based encryption schemes can be weakened, by forcing the adversary to *select* the challenge identity $\mathsf{id}_{\mathsf{ch}} \in \mathcal{ID}$ at the first stage of the previous experiment. In some sense, the adversary commits to the identity he will try to attack in the future.

Definition 2 (IND-sID-CPA). *Let $\Pi = (\mathsf{Setup}, \mathsf{Extract}, \mathsf{Encrypt}, \mathsf{Decrypt})$ be an identity-based encryption scheme. Let $k \in \mathbb{N}$. Let $\mathcal{A} = (\mathcal{A}_{init}, \mathcal{A}_f, \mathcal{A}_g)$ be an adversary that runs in three stages with access to an extraction oracle $\mathcal{O}_{\mathsf{Extract}}(\cdot)$. We consider the following random experiments:*

$$\underline{\boldsymbol{Exp}^{\mathsf{ind\text{-}sid\text{-}cpa}}_{\Pi,\mathcal{A}}(k,\mathcal{ID})}$$
$(\mathsf{id}_{ch}, st) \leftarrow \mathcal{A}_{init}(1^k, \mathcal{ID})$
$\mathsf{params} \leftarrow \mathsf{IBE.Setup}(1^k)$
$(m_0, m_1, st') \leftarrow \mathcal{A}_f^{\mathcal{O}_{\mathsf{Extract}}(\cdot)}(1^k, st)$
$b \xleftarrow{\$} \{0,1\}$
$c \leftarrow \Pi.\mathsf{Encrypt}(1^k, \mathsf{params}, \mathsf{id}_{ch}, m_b)$
$b' \leftarrow \mathcal{A}_g^{\mathcal{O}_{\mathsf{Extract}}(\cdot)}(1^k, c, st')$
Return $(b' = b)$

During the two stages, $\mathcal{A}_f$ and $\mathcal{A}_g$ run under the restriction that they do not query their extraction oracle on id_{ch}. The advantage *of $\mathcal{A}$ is defined as*

$$\mathsf{Adv}^{\mathsf{ind\text{-}sid\text{-}cpa}}_{\Pi,\mathcal{A}}(k,\mathcal{ID}) = \left|\Pr\left[\boldsymbol{Exp}^{\mathsf{ind\text{-}sid\text{-}cpa}}_{\Pi,\mathcal{A}}(k,\mathcal{ID}) = 1\right] - \frac{1}{2}\right|.$$

The scheme Π is said to be indistinguishable under a chosen plaintext attack for selective identity *if the function $\mathsf{Adv}^{\mathsf{ind\text{-}sid\text{-}cpa}}_{\Pi,\mathcal{A}}$ is negligible for any adversary $\mathcal{A}$ whose time complexity is polynomial in k.*

In contrast to this weakened *selective* security notion for identity-based encryption, we will sometimes refer to the standard IND-CPA security notion as *full* security.

Anonymity. The notion of anonymity corresponds to the notion of *key-privacy* for public key encryption [BBDP01]. Roughly speaking, it ensures that a ciphertext does not leak any information on the identity of the recipient. Halevi gave in [Hal05] a simple sufficient condition for public-key encryption that provides data privacy to reach key-privacy. Essentially, this condition means that a random encryption of a random message is independent of the public key. Abdalla *et al.* extended this condition to identity-based encryption in [A+08]. Anonymity for IBE was formally defined in [A+08] for the first time, in an *adaptive* scenario: the attacker chooses the two identities it wants to be challenged on at the end of its first stage (where it can query for secret key extractions), just before querying the challenge ciphertext.

Definition 3 (ANO-CPA). *Let $\Pi = (\mathsf{Setup}, \mathsf{Extract}, \mathsf{Encrypt}, \mathsf{Decrypt})$ be an identity-based encryption scheme. Let $k \in \mathbb{N}$ and let $\mathcal{ID} = \mathcal{ID}(k)$ be a set of identities. Let $\mathcal{D} = (\mathcal{D}_f, \mathcal{D}_g)$ be an adversary that runs in two stages with access to an extraction oracle $\mathcal{O}_{\mathsf{Extract}}(\cdot)$. We consider the following random experiments:*

$$\underline{\boldsymbol{Exp}^{\mathsf{ano\text{-}cpa}}_{\Pi,\mathcal{D}}(k,\mathcal{ID})}$$
$(\mathsf{params}, \mathsf{msk}) \leftarrow \Pi.\mathsf{Setup}(1^k)$
$(m, \mathsf{id}_0, \mathsf{id}_1, st) \leftarrow \mathcal{D}_f^{\mathcal{O}_{\mathsf{Extract}}(\cdot)}(1^k, \mathsf{params})$
$\tilde{b} \xleftarrow{\$} \{0,1\}$
$c \leftarrow \Pi.\mathsf{Encrypt}(1^k, \mathsf{params}, \mathsf{id}_{\tilde{b}}, m)$
$\tilde{b}' \leftarrow \mathcal{D}_g^{\mathcal{O}_{\mathsf{Extract}}(\cdot)}(1^k, c, st)$
Return $(\tilde{b}' = \tilde{b})$

During the two stages, $\mathcal{D}_f$ and $\mathcal{D}_g$ run under the restriction that they do not query their extraction oracle on $\mathsf{id}_0, \mathsf{id}_1$. The advantage *of $\mathcal{D}$ is defined as*

$$\mathsf{Adv}^{\mathsf{ano\text{-}cpa}}_{\Pi,\mathcal{D}}(k,\mathcal{ID}) = \left|\Pr\left[\boldsymbol{Exp}^{\mathsf{ano\text{-}cpa}}_{\Pi,\mathcal{D}}(k,\mathcal{ID}) = 1\right] - \frac{1}{2}\right|.$$

The scheme Π is said to be anonymous under a chosen plaintext attack *if the function $\mathsf{Adv}^{\mathsf{ano\text{-}cpa}}_{\Pi,\mathcal{D}}$ is negligible for any adversary $\mathcal{D}$ whose time complexity is polynomial in k.*

Again, this notion of adaptive (or full) anonymity, ANO-CPA, can be weakened if the adversary is forced to select the two challenge identities at the first stage of the attack. The resulting notion of *selective anonymity* is formally defined as follows.

Definition 4 (ANO-sID-CPA). *Let $\Pi = (\mathsf{Setup}, \mathsf{Extract}, \mathsf{Encrypt}, \mathsf{Decrypt})$ be an identity-based encryption scheme. Let $k \in \mathbb{N}$. Let $\mathcal{D} = (\mathcal{D}_{init}, \mathcal{D}_f, \mathcal{D}_g)$ be an adversary that runs in three stages with access to an extraction oracle $\mathcal{O}_{\mathsf{Extract}}(\cdot)$. We consider the following random experiments:*

$$
\begin{array}{l}
\underline{\boldsymbol{Exp}^{\mathsf{ano\text{-}sid\text{-}cpa}}_{\Pi,\mathcal{D}}(k, \mathcal{ID})} \\
\quad (\mathsf{id}_0, \mathsf{id}_1, st) \leftarrow \mathcal{D}_{init}(1^k, \mathcal{ID}) \\
\quad \mathsf{params} \leftarrow \mathsf{IBE.Setup}(1^k) \\
\quad (m, st') \leftarrow \mathcal{D}_f^{\mathcal{O}_{\mathsf{Extract}}(\cdot)}(1^k, st) \\
\quad \tilde{b} \xleftarrow{\$} \{0,1\} \\
\quad c \leftarrow \Pi.\mathsf{Encrypt}(1^k, \mathsf{params}, \mathsf{id}_{\tilde{b}}, m) \\
\quad \tilde{b}' \leftarrow \mathcal{D}_g^{\mathcal{O}_{\mathsf{Extract}}(\cdot)}(1^k, c, st') \\
\quad \textit{Return } (\tilde{b}' = \tilde{b})
\end{array}
$$

During the two last stages, $\mathcal{D}_f$ and $\mathcal{D}_g$ run under the restriction that they do not query their extraction oracle on $\mathsf{id}_0, \mathsf{id}_1$. The advantage *of $\mathcal{D}$ is defined as*

$$
\mathsf{Adv}^{\mathsf{ano\text{-}sid\text{-}cpa}}_{\Pi,\mathcal{D}}(k, \mathcal{ID}) = \left| \Pr\left[\boldsymbol{Exp}^{\mathsf{ano\text{-}sid\text{-}cpa}}_{\Pi,\mathcal{D}}(k, \mathcal{ID}) = 1 \right] - \frac{1}{2} \right| .
$$

The scheme Π is said to be anonymous under a selective identity chosen plaintext attack *if the function $\mathsf{Adv}^{\mathsf{ano\text{-}sid\text{-}cpa}}_{\Pi,\mathcal{D}}$ is negligible for any adversary $\mathcal{D}$ whose time complexity is polynomial in k.*

On the Size of $\mathcal{ID}$. If the size of the space $\mathcal{ID}$ of possible identities is polynomial in the security parameter k, then the selective and the adaptive definitions of both indistinguishability and anonymity are polynomially equivalent. Indeed, an adaptive attacker can be turned into a selective attacker that guesses in the first stage the identity/ies that the adaptive attacker will choose later. The guess will be correct with non-negligible probability.

Therefore, in the rest of the paper we assume that the size of $\mathcal{ID}$ is at least exponential in k, and so $\alpha/\#\mathcal{ID}$ is a negligible function with respect to k, for any constant α.

On the techniques to prove adaptive security. Boneh and Boyen [BB04b] and Waters [Wat05] were the first to prove full semantic security in the standard model. Their security proofs followed what Waters [Wat09] calls a partition strategy: the space of identities is partitioned into the set of identities for which a valid secret key can be simulated and those for which a valid challenge ciphertext can be simulated and then it must be argued that the probability that the adversary asks his queries according to the partition is large enough. On the other hand, Gentry [Gen06] was able to prove adaptive semantic security and anonymity using completely different techniques, at the price of basing security on a computational hypothesis which depends on the maximum number of extraction queries of the adversary. Finally, Waters proposed in [Wat09] a new methodology to prove adaptive security, called *dual system encryption*, which has been successfully applied to prove adaptive security not only of IBE and HIBE schemes, but also of schemes with even more functionality like attribute-based encryption or predicate-based encryption schemes [L+10]. Recently, [DIP10] designed a HIBE scheme adaptively anonymous and semantically secure using these techniques.

3 Relations among IND-sID-CPA, IND-CPA, ANO-sID-CPA and ANO-CPA

This section contains the main results of the paper. We prove different relations between the notions of indistinguishability and anonymity introduced in the previous section. These relations are depicted in Table 3.

Anonymity / Indistinguishability	ANO-sID-CPA	ANO-CPA
IND-sID-CPA	$\not\Rightarrow$ IND-CPA (Thm. 1), $\not\Rightarrow$ ANO-CPA (Thm. 2)	$\Rightarrow$ IND-CPA (Thm. 3), $\Rightarrow$ ANO-CPA (trivial)
IND-CPA	$\not\Rightarrow$ ANO-CPA (Thm. 2), $\Rightarrow$ IND-CPA (trivial)	$\Rightarrow$ ANO-CPA (trivial), $\Rightarrow$ IND-CPA (trivial)

Table 1. Taxonomy of the notions of IND-sID-CPA, IND-CPA, ANO-sID-CPA and ANO-CPA for IBE.

Essentially, there are two negative results and one positive result. The first negative result (Theorem 1) states that an IBE scheme which is both IND-sID-CPA and ANO-sID-CPA secure must not necessarily be IND-CPA secure. Then, we also prove in Theorem 2 that an IBE scheme which is both IND-CPA and ANO-sID-CPA secure must not necessarily be ANO-CPA secure. Of course, this negative result propagates up in the table, and also holds if the IBE scheme is only IND-sID-CPA and ANO-sID-CPA secure. Finally, our positive result (Theorem 3, which may be considered as the most significant contribution of our study) states that an IBE scheme which is both IND-sID-CPA and ANO-CPA secure must necessarily be IND-CPA secure, as well.

Therefore, the situation (as well as the table) is not symmetric. On the one hand, if an IBE scheme is fully (adaptively) anonymous, then selective and adaptive indistinguishability are equivalent (Theorem 3). On the other hand, even if an IBE scheme is fully (adaptively) secure in terms of indistinguishability, there is still a separation between selective and adaptive anonymity (Theorem 2).

It is not hard to see that all the proofs in this paper can be adapted to the scenario of chosen-ciphertext attacks (CCA), in which adversaries can request decryption of ciphertexts chosen by them (different to the challenge ciphertext). Therefore, by replacing CPA with CCA everywhere in this section we obtain exactly the same relations among the notions of IND-sID-CCA, IND-CCA, ANO-sID-CCA and ANO-CCA.

3.1 Negative Results

A necessary and implicit assumption in the proof of our two negative results is the existence of identity-based encryption schemes which are IND-sID-CPA and ANO-sID-CPA secure at the same time (for Theorem 1) and are IND-CPA and ANO-sID-CPA secure at the same time (for Theorem 2). Some examples of such schemes can be found in [BW06,BW07,ABB10a] for the first case, or in [Gen06,CKRS09,DIP10] for the second case.

The proof for the first negative result follows the same argument as the proof in [Gal06] for the separation between the notions of IND-sID-CPA and IND-CPA. Actually, we could prove that simultaneous IND-sID-CPA and ANO-sID-CPA (selective) security does not even imply one-wayness against full (non-selective) attackers.

Theorem 1. *There exist identity-based encryption schemes that are secure under* IND-sID-CPA *and* ANO-sID-CPA *attacks, but are not secure under* IND-CPA *attacks.*

Proof. Let $\Pi = (\mathsf{Setup}, \mathsf{Extract}, \mathsf{Encrypt}, \mathsf{Decrypt})$ be an IBE scheme with set of identities $\mathcal{ID}$, which is at the same time IND-sID-CPA and ANO-sID-CPA secure. Let us consider the following IBE scheme Π'.

- $\Pi'.\mathsf{Setup}(1^k)$ first runs $\Pi.\mathsf{Setup}(1^k) \rightarrow (\mathsf{params}, \mathsf{msk})$. Then an identity $\mathsf{id}^* \in \mathcal{ID}$ is chosen at random. The public parameters of Π' are $\mathsf{params}' = (\mathsf{params}, \mathsf{id}^*)$. The master secret key is msk.
- $\Pi'.\mathsf{Extract}$ works exactly as $\Pi.\mathsf{Extract}$.
- $\Pi'.\mathsf{Encrypt}(1^k, \mathsf{params}', \mathsf{id}, m) = \begin{cases} (0, \Pi.\mathsf{Encrypt}(1^k, \mathsf{params}, \mathsf{id}, m)), & \text{if } \mathsf{id} \neq \mathsf{id}^* \\ (1, m), & \text{if } \mathsf{id} = \mathsf{id}^* \end{cases}$
- $\Pi'.\mathsf{Decrypt}$ runs $\Pi.\mathsf{Decrypt}$ if the first bit of the ciphertext is 0. Otherwise, if the first bit of the ciphertext is 1 (when the intended identity is id^*), this bit is removed and the rest of the ciphertext is output as the plaintext.

The scheme Π' inherits both selective properties (IND-sID-CPA and ANO-sID-CPA security) of scheme Π, as long as the special identity id^* chosen in $\Pi'.\mathsf{Setup}$ is different from the identities $\mathsf{id}_{\mathsf{ch}}, \mathsf{id}_0, \mathsf{id}_1$ chosen by the selective attackers in the first stage. This happens with overwhelming probabilities $1 - \frac{1}{\#\mathcal{ID}}$ (for IND-sID-CPA attackers) and $1 - \frac{2}{\#\mathcal{ID}}$ (for ANO-sID-CPA attackers). In this case, a successful attack against Π' can be immediately turned into a successful attack against Π.

But clearly Π' does not enjoy the indistinguishability property against adaptive (non-selective) attackers. Such an attacker $\mathcal{A}$ receives the public parameters $\mathsf{params}' = (\mathsf{params}, \mathsf{id}^*)$ of the scheme and can choose the challenge identity as $\mathsf{id}_{\mathsf{ch}} = \mathsf{id}^*$, along with two arbitrary messages m_0, m_1. The challenge ciphertext will then be $c = (1, m_b)$, and so $\mathcal{A}$ will always guess the correct value of the bit b.

Therefore, we have described an IBE scheme Π' which is at the same time IND-sID-CPA and ANO-sID-CPA secure, but is not IND-CPA secure. □

Now we prove a stronger negative result: if we strengthen the semantic security notion, from IND-sID-CPA to IND-CPA, and we keep the selective anonymity notion of ANO-sID-CPA, then we cannot guarantee that the scheme achieves the full anonymity notion of ANO-CPA.

Theorem 2. *There exist identity-based encryption schemes that are secure under* IND-CPA *and* ANO-sID-CPA *attacks, but are not secure under* ANO-CPA *attacks.*

Proof. Let $\Pi = (\mathsf{Setup}, \mathsf{Extract}, \mathsf{Encrypt}, \mathsf{Decrypt})$ be an IBE scheme with set of identities $\mathcal{ID}$, which is at the same time IND-CPA and ANO-sID-CPA secure. Let us consider the following IBE scheme Π'.

- $\Pi'.\mathsf{Setup}(1^k)$ first runs $\Pi.\mathsf{Setup}(1^k) \rightarrow (\mathsf{params}, \mathsf{msk})$. Then an identity $\mathsf{id}^* \in \mathcal{ID}$ is chosen at random. The public parameters of Π' are $\mathsf{params}' = (\mathsf{params}, \mathsf{id}^*)$. The master secret key is msk.
- $\Pi'.\mathsf{Extract}$ works exactly as $\Pi.\mathsf{Extract}$.
- $\Pi'.\mathsf{Encrypt}(1^k, \mathsf{params}', \mathsf{id}, m) = \begin{cases} (0, \Pi.\mathsf{Encrypt}(1^k, \mathsf{params}, \mathsf{id}, m)), & \text{if } \mathsf{id} \neq \mathsf{id}^* \\ (1, \Pi.\mathsf{Encrypt}(1^k, \mathsf{params}, \mathsf{id}, m)), & \text{if } \mathsf{id} = \mathsf{id}^* \end{cases}$
- $\Pi'.\mathsf{Decrypt}$ just ignores the first bit of the ciphertext, and runs $\Pi.\mathsf{Decrypt}$.

Regarding indistinguishability, the scheme Π' clearly inherits the IND-CPA property of Π. Regarding anonymity, the scheme Π' inherits the ANO-sID-CPA property of Π as long as the special identity id^* chosen in $\Pi'.\mathsf{Setup}$ is not one of the two identities $\mathsf{id}_0, \mathsf{id}_1$ chosen previously by the ANO-sID-CPA attacker. This happens with overwhelming probability $1 - \frac{2}{\#\mathcal{ID}}$.

However, it is easy to see that the scheme Π' is not secure against full ANO-CPA attackers. Such an attacker $\mathcal{D}$ will see $\mathsf{params}' = (\mathsf{params}, \mathsf{id}^*)$ and then will choose an arbitrary identity $\mathsf{id}_0 \neq \mathsf{id}^*$ and $\mathsf{id}_1 = \mathsf{id}^*$. Later, when receiving the challenge ciphertext $c = \Pi'.\mathsf{Encrypt}(1^k, \mathsf{params}', \mathsf{id}_b, m)$, the attacker $\mathcal{D}$ obtains exactly the value of the bit b by looking at the first bit of c.

Summing up, we have an IBE scheme Π' which is IND-CPA and ANO-sID-CPA secure, but is not ANO-CPA secure. □

3.2 Positive Result

However, if we strengthen the anonymity notion, from ANO-sID-CPA to ANO-CPA, then the weaker (selective) indistinguishability notion of IND-sID-CPA becomes equivalent to the adaptive notion of

IND-CPA. The following theorem ensures that an identity based encryption scheme which is selectively secure (IND-sID-CPA) and fully anonymous (ANO-CPA) in a chosen plaintext scenario is also fully secure (IND-CPA).

Theorem 3. *Let k be an integer and let $\mathcal{ID} = \mathcal{ID}(k)$ be a set of possible identities. For any IND-CPA adversary $\mathcal{A}$ against an identity based encryption scheme Π, there exists an IND-sID-CPA adversary $\mathcal{A}'$ and an ANO-CPA adversary $\mathcal{D}$ such that*

$$\mathit{Adv}^{\mathit{ind\text{-}cpa}}_{\Pi,\mathcal{A}}(k,\mathcal{ID}) \leq \frac{q_E+1}{\#\mathcal{ID}} + 2\cdot \mathit{Adv}^{\mathit{ano\text{-}cpa}}_{\Pi,\mathcal{D}}(k,\mathcal{ID}) + \mathit{Adv}^{\mathit{ind\text{-}sid\text{-}cpa}}_{\Pi,\mathcal{A}'}(k,\mathcal{ID})$$

where q_E denotes $\mathcal{A}$'s number of queries to its extraction oracle, and $\#\mathcal{ID}$ is the cardinality of the set $\mathcal{ID}$.

Proof. Let $\mathcal{A}$ be a successful IND-CPA adversary against the identity-based encryption Π. Let us consider the following sequence of games:

Game_0 This game is the real game between an IND-CPA adversary $\mathcal{A}$ against the indistinguishability of Π with chosen plaintexts and its challenger $\mathcal{C}$. Precisely, $\mathcal{C}$ runs $\Pi.\mathsf{Setup}(1^k)$ and gives the resulting parameters to $\mathcal{A}$, which engages in a series of queries to obtain the secret keys corresponding to identities of its choice. The challenger answers these queries with $\Pi.\mathsf{Extract}(\cdot)$. When the attacker outputs two messages m_0 and m_1, and an identity $\mathsf{id}_{\mathsf{ch}}$ (that he has not queried to the secret key extraction oracle), the challenger picks a bit $b \in \{0,1\}$ at random and runs $\Pi.\mathsf{Encrypt}(1^k, \mathsf{params}, \mathsf{id}_{\mathsf{ch}}, m_b)$ and gives the corresponding ciphertext c_{ch} to $\mathcal{A}$. Once again, $\mathcal{A}$ makes some extraction queries (not for $\mathsf{id}_{\mathsf{ch}}$) and eventually outputs a bit b'. We denote as Guess_0 the event $b = b'$. We have

$$\mathsf{Adv}^{\mathsf{ind\text{-}cpa}}_{\Pi,\mathcal{A}}(k,\mathcal{ID}) = \left|\Pr[\mathsf{Guess}_0] - \frac{1}{2}\right|.$$

Game_1 In this game, the challenger is modified in the following way: before all, it picks at random an identity $\mathsf{id}^\star \in \mathcal{ID}$. It then behaves exactly as in the previous game, except that the simulation aborts whenever $\mathcal{A}$ asks an extraction query (at any stage) for $\mathsf{id}^\star$, or if $\mathcal{A}$ chooses $\mathsf{id}_{\mathsf{ch}} = \mathsf{id}^\star$; this might happen with probability $(q_E+1)/\#\mathcal{ID}$. If we denote as Guess_1 the event $b = b'$, then

$$|\Pr[\mathsf{Guess}_0] - \Pr[\mathsf{Guess}_1]| \leq \frac{q_E+1}{\#\mathcal{ID}}.$$

Game_2 In this last game, the only modification concerns the challenge ciphertext: the challenger sends to the attacker the challenge ciphertext $c^\star = \Pi.\mathsf{Encrypt}(1^k, \mathsf{params}, \mathsf{id}^\star, m_b)$ instead of a ciphertext c_{ch} intended to $\mathsf{id}_{\mathsf{ch}}$. We denote as Guess_2 the event $b = b'$.

We will now show:

Fact 1 *There exists an ANO-CPA attacker $\mathcal{D}$ such that*

$$|\Pr[\mathit{Guess}_1] - \Pr[\mathit{Guess}_2]| = 2\cdot \mathit{Adv}^{\mathit{ano\text{-}cpa}}_{\Pi,\mathcal{D}}(k,\mathcal{ID}).$$

Proof. With the IND-CPA adversary $\mathcal{A}$ as a black-box oracle, the adversary $\mathcal{D}$ is designed as follows: it picks at random an identity $\mathsf{id}^\star$ that it stores internally. Then $\mathcal{D}$ receives some global parameters params that it forwards to $\mathcal{A}$. After that, $\mathcal{D}$ answers $\mathcal{A}$'s extraction queries with the help of its own extraction oracle.

When $\mathcal{A}$ outputs two messages m_0, m_1 and an identity $\mathsf{id}_{\mathsf{ch}}$, $\mathcal{D}$ picks at random a bit $d \in \{0,1\}$ and sends to its challenger m_d as well as the identities $\mathsf{id}_0 = \mathsf{id}^\star$ and $\mathsf{id}_1 = \mathsf{id}_{\mathsf{ch}}$. Its challenger replies with the ciphertext $c^\star = \Pi.\mathsf{Encrypt}(1^k, \mathsf{params}, \mathsf{id}_{\tilde{b}}, m_d)$ for the unknown challenge bit $\tilde{b}$. $\mathcal{D}$ forwards this ciphertext $c^\star$ to $\mathcal{A}$, which outputs a bit d' after a number of queries to the extraction oracle whose

answers are obtained with $\mathcal{D}$'s oracle (note that id_1 cannot be queried by definition of the attacker $\mathcal{A}$ and that id_0 cannot be queried by definition of games 1 and 2). If $d' = d$, $\mathcal{D}$ outputs 1. Otherwise, $\mathcal{D}$ outputs 0.

We have the following equalities:

$$\frac{1}{2} + \mathsf{Adv}^{\mathsf{ano\text{-}cpa}}_{\Pi,\mathcal{D}}(k, \mathcal{ID}) = \Pr[D \text{ wins}] = \frac{1}{2}\Pr[d' \neq d | \tilde{b} = 0] + \frac{1}{2}\Pr[d' = d | \tilde{b} = 1] =$$

$$= \frac{1}{2}\left(1 - \Pr[\mathsf{Guess}_2]\right) + \frac{1}{2}\Pr[\mathsf{Guess}_1] = \frac{1}{2} + \frac{1}{2}\left(\Pr[\mathsf{Guess}_1] - \Pr[\mathsf{Guess}_2]\right)$$

The equalities hold because when $\tilde{b} = 1$ Game 1 is perfectly simulated, and when $\tilde{b} = 0$ then Game 2 is perfectly simulated. □

Fact 2 *There exists an IND-sID-CPA attacker $\mathcal{A}'$ such that* $\mathsf{Adv}^{\mathsf{ind\text{-}sid\text{-}cpa}}_{\Pi,\mathcal{A}'}(k, \mathcal{ID}) = \left|\Pr[\mathsf{Guess}_2] - \frac{1}{2}\right|$.

Proof. The IND-sID-CPA attacker $\mathcal{A}'$ works as follows: in Game 2, it picks at random an identity $\mathsf{id}^\star$ that it outputs as the identity he wants to attack. After that, $\mathcal{A}'$ forwards to $\mathcal{A}$ the global parameters params it receives. $\mathcal{A}'$ answers $\mathcal{A}$'s extraction queries with its own oracle, and when $\mathcal{A}$ outputs at the end of its first stage a triple $(m_0, m_1, \mathsf{id}_{\mathsf{ch}})$, $\mathcal{A}'$ extracts m_0 and m_1 and gives these messages to its challenger. It then receives a challenge ciphertext $c^\star = \Pi.\mathsf{Encrypt}(1^k, \mathsf{params}, \mathsf{id}^\star, m_{d^\star})$ for the unknown challenge bit $d^\star$. $\mathcal{A}'$ forwards this challenge to $\mathcal{A}$, and then answers its extraction queries with its own oracle. Eventually, $\mathcal{A}$ outputs a bit d' that $\mathcal{A}'$ outputs as well. Game 2 is perfectly simulated and therefore $\mathsf{Adv}^{\mathsf{ind\text{-}sid\text{-}cpa}}_{\Pi,\mathcal{A}'}(k, \mathcal{ID}) = \left|\Pr[\mathsf{Guess}_2] - \frac{1}{2}\right|$. □

A direct calculation from Fact 1 and Fact 2 gives the bound in the statement of the theorem. Regarding the term $\frac{q_E}{\#\mathcal{ID}}$, it is negligible in k if the size of $\mathcal{ID}$ is at least exponential in k, which is always the case (otherwise, as we have discussed in the introduction, selective and full security are polynomially equivalent).

□

4 Extension to HIBE and Applications

Hierarchical identity-based encryption (HIBE for short) [HL02,GS02,GH09] is an extension of identity-based encryption, where identities are organized in hierarchies: the secret key of an identity id_1 can be used to derive a valid secret key for any identity id_2, as long as id_1 is a prefix of id_2.

The definitions in Section 2 and the result in Section 3 can be adapted to the HIBE scenario. The main difference is in the description of Game 1, in the proof of Theorem 3: once the random identity $\mathsf{id}^\star \in \mathcal{ID}$ is chosen, the challenger aborts if the adversary makes a secret key query either for $\mathsf{id}^\star$ or for any prefix of $\mathsf{id}^\star$. If the HIBE scheme admits a hierarchy with ℓ levels, then the probability of an abort is bounded by $\frac{\ell \cdot q_E}{\#\mathcal{ID}}$.

Our positive result may be useful to simplify the proofs of some existing (H)IBE schemes that are adaptively secure. For instance, we leave as an open question to study whether it can help to get a simpler proof or scheme in the case of [DIP10]. Therein, De Caro *et al.* use the dual system encryption technique of Waters to obtain a fully secure and fully anonymous HIBE. The proof of security is established by means of a sequence of games. Very roughly, for those familiar with the methodology of Waters, except for the last three games, the only difference between one game and the next is the fact that the challenge ciphertext or some secret key is replaced with its semifunctional counterpart. De Caro *et al.* prove that no adversary can distinguish between Game_i and Game_{i+1}, provided some computational assumption (Assumption 1 in the paper) holds. Finally, in the game before the last, the challenge ciphertext is replaced by some string which is independent from the message and in the last game by a string which is independent from the challenge identity. The proof that no adversary can distinguish whether he is on one of these two games or in the previous one is based on two different computational assumptions (Assumptions 2 and 3 of the paper).

Using our approach, if one could argue independently that the scheme is IND-sID-CPA secure, full anonymity would imply full security. Potentially this could result in a scheme based on less computational assumptions — since probably one of assumptions 2 and 3 could be dropped or they could be recombined—, although arguably this would depend on the hypothesis needed to prove selective security.

The same argument could be applied if, for instance, the HIBE scheme of Boyen-Waters [BW06], which is only selectively anonymous and selectively semantically secure, could be proven anonymous against an adaptive adversary, for example using the new dual encryption techniques of Waters [Wat09].

5 Conclusion

This paper contains a theoretical study of the relations between selective and adaptive security properties for identity-based encryption schemes which enjoy at the same time some level of anonymity and semantic security.

The security analysis of the anonymous identity-based encryption schemes that exist in the literature seem to suggest that proving adaptive anonymity is as hard as proving adaptive semantic security. Indeed, either semantic security and anonymity are both proved in the selective model [BW06,BW07,SKOS09,Duc10,ABB10a] or they are both proved in the adaptive model [Gen06,CKRS09,DIP10]. This probably responds to the fact that similar challenges appear when proving full anonymity and full semantic security, namely the problem that the partition strategy (which is really useful in the selective case) is much harder to apply in the adaptive case.

Our study suggests that another approach to proving that a scheme is fully anonymous and fully secure is possible. Once adaptive anonymity is proved for a scheme, then semantic security can be proved in a selective scenario. We believe that these theoretical relations may have an impact in the design or in the analysis of anonymous (hierarchical) identity-based encryption schemes.

Finally, we emphasize that the symmetric is not true: if a scheme is proved to be semantically secure under adaptive attacks, then proving selective anonymity is not enough to ensure adaptive anonymity.

Acknowledgements

This work was started while Fabien Laguillaumie was visiting the Universitat Politècnica de Catalunya. The work of Javier Herranz is supported by a Ramón y Cajal grant, partially funded by the European Social Fund (ESF) of the Spanish MICINN Ministry. Carla Ràfols holds an FPI grant of the Spanish MICINN Ministry. The work of both these authors is partially supported by the Spanish MICINN Ministry under project MTM2009-07694. The work of Fabien Laguillaumie is supported by the French ANR-07- TCOM-013-04 PACE Project.

References

[A+08] M. Abdalla, M. Bellare, D. Catalano, E. Kiltz, T. Kohno, T. Lange, J. Malone-Lee, G. Neven, P. Paillier and H. Shi. Searchable encryption revisited: consistency properties, relation to anonymous IBE, and extensions. *J. Cryptology*, Vol. 21, No. 3, 350–391 (2008).

[ABB10a] S. Agrawal, D. Boneh and X. Boyen. Efficient lattice (H)IBE in the standard model. *Proceedings of Eurocrypt 2010*, Springer LNCS **6110**, 553–572 (2010).

[ABB10b] S. Agrawal, D. Boneh and Xavier Boyen. Lattice basis delegation in fixed dimension and shorter ciphertext hierarchical IBE. *Proceedings of Crypto 2010* Springer LNCS **6223**, 98–115 (2010).

[AG09] G. Ateniese and P. Gasti. Universally anonymous IBE based on the quadratic residuosity assumption. *Proceedings of CT-RSA 2009*, Springer LNCS **5473**, 32–47 (2009).

[BBDP01] M. Bellare, A. Boldyreva, A. Desai and D. Pointcheval. Key-privacy in public-key encryption. *Proceedings of Asiacrypt 2001*, Springer LNCS **2248**, 566–582 (2001).

[BB04a] D. Boneh and X. Boyen. Efficient selective-ID secure identity based encryption without random oracles. *Proceedings of Eurocrypt 2004*, Springer LNCS **3027**, 223–238 (2004).

[BB04b] D. Boneh and X. Boyen. Secure identity based encryption without random oracles. *Proceedings of Crypto 2004*, Springer LNCS **3152**, 443–459 (2004).

[BDOP04] D. Boneh, G. Di Crescenzo, R. Ostrovsky, and G. Persiano. Public key encryption with keyword search. *Proceedings of Eurocrypt 2004*, Springer LNCS **3027**, 506–522 (2004).

[BF03] D. Boneh and M. Franklin. Identity based encryption from the Weil pairing. *SIAM J. of Computing*, Vol. 32, No. 3, 586–615 (2003).

[BW06] X. Boyen and B. Waters. Anonymous hierarchical identity-based encryption (without random oracles). *Proceedings of Crypto 2006*, Springer LNCS **4117**, 290–307 (2006).

[BW07] D. Boneh and B. Waters. Conjunctive, subset, and range queries on encrypted data. *Proceedings of TCC 2007*, Springer LNCS **4392**, 535–554 (2007).

[CHKP10] D. Cash, D. Hofheinz, E. Kiltz and C. Peikert. Bonsai trees, or how to delegate a lattice basis. *Proceedings of Eurocrypt 2010*, Springer LNCS **6110**, 523–552 (2010).

[CKRS09] J. Camenisch, M. Kohlweiss, A. Rial and C. Sheedy. Blind and anonymous identity-based encryption and authorised private searches on public key encrypted data. *Proceedings of PKC 2009*, Spinger LNCS **5443**, 196-214 (2009).

[CS05] S. Chatterjee and P. Sarkar. Trading time for space: towards an efficient IBE scheme with short(er) public parameters in the standard model. *Proceedings of ICISC 2005*, Springer LNCS **3935**, 424–440 (2006).

[Coc01] C. Cocks. An identity based encryption scheme based on quadratic residues. *Proceedings of IMA Cryptography and Coding 2001*, Springer LNCS **2260**, 360–363 (2001).

[DIP10] A. De Caro, V. Iovino, and G. Persiano. Fully secure anonymous HIBE and secret-key anonymous IBE with short ciphertexts. To appear in *Proceedings of Pairing 2010*, Springer LNCS. Preliminary version available Cryptology ePrint Archive, Report 2010/197 (2010) http://eprint.iacr.org/.

[Duc10] L. Ducas. Anonymity from asymmetry: new constructions for anonymous HIBE. *Proceedings of CT-RSA 2010*, Springer LNCS **5985**, 148–164 (2010).

[Gal06] D. Galindo. A separation between selective and full-identity security notions for identity-based encryption. *Proceedings of ICCSA 2006*, Springer LNCS **3982**, 318–326 (2006).

[Gen06] C. Gentry. Practical identity-based encryption without random oracles. *Proceedings of Eurocrypt 2006*, Springer LNCS **4004**, 445–464 (2006).

[GH09] C. Gentry and S. Halevi. Hierarchical identity based encryption with polynomially many levels. *Proceedings of TCC 2009*, Springer LNCS **5444**, 437–456 (2009).

[GS02] C. Gentry and A. Silverberg. Hierarchical ID-based cryptography. *Proceedings of Asiacrypt 2002*, Springer LNCS **2501**, 548–566 (2002).

[Hal05] S. Halevi. A sufficient condition for key-privacy. Cryptology ePrint Archive, Report 2005/005 (2005). http://eprint.iacr.org/.

[HL02] J. Horwitz and B. Lynn. Toward hierarchical identity-based encryption. *Proceedings of Eurocrypt 2002*, Springer LNCS **2332**, 466–481 (2002).

[L+10] A. Lewko, T. Okamoto, A. Sahai, K. Takashima and B. Waters. Fully secure functional encryption: attribute-based encryption and (hierarchical) inner product encryption. *Proceedings of Eurocrypt 2010*, Springer LNCS **6110**, 62–91 (2010).

[LW10] A. B. Lewko and B. Waters. New techniques for dual system encryption and fully secure hibe with short ciphertexts. *Proceedings of TCC 2010*, Springer LNCS **5978**, 455–479 (2010).

[Nac05] D. Naccache. Secure and practical identity-based encryption. Cryptology ePrint Archive, Report 2005/369 (2005). http://eprint.iacr.org/.

[SKOS09] J. H. Seo, T. Kobayashi, M. Ohkubo, and K. Suzuki. Anonymous hierarchical identity-based encryption with constant size ciphertexts. *Proceedings of PKC 2009*, Spinger LNCS **5443**, 215–234 (2009).

[Sha85] A. Shamir. Identity-based cryptosystems and signature schemes. *Proceedings of Crypto 1984*. Springer LNCS **196**, 47–53 (1985).

[Wat05] B. Waters. Efficient identity-based encryption without random oracles. *Proceedings of Eurocrypt 2005*, Springer LNCS **3494**, 114–127 (2005).

[Wat09] B. Waters. Dual system encryption: realizing fully secure IBE and HIBE under simple assumptions. *Proceedings of Crypto 2009*, Springer LNCS **5677**, 619–636 (2009).