Boneh-Gentry-Hamburg's Identity-based Encryption Scheme Revisited

George Teşeleanu, Ferucio Laurenţiu Ţiplea, Sorin Iftene, Anca-Maria Nica

Department of Computer Science, "Al.I.Cuza" University of Iaşi
700506 Iaşi, Romania,
{george.teseleanu,ferucio.tiplea,siftene,anca.nica}@info.uaic.ro

Abstract. *BasicIBE* and *AnonIBE* are two space-efficient identity-based encryption (IBE) schemes based on quadratic residues, proposed by Boneh, Gentry, and Hamburg, and closely related to Cocks' IBE scheme. *BasicIBE* is secure in the random oracle model under the quadratic residuosity assumption, while *AnonIBE* is secure in the standard model under the interactive quadratic residuosity assumption. In this paper we revise the *BasicIBE* scheme and we show that if the requirements for the deterministic algorithms used to output encryption and decryption polynomials are slightly changed, then the scheme's security margin can be slightly improved.

1 Introduction

Identity-based cryptography was proposed in 1984 by Adi Shamir [5] who formulated its basic principles. The first *identity-based encryption* (IBE) scheme was proposed by Boneh and Franklin [2] and is based on bilinear maps. Shortly, Cocks [4] proposed another IBE scheme based on the standard quadratic residuosity problem[1] modulo an RSA composite n. Cocks' scheme encrypts a bit by two integers modulo n such that the bit is recovered as the Jacobi symbol of one of these two integers together with the private key. Although the scheme is very elegant and quite fast, its main disadvantage is ciphertext expansion: a bit of message requires $2 \log n$ bits of ciphertext. In [3], Boneh, Gentry, and Hamburg proposed two abstract IBE schemes with short ciphertexts that are related to Cocks' scheme. The first scheme, named *BasicIBE*, is IND-ID-CPA[1] secure in the random oracle model (ROM) under the quadratic residuosity assumption, while the second one, named *AnonIBE*, is ANON-IND-ID-CPA[1] secure in the standard model under the interactive quadratic residuosity assumption[1]. Both security results are obtained by providing upper bounds on the advantage of an efficient adversary against the corresponding IBE scheme.

In order to provide a tighter upper bound for the *BasicIBE* scheme we slightly change the requirements for the deterministic algorithms $\mathcal{Q}$ used to output encryption and decryption polynomials. The concrete instantiation of $\mathcal{Q}$ provided in [3] satisfies the new set of restrictions. Thus, without changing the instantiation of *BasicIBE* we obtain a marginally better security margin.

Structure of the paper. We introduce notations and definitions used throughout the paper in Section 2. In Section 3 we describe the *BasicIBE* scheme and we provide the security margin proved in [3]. We reassess *BasicIBE*'s security proof in Section 4. We conclude in Section 5.

2 Preliminaries

Notations. Throughout the paper, λ will denote a security parameter. The action of selecting a random element x from a sample space X is denoted by $x \xleftarrow{\$} X$. We denote by $x \leftarrow y$ the assignment of value y to variable x. The probability that event E happens is denoted by $Pr[E]$. The Jacobi symbol of an integer a modulo an integer n is denoted by $\left(\frac{a}{n}\right)$. J_n stands for the set of integers in $\mathbb{Z}_n^*$ whose Jacobi symbol is 1, QR_n denotes the set of quadratic residues in $\mathbb{Z}_n^*$ and $SQRT_n(a)$ is the set of square roots of a modulo

[1] We refer the reader to Section 2 for a definition of the concept.

n. $\mathbb{Z}_n[x]$ is the ring of polynomials over $\mathbb{Z}_n$. We denote by PPT algorithm a probabilistic polynomial-time algorithm. By $\text{RSAGen}(\lambda)$, we understand a PPT algorithm that generates two equal size primes p and q larger then 2^λ.

2.1 Security Assumptions

Definition 1 (Pseudorandom Function - PRF). *A function $F : \{0,1\}^n \times \{0,1\}^\lambda \to \{0,1\}^m$ is a PRF if:*

- *Given a key $K \in \{0,1\}^\lambda$ and an input $X \in \{0,1\}^n$ there is an efficient algorithm to compute $F_K(X) = F(X,K)$.*
- *For any algorithm A, the PRF-advantage of A, defined as*

$$\text{PRFAdv}_{A,F}(\lambda) = \left| Pr[A^{F_K(\cdot)} = 1 | K \xleftarrow{\$} \{0,1\}^\lambda] - Pr[A^{F(\cdot)} = 1 | F \xleftarrow{\$} \mathcal{F}] \right|$$

is negligible for any PPT algorithm A, where $\mathcal{F} = \{F : \{0,1\}^n \to \{0,1\}^m\}$.

Definition 2 (Quadratic Residuosity - QR). *Let $n = pq$, where $(p,q) \leftarrow \text{RSAGen}(\lambda)$ and let A be a PPT algorithm which returns 1 on input (x,n) if x is a quadratic residue modulo n. We define the advantage*

$$\text{QRAdv}_{A,\text{RSAGen}}(\lambda) = \left| Pr[A(x,n) = 1 | x \xleftarrow{\$} QR_n] - Pr[A(x,n) = 1 | x \xleftarrow{\$} J_n \setminus QR_n] \right|$$

The Quadratic Residuosity assumption states that for any efficient PPT algorithm A the advantage $\text{QRAdv}_{A,\text{RSAGen}}(\lambda)$ is negligible.

2.2 Identity-based encryption

An IBE scheme consists of four PPT algorithms: *Setup*, *Extract*, *Encrypt*, and *Decrypt*. The first one takes as input a security parameter and outputs the system public parameters together with a master key. The *Extract* algorithm takes as input an identity ID together with the public parameters and the master key and outputs a private key associated to ID. The *Encrypt* algorithm, starting with a message m, an identity ID, and the public parameters, encrypts m into some ciphertext c (the encryption key is ID or some binary string derived from ID). The last algorithm decrypts c into m by using the private key associated to ID.

Definition 3 (Anonymity and Indistinguishability under Selective Identity and Chosen Plaintext Attacks - ANON-IND-ID-CPA). *The ANON-IND-ID-CPA security of an IBE scheme $\mathcal{S}$ is formulated by means of the following game between a challenger C and an adversary A:*

Setup(λ): The challenger C generates the public parameters PP and sends them to adversary A, while keeping the master key msk to himself.

Queries: The adversary issues a finite number of adaptive queries. A query can be one of the following types:

- *Private key query. When A requests a query for an identity, the challenger runs the $Extract$ algorithm and returns the resulting private key to A.*
- *Encryption query. Adversary A can issue only one query of this type. He sends C two pairs (ID_0, m_0) and (ID_1, m_1) consisting of two equal length plaintexts m_0 and m_1 and two identities ID_0 and ID_1. The challenger flips a coin $b \in \{0,1\}$ and encrypts m_b using ID_b. The resulting ciphertext c is sent to the adversary. The following restrictions are in place: private key queries for ID_0 and ID_1 must never be issued.*

Guess: In this phase, the adversary outputs a guess $b' \in \{0,1\}$. He wins the game, if $b' = b$.

The advantage of an adversary A attacking an IBE scheme is defined as

$$\text{IBEAdv}_{A,\mathcal{S}}(\lambda) = |Pr[b = b'] - 1/2|$$

where the probability is computed over the random bits used by C and A. An IBE scheme is ANON-IND-ID-CPA secure, if for any PPT adversary A the advantage $\text{IBEAdv}_{A,\mathcal{S}}(\lambda)$ is negligible. If we consider $ID_0 = ID_1$ in the above game, we obtain the concept of IND-ID-CPA security.

3 Boneh-Gentry-Hamburg's *BasicIBE* Scheme

Let n be an RSA composite and r the private key of the recipient. Cocks' IBE scheme [4] encrypts a bit $m \in \{\pm 1\}$ by two integers c_1 and c_2 such that either the Jacobi symbol of $(c_1 + 2r)$ or the Jacobi symbol of $(c_2 + 2r)$ modulo n is m. The scheme is IND-ID-CPA secure in ROM under the QR assumption.

Despite its elegance, Cocks' scheme produces large ciphertext: $2 \log n$ bits are used to encrypt just one bit. Moreover, it is not anonymous [1]. In 2007, Boneh, Gentry and Hamburg proposed two space efficient IBE schemes related to Cocks' scheme, whose security is also based on the QR problem [3]. Additionally, one of the schemes is anonymous. Below we describe the *BasicIBE* scheme.

Setup(λ): Let $n = pq$, where $(p, q) \leftarrow \text{RSAGen}(\lambda)$. Generate $u \in J_n \setminus QR_n$, and choose a hash function $h : \{0,1\}^* \times \{1, \ldots, \ell\} \to J_n$, for some integer $\ell \geq 1$. Choose a deterministic algorithm $\mathcal{Q}$, that takes as input n and any $R, S \in \mathbb{Z}_n^*$ and outputs two polynomials $f, g \in \mathbb{Z}_n[x]$. Output the public parameters $PP = (n, u, h, \ell, \mathcal{Q})$. The master key is $msk = (p, q, K)$, where K is a random key for a pseudorandom function $F_K : \{0,1\}^* \times \{1, \ldots, \ell\} \to \{0, 1, 2, 3\}$.

Extract(msk, ID, ℓ): For each $j \in \{1, \ldots, \ell\}$, let $R_j = h(ID, j)$ and $i_j = F_K(ID, j)$. If r_0, r_1, r_2, r_3 is a fixed total ordering of the square roots of R_j or uR_j (depending on which of them is a quadratic residue), then the private key is $r = (r_{i_1}, \ldots, r_{i_\ell})$.

Encrypt(PP, ID, m): Let $m = m_1 \ldots m_\ell \in \{\pm 1\}^\ell$ be an ℓ-bit sequence. Generate at random $s \in \mathbb{Z}_n^*$ and set $S = s^2 \bmod n$. For each $j \in \{1, \ldots, \ell\}$, let $R_j = h(ID, j)$. Use algorithm $\mathcal{Q}$ to compute the polynomials $(f_j, g_j) \leftarrow \mathcal{Q}(n, R_j, S)$ and $(\bar{f}_j, \bar{g}_j) \leftarrow \mathcal{Q}(n, uR_j, S)$. Set $c_j \leftarrow m_j \cdot \left(\frac{g_j(s)}{n}\right)$ and $\bar{c}_j \leftarrow m_j \cdot \left(\frac{\bar{g}_j(s)}{n}\right)$. Output $(c, \bar{c}, S)$, where $c = c_1 \cdots c_\ell$ and $\bar{c} = \bar{c}_1 \cdots \bar{c}_\ell$.

Decrypt$(c, \bar{c}, S, r, PP)$: For each $j \in \{1, \ldots, \ell\}$, let $R_j = h(ID, j)$. If $R_j = r_j^2$, then $(f'_j, g'_j) = \mathcal{Q}(n, R_j, S)$, else $(f'_j, g'_j) = \mathcal{Q}(n, uR_j, S)$. Compute $m_j = c_j \cdot \left(\frac{f'_j(r_{i_j})}{n}\right)$. Output the message $m = m_1 \cdots m_\ell$.

Before stating the security result from [3], we first need to introduce some properties that need to be satisfied by $\mathcal{Q}$.

Definition 4 (IBE Compatible). *Let n be a positive integer and $R, S \in \mathbb{Z}_n^*$. Let $\mathcal{Q}$ be a deterministic algorithm that takes as input n, R, S and outputs two polynomials $f, g \in \mathbb{Z}_n[x]$. Algorithm $\mathcal{Q}$ is IBE Compatible if the following conditions are satisfied*

1. *If $R, S \in QR_n$, then $f(r)g(s) \in QR_n$, where $r \in SQRT_n(R)$ and $s \in SQRT_n(S)$.*
2. *If $R \in QR_n$ and $S \in J_n \setminus QR_n$, then $\left(\frac{f(r)}{n}\right)$ is uniformly distributed in $\{\pm 1\}$, where $r \xleftarrow{\$} SQRT_n(R)$.*

Definition 4 is introduced in [3] in another form. If $R \in QR_n$, the authors require that $f(r)f(-r)S \in QR_n$ for all $r \in SQRT_n(R)$. In [3, Lemma 3.3] it is proven that if $r \xleftarrow{\$} SQRT_n(R)$ and $S \in J_n \setminus QR_n$, then $f(r)f(-r)S \in QR_n$ implies Definition 4, Condition 2.

In the security proof of *BasicIBE*, Boneh, Gentry and Hamburg [3] use that $\left(\frac{f(r)}{n}\right)$ is uniformly distributed in $\{\pm 1\}$ and not that $f(r)f(-r)S \in QR_n$. Thus, Definition 4, Condition 2 captures the exact security requirement for $\mathcal{Q}$.

If $\mathcal{Q}$ is IBE compatible, then Definition 4, Condition 1 guarantees the soundness of decryption. As with respect to security, the following result is proven in [3].

Theorem 1. *If the QR assumption holds and $\mathcal{Q}$ is IBE compatible, then the BasicIBE scheme is IND-ID-CPA secure in ROM. Formally, for any efficient PPT adversary A there exist efficient PPT algorithms B_1 and B_2 such that*

$$\text{IBEAdv}_{A,BasicIBE}(\lambda) \leq \text{PRFAdv}_{B_1,F}(\lambda) + 2\text{QRAdv}_{B_2,\text{RSAGen}}(\lambda).$$

A few words are in place about Theorem 1. The proof of this theorem as it is in [3], exploits the fact that h outputs truly random elements from J_n and replaces h with $H(ID, j) = u^{a_j} v_j^2$, where $a_j \xleftarrow{\$} \{0, 1\}$, $v_j \xleftarrow{\$} \mathbb{Z}_n^*$. The initial IND-ID-CPA game is successively changed into another game where the challenge ciphertext is created by decrypting the message (that is, by encrypting it by f's instead of g's). In order to have $R_j, uR_j \in QR_n$ and $S \in J_n \setminus QR_n$, the QR assumption is used two times, which gives rise to the factor $2\mathrm{QRAdv}_{B_2,\mathrm{RSAGen}}(\lambda)$. Moreover, to ensure that r_j is a random square root of R_j or uR_j, F_K is replaced by a truly random function, and this gives rise to the factor $\mathrm{PRFAdv}_{B_1,F}(\lambda)$.

Remark 1. We emphasize that the *BasicIBE* scheme is an abstract IBE scheme because no concrete algorithm $\mathcal{Q}$ is presented. In [3], the method proposed to construct the polynomials f and g is based on the congruence given by

$$Rx^2 + Sy^2 \equiv 1 \bmod n,$$

where n is an integer and $R, S \in \mathbb{Z}_n^*$.

Any solution (x_0, y_0) to the above equation gives rise to two polynomials

$$f(r) = x_0 r + 1 \bmod n \qquad g(s) = 2(y_0 s + 1) \bmod n$$

that satisfy the conditions from Definition 4.

If we instantiate $\mathcal{Q}$ as in Remark 1, then the encryptor must find solutions to 2ℓ congruences, while the decryptor must find solutions to ℓ congruences. Boneh, Gentry, and Hamburg [3] have proposed the following *Combining Lemma* in order to reduce the number of congruences to be solved by the encryptor.

Lemma 1. *If (x_1, y_1) is a solution to the congruence $R_1x^2 + Sy^2 \equiv 1 \bmod n$ and (x_2, y_2) is a solution to the congruence $R_2x^2 + Sy^2 \equiv 1 \bmod n$, then $(x_{1,2}, y_{1,2})$ is a solution to the congruence $R_1R_2x^2 + Sy^2 \equiv 1 \bmod n$, where*

$$x_{1,2} = \frac{x_1x_2}{Sy_1y_2 + 1} \bmod n \quad \textit{and} \quad y_{1,2} = \frac{y_1 + y_2}{Sy_1y_2 + 1} \bmod n,$$

provided that $(Sy_1y_2 + 1, n) = 1$.

Using this result, the encryptor first finds solutions to $ux^2 + Sy^2 \equiv 1 \bmod n$ and $R_jx^2 + Sy^2 \equiv 1 \bmod n$, for all $1 \le j \le \ell$, and then combines them to obtain solutions to $uR_jx^2 + Sy^2 \equiv 1 \bmod n$, for all $1 \le j \le \ell$. Therefore, the encryptor needs to find solutions to $\ell + 1$ congruences, instead of 2ℓ. It is to be remarked that, the proposed method does not affect the security of the scheme.

4 A New Security Analysis for *BasicIBE*

In [3], the authors only require $\mathcal{Q}$ to be IBE compatible, but the algorithm proposed by them satisfies one more condition, that is captured in the following definition. It is easy to see that g satisfies the Condition 3, since $Rx^2 + Sy^2 = 1$ is symmetric.

Definition 5 (Extended IBE Compatible). *Let n be a positive integer and $R, S \in \mathbb{Z}_n^*$. Let $\mathcal{Q}$ be a deterministic algorithm that takes as input n, R, S and outputs two polynomials $f, g \in \mathbb{Z}_n[x]$. Algorithm $\mathcal{Q}$ is Extended IBE Compatible if it is IBE compatible and the following condition is satisfied*

3. *If $R \in J_n \setminus QR_n$ and $S \in QR_n$, then $\left(\frac{g(s)}{n}\right)$ is uniformly distributed in $\{\pm 1\}$, where $s \xleftarrow{\$} SQRT_n(S)$.*

Using the extra property satisfied by $\mathcal{Q}$[2], we slightly modify the security proof of Theorem 1 and we obtain a tighter upper bound.

[2] captured in Definition 5

Theorem 2. *If the* QR *assumption holds and $\mathcal{Q}$ is extended IBE compatible, then the BasicIBE scheme is* IND-ID-CPA *secure in ROM. Formally, for any efficient PPT adversary A there exist efficient PPT algorithms B_1 and B_2 such that*

$$\text{IBEAdv}_{A,BasicIBE}(\lambda) \leq \text{PRFAdv}_{B_1,F}(\lambda) + \text{QRAdv}_{B_2,\text{RSAGen}}(\lambda).$$

Proof. Let A be an IND-ID-CPA adversary for $BasicIBE$. We prove that his advantage is negligible. We present the proof as a sequence of games. Let W_i be the event that A wins game i.

Game 0. The first game is identical to the IND-ID-CPA game[3]. Thus, we have

$$|P[W_0] - 1/2| = \text{IBEAdv}_{A,BasicIBE}(\lambda) \tag{1}$$

Game 1. In this game, F_K is replaced by a truly random function $f : \{0,1\}^* \times \{1,\ldots,\ell\} \to \{0,1,2,3\}$. Adversary A will not notice the difference, since F_K is a pseudorandom function. Formally, there exists an algorithm B_1 such that

$$|Pr[W_0] - Pr[W_1]| = \text{PRFAdv}_{B_1,F}(\lambda). \tag{2}$$

Game 2. We slightly modify how the challenger answers hash queries. Thus, for a query $h(ID, j)$ it first chooses $a_j \xleftarrow{\$} \{0,1\}$ and $v_j \xleftarrow{\$} \mathbb{Z}_n^*$. Then, it outputs $h(ID, j) = u^{a_j} v_j^2$. It is easy to see that the challenger implements a random function.

Let $R_j = h(ID, j)$ for some (ID, j). The challenger also has to answer private key extraction queries. Thus, when an extraction query for ID is received, the challenger answers with $R_j^{1/2} = v_j$ if $a_j = 0$ or $(uR_j)^{1/2} = uv_j$ if $a_j = 1$, where $1 \leq j \leq \ell$. Since a_j and v_j are random elements, the challenger outputs a random square root of either R_j or uR_j. Thus, *Game 1* and *Game 2* are identical from A's point of view. Note that in this case the challenger can answer private key queries without knowing the factorization of n. Formally, we have

$$Pr[W_1] = Pr[W_2]. \tag{3}$$

Game 3. By maintaining a list with all the hash queries, the challenger can decide if $R_j \in QR_n$ or $uR_j \in QR_n$. We change the encryption algorithm as follows

- If $R_j \in QR_n$, then $c_j \leftarrow m_j \cdot \left(\frac{f_j(r_{i_j})}{n}\right)$ and $\bar{c}_j \xleftarrow{\$} \{\pm 1\}$;
- Else $c_j \xleftarrow{\$} \{\pm 1\}$ and $\bar{c}_j \leftarrow m_j \cdot \left(\frac{\bar{f}_j(r_{i_j})}{n}\right)$.

We will show that the ciphertext has the same distribution as in *Game 2*. First, recall that $S \in QR_n$.

If $R_j \in QR_n$, then according to Definition 4, Condition 1 c_j has the same value as in *Game 2*. Since $uR_j \in J_n \setminus QR_n$, Definition 5, Condition 3 assures us that $\left(\frac{\bar{g}_j(s)}{n}\right)$ is uniformly distributed in $\{\pm 1\}$. Thus $\bar{c}_j$ chosen in this game has the same distribution as in *Game 2*.

A similar discussion for the case $R_j \in J_n \setminus QR_n$ shows that c_j and $\bar{c}_j$ have the same distribution as in *Game 2*.

Since these are the only changes between *Game 2* and *Game 3*, A will not notice the difference assuming $\mathcal{Q}$ is extended IBE compatible. Formally, this means that

$$Pr[W_2] = Pr[W_3]. \tag{4}$$

Game 4. In this game the challenger chooses $S \in J_n \setminus QR_n$. Since this is the only change between *Game 3* and *Game 4*, A will not notice the difference assuming the QR assumption holds. Formally, this means that there exists an algorithm B_2 such that

$$|Pr[W_3] - Pr[W_4]| = \text{QRAdv}_{B_2,\text{RSAGen}}(\lambda). \tag{5}$$

[3] as in Definition 3

Game 5. To make the challenge ciphertext independent of the challenge bit b, we slightly change *Game 4*. Thus, the challenger chooses $c_j \xleftarrow{\$} \{\pm 1\}$, for $j \in \{1, \ldots, \ell\}$.

We will show that *Game 4* and *Game 5* are identical. To do that, we prove that the ciphertext distribution remains the same as in *Game 4*.

Due to the change made in *Game 4*, we have that $S \in J_n \setminus QR_n$. If $R_j \in QR_n$, then according to Definition 4, Condition 2 $\left(\frac{f_j(r_{i_j})}{n}\right)$ is uniformly distributed in $\{\pm 1\}$. Thus, c_j has the same distribution in both games. The case $uR_j \in QR_n$ is treated in a similar way.

Since these are the only changes between *Game 4* and *Game 5*, A will not notice the difference assuming $\mathcal{Q}$ is extended IBE compatible. Formally, this means that

$$Pr[W_4] = Pr[W_5]. \tag{6}$$

In this game, c and $\bar{c}$ are independent of the challenge bit. Thus, we have

$$Pr[W_5] = 1/2. \tag{7}$$

Finally, the statement is proven by combining the equalities (1) – (7). □

Remark 2. By tweaking the proof of Theorem 2, we obtain the same upper bound for the ANON-IND-ID-CPA security[4] of the $AnonIBE$ scheme. Note that the deterministic algorithm used by the $AnonIBE$ scheme also satisfies the extra property stated in Definition 5.

5 Conclusions

Boneh, Gentry, and Hamburg have proposed in [3] two IBE schemes related to Cocks' IBE scheme, called $BasicIBE$ and $AnonIBE$. Compared to $BasicIBE$ and Cocks' IBE scheme, $AnonIBE$ also provides anonymity of identity. These two schemes are more space efficient than Cocks' IBE scheme, but the concrete instantiations are less time efficient.

In this paper we have revisited the $BasicIBE$ abstract scheme, by slightly modifying the requirements for the deterministic algorithm. These requirements are already fulfilled by the concrete instantiation proposed in [3] (however, the authors of [3] did not use them). By using a different proof approach, we managed to obtain a tighter security margin for $BasicIBE$.

References

1. Giuseppe Ateniese and Paolo Gasti. Universally Anonymous IBE Based on the Quadratic Residuosity Assumption. In *CT-RSA 2009*, volume 5473 of *Lecture Notes in Computer Science*, pages 32–47. Springer, 2009.
2. Dan Boneh and Matt Franklin. Identity-Based Encryption from the Weil Pairing. In *CRYPTO 2001*, volume 2139 of *Lecture Notes in Computer Science*, pages 213–229. Springer, 2001.
3. Dan Boneh, Craig Gentry, and Michael Hamburg. Space-Efficient Identity Based Encryption Without Pairings. In *FOCS 2007*, pages 647–657. IEEE, 2007.
4. Clifford Cocks. An Identity Based Encryption Scheme Based on Quadratic Residues. In *IMA 2001*, volume 2260 of *Lecture Notes in Computer Science*, pages 360–363. Springer, 2001.
5. Adi Shamir. Identity-Based Cryptosystems and Signature Schemes. In *CRYPTO 1984*, volume 196 of *Lecture Notes in Computer Science*, pages 47–53. Springer, 1985.

[4] in ROM