Efficient Identity-Based Encryption Without Random Oracles Preliminary Version

Brent R. Waters

Abstract

We present the first efficient Identity-Based Encryption (IBE) scheme that is secure in the full model without random oracles. We first present our IBE construction and reduce the security of our scheme to the Decisional Bilinear Diffie-Hellman problem.

1 Introduction

Identity-Based Encryption allows for a party to encrypt a message using the recipient's identity as a public key. The ability to use identities as public keys avoids the need to distribute public key certificates. This can be very useful in applications such as email where the recipient is often off-line when the message has been sent.

The first efficient and secure method for Identity-Based Encryption was put forth by Boneh and Franklin [3]. They proposed a solution using efficiently computable bilinear maps that was shown to be secure in the Random Oracle model. Since then, there have been schemes shown to be secure without random oracles, but in a weaker model of security know as the Selective-ID model [4, 1]. Most recently, Boneh and Boyen [2] described a scheme that was proved to be secure in full model without random oracles; the possibility of such a scheme was to that point an open problem. However, their scheme is too inefficient to be of practical use.

We present the first efficient Identity-Based Encryption scheme that is secure in the full model without random oracles. The proof of our scheme makes use of an algebraic method first used by Boneh and Boyen [1] and the security of our scheme reduces to the Bilinear Decisional Diffie-Hellman assumption.

We then describe our Forward-Secure Signature scheme; our scheme makes use of techniques we used for IBE along with a Binary Encryption Tree as described in Canetti et. al. [4].

1.1 Related Work

Shamir [9] first presented the idea of Identity-Based Encryption as a challenge to the research community. However, the first secure and efficient scheme by Boneh and Franklin[3] did not appear until much later. The authors took a novel approach in using efficiently computable bilinear maps in order to achieve their result.

Canetti *et. al.* [4] described a somewhat weaker model of security for Identity-Based Encryption that they labeled the *Selective-ID* model. In the Selective-ID model the adversary must first declare which ciphertext it wishes to be challenged on before the global parameters are generated. The authors proceed to described a secure scheme under this model without random oracles. Boneh and Boyen [1] improved upon this result by describing an efficient scheme in the Selective-ID model.

Finally, Boneh and Boyen [2] describe a scheme that is fully secure without random oracles. One property of their security reduction is that in the simulator they construct there exist specific bits, which if set in the challenge identity result in the simulator needing to abort. This results in the need to apply an error-correcting code to identities in order to guarantee an adequate Hamming separation between different identities. In contrast, in our reduction there are no specific identity bits that correspond to an identity being able to serve as a challenge. Therefore, we are able to avoid applying an error-correcting code to identities and can construct an efficient scheme.

1.2 Organization

We organize the rest of the paper as follows. In Section 2 we give the definitions of IBE definition of security. In Section 3 we describe the number theoretic assumptions we will be using. In Section 4 we present the construction of our IBE scheme and follow with a proof of security in Section 5. Finally, we conclude in Section 7.

2 Security Definitions

In this section we present the definition for a semantically secure Identity-Based Encryption. This definition was first described by Boneh and Franklin [3].

Consider the following game played by an adversary. The game has four distinct phases which are:

Setup The challenger generates the master public parameters and gives them to the adversary.

Phase 1 The adversary is allowed to make multiple queries for a private key, v, where v is an identity specified by the adversary. The adversary can repeat this multiple times for different identities.

Challenge The adversary submits a public key, v^* , and two messages M_0 and M_1 . The adversaries choice of v^* is restricted to the identities that he did not request a private key for in Phase 1. The challenger flips a fair binary $coin, \gamma$, and returns an encryption of M_{γ} under the public key v^* .

Phase 2 Phase 1 is repeated with the restriction that the adversary cannot request the private key for v^* .

Guess The adversary submits a guess, γ' , of γ .

Definition 1 (IBE Semantic Security). An Identity-Based Encryption scheme is semantically secure if for all computationally bounded adversaries, \mathcal{A} , \mathcal{A} has a most a negligible advantage in succeeding in the above game.

3 Complexity Assumptions

Let \mathcal{G} be a group of prime order p with an admissible bilinear map, e, into \mathcal{G}_1 and g be a generator of \mathcal{G} .

3.1 Decisional Bilinear Diffie-Hellman (BLDH) Assumption

The challenger chooses $a, b, c, z \in \mathbb{Z}_p$ at random and then it flips a fair binary coin β . If $\beta = 1$ it outputs the tuple $(g, A = g^a, B = g^b, C = g^c, Z = e(g, g)^{abc})$. Otherwise, if $\beta = 0$, the challenger outputs the tuple $(g, A = g^a, B = g^b, C = g^c, Z = e(g, g)^z)$. The adversary must then output a guess β' of β .

The assumption is that no computationally bounded adversary can win the game with more than a negligible advantage.

4 Construction

Let \mathcal{G} be a group of prime order, p, for which there exists an efficiently computable bilinear map into \mathcal{G}_1 . Additionally, let $e: \mathcal{G} \times \mathcal{G} \to \mathcal{G}_1$ denote the bilinear map and g be the corresponding generator. The size of the group is determined by the security parameter. Identities will be represented as bitstrings of length n. We can also let identities be bitstrings of arbitrary length and n be the output length of a collision-resistant hash function, $H: \{0, 1\}^* \to \{0, 1\}^n$.

We now describe the construction of our scheme.

Setup The system parameters are generated as follows. A secret $\alpha \in \mathbb{Z}_p$ is chosen at random. We set the value $g_1 = g^{\alpha}$ and choose g_2 randomly in \mathcal{G} . Additionally, the simulator chooses a random value $u' \in \mathcal{G}$ and a random *n*-length vector $U = (u_i)$, whose elements are chosen at random from \mathcal{G} . The published public parameters are g_1, g_2, u' , and U. The master secret is g_2^{α} .

Key Generation Let v be an n bit string representing an identity and let v_i denote the *ith* bit of v and let $\mathcal{V} \subseteq \{0, \ldots, n\}$ be the set of all i for which $v_i = 1$. (That is \mathcal{V} is the set of indicies for which the bitstring v is set to 1.) A secret key for v is generated as follows. First, a random $r \in \mathbb{Z}_p$ is chosen. Then the private key is constructed as:

$$d_v = \left(g_2^{\alpha}\left(u'\prod_{i\in\mathcal{V}}u_i\right)^r, g^r\right).$$

Encryption A message $M \in \mathcal{G}_1$ is encrypted for an identity v as follows. A value $t \in \mathbb{Z}_p$ is chosen at random. The ciphertext is then constructed as

$$C = \left(e(g_1, g_2)^t M, g^t, \left(u' \prod_{i \in \mathcal{V}} u_i \right)^t \right).$$

Decryption Let $C = (c_1, c_2, c_3)$ be a valid encryption of M under the identity v. Then C can be decrypted by $d_v = (d_1, d_2)$ as:

$$c_{1}\frac{e(d_{2},c_{3})}{e(d_{1},c_{2})} = \left(e(g_{1},g_{2})^{t}M\right)\frac{e(g^{r},\left(u'\prod_{i\in\mathcal{V}}u_{i}\right)^{t})}{e(g_{2}^{\alpha}\left(u'\prod_{i\in\mathcal{V}}u_{i}\right)^{r},g^{t})} = \left(e(g_{1},g_{2})^{t}M\right)\frac{e(g,\left(u'\prod_{i\in\mathcal{V}}u_{i}\right)^{rt})}{e(g_{1},g_{2})^{t}e(\left(u'\prod_{i\in\mathcal{V}}u_{i}\right)^{rt},g)} = M$$

4.1 Efficiency

If the value of $e(g_1, g_2)$ is cached then encryption requires on average $\frac{n}{2}$ (and at most n) group operations in \mathcal{G} , two exponentiations in \mathcal{G} , one exponentiation in \mathcal{G}_1 , and one group operation in \mathcal{G}_1 .

Decryption requires two bilinear map computations, one group operation in \mathcal{G}_1 and one inversion in \mathcal{G}_1 .

5 IBE Security

We now prove the security of our scheme.

Theorem 1. Suppose there is an adversary \mathcal{A} that can break our scheme with advantage ϵ and makes at most q queries. Then we can construct a simulator that breaks the BLDDH game with advantage $\frac{\epsilon}{14nq}$.

Proof. We construct the simulator in the following manner. The simulator will take BLDDH challenge 5-tuple (g, A, B, C, Z) and outputs a guess, β' , as to whether the challenge is a BLDDH 5-tuple. The simulator runs \mathcal{A} executing the following steps.

Setup The simulator first sets an integer, m = 6q, and chooses an integer, k, uniformly at random between 0 and n. It then chooses a random n-length vector, $X = (x_i)$, where the elements of X are chosen uniformly at random from the integers between 0 and m - 1 and a value, x', chosen uniformly at random between 0 and m - 1. Additionally, the simulator chooses a random $y' \in \mathbb{Z}_p$ and an n-length vector, $Y = (y_i)$, where the elements of Y are chosen at random in \mathbb{Z}_p . These values are all kept internal to the simulator.

Again, for an identity v we will let $\mathcal{V} \subseteq \{0, \ldots, n\}$ be the set of all i for which $v_i = 1$ First, we define $F(v) = (p - mk) + x' + \sum_{i \in \mathcal{V}} x_i$. Next, we define $J(v) = y' + \sum_{i \in \mathcal{V}} y_i$. Finally, we define a binary function K(v) as

$$K(v) = \begin{cases} 0, & \text{if } x' + \sum_{i \in \mathcal{V}} x_i \equiv 0 \pmod{m} \\ 1, & \text{otherwise.} \end{cases}$$

The simulator assigns $g_1 = A$ and $g_2 = B$. It then assigns $u' = g_2^{p-km+x'}g^{y'}$ and $u_i = g_2^{x_i}g^{y_i}$. We note that from the perspective of the adversary all of these values all chosen uniformly at random as in the scheme described in the previous section.

Phase 1 The adversary, \mathcal{A} , will issue private key queries. Suppose the adversary issues a query for identity v. We first check if K(v) = 0. If this is the cases the simulator aborts the simulation and randomly outputs its guess β' .

Otherwise, the simulator chooses a random $r \in \mathbb{Z}_p$. Using the algebraic methods first described by Boneh and Boyen [1] it constructs the private key, d, as

$$d = (d_0, d_1) = \left(g_1^{\frac{-J(v)}{F(v)}} (u' \prod_{i \in \mathcal{V}} u_i)^r, g_1^{\frac{-1}{F(v)}} g^r\right).$$

 $T(\alpha)$

Let $\tilde{r} = r - \frac{a}{F(v)}$. Then we have

$$\begin{array}{rcl} d_{0} & = & g_{1}^{\frac{-J(v)}{F(v)}}(u'\prod_{i\in v}u_{i})^{r} \\ & = & g_{1}^{\frac{-J(v)}{F(v)}}(g_{2}^{F(v)}g^{J(v)})^{r} \\ & = & g_{2}^{a}(g_{2}^{F(v)}g^{J(v)})^{-\frac{a}{F(v)}}(g_{2}^{F(v)}g^{J(v)})^{r} \\ & = & g_{2}^{a}(u'\prod_{i\in \mathcal{V}}u_{i})^{r-\frac{a}{F(v)}} \\ & = & g_{2}^{a}(u'\prod_{i\in \mathcal{V}}u_{i})^{\tilde{r}}. \end{array}$$

Aditionally, we have

$$d_1 = g_1^{\frac{-1}{F(v)}}g'_1$$
$$= g^{r-\frac{a}{F(v)}}$$
$$= g^{\tilde{r}}.$$

This simulator will be able to perform this computation iff $F(v) \neq 0 \mod p$. It is sufficient to be in the non abort case where $K(v) \neq 0$. (We can assume p > nm for any reasonable values of p, n, and m. Therefore, $F(v) \equiv 0 \mod p$ only if F(v) = p and we can restrict ourselves to testing the sufficient condition that $K(v) \neq 0$.)

Challenge The adversary next will submit two messages $M_0, M_1 \in \mathcal{G}_1$ and an identity, v. If $x' + \sum_{i \in \mathcal{V}} x_i \neq km$ the simulator will abort and submit a random guess for β' . Otherwise, we have $F(v) \equiv 0 \pmod{p}$ and the simulator will flip a fair coin, γ , and construct the ciphertext

$$T = (ZM_{\gamma}, C, C^{J(v)}).$$

Suppose that the simulator was given a BLDDH tuple, that is $Z = e(g, g)^{abc}$. Then we have

$$T = \left(e(g,g)^{abc}M_{\gamma}, g^c, g^{cJ(v)}\right) = \left(e(g_1,g_2)^c M_{\gamma}, g^c, (u'\prod_{i\in\mathcal{V}}u_i)^c\right).$$

We see that T is a valid encryption of M_{γ} .

Otherwise, we have that Z is a random element of \mathcal{G} . In that case the ciphertext will give no information about the simulator's choice in γ .

Phase 2 The simulator repeats the same method it used in Phase 1.

Guess Finally, the adversary \mathcal{A} outputs a guess γ' of γ . If $\gamma' = \gamma$ then the simulator outputs a guess of $\beta' = 1$, otherwise it outputs $\beta' = 0$.

We let **abort** denote the event that the simulator aborted. If the simulator aborts it will take a random guess, β' at β . Therefore, we have that $\Pr[\beta' = \beta | \texttt{abort}] = \frac{1}{2}$.

Suppose now that the simulator did not abort. If $\beta = 0$, that is Z was randomly chosen member of \mathcal{G} , then the adversary will guess γ correctly with probability of $\frac{1}{2}$. Therefore, we have $\Pr[\beta' = \beta | \overline{\mathtt{abort}} \land \beta = 0] = \frac{1}{2}$.

If the simulator did not abort and it was given a valid tuple ($\beta = 1$) then it simulates the experiment perfectly and \mathcal{A} will guess γ correctly with advantage ϵ . Since the simulator outputs $\beta' = 1$ only when \mathcal{A} is correct in we have $\Pr[\beta' = \beta | \overline{abort} \land \beta = 1] = \frac{1}{2} + \epsilon$.

Putting this all together we calculate the advantage of the simulator as

$$\begin{aligned} &\Pr[\texttt{abort}]\Pr[\beta' = \beta|\texttt{abort}] + \Pr[\texttt{abort}](\Pr[\beta = 0]\Pr[\beta' = \beta|\texttt{abort} \land \beta = 0] + \\ &\Pr[\beta = 1]\Pr[\beta' = \beta|\texttt{abort} \land \beta = 1]) - \frac{1}{2} \\ &= \Pr[\texttt{abort}]\frac{1}{2} + \Pr[\texttt{abort}](\frac{1}{2}\frac{1}{2} + \frac{1}{2}(\frac{1}{2} + \epsilon)) - \frac{1}{2} \\ &= \Pr[\texttt{abort}]\frac{\epsilon}{2}. \end{aligned}$$

We now need to figure out the probability that the simulator aborts. We introduce the following notation: let $abort_1$ be the event that the simulator aborts in Phase 1, $abort_c$ be the event that the simulator aborts in the challenge phase given that it did not abort in Phase 1, and $abort_2$ be the event that the simulator aborts in Phase 2 given that it did not abort before.

Suppose that \mathcal{A} makes at most q_1 queries in Phase 1 and q_2 queries in Phase 2. Let v_1, \ldots, v_{q_1} denote the queries made in Phase 1, $v_{q_1+1}, \ldots, v_{q_1+q_2}$ denote the queries made in Phase 2, and v^* denote the challenge query. We calculate the probability of the individual abort events as follows.

$$\Pr[\texttt{abort}_1] = \Pr[\bigvee_{i=1}^{q_1} K(v_i) = 0]$$
(1a)

$$\leq \sum_{i=1}^{q_1} \Pr[K(v_i) = 0]$$
 (1b)

$$=\sum_{i=1}^{q_1} \frac{1}{m} \tag{1c}$$

$$=\frac{q_1}{m} \tag{1d}$$

$$\Pr[\overline{\mathtt{abort}_{\mathtt{c}}}] = \Pr[\sum_{i=1}^{n} x_{i,v_i^*} = km | \bigwedge_{i=1}^{q_1} K(v_i) = 1]$$
(2a)

$$= \frac{1}{n} \Pr[K(v^*) = 0 | \bigwedge_{i=1}^{q_1} K(v_i) = 1]$$
(2b)

$$= \frac{1}{n} \frac{\Pr[K(v^*) = 0]}{\Pr[\bigwedge_{i=1}^{q_1} K(v_i) = 1]} \Pr[\bigwedge_{i=1}^{q_1} K(v_i) = 1 | K(v^*) = 0]$$
(2c)

$$\geq \frac{1}{n} \Pr[K(v^*) = 0] \Pr[\bigwedge_{i=1}^{q_1} K(v_i) = 1 | K(v^*) = 0]$$
(2d)

$$= \frac{1}{nm} \Pr[\bigwedge_{i=1}^{q_1} K(v_i) = 1 | K(v^*) = 0]$$
(2e)

$$= \frac{1}{nm} (1 - \Pr[\bigvee_{i=1}^{q_1} K(v_i) = 0 | K(v^*) = 0])$$
(2f)

$$\geq \frac{1}{nm} (1 - \sum_{i=1}^{q_1} \Pr[K(v_i) = 0 | K(v^*) = 0])$$
(2g)

$$=\frac{1}{nm}(1-\sum_{i=1}^{q_1}\frac{1}{m})$$
(2h)

$$=\frac{1}{nm}(1-\frac{q_1}{m})\tag{2i}$$

We calculate Equation 2b as follows. We have that $\sum_{i=1}^{n} x_{i,v_i^*} = km$ iff $\sum_{i=1}^{n} x_{i,v_i^*} = wm$, that is $K(v_i^*) = 0$, for some $0 \le w < n$ and w = k. Since k is chosen uniformly at random between 0 and n-1 the probability that w = k is $\frac{1}{n+1} \sim \frac{1}{n}$; this is independent of the condition $\bigwedge_{i=1}^{q_1} K(v_i) = 1$.

We derive Equation 2h from the following observation. Any pair of different idenities, (v, v') will differ by at least one bit and therefore for at least one i, x_i will be included in the function K for one and not the other. Due to the uniform distribution of members of X over \mathbb{Z}_m , the probability that K(v) = 0 is pair-wise independent from the probability that K(v') = 0.

$$\Pr[\texttt{abort}_2] = \Pr[\bigvee_{i=1}^{q_2} K(v_{q_1+i}) = 0 | K(v^*) = 0 \land (\bigwedge_{j=1}^{q_1} K(v_j) = 1)]$$
(3a)

$$\leq \sum_{i=1}^{q_2} \Pr[K(v_{q_1+i}) = 0 | K(v^*) = 0 \land (\bigwedge_{j=1}^{q_1} K(v_j) = 1)]$$
(3b)

$$=\sum_{i=1}^{q_2} \frac{\Pr[K(v_{q_1+i})=0|K(v^*)=0]}{\Pr[\bigwedge_{j=1}^{q_1} K(v_j)=1|K(v^*)=0]} \Pr[\bigwedge_{j=1}^{q_1} K(v_j)=1|K(v^*)=0 \land K(v_{q_1+i})=0]$$
(3c)

$$\leq \sum_{i=1}^{q_2} \frac{\Pr[K(v_{q_1+i}) = 0 | K(v^*) = 0]}{\Pr[\bigwedge_{j=1}^{q_1} K(v_j) = 1 | K(v^*) = 0]}$$
(3d)

$$\leq \sum_{i=1}^{q_2} \frac{\Pr[K(v_{q_1+i}) = 0 | K(v^*) = 0]}{1 - \frac{q_1}{m}}$$
(3e)

$$=\sum_{i=1}^{q_2} \frac{1}{m(1-\frac{q_1}{m})} \tag{3f}$$

$$=rac{q_2}{m(1-rac{q_1}{m})}$$
 (3g)

We derive Equation 3e from fact that $m > q \ge q_1$.

We now put it all together. We lower bound the probability that the simulator will abort by calculating a lower bound for $\Pr[abort_1]\Pr[abort_c]\Pr[abort_2]$. (Recall, that the event of the not aborting at later stages was conditioned on the simulator not aborting earlier.)

$$\Pr[\texttt{abort}_1]\Pr[\texttt{abort}_c]\Pr[\texttt{abort}_2] \ge (1 - \frac{q_1}{m})\frac{1}{nm}(1 - \frac{q_1}{m})(1 - \frac{q_2}{m(1 - \frac{q_1}{m})})$$
(4a)

$$\geq (1 - \frac{q}{m})\frac{1}{nm}(1 - \frac{q}{m})(1 - \frac{q}{m(1 - \frac{q}{m})})$$
(4b)

$$= \frac{1}{n} \left(\frac{(1 - \frac{q}{m})^2}{m} - \frac{q(1 - \frac{q}{m})}{m^2} \right)$$
(4c)

$$\geq \frac{1}{n} \left(\frac{(1 - \frac{q}{m})^2}{m} - \frac{q}{m^2} \right) \tag{4d}$$

$$\geq \frac{1}{n} \left(\frac{1 - 2\frac{q}{m}}{m} - \frac{q}{m^2} \right) \tag{4e}$$

$$=\frac{1}{n}\left(\frac{1}{m}-\frac{3q}{m^2}\right)\tag{4f}$$

The simulator will set m to maximize $\frac{1}{m} - \frac{3q}{m^2}$. If we take the derivative with respect to m we get $\frac{6q}{m^3} - \frac{1}{m^2}$. We solve this function to find that it is maximized when m = 6q. Plugging this back in we find that $\Pr[\texttt{abort}] \geq \frac{1}{7nq}$. Therefore, if adversary that can break our scheme with advantage ϵ , then it we can construct an adversary that wins the BLDDH game with advantage $\frac{\epsilon}{14nq}$.

5.1 Hierarchical IBE and CCA Security

We could use our scheme to build a ℓ -deep Hierarchical IBE encryption scheme [6, 8] by creating an ℓ by n matrix U of public parameters. The security reduction of our scheme would be $\left(\frac{\epsilon}{14nq}\right)^{\ell}$ and thus would only be suitable for small values of ℓ . The security reduction results from the fact that the simulator must be "right" for all ℓ levels in the hierarchy in any particular run.

However, a hierarchy of small depth might be useful in practice as hierarchies in certificatebased public key infrastructures are typically not very deep. In particular if we use a hierarchy of two levels we can apply the result of Canetti et. al. [5] to achieve CCA security in our scheme.

6 Signatures

We also note that a signature scheme that is secure against existential forgery under an adaptive chosen-message attack [7] can be built without using random oracles using our same techniques. The security of this scheme would rely on the Computational Diffie-Hellman assumption.

7 Conclusion

We presented the first efficient Identity-Based Encryption scheme that is secure in the full model without random oracles. We proved our they security of our scheme by reducing it to the Bilinear Decisional Diffie-Hellman problem.

8 Acknowledgements

We would like to thank Dan Boneh for giving helpful suggestions.

References

- Dan Boneh and Xavier Boyen. Efficient selective-id secure identity based encryption without random oracles. In Proceedings of the International Conference on Advances in Cryptology (EUROCRYPT '04), Lecture Notes in Computer Science. Springer Verlag, 2004.
- [2] Dan Boneh and Xavier Boyen. Secure identity based encryption without random oracles. In *Proceedings of the Advances in Cryptology (CRYPTO '04)*, 2004.
- [3] Dan Boneh and Matthew K. Franklin. Identity-based encryption from the weil pairing. In Proceedings of the 21st Annual International Cryptology Conference on Advances in Cryptology, pages 213–229. Springer-Verlag, 2001.
- [4] Ran Canetti, Shai Halevi, and Jonathan Katz. A forward-secure public-key encryption scheme. In *Proceedings of Eurocrypt 2003*. Springer-Verlag, 2003.
- [5] Ran Canetti, Shai Halevi, and Jonathan Katz. Chosen-ciphertext security from identitybased encryption. In *Proceedings of Eurocrypt 2004*. Springer-Verlag, 2004.
- [6] Craig Gentry and Alice Silverberg. Hierarchical id-based cryptography. In Proceedings of the 8th International Conference on the Theory and Application of Cryptology and Information Security, pages 548–566. Springer-Verlag, 2002.
- [7] Shafi Goldwasser, Silvio Micali, and Ronald L. Rivest. A digital signature scheme secure against adaptive chosen-message attacks. *SIAM J. Comput.*, 17(2):281–308, 1988.
- [8] Jeremy Horwitz and Ben Lynn. Toward hierarchical identity-based encryption. In Advances in Cryptology: EUROCRYPT 2002, pages 466–481, 2002.
- [9] Adi Shamir. Identity-based cryptosystems and signature schemes. In Proceedings of CRYPTO 84 on Advances in cryptology, pages 47–53. Springer-Verlag New York, Inc., 1985.