Journal of Zhejiang University-SCIENCE C (Computers & Electronics)
ISSN 1869-1951 (Print); ISSN 1869-196X (Online)
www.zju.edu.cn/jzus; www.springerlink.com
E-mail: jzus@zju.edu.cn

CCA2 secure biometric identity based encryption with constant-size ciphertext*

Yang YANG[†1,2], Yu-pu HU[1], Le-you ZHANG[3], Chun-hui SUN[1]
([1]*Department of Communication Engineering, Xidian University, Xi'an 710071, China*)
([2]*Department of Mathematics and Computer Science, Fuzhou University, Fuzhou 350002, China*)
([3]*Department of Mathematics Science, Xidian University, Xi'an 710071, China*)
[†]E-mail: yang.yang.research@gmail.com

Received Dec. 11, 2010; Revision accepted May 26, 2011; Crosschecked Sept. 1, 2011

Abstract: We propose a new biometric identity based encryption scheme (Bio-IBE), in which user biometric information is used to generate the public key with a fuzzy extractor. This is the first Bio-IBE scheme that achieves constant size ciphertext. This is also a scheme that is secure against the adaptive chosen ciphertext attack (CCA2). Details are presented along with a discussion of Shamir's threshold secret sharing and fuzzy extraction of biometrics, which is based on error correction codes. We also define a security model and prove that the security of the proposed scheme is reduced to the decisional bilinear Diffie-Hellman (DBDH) assumption. The comparison shows that the proposed scheme has better efficiency and stronger security compared with the available Bio-IBE schemes.

Key words: Public key cryptography, Identity-based, Data security, Biometric, Encryption
doi:10.1631/jzus.C1000429 **Document code:** A **CLC number:** TP309

1 Introduction

Identity based encryption (Shamir, 1984) means a public key cryptographic system that allows the user to choose his/her telephone number or email address as the public key, instead of generating a random pair of public/private keys. A private key generator (PKG) calculates the private key from the user identity and the master secret key, and then distributes the private key to the participant. The main disadvantage of these schemes is that human biometric characteristics (such as palm prints, fingerprints, speech-sounds, and iris scans) are not allowed to be used as identities. Since a biometric sample is often disturbed by noises or has distortion when sampled, common identity based schemes cannot be used.

A new concept 'fuzzy identity based encryption' was proposed by Sahai and Waters (2005), in which identities are regarded as a set of attributes rather than a string of social characters, as in IBE schemes. Burnett *et al.* (2007) proposed the concept of biometric identity based signature (Bio-IBS), where they used biometric data to construct the public key, but they proposed no concrete scheme. Sarier (2008) proposed the first biometric identity based encryption (Bio-IBE) scheme. Then, Sarier (2010) constructed generic Bio-IBE schemes that require no bilinear pairing operation. Later on, Sarier (2011) modified the scheme by introducing an IBS scheme to resist a new denial of service (DoS) attack. The chief shortcoming of these available schemes (Sarier, 2008; 2010; 2011) is that the size of ciphertext is linear with the user identity and requires a large amount of pairing operations in the decryption phase. The research in this paper is motivated by the observation of these shortcomings.

* Project supported by the National Basic Research Program (973) of China (No. 2007CB311201) and the National Natural Science Foundation of China (Nos. 60970119, 60803149, 60833008, and 61072067)

To improve efficiency, this paper presents a new Bio-IBE scheme that has several desirable properties:

1. Constant-size ciphertext. Unlike other available Bio-IBE schemes (Sarier, 2008; 2010; 2011) where the sizes of ciphertext are linear with the attribute vector of the receiver, this proposal is the first scheme that achieves constant-size ciphertext. The communication overhead is greatly reduced in the proposed scheme.

2. Efficient decryption algorithm. This scheme is the first scheme that requires constant bilinear pairing operations (i.e., only two pairings) in the decryption algorithm, while the pairing operations in other schemes are linear with the error tolerant parameter.

3. Simpler key generation algorithm. The key generation algorithm is more efficient since the costly MapToPoint hash function is replaced by an ordinary hash function.

4. Decisional bilinear Diffie-Hellman (DBDH) assumption. This proposed scheme is proved secure and its security is reduced to the DBDH assumption. The hardness of DBDH assumption is stronger than that of k-BDHI assumption, which is the basis for other schemes (Sarier, 2008; 2011).

5. Chosen ciphertext security. This is a scheme that is secure against the adaptive chosen ciphertext attack (CCA2), while many of other Bio-IBE schemes with concrete construction and detailed proof are secure only against the chosen plaintext attack (CPA).

2 Preliminaries

2.1 Bilinear map

Let G and G_1 be two (multiplicative) cyclic groups of prime order p. Let g be a generator of G. Bilinear map $\hat{e}$ is a map $\hat{e}: G \times G \rightarrow G_1$ with the following properties:

1. Bilinearity: $\hat{e}(u^a, v^b) = \hat{e}(u, v)^{ab}$, for all $u, v \in G, a, b \in \mathbb{Z}_p$;

2. Non-degeneracy: $\hat{e}(g, g) \neq 1_{G_1}$;

3. Computability: there exists an efficient algorithm to calculate $\hat{e}(u, v)$, for all $u, v \in G$.

The Tate pairing and modified Weil pairing (Boneh and Franklin, 2001) are maps of this kind.

2.2 Hardness assumption

Security of the proposed Bio-IBE scheme is reduced to the hardness of DBDH problem.

Definition 1 (DBDH problem) Given a group G of prime order p with generator g and elements $g^a, g^b, g^c \in G$ for some uniformly chosen $a, b, c \in \mathbb{Z}_p$ and $Z \in G_1$ as input, the DBDH problem is to decide whether Z is equal to $\hat{e}(g, g)^{abc}$ or not.

An algorithm $\mathcal{C}$ has advantage ε in solving the DBDH problem if

$$\begin{aligned}&|\Pr[\mathcal{C}(\hat{e}(g,g)^{abc}, g, g^a, g^b, g^c) = 1] \\ &-\Pr[\mathcal{C}(Z, g, g^a, g^b, g^c) = 1]| \geq \varepsilon,\end{aligned}$$

where g, a, b, c, Z are randomly chosen.

Definition 2 If there is no adversary that can solve the DBDH problem with an advantage ε running in time t, we say that (t, ε) DBDH assumption holds.

2.3 Shamir's threshold secret sharing

In Shamir's threshold secret sharing scheme, a secret is divided into several parts that are disseminated to different participants. A certain number of the parts are required for reconstructing the secret. Let $s \in \mathrm{GF}(p)$ be the secret to be shared and the dealer selects a polynomial $f(x)$ of degree $d-1$ with $f(0) = s$, i.e.,

$$f(x) = s + \sum_{i=1}^{d-1} a_i x^i \pmod{p}.$$

If we assign each participant P_i a unique element α_i, the dealer sends P_i the secret share $s_i = f(\alpha_i)$. A participant group $\mathcal{S}$ with $|\mathcal{S}| \geq d$ can recover the secret s by computing

$$f(x) = \sum_{P_i \in \mathcal{S}} \Delta_{\alpha_i, \mathcal{S}}(x) f(\alpha_i) = \sum_{P_i \in \mathcal{S}} \Delta_{\alpha_i, \mathcal{S}}(x) s_i,$$

where

$$\Delta_{\alpha_i, \mathcal{S}}(x) = \prod_{P_j \in \mathcal{S}, j \neq i} \frac{x - \alpha_j}{\alpha_i - \alpha_j} \pmod{p}.$$

On the other hand, a group $\mathcal{T}$ of the participants cannot obtain any information about the secret s if $|\mathcal{T}| < d$.

2.4 Fuzzy extraction from biometrics

The word 'biometrics' derives from the Greek words 'metrikos' (measure) and 'bios' (life). It indicates a science involving the analysis of biologic characteristics. However, using biometric measurement as a basis for keys is problematic, because biometric data are not perfectly reproducible. For a Bio-IBE scheme, the biometrics is processed as follows (Sarier, 2008) to overcome this problem:

1. The biometrics of the user is extracted with a sensor to obtain the raw biometric data.

2. The feature extractor is exerted on the raw biometric information to obtain the feature vector (i.e., the attributes). Then, each attribute is assigned a unique $\mu_i \in \mathbb{Z}_p^*$ to construct the identity $\omega = (\mu_1, \mu_2, \cdots, \mu_n)$. Here n represents the size of attributes.

3. Each feature forming the feature vector is quantized to generate the binary template b (i.e., the biometric template).

4. The fuzzy extractor is used to generate a unique string ID via error-correction codes from the binary template b of the user such that an error tolerant t is allowed. In other words, we obtain the same string ID although the fuzzy extractor is applied on a disparate b' that satisfies $\mathrm{dis}(b', b) < t$. Here, t is the error tolerance threshold of the fuzzy extractor and dis() is the distance metric measuring the difference in biometric reading.

The fuzzy extractor mentioned above is based on the fuzzy commitment scheme (Juels and Wattenberg, 1999). The fuzzy extractor (Dodis *et al.*, 2008) is described as follows.

Let $\mathcal{M} = \{0, 1\}^n$ be a metric space consisting of biometric data points with a distance function dis: $\mathcal{M} \times \mathcal{M} \rightarrow \mathbb{Z}^+$. Here, dis measures the distance between biometric templates $b, b' \in \mathcal{M}$. An $(\mathcal{M}, l, t)$ fuzzy extractor consists of two functions Gen and Rep.

Gen: The Gen function takes $b \in \mathcal{M}$ as the input and outputs a string ID and public string PAR. PAR is used by the Rep algorithm to regenerate ID from b', where $\mathrm{dis}(b, b') \leq t$.

Rep: The Rep function allows the recovery of ID from PAR and any b' such that $\mathrm{dis}(b, b') \leq t$. In other words, if $\mathrm{Gen}(b) \rightarrow (\mathrm{ID}, \mathrm{PAR})$, then $\mathrm{Rep}(b', \mathrm{PAR}) \rightarrow \mathrm{ID}$, $\forall b, b' \in \mathcal{M}$ with $\mathrm{dis}(b, b') \leq t$.

Dodis *et al.* (2008) described a concrete fuzzy extractor under the Hamming distance metric for the space $\mathcal{M} = \{0, 1\}^n$ and a collision resistant hash function $H : \{0, 1\}^n \rightarrow \{0, 1\}^l$.

The Gen function takes b as the input, and then returns ID$= H(b)$ and public parameter $\mathrm{PAR} = b \oplus C_{\mathrm{e}}(\mathrm{ID})$, where C_{e} is an encoding function.

The Rep function takes PAR and the biometrics b' as input and calculates

$$\begin{aligned} \mathrm{ID}' &= C_{\mathrm{d}}(b' \oplus \mathrm{PAR}) = C_{\mathrm{d}}(b' \oplus b \oplus C_{\mathrm{e}}(\mathrm{ID})) \\ &= C_{\mathrm{d}}(e' \oplus C_{\mathrm{e}}(\mathrm{ID})), \end{aligned}$$

where $e' = b' \oplus b$. Then, if $\mathrm{dis}(b, b') \leq t$, ID$'$ =ID. Here, C_{d} is the decoding function corresponding to C_{e} and corrects error up to threshold t.

3 Formal definition of Bio-IBE

3.1 Bio-IBE scheme

A Bio-IBE consists of the following four algorithms: Setup, Extraction, Encryption, and Decryption. They are described as follows.

Setup: Given a security parameter k and error tolerant parameter d, the algorithm generates the master secret key MK and the public parameters PK of the system.

Extraction: Given an identity ω and master secret key MK, the algorithm returns the corresponding private key K_ω.

Encryption: Given the public parameters PK, identity ω', and a message M, the algorithm outputs the ciphertext C.

Decryption: Given a private key K_ω and the ciphertext C that is encrypted with identity ω', the algorithm outputs plaintext M if $|\omega' \cap \omega| \geq d$; otherwise, it aborts.

3.2 Security model

Definition 3 A Bio-IBE is indistinguishable against the chosen ciphertext attack (IND-sID-CCA2) if no attacker has a non-negligible advantage in the following game.

Initialization: Adversary $\mathcal{A}$ outputs a challenge identity ω^*.

Setup: Challenger $\mathcal{C}$ runs the Setup algorithm and sends adversary $\mathcal{A}$ the public parameters PK.

Phase 1: Adversary $\mathcal{A}$ issues private key extraction queries and decryption queries.

1. Extraction queries: $\mathcal{A}$ issues private key extraction queries for identity γ_j, where $|\gamma_j \cap \omega^*| < d$. In response, $\mathcal{C}$ runs the Extraction algorithm to obtain the private key K_{γ_j} and sends it to $\mathcal{A}$.

2. Decryption queries: $\mathcal{A}$ issues decryption queries on ciphertext C and an identity γ_j, where

$|\gamma_j \cap \gamma'_j| \geq d$. In response, $\mathcal{C}$ runs the Extraction algorithm to obtain the private key K_{γ_j}, and then runs the Decryption algorithm to obtain the plaintext M, which is forwarded to $\mathcal{A}$.

Challenge: Adversary $\mathcal{A}$ outputs two messages M_0, M_1 to challenge. Challenger $\mathcal{C}$ randomly chooses $\beta \in \{0,1\}$ and encrypts M_β with identity ω^*. The ciphertext is returned to $\mathcal{A}$.

Phase 2: $\mathcal{A}$ issues private key extraction queries and decryption queries as in phase 1.

Guess: Adversary $\mathcal{A}$ outputs a guess $\beta' \in \{0,1\}$ and wins the game if $\beta' = \beta$. Adversary $\mathcal{A}$'s advantage in attacking the above game is defined as

$$\mathrm{Adv}_{\mathcal{A}} = |\Pr[\beta' = \beta] - 1/2|.$$

Definition 4 We say that a Bio-IBE is (t, ϵ, q_E, q_D)-IND-sID-CCA2 secure if all t time adversaries making at most q_E private key extraction queries, q_D decryption queries have advantage at most ϵ in winning the above game.

FID and sID security models: The distinction between these two models was clearly explained in Boneh and Boyen (2004). In the full identity (FID) security model, the adversary can issue adaptive chosen identity queries. After the query phase, the adversary chooses the identity it wishes to attack, and asks for a challenge for this identity. The selective identity (sID) security model is slightly weaker than the FID model. In the sID security model, the adversary must commit to the identity it intends to attack ahead of the Setup phase. Then, the adversary can still issue adaptive chosen ciphertext and adaptive chosen identity queries. As shown in Boneh and Boyen (2004), any sID secure IBE is also a FID secure IBE with somewhat inefficient reduction.

4 An efficient Bio-IBE scheme

In this section, notations are defined and an efficient Bio-IBE scheme is proposed. The proposed scheme involves three roles: the private key generator (PKG), a sender, and a receiver. In the encryption phase, the sender receives the biometric data of the receiver and the corresponding parameter PAR. The sender extracts the features and calculates the biometric string ID with a fuzzy extractor. In the decryption phase, the proposed scheme allows for error tolerance due to the noisy nature of biometrics. As in Sarier (2008), we suppose that if $|\omega \cap \omega'| \geq d$, then $\mathrm{dis}(b, b') \leq t$ and $\mathrm{ID} = \mathrm{ID}'$. In this way, the receiver can decrypt the ciphertext encrypted with ω' using the private key corresponding to ω if ω and ω' are within a certain distance d of each other.

4.1 Notations

G: a multiplicative cyclic group of prime order p.

G_1: a multiplicative cyclic group of prime order p.

g: a generator of G with order p.

$\hat{e}$: a bilinear pairing map, where $\hat{e}: G \times G \to G_1$.

M: a plaintext of message.

C_{e}: the encoding function of error correction code.

C_{d}: the decoding function of error correction code.

d: the error tolerant parameter, which represents the distance that is allowed between two biometric attributes for a successful decryption.

H_1: a collusion resistant hash function, where $H_1: \mathbb{Z}_p^* \times \{0,1\}^* \to \mathbb{Z}_p^*$.

4.2 Setup phase

Given a security parameter k_0, PKG generates two multiplicative cyclic groups G and G_1 of prime order $p > 2^{k_0}$ and picks a random generator $g \in G$. PKG randomly selects $g_1 \in G$, $s \in \mathbb{Z}_p^*$ and computes $g_2 = g^s$. Choose an error tolerant parameter $d \in \mathbb{Z}^+$. Then, PKG publishes the system's public parameters $\{g, g_1, g_2, d\}$ and keeps the master secret key s secret.

4.3 Extraction phase

First, the user biometric attributes $\omega = (\mu_1, \mu_2, \cdots, \mu_n)$ are extracted from the raw biometric information with a feature extractor, where each $\mu_i \in \mathbb{Z}_p^*$. Then, calculate the biometric string ID$=$ $H(b)$ from the biometric template b. Given the user's ω and ID, the PKG extracts the private key as follows:

1. Choose a random $d-1$ degree polynomial $p(x) = a_0 + a_1 x + \ldots + a_{d-1} x^{d-1}$ such that $p(0) = a_0 = s$.

2. For $\mu_i \in \omega$, compute $d_{i,1} = (g_1 \cdot g^{H_1(\omega, \mathrm{ID})})^{p(\mu_i)} = (g_1 \cdot g^{h^{\mathrm{ID}}})^{p(\mu_i)}$, $d_{i,2} = g^{p(\mu_i)}$. Calculate PAR$=$Gen(b, ID).

3. PKG confidentially sends $K_\omega = \{d_{i,1}, d_{i,2}\}_{\mu_i \in \omega}$ to the user and publishes PAR.

4.4 Encryption phase

The sender receives biometric data of the designated receiver together with the public parameter PAR. The sender extracts the feature attributes ω' and computes $\mathrm{ID}' = \mathrm{Rep}(b', \mathrm{PAR})$. Here, if $\mathrm{dis}(b', b) < t$, then $\mathrm{ID} = \mathrm{ID}'$. Given ID', ω', and M, the sender performs the following:

1. Select a random $r \in \mathbb{Z}_p^*$ and calculate $C_1 = g^r, C_2 = (g^{H_1(\omega', \mathrm{ID}')})^r = (g^{h^{\mathrm{ID}'}})^r, C_3 = \hat{e}(g_1, g_2)^r M$.

2. Send the ciphertext $C = (\omega', C_1, C_2, C_3)$ to U.

4.5 Decryption phase

Given a ciphetrtext $C = (\omega', C_1, C_2, C_3)$, a receiver with private key K_ω decrypts C as follows:

1. If $|\omega \cap \omega'| < d$, U aborts.

2. If $|\omega \cap \omega'| \geq d$, U chooses any set that satisfies $S \subseteq \omega \cap \omega'$ and $|S| = d$ and calculates

$$M = C_3 \cdot \frac{\hat{e}(C_2, \prod_{\mu_j \in S} (d_{i,2})^{\Delta_{\mu_j, S}(0)})}{\hat{e}(C_1, \prod_{\mu_j \in S} (d_{i,1})^{\Delta_{\mu_j, S}(0)})}.$$

The plaintext M can be recovered since

$$\begin{aligned}
& C_3 \cdot \frac{\hat{e}(C_2, \prod_{\mu_j \in S} (d_{i,2})^{\Delta_{S,\mu_j}(0)})}{\hat{e}(C_1, \prod_{\mu_j \in S} (d_{i,1})^{\Delta_{S,\mu_j}(0)})} \\
= & \ C_3 \cdot \frac{\hat{e}((g^{H_1(\omega', \mathrm{ID}')})^r, \prod_{\mu_j \in S} (g^{p(\mu_i)})^{\Delta_{S,\mu_j}(0)})}{\hat{e}(g^r, \prod_{\mu_j \in S} ((g_1 \cdot g^{H_1(\omega, \mathrm{ID})})^{p(\mu_i)})^{\Delta_{S,\mu_j}(0)})} \\
= & \ \frac{\hat{e}((g^{H_1(\omega', \mathrm{ID}')})^r, g^s)}{\hat{e}(g^r, (g_1 \cdot g^{H_1(\omega, \mathrm{ID})})^s)} \cdot M\hat{e}(g_1, g_2)^r \\
= & \ \frac{\hat{e}((g^{H_1(\omega', \mathrm{ID}')})^r, g^s)}{\hat{e}((g^{H_1(\omega, \mathrm{ID})})^s, g^r) \cdot \hat{e}(g_1^s, g^r)} \cdot M\hat{e}(g_1, g_2)^r \\
= & \ \frac{\hat{e}((g^{H_1(\omega', \mathrm{ID}')})^r, g^s)}{\hat{e}((g^{H_1(\omega, \mathrm{ID})})^s, g^r) \cdot \hat{e}(g_1, g^s)^r} \cdot M\hat{e}(g_1, g_2)^r \\
= & \ \frac{\hat{e}((g^{h^{\mathrm{ID}'}})^r, g^s)}{\hat{e}((g^{h^{\mathrm{ID}}})^s, g^r) \cdot \hat{e}(g_1, g^s)^r} \cdot M\hat{e}(g_1, g_2)^r \\
= & \ M.
\end{aligned}$$

The last equation holds since $\mathrm{ID}' = \mathrm{ID}$ when $|\omega \cap \omega'| \geq d$ and $\mathrm{dis}(b, b') < t$.

5 Security analysis

Theorem 1 Suppose that the (t', ε') DBDH assumption holds in G. Then the constructed Bio-IBE scheme is $(t, \varepsilon, q_E, q_D)$ IND-sID-CCA2 secure for $\varepsilon' = \varepsilon, t' = t + d(t_{\mathrm{Mul}} + t_{\mathrm{Exp}})q_E$, where t_{Mul} is the time for a multiplication, t_{Exp} is the time for an exponentiation, and d is the error tolerant parameter.

Proof Suppose that there exists a $(t, \varepsilon, q_E, q_D)$ adversary $\mathcal{A}$ against the proposed scheme. Then an algorithm $\mathcal{C}$ can be constructed to solve the DBDH problem in time t' and with a probability ε'. Challenger $\mathcal{C}$ is given a tuple (g, g^a, g^b, g^c, Z) of the DBDH problem, where Z either equals $\hat{e}(g, g)^{abc}$ or is a random element in G_1. The game proceeds as follows.

Initialization: A challenge identity $\omega^* = (\mu_1^*, \mu_2^*, \cdots, \mu_n^*)$ is chosen by adversary $\mathcal{A}$.

Setup: The challenger $\mathcal{C}$ sets $g_1 = g^a, g_2 = g^b$ and chooses error tolerant parameter $d \in \mathbb{Z}^+$. Then $\mathcal{C}$ returns the public parameters PK$= (g, g_1, g_2, d)$ to adversary $\mathcal{A}$.

Hash queries: Adversary $\mathcal{A}$ is allowed to issue a hash query in any phase. Upon receiving a query ω_i, if there exists (ω_i, l_i, g^{h_i}) in H-list, return g^{h_i}. If $\omega_i = \omega^*$, choose $l^* \in \mathbb{Z}_p$ at random and set $g^{h^*} = g^{l^*}$. Otherwise, randomly select $l_i \in \mathbb{Z}_p$ and compute $g^{h_i} = g^{l_i}/g_1$.

Phase 1: Adversary issues private key extraction queries and decryption queries in this phase.

1. Extraction queries: When a private key query is received for $\gamma_j = (\mu_1, \mu_2, \cdots, \mu_n)$, where $|\gamma_j \cap \omega^*| < d$, challenger $\mathcal{C}$ sets $\Gamma = \gamma_j \cap \omega^*$ and sets Γ' to be any set that satisfies $\Gamma \subseteq \Gamma' \subseteq \gamma_j$, $|\Gamma'| = d - 1$. Let $S = \Gamma' \cup \{0\}$. Run the above hash query to obtain (γ_j, l_j, g^{h_j}) in H-list.

(1) For every $\mu_i \in \Gamma'$, pick $\lambda_i \in \mathbb{Z}_p$ at random and compute $(d_{i,1}, d_{i,2}) = ((g_1 g^{h_j})^{\lambda_i}, g^{\lambda_i})$. For a random polynomial $p(\cdot)$ of degree $d-1$ over $\mathbb{Z}_p$ with $p(0) = b$, define $\lambda_i = p(\mu_i)$. Thus, challenger $\mathcal{C}$ has successfully constructed $(d_{i,1}, d_{i,2})$ for $\mu_i \in \Gamma'$.

(2) For every $\mu_i \in \gamma_j \setminus \Gamma'$, compute

$$\begin{aligned}
d_{i,1} &= g_2^{\Delta_{0,S}(\mu_i) l_j} (\prod_{\mu_k \in \Gamma'} (g_1 \cdot g^{h_j})^{\Delta_{\mu_k, S}(\mu_i) \lambda_k}), \\
d_{i,2} &= g_2^{\Delta_{0,S}(\mu_i)} (\prod_{\mu_k \in \Gamma'} g^{\Delta_{\mu_k, S}(\mu_i) \lambda_k}).
\end{aligned}$$

Note that $g_1 \cdot g^{h_j} = g^{l_j}$ for $\gamma_j \neq \omega^*$. Then,

$$\begin{aligned}
d_{i,1} &= g^{l_j \Delta_{0,S}(\mu_i) b} (g^{l_j (\sum_{\mu_k \in \Gamma'} \Delta_{\mu_k, S}(\mu_i) p(\mu_k))}) \\
&= g^{l_j (\Delta_{0,S}(\mu_i) p(0) + \sum_{\mu_k \in \Gamma'} \Delta_{\mu_k, S}(\mu_i) p(\mu_k))} \\
&= g^{l_j p(\mu_i)} = (g_1 \cdot g^{h_j})^{p(\mu_i)} \\
&= (g_1 \cdot g^{H_1(\gamma_j, \mathrm{ID})})^{p(\mu_i)}, \\
d_{i,2} &= g^{\Delta_{0,S}(\mu_i) b} (g^{\sum_{\mu_k \in \Gamma'} \Delta_{\mu_k, S}(\mu_i) p(\mu_k)}) \\
&= g^{\Delta_{0,S}(\mu_i) p(0) + \sum_{\mu_k \in \Gamma'} \Delta_{\mu_k, S}(\mu_i) p(\mu_k)} \\
&= g^{p(\mu_i)}.
\end{aligned}$$

Thus, challenger $\mathcal{C}$ has successfully simulated the private key of $\gamma_j = (\mu_1, \mu_2, \cdots, \mu_n)$.

2. Decryption queries: To answer the decryption query on $C = (\gamma'_j, C_1, C_2, C_3)$ and an identity γ_j where $|\gamma_j \cap \gamma'_j| \geq d$, challenger $\mathcal{C}$ operates as follows.

(1) Run the above extraction algorithm to construct the private key $K_{\gamma_j} = (d_{i,1}, d_{i,2})_{\mu_i \in \gamma_j}$.

(2) Choose any set S that satisfies $S \subseteq \gamma_j \cap \gamma'_j$ and $|S| = d$. Then compute the plaintext M and send it to $\mathcal{A}$, where M is computed as follows:

$$M = \frac{\hat{e}(C_2, \prod_{\mu_i \in S} (d_{i,2})^{\Delta_{\mu_i,S}(0)})}{\hat{e}(C_1, \prod_{\mu_i \in S} (d_{i,1})^{\Delta_{\mu_i,S}(0)})} \cdot C_3.$$

Challenge: Adversary $\mathcal{A}$ outputs two messages M_0, M_1 to challenge. Challenger $\mathcal{C}$ randomly chooses $\beta \in \{0, 1\}$ and encrypts M_β with identity ω^*, which is derived from b^*. The ciphertext is returned to $\mathcal{A}$:

$$C^* = (\omega^*, C_1^*, C_2^*, C_3^*) = (\omega^*, g^c, (g^c)^{l^*}, Z \cdot M_\beta).$$

1. If $Z = \hat{e}(g, g)^{abc}$ (i.e., when the input tuple is sampled from $\mathcal{P}_{\text{BDHE}}$), C^* is a valid encryption of M_β since

$$C_1^* = g^c,$$

$$C_2^* = (g^c)^{l^*} = (g^{l^*})^c = (g^{h^*})^c = (g^{H(\omega^*, \text{ID})})^c,$$

$$C_3^* = Z \cdot M_\beta = \hat{e}(g, g)^{abc} \cdot M_\beta = \hat{e}(g_1, g_2)^c \cdot M_\beta.$$

2. If Z is uniform in G_1 (i.e., when the input tuple is sampled from $\mathcal{R}_{\text{BDHE}}$), C_3^* is independent of M_β. Thus, C^* is independent of β in the adversary's view.

Phase 2: Adversary $\mathcal{A}$ issues private key extraction queries and decryption queries as in phase 1.

Guess: Adversary $\mathcal{A}$ finally outputs a guess $\beta' \in \{0, 1\}$. Challenger $\mathcal{C}$ concludes the game as follows. If $\beta' = \beta$, $\mathcal{C}$ outputs 1 indicating that $Z = \hat{e}(g, g)^{abc}$. Otherwise, it outputs 0 indicating that Z is random in G_1.

Probability analysis:

1. If the input is a tuple sampled from $\mathcal{P}_{\text{BDHE}}$ (i.e., $Z = \hat{e}(g, g)^{abc}$), then $\mathcal{A}$'s view is identical to the view in a real attack. Thus, $|\Pr[\beta = \beta'] - 1/2| \geq \varepsilon$.

2. If the input is a tuple sampled from $\mathcal{R}_{\text{BDHE}}$ (i.e., Z is uniform in G_1), then $|\Pr[\beta = \beta']| = 1/2$.

3. Therefore,

$$|\Pr[\mathcal{C}(\hat{e}(g, g)^{abc}, g, g^a, g^b, g^c) = 0]$$
$$-\Pr[\mathcal{C}(Z, g, g^a, g^b, g^c) = 0]| \geq |(\frac{1}{2} \pm \varepsilon) - \frac{1}{2}| = \varepsilon.$$

Time analysis: The running time of the simulation is dominated by the multiplication and exponent operation in the extraction query phase. Then we have $t' = t + d(t_{\text{Mul}} + t_{\text{Exp}})q_E$.

6 An improved Bio-IBE scheme to resist DoS attacks

Recently, Sarier (2001) presented a new DoS attack in which an adversary maliciously changes the public string PAR and brings the sender to use an incorrect public key for encryption. Then the receiver of ciphertext cannot obtain the right plaintext upon decryption. To ensure the validity and integrity of PAR, we use the digital signature that is signed by PKG to make our scheme immune against the DoS attack. Then the signature σ and PAR should be stored together. The efficient signature scheme used in our modified scheme is based on the scheme in Cha and Cheon (2003), which has its security reduced to computational Diffi-Hellman (CDH) assumption.

Setup: The public parameters (g, g_1, g_2) and master secret key s are chosen as in Bio-IBE other than two hash functions $H_2 : \{0, 1\}^* \times G \to \mathbb{Z}_p$ and $H_3 : \{0, 1\} \to G$.

Extraction: Given PKG's identity ID, the algorithm computes $K_{\text{ID}} = H_3(\text{ID})^s$ and outputs it as the private key associated to ID.

Signature: Given a secret key K_{ID} and a public string PAR, the algorithm picks a random $r \in \mathbb{Z}_p$ and outputs a signature $\sigma = (U, V)$ where $U = H_3(\text{ID})^r$ and $V = (K_{\text{ID}})^{r + H_2(\text{PAR}, U)}$.

Verification: To verify a signature $\sigma = (U, V)$ of PAR for PKG's identity ID, check whether

$$\hat{e}(g, V) = \hat{e}(g_2, U)\hat{e}(g_2, H_3(\text{ID})^{H_2(\text{PAR}, U)}).$$

Correctness:

$$\begin{aligned} \hat{e}(g, V) &= \hat{e}(g, (H_3(\text{ID})^s)^{r + H_2(\text{PAR}, U)}) \\ &= \hat{e}(g^s, H_3(\text{ID})^{r + H_2(\text{PAR}, U)}) \\ &= \hat{e}(g_2, H_3(\text{ID})^r \cdot H_3(\text{ID})^{H_2(\text{PAR}, U)}) \\ &= \hat{e}(g_2, U \cdot H_3(\text{ID})^{H_2(\text{PAR}, U)}) \\ &= \hat{e}(g_2, U)\hat{e}(g_2, H_3(\text{ID})^{H_2(\text{PAR}, U)}). \end{aligned}$$

7 Efficiency discussions and comparisons

7.1 Comparisons with Bio-IBE schemes

The proposed scheme and improved scheme are compared with those proposed by Sarier (2008; 2011) (Table 1). Since the scheme in Sarier (2011) is an improved version of that in Sarier (2008) (the scheme in Sarier (2011) integrates an IBS to the Bio-IBE scheme in Sarier (2008) to resist DoS attacks), one can just compare our scheme with that of Sarier (2008) in detail as follows:

1. The size of the public parameter and private key in Sarier (2008) is smaller than that of the proposed scheme. Since the private key is shorter in Sarier (2008), the cost of private key generation is n exponential operations less than the cost in the proposed scheme.

2. The size of ciphertext in Sarier (2008) is $n|G| + n_m$, where n is the length of user identity and n_m is the size of the message. This scheme has constant size ciphertext which consists of only two elements in group G and one element in group G_1. Thus, the network traffic load is decreased and the bandwidth at the user is reduced.

3. The cost of encryption is reduced in this scheme compared with that in Sarier (2008). The scheme in Sarier (2008) requires n exponential operations and one pairing operation, while ours requires only three exponential operations and no pairing operation.

4. The cost of decryption in this scheme is less than that in Sarier (2008). Measured in terms of the pairing operation that is the most time-consuming operation, the scheme in Sarier (2008) requires d pairing operations while only two pairing operations are required in our scheme. Thus, our scheme is more efficient in computation.

5. This scheme has its security reduced to DBDH assumption, while the security of the scheme in Sarier (2008) is reduced to k-BDHI assumption. As shown in Cheon (2006), the hardness of DBDH assumption is stronger than that of k-BDHI assumption.

To sum up, this scheme achieves higher efficiency and stronger security than the scheme in Sarier (2008).

7.2 Comparisons with fuzzy schemes

Due to the similarity between Bio-IBE and fuzzy IBE, the proposed scheme is compared with the existing fuzzy IBE schemes in Table 2. It is obvious that the new scheme achieves higher efficiency than the existing fuzzy IBE schemes. The detailed comparison is as follows:

1. The size of public parameters of the suggested scheme is a constant. However, in Sahai and Waters (2005), Fang *et al.* (2008), and Li *et al.* (2009), the size of public parameters grows with either the total number of users $|u|$ in the system or the length n of identity.

2. The size of the private key in Sahai and Waters (2005), Baek *et al.* (2007), Fang *et al.* (2008), Li *et al.* (2009), and Ren *et al.* (2010) and the proposed scheme grows with the length n of user identity. Then, the cost of generating the private key also grows with n.

In Sahai and Waters (2005), Baek *et al.* (2007), Fang *et al.* (2008), Li *et al.* (2009), and Ren *et al.* (2010), the size of ciphertext grows with the length n of user identity. Furthermore, Fang *et al.* (2008) added the length n_m of plaintext m. In this scheme, the size of ciphertext is a constant which consists only of two elements in group G and one element in group G_1. This performance is notably superior to that of other schemes.

3. In Sahai and Waters (2005), Baek *et al.* (2007), Fang *et al.* (2008), Li *et al.* (2009), and Ren *et al.* (2010), the cost of encryption is proportional to the length n of user identity. The suggested scheme requires merely three exponential operations.

4. In Sahai and Waters (2005), Baek *et al.* (2007), Fang *et al.* (2008), Li *et al.* (2009), and Ren *et al.* (2010), a large amount of pairing operations are required in the decryption process. The proposed scheme needs only two pairing operations to complete the decryption process.

5. The security of this scheme is reduced to the DBDH assumption and the security of schemes in Sahai and Waters (2005), Fang *et al.* (2008), Li *et al.* (2009), and Ren *et al.* (2010) is reduced to MBDH, mBDDH, k-BDH, and q-TBDHE assumptions. As shown in Cheon (2006), the hardness of DBDH assumption is stronger than that of other assumptions mentioned above.

Table 1 Comparison of various Bio-IBE schemes

	Value/Description			
	Sarier (2008)	Sarier (2011)	Our scheme	Improved scheme
Based on biometric	Yes	Yes	Yes	Yes
Size of public key	$2\|G\|$	$2\|G\|$	$3\|G\|$	$3\|G\|$
Size of private key	$n\|G\|$	$n\|G\|$	$2n\|G\|$	$2n\|G\|$
Size of ciphertext	$n\|G\|+n_m$	$n\|G\|+n_m$	$2\|G\|+\|G_1\|$	$2\|G\|+\|G_1\|$
Cost of key generation	nt_{Exp}	nt_{Exp}	$2nt_{\mathrm{Exp}}$	$2nt_{\mathrm{Exp}}$
Cost of encryption	$t_{\mathrm{Pair}}+nt_{\mathrm{Exp}}$	$t_{\mathrm{Pair}}+nt_{\mathrm{Exp}}$	$3t_{\mathrm{Exp}}$	$3t_{\mathrm{Exp}}$
Cost of decryption	dt_{Pair}	dt_{Pair}	$2t_{\mathrm{Pair}}$	$2t_{\mathrm{Pair}}$
Hardness assumption	k-BDHI	k-BDHI	DBDH	DBDH
DoS attack	No	Yes	No	Yes
Security model	sID	sID	sID	sID

Table 2 Comparison of various fuzzy IBE schemes

	Value/Description					
	Sahai and Waters (2005)	Baek *et al.* (2007)	Fang *et al.* (2008)	Li *et al.* (2009)	Ren *et al.* (2010)	Our scheme
Size of public key	$u\|G\|+\|G_1\|$	$3\|G\|$	$(n+2)\|G\|$	$4n\|G\|+(d+1)\|G_1\|$	$7\|G\|$	$3\|G\|$
Size of private key	$n\|G\|$	$2n\|G\|$	$3n\|G\|$	$(nd+2n+1)\|G_1\|$	$n\|G\|$	$2n\|G\|$
Size of ciphertext	$n\|G\|+\|G_1\|$	$(n+1)\|G\|+\|G_1\|$	$(n+1)\|G\|+n_m$	$(2n+2)\|G\|$	$n\|G\|+3\|G_1\|$	$2\|G\|+\|G_1\|$
Cost of key generation	nt_{Exp}	$2nt_{\mathrm{Exp}}$	$4nt_{\mathrm{Exp}}$	$(nd+2n+1)t_{\mathrm{Exp}}$	$3nt_{\mathrm{Exp}}$	$2nt_{\mathrm{Exp}}$
Cost of encryption	$(n+1)t_{\mathrm{Exp}}$	$(n+2)t_{\mathrm{Exp}}$	$(2n+2)t_{\mathrm{Exp}}$	$(2n+2)t_{\mathrm{Exp}}$	$(3n+5)t_{\mathrm{Exp}}$	$3t_{\mathrm{Exp}}$
Cost of decryption	dt_{Pair}	$(d+1)t_{\mathrm{Pair}}$	$2dt_{\mathrm{Pair}}$	$(n+d+1)t_{\mathrm{Pair}}$	$(n+1)t_{\mathrm{Pair}}$	$2t_{\mathrm{Pair}}$
Hardness assumption	MBDH	DBDH	mBDDH	k-BDH	q-TBDHE	DBDH
Security model	sID	sID	sID	sID	FID	sID

6. All of the schemes in comparison are proved secure in the sID security model except the scheme in Ren *et al.* (2010). Although the FID security model is slightly stronger than the sID model, our scheme is much more efficient than the scheme in Ren *et al.* (2010). In Ren *et al.* (2010), the size of ciphertext and the cost of encryption and decryption grow linearly with the length n of user identity. While in our scheme, the size and computation cost of ciphertext are constant. Moreover, the security of the suggested scheme is reduced to DBDH assumption while the scheme in Ren *et al.* (2010) is reduced to q-TBDHE assumption. As shown in Cheon (2006), the hardness of DBDH assumption is stronger than that of q-TBDHE assumption.

8 Conclusions

In this paper, we construct a new efficient biometric identity based scheme which is secure against the adaptive chosen ciphertext attack (CCA2) and which is also the first scheme achieving constant size ciphertext. This new scheme has many other advantages over those existing ones. First, the hash function is a regular one rather than the MapToPoint function for the encryption scheme. Second, the security of the proposed scheme is reduced to the hardness of DBDH assumption rather than some strong assumptions. Finally, only two pairing operations are executed in the decryption phase. Thus, the communication overhead of the network is decreased and the computation at the receiver is greatly reduced.

References

Baek, J., Susilo, W., Zhou, J.Y., 2007. New Constructions of Fuzzy Identity-Based Encryption. Proc. 2nd ACM Symp. on Information Computer and Communications Security, p.368-370. [doi:10.1145/1229285.1229330]

Boneh, D., Boyen, X., 2004. Efficient Selective-ID Secure Identity-Based Encryption without Random Oracles. Proc. EUROCRYPT, p.223-238.

Boneh, D., Franklin, M.K., 2001. Identity-Based Encryption from the Weil Pairing. Proc. CRYPTO, p.213-229.

Burnett, A., Byrne, F., Dowling, T., Duffy, A., 2007. A biometric identity based signature scheme. *Int. J. Network Secur.*, **5**(3):317-326.

Cha, J.C, Cheon, J.H., 2003. An Identity-Based Signature from Gap Diffie-Hellman Groups. Proc. Public Key Cryptography, p.18-30.

Cheon, J.H., 2006. Security Analysis of the Strong Diffie-Hellman Problem. Proc. EUROCRYPT, p.1-11.

Dodis, Y., Ostrovsky, R., Reyzin, L., Smith, A., 2008. Fuzzy extractors: how to generate strong keys from biometrics and other noisy data. *SIAM J. Comput.*, **38**(1):97-139. [doi:10.1137/060651380]

Fang, L., Wang, J., Ren, Y., Xia, J., Bian, S., 2008. Chosen ciphertext secure fuzzy identity based encryption without ROM. *J. Shanghai Jiao Tong Univ. (Sci.)*, **13**(6):646-650. [doi:10.1007/s12204-008-0646-y]

Juels, A., Wattenberg, M., 1999. A Fuzzy Commitment Scheme. ACM Conf. on Computer and Communications Security, p.28-36.

Li, X.M., Yang, B., Guo, Y.B., 2009. Fuzzy Identity Based Encryption Scheme with Some Assigned Attributes. Proc. 5th Int. Conf. on Information Assurance and Security, p.133-136. [doi:10.1109/IAS.2009.145]

Ren, Y.L., Gu, D.W., Wang, S.Z., Zhang, X.P., 2010. New fuzzy identity-based encryption in the standard model. *Informatica*, **21**(3):393-407.

Sahai, A., Waters, B., 2005. Fuzzy Identity-Based Encryption. Proc. EUROCRYPT, p.457-473.

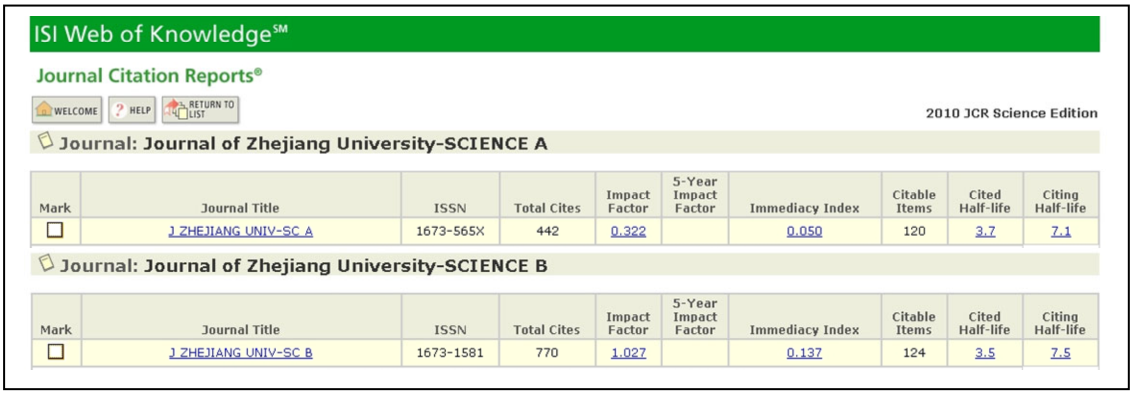
Sarier, N.D., 2008. A New Biometric Identity Based Encryption Scheme. Proc. ICYCS, p.2061-2066.

Sarier, N.D., 2010. Generic Constructions of Biometric Identity Based Encryption Systems. Proc. WISTP, p.90-105.

Sarier, N.D., 2011. A new biometric identity based encryption scheme secure against DoS attacks. *Secur. Commun. Networks*, **4**(1):23-32. [doi:10.1002/sec.162]

Shamir, A., 1984. Identity-Based Cryptosystems and Signature Schemes. Proc. Crypto, p.47-53.

2010 JCR of Thomson Reuters for *JZUS-A* and *JZUS-B*



ISI Web of Knowledge℠

Journal Citation Reports®

2010 JCR Science Edition

Journal: Journal of Zhejiang University-SCIENCE A

Mark	Journal Title	ISSN	Total Cites	Impact Factor	5-Year Impact Factor	Immediacy Index	Citable Items	Cited Half-life	Citing Half-life
□	J ZHEJIANG UNIV-SC A	1673-565X	442	0.322		0.050	120	3.7	7.1

Journal: Journal of Zhejiang University-SCIENCE B

Mark	Journal Title	ISSN	Total Cites	Impact Factor	5-Year Impact Factor	Immediacy Index	Citable Items	Cited Half-life	Citing Half-life
□	J ZHEJIANG UNIV-SC B	1673-1581	770	1.027		0.137	124	3.5	7.5

JZUS-A is an international “Applied Physics & Engineering” reviewed-Journal, covering research in Applied Physics, Mechanical and Civil Engineering, Environmental Science and Energy, Materials Science, and Chemical Engineering. *JZUS-B* is an international “Biomedicine & Biotechnology” reviewed-Journal, covering research in Biomedicine, Biochemistry, and Biotechnology. *JZUS-A* and *JZUS-B* were covered by SCI-E in 2007 and 2008, respectively.