SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

February 2024, Vol. 67, Iss. 2, 122101:1–122101:14
https://doi.org/10.1007/s11432-022-3706-7

Fault-tolerant identity-based encryption from SM9

Xiaohong LIU[1], Xinyi HUANG[2*], Zhaohui CHENG[3] & Wei WU[4]

[1]*Fujian Provincial Key Laboratory of Network Security and Cryptology, College of Computer and Cyber Security, Fujian Normal University, Fuzhou* 350117*, China;*
[2]*Artificial Intelligence Thrust, Information Hub, The Hong Kong University of Science and Technology (Guangzhou), Guangzhou* 511455*, China;*
[3]*Olym Information Security Technology Ltd., Shenzhen* 518052*, China;*
[4]*Center for Applied Mathematics of Fujian Province, School of Mathematics and Statistics, Fujian Normal University, Fuzhou* 350117*, China*

Received 15 August 2022/Revised 24 October 2022/Accepted 21 February 2023/Published online 22 January 2024

Abstract This paper initiates the formal study of attribute-based encryption within the framework of SM9, the Chinese National Cryptography Standard for Identity-Based Cryptography, by presenting two new fault-tolerant identity-based encryption (FIBE) schemes. Our first scheme uses the same private-key/ciphertext structure as the original SM9 algorithm and operates in a small attribute universe. As a result, it can be effectively and smoothly integrated into the information systems using SM9. In the random oracle model, we prove that our scheme is ciphertext-indistinguishable against fuzzy selective-identity and chosen-plaintext attacks under the $(k+3)$-DBDHI assumption. Our second design is a large universe FIBE scheme based on SM9 that is ciphertext-indistinguishable against chosen-plaintext attacks in the random oracle model under the (f, g)-GDDHE assumption. Finally, we compare the communication and computing costs of our schemes to those of other classical ones. The comparison shows that our schemes have comparable performance as others. We believe that our findings will accelerate the applications of SM9 in modern information systems such as cloud computing and blockchain.

Keywords attribute-based encryption, identity-based encryption, fault-tolerant, SM9

1 Introduction

Data confidentiality has always been one of the most important goals in information systems [1–3]. Encryption technology is a common method for ensuring data confidentiality. In recent years, biometric measurements (such as facial patterns, fingerprints, eye irises, and voice patterns) have been used as a public key to encrypt data. This is a well-studied area in cryptography, i.e., identity-based encryption (IBE) [4]. The public key in IBE is one's identity (such as phone number, email address, ID number, and equipment identification) or a bitstring calculated from the identity through a cryptographic algorithm. As a result, there is no need for a certificate to ensure the authenticity of the public key. Unlike telephone numbers or email addresses, biometric measurements are inherent, unique, unforgeable, difficult to copy or transfer, and do not need to be stored. We can provide our biometric measurements to others at any time. Furthermore, we can obtain the private key corresponding to the biometric measurements from a trusted third party without verifying to the third party that we are indeed qualified to use the biometric measurements. Due to these features, using biometric measurements to encrypt data has many important advantages over traditional IBE.

However, biometric measurements are noisy, and the recognition and extraction of the same biometric measurements at different times are rarely identical. Therefore, applying biometrics to traditional IBE encryption will result in decryption failure. In this case, the encryption scheme we need must have fault tolerance, i.e., a secret key corresponding to a biometric measurement can decrypt ciphertexts encrypted with relatively close but different measurements.

To address this issue, based on IBE, Sahai and Waters [5] introduced a set of descriptive attributes to represent the identity in IBE, termed fuzzy (fault-tolerant) identity-based encryption (FIBE). As an

* Corresponding author (email: xinyi@ust.hk)

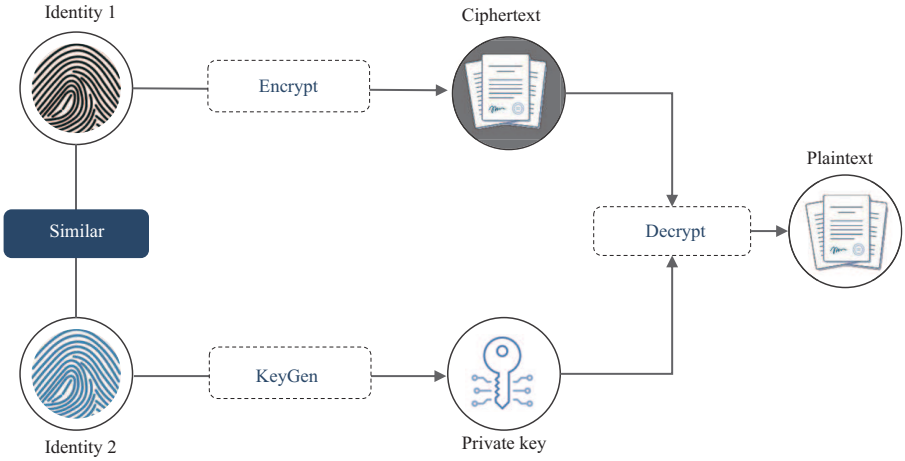


Figure 1 (Color online) An overview of FIBE systems.

improvement and generalization of IBE, FIBE can provide fault-tolerant capability; that is, one can decrypt a ciphertext containing an identity ID' using a private key containing another identity ID, as long as these two identities pass some measurement within a certain range (Figure 1). As a result, FIBE is often used in the IBE systems that use biometric measurements as identities, such as the common face recognition technology, iris recognition technology, fingerprint recognition technology (in electronic passports), staff attendance, mobile payment, and intelligent lock. In addition, as the prototype of attribute-based encryption (ABE) [6], FIBE can allow one to express how to share data more flexibly. With this feature, FIBE and its variants have found several applications in fine-grained access control [7], privacy protection [8], e-healthcare cloud system [9], and keyword search system [10]. Therefore, the research on FIBE is of great significance in both theory and practice.

Recently, China has issued many cryptography algorithms to satisfy the strategic demand for independent and safe controllability. Among these algorithms, SM9 [11] refers to the series of identity-based cryptography algorithms that have become the international standard. However, SM9 aims to meet the common basic security requirements (i.e., data confidentiality and integrity) of information systems. More particularly, it does not support fault tolerance or threshold access control, which seriously hinders the practical application of SM9. So far, there are no FIBE schemes based on SM9 in publicly accessible venues.

1.1 Our contribution

This paper initiates the study of FIBE in the framework of SM9, aiming to build more functional domestic cryptography algorithms. In the **Setup** stage, we define an attribute universe, as well as the corresponding attribute private key and attribute public key. In the **KeyGen** stage, we first use a subset of the attribute universe introduced during the establishment of the system to represent one's identity. Then we calculate the private key component associated with each attribute. We apply Shamir's secret sharing technology to distribute a portion of the system master key to each private key component, allowing our scheme to be fault-tolerant. To be specific, we only need a subset of the private key to perform the decryption, with the size of the subset equal to the fault tolerance value. In addition, because random polynomials are selected differently when generating each user's private key, our scheme can resist collusion attacks. That is, if no single user can decrypt a ciphertext, multiple users who collude will also fail to decrypt. Our ciphertext structure is completely consistent with that of SM9 at the **Encrypt** stage. As a result, our scheme is fully compatible with existing information systems using SM9.

We present two FIBE schemes from SM9. The first scheme falls in the field of a small universe, where the attribute universe is specified at the **Setup** stage. Under the decisional $(k+3)$-bilinear Diffie-Hellman inversion ($(k+3)$-DBDHI) assumption, the scheme we introduce is ciphertext-indistinguishable against fuzzy selective-identity and chosen-plaintext attacks (IND-FSID-CPA) in the random oracle (RO) model. Furthermore, we propose a large universe SM9 FIBE design in which the attribute universe is Z_p^*, and the size of public parameters is only proportional to the maximum number of attributes contained in an identity we use when encrypting or generating the private key. We claim that the design has

ciphertext-indistinguishable against chosen-plaintext attacks (IND-CPA) security in the RO model if the generic decisional Diffie-Hellman exponent ((f, g)-GDDHE) problem is difficult. In particular, the security of both schemes can be upgraded to the scenario of chosen-ciphertext attacks (CCA) using FO transformation techniques [12]. Finally, we conduct extensive experiments to demonstrate that both schemes are feasible.

1.2 Related work

IBE. PKI (public-key infrastructure) is the traditional way to deploy public-key encryption. In PKI, there exists a third party called CA, whose major function is managing public-key certificates. Before data encryption, one must first check whether the recipient's public-key certificate is valid. While PKI solves the issue of public-key authentication, it brings a significant burden to the system because of certificate management. The IBE proposed by Shamir [4] can effectively eliminate the issue of certificate management. When encrypting a message in the new framework, one can use her identity (e.g., phone number and mail address) as a public key. Consequently, public-key certificates are no longer required to be maintained. Subsequently, a practical IBE design using bilinear mapping was described in [13], resulting in a new development in identity-based cryptography. Nevertheless, all these schemes work in the RO model. An enormous amount of studies have contributed greatly to the construction of IBE without the RO model. As a consequence, Canetti et al. [14], in 2003, gave a weak definition of security (namely, selective-identity security) and proposed an IBE scheme of selective-identity security without the RO. In 2004, Boneh and Boyen [15] put forward two IBE schemes of selective-identity security without the RO model. In the same year, they [16] also proposed a completely secure IBE scheme without the RO model. Waters [17] proposed a dual system encryption technology in 2009, which can be applied to build and prove a completely secure IBE system. In 2017, Döttling and Garg [18] provided fully secure IBE and hierarchical IBE schemes, and the highlight of their schemes is that the bilinear mapping is not used. Based on this work, they [19] also put forward a general construction of fully secure IBE with reference to selective-identity secure IBE. In recent years, IBE research mainly focused on its applications in various situations, such as wireless sensor networks [20], large-scale 5G [21], fog computing [22], and cloud computing [23].

FIBE. The concept of FIBE was first proposed by Sahai and Waters [5]. Soon after, based on the construction of Sahai-Waters, Pirretti et al. [24] suggested a secure and highly efficient FIBE scheme using the RO model. In 2007, a new and more efficient FIBE construction using the RO model was given in [25]. However, all of these schemes have CPA security. Based on [25], Shi et al. [26] proposed an FIBE construction in 2009, which achieves CCA security in the RO model. In 2010, Ren et al. [27] put forward an FIBE scheme without the RO model, which, however, has been proved to be insecure in [28, 29]. In 2016, Mao et al. [30] presented a completely secure FIBE design without the RO model with a shorter public parameter. Nowadays, FIBE schemes are applied in many situations, including the Internet of medical things [31, 32] and the opportunistic network [33].

SM9. SM9 refers to identity-based cryptographic algorithms and schemes, which are both Chinese and international standards. In 2018, Cheng [34] presented the formal security proof of the SM9 key agreement and encryption algorithm. In 2019, Shi et al. [35] put forward a ciphertext-policy ABE (CP-ABE) scheme from SM9. In 2020, Sun et al. [36] explored the user revocation issue in SM9 and proposed a secure user revocation scheme with robustness. Mu et al. [37] presented a secure two-party signing protocol based on SM9 by utilizing the additive homomorphic property from [38]. Ji et al. [39] also suggested an efficient CP-ABE design from SM9 in 2021, which serves as a dispatching and control cloud. Lai et al. [40] introduced an identity-based signature scheme from SM9 that supports online/offline operations and is about 99% faster than the original SM9 identity-based signature scheme when a signature is generated. In the same year, they [41] proved the security of the SM9 digital signature algorithm under a new complexity assumption and proposed a new SM9 key encapsulation algorithm. Nevertheless, there is currently no design for FIBE based on SM9.

1.3 Organization

Section 2 mainly introduces some preparatory knowledge that this paper uses, including bilinear groups and security assumptions, which is followed by the formal definition of FIBE. Our first FIBE construction based on SM9 and its security proof are placed in Section 3. Our second construction and its security

proof are placed in Section 4. The performance analysis for both schemes is carried out in Section 5. Section 6 summarizes this paper.

2 Background

We first present a concise review about background knowledge, such as bilinear groups, relevant complexity assumptions, and formal definitions of FIBE.

2.1 Bilinear groups

Let a set BP consist of five elements $\{G_1, G_2, G_T, e, p\}$, in which G_1, G_2 are two additive cyclic groups, G_T is a multiplicative cyclic group, and p is a prime number that represents the order of a group. We call this set BP a bilinear group, if there exists an efficiently computable mapping $e : G_1 \times G_2 \to G_T$, that satisfies two operational principles [42].

(1) Bilinearity: for all $P_1 \in G_1, P_2 \in G_2$ and $a, b \in Z_p$, we have $e(aP_1, bP_2) = e(P_1, P_2)^{ab}$.

(2) Non-degeneracy: $e(P_1, P_2) \neq 1_{G_T}$.

2.2 Complexity assumptions

Let a bilinear group BP$= (G_1, G_2, G_T, e, p)$, in which G_1 and G_2 are generated by P and Q respectively.

Definition 1 (Decisional bilinear Diffie-Hellman inversion (q-DBDHI) problem [42])**.** Given $(q + 2)$ elements $(P, Q, aQ, a^2Q, \ldots, a^qQ)$, and a random group element $T \in G_T$, decide whether $T = e(P, Q)^{\frac{1}{a}}$ or not.

Definition 2 (General decisional Diffie-Hellman exponent ((f, g)-GDDHE) problem [43])**.** Given

$$P_0, aP_0, a^2P_0, \ldots, a^kP_0, a^2f(a)P_0, ba^2f(a)P_0,$$

$$H_0, aH_0, a^2H_0, \ldots, a^{2m}H_0, bg(a)H_0,$$

and a random group element $T \in G_T$, where $H_0 \in G_1$, $P_0 \in G_2$, and $f(x)$ and $g(x)$ are coprime polynomials of degree k and m respectively, decide whether $T = e(H_0, P_0)^{baf(a)}$ or not.

2.3 Definition and security model of FIBE

Generally speaking, an FIBE scheme should be constructed from the following four procedures [5], which is described in Figure 2.

Setup(λ)**.** Providing a security parameter λ, this procedure generates the master secret key msk and the public parameters params containing the attribute universe U and fault tolerance value d.

KeyGen(msk, w)**.** Providing the master secret key msk and an identity w, this procedure generates the private key D_w.

Encrypt(params, w, M)**.** Providing the public params, an identity w and a plaintext message M, this procedure generates the ciphertext C.

Decrypt$(\text{params}, D_{w'}, C)$**.** Providing the public params, a private key $D_{w'}$ containing an identity w', and a ciphertext C containing another identity w, this procedure decrypts this ciphertext and generates the corresponding plaintext M or $\perp$ (an error symbol).

Correctness. An FIBE construction is correct when provided any message M, any identity w and w' satisfying $|w \cap w'| \geqslant d$, the following equation holds:

$$\textbf{Decrypt}\left(\textbf{Encrypt}(M, w, \textbf{Setup}(\lambda)), \textbf{KeyGen}(w', \text{msk})\right) = M.$$

The first security requirement of FIBE is proposed in [5], i.e., the IND-FSID-CPA security model. It adds an additional initialization phase than the traditional IND-CPA security model. As shown in Figure 3, there exist an attacker and a challenger who interact with each other to complete the following game.

Initialization. The attacker first claims that the challenge identity α is the attack target.

Setup. The challenger first runs the **Setup** procedure, and then transmits the public params to the attacker.

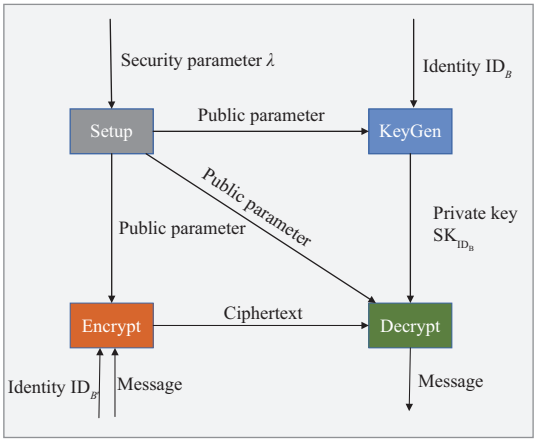


Figure 2 (Color online) Work-flow of FIBE.

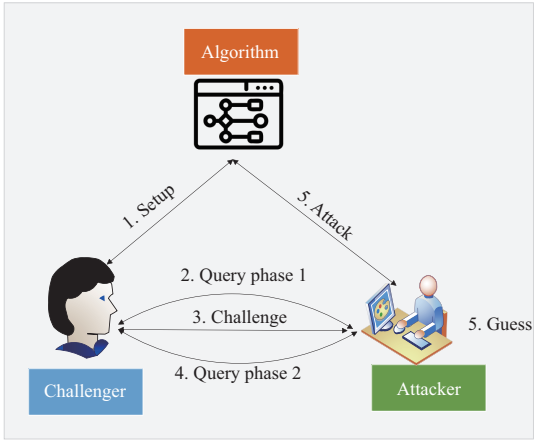


Figure 3 (Color online) Game of IND-CPA.

Query phase 1. At this stage, the private key related to any identity γ_j that meets $|\gamma_j \cap \alpha| < d$ for all j can be adaptively queried by the attacker.

Challenge. At this stage, the Challenger will first receive two equal length messages M_0, M_1 from the attacker. Then it selects randomly $v \in \{0,1\}$, and encrypts the message M_v with the challenge identity α to obtain the ciphertext. Finally, it delivers the resulting ciphertext to the attacker.

Query phase 2. Repeat **Query phase 1**.

Guess. The attacker gives the guess $v' \in \{0,1\}$.

Definition 3. If the advantage of all PPT attackers in the interactive game is negligible, then we state an FIBE is IND-FSID-CPA secure, where the advantage can be calculated as $\text{Adv}=|\Pr(v'=v)-\frac{1}{2}|$.

3 FIBE based on SM9: small universe

This section focuses on the construction of an efficient FIBE design within the framework of SM9, which is selective-identity secure under the Decisional $(k+3)$-BDHI assumption (defined in Subsection 2.2). As in [5], we can also calculate the minimum set overlap to represent the fault tolerance value d, that is, if at least d common attributes are contained in both the private key components and the ciphertext components, we can perform decryption. Otherwise, the decryption would fail. The complete overview of our scheme is provided below.

3.1 Construction

Let a bilinear group BP$=\{G_1, G_2, G_T, e, N\}$[1)], and a hash function $H_1 : \{0,1\}^* \to Z_N^*$. In addition, if G_1 is generated by P, and G_2 is generated by Q, then there exists a valid openly computable isomorphic mapping ψ: $G_2 \to G_1$[2)], that satisfies $\psi(Q) = P$. Their choices are exactly the same as those in SM9. Pick a one-byte appendix hid, and select the first k elements in Z_N^* to form the attribute universe U, that is, $U = \{1, 2, \ldots, k\}$ and $|U| = k$. The Lagrange coefficient $\Delta_{x_i,S}(x)$ for $x_i \in Z_N^*$ and $S \subseteq Z_N^*$ is defined as

$$\Delta_{x_i,S}(x) = \prod_{x_l \in S, x_l \neq x_i} \frac{x - x_l}{x_i - x_l}.$$

Setup$(\lambda) \to (\text{mpk}, \text{msk})$. This algorithm picks the random generator $P_1 \in G_1, P_2 \in G_2$ where $\psi(P_2) = P_1$, and $a \in Z_N$. It computes $P_{\text{pub}} = aP_2$, sets $g = e(P_1, P_{\text{pub}})$, chooses $t_1, t_2, \ldots, t_k$ randomly from Z_N,

1) In SM9, N is a prime number, representing the order of a group.

2) The computable isomorphic map is not required by the scheme but is used in the security proof with the chosen complexity assumption.

and computes

$$\begin{aligned} T_1 &= t_1(H_1(\text{“1”}\|\text{hid}, N)P_1 + \psi(P_{\text{pub}})),\\ T_2 &= t_2(H_1(\text{“2”}\|\text{hid}, N)P_1 + \psi(P_{\text{pub}})),\\ &\vdots\\ T_k &= t_k(H_1(\text{“}k\text{”}\|\text{hid}, N)P_1 + \psi(P_{\text{pub}})). \end{aligned}$$

The public parameters $\text{mpk} = \{P_1, P_2, P_{\text{pub}}, g, T_1, T_2, \ldots, T_k\}$, and the master private key $\text{msk} = \{t_1, t_2, \ldots, t_k, a\}$.

KeyGen$(\text{mpk}, \text{msk}, w) \to \text{sk}_w$. Provided an identity $w \subseteq U$, this algorithm picks a random polynomial q of degree $d-1$, that is $q(x) = a + a_1x + a_2x^2 + \cdots + a_{d-1}x^{d-1}$, which satisfies $q(0) = a$. For every $i \in w$, the algorithm computes

$$D_i = \left[\frac{1}{H_1(\text{“}i\text{”}\|\text{hid}, N) + a} \cdot \frac{q(i)}{t_i}\right] P_2.$$

All private key components $D_w = (D_i)_{i\in w}$ are given to the identity owner.

Encrypt$(\text{mpk}, M, w') \to \text{CT}$. Initially, to generate a ciphertext that was encrypted under an identity w' and a message M, this algorithm picks randomly $s \in Z_N$ and calculates

$$\text{CT} = (w', C = Mg^s, \{C_i = sT_i\}_{i\in w'}).$$

Decrypt$(\text{params}, D_w, \text{CT}) \to M$. Given a private key D_w containing an identity w and a ciphertext CT containing another identity w', if $\mid w \cap w' \mid \geqslant d$, this algorithm chooses a random subset S containing d elements from $w \cap w'$ and computes:

$$M = \frac{C}{\prod_{i\in S}(e(C_i, D_i))^{\Delta_{i,S}(0)}}.$$

The following derivation gives the verification of the **correctness** of the decryption algorithm:

$$\begin{aligned} M &= \frac{C}{\prod_{i\in S} e(C_i, D_i)^{\Delta_{i,S}(0)}}\\ &= \frac{C}{\prod_{i\in S} e(st_i(H_1(\text{“}i\text{”}\|\text{hid}, N)P_1 + aP_1), \frac{q(i)P_2}{(H_1(\text{“}i\text{”}\|\text{hid},N)+a)t_i})^{\Delta_{i,S}(0)}}\\ &= \frac{C}{\prod_{i\in S} e(sP_1, q(i)P_2)^{\Delta_{i,S}(0)}} = \frac{C}{e(sP_1, P_2)^{\sum_{i\in S} q(i)\Delta_{i,S}(0)}}\\ &= \frac{C}{e(sP_1, P_2)^a} = \frac{C}{e(P_1, P_{\text{pub}})^s} = \frac{C}{g^s} = M. \end{aligned}$$

3.2 Security analysis

This part focuses on a formal security analysis of our FIBE design based on SM9. The proof is motivated by those in [5, 42].

Theorem 1. Our FIBE design on the basis of SM9 has IND-FSID-CPA security under the RO model if the $(k+3)$-DBDHI problem is difficult.

Proof. Assume that an attacker $\mathcal{A}$ can win the IND-FSID-CPA security game with the advantage of at least ε after a series of private key queries. Then we can construct a challenger $\mathcal{B}$ to figure out the $(k+3)$-DBDHI problem.

Suppose that provided a $(k+3)$-DBDHI description $(P, Q, aQ, a^2Q, \ldots, a^{k+3}Q, T)$, where $P \in G_1$, and $Q \in G_2$, $\mathcal{B}$ needs to determine whether $T = e(P, Q)^{\frac{1}{a}}$. We define the universe U as $\{1, 2, \ldots, k\}$.

Init. $\mathcal{A}$ sends the challenge identity, α, to $\mathcal{B}$.

Setup. $\mathcal{B}$ constructs the system parameters through the following procedures.

- $\mathcal{B}$ selects k different random numbers $h_1, h_2, \ldots, h_k \in Z_N$, and expands $f(z) = \prod_{i=1}^{k}(z + h_i)$ to gain $c_0, c_1, \ldots, c_k \in Z_N$ satisfying $f(z) = \sum_{i=0}^{k} c_i z^i$.
- $\mathcal{B}$ calculates generators $P_2 = \sum_{i=0}^{k} c_i a^i Q = \sum_{i=0}^{k} c_i (a^i Q)$ and $P_1 = \psi(P_2)$, where $\psi(Q) = P$.

• The public key P_{pub}, g are set as $P_{\text{pub}} = aP_2 = af(a)Q = \sum_{i=0}^{k} c_i(a^{i+1}Q)$, and $g = e(P_1, P_{\text{pub}})$.

• $\mathcal{A}$ can query the hash function as follows:

When $\mathcal{A}$ asks for a query $i \in [1,k]$ to H_1, $\mathcal{B}$ answers $H_1(\text{“}i\text{”}\|\text{hid}, N) = h_i$, and then stores (i, h_i) in a list L.

• For all $i \in \alpha$, $\mathcal{B}$ computes $T_i = \beta_i h_i a^2 P_1 + \beta_i a^3 P_1$, where β_i is selected randomly from Z_N and implicitly have $t_i = a^2\beta_i$. For all $i \notin \alpha$, it computes $T_i = \omega_i h_i P_1 + \omega_i a P_1$, where ω_i is selected randomly from Z_N and implicitly have $t_i = w_i$. Here, aP_1, a^2P_1, a^3P_1 can be computed by the above parameters.

Finally $\mathcal{B}$ sends $\mathcal{A}$ mpk $= \{P_1, P_2, P_{\text{pub}}, g, T_1, T_2, \ldots, T_k\}$ as params. Note that all parameters in the scheme construction are randomly chosen from the view of $\mathcal{A}$. For $\forall i \in [1,k]$, $\mathcal{B}$ can expand $f_i(z) = \frac{f(z)}{z+h_i} = \sum_{i=0}^{k-1} d_i z^i$, and compute $\sum_{i=0}^{k-1} d_i(a^iQ) = \frac{f(a)}{a+h_i}Q = \frac{1}{a+h_i}P_2$. Therefore, although $\mathcal{B}$ does not know α, when given h_i, $\frac{1}{a+h_i}P_2$ is computable.

Phase 1. $\mathcal{A}$ can adaptively query the private key related with any identity γ that meets the condition $|\gamma \cap \alpha| < d$. First, we define the set

$$\Gamma = \gamma \cap \alpha.$$

Then, we select a set Γ' to ensure that

$$\Gamma \subseteq \Gamma' \subseteq \gamma, \quad \text{and} \quad |\Gamma'| = d - 1.$$

Finally, let

$$S = \Gamma' \cup \{0\}.$$

For $i \in \Gamma'$, $\mathcal{B}$ can calculate in the following two cases.

• If $i \in \Gamma$, $\mathcal{B}$ picks $\lambda_i \in Z_N$ at random and computes $D_i = \frac{\lambda_i}{(h_i+a)\beta_i}P_2$.

• Otherwise, $\mathcal{B}$ picks $s_i \in Z_N$ at random and computes $D_i = \frac{s_i}{(h_i+a)w_i}P_2$.

The calculation is correct because a random $d-1$ degree polynomial $q(x)$ is implicitly selected, where

$$q(i) = \begin{cases} a, & i = 0, \\ a^2\lambda_i, & i \in \Gamma, \\ s_i, & i \in \Gamma' - \Gamma. \end{cases}$$

For $i \notin \Gamma'$, $\mathcal{B}$ can also calculate

$$D_i = \frac{1}{(h_i+a)w_i}P_2 \sum_{j \in \Gamma' - \Gamma} s_j \Delta_{j,S}(i) + \frac{a^2}{(h_i+a)w_i}P_2 \sum_{j \in \Gamma} \lambda_j \Delta_{j,S}(i) + \frac{a}{(h_i+a)w_i}P_2 \Delta_{0,S}(i).$$

where $\frac{a^2}{h_i+a}P_2 = \frac{a^2}{h_i+a}f(a)Q = a^2 f_i(a)Q = \sum_{i=2}^{k+1} d_{i-2}(a^iQ)$ and $\frac{a}{h_i+a}P_2 = \frac{a}{h_i+a}f(a)Q = af_i(a)Q = \sum_{i=1}^{k} d_{i-2}(a^iQ)$.

Therefore, for $i \notin \Gamma'$, $\mathcal{B}$ is able to calculate $D_i = \frac{q(i)}{(h_i+a)w_i}P_2$ using interpolation.

Consequently, $\mathcal{B}$ can construct a private key with the same distribution as that from the real construction.

Challenge. $\mathcal{A}$ first sends M_0 and M_1 to $\mathcal{B}$ as the challenge message. Then $\mathcal{B}$ selects randomly $v \in \{0,1\}$ and computes as the following.

$\mathcal{B}$ randomly selects $t \in Z_N$, and sets

$$\text{CT} = (\alpha, C = M_v T', \{C_i = t\beta_i(h_i P_1 + aP_1)\}_{i \in \alpha}),$$

where $T' = [e(P_1 + c_0 P, \sum_{i=0}^{k-1} c_{i+1}a^iQ) \cdot T^{{c_0}^2}]^t$.

In the case $T = e(P,Q)^{\frac{1}{a}}$, we set $s = \frac{t}{a^2}$ and have $C_i = sa^2\beta_i(h_iP_1 + aP_1) = sT_i$ and $T' = [e(P_1 + c_0P, \sum_{i=0}^{k-1} c_{i+1}a^iQ) \cdot T^{{c_0}^2}]^t = e(P_1, P_2)^{as} = g^s$. Hence the ciphertext created above is an encryption of message M_v under the public key α.

Otherwise, because T is random, no information of the message M_v can be leaked by this ciphertext.

Phase 2. Repeat **Phase 1**.

Guess. $\mathcal{A}$ ultimately submits a guess v' about v. $\mathcal{B}$ then guesses $T = e(P,Q)^{\frac{1}{a}}$ if $v' = v$. Otherwise, $\mathcal{B}$ believes T is random.

For ciphertext $T' = g^s$, the calculation is as follows:

$$\begin{aligned}
T' &= \left[e\left(P_1 + c_0 P, \sum_{i=0}^{k-1} c_{i+1} a^i Q \right) \cdot T^{{c_0}^2} \right]^t \\
&= e\left((P_1 + c_0 P), \sum_{i=0}^{k-1} c_{i+1} a^i Q \right)^t \cdot e(P,Q)^{\frac{{c_0}^2 t}{a}} \\
&= e\left(tP_1, \sum_{i=0}^{k-1} c_{i+1} a^i Q \right) \cdot e\left(c_0 tP, \sum_{i=0}^{k-1} c_{i+1} a^i Q \right) \cdot e(P,Q)^{\frac{{c_0}^2 t}{a}} \\
&= e(P,Q)^{f(a)t \sum_{i=0}^{k-1} c_{i+1} a^i + c_0 t \sum_{i=0}^{k-1} c_{i+1} a^i + \frac{{c_0}^2 t}{a}} \\
&= e(P,Q)^{\frac{f^2(a)t}{a}}.
\end{aligned}$$

Let us further explain the equation $e(P,Q)^{f(a)t \sum_{i=0}^{k-1} c_{i+1} a^i + c_0 t \sum_{i=0}^{k-1} c_{i+1} a^i + \frac{{c_0}^2 t}{a}} = e(P,Q)^{\frac{f^2(a)t}{a}}$. Because $\frac{f^2(a)t}{a} = \frac{f(a)}{a} \cdot f(a)t = (\sum_{i=0}^{k-1} c_{i+1} a^i + \frac{c_0}{a}) f(a)t = f(a)t \sum_{i=0}^{k-1} c_{i+1} a^i + c_0 t \cdot \frac{f(a)}{a} = f(a)t \sum_{i=0}^{k-1} c_{i+1} a^i + c_0 t \sum_{i=0}^{k-1} c_{i+1} a^i + \frac{{c_0}^2 t}{a}$, so the equation holds.

As a result, $T' = e(P,Q)^{\frac{f^2(a)t}{a}} = e(f(a)P, f(a)Q)^{\frac{t}{a}} = e(P_1, P_2)^{\frac{t}{a}} = e(P_1, P_2)^{as} = g^s$.

Finally, we evaluate the advantages of $\mathcal{B}$ solving the $(k+3)$-DBDHI problem.

• When $T = e(P,Q)^{\frac{1}{a}}$, the advantage of $\mathcal{A}$ correctly guessing v is ε by definition. As a result, if $\mathcal{A}$ outputs $v' = v$, the probability that $\mathcal{B}$ guesses $T = e(P,Q)^{\frac{1}{a}}$ is at least $\varepsilon + \frac{1}{2}$.

• When $T \neq e(P,Q)^{\frac{1}{a}}$, $\mathcal{A}$ gains no information about M_v. As a result, if $\mathcal{A}$ output $v' \neq v$, the probability that $\mathcal{B}$ guesses $T = e(P,Q)^{\frac{1}{a}}$ is $\frac{1}{2}$.

To sum up, the advantage of $\mathcal{B}$ figuring out this $(k+3)$-DBDHI problem is not less than $\frac{1}{2}(\frac{1}{2} + \varepsilon) + \frac{1}{2} \cdot \frac{1}{2} - \frac{1}{2} = \frac{1}{2}\varepsilon$. This completes our proof.

Remarks. Our scheme has CPA security under the RO model. So we can achieve the CCA security by using the FO transformation techniques [12].

4 FIBE based on SM9: large universe

The above scheme only supports a small attribute universe, and the number of the public params is proportional to the number of attributes in the universe. However, in practical applications, both mpk and msk must be updated if the number of attributes actually required is greater than the predefined bound. To solve this issue, FIBE with a large universe is introduced. Combining the construction idea of the previous scheme with the Boneh and Boyen algebraic structure used in [15], we propose our large universe construction, in which the universe $U = Z_N^*$, and the size of the public params is only proportional to n, which we fix as the maximum size of identity we use when encrypting.

4.1 Description

Let a bilinear group BP= $\{G_1, G_2, G_T, e, N\}$, and a hash function $H_1 : \{0,1\}^* \to Z_N^*$. Their choices are exactly the same as those in SM9. Pick a one-byte appendix hid. The definition of the Lagrange coefficient $\Delta_{x_i,S}(x)$ is the same as that in Subsection 3.1.

Setup$(\lambda) \to (\text{mpk}, \text{msk})$. This algorithm picks randomly the generator $P_1 \in G_1, P_2 \in G_2$, and $a \in Z_N^*$. It computes $P_{\text{pub}} = aP_1$ and $g = e(P_{\text{pub}}, P_2)$. We define a set N' as $N' = \{1, 2, \ldots, n+1\}$ and a function $T(X)$ as

$$T(X) = \sum_{i=1}^{n+1} \Delta_{i,N'}(X) t_i.$$

Here $t_1, t_2, \ldots, t_{n+1}$ are picked randomly from G_2. The public key mpk= $\{P_1, P_2, P_{\text{pub}}, g, t_1, t_2, \ldots, t_{n+1}\}$ and the master secret key msk $= a$.

KeyGen$(\text{mpk}, \text{msk}, w) \to \text{sk}_w$. Initially, provided an identity $w \subseteq U$, the algorithm chooses a random polynomial q of degree $d-1$, which satisfies $q(0) = a$, and for all $i \in w$, computes $h_i = H_1(\text{“}i\text{”} \parallel \text{hid}, N)$.

Then this algorithm picks $r_i \in Z_N$ at random, and computes

$$D_i = \frac{q(i)}{h_i + a} P_2 + r_i T(i), \quad d_i = r_i(h_i P_1 + P_{\text{pub}}).$$

All private key components $D_w = \{D_i, d_i\}_{i \in w}$ are given to the identity owner.

Encrypt$(\text{mpk}, M, w') \rightarrow$CT. Initially, to generate a ciphertext of message M and identity w', this algorithm picks randomly $s \in Z_N$, and computes

$$\text{CT} = (w', C = Mg^s, \{C_i = s(H_1(\text{“}i\text{”} \parallel \text{hid}, N)P_1 + P_{\text{pub}}), C_i' = sT(i)\}_{i \in w'}).$$

Decrypt$(\text{params}, D_w, \text{CT}) \rightarrow M$. Given a ciphertext CT containing an identity w' and a private key D_w containing another identity w, if $\mid w \cap w' \mid \geqslant d$, this algorithm chooses any a subset S containing d elements from $w \cap w'$ and computes

$$M = C \prod_{i \in S} \left(\frac{e(d_i, C_i')}{e(C_i, D_i)} \right)^{\Delta_{i,S}(0)}.$$

The following derivation gives the verification of the **correctness** of the decryption algorithm.

$$\begin{aligned} M &= C \prod_{i \in S} \left(\frac{e(d_i, C_i')}{e(C_i, D_i)} \right)^{\Delta_{i,S}(0)} = Mg^s \prod_{i \in S} \left(\frac{e(d_i, C_i')}{e(C_i, D_i)} \right)^{\Delta_{i,S}(0)} \\ &= Mg^s \prod_{i \in S} \left(\frac{e(r_i(h_i P_1 + P_{\text{pub}}), sT(i))}{e(s(h_i P_1 + P_{\text{pub}}), \frac{q(i)}{h_i + a} P_2 + r_i T(i))} \right)^{\Delta_{i,S}(0)} \\ &= Mg^s \prod_{i \in S} \frac{1}{e(P_1, P_2)^{sq(i)\Delta_{i,S}(0)}} = Mg^s \frac{1}{e(P_1, P_2)^{\sum_{i \in S} sq(i)\Delta_{i,S}(0)}} \\ &= Mg^s \frac{1}{e(P_1, P_2)^{sa}} = Mg^s \frac{1}{g^s} = M. \end{aligned}$$

4.2 Proof of security

In this subsection, we will analyze the security of our large universe construction in the IND-CPA security game.

Theorem 2. If the (f, g)-GDDHE assumption holds, our large universe construction has IND-CPA security under the RO model.

Proof. Assume that an attacker $\mathcal{A}$ can win the IND-CPA security game with the advantage of at least ε after a series of private key queries. Next, we can construct a challenger $\mathcal{B}$ to figure out the (f, g)-GDDHE problem.

First of all, we set the size of the identity used for encryption as m, and the total number of times the RO is queried as k. Then assuming that $\mathcal{B}$ is given a (f, g)-GDDHE description

$$P_0, aP_0, a^2 P_0 \ldots, a^k P_0, a^2 f(a) P_0, ba^2 f(a) P_0,$$

$$H_0, aH_0, a^2 H_0, \ldots, a^{m+1} H_0, bg(a) H_0,$$

and a random group element $T \in G_T$, where $f(x)$ and $g(x)$ are coprime polynomials of degree t and m respectively. Finally, $\mathcal{B}$ needs to decide whether $T = e(H_0, P_0)^{baf(a)}$.

Setup. The system parameters are constructed as follows.

- $\mathcal{B}$ selects $(k + m)$ different random numbers $h_1, h_2, \ldots, h_k, \ldots, h_{k+m} \in Z_N$, and defines $f(x) = \prod_{i=1}^{k}(x + h_i)$ and $g(x) = \prod_{i=k+1}^{k+m}(x + h_i)$. For $\forall i \in [1, k]$, $\mathcal{B}$ defines $f_i(x) = \frac{f(x)}{x + h_i}$. For $\forall i \in [k+1, k+m]$, $\mathcal{B}$ defines $g_i(x) = \frac{g(x)}{x + h_i}$.
- $\mathcal{B}$ sets generators $P_2 = f(a)P_0$, and $P_1 = \prod_{i=k+1}^{k+m}(a + h_i)H_0 = g(a)H_0$.
- The public key P_{pub}, g are set as $P_{\text{pub}} = a \prod_{i=k+1}^{k+m}(a + h_i)H_0$, and $g = e(P_{\text{pub}}, P_2)$.
- $\mathcal{B}$ randomly selects a n degree polynomial $c(X)$ and sets $t_i = c(i)P_2$ for $1 \leqslant i \leqslant n + 1$. Note that since $c(X)$ is selected randomly, all t_i will also be randomly chosen from the view of $\mathcal{A}$. In addition, we implicitly have $T(i) = c(i)P_2$.

Finally, $\mathcal{B}$ sends $\mathcal{A}$ the public params $\text{mpk} = (P_1, P_2, P_{\text{pub}}, g, t_1, t_2, \ldots, t_{n+1})$, and makes the master private key a secret.

Although $\mathcal{B}$ does not know a, when given $h_i (i \in [1,k])$, $\frac{1}{a+h_i}P_2$ and $\frac{a}{a+h_i}P_2$ are computable.

Phase 1. $\mathcal{A}$ first adaptively makes requests for an RO H_1, which is under control of $\mathcal{B}$ according to the following procedure:

When $\mathcal{A}$ asks for a query $i \in [1,k]$ to H_1, $\mathcal{B}$ answers $H_1(\text{“}i\text{”}||\text{hid}, N) = h_i$, and then stores (i, h_i) in a list L.

Then, $\mathcal{A}$ adaptively makes requests for a private key containing any identity γ. We first define a set Γ, which satisfies

$$\Gamma \subseteq \gamma \quad \text{and} \quad |\Gamma| = d - 1.$$

Let

$$S = \Gamma \cup \{0\}.$$

If $i \in \Gamma$, $\mathcal{B}$ picks $r_i \in Z_N$ at random and computes

$$D_i = \frac{q(i)}{h_i + a}P_2 + r_i T(i), \quad d_i = r_i(h_i P_1 + P_{\text{pub}}).$$

The calculation is correct because a $d-1$ degree polynomial $q(x)$ satisfying $q(0) = a$ is randomly and implicitly chosen.

If $i \in \gamma - \Gamma$, $\mathcal{B}$ picks $r_i' \in Z_N$ at random. To be consistent with the $q(x)$ we implicitly selected, it first computes $q(j)$ using the polynomial $q(x)$ which is chosen above, then computes

$$D_i = \frac{\sum_{j\in\Gamma} q(j)\Delta_{j,S}(i)}{h_i + a}P_2 + \frac{a\Delta_{0,S}(i)}{h_i + a}P_2 + r_i' T(i),$$

and

$$d_i = r_i' \left(h_i P_1 + P_{\text{pub}}\right).$$

Here, $\frac{1}{a+h_i}P_2$ and $\frac{a}{a+h_i}P_2$ can be calculated by the above parameters. Therefore, the private key containing the identity γ can be generated in the above way, and its distribution is the same as that in the original scheme.

Challenge. $\mathcal{A}$ submits two challenge messages M_0, M_1 and a challenge identity α, such that $|\alpha \cap \gamma| < d$ for all attribute sets γ that have been queried for the private key. $\mathcal{B}$ randomly picks $v \in \{0,1\}$ and sets the ciphertext as $\text{CT} = \left(\alpha, C, \{C_i, C_i'\}_{i\in\alpha}\right)$.

Here

$$C = M_v T^{\prod_{i=k+1}^{k+m} h_i} \cdot e\left(\frac{1}{a}\left(\prod_{i=k+1}^{k+m}(h_i + a) - \prod_{i=k+1}^{k+m} h_i\right) H_0, ba^2 f(a) P_0\right),$$

$$C_i = b(h_i P_1 + P_{\text{pub}}),$$

$$C_i' = \text{bc}(i) P_2.$$

In the case $T = e(H_0, P_0)^{baf(a)}$, we set $s = b$, and have

$$\begin{aligned}
C &= M_v T^{\prod_{i=k+1}^{k+m} h_i} \cdot e\left(\frac{1}{a}\left(\prod_{i=k+1}^{k+m}(h_i + a) - \prod_{i=k+1}^{k+m} h_i\right) H_0, ba^2 f(a) P_0\right) \\
&= M_v e(H_0, P_0)^{baf(a)\prod_{i=k+1}^{k+m} h_i} \cdot e\left(\frac{1}{a}\left(\prod_{i=k+1}^{k+m}(h_i + a) - \prod_{i=k+1}^{k+m} h_i\right) H_0, ba^2 f(a) P_0\right) \\
&= M_v e\left(H_0, P_0\right)^{baf(a)\prod_{i=k+1}^{k+m}(h_i+a)} \\
&= M_v e\left(a\prod_{i=k+1}^{k+m}(h_i + a) H_0, f(a) P_0\right)^b \\
&= M_v e\left(P_{\text{pub}}, P_2\right)^b \\
&= M_v g^s,
\end{aligned}$$

$$C_i = b(h_i P_1 + P_{\text{pub}}) = s(h_i P_1 + P_{\text{pub}}),$$

and

$$C_i' = bc(i)P_2 = sc(i)P_2 = sT(i).$$

Here $\frac{1}{a}(\prod_{i=k+1}^{k+m}(h_i + a) - \prod_{i=k+1}^{k+m} h_i)$ is a polynomial of degree $m-1$, and $\frac{1}{a}(\prod_{i=k+1}^{k+m}(h_i + a) - \prod_{i=k+1}^{k+m} h_i)H_0$ can be calculated through the problem instance. Hence the ciphertext created above is encrypted under the identity α and message M_v.

Otherwise, $T \neq e(H_0, P_0)^{baf(a)}$. Because T is random, no information of the message M_v can be leaked by this ciphertext.

Phase 2. Repeat **Phase 1**, but $\mathcal{A}$ cannot query the private key of the attribute set γ that meets $|\alpha \cap \gamma| \geqslant d$.

Guess. $\mathcal{A}$ will ultimately produce a guess v' about v, and $\mathcal{B}$ then guesses $T = e(H_0, P_0)^{baf(a)}$ if $v' = v$. Otherwise, $\mathcal{B}$ believes T is random.

Finally, we evaluate the advantages of $\mathcal{B}$ solving the (f, g)-GDDHE problem.

Since $\mathcal{B}'$s advantage in solving difficult problems is related to the probability of not aborting the game and breaking the challenge ciphertext, we calculate this advantage in the following two steps.

• **Probability of not aborting.** Suppose that $\mathcal{A}$ has asked for q private key queries in the **Phase 1** and **Phase 2**. This game does not abort if and only if $|\alpha \cap \gamma_i| < d$, where α represents the challenge attribute set, and γ_i $(i = 1, 2, \ldots, q)$ represents the attribute set queried in the **Phase 1** and **Phase 2**. Therefore, we have

$$P(\overline{\text{abort}}) = \left(\frac{C_n^0 C_{N-n}^n + C_n^1 C_{N-n}^{n-1} + \cdots + C_n^{d-1} C_{N-n}^{n-d+1}}{C_N^n}\right)^q.$$

In this formula,

$$\frac{C_n^0 C_{N-n}^n}{C_N^n} = \frac{(N-n)!}{n!(N-2n)!} \cdot \frac{n!(N-n)!}{N!} = \frac{(N-2n+1)(N-2n+2)\cdots(N-n)}{(N-n+1)(N-n+2)\cdots N} \to 1.$$

$$\begin{aligned}\frac{C_n^r C_{N-n}^{n-r}}{C_N^n} &= \frac{n!(N-n)!(N-n)!n!}{r!(n-r)!(n-r)!(N-2n+r)!N!} = \frac{(n!)^2}{r!((n-r)!)^2} \cdot \frac{(N-n)!(N-n)!}{(N-2n+r)!N!} \\ &= \frac{(n!)^2}{r!((n-r)!)^2} \cdot \frac{(N-2n+r+1)(N-2n+r+2)\cdots(N-n)}{(N-n+1)(N-n+2)\cdots N} \to 0 \ (0 < r < d).\end{aligned}$$

As a result, the probability of not aborting this game is close to 1.

• **Probability that breaking the challenge ciphertext.** If $T = e(H_0, P_0)^{baf(a)}$, $\mathcal{A}$'s probability of correctly guessing the encrypted ciphertext is at least $\varepsilon + \frac{1}{2}$ by definition. Otherwise, $\mathcal{A}$ gains no information about M_v, $\mathcal{A}$'s probability of guessing the encrypted ciphertext correctly is $\frac{1}{2}$. As a consequence, $\mathcal{A}$'s advantage of breaking the challenge ciphertext is not less than $\frac{1}{2} \cdot (\varepsilon + \frac{1}{2}) + \frac{1}{2} \cdot \frac{1}{2} - \frac{1}{2} = \frac{1}{2}\varepsilon$.

To sum up, the overall advantage of $\mathcal{B}$ figuring out the (f, g)-GDDHE problem is not less than $1 \cdot \frac{1}{2}\varepsilon = \frac{1}{2}\varepsilon$. This completes our proof.

5 Performance analysis

The performance of our schemes is evaluated by comparing with the most classic Sahai-Waters schemes, which we mark with SW.

5.1 Theoretical analysis

In this subsection, we evaluate the communication and computation cost theoretically. The result of communication cost is listed in Table 1. Here, $|G_i|$ shows the bit length of a member of the group G_i. k, m represent the number of attributes contained in a universe and identity respectively. We notice that our small universe case is completely consistent with the small universe case of SW when it comes to communication cost. Because of the characteristics of SM9 private-key structure, our large universe

Table 1 Comparison of communication overhead of our schemes and SW's schemes

	Size of params	Size of private key	Size of ciphertext	Secure model	Difficult problem
SW's small case	$k\|G_1\| + \|G_T\|$	$m\|G_1\|$	$m\|G_1\| + \|G_T\|$	IND-FSID-CPA	MBDH
Our small case	$k\|G_1\| + \|G_T\| + \|G_2\|$	$m\|G_2\|$	$m\|G_1\| + \|G_T\|$	IND-FSID-CPA	DBHI
SW's large case	$(m+2)\|G_1\|$	$2m\|G_1\|$	$(m+1)\|G_1\| + \|G_T\|$	IND-FSID-CPA	MBDH
Our large case	$(m+2)\|G_2\| + \|G_T\|$	$m(\|G_1\| + \|G_2\|)$	$m(\|G_1\| + \|G_2\|) + \|G_T\|$	IND-CPA	GDDHE

Table 2 Comparison of computation overhead of our schemes and SW's schemes

	Cost of extract	Cost of encrypt	Cost of decrypt
SW's small case	$m \cdot \mathrm{sm1}$	$m \cdot \mathrm{sm1} + T_e$	$d \cdot T_p + T_i$
Our small case	$m \cdot \mathrm{sm2}$	$m \cdot \mathrm{sm1} + T_e$	$d \cdot T_p + T_i$
SW's large case	$3m \cdot \mathrm{sm1}$	$(m+1) \cdot \mathrm{sm1} + T_e + T_p$	$d(2T_p + T_e + T_i)$
Our large case	$2m(\mathrm{sm1} + \mathrm{sm2})$	$m(2\mathrm{sm1} + \mathrm{sm2}) + T_e$	$d(2T_p + T_e + T_i)$

Table 3 Comparison of simulation time of our schemes and SW's schemes (ms)

	Setup	KeyGen	Encrypt	Decrypt
SW's small case	14	1	1	90
Our small case	16	2	1	90
SW's large case	0.8	18	31	250
Our large case	13	22	19	250

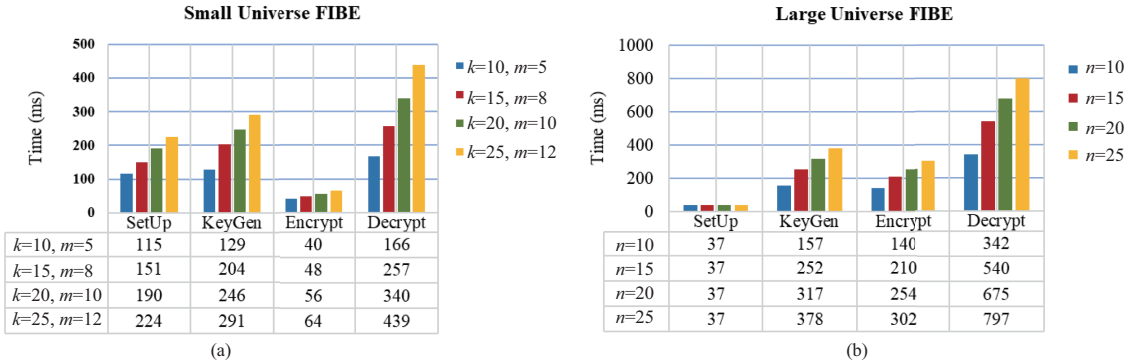
design has m more elements of group G_2 than the large universe case of SW in terms of ciphertext structure.

The evaluation of computation cost is listed in Table 2. Since the bilinear group used in the comparison schemes is a multiplicative group and ours use additive group, for the convenience of comparison, we define the following symbols: sm$i$ $(i = 1, 2)$ represents the time for calculating a scalar multiplication in the additive group G_i $(i = 1, 2)$ of our schemes, but also represents the time for calculating an exponential operation in the multiplicative group of the SW schemes; T_e, T_i denote the time for computing an exponential and inverse in G_T respectively; T_p denotes the time for computing a single pairing; d denotes the specified fault tolerance value. We omit the multiplication operation in G_T and the addition operation in G_i $(i = 1, 2)$ of our schemes, the multiplication operation in G_1 of the SW schemes and the hash operation in the comparison. From Table 2, we can see that our small universe case is completely equivalent to the SW small universe case in terms of computational cost. Our large universe case has m times and $2m$ times more multiplication operation time than the SW large universe case when generating the private key and ciphertext respectively. But in the encryption stage, our scheme does not need any pairing operations. The decryption cost remains the same.

5.2 Evaluation

In this subsection, we analyze the running time through a programming implementation. For the sake of testing, the ψ function in our scheme is ignored. The device used in the test environment is a notebook with the following characteristics: Intel(R) Core(TM) i5-11300H@3.10 GHz 3.11 GHz CPU, 64-bit Windows 10 operating system, 16.0 GB memory. The cryptography library we use is the Multiprecision Integer and Rational Arithmetic C/C++ Library (Miracl) [44]. The programming language used is C++, and the average value is taken after running the program for 100 times. Because the pairing operations are different, we first compare the running time of each sub-algorithm of our schemes and SW schemes in a symmetrical setting at 128-bit security. We translate our schemes to the symmetric setting, namely, we assume $G_1 = G_2$ in our scheme. We note that in the symmetric setting, the hardness assumption and security proof of our schemes are also true. We use the same Tate pairing embedding degree 2 and the super singular curve. In the small universe construction, we assume $k = 25$, $m = 12$, and in the large universe construction, we assume $n = 25$ (n represents the maximum size identity). The results are demonstrated in Table 3, and it is apparent that the time consumption of our schemes is similar to that of SW schemes. It is consistent with the theoretical analysis.

Then, in order to effectively integrate with existing information systems using SM9, we also test the running time of each sub-algorithm of our schemes with different parameters at 256-bit security. We use the same R-ate pairing operation and the BN curve as SM9 IBE algorithm. From Figure 4, it is



	SetUp	KeyGen	Encrypt	Decrypt
k=10, m=5	115	129	40	166
k=15, m=8	151	204	48	257
k=20, m=10	190	246	56	340
k=25, m=12	224	291	64	439

(a)

	SetUp	KeyGen	Encrypt	Decrypt
n=10	37	157	140	342
n=15	37	252	210	540
n=20	37	317	254	675
n=25	37	378	302	797

(b)

Figure 4 (Color online) Running time of each sub-algorithm of our schemes under different parameters. (a) In the small universe case, the parameters ($k = 10$, $m = 5$), ($k = 15$, $m = 8$), ($k = 20$, $m = 10$), and ($k = 25$, $m = 12$) correspond to the system of 126, 6435, 167960, 67603900 users respectively; (b) in the large universe case, the parameters n represents the maximum number of attributes included in an identity when encrypting.

obvious that in the small universe case, even when $k = 25$, $m = 12$, the encryption takes only 64 ms and the decryption time is only 439 ms. In the large universe case, even when $n = 25$, our encryption time is only 302 ms and the decryption time is 797 ms. With further optimization including implementing field operations with assembly code and adopting faster method for multi-pairing computation[3)], the implementation could be at least an order of magnitude faster. In general, our scheme is feasible in practical application.

6 Conclusion and future work

We design two novel FIBE schemes based on SM9 to promote the application of SM9 in modern information systems such as cloud computing and blockchain. The first scheme works in a small attribute universe setting, and the second one accommodates a large attribute universe. Both schemes can be proved to achieve IND-FSID-CPA security and IND-CPA security by reducing them to $(k + 3)$-DBDHI and (f, g)-GDDHE hardness assumptions, respectively. The performance analysis indicates that both schemes are comparable to other existing schemes.

There are several future work directions. First, the decryption operation in our design is expensive in terms of communication and computation costs. Therefore, further research into how to create a more efficient FIBE scheme rooted in SM9 by optimizing or outsourcing decryption is a problem that needs further investigation. Second, although our second scheme implements the property of a large universe, the construction of the scheme depends on the maximum number of attributes included in an identity when encrypting. Therefore, it is also worthwhile to design a completely large universe FIBE scheme based on SM9, where the setting of public parameters is unaffected by the size of the attribute universe or identity at all.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 62032005, 62372108).

References

1 Lu S Q, Zheng J H, Cao Z F, et al. A survey on cryptographic techniques for protecting big data security: present and forthcoming. Sci China Inf Sci, 2022, 65: 201301
2 Guo X J, Li J, Liu Z L, et al. Labrador: towards fair and auditable data sharing in cloud computing with long-term privacy. Sci China Inf Sci, 2022, 65: 152106
3 Zhao Y, Xu K, Li Q, et al. Intelligent networking in adversarial environment: challenges and opportunities. Sci China Inf Sci, 2022, 65: 170301
4 Shamir A. Identity-based cryptosystems and signature schemes. In: Proceedings of the Workshop on the Theory and Application of Cryptographic Techniques, 1985. 47–53
5 Sahai A, Waters B. Fuzzy identity-based encryption. In: Proceedings of the Annual International Conference on the Theory and Applications of Cryptographic Techniques, Berlin, 2005. 457–473
6 Goyal V, Pandey O, Sahai A, et al. Attribute-based encryption for fine-grained access control of encrypted data. In: Proceedings of the 13th ACM Conference on Computer and Communications Security, New York, 2006. 89–98
7 Ghopur D, Ma J F, Ma X D, et al. Puncturable ciphertext-policy attribute-based encryption scheme for efficient and flexible user revocation. Sci China Inf Sci, 2023, 66: 172104
8 Guan Z T, Yang W T, Zhu L H, et al. Achieving adaptively secure data access control with privacy protection for lightweight IoT devices. Sci China Inf Sci, 2021, 64: 162301

3) https://eprint.iacr.org/2019/077.pdf.

9 Zhang K, Li H, Ma J F, et al. Efficient large-universe multi-authority ciphertext-policy attribute-based encryption with white-box traceability. Sci China Inf Sci, 2018, 61: 032102
10 Yao L S, Hou L, Weng J, et al. Provably secure attribute-based authenticated encryption with keyword search from ideal lattices. Sci China Inf Sci, 2024, 67: 119101
11 Cheng Z H. The SM9 cryptographic schemes. 2017. https://eprint.iacr.org/2017/117.pdf
12 Fujisaki E, Okamoto T. Secure integration of asymmetric and symmetric encryption schemes. In: Proceedings of the Advances in Cryptology, Berlin, 1999. 537–554
13 Boneh D, Franklin M. Identity-based encryption from the weil pairing. In: Proceedings of the Annual International Cryptology Conference, Berlin, 2001. 213–229
14 Canetti R, Halevi S, Katz J. A forward-secure public-key encryption scheme. In: Proceedings of the Conference on the Theory and Applications of Cryptographic Techniques, Berlin, 2003. 255–271
15 Boneh D, Boyen X. Efficient selective-ID secure identity-based encryption without random oracles. In: Proceedings of the Theory and Applications of Cryptographic Techniques, Berlin, 2004. 223–238
16 Boneh D, Boyen X. Secure identity based encryption without random oracles. In: Proceedings of the Annual International Cryptology Conference, Berlin, 2004. 443–459
17 Waters B. Dual system encryption: realizing fully secure IBE and HIBE under simple assumptions. In: Proceedings of the Annual International Cryptology Conference, Berlin, 2009. 619–636
18 Döttling N, Garg S. Identity-based encryption from the diffie-hellman assumption. In: Proceedings of the Annual International Cryptology Conference, 2017. 537–569
19 Döttling N, Garg S. From selective IBE to full IBE and selective HIBE. In: Proceedings of the Theory of Cryptography Conference, Berlin, 2017. 372–408
20 Cao C H, Tang Y N, Huang D Y, et al. IIBE: an improved identity-based encryption algorithm for WSN security. Secur Commun Netw, 2021, 2021: 1–8
21 Gupta R K, Almuzaini K K, Pateriya R K, et al. An improved secure key generation using enhanced identity-based encryption for cloud computing in large-scale 5G. Wirel Commun Mob Com, 2022, 2022: 1–14
22 Farjana N, Roy S, Mahi M, et al. An identity-based encryption scheme for data security in fog computing. In: Proceedings of the International Joint Conference on Computational Intelligence, Berlin, 2020. 215–226
23 Qin B D, Liu X M, Wei Z, et al. Space efficient revocable IBE for mobile devices in cloud computing. Sci China Inf Sci, 2020, 63: 139110
24 Pirretti M, Traynor P, McDaniel P, et al. Secure attribute-based systems. In: Proceedings of the 13th ACM Conference on Computer and Communications Security, 2006. 99–112
25 Baek J, Susilo W, Zhou J. New constructions of fuzzy identity-based encryption. In: Proceedings of the 2nd ACM Symposium on Information, Computer and Communications Security, New York, 2007. 368–370
26 Shi W B, Jang I, Yoo H S. Chosen ciphertext secure fuzzy identity-based encryption scheme with short ciphertext. In: Proceedings of the 4th International Conference on Computer Sciences and Convergence Information Technology, Washington, 2009. 1036–1040
27 Ren Y L, Gu D W, Wang S Z, et al. New fuzzy identity-based encryption in the standard model. Informatica, 2010, 21: 393–407
28 Tian M M, Huang L S, Yang W. Security analysis of a fuzzy identity-based encryption scheme. J Circ Syst Comp, 2014, 23: 1450033
29 Wang X A, Yang X, Zhang M, et al. Cryptanalysis of a fuzzy identity based encryption scheme in the standard model. Informatica, 2012, 23: 299–314
30 Mao Y J, Li J, Chen M R, et al. Fully secure fuzzy identity-based encryption for secure IoT communications. Comput Stand Interfaces, 2016, 44: 117–121
31 Aggarwal M, Zubair M, Unal D, et al. A testbed implementation of a biometric identity-based encryption for IoMT-enabled healthcare system. In: Proceedings of the 5th International Conference on Future Networks & Distributed Systems, New York, 2021. 58–63
32 Aggarwal M, Zubair M, Unal D, et al. Fuzzy identification-based encryption for healthcare user face authentication. J Emergency Med Trauma Acute Care, 2022, 2022
33 Bai Y, Xu J B. Access control scheme based on fuzzy identity in opportunistic network. Procedia Comput Sci, 2018, 131: 1122–1127
34 Cheng Z H. Security analysis of SM9 key agreement and encryption. In: Proceedings of the International Conference on Information Security and Cryptology, Berlin, 2018. 3–25
35 Shi Y, Ma Z Y, Qin R F, et al. Implementation of an attribute-based encryption scheme based on SM9. Appl Sci, 2019, 9: 3074
36 Sun S Z, Ma H, Zhang R, et al. Server-aided immediate and robust user revocation mechanism for SM9. Cybersecurity, 2020, 3: 12
37 Mu Y H, Xu H X, Li P L, et al. Secure two-party SM9 signing. Sci China Inf Sci, 2020, 63: 189101
38 Paillier P. Public-key cryptosystems based on composite degree residuosity classes. In: Proceedings of the Advances in Cryptology, Berlin, 1999. 223–238
39 Ji H H, Zhang H J, Shao L S, et al. An efficient attribute-based encryption scheme based on SM9 encryption algorithm for dispatching and control cloud. Connection Sci, 2021, 33: 1094–1115
40 Lai J C, Huang X Y, He D B, et al. Provably secure online/offline identity-based signature scheme based on SM9. Comput J, 2022, 65: 1692–1701
41 Lai J C, Huang X Y, He D B, et al. Security analysis of uppercase SM9 digital signature and key encapsulation (in Chinese). Sci Sin Inform, 2021, 51: 1900–1913
42 Chen L Q, Cheng Z H. Security proof of Sakai-Kasahara's identity-based encryption scheme. In: Proceedings of the IMA International Conference on Cryptography and Coding, Berlin, 2005. 442–459
43 Delerablee C. Identity-based broadcast encryption with constant size ciphertexts and private keys. In: Proceedings of the Advances in Cryptology, Berlin, 2007. 200–215
44 Scott M. Miracl-a multiprecision integer and rational arithmetic C/C++ library. http://www.shamus.ie