Neyire Deniz Sarier (2010). Generic Constructions of Biometric Identity Based Encryption Systems. In *Information Security Theory and Practices. Security and Privacy of Pervasive Systems and Smart Devices*, Pierangela Samarati, Michael Tunstall, Joachim Posegga, Konstantinos Markantonakis & Damien Sauveron, editors, volume 6033 of *Lecture Notes in Computer Science*, 90–105. Springer-Verlag. ISBN 3-642-12367-8. ISSN 0302-9743. URL https://dx.doi.org/10.1007/978-3-642-12368-9_7. *BestStudentPaperAward.*

Generic Constructions of Biometric Identity Based Encryption Systems

Neyire Deniz Sarier

Bonn-Aachen International Center for Information Technology
Computer Security Group
Dahlmannstr. 2, D-53113 Bonn Germany,
denizsarier@yahoo.com

Abstract. In this paper, we present a novel framework for the generic construction of biometric Identity Based Encryption (IBE) schemes, which do not require bilinear pairings and result in more efficient schemes than existing fuzzy IBE systems implemented for biometric identities. Also, we analyze the security properties that are specific to biometric IBE namely anonymity and identity privacy. Considering these notions, we present generic constructions for biometric IBE and ID-KEM based on weakly secure anonymous IBE schemes, error correcting codes and generic conversion schemes. Finally, we describe concrete applications of our framework and compare them to the existing fuzzy IBE systems in terms of time complexity.
Keywords: Biometric IBE, Generic Construction, IND-CCA, Anonymity, Error Correcting Codes

1 Introduction

In Eurocrypt'04, Sahai and Waters proposed a new Identity Based Encryption (IBE) system called fuzzy IBE, which provides error tolerance property for IBE in order to use biometric attributes as the identity instead of an arbitrary string like an email address. This new system combines the advantages of IBE with using biometrics as an identity, where IBE avoids the need for an online Public Key Infrastructure (PKI), which is the most inefficient and costly part of public key encryption. The use of biometrics as the identity in the framework of IBE simplifies the process of key generation at the Private Key Generator (PKG). Since biometric information is unique, unforgettable and non-transferable, the user only needs to provide his biometrics at the PKG to obtain his secret key instead of presenting special documents and credentials to convince the PKG about his identity. Also, biometrics is attached to the user, hence the user does not need to remember any password, to use any public key or even an e-mail address since the public key of the user is always with him to be used for encryption during an ad hoc meeting. Finally, biometric data could be easily integrated with fuzzy IBE due to its error tolerance property, which is required for the noisy nature of biometrics. The main feature of fuzzy IBE is the construction of the

secret key based on the biometric data of the user which can decrypt a ciphertext encrypted with a slightly different measurement of the same biometrics. Specifically, fuzzy IBE allows for error tolerance in the decryption stage, where a ciphertext encrypted with the biometrics w could be decrypted by the receiver using the private key corresponding to the biometrics w', provided that w and w' are within a certain distance of each other. Besides, fuzzy IBE could be applied in the context of Attribute-Based Encryption (ABE), where the sender encrypts data using a set of attributes such as {university, faculty, department} and the ciphertext could only be decrypted if the receiver has the secret key associated to all of these attributes or sufficient number of them. In current fuzzy IBE schemes, the private key components are generated by combining the values of a unique polynomial evaluated on each attribute with the master secret key. This way, different users, each having some portion of the secret keys associated to the attributes of a given ciphertext c cannot collude to decrypt c, which is defined as collusion resistance. The basic fuzzy IBE schemes guarantee a weak level of security for identity based setting i.e. Indistinguishability against Chosen Plaintext Attack (ID-IND-CPA), but they could be combined with well-known generic conversion systems to obtain a high level of security i.e. Indistinguishability against Chosen Ciphertext Attack (ID-IND-CCA). Besides, the biometrics is considered as public information, hence the compromise of the biometrics does not affect the security of the system. Thus, in existing systems, biometrics w of the receiver is sent together with the corresponding ciphertext, which could effect the privacy of the user's actions negatively.

1.1 Related Work

The first fuzzy IBE scheme is described by Sahai and Waters in [18], where the size of the public parameters is linear in the number of the attributes of the system or the number of attributes (or features) of a user. More efficient fuzzy IBE [2, 13], Attribute-Based Encryption (ABE) [17] and biometric IBE [19] schemes are achieved with short public parameter size by employing the Random Oracle Model (ROM). To achieve ID-IND-CCA security, these schemes could be combined with well known generic conversion schemes such as Fujisaki-Okamoto [12] or REACT [16]. The signature analogue of fuzzy IBE, i.e. fuzzy IBS is first defined in [24] and further improved in [20]. Similarly, a threshold Attribute Based Signature (t-ABS) scheme and its extension to threshold attribute based anonymous credential systems is presented in [21], where the authors also define the security notions of weak/full signer attribute privacy for t-ABS. The only work that considers privacy of biometric attributes in fuzzy IBE is the master thesis of [23], which adapts the Boneh-Franklin IBE scheme [4] to function as an error tolerant IBE scheme, where the IBE scheme in [4] is anonymous. Recently, other anonymous IBE schemes [5, 1] for the standard IBE setting (i.e. non-biometric identities) are described, which do not require bilinear pairings and their security is based on the standard quadratic residuosity problem. Besides, these schemes encrypt a message bit by bit, thus they can be used to encrypt short session keys due to the large bandwidth consumption. To achieve ID-IND-CCA security,

the schemes can implement the KEM/DEM construction of Bentahar et al. [3], which takes as input a weakly secure IBE scheme and a hash function to output an ID-IND-CCA secure KEM that is combined with an IND-CCA secure DEM.

1.2 Motivation and Contributions

Currently, the secrecy of biometric data is viewed with skepticism since it is very easy to obtain biological information such as fingerprint, iris or face data through fingerprint marking or using a camcorder. However, biometrics is a sensitive information, as in the case of biometric remote authentication, it should not be easy to obtain the biometric data by compromising the central server, where the biometrics of each user is often associated with his personal information. In particular, a user could use its biometrics on a number of applications such as identification, authentication, signing, etc. Thus, the secrecy of identity-biometrics relation should be maintained, which is defined as identity privacy [8, 22]. Current fuzzy IBE and biometric IBE systems do not consider anonymity and privacy of user biometrics at the same time, hence, it is vital to describe an efficient and anonymous error-tolerant encryption system for biometric identities in order to avoid traceability of the user's actions. Although the fuzzy IBE scheme of [23] provides anonymity, the scheme combines each biometric attribute with the identity (i.e. Name, e-mail address) of the user to avoid the collusion attacks. This approach is not only against identity privacy but also against the main principle of fuzzy IBE or biometric IBE, where the identity of the user should only consist of his biometric data.

The contributions of this paper are twofold. First, we analyse the security properties of biometric IBE schemes and present a new method for preventing collusion attacks. Next, we present generic constructions for biometric IBE and ID-KEMs that provide either entropic security or ID-IND-CCA security depending on the primitives used, which do not require bilinear pairings. For this purpose, we combine fuzzy sketches, error correcting codes and/or modify well known generic conversion schemes to function in the error-tolerant setting. Also, we will describe concrete applications of our generic constructions using anonymous IBE schemes [5, 1] that encrypt each message bit by bit and do not depend on bilinear pairings. To avoid collusion attacks and to guarantee the security notions that we present, the anonymous IBE schemes are modified according to our new method, thus, we achieve more efficient biometric IBE schemes compared to current fuzzy IBE systems implemented for biometric identities in the ROM.

2 Preliminaries and Definitions

2.1 Fuzzy Identity Based Encryption

In [18, 2], the generic fuzzy IBE scheme is defined as follows.

- Setup(l): Given a security parameter l, the Private Key Generator (PKG) generates the master secret key ms and the public parameters $params$.

– Key Generation(w, ms): Given a user's identity $w = (w_1, ..., w_n)$ and ms, the PKG returns the corresponding private key D_w.
– Encrypt(w', m): A probabilistic algorithm that takes as input an identity $w' = (w'_1, ..., w'_n)$, $params$ and a message $m \in M$, outputs the ciphertext U.
– Decrypt(U, D_w): A deterministic algorithm that given the private key D_w and a ciphertext U encrypted with w' such that $|w \cap w'| \geq d$, returns either m or $\perp$. Here d denotes the error tolerance parameter of the scheme.

The security of a biometric IBE scheme is defined using the following game between an adversary and a challenger.

Experiment ID-IND-ATK(l, IBE, A=(A_1, A_2))
$(params, ms) \leftarrow$Setup(l)
$(s, w^*, m_0, m_1) \leftarrow A_1^O(params)$ with $|m_0| = |m_1|$
$b \overset{R}{\leftarrow} \{0, 1\}$, $U^* \leftarrow$ Encrypt$(w^*, params, m_b)$
$b' \leftarrow A_2^O(w^*, U^*, params)$
If $b' = b$ return 1 else return 0

The advantage of the attacker A is $Adv_{A,\mathsf{IBE}}^{\mathsf{ID\text{-}IND\text{-}ATK}} = |Pr[b' = b] - \frac{1}{2}|$. Hence, a biometric IBE scheme is ID-IND-ATK secure if the advantage of A is negligible in the security parameter l. If ATK = CCA, then A has access to a decryption oracle in addition to the encryption and private key extraction oracles available to A when ATK = CPA.

2.2 Error Correcting Codes and Fuzzy Sketches

Let $\mathcal{H} = \{0, 1\}^N = \mathbb{F}_2^N$ be the Hamming space of length N ,where $\mathbb{F}_2 = \{0, 1\}$. An Error Correcting Code (ECC) over $\mathcal{H}$ is a subset $C \subset \mathcal{H}$, where elements of C are called as codewords. An (N, S, d) binary linear error correcting code C is a vector subspace of $\mathbb{F}_2^N$. When C contains 2^k codewords, then C is denoted as $[N, k, t]$, where t is the correction capacity of C.

The main idea of fuzzy sketches is given a public data PAR $= c \oplus b$, one tries to correct the corrupted codeword PAR $\oplus\, b' = c \oplus (b \oplus b')$. If the Hamming distance $\mathsf{dis}_{\mathcal{H}}$ (b,b') is small, recovering c from PAR $\oplus\, b'$ is possible [7]. An important requirement for such a scheme is that the value PAR should not reveal too much information about the biometric template b, which is obtained as described in section 2.4.

2.3 Robust Sketch and Robust Fuzzy Extractors

Since the correction is performed by combining the biometrics b' with the public value PAR of the signer, the presence of an active adversary who maliciously alters the public string PAR leads an adversary even to obtain the secret b' entirely depending on the utilized sketch or fuzzy extractor [6]. This attack can be avoided by using a robust fuzzy extractor, which is resilient to modification of the public value PAR [6]. The generic robust fuzzy sketch described in [6] replaces the value PAR with PAR$^* = \langle \mathsf{PAR}, H(b, \mathsf{PAR}) \rangle$, where H is a hash function. By

applying a strong extractor, one can convert any robust sketch to a robust fuzzy extractor. The formal definition of fuzzy extractors is presented in Appendix A.

2.4 Collusion Attacks

Any biometric IBE/IBS scheme requires the biometric measurement of the receiver or the signer, respectively. For this purpose, the biometrics of the user is captured using a sensor and the raw biometric data is further processed to extract the feature vector and to obtain the biometric template b of the user. In a biometric encryption scheme, feature extraction is applied on the raw biometric data to obtain the feature vector (or the attributes) and then, each attribute is associated with a unique integer $w_i \in \mathbb{Z}_p^*$ to form the identity $w = (w_1, ..., w_n)$ [18, 2]. Here, n denotes the size of the attributes of each user. Since some of the attributes could be common in some users, a unique polynomial is selected for each user and included in the key generation algorithm to bind the private key to the user. This way, different users cannot collude in order to decrypt a ciphertext that should only be decrypted by the real receiver.

In the biometric cryptosystems such as BIO-IBS [9] and BIO-IBE of [19], the biometric template b is computed using the feature vector and the hash of b is used as the identity ID. Here, the template b is assumed to be a fixed length binary string, so each feature forming the original biometric template (namely the feature vector) are quantized to generate multiple bits per feature that are concatenated to obtain the binary template b. Particularly, the framework for biometric template generation consists of (1) extracting features; (2) quantization and coding per feature and concatenating the output codes; (3) applying error correction coding (ECC) and hashing [10]. During this process, many quantizers produce and use side-information, which could be published to be used later in the reconstruction of the binary template b'.

As different from existing fuzzy IBE systems, the BIO-IBE [19] requires the use of the biometric template b obtained from the feature vector of the user, where feature extraction is the most costly part of the biometric template generation. Since feature extraction is already performed in any fuzzy IBE scheme, one can easily apply a robust fuzzy extractor on the feature vector to bind the private key components to the user's identity and thus avoid collusion attacks. Instead of choosing a unique polynomial for each user, we use the robust fuzzy extractor to obtain a unique biometric string ID via error correction codes from the biometric template b of the user in such a way that an error tolerance t is allowed. In other words, we will obtain the same biometric string ID even if the fuzzy extractor is applied on a different b' such that $\mathsf{dis}_{\mathcal{H}}(\mathsf{b},\mathsf{b}') < t$. Here, $\mathsf{dis}()$ is the distance metric used to measure the variation in the biometric reading and t is the error tolerance parameter of the fuzzy extractor.

In the anonymous fuzzy IBE scheme of [23], collusion attacks are avoided by combining each biometric feature w_i with the identity (i.e. Name,e-mail) of the user. However, this approach is against the nature of fuzzy IBE, where the identities should only consist of the biometric data of the user. Besides, an important privacy property that we will present in the next section is not satisfied

despite the anonymity of the scheme. One can correct this fuzzy IBE scheme with a similar approach introduced in [19], namely, the identity is obtained from the biometric information of the user using a feature extraction algorithm followed by a fuzzy extraction process, where the result of the former procedure (i.e. $w = (w_1, ..., w_n)$) is combined with the output of the latter (i.e. ID) to obtain the biometric attribute set $\mathsf{BID} = \langle H(w_1, \mathsf{ID}), ..., H(w_n, \mathsf{ID})\rangle$ to be used in the key generation phase. This way, the privacy of biometric-identity relation and the resistance against collusion attacks is maintained. Here, H is an cryptographic hash function.

3 A Generic Construction based on Robust Sketch

The first idea for an efficient biometric IBE scheme without using bilinear pairings is to combine any IBE scheme with an Error Correcting Code (ECC) and a robust sketch. Particularly, given IBE=(KeyGen, Enc,Dec) is an $\mathsf{ID\text{-}IND\text{-}CPA}$ secure IBE scheme that encrypts a message (i.e. the codeword c) bit by bit, an $\mathsf{ECC}()$ with correction capacity t and a robust sketch of [6] with $\mathsf{PAR}^* = \langle \mathsf{PAR}, H(m, \mathsf{PAR})\rangle$, the construction is described as follows.

- $\mathsf{Setup}(l)$: Given a security parameter l, the PKG generates the master secret key ms and the public parameters of the system.
- $\mathsf{KeyGen}(w, ms)$: Given a user's biometric feature vector w and ms, it returns the corresponding private key D_w.
- $\mathsf{Encrypt}(w', m)$: An algorithm that takes as input a biometric vector w', Enc algorithm of the IBE, a message $m \in M$ and a robust sketch, outputs the ciphertext $\langle U, V, W\rangle = \langle \mathsf{Enc}^{\mathsf{IBE}}_{w'}(c; r), \mathsf{PAR}^*\rangle = \langle \mathsf{Enc}^{\mathsf{IBE}}_{w'}(c), \mathsf{PAR}, H(m, \mathsf{PAR})\rangle$, where $\mathsf{PAR} = c \oplus m$ for a random codeword $c \in C$. For the case that the underlying IBE scheme is not anonymous, the biometric vector w' of the receiver is attached to the ciphertext.
- $\mathsf{Decrypt}(D_w)$: A deterministic algorithm that given the private key D_w of the Dec algorithm, an error correcting procedure $\mathsf{ECC}()$ and a ciphertext encrypted with w' such that $d \leq |w \cap w'|$, the algorithm computes $c' = \mathsf{Dec}^{\mathsf{IBE}}_{w}(U)$ and corrects the error via c=$\mathsf{ECC}(c')$. Next, $m = c \oplus V$ is obtained and if $W = H(m, V)$, m is returned, else $\perp$ is returned.

3.1 Entropic Security vs. Indistinguishability

As it is noted in [11], semantic security cannot be achieved for fuzzy sketches, when the adversary generates the two strings m_1, m_2 such that $|m_1| = |m_2|$ and thus knows that the challenge ciphertext is the encryption of one of m_1, m_2, the adversary can easily distinguish by computing $m_i \oplus V$ and verifying $W = H(m_i, V)$ from the challenge ciphertext. Thus, fuzzy sketches guarantee entropic security, which is weaker than Shannon security and assumes that the adversary is sufficiently uncertain about the challenge message.

In the context of our generic conversion, integration of a robust sketch can only satisfy OW-CCA (Onewayness against Chosen Ciphertext Attack) in the

ROM, which is a weak security level. Therefore, we present constructions that provide IND-CCA security in the following sections.

4 Security Properties

In addition to the standard security level (IND-CCA) that any encryption scheme should achieve, biometric IBE schemes have to guarantee the following properties that are particularly important for biometric cryptosystems, since a user could use its biometrics on a number of applications such as identification, authentication, signing, etc. Thus, the traceability of the user's actions should be prevented through the anonymity of the ciphertexts and the secrecy of the identity-biometrics relation.

4.1 Anonymity

Informally, Recipient Anonymity (RA) or key privacy means that the adversary must be unable to decide whether a ciphertext was encrypted for a chosen identity, or for a random identity. In other words, an adversary cannot tell who the recipient is by looking at the ciphertext, which could be used to thwart traffic analysis. If the ciphertext could be anonymized by anyone using the public key of the recipient, i.e. not just by the encryptor, the encryption scheme is defined as universally anonymous. In current fuzzy IBE systems, the biometric vector w of the receiver is attached to the ciphertext since set overlap is used as the distance metric between the identities w and w'. Hence, a different system should be designed to achieve anonymity. The formal definition is as follows:

Experiment $Exp_A^{\mathsf{IBE\text{-}RA\text{-}CPA}}$
$(ms, params) \leftarrow \mathsf{Setup}(l)$
$(w_0, w_1, s, m) \leftarrow A(params)$ s.t. $|w_0 \cap w_1| < d$
$b \overset{\mathrm{R}}{\leftarrow} \{0, 1\}$
$U^* \leftarrow \mathsf{Encrypt}(m, params, w_b)$
$b' \leftarrow A(s, U^*, params)$
If $b' = b$ return 1 else return 0

The advantage of the attacker A is $Adv_{A,\mathsf{IBE}}^{\mathsf{IBE\text{-}RA\text{-}CPA}} = |Pr[b' = b] - \frac{1}{2}|$. A biometric IBE scheme IBE is said to be IBE-RA-CPA-secure if the respective advantage function is negligible for all polynomial time adversaries (PTAs) A.

4.2 Identity Privacy

For biometric authentication, this notion guarantees the privacy of the sensitive relationship between the user identity (i.e. ID= Name or e-mail address) and its biometrics against a malicious service provider or a malicious database [8, 22]. For biometric IBE setting, this notion can be adapted for having privacy even against the trusted authority (PKG) or the encryptor. Thus, identity privacy is a stronger notion than anonymity, namely, identity privacy implies anonymity,

which is shown in the following lemma. The privacy of biometrics-identity relation is achieved for many fuzzy IBE systems, which depend only on biometric identities except for the fuzzy IBE scheme in [23], which combines the identity (i.e. Name) of the receiver with his biometric features to avoid collusion attacks. This approach is not only against identity privacy but also against the main principle of fuzzy IBE. However, this scheme could be corrected using our method described in section 2.4. This notion is formally defined as follows:

Experiment $Exp_A^{\mathsf{IBE\text{-}IP\text{-}CPA}}$
$ms, params \leftarrow \mathsf{Setup}(l)$
$(s, m, ID, w_0, w_1) \leftarrow A(params)$ s.t. $|w_0 \cap w_1| < d$
$b \overset{\mathrm{R}}{\leftarrow} \{0,1\}$, $U^* \leftarrow \mathsf{Encrypt}(m, params, w_b, ID)$
$b' \leftarrow A(s, U^*, params)$
If $b' = b$ return 1 else return 0

The advantage of the attacker A is $Adv_{A,\mathsf{IBE}}^{\mathsf{IBE\text{-}IP\text{-}CPA}} = |Pr[b' = b] - \frac{1}{2}|$. A biometric IBE scheme IBE is said to be IBE-RA-CPA-secure if the respective advantage function is negligible for all PTAs A.

Lemma 1. *Identity privacy implies anonymity.*

Proof. Assume that a given biometric IBE scheme is not anonymous, then using this scheme we construct another biometric IBE scheme that does not guarantee identity privacy.

Given any biometric IBE scheme which is not anonymous and has the encryption algorithm Encrypt, define a new biometric IBE scheme with an encryption algorithm that appends the identity information (Name, e-mail) to the ciphertext. Since the new biometric IBE scheme is also not anonymous, the link between the identity and biometrics is not kept secret. Thus identity privacy is not satisfied.

5 Generic Construction of Biometric IBE

In this section, we described generic constructions converting any one way secure IBE scheme that encrypts a message bit by bit into an ID-IND-CCA secure encryption scheme in the error-tolerant setting. Due to page limitations, the proofs will be presented in the full version of the paper.

5.1 Based on Fujisaki-Okamoto (FO) Conversion

Fujisaki and Okamoto proposed a simple conversion scheme called as a hybrid scheme ε^{hy} from weak asymmetric-key encryption (AK) and symmetric-key encryption (SK) schemes into a public-key encryption scheme which is secure in the sense of IND-CCA. Basically, ε^{hy} is defined in [12] as follows.

$$\varepsilon^{hy}(m;\sigma) = \langle AE_{pk}(\sigma; H(\sigma, m)) || SE_{G(\sigma)}(m) \rangle$$

In ε^{hy} , σ is generated at random, H and G are two cryptographic hash functions with H: AKMS x SKMS $\rightarrow$ COINS and G: AKMS $\rightarrow$ SKS ,where AKMS denotes asymmetric-key message space, SKMS denotes symmetric-key message space, and SKS is the symmetric-key space. The idea is , first encrypting the redundancy σ with the random coin $H(\sigma, m)$ under public key pk using the probabilistic scheme AE and then encrypting the message under the symmetric key $G(\sigma)$ using the scheme SK. In [12], it is proven that if AE is an one-way encryption scheme, then ε^{hy} is IND-CCA secure in the ROM. However, it is shown that if AE scheme satisfies IND-CPA security, then there is a significant improvement in the security reduction, where IND-CPA implies also one-way encryption [4]. Finally, in [14], the authors describe the FO conversion for IBE encryption, which we will extend for our setting as below.

According to our framework, we present an ID-IND-CCA secure application that works in error-tolerant IBE setting as follows. Here $c \in C$ is a random codeword and IBE=(KeyGen, Enc,Dec) is an ID-IND-CPA secure IBE scheme that encrypts a message (i.e. the codeword c) bit by bit.

- Setup(l): Given a security parameter l, the PKG generates the master secret key ms and the public parameters of the system.
- KeyGen(w, ms): Given a user's biometric feature vector w and ms, it returns the corresponding private key D_w.
- Encrypt(w', m): A probabilistic algorithm that takes as input biometrics w', Enc algorithm of the IBE scheme, a message $m \in M$ and a random codeword $c \in C$, outputs the ciphertext $\langle U, V, W\rangle = \langle \mathsf{Enc}^{\mathsf{IBE}}_{w'}(c; H_1(\sigma, m)), H_2(c) \oplus \sigma, H_3(\sigma) \oplus m\rangle$. For the case that the underlying IBE scheme is not anonymous, the biometric vector w' of the receiver is attached to the ciphertext.
- Decrypt(D_w): A deterministic algorithm that given the private key D_w of the Dec algorithm, an error correcting procedure ECC() and a ciphertext encrypted with w' such that $|w \cap w'| \geq d$, first computes $c' = \mathsf{Dec}^{\mathsf{IBE}}_{w}(U)$ and error corrects c=ECC(c'). Next, $\sigma = H_2(c) \oplus V$ and $m = H_3(\sigma) \oplus W$ is obtained. Finally, by computing $H_1(\sigma, m)$ and using it in reencryption, the correctness is checked and m is returned.

5.2 Based on REACT

As it is noted in [16], Fujisaki-Okamoto transformation converts any one-way cryptosystem into a CCA secure encryption scheme, but it is not optimal due to the re-encryption operation during the decryption phase. In [16], an efficient and IND-CCA secure generic conversion scheme is presented, which takes as input a OW-PCA secure encryption scheme and avoids the disadvantages of FO transformation via

$$\varepsilon^{hy}(m; R) = \langle AE_{pk}(R)||SE_{G(R)}(m)||H(R, m, AE_{pk}(R), SE_{G(R)}(m))\rangle$$

Similar to FO, REACT is also implemented for IBE in [14]. When used in biometric IBE setting, one should modify REACT for IBE as

$$\langle U, V, W \rangle = \langle \mathsf{Enc}_{w'}^{\mathsf{IBE}}(c), G(c) \oplus m, H(c, m, U, V) \rangle$$

Thus, the only difference to the FO transformation adapted to the error-tolerant setting occurs in the decryption stage where only one hash computation, i.e. $H(c, m, U, V)$ is verified instead of a full reencryption. Here, $c \in C$ denotes a random codeword.

6 A Generic Biometric ID-KEM Construction

A Key Encapsulation Mechanism (KEM) consists of three algorithms: Key generation, encryption and decryption algorithms, where a KEM outputs a random session key to be used by the Data Encapsulation Mechanism (DEM) in the symmetric encryption. Current identity-based KEM's [3] are not suitable for error prone identities, thus we present a generic construction for a biometric ID-KEM that takes any IBE scheme IBE=(KeyGen, Enc,Dec) which encrypts a message bit by bit. Here, H_1, H_2 and H denote cryptographic hash functions with $H_1 : \{0,1\}^* \rightarrow \{0,1\}^{z_1}$, $H_2 : \{0,1\}^* \rightarrow \{0,1\}^{z_2}$ and $H : \{0,1\}^* \rightarrow \{0,1\}^{z}$.

- $\mathsf{Setup}(l)$: Given a security parameter l, the PKG generates the master secret key ms and the public parameters of the system.
- $\mathsf{KeyGen}(w, ms)$: Given a user's biometric feature vector w and ms, and returns the corresponding private key D_w.
- $\mathsf{Encapsulate}(w', m)$: An algorithm that takes as input a biometric vector w', Enc algorithm of the IBE scheme, a message $m \in M$ and a random codeword $c \in C$, it returns $\langle U, V, K \rangle = \langle \mathsf{Enc}_{w'}^{\mathsf{IBE}}(c; H_1(m)), H(c) \oplus m, H_2(m) \rangle$. Here, $K \in \mathbb{K}_{\mathsf{ID\text{-}KEM}}$ is an encapsulation key from the key space of the ID-KEM. For the case that the underlying IBE scheme is not anonymous, the biometric vector w' of the receiver is attached to the ciphertext.
- $\mathsf{Decapsulate}(D_w)$: A deterministic algorithm that given the private key D_w of the Dec algorithm, an error correcting procedure ECC() and a ciphertext (U, V) encrypted with w' such that $d \leq |w \cap w'|$, it computes $c' = \mathsf{Dec}_w^{\mathsf{IBE}}(U)$ and corrects the error via c=$\mathsf{ECC}(c')$. Next, $m = H(c) \oplus V$ is obtained and $H_1(m)$ is computed to verify $U = \mathsf{Enc}_{w'}^{\mathsf{IBE}}(c; H_1(m))$. Finally, the algorithm returns either the encapsulated key $H_2(m)$, else $\perp$ is returned.

The security of a biometric ID-KEM is defined using the following game between an adversary and a challenger.

Experiment ID-IND-CCA(l, ID-KEM, A)
$(params, ms) \leftarrow \mathsf{Setup}(l)$
$(s, w^*) \leftarrow A_1^O(params)$
$(K_0, U^*, V^*) \leftarrow \mathsf{Encapsulate}(w^*, params, c, m)$
$(K_1) \overset{\mathrm{R}}{\leftarrow} \mathbb{K}_{\mathsf{ID\text{-}KEM}}$
$b' \leftarrow A_2^O(w^*, U^*, V^*, K_b)$
If $b' = b$ return 1 else return 0

The advantage of the attacker A is $Adv^{\mathsf{ID\text{-}IND\text{-}CCA}}_{A,\mathsf{ID\text{-}KEM}} = |Pr[b' = b] - \frac{1}{2}|$. Hence, a biometric ID-KEM is ID-IND-CCA secure if the advantage of A is negligible in the security parameter l.

Theorem 1. *Suppose there exists a CCA adversary A which distinguishes ID-KEM with advantage ϵ in less than $q_{H_{12}}, q_H$ random oracle calls q_D decryption queries. Then there exists an algorithm R which inverts IBE with probability $\epsilon' \geq \frac{2\epsilon}{q_{H_{12}}+q_H+q_D} - \frac{q_D}{2^z}$.*

Proof. Given an ID-IND-CCA secure biometric ID-KEM, the goal of the reduction algorithm R is to invert the ID-OW-CPA secure IBE scheme using an adversary A running against ID-KEM.

The challenger of R outputs the public parameters of IBE, which is passed to the adversary A in order to simulate the setup phase of ID-KEM.

A responds with the challenge biometric identity w^*, which is relayed to the challenger of R, which returns the encryption U^* of a random message $c^* \in C$. R outputs U^* together with a randomly chosen V^* and a random key K_0 to simulate the challenge phase of ID-KEM and answers the random oracle and decryption queries of A as follows.

1. H_1-queries: On each new input m, R picks random h_1 and h_2 from the ranges of H_1 and H_2 , returns h_1 to A and inserts the tuple (m, h_1, h_2) to the H_{12}List.
2. H_2-queries: On each new input m, R picks random h_1 and h_2 from the ranges of H_1 and H_2 , returns h_2 to A and inserts the tuple (m, h_1, h_2) to the H_{12}List.
3. H-queries: On each new input (c), R returns a random h and adds the tuple (c, h) to the HList.
4. Private Key Extraction queries: For any identity w such that $|w \cap w^*| < d$, the extration query is passed to the challenger of R.
5. Decryption queries: On each new input (w, U, V),
   - If $|w \cap w^*| < d$, R runs the private key extraction oracle and answers A as the real decapsulation oracle would.
   - If $|w \cap w^*| \geq d$ but $(U, V) \neq (U^*, V^*)$, R computes for each pair in the HList $m = H(c) \oplus V$. Next, R checks for each computed m, $\mathsf{Enc}^{\mathsf{IBE}}_w(c; H_1(m)) = U$ using the simulation of H_1 as above. If the check is successful, then R simulates the H_2 oracle to return $H_2(m)$. If not, R returns reject.

Finally, A outputs its guess b'. R will pick at random an entry from HList and returns this to the challenger. Similar to the computation of the reduction cost of theorem 6 in [3], we obtain $\epsilon' \geq \frac{2\epsilon}{q_{H_{12}}+q_H+q_D} - \frac{q_D}{2^z}$

7 Applications

In this section, we present two concrete applications based on the anonymous IBE schemes of [5, 1], which do not require bilinear pairings and encrypt a message bit by bit. Thus, they could be used as an input to our generic constructions with the following modifications to avoid collusion attacks.

7.1 Based on the scheme of Boneh et al. [5]

The first space efficient IBE scheme is introduced in [5], which is ID-IND-CPA secure in the standard model based on the difficulty of the Quadratic Residuosity (QR) problem and the encryption of a n-bit message results in a single element in $\mathbb{Z}/N\mathbb{Z}$ plus $n+1$ additional bits.

1.Setup(l): Generate two primes (p, q) and compute $N = pq$, where N is a RSA composite. Select a random $u \overset{R}{\leftarrow} J(N)/\mathrm{QR}(N)$. Here, $J(N)$ denotes the set $\{x \in \mathbb{Z}/N\mathbb{Z} : (\frac{x}{N}) = 1\}$, where $(\frac{x}{N})$ is the Jacobi symbol of x in $\mathbb{Z}/N\mathbb{Z}$. Also, $\mathrm{QR}(N)$ is the set of quadratic residues in $J(N)$. The public parameters are $params = (N, u, H)$, where H is a hash function $H : \mathcal{BID} \rightarrow J(N)$, where $\mathcal{BID} = \mathcal{W} \times \mathcal{ID}$. We assume that the features $w_j \in \mathcal{W}$ are ordered as in [15]. The master key is $msk = (p, q, K)$, namely the factorization of N together with a random key K for a pseudorandom function $F_K : \mathcal{W} \times \mathcal{ID} \rightarrow \{0, 1, 2, 3\}$.

2.Keygen(msk, w): It takes as input msk, a biometric vector w with length n. The algorithm outputs a private key $D_{\mathsf{BID}} = (r_1, ..., r_n)$ for decrypting encryptions of n-bit messages as follows. For j= 1,...,n do:

- $R_j \leftarrow H(w_j, \mathsf{ID}) \in J(N)$ and $t \leftarrow F_K\ (w_j, \mathsf{ID}) \in \{0, 1, 2, 3\}$
- let $a \in \{0, 1\}$ such that $u^a R_j \in \mathrm{QR}(N)$
- let z_0, z_1, z_2, z_3 be the four square roots of $u^a R_j \in \mathbb{Z}/\ N\mathbb{Z}$ and set $r_j \leftarrow z_t$

3.Enc($params, w', c$):The encryption algorithm that takes as input biometrics w' of the receiver, $params$ and a codeword $c = c_1...c_n \in C$. It generates a random $s \in \mathbb{Z}/N\mathbb{Z}$ and sets $S \leftarrow s^2 \bmod N$. Then, $Q'(N, u, 1, S)$ is computed to obtain the polynomial τ and $k \leftarrow (\frac{\tau(s)}{N})$. Here, Q' is a deterministic algorithm that satisfies some properties [5] and takes as inputs (N, u, R_j, S), where $N \in \mathbb{Z}^+$, and $u, R_j, S \in \mathbb{Z}/N\mathbb{Z}$. It outputs polynomials $f_j, \bar{f}_j, g_j, \tau \in \mathbb{Z}/N\mathbb{Z}[x]$. Finally, for $j = 1, ..., n$ do:

- Compute $R_j \leftarrow H(\ w_j, \mathsf{ID})$ and run $Q'(N, u, R_j, S)$ to obtain g_j
- Compute $e_j = c_j \cdot (\frac{g_j(s)}{N})$

The ciphertext is $U = (S, k, e, \mathcal{L})$, where $e = e_1...e_n$ and $\mathcal{L}$ is a label that contains information about how each e_j is associated to the index of $w_j \in \mathcal{W}$.

4.Dec(U, D_{BID}): The decryption algorithm takes as input the ciphertext U and the private key $D_{\mathsf{BID}} = (r_1, ..., r_n)$ and recovers $c' = c'_1...c'_n$ as follows. For $j = 1, ..., n$, set $R_j \leftarrow H(w_j, \mathsf{ID})$ and run $Q'(N, u, R_j, S)$ to obtain f_j, $\bar{f}_j$

$$\text{If } r_j^2 = R_j \text{ set } c_j \leftarrow e_j \cdot (\tfrac{f_j(r_j)}{N}), \text{ else if } r_j^2 = uR_j \text{ set } c_j \leftarrow e_j \cdot k \cdot (\tfrac{\bar{f}_j(r_j)}{N})$$

The security of the Anonymous IBE depends on the difficulty of the interactive quadratic residuosity (IQR) problem in the standard model and the

encryption of a binary string results as a ciphertext of size $log_2 N + n + 1$, where N is a RSA modulus and n is length of c and w. In case that the number of biometric features are less than the length of the codeword, than either the extraction algorithm could be used to extract more features or the technique in [5] is used which computes the hash of the unique identity of the receiver together with the indices $j = 1, ..., n$. Since the modified scheme is also secure in the sense of ID-IND-CPA, it is input to one of our generic constructions to obtain either an ID-IND-CCA secure encryption scheme or an ID-IND-CCA secure KEM.

The main drawback of the scheme of Boneh et al [5] is its inefficiency since the complexity is quartic in the security parameter. Recently, Ateniese and Gasti [1] proposed an efficient and universally anonymous IBE scheme based on the quadratic residuosity assumption in the ROM. Similar to the modification presented above, if the key generation of the scheme in [1] is adapted for biometric identities, we are able to integrate this modified IBE scheme into one of our generic constructions, which is described as a concrete example in Appendix B.

8 Comparison

To show the efficiency of our constructions, we will compare our results to the existing fuzzy IBE schemes secure in the ROM. In [1], the authors implement different anonymous IBE schemes to present the average times of encryption of a short session key. Using these values presented in [1], we compare our results to any pairing based fuzzy IBE system in Table 1. For simplicity, we use different variables to represent the approximate times, where x and y denote the encryption and decryption times for Boneh-Franklin IBE [4] scheme implemented for a unique identity such as an e-mail address. Specifically, x is the time to compute two exponentiations within their respective groups if the bilinear pairing is precomputed and y is the time for one pairing computation, which is the dominant operation in terms of computation cost. For fuzzy IBE systems, since the identity is represented as a feature vector of length n such that $20 < n < 100$ depending on the biometric modality, the required times are computed as multiples of x and y. When compared to the exact times of the scheme [1] that we implement for our generic construction, the encryption of a message of the same size requires approximately $4x$, whereas the decryption time is again y. Finally, for our construction, a fuzzy extraction procedure FE for encryption and an error correcting procedure ECC for decryption stage is required, which can be efficiently implemented as in [9]. Finally, d is the error tolerance parameter.

9 Conclusion

In this paper, we present generic constructions for biometric IBE schemes and describe the relevant security notions. In order to provide anonymity, we assume that biometrics consists of an ordered set of features as in face biometrics [15]. An interesting future work can be the design of generic constructions of anonymous

Table 1. Comparison of time complexity

	Encryption time	Decryption time
Boneh-Franklin IBE* [4]	x	y
Pairing based Systems†	nx	dy
Our Construction‡	$4x$ + FE	y+ ECC

∗:non-biometric identities; †:for biometric identities; ‡:Based on the scheme of [1];

biometric IBE schemes, where biometrics can be represented as an unordered set of features, which is the case for some biometric modalities.

References

1. G. Ateniese and P. Gasti. Universally Anonymous IBE Based on the Quadratic Residuosity Assumption. In *CT-RSA'09*, volume 5473 of *LNCS*, pages 32–47. Springer, 2009.
2. J. Baek, W. Susilo, and J. Zhou. New constructions of fuzzy identity-based encryption. In *ASIACCS'07*, pages 368–370. ACM, 2007.
3. K. Bentahar, P. Farshim, J. Malone-Lee, and N. P. Smart. Generic Constructions of Identity-Based and Certificateless KEMs. *J. Cryptology*, 21(2):178–199, 2008.
4. D. Boneh and M. K. Franklin. Identity-Based Encryption from the Weil Pairing. *SIAM J. Comput.*, 32(3):586–615, 2003.
5. D. Boneh, C. Gentry, and M. Hamburg. Space-Efficient Identity Based Encryption Without Pairings. In *FOCS'07*, pages 647–657. IEEE, 2007.
6. X. Boyen, Y. Dodis, J. Katz, R. Ostrovsky, and A. Smith. Secure Remote Authentication Using Biometric Data. In *EUROCRYPT'05*, volume 3494 of *LNCS*, pages 147–163. Springer, 2005.
7. J. Bringer, H. Chabanne, G. D. Cohen, B. Kindarji, and G. Zémor. Optimal Iris Fuzzy Sketches. In *BTAS'07*, pages 1–6. IEEE, 2007.
8. J. Bringer, H. Chabanne, M. Izabachène, D. Pointcheval, Q. Tang, and S. Zimmer. An Application of the Goldwasser-Micali Cryptosystem to Biometric Authentication. In *ACISP'07*, volume 4586 of *LNCS*, pages 96–106. Springer, 2007.
9. A. Burnett, F. Byrne, T. Dowling, and A. Duffy. A Biometric Identity Based Signature Scheme. *International Journal of Network Security*, 5(3):317–326, 2007.
10. C. Chen, R. N. J. Veldhuis, T. A. M. Kevenaar, and A. H. M. Akkermans. Multi-bits biometric string generation based on the likelyhood ratio. In *BTAS'07*, pages 1–6. IEEE, 2007.
11. Y. Dodis and A. Smith. Correcting errors without leaking partial information. In *STOC'05*, pages 654–663. ACM, 2005.
12. E. Fujisaki and T. Okamoto. Secure integration of asymmetric and symmetric encryption schemes. In *Crypto'99*, volume 1666 of *LNCS*, pages 537–554. Springer, 1999.
13. J. Furukawa, N. Attrapadung, R. Sakai, and G. Hanaoka. A Fuzzy ID-Based Encryption Efficient When Error Rate Is Low. In *INDOCRYPT'08*, volume 5365 of *LNCS*, pages 116–129. Springer, 2008.

14. T. Kitagawa, Yang P, G. Hanaoka, R. Zhang, H. Watanabe, K. Matsuura, and H. Imai. Generic Transforms to Acquire CCA-Security for Identity Based Encryption: The Cases of FOpkc and REACT. In *ACISP'06*, volume 4058 of *LNCS*. Springer, 2006.
15. Q. Li, Y. Sutcu, and N. D. Memon. Secure Sketch for Biometric Templates. In *ASIACRYPT'06*, volume 4284 of *LNCS*, pages 99–113. Springer, 2006.
16. T. Okamoto and D. Pointcheval. REACT: Rapid Enhanced-Security Asymmetric Cryptosystem Transform. In *CT-RSA'01*, volume 2020 of *LNCS*, pages 159–175. Springer, 2001.
17. M. Pirretti, P. Traynor, P. McDaniel, and B. Waters. Secure Attribute-Based Systems. In *ACM CCS'06*, pages 99–112. ACM, 2006.
18. A. Sahai and B. Waters. Fuzzy Identity-Based Encryption. In *EUROCRYPT'05*, volume 3494 of *LNCS*, pages 457–473. Springer, 2005.
19. N. D. Sarier. A New Biometric Identity Based Encryption Scheme. In *TrustCom'08*, pages 2061–2066. IEEE, 2008.
20. N. D. Sarier. Biometric Identity Based Signature Revisited. In *EuroPKI'09*. Springer, to appear, 2009.
21. S. F. Shahandashti and R. Safavi-Naini. Threshold Attribute-Based Signatures and Their Application to Anonymous Credential Systems. In *AFRICACRYPT'09*, volume 5580 of *LNCS*, pages 198–216. Springer, 2009.
22. Q. Tang, J. Bringer, H. Chabanne, and D. Pointcheval. A Formal Study of the Privacy Concerns in Biometric-Based Remote Authentication Schemes. In *ISPEC'08*, volume 4991 of *LNCS*, pages 56–70. Springer, 2008.
23. P.P. van Liesdonk. Anonymous and Fuzzy Identity-Based Encryption. Master's thesis, Technische Universiteit Eindhoven, 2007.
24. P. Yang, Z. Cao, and X. Dong. Fuzzy Identity Based Signature. Cryptology ePrint Archive, Report 2008/002, 2008. http://eprint.iacr.org/.

Acknowledgement

The author is grateful to her supervisor Prof. Dr. Joachim von zur Gathen for his valuable support, encouragement and guidance.

Appendix A: Fuzzy Extractor

Formally, an $(\mathcal{M}, l, t)$ fuzzy extractor is defined as follows [9]. Let $\mathcal{M} = \{0,1\}^v$ be a finite dimensional metric space with a distance function $\mathbf{dis} : \mathcal{M} \times \mathcal{M} \rightarrow \mathbb{Z}^+$. Here, $b \in \mathcal{M}$ and **dis** measures the distance between b and b', where $b, b' \in \mathcal{M}$. An $(\mathcal{M}, l, t)$ fuzzy extractor consists of two functions **Gen** and **Rep**.

- **Gen**: A probabilistic generation procedure that takes as input $b \in \mathcal{M}$ and outputs an biometric identity string $ID \in \{0,1\}^l$ and a public parameter PAR, that is used by the **Rep** function to regenerate the same biometric string ID from b' such that $\mathbf{dis}(b, b') \leq t$.
- **Rep**: A deterministic reproduction procedure that takes as input b' and the publicly available value PAR, and outputs ID if $\mathbf{dis}(b, b') \leq t$.

In [9], the authors describe a concrete fuzzy extractor using a $[n, k, 2t{+}1]$ BCH error correction code, Hamming Distance metric and a one-way hash function $H : \{0,1\}^n \rightarrow \{0,1\}^l$. Specifically,

- The **Gen** function takes the biometrics b as input and returns $ID = H(b)$ and public parameter $PAR = b \oplus C_e(ID)$, where C_e is a one-to-one encoding function.
- The **Rep** function takes a biometric b' and PAR as input and computes $ID' = C_d(b' \oplus PAR) = C_d(b \oplus b' \oplus C_e(ID))$. $ID = ID'$ if and only if $\mathbf{dis}(b, b') \leq t$. Here C_d is the decoding function that corrects the errors upto the threshold t.

Appendix B: A Concrete Application

The second application of our generic construction is based on the scheme of [1], whose security relies on the quadratic residuosity assumption in the ROM. Similar to the scheme of [5], an n-bit message (n is the length of c) is encrypted bit by bit resulting in a ciphertext of $2n(120{+}1024)$ bits if necessary optimizations are applied. Thus, it could be used as an input to our generic constructions with the following modifications to avoid collision attacks.

1.$\mathsf{Setup}(k_0)$: Let H be a full domain hash function $H : \mathcal{BID} \rightarrow \mathbb{Z}_N^*[{+}1]$ with $\mathcal{BID} = \mathcal{W} \times \mathcal{ID}$ and k_0 a security parameter. Generate two primes (p, q) and compute $N = pq$, where N is a k_0-bit Blum integer and p, q are two $k_0/2$-bit primes each congruent to 3 modulo 4. The public parameters are $params = (N, H)$ and the master secret key is $msk = (p, q)$. We assume that the features $w_j \in \mathcal{W}$ are ordered as in [15].

2.$\mathsf{Keygen}(msk, w, n)$: It takes as input msk, a biometric vector w with length n. The algorithm outputs a private key $D_{\mathsf{BID}} = (r_1, ..., r_n)$ for decrypting encryptions of n-bit messages as follows. For $j{=}$ 1,...,n do:

- $a_j \leftarrow H(w_j, \mathsf{ID})$ Thus, the jacobi symbol $(\frac{x}{N}) = +1\}$.
- let $r_j \in \mathbb{Z}_N^*$ such that $r_j^2 \equiv a_j \bmod (N)$ or $r_j^2 \equiv -a_j \bmod (N)$

3.$\mathsf{Enc}(params, w', c)$:The encryption algorithm that takes as input biometrics w' of the receiver, $params$ and a message $c = c_1...c_n \in C$. For $j = 1, ..., n$, choose at random $t_j, v_j \in \mathbb{Z}_N^*$ such that $(\frac{t_j}{N}) = (\frac{v_j}{N} = c_j\}$ and compute $a_j = H(\ w_j, \mathsf{ID})$. Next, compute $(f_j, g_j) = (t_j + \frac{a_j}{t_j}, v_j - \frac{a_j}{v_j})$ and mask the ciphertext using one of the constructions in [1]. Next, the encryptor sends $U = (Z_1^j, \alpha_1^j, ..., \alpha_l^j)||(Z_2^j, \beta_1^j, ..., \beta_l^j||\mathcal{L})||MID_c$, where $\mathcal{L}$ is a label that contains information about how each component of the ciphertext is associated to the index of $w_j \in \mathcal{W}$ and MID_c is a message identifier for c.

4.$\mathsf{Dec}(U, D_{\mathsf{BID}})$: The decryption algorithm is performed as described in [1].