Privacy-Preserving Biometric-Based Remote User Authentication

Yangguang Tian, Yingjiu Li, Ximeng Liu, Robert H. Deng, Binanda Sengupta

School of Information Systems, Singapore Management University, Singapore
{yggtian, yjli, xmliu, robertdeng, binandas}@smu.edu.sg

Abstract

Biometric-based remote user authentication (BRUA) is a useful primitive that allows an authorized user to remotely authenticate to a cloud server using biometrics. However, the existing BRUA solutions in the client-server setting lack certain privacy considerations. For example, authorized user’s multiple sessions should not be linked while her identity remains anonymous to cloud server. In this work, we introduce a new framework for biometric-based remote user authentication, such that authorized users authenticate to an honest-but-curious server in an anonymous and unlinkable manner. In particular, we employ two non-colluding cloud servers to perform the complex biometrics matching. We formalize two new security models, including biometrics privacy and user privacy, for the proposed framework, and prove the security of the proposed framework in the standard model.

Keywords: Remote user authentication, Unlinkability, Biometrics privacy, Biometrics matching

1 Introduction

Biometric-based user authentication has been widely used in many real-life applications, such as mobile security, financial transactions and identification checks [1]. There are some attractive features using biometrics over conventional password. For example, people need to remember many secure passwords for many different accounts and update passwords frequently for security reasons. By contrast, biometrics is permanently and uniquely associated with an individual, so the individual can use biometrics for user authentication.

Biometric-based user authentication also leads to some security and privacy concerns. First, biometrics is not revocable. If biometrics is compromised, then the user may lose her security forever, especially for the single-factor biometric-based user authentication. Second, authorized users may concern the privacy of biometrics stored on the authentication server. Therefore, no biometrics should be stored in plaintext, because biometrics may contain a wealth of personal information (e.g., DNA).

To protect biometrics information, there are mainly three methods in the literature: non-invertible transform [2], fuzzy extractors [3] and homomorphic cryptosystem [4]. The noninvertible transform relies on a static secret key, which is always available at the time of authentication to transform the requested biometrics for user authentication. This method is actually a two-factor (i.e., biometrics plus secret key) user authentication, and is not scalable for cross-platform setting, because in practice users may own several devices (e.g., smart-phone, pad and tablet) and access to the same service provider from various platforms. The fuzzy extractors based user authentication [5-6] is a single-factor user authentication. However, deriving a secret key from biometrics and other noisy data with high stability and entropy simultaneously is a non-trivial task.

Using homomorphic encryption [4] to protect biometrics information is a promising approach when designing biometric-based user authentication. In particular, if homomorphic encryption is deployed, then authentication servers do not need to interact with an authorized user when performing user authentications. We stress that the authentication server in cloud can perform complex mathematical computations (i.e., biometrics matching) in the encrypted format, because cloud computing provides ubiquitous, dynamic, scalable and on-demand services. That is, the cloud-based biometrics can facilitate efficient biometrics matching for user authentication. In this work, we focus on biometric-based remote user authentication (BRUA) using homomorphic encryption, where authorized users wish to remotely authenticate to an authentication server using encrypted biometrics.

The privacy should be preserved not only for biometrics information, but also for non-biometrics information such as identity, behaviour and interaction history. Identity-concealment is an important privacy property and is mandated or recommended by widely standardized and deployed cryptographic protocols, such as TLS1.3 and QUIC [7]. Identity-concealment means that the transcript of protocol execution should not leak authorized user’s real identity. Moreover, unlinkability is also desired, in which multiple sessions
of the same authorized user cannot be linked by the authentication server. The main goal of this work is to design an identity-concealed and unlinkable BRUA using homomorphic encryption.

Homomorphic encryption can be used to encrypt identity information of authorized users during the protocol execution. However, if the same anonymous user authenticates twice to an authentication server, then the authentication server can trivially link the anonymous authenticated user to a specific record in his database which stores all enrolled users’ records. We further notice that such kind of unlinkability between authorized user and database record is an important feature for sensitive IT infrastructure such as personal record management systems [8].

Since biometrics matching of BRUA may handle various kinds of distance calculations (e.g., Euclidean distance [9], Hamming distance [5] or Chebyshev distance [6]), a suitable homomorphic encryption primitive is critical to the success of user authentication. Fully homomorphic encryption can easily support all aforementioned distance calculations. Because it enables addition and multiplication simultaneously (on encrypted biometrics) when performing biometrics matching. However, it is still not practical in real-world environment due to its computational cost and system complexity [10].

Instead of fully homomorphic encryption, we rely on partial homomorphic encryption such as Paillier cryptosystem. However, Paillier cryptosystem only supports the additive operations over encrypted biometrics. Sometimes, the multiplicative operations are required when considering the Euclidean distance based biometrics matching. Therefore, how to exploit the Paillier cryptosystem to support complex mathematical operations for biometrics matching is our first challenge task.

Second, biometrics are typically encrypted under user’s own public keys and stored in the authentication server. Since biometrics matching takes different ciphertexts under the same public key as input, the authentication server must transform the ciphertexts under different public keys into the ciphertexts under the same public key. Such transformation is easy when authorized users are identified. However, this contradicts to the user privacy we desired. Hence, achieving an anonymous and unlinkable user authentication is a rather challenging task.

1.1 This Work

In this work, we introduce the notion of privacy-preserving biometric-based remote user authentication (PriBioAuth), allowing authorized users to remotely authenticate to an authentication server using encrypted biometrics. Our proposed solution employs two (non-colluding) honest-but-curious cloud servers in the system [11-12], one acts as authentication server, while the other one acts as a dedicated computational server which works with authentication server to assist certain biometrics matching.

For the anonymous and unlinkable PriBioAuth, we first propose an anonymous key transformation (AKeyTrans) protocol, such that authentication server performs the key transformation in an anonymous manner. Meanwhile, inspired by the concept of oblivious access control [13-14], we allow authenticated users authenticate themselves to an authentication server obliviously. Putting anonymous key transformation and oblivious access control together, the proposed PriBioAuth can achieve the claimed user privacy. Our overall contributions can be summarized as follows.

Security and privacy guarantee. We provide the formal security requirements for privacy-preserving biometric-based remote user authentication protocols. We formalize two formal security models which include various kinds of security and privacy properties, such as biometrics privacy, obliviousness of access control, identity-concealment (anonymity) and unlinkability.

Practical and flexible constructions. We present two practical solutions for biometric-based remote user authentication using two non-colluding cloud servers, an authentication server and a computational server. First, we present a basic construction with biometrics privacy, which is useful if authorized users wish to be recognized by authentications servers. We also present a PriBioAuth construction, in which both biometrics privacy and user privacy are addressed simultaneously.

Secure biometrics matching. The authentication server in conjunction with the computational server can perform various kinds of mathematical computations for biometrics matching. We provide a set of secure multi-party computation (SMC) sub-protocols to guarantee the success of biometric-based remote user authentication, including less than, equivalent testing and multiplicative computation protocols. In particular, no user interaction is required for biometrics matching.

Scalability of use. It is easy to employ our solutions in a cross-platform setting. Because the proposed solutions are a single-factor user authentication without generating extra secret keys at the time of authentication.
biometrics template and identity (or pseudonym). However, we assume biometrics is a secret information in this work.

Homomorphic encryption (see below) is a suitable cryptographic tool to protect biometrics instead of non-invertible transform and fuzzy extractors. In particular, it supports the secure multi-party computations (SMC) on encrypted biometrics for biometrics matching. Note that some well-known works [17-21] have used the Paillier cryptosystem as encryption primitive to protect user’s biometrics. For example, Huang et al. [19] proposed a flexible biometric-based identification framework. They use the garbled circuit to efficiently and obliviously perform biometrics matching and retrieve the outcome of results. However, the authentication server should interact with authorized user to finalize the biometrics matching.

Bringer et al. [15] proposed a biometric-based user authentication protocol using Goldwasser-Micali (GM) cryptosystem [22]. Note that the GM cryptosystem takes the binary string (such as Iris [23]) as input. To allow Paillier cryptosystem process the binary input, Schoenmakers and Tuyls [24] proposed a generic framework, such that the underlying Paillier cryptosystem [25] can process binary string for biometrics matching. That is, the Paillier cryptosystem can handle bits strings using their proposed binary conversion.

Homomorphic encryption. Homomorphic encryption (HE) is a well-known approach for privacy-preserving secure multiparty computation. There are mainly two types of HE system in the literature: one is full FE, and the other is partial HE. The latter type consists of additive homomorphic encryption and multiplicative homomorphic encryption separately, while the former type can support both addition and multiplication over ciphertext simultaneously. We omit the somewhat HE for simplicity.

Gentry [4] proposed the first full HE scheme based on lattice-based cryptography. While a number of following works [10, 26] have been proposed afterwards, it is still not practical to implement in real-life applications. The partial homomorphic encryption is often considered as a suitable alternative in practice. For example, Paillier cryptosystem [25] supports addition over ciphertext, while ElGamal cryptosystem [27] supports multiplication over ciphertext.

Based on the practical Paillier cryptosystem, Peter et al. [11] proposed an efficient outsourcing SMC protocol which is proven to be secure in the honest-but-curious model. In particular, their proposed method can be used for privacy-preserving face authentication. Later on, Liu et al. [12] proposed an efficient outsourcing toolkits for SMC protocols. To support various computations (e.g., multiplication, less than and division) in cloud, Liu et al. proposed a new cryptographic primitive: distributed two trapdoors public key cryptosystem (DT-PKC) (which is an extension from [28]). This work aims to exploit some inherent features of DT-PKC for user authentications. In particular, we discover that such kind of homomorphic cryptosystems [11, 12, 28] have a “key privacy” [29] property.

The remainder of this paper is structured as follows: In the next Section, we formalize the system model and the threat models (namely, biometrics privacy and user privacy). In Section 3, we describe some preliminaries which will be used in our proposed constructions, and present the proposed constructions in Section 4. We then present our security analysis and performance analysis in Section 5 and 6 respectively. The paper is concluded in Section 7.

2 Security Model

In this section, we present the corresponding models for privacy-preserving biometric-based remote user authentications. As mentioned in the introduction, a user authentication should achieve several security and privacy goals: biometrics privacy and user privacy. We present the commonly used notations in Figure 1.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{pk}_i, \text{sk}_i$</td>
<td>User i’s public/secret key</td>
</tr>
<tr>
<td>ID$_i$</td>
<td>Identity of user $i$</td>
</tr>
<tr>
<td>dist$(x, y)$</td>
<td>Distance between vector $x$ and vector $y$</td>
</tr>
<tr>
<td>$t \in \mathbb{R}^+$</td>
<td>Threshold value (positive real number)</td>
</tr>
<tr>
<td>$B$</td>
<td>Plain biometrics</td>
</tr>
<tr>
<td>$C$</td>
<td>Reference biometrics</td>
</tr>
<tr>
<td>$\overline{N}$</td>
<td>Dimension of biometrics</td>
</tr>
<tr>
<td>$\mathbb{Z}_n$</td>
<td>Finite field</td>
</tr>
<tr>
<td>$n$</td>
<td>Number of users</td>
</tr>
<tr>
<td>$k$</td>
<td>Number of secret credentials</td>
</tr>
<tr>
<td>$[x]$</td>
<td>Encryption on $x$ under the public key $\text{pk}$</td>
</tr>
<tr>
<td>$(\overline{N}, g)$</td>
<td>Public parameters in DT-PKC</td>
</tr>
<tr>
<td>$S$</td>
<td>Splitting technique in DT-PKC</td>
</tr>
<tr>
<td>$\text{Enc}$</td>
<td>Encryption algorithm in DT-PKC</td>
</tr>
<tr>
<td>$\text{Dec}$</td>
<td>Decryption algorithm in DT-PKC</td>
</tr>
<tr>
<td>$\text{PD}(1/2)$</td>
<td>Partial decryption algorithm in DT-PKC</td>
</tr>
</tbody>
</table>

Figure 1. Summary of notations

2.1 System Model

We present a biometric-based remote user authentication system (see Figure 2) involving three types of entities: key generation center (KGC), requested user (RU) and authentication server CP (which may consist of an additional computational cloud server (CSP)). We then define a biometric-based remote user authentication which consists of the following algorithms:

Setup. The KGC takes the security parameter $O$ as input, outputs a master public/secret key pair $(\text{mpk}, \text{msk})$. In addition, KGC outputs a set of credentials $\{\text{msk}_i\}^k_{i=1}$, and distributes them to respective CP and CSP through a secure channel.
Figure 2. Biometric-based remote user authentication under consideration

**KeyGen.** User takes master public key $mpk$ as input, outputs a public/secret key pair $(pk, sk)$.

**Registration.** User enrolls her identity ID along with a reference biometrics $C$ to CP. There may exist an interactive algorithm between the CP and a CSP, in cloud. Note that user becomes a RU after registration, and the binding between user’s real identity ID and her public key $pk$ is authenticated by a certificate $cert$ issued by KGC.

**Authentication.** RU sends her identity ID and a candidate biometrics $C'$ to the cloud server CP, then CP accept it if and only if $\text{dist}(C', C) \leq t$. There may exist an interactive algorithm between CP and CSP in cloud. Note that both the reference and candidate biometrics are in an encrypted format, more specifically, they are encrypted under user’s own public key.

### 2.2 Threat Model

As mentioned in the introduction, a biometric-based remote user authentication should achieve biometrics privacy and user privacy.

#### 2.2.1 Biometrics Privacy

Informally, an adversary attempts to learn user’s plain biometrics. Below is the formal biometrics privacy game between an adversary $A$ and a simulator $S$.

- **Setup:** $S$ first generates public/secret key pairs $(pk_i, sk_i)$ ($i \in [1, n]$) for $n$ users and $m$ servers respectively in the system. In addition, $S$ generates a set of secret credentials $\{sk^{1/j}_{i,j} \}_{j=1}^k$ for $k (k \leq m)$ servers. $S$ also generates user’s plain biometrics $\{B_i\}$ and their corresponding reference biometrics $\{C_i\}$, and returns all reference biometrics to $A$. $S$ eventually tosses a random coin $b$ which will be used later in the game.

- **Training:** $A$ can make the following queries in arbitrary sequence to $S$.
  - **Send:** If $A$ issues a send query in the form of $(ID, i, msg)$ (resp. $(CP, i, msg)$) to simulate a network message for the $i$-th session of user ID (resp. server CP), then $S$ would simulate the reaction of instance oracle $\Pi_{ID}$ (resp. $\Pi_{CP}$) upon receiving message $msg$, and return to $A$ the response that $\Pi_{ID}$ (resp. $\Pi_{CP}$) would generate (we denote the $i$-th session established by user ID as instance oracle $\Pi_{ID}^i$). If $A$ issues a Send query in the form of $(ID', 'start')$ (resp. $(CP', 'start')$), then $S$ creates a new instance oracle $\Pi_{ID'}^i$ (resp. $\Pi_{CP'}^i$) and returns to $A$ the first protocol message.
  - **Secret Key Reveal:** If $A$ issues a Secret Key Reveal (or corrupt, for short) query to user $i$, then $S$ returns user $i$’s secret key $sk_i$ to $A$. Note that $A$ is allowed to issue at most $n-1$ Secret Key Reveal queries to $S$. We denote the honest (i.e., uncorrupted) user set as $U^\prime$.
  - **Secret Credential Reveal:** If $A$ issues a credential reveal query to the CP, then $S$ returns CP’s secret credential $sk_{ID}$ to $A$.

- **Challenge:** $A$ randomly chooses two challenge biometrics $(B_0, B_1) (\not\in \{B_i\})$ of a challenge user ID, and sends the challenge biometrics to $S$. $S$ simulates the reference biometrics of user $U_i$ by either $C'_b = F(pk_b, B_0)$ if $b = 0$ or $C'_b = F(pk_b, B_1)$ if $b = 1$. Note that $A$ is allowed to reveal $k$-1 secret credentials (by corrupting servers), and $F$ denotes a probabilistic algorithm. Finally, $A$ outputs $b'$ as its guess for $b$. If $b' = b$, then $S$ outputs 1; Otherwise, $S$ outputs 0. We define the advantage of an adversary $A$ in the above game as

$$Adv_A(O, k) = |Pr[S \rightarrow 1] - 1/2|.$$  \hspace{1cm} (1)

**Definition 1.** We say that a PriBioAuth scheme has biometrics privacy if for any probabilistic polynomial-time (PPT) $A$, $Adv_A(O, k)$ is a negligible function of the security parameter $O$.

#### 2.2.2 User Privacy

Informally, an adversary attempts to identify the users involved in a biometric-based remote user authentication protocol. Below is the formal user privacy game between an adversary $A$ and a simulator $S$.

**Setup.** $S$ first generates public/secret key pairs $(pk_i, sk_i)$ ($i \in [1, n]$) for $n$ users and $m$ servers respectively in the system. In addition, $S$ generates a set of secret credentials $\{sk^{1/j}_{i,j} \}_{j=1}^k$ for $k (k \leq m)$ servers. $S$ also
generates user’s plain biometrics \( \{ B_i \} \) and their corresponding reference biometrics \( \{ C_i \} \), and returns all public information (including \( \{ C_i \} \)) to \( A \). \( S \) eventually tosses a random coin \( b \) which will be used later in the game.

**Training.** \( A \) is allowed to issue send query, at most \( n-2 \) Secret Key Reveal and \( k-1 \) Secret Credential Reveal queries to \( S \). We denote the honest (i.e., uncorrupted) user set as \( U' \).

**Challenge.** \( A \) randomly selects two users \( ID_b, ID_j \in U' \) as challenge candidates, then \( S \) removes them from \( U' \) and simulates \( ID^*_b \) to \( A \) by either \( ID^*_b = ID_i \) if \( b = 0 \) or \( ID^*_b = U_j \) if \( b = 1 \).

\[
A \leftrightarrow ID^*_b = \begin{cases} ID_i & b = 0 \\ ID_j & b = 1 \end{cases}
\]

Let \( A \) interact with \( ID^*_b \). Finally, \( A \) outputs \( b' \) as its guess for \( b \). If \( b' = b \), then \( S \) outputs 1; Otherwise, \( S \) outputs 0. We define the advantage of an adversary \( A \) in the above game as

\[
Adv_A(O, k) = |Pr[S \rightarrow 1] - 1/2|.
\]

**Definition 2.** We say that a PriBioAuth scheme has user privacy if for any PPT \( A \), \( Adv_A(O, k) \) is a negligible function of the security parameter \( O \).

**Remark.** We assume a passive adversary, who is able to monitor or eavesdrop (except modify or tamper) all transcripts transmitted on the network. We consider an honest-but-curious model in this work, which is formalized by some existing works [11, 19, 21]. Specifically, the request user and the authentication server are assumed to execute the protocol as specified, just try to learn additional information from the transcript and intermediate results during protocol execution.

### 3 Preliminaries

We briefly present some secure computation protocols described in [12], which will be used in our proposed user authentication constructions. We just mention their functionalities for simplicity.

**Secure Less Than Protocol (SLT).** We assume two encrypted integers \([x]\) and \([y]\), the SLT protocol will provide an encrypted results \([u]\), which can be used to determine the relationship between the plaintexts of two encrypted integers (i.e., \( x > y \) or \( x \leq y \)). As a result, \( u = 0 \) indicates \( x > y \), and \( u = 1 \) indicates \( x \leq y \).

**Secure Equivalent Testing Protocol (SEQ).** Given two encrypted integers \([x]\) and \([y]\), SEQ will provide the encrypted results \([f]\) to determine whether the plaintext of the two encrypted integers are equivalent (i.e., \( x = y \)). As a result, \( f = 1 \) indicates \( x = y \), and \( f = 0 \) indicates \( x \neq y \).

**Secure Multiplicative Computation Protocol (SMT).** Given two encrypted integers \([x]\) and \([y]\) as input, the SMT can generate the result \([x \cdot y]\) by using two non-colluding cloud servers CP and CSP.

### 3.1 Secure Euclidean Distance Computation Protocol (SEDC)

We present the proposed secure Euclidean distance computation protocol. We use Fingerprints as the candidate of biometrics, which is represented by FingerCode. The Finger-Code [30] is typically a \( N \)-dimensional (e.g., \( N=640 \)) feature vector, and each entry is a 8-bit integer. The Euclidean distance \( d = \text{dist} (B, B') \) between reference biometrics \( B = (v_1, \ldots, v_n) \) and candidate biometrics \( B' = (v'_1, \ldots, v'_n) \) is calculated as

\[
\begin{align*}
&d = \sum_{j=1}^{N} (v_j - v'_j)^2 \\
&= (v_1 - v'_1)^2 + (v_2 - v'_2)^2 + \cdots + (v_n - v'_n)^2 \\
&= \sum_{j=1}^{N} v_j^2 + \sum_{j=1}^{N} (-2v_j \cdot v'_j)^2 + \sum_{j=1}^{N} v'_j^2.
\end{align*}
\]

Note that the CP and CSP perform the biometrics matching between encrypted vectors \([B] = ([v_j])_{j=1}^{N}\) and \([B'] = ([v'_j])_{j=1}^{N}\) as shown in Figure 3. In particular,

\[
h = m \cdot m' = \sum_{j=1}^{N} ((v_j - v'_j) \cdot r_{j,1}) (v_j - v'_j) \cdot r_{j,2}), \quad \text{and CP}
\]

performs the following calculation

\[
d = H \cdot S_1 \cdot S_2 \cdot S_3
\]

\[
\begin{align*}
&= \left[ \sum_{j=1}^{N} ((v_j - v'_j)^2 \cdot r_{j,1}) ] (v_j - v'_j)^2 \cdot r_{j,2} ] \\
&- \sum_{j=1}^{N} r_{j,1} \cdot (v_j - v'_j)^2 + \sum_{j=1}^{N} r_{j,2} \cdot (v_j - v'_j) \\
&+ r_{j,1} \cdot r_{j,2} \right] \right] = \sum_{j=1}^{N} (v_j - v'_j)^2
\end{align*}
\]

### 3.2 Another Look of DT-PKC

The underlying DT-PKC is the main building block of the proposed constructions. We discover that the DT-PKC has an inherent feature: “key privacy”, which is introduced by Bellare et al. [29]. It means that an adversary in possession of a ciphertext cannot tell which specific key, out of a set of known public keys, is the one under which the ciphertext was created. In particular, they formalized a new model: “indistinguishability of keys” (IK). We formally prove the DT-PKC cryptosystem is secure in the IK-CPA model, in addition to its IND-CPA security [12]. We believe that both BCP [28] and its variant DT-PKC
cryptosystem have such implicit property.

3.2.1 Security Model of Key Privacy

Definition 3. The IK-CPA experiment between an adversary $A$ and a simulator $S$ is defined below [29].

$$
\text{Experiment } \text{Exp}^{\text{IK-CPA}}(O) = \langle \text{KeyGen}(1^2), \langle (pk_0, sk_0), (pk_1, sk_1) \rangle \leftarrow \text{KeyGen}(1^2), (msg^*, st) \leftarrow A(\text{find}, pk_0, pk_1), C^* \leftarrow \text{Enc}_{pk_0}(msg^*) \rangle
$$

$$
b' = A(\text{guess}, st, C^*)$$

If $b' = b, \text{return } 1$; else, return 0.

Note that $st$ denotes some state information. We define the advantage of the adversary as

$$
\text{Adv}^{\text{IK-CPA}}_{PE}(O) = |\Pr[O \rightarrow 1] - 1/2|.
$$

Definition 4. An encryption scheme $(\text{PE}, \text{KeyGen}, \text{Enc}, \text{Dec})$ is said to be IK-CPA secure if $\text{Adv}^{\text{IK-CPA}}_{PE}(O)$ is negligible in $O$ for any PPT adversary $A$.

3.2.2 Security of DT-PKC

We prove the DT-PKC is IK-CPA secure if the underlying DDH assumption holds in group $\mathbb{Z}_N^*$. In particular, we assume the factorization of the modulus $N$ is hard, or the DDH assumption over $\mathbb{Z}_N^*$ turns out to be easy (refer to Theorem 4 in [28] for detailed relations).

Theorem 1. The DT-PKC achieves IK-CPA security if the DDH assumption holds in $\mathbb{Z}_N^*$.

Proof. Assume that there exists a PPT adversary $A$ breaking the IK-CPA security of the DT-PKC scheme, then we can construct an algorithm $S$ to break the DDH assumption over $\mathbb{Z}_N^*$. The algorithm $S$ has almost the same time complexity with $A$.

The adversary $S$ will use $A$ as a subroutine (see Figure 4). $S$ first generates another DH tuple ($(X_1, Y_1, Z_1)$ mod $N^2$) which has the same property and distribution as its own challenge tuple $(X, Y, Z)$ using DDH random self-reducibility [31]. That means if its challenge is a real DH tuple, then it is the computed tuple; Otherwise, it is a random tuple in $\mathbb{Z}_N^*$. Then using its challenge and computed tuples, $S$ outputs two challenge public keys for $A$ (in the find stage). At the end of find stage, $A$ submits a challenge message $msg^*$ and some state information $st$ to $S$. $S$ takes challenge message $msg^*$ and $st$ as input, outputs a challenge ciphertext which is an encryption of $msg^*$ under $pk_b$ according to the bit $b$. In addition, $S$ randomly chooses $msk_{0,1}(i = \{1, 2\})$ from the interval $[1, N(N - 1)/2]$ as secret credentials.

$$
\text{Adversary } S(N, g, X, Y, Z)
$$

$\begin{align*}
& u, v, w \leftarrow \mathbb{Z}_N^* \\
& X_0 = X, Y_0 = Y, T_0 = T \\
& X_1 = X_0 \cdot g^v \mod N^2; Y_1 = Y_0 \cdot g^w \mod N^2 \\
& T_1 = T^u \cdot X_1 \cdot Y_1 \cdot g^{uw} \mod N^2 \\
& pk_0 = (N, g, X_0); pk_1 = (N, g, X_1) \\
& (msg^*, st) \leftarrow A(\text{find}, pk_0, pk_1) \\
& msk = \begin{cases} 
0 & \text{if } b = 0 \\
1 & \text{if } b = 1
\end{cases} \\
& \text{mask}(i) \leftarrow \begin{cases} 
1/N - 1/2 & \text{if } b' = b, \text{return } 1; \text{else, return } 0.
\end{cases}
\end{align*}
$$

Figure 4. Description of adversary $S$

We then analyze the behaviour of $S$ on $\text{Exp}_S^{\text{DDH-REAL}}$ and $\text{Exp}_S^{\text{DDH-RAND}}$ respectively. In the $\text{Exp}_S^{\text{DDH-REAL}}$, the input $(X, Y, Z)$ mod $N^2$ to $S$ satisfy $T = g^{wy} \mod N^2$ where $x, y, w \in \mathbb{Z}_N^*$. Notice that the computed tuple $(X_1, Y_1, Z_1)$ mod $N^2$ is also valid and they are uniformly and independently distributed in interval $[1, N^2]$, because $X_1 = g^{xu} \mod N^2$, $Y_1 = g^{xy} \mod N^2$, $Z_1 = g^{(xu+wy)} \mod N^2$ and $u, v, w$ are random element in $\mathbb{Z}_N^*$. Thus, the $X_0, X_1$ have the proper distribution of two challenge public keys, and the challenge ciphertext is distributed exactly like a real DT-PKC encryption of message under public key $pk_b$. Meanwhile, the randomly chosen secret credential is statistically indistinguishable from a real secret credential (from interval $[1, \lambda \cdot N]$) from $A$’s point of view. Therefore, we have

$$
\begin{align*}
\text{Pr}[\text{Exp}_S^{\text{DDH-REAL}}(O) = 1] &= 1/2 \cdot \text{Pr}[\text{Exp}_S^{\text{IK-CPA}*}(O) = 1] \\
&= 1/2 \cdot (1 - \text{Pr}[\text{Exp}_S^{\text{IK-CPA}*}(O) = 1]) \\
&= 1/2 + 1/2 \cdot \text{Adv}^{\text{IK-CPA}}_{PE}(O).
\end{align*}
$$

As for $\text{Exp}_S^{\text{DDH-RAND}}(O)$, the input $(X, Y, Z)$ mod $N^2$ to $S$ in Figure 4 are all uniformly distributed in group $\mathbb{Z}_N^*$. Therefore, the corresponding computed values
above are all uniformly and independently distributed over \(\mathbb{Z}_N^\ast\). In particular, the challenge ciphertext is also random elements in \(\mathbb{Z}_N^\ast\), and independent of bit \(b\). Hence we have

\[
\Pr[\operatorname{Exp}^\text{DDH-RAND}_S(0) = 1] \leq 1/2 + 1/2^{O(1)}.
\]

The last term indicates that the random input \((X, Y, Z)\) to \(S\) happen to have the distribution of a valid Diffie-Hellman tuple, which has a negligible probability \(1/2^{O(1)}\) since \(2^{O(1)} < \lambda < 2^\lambda\). Also note the randomly chosen secret credential and real secret credential has a negligible \(1/2^\ell\) (\(\ell\) is the half bit-length of the modulus \(N\)) statistical distance, and the randomly chosen secret credentials are consistent values to both \(\operatorname{Exp}^\text{DDH-REAL}_S\) and \(\operatorname{Exp}^\text{DDH-RAND}_S\). By combing all equations above, we have

\[
\mathbf{Adv}^\text{DDH}_S(0) = \Pr[\operatorname{Exp}^\text{DDH-REAL}_S(0) = 1] + \Pr[\operatorname{Exp}^\text{DDH-RAND}_S(0) = 1] \\
\geq 1/2 - \Pr[\operatorname{Exp}^\text{IK-CPA}_4(0) - 1/2^{O(1)} - 1/2^\ell].
\]

### 4 Proposed Construction

**Basic construction.** We present a basic construction to show how secure computational sub-protocols are executed in the BRUA. We focus on biometrics privacy only in the basic construction.

- **Setup.** KGC takes the security parameter as input, outputs master public/secret key pair \((\text{mpk}, \text{msk})\). In addition, KGC outputs two secret credentials \((\text{msk}^{(1)}; \text{msk}^{(2)}) \leftarrow S(\text{msk})\), and distributes them to CP and CSP respectively.

- **KeyGen.** User takes master public key \(\text{mpk}\) as input, outputs a public/secret key pair \((pk, sk)\).

- **Registration.** User registers her identity ID along with a reference biometric \([B]\) (i.e., \(\text{Enc}_{\text{sk}}(B)\)) to CP, where \(B = (v_1, \ldots, v_N)\).

- **Authentication.** The interaction among RU, CP and CSP is described as follows.
  - RU \(\rightarrow\) CP: RU sends her identity ID and candidate biometrics \([B']\) (i.e., \(\text{Enc}_{\text{sk}}(B')\)) to CP;
  - CP \(\leftrightarrow\) CSP:
    1. Run the SEDC protocol to obtain a Euclidean distance \([d] = \|\text{dist}(B, B')\|\);
    2. Run the modified SLT protocol which takes \([d]\) and \([z]\) as input, outputs the relationship \([u]\) between the plaintext of two encrypted data. Note that \(u = 1\) if \(d \leq t\); Otherwise, \(u = 0\).
  - CP: CP chooses a nonce \(r\) and computes \([u + r]\);
  - Then CP decrypts it using secret credential \(sk^{(1)}\), sends partially decrypted ciphertext and \([u + r]\) to CSP; While CSP obtains \(u + r\) using secret credential \(sk^{(2)}\) and encrypts it under public key of CP \((pk)\). Eventually, CP accepts RU if \((u + r) - r = 1\); Otherwise, CP outputs \(\perp\).

**Proposed PriBioAuth Construction.** We now present our privacy-preserving biometric-based remote user authentication (PriBioAuth) framework. KGC first generates two secret credentials and distributes them to CP and CSP respectively. A RU encrypts ID and biometrics using her own public key, and sends them to CP for Registration. As for Authentication, RU sends encrypted ID and candidate biometrics to CP, while CP accepts RU iff the candidate biometrics is “close enough” to RU’s reference biometrics. In particular, we assume that CP stores a set of encrypted identities and biometrics information after Registration.

**Problem Statement.** In the Authentication stage, the candidate biometrics should be compared with reference biometrics in database. First problem is that a set of enrolled biometrics are not under same public key, but the underlying DT-PKC requires homomorphic operations under the same public key. Another problem is that CP should perform biometrics matching between one record in database and candidate identity/biometrics. In other words, CP can trivially link the anonymous authenticated RU to a specific record in the database.

**Design Rational and Overview.** To address the above problems, we first need an additional procedure to fix these “various” encrypted data prior to the actual biometrics matching between CP and CSP. Specifically, CP partially decrypts reference data using distributed secret credential, and sends them to CSP for full decryption on reference data. Then CSP randomly chooses a “dummy” public key \(pk^*\) such that \(pk^* \neq \{pk_s, pk_{CP}\}\), and re-encrypts data using \(pk^*\) (the corresponding secret key \(sk^*\) is unknown to all RU, CP and CSP, while \(pk^*\) is known to all users).

After anonymous key transformation (AKeyTrans) during Authentication, CP and CSP run the corresponding SLT and SEQ protocols on candidate identity and reference identity. Consequently, CP and CSP run the SEDC and SLT protocols to obtain the relationship between candidate biometrics and reference biometrics. If both SEQ protocol and SLT protocol output “\(\{1\}\)” (encryption under public key \(pk^*\)), then CP authenticates a requested user RU.

To achieve the claimed user privacy, CP will go through all records in database when authenticating a RU. More precisely, CP obtains a set of individual encrypted results \([\{0\}^*, \{1\}^*, \ldots, \{0\}^*\) after going through the entire database; then CP can obtain the encrypted final results \([1] = \{0\}, \{1\}^*, \ldots, \{0\}^*\). After interacting with CSP, CP outputs the plain authentication results “1” iff the candidate identity/biometrics is matching one of records in database. We present the detailed PriBioAuth.
framework below, note that both Setup and KeyGen algorithm follow the basic construction.

- **Registration.** User randomly chooses a nonce $r$ first, then computes reference identity $[ID]$, biometrics $[B]$ (i.e., $Enc_p(B)$), and two encrypted nonces $[r]$, $[r]_*$ (the second one is using public key $pk*$). Eventually, user sends her identity ID and all encrypted values to CP. In particular, CP and CSP perform the AKeyTrans protocol as described in Figure 5.

  Note that $B = (v_1, \ldots, v_N) = \{v_i\},$ and CP holds a set of transformed reference identity/biometrics $([ID], [ID]_*, [B]_*)$ under public key $pk*$.

- **Authentication.** RU generates the candidate request using the same method described above, and sends a message $([ID], [B]_*, [r]_*)$ as authentication Request to CP. Then CP and CSP take one record in database $([ID]_*, [B]_*)$ as reference input, and perform user authentication as specified in Figure 6. Eventually, CP accept RU if the final results is “1”; Otherwise, CP outputs “ ⊥”.

![Figure 5. AKeyTrans Protocol under Public Key pk-*](image)

**Figure 6.** Authentication with corresponding sub-computational protocols

5 Security Analysis

Due to the page limit, the detailed security proof and the subsequent proof are deferred to the full version of this work.

**Theorem 2.** The proposed PriBioAuth framework has biometrics privacy if the underlying DT-PKC is semantically (IND-CPA) secure.

**Theorem 3.** The proposed PriBioAuth framework has user privacy if the underlying DT-PKC is IK-CPA secure.

6 Evaluations

In this section, we present both complexity analysis and performance analysis on proposed solutions.

6.1 Complexity Analysis

We first present a comprehensive complexity analysis between the most relevant work [11], our basic construction and PriBioAuth construction in terms of storage costs and computational costs (see Figure 7).
Storage cost. Let $L_{ID}$ denote the length of identity; $L_B$ denote the length of biometrics $B$; $L_{Z_{x'}}$ denote the length of element in $\mathbb{Z}_{x'}$; $L_{Z_{x''}}$ denote the length of element in $\mathbb{Z}_{x''}$. The standard length of encryption is $2 \cdot L_{x''}$, and we denote the storage cost of $N$-dimensional biometrics is $2 \cdot N \cdot L_{Z_{x''}}$. In [11], server $C$ stores all (uploaded) users’ ciphertext $(2n \cdot N \cdot L_{Z_{x''}})$, respective public keys $(n \cdot L_{Z_{x'}})$ and the sum of public keys $L_{Z_{x'}}$. Similarly, another server $S$ stores master secret key $(2 \cdot L_{x''})$, respective public keys and the sum of public keys.

In our basic construction, CP stores all registered users’ ciphertext $(2n \cdot N \cdot L_{Z_{x''}})$, identities $(n \cdot L_{ID})$, secret credential $(L_{Z_{x'}})$ and his own secret key $(L_{Z_{x'}})$. While CSP stores another secret credential and dummy public key $(L_{Z_{x'}})$. For the PriBioAuth construction, CP stores all registered users’ ciphertext $(2n \cdot N \cdot L_{Z_{x''}} + 2n \cdot L_{Z_{x'}})$ for biometrics and identity, plain identity, secret credential and his own secret key.

Computational cost. The time-complexity relies on the size of public parameter $N$, the number of records in database $n$, the number of addition, multiplication and exponentiation operations. Let $O(N)$ be a linear time algorithm, $O(N^\alpha)$ denotes a polynomial time algorithm for constant $\alpha$ and sets $\alpha = 3$ with respect to the exponentiation. Note that the Retrieval means that CP retrieves the outcome of authentication from CSP. We stress that the action of RU (e.g., a resource-limited device without storing any secret keys) is just Paillier encryption on ID and plain biometrics, while CP and CSP in cloud collaboratively run the corresponding sub-protocols without interacting with RU.

6.2 Performance Analysis

This experiment was run on virtual machines (3.6 GHz single-core processor and 6 GB RAM memory). The experiment assumes that user’s biometric information has been converted into the format needed, because the representation (depends on the feature extraction algorithms) of biometric data may vary. The running time and communication cost mainly depend on the bit length of $N$. Two extra factors are also needed to be considered, one is the vector dimension $N$, and the other one is the number of users $n$ when evaluating the proposed PriBioAuth framework. The comprehensive performance analysis is presented in Figure 8.

(1) SMT, SLT and SEQ sub-protocols. The SMT, SLT and SEQ sub-protocols are supportive materials which will be used in our proposed basic and PriBioAuth construction, we analyze its performance at Figure 8(a) and Figure 8(b) respectively. We observe that both running time and communication cost increase with respect to bit length $N$.

This is because if $N$ increases, then any basic operations (modular multiplication and exponential) increases. As a result, more bits need to be transmitted between CP and CSP.

(2) SEDC sub-protocol. The SEDC sub-protocol is essential for the efficiency of our proposed constructions. We analyze its performance at Figure 8(c) and Figure 8(d) respectively, and we observe that both the running time (see left coordinate) and communication cost (see right coordinate) increase with respect to bit length $N$ and vector dimension $N$. In particular, the vector dimension of biometrics here ranges from 100 to 500, and each vector is a 8-bit integer. Note that the efficiency of SEDC subprotocol is linear in the dimension of extracted feature vectors $N$.

(3) Basic construction. The running time and communication cost will increase with respect to vector dimension $N$ and bit length $N$. In Figure 8, we
show its performance at respective stages based on various size (100-500) of vector dimension. We observe that its performance is linear in the vector dimension \( N \), and the **Authentication** stage takes more running time and communication cost than **Registration**, because the corresponding SEDC subprotocol is required for biometrics matching between CP and CSP. In Figure 8(f), we show its full performance at respective stages based on bit length \( N \). We notice that its performance is also linear in the bit length \( N \) because some operations with respect to computational sub-protocols are increased.

(4) PriBioAuth construction. From Figure 8(g) to Figure 8(i), we observe that the running time and communication cost will increase with respect to vector dimension \( N \), number of users \( n \) (10-50) and bit length \( N \). We also observe that the PriBioAuth construction is linear in these factors. In particular, CP and CSP in **Authentication** stage perform more cryptographic operations than **Registration** stage, because the corresponding computational sub-protocols are required.

**Figure 8.** Evaluation findings of proposed constructions and their corresponding sub-protocols.
7 Conclusion

In this paper, we proposed a new framework of privacy-preserving biometric-based remote user authentication using homomorphic encryption. We also defined the new formal security models for biometrics privacy and user privacy, and proved the security of the proposed framework in the standard model. We leave the construction of biometric-based remote user authentication without going through the whole database as our future work, such that the time-complexity is not linear in the number of enrolled users.

Acknowledgements

The work is supported by the Singapore National Research Foundation under NCR Award Number NRF2014NCR-NCR001-012, and is partially sponsored by the National Science Foundation of China under Grant No. 61872264. It is also supported by AXA Research Fund.

References


Biographies

**Yangguang Tian** received his Ph.D degree in Applied Cryptography from University of Wollongong, Australia. He is a research fellow at the School of Information Systems, Singapore Management University. His research interests include user authentication, privacy protection, applied cryptography and network security.

**Yingjiu Li** is currently an Associate Professor in the School of Information Systems at Singapore Management University. His research interests include IoT Security and Privacy, Mobile and System Security, Applied Cryptography and Cloud Security, and Data Application Security and Privacy.

**Ximeng Liu** received the B.Sc. and Ph.D. degrees in electronic engineering from Xidian University, Xi’an, China. He is a research fellow at School of Information System, Singapore Management University, Singapore. His research interests include cloud security, applied cryptography and big data security.

**Robert H. Deng** is AXA Chair Professor of Cybersecurity and Director of the Secure Mobile Centre, School of Information Systems, Singapore Management University. His research interests are in the areas of data security and privacy, cloud security and Internet of Things security. He is an IEEE Fellow.

**Binanda Sengupta** received his B.E., M.S. and Ph.D. degrees in Computer Science from Jadavpur University, Indian Institute of Technology Kharagpur and Indian Statistical Institute, Kolkata, India, respectively. His broad research area includes applied cryptography, cloud computing, security and privacy. He is currently a postdoctoral research fellow at the School of Information Systems, Singapore Management University.