2017

Fuzzy Extractors for Biometric Identification

Nan LI
Surya Nepal, nl864@uowmail.edu.au

Fuchun Guo
University of Wollongong, fuchun@uow.edu.au

Yi Mu
University of Wollongong, ymu@uow.edu.au

Willy Susilo
University of Wollongong, wsusilo@uow.edu.au

Sanjay K. Nepal
University of Waterloo

Recommended Citation

LI, Nan; Guo, Fuchun; Mu, Yi; Susilo, Willy; and Nepal, Sanjay K., "Fuzzy Extractors for Biometric Identification" (2017). *Faculty of Engineering and Information Sciences - Papers: Part B*. 697.
https://ro.uow.edu.au/eispapers1/697

Fuzzy Extractors for Biometric Identification

Abstract

2017 IEEE. Fuzzy extractor provides key generation from biometrics and other noisy data. The generated key is seamlessly usable for any cryptographic applications because its information entropy is sufficient for security. Biometric authentication offers natural and passwordless user authentication in various systems where fuzzy extractors can be used for biometric information security. Typically, a biometric system operates in two modes: verification and identification. However, existing fuzzy extractors does not support efficient user identification. In this paper, we propose a succinct fuzzy extractor scheme which enables efficient biometric identification as well as verification that it satisfies the security requirements. We show that the proposed scheme can be easily used in both verification and identification modes. To the best of our knowledge, we propose the first fuzzy extractor based biometric identification protocol. The proposed protocol is able to identify a user with constant computational cost rather than linear-time computation required by other fuzzy extractor schemes. We also provide security analysis of proposed schemes to show their security levels. The implementation shows that the performance of proposed identification protocol is constant and it is close to that of verification protocols.

Disciplines

Engineering | Science and Technology Studies

Publication Details

LI, N., Guo, F., Mu, Y., Susilo, W. & Nepal, S. (2017). Fuzzy Extractors for Biometric Identification. 37th IEEE Internaitonal Conference on Distributed Computing Systems (ICDCS 2017) (pp. 667-677). United States: IEEE.

Fuzzy Extractors for Biometric Identification

Nan Li, Surya Nepal
DATA61
CSIRO
Marsfield, Australia
Email: {nan.li, surya.nepal}@data61.csiro.au

Fuchun Guo, Yi Mu and Willy Susilo
Centre for Computer and Information Security Research
School of Computing and Information Technology
University of Wollongong, Wollongong, Australia
Email: {fuchun, ymu, wsusilo}@uow.edu.au

***Abstract*—Fuzzy extractor provides key generation from biometrics and other noisy data. The generated key is seamlessly usable for any cryptographic applications because its information entropy is sufficient for security. Biometric authentication offers natural and passwordless user authentication in various systems where fuzzy extractors can be used for biometric information security. Typically, a biometric system operates in two modes: verification and identification. However, existing fuzzy extractors does not support efficient user identification. In this paper, we propose a succinct fuzzy extractor scheme which enables efficient biometric identification as well as verification that it satisfies the security requirements. We show that the proposed scheme can be easily used in both verification and identification modes. To the best of our knowledge, we propose the first fuzzy extractor based biometric identification protocol. The proposed protocol is able to identify a user with constant computational cost rather than linear-time computation required by other fuzzy extractor schemes. We also provide security analysis of proposed schemes to show their security levels. The implementation shows that the performance of proposed identification protocol is constant and it is close to that of verification protocols.**

I. Introduction

Biometric system has been widely used in various applications, such as mobile security, e-payment and identification check [1]. A biometric system is essentially based on the pattern recognition techniques which extract user's biometric features, and then compare with previously stored biometric templates. For example, national police office may collect citizen's biometric information like fingerprint and facial information for the security check. Presenting a finger on biometric device, it usually transmits the biometric information to an image, and apply recognition algorithms to extract features. Then, the verification process searches a matching template from the backend database and makes a decision.

There are some attractive features by using biometrics for user authentication. First, biometric is naturally with people. In practice, people are likely to have many different accounts for emails, shopping, education, etc. A problem is to respectively create and remember different secure passwords for these accounts. Because secure password (e.g., j4U-8x7AK5.#o0) is hard to remember [2], people can use their biometric instead of password to perform authentication. Second, biometrics can provide high level uniqueness and security. Although there exists attacks (e.g., [3]) against the security of biometric information, some biometrics (e.g., iris) still remain in high security [4].

On the other hand, using biometric systems leads some other security and privacy issues. First, biometric is usually hard to be modified, so that it is not revocable once it has been compromised. If an adversary steals user's biometric information, the user may lose the security forever. Biohasing is a tool which offers revocability, while it requires multi-factors for authentication [5]. Second, the accuracy of recognition significantly impacts the decision of biometric systems. For example, low accuracy biometric devices may provide inaccurate information that an illegal user is able to pass the authentication. Third, privacy sensitive users concern the security of stored biometric information on the authentication server. Certainly, no biometric (template) information should be stored in plaintext. It is needed to employ proper security protections on such data. Note that data encryption cannot prevent insider attacks from the authentication server.

Dodis et al. [6], [7] introduced the notion of *secure sketches* and *fuzzy extractors*. Taking biometric information Bio as input, a secure sketch scheme produces public information which does not reveal the input. The public information can be used to reproduce the original input Bio if a close biometric information Bio' is presented. Fuzzy extractors take noisy input, including biometric information, and generate a nearly uniform string (with some public helper data) to be used directly in cryptographic applications. For instance, the output string can be used as a private key in public key based cryptographic schemes. Also, the generated public helper data does not significantly leak the information of input, i.e biometric information. By giving the helper data and some biometric information close to the original input, the same string is recoverable. Therefore, fuzzy extractors are able to protect the security and privacy of user's biometric information.

A. Motivations and Contributions

Biometric systems typically operate in two modes: verification mode and identification mode [8]. In the verification mode, a user claims an identity and provides the biometric for verification. The authentication server retrieves the user's record from the database. User and authentication server runs biometric authentication protocol to check the validity of user. Particularly, in fuzzy extractor based biometric authentication protocols [9], [10], [11], the authentication server retrieves user's helper data and sends it to the user with some challenges. If the user can recover a valid secret string and responds the challenges, then the user is authenticated. In the identification mode, a user provides the biometric information and authentication server needs to decide whether the user is valid. This mode is close to the verification mode, but the difference is that the verification mode requires a claimed identity with biometric. Usually, the verification mode conducts 1-to-1 mapping, while identification performs 1-to-N comparison. Note that the terms "verification" and "authentication" are interchanging in this paper that they both indicate a protocol runs in the verification mode of biometric systems.

Biometric identification has been used in some scenarios such as criminal watching-list and identity management systems, because it is an intuitive method to solve the problems. Fuzzy extractors provide security protections for biometric systems, specifically the biometric information. In user verification, existing fuzzy extractor schemes are able to achieve a proper security level and they can be used in authentication protocols. However, fuzzy extractor based user identification could be inefficient. Consider that user's identity is unknown, the authentication server has to conduct exhaustive search to find the identity. That is, the fuzzy extractor and challenge-response verification will be performed by many times. Moreover, fuzzy extractor based protocols usually requires heavy computations like public key based cryptographic operations. Note that it is possible to apply lightweight symmetric key based cryptography with fuzzy extractors, while the communication cost (for helper data transmission) is still an issue. Therefore, fuzzy extractors have not been used for user identification.

We focus on the gap between fuzzy extractors and other approaches for biometric identification. Fuzzy extractor based protocols need valid helper data to recover the secret. Without providing user's identity (in identification mode), the server needs a computationally exhaustive search rather than simple lookup. Hence, the computational time is linear, i.e, $O(n)$. It significantly influences identification performance and restricts the use of fuzzy extractors in biometric systems. This paper proposes a new fuzzy extractor scheme which is succinct and it reduces the computational time to constant during identification. We show that the proposed scheme can be used in both verification and identification modes. To the best of our knowledge, we propose the first fuzzy extractor based biometric identification protocol. Also, we provide security analysis for the proposed schemes and protocols.

II. Preliminaries

In this section, we review some background of fuzzy extractors and describe notations used throughout this paper.

A. Fuzzy Extractor

Fuzzy extractor converts nonuniform data to uniformly random strings which can be used in cryptographic applications. A typical application of fuzzy extractor is to extract reproducible string from biometric information. The string then is considered as a secret to authenticate people. This section briefly reviews the fuzzy extractor introduced in [7].

1) Metric Space: A metric space is a set $\mathcal{M}$ with a distance function $\mathsf{dis} : \mathcal{M} \times \mathcal{M} \to \mathbb{R}^+ = [0, \infty)$ which obeys various properties, such as the triangle inequality $\mathsf{dis}(x, z) \leq \mathsf{dis}(x, y) + \mathsf{dis}(y, z)$ and symmetry $\mathsf{dis}(x, y) = \mathsf{dis}(y, x)$. In previous work, Hamming distance, set difference and edit distance have been used to construct fuzzy extractors.

2) Min-Entropy, Average Min-Entropy and Statistical Distance: The security of fuzzy extractors considers the entropy of output strings. An adversary who attempts to predict a random value is to guess the most likely value. The min-entropy $\mathbf{H}_\infty(A)$ of a random variable A is

$$\mathbf{H}_\infty(A) = -\log(\max_a \Pr[A = a]),$$

where $\max_a \Pr[A = a]$ is the predictability of A. For conditional distributions, we use the notion of average min-entropy. It is the logarithm of predictability of A if a value b of random variable B is given. The average min-entropy $\tilde{\mathbf{H}}_\infty(A|B)$ of A given B is defined as following.

$$\begin{aligned}\tilde{\mathbf{H}}_\infty(A|B) &= -\log\left(\mathbb{E}_{b \leftarrow B}\left[\max_a \Pr[A = a|B = b]\right]\right) \\ &= -\log\left(\mathbb{E}_{b \leftarrow B}\left[2^{-\mathbf{H}_\infty(A|B=b)}\right]\right).\end{aligned}$$

The statistical distance between two probability distributions A_1 and A_2 is

$$\mathbf{SD}(A_1, A_2) = \frac{1}{2}\sum_u \left|\Pr(A_1 = u) - \Pr(A_2 = u)\right|.$$

The security of a fuzzy extractor usually considers the statistical distance between a given distribution (from the extractor) and a uniform distribution U.

3) Secure Sketches and fuzzy Extractors: Secure sketch is a building block of fuzzy extractors. A secure sketch scheme takes as input noisy information ω, such as biometric information, then outputs a sketch s which is an auxiliary string. Note that secure sketches and fuzzy extractors are applicable to various noisy data other than biometric information. Secure sketch schemes normally use error correcting techniques to recover ω under s if and only if the given input ω' is close to ω. The sketch s can be published since it does not reveal much information about ω.

Definition 1. *A secure sketch consists of two randomized procedures* $(\mathsf{SS}, \mathsf{Rec})$ *with the following properties.*

- *The sketch* SS *on input* $\omega \in \mathcal{M}$ *outputs a sketch* $s \in \{0,1\}^*$.
- *The function* Rec *on input an element* $\omega' \in \mathcal{M}$ *and a sketch* $s \in \{0,1\}^*$ *outputs* ω *if* $\mathsf{dis}(\omega, \omega') \leq t$, *where* t *is a threshold.*

Fuzzy extractors extract some randomness from a noisy input $\omega \in \mathcal{M}$. Then, it can also be recovered from a given input ω' if ω and ω' are close. The difference is that fuzzy extractors return a uniform string, but secure sketch returns a non-uniform string.

Definition 2. *A fuzzy extractor consists of two randomized procedures* $(\mathsf{Gen}, \mathsf{Rep})$ *with the following properties.*

- *The generation function* Gen *on input* $x \in \mathcal{M}$ *outputs a string* $R \in \{0,1\}^\ell$ *and helper data* $P \in \{0,1\}^*$, *such that*
$$\mathsf{Gen}(x) \rightarrow (R, P).$$
- *The reproduction procedure* Rep *on input an element* $x' \in \mathcal{M}$ *and helper data* $P \in \{0,1\}^*$ *outputs* R, *such that*
$$\mathsf{Rep}(x', P) \rightarrow R \text{ if } \mathsf{dis}(x, x') \leq t.$$

Because secure sketch can reconstruct the original input from some given noisy data, it can be used to construct fuzzy extractor schemes. Generally speaking, a fuzzy extractor can be derived by using a secure sketch with a strong randomness extractor. We now review a generic fuzzy extractor construction from a secure sketch.

- Gen*: Let* SS *be a secure sketch and* Ext *be a strong extractor. Given an input* x, $\mathsf{Gen}(x; r_1, r_2) \rightarrow (P, R)$, *such that*
$$P = (\mathsf{SS}(x; r_1), r_2), \ R = \mathsf{Ext}(x; r_2).$$
- Rep*: Given an noisy input* x' *and* P, *recover the original input* $x = \mathsf{Rec}(x', \mathsf{SS}(x; r_1))$, *then compute* $R = \mathsf{Ext}(x; r_2)$.

B. L^p Norms

Let V be a vector space, norm is a function $|| \cdot ||$ which assigns a strictly positive real number to a vector on V. Given a distance function dis in metric space, if $|| \cdot ||$ is a norm on V, then $\mathsf{dis}(x - x') = ||x - x'||$ is a metric on V.

L^p norm has been widely used in addressing pattern recognition problems. It can be used to measure the distance between two pieces of biometric information. In Euclidean space, the associate norm is called L^2 norm ($p = 2$) which is a special case of L^p norm. Generally, the L^p norm is

$$||\mathbf{x}|| = \left(\sum_{i=1}^{n} |x_i|^p \right)^{\frac{1}{p}}.$$

If we consider the case $p \rightarrow \infty$, the norm $L^\infty = \max_i(|x_i|)$ is called maximum norm. Chebyshev distance is an example of using the maximum norm.

Definition 3. *Given two vectors* $\mathbf{x} = (x_1, \cdots, x_n)$ *and* $\mathbf{y} = (y_1, \cdots, y_n)$, *their Chebyshev distance is defined as*

$$\mathsf{dis}(\mathbf{x}, \mathbf{y}) = \max_i(|x_i - y_i|).$$

We say that $\mathbf{x}$ *and* $\mathbf{y}$ *are close if* $\mathsf{dis}(\mathbf{x}, \mathbf{y}) \leq t$, *where* $t \in \mathbb{R}^+$ *is a threshold.*

To describe the construction of proposed schemes and protocols, we give some notations in TABLE I.

Bio:	user's biometric information.
BioD:	a trusted biometric device.
DB:	a database stores public information like helper data.
$\mathbf{s}{\approx}\mathbf{s}'$:	two vectors $\mathbf{s}$ and $\mathbf{s}'$ are close according to some measurement.
	integer to a variable r.
KeyGen:	a key generation algorithm of digital signatures.
Sign:	a signing algorithm of digital signatures.
Verify:	a verification algorithm of distal signatures.
$\mathsf{dis}(\mathbf{x}, \mathbf{y})$:	a function returns distance between $\mathbf{x}$ and $\mathbf{y}$.

Table I
NOTATIONS.

III. BIOMETRIC IDENTIFICATION

Biometric information contains unique biological characteristics of individuals. It enables people identification and performs access control tasks. A user provides his biometric template during the registration phase called enrollment. When an identification is required, the user provides again his biometric information to a trusted biometric device and a new biometric template can be extracted. The user is identified if and only if the new biometric template is close to one of the stored templates. That is, the differences between two templates are smaller than a threshold. To compare two templates, an authentication server (AS) can search the stored biometric templates and check if the received template is close to any of the records.

User privacy and biometric template security are important problems. The server has to guarantee the security of the system, otherwise a user's privacy and biometric information security will be compromised. Unfortunately, it is hard to create a fully trusted system and a client may not trust remote servers.

Fuzzy extractor is a candidate to solve the security and privacy issues regarding biometric template. Biometric template is converted to some public helper data which does not need to be securely stored. Also, the stored data leaks negligible information about biometric template. Note that it differs from the biohashing which usually requires additional password or token to perform verification [5].

Biometric identification protocol consists of the following algorithm and protocols: System Setup ($\mathsf{SysSetup}$), User enrollment ($\mathsf{UserEnro}$) and Biometric Identification ($\mathsf{BioIden}$).

- SysSetup: Taking as input a security parameter λ, it generates system public parameters $params$.
- UserEnro: Taking as input a user's identity ID and user's biometric information Bio, it generates a pair of public and private keys (pk, sk) and some helper data P. The public information (ID, pk, P) is given to the authentication server, while the private key sk is discarded immediately. This protocol is played by a user, biometric device and authentication server.
- BioIden: Taking as input user's biometric information Bio, it outputs and user's identity if the user is identified, otherwise it outputs $\perp$.

We briefly describe enrollment and identification phases of fuzzy extractor based biometric identification. In enrollment phase, a user provides his biometric information Bio to a biometric device. The device extracts some helper data P and a string sk which is the user's private key. Then the user generates the corresponding public key pk and gives it to the server. Note that the server only stores (ID, P, pk) that the string sk is unknown. In identification phase, a user plays a challenge-response protocol (BioIden) with the authentication server. At the end of protocol run, user's identity is revealed or the identification failed.

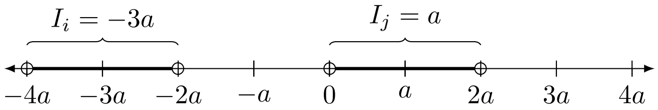
IV. Succinct Fuzzy Extractor

This section describes the concrete constructions of proposed secure sketch and fuzzy extractor schemes.

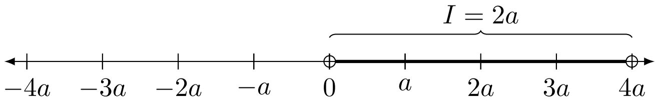
A. Number Line

Chebyshev distance is applied in our proposed fuzzy extractor to recover the original input from a noisy input. It is a new approach (for fuzzy extractors) different from the previous methods, such as the Hamming distance and setting distance. To use the Chebyshev distance, we firstly define a number line to express vector elements of biometric information. Each element of a vector is a point of the following number line.

Definition 4. *We define a number line L_a as follows*



where $a \in \mathbb{R}^+$ is a unit and 0 is the middle point of L_a. Let points $(\dots, -3a, -a, a, 3a, \dots)$ be odd points and $(\dots, -2a, 0, 2a, \dots)$ be even points. We define an interval on above number line as $(b, b+2a)$, where b is an even point. I is the identifier of an interval, where $I = \dots, -3a, -a, a, 3a, \dots$. Indeed, an interval is identified by its middle (even or odd) point. More generally, an interval of L_a is $(b, b+ka)$ that it contains k units, where $k \in \{2, 4, 6, \dots\}$. For example, the following number line L_a is divided by intervals where $k = 4$.



B. Secure Sketch Based on the Maximun Norm

A secure sketch scheme consists of two algorithms: Sketch (SS) and Reconstruction (Rec). We also show the system setup Setup for the proposed secure sketch.

- Setup: Let L_a be a number line as in *Definition 4*. We assume that all points on L_a are integers and L_a has exactly v intervals, where $v > 1$. Note that it is not necessary to have this assumption, but it simplifies the security analysis. For each interval $(b, b+ka)$, there are $ka - 1$ points such that $b+1, \dots, b+ka-2, b+ka-1$. The maximum acceptable Chebyshev distance (threshold) is $t < \frac{ka}{2}$.
- $\mathsf{SS}(\mathbf{x}) \to \mathbf{s}$: Assume that user's biometric information has been encoded into a vector $\mathbf{x} = (x_1, \dots, x_n)$, where x_i is a point of L_a.
  - For all $x_i \in \{x_1, \dots, x_n\}$, move x_i by s_i to the closest interval identifier I_i, such that
  $$I_i = x_i + s_i,$$
  where $|s_i| \leq \frac{ka}{2}$. It sets and returns the extracted sketch $\mathbf{s} = (s_1, \dots, s_n)$, which can be published.
  - Special case 1: If x_i is a point (e.g., $x_i = 0$) which does not in any interval, it flips a coin c and moves x_i to the closest left interval identifier if $c = 0$, otherwise moves to the right.
  - Special case 2: If x_i is the largest point, it is moved to either $x_i - \frac{ka}{2}$ or $-x_i + \frac{ka}{2}$ based on the coin c. It is similar if x_i is the smallest point. That is, L_a can be considered as a ring.
- $\mathsf{Rec}(\mathbf{y}, \mathbf{s}) \to \mathbf{z}$: Taking as input a user's biometric information which has been encoded into a vector $\mathbf{y} = (y_1, \dots, y_n)$, where y_i is a point of L_a, and a sketch $\mathbf{s}$, the recovery procedure $\mathsf{Rec}(\mathbf{y}, \mathbf{s})$ is as follows.
  - For all $y_i \in \{y_1, \dots, y_n\}$ and $s_i \in \{s_1, \dots, s_n\}$, compute
  $$y'_i = y_i + s_i.$$
  if $y'_i > \frac{kav}{2}$, then calculate $y'_i = y'_i - ka$.
  if $y'_i < -\frac{kav}{2}$, then calculate $y'_i = y'_i + ka$.
  - For all $y'_i \in \{y'_1, \dots, y'_n\}$, find the identifier I_i of interval which contains y'_i.
  If $|I_i - y'_i| > t$, it aborts the algorithm and returns $\perp$, otherwise, it computes, for $i = 1, \dots, n$,
  $$z_i = I_i - s_i.$$
  Then, it returns the vector $\mathbf{z} = (z_1, \dots, z_n)$.

We now show the correctness of above secure sketch scheme.

Theorem 1 (Correctness)**.** *The proposed secure sketch scheme recovers the vector* $\mathbf{y}$ *to* $\mathbf{x}$ *by given a sketch* $\mathbf{s}$ *if and only if* $\mathsf{SS}(\mathbf{x}) \to \mathbf{s}$ *and* $\mathsf{dis}(\mathbf{x}, \mathbf{y}) \le t$*, where* t *is the maximum acceptable Chebyshev distance.*

Proof: Given two n-dimensional vectors $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$, such that x_i, y_i are the points of the number line L_a, t is defined as in Setup. Let $\mathbf{x}$ be the input vector of SS and $\mathbf{y}$ be the input of Rec. $|s_i| \le \frac{ka}{2}$ is the movement from x_i to its interval identifier I_i. If $\mathsf{dis}(\mathbf{x}, \mathbf{y}) \le t$, for all $x_i \in \mathbf{x}$, $y_i \in \mathbf{y}$, $|x_i - y_i| \le t$ holds by the definition of Chebyshev distance. Then, we have the following equations.

$$\begin{cases} I_i = x_i + s_i, \\ y'_i = y_i + s_i, \\ I_i - t \le y'_i \le I_i + t. \end{cases}$$

Obviously, the point y'_i is in the interval $\left(I_i - \frac{ka}{2}, I_i + \frac{ka}{2}\right)$ because $t < \frac{ka}{2}$. According to the recovery procedure, the output point z_i can be computed as

$$z_i = I_i - s_i = x_i.$$

Therefore, the original input $\mathbf{x}$ can be recovered by using Rec.

Assume that $\mathsf{dis}(\mathbf{x}, \mathbf{y}) > t$, for at least one pair (x_i, y_i), we have the following inequalities.

$$y_i + s_i < x_i + s_i - t,$$

or

$$y_i + s_i > x_i + s_i + t.$$

The computed point $y'_i \notin [I_i - t, I_i + t]$ and $z_i \ne x_i$, so that $\mathbf{x}$ cannot be recovered by given $\mathbf{y}$ if the distance between $\mathbf{x}$ and $\mathbf{y}$ exceeds the threshold t. Thus, if and only if $\mathsf{dis}(\mathbf{x}, \mathbf{y}) \le t$ and $\mathsf{SS}(\mathbf{x}) \to \mathbf{s}$, $\mathbf{x}$ can be recovered by using $\mathbf{y}$. ■

C. Fuzzy Extractor Scheme

A secure sketch scheme can be converted to a fuzzy extractor scheme by using the generic construction introduced in Section II. This paper adopts a variant generic construction because of some weaknesses.

Boyen et al. [10] pointed out a problem of normal secure sketches and fuzzy extractors. Public helper data is stored and transmitted without any security protection mechanisms. An active adversary can modify the helper data and no security guarantees are provided in this case. Hence, they introduced a concept of *robust secure sketch* that such schemes can detect the modification of helper data during the recovery procedure. A generic construction of robust secure sketch (in the random oracle model) is also provided for any secure sketch which satisfies certain technical properties. We now review the proposed generic construction in [10].

- Setup: Let $H : \{0,1\}^* \to \{0,1\}^l$ be a hash function. The setup procedure is the same as a given well-formed secure sketch $(\mathsf{SS}', \mathsf{Rec}')$.
- $\mathsf{SS}(x) \to s$: Take as input some biometric information x and the sketching algorithm SS', it computes as follows.
  1) $\mathsf{SS}'(x) \to s'$,
  2) $h = H(x, s')$,
  3) return $s = (s', h)$.
- $\mathsf{Rec}(y, s) \to x$: Take as input some biometric information y, the helper data s and the recovery algorithm Rec', it computes as follows.
  1) $\mathsf{Rec}'(y, s') \to x'$,
  2) if $x' = \perp$, return $\perp$,
  3) if $H(x', s') \ne h$, return $\perp$, otherwise, return $x = x'$.

Note that parameters x, y, s and s' are vectors.

We also applies the above generic construction to convert the proposed secure sketch scheme.

- Setup: Let L_a be a number line as in *Definition 4*. $H : \{0,1\}^* \to \{0,1\}^l$ is a collision-resistant cryptographic hash function. L_a is assumed to contain exactly v intervals. The range of the number line L_a is $[-\frac{kav}{2}, \frac{kav}{2}]$, where $k = 2, 4, 6, \ldots$. The threshold is t, such that $t < \frac{ka}{2}$.
- SS: Take as input some biometric information $\mathbf{x}$, it runs the sketch algorithm of Section IV-B and gains the output a vector $(s_1, \ldots, s_n)$. Then, it computes $h = H(\mathbf{x}, s_1, \ldots, s_n)$ and returns $\mathbf{s} = (s_1, \ldots, s_n, h)$.
- Rec: Take as input some biometric information $\mathbf{y}$ and a sketch s, such that $s = (s_1, \ldots, s_n, h)$, it runs the reconstruction Rec in Section IV-B and returns $\mathbf{z}$ if $h = H(\mathbf{z}, s_1, \ldots, s_n)$.

Now, we can obtain a fuzzy extractor scheme. The proposed fuzzy extractor consists of three algorithms: System Setup (Setup), Generation procedure (Gen) and Reproduction Procedure (Rep).

- Setup: Let Ext be a strong extractor. It runs the same system setup of robust secure sketch.
- Gen: Given biometric information $\mathbf{x}$, it selects an ℓ-bit random string $r \in \{0,1\}^\ell$ and runs SS algorithm of robust secure sketch and obtains $\mathbf{s}$. Then, it computes $R = Ext(\mathbf{x}, r)$ and sets the public helper data $P = (\mathbf{s}, r)$. The algorithm returns (R, P).
- Rep: Given biometric information $\mathbf{y}$ and helper data P, the algorithm runs Rec algorithm of robust secure sketch and obtains the output $\mathbf{z}$. Then, it reproduces the string R by computing $E(\mathbf{z}, r)$.

The correctness of proposed fuzzy extractor scheme holds because the underlying secure sketch is correct. We analyze the security of proposed secure sketch and fuzzy extractor in Section VI-A.

V. Proposed Biometric Identification Protocol

In this section, we show that the proposed fuzzy extractor can be used to construct an efficient biometric identification

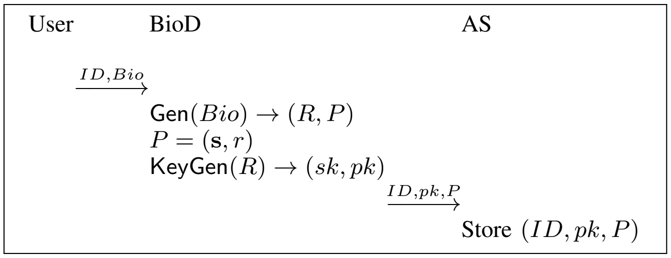


Figure 1. User enrollment.

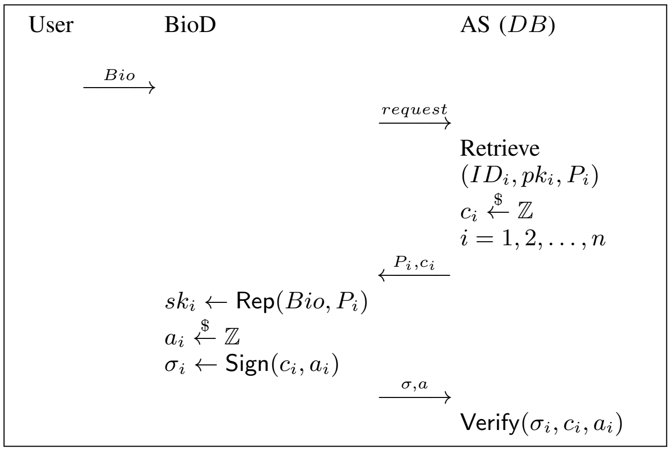

User
BioD
AS (DB)
$\xrightarrow{Bio}$
$\xrightarrow{request}$
Retrieve
(ID_i, pk_i, P_i)
$c_i \xleftarrow{\$} \mathbb{Z}$
$i = 1, 2, \ldots, n$
$\xleftarrow{P_i, c_i}$
$sk_i \leftarrow \mathsf{Rep}(Bio, P_i)$
$a_i \xleftarrow{\$} \mathbb{Z}$
$\sigma_i \leftarrow \mathsf{Sign}(c_i, a_i)$
$\xrightarrow{\sigma,a}$
$\mathsf{Verify}(\sigma_i, c_i, a_i)$

Figure 2. Fuzzy extractor based biometric identification protocol in normal approach.

protocol. Typically, a fuzzy extractor based biometric identification protocol (e.g., Fig. 2) may need to be performed $O(n)$ times for identification. For example, an authentication server has to exhaustively test records of helper data, because it does not know which record is with respect to the user. This is also the major difference from the biometric verification where an identifier like user's identity is provided. Note that $request$ is to retrieve helper data without sacrificing user's biometric information. However, we apply the proposed fuzzy extractor and secure sketch in a slightly different manner. To identify a user, the proposed protocol firstly compares the received sketch with the records in database. It avoids to conduct heavy computations (for public key cryptography) of the protocol. In our protocol, the computational time for user identification is constant. We now provide the details of our protocol (Fig. 3) as follows.

- $\mathsf{SysSetup}$: The authentication server chooses a security parameter λ and generates a number line L_a and the maximum acceptable Chebyshev distance t. It chooses a collision-resistant cryptographic hash function $H : \{0,1\}^* \rightarrow \{0,1\}^l$ and a strong extractor Ext. Then, it sets $params = (L_a, t, H, Ext)$ and publishes $params$.
- $\mathsf{UserEnro}$: To register a user (Fig. 1), this protocol runs as follows among the user, biometric device and authentication server.

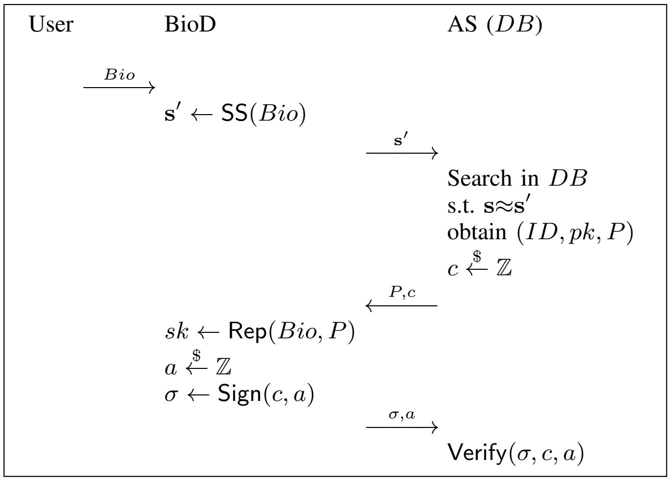

User
BioD
AS (DB)
$\xrightarrow{Bio}$
$\mathbf{s}' \leftarrow \mathsf{SS}(Bio)$
$\xrightarrow{\mathbf{s}'}$
Search in DB
s.t. $\mathbf{s} \approx \mathbf{s}'$
obtain (ID, pk, P)
$c \xleftarrow{\$} \mathbb{Z}$
$\xleftarrow{P,c}$
$sk \leftarrow \mathsf{Rep}(Bio, P)$
$a \xleftarrow{\$} \mathbb{Z}$
$\sigma \leftarrow \mathsf{Sign}(c, a)$
$\xrightarrow{\sigma,a}$
$\mathsf{Verify}(\sigma, c, a)$

Figure 3. Proposed biometric identification protocol.

User→BioD: The user presents its identity ID and biometric information Bio to the biometric device. Note that the biometric device is connected with other devices which can take input for identity information and it communicates with the authentication server.
BioD→AS: Upon receiving (ID, Bio), BioD runs the algorithm Gen and outputs (R, P), where R is a secret string and $P = (\mathbf{s}, r)$ is helper data. Let sk be the private key of a signature scheme. It derives the corresponding public key pk by running the KeyGen. BioD sends (ID, pk, P) to AS and erases (ID, Bio, sk) immediately.
AS: It inserts the record (ID, pk, P) into the DB.

- $\mathsf{BioIden}$: To identify a user, biometric device, user and authentication server interact as follows.
User→BioD: The user present its biometric information Bio to BioD.
BioD→AS: Upon receiving a user's biometric information Bio, BioD runs the algorithm SS and outputs a sketch $\mathbf{s}'$. It sends $\mathbf{s}'$ to AS.
AS→BioD: Upon receiving the sketch $\mathbf{s}'$, AS searches $\mathbf{s}$ from the database, such that for all $s_i \in \mathbf{s}$, $s'_i \in \mathbf{s}'$, it satisfies any of the following conditions.

$$s_i > 0, s'_i > 0 : |s_i - s'_i| \in [0, t] \quad (1)$$

$$s_i \leq 0, s'_i \leq 0 : |s_i - s'_i| \in [0, t] \quad (2)$$

$$s_i > 0, s'_i \leq 0 : |s_i - s'_i - ka| \notin (t, ka - t) \quad (3)$$

$$s_i \leq 0, s'_i > 0 : |s_i - s'_i + ka| \notin (t, ka - t) \quad (4)$$

Note that the above computations can be avoided by performing some pre-computations, i.e, the server only needs to check whether s'_i is in the specific range. Then, AS retrieves the record (ID, pk, P) and randomly chooses a challenge $c \in \mathbb{Z}$. AS sends (P, c) to BioD.

If there does not exist a valid record, the identification failed and outputs $\perp$.

BioD→AS: Upon receiving (P, c), BioD runs the algorithm Rep to recover the private key sk and chooses a random nonce $a \in \mathbb{Z}_p$. It generates a signature σ of the message (c, a) by using the signature algorithm Sign. BioD responds (σ, a) to AS.

AS: Upon receiving the response (σ, a), AS verifies the response by running the signature verification algorithm Verify. If the σ is valid, the identification succeeds, otherwise it is failed.

During the identification process, AS needs to search in the database for the corresponding helper data. We show that the computational time of this process is constant. In other words, AS does not need to exhaustively conduct "compute-then-compare" mode.

Theorem 2 (Correctness of Conditions)**.** *Given two biometric information (Bio, Bio'), a number line L_a with parameters (v, k, a) and two sketches $(s_i, s_i') \in (\mathsf{s}, \mathsf{s}')$, where $\mathsf{SS}(Bio) \rightarrow \mathsf{s}$ and $\mathsf{SS}(Bio') \rightarrow \mathsf{s}$, (s_i, s_i') satisfies one of the conditions $(1) - (4)$ if Bio and Bio' are close.*

Proof: Since Bio and Bio' are close, for all pairs of $(x_i, x_i') \in (Bio, Bio')$, $|x_i - x_i'| \leq t$. Clearly, if points x_i and x_i' move towards the same direction, the movement s_i and s_i' are respectively the distances to the same interval identifier I_i. Based on the sketch algorithm, we have the equation

$$x_i + s_i = x_i' + s_i'.$$

So that we have,

- If $s_i > 0, s_i' > 0$, then $|s_i - s_i'| = |x_i' - x_i| \in [0, t]$.
- If $s_i \leq 0, s_i' \leq 0$, then $|s_i - s_i'| = |x_i' - x_i| \in [0, t]$.

Another case is that two points are moved to different direction during the sketch procedure.

In condition (3), since $s_i > 0, s_i' \leq 0$, then there are two cases.

- If x_i and x_i' are in the same interval, then they are moved to the same identifier I_i. Then,

  $$s_i - s_i' = |s_i - s_i'| = |x_i - x_i'| \in [0, t].$$

  So that

  $$|s_i - s_i' - ka| = ka - (s_i - s_i') \in [ka - t, ka].$$

- If x_i and x_i' are in different intervals I_i and I_i', respectively, we have $I_i = I_i' \pm ka$, that is,

  $$x_i + s_i = x_i' + s_i' \pm ka.$$

  If $x_i + s_i = x_i' + s_i' + ka$, we have $x_i' \leq x_i$. Then,

  $$\begin{aligned} |s_i - s_i' - ka| &= ka - (s_i - s_i') \\ &= ka - (x_i' - x_i + ka) \\ &= x_i - x_i'. \end{aligned}$$

  So that $|s_i - s_i' - ka| \in [0, t]$. Such range is also $[0, t]$, if $x_i + s_i = x_i' + s_i' - ka$. Because the range of s_i and s_i' are $(0, \frac{ka}{2}]$ and $[-\frac{ka}{2}, 0]$, the value $|s_i - s_i' - ka|$ is in $[0, ka]$. Therefore, $|s_i - s_i' - ka| \notin (t, ka - t)$ and condition (3) holds.

If $s_i \leq 0, s_i' > 0$, the proof is similar to the above. We also have that $|s_i - s_i' + ka| \notin (t, ka - t)$ and condition (4) holds.

Finally, if two biometric information Bio and Bio' are close, the corresponding sketch elements satisfy one of conditions $(1) - (4)$. ■

It is important to notice that two sketches which satisfy these conditions does not result in two close biometric information. Because the same movement s could be performed in different intervals. Two biometric points x_i and x_i' (in different intervals), such that $\mathsf{dis}(x_i, x_i') > a$, have a probability to obtain the sketches which satisfy one of conditions $(1) - (4)$. Now, we show this probability theoretically. Assume that a number line L_a has exactly v intervals. Given a biometric point, there are at most $2t+1$ valid close points, while the total number of "close" points are $(2t+1)v$. That means there are many false close points in other intervals. If we consider the distribution of biometric points is uniformly at random, each point is "close" to another point with the probability $\frac{(2t+1)v}{kav}$. Note that biometric information usually contains many points, say n. Let the event E be two pieces of biometric information output a false close. We have the probability of E

$$\begin{aligned} \Pr[E] &= \left(\frac{(2t+1)v}{kav}\right)^n - \left(\frac{2t+1}{kav}\right)^n \\ &= \frac{(2t+1)^n(v^n - 1)}{(kav)^n} \\ &< \left(\frac{(2t+1)v}{kav}\right)^n \\ &= \left(\frac{2t+1}{ka}\right)^n. \end{aligned}$$

Therefore, the probability $\Pr[E]$ is negligible and the sketch comparison can be used to find helper data.

VI. Security Models and Analysis

The security of biometric information is important in biometric authentication protocols. To analyze the security of the proposed schemes, we firstly describe the security models, including the capability of adversaries and aim of attacks, etc. Secondly, we will formally prove the security of proposed schemes. Theoretical (provably secure) analysis is important and it shows the security level of schemes. Note that system failure due to effects, such as denial-of-service and intrusion, are beyond the scope of this paper.

A. Security of Fuzzy Extractors

The proposed secure sketch is used as a building block of our fuzzy extractors. To show the security level of secure sketch scheme, we consider an adversary who aims to recover the input from a given sketch. That is, we assume that an adversary can obtain and manipulate sketches in some manners. Formally, the advantage of this adversary is captured by using information entropy.

Definition 5 ([7])**.** *A secure sketch is $(\mathcal{M}, m, \tilde{m}, t)$-secure if for any distribution W over metric space $\mathcal{M}$ with min-entropy m, an adversary has an advantage at most $2^{-\tilde{m}}$, where $\tilde{m} \leq \tilde{\mathbf{H}}_{\infty}(W|SS(W))$, to recover the value of W.*

The security of fuzzy extractors is considered based on the statistical indistinguishability of two distributions.

Definition 6 ([7])**.** *A fuzzy extractor is $(\mathcal{M}, m, \ell, t, \epsilon)$-secure if for any input distribution W over metric space $\mathcal{M}$ with min-entropy m, the output string (excludes helper data) is in distribution R, the statistical distance between R and uniform distribution U_ℓ is negligible even if the helper data P is given, that is $\mathbf{SD}((R, P), (U_\ell, P)) \leq \epsilon$, where ϵ is negligible.*

An adversary against fuzzy extractors is able to read all public information, especially the helper data of users. Then, the adversary aims to either recover the user's biometric information or reproduce the secret string which is an output of fuzzy extractor. It is clear that if an adversary is able to reconstruct biometric information, then the secret string is recoverable. Hence, we consider the adversary who aims to find the secret string. We now show the security analysis of our proposed secure sketch and fuzzy extractor scheme.

Theorem 3 (Security)**.** *The proposed* (SS, Rec) *is an $(L_a, n\log kav, n\log v, t)$-secure sketch, where v is the number of intervals of L_a. The entropy loss is $n\log ka$ and the storage is $n\log(ka+1)$. Both the sketch algorithm* SS *and recovery algorithm* Rec *run in polynomial time in n, k, a and v.*

Proof: To recover a point $x_i \in \mathbf{x}$, an adversary needs to obtain the interval index I_i of x_i and the corresponding movement s_i. Since the helper data s is known to the adversary, it is easy to recover a point if its interval is given. Assume that the distribution of encoded points is uniformly on L_a [12], the best strategy of an adversary is to guess the interval of x_i. Let the parameters of L_a be (k, v, a) Since that there are v intervals on the number line L_a, the number of points on L_a is kav, because $-\frac{kav}{2}$ is considered the same as the point $\frac{kav}{2}$. The min-entropy m of input information $\mathbf{x} = (x_1, \dots, x_n)$ is $n\log kav$. Given a helper data $s_i \in s$, the probability of recovering the corresponding x_i is $\Pr[X = x_i|S = s_i]$.

In the sketch algorithm, a point x_i, where x_i is not an even point, has the only way of movement. Recall that the number line L_a is in range $[-\frac{kav}{2}, \frac{kav}{2}]$. The movement s_i is deterministic if such x_i is given. For even points, they can be randomly moved to either the closest right or left interval identifier. Let (x_i, s_i) be a pair of input point and the corresponding movement. Then, we have

$$\Pr[S = s_i|X = x_i] = \begin{cases} 1 & \text{if } x_i \neq \frac{ka(2\alpha - v)}{2},\ \alpha = 0, \dots, v; \\ \frac{1}{2} & \text{otherwise.} \end{cases}$$

On the other hand, the value of x_i is probabilistic if s_i is given. The probability of choosing x_i is $\Pr[X = x_i] = \frac{1}{kav}$. To discuss the security of the secure sketch, it is needed to calculate the average min-entropy $\tilde{\mathbf{H}}_{\infty}(X|S)$ which is related to the maximum probability of $\Pr[X = x_i|S = s_i]$. There are two cases ($|s_i| = \frac{ka}{2}$; $|s_i| \neq \frac{ka}{2}$) for all s_i. We show the probabilities of both cases as follows.

Case 1 ($|s_i| \neq \frac{ka}{2}$):

$$\begin{aligned} & \max_{x_i} \Pr[X = x_i|S = s_i] \\ = & \frac{\Pr[S = s_i|X = x_i]\Pr[X = x_i]}{\sum_{j=1}^{n} \Pr[S = s_i|X = x_j]\Pr[X = x_j]} \\ = & \frac{1 \times \frac{1}{kav}}{\frac{1}{kav}(\Pr[S = s_i|X = x_1] + \cdots + \Pr[S = s_i|X = x_n]} \\ = & \frac{1}{\underbrace{(0 + 0 + \cdots + 0 + 1) + \cdots + (0 + 0 + \cdots + 0 + 1)}_{v}} \\ = & \frac{1}{v}. \end{aligned}$$

Case 2 ($|s_i| = \frac{ka}{2}$):

$$\begin{aligned} & \max_{x_i} \Pr[X = x_i|S = s_i] \\ = & \frac{\frac{1}{2} \times \frac{1}{kav}}{\frac{1}{kav}(\Pr[S = s_i|X = x_1] + \cdots + \Pr[S = s_i|X = x_n]} \\ = & \frac{1}{\underbrace{(0 + 0 + \cdots + 0 + \frac{1}{2}) + \cdots + (0 + 0 + \cdots + 0 + \frac{1}{2})}_{v}} \\ = & \frac{1}{v}. \end{aligned}$$

According to the above equations, the maximum probabilities of recovering x_i, when s_i is given, are the same for all $s_i \in [-\frac{ka}{2}, \frac{ka}{2}]$. Because the elements of vector $\mathbf{x} = (x_1, \dots, x_n)$ are independent, the average min-entropy of the proposed secure sketch can be computed as

$$\begin{aligned} \tilde{\mathbf{H}}_{\infty}(X|S) &= -n\log\left(\mathbb{E}_{s_i \leftarrow S}\left[\max_{x_i} \Pr[X = x_i|S = s_i]\right]\right) \\ &= n\log v. \end{aligned}$$

The entropy loss of our secure sketch is

$$m - \tilde{\mathbf{H}}(X|S) = n\log kav - n\log v = n\log ka.$$

The outputs size of the sketch $\mathbf{s}$ is $n\log(ka+1)$. ■

According to the analysis, the security of our secure sketch is related to the amount v of intervals and the number

n of points of vector $\mathbf{x}$. The security can be enhanced by increasing v and n.

The security of proposed robust secure sketch depends on the security of generic construction and the underlying secure sketch. Because the basic secure sketch scheme is proved secure, the proposed robust secure sketch is secure followed by the security analysis of [10].

Theorem 4. *The proposed fuzzy extractor scheme is secure if the proposed secure sketch is secure.*

Proof: The proof is obvious. We use the generic construction [7] to derive the fuzzy extractor scheme from the proposed secure sketch. The generic construction guarantees the security of such fuzzy extractors. Since that the secure sketch scheme is $(L_a, n \log kav, n \log v, t)$ secure, the obtained fuzzy extractor scheme is also secure. In other words, the output sting R is indistinguishable with a string R' which is generated uniformly at random. ■

B. Security of Biometric Identification Protocols

In this section, we describe the security model of fuzzy extractor based biometric identification protocols. Firstly, we show the capabilities of adversaries by giving particular access to resources.

- An adversary is able to eavesdrop communication channel between the authentication server and biometric device.
- An adversary is able to manipulate on interactive messages over the communication channel. For instance, a message can be modified, injected or deleted.
- An (insider) adversary is able to access public helper data stored on the authentication server.

In analysis, biometric device is assumed to be secure with temper-resistant protections. If a biometric device is compromised, then any user who operate on the machine is under attack. Also, user's biometric information is unknown to an adversary because it plays as a secret key like password in single-factor authentication systems. Note that if biometric is used with other factors, such as token, the assumption can be relaxed [13], [14].

An attacker against biometric identification protocols aims to be identified as the target user without its biometric information. Biometric identification is different from traditional credential based identification. False accept is likely to occur due to various reasons. For example, the accuracy of biometric extraction can significantly influence the identification result. In particular, face recognition could have difficulties to distinguish twins. However, these issues can be relieved by using multiple types of biometrics, such as fingerprint and iris.

Theorem 5. *The proposed biometric identification protocol is secure if the underlying secure sketch, fuzzy extractor and digital signature scheme are secure.*

Proof: The protocol is played by user, biometric device and the authentication server. Assume that user has registered by using the protocol $\mathsf{UserEnro}$. A record (ID, pk, P) is stored in the database. If the employed digital signature is secure, then the private key sk, which is the secret string of fuzzy extractor's output, is not recoverable because the proposed fuzzy extractor is secure. During the identification, biometric device runs SS algorithm and obtains a user's sketch $\mathbf{s}$. Since that the proposed secure sketch is proved secure, an adversary cannot reveal the information of user's biometric. Upon receiving a sketch $\mathbf{s}'$, the server identifies a user without knowing biometric information. By using a secure digital signature scheme, if and only if the user can recover the private key sk, it can generate a valid response. So that, the proposed biometric identification protocol is secure. ■

VII. Implementations and Performances

To evaluate the performance, we implement the proposed secure sketch scheme, fuzzy extractor and identification protocol. The implementation is conducted by using Python on Linux computer (CPU: Intel Core i5-5300U@2.3GHz; Memory: 2GB; Virtual Machine). The conducted performance test aims to show the (speed) performance comparison between the proposed protocol and the normal approach (e.g., Fig. 2). The test assumes that user biometric data has been converted into the format needed, because the data conversion is exactly the same in two protocols. Note that both the proposed protocol and the normal approach use the same format of data as input. Also, the representation (depends on feature extraction algorithms) of biometric data could be vary. Without loss of generality, we use simulated data which is independent from any type of biometric. It is clear to show the speed difference between two protocols.

Table II introduces the parameters used in the implementation. The number line L_a consists of three parameters a, k and v. For an interval, there are at least 2 units, that is $k = 2$. However, this setting cannot achieve constant identification in the protocol. According to the probability of false close biometric information, the value of k should be $k \in \{4, 6, \ldots\}$. The maximum acceptable Chebyshev distance t (the threshold) is set to a for the simplicity. The implementation tested different size of input. The dimension n of input data is selected from $1,000$ to $31,000$ and representation range of an element is $[-100000, 100000]$. The result shows that dimensions have negligible impact to the protocol performance. Assume that biometric features have been extracted. In verification mode, one protocol execution for user verification needs 99 milliseconds ($n = 5000$). Note that feature extraction time could vary according to different biometrics. Also, a fuzzy extractor is hard to handle all types of biometrics [15]. The proposed protocol can be used with some biometrics such as face and fingerprint.

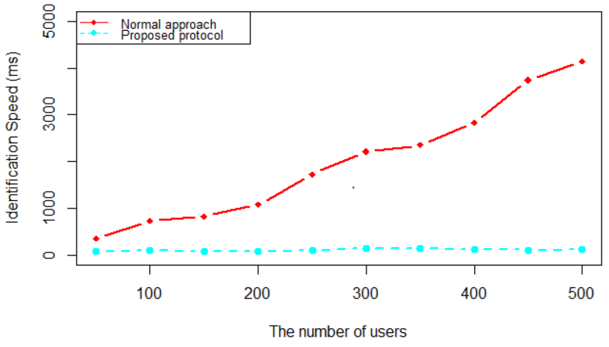


Figure 4. Speed of identification protocol.

In identification mode, the identification speed is significantly improved. Fig. 4 shows the comparison between the proposed protocol and normal method of fuzzy extractor based identification protocols. In the experiment, we assume that the database of helper data has been downloaded, so that the network transmission time is omitted. Because the proposed protocol uses the secure sketch scheme to identify helper data, only one digital signature computation is needed. It reduces the computational time to nearly constant. We obtains that the identification time is around 110 milliseconds which is close to the speed in verification mode. Therefore, the proposed identification protocol offers efficient user identification.

Parameter	Value
a	100
k	4
v	500
n	1000 - $31,000$
t	100
Rep. Range	$[-100000, 100000]$
$\tilde{m}$	$\approx 44,829$ bits ($n = 5,000$)
Storage	$\approx 45,000$ bits ($n = 5,000$)
Random Extractor	SHA256
Signature scheme	Digital Signature Algorithm (DSA)

Table II
IMPLEMENTATION PARAMETERS OF OUR PROTOCOL.

VIII. Related Work

Juels and Wattenberg [16] introduced a new type of cryptographic primitive called *fuzzy commitment scheme*. It combines both the techniques of cryptography and error correcting codes. The proposed commitment scheme is based on the Hamming distance and it is a basis of later code-offset sketch schemes. Juels and Sudan proposed a fuzzy vault scheme in [17], which is using the set difference to measure given biometric information. The fuzzy vault scheme randomly creates a secret k degree polynomial $p(x)$ during the sketch generation procedure. Given valid biometric information, a user can regenerate the polynomial. Assume that the amount of chaff points is large enough, an adversary has a negligible probability to recover the polynomial by given the set of chaff points.

Boyen [9], [18] showed that many constructions of fuzzy extractors are not secure against specific attacks. If a user has multiple sketches from the same sketch scheme, his biometric information can be leaked. The corresponding issue is called *reusability* of fuzzy extractors. The notion of *adaptive chosen perturbation attacks* was introduced and both of outsider and insider attackers were discussed. Then, a generic fuzzy sketch based on permutation groups was proposed. Note that the construction is provably secure in the random oracle model.

Apart from the above attacks, another type of active attack was considered. Boyen et al. [10] pointed out that the security of secure sketches and fuzzy extractors cannot be guaranteed if the public helper data was modified. So that they introduced a new concept called *robust sketches*. It can detect the modified auxiliary data and abort the algorithm immediately. They provided a generic conversion from a secure sketch, which satisfies certain technical properties, to a robust sketch in the random oracle model. The first construction of robust fuzzy extractors secure in the standard model was proposed by Dodis et al. [19], [20].

In terms of metric spaces, there are two directions to construct fuzzy extractors. The mentioned references above focused on discrete metric spaces, meanwhile some other work utilized continuous metric spaces. Linnartz and Tuyls [21] proposed a biometric authentication system which considered the continuous space $\mathbb{R}^n$. The security analysis of secure sketches which are based on continuous domains is different from secure sketches use discrete domains. Li et al. [22] studied the problems of secure sketches related to continuous domains and introduced a new concept of *relative entropy loss*.

IX. Conclusion

Fuzzy extractors protect the security of biometric information. In verification mode, user claims its identity and retrieves the helper data to recover the secret. However, a gap between fuzzy extractor based biometric system and other approaches is that there is no construction for efficient biometric identification. In scenarios where user's identity is unknown, the computational time for identification is $O(n)$. We proposed a set of schemes, including secure sketch, fuzzy extractor and an identification protocol. The experiment shows that our fuzzy extractor based identification protocol runs in around 110 ms which is close to the speed (99 ms) of verification.

The security of proposed protocol underlies on the security of secure sketch and fuzzy extractor scheme. We theoretically analyzed the security level of these schemes and showed that they are provably secure.

REFERENCES

[1] A. K. Jain, K. Nandakumar, and A. Ross, "50 years of biometric research: Accomplishments, challenges, and opportunities," *Pattern Recognition Letters*, vol. 79, pp. 80–105, 2016.

[2] U. Uludag, S. Pankanti, S. Prabhakar, and A. K. Jain, "Biometric cryptosystems: Issues and challenges," *Proceedings of the IEEE*, vol. 92, no. 6, pp. 948–960, 2004.

[3] U. Uludag and A. K. Jain, "Attacks on biometric systems: a case study in fingerprints," *Proceedings of SPIE*, E. J. D. III and P. W. Wong, Eds., vol. 5306. SPIE, 2004, pp. 622–633.

[4] C. S. Chin, A. T. B. Jin, and D. N. C. Ling, "High security iris verification system based on random secret integration," *Computer Vision and Image Understanding*, vol. 102, no. 2, pp. 169–177, 2006.

[5] A. K. Jain, K. Nandakumar, and A. Nagar, "Biometric template security," *EURASIP J. Adv. Sig. Proc.*, vol. 2008, 2008.

[6] Y. Dodis, L. Reyzin, and A. D. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," *Proceedings of EUROCRYPT 2004*, ser. LNCS, C. Cachin and J. Camenisch, Eds., vol. 3027. Springer, 2004, pp. 523–540.

[7] Y. Dodis, R. Ostrovsky, L. Reyzin, and A. D. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," *SIAM J. Comput.*, vol. 38, no. 1, pp. 97–139, 2008.

[8] S. Prabhakar, S. Pankanti, and A. K. Jain, "Biometric recognition: Security and privacy concerns," *IEEE Security & Privacy*, vol. 1, no. 2, pp. 33–42, 2003.

[9] X. Boyen, "Reusable cryptographic fuzzy extractors," in *Proceedings CCS 2004*, V. Atluri, B. Pfitzmann, and P. D. McDaniel, Eds. ACM, 2004, pp. 82–91.

[10] X. Boyen, Y. Dodis, J. Katz, R. Ostrovsky, and A. D. Smith, "Secure remote authentication using biometric data," in *Proceedings of EUROCRYPT 2005*, ser. LNCS, R. Cramer, Ed., vol. 3494. Springer, 2005, pp. 147–163.

[11] I. Buhan, J. Doumen, P. Hartel, and R. Veldhuis, "Secure ad-hoc pairing with biometrics: Safe," in *First International Workshop on Security for Spontaneous Interaction*. Innsbruck, Austria: Ubicomp 2007 Workshop Proceedings, 2007, pp. 450–456.

[12] L. Reyzin, "Entropy loss is maximal for uniform inputs," *CS Department, Boston University*, vol. Technical Report BUCS-TR-2007-011, 2007.

[13] A. T. B. Jin, D. N. C. Ling, and A. Goh, "Biohashing: two factor authentication featuring fingerprint data and tokenised random number," *Pattern Recognition*, vol. 37, no. 11, pp. 2245–2255, 2004.

[14] X. Huang, Y. Xiang, A. Chonka, J. Zhou, and R. H. Deng, "A generic framework for three-factor authentication: Preserving security and privacy in distributed systems," *IEEE Trans. Parallel Distrib. Syst.*, vol. 22, no. 8, pp. 1390–1397, 2011.

[15] B. Fuller, L. Reyzin, and A. D. Smith, "When are fuzzy extractors possible?" in *Proceedings of ASIACRYPT 2016*, ser. LNCS, J. H. Cheon and T. Takagi, Eds., vol. 10031, 2016, pp. 277–306.

[16] A. Juels and M. Wattenberg, "A fuzzy commitment scheme," in *Proceedings of CCS '99*, J. Motiwalla and G. Tsudik, Eds. ACM, 1999, pp. 28–36.

[17] A. Juels and M. Sudan, "A fuzzy vault scheme," *Des. Codes Cryptography*, vol. 38, no. 2, pp. 237–257, 2006.

[18] R. Canetti, B. Fuller, O. Paneth, L. Reyzin, and A. D. Smith, "Reusable fuzzy extractors for low-entropy distributions," in *Proceedings of EUROCRYPT 2016*, ser. LNCS, M. Fischlin and J. Coron, Eds., vol. 9665. Springer, 2016, pp. 117–146.

[19] Y. Dodis, J. Katz, L. Reyzin, and A. D. Smith, "Robust fuzzy extractors and authenticated key agreement from close secrets," in *Proceedings of CRYPTO 2006*, ser. LNCS, C. Dwork, Ed., vol. 4117. Springer, 2006, pp. 232–250.

[20] Y. Dodis, B. Kanukurthi, J. Katz, L. Reyzin, and A. D. Smith, "Robust fuzzy extractors and authenticated key agreement from close secrets," *IEEE Trans. Information Theory*, vol. 58, no. 9, pp. 6207–6222, 2012.

[21] J. M. G. Linnartz and P. Tuyls, "New shielding functions to enhance privacy and prevent misuse of biometric templates," in *Proceedings of AVBPA 2003*, ser. LNCS, J. Kittler and M. S. Nixon, Eds., vol. 2688. Springer, 2003, pp. 393–402.

[22] Q. Li, Y. Sutcu, and N. D. Memon, "Secure sketch for biometric templates," in *Proceedings of ASIACRYPT 2006*, ser. LNCS, X. Lai and K. Chen, Eds., vol. 4284. Springer, 2006, pp. 99–113.