Josai Mathematical Monographs
vol. 7 (2014), pp. 55-65

A relation between irreversibility and unlinkability for biometric template protection algorithms

Manabu Inuma

Abstract. For biometric recognition systems, privacy protection of enrolled users' biometric information, which are called biometric templates, is a critical problem. Recently, various template protection algorithms have been proposed and many related previous works have discussed security notions to evaluate the protection performance of these protection algorithms. Irreversibility and unlinkability are important security notions discussed in many related previous works. In this paper, we prove that unlinkability is a stronger security notion than irreversibility.

1. Introduction

Biometrics is a technique which automatically recognizes an individual by using his/her physical or behavioral characteristics such as fingerprints, face, vein pattern, (on-line or off-line) handwriting, or gait. A biometric recognition system stores biometric features extracted from each user's biometric characteristic. The stored biometric features of each user is called a (*biometric*) *template*. During verification, it compares freshly extracted biometric features with stored biometric features and decides whether these two biometric feature sets originate in the same user or not. Biometric features extracted from a user's biometric characteristic are strongly linked to the user and almost unchangeable during his/her lifetime. Once biometric features of a user are leaked together with the user's identity, he/she will face a severe risk of identity theft. Moreover, biometric features often contain sensitive privacy information about the user. To solve these security problems, some traditional biometric authentication system utilizes a symmetric-key or public-key encryption scheme $(\mathrm{Enc}, \mathrm{Dec})$, where Enc and Dec are the encryption and decryption algorithms, respectively. During enrollment, the system encrypts each user's biometric features x into a cyphertext $\mathrm{Enc}(x)$ and stores it in some storage device, and, during verification, decrypts $\mathrm{Enc}(x)$ into the original biometric features $x = \mathrm{Dec}(\mathrm{Enc}(x))$ and compares x with freshly extracted biometric features x'. However, such a traditional system has the problem that the adversary who knows all algorithms and all keys utilized in the system can easily recover the original biometric features x from a cyphertext $\mathrm{Enc}(x)$, even if he does not present biometric features x' sufficiently close to x. For example, a malicious administrator of the biometric system might recover user's biometric features and abuse them.

A *biometric template protection* (*BTP*) *algorithm* (see Def. 2 in Sect. 4) is a primitive mechanism for constructing a system in which only the individual presenting biometric features sufficiently similar to the enrolled biometric features can recover the enrolled features from the protected template and can be successfully verified, or which completes the verification without revealing the enrolled features. Recently, many BTP algorithms (cf. [**2, 4, 8, 9, 12, 13, 16, 17, 18**]) have been proposed and various security notions (cf. [**1, 2, 4, 5, 6, 7, 14, 16**]) have been discussed in some related previous works. Two important security notions, irreversibility and unlinkability, are addressed in most related previous works. Inuma, Otsuka [**5**] show that unlinkability is a stronger notion than irreversibility, namely unlinkability implies irreversibility. However, they do not give the proof of this statement due to limitation of pages and will describe the proof in the full version paper. In this paper, as a preparation for the full version paper, we introduce simpler formalizations of irreversibility and unlinkability for BTP algorithms (see Def. 5 in Sect. 5 and Def. 7 in Sect. 6) and give the proof for the relationship between the security notions (see Theorem 9 in Sect. 7).

We introduce the *minimum entropy* (cf. [**3, 4**]) for a biometric recognition system, which is defined as the amount of information unavailable to the computationally unbounded adversary who attempts to guess biometric features which are decided, by the biometric recognition system, to match biometric features extracted from a randomly chosen user's biometric characteristic (see Def. 1 in Sect. 3). Our formalization of the security notions employs the minimum entropy as the security parameter. In the real world, since most existing biometric recognition systems do not have sufficiently large minimum entropy[1] and many large databases of biometric samples are available to the public, the adversary who obtains a protected template can recover the original biometric features by exhaustive database search. To compensate the insufficiency of biometric recognition systems, some previous works (cf. [**1, 5, 6, 15**]) require secrecy of protected templates. For example, the International Standard ISO/IEC 24745 [**6**] and Inuma, Otsuka [**5**] require that each protected template should be decomposed into two data, a *pseudonymous identifier* (*PI*) and an *auxiliary data* (*AD*), and these two data should be (logically or physically) separately stored. However, in this paper, we do not address insufficiency of the entropy for existing modalities and the separation of protected templates. We only focus on mathematically proving the relationship between irreversibility and unlinkability.

[1] Surprisingly, Une, Otsuka, Imai [**11**] report that the success probability of a strong presentation attack, which is called *wolf attack*, to a fingerprint recognition system [**13**] is larger than $2^{-0.8}$, namely the minimum entropy of the fingerprint recognition system is smaller than 0.8.

2. Notions and Preliminaries

For any random variable X on a (finite) set $\mathcal{M}$, the notation $x \leftarrow X$ denotes an event that x is chosen according to X. For any set T, the notation $t \leftarrow T$ denotes an event that t is chosen from the set T uniformly at random, namely T can be regarded as a random variable representing the uniform distribution on T. For any deterministic or randomized algorithm A on $\mathcal{M}$, let $A(X)$ be a random variable representing the distribution of outputs of A whose input $x \in \mathcal{M}$ is chosen according to X, namely $\Pr[a \leftarrow A(X)] = \Pr[x \leftarrow X,\ a \leftarrow A(x)]$. For any function $f : \mathcal{M} \to \mathbf{R}$, the notation $\mathop{\mathrm{E}}_{x \leftarrow X} f(x)$ denotes the expected value of f under the condition that x is chosen according to X, namely $\mathop{\mathrm{E}}_{x \leftarrow X} f(x) = \sum_{x \in \mathcal{M}} Pr[X = x] f(x)$.

3. Biometric systems and entropy

Let $\mathcal{U}$ be a finite set consisting of all users who have biometric characteristics utilized in BTP algorithms. Assume that each user $u \in \mathcal{U}$ has his/her own biometric characteristic b_u and therefore, in the following, we identify u with b_u and use the notation u instead of b_u, namely, the set $\mathcal{U}$ can be regarded as a set consisting of all individuals' biometric characteristics (e.g., a right index finger, a face and so on).

When a biometric characteristic $u \in \mathcal{U}$ is presented to the sensor, biometric samples (e.g., fingerprint images, face images and so on) are captured from the presented characteristic and a set of biometric features is extracted from the captured biometric samples. We assume that each set of biometric features is represented as an element x of a finite set $\mathcal{M}$ and called a *feature element.* Since two feature elements generated from a characteristic u are rarely identical, we can regard the above feature element extraction procedure EXT as a randomized algorithm which takes as input a biometric characteristic $u \in \mathcal{U}$ and returns a feature element $x \in \mathcal{M}$.

A *comparison algorithm* CMP is a deterministic algorithm which takes as input two feature elements $x, x' \in \mathcal{M}$ and returns "*match*" if, in some manner, x and x' are decided to be extracted from the same user's characteristic, and otherwise returns "*non-match*". We call the tuple $\Theta = (\mathcal{U}, \mathcal{M}, \mathrm{EXT}, \mathrm{CMP})$ a *biometric system.* In this paper, we assume that the adversary knows the whole information about Θ.

Here we introduce the *minimum entropy* (cf. [**3, 4**]) for a biometric system.

Definition 1 (Minimum entropy (cf. [**3, 4**])). *For any biometric system Θ, the* minimum entropy $H_\infty(\Theta)$ *of* Θ *is defined by*

$$H_\infty(\Theta) = \min_{y \in \mathcal{M}} \left(-\log_2 \mathop{\mathrm{E}}_{x \leftarrow \mathrm{EXT}(\mathcal{U})} \Pr\left[\mathrm{CMP}(x, y) = \text{“}match\text{”} \mid x \leftarrow \mathrm{EXT}(\mathcal{U})\right] \right) \quad (1)$$

$$= -\log_2 \max_{y \in \mathcal{M}} \left(\mathop{\mathrm{E}}_{x \leftarrow \mathrm{EXT}(\mathcal{U})} \Pr[\mathrm{CMP}(x,y) = \text{“}match\text{”} \mid x \leftarrow \mathrm{EXT}(\mathcal{U})] \right) .$$

Une, Otsuka, Imai [**11, 19**] propose a security measure for a biometric system, the *wolf attack probability*, which is defined as the maximum success probability of the *wolf adversaries* who attempt to impersonate a user by presenting a feature element which matches as many users' feature elements as possible, where the adversaries are allowed to take as input only information independent of the user to be impersonated. For any biometric system Θ, relaxing computational limitations of the wolf adversaries, we have $H_\infty(\Theta) = -\log_2 WAP_\Theta$, where WAP_Θ denotes the wolf attack probability of Θ.

4. Biometric template protection algorithms

In this section, we give an explicit formulation of *biometric template protection (BTP) algorithms.*

Definition 2 (BTP algorithms [**5**]). *A biometric template protection (BTP) algorithm* $\Pi = (\mathrm{Gen}, \mathrm{Prt}, \mathrm{Rcg})$ *associated with the biometric system* $\Theta = (\mathcal{U}, \mathcal{M}, \mathrm{EXT}, \mathrm{CMP})$ *of minimum entropy* k *consists of three polynomial-time (in* k*) algorithms. The* parameter generation algorithm Gen *takes as input the security parameter* k *and returns a set of some common parameters* $\mathfrak{p}$[2] *including the security parameter* k*. The* protection algorithm Prt *is a randomized algorithm which takes as input a feature element* $x \in \mathcal{M}$ *and the common parameters* $\mathfrak{p}$*, and returns a* protected template Ξ*. The* recognition algorithm Rcg *is a deterministic algorithm which takes as input a new feature element* $x' \in \mathcal{M}$*, a protected template* Ξ *and the common parameters* $\mathfrak{p}$*, and returns either “match” or “non-match”.*

Definition 3 (Correctness). *A BTP algorithm* $\Pi = (\mathrm{Gen}, \mathrm{Prt}, \mathrm{Rcg})$ *associated with a biometric system* $\Theta = (\mathcal{U}, \mathcal{M}, \mathrm{EXT}, \mathrm{CMP})$ *is said to be* correct *if, for any* $x, x' \in \mathcal{M}$*,* $\mathrm{CMP}(x,x') =$ *“match” implies* $\mathrm{Rcg}(\mathrm{Prt}(x), x') =$ *“match”.*

Definition 4 (Validity). *A BTP algorithm* $\Pi = (\mathrm{Gen}, \mathrm{Prt}, \mathrm{Rcg})$ *associated with a biometric system* $\Theta = (\mathcal{U}, \mathcal{M}, \mathrm{EXT}, \mathrm{CMP})$ *is said to be* valid *if, for any* $x, x' \in \mathcal{M}$*,* $\mathrm{Rcg}(\mathrm{Prt}(x), x') =$ *“match” implies* $\mathrm{CMP}(x,x') =$ *“match”.*

5. Irreversibility

Let $\Pi = (\mathrm{Gen}, \mathrm{Prt}, \mathrm{Rcg})$ be a BTP algorithm. When the adversary obtains a protected template, he might be able to recover a feature element close to the

[2] Here $\mathfrak{p}$ may be just the security parameter k or include some additional information such as setup information. For example, in a fuzzy commitment scheme [**9**], $\mathfrak{p}$ includes, in addition to k, the parity check matrix or the generator matrix of the employed linear error-correcting code and a hash function used to create public keys.

original feature element, create a forgery from the recovered feature element, and impersonate the user by presenting the forgery to the system. The security notion, *irreversibility*, requires that such recovering of feature elements from protected templates is hard, namely for any polynomial-time adversary given a protected template Ξ, it is computationally hard to recover a feature element $y \in \mathcal{M}$ such that $\mathrm{Rcg}(\Xi, y) =$"*match*". To formalize the security notion, irreversibility, we will define *Irreversibility* (*IRR*) *Game* between the challenger Ch and the adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$, where $\mathcal{A}_1$ is a probabilistic polynomial-time (ppt) algorithm which takes as input the common parameters $\mathfrak{p}$ and sends a state $\mathfrak{st}$ to another ppt algorithm $\mathcal{A}_2$, and $\mathcal{A}_2$ takes as input the state $\mathfrak{st}$ and a protected template Ξ, and attempts to guess a feature element satisfying $\mathrm{Rcg}(\Xi, y) =$"*match*".

The following formalization of IRR Game is almost the same as $\{\mathrm{PI}, \mathrm{AD}\}$-pseudo authorized leakage irreversibility game defined in [**5**] and is simpler than irreversibility game defined in [**14**] in which the adversary obtains two protected templates generated from two feature elements sufficiently close to each other and attempts to recover either of the feature elements.

Definition 5 (Irreversibility (IRR) Game (cf. [**5, 14**])).

Setup. *A BTP algorithm* $\Pi = (\mathrm{Gen}, \mathrm{Prt}, \mathrm{Rcg})$ *associated with the biometric system* $\Theta = (\mathcal{U}, \mathcal{M}, \mathrm{EXT}, \mathrm{CMP})$ *of minimum entropy k is set up.*

Step 1. *The challenger* Ch *inputs the security parameter k into* Gen*, receives the parameters $\mathfrak{p}$ output from* Gen*, and sends $\mathfrak{p}$ to the adversary $\mathcal{A}_1$.*

Step 2. *The adversary $\mathcal{A}_1$ receives $\mathfrak{p}$ and sends a state $\mathfrak{st}$ to $\mathcal{A}_2$.*

Step 3. *The challenger* Ch *chooses a biometric characteristic $u \in \mathcal{U}$ uniformly at random, obtains a protected template $\Xi \leftarrow \mathrm{Prt}(\mathrm{EXT}(u))$, and sends Ξ to the adversary $\mathcal{A}_2$.*

Step 4. *The adversary $\mathcal{A}_2$ receives the state $\mathfrak{st}$ and the protected template Ξ from $\mathcal{A}_1$ and* Ch*, respectively, and returns $y \in \mathcal{M}$.*

If $\mathrm{Rcg}(\Xi, y) =$ *"match", then the adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ wins.*

Here we define the advantage $\mathrm{Adv}_{\Pi,\mathcal{A}}^{\mathrm{IRR}}(k)$ of the adversary $\mathcal{A}$ in the above IRR Game by

$$\mathrm{Adv}_{\Pi,\mathcal{A}}^{\mathrm{IRR}}(k) = \Pr[\mathcal{A} \text{ wins in IRR Game}] - \frac{1}{2^k}$$

Definition 6 (Irreversibility (cf. [**5, 10**])). *Fix a positive function $\varepsilon(k)$ of k. We say a BTP algorithm Π is $\varepsilon(k)$-irreversible if* $\mathrm{Adv}_{\Pi,\mathcal{A}}^{\mathrm{IRR}}(k) < \varepsilon(k)$ *for any ppt adversary $\mathcal{A}$ in IRR Game.*

6. Unlinkability

If the adversary given two protected templates can determine whether the protected templates are generated from the same characteristic or not, the leakage of multiple protected templates together with access logs might cause many privacy problems. The security notion, *unlinkability*, requires that such a linking attack is hard, namely for any polynomial-time adversary given two protected template Ξ and Ξ', it is computationally hard to determine whether the protected templates originate from the same characteristic or not. To formalize the security notion, unlinkability, we will define *Unlinkability* (*UNLINK*) *Game* between the challenger Ch and the adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$.

Definition 7 (Unlinkability (UNLINK) Game (cf. [**5, 14**])).

Setup. *A BTP algorithm* $\Pi = (\mathrm{Gen}, \mathrm{Prt}, \mathrm{Rcg})$ *associated with the biometric system* $\Theta = (\mathcal{U}, \mathcal{M}, \mathrm{EXT}, \mathrm{CMP})$ *of minimum entropy k is set up.*

Step 1. *The challenger* Ch *inputs the security parameter k into* Gen*, receives the parameters* $\mathfrak{p}$ *output from* Gen*, and sends* $\mathfrak{p}$ *to the adversary* $\mathcal{A}_1$.

Step 2. *The adversary* $\mathcal{A}_1$ *receives* $\mathfrak{p}$ *and sends a state* $\mathfrak{st}$ *to* $\mathcal{A}_2$.

Step 3. *The challenger* Ch *chooses a biometric characteristic* $u \in \mathcal{U}$ *uniformly at random, obtains a feature element* $x \leftarrow \mathrm{EXT}(u)$*, and moreover obtains a protected template* $\Xi \leftarrow \mathrm{Prt}(x)$.
Then Ch *flips the random coin* $b \in \{0, 1\}$.
If $b = 0$*, then* Ch *obtains another protected template* $\Xi' \leftarrow \mathrm{Prt}(x)$ *by inputting the same feature element* x.
If $b = 1$*, then* Ch *again chooses a biometric characteristic* $v \in \mathcal{U}$ *uniformly at random and obtains a protected template* $\Xi' \leftarrow \mathrm{Prt}(\mathrm{EXT}(v))$.
The challenger sends (Ξ, Ξ') *to the adversary* $\mathcal{A}_2$.

Step 4. *The adversary* $\mathcal{A}_2$ *receives the state* $\mathfrak{st}$ *and* (Ξ, Ξ') *from* $\mathcal{A}_1$ *and* Ch*, respectively, and returns* $b' \in \{0, 1\}$ *as a guess of* b.

If $b' = b$*, then the adversary* $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ *wins.*

The adversary in the above UNLINK Game is weaker than that in $\{\mathrm{PI}, \mathrm{AD}\}$-unlink game defined in [**5**] and is stronger than that in indistinguishability game defined in [**14**]. In $\{\mathrm{PI}, \mathrm{AD}\}$-unlink game defined in [**5**], original feature elements are chosen by the adversary. In indistinguishability game defined in [**14**], if $b = 0$, the second protected template Ξ' is generated from a feature element x' sufficiently close to x.

The advantage $\mathrm{Adv}^{\mathrm{UNLINK}}_{\Pi,\mathcal{A}}(k)$ of the adversary $\mathcal{A}$ in the above UNLINK Game is defined by

$$\mathrm{Adv}^{\mathrm{UNLINK}}_{\Pi,\mathcal{A}}(k) = \left|2\Pr\left[\mathcal{A} \text{ wins in UNLINK Game}\right] - 1\right| \tag{2}$$

Definition 8 (Unlinkability). *Fix a positive function $\varepsilon(k)$ of k. We say that a BTP algorithm Π is $\varepsilon(k)$-unlinkable if $\mathrm{Adv}^{UNLINK}_{\Pi,\mathcal{A}}(k) < \varepsilon(k)$ for any ppt adversary $\mathcal{A}$ in UNLINK Game.*

7. Relation between Irreversibilty and Unlinkabilty

In this section, we show that unlinkability is a stronger notion than irreversibility for any correct and valid BTP algorithm.

Theorem 9. *Fix a positive function $\varepsilon(k)$ of k. For any correct and valid BTP algorithm Π associated with the biometric system Θ of minimum entropy k, if Π is $\varepsilon(k)$-unlinkable, then Π is $\varepsilon(k)$-irreversible.*

Proof. It is sufficient to show that if there exists an adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ in IRR Game satisfying $\mathrm{Adv}^{\mathrm{IRR}}_{\Pi,\mathcal{A}}(k) \geq \varepsilon(k)$, then there exists an adversary $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2)$ in UNLINK Game satisfying $\mathrm{Adv}^{\mathrm{UNLINK}}_{\Pi,\mathcal{B}}(k) \geq \varepsilon(k)$.

Define the adversary $\mathcal{B} = (\mathcal{B}_1, \mathcal{B}_2)$ by using the adversary $\mathcal{A} = (\mathcal{A}_1, \mathcal{A}_2)$ as follows.

The adversary $\mathcal{B}_1$ The adversary $\mathcal{B}_1$ receives $\mathfrak{p}$ from the challenger Ch, obtains a state $\mathfrak{st} \leftarrow \mathcal{A}_1(\mathfrak{p})$, and sends the state $\mathfrak{st}$ to the adversary $\mathcal{B}_2$.

The adversary $\mathcal{B}_2$ The adversary $\mathcal{B}_2$ receives the state $\mathfrak{st}$ and (Ξ, Ξ') from $\mathcal{B}_1$ and Ch, respectively. Then $\mathcal{B}_2$ obtains a feature element $y \leftarrow \mathcal{A}_2(\Xi, \mathfrak{st})$.
If $\mathrm{Rcg}(\Xi, y) =$"*match*" and $\mathrm{Rcg}(\Xi', y) =$"*match*", then $\mathcal{B}_2$ returns $b' = 0$.
If $\mathrm{Rcg}(\Xi, y) =$"*match*" and $\mathrm{Rcg}(\Xi', y) =$"*non-match*", then $\mathcal{B}_2$ returns $b' = 1$.
If $\mathrm{Rcg}(\Xi, y) =$"*non-match*", then $\mathcal{B}_2$ returns $b' \in \{0, 1\}$ uniformly at random.

When $b = 0$, there are the following two cases in which the adversary $\mathcal{B}$ correctly returns $b' = 0$.

Case 1. $\mathcal{A}_2$ guesses a feature element y satisfying $\mathrm{Rcg}(\Xi, y) =$"*match*".
In this case, from validity of Π, for the original feature element x of Ξ, $\mathrm{CMP}(x, y) =$"*match*". Then, from correctness of Π, $\mathrm{Rcg}(\Xi', y) =$"*match*" since Ξ' is also generated from x.

Case 2. $\mathcal{A}_2$ guesses a feature element y satisfying $\mathrm{Rcg}(\Xi, y) =$"*non-match*" and $\mathcal{B}_2$ chooses $b' = 0$ from $\{0, 1\}$ with probability $\dfrac{1}{2}$.

Therefore, we have

$$\begin{aligned}
&\Pr[\mathcal{B} \text{ returns } b'=0 \,|\, b=0]\\
=&\mathop{\mathrm{E}}_{\substack{\mathfrak{st}\leftarrow\mathcal{A}_1(\mathrm{Gen}(k))\\ \Xi\leftarrow\mathrm{Prt}(\mathrm{EXT}(\mathcal{U}))}}\left(\Pr\left[\begin{array}{c} y\leftarrow\mathcal{A}_2(\Xi,\mathfrak{st}),\\ \mathrm{Rcg}(\Xi,y)=\text{“}\textit{match}\text{”}\end{array}\middle|\begin{array}{c}\mathfrak{st}\leftarrow\mathcal{A}_1(\mathrm{Gen}(k))\\ \Xi\leftarrow\mathrm{Prt}(\mathrm{EXT}(\mathcal{U}))\end{array}\right]\right.\\
&\left.+\Pr\left[\begin{array}{c} y\leftarrow\mathcal{A}_2(\Xi,\mathfrak{st}),\\ \mathrm{Rcg}(\Xi,y)=\text{“}\textit{non-match}\text{”},\\ b'=0\leftarrow\{0,1\}\end{array}\middle|\begin{array}{c}\mathfrak{st}\leftarrow\mathcal{A}_1(\mathrm{Gen}(k))\\ \Xi\leftarrow\mathrm{Prt}(\mathrm{EXT}(\mathcal{U}))\end{array}\right]\right)\\
=&\Pr[\mathcal{A} \text{ wins in IRR Game}]+\frac{1}{2}\Pr[\mathcal{A} \text{ loses in IRR Game}]\\
=&\frac{1}{2}\Big(1+\Pr[\mathcal{A} \text{ wins in IRR Game}]\Big)
\end{aligned}$$

When $b=1$, there are the following two cases in which the adversary $\mathcal{B}$ correctly returns $b'=1$.

Case 1. $\mathcal{A}_2$ guesses a feature element y satisfying $\mathrm{Rcg}(\Xi,y)=$“*match*” and $\mathrm{Rcg}(\Xi',y)=$“*non-match*”.

Case 2. $\mathcal{A}_2$ guesses a feature element y satisfying $\mathrm{Rcg}(\Xi,y)=$“*non-match*” and $\mathcal{B}_2$ chooses $b'=1$ from $\{0,1\}$ with probability $\frac{1}{2}$.

Therefore, we have

$$\begin{aligned}
&\Pr[\mathcal{B} \text{ returns } b'=1 \,|\, b=1]\\
=&\mathop{\mathrm{E}}_{\substack{\mathfrak{st}\leftarrow\mathcal{A}_1(\mathrm{Gen}(k))\\ \Xi\leftarrow\mathrm{Prt}(\mathrm{EXT}(\mathcal{U}))\\ \Xi'\leftarrow\mathrm{Prt}(\mathrm{EXT}(\mathcal{U}))}}\left(\Pr\left[\begin{array}{c} y\leftarrow\mathcal{A}_2(\Xi,\mathfrak{st}),\\ \mathrm{Rcg}(\Xi,y)=\text{“}\textit{match}\text{”},\\ \mathrm{Rcg}(\Xi',y)=\text{“}\textit{non-match}\text{”}\end{array}\middle|\begin{array}{c}\mathfrak{st}\leftarrow\mathcal{A}_1(\mathrm{Gen}(k))\\ \Xi\leftarrow\mathrm{Prt}(\mathrm{EXT}(\mathcal{U}))\\ \Xi'\leftarrow\mathrm{Prt}(\mathrm{EXT}(\mathcal{U}))\end{array}\right]\right.\\
&\left.+\Pr\left[\begin{array}{c} y\leftarrow\mathcal{A}_2(\Xi,\mathfrak{st}),\\ \mathrm{Rcg}(\Xi,y)=\text{“}\textit{non-match}\text{”},\\ b'=1\leftarrow\{0,1\}\end{array}\middle|\begin{array}{c}\mathfrak{st}\leftarrow\mathcal{A}_1(\mathrm{Gen}(k))\\ \Xi\leftarrow\mathrm{Prt}(\mathrm{EXT}(\mathcal{U}))\\ \Xi'\leftarrow\mathrm{Prt}(\mathrm{EXT}(\mathcal{U}))\end{array}\right]\right)\\
=&1-\mathop{\mathrm{E}}_{\substack{\mathfrak{st}\leftarrow\mathcal{A}_1(\mathrm{Gen}(k))\\ \Xi\leftarrow\mathrm{Prt}(\mathrm{EXT}(\mathcal{U}))\\ \Xi'\leftarrow\mathrm{Prt}(\mathrm{EXT}(\mathcal{U}))}}\Pr\left[\begin{array}{c} y\leftarrow\mathcal{A}_2(\Xi,\mathfrak{st}),\\ \mathrm{Rcg}(\Xi,y)=\text{“}\textit{non-match}\text{”}\\ \text{or}\\ \mathrm{Rcg}(\Xi',y)=\text{“}\textit{match}\text{”}\end{array}\middle|\begin{array}{c}\mathfrak{st}\leftarrow\mathcal{A}_1(\mathrm{Gen}(k))\\ \Xi\leftarrow\mathrm{Prt}(\mathrm{EXT}(\mathcal{U}))\\ \Xi'\leftarrow\mathrm{Prt}(\mathrm{EXT}(\mathcal{U}))\end{array}\right]\\
&+\frac{1}{2}\Pr[\mathcal{A} \text{ loses in IRR Game}]\\
\geq&1-\left(\mathop{\mathrm{E}}_{\substack{\mathfrak{st}\leftarrow\mathcal{A}_1(\mathrm{Gen}(k))\\ \Xi\leftarrow\mathrm{Prt}(\mathrm{EXT}(\mathcal{U}))}}\left(\Pr\left[\begin{array}{c} y\leftarrow\mathcal{A}_2(\Xi,\mathfrak{st}),\\ \mathrm{Rcg}(\Xi,y)=\text{“}\textit{non-match}\text{”}\end{array}\middle|\begin{array}{c}\mathfrak{st}\leftarrow\mathcal{A}_1(\mathrm{Gen}(k))\\ \Xi\leftarrow\mathrm{Prt}(\mathrm{EXT}(\mathcal{U}))\end{array}\right]\right.\right.
\end{aligned}$$

$$+ \mathop{\mathrm{E}}_{\substack{\mathfrak{st} \leftarrow \mathcal{A}_1(\mathrm{Gen}(k)) \\ \Xi \leftarrow \mathrm{Prt}(\mathrm{EXT}(\mathcal{U})) \\ \Xi' \leftarrow \mathrm{Prt}(\mathrm{EXT}(\mathcal{U}))}} \Pr\left[\begin{array}{c} y \leftarrow \mathcal{A}_2(\Xi, \mathfrak{st}), \\ \mathrm{Rcg}(\Xi', y) = \text{“}\mathit{match}\text{”} \end{array} \middle| \begin{array}{c} \mathfrak{st} \leftarrow \mathcal{A}_1(\mathrm{Gen}(k)) \\ \Xi \leftarrow \mathrm{Prt}(\mathrm{EXT}(\mathcal{U})) \\ \Xi' \leftarrow \mathrm{Prt}(\mathrm{EXT}(\mathcal{U})) \end{array}\right]\Bigg)$$

$$+\frac{1}{2}\Pr[\mathcal{A} \text{ loses in IRR Game}]$$

$$= \Pr[\mathcal{A} \text{ wins in IRR Game}]$$

$$- \mathop{\mathrm{E}}_{\substack{\mathfrak{st} \leftarrow \mathcal{A}_1(\mathrm{Gen}(k)) \\ \Xi \leftarrow \mathrm{Prt}(\mathrm{EXT}(\mathcal{U})) \\ x' \leftarrow \mathrm{EXT}(\mathcal{U})}} \Pr\left[\begin{array}{c} y \leftarrow \mathcal{A}_2(\Xi, \mathfrak{st}), \\ \mathrm{CMP}(x', y) = \text{“}\mathit{match}\text{”} \end{array} \middle| \begin{array}{c} \mathfrak{st} \leftarrow \mathcal{A}_1(\mathrm{Gen}(k)) \\ \Xi \leftarrow \mathrm{Prt}(\mathrm{EXT}(\mathcal{U})) \\ x' \leftarrow \mathrm{EXT}(\mathcal{U}) \end{array}\right]\Bigg)$$

$$+\frac{1}{2}\Big(1 - \Pr[\mathcal{A} \text{ wins in IRR Game}]\Big)$$

where x' is the original feature element of Ξ'. Since $\mathcal{A}_2$ takes as input Ξ and $\mathfrak{st}$ which are information independent of the event $x' \leftarrow \mathrm{EXT}(\mathcal{U})$, the probability that $\mathcal{A}_2$ guess a feature element y satisfying $\mathrm{CMP}(x', x) =$“*match*” is less than or equals to $\frac{1}{2^k}$ from Definition 1 of the minimum entropy. Then we have

$$\Pr[\mathcal{B} \text{ returns } b' = 1 \,\Big|\, b = 1] \geq \frac{1}{2}\Big(1 + \Pr[\mathcal{A} \text{ wins in IRR Game}]\Big) - \frac{1}{2^k}$$

Consequently we have

$$\begin{aligned} & 2\Pr[\mathcal{B} \text{ in UNLINK Game wins}] - 1 \\ =& 2\Big(\Pr[b=0]\Pr[b'=0 \mid b=0] + \Pr[b=1]\Pr[b'=1 \mid b=1]\Big) - 1 \\ =& \Pr[b'=0 \mid b=0] + \Pr[b'=1 \mid b=1] - 1 \\ \geq& \frac{1}{2}\Big(1 + \Pr[\mathcal{A} \text{ wins in IRR Game}]\Big) + \frac{1}{2}\Big(1 + \Pr[\mathcal{A} \text{ wins in IRR Game}]\Big) - \frac{1}{2^k} - 1 \\ =& \Pr[\mathcal{A} \text{ wins in IRR Game}] - \frac{1}{2^k} = \mathrm{Adv}^{\mathrm{IRR}}_{\Pi,\mathcal{A}}(k) \geq \varepsilon(k) \end{aligned}$$

Therefore we have

$$\mathrm{Adv}^{\mathrm{UNLINK}}_{\Pi,\mathcal{B}}(k) = \big|2\Pr[\mathcal{B} \text{ in UNLINK Game wins}] - 1\big| \geq \varepsilon(k)$$

Hence, the result follows. □

In general, unlinkability is properly stronger than irreversibility. For example, under the assumption that $\mathrm{EXT}(\mathcal{U})$ is a uniform distribution on $\mathcal{M}$, a fuzzy commitment scheme [**9**] achieves sufficiently strong irreversibility but cannot achieve similarly strong unlinkability. Namely, for any ppt IRR Game adversary, the adversary's advantage is negligible, but there exists a ppt UNLINK Game adversary whose advantage is not negligible (cf. [**14**]).

8. Conclusions

In this paper, we introduce explicit formalizations of two important security notions, irreversibility and unlinkability, for BTP algorithms and prove that unlinkability is a stronger security notion than irreversibility. For simplicity, we assume that BTP algorithms are strictly correct and valid. In general, a BTP algorithm is correct and nearly valid, or valid and nearly correct. As a future work, we need to give the proof under such a general assumption.

References

[1] A. Nagar, K. Nandakumarand, and A. K. Jain. Biometric template transformation: A security analysis. In *Proc. SPIE, Electronic Imaging, Media Forensics and Security XII*, 2010.

[2] Xavier Boyen. Reusable cryptographic fuzzy extractors. In *Proceedings of the 11th ACM conference on Computer and communications security*, pages 82–91. ACM, 2004.

[3] Benny Chor and Oded Goldreich. Unbiased bits from sources of weak randomness and probabilistic communication complexity. *SIAM Journal on Computing*, 17(2):230–261, 1988.

[4] Y Dodis, L Reyzin, and A Smith. Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. *Advances in cryptology-Eurocrypt 2004*, pages 523–540, 2004.

[5] Manabu Inuma and Akira Otsuka. Relations among Security Metrics for Template Protection Algorithms. In *Biometrics: Theory, Applications, and Systems, 2013. BTAS 2013. IEEE 6th International Conference on*, pages 1–8, 2013.

[6] ISO/IEC 24745: Information technology – Security techniques – Biometric information protection, 2011.

[7] Anil K Jain, Karthik Nandakumar, and Abhishek Nagar. Biometric template security. *EURASIP Journal on Advances in Signal Processing*, 2008.

[8] A Juels and M Sudan. A fuzzy vault scheme. In *Information Theory, 2002. Proceedings. 2002 IEEE International Symposium on*, 2002.

[9] Ari Juels and Martin Wattenberg. A fuzzy commitment scheme. In *CCS '99: Proceedings of the 6th ACM conference on Computer and communications security*. ACM Request Permissions, November 1999.

[10] K. Simoens, B. Yang, X. Zhou, F. Beato, C. Busch, E. Newton, and B. Preneel. Criteria towards metrics for benchmakring template protection algoritms. In *Proc. of the 5th IAPR International Conference on Biometrics* (*ICB 2012*), 2012.

[11] UNE Masashi and Akira Otsuka. Wolf attack probability: A theoretical security measure in biometric authentication systems. *IEICE transactions on information and systems*, 91(5):1380–1389, 2008.

[12] Nalini Ratha, Sharat Chikkerur, Jonathan Connell, and Ruud Bolle. Generating Cancelable Fingerprint Templates. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4):561–572, 2007.

[13] N.K Ratha, J.H Connell, and R.M Bolle. Enhancing security and privacy in biometrics-based authentication systems. *IBM Systems Journal*, 40(3):614–634, 2001.

[14] Koen Simoens, Pim Tuyls, and Bart Preneel. Privacy Weaknesses in Biometric Sketches. In *30th IEEE Symposium on Security and Privacy (SP)*, pages 188–203. IEEE, May 2009.

[15] Kenta Takahashi and S Hirata. Parameter management schemes for cancelable biometrics. In *Computational Intelligence in Biometrics and Identity Management (CIBIM), 2011 IEEE*

Workshop on, pages 145–151. IEEE, 2011.

[16] Kenta Takahashi and Shinji Hirata. Generating provably secure cancelable fingerprint templates based on correlation-invariant random filtering. In *Biometrics: Theory, Applications, and Systems, 2009. BTAS '09. IEEE 3rd International Conference on*, pages 1–6, 2009.

[17] Kenta Takahashi and Shinji Hirata. Cancelable biometrics with provable security and its application to fingerprint verification. *IEICE transactions on fundamentals of electronics, communications and computer sciences*, 94(1):233–244, 2011.

[18] Kenta Takahashi and K Naganuma. Unconditionally provably secure cancellable biometrics based on a quotient polynomial ring. *Biometrics, IET*, 1(1):63–71, 2012.

[19] Masashi Une, Akira Otsuka, and Hideki Imai. Wolf attack probability: a new security measure in biometric authentication systems. In *Advances in Biometrics*, pages 396–406. Springer, 2007.

Manabu Inuma

Department of Mathematics, Faculty of Science, Josai University
Keyakidai 1-1, Sakado, Saitama, 350-0295, Japan
inuma@josai.ac.jp