IEICE TRANS. INF. & SYST., VOL.E103–D, NO.5 MAY 2020

PAPER

Security Evaluation of Negative Iris Recognition

Osama OUDA[†,††a)], **Slim CHAOUI**[†,†††], *Nonmembers*, *and* **Norimichi TSUMURA**[††††], *Member*

SUMMARY Biometric template protection techniques have been proposed to address security and privacy issues inherent to biometric-based authentication systems. However, it has been shown that the robustness of most of such techniques against reversibility and linkability attacks are overestimated. Thus, a thorough security analysis of recently proposed template protection schemes has to be carried out. Negative iris recognition is an interesting iris template protection scheme based on the concept of negative databases. In this paper, we present a comprehensive security analysis of this scheme in order to validate its practical usefulness. Although the authors of negative iris recognition claim that their scheme possesses both irreversibility and unlinkability, we demonstrate that more than 75% of the original iris-code bits can be recovered using a single protected template. Moreover, we show that the negative iris recognition scheme is vulnerable to attacks via record multiplicity where an adversary can combine several transformed templates to recover more proportion of the original iris-code. Finally, we demonstrate that the scheme does not possess unlinkability. The experimental results, on the CASIA-IrisV3 Interval public database, support our theory and confirm that the negative iris recognition scheme is susceptible to reversibility, linkability, and record multiplicity attacks.

key words: *biometric template protection, negative iris recognition, irreversibility, attacks via record multiplicity, linkability attacks*

1. Introduction

With the increased deployment of biometric-based authentication systems, protecting stored and/or transmitted biometric data against unauthorized disclosure, eavesdropping, malicious or accidental alteration has become a key requirement to secure non-revocable biometric templates and safeguard users' privacy. As a result, several biometric template protection schemes have been proposed over the past few years [1]. These schemes can be broadly classified into two main categories; namely, cancelable biometrics and biometric cryptosystems. The main goal of both categories is to authenticate/identify individuals without leaking much information about their inherent biometric data via deriving irreversible identities from such data. Schemes in both categories should not deteriorate the recognition accuracy of the underlying biometric recognition system. Moreover, such schemes should resist various types of security and privacy attacks.

Although authors of most existing biometric template protection schemes payed much attention to prove that their proposed schemes preserve recognition accuracy, the security of such schemes is ignored or superficially discussed and is therefore suspected to be highly overestimated [2], [3]. A rigorous security analysis of existing schemes is thus required in order to ensure that they can resist reversibility and linkability attacks before they are deployed in real-world applications that require high security. Some recent studies [4], [5] have already shown that several template protection schemes are vulnerable to different types of attacks. As a result, remedies to the discovered defects have been proposed to enhance the security of some flawed schemes [6]–[8].

Fuzzy commitment [9] and fuzzy vault [10] schemes are two popular biometric cryptosystems that utilize error-correcting codes to handle intra-class variations in biometric data. The fuzzy commitment scheme (FCS) follows a key binding approach in which a biometric key is generated via XOR-ing a randomly selected codeword with the binary string derived from the biometric sample. Although the FCS has been successfully applied to several biometric modalities such as iris [11], fingerprint [12], face [13], and online signature [14], it has been shown that it is vulnerable to the decodability attack [15] which facilitates matching protected templates across different applications. Kelkboom et al. [7] suggested to implement a random bit shuffling process on the binary biometric vector before applying the FCS to it in order to prevent cross-matching. However, Tams [16] showed that this countermeasure can not completely prevent attackers from cross-matching biometric keys generated using the FCS.

Unlike the FCS, the fuzzy vault scheme (FVS) is suitable for protecting unordered sets of biometric features and thereby it can be applied to various biometric modalities without the need for preprocessing steps (such as binarization). The idea of the FVS is to use biometric features to lock a random key encoded as the coefficients of a selected polynomial p by projecting these features onto p and adding chaff points in order to obscure the projected biometric features. The FVS is one of the most accepted and well-studied [17], [18] approaches for binding biometric data with cryptographic keys. However, several studies showed that it is vulnerable to various attacks such as brute-

Manuscript received October 11, 2019.
Manuscript revised January 7, 2020.
Manuscript publicized January 29, 2020.

[†]The authors are with the College of Computer and Information Sciences, Jouf University, Sakaka, Saudi Arabia.
[††]The author is with the Department of Information Technology, Faculty of Computer and Information Sciences, Mansoura University, Mansoura 35516, Egypt.
[†††]The author is with Sfax University, SETIT-Lab, Sfax, Tunisia.
[††††]The author is with the Department of Information and Computer Sciences, Chiba University, Chiba-shi, 263–0022, Japan.
a) E-mail: omalsayed@ju.edu.sa

DOI: 10.1587/transinf.2019EDP7276

force [19], correlation [20], record multiplicity [21], and collusion [22] attacks.

In addition to biometric cryptosystems, cancelable biometrics (CB) schemes offer a means for protecting biometric data through the application of non-invertible transforms that can derive several different distorted versions of the same biometric sample. BioHashing [23] is a well-known scheme that has been applied to a variety of biometrics modalities [24]–[26]. The basic idea behind BioHashing is to apply user-specific random projections to biometric features in order to obtain cancelable templates. CB schemes based on the basic concept of BioHashing suffer from several vulnerabilities that would enable attackers to launch reversibility [27] and pre-image attacks [5]. BioEncoding [28], [29] is another CB scheme that can be used to protect binary biometric data such as iris-codes. This transform divides a binary template into words of fixed size and then apply random Boolean functions to each word so that each word is mapped to a single random bit. The protected template consists of the resulting bits. Lacharme [30] pointed out that BioEncoding is susceptible to correlation and invertibility attacks since it is possible to reconstruct several pre-images from the protected template. Ouda et al. [31] suggested to randomly shuffle bits of the iris-code using an application-specific permutation prior to applying the BioEncoding transformation in order to thwart correlation attacks as well as to ensure diversity.

Rathgeb et al. [32] proposed a CB iris biometric template protection scheme based on Bloom filters. This scheme divides the iris-code into a number of sub-matrices and transforms each sub-matrix into a Bloom filter using predetermined application-specific secret value. Hermans et al. [4] presented a simple attack that illustrates that this scheme does not achieve unlinkability. Gomez-Barrero et al. [6] proposed to integrate an additional structure-preserving feature re-arrangement step to improve the robustness of the original scheme in [32] against cross-matching attack.

A novel cancelable iris biometric technique, referred to as the negative iris recognition scheme, has been recently proposed by Zhao et al. [33] to protect iris-codes. This scheme utilizes the concept of negative database (NDB) [34], which is a new technique for privacy preservation, to generate revocable iris templates. The authors claim that their scheme is secure against possible reversibility and linkability attacks since reversing the NDB has been demonstrated to be an NP-hard problem [35].

In this paper, we present a thorough security analysis of the negative iris recognition scheme and demonstrate that it is vulnerable to reversibility, linkability and record multiplicity attacks. In order to confirm the presented analysis, we tested the proposed attacks on a publicly available iris images dataset. The obtained experimental results illustrate the effectiveness of the proposed attacks and show that, contrary to the claims stated in [33], the scheme does not satisfy irreversibility and unlinkability requirements.

The remainder of this paper is structured as follows. Section 2 provides a brief overview of the negative iris recognition scheme. Section 3 presents our security analysis and describes three different attacks against the negative iris scheme. The experimental results are presented in Sect. 4. Finally, Sect. 5 concludes the paper.

2. Negative Iris Recognition

The negative iris recognition scheme is based on the concept of negative database which was firstly proposed by Esponda et al. [36] to protect data privacy. Rather than storing the actual data, the basic idea behind NDB is to store elements in the complementary set of the original database. Consider a set $U = \{0,1\}^n$ and a database $DB \subset U$. An NDB of DB covers only the elements in $U - DB$ and is said to be complete if $NDB = U - DB$; otherwise, it is incomplete. Typically, a negative database is expressed in a compressed form using the notation $'*'$ that can represent both 0 and 1. Thus, for each entry $\mathbf{t} \in \{0,1,*\}^n$ in NDB, the positions that have the value $'*'$ are called unspecified positions whereas positions that have the values 0 or 1 are called specified positions. Table 1 demonstrates different examples of complete (NDB_1 and NDB_2) and incomplete (NDB_3 and NDB_4) negative databases of a database containing two 3-bit entries.

Several algorithms for generating *NDBs* have been proposed [37], [38] in the literature. However, the irreversibility of most of these algorithms cannot be ensured or controlled. The p-hidden algorithm proposed by Liu et al. [37] has been demonstrated to have strong security against reversibility attacks [35]. Therefore, Zhao et al. [33] chose this algorithm to implement their negative iris recognition scheme. The p-hidden algorithm can generate a hard-to-reverse negative database ($NDB_{\mathbf{s}}$) with $m = n \times r$ entries from a database with a single n-bit binary string ($\mathbf{s}$), where r is a parameter that controls the number of entries in $NDB_{\mathbf{s}}$ (often set to 6.5). At each iteration of the p-hidden procedure, one entry $\mathbf{t}$, with three randomly selected specified positions and $n-3$ unspecified positions, is generated and added to $NDB_{\mathbf{s}}$. The values at the specified positions in $\mathbf{t}$ differ from the values of the corresponding positions in $\mathbf{s}$ according to two probability parameters, p_1 and p_2. Specifically, values at the specified positions of $\mathbf{s}$ and $\mathbf{t}$ could differ at one, two, or three positions with probabilities p_1, p_2, or $p_3 = 1 - p_1 - p_2$, respectively. Since the positions that have different values across $\mathbf{s}$ and $\mathbf{t}$ are selected randomly and with different probabilities, it would be difficult for an attacker to reverse the generated NDB and recover the original DB.

In order to generate hard-to-reverse negative databases,

Table 1 Examples of complete and incomplete NDBs.

DB	$U - DB$	NDB_1	NDB_2	NDB_3	NDB_4
001	000	1^{**}	$^{**}0$	1^*0	0^*0
011	010	0^*0	1^*1	000	11^*
	100				101
	101				
	110				
	111				

Liu et al. [37] demonstrated that p_1 and p_2 should satisfy $4p_1 + 2p - 2 > 3$ and $p_1 + p_2 < 1$. Besides, they analyzed the hardness level of reversing *NDBs* generated by the p-hidden algorithm for different values of p_1, p_2, and r and they found that the highest security level can be achieved when $p_1 = 0.80, p_2 = 0.14$, and $r = 6.5$. Thereby, the same parameter settings have been adopted in the implementation of the negative iris recognition scheme [33].

3. Attacks on Negative Iris Recognition

The security of the negative iris recognition scheme is based on the fact that reversing the p-hidden-NDB to recover the protected template is equivalent to solving the 3-SAT problem that has at least one solution. However, in the context of biometric systems, the authentication process succeeds if the sample presented to the system during verification is sufficiently similar to the stored template. In other words, the adversary does not need to obtain an exact version (solution) of the stored template in order to spoof the authentication system. Typically, the standard deployed matching threshold for iris recognition systems is 0.32 fractional Hamming distance [39]. That is, it is sufficient for the adversary to recover at least 68% of the iris-code bits in order to break into iris-code based authentication systems. In this section, we show that although it might be difficult to completely recover the original iris-code from the p-hidden-NDB, it is possible to obtain a bit-string that is similar enough to spoof the authentication system. Moreover, the vulnerability of the scheme to attacks via record multiplicity as well as linkage attacks is analyzed.

3.1 Reversibility Attacks

In this section, we discuss the robustness of the negative iris recognition scheme against reversibility, a.k.a. invertibility, attacks. As described in the previous section, for a *DB* containing a single iris-code $\mathbf{x} \in \{0,1\}^n$, a corresponding negative database, $NDB_{\mathbf{x}}$, of $m = r \times n$ entries, $\mathbf{t}_i \in \{0,1,*\}^n, i = 1,2,\ldots,m$, can be generated using the $p-$hidden algorithm. Let $\mathbf{t}_i(k)$, the bit at position k of an entry $\mathbf{t}_i$ in $\text{NDB}_{\mathbf{x}}$, be a specified bit. According to the p-hidden algorithm, one, two, or three of the (three) specified bits of any entry in $NDB_{\mathbf{x}}$ are flipped with probabilities p_1, p_2, and $p_3 = 1 - (p_1 + p_2)$, respectively. Thus, the probability, P_{diff}, that $\mathbf{t}_i(k)$ is a flipped version of $\mathbf{x}(k)$ is given as:

$$P_{diff} = P(\mathbf{x}(k) \neq \mathbf{t}_i(k)) = \frac{p_1 + 2p_2 + 3p_3}{3}. \quad (1)$$

Similarly, the probability, P_{same}, that $\mathbf{t}_i(k)$ is not a flipped version of $\mathbf{x}(k)$, i.e. $\mathbf{x}(k) = \mathbf{t}_i(k)$, is given as:

$$P_{same} = P(\mathbf{x}(k) = \mathbf{y}_i(k)) = \frac{2p_1 + p_2}{3}. \quad (2)$$

Reversing $\text{NDB}_{\mathbf{x}}$, and hence recovering the original iris-code $\mathbf{x}$, would be difficult if $P_{same} = P_{diff}$. However, as pointed out in [37], the p-hidden algorithm can generate hard-to-reverse NDBs if and only if the conditions $4p_1 + 2p_2 > 3$ and $p_1 + p_2 < 1$ are satisfied. Consequently, form (3), we have $0.5 < P_{same} < 2/3$; that is, $P_{same} > P_{diff}$. Moreover, it has been shown that the security of NDBs generated using the p-hidden algorithm increases when p_1 increases, i.e. $p_1 >> p_2$. Specifically, the results reported in [37] have shown that the highest hardness level of reversing NDBs generated using the p-hidden algorithm is achieved at $p_1 = 0.80, p_2 = 0.14$, and $r = 6.5$. This implies that, under these recommended settings, the probability that any bit at a specified position k in $\text{NDB}_{\mathbf{x}}$ is equal to $\mathbf{x}(k)$ is 0.58.

Algorithm 1 Reversing p-hidden based negative iris-codes

Input: a negative iris-code $\text{NDB}_{\mathbf{x}}$
Output: recovered iris-code $\hat{\mathbf{x}}$

```
for k ← 1 to n do
    S_k ← 0
    sum_k ← 0
    for i ← 1 to m do
        if t_i(k) ≠' *' then
            sum_k = sum_k + t_i(k)
            S_k = S_k + 1
        end if
    end for
    α = sum_k / S_k
    if α > 0.5 then
        x̂(k) = 1
    else
        x̂(k) = 0
    end if
end for
return x̂
```

The above parameter settings were adopted in [33] to secure iris-codes using negative databases. Here we show that under these parameter settings, the attackers can recover more than 75% of $\mathbf{x}$ using a single NDB. Let S_k be the number of times each position k in $NDB_{\mathbf{x}}$ is selected as a specified position, and let sum_k be the sum of all bits at that position. Algorithm 1 demonstrates a simple yet powerful method that can be followed by the adversary to recover the value of $\mathbf{x}(k)$. At each position k, he/she can simply compute $\alpha = sum_k/S_k$ and then use the following rule:

$$\hat{\mathbf{x}}(k) = \begin{cases} 1 & \alpha > 0.5 \\ 0 & \text{otherswise,} \end{cases}$$

where $\hat{\mathbf{x}}(k)$ is the guessed value of $\mathbf{x}(k)$. Obviously, the expectation of $\alpha = P_{same}$. Therefore, the probability that $\mathbf{x}(k)$ is guessed correctly using the above rule is:

$$\begin{aligned} P_{correct} &= P(\mathbf{x}(k) = \hat{\mathbf{x}}(k)) \\ &= \sum_{i=\lfloor \frac{S_k}{2} \rfloor + 1}^{S_k} \binom{S_k}{i} \times P_{same}^i \times (1 - P_{same})^{S_k - i}. \end{aligned} \quad (3)$$

Let γ denote the proportion of correctly guessed iris-code bits. Thus, the expectation of γ, $E(\gamma)$, is equal to $P_{correct}$. Figure 1 shows the relationship between $E(\gamma)$ and P_{same}. As it can be seen in the figure, $E(\gamma)$ increases with P_{same} and has its minimum value when P_{same} is approximately equal to 0.5. Since smaller values of $E(\gamma)$ implies

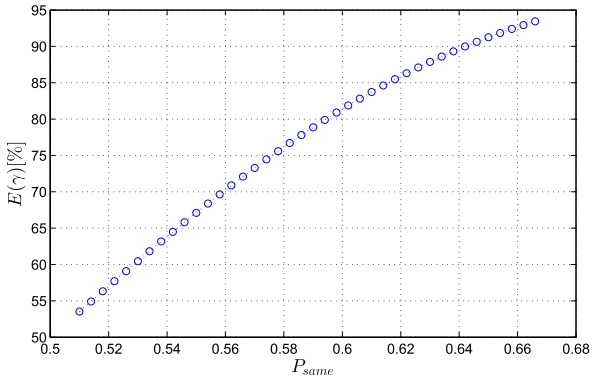

Fig. 1 The relationship between the expectation of the percentage of correctly guessed iris-code bits and P_{same}.

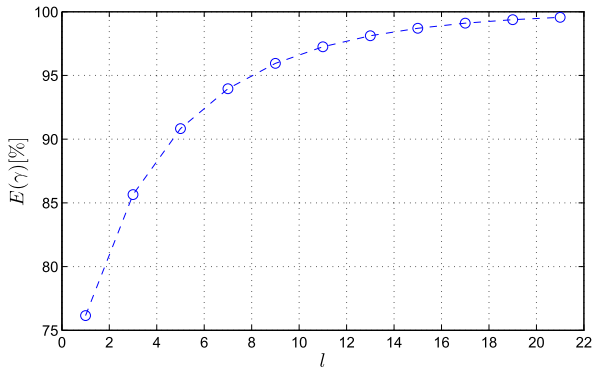

Fig. 2 The relationship between the expected percentage of correctly recovered iris-code bits and the number of compromised *NDBs* used in ARM.

higher irreversibility levels of $NDB_{\mathbf{s}}$, one should consider values of p_1 and p_2 that give small values of P_{same}, i.e. $p_1 \approx p_2 \approx 0.5$. However, as discussed in [37], with such values of p_1 and p_2, it is not possible to generate hard-to-reverse negative databases. On the other hand, under the recommended values of p_1 and p_2 ($p_1 = 0.80, p_2 = 0.14$), attackers can recover more than 75% of the original iris-code (see Fig. 1, $P_{same} = 0.58$). Precisely, from (4), the expected percentage of correctly recovered iris-code bits is 76.15% at $P_{same} = 0.58$ which is a sufficient percentage for launching spoofing (pre-image) attacks against iris recognition systems since the standard.

Obviously, the time complexity of computing $\hat{\mathbf{x}}$ using the attack described in Algorithm 1 depends primarily on the iris-code length as well as the number of entries in the negative iris database. Thus, the required time complexity is $O(mn)$. This poses a serious threat to the negative iris recognition scheme even if it adopts had-to-reverse NDBs.

3.2 Attacks via Record Multiplicity

The analysis presented in the previous subsection assumes that the adversary has an access to a single enrollment NDB. However, attackers might be able to collect multiple enrollment templates and try to combine them somehow to recover the original iris-code. This type of attacks is commonly referred to as attacks via record multiplicity (ARM) [40]. In this subsection, we investigate the possibility of retrieving the original iris-code via combining different numbers of compromised NDBs.

Assume that the attacker could have access to l NDBs, $\{NDB_{\mathbf{x}_i}\}_{i=1}^{l}$, generated from the same iris-code $\mathbf{x}$. The attacker can follow a straightforward strategy to retrieve $\mathbf{x}$ by firstly reversing each compromised template separately following the procedure described in the previous subsection to obtain $\{\hat{\mathbf{x}}_i\}_{i=1}^{l}$. Then, he/she can consider bits at the same position k across the l recovered bit-strings as an l-bit binary repetition code, of the kth bit of $\mathbf{x}$, that might have been corrupted due to the applied cancelable negative database transformation. A simple majority decision for each l-bit codeword can be utilized to decide whether to consider the correct bit value is zero or one. After decoding all bits, a recovered version, $\hat{\mathbf{x}}$, of $\mathbf{x}$ can be obtained.

As discussed in the previous subsection, the probability of correctly guessing each bit of $\mathbf{x}$, $P_{correct}$, is 0.7615 given a single compromised template (under the parameter settings adopted in [33]). This implies that the number of bits in $\mathbf{x}$ that can be retrieved correctly will increase as l increases (since $P_{correct} \gg 0.5$). The probability, P_{decode}, of correctly decoding each l-bit word composed from bits at the same position k across the bit-strings $\{\hat{\mathbf{x}}_i\}_{i=1}^{l}$ can be found as follow:

$$P_{decode} = \sum_{i=\lfloor \frac{l}{2} \rfloor + 1}^{l} \binom{l}{i} \times P_{correct}^{i} \times (1 - P_{correct})^{l-i}. \qquad (4)$$

The relationship between $E(\gamma)$ and l, under the recommended parameter settings, is depicted in Fig. 2. It is clear from this figure that the percentage of correctly recovered bits of $\mathbf{x}$ jumps from approximately 76% to more than 85% at $l = 3$. Obviously, this increase in percentage continues as l increases. Thus, we can conclude that the p-hidden based negative iris-codes are susceptible to ARM.

3.3 Linkage Attacks

Unlinkability is another major security requirement that is expected to be fulfilled by biometric template protection schemes. A template protection scheme is said to satisfy unlinkability if different secured templates that are generated from the same biometric data are not linkable across applications. Zhao et al. [33] claim that attackers cannot determine whether any two different p-hidden-NDBs are generated from the same iris data (and thereby their negative iris recognition scheme possesses unlinkability) since this is equivalent to the problem of determining whether two satisfiable SAT instances have the same solution. As we mentioned earlier, in the context of biometric systems, the attacker does not need to find an exact version (solution) of the original biometric data in order to violate users' privacy. In this section, we show how an attacker can utilize Algorithm 1 to launch a linkage attack on Zhao et al.'s negative

iris recognition scheme given two negative iris databases.

Let $NDB_{\mathbf{x}}$ and $NDB_{\mathbf{y}}$ be two negative iris databases stored in two different applications. The attacker can use Algorithm 1 to obtain $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ and then find the Hamming distance between them, $d_H(\hat{\mathbf{x}}, \hat{\mathbf{y}})$. Let X be the random variable that denotes the Hamming distance between $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$. Based on the analysis presented in Sect. 3-A, the expected Hamming distance, $E(d_H(\mathbf{x}, \hat{\mathbf{x}}))$, between the original, $\mathbf{x}$, and recovered, $\hat{\mathbf{x}}$, strings can be found using γ. Specifically, if $NDB_{\mathbf{x}}$ and $NDB_{\mathbf{y}}$ are generated from the same iris-code (i.e. $\mathbf{x} = \mathbf{y}$), then $E(d_H(\mathbf{x}, \hat{\mathbf{x}})) = E(d_H(\mathbf{x}, \hat{\mathbf{y}})) = 1 - E(\gamma)$. In this case, X follows a binomial distribution, $B(n, p)$, with parameters n and $p = 2 \times P_{correct} \times (1 - P_{correct})$. On the other hand, if $NDB_{\mathbf{x}}$ and $NDB_{\mathbf{y}}$ are generated from iris-codes belonging to two different classes(eyes), then X satisfies $B(n, 0.5)$.

Figure 3 depicts the relationship between $E(\gamma)$ and $E(X)$. The figure clearly shows that $E(X)$ increases as $E(\gamma)$ decreases. This implies that as the probability of guessing bits of the original iris-code (i.e. $P_{correct}$) increases, the Hamming distance between reversed negative databases generated from the same iris-code decreases. Specifically, under the parameter settings adopted in the negative iris recognition scheme [33], the expected Hamming distance between strings recovered from two protected negative databases of the same iris-code is 0.36, compared to 0.50 for strings obtained from negative databases generated from different iris-codes. It is worth noting that the above analysis assumes that $\mathbf{x} = \mathbf{y}$. In practice, however, iris-codes generated from the same eye are not identical. Therefore, the relationship between $d_H(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ and $d_H(\mathbf{x}, \mathbf{y})$ should be investigated.

Our goal is to find the expectation of $d_H(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ at any value of $d_H(\mathbf{x}, \mathbf{y})$. Let us consider that each bit x' in $\hat{\mathbf{x}}$ is equal to the corresponding bit y' in $\hat{\mathbf{y}}$ with probability $(1-q)$ (see Fig. 4 (a)). Then, the expectation of $d_H(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ is equal to q. To find q, we need to find the values of two other probabilities, q_1 and q_2 (see Fig. 4 (b)), defined as follows:

$$\begin{aligned} q_1 &= P(x \neq x') = P(y \neq y'), \\ q_2 &= P(x \neq y), \end{aligned} \tag{5}$$

where x, y, x', and y' denote bits at the same position in the strings $\mathbf{x}, \mathbf{y}, \hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$, respectively. Given $d_H(\mathbf{x}, \mathbf{y})$ and recalling from the previous subsection that $q_1 = 1 - P_{correct}$, q can be computed as follows (see Fig. 4 (b)):

$$q = q_1^2 q_2 + (1 - q_1)^2 q_2 + 2(1 - q_1)(1 - q_2) q_1. \tag{6}$$

Figure 5 depicts the relationship between $d_H(\mathbf{x}, \mathbf{y})$ and $E(d_H(\hat{\mathbf{x}}, \hat{\mathbf{y}}))$. This figure illustrates that the expectation of the Hamming distance between the recovered iris-codes increases linearly with the Hamming distance between the original codes. The figure also shows that even if the Hamming distances between similar (generated from the same eye) unprotected iris-codes are relatively large (≤ 0.3), the expectation of the Hamming distances between their corresponding recovered bit-strings is distinguishable from those obtained from matching iris-codes generated from different eyes. That is, negative iris-codes based on the p-hidden algorithm are vulnerable to linkage attacks since, as opposed

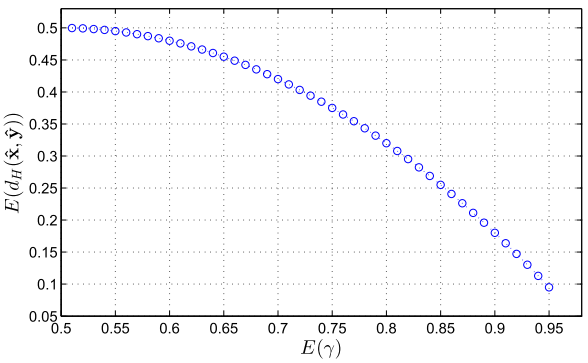

Fig. 3 The relationship between the expected proportion of correctly recovered iris-code bits, $E(\gamma)$, and the Hamming distance between two reversed iris-codes $E(d_H(\hat{\mathbf{x}}, \hat{\mathbf{y}}))$ using Algorithm 1.

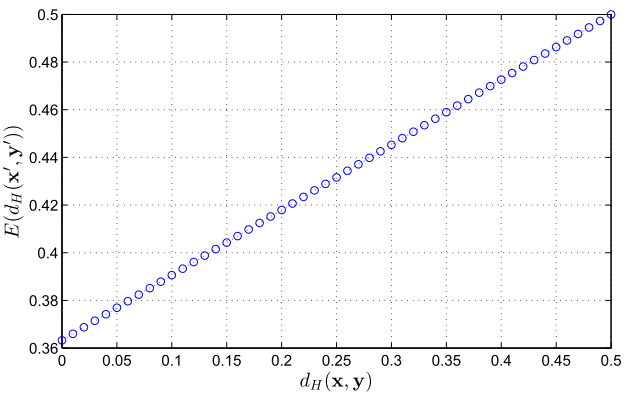

Fig. 5 The relationship between the Hamming distances between original iris-codes and the Hamming distances between the corresponding codes recovered using Algorithm 1.

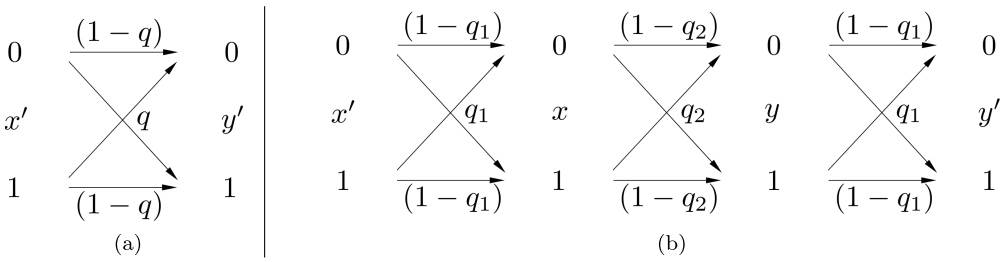

Fig. 4 Modeling bit transitions between two iris-codes recovered from two different NDBs.

to what is claimed in [33], attackers can easily determine whether two p-hidden-NDBs are generated from the same class or not.

4. Experimental Results

We investigated the effectiveness of the proposed attacks against negative iris recognition on the widely used iris dataset CASIA-IrisV3-Interval [41]. This dataset consists of 2639 iris images captured from 295 eyes (classes) of 249 different people. Iris-codes were obtained from all images in the dataset using the method proposed by Masek in [42]. In order to generate the protected iris-codes, we have implemented the p-hidden-based negative database algorithm described in [33], [40] using MATLAB R2017a running on a 2.4GHz 64bit Windows 10 with 8GB memory. The correctness of our implementation has been verified by evaluating the recognition accuracy of the protected iris recognition system adopting the parameter values tested in [33] and matching the obtained results with the results reported in [33].

4.1 Reversibility

In this experiment, we investigated the average fraction of recovered bits obtained using our proposed reversibility attack. Cancelable iris-codes were obtained by applying the p-hidden NDB algorithm to the iris-codes generated from all images in the utilized dataset. Although Liu et al. [37] suggested to set $p_1 = 0.80$ and $p_2 = 0.14$ in order to achieve high security, we evaluated the proposed reversibility attack using different values of p_1 and p_2 (including the recommended values). All tested values satisfy the two conditions ($p_1 + p_2 < 1$ and $4p_1 + 2p_2 > 3$) that are required to generate hard-to-reverse NDBs. In all experiments, we set r to 6.5 as suggested in [33], [37]. The tested values of p_1 and p_2 along with the corresponding values of P_{same} are listed in Table 2.

The proposed reversibility attack was then employed to recover the original iris-codes from the cancelable NDBs. The Hamming distance between each iris-code and the bit-string obtained from reversing its corresponding cancelable NDB was calculated and the average distance was obtained for all iris-codes. For each set of chosen parameters, experiments were repeated 100 times to check the reliability of our results. Figure 6 shows the average fraction of recovered bits obtained experimentally and analytically (Eq. 3) at the tested values of p_1 and p_2.

The experimentally obtained average values are displayed as squares with error bars that show the standard deviation for each average value. The small error bars indicate the reliability of the average value as a representative number for the obtained experimental results. The figure clearly shows that the obtained experimental results are almost identical to the analytical results. The figure also illustrates that the average fraction of recovered bits increases as the value of P_{same}. Moreover, the average fraction of recovered bits is larger than 0.75 at the recommended values of p_1 and p_2 ($p_1 = 0.80, p_2 = 0.14, P_{same} = 0.58$).

Table 2 Tested values of p_1 and p_2.

p_1	p_2	$p_1 + p_2$	$4p_1 + 2p_2$	P_{same}
0.70	0.11	0.81	3.02	0.50
0.70	0.13	0.83	3.06	0.51
0.70	0.16	0.86	3.12	0.52
0.70	0.19	0.89	3.18	0.53
0.70	0.22	0.92	3.24	0.54
0.70	0.25	0.95	3.30	0.55
0.70	0.28	0.98	3.36	0.56
0.80	0.11	0.91	3.42	0.57
0.80	0.14	0.94	3.48	0.58
0.80	0.17	0.97	3.54	0.59
0.90	0.00	0.90	3.60	0.60
0.90	0.03	0.93	3.66	0.61
0.90	0.06	0.96	3.72	0.62
0.90	0.09	0.99	3.78	0.63
0.94	0.04	0.98	3.84	0.64
0.96	0.03	0.99	3.90	0.65
0.99	0.00	0.99	3.96	0.66

4.2 Attacks via Record Multiplicity

In this experiment, we investigate the effectiveness of combining multiple iris negative databases generated from the same iris sample in order to recover the original iris-code. For each iris-code in the utilized dataset, 21 different NDBs were generated using the p-hidden algorithm employing the recommended parameter settings ($p_1 = 0.80$, $p_2 = 0.14$, and $r = 6.5$). The average fractions of correctly recovered bits resulting from combining different numbers of iris NDBS (ranging from 3 to 21 step 2) were calculated as described in Sect. 3-B. Figure 7 shows the results obtained experimentally as well as analytically. Obviously, the experimental results confirm the analytical result and shows that the fraction of correctly recovered bits increases as the number (l) of the combined NDBs increases.

4.3 Linkability

In this experiment, we investigate the robustness of Zaho

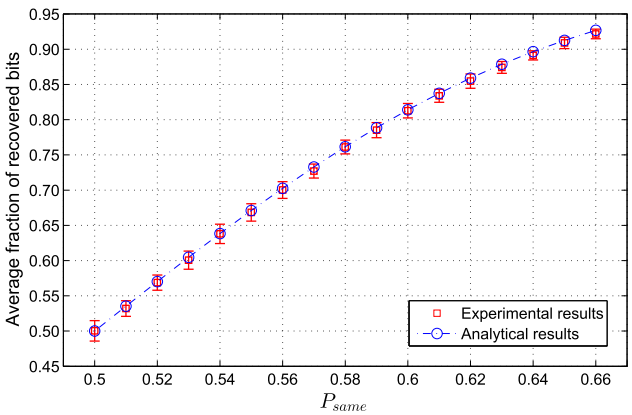

Fig. 6 Average fraction of correctly recovered bits obtained experimentally and analytically for the proposed reversibility attack.

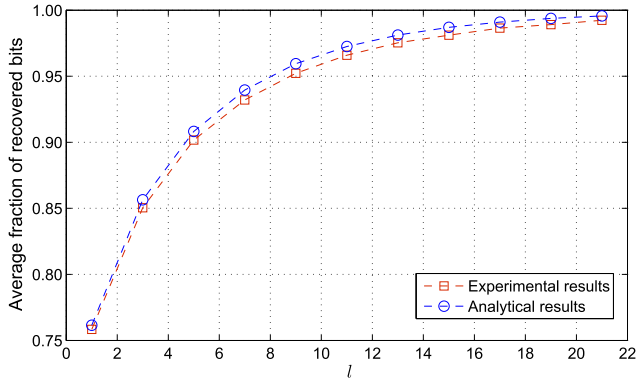

Fig. 7 Average fraction of correctly recovered bits obtained experimentally and analytically for the proposed record multiplicity attack.

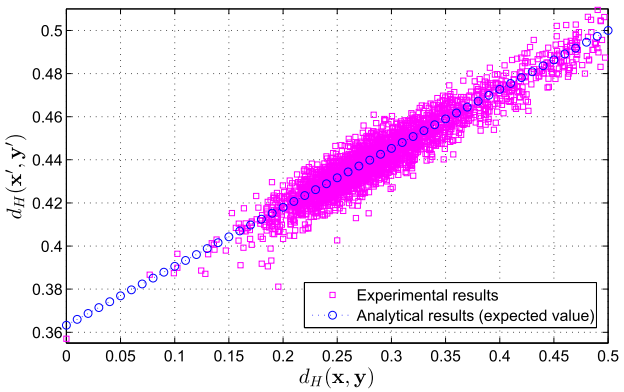

Fig. 8 The relationship between the Hamming distances between original iris-codes and the Hamming distances between the corresponding codes recovered using Algorithm 1.

et al.'s negative iris recognition scheme against linkability attacks. For each class in the adopted iris dataset, we randomly chose two different iris images and obtained the protected iris NDBs by applying the p-hidden algorithm to their corresponding iris-codes. Then, both NDBs were reversed using the algorithm described in Sect. 3-A and the Hamming distance between the recovered bit-strings was calculated.

The relationship between the Hamming distances between the original pairs of iris-codes (belonging to the same class) and the Hamming distances between the corresponding recovered bit-strings is depicted in Fig. 8. We can observe from the figure that the obtained results supports the analytical result (blue circles) and indicates that if the Hamming distance between two iris-codes is less than 0.25 (average Hamming distance between iris-codes generated from the same eye), the Hamming distance between their corresponding recovered bit-strings is less than 0.44. This can give attackers a valuable clue for guessing whether two iris-codes are belonging to the same class or not.

5. Conclusion

In this paper, we presented a security analysis of the negative iris recognition scheme that has recently been proposed to secure iris-codes utilizing the concept of negative database. In contrast to what is claimed by its authors, we have shown that this scheme is vulnerable to reversibility attacks. Precisely, we demonstrated a simple yet powerful approach that can recover more than 75% of the original iris-code using a single protected template. The recovered code can be utilized to launch a pre-image attack to spoof iris-code based authentication systems. Besides, we have also shown that attackers can combine multiple negative iris-codes to recover larger proportion of the original iris-code. Precisely, the obtained results demonstrated that it is possible to recover more than 85% of the original iris-code using a record-multiplicity attacks that combines only three compromised NDBs. Moreover, in contrast to what is claimed by Zhao et al., the presented security analysis pointed out that the negative iris recognition scheme is vulnerable to linkage attacks since the attacker can easily determine whether two p-hidden-NDBs are generated from the same iris template or not. Overall, we conclude that the scheme does not possess irreversibility and unlinkability requirements that should be satisfied by secure biometric template protection schemes.

Acknowledgments

This work was supported by Jouf University, Sakaka, Al Jouf, Saudi Arabia under Grant 40/342.

References

[1] P. Campisi, Security and Privacy in Biometrics, Springer, 2013.
[2] A.K. Jain, K. Nandakumar, and A. Nagar, "Biometric template security," EURASIP Journal on advances in signal processing, vol.2008, p.113, 2008.
[3] C. Rathgeb and A. Uhl, "A survey on biometric cryptosystems and cancelable biometrics," EURASIP Journal on Information Security, vol.2011, no.1, p.3, 2011.
[4] J. Hermans, B. Mennink, and R. Peeters, "When a bloom filter is a doom filter: Security assessment of a novel iris biometric template protection system," 2014 IEEE International Conference of the Biometrics Special Interest Group (BIOSIG), pp.1–6, 2014.
[5] P. Lacharme, E. Cherrier, and C. Rosenberger, "Preimage attack on biohashing," 2013 International Conference on Security and Cryptography (SECRYPT), pp.363–370, IEEE, 2013.
[6] M. Gomez-Barrero, C. Rathgeb, J. Galbally, C. Busch, and J. Fierrez, "Unlinkable and irreversible biometric template protection based on bloom filters," Information Sciences, vol.370, pp.18–32, 2016.
[7] E.J.C. Kelkboom, J. Breebaart, T.A.M. Kevenaar, I. Buhan, and R.N.J. Veldhuis, "Preventing the decodability attack based cross-matching in a fuzzy commitment scheme," IEEE Transactions on Information Forensics and Security, vol.6, no.1, pp.107–121, March 2011.
[8] K. Nandakumar, A. Nagar, and A.K. Jain, "Hardening fingerprint fuzzy vault using password," International conference on Biometrics, pp.927–937, Springer, 2007.
[9] A. Juels and M. Wattenberg, "A fuzzy commitment scheme," ACM Conference on Computer and Communications Security, pp.28–36, 1999.
[10] A. Juels and M. Sudan, "A fuzzy vault scheme," Des. Codes Cryptography, vol.38, no.2, pp.237–257, 2006.
[11] F. Hao, R. Anderson, and J. Daugman, "Combining crypto with biometrics effectively," IEEE Trans. Computers, vol.55, no.9,

pp.1081–1088, 2006.
[12] K. Nandakumar, "A fingerprint cryptosystem based on minutiae phase spectrum," 2010 IEEE International Workshop on Information Forensics and Security, pp.1–6, Dec. 2010.
[13] H. Lu, K. Martin, F. Bui, K.N. Plataniotis, and D. Hatzinakos, "Face recognition with biometric encryption for privacy-enhancing self-exclusion," 2009 16th International Conference on Digital Signal Processing, pp.1–8, July 2009.
[14] E. Maiorana and C. Ercole, "Secure biometric authentication system architecture using error correcting codes and distributed cryptography," Proceedings of Gruppo Nazionale Telecomunicazioni e Teoria dell' Informazione, pp.1–12, 2007.
[15] X. Zhou, A. Kuijper, R. Veldhuis, and C. Busch, "Quantifying privacy and security of biometric fuzzy commitment," 2011 International Joint Conference on Biometrics (IJCB), pp.1–8, IEEE, 2011.
[16] B. Tams, "Decodability attack against the fuzzy commitment scheme with public feature transforms," arXiv preprint arXiv: 1406.1154, 2014.
[17] K. Koptyra and M.R. Ogiela, "Multiply information coding and hiding using fuzzy vault," Soft Computing, vol.23, no.12, pp.4357–4366, 2019.
[18] L. You and T. Wang, "A novel fuzzy vault scheme based on fingerprint and finger vein feature fusion," Soft Computing, vol.23, no.11, pp.3843–3851, 2019.
[19] P. Mihailescu, "The fuzzy vault for fingerprints is vulnerable to brute force attack," arXiv preprint arXiv:0708.2974, 2007.
[20] A. Kholmatov and B. Yanikoglu, "Realization of correlation attack against the fuzzy vault scheme," Security, Forensics, Steganography, and Watermarking of Multimedia Contents X, p.68190O, International Society for Optics and Photonics, 2008.
[21] W.J. Scheirer and T.E. Boult, "Cracking fuzzy vaults and biometric encryption," 2007 Biometrics Symposium, pp.1–6, IEEE, 2007.
[22] H.T. Poon and A. Miri, "A collusion attack on the fuzzy vault scheme," ISC Int. J. Inform. Secur, vol.1, no.1, pp.27–34, 2009.
[23] 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.
[24] T. Connie, A. Teoh, M. Goh, and D. Ngo, "Palmhashing: a novel approach for cancelable biometrics," Information processing letters, vol.93, no.1, pp.1–5, 2005.
[25] D.C.L. Ngo, A.B.J. Teoh, and A. Goh, "Biometric hash: high-confidence face recognition," IEEE Transactions on Circuits and Systems for Video Technology, vol.16, no.6, pp.771–775, 2006.
[26] 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.
[27] Y. Lee, Y. Chung, and K. Moon, "Inverse operation and preimage attack on biohashing," 2009 IEEE Workshop on Computational Intelligence in Biometrics: Theory, Algorithms, and Applications, pp.92–97, IEEE, 2009.
[28] O. Ouda, N. Tsumura, and T. Nakaguchi, "Tokenless cancelable biometrics scheme for protecting iris codes," International Conference on Pattern Recognition, pp.882–885, 2010.
[29] O. Ouda, N. Tsumura, and T. Nakaguchi, "Bioencoding: A reliable tokenless cancelable biometrics scheme for protecting iriscodes," IEICE Transactions on Information and Systems, vol.E93-D, no.7, pp.1878–1888, 2010.
[30] P. Lacharme, "Analysis of the iriscodes bioencoding scheme," Int. J. Comput. Sci. Softw. Eng. (IJCSSE 2012), vol.6, no.5, pp.315–321, 2012.
[31] O. Ouda, N. Tsumura, and T. Nakaguchi, "On the security of bioencoding based cancelable biometrics," IEICE Transactions on Information and Systems, vol.E94-D, no.9, pp.1768–1777, 2011.
[32] C. Rathgeb, F. Breitinger, C. Busch, and H. Baier, "On application of bloom filters to iris biometrics," IET Biometrics, vol.3, no.4, pp.207–218, 2014.
[33] D. Zhao, W. Luo, R. Liu, and L. Yue, "Negative iris recognition," IEEE Transactions on Dependable and Secure Computing, vol.15, no.1, pp.112–125, 2018.
[34] F. Esponda, E.S. Ackley, P. Helman, H. Jia, and S. Forrest, "Protecting data privacy through hard-to-reverse negative databases," International Journal of Information Security, vol.6, no.6, pp.403–415, 2007.
[35] F. Esponda, S. Forrest, and P. Helman, "Negative representations of information," International Journal of Information Security, vol.8, no.5, pp.331–345, 2009.
[36] F. Esponda, E.S. Ackley, S. Forrest, and P. Helman, "Online negative databases," International Conference on Artificial Immune Systems, pp.175–188, Springer, 2004.
[37] R. Liu, W. Luo, and L. Yue, "The p-hidden algorithm: Hiding single databases more deeply," Immune Computation, vol.2, no.1, pp.43–55, 2014.
[38] G. Danezis, C. Diaz, S. Faust, E. Käsper, C. Troncoso, and B. Preneel, "Efficient negative databases from cryptographic hash functions," International Conference on Information Security, pp.423–436, Springer, 2007.
[39] J. Daugman, "Evolving methods in iris recognition," IEEE International Conference on Biometrics: Theory, Applications, and Systems,(BTAS07), http://www.cse.nd.edu/BTAS_07/John_Daugman_BTAS.pdf. Accessed Sept, 2016.
[40] C. Li and J. Hu, "Attacks via record multiplicity on cancelable biometrics templates," Concurrency and Computation:Practice and Experience, vol.26, no.8, pp.1593–1605, 2013.
[41] The CASIA Iris Image Database.
[42] L. Masek and P. Kovesi, "Matlab source code for biometric identification system based on iris patterns," The School of Comput. Sci. Software Eng.,The Univ. Western Australia, 2003.

Osama Ouda received the bachelor's degree in Computer Science from College of Computer and Information Sciences, Mansoura University, Mansoura, Egypt, in 2000, and the master's degrees in Computer Science from College of Computer and Information Sciences, Ain-Shams University, Cairo, Egypt, in 2007, and the Ph.D. degree from the Graduate School of Advanced Integration Science, Chiba University, Chiba, Japan, in 2011. Since 2011, he has been an Assistant Professor with the Department of Computer Science, Mansoura University, Egypt. From 2013 to 2014, he was a Postdoctoral Research Scientist with the iPRoBe Lab, Department of Computer Science and Engineering at Michigan State University, Lansing, USA. In 2017, he was appointed as Associate Professor at the Department of Information Technology, Mansoura University, Egypt. He is currently an Associate Professor with the College of Computer and Information Sciences, Jouf University, Sakaka, Saudi Arabia. His research interests include biometric template security, pattern recognition, and machine learning. He is a member of IEEE since 2011. He is serving as a Reviewer for different journals and conferences.

Slim Chaoui received his Dipl.-Ing. degree in electrical engineering from the Technical University of Braunschweig, Germany, in 1997 and the PhD degree in telecommunications from the Technical University of Darmstadt, Germany in 2003. Currently, he is an Associate Professor with the Jouf University, College of Computer and Information Sciences, KSA. He is a member of SETIT laboratory, Sfax University, Tunisia. His research interests are focused in mobile communications, cooperative communications, network coding, lossless image compression and joint source channel coding/decoding.

Norimichi Tsumura was born in Wakayama, Japan, on April 1967. He received the B.E., M.E. and D.E. in applied physics from Osaka University in 1990, 1992 and 1995, respectively. He moved to the Department of Information and Image Sciences, Chiba University in April 1995, as an assistant professor. He is currently an associate professor in the Department of Information and Computer Sciences, Chiba University, since February 2002, and a researcher in PRESTO, Japan Science and Technology Corporation (JST), since December 2001. He was a visiting associate professor in University of Rochester from March 1999 to January 2000. He was a visiting researcher in Columbia University from April 2012 to March 2013. He is the fellow of the IS&T (Society for Imaging Science and Technology). He received the Optics Prize for Young Scientists (The Optical Society of Japan) in 1995, Applied Optics Prize for excellent research and presentation (The Japan Society of Applied Optics) in 2000, and the Charles E. Ives Award (Journal Award: IS&T) in 2002. He is interested in the color image processing, material appearance, computer vision, computer graphics and biomedical optics.