Security Analysis on Locality-Sensitive Hashing-Based Biometric Template Protection Schemes

Seunghun Paik
whitesoonguh@hanyang.ac.kr
Sunpill Kim
ksp0352@hanyang.ac.kr
Jae Hong Seo*
jaehongseo@hanyang.ac.kr

Department of Mathematics
Research Institute for Natural Sciences
Hanyang University
Seoul, Republic of Korea

Abstract

Designing an efficient and secure biometric template protection (BTP) scheme is a long-lasting challenge, and locality-sensitive hashing (LSH) is one of the promising building blocks for designing secure BTP schemes. We find that many existing LSH-based BTP schemes are designed with an identical structure, and thus we formulate such a structure as *locality-sensitive predicate* to capture its key properties. This enables us to analyze the security of a wide range of LSH-based BTPs. Based on this idea, we propose a novel method that recovers feature templates from templates protected by several LSH-based BTP schemes. In particular, the recovered templates by ours have a higher purity than those recovered by the other methods in the sense that ours recovers a close template to the original template. Recovering closer templates has several advantages over the previous methods. First, we successfully cryptanalyze a recent LSH-based BTP scheme for the first time, which was not cryptanalyzed by the previous methods. Second, by combining existing face reconstruction methods, we successfully reconstruct the face image that resembles the original face image (*e.g.*, LFW dataset). This property has not been achieved by previous attack methods. To clearly show it, we evaluate the true accept ratio (TAR) of reconstructed face images when different face images of the same identities are enrolled. Ours achieves a similar TAR (around -0.3%∼-1.4%) to the (unprotected) recognition system, but the others achieve a much lower TAR (around -84%∼-20%). To facilitate future research, our implementation code is available on github.

1 Introduction

With the advancement of deep learning algorithms, biometric authentication systems have shown great success in practice. Since these authentication systems exploit the biometric information of each individual, it is crucial to keep such sensitive information private. However, restoring the biometric information of the feature vector is not impossible because of reconstruction attacks [7, 19, 23]. These attacks demonstrated that original biometrics can be recovered from the feature vector if the adversary is permitted to oracle access to the target

* Corresponding author.

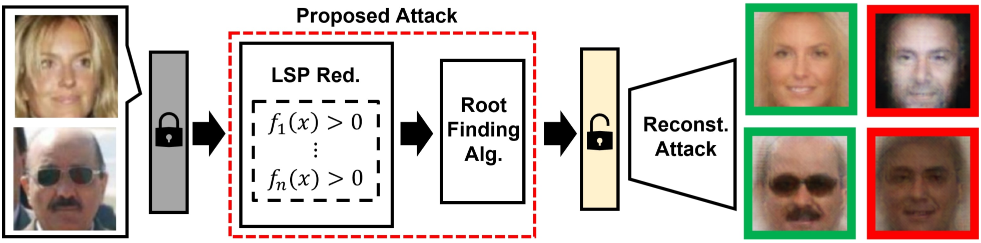

Figure 1: Overview of our irreversibility attack on the LSH-based BTPs. Our attack consists of two parts: (1) reducing the given problem into solving a system of inequalities by locality-sensitive predicate (LSP) and (2) applying the root-finding algorithm. The retrieved feature vector can be exploited as the input for reconstruction attacks, such as NbNet [19]. Images in green boxes are reconstructed by the proposed attack, and those in red boxes are by the previous attack [5]. The target BTP is [17].

feature extractor as a black box, *i.e.*, the adversary can obtain the feature vector of any image from the target feature extractor without knowing internal values such as model parameters.

The necessity of protecting biometric templates has been discussed a lot, and the notion of biometric template protection (BTP) is well established. There are three standard security notions [1] as follows: irreversibility, revocability, and unlinkability. Putting aside these security notions, in practice, it is also important to minimize the performance degradation caused by BTP. A notable design methodology for attaining these requirements is to employ locality-sensitive hashing (LSH), which is a family of hash functions that causes a collision with high probability when two input vectors in a metric space are close enough. LSH-based BTP schemes have practical benefits in terms of efficiency and accuracy.

In this study, we show that the recent LSH-based BTP constructions [14, 17] do not satisfy the irreversibility in a threat model where a polynomial number of black-box oracle accesses to the target feature extractor is allowed to the adversary. More precisely, along with the existing reconstruction attack methods, we successfully reconstruct the biometrics from the stolen template protected by LSH-based BTP, which is considerably similar to that from the enrolled identity. Furthermore, we point out that previous attacks are insufficient to reconstruct meaningful biometrics from their attack results. We observe that although previous attacks [5, 9] succeeded in finding the feature vector whose corresponding protected template is close enough to the stolen one, the recovered biometrics are quite different from those of the enrolled identity. In Section 5, we will give some examples of the biometrics that can impersonate the target system, even though they are quite different from those used to generate protected templates.

Our contribution can be summarized as follows: (1) We introduce a novel point of view for understanding LSH, called the *locality-sensitive predicate*, which yields a general attack methodology to recover the template from the protected template by LSH-based BTP schemes. (2) We point out that two previous attacks [5, 9] could not fully break irreversibility in the sense of the applicability of reconstruction attacks. We note that the proposed attack shares the same threat model as these two previous attacks. (3) By combining ours with a famous reconstruction attack NbNet [19], we verify that our attack methodology can fully break the irreversibility of three recent LSH-based BTPs: GRP-IoM, URP-IoM [14] and the BTP proposed by [17]. To the best of our knowledge, ours is the first attack on breaking the irreversibility of the latter one. The overview of the proposed attack is illustrated in Figure 1.

2 Related Works

LSH-based BTPs. One of the most remarkable LSH-based BTPs is the index-of-maximum (IoM) hashing proposed by [14]. The author of [14] proposed two LSH constructions, GRP-IoM and URP-IoM, based on their ranking-based strategy: transform the given input vector with internal randomness and output the index of the highest entry of the transformed vector. By using each of them, they designed two LSH-based BTPs and claimed that their schemes satisfy irreversibility, showing better performance (EER) than the previous BTP constructions. After this, several variants using the ranking-based strategy were proposed [2, 16, 18].

Recently, another type of LSH [17] has been proposed. They first showed that random projection-based LSH schemes were insecure against genetic algorithms and noted that this vulnerability stems from the structural limitations of previous LSH schemes. To circumvent this situation, they proposed a novel type of construction by exploiting the binary expression of integers. They argued that, owing to their binary representation strategy, the proposed BTP is secure against genetic algorithm-based attacks, showing much less performance degradation than in previous studies.

Previous Attack Methodologies. Despite such numerous LSH-based BTP proposals, several attacks have already been reported. They attempted to find a pre-image of the stolen protected template and extract the feature vector, and we classify each of them into *optimization-based approaches* and *genetic algorithm-based approaches*. In optimization-based attacks, a notable methodology was presented by [9]. The authors proposed an efficient attack targeting IoM-based BTP schemes. Their key idea is to construct a system of linear inequalities from the stolen template. This system of inequalities can be solved efficiently, so their algorithm can impersonate the target BTP within a few seconds. However, due to the existence of BTP proposals that cannot be converted to a linear inequality, their method cannot be well generalized to other schemes. Recently, a variant [8] of this attack was proposed, but this method also suffers from the same problem.

On the other hand, there were several attacks using the genetic algorithm [5, 6, 17, 24]. They attempted to impersonate a target BTP scheme by carefully designed genetic algorithms. Unlike optimization-based approaches, genetic algorithm-based approaches do not rely on the structure of the target LSH-based protection scheme. In particular, [6] reported that this genetic algorithm-based attack can be applicable to other BTP constructions, *e.g.*, Bloom filter-based approach [22]. However, this type of attack requires much more time compared to optimization-based approaches. Furthermore, we observe that the retrieved feature vector from the genetic algorithm is often irrelevant to the original one, although such a found feature vector that can be used for impersonation. In the right hand side of the Table 1, we provide a summary of each attack and a comparison with our proposed attack algorithm.

3 LSH-based BTPs

The BTP is an extension of the usual biometric authentication system, which has an additional process called *transformation* T to protect the biometric templates generated by the feature extractor Ext. Basically, it is expected that the feature vectors extracted from the same identity are close enough. There are three security notions for BTPs: *irreversibility, revocability* and *unlinkability*. Irreversibility means that it should be computationally infeasible to recover the original biometrics from the protected template.

LSH-based BTP	Projection	Many-to-One	Additional Transf.
BioHashing [13]	Rand. Proj.	Sign	N/A
GRP-IoM [14]	Rand. Proj.	Argmax	N/A
URP-IoM [14]	Rand. Perm.	Argmax	Hadamard Prod.
IFO [16]	Rand. Perm.	Argmax	Hadamard Prod.
ABH [17]	Rand. Proj.	Sign	Binary Repr.
IMM [18]	Rand. Proj.	Argmax/min	N/A

Attack	Type	Target	Gen	Rec
[9]	Opt	[14]	X	O
[8]		[13]	X	O
[5]	GA	[14], [22]	O	X
[6]		[13], [14], [16],[22]	O	X
[22]		[14], [13]	O	X
[17]		[14], [13]	O	X
Ours	Opt	[14], [17]	O	O

Table 1: Left: Analysis of each known LSH-based BTP construction with respect to the type of projection, many-to-one function, and additional transform. A detailed explanation for the analysis of LSH will be presented in Section 4. Right: Comparison between the proposed algorithm and previous attack methods in terms of generalization (Gen) and recovering similar images to the original images (Rec). “Opt" and “GA" stand for optimization-based and genetic algorithm-based attacks, respectively.

In order to satisfy the *revocability*, the ability to revoke the stolen template without security loss, the transformation T takes an additional randomness R. This randomness is enclosed in the protected template. That is, the target system stores a pair $(T(x,R),R)$, where x is the biometric template extracted from the given biometrics.

Furthermore, in many concrete constructions, we expect two additional properties to be satisfied. First, for the sake of *irreversibility*, T should be difficult to invert, even when randomness R is disclosed to the public. Although some BTP schemes [2, 12, 13, 21] utilize the randomness R as a user-specific token, we slightly modify them by treating R as public data because of the aforementioned reason. Second, for the sake of better accuracy, the transformation T should almost preserve the distance for each randomness R such that close vectors map to close vectors. This is because the authentication process with BTP is conducted by determining whether the matching score between the stored protected template and the output of T from the queried biometrics and the stored randomness is close enough or not.

One methodology of designing T that satisfies the second property is to utilize LSH, which is a collection of “similarity preserving” functions. Every function in the LSH guarantees a certain collision probability on outputs if the given two inputs are sufficiently close. Since its main functionality coincides with the desirable properties of BTP schemes, several efforts have been made to design effective LSH-based BTPs. Due to space constraints, formal definitions are relegated to the supplementary material.

4 Proposed Method

We now introduce our key observations from known LSH-based BTP constructions. From them, we propose a novel point of view on analyzing LSH with a component called a locality-sensitive predicate (LSP). This component gives a general method to reformulate the given LSH, thus leading to vulnerability against our proposed attack methodology.

4.1 Understanding LSH as a Composition of Predicates

From the investigation of LSH-based BTP proposals, we observe that their construction methodology shares the following two steps: (random) projection and many-to-one transform. When calculating the hashed value from the given input, they first project the given input into another space, utilizing the randomness. After this, they apply the many-to-one function to facilitate the collision. Additional transforms can be applied if necessary. In the left hand side of the Table 1, we represent each LSH-based BTP proposal in our observation.

We figure out that each entry of the hashed value may leak some information about the input vector by using the structure of the many-to-one transformation. For example, by observing the output of the sign function, we can determine whether the angle between the input vector and the vectors exploited in random projection is less than 90° or not. In addition, the argmax function leaks information about the index whose corresponding random vector is closest to the input feature vector. From this, we can obtain a bundle of inequalities with respect to the input vector, whose solution space is equivalent to the pre-image of the many-to-one transform. We note that the construction of such inequalities can be well harmonized with an additional transform such as the Hadamard product or sinusoidal function. We call each inequality a locality-sensitive predicate (LSP). Due to space constraints, we present concrete examples of the proposed reformulation in the supplementary material.

Algorithm 1 Proposed Attack

```
Require: Target LSH $T$, Compromised Template $(\mathbf{s}, \mathbf{R})$
  Convert $T$ as LSPs $f_1^T, f_2^T, \cdots, f_N^T$
  Set $F(\cdot) \leftarrow (f_1^T(\cdot), \dots, f_N^T(\cdot))$
  Set $G(\cdot) \leftarrow \left|\left|F(\cdot) + |F(\cdot)|\right|\right|_1$
  Solve the equation $G(x) = 0$ using Root-Finding Alg.
  Return the candidate feature vector $\hat{\mathbf{x}}$
```

4.2 Proposed Attack Algorithm

Before providing the detailed attack method, we first specify the adversary's capabilities. We assume that (1) the adversary can obtain a protected template from the target BTP's database, (2) the internal structure of the transformation algorithm is publicly known, and (3) the adversary can access the target system by approaching it as a black-box system; the adversary can obtain a protected template from the queried biometrics. In order to exclude unrealistic adversaries, we restrict the number of queries made by the adversary to a polynomial of the dimension of the protected template. Under these capabilities, the main goal of the adversary is to reconstruct the biometrics used for generating the stolen protected template, *i.e.*, to break the irreversibility of the given system. We note that the adversary can utilize the reconstruction attacks to recover the biometrics from the pre-image of the transform T.

The strategy of the adversary can be summarized as follows: First, the adversary reformulates the given LSH scheme into a mathematically equivalent form of LSPs. Then, the adversary converts these LSPs into a system of inequalities. By reducing these inequalities to the root-finding problem, the adversary can obtain a pre-image of the given template from numerical root-finding algorithms. Finally, the adversary performs a reconstruction attack to obtain the biometrics that might resemble the biometrics from the stolen template's identity.

More precisely, let us denote the transformation algorithm of the target BTP as T, and suppose that the protected (s, R) is compromised by the adversary. In this setting, the goal of the adversary is to find a biometric template $\hat{x}$ satisfying $T(\hat{x}, R) = s$. Under our assumption, the adversary can reformulate T as LSPs $f_1^T, \dots, f_N^T$ for some positive integer N. Finding the pre-image of s is equivalent to solving the system of inequalities $f_i^T(x) > 0$, $\forall i = 1, \dots, N$. For simplicity, we define a function F such that $F(x) = (f_1^T(x), \dots, f_N^T(x))$. Finally, let us denote $G(x) = \left|\left|F(x) - |F(x)|\right|\right|_1$, where $||\cdot||_1$ denotes the L_1 norm, and $|F(x)|$ is the vector obtained by taking the absolute value for all entries of $F(x)$. Then we can observe that

$G(x) = 0$ if and only if every entry of $F(x)$ is positive. Therefore, the adversary can apply an appropriate root-finding algorithm, such as Newton's method, to obtain a candidate feature vector $\tilde{x}$. The formal description of our attack algorithm is provided in Algorithm 1.

5 Experimental Analysis

We now demonstrate our proposed attack method against other known LSH-based BTPs, showing that they are *reversible* within the system parameters claimed by their authors. All experiments were conducted in an environment with a CPU i7-11700k and 64GB RAM. We implement all algorithms in PyTorch [20] and the source code is publicly available at github.

5.1 Experimental Setting

Throughout this paper, we use the true accept ratio (TAR) and the false accept ratio (FAR) for the performance evaluation of biometric authentication systems (with protection). Motivated by [19], we consider two scenarios, Type-1 and Type-2 attacks, in order to evaluate each attack method with respect to impersonation and irreversibility. In the Type-1 attack, the adversary attempts to impersonate the target LSH-based BTP by reconstructing the biometric from the compromised protected template. On the other hand, the Type-2 attack considers the scenario that the adversary attempts to reconstruct the biometrics similar to that from the enrolled identity. In contrast to Type-1 attack that aims to impersonate the target system, the adversary is asked to find the biometrics that can be exploited to impersonate other (potentially unprotected) systems, which corresponds to breaking the irreversibility of the given LSH-based BTP. For this, we define the Type-2 attack success rate (ASR_2) by evaluating the probability that the adversary succeeds in impersonating the unprotected biometric authentication system via retrieved biometrics from a stolen protected template.

For the face image dataset, we use the LFW [11] benchmark dataset, which is widely used to evaluate the performance of face recognition systems. For the *extraction algorithm*, we used a pre-trained ArcFace [3] ResNet50 model trained by the MS1M-RetinaFace dataset [4, 10], and the corresponding model achieved TAR@FAR=99.70%@1e-3 on LFW benchmark, with a threshold of $0.2082 \approx 77.98°$.

On the implementation of LSH-based BTPs, we selected three well-known proposals: GRP-IoM, URP-IoM [14], and ABH[1] [17], as *transformations*. Because [14, 17] has no official source code, we implemented them from the pseudocode provided by their papers and evaluated the attack success rate. For each BTP, we select the parameter setting that shows the best benchmark performance as the target scheme, with the verification thresholds $\tau_{\mathrm{GRP}} = 0.137$, $\tau_{\mathrm{URP}} = 0.038$ and $\tau_{\mathrm{ABH}} = 0.7$ for each BTP, respectively. The precise parameter settings are given in the supplementary material.

5.2 Attack Results

Prior to evaluate the proposed attack method, we first implement a variant of NbNet [19] suggested by [15][2]. Using this, our experiment is conducted by the following procedure: First, we load pairs of images with the same identity from LFW. For each pair, we choose one of the images and create a protected template using the target LSH-based BTP. Now, we

[1] In fact, the authors of [17] did not name their BTP proposal, so we call it an advanced BioHashing (ABH).

[2] For more detailed information, we recommend the reader check our implementation code in github.

Method	Genetic Algorithm [5]			Ghammam et al. [9]	Proposed Algorithm		
Target LSH	ABH	GRP-IoM	URP-IoM	URP-IoM	ABH	GRP-IoM	URP-IoM
TAR@FAR	99.70%@1e-3	99.63%@3e-4	99.03%@7e-4	99.03%@7e-4	99.70%@1e-3	99.63%@3e-4	99.03%@7e-4
ASR$_1$	2%	100%	75%	0.4%	**100%**	**100%**	**98.7%**
ASR$_2$	29%	79%	15%	0.07%	**99.46%**	**98.90%**	**97.63%**
Average Angle (Type-1)	73.50°	65.02°	78.26°	88.87°	31.54°	42.47°	43.30°
Average Angle (Type-2)	79.82°	73.85°	82.24°	88.11°	54.84°	60.35°	60.92°
Average Matching Score	0.02	0.274	0.06	0.011	0.999	0.408	0.325

Table 2: Comparison of the proposed algorithm and the previous methods, a genetic algorithm-based method [5] and an optimization-based method [9], with respect to the attack success rates of Type-1,2 attacks against each LSH-based BTP.

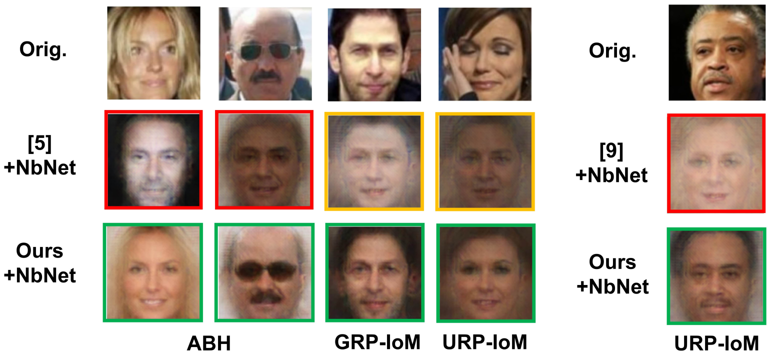

Figure 2: Comparison of the proposed algorithm and the previous methods, a genetic algorithm-based method [5] and an optimization-based method [9], with respect to reconstructing facial images from the templates protected by each LSH-based BTP. Green box: success on both Type-1 and Type-2 attacks. Yellow box: success on Type-1 attack only. Red box: failure on both types of attacks.

apply the proposed attack to the protected template and obtain a reconstructed feature vector. Finally, we recover a facial image using the pretrained NbNet.

In order to compare our method against previously reported ones, we implemented two previous methods, a genetic algorithm-based method [5] and an optimization-based method [9], by following pseudocodes in their original papers. We note that due to the excessive computational cost of the genetic algorithm, we only tested 100 pairs of images randomly sampled from the LFW dataset on evaluating [5]. In addition, we note that our method can be understood as a generalization of [9] by considering the system of linear constraints made by their attack as our LSPs. Thus, we only conducted URP-IoM on evaluating theirs, which cannot be expressed as a system of linear constraints because of the Hadamard product as an additional transform. Detailed parameters for conducting each attack method, including the choice of root-finding algorithm in ours, are provided in the supplementary material.

In Table 2, we report the Type-1,2 attack success rates of each method, along with the average angles between the enrolled biometrics and the reconstructed ones in each scenario. Furthermore, in the Type-1 attack scenario, we provide the average matching score between the stolen protected template and the template from the reconstructed biometrics under the same randomness. Our attack achieves 100%, 100% and 98.70% Type-1 attack success rates and 99.46%, 98.90%, and 97.63% Type-2 attack success rates for ABH, GRP-IoM, and URP-IoM, respectively. In contrast to ours, attack success rates measured by each previous method are not as high as ours. In addition, as shown in the last three rows in this table,

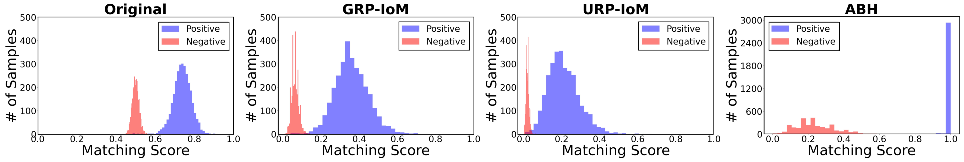

Figure 3: The distribution of the matching score between each positive and negative pair after applying each LSH-based BTP, respectively. The “Original" refers to the result of the unprotected biometric authentication system. For comparison, we scale each matching score so that it lies in the range [0, 1].

the feature vector from the reconstructed biometrics is close enough to the enrolled one, surpassing the verification threshold by a large amount for each attack scenario. On the other hand, although previous methods show the ASR_2 better than FAR, they tend to recover the biometrics whose angular distance between enrolled one lies near the verification threshold.

To investigate whether our attack method actually led to the leakage of biometric information, we visualize the reconstructed facial images from the protected template in Figure 2. The reconstructed images from ours accurately capture the characteristics of the corresponding image so that they can be exploited to succeed in both Type-1 and Type-2 attacks. On the other hand, although retrieved images from [5] can be used to impersonate the target system, they are not similar enough to break the irreversibility (Type-2 attack). In addition, those from [9] failed to succeed in both types of attacks.

5.3 Analysis on the Previous Methods

In this section, we give our analyses on the reasons why previous methods, genetic algorithm-based method [5] and optimization-based method [9], failed to succeed in the Type-2 attack.

Genetic Algorithm-based Method We focus on the relationship between the matching score from hashed values and the angle from feature vectors. More precisely, we compare the angle distribution of positive and negative pairs in LFW and the matching scores of the hashed value for each LSH-based BTP. We illustrate each distribution on Figure 3. We can figure out that after applying each IoM-type LSH, the overall distribution shifts to the left in general, whereas ABH makes each distribution much more distinctive, so almost all of the positive pairs have the matching score 1.[3]

This result can be interpreted as follows: Although the distributions of matching score between positive and negative pairs are well distinguished, the relation between matching score in protected templates and identities bound to each template is weaker than that between the angle of feature vectors and identities. From this, one can infer that there might be a significant difference in the difficulty of the impersonation on IoM-type BTP and of extracting the biometric information. On the other hand, we observe that the genetic algorithm fails to find the pre-image of the given protected template. This is because genetic algorithm-based attacks do not utilize the structure of the target LSH rather than solving a system of constraints as in optimization-based methods. Therefore, although genetic algorithm-based attacks can be applied to various types of LSH-based BTP proposals, they are insufficient for the leakage of biometric information because of the aforementioned reasons.

Optimization-based Method One may notice that the result from our experiment is quite different from that reported by [9]. This is because of the differences in the experimental

[3] This is the reason why ABH is believed to be immune against genetic algorithm-based attacks, as [17] argued.

setting: in contrast to [9], the adversary of our attack scenario is permitted to exploit only one template to recover the pre-image. Furthermore, our adversary is asked to reconstruct the biometrics from the pre-image. Thus, each evaluation in our experiment includes the use of pre-trained NbNet, whereas the adversary in [9] was only requested to retrieve the feature vector. We note that the algorithm used in our experiment also finds the vector well that can impersonate the target URP-IoM system if it itself can be inserted directly into the system. More precisely, the average matching score between the protected template and the found pre-image is 0.462, which is much higher than the verification threshold $\tau_{URP} = 0.038$ in our parameter setting.

Putting aside these differences, we further analyze the reason why [9] does not work well in our setting. To this end, we investigate the trick used to construct a linear constraint system from URP-IoM. Since URP-IoM utilizes Hadamard product before applying the argmax function, when we construct a constraints system, we always encounter constraints of the form $\left(\prod_{i\in I} x_i\right) - \left(\prod_{j\in J} x_j\right) \leq 0$, where I, J are index sets, and x_k is a variable for all $k \in I \cup J$. To convert these non-linear constraints into linear ones, their key idea is to make use of the property of the logarithm. More precisely, if we apply the logarithm to the above constraint, we obtain $\sum_{i\in I} \log x_i \leq \sum_{j\in J} \log x_j$. Therefore, by taking $z_k = \log x_k$ for all $k \in I \cup J$, we obtain the linear constraint $\sum_{i\in I} z_i \leq \sum_{j\in J} z_j$, thus the given system can be solved by linear programming.

However, this type of conversion severely harms the size of the solution space. This is because the logarithm function is only defined over positive real numbers, so every component of the obtained vector after solving the converted system must be positive. From this, one can observe that the solution space drastically shrinks as the dimension increases. Precisely, for d-dimensional Euclidean space, the corresponding solution space is only 2^{-d} of the overall space. Thus, it can be expected that although the pre-image of the hashed value via URP-IoM can be calculated, this pre-image will be far apart from the feature vector of enrolled biometrics. We conclude that this is the main reason that [9] shows relatively poor performance on breaking URP-IoM.

6 Conclusion

Satisfying irreversibility is an important security goal for BTP because of practical threats such as [7, 19], which restore users' biometrics from unprotected biometric templates. In this research, we present a general irreversibility attack methodology against a notable BTP approach called LSH-based BTPs. We consider the adversary that can compromise one protected template in the database of the target system, with polynomial numbers of oracle accesses to the target feature extractor in the black-box model. Although there are several approaches for breaking LSH-based BTPs: optimization-based [9] and genetic algorithm-based [5, 6, 17, 24] under the same scenario, we point out that they cannot be generalized well to other LSHs or are insufficient to recover face images although their attack can be used to impersonate the target system. We validate our methodology by applying it to previous famous LSH-based BTPs [14, 17], showing that these BTPs are *reversible* in the sense that we can successfully restore face images that resemble the original ones from their protected templates. Our results suggest that major repairs on LSH-based BTPs are necessary.

Acknowledgement. This work was supported by Samsung Science and Technology Foundation (SSTF) grant funded by Samsung (Grant No. 20230000002564, Secure Multi-party Approximate Computation).

References

[1] Christoph Busch. ISO 24745-biometric template protection. In *The first International Biometric Performance Testing Conference, Match*, 2010.

[2] Jiandong Cui and Andrew Beng Jin Teoh. Deep index-of-maximum hashing for face template protection. In *2020 5th International Conference on Computer and Communication Systems (ICCCS)*, pages 413–418. IEEE, 2020.

[3] Jiankang Deng, Jia Guo, Niannan Xue, and Stefanos Zafeiriou. Arcface: Additive angular margin loss for deep face recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4690–4699, 2019.

[4] Jiankang Deng, Jia Guo, Yuxiang Zhou, Jinke Yu, Irene Kotsia, and Stefanos Zafeiriou. Retinaface: Single-stage dense face localisation in the wild. *arXiv preprint arXiv:1905.00641*, 2019.

[5] Xingbo Dong, Zhe Jin, and Andrew Teoh Beng Jin. A genetic algorithm enabled similarity-based attack on cancellable biometrics. In *2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS)*, pages 1–8. IEEE, 2019.

[6] Xingbo Dong, Jaewoo Park, Zhe Jin, Andrew Beng Jin Teoh, Massimo Tistarelli, and KokSheik Wong. On the risk of cancelable biometrics. *arXiv preprint arXiv:1910.07770*, 2019.

[7] Chi Nhan Duong, Thanh-Dat Truong, Khoa Luu, Kha Gia Quach, Hung Bui, and Kaushik Roy. Vec2face: Unveil human faces from their blackbox features in face recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6132–6141, 2020.

[8] Axel Durbet, Paul-Marie Grollemund, Pascal Lafourcade, Denis Migdal, and Kevin Thiry-Atighehchi. Authentication attacks on projection-based cancelable biometric schemes. In *19th International Conference on Security and Cryptography*, pages 568–573. SCITEPRESS-Science and Technology Publications, 2022.

[9] Loubna Ghammam, Koray Karabina, Patrick Lacharme, and Kevin Thiry-Atighehchi. A cryptanalysis of two cancelable biometric schemes based on index-of-max hashing. *IEEE Transactions on Information Forensics and Security*, 15:2869–2880, 2020.

[10] Yandong Guo, Lei Zhang, Yuxiao Hu, Xiaodong He, and Jianfeng Gao. Ms-celeb-1m: A dataset and benchmark for large-scale face recognition. In *European conference on computer vision*, pages 87–102. Springer, 2016.

[11] Gary B Huang, Marwan Mattar, Tamara Berg, and Eric Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. In *Workshop on faces in'Real-Life'Images: detection, alignment, and recognition*, 2008.

[12] Yubing Jiang, Peisong Shen, Li Zeng, Xiaojie Zhu, Di Jiang, and Chi Chen. Cancelable biometric schemes for euclidean metric and cosine metric. *Cybersecurity*, 6(1):1–20, 2023.

[13] Andrew Teoh Beng Jin, David Ngo Chek Ling, and Alwyn Goh. Biohashing: two factor authentication featuring fingerprint data and tokenised random number. *Pattern recognition*, 37(11):2245–2255, 2004.

[14] Zhe Jin, Jung Yeon Hwang, Yen-Lung Lai, Soohyung Kim, and Andrew Beng Jin Teoh. Ranking-based locality sensitive hashing-enabled cancelable biometrics: Index-of-max hashing. *IEEE Transactions on Information Forensics and Security*, 13(2):393–407, 2017.

[15] Sunpill Kim, Yunseong Jeong, Jinsu Kim, Jungkon Kim, Hyung Tae Lee, and Jae Hong Seo. Ironmask: Modular architecture for protecting deep face template. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 16125–16134, 2021.

[16] Yen-Lung Lai, Zhe Jin, Andrew Beng Jin Teoh, Bok-Min Goi, Wun-She Yap, Tong-Yuen Chai, and Christian Rathgeb. Cancellable iris template generation based on indexing-first-one hashing. *Pattern Recognition*, 64:105–117, 2017.

[17] Yenlung Lai, Zhe Jin, KokSheik Wong, and Massimo Tistarelli. Efficient known-sample attack for distance-preserving hashing biometric template protection schemes. *IEEE Transactions on Information Forensics and Security*, 16:3170–3185, 2021.

[18] Yuxing Li, Liaojun Pang, Heng Zhao, Zhicheng Cao, Eryun Liu, and Jie Tian. Indexing-min–max hashing: Relaxing the security–performance tradeoff for cancelable fingerprint templates. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 52(10):6314–6325, 2022.

[19] Guangcan Mai, Kai Cao, Pong C Yuen, and Anil K Jain. On the reconstruction of face images from deep face templates. *IEEE transactions on pattern analysis and machine intelligence*, 41(5):1188–1202, 2018.

[20] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.

[21] João Ribeiro Pinto, Miguel V Correia, and Jaime S Cardoso. Secure triplet loss: Achieving cancelability and non-linkability in end-to-end deep biometrics. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 3(2):180–189, 2020.

[22] Christian Rathgeb, Frank Breitinger, and Christoph Busch. Alignment-free cancelable iris biometric templates based on adaptive bloom filters. In *2013 international conference on biometrics (ICB)*, pages 1–8. IEEE, 2013.

[23] Thanh-Dat Truong, Chi Nhan Duong, Ngan Le, Marios Savvides, and Khoa Luu. Vec2face-v2: Unveil human faces from their blackbox features via attention-based network in face recognition. *arXiv preprint arXiv:2209.04920*, 2022.

[24] Hanrui Wang, Xingbo Dong, Zhe Jin, Andrew Beng Jin Teoh, and Massimo Tistarelli. Interpretable security analysis of cancellable biometrics using constrained-optimized similarity-based attack. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 70–77, 2021.