# Authentication Attacks on Projection-based Cancelable Biometric Schemes

DURBET Axel[1], GROLLEMUND Paul-Marie[2], LAFOURCADE Pascal[1], MIGDAL Denis[1] and THIRY-ATIGHEHCHI[1]

[1] *Université Clermont-Auvergne, CNRS, Mines de Saint-Étienne, LIMOS, France*

[2] *Université Clermont-Auvergne, CNRS, LMBP, France*

Keywords: Cancelable biometrics; Local-Sensitive Hash; Sobel filter; Reversibility attacks; Biohash

Abstract: Cancelable biometric schemes aim at generating secure biometric templates by combining user specific tokens, such as password, stored secret or salt, along with biometric data. This type of transformation is constructed as a composition of a biometric transformation with a feature extraction algorithm. The security requirements of cancelable biometric schemes concern the irreversibility, unlinkability and revocability of templates, without losing in accuracy of comparison. While several schemes were recently attacked regarding these requirements, full reversibility of such a composition in order to produce colliding biometric characteristics, and specifically presentation attacks, were never demonstrated to the best of our knowledge. In this paper, we formalize these attacks for a traditional cancelable scheme with the help of integer linear programming (ILP) and quadratically constrained quadratic programming (QCQP). Solving these optimization problems allows an adversary to slightly alter its fingerprint image in order to impersonate any individual. Moreover, in an even more severe scenario, it is possible to simultaneously impersonate several individuals.

## 1 Introduction

Biometric authentication is more and more used in daily life and are commonly integrated on many smart objects and devices, *e.g.*, computer, smartphone, USB drive, passport. Since biometrics is more convenient and quicker to use, and biometric characteristics cannot be lost or forgotten, biometric authentication solutions are in general preferred over their password or physical token counterparts. Despite their many advantages, biometric solutions are not exempt from vulnerabilities. As biometric-based technologies are deployed at a larger scale, centralized biometric databases and devices become natural targets in cyber attacks. These cyber attacks have the potential to be harmful on the long term if they lead to the theft of biometric data. Therfore, a biometric data may actually be vulnerable to impersonation attacks and privacy leakage.

Several criteria essential to biometric authentication systems have been identified in ISO/IEC 24745 (ISO, 2011) and ISO/IEC 30136 (ISO, 2018): Irreversibility, unlinkability, revocability and performance preservation of templates.

- *Irreversibility* prevents from finding the original person's biometric data from the protected template.
- *Unlinkability* prevents cross-matching attacks or, in other words, the linkability between two digital identities, *i.e.*, two biometric templates.
- *Revocability* requires the scheme to be able to generate new protected templates in case of compromission of the biometric database.
- The last criteria, *performance preservation*, stipulates that recognition accuracy of protected templates should not be degraded compared to the original data.

Fulfilling this set of criteria is now necessary to comply with the *privacy* principles of the GDPR.

Faced with the mentioned vulnerabilities and requirements, the community has proposed primitives dedicated to biometrics, so called biometric template protection (BTP) schemes. Examples of such primitives include cancelable biometrics (see (Jin et al., 2004; Sutcu et al., 2005)), biometric cryptosystems (*e.g.*, fuzzy vault (Juels, 2006), fuzzy extractors (Dodis et al., 2004)), and hybrid biometrics (Bringer et al., 2008; Jain and Nandakumar, 2012). In this paper, we focus on cancelable biometrics (CB) which is an example of BTP scheme claimed to meet the four criterias. For more details on BTP schemes, the reader is referred to two surveys (Nandakumar and Jain, 2015) and (Natgunanathan et al., 2016). In CB, a biometric template is computed through a process where the inputs are biometric data (*e.g.*, biometric image) of a user and a user specific token (*e.g.*, a random key, seed, salt, or password). A CB scheme generally consists of a sequence of processes (an extraction of features followed by a parameterized transformation) that produces the biometric templates, and a matcher to generate a matching score between the templates. With a CB scheme, templates can be revoked, changed, and renewed by changing user specific tokens. Even though user tokens in CB may be considered as secret, the security of a two-factor authentication system should not be reduced to a single factor. Cryptanalysis of CB schemes with strong adversarial models commonly assume that the attacker knows both the biometric template and token of the user. This assumption is plausible in practice because a user token may have low entropy (*e.g.*, a weak password), or it may just have been compromised by an attacker. This stolen-key scenario is also known as the stolen-token scenario (Teoh et al., 2008).

Ratha *et al.* (Ratha et al., 2001) were the first to introduce CB in the case of face recognition. Since then, several CB schemes have been proposed, including the popular Biohashing algorithm (Jin et al., 2004) applied on many modalities such as fingerprints, face, and iris. CB schemes offer several advantages such as efficient implementation, high matching accuracy, and revocability. However, several attacks on a variety of CB schemes have been proposed: attacks against privacy by approximating feature vectors or linking several templates of an individual, and authentication attacks by elevating the false acceptance rate (FAR). We refer the reader to (Nagar et al., 2010; Topcu et al., 2016) for attacks on biohashing type schemes, (Quan et al., 2008; Li and Hu, 2014) for attacks using the Attack via Record Multiplicity (ARM) technique, (Lacharme et al., 2013; Dong et al., 2019a) for attacks using genetic algorithms, as well as attacks using constrained programming on CB schemes built upon ranking based hashing (Ghammam et al., 2020).

Authentication attacks using genetic algorithms have been proposed in (Dong et al., 2019a; Rozsa et al., 2015). Their objective is to find the right parameters for generating fingerprint images in order to elevate FAR rates. In the case of the fingerprint modality, strategies making use of both hill climbing attacks and genetic algorithms have also been proposed in (Dong et al., 2019b; Wang et al., 2021).

**Contributions.** In this paper, we propose reversibility attacks against some projection-based CB schemes, such as the BioHashing (Jin et al., 2004). The particularity of our attacks, as opposed to previous works, is that we reverse the complete sequence of treatments including the *feature extraction* algorithm. This allows us to construct impostor fingerprint images, thus enabling authentication (or presentation) attacks. In our authentication attacks, an adversary, who already has the knowledge of a user's specific token and has at least one fingerprint template of the same user, tries to alter his own fingerprint image such that the adversary can now use its own altered biometrics and the stolen token to be falsely authenticated as a legitimate user. The considered CB schemes are built upon uniform random projection (URP) and a feature extractor such as Sobel or Gabor filter. To perform our attacks, we use Integer Linear Programming (ILP) as well as quadratically constrained quadratic programming (QCQP). Constrained optimization with linear programs has been previously used in the cryptanalysis of other schemes; see (Ghammam et al., 2020; Topcu et al., 2016).

We can state our results as follows:

**1) Simple authentication attacks.** A complete reversal methodology of some projection-based CB schemes, including the BioHash algorithm, is proposed. The main ideas are to solve an integer linear program and a quadratically constrained quadratic program to reverse both the projection and the feature extraction. The solution provided by a solver (*e.g.*, Gurobi) is a fingerprint image of the attacker whose the amount of changes is minimized. Practical resolutions are provided for tiny synthetic images.

**2) One fingerprint image for several impersonations.** The first attack is extended to produce a fingerprint image that impersonates the identity of several users. Our formalized constrained problems and experimentations on tiny synthetic images show that an adversary can alter its own fingerprint image to be authenticated as any of several legitimate users. To reach this objective, two different attacks are proposed:

- The first strategy for the attacker is to collect the pairs of (token, template) of the target users to enlarge the set of constraints of a QCQP program. The solution sought is a single altered fingerprint image of the attacker such that, when combined with the distinct stolen tokens, the generated templates match exactly the stolen templates of the respective users. Impersonating a large number of target users under this approach imposes a due acceptance of a larger number of changes in the altered fingerprint image of the attacker.
- The second proposed strategy does not require the knowledge of the tokens and consists in generat-

ing a template which is an average (barycentric) template of the target users. Then, the attacker formalizes a set of constraints using this template and her token. She solves it to find a fingerprint image as close as possible to her own. If the target templates lie in a ball of radius two times the decision threshold (in the template space), her altered fingerprint image enables an authentication attack for multiple users. In other words, her altered image is a "master print" for these target users.

**Outline.** Some background information and the adversarial models are presented in Section 2. Section 3 provides our simple authentication attacks. Section 4 introduces an attack not relying on the knowledge of the passwords. Then, in Section 5, it is shown how to impersonate several users with different passwords. Finally, experimental evaluations and future works are discussed in Section 6 and Section 7 respectively.

# 2 Background

Cancelable biometric schemes generate secure biometric templates by combining user specific tokens such as password with his biometric data such as fingerprint. The goal is to create templates meeting the four aforementioned criteria, *i.e.*, irreversible, unlinkable, and revocable templates, with high accuracy of comparison. Biometric templates in CB schemes are constructed in two steps: (*i*) *Feature extraction*: A feature vector is derived from a biometric image; (*ii*) *Transformation*: A user specific token is used to transform the user's feature vector to a template.

In the following, we let $(\mathcal{M}_I, D_I)$, $(\mathcal{M}_F, D_F)$ and $(\mathcal{M}_T, D_T)$ be three metric spaces, where $\mathcal{M}_I$, $\mathcal{M}_F$ and $\mathcal{M}_T$ represent the fingerprint image space, the feature space and the template space, respectively; and $D_I$, $D_F$ and $D_T$ are the respective distance functions. Note that $D_I$ and $D_F$ are instantiated with the Euclidean distance, while $D_T$ is instantiated with the Hamming distance.

## 2.1 Feature Extraction with Sobel Filtering

Let $\mathcal{U}$ be the set of users of the biometric system. We identify a user with its biometric characteristic, and define a function $\mathcal{BC}(\cdot)$ that takes a biometric characteristic $usr \in \mathcal{U}$ as input, and outputs a digital representation of biometric data $I$; for instance, the scan image of a fingerprint. Note that for two different computations of $I = \mathcal{BC}(usr)$ and $I' = \mathcal{BC}(usr)$ (*e.g.*, at different times, or different devices), we may have $I \neq I'$ due to the inherent noise in the measurement of biometric data.

**Definition 2.1.** *A biometric feature extraction scheme is a pair of deterministic polynomial time algorithms* $\Pi := (E, V)$, *where:*

- $E$ *is the feature extractor of the system, that takes biometric data* $I$ *as input, and returns a feature vector* $F \in \mathcal{M}_F$.
- $V$ *is the verifier of the system, that takes two feature vectors* $F = E(I)$, $F' = E(I')$, *and a threshold* $\tau$ *as input, and returns True if* $D(F, F') \leq \tau$, *and returns False if* $D(F, F') > \tau$.

**Sobel Filter.** An example of feature extraction is the Sobel filtering (Vincent and Folorunso, 2009). Sobel Filter is usually used for edge detection. The resulting image is obtained by computing two convolutions given by the following matrices:

$$G_1 = \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix} \text{ and } G_2 = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}.$$

We denote by $*$ the operator of convolution and by $I$ the matrix of the image in shades of gray. Note that pixels at the edges of the image are ignored and their values are set to 0 in the corresponding matrix $I$. The horizontal and vertical gradients, $G_x$ and $G_y$, are computed as follows $G_x = G_1 * I$ and $G_y = G_2 * I$ with $*$ the sign of convolution 2.2 from (Stockman and Shapiro, 2001). Then, the matrix of the output image $S$ is computed as $\|G_x + G_y\|_2$ where $\|\cdot\|_2$ denotes the Euclidean norm. However, the norm does not apply in the usual way. In fact, in this case it applies coordinate by coordinate. For example, the first coordinate of $S$ is $S_{1,1} = \sqrt[2]{G^2_{x_{1,1}} + G^2_{y_{1,1}}}$.

**Definition 2.2** (Convolution $*$). *The general expression of a matrix convolution is*

$$\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} * \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mn} \end{bmatrix}$$
$$= \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} x_{(m-i)(n-j)} y_{(1+i)(1+j)}$$

Figure 1 shows an example of fingerprint input with its corresponding output by the filter.

## 2.2 Generation of Templates with URP

**Definition 2.3.** *Let* $\mathcal{K}$ *be the token (seed) space, representing the set of tokens to be assigned to users. A*

Figure 1: Left: Fingerprint image. Right: Resulting image after Sobel filter.

*cancelable biometric scheme is a pair of deterministic polynomial time algorithms* $\Xi := (\mathcal{T}, \mathcal{V})$*, where:*

- $\mathcal{T}$ *is the transformation of the system, that takes a feature vector* $F \in \mathcal{M}_F$ *and the token parameter* $P$ *as input, and returns a biometric template* $T = \mathcal{T}(P,F) \in \mathcal{M}_T$.
- $\mathcal{V}$ *is the verifier of the system, that takes two biometric templates* $T = \mathcal{T}(P,F)$, $T' = \mathcal{T}(P',F')$, *and a threshold* $\tau_T$ *as input; and returns* $True$ *if* $D_T(T,T') \leq \tau_T$, *and returns* $False$ *if* $D_T(T,T') > \tau_T$.

The attacked CB instantiation, described in Algorithm 1, is based on a uniform random projection (*URP*). Such a projection serves as an embedding of a high-dimensional space into a space of much lower dimension while preserving approximately the distances between all pairs of points. This type of dimensionality reduction is characterized by the Johnson–Lindenstrauss lemma 2.1 (Johnson, 1984). Algorithm 1 assumes the second factor, *i.e.*, the token, is a password and output a *Biometric Compressed Vector* (BCV).

**Lemme 2.1** (Johnson-Lindenstrauss)**.** *Given* $0 < \varepsilon < 1$, *a set* $X$ *of* $m$ *point in* $\mathbb{R}^N$, *and a number* $n > 8\left(\frac{ln(m)}{\varepsilon^2}\right)$, *there is a linear map* $f : \mathbb{R}^N \mapsto \mathbb{R}^n$ *such that for all* $u,v \in \mathbb{X}$ *:*

$$(1-\varepsilon)\,||u-v|| \leq ||f(u) - f(v)|| \leq (1+\varepsilon)\,||u-v||$$

**Remark 2.2.1.** *Biohashing instantiation (Jin et al., 2004) is based on the same type of projection, except that an additional step of orthonormalization of the family V by Gram-Schmidt is performed. This skipped step affects neither the recognition accuracy nor the feasibility of the attacks. However, their running times are reduced. Indeed, experiments over FVC-2002-DB1 using the URP-Sobel scheme yield a decision threshold at* 225 *for an EER equal to* 0.29%. *However, in the case of Biohashing, the same experiments yield a decision threshold at* 224 *for an EER equal to* 0.27%. *Therefore, the recognition accuracy results are pretty similar whether or not orthonormalization is performed.*

**Algorithm 1** [URP-SOBEL]

**Inputs :** biometric data $I$; token parameter $P$
**Output :** $T = (t_1, \ldots, t_m)$
$T$ is a BCV (Biometric Compressed Vector)

Apply Sobel filter on $I$ to produce an $n$-sized feature vector: $F = (f_1, \ldots, f_n)$.
Generate with the token $P$ a family $V$ of $m$ pseudorandom vectors $V_1, \ldots, V_m$ of size $n$ according to a uniform law $\mathcal{U}([-0.5, 0.5])$.
Arrange the family $V$ as a matrix $M$ of size $n \times m$.
Compute $T$ as the matrix-vector product $F \times M$.
**for** $t_i$ in $T$ **do**
    **if** $t_i < 0$ **then** $t_i = 0$ **else** $t_i = 1$
**end for**
**return** $T$

## 2.3 Attack Models and Objectives

We perform an authentication attack and, we are able to get access to this system in the name of the targeted person.

To perform this attack some information are needed:

- The password of our target.
- The original biohash of the target.
- Knowledge over the attacked system:
  - How to get the matrix from the password.
  - The value of the quantization that was used to create the BCV.

We show that anybody can perform a simple authentication attack or a one fingerprint image for several impersonations attack by building a template preimage if he knows the above information.

The informal definitions of (Ghammam et al., 2020) are tailored for the rest of the paper. Let $I \in \mathcal{M}_I$ be a fingerprint image, and let $T = \Xi.\mathcal{T}(P, E(I)) \in \mathcal{M}_T$ be the template generated from $I$ and the secret parameter $P$. In our authentication attack, an adversary is given $T$, $P$, and a threshold value $\tau_T$, and the adversary tries to find a fingerprint image $I^* \in \mathcal{M}_I$ such that for $T^* = \Xi.\mathcal{T}(P, E(x^*))$, $T^*$ is exactly the same as $T$, or $T^*$ is close to $T$ with respect to the distance function over $\mathcal{M}_T$ and the threshold value $\tau_T$. In this case, we say that $I^*$ is a $\tau_T$-nearby-template preimage (or simply a nearby-template preimage, when $\tau_B$ is clear from the context) of the template $T$.

A strategy for the adversary which have stolen the secret parameter $P$ is to alter her fingerprint image $I_A$ such that $P$ along with her extracted feature vector $F_A$ enable the generation of the exact template $T$. This motivates the notion of *template fingerprint preimage* defined below.

**Definition 2.4** (Template fingerprint preimage)**.** *Let $I \in \mathcal{M}_I$ be a fingerprint image, and $T = \Xi.\mathcal{T}(P, \Pi.E(I)) \in \mathcal{M}_T$ a template for some secret parameter $P$. A template preimage of $T$ with respect to $P$ is a fingerprint image $I^*$ such that $T = \Xi.\mathcal{T}(P, \Pi.E(I^*))$.*

Another authentication attack consists in generating a fingerprint image that yields the exact templates of two distinct users with their corresponding stolen tokens. More formally, we have the following definition:

**Definition 2.5** (Two-template fingerprint preimage)**.** *Let $I_1$ and $I_2 \in \mathcal{M}_I$ be two fingerprint images of distinct users, and two templates $T_1 = \Xi.\mathcal{T}(P_1, \Pi.E(I_1)) \in \mathcal{M}_T$ and $T_2 = \Xi.\mathcal{T}(P_2, \Pi.E(I_2)) \in \mathcal{M}_T$ for distinct secret parameters $P_1$ and $P_2$. A two-template preimage of the pair $(T_1, T_2)$ with respect to the pair $(P_1, P_2)$ is a fingerprint image $I^*$ such that $T_1 = \Xi.\mathcal{T}(P_1, \Pi.E(I^*))$ and $T_2 = \Xi.\mathcal{T}(P_2, \Pi.E(I^*))$.*

To capture the case of multi-collisions, this last definition can be generalized to a notion of a $n$-template fingerprint preimage.

**Definition 2.6** ($n$-template fingerprint preimage)**.** *Let $I_1, \ldots, I_n \in \mathcal{M}_I$ be $n$ fingerprint images of distinct users, and $n$ templates $T_i = \Xi.\mathcal{T}(P_i, \Pi.E(I_i)) \in \mathcal{M}_T$ for distinct secret parameters $P_i$ $\forall i \in \{0, \ldots, n\}$. A $n$-template preimage of $(I_1, \ldots, I_n)$ with respect to $(P_1, \ldots, P_n)$ is a fingerprint image $I^*$ such that:*

$$\forall i \in \{0, \ldots, n\}, T_i = \Xi.\mathcal{T}(P_i, \Pi.E(I^*)).$$

# 3 Simple Authentication Attack

## 3.1 Overview

There are two ways of performing this attack. The first one includes two steps described in Section 3.1.1. First, given an attacker feature vector, we seek the slightest modification of it such that its transformation by $\Xi$ yields exactly the template of the victim. Then, using the filter constraints of the convolution, we seek the slightest variation of the attacker's image such that the filtering of this variation produces exactly the modified feature vector. The second approach described in Section 3.1.2 consists in generating all constraints at once and directly generating the modified attacker's image.

### 3.1.1 First Approach

The attack takes as input the following parameters:

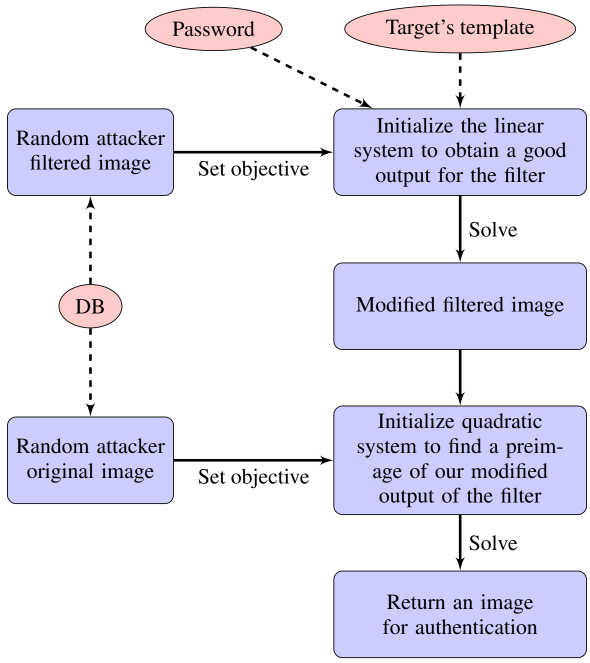


Figure 2: Principle of the attack's first approach.

- The target's password ($P_t$).
- The target's template ($T_t$).
- The attacker's image ($I_A$).

This attack computes and uses following information:

1. Attacker's feature ($F_A$).
2. Modified attacker's feature ($F'_A$).

The output is a modified attacker's image $X$ which matches the target template.

First, the attacker uses $I_A$ to compute $F_A$ using filter. Then, with $P_t$ and $T_t$, the attacker modifies image's feature to match the target template $F'_A$. As described in Section 3.2.1, it is done by solving an under-constraint linear system and seeking the nearest modified feature which matches the target template. After that, using $F'_A$ and $I_A$, the attacker modifies its image to match the modified feature. As described in Section 3.2.2, it is done by solving an under-constraint quadratic system and seeking the nearest modified image which matches the feature.

Figure 2 gives an overview of this first method step by step, where inputs are in circles and different steps in boxes.

### 3.1.2 Second Approach

The attack takes as input the same parameters ($P_t$, $T_t$ and $I_A$). The output is a modified attacker's image $X$ which matches the target template.

The main idea is to merge both steps described

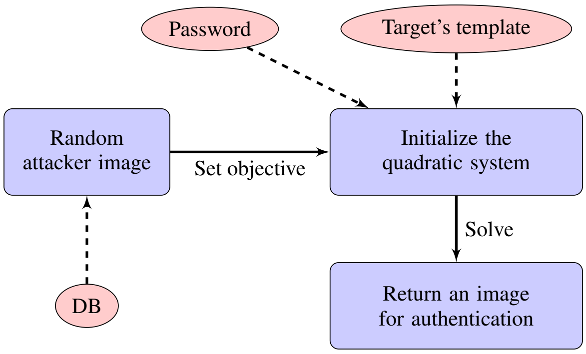


Figure 3: Principle of the attack's second approach.

in Section 3.1.1. A unique constrained quadratic system is solved to find the nearest modified image which matches the template (see Figure 3).

## 3.2 Program Formulation for the Two-Phase Approach

As explained, we proceed in two steps.

### 3.2.1 Getting a Correct Output for the Filter

For this part, we assume that we are after the filter. We see how to inverse the filter later.

We want to reverse target's template by using the password. To do that, let $X = (x_0, \ldots, x_n)$, $M$ the projection matrix derivated from target's password and $f$ the quantization function which takes $XM$ to create a binary template.

We know the projection matrix and the image we need to get for the target client. Thus, one can seek to calculate a pre-image of the projected vector by solving a system under constraints.

**Remark 3.2.1.** *This attack works for many projections system such as Biohash.*

Let us write it more formally. Let $T = (t_1, \ldots, t_m)$ the biometric template, $n$ the size of BCV and

$$M = \begin{bmatrix} a_{1,1} & \cdots & a_{0,n} \\ \vdots & \ddots & \vdots \\ a_{m,0} & \cdots & a_{m,n} \end{bmatrix}.$$

Let $\mathcal{K}_1$ be all indices where the template is equal to 0 and $\mathcal{K}_2$ all other indices. So, we seek a solution to the following system:

$$\begin{cases} XM_i < 0, \forall i \in \mathcal{K}_1 \\ XM_i \geq 0, \forall j \in \mathcal{K}_2 \\ x_i \in \mathbb{R}^+, \forall i \in (\mathcal{K}_1 \cup \mathcal{K}_2) \end{cases} \tag{1}$$

With $M_i$ the $i$-th column of $M$. We seek to minimize the distance between $F$ and $F_A$. By doing so, the attacker can be authenticated by modifying the smallest number of information of his own biometric feature vector.

This part of the attack solves the following problem. By taking $F_A = (o_1, \ldots, o_n)$ the attacker's biometric feature, $M$ the projection matrix we have:

- Minimize: $\|X - F_A\|^2$
- Under the following constraints:

$$\begin{cases} XM_i < 0, \forall i \in \mathcal{K}_1 \\ XM_i \geq 0, \forall j \in \mathcal{K}_2 \\ x_i \in \mathbb{R}^+, \forall i \in (\mathcal{K}_1 \cup \mathcal{K}_2) \end{cases} \tag{2}$$

With $M_i$ the $i$-th column of $M$.

### 3.2.2 Get a Preimage to Avoid Filter Effect

The filter leads to a loss of information. But we can write a quadratic system to create a collision and get a correct preimage. Let the image matrix be

$$I = \begin{bmatrix} o_{0,0} & \cdots & o_{0,\text{width-1}} \\ \vdots & \ddots & \vdots \\ o_{\text{length-1},0} & \cdots & o_{\text{length-1},\text{width-1}} \end{bmatrix}$$

Applying the filter to that formal matrix yields a new matrix $D$ which has quadratic components. But, we know that $D$ must be equal to $F_A$. Thus, we can solve a quadratic system with (length $\times$ width) equations and (length $\times$ width) variable to find a preimage.

Let $I_A = (o_{i,j})$ denote the attacker's original image, $F_A = (a_{i,j})$ its modified feature, $I' = (x'_{i,j})$ the modified original image and $X = (x_{i,j})$ its augmented form. We consider the augmented form as the original matrix where zeroes are added all around the matrix to compute the convolution without overflowing.

In the case of Sobel filter, we solve the following problem:

- Minimize: $\sum_{i,j} (o_{i,j} - x_{i,j})^2$
- Subject to the following constraints:

$$\begin{cases} \alpha_{i,j} = x_{(i-1,j-1)} + 2x_{(i,j-1)} + x_{(i+1,j-1)} \\ \qquad - x_{(i-1,j+1)} - 2x_{(i,j+1)} - x_{(i+1,j+1)} \\ \beta_{i,j} = x_{(i-1,j-1)} + 2x_{(i-1,j)} + x_{(i-1,j+1)} \\ \qquad - x_{(i+1,j-1)} - 2x_{(i+1,j)} - x_{(i+1,j+1)} \\ a_{i,j}^2 = \alpha_{i,j}^2 + \beta_{i,j}^2, \forall (i,j) \\ x_{i,j} = 0 \text{ if } i = 0 \text{ or } i = length + 1 \\ x_{i,j} = 0 \text{ if } j = 0 \text{ or } j = width + 1 \\ x_{i,j} \in [\![0, 255]\!], \forall (i,j) \end{cases} \tag{3}$$

Using the notations of Section 2.1, we obtain:

- Minimize: $\|X - I_A\|^2$
- Under the following constraints:

$$\begin{cases} (F_A)^2 = \left[(G_1 * X)^2 + (G_2 * X)^2\right] \\ x_{i,j} \in \llbracket 0,255 \rrbracket, \forall (i,j) \end{cases} \tag{4}$$

## 3.3 Formulation as a Single Program

Yet another method is to merge both systems to create a new quadratic system. In this case, we avoid some problems such as having an intermediate feature vector which is not in the range of the filter function.

Assume that $I_A = (o_{i,j})_{n\times m}$ is the attacker's original image, $I' = (x'_{i,j})_{n\times m}$ the modified original image and $X = (x_{i,j})_{n\times m}$ its augmented form. Let $\mathcal{K}_1$ be all indices where the template is equal to 0 and $\mathcal{K}_2$ all other indices. Let $M = (a_{i,j})_{(n*m)\times \ell}$ be the projection matrix. Let $Y_{flat}$ be the flattened form of the matrix $Y$ where rows are concatenated in a single vector.

Thus, using the notations from the sections 3.1.1 and 3.1.2 we define the following problem for Sobel filter:

- Minimize: $\|X - I_A\|^2$
- Under the following constraints:

$$\begin{cases} Y^2 = \left[(G_1 * X)^2 + (G_2 * X)^2\right] \\ Y_{flat} M_i < 0, \forall i \in \mathcal{K}_1 \\ Y_{flat} M_j \geq 0, \forall j \in \mathcal{K}_2 \\ x_{i,j} \in \llbracket 0,255 \rrbracket, \forall (i,j) \end{cases} \tag{5}$$

Where the squaring stands for the coordinate by coordinate squaring (*i.e.* Hadamard squaring) and $M_i$ the $i$-th column of $M$.

# 4 Multiple Authentications Attack

The goal of this attack is to find an image that can impersonate several victims. The attacker computes an image whose derived template is kind of a barycenter of all the targeted templates This is possible only if these templates are not too far with respect to a threshold. A specificity of this attack is that the passwords of the victims are not required.

## 4.1 Overview

Let $\mu$ denote the set of the victims. The attack takes as input the following parameters:

1. The target's templates $(T_t)_{i\in\mu}$.
2. The attacker's image $(I_A)$.

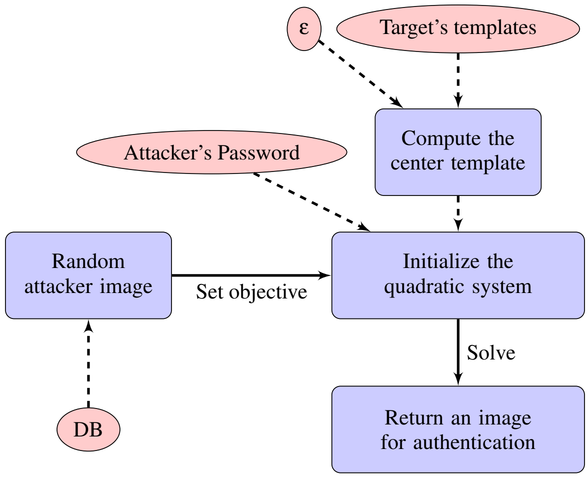


Figure 4: Principle of the multiple authentications attack.

3. The attacker's password $(P_A)$.
4. The value of $\varepsilon$ the decision threshold.

The output is a modified attacker's image $X$ which matches the modified template.

First, with respect to all the targeted templates, we seek a template $T$ such that they are in a ball centered in $T$ and of radius $\varepsilon$. If a center[1] does not exist, a subset of the targeted templates for which the center exists is considered.

Then, a quadratic system and a function to minimize can be built as explained in Section 4.2. Thus, solving this problem gives us the modified image for multiple authentications with the same password. We present an overview of this attack in Figure 4.

## 4.2 Program Formulation

Let $M$ be the projection matrix and $T$ the template at the center of the ball as defined in Section 4.1. Assume that $\mathcal{K}_1$ is the list of all indices where $T_i$ is equal to 0 and $\mathcal{K}_2$ is the list of the remaining indices. The other notations are the same as in Section 3.3. The problem can be defined like this:

- Minimize: $\|X - I_A\|^2$
- Under the following constraints where $M_k$ is the $k$-th column of $M$:

$$\begin{cases} Y^2 = \left[(G_1 * X)^2 + (G_2 * X)^2\right] \\ Y_{flat} M_j < 0, \forall j \in \mathcal{K}_1 \\ Y_{flat} M_k \geq 0, \forall k \in \mathcal{K}_2 \\ x_{i,j} \in \llbracket 0,255 \rrbracket, \forall (i,j) \end{cases} \tag{6}$$

With $M_i$ the $i$-th column of $M$.

[1] Several centers may exist and, with more than 2 templates the existence of at least one center is not ensured.

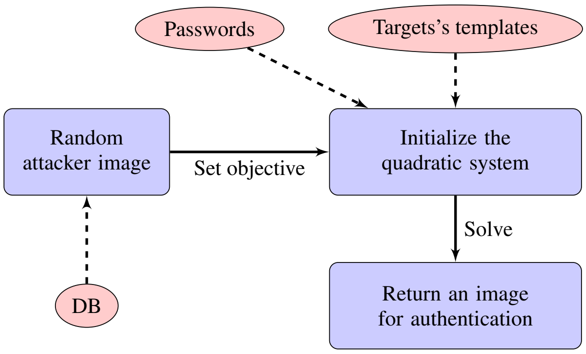

Figure 5: Principle of the attack's second approach.

# 5 Multiple Collisions Attack

In this attack, the attacker knows the templates and passwords of the victims. Then, his goal is to use all these information to generate one image that allows her to impersonate all the victims using their own password.

## 5.1 Overview

The attack takes as input the following parameters:

1. The target's templates $(T_t)_{i\in\mu}$.
2. The attacker's image $(I_A)$.
3. The target's passwords $(P)_{i\in\mu}$.

The output is a modified attacker's image $X$ which matches all templates for the corresponding password.

We define a quadratic system and a function to minimize as explained in Section 5.2. Thus, solving this problem gives us the modified image for multiple authentications for each password. An overview of this attack is depicted in Figure 5.

## 5.2 Program Formulation

Let $M_i$ be the projection matrix for the $i$-th user. Assume that $(\mathcal{K}_1)_i$ is the list of all indices where $(T_t)_i$ is equal to 0 and $(\mathcal{K}_2)_i$ all other indices. The other notations are the same as in Section 3.3. The problem can be defined like this:

- Minimize: $\|X - I_A\|^2$
- Under the following constraints where $(M_i)_j$ is the $j$-th column of $M_i$:

$$\begin{cases} Y^2 = \left[(G_1 * X)^2 + (G_2 * X)^2\right] \\ Y_{flat}(M_i)_j < 0, \forall i \in \mu, \forall j \in (\mathcal{K}_1)_i \\ Y_{flat}(M_i)_k \geq 0, \forall i \in \mu, \forall k \in (\mathcal{K}_2)_i \\ x_{i,j} \in [\![0,255]\!], \forall (i,j) \end{cases} \tag{7}$$

As matrices $M_i$ are fully random, the probability of them forming an indexed family of linearly dependent vectors is negligible, thus making the system solvable. Assume that $L(V_1,\ldots,V_k)$ is the event that $(V_1,\ldots,V_k)$ is an indexed family of linearly independent vectors, with $n$ the size of vector and $\eta$ the number of precision bits for our numbers. It can be shown that

$$P(L(V_1,\ldots,V_k)) = \frac{\prod_{i=2}^{k} 2^{\eta(n-i+1)} - 1}{\prod_{i=2}^{k} 2^{\eta(n-i+1)}}.$$

Since this probability is near 1, the usurpation of $\left\lfloor \frac{n}{w} \right\rfloor$ persons with $w$ the size of the template is a likely event.

**Remark 5.2.1.** *A variant of this attack could be achieved without the users' passwords. The attacker just has to replace the passwords of the victims by distinct random strings. Thus, she obtains an image that allows her to impersonate several people. She merely chooses one victim by using its corresponding string. However, it may be possible that its string lead to an infeasible model and so another must be chosen.*

# 6 Reversibility Attack Evaluation

We evaluate the impact of our authentication attack with the second approach (3.1.2) through our Python implementation. The Gurobi Python interface (Gurobi 9.1.2) is used to solve the non-convex quadratically constrained programs, on a computer running on Debian 11, with an EPYC 7F72 dual processor (48 cores) and 256GB of RAM. The focus is only done on the results of this attack because its practicality implies the practicality of the others.

We have launched resolutions of the constrained programs 50 times, each with a time limit of 150 seconds. Table 1 reports the running times for the different settings along with the amount of changes done in the attacker fingerprint, by means of the Euclidian distance.

In Table 1, we remark that with a 4×4-pixel image and a 50-bit template, the hard cap of 150 seconds starts to be insufficient to solve the system and optimize the criterion. However, the experiments are encouraging given that we face an NP-hard problem (Sahni, 1974). By setting the hard cap to 500 seconds, we are able to solve the system with a 10×10-pixel image and a better ratio amount of changes over image size.

| Image Size | Template Size | Mean Distance | Mean Time (s) |
|---|---|---|---|
| $2\times2$ | 20 | 58 | 0.04 |
| $2\times3$ | | 65 | 27.92 |
| $3\times3$ | | 71 | 120.18 |
| $4\times3$ | | 95 | 135.56 |
| $4\times4$ | | 152 | 140.83 |
| $2\times2$ | 30 | 81 | 1.58 |
| $2\times3$ | | 83 | 77.24 |
| $3\times3$ | | 76 | 129.36 |
| $4\times3$ | | 102 | 138.0 |
| $4\times4$ | | 153 | 143.18 |
| $2\times2$ | 40 | 119 | 0.12 |
| $2\times3$ | | 102 | 33.1 |
| $3\times3$ | | 121 | 144.0 |
| $4\times3$ | | 133 | 146.81 |
| $4\times4$ | | 168 | 146.60 |
| $2\times2$ | 50 | 99 | 0.14 |
| $2\times3$ | | 117 | 32.76 |
| $3\times3$ | | 133 | 150.0 |
| $4\times3$ | | 144 | 146.67 |
| $4\times4$ | | 177 | 150.0 |

Table 1: Summary of the experiments.

## 7 Concluding Remarks

In this paper, we present several authentication attacks on a popular CB scheme consisting in a composition of a kernel-based filter with a projection-based transformation, in the stolen token scenario. Their particularity is to completely reverse a CB scheme. to impersonate any or several users. To the best of our knowledge, this is the first time that attacks are conducted on a complete chain of treatments, including a non-linear filter. The proposed methodology is to formalize the attacks as constrained optimization problems. As long as the attacker has access to one or several templates with the corresponding passwords, our attacks can be performed. In addition, we present two ways for the attacker to impersonate several legitimate persons Some attacks proposed do not need any token from the clients. Our practical experiments show that the modification of the attacker's image is minimal over small images. The next step is to perform these attacks on larger images and look for the limit of the number of people that can be impersonated at the same time. Future work will be focused on finding optimizations and relaxations of the systems to ensure the scaling of our attacks.

## Acknowledgement

The authors acknowledges the support of the French Agence Nationale de la Recherche (ANR), under grant ANR-20-CE39-0005 (project PRIVABIO).

## REFERENCES

(2011). ISO/IEC24745:2011: Information technology – Security techniques – Biometric information protection. Standard, International Organization for Standardization.

(2018). ISO/IEC30136:2018(E): Information technology – Performance testing of biometrictemplate protection scheme. Standard, International Organization for Standardization.

Bringer, J., Chabanne, H., and Kindarji, B. (2008). The best of both worlds: Applying secure sketches to cancelable biometrics. *Sci. Comput. Program.*, 74(1-2):43–51.

Dodis, Y., Reyzin, L., and Smith, A. (2004). Fuzzy extractors: How to generate strong keys from biometrics and other noisy data. In Cachin, C. and Camenisch, J. L., editors, *Advances in Cryptology - EUROCRYPT 2004*, pages 523–540, Berlin, Heidelberg. Springer Berlin Heidelberg.

Dong, X., Jin, Z., and Jin, A. T. B. (2019a). 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.

Dong, X., Jin, Z., Teoh, A. B. J., Tistarelli, M., and Wong, K. (2019b). On the reliability of cancelable biometrics: Revisit the irreversibility. *CoRR*, abs/1910.07770.

Ghammam, L., Karabina, K., Lacharme, P., and Thiry-Atighehchi, K. (2020). A cryptanalysis of two cancelable biometric schemes based on index-of-max hashing. *IEEE Transactions on Information Forensics and Security*, PP:1–12.

Jain, A. K. and Nandakumar, K. (2012). Biometric authentication: System security and user privacy. *Computer*, 45:87–92.

Jin, A. T. B., Ling, D. N. C., and Goh, A. (2004). Biohashing: two factor authentication featuring fingerprint data and tokenised random number. *Pattern Recognition*, 37(11):2245–2255.

Johnson, W. (1984). Extensions of lipschitz mappings into hilbert space. *Contemporary mathematics*, 26:189–206.

Juels, A. (2006). A fuzzy vault scheme. *Designs, Codes and Cryptography*, 38:237–257.

Lacharme, P., Cherrier, E., and Rosenberger, C. (2013). Preimage attack on biohashing. In *2013 International Conference on Security and Cryptography (SECRYPT)*, pages 1–8.

Li, C. and Hu, J. (2014). Attacks via record multiplicity on cancelable biometrics templates. *Concurrency Computation: Practice and Experience*, pages 1593–1605.

Nagar, A., Nandakumar, K., and Jain, A. K. (2010). Biometric template transformation: a security analysis. In Memon, N. D., Dittmann, J., Alattar, A. M., and Delp, E. J., editors, *Media Forensics and Security*, volume 7541 of *SPIE Proceedings*, page 75410O. SPIE.

Nandakumar, K. and Jain, A. K. (2015). Biometric template protection: Bridging the performance gap between theory and practice. *IEEE Signal Processing Magazine*, 32:88–100.

Natgunanathan, I., Mehmood, A., Xiang, Y., Beliakov, G., and Yearwood, J. (2016). Protection of privacy in biometric data. *IEEE Access*, 4:880–892.

Quan, F., Fei, S., Anni, C., and Feifei, Z. (2008). Cracking cancelable fingerprint template of ratha. In *2008 International Symposium on Computer Science and Computational Technology*, volume 2, pages 572–575.

Ratha, N. K., Connell, J. H., and Bolle, R. M. (2001). Enhancing security and privacy in biometrics-based authentication system. *IBM Systems J.*, 37(11):2245–2255.

Rozsa, A., Glock, A. E., and Boult, T. E. (2015). Genetic algorithm attack on minutiae-based fingerprint authentication and protected template fingerprint systems. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.

Sahni, S. (1974). Computationally related problems. *SIAM Journal on Computing*, 3(4):262–279.

Stockman, G. and Shapiro, L. G. (2001). Upper Saddle River, NJ, USA.

Sutcu, Y., Sencar, H. T., and Memon, N. (2005). A secure biometric authentication scheme based on robust hashing. In *Proceedings of the 7th Workshop on Multimedia and Security*, page 111–116, New York, NY, USA. Association for Computing Machinery.

Teoh, A. B. J., Yip, W. K., and Lee, S. (2008). Cancellable biometrics and annotations on BioHash. *Pattern Recognition*, 41(6):2034–2044.

Topcu, B., Karabat, C., Azadmanesh, M., and Erdogan, H. (2016). Practical security and privacy attacks against biometric hashing using sparse recovery. *EURASIP Journal on Advances in Signal Processing*, 2016(1):100.

Vincent, O. and Folorunso, O. (2009). A descriptive algorithm for sobel image edge detection.

Wang, H., Dong, X., Jin, Z., Teoh, A., and Tistarelli, M. (2021). Interpretable security analysis of cancellable biometrics using constrained-optimized similarity-based attack. In *Proceedings - 2021 IEEE Winter Conference on Applications of Computer Vision Workshops, WACVW 2021*, pages 70–77.