A Biometric Template Security Approach to Fingerprints Based on Polynomial Transformations

Ramon Santana

Abstract—The use of biometric identifiers in the field of information security, access control to resources, authentication in ATMs and banking among others, are of great concern because of the safety of biometric data. In the general architecture of a biometric system have been detected eight vulnerabilities, six of them allow obtaining minutiae template in plain text. The main consequence of obtaining minutiae templates is the loss of biometric identifier for life. To mitigate these vulnerabilities several models to protect minutiae templates have been proposed. Several vulnerabilities in the cryptographic security of these models allow to obtain biometric data in plain text. In order to increase the cryptographic security and ease of reversibility, a minutiae templates protection model is proposed. The model aims to make the cryptographic protection and facilitate the reversibility of data using two levels of security. The first level of security is the data transformation level. In this level generates invariant data to rotation and translation, further transformation is irreversible. The second level of security is the evaluation level, where the encryption key is generated and data is evaluated using a defined evaluation function. The model is aimed at mitigating known vulnerabilities of the proposed models, basing its security on the impossibility of the polynomial reconstruction.

Keywords—Fingerprint, template protection, bio-cryptography, minutiae protection.

I. INTRODUCTION

The use of biometric identifiers to control access to protected resources increases safety considerably. The main reason is that physical or behavioral characteristics are inherent to an individual, which are more difficult to steal, lose or guess than traditional identifiers. The fingerprint is considered the most widely used biometric identifier for recognizing people. This is mainly due to that the acquisition process of the biometric feature is minimally invasive [1].

Biometrics is the measurement of biological data [2], the term is commonly used to refer to the recognition of a person by physical characteristics such as fingerprint, face, iris; or behavior characteristics as signature and the way they walk. Today biometrics has a great scope in criminal, government and commercial systems [3], gaining wide acceptance as one of the most effective technologies for people authentication in a wide range of informatics applications. A biometric system is essentially a pattern recognition system that operates from the acquisition of biometric data from an individual, extracts a set of characteristics of the data captured and compared with the data stored. Depending on the context, it can be used for verification or biometric identification. The verification process validates the identity of an individual by comparing the obtained sample with the one stored in the database. The identification process compares the acquired fingerprint with all samples stored in the database. This process is a critical component in the implementation of negative recognition, which prevents a person from having multiple identities [2].

The implementation of biometrics in civil and governmental sectors as a means of security through public and / or private networks, has generated more concern for the security of the biometric data. Analyzes made by several authors [2], [4], [5] detected eight points of vulnerability in the overall architecture of a biometric system in which it is possible to obtain the biometric trait, as shown in Fig. 1.

The vulnerability points in the general architecture of an automatic fingerprint identification, which are of interest for this research are those by which it is possible to obtain the minutiae template partially or completely. These are:

1) The of biometric features extractor.
2) The communication channel between the extractor and the biometrics comparator.
3) The comparator of biometric features.
4) The communication channel between the comparator and the biometrics features database.
5) The biometric database.

The main reason for this concern is that the fingerprint, as biometric identifier, is unique for lifetime and can not be canceled or changed as a personal password. If an attacker obtains the minutiae belonging to a fingerprint, this means the loss of the identifier for the lifetime. This is because of the leak of a minutiae template, either partially or completely, allows the reconstruction of the corresponding fingerprint, obtaining an impression as proposed in [3].

The easiest way to protect the stored biometric data would be using the traditional cryptography [4], [6] however, the properties of the methods used by these methods obstruct the process of comparing minutiae in a protected domain. This is mainly due to that small changes in the data set to encrypt cause large changes in the set of encrypted data. Fingerprint samples change due to various factors such as translation, rotation, overlap and nonlinear deformation experienced by the finger when making contact with a surface.

The main drawback of using conventional cryptographic methods such as AES, RSA, triple DES, among others, to protect biometric data lies in the loss of biometric performance during the minutiae template matching in the protected domain. Therefore, it is necessary to decode the set of biometric data before making the feature comparison process. During this time, the biometric features are in plain text. Different attacks using Trojan viruses or hardware malfunction during the comparison process are some of the factors that allow to obtain the minutiae during the comparison process.

R. Santana is with the Department of Biometrics, Personal Identification and Digital Security Center, Habana, HA, 10800 CU (e-mail: rsfernandez@uci.cu).
In this research, "Section II background work" authors analyzed proposed for the protection of templates fingerprint minutiae, strengths and weaknesses, in Section III a method of protecting minutiae templates based on the problem of reconstruction work proposed analyzes polynomial, in Section IV the cryptographic strength of the proposed model is analyzed.

II. BACKGROUND WORK

Minutiae templates protection models base their operation in the transformation of data, masking of the original data or a combination of both. In the case of biometric crypto-systems have been developed two main approaches, models for generation of cryptographic keys and key partnership models. The main objective in both cases is to decrease the chances of obtaining minutiae templates, partially or totally, in attacks performed to automated fingerprint identification systems while facilitating the comparison in the protected domain. The models summarized below constitute the cryptographic base of the fingerprint minutiae template protection process.

A. Fuzzy Vault

This model is a cryptographic construction based on fuzzy compromise proposed in [7]. It is a biometric crypto-system designed to perform encryption of disorderly sets. The fuzzy vault model, first proposed in [8], was conceived as an encryption method tolerant to fault. The model proposed in [8] is composed by two methods, an encoding method and a decoding method of the fuzzy vault. The procedure performed to encode the biometric data is to create a generalized Reed-Solomon key word, representing the secret (along with the corresponding polynomial p where k represents the coefficients of the polynomial). The X coordinates corresponding to the original data set A are evaluated at p ← k. To hide the result of this operation, a set of garbage points or mockery spots are generated in the way (x,y) and mixed randomly. As premised, in the generation of garbage points, they should be selected in the way that do not to intersect in set A or the polynomial p. The method for decoding data contained in a fuzzy vault takes as input the sample set B at the vault Va and consists in determining the codeword encoding the secret k. It is performed k' ← p (inverse encryption procedure) to denote the conversion of a polynomial of max degree k in f° secret. Denoted (x,y) (b_i, y) R as the projection of the encrypted set R in the coordinate x (b_i). If there is a pair (b_i, y) belonging to the set R for any value of y, then (b_i, y) = (x, y), if it is not assigned the null value to the point (x, y). If successful, the secret k' is obtained as a result, which must be equal the original if the test set B is similar to the original set A. According to several authors [4], [8], [9], the cryptographic security of this model is based on the computational difficulty of solving the problem of reconstructing the polynomial and the number of garbage points that are added to mask the original points.

B. Cancellable Templates

This pattern consists in a repeated and intentional distortion of the biometric signal based on a transformation [10]. The transformation has the fundamental property of non invertibility of data. This type of transformation can be applied in both the domain of the signal and the extracted features of the biometric feature. The procedure for the transformation involves mapping the original characteristics S into S’ so that it cannot be recovered S from S’. The function used to map the biometric features has the property of one to many and various functions can be used to perform the transformation of the other two components of a minutiae (y, σ).

In [11] and [12] another approach of the cancellable templates model for protecting fingerprint minutiae templates is described. In this approach the analysis is performed in the domain of the extracted features and not in the signal domain. For the encryption of features is used a oneway function or non-invertible function, with the one to many property. The alignment process is performed by detecting the parabolic and triangular symmetry associated to the singular points of the fingerprint.

In this approach are described 3 types of transformations for encrypting data, Cartesian, polar and functional transformations. The Cartesian transformation maps the minutiae in rectangular coordinates using as reference one of the singular points, orienting the x axis in the same direction as the singularity and dividing the rectangular area in cells or sub areas of fixed size. This transformation consists in the cell change of the minutiae and rotations can be made.
in multiples of 90 degrees after transformation. The cells mapping is made from a mapping matrix $M$, so that the process can be denoted as $C' = CM$, where $C'$ is the transformed set and $C$ is the original set.

The radial or polar transformation consists of mapping the original minutiae in the polar coordinates space with referenced to the core singularity. For mapping the minutiae, the space is divided into polar sectors and minutiae are changed to different sectors to alter the position and angle value. The mapping is performed considering the translation key $1 \times LS$ where $L$ is the number of levels and $S$ represents the angle. The transformation function can be described as $C' = C + M$. The functional transformation consists in evaluating the minutiae in a parametric function, softened locally but not globally and governed by a key. The function has three restrictions as follows:

1) The transformation must be locally softened. This ensures that small changes in the position of the minutiae before transformation leads to small changes in the position of the transformed data.
2) The transformation should not be globally softened. This ensures that the original and transformed data are not highly correlated, to ensure the cryptographic security of the model.
3) The data transformation must ensure that the distance between the original and processed data is greater than that accepted by the comparison algorithm.

The encoding process using cancellable templates is performed in each authentication and every biometric enrollment in the system. If a protected template is compromised, it is possible to change the transformation function to generate a new template from the biometric data of the user. Thus, even if the protected template and the transformation function are known, the original biometric data can not be recovered.

C. Biohashing

This model consists in the representation and transformation of a set of data extracted from the minutiae starting from a reference point, using the extraction technique proposed in [13]. This model is exclusively applied to protection texture features of the fingerprint and comprises a rendering method and filtering method. The method used for representing information is called FingerCode, which consists of three basic steps:

1) Determine the reference framework in the fingerprints image.
2) Filter the image in 8 different directions using the Gabor filter bank.
3) Calculate the standard deviation of gray values in sectors around the reference point.

The filtering of the characteristics generates a set of disks that contain the information to be filtered to form a fixed length vector that represents the biometric fingerprint hash under analysis. Calculating the standard deviation in these filters defines the feature vector components. Another approach of the protection model is described in [14] called biometric hash. The main contribution of this approach, in relation to the previous model, is to eliminate the dependence of the core of the fingerprint as a reference point. In this case each minutia is represented by its FingerCode and to protect each FingerCode, the creation of the biometric hash is performed, which is described by the following steps:

1) Features calculation.
   a) The minutiae template is extracted from the image.
   b) For each minutia its Finger Code is calculated, the result is called MinuCode.
   c) The BioHashing of each Minucode is obtained.
2) Features comparison
   a) Deformations caused by the rotation are corrected.
   b) The BioHashing of the new Minucode is processed.
   c) The process of local comparison is made between the two templates.

Another approach for obtaining the biometric hash is described in [15]. In this work two descriptors are proposed, the first-one based on texture to capture the ridges flow pattern and another descriptor based on minutiae, for each minutia relationship with its neighborhood. The feature extraction is performed similar to the previously proposed, the variation resides in the use of the k-neighborhood (K-Plet) with center in the minutia that is being analyzed for the based on minutia descriptor. This allows the local representation of information among minutiae and it is selected to verify if the comparison of the based on texture descriptor is globally consistent. To minimize the impact of changes in the minutia making up the K-Plet structure, a comparison of structures is made using a proposed alignment technique.

III. PROPOSED MINUTIAE TEMPLATE PROTECTION SCHEME

In this research a model is proposed to protect the biometric data contained in the fingerprint minutiae templates composed of two levels of security. The inputs for the model are the minutiae templates in plain text and the outputs are the minutiae templates in the protected domain. The security levels of the model are:

1) The first security level of the model consists in the transformation of the information contained in the minutiae templates. The minutiae are represented using a minutiae structures based method and from this the extraction of identifiable characteristics is performed.
2) The second security level consists in evaluating the extracted features in an invertible function. This function is created from a seed given by the user.

The model also provides the inclusion or adaptation of a feature comparison method. This method should take into account the intra-user variations that fingerprints have besides detecting with a high level of certainty the genuine characteristics and aggregate ones in the input data. The fingerprint minutiae templates protection polynomial protection model is composed of three methods:

1) Method of representation and extraction of identifiable characteristics
2) Encryption method of identifiable characteristics.
3) Comparison method of identifiable characteristics. The representation and identifiable characteristics extraction component, makes the representation of the information contained in minutiae templates through the complex structure, which is a contribution of the research. This component receives as input the minutiae template in plain text and returns a set of transformed features derived from the minutiae, that allow to identify a person. The identifiable characteristics encryption component, performs the features encoding and enables the revocation of protected templates. This component receives as input and the transformed features returns the coded features. The comparison component, identifiable characteristics, calculates the similarity index between two protected templates. The component takes as input two sets of coded characteristics or protected templates and performs the comparison on two levels. The first level is the comparison of the primary structures, returning the similarity index between them. The second level is the comparison of the secondary structures, returning the similarity index between them. Finally, the consolidation of the resulting data from both stages is performed, both locally and globally.

A. Method for Representing and Extracting of Identifiable Characteristics

A method for representing the information contained in the minutiae, consists in the analysis and processing of the information contained in a minutia (x, y, σ, t). As essential property of this transformation, as part of the research, the resulting characteristics must be discriminative enough to identify a person. The representation and identifiable characteristics extraction component proposed as part of the model, should meet the following restrictions:
  1) Invariant to rotation and translation.
  2) Resistant to nonlinear deformation.
  3) Resistant to partial overlapping.
  4) Irreversible or one-way transformation.

As part of this research it is recommended to use the complex structure for the representation and the information extraction; as it provides local and global information about the fingerprint, that can be used during the comparison process. The following paragraphs, is described in a general way, the composition of a complex structure, its relationships and the kind of analysis that is performed with its use. A complex structure is characterized by the union of two minutiae structures widely studied in the literature:
  1) The n-nearest neighbors structure
  2) Minutiae triplets.

The n-nearest neighbors structure is used in the stage of representation and extracting information to globally characterize the fingerprint, establishing relationships among the n closest minutiae to a reference minutia. In the stage of comparing, this structure is used to obtain the relationship among the minutiae, allowing to detect which minutiae belong to the original set (sample set). Minutiae triplets are used by the representation and information extraction method for local analysis of the fingerprint. Through it, the description of the relationship established among three minutiae belonging to the complex structure is made. Of each triplet, a set of information is extracted, characterizing and locally identifying the structure. This identifying information initially transformed, is used as input in the encryption method. The representation and information extraction method is considered the first security level of the protection model called the transformation level. In general, this component model consists of:
  1) Minutiae structures creation.
  2) Extraction of the identifying information.
  3) Descriptors extraction.
  4) Transformed information classification.

B. Encryption Method of Identifiable Characteristics

As a second component of the model an encryption method of identifiable characteristics extracted in the previous component is proposed. For this, it is necessary to carry out:
  1) Generation of the encryption key.
  2) Creation of the encryption function.
  3) Data evaluation.

Due to the high variability of the fingerprints within and between user, the proposal is to perform the encryption process using a polynomial as a transformation function. For the construction of the polynomial should be considered:
  1) The comparison between the original feature set x and the transformed feature set f(x) can not be greater than the similarity threshold u defined in the comparison method.

\[
C(f_x), x > u
\]  

(1)

2) The generated polynomial must be of degree \( n > 3 \)
3) Several protected templates, generated from the same set of biometric data, can not be positive in a cross comparison.
4) Data to be stored for comparison will only be the image function, eliminating everything else

This way, is possible to ensure that the data transformation is irreversible and can not be correlated by a values multiplicity attack. The safety of the method is based in the impossibility of the polynomial reconstruction, which is an NP problem.

C. Comparison of Identifiable Characteristics

In order to perform biometric recognition in the protected domain, a method for comparing the identifiable characteristics in the protected domain is developed. The method used for comparison can be specifically designed for this process from the extracted features, or can be adapted for comparison from an existing method.

The protected templates comparison component in the protected domain is decomposed into:
  1) Creation of protected structures.
  2) Comparison of protected structures.
     a) Calculation of the central angles similarity.
     b) Local comparison.
     c) Global comparison.
  3) Data consolidation.
As premises for developing a protected templates comparison component using this model are defined:

1) The structures are treated as minutiae for the purpose of comparison, analyzing the similarity index between two structures.
2) The analysis of the information contained in the structures must be performed locally and globally.
3) Possibility to obtain which data match the original template and which data are introduced by the intra-user variations.
4) Data consolidation should reflect global and local analysis, providing greater emphasis on local analysis.

Additionally, it is necessary to carry out the calculation of the similarity between complex structures based on decision thresholds for greater accuracy in the comparison process globally. Depending on the complex structures selected for representing and extracting identifying information from the minutiae, it is proposed to establish a similarity threshold per selected feature. This allows data discrimination both locally and globally, increasing the accuracy of the comparison method.

IV. SECURITY OF THE PROPOSED MODEL

Fingerprint minutiae templates protection models must ensure safety of the protected data to different types of attacks. On the other hand, the revocability of protected templates should ensure that it is possible to generate more than one protected template from the same biometric feature and that is not possible to correlate two templates to obtain the original biometric data. To validate the proposed model, the cryptographic security of the pioneering models is analyzed as well as the revocability scheme. To do so, an analysis of the cryptographic security of the proposed model compared to pioneers models is performed, based on executing various attacks presented below:

1) Strength bits attacks
2) Brute Force attacks
3) Correlation attacks
4) Previous image attacks

A. Cryptographic Security

Initially, the cryptographic security of the pioneering models is analyzed in comparison with the proposed model. To do so, the number of strength bits present in each one is calculated. For the calculation, the N number of minutiae must coincide regarding the total of m minutiae to break the transformation is analyzed. This is an important aspect because it expresses the security provided by the model to the protected data.

1) Fuzzy Vault

This protection method was proposed to perform the encryption fingerprint minutiae templates. The authors propose that the method has a security level of 85 bits for a system with strong personal entropy. To validate this level of security, they argue that to unlock a fuzzy vault is necessary to answer 29 questions out of 32 that were properly insured originally. This means, in biometric terms, that out of 32 minutiae that are in the original set, 29 have to compare positively with the sample set. Given that this protection method is proposed for applications that do not contain a high number of users, then the main problem is the amount of data to be compared by the comparison method. This causes the cryptographic security model to be variable, depending on the amount of comparison data. For example, the worst scenario would be finding a match of 1 in 6 million records. In this example, the security of the method is estimated to be 33 strength bits because it depends on the interpolation of a set of original features in the vault.

1) Cancellable templates

In [11] the authors describe the cancellable templates model and perform an analysis of the security of each of the proposed transformations. The analysis is performed in the case of functional transformations because they are the most similar to the transformation proposed by the author. To perform the theoretical analysis is taken into account that 8 bits per minutiae are encoded, the minimum number of minutiae to compare positively is 15 and the total amount is 35 minutiae.

In this theoretical analysis a set of parameters about the alignment process is assumed. It is not necessary to estimate these parameters in the model proposed by the author because the model is invariant to rotation and translation. To calculate the strength of the method, the used function is:

\[ p = 8m - \log_2(N/m) \]  

The obtained results indicate that the method offers a 66 bit security.

1) Biohashing

This model contains significant differences when compared to fuzzy vault models and cancellable templates models. The input of this model is the image of the fingerprint, from which the FingerCode [13] is extracted, the Biocode is calculated and finally the biometric hash is obtained as proposed in [16]. The output of the model is a fixed-length binary vector that characterizes the fingerprint and it is used in the comparison process for recognizing a person.

In the analysis made by the authors mentioned before, it is not textually exposed the expression used to calculate the cryptographic security of the biometric hash model. An analysis of the model is made, based on the premise that it is needed the image of the fingerprint to perform the encryption, guaranteeing a safety of 384 bits of security.

B. Proposed Model

The validation of the increase in the cryptographic security of the proposed model is based on the comparative analysis of cryptographic security among different models such as fuzzy vault, cancellable templates biometric hash. For each transformation are encoded at least 32 bits of information, invariant to rotation and translation, the average amount of complex structures is 40 templates and at least must be compared 12 complex structures to obtain a similarity index greater or equal to the one proposed by the researcher. To calculate this probability, the proposed function is:

\[ p = 32m - \log_2(N/m) \]
Where \( p \) represent the scheme bits of strength. This expression calculates that the method has 354 bits of strength, which is considered a high level of security compared to the pioneering models. It is noteworthy that the strength is in correspondence with the number of bits to be encoded and the amount of complex structures formed. In the case the proposed model in this research, the amount of encoded bits is larger and the amount of structures are in correspondence with the number of minutiae contained in the plain text template. The amount of information to encode is three times bigger because from a single minutia only the coordinates \((x, y)\) and the \(\sigma\) angle are encoded.

**C. Brute Force Attacks**

This type of attack consists in guessing a finite set of identifiable characteristics, the sufficient amount to identify a person. Usually, the difficulty of performing a brute force attack is expressed in number of operations necessary for successfully reconstruct the biometric template. To estimate the strength of the proposed model, a theoretical analysis of different brute force attacks to fuzzy vaults models, cancellable templates and the proposed model is performed. To perform this attack, a randomly generated template is sent and the satisfaction criteria is evaluated. This criterion consists in checking how many elements of the generated dataset match the elements of the protected dataset. The attack ends when the criteria are met, indicating that the security of the model has been broken and obtaining a dataset which enables to impersonate a person.

1) Fuzzy Vault

In [16] an implementation of fuzzy vault model is performed and its cryptographic security is calculated using a key of 144 bits, of which 128 are used for encryption and 16 for code error correction. For this, it is calculated the combination of the number of elements which are real in the vault with the number of combinations of elements as shown in function:

\[
C(\text{totalelements}, \text{elementscombinations})
\]

This attack aims to identify genuine points and garbage points inside the vault, to find a polynomial interpolation to obtain the original data. To consider a successful attack on a vault with 18 original points and 200 garbage points, is estimated that an average of \(5.3 \times 10^{10}\) attempts are required to find the amount of original points for a positive comparison.

1) Cancellable templates

In the case of the cancellable templates model, it is only analyzed the functional transformations because they are the most similar to the transformation proposed by the author. The minimum number of minutiae to compare positively is 15 and the total number of minutiae is 35, which substituted in the above expression results in approximately \(3 \times 10^9\) attempts. Additionally, the statement discussed in [10] about the comparison methods to use, it is estimated that the probability of successfully making a brute force attack can be calculated by the function:

\[
G = \frac{N}{K \times d}
\]

where \( N \) represents the amount of minutiae that the template has, \( K \) and \( d \) represent the possible values that the coordinates and orientation can take. The probability to get a template that can be compared is estimated at 0.03125 per cent.

1) Proposed Model

In the case of the brute force attack analyzed by [17] in relation to the proposed model, it is estimated that in a template of 18 protected structures:

- Must be at least 6 elements of a triplet.
- Must be at least 190 triplets.

For this, it is calculated the combination of \(C(190, 6)\), resulting in \(6033463255 \approx 6.0 \times 10^{10}\) probabilities of finding a protected template that compares, considering that elements can be repeated within a triplet. In a perfect scenario where must be compared all the elements of a triplet and all the triplets formed in the structures as shown below:

- The 9 elements of a triplet
- The 342 triplets formed in the 18 structures.

The combination \(C(342, 9)\) is calculated, resulting in \(158625578809472060 \approx 1.5 \times 10^{17}\). This scenario is the one that best fits the one calculated by [17] so that in comparison, it could be stated that the probability of success in this type of attack in comparison with the proposed method is higher. In relation to the analysis made in the cancellable templates, the proposed model has the same amount of minutiae and complex structures, however, the complex structure contains more information than a minutia. In this case, a complex structure consists of 19 triplets which at the same time are composed of the data belonging to the three inner angles, three variations of the adjacent angles to a side and 3 sides. These are the data on which the comparison method is based to establish the similarity index between two protected templates. To calculate the probability of obtaining a set of complex structures that match the original set, the expression established is:

\[
G = N(a_i \times d a \times d)
\]

where \( N \) represents the amount of complex structures, \( a_i \) and \( da \) represent the possible values of the inner angles and the difference of adjacent angles respectively and \( d \) represents the possible values of the sides. As a result, it is obtained a chance of making a successful attack of \(0.4 \times 10^{25}\). Therefore, it can be stated that the probability of obtaining a trait using this attack is considerably low.

**V. CONCLUSION AND FUTURE WORK**

The protection of the biometric data used for people recognition is one of the most active research topics in the area of biometric identification. The minutiae templates protection models proposed so far, present a set of vulnerabilities that allow obtaining characteristics in plain text. The polynomial model protection for the protection of fingerprint minutiae templates, improves the cryptographic security of current models as long as only the result of the evaluation of each of the original points is stored. As future work, the implementation of the polynomial protection method will be made using various algorithms proposed by the author.
and a characteristics comparison method in the protected domain that meets the performance requirement is going to be proposed.

REFERENCES