ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 1, January 2013

Simple Hash Functions for Biometric Template Authentication

M.Mani Roja, Dr. Sudhir Sawarkar

Abstract*—The safety issues surrounding databases are mainly about safeguarding large and valuable collections of personally identifying information. If these databases are part of an important security system, then they are natural targets for attack, theft, compromise, and malicious or fraudulent use. When the reference information, captured during the enrollment phase, is not properly protected some privacy problems arise.This paper proposes an efficient authentication approach to protect the database templates using K means algorithm and signular values for the generation of hash values. These hash values will be used to verify the authenticity of stored templates.

Index Terms—**Hash Functions, K Means Algorithm, Singular Value Decomposition and Mean Square Error.**

I. INTRODUCTION

With the widespread deployment of biometric systems in various applications, there are increasing concerns about the security and privacy of biometric technology. Here, the term security is used to represent the overall reliability of the system, rather than just the simplistic notion of increased authentication accuracy brought about by the use of biometrics for verification/identification. Proper use of cryptography greatly reduces the risks in biometric system as the hackers have to find both the secret key and template. Unlike passwords and tokens, compromised biometric templates cannot be revoked and reissued[1]. Due to intra-user variability in the acquired biometric traits, ensuring the security of the template while maintaining the recognition performance is a challenging task. The security analysis of traditional password-based authentication schemes can be based on simple parameters such as

- Minimum length of passwords
- The password change period
- Inclusion of special characters

The security of card/token based system can be analyzed based on the parameters

- Illegal utilization of the token if the token is lost and found by intruder
- Ability to generate a token
- Ability to forge a token

A. Challenges in Biometric Security

Biometric systems are more complicated than the conventional schemes using password and token because of the following reasons.

- During every acquisition of the biometric data, there is a minor variation of the biometric is possible.
- Biometrics may need very good image enhancement schemes if the quality of the captured biometric sample is poor.

Proper use of cryptography greatly reduces the risks in biometric system as the hackers have to find both the secret key and template. Unlike passwords and tokens, compromised biometric templates cannot be revoked and reissued. Due to intra-user variability in the acquired biometric traits, ensuring the security of the template while maintaining the recognition performance is a challenging task [1].

B. Attacks on Biometric Templates

The biometric templates can undergo four possible vulnerabilities. They are

- An intruder who wants to gain unauthorized access can replace the existing template
- Unauthorized access can be gained using a physical spoof.
- The template can be stolen and can be replayed later to gain access.
- The templates can be used for cross-matching across different databases to covertly track a person without his/her consent.

Due to these reasons, the raw biometric images should not be stored in plaintext form and fool-proof techniques are required to securely store the templates such that both the security of the application and the users' privacy are not

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 1, January 2013

compromised by adversary attacks.

II. LITERATURE REVIEW

Almost all the commercial biometric systems secure the stored templates by encrypting those using standard cryptographic techniques. Either a public key cryptosystem like RSA or a symmetric key cipher like AES [1] is commonly used for template encryption. Since the above cryptosystems are generic, they can be directly applied to any biometric template and the encrypted templates are secure as long as the decryption key is secure. However, encryption is not a good solution for biometric template protection due to two main reasons. Firstly, encryption is not a smooth function and a small difference in the values of the feature sets extracted from the raw biometric data would lead to a very large difference in the resulting encrypted features [1]. Due to this reason, one cannot store a biometric template in an encrypted form and then perform matching in the encrypted domain. Hence, for every authentication attempt,

- The template is decrypted,
- Matching is performed between the query and decrypted template and
- The decrypted template is then removed from memory.

Thus, the template gets exposed during every authentication attempt. Secondly, the security of the encryption scheme depends on the decryption key. Hence, the decryption key needs to be securely stored in the system and if the key is compromised, the template is no longer secure. Because of these two reasons, standard encryption algorithms alone are not adequate for securing biometric templates and techniques that are designed to specifically account for the intra-user variability in the biometric data are needed.

Adler [2] used a "Hill Climbing Attack" to generate a face image from a face template. Hill [3] describes a masquerade attack wherein the fingerprint structure is determined using the minutiae template alone. Ross et al. [4] propose another technique to elicit the fingerprint structure from the minutiae template. They use Gabor-like filters suggested by Cappelli et al. [5] to generate fingerprints. Feng et al. [6] have also proposed a similar technique by modeling a fingerprint image as a 2D Frequency Modulation (FM) signal whose phase consists of the continuous part and the spiral part, which corresponds to minutiae. Vetro et al. [7] have discussed the application of distributed source coding techniques to biometric security, by using a Slepian-Wolf coding system to provide a secure means of storing biometric data that provides robust biometric authentication for genuine users and guards against attacks from imposters. Wang et al. [8] have presented a theoretical framework for the analysis of privacy and security tradeoffs in secure biometric authentication systems.

Yeung and Pankanti [9] describe an invisible fragile watermarking technique to detect regions in a fingerprint image that have been tampered by an attacker. Jain and Uludag [58] suggest the use of steganography principles to hide biometric data (e.g., fingerprint minutiae) in host images (e.g., faces). Ferri et al. [10] propose an algorithm to embed dynamic signature features into face images present on ID cards. Ferri et al. report that any modification of the face image can be detected, thereby disallowing the use of fake ID cards. Ratha et al. [11] propose the use of distortion functions to generate biometric data that can be cancelled if necessary. Uludag et al. [12] convert fingerprint templates (minutiae data) into point lists in 2D space, which implicitly hide a given secret (e.g., a 128-bit key). Mohapatra et al. [13] proposed a Biometric encryption method neither the key nor the original trait is stored, rather BE called biometric encrypted template is stored that contains the original template and as well as the key. Chander Kant et al. [14] presented a more secure system by use of steganography. Here the secret key (which is in the form of pixel intensities) will be merged in the picture itself while encoding, and at decoding end only the authentic user, who is aware of this specific, will be allowed to decode the encrypted image. Manvjeet Kaur et al.[64] develop a system to encrypt and decrypt the biometric image using helper data of a fingerprint and password to make it secure so that even if someone gains access to the encrypted image stored in the database, he will not able to reproduce the original image from it and it will be useless for him.

III. HASH FUNCTIONS FOR TEMPLATE PROTECTION

In cryptographic approach of template protection, since the matching can not be done in the encrypted form, we have to decrypt the templates before every authentication process. In large scale authentication, it very difficult to do this process in regular fashion. Hence we propose the Hash function based authentication approach. We generate a fixed block of data for each biometric template in the database.. On a regular interval basis, the hash values for each database templates will be calculated and verified with the stored Hash values. Any change in the hash value indicates that the template has been modified or replaced so that we can take the necessary actions.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 1, January 2013

A cryptographic hash function h takes as input a message (M) of arbitrary length and produces as output a message digest (Hash) of fixed length (H) in the form H = H (M). This hash value is appended to the message at the source. The receiver authenticates this message by recalculating the n hash value [15].

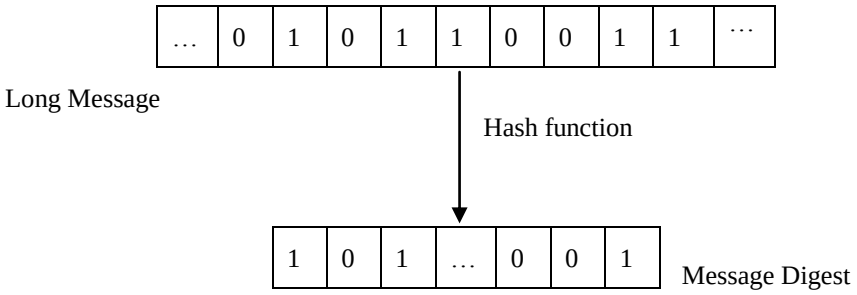

Fig. 1. Generation of hash function

A. Requirements of a Hash function

The purpose of a Hash function is to produce a 'fingerprint' of a file, message, or block of data. To be useful for authentication applications, a hash function H must have the following properties

- H can be applied to a block of data of any size
- H produces a fixed length output.
- H(M) is relatively easy to compute for any given message M, making both hardware and software implementations practical
- For any given value h, it is computationally infeasible to find M such that H (M) =h. This is known as one way property.
- For any given message M, it is computationally infeasible to find Y ≠ M such that H (Y) = H(M)

B. Simple hash Function

All hash functions operate using the following general principles. The input is viewed as a sequence of n bit blocks. The input is processed one block at a time to produce n bit Hash function. One of the simplest hash functions is the bit – by- bit exclusive OR of every block. This can be expressed as follows

$$C_i = b_{i1} \oplus b_{i2} \oplus b_{i3} \ldots \oplus b_{im}$$

where

$C_i = i^{th}$ *bit of the hash code*

$m =$ *number of* $n-$*bit blocks in the input*

$b_{ij} = i^{th}$ *bit in* j^{th} *block*

$\oplus =$ *XOR operation* (6.28)

This operation produces a simple parity for each bit position and is known as longitudinal redundancy check.
In our work, we have proposed the use of centroids values and singular values as hash values for the authentication of Biometric Templates. This hash values are of smaller size than the original image. Every image will have its unique hash value. Once, the enrollment process is over, we extract the hash value form the every template present in the database and these values are stored in the MySQL server. On regular basis, these hash values will be recalculated from the database templates, and will be compared with the hash valued of the associated templates in the server. If any one modified the template or deleted and replaced the templates, the hash value generated will not match with the stored hash values, we can come to know that the system is having security threats and we can take the necessary precautions. The procedure for extracting the hash values is given below

C. Kmeans based Hash value for authentication

K-means is one of the simplest unsupervised learning algorithms that solve the well known clustering problem. The procedure follows a simple and easy way to classify a given data set through a certain number of clusters (assume k clusters) fixed a priori. The main idea is to define k centroids, one for each cluster. Since the end result is a function of the centroid location, the location of the centroids is of utmost importance. Hence, a variation in the location of the centroids may give rise to a different result. So, the better choice is to place them as much as possible far away from each other. The next step is to take each point belonging to a given data set and associate it to the

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 1, January 2013

nearest centroid. When no point is pending, the first step is completed and an early group age is done. At this point we need to re-calculate k new centroids as barycenters of the clusters resulting from the previous step. After we have these k new centroids, a new binding has to be done between the same data set points and the nearest new centroid. A loop has been generated. As a result of this loop we may notice that the k centroids change their location step by step until no more changes can be done [16].

The steps for generating Hash values using K means approach are

- Place k points into the space represented by the objects that are being clustered. These points represent initial group centroids.
- Assign each object to the group that has the closest centroid.
- When all objects have been assigned, recalculate the positions of the k centroids
- Arrange the centroids as $N_1 N_2 N_3 \ldots N_k$. If there are 8 clusters, there will be 8 centoids vales.
- These k centroids will be stored in password protected MySQL server

D. SVD based Hash value for authentication

The singular value decomposition of image A is a decomposition of the form [17]

$$A = UDV^T \tag{3.8}$$

Where A is m x n matrix, U and V are orthogonal matrices. D is a diagonal matrix of singular values; the singular values $\sigma_1 \geq \sigma_2 \geq \sigma_3 \ldots\ldots \sigma_n \geq 0$ appear in descending order along the main diagonal of D. The singular values are obtained by taking the square root the of Eigen values of AAT and ATA. Hence the image A can be represented as

$$A = [U_1, U_2, \ldots U_N] \begin{bmatrix} \sigma_1 & 0 \ldots 0 \\ 0 & \sigma_2 \ldots 0 \\ \ldots\ldots & \\ 0 & 0 \ldots . \sigma_N \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \\ . \\ . \\ V_N^T \end{bmatrix} \tag{3.9}$$

The U and V vector are calculated as the Eigen vectors of AAT and ATA respectively. The square roots of the Eigen values are the singular values along the diagonal of the matrix D. Since our signature is resized to 128 x 256, we get a total of 128 singular values out of which we consider the first few values as hash values.

- Find out the singular values of the templates.
- Consider the top ‘k’ singular values.
- These k singular values will be stored as the reference hash values

E. Authentication procedure

During database authentication procedure, the hash values will be recalculated for the templates present in the database. let us consider these values X’as

$$X = \{x_1, x_2, x_3 \ldots . x_k\}$$

and he hash values stored in the server Y as

$$Y = \{y_1, y_2, y_3 \ldots . y_k\}$$

During the database authentication process, the mean square error between these two values will be calculated as

$$MSE = \sqrt{\sum_{i=1}^{k} (X_i - Y_i)^2} \tag{3.10}$$

If the value is less than the predefined threshold, then the template is not altered otherwise the template has been modified. The advantage of this method is that, the database templates are in their original form. Hence no need to do the process of keeping encrypted templates in the database and decrypt the templates during every authentication process. In this approach, only reference hash values are stored in MySQL server. On regular intervals, checking is done to find out any attacks on the database. Hence, no need to bring the database from the server to the workspace during every authentication. So this method performs faster than the previous methods.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 1, January 2013

The drawback of these methods is that the database is exposed to the adversary. To overcome this drawback, we are doing the regular checking; hence we can find out the database attack easily.

IV. RESULTS AND ANALYSIS

The first result in fig shows that authentication process when the comparison is done with the same template. The template has been declared as the original template.

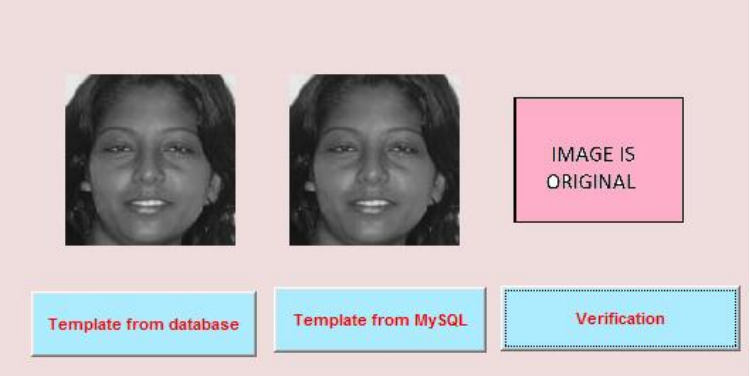

Fig.2. Successful Verification of Template

In the second case, the user template is compared with another sample of same user. But we were able to find that the template has been altered .

Fig 3. Detection of alteration with another sample of same user

In the third result, the user image is compared with some other user image. In both the case, our system identified the forgery and alerted the system. Detect that the original template has been replace.

Fig 3. Detection of alteration with sample of different user

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 1, January 2013

The mse values calculated for hash based protection schemes are listed in table. From the table, it is very clear that, any biometric attack on database templates will be noticed immediately since the mse values are two high

Table. 1. MSE values in Hash based authentication

Template compared	Kmeans hash	SVD hash
With original sample	0	0
With the sample of same user	367	3679
With sample of different user	530	7724

V. CONCLUSION

We have implemented a hash based authentication approach for the protection of database templates. Reference hash values were generated using K means approach and SVD approach and stored in MySQL server. At regular intervals, the hash values for the database templates were calculated and compared with the stored hash values. We were able to get successful results using both the approaches since we were able to detect the variations even if one user template is compared with another sample of same user.

ACKNOWLEDGMENT

The authors would like to thanks B.E EXTC students of TSEC to provide their facial images for testing.

REFERENCES

[1] Anil K. Jain, Karthik Nandakumar, and Abhishek Nagar, “Biometric Template Security”, EURASIP Journal on Advances in Signal Processing, Volume 2008, Article ID 579416.

[2] U. Uludag, S. Pankanti, S. Prabhakar, and A. K. Jain, “Biometric cryptosystems: issues and challenges,” Proceedings of the IEEE, vol. 92, no. 6, pp. 948–960, 2004.

[3] Advanced Encryption Standard, November 2001.

[4] A. Adler, “Can images be regenerated from biometric templates?” in Biometrics Consortium Conference, (Arlington, VA), September 2003.

[5] C. J. Hill, “Risk of masquerade arising from the storage of biometrics,” B.S. Thesis, Australian National University, November 2001, http://chris.fornax.net/biometrics.html.

[6] A. Ross, J. Shah, and A. K. Jain, “Towards reconstructing fingerprints from minutiae points,” in Proc. SPIE, Biometric Technology for Human Identification II, Vol. 5779, pp. 68–80, (Orlando, FL), March 2005.

[7] R. Cappelli, R. Erol, D. Maio, and D. Maltoni, “Synthetic fingerprint-image generation,” in Proc. Int’l Conf. Pattern Recognition (ICPR), Vol. 3, pp. 475–478, (Barcelona, Spain), September 2000.

[8] J. Feng, and A. K. Jain, "FM Model Based Fingerprint Reconstruction from Minutiae Template", Proc. International Conference on Biometrics (ICB), June, 2009.

[9] A. Vetro, S. C. Draper, S. Rane and J. Yedidia, Distributed Source Coding: Theory, Algorithms, and Applications, P. L. Dragotti and M. Gastpar (editors), Academic Press, 2009, pp. 293-324.

[10] N. Ratha, J. Connell, and R. bolle, “Enhancing security and privacy in biometrics-based authentication systems,” IBM Systems Journal, Vol. 40, No. 3, pp. 614–634, 2001.

[11] Y. Wang, S. Rane, S. C. Draper and P. Ishwar, "A theoretical analysis of authentication, privacy and reusability across secure biometric systems," to appear in IEEE Trans. Inform. Forensics Security.

[12] M. Yeung and S. Pankanti, “Verification watermarks on fingerprint recognition and retrieval,” in Proc. SPIE, Security and Watermarking of Multimedia Contents, Vol. 3657, pp. 66–78, (San Jose, USA), January 1999.

[13] U. Uludag and A. K. Jain, “Attacks on biometric systems: a case study in fingerprints,” in Proc. SPIE, Security, Steganography and Watermarking of Multimedia Contents VI, Vol. 5306, pp. 622–633, (San Jose, CA), January 2004.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 1, January 2013

[14] L. C. Ferri, A. Mayerhofer, M. Frank, C. Vielhauer, and R. Steinmetz, "Biometric authentication for ID cards with hologram watermarks," in Proc. SPIE, Security and Watermarking of Multimedia Contents IV, Vol. 4675, pp. 629–640, (Bellingham, WA), January 2002.

[15] N. Ratha, J. H. Connell, and R. M. Bolle, "An analysis of minutiae matching strength," in Proc. Audio and Video-based Biometric Person Authentication (AVBPA), pp. 223–228, (Halmstad, Sweden), June 2001.

[16] A. K. Jain and U. Uludag, "Hiding biometric data," IEEE Trans. Pattern Anal. Mach. Intelligence, Vol. 25, No. 11, pp. 1493–1498, 2003.

[17] A.K.Mohapatra, Madhvi Sandhu, Biometric Template Encryption, Published in International Journal of Advanced Engineering & Application, Jan. 2010.

[18] Chander Kant, Ranjender Nath & Sheetal Chaudhary Biometrics Security using Steganography", International Journal of Security, Volume (2) : Issue (1).

[19] Manvjeet Kaur, Sanjeev Sofat, Deepak Saraswat, Template and Database Security in Biometrics Systems: A Challenging Task International Journal of Computer Applications (0975 – 8887) Volume 4 – No.5, July 2010.

AUTHOR BIOGRAPHY

M. Mani Roja was born in Tirunelveli (T.N.) in India on June 19, 1969. She has received B.E. in Electronics & Communication Engineering from GCE Tirunelveli, Madurai Kamraj University in 1990, and M.E. in Electronics from Mumbai University in 2002. Her employment experience includes 22 years as an educationist at Thadomal Shahani Engineering College (TSEC), Mumbai University. She holds the post of an Associate Professor in TSEC. Her special fields of interest include Image Processing and Data Encryption. Currently, she is pursuing her PhD from Sant Gadge Baba Amravati University. She has over 20 papers in National / International Conferences and Journals to her credit. Ms. M.Mani Roja is a member of IETE, ISTE, IACSIT and ACM.

Sudhir Sawarkar was born in Amravati, Maharashtra in India on October, 1966. He received his BE (Electronics) and ME (Electronics) from Sant Gadge Baba Amravati University, India in 1988 and 1995 respectively. He received his PhD degree in 2007 from Dr. Babasaheb Ambedkar Technological University, Lonere, and Maharastra India. He is currently working as a Principal of Datta Meghe college of Engineering, Navi Mumbai. His special fields of interest include Image Processing and Neural networks. His employment experience includes 25 years in teaching. He has published more than 75 research papers in national / international journals /conferences. He has guided many ME dissertations. He is a recognized PhD supervisor in Amravati University and many other universities.