Biometric Cryptography with Deep Learning

Rati Kumari M20CS012[**]

[*] Department of Computer Science and Engineering, Indian Institute of Technology, Jodhpur, India

Abstract- Biometrics have been used by humans since the stone age, where handprints are used to represent the locality of a tribe. With the growth of the Internet, bio-metrics are now considered as reliable resources for user authentication. The conventional system requires a password for each device and the need to remember and update frequently makes it difficult for the end users. Public Key Cryptography (PKC) and bio-metric provides secure encryption and authentication. As bio-metrics are unique and time invariant this makes it reliable for verification and authentication for e-commerce, banking, medical records,etc. Bio-metric is identifying individuals by both physical and behavioural characteristics. The physiological characteristics are fingerprints, iris, retina, face, palm print, etc. and behavioural characteristics consist of signature, lip movement, speech, gait, gesture,etc. Multimodal biometric combines different biometric traits to enhance the security and robustness. The biometric along with the concepts of deep learning and homomorphic encryption can enhance security.

Index Terms- Biometric Authentication, homomorphic encryption, cancelable biometric, multimodal biometrics.

I. Introduction

Biometrics have been the area of research because of their features to identify individuals on the basis of both physical and behavioural characteristics. Features like face, iris, fingerprints, handprints, DNA, hand geometry, odor, retina, etc. are unique physiological characteristics whereas keystroke, signature, speech, gait, gesture comes under behavioural characteristics. Biometric encryption is a process where a random number is uniquely binded with the biometric template and authenticates users on the basis of their unique physical and behavioural characteristics.

Traditionally, passwords or token based systems have been used for encryption and authentication but biometrics are more convenient. But this convenience comes with the cost of safety of biometrics. The passwords once corrupted can be replaced with the new one but this is not the case with biometrics as we all have only a single copy available. So, to protect the privacy of users instead of storing biometric data, the biometric templates are considered and protected under ISO/IEC standard 24745. From biometrics scanner like fingerprint scanner the image is obtained and digital image analysis is performed. Afterwards the important features are extracted and the generated templates are stored on the server. The steps involved in biometric system are as follows:

Data Acquisition: Biometric data is obtained from the biometric scanner like fingerprint scanner, iris scanner,etc..

Feature Extraction: Digital Image Analysis is performed on the captured image and features are extracted.

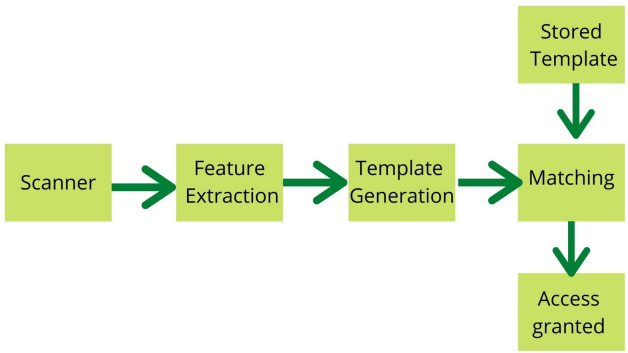

Figure 1. General working of biometric system

Template and Matching: Templates are generated for the extracted features and are stored in the enrollment phase. In the authentication phase the templates are matched with the existing stored templates and corresponding score is obtained.

Decision: As the biometrics are fuzzy so if the matching score is within the threshold limit then it is approved otherwise not.

Various terms involved in biometric cryptography are:

Cancelable Biometric: Cancelable biometric is like OWF (One Way Function) where if the biometric templates are corrupted the new template is generated and stored on the server. Biometrics are also stored after adding some randomness known as salting to ensure safety of templates. These templates follow properties like diversity, revocability, irreversibility and performance.

Biometric Cryptosystem: Biometric Cryptosystems involve adding key generation or some key binding with biometric template to prevent forgery.

Multimodal Biometrics: Using biometrics combination is much more effective like using fusion of Iris, fingerprints, handprints etc.than single biometric. This makes the system more secure and difficult for adversaries to obtain all biometrics.

Fuzzy Extractor: Biometrics are highly volatile unlike passwords where exact match is done. To overcome this fuzziness, threshold is taken and templates are matched with the threshold level. Even multiple instance / sample / sensors / algorithms are taken to enhance the accuracy of biometric templates and avoid spoofing attacks. Fusion of biometric is done at sensors, features, decision and score level.

Homomorphic Encryption: It is a different concept where the operations like addition and multiplication are performed on the encrypted data. It is of three types:

Partial: allows users to perform either addition or multiplication with an unlimited number of times on encrypted data.

Somewhat: allows both addition and multiplication but with a limited number of times on encrypted data.

Fully: allows both addition and multiplication with an unlimited number of times on encrypted data.

This encryption is useful in cloud computations where the computations are performed on the encrypted data and the computations are again encrypted and returned back to the end user. Various encryptions like ElGamal $(multiplication\ mod\ p)$, Paillier $(add\ mod\ N)$ use homomorphic concepts. It uses the public key cryptography concept with key generation, encryption and decryption steps.

Plain text $= \{0,1\}^n$
Cipher text $= \{0,1\}^n$
$(pk, sk) \leftarrow KeyGen(\$)$
$c \leftarrow Enc(pk, b)$
$b \leftarrow Dec(sk, c)$

Homomorphic encryption follow addition and multiplication property:

$$Enc(b^0) + Enc(b^1) = Enc(b^0 + b^1 mod\ n)$$
$$Enc(b^0) \times Enc(b^1) = Enc(b^0 \times b^1\ mod\ n)$$

or,

$$Dec(Enc(b^0) + Enc(b^1) = b^0 + b^1$$
$$Dec(Enc(b^0) \times k) = b^0 k$$

$$(pk, Enc(pk, 0)) \Leftrightarrow (pk, Enc(pk, 1))$$

means distribution is indistinguishability by efficient polynomial time algorithms.

If the adversary hacks the database server then she can easily access the encrypted templates. With the advancements in computing power and deep learning, it is possible to decode the template and generate actual iris or fingerprints from the extracted features. Some brief about deep learning concepts are:

CNN: Convolution Neural Network extracts the features from high quality images and converts them into lower dimension without losing important characteristics. It learns weights and biases when trained over a large dataset. CNNs are extensively used in face recognition and image classification.

Autoencoder: An autoencoder is a kind of artificial neural network that learns efficient data codings in an unsupervised manner. The main aim of an autoencoder is to learn important features from the data (encoding), and are typically used for dimensionality reduction.

GANs: Generative Adversarial Network is an unsupervised learning model in deep learning that learns the underlying features or patterns in input data in such a way that it can be used to regenerate new output that seems to be drawn from the original dataset. It works like a game where the generator generates the new output and the distinguisher differentiates between the fake and original image.

Transfer Learning: Transfer learning concept is used to utilize the knowledge obtained with training on one dataset to another dataset in the same domain. This concept is used to learn the features when trained on a large dataset. Now these learnt features are fine tuned on the new dataset in the same domain. Examples can be like learning important features from the face dataset and then fine tuning it for emotion detection. Transfer learning should answer questions like: What to transfer? When to transfer? How to transfer?

Few Shot Learning: As the name suggests here the training dataset is limited unlike other concepts where training dataset is large, features are learnt from the training dataset. Here, the model learns the discriminative features from the local patches of image followed by aggregating features for classifying data.

Machine learning with encrypted data is an important area of cryptography. As deep learning demands high computational power, it is not easily available. For this cloud computing is used where computations are sent to the cloud. As the third party can not be trusted so encrypted data is sent to the cloud and computations are done on encrypted data. These encrypted computations are sent back to the user where the data is decrypted using the user's private key.

The cryptographers have to be one step ahead of adversaries in order to protect the fundamental rights of privacy (GDPR) and simultaneously provide secure encryption and authentication. To provide secure encryption numerous steps are taken like instead of storing the template on the server it is stored locally, using PKC, homomorphic encryption, ZKP,etc.

The structure of this paper is as follows. Section 2 contains literature review that explains important work done in the field of cryptography. followed by Section 3 containing Analysis. Section 4 contains Conclusion, Acknowledgements and References.

II. LITERATURE REVIEW

Various encryption techniques have been developed involving biometrics from single biometric to multimodal biometric. Numerous steps are taken to provide comfort to the end user along with giving a secure encryption mechanism where users need not to remember passwords or tokens. Laptops, smartphones are designed with biometrics to provide secure authentication. The biometric has done rapid progress in the area of authentication. In this section the important work done in the biometric field is reviewed.

Mahesh Kumar Moramudi [1] performed an experiment involving Iris as biometrics and Homomorphic Encryption. As Iris is independent of hereditary DNA characteristics and is unique and stable so it is considered as a reliable biometric template. As most of the cryptosystem involves cancelable biometrics hence the biometric templates used by it are not accurate, so this paper highlights the use of highly accurate biometric by generating rotational invariant Iris template with reduced computation time. The proposed method used 2560 dimensional iris features with error rate of 0.19% for CASIA-V 1.0 dataset.

Jia-Chng Loh [2] worked on enabling cross-organizational biometric authentication service through secure sharing of biometric templates. This comprises distance recoverable encryption and secure distance computation. Paper explains the encrypt-then-split mechanism in which each organization holds only an encrypted partial biometric database. This minimizes the risk of template reconstruction if the encrypted partial database is recovered due to a leak of the encryption key. The benefit of PBio is that encrypted partial templates allow quicker rejection for non-matching instances.

Deshpande [3] experimented to enhance the accuracy of biometric recognition.. An optimum algorithm to appropriately choose weights has been suggested, which iteratively enhances the system accuracy and effectiveness of the system..

Marta [4] experimented with the fusion of Multi-biometric templates and applied Homomorphic probabilistic Encryption.

Minaee [5] did a survey on Biometrics Recognition systems using deep learning concepts and discussed the state of the art performance achieved using different concepts like CNN, GANs etc. by analyzing 120 promising works.

Durak [6] implemented a practical biometric authentication mechanism based on 3D Fingervein. In this work the fingervein of the user is obtained and important features are extracted and corresponding biometric templates are generated. These templates are stored securely on the server. In the authentication phase the biometric templates are matched and granted access if certain thresholds are fulfilled. This authentication system was successfully implemented in the Hospital.

Hataichanok [7] explains the concept of Multi-Modal Behavioural Biometric Authentication for Mobile Devices, where behaviour of individuals is analysed and used for continuous transparent authentication purposes.The results revealed that behavioral biometrics like observing keystroke dynamics and linguistic profiling could discriminate users with overall error rates of 20%, 20%, and 22%, respectively. Fusion methods like simple sum and weighted average are used, and this improved the classification performance with an overall EER 8%.

Ning Yu [8] experimented on GANs and its involvement in fingerprint matching. The results showed that even minor changes in training results in generation of different fingerprints. It is observed that the learnt fingerprint outperforms various models and gives better results.

Shashank Agrawal [9] and his team worked on Biometric enabled threshold authentication where biometric templates are stored on multiple devices of the user instead of storing it on the server. This paper has extensively used the latest encryption methods like homomorphic encryption, ZKP,etc.

III. ANALYSIS

Instead of using single biometric, multimodal biometrics are more effective for security purposes. Cancelable biometrics also enhance the use of biometric templates as the biometrics templates are stored on server instead of storing actual biometrics. When the biometric templates are stored on the server the adversaries can obtain these templates because of increased computation power. As the computations are performed on encrypted data in homomorphic encryption, the probability to guess the data by adversary is negligible. So, homomorphic encryption is popular encryption that is used to enhance security.

Another concern can be the advancement in deep learning concepts. Generative Adversarial Network GANs are highly popular because of their success in generating photo realistic images. This concept of generating similar images can be used to generate fake biometric templates. Since, the biometrics are extracted with fuzzy extractors and are matched with threshold values. If fake biometric matches the threshold level then it can lead to security breach.

IV. CONCLUSION

Biometric encryption and authentication is a vast area of research with the main aim to provide better facilities and protect the privacy of the user. Biometrics have shifted the conventional system of authentication where we need not to carry any token or remember the password. Hamming distance, Euclidean distance and other distance measures are used for finding distance between the biometrics Usually, different distance measures are used for different biometrics.

Some concerns can be like with the use of deep learning concepts

if the Isis or fingerprints can be regenerated then it can also be possible to generate fake templates that are close to the real one by using Generative Adversarial Network GANs. So, to avoid any misuse the cryptographers are already prepared by using Zero Knowledge Proof (ZKP) where there is no need to store any information.

Acknowledgment

Biometric encryption has developed a lot with the growth of the Internet and smartphones. I came to know about various terms and the algorithm behind secure encryptions. I would like to thank Dr. Somitra Kumar Sanadhya for continuous guidance and providing me the opportunity to explore this area and learn new concepts.

References

[1] Morampudi, M.K., Prasad, M.V.N.K. & Raju, U.S.N. Privacy-preserving iris authentication using fully homomorphic encryption. Multimed Tools Appl 79, 19215–19237 (2020). https://doi.org/10.1007/s11042-020-08680-5

[2] PBio: Enabling Cross-organizational Biometric Authentication Service through Secure Sharing of Biometric Templates, Jia-Chng Loh* , Geong-Sen Poh† , Jason H. M. Ying* , Jia Xu† , Hoon Wei Lim† NUS-Singtel Cyber Security Lab, https://eprint.iacr.org/2020/1381.pdf

[3] Deshpande, P.D., Mukherji, P. & Tavildar, A.S. Accuracy enhancement of biometric recognition using iterative weights optimization algorithm. EURASIP J. on Info. Security 2019, 6 (2019). https://doi.org/10.1186/s13635-019-0089-z

[4] Marta Gomez-Barrero, Emanuele Maiorana, Javier Galbally, Patrizio Campisi, Julian Fierrez,Multi-biometric template protection based on Homomorphic Encryption, Pattern Recognition,Volume 67,2017,Pages 149-163,ISSN 0031-3203, https://doi.org/10.1016/j.patcog.2017.01.024.

[5] Minaee, Shervin & Abdolrashidi, Amirali & Su, Hang & Bennamoun, Mohammed & Zhang, David. (2019). Biometric Recognition Using Deep Learning: A Survey. https://arxiv.org/pdf/1912.00271.pdf

[6] Durak, F. & Huguenin-Dumittan, Loïs & Vaudenay, Serge. (2020). \mathsf {BioLocker}$$: A Practical Biometric Authentication Mechanism Based on 3D Fingervein. 10.1007/978-3-030-57878-7_4, https://eprint.iacr.org/2020/453.pdf.

[7] Saevanee, Hataichanok & Clarke, Nathan & Furnell, Steven. (2012). Multi-Modal Behavioural Biometric Authentication for Mobile Devices. IFIP Advances in Information and Communication Technology. 376. 10.1007/978-3-642-30436-1_38.

[8] Yu, Ning & Davis, Larry & Fritz, Mario. (2019). Attributing Fake Images to GANs: Learning and Analyzing GAN Fingerprints. 7555-7565. 10.1109/ICCV.2019.00765.

[9] Shashank Agrawal and Saikrishna Badrinarayanan and Payman Mohassel and Pratyay Mukherjee and Sikhar Patranabis, Biometric Enabled Threshold Authentication, Cryptology ePrint Archive, Report 2020/679, 2020, https://eprint.iacr.org/2020/679

[10] Minaee, Shervin & Azimi, Elham & Abdolrashidi, Amirali. (2019). FingerNet: Pushing The Limits of Fingerprint Recognition Using Convolutional Neural Network.