Biometric Encryption in the Medical and Healthcare Industry

Group - Shinchan

Bindu Yaddula

Radha Gude

Harsha Ramireddy

Neeharika Jasthi

Alekhya Jillela

Joshua Koni

School of Computing and Engineering: University of Missouri – Kansas City

CS5596A Computer Security I: Cryptography

Prof. Sravya Chirandas

November 13, 2022

TABLE OF CONTENTS

No table of contents entries found.

Abstract

From the study it can be said that biometrics have played a crucial role in the healthcare and medical organization. Biometric recognition is done on the basis of pattern recognition technique that basically differentiates the persons on the basis of the fingerprint, face, iris, and voice and hand geometry. This method is used for confirmation and identification of the person. The technology of biometric was proven to be a compiled and secured method that prevents healthcare informatics from getting into the wrong hands thus preventing occurrence of any misconduct. With the use of this technology that helps in protecting the confidentiality and privacy of the patients. Machine learning is the traditional method used for biometric encryption. This process has been further advanced by the use of deep learning that makes the process fast and secure.

Introduction

Biometric recognition that refers to individual recognition based on psychological, behavioral and anatomical characteristics. Unlike the methods of usual identification that is focused on the belongings of the person (token, key, card) or knowing of the persons (PIN, Password), the process of biometrics which allows individual identification on the basis of who the person actually is. Biometric recognition is done on the basis of the technique of pattern recognition that differentiates individuals on the basis of a feature vector that is derived from the behavioral or psychological characteristics of the individual. These characteristics include fingerprint, retina, odor, face, iris, hand geometry, voice or palm print. The method of biometric is used for the confirmation or identification of the identity of the person. The methods of identification are used for the determination of the subject identity on the basis of biometric sample comparison of the biometrics samples that are obtained from the record set stored in the database.

The method of authentication is used for the identity confirmation that the individual claimed. Confirmation is done only on the basis of the biometric feature that is stored corresponding to the identity claimed.

Discussion

Biometric Technology

Accuracy in the technology of biometrics which depends on the system's ability in order to obtain a better quality image initially of the features of biometrics along with the matching of the individuals with original templates. The "False Rejection Rate (FRR)" and "False Acceptance Rate (FAR)" are basically affected due to placing incorrectly the features of biometrics, humidity, dirt, and other changes. Biometric technology has been used widely in the healthcare industry. This method is used to restrict or secure access to the medical facilities, protecting and managing the patient's identity, and confidential information, and helps in fraud reduction in the program of healthcare. A mechanism is provided by biometric technology for the purpose of identification or verification of the identity that depends upon the requirement of the organization for the protection of the information and resources.

The biometric provides the services of security that considers the flowing of sensitive information that are existing in the applications of large software and the resources that are required for the management of complex information which are accessed by thousands or hundreds of remote users. However, the technology of biometrics has proven to provide a secure and compiling method that restricts the access to health informatics and health facilities thus preventing the occurrence of any kind of misconduct. This method is quite useful for protecting the privacy of the patients and also their confidentiality. Several features of security are provided by the

biometric technology like the authentication process that is user-friendly and fast and controls access along with the ability for sensitive information encryption for both the environment of shared care and local applicants. The main focus of the study is exploring biometric technology uses and its security application in healthcare organizations. This focuses on secure architecture on the basis of biometric technology for access controlling and protecting the information that are sensitive enough in the electronic records of health in the environment of shared care.

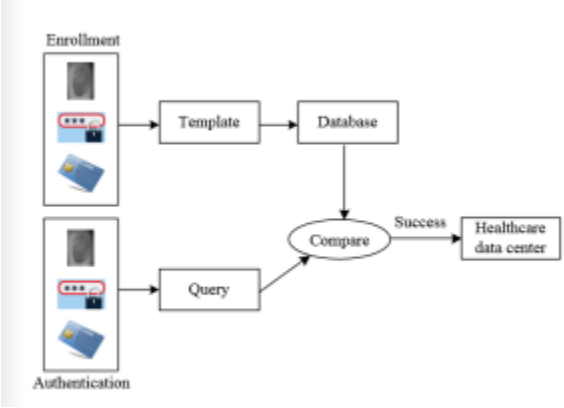

Figure 1: Framework for authentication system based on biometry

Biometric Measures

The automated method for the purpose of verifying and the identity recognition of an individual that are used mainly on two categories: Behavioral biometrics, Psychological biometrics. The behavioral biometrics that focuses on the actions of human beings. These measures of biometrics basically provide solutions that are fully proven covering the whole population and the new measures of biometrics that have proposed lip-print, dental radiography, tongue print and mouse dynamics. The measures of biometric that possess various characteristics that are used for several applications. The advantages of biometric measures are higher accuracy potential, imposters' resistance, stability on a long-term basis and fast proceeding. Some of the biometric measures are as follows:

- Universality: The particular biometric measure is applied on the whole population of the users. For the purpose of learning it is considered as consistent data for avoiding some of the issues of learning like bad training or overfitting.
- Uniqueness: The ability for successful discrimination of the population. This depends on the classify information ability. The data are not separable. However, some of the data can be separable non-linearly. The method of kernel that has solved various problems that have accounted for their criteria of k- separability.
- Cost-efficiency: The entire process is required to be cost efficient.

Biometric System and Machine Learning

The subject of machine learning that basically uses computers to stimulate the activities of human learning. The biometric system is the subject that studies about the features of biometric to stimulate the learning tasks for the identification of the individual. This field is developing continuously, many advancements are made in machine learning for the pattern recognition in biometric (Ortiz *et al*. 2018[5]). Some of the approaches of machine learning that are categorized into three types: Supervised learning, reinforcement learning and unsupervised learning, on the basis of classification, identification, clustering, recognition tasks and dimensionality reduction that are required for the development of the systems of biometrics.

- **Unsupervised Learning**: In this process a machine that basically receives input sequences in the form of x1, x2, x3, Where x denotes the input in the form of fingerprint edges numbers, graph representation of hand vein, eye distance, retina image and color and X= x is the set of samples. In this type of unsupervised learning the inputs are received simply by the machines and input representations are built and used for the decision making purpose, future input predictions, effective communications of the inputs on other

machines. This focuses on the pattern finding of the above data and beyond this it is considered as unstructured noise that is pure. Its main goals are dimensionality task reduction and clustering. The goal can only be achieved by using various algorithms and approaches as follows: k- mean, algorithm of expectation- maximization, approaches of Hebbian Learning and models of Gaussian Mixture. For the application of biometric, it mainly performs individual protection of data by the encryption of the biometric information, extraction of biometric data, fusion of feature level, behavioral detection of patterns.

- **Supervised Learning**: The supervised learning that provides a desired output sequence like y1, y2, y3... and the main focus of the machine is generating the correct output for the given input. It basically serves the final stage in the recognition system on the basis of biometrics. The main focus of supervised learning is on regression and classification. It was proved that supervised learning is utilized for fusion of biometric modalities, data classification of biometric and reliable regression and the multi- biometric system that is secure and successful (Saraswat *et al.* 2022[1]). From modern techniques, interesting results are obtained. The architecture of "Deep Neural Net (DNN)" is a developed method that involves the process of identity verification with an accuracy of 97. 35%. This method of learning leverages a large dataset level of the faces to obtain the representation of the faces in a generalized way.
- **Reinforcement Learning**: In this type of learning that involves the interaction of machines with the environment by action production a1, a2, a3 ... The environment state is affected by these actions, which results in scalar rewards received by the machines. This focuses on control system learning for the maximization of the numeric value that serves the objective

of the long- term. This learning method is based on the idea that if improvement is required for a particular action in the affair states, then action producing tendency is strengthened. Based on approaches of reinforcement learning that mainly focuses on task classification, training continuously by punishing signal or feedback reward, feature extraction and identification of discriminant or dominant features. This learning system is more versatile in comparison to the other two. However, it is only limited to the problems of low dimensionality. DRL was proven useful for problem solving purposes. However it is required to address various problems before the application of these techniques in the wider complexity of the issues of the real world.

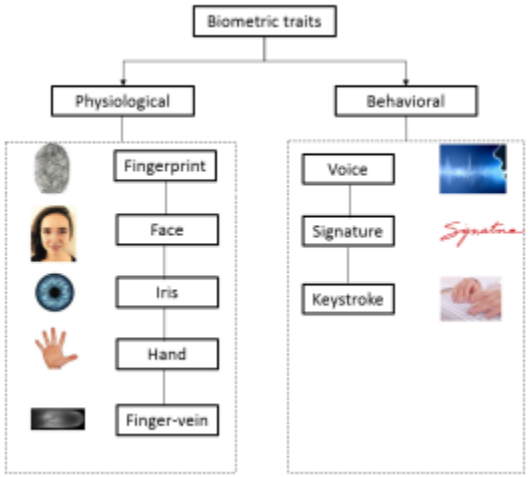

Figure 2: Biometric Traits classification

Proposed Cancellable Biometric Cryptosystem Based on Iris Images

Data encryption and data decryption are the two stages of the proposed iris-based cancellable biometric cryptosystem scheme's operation which is illustrated in fig 3.

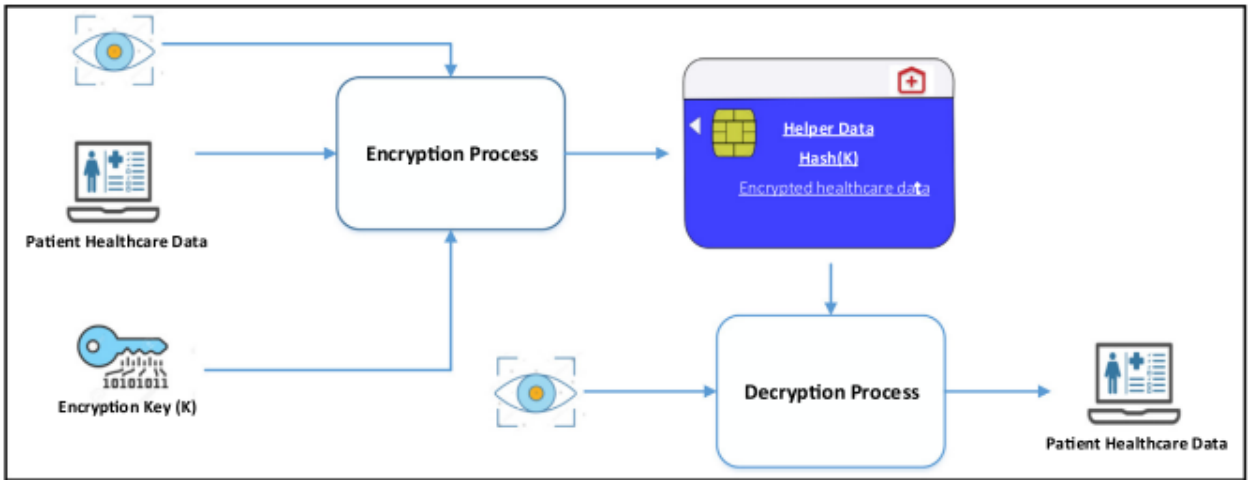

Figure 3: Architecture of iris-based cancellable biometric cryptosystem.

Encryption Process

In data encryption, the image of the iris, the patient's health data and the encryption key is taken as input and generates helper data, an encrypted version of the patient's data, and a hash of the encryption key, and all are stored in the patient's smartcard. The encryption key is linked to a cancelable transform of the iris-template using the fuzzy commitment technique of biometric cryptosystems, which prevents it from being utilized by unauthorized parties. The encryption process is displayed in fig 4.

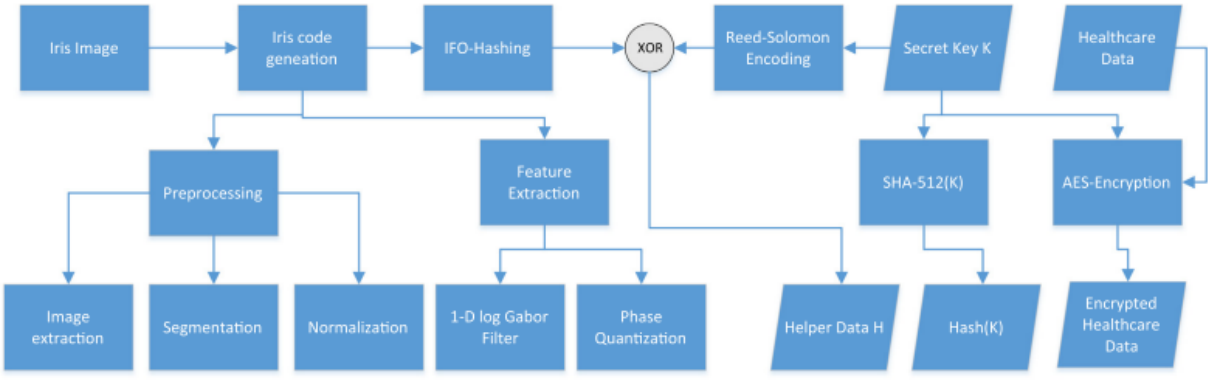

Figure 4: Encryption process

The encryption is bifurcated into 4 stages as shown below.

I. Generating the Iris Code:

Using a segmentation technique based on the Hough transform, the circular region of the iris and pupil is first separated from the rest of the input eye image. After that, a normalization process based on Daugman's rubber sheet model is used to the retrieved iris region. A 1D log-Gabor wavelet is convolved with the normalized iris pattern to encode the information. In the end, the phase data generated by 1D Log-Gabor filters is subjected to four level quantization to provide a bit-wise biometric template including the distinct iris model. The process is displayed in fig 5. The method for code generation is adopted from [7].

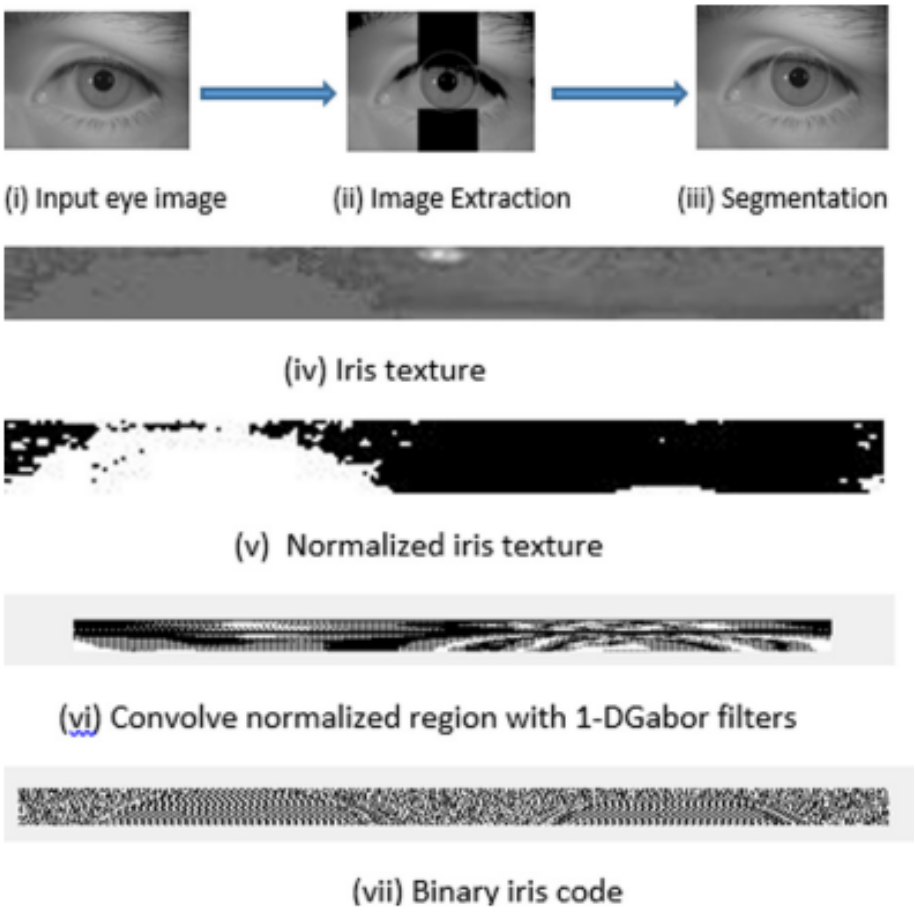

Figure 5: Generating the iris code

II. Generating the Cancellable Template:

The cancellable iris template creation is accomplished using the Indexing-First-One (IFO) hashing technique [8]. The generated iris code, which has a length of (n1 x n2), is fed into the IFO hashing method. The iris code's rows are put through "P" random permutations. A Hadamard product code is derived by multiplying all of the randomly permuted iris codes. The first binary "1" is then sought after using a k-size window in each row of the product code, and its index is recorded. Applying a modulo threshold function yields the hash code value. By iterating the similar technique 'm' times, an IFO hash code of length (n1 x m) can be generated.

III. Key Binding:

The encryption technique used is symmetric key encryption based, which uses the same key for both encryption and decryption. Secret key k is used to encrypt the patient's medical information. The template iris code I_p^T, as described in Section I above, is generated using the patient's (p) input eye image. The patient's (p) cancelable template IFO hash code C_p^T is created using the template Iris code I_p^T as input to the IFO hashing function.

$$C_p^T = \text{IFOHashing}(I_p^T)$$

The Reed-Solomon (n, k, m) encoding technique is used to encode the secret encryption key, K [9]. Helper data H is derived when the encoder's output is XORed using the IFO hash.

$$H = \text{RSEncoder}(K) \oplus C_p^T$$

IV. Encryption:

The secret key K and the AES algorithm [10] are used to encrypt the patient's medical information.

E = AES(K, HelathcareData)

Additionally, the secret key K is subjected to the SHA512 algorithm [11], resulting in the secret key K's cryptographic hash. The encryption process output results are encrypted healthcare data E, a hash of the secret encryption key K, and Helper data H.

Decryption Process

In data decryption, The image of the patient's iris and the smartcard data is used as an input and retrieves the encrypted data without the need for the encryption key. This is achieved by creating a decryption key from the given input and the secret key is not stored either on the smartcard and is not provided by the patient either. The data decryption is displayed in fig 6.

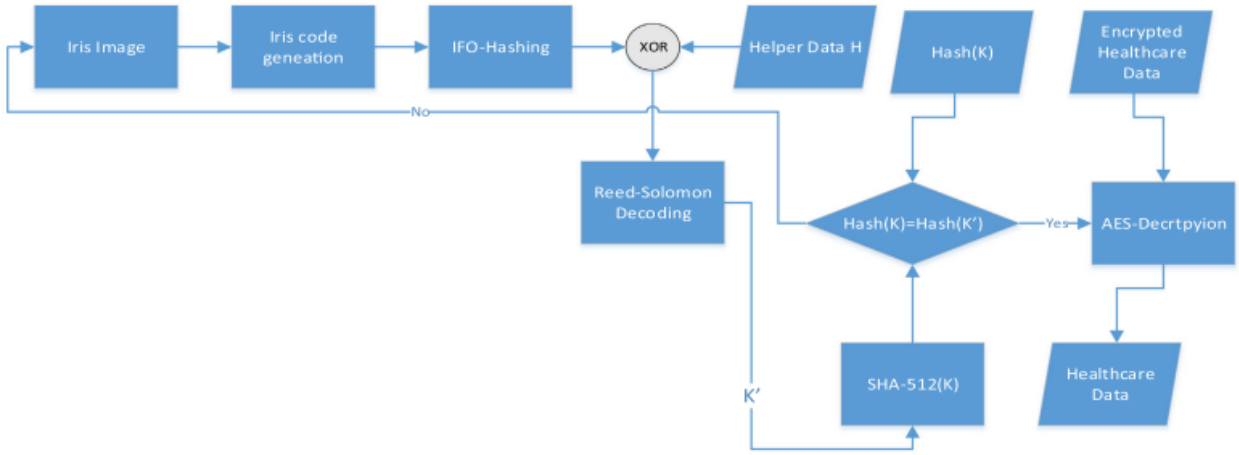

Figure 6: Decryption process

Decryption process is bifurcated into 3 stages: Generating iris code, Generating cancellable template, and Key unbinding. Generating iris code and cancellable template are same as in the encryption process. The query iris code I_p^Q is created using the patient's (p) query eye image. The

patient's (p) cancelable query IFO hash code $C_p^{\cap}$ is created using the query Iris code $I_p^{\cap}$ as input to the IFO hashing function.

$$C_p^{\cap} = \text{IFOHashing}(I_p^{\cap})$$

I. Key Unbinding:

On the helper data and IFO hash code, an XOR operation is carried out. The operation's output is decoded using the same block code that was employed during encoding. It's possible that the message that has been decoded is secret key, K.

$$K' = \text{RSDecoder}(C_p^{\cap} \oplus H)$$

Key K stored hash is compared to the computed cryptographic hash K'.

$$\text{SHA512}(K) = \text{SHA512}(K')$$

If both K and K' are identical, it is possible to use K' to decrypt the card's encrypted healthcare data E.

$$\text{HealthCareData} = \text{AES}(K', E)$$

The Table 1. The notations used here are displayed for reference.

Notations	Description
K	Secret Key
I_p^T	template iris code of patient P
C_p^T	IFO hash code of patient P
H	helper data
E	encrypted healthcare data
I_p^Q	Query Iris code of patient P
C_p^Q	query IFO hash code of patient P

Table 1. Notations and their descriptions

Experimented Results

CASIA-IrisV3-Lamp database is used in this experiment. The data is comprised of 20 left and right iris images of 411 individuals of each eye as features (40 x 411 = 16440 total images). The first and third left eye iris images of the first 100 individuals are used for model training and model testing purposes. False Rejection Rate is computed by using first and third images of same user as template and query respectively. False Acceptance Rate is computed by using first image of each individual and the third image of the rest of the individuals as template and query respectively. The accuracy of the model is calculated when FRR = FAR which explains the cryptosystem's performance.

RS code is comprised of three arguments n,k,m i.e; RS(n,k,m) where n = length of the block, k = length of the message, m = no of bits of each message. Table 2 shows the system performance for various RS codes. FRR is low and FAR is high for key lengths that are small and FAR is low and FRR is high for key lengths that are large.

System performance.

Coding Scheme	Key length(bits)	RS codeword length (bits)	FAR %	FRR%
RS (31,9,5)	45	155	0.8	2
RS(63,11,6)	66	378	0	3.5
RS (63,17,6)	102	378	0	5
RS (127,36,7)	252	889	0	7

Table 2. System performance

With a key length of 252, the maximum theoretical error-correcting capability for RS (127, 36, 7) is t = n-k/2 = 45.5 bits. Unlike the theoretical value, test results gave a max t value of 38 bits. The FRR and FAR are calculated for various values of t by keeping n = 127 unchanged, as depicted in Fig. 7. It is apparent that FRR is high, FAR is low for lower values of t.

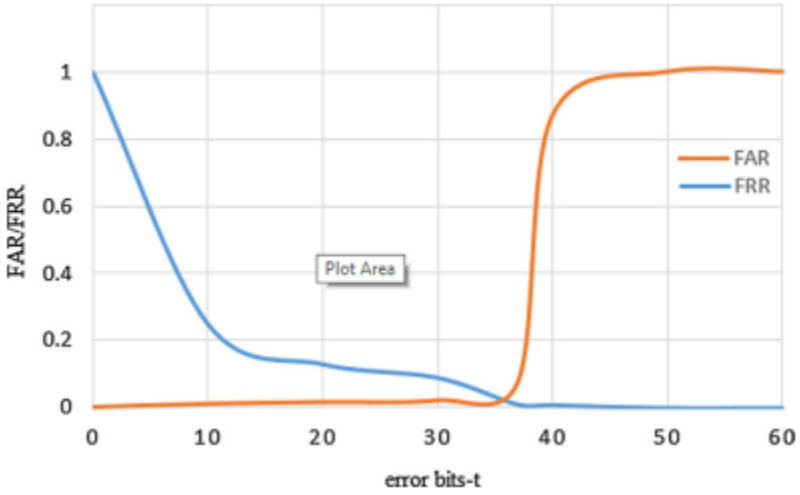

Figure 7: System's error correcting capability for n =127

Our proposed approach offers a superior FAR and FRR value with higher key lengths for use in practice with various cryptographic algorithms. As opposed to Yang et al[3] .'s key length of 63 bits with FAR of 0% and FRR of 63%, our suggested scheme's maximum key length is 252 bits with FAR of 0% and FRR of 7%. The comparison is shown in table 3.

Performance Comparison.

Our Proposed Scheme			W. Yang et al. [9]		
Key Length	FAR (%)	FRR (%)	Key Length	FAR (%)	FRR (%)
45	0.8	2	13	0	8
66	0	3.5	37	0	20
102	0	5	47	0	24
252	0	7	63	0	63

Table 3: Performance comparison of our system with W.Yang et al. [3]

Time Complexity:

In our proposed method, the patient healthcare data is encrypted with the AES technique and the hash of the encryption key is calculated with the SHA512 algorithm. Blocks in AES are always 128 bits in size. In the event that our message consists of "m" plain text blocks, the AES algorithm's time complexity is O(m). The SHA512 calculates the hash in roughly the same amount of time if fixed input sizes. So, the time complexity would be O (1).

Challenges for Electronic Health Record Security

The records of electronic health can be secured which is a costly and complex activity, where the information is maintained by the multiple actors. The security needs that are recognized globally are confidentiality, availability, accountability and integrity. The information availability is the key factor towards the electronic system of health records, the users with the information accessing rights should only be allowed to perform the duties. The key concept of “need- to- know" is applied here. Under this the users are allowed basically with the EHR of the patients to collect the required information to perform tasks keeping in mind the security and access policies of the

health organization (Flores Zuniga *et al.* 2010[4]). User authentication is an important step that guarantees that the access to information is modified and added by only that individual who has the privilege to do so. Appropriate balance between the information availability and security requirement is the main goal of the complex environment like healthcare. The addition of excessive mechanisms for security may result in less efficiency, be less user friendly and require more time are the important factors that are required for consideration. The important factor is confidentiality that drives the relationship between the doctor and the patients but is also concerned about the protection of the privacy of the patients. There are threats of security breaches towards the integrity of the electronic records of health and reliable information are provided for accountability purposes.

Conclusion

From the study it can be concluded that biometric encryption plays a crucial role in the healthcare and medical industry. It basically helps in the identification of the people on the basis of fingerprints, iris, retina, odor, face hand geometry, voice or palm print. The biometric method is basically used for confirmation and identification of the identity of a person. The subject identity is identified by the biometric sample comparisons that are stored in the form of data record sets. The technology of biometrics has proven to provide a secure and compiling method that prevents the access of health facilities and health information thus preventing the occurrence of any type of misconduct. It allows access of information only to the person who has the rights to access the information. This method helps in protecting the privacy and confidentiality of the patients. Machine learning has been used for the purpose of biometrics encryption.

References

[1]Saraswat, D., Ladhiya, K., Bhattacharya, P., & Zuhair, M. (2022, April). PHBio: A Pallier Homomorphic Biometric Encryption Scheme in Healthcare 4.0 Ecosystems. In *2022 3rd International Conference on Intelligent Engineering and Management (ICIEM)* (pp. 306-312). IEEE.

[2]Kausar, F. (2021). Iris based cancelable biometric cryptosystem for secure healthcare smart card. *Egyptian Informatics Journal*, *22*(4), 447-453.

[3]Yang, W., Wang, S., Hu, J., Zheng, G., Chaudhry, J., Adi, E., & Valli, C. (2018). Securing mobile healthcare data: a smart card based cancelable finger-vein bio-cryptosystem. *IEEE Access*, *6*, 36939-36947.

[4]Flores Zuniga, A. E., Win, K. T., & Susilo, W. (2010). Biometrics for electronic health records. *Journal of medical systems*, *34*(5), 975-983.

[5]Ortiz, N., Hernández, R. D., Jimenez, R., Mauledeoux, M., & Avilés, O. (2018). Survey of biometric pattern recognition via machine learning techniques. *Contemporary Engineering Sciences*, *11*(34), 1677-1694.

[6]Li, X., Jiang, Y., Chen, M., & Li, F. (2018). Research on iris image encryption based on deep learning. *EURASIP Journal on Image and Video Processing*, *2018*(1), 1-10.

[7]Masek, L. (2003). *Recognition of human iris patterns for biometric identification* (Doctoral dissertation, Master's thesis, University of Western Australia).

[8]Lai, Y. L., Jin, Z., Teoh, A. B. J., Goi, B. M., Yap, W. S., Chai, T. Y., & Rathgeb, C. (2017). Cancellable iris template generation based on indexing-first-one hashing. *Pattern Recognition*, *64*, 105-117.

[9]Plank, J. S. (1997). A tutorial on Reed–Solomon coding for fault-tolerance in RAID-like systems. *Software: Practice and Experience*, *27*(9), 995-1012.

[10]FIPS PUB 197, Advanced Encryption Standard (AES), (accessed online: March 1, 2020).

[11]FIPS PUB 180-4, Secure Hash Standard (SHS), [accessed online: Feb 24, 2020].