Privacy Preserving-Biometric Identification Scheme Over Cloud Encrypted Data

[1]R.Mytheli, [2] R.KaviPriya
[1] PG Scholar, Department of CSE, Kuppam Engineering College, Kuppam, Chittor District, A.P
[2] Assistant Professor, Department of CSE, Kuppam Engineering College, Kuppam, Chittor District, A.P

Abstract-

Biometric identification has become increasingly popular in recent years. With the development of cloud computing, database owners are motivated to outsource the large size of biometric data and identification tasks to the cloud to get rid of the expensive storage and computation costs, which however brings potential threats to users' privacy. In this paper, we propose an efficient and privacy-preserving biometric identification outsourcing scheme. Specifically, the biometric data is encrypted and outsourced to the cloud server. To execute a biometric identification, the database owner encrypts the query data and submits it to the cloud. The cloud performs identification operations over the encrypted database and returns the result to the database owner. A thorough security analysis indicates the proposed scheme is secure even if attackers can forge identification requests and collude with the cloud. Compared with previous protocols, experimental results show the proposed scheme achieves a better performance in both preparation and identification procedures.

Index terms **–** Biometric Authentication, Finger Recognition, Finger Extraction, Encryption

I. INTRODUCTION

Today Cloud Computing is becoming a hot trend in IT industries. Most of the enterprises are using cloud for storing and maintaining their huge data on cloud servers. But security of critical data over the cloud has become a concern for both cloud service users and providers. Traditional authentication mechanism like password, key generation, encryption mechanism has failed. Hackers are able to crack these passwords. So, the data is not secure until we have a secure mechanism to protect the data from intruders and hackers. In this paper, we are presenting a secure authentication mechanism unlike password or key which can't be hacked easily. Biometrics is an automatic identification of a person by using certain physiological features associated with the person. Biometrics data is unique for every individual. So our project aims at using Biometric data of user for the authentication process. Biometric System is a combination of sensors, feature extractor and matching modules which implements biometric recognition algorithms. The sensors scan the biometric trait of the user and produce its digital representation. A quality check is generally performed to ensure that the acquired biometric sample is reliable and can be processed by the subsequent feature extraction and matching modules. The feature extraction module will discard the useless and extraneous data present in the taken sample and extracts useful information called features that can be used for matching. During matching, the query biometric sample is matched with the reference information which is stored in the database to establish the identity associated with the query. This operation is done in two stages, first is the Enrolment and second is the recognition. In Enrolment stage the biometric information of the person is stored in the database. We are implementing our project to match fingerprint data of user for authentication in cloud. We will store the users fingerprint data in compressed form on a cloud database for the time and use that for matching whenever a user tries to login the next time. We are using Biometric scanner to extract fingerprint of user. Fingerprint data will be transmitted in the compressed from for security of users Biometric data. There is a matching module to match the fingerprints against the one stored on the database. If the fingerprint matches, it will allow the registered user to login. Since it will be an overhead for the cloud service providers, our project aims at creating a separate web client between user and cloud service provider to provide a secure service of Biometric Authentication

This paper is organized in five sections. After this introduction, in Section II, literature survey discussed of the paper, section III about the System Analysis, Section IV about System Design, as well as the novel feature of the proposed method. Finally, Sections V and VI provide the simulation results and the conclusions, respectively.

II. LITERATURE SURVEY

This section of Literature Survey eventually reveals home facts of Biometric Authentication based on the analysis of many authors work as follows :

Chandra Shekhar Vorugunti [1] has introduced a new concept of BioAaaS to maintain secure authentication. Based on SAAS model of Cloud it provides a light weight and secure authentication mechanism. It contains two steps for authentication. First is Enrolment and next is

Verification. In Enrolment process the biometric data is converted into a binary form. The feature extractor then converts the binary string into a set of features. In verification process same
process will be processed when the user logins to the cloud. The matching module matches the features of the stored data and login data. Thus they have provided a service to do heavy weight cryptographic encryption and decryption operation on user's biometric data.

D J Craft [2] reports on fast hardware implementation of lossless data compression algorithms. It proposed Adaptive Lempel-Ziv Algorithm (LZ1 & LZ2). LZ algorithm are symbol based that is they operate one data one character at a time. They achieve compression by locating frequency occurring sequences of such symbol in input data stream. ALDC have two extensions as BLDC & CLDC. BLDC pre-processing works well on only bitmapped image data. CLDC is combination of ALDC & BLDC. The main difference between LZ1 & LZ2 is in the data structure employed & the way reference to sequence are coded. Cong Li et al. [3] proposed Burrow Wheeler Transformation based DNA sequence data multi-compression using Open MP & MPI. They proposed data compression (DNA sequence) using fewer bits rather than encoded data to represent information. BWT based DNA compression includes few steps.

First DNA sequence data is encoded with 0/1 which has 4 characters. Then BWT transformation is performed over it. Again MTF transformation is performed. Then we compress data with classical algorithm.

Kiran Kumar K et al. [4] have described that there are two properties of fingerprint namely uniqueness and permanence that are used for identification and verification. These properties are judged by minutae and ridges. The method used in this paper has 8 stages. They are gray-level fingerprint image, binarization, thinning, minutae extraction, false minutae, matching scores, ridges extraction, minutae and ridge score fused using strength factor. The block filters preserve the outermost pixels along each ridge.

Jeff Collier [5] proposed a system for developing a software that would address the challenges such as in-memory map/reduce. It also deals with the node that has ability to leave and re-join the cloud by applying compression and image processing algorithms.

Hu Chun et al. [6] have proposed a situation where biometric data is kept encrypted in whole process of transmission and matching. It uses two approaches homomorphism encryption and garbled circuit. It provides highly computing capability.

Surender Sharma et al. [7] have introduced health care monitoring system application that provides the patients with necessary healthcare information yet it also gives a chance to threats of intervention that would make the critical data insecure. They have used Body Area wireless sensor network as monitoring component. The cloud based HMA was therefore developed using master slave like pattern, where the master could have generic functions while slave would have functionalities specific to the medical condition. Thus they have utilized biometric encryption for providing protection to the data. Here the user's biometric characteristics work as decryption key here fuzzy extractor scheme has been used to convert the scanned fingerprint data to some random string and an helper string to apply cryptographic techniques. This framework accomplishes both the goals, secure access and data protection.

Krishnaraj Madhavji Sunjiv Soyjaudah [8] discussed about eight points of vulnerabilities that can be hacked. In cloud data is moved dynamically so security is a major concern and there are the problems which arise in the management of biometric data. So a cancellable biometric authentication system is proposed by him.

Cancellable Biometric Authentication is a concept in which the original image is first distorted then shared on cloud. This distorted biometric image is used for authentication. This provides security and privacy as the original biometric are never revealed to authentication server. Data Hiding is also done using this technique to overcome the replay attack. This is done by secretly embedding the private information in biometric image.

Dr. Anandhakumar P et al. [9] has addressed the issues that arouse during storing different documents and files and photo contents on Cloud. Huge amount of photos are maintained by cloud providers. Huffman coding cannot achieve high levels of compression also all the binary strings or codes in the encoded data are of different lengths. So it is difficult to decode. The representative signal(RS) based approach is suitable only when images are highly correlated to each other so it fails badly in case of illumination changes. To overcome these drawbacks LZ-77 algorithm is proposed in this paper. Lz-77 replaces the repeatedly occurring data with reference to single copy which is already existing in an uncompressed data stream. It uses length-distance pair to encode the match. As compression is in
cloud environment, k-means algorithm is used to transfer up by using Map-Reduce concept. They also proposed an idea for effective compression of photo albums which also reduces the time complexity.

III. SYSTEM ANALYSIS

This section introduces the system model, attack model, design goals and the notations used in the following sections.

A. SYSTEM MODEL

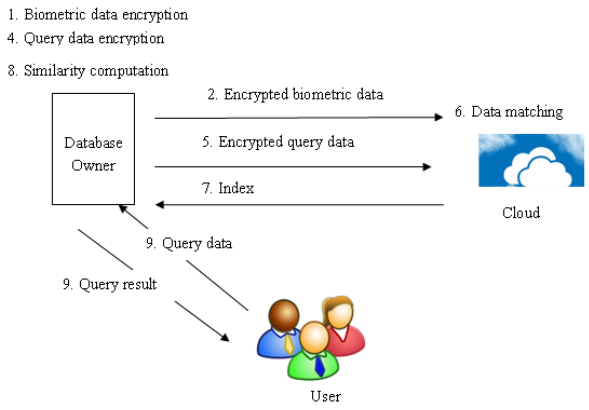

Figure 1: System model

As shown in Fig.1, three types of entities are involved in the system including the database owner, users and the cloud. The database owner holds a large size of biometric data (i.e., fingerprints, irises, voice, and facial patterns etc.), which is encrypted and transmitted to the cloud for storage. When a user wants to identify himself/herself, a query request is be sent to the database owner. After receiving the request, the database owner generates a ciphertext for the biometric trait and then transmits the ciphertext to the cloud for identification. The cloud server figures out the best match for the encrypted query and returns the related index to the database owner. Finally, the database owner computes the similarity between the query data and the biometric data associated with the index, and returns the query result to the user.

In our scheme, we assume that the biometric data has been processed such that its representation can be used to execute biometric match. Without loss of generality, similar to [17], [18], we target fingerprints and use Finger Codes [19] to represent the fingerprints. More specifically, a Finger Code consists of n elements and each element is a l-bit integer (typically n = 640 and l = 8). Given two Finger Codes x = [x1; x2; _ _ _ ; xn] and y = [y1; y2; _ _ _ ; yn], if their Euclidean distance is below a threshold _, they are usually considered as a good match, which means the two fingerprints are considered from the same person.

B. ATTACK MODEL

First of all, the cloud server is considered to be "honest but
curious" as described in [13]–[15], [17]. The cloud strictly follows the designed protocol, but makes efforts to reveal privacy from both the database owner and the user. We assume that an attacker can observe all the data stored in the cloud including the encrypted biometric database, encrypted queries and matching results. Moreover, the attacker can act as a user to construct arbitrary queries. Thus, we categorize the attack model into three levels as follows:

- Level 1: Attackers can only observe the encrypted data stored in the cloud. This follows the well-known cipher text-only attack model [20].
- Level 2: In addition to the encrypted data stored in the cloud, attackers are able to get a set of biometric traits in the database D but do not know the corresponding ciphertexts in the database C , which is similar to the known-candidate attack model [21].
- Level 3: Besides all the abilities in level-2, attackers in level-3 can be valid users. Thus, attackers can forge as many identification queries as possible and obtain the corresponding ciphertexts. This attack follows the known-plaintext attack model [20].

A biometric identification scheme is secure if it can resist the level attack. Note that that if the proposed scheme can resist level-2 and level-3 attacks, it does not mean that the attacker can both be the valid user and observe some plaintexts of the biometric database simultaneously. This sophisticated attack is too strong and no effective methods are designed to defend against this kind of attack [14]. In this paper, we focus on the collusion attack between a malicious user and the cloud server. The relationship between the plaintexts of the biometric database and the ciphertexts is not known to the attacker, which is similar to the attack model proposed in [14].

C. DESIGN GOALS

In order to achieve practicality, both security and efficiency are considered in the proposed scheme. To be more specific, design goals of the proposed scheme are described as follows:

- Efficiency: Computational costs should be as low as possible at both the database owner side and the user side. To gain high efficiency, most biometric identification operations should be executed in the cloud.
- Security: During the identification process, the privacy of biometric data should be protected. Attackers and the semi-honest cloud should learn nothing about the sensitive informatio

D. NOTATIONS

bi - the i-th sample fingercode,denoted as an n-dimensional vector $b_i=[b_{i1},b_{i2},,,,,b_{in}]$.

Bi – the extended sample fingercode of bi, denoted as an (n+1) dimensional vector $B_i=[b_{i1},b_{i2},,,,bi_{(n+1)}]$, where $b_{i(n+1)}=-0.5(b_{i1}{}^2+bi22+….+b^2im]$.

b_c- the query FingerCode denoted as an n- dimensional vector $b_c=[b_{c1},b_{c2},,,,,b_{cn}]$.

Bc – the extended sample fingercode of bi, denoted as an (n+1) dimensional vector $B_c=[b_{c1},b_{c2},,,,bc_{(n+1)}]$, where $b_{c(n+1)}=1$.

III. SECURITY ANALYSIS OF YUAN AND YU'S SCHEME

In this section, we firstly describe Yuan and Yu's scheme and then give the security analysis about their scheme. To facilitate understanding of the scheme, we use ∗

to denote the elements multiplication operations, and use × to denote the matrices or vectors multiplication operations.

A. YUAN AND YU'S SCHEME

Step 1: The database owner randomly generates an (n+1) x (n + 1) matrix A where $HxA^T=1$ and Ai is a row vector in A,1< i< (n+1). Then, the database owner generates a corresponding matrix $Di = [A_i^T * b_{i1}, \dots A_2^T * b_{i2} \dots A_{n+1}^T * b_{i(n+1)}$ to hide Bi. After the database owner performs the following operations

$$Ci = M_1 \times D_i \times M_2 \quad (1)$$

$$C_h = H \times M_i^{-1} \quad (2)$$

$$. C_r = M_3^{-1} \times R^T \quad (3)$$

Subsequently, the database owner uploads (Ci; Ch; Cr; Ii) to the cloud, where Ii is the index of Bi.

Step 2: After Step 1 is executed, the cloud has stored many tuples in its database C. When a user requests to identify his/her identity, he/she extends bi and then submits the extended query Bi to the database owner. On receiving the request from the user, the database owner generates a random (n + 1) x (n + 1). matrix E such that Ei x RT = 1, where Ei is a row vector in matrix E and 1 < i _< (n + 1). The database owner then generates a corresponding matrix $F_C = [E_{T1} \times b_{c1}, E_{T2} \times b_{C2} \dots . E_{Tn}$.to hide the query Finger Code Bc. The Database owner then performs the following operations:

$$C_f = M_2^{-1} \times F_c \times M_3 \quad (4)$$

Then, the database owner uploads Cf to the cloud.

Step 3: On receiving Cf , the cloud begins to search for the best match. Specifically, the cloud computes Pi = ChxCixCfxCr for all encrypted biometric database to compare the Euclidean distances between bc and bi. Other details are eliminated since they are irrelevant for the security analysis we will describe.

B. SECURITY ANALYSIS OF YUAN AND YU'S SCHEME

In level-3 attack, an attacker has the ability to select query FingerCodes Γ of his/her interest as inputs and then tries to recover the privacy of Bi . Specifically, the attacker can compute the secret key M2 by performing the following equation:

$$C_f \times C_r = M_2^{-1} \times F_c \times M_3 \times M_3^{-1} \times R^T$$
$$= M_2^{-1} \times F_C \times R^T$$
$$= M_2^{-1} \times B_C^T \quad (5)$$

In equation 5, Cf is an (n + 1) _ (n + 1) matrix and Cr is an (n + 1)-dimensional vector which are both known to the attacker. Bc is an (n + 1)-dimensional vector which can be constructed by the attacker.

IV. A NOVEL BIOMETRIC IDENTIFICATION SCHEME

In this section, we show the details of the proposed biometric identification scheme.

A. Overview

We construct a novel biometric identification scheme to address the weakness of Yuan and Yu's scheme [13]. To achieve a higher level of privacy protection, a new retrieval way is constructed to resist the level-3 attack. Moreover, we also reconstruct the ciphertext to reduce the amount of uploaded data and improve the efficiency both in the preparation and identification procedures. In the remaining part of this section, we will introduce the preparation process and the identification process.

B. Preparation Process

In the preparation process, bi is the i-th sample feature vector derived from the fingerprint image using a feature extraction algorithm [19]. To be more specific, bi is an n-dimensional vector with l bits of each element where n = 640 and l = 8. For ease of identification, bi is extended by adding an (n+ 1)-th element as Bi. Then, the database owner encrypts Bi with the secret key M1 as follows:

$$C_i = B_i \times M_1 \quad (6)$$

The database owner further performs the following operation:

$$C_h = M_2^{-1} \times H^T \quad (7)$$

Each FingerCode Bi is associated with an index Ii. After execute the encryption operations, the database owner uploads (Ci;Ch; Ii) to the cloud.

C. Identification Process

The identification process includes the following steps:

Step 1: When a user has a query fingerprint to be identified, he/she first gets the query FingerCode bc derived from the query fingerprint image. The FingerCode bc is also an n dimensional vector. Then, the user sends bc to the database owner.

Step 2: After receiving bc, the database owner extends bc to Bc by adding an (n + 1)-th element equals to 1. Then the database owner randomly generates an (n + 1) x (n + 1) matrix E.

Step 3: After receiving Cf from the database owner, the cloud begins to search the FingerCode which has the minimum Euclidean distance with the query FingerCode Bc. Pi denotes the relative distance between Bi and Bc

Step 4: After receiving the index Ii, the database owner gets the corresponding sample FingerCode bi in the database D and calculates the accurate Euclidean distance between bi

Step 5: Finally, the database owner returns the identification result to the user.

V. SECURITY ANALYSIS

In this part, we first prove that our scheme is secure under level-2 and level-3 attacks, and then we will show the

proposed scheme can resist the attack proposed by Zhu et al [18].

A. Security Analysis Under Level-2 Attack

According to the attack scenario 2, an attacker can obtain some plaintexts of the biometric database, but does not know the corresponding ciphertexts.

B. Security Analysis under Level-3 Attack

In the level-3 attack, besides the knowledge of encrypted data in the cloud, the attacker can forge a large number of query FingerCodes as inputs. In the following, we will show the proposed scheme is secure by proving that the secret keys cannot be recovered.

C. Security Analysis under the Attack Proposed by ZHU ET AL.

Zhu et al. [18] showed an attack for Yuan and Yu's scheme. In their attack, the attacker observes the cloud and gets the values of relative distance.

VI. PERFORMANCE ANALYSIS

To evaluate the performance of the proposed scheme, we implement a cloud-based privacy-preserving fingerprint identification system. For the cloud, we use 2 nodes with 6-core 2.10 GHz Intel Xeous CPU and 32GB memory. We utilize a laptop with an Intel Core 2.40 GHz CPU and 8G. Similar to [13] and [14], the query FingerCodes are randomly selected from the database which is constructed with random 640- entry vectors.

A. Complexity Analysis

In this work, each matrix multiplication costs O(n3), where n denotes the dimension of a FingerCode, and the sorting cost of fuzzy Euclidean distances has time complexity of O(mlogm). As illustrated in Table 2, our scheme has lower complexities in the preparation phase. That is, more computation and bandwidth costs can be saved for the database owner. In the identification phase, the computation complexity of our scheme is lower than that in [14]. The reason is that our scheme performs vector-matrix multiplication operations to find the close match, while [14] needs to execute matrix matrix multiplication operations. Although the complexity of our scheme is the same as that in [13], we emphasize that [13] sacrifices the substantial security to achieve such fast computation of Pi. Moreover, our scheme executes fewer multiplication operations, and thus obtains better performance.

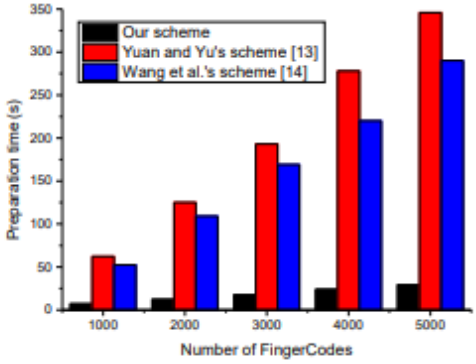

Figure 2: Time costs in the preparation phase

B. Experimental Evaluation

Preparation phase. Fig. 2 and Fig. 3 show the computation and communication costs in the preparation phase with the number of FingerCodes varying from 1000 to 5000. As shown in Fig.2, in our scheme, registering 5000 FingerCodes needs 29.37s, which can save about 88.85% and 90.58% time cost compared with [13] and [14] respectively. The reason is when encrypting a sample FingerCode, in our scheme, only one matrix is needed which leads to fewer matrix multiplication operations. Fig. 3 shows the bandwidth costs of the three schemes. Since the data outsourced to the cloud is in the form of vectors in comparison with matrices in the other two schemes, the communication cost in our scheme is much less than [13], [14]. Identification phase. Fig.4 and Fig. 5 show the computation and communication costs in the identification phase with the number of FingerCodes ranges from 1000 to 5000. As demonstrated in Fig. 4, all schemes grow linearly as the size of database increases. As in our scheme fewer matrix multiplication operations are used than [13], it can save about 56% time cost. Compared with [14], the identification time can be saved as much as 84.75%, since the vector-matrix multiplication rather than the matrix-matrix multiplication operation is executed. The bandwidth costs of the three schemes, as shown in Fig. 5, are almost the same. The reason is that all schemes need to transmit a matrix in the identification phase.

Figure 3: Bandwidth costs in the Preparation phase

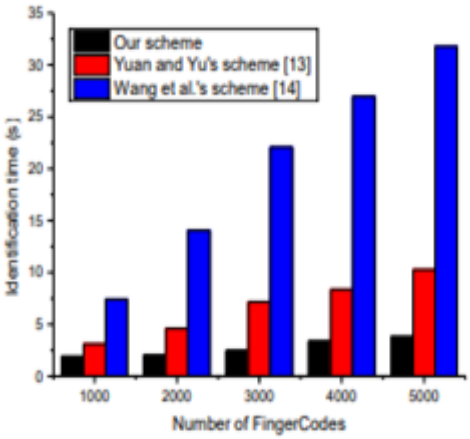

Figure 4: Time costs in the identification phase

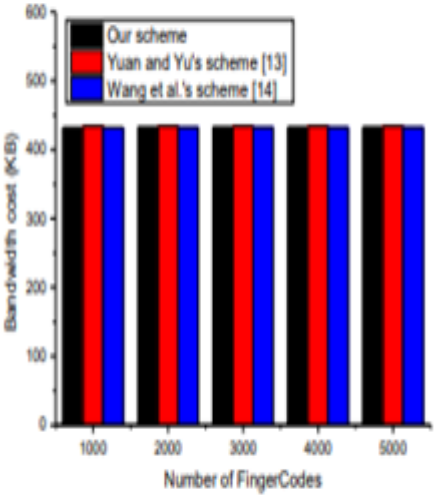

Figure 5: Bandwidth costs in the identification phase

VII. CONCLUSION

In this paper, we proposed a novel privacy-preserving biometric identification scheme in the cloud computing. To realize the efficiency and secure requirements, we have designed a new encryption algorithm and cloud authentication certification. The detailed analysis shows it can resist the potential attacks. Besides, through performance evaluations, we further demonstrated the proposed scheme meets the efficiency need well.

REFERENCESS

[1] A. Jain, L. Hong and S. Pankanti, "Biometric identification," Communications
of the ACM, vol. 43, no. 2, pp. 90-98, 2000.

[2] R. Allen, P. Sankar and S. Prabhakar, "Fingerprint identification technology," Biometric Systems, pp. 22-61, 2005.

[3] J. de Mira, H. Neto, E. Neves, et al., "Biometric-oriented Iris IdentificationBased on Mathematical Morphology," Journal of Signal Processing Systems,vol. 80, no. 2, pp. 181-195, 2015.

[4] S. Romdhani, V. Blanz and T. Vetter, "Face identification by fitting a3d morphable model using linear shape and texture error functions," inEuropean Conference on Computer Vision, pp. 3-19, 2002.

[5] Y. Xiao, V. Rayi, B. Sun, X. Du, F. Hu, and M. Galloway, "A survey of keymanagement schemes in wireless sensor networks," Journal of ComputerCommunications, vol. 30, no. 11-12, pp. 2314-2341, 2007.

[6] X. Du, Y. Xiao, M. Guizani, and H. H. Chen, "An effective key managementscheme for heterogeneous sensor networks," Ad Hoc Networks, vol. 5, no.1, pp. 24-34, 2007.

[7] X. Du and H. H. Chen, "Security in wireless sensor networks," IEEEWireless Communications Magazine, vol. 15, no. 4, pp. 60-66, 2008.

[8] X. Hei, and X. Du, "Biometric-based two-level secure access control for implantablemedical devices during emergency," in Proc. of IEEE INFOCOM2011, pp. 346-350, 2011.

ABOUT AUTHORS

1. Miss R.Mytheli received B.Tech degree from Kuppam Engineering College, Kuppam, Chittoor Dist, A.P. Currently She is Pursuing M.Tech degree in Department of CSE in Kuppam Engineering College, Kuppam, Chittoor Dist, A.P.

Her interested areas are Cloud Computing etc.

2. Mrs R.Kavi Priya received B.Tech degree from Hindustan College of engineering, Chennai, and India. and M.tech degree received from Jayam College of Engineering and technology and India.

Her interested areas are wireless sensor networks.