# Application of Improved Encryption Algorithm in Information Security Protection of Cloud Computing Users

Aimin Niu

ShanDong Agriculture and Engineering University, Jinan, 350001, China

hnam74@163.com

**Abstract.** In order to prevent the user name and password from being stolen, this paper puts forward the application research of improved encryption algorithm in cloud computing user information security protection. In this paper, the privacy protection method based on face recognition mainly uses Gabor filter to extract feature information, ICA-PCA method to reduce dimension, and then uses the improved nearest neighbor algorithm based on user privacy to identify identity. Because both the client and the server can't accurately obtain each other's information, it fully guarantees the safety and zero leakage of user information. Compared with traditional encryption methods, this method realizes linear complexity and reduces communication complexity effectively. In order to verify the effectiveness of the algorithm, this paper uses ORL face database to carry out experiments, and compares with the existing algorithms. The results show that the method in this paper can carry out identity authentication and identification more efficiently and effectively, which provides a strong security guarantee for the privacy of users in the cloud computing digital library under the big data environment.

**Keywords:** cloud computing; Information security; Privacy protection; Random encryption

## 1 Introduction

After the network entered the Web2.0 era, the concept of cloud computing was like fire. In 2006, Eric Schmidt first explained what "cloud computing" was at the search engine conference. There are many definitions of cloud computing, but they are all similar. The so-called cloud computing is a dynamic and scalable computing model developed from distributed processing, parallel processing and grid computing, which uses virtualization technology to establish a unified resource pool of infrastructure, services, applications and information, and effectively organizes and operates various infrastructure resource pools with distributed technology [1]. Cloud computing can provide users with low-cost, high-performance, rapid configuration and quantitative computing services. However, while cloud computing brings benefits of economies of scale, high


G. Guan et al. (eds.), *Proceedings of the 2023 3rd International Conference on Education, Information Management and Service Science (EIMSS 2023)*, Atlantis Highlights in Computer Sciences 16,
https://doi.org/10.2991/978-94-6463-264-4_42

application and high availability to users, its virtualization technology, resource sharing, distributed data storage system and other characteristics make it face great threats in security. When user data is stored in the cloud, how to realize the service security, data security and user privacy security of cloud computing? Whether the disclosure of user data will pose a great threat to individuals, enterprises and even the country has become an important topic in cloud security research at this stage, and it has attracted more and more attention from people in the industry. In recent years, many IT companies have invested in cloud computing. However, during the application of cloud computing services, due to the lack of stability in the operation of cloud computing services deployed and developed by enterprises, cloud computing security issues are constantly emerging. In 2009, Google's cloud platform broke down and customer personal information was leaked [2]; At the end of 2010, there was a database script error in Microsoft Hotmail, and 17,000 real accounts were deleted. It took Microsoft a week to fully recover these accounts. On April 21, 2011, Amazon Cloud Computing Center went down, which caused the cloud service to be interrupted for nearly 4 days, and many services relying on the cloud computing center stopped working. In July, 2012, Dropbox's account was attacked, and hackers released bad information [3]. As a new IT resource service model, cloud computing has a shared IT resource pool composed of a large number of computer resources. All the data that users want to use will be uploaded to the data resource pool. We can regard it as a huge data center, and then integrate these data to provide users with services such as computing and storage, which greatly improves the efficiency of resource use. After the cloud users upload the data they provide to the cloud, the data management will no longer be controlled by the cloud users. At this time, the user data is in an uncontrollable domain, which will make the users lack enough trust in the cloud service providers, and of course they don't want their data to be seen by others. However, in the process of data transmission to the data resource pool, due to the abuse of cloud computing, problems caused by sharing technology, account and service hijacking, unknown security environment and other unsafe factors, data is lost or leaked, resulting in data being destroyed and losing its integrity and confidentiality. The essence of cloud computing is to distribute computing tasks on a resource pool composed of a large number of computers. The integrity and confidentiality of data directly affect whether computing tasks can be carried out smoothly. In this paper, a method of combining biometric identification technology with random variable encryption is proposed to ensure the absolute safety and zero leakage of users' private information, which provides a reference for the research on privacy protection of users in cloud computing digital libraries [4].

## 2 Methods

### 2.1 User privacy protection system based on biometric identification

In the cloud computing environment, users are both service providers and service users, each node plays the role of client and server at the same time, and all participants have equal status. In the open mode of network environment, the traditional user name and password are easily stolen and embezzled, threatening the security of personal account

information, especially sensitive information (bank accounts, medical records, insurance information, etc.) [5]. Biometric identification has gradually become a common means of identity authentication because of its good characteristics such as data characteristics that are not easy to be lost and stolen. At present, commonly used biometric features include fingerprints, faces, irises and so on. Compared with other biometric methods, face recognition is non-invasive, direct, friendly, natural and convenient, and has become the mainstream of biometric identification. This paper will study the user privacy protection system of identity authentication based on face recognition [6].

Biometric identification system includes two modules: registration stage and identification stage. In the registration stage, the user's biometric information is registered first, then the user's feature data is extracted, a feature template is created and stored in the database. The identification stage is similar to the registration stage. First, the user's biometric information is obtained, feature data is extracted, and then matched with the feature template in the database to verify the user's identity.

In order to carry out identity authentication based on facial features, it is necessary to detect and preprocess the face in the input image, accurately locate the position information of the face, and remove the noise information to improve the operation efficiency of feature extraction. The specific process is shown in Figure 1. In this paper, the face detection function based on Adaboost algorithm provided by OpenCV is used for face detection and preprocessing [7].

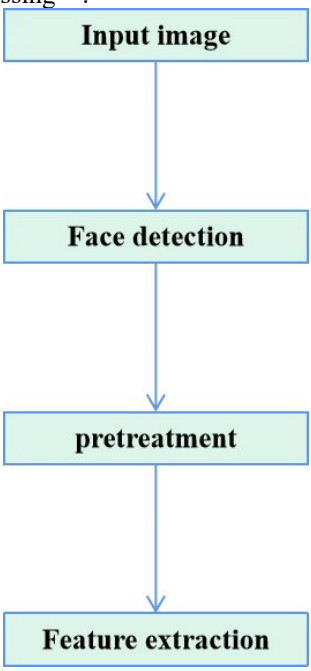


**Fig. 1.** Face feature extraction process

### 2.2 Gabor filtering feature extraction

In the process of face recognition, feature extraction is a key step, and its effectiveness is directly related to the accuracy and efficiency of classification and recognition. In this paper, Gabor filter is used to extract features. Because, among many feature extraction methods, Gabor filtering can extract image features in different frequency domains, different scales and different directions, and can better simulate the feelings of biological visual neurons. In addition, Gabor wavelet is not affected by image gray value and is insensitive to illumination. In this paper, 5-scale and 8-direction filter banks are used for feature extraction. A face image is filtered to get 40 groups of feature vectors with the same size as the original image but completely different properties, which are reorganized in columns. The image size selected in this paper is 92*112. After Gabor filtering, the image becomes a one-dimensional column vector of 92 * 112 * 40 = 412,160 [8].

### 2.3 Feature dimension reduction based on PCA-ICA

Because the convolution of Gabor kernel function of Gabor filter in different scales and directions makes the dimension of image multiply, which easily leads to the disaster of feature dimension and increases the computational complexity. Therefore, it is necessary to reduce the dimension of feature vector and project high-dimensional data into a lower-dimensional space, thus reducing the computational complexity and improving the operational efficiency. The traditional principal component analysis (PCA) method has a good classification effect on data with Gaussian distribution. However, if the data distribution presents non-Gaussian characteristics, simply using the traditional PCA method will often cause the neglected principal component to contain important nonlinear information of the system, which will lead to the inaccurate classification and identification and the inability to correctly identify users. However, using ICA method alone will increase the number of iterations when the amount of data is huge, which will increase the complexity of the algorithm. Therefore, this paper proposes a feature extraction method of PCA-ICA. Firstly, PCA is used to reduce the dimension and remove the second-order correlation. Then further use ICA to extract features, so that the data information can converge quickly and achieve the best possible separability [9].

The specific algorithm steps of the feature extraction method based on PCA-ICA are as follows:

(1) firstly, the original feature matrix is standardized and normalized to obtain a new matrix F ";

(2) F "is extracted by PCA method, from which feature vectors corresponding to the largest top m feature values are extracted to form a new feature space A;

(3) projecting the original data information into the feature space to obtain a new matrix $Y_{s\times n}$;

(4) Using ICA algorithm, further feature extraction of $Y_{s\times n}$ is carried out, and the optimal separation matrix W is solved first, and then $Y_{s\times n}$ is further projected into W space.

### 2.4 Identification algorithm based on random selection of feature vectors

In order to protect the user's feature information from being leaked, most of the existing researches use encryption algorithms to encrypt feature vectors, among which homomorphic encryption is the most. However, the encrypted feature vectors may still be stolen in the network communication process, and they cannot be completely avoided being interpreted by the server. In order to solve this problem, the solution of this paper is to realize that the client knows nothing about the real data of the server, and the server knows nothing about the real data of the client, so as to ensure that the user's private information has no rules to follow in the network transmission process and ensure the absolute security of the information. Therefore, an identity identification algorithm based on random selection of feature vectors is proposed to ensure zero information leakage, and the algorithm complexity and communication complexity of this method are greatly reduced compared with homomorphic encryption algorithm because of omitting the tedious process of encryption and decryption.

## 3 Results and analysis

In this paper, the international universal ORL face database founded by AT&T Laboratory of A University is selected as the experimental class library. There are 40 objects of different ages, sexes and postures in the database, each of which includes 10 images, with a total of 400 facial images. Each image is a gray image with a size of 92×112, and its background is black. Some objects include changes in postures, expressions, facial ornaments and so on. It is recognized by people in resolution and is the most used one at present. In this paper, eight images of each sample are selected as training samples, and the remaining two are used as test samples, so there are 320 sample images in the training sample database.

### 3.1 Operational efficiency

Without losing generality, under normal circumstances, the safety parameter length is set to 1024 bits, the multiplication calculation time is recorded as $T_m$, and the exponential calculation time is recorded as $T_e$, and $T_e \approx 240T_m$. Table 1 shows the calculation time of this method compared with other encryption steps. As can be seen from the table, compared with the homomorphic encryption algorithm of cryptography, this method greatly reduces the operation time and improves the efficiency.

**Table 1.** Comparison of calculation time of each step of encryption

| | A method | The method in this paper |
|---|---|---|
| 5.1 | $O(T_m+T_e)$ | O(1) |
| 5.2 | $O(T_m+T_e)$ | O(1) |
| 5.3 | $O(\log_2 n(T_m+T_e))$ | $O(T_m)$ |

### 3.2 Recognition rate

The following table shows the performance comparison between several existing research algorithms and the algorithm in this paper. From Table 2, it is not difficult to find that this paper not only reduces the operation time and improves the efficiency, but also ensures the accuracy without reducing the recognition rate, so the method in this paper is effective.

**Table 2.** Performance comparison of five schemes

| Scheme | Encryption protocol | recognition algorithm | discrimination | Communication complexity | experiment condition |
|---|---|---|---|---|---|
| Method B | Paillier+ GDK | PCA | 92% | 6.89M | ORL face Each library has 8. Sample training, 2 sample tests |
| Method C | HE+GC | PCA | 92% | 2.85M | |
| Method A | FHE | Gabor+PCA Gabor+ | 94% | 4.11M | |
| The method in this paper | - | PCA-ICA | 95% | 2.49M | |

### 3.3 The effectiveness of digital library applications

In order to further study the application of privacy protection system based on biometrics in digital library, this paper built a small face database, selected 10 volunteers, and took 10 photos in different lighting, different angles and different expressions, totaling 100 images, and normalized the image size. The processed image size was 112*92.

In this paper, the public cloud platform, digital library alliance of universities in A province, is selected as an example to verify the effectiveness of the privacy protection method based on biometrics proposed in this paper.

A complete privacy protection system of digital library based on biometrics includes two major processes: training process and identification process. The training process is mainly divided into three stages. In the first stage, each volunteer selected five images to form a face database. In the second stage, Gabor filter analysis and PCA-ICA algorithm are carried out on the face database to reduce the dimension of features, and the feature face space is constructed. Setting the feature dimension ratio to 0.9 is the best dimension ratio based on many experiments. In the third stage, the face image is projected into the feature face space and saved. The identification process is also divided into three stages.

In the first stage, the face image is preprocessed and feature extracted. In the second stage, the test image is projected into the feature space by using the random vector projection method in this paper. In the third stage, according to the nearest neighbor algorithm, Euclidean distance is selected to identify the image. The test results of

selected individuals in this paper show that the average recognition rate is as high as 99.5% and the response time is very short, with an average of 0.735 seconds. Therefore, the privacy protection method based on biometrics proposed in this paper is effective and feasible for the privacy protection of users in digital libraries [10].

## 4 Conclusion

With the advent of the era of big data, cloud computing technology has been widely used in various fields, and information security based on cloud computing has also received increasing research attention. In view of the fact that user name password is easy to be stolen, a face recognition method based on Gabor filtering and PCA-ICA is proposed for identity authentication. In order to ensure zero information leakage, a new linear encryption algorithm is proposed, and the feature vectors are randomized and encrypted. Experiments show that this method not only ensures the recognition rate of user identity authentication, but also greatly reduces the algorithm complexity and communication complexity compared with homomorphic encryption algorithm because of omitting the tedious process of encryption and decryption, which has certain guiding significance for the research of user information security in big data environment. The deficiency of this paper is that the user sample information of digital library is not actually selected for case study, which will be discussed in detail as the next step.

## References

1. Yang, C. Y. , Ling, Y. , & Li, X. . (2021). Research on information encryption algorithm under the power network communication security model. Journal of Physics: Conference Series, 1852(3), 032007 (7pp).
2. Sharma, J. , Kim, D. , Lee, A. , & Seo, D. . (2021). On differential privacy-based framework for enhancing user data privacy in mobile edge computing environment. IEEE Access, PP(99), 1-1.
3. Chen, Z. , Wu, A. , Li, Y. , Xing, Q. , & Geng, S. . (2021). Blockchain-enabled public key encryption with multi-keyword search in cloud computing. Security and Communication Networks, 2021(2), 1-11.
4. Lina, G. E. , Yugu, H. U. , Zhang, G. , & Chen, Y. . (2021). Reverse hybrid access control scheme based on object attribute matching in cloud computing environment. Journal of Computer Applications, 41(6), 1604-1610.
5. Qerom, M. , Ahamad, D. , Akhtar, M. M. , & Hameed, S. A. . (2021). Provably secure authentication approach for data security in cloud using hashing, encryption, and chebyshev-based authentication. International Journal of Electronic Security and Digital Forensics, 1(1), 1.
6. Liu, L. , Gao, M. , Zhang, Y. , & Wang, Y. . (2022). Application of machine learning in intelligent encryption for digital information of real-time image text under big data. EURASIP Journal on Wireless Communications and Networking, 2022(1), 1-16.

7. Thabit, F. , Alhomdy, S. , & Jagtap, S. B. . (2021). A novel technique of data security in cloud computing based on two layer of cryptography algorithm using genetics algorithm and asymmetric cryptography. Journal of Parallel and Distributed Computing, 10(4), 93-96.
8. Chouhan, J. , Rai, M. , & Sabri, M. S. . (2021). A literature survey of different data encryption algorithm for secure health care system in cloud. SSRN Electronic Journal, 10(3), 1-7.
9. Shabbir, M. , Shabbir, A. , Iwendi, C. , Javed, A. R. , & Lin, C. W. . (2021). Enhancing security of health information using modular encryption standard in mobile cloud computing. IEEE Access, PP(99), 1-1.
10. Huang, F. , Yong, M. , & Jian, Z. . (2021). Genetic algorithm-based power system information security risk assessment method. Journal of Physics: Conference Series, 1852(2), 022063 (7pp).