Research on Secure Storage of Network Data Based on Cloud Computing Technology

Jinfeng Zhu

(Corresponding author: Jinfeng Zhu)

School of Computer Science and Engineering, Sanjiang University
No. 310, Longxi Road, Tiexin Bridge, Yuhuatai District, Nanjing, Jiangsu 210012, China
Email: ji18094@163.com

(Received July 10, 2020; Revised and Accepted Aug. 10, 2021; First Online Nov. 9, 2021)

Abstract

Cloud storage technology brought by cloud computing relieves the storage performance pressure of independent computers, but it also puts forward requirements for data storage security. This paper gave brief introductions to the cloud computing-based cloud storage, asymmetric encryption algorithm, the Advanced Encryption Standard (AES) algorithm, and an asymmetric encryption algorithm. This SM2 algorithm was used to ensure the security of cloud storage data. A hybrid encryption technology combining the AES algorithm and the SM2 algorithm was proposed. Finally, we tested the performance of the three encryption algorithms in cloud storage in the laboratory's local area network (LAN). The results showed that the AES algorithm spent the shortest time encrypting and decrypting data. The SM2 algorithm spent the longest time, and the time consumed by the AES+SM2 algorithm was significantly shorter than that of the SM2 algorithm and slightly longer than the AES algorithm. Considering that the complexity of the ciphertext decoded by the AES algorithm in the encryption performance experiment was the highest, followed by the SM2 algorithm and the AES+SM2 algorithm, the defect that the time consumed by the hybrid encryption technology was slightly longer than that by the AES algorithm in encryption and decryption was almost ignored. The AES algorithm failed in recognizing the tampering of ciphertext data in the cloud storage process, but the SM2 algorithm and the AES+SM2 algorithm succeeded.

Keywords: Asymmetric Encryption; Cloud Computing; Data Storage; Symmetric Encryption

1 Introduction

Due to the convenience of the Internet, people use the Internet more frequently for shopping, watching videos, and even banking [10]. Facing the increasing big data on the Internet, a single computer device has been difficult to meet the corresponding computing needs [3]. Cloud computing brought by high-speed and low latency communication technology provides an effective solution for the efficient processing of Internet big data [12]. When users use cloud computing to process big data, they no longer rely on the computing resources of their devices but upload the processing tasks to the cloud and use the distributed computing resources in the cloud to solve the tasks uploaded by users [9]. Cloud storage is an expanding service of cloud computing, which also uses distributed storage resources to store data in the cloud. Similarly, due to the good scalability, the cloud storage space is large enough, and the performance requirements for client equipment are low.

Cloud computing and cloud storage can provide users with computing and storage resources that are not limited by the performance of devices; however, important privacy information will also be transmitted in the process of use. Also, the open sharing characteristic of cloud services makes the stored data more vulnerable to intrusion and theft. Yoon *et al.* [14] proposed a density-based data encryption scheme and a database outsourcing query processing algorithm. In the performance analysis, compared with the existing schemes, the proposed scheme had better query processing performance and ensured the user's privacy. Xue and Ren [13] proposed an improved, efficient data encryption method, which was based on the ciphertext policy attribute-based encryption and used the fixed-length ciphertext to control the time cost. The simulation results showed that the improved algorithm had high reliability. Due to the traditional cloud computing-based information transmission mechanism and the problems of large errors and low security existing in and the Internet of things, Ding et al. [2] proposed a network resource management algorithm and verified the effectiveness of the algorithm through experiments.

2 Cloud Computing and Cloud Storage

Cloud computing [4] is a distributed data computing mode, which uses the data sharing function of the Internet to process a large number of data in a distributed way. These processes can be carried out in parallel; thus, it has higher computing efficiency. As cloud computing technology is based on the data sharing function on the Internet, although it makes the data processing more efficient, the openness brought by data sharing makes the data more vulnerable to attack. If they are public data, the loss will not be great; but if they are private data very important for users, the loss will be very serious.

Cloud storage is based on cloud computing technology, which not only improves the storage capacity but also inherits the characteristics of data sharing and data security protection in cloud computing [11]. The basic architecture of the cloud computing-based network storage system is shown in Figure 1. Users use the client to register or log in directly in the login interface and connect with the control center in the cloud server through the corresponding user interface to carry out corresponding system operations, including information registration, user login, data access, etc. After receiving the system operation information from the user interface, the control center obtains the corresponding service data from the node cluster, and the data of the node cluster come from the data pool composed of node clusters. Data pool is cloud computing space, usually composed of node clusters, and it can store data.

3 Storage Data Encryption

The growing big data contains a large number of valuable laws. The mining and analysis of the laws can provide effective guidance for activities such as production and operation. But these big data also contain a lot of privacy information, which would cause serious losses if they were used by lawbreakers. Therefore, in the cloud storage process, in addition to the need for identity authentication in the transmission process, storing data after encryption is also needed.

There are many kinds of data encryption algorithms in the cloud storage process, but the steps and models used in information transmission and storage encryption are nearly the same. Figure 2 is the basic model of transmission information encryption technology. First, at the sending node, the plaintext information to be sent is encrypted by the key generated by the encryption algorithm to obtain the ciphertext; then, the ciphertext is transmitted to the receiving node for storage through the secure channel of the Internet, and the key is transmitted to the receiving node through the non-secure channel [5]; the receiving node receives the ciphertext and stores it, and the ciphertext needs to be decoded using the key to obtain the plaintext before reading. In the basic model, the key and ciphertext are transmitted through different channels to reduce the probability that they are stolen at the same time. Ciphertext, the main body of data transmission, is further protected through the secure channel. As the key of encryption and decryption, the key itself does not have a specific meaning. As long as the key is not intercepted at the same time as the ciphertext, the plaintext will not be disclosed; therefore, the key can be transmitted by the non-secure channel to save the resource of the secure channel.

3.1 Advanced Encryption Standard Algorithm

It is seen from the previous description that cloud storage is based on cloud computing technology and uploads the data to be stored to the cloud composed of server node clusters. In order to prevent data leakage due to the data sharing characteristic in cloud storage and strengthen the security of data storage, the data are encrypted by an encryption algorithm and stored in the cloud [7].

Advanced encryption standard (AES) algorithm [1] is a kind of symmetric encryption technology. Considering that a 128-bit key is enough, it is usually used as the key of the AES algorithm. The encryption process of the AES algorithm is as follows.

1) The plaintext that needs to be encrypted is grouped, and the length of each group is 128 bytes. When it is less than 128 bytes, the group is supplemented.
2) The plaintext (128 bytes/group) and the key with the same number of bytes are put into two 4×4 matrices.
3) The plaintext matrix is encrypted by the key matrix.
4) The encrypted ciphertext matrix is processed by byte replacement, row displacement, and column mixing.
5) The key matrix is extended by using the key schedule function, and then the ciphertext matrix is encrypted again by using the extended key matrix.
6) Steps 4 and 5 are repeated ten times. Column mixing operation is not carried out in the tenth repetition. Finally, the ciphertext is obtained.

The decryption process of the AES algorithm is the reverse operation of the encryption process. In addition to the reverse operation of steps, the encryption of the information matrix is also reverse, including reverse byte replacement, reverse row displacement, reverse column mixing, etc.

3.2 SM2 Algorithm

SM2 algorithm [6] is a kind of asymmetric encryption technology, which uses a basic elliptic curve to generate key pairs. The encryption process is as follows.

1) Integer k generates randomly in $[1, n-1]$, where n refers to the order of base point G.

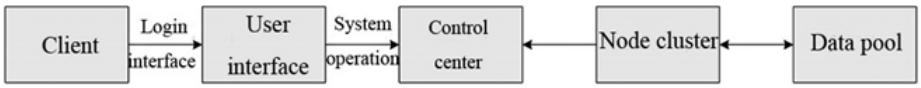

Figure 1: The basic architecture of the cloud computing-based network storage system

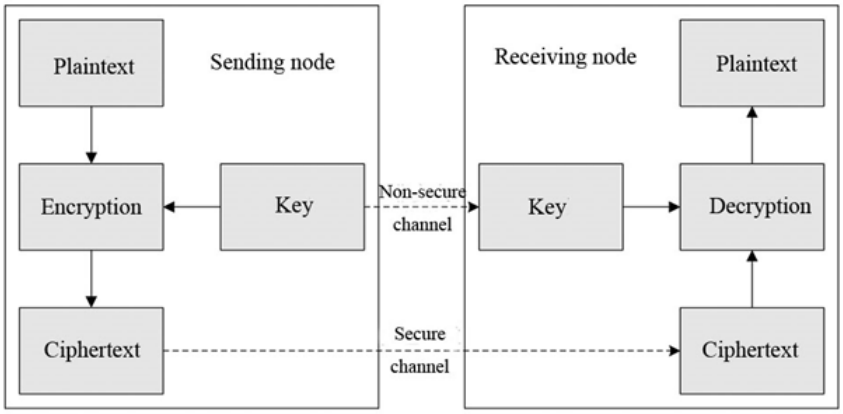

Figure 2: The basic model of information encryption technology in cloud storage

2) Points C_1, S, and Q in the elliptic curve are calculated according to Equation (1), and the formula is:

$$\begin{cases} C_1 = [k]G = (x_1, y_1) \\ S = [h]P_B \\ Q = [k]P_B = (x_2, y_2) \end{cases} \tag{1}$$

where h is a cofactor; (x_1, y_1) and (x_2, y_2) are coordinates of C_1 and Q, which need to be converted into bit string in the following steps, and P_B is a public key.

3) Bit string t of the key data is calculated:

$$t = DF(x_2||y_2, klen). \tag{2}$$

The decryption process is as follows.

1) Whether the equation of the elliptic curve is true or not is verified according to the coordinate point C_1 converted by bit string C_1 separated from ciphertext C. If it is true, decryption will continue; if it is not true, it means that the ciphertext is not encrypted by the SM2 algorithm, and encryption is no longer carried out.

2) Coordinate point C_1 is substituted into:

$$S = [h]C_1. \tag{3}$$

Whether point S is an infinite point in the elliptic curve is tested. If it is an infinite point in the elliptic curve, decryption will stop; if not, decryption will continue.

3) Key bit string t' used for decryption is obtained based on coordinate point C_1 and private key d_B. The calculation formula is:

$$\begin{cases} (x_2, y_2) = d_B C_1 \\ t' = KDF(x_2||y_2, klen). \end{cases} \tag{4}$$

4) C_2 in ciphertext C is decoded using t' to obtain plain text M'. The hash value of M' is calculated using the hash function of the SM3 algorithm. If the hash value of M' is the same with C_3, it indicates the information has not been tampered, and the decryption is successful. If the hash value is not the same with C_3, it indicates the information has been tampered, and the decryption fails.

3.3 Hybrid Encryption Algorithm

Although the AES algorithm has high efficiency in encryption and decryption, which is suitable for the big data environment under cloud computing, the single key is easier to be intercepted during transmission, and the receiver can not confirm whether the transmission data has been tampered with [8]. Only the public key will participate in the transmission of the SM2 algorithm; thus, even if the public key is intercepted, the third party can not decrypt by virtue of the public key, and the integrity of information can be verified by calculating the hash value of the information by the hash function, which greatly improves the security of transmission and storage. However, as asymmetric encryption technology, the SM2 algorithm is not as efficient as the AES algorithm in key generation, encryption, and decryption; thus, it is not suitable for the big data environment of cloud computing.

In order to improve the efficiency of encryption as much as possible on the premise of ensuring the storage security of transmission data, this paper uses hybrid encryption technology to encrypt the network data in the cloud computing environment.

The flow of hybrid encryption technology used in the cloud computing-based data storage process is shown in Figure 3, and the specific steps are as follows.

1) After the user enters the plaintext information in the client, the client will use a hash function, the SM3 algorithm, to calculate the hash value Z_1 of the plaintext and merge it with the plaintext.

2) The client uses the key of the AES algorithm to encrypt the merged information to obtain ciphertext C_1. The key of the AES algorithm was processed by SM2 encryption using the public key sent by the cloud server to obtain ciphertext C_2.

3) C_1 and C_2 are merged and uploaded to the cloud server.

4) After receiving the merged information, the cloud server splits it into C_1 and C_2.

5) The server uses the cloud private key pair to decrypt C_2 to obtain the key of the AES algorithm. C_1 is decoded by the key to obtain the ciphertext and the merged information of Z_1.

6) The hash value Z_2 is obtained by performing a hash calculation on the ciphertext using the SM3 algorithm. Z_1 is compared with Z_2. If they are not consistent, it indicates that the transmission data have problems, and the cloud server will not store the data; if they are consistent, the server will store the ciphertext.

In the above process, Steps 1, 2, and 3 are carried out in the user client, and Steps 4, 5, and 6 are carried out in the cloud server. Also, the decryption operation of the regular cloud storage service before storing the ciphertext in the cloud mainly aims to obtain the hash value of the decoded plaintext to verify whether there is a problem with the ciphertext information. The plaintext obtained in the verification process will not be stored in the server, and the ciphertext will be stored only after the verification is passed. Theoretically, there will be no plaintext in the cloud server.

4 Experimental Analysis

4.1 Experimental Environment

The experiment was carried out in the LAN of the laboratory, in which servers A and B were used as the storage nodes of cloud service, server C was used as the server for key generation, encryption, and decryption of cloud storage service, and a personal computer was used as the client. The server as the storage node had a 2 TB capacity. The server for key generation, encryption, and decryption of the cloud storage service had 16 G memory, a Core I7 processor, and a 64-bit Windows operating system. The personal computer was configured with 6 GB memory, a 64-bit Windows operating system, and a Core i7 processor.

4.2 Experimental Methods

1) One hundred packets in sizes of 5 MB, 15 MB, 25 MB, 35 MB, and 45 MB were set and uploaded to the cloud server using the personal computer client. In the uploading process, the data were encrypted by AES, SM2, and AES + SM2 algorithms. The time consumed by the three encryption algorithms in encryption and decryption was detected.

2) Another one hundred packets in sizes of 5 MB, 15 MB, 25 MB, 35 MB, and 45 MB were also set and uploaded to the cloud server by the personal computer client. Before the ciphertext was stored in the cloud server, the ciphertext was decoded by another server to simulate the situation that the information was cracked by a third party. The maximum cracking time was set as 60 min. The cracked ciphertext was compared with the original text to obtain the decryption integrity.

3) One hundred packets in sizes the same as in the previous two experiments were set. Before uploading to the cloud server through the client, the packets were uploaded to the third-party server. Without the corresponding key, the ciphertext decryption under the three encryption algorithms was difficult and could not be cracked in a short time; moreover, the cloud server would not store the plaintext after decoding the ciphertext. Therefore, the ciphertext was tampered with by adding 5-bit ciphertext in the original ciphertext packet and uploaded to the cloud server for storage to verify whether the three encryption algorithms could identify the tampered ciphertext data.

4.3 Experimental Results

In the cloud computing-based cloud storage service, the encryption of information is very important. In addition to ensuring the security of encryption, the efficiency of encryption and decryption is also an important factor. Figure 4 shows the encryption and decryption time of three encryption technologies in cloud computing. First of all, it was seen that the time consumed by the three encryption technologies increased with the increase of the data packet size, i.e., the increase of the amount of data to be processed. The decryption and encryption time of different algorithms differed little. The AES algorithm only used a single key, which was more efficient in key generation, encryption, and decryption. The key pair

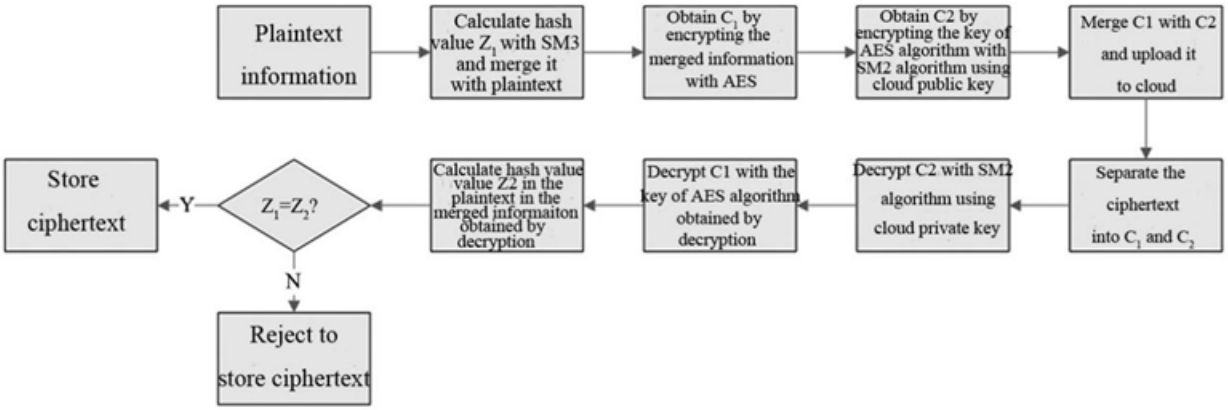

Figure 3: The flow chart of hybrid encryption technology

generation of the SM2 algorithm was more complex and time-consuming, and its encryption process and decryption process were not as reversible as the AES algorithm; thus, it was inefficient in the face of big data. The AES + SM2 hybrid algorithm used the SM2 algorithm for re-encryption based on the AES algorithm; thus, the time consumed was slightly longer than the AES algorithm; however, the SM2 algorithm was only used for encrypting the key of the AES algorithm, and the length of the key was significantly smaller than the plaintext. Thus, the time consumed by the hybrid algorithm was shorter than the single SM2 algorithm.

In the cloud computing-based cloud storage service, the confidentiality of encrypted information is very important. Figure 5 shows the decryption integrity of the data packet 60 minutes after being cracked by a third party under three encryption technologies. It was seen from Figure 4 that the decryption integrity of data packets under the three encryption technologies decreased with the increase of the data packet size. The increase of the data packet size meant that the amount of data increased, and the amount of calculation also increased during cracking; thus, the decryption integrity decreased under the limited decryption time. Comparing the data integrity of packets of the same size being processed by different encryption technologies, the AES algorithm was the highest, followed by the SM2 algorithm and the AES + SM2 algorithm. The reason for the above result is as follows. The AES algorithm only used single key encryption; thus, the encrypted ciphertext presented a relatively fixed law. The SM2 algorithm used public-key encryption; thus, the regularity of encrypted ciphertext was greatly reduced. The AES + SM2 algorithm not only encrypted the plaintext but also added the ciphertext to the key of the AES algorithm after asymmetric encryption, reducing the regularity of the ciphertext and improving the decryption difficulty.

The ciphertext packets that needed to be uploaded to the cloud server for storage were tampered with by a third-party server, and the detection performance of the three encryption technologies on the tampered data was tested. The results are shown in Table 1. It was seen from Table 1 that when the AES algorithm was used as the encryption technology for cloud storage, the tampered ciphertext packets of any size were successfully stored in the cloud server; the cloud storage service using the SM2 technology and the AES + SM2 encryption technology effectively detected and rejected the tampered ciphertext packets. The reason for the above results is as follows. The AES algorithm used single key encryption and decryption without an additional verification mechanism. Even if the encrypted ciphertext was decrypted to plaintext, if there was no actual plaintext for comparison, it was impossible to confirm whether the data has been tampered with. If the plaintext was sent for verification, then encryption technology was meaningless. The SM2 algorithm could verify by taking the hash value of the plaintext as a comparison basis. The AES+SM2 algorithm combined the two algorithms, thereby inheriting the data integrity detection mechanism in the SM2 algorithm. In that mechanism, it did not need to send plaintext as a comparison but used hash values representing the complete plaintext to ensure the security of plaintext.

5 Conclusion

This paper gave brief introductions to the cloud computing-based cloud storage, a symmetric encryption algorithm, the AES algorithm, and an asymmetric encryption algorithm, the SM2 algorithm. Combining the two algorithms, the AES + SM2 hybrid encryption technology was proposed. Finally, the performance of the three encryption algorithms in cloud storage was tested in the LAN of the laboratory. The results are as follows.

1) With the increase of the cloud storage packet size, the encryption and decryption time of the three encryption algorithms decreased; when processing the packet of the same size, the AES algorithm had the shortest encryption and decryption time, the

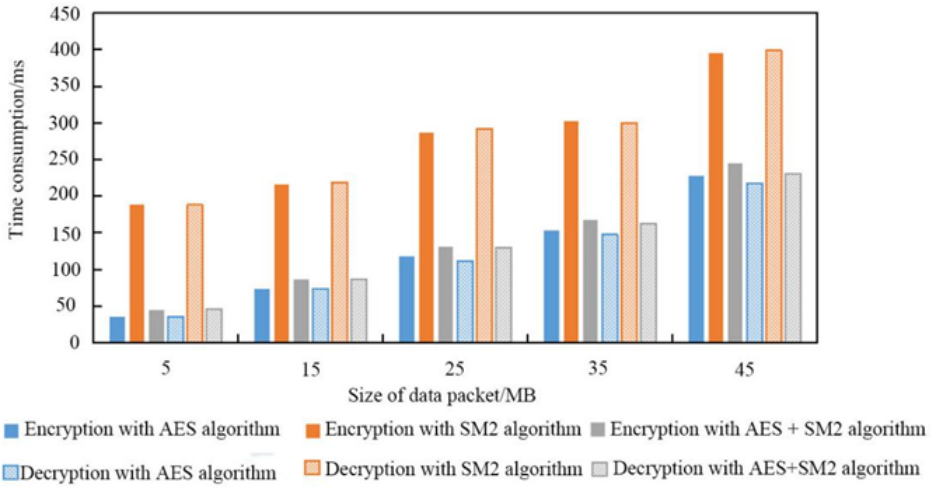

Figure 4: Time consumed by three encryption technologies in encryption and decryption in cloud computing

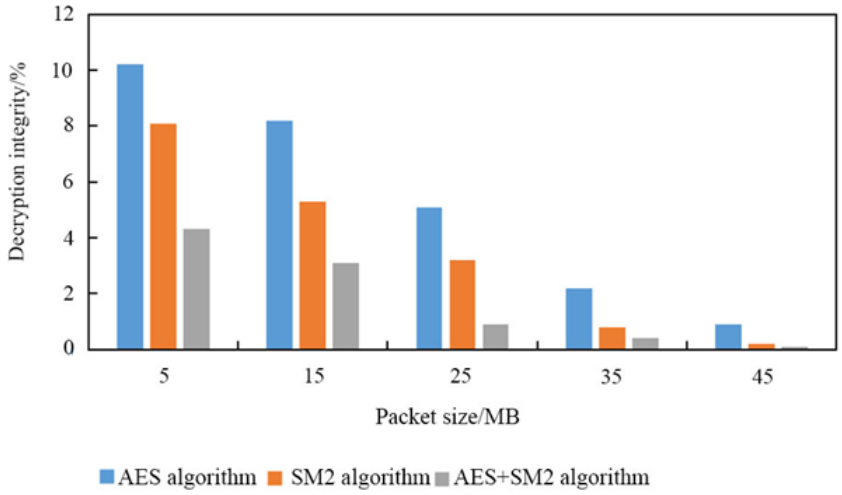

Figure 5: Security of three encryption technologies in cloud computing

Table 1: Detection performance of three encryption technologies for tampered ciphertext

	5 MB	15 MB	25 MB	35 MB	45 MB
AES	Storage succeeds	Storage succeeds	Storage succeeds	Storage succeeds	Storage succeeds
SM2	Storage fails	Storage fails	Storage fails	Storage fails	Storage fails
AES+SM2	Storage fails	Storage fails	Storage fails	Storage fails	Storage fails

SM2algorithm had the longest encryption and decryption time, and the time of the AES + SM2 algorithm was between the two other two algorithms.

2) With the increase of the data packet size, the integrity of data under the three encryption technologies decreased, and when processing the data packet of same size, the integrity of data under the AES algorithm was the highest, followed by the SM2 algorithm and the AES + SM2 algorithm.

3) In the face of tampered data, no matter how big the data packet was, the cloud storage service under the AES encryption technology could not make effective identification, but the cloud storage service under SM2 and AES + SM2 encryption technologies could make effective identification.

References

[1] A. Banushri, R. A. Karthika, "A Survey on data security using file hierarchy attribute-based encryption in cloud computing environment," *Journal of Advanced Research in Dynamical & Control Systems*, vol. 2017, no. 4, pp. 144-149, 2017.

[2] L. Ding, Z. Wang, X. Wang, D. Wu, "Security information transmission algorithms for IoT based on cloud computing," *Computer Communications*, vol. 155, pp. 32-39, 2020.

[3] G. M. Kiran, N. Nalini, "Enhanced security-aware technique and ontology data access control in cloud computing," *International Journal of Communication Systems*, vol. 33, no. 15, pp. e4554, 2020.

[4] K. S. Kumar, K. Shankar, M. Ilayaraja, M. Rajesh, "Sensitive data security in cloud computing aid of different encryption techniques," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 9, no. 18, pp. 2888-2899, 2018.

[5] R. Manoharan, "Sensitive data security in cloud computing aid of different encryption techniques," *Journal of Advanced Research in Dynamical and Control Systems*, vol. 18, pp. 2888-2899, 2017.

[6] P. R. More, S. Y. Gaikwad, "An advanced mechanism for secure data sharing in cloud computing using revocable storage identity based encryption," *International Journal of Engineering Business Management*, vol. 1, no. 1, pp. 12-14, 2017.

[7] V. Naresh, M. Anudeep, M. Saipraneeth, A. S. Reddy, V. Navya, "Encryption-based secure and efficient access control to out sourced data in cloud computing," *International Journal of Engineering & Technology*, vol. 7, no. 2, pp. 315, 2018.

[8] T. Paka, S. Divya, "Data storage security and privacy in mobile cloud computing using hierarchical attribute based encryption (HABE)," *International Journal of Computer Sciences and Engineering*, vol. 7, no. 6, pp. 750-754, 2019.

[9] K. Rangasami, S. Vagdevi, "Comparative study of homomorphic encryption methods for secured data operations in cloud computing," in *International Conference on Electrical*, pp. 1-6, 2017.

[10] P. Sawle, T. Baraskar, "Survey on data classification and data encryption techniques used in cloud computing," *International Journal of Computer Applications*, vol. 135, no. 12, pp. 35-40, 2016.

[11] M T. Sultan, K. N. Yasen, "Homomorphic encryption implementation to ensure data security in cloud computing," *Journal of Theoretical and Applied Information Technology*, vol. 96, no. 7, pp. 1826-1836, 2018.

[12] L. Teng, H. Li, S. Yin, Y. Sun, "A modified advanced encryption standard for data security," *International Journal of Network Security*, vol. 22, no. 1, pp. 112-117, 2020.

[13] S. Xue, C. Ren, "Security protection of system sharing data with Improved CP-ABE encryption algorithm under cloud computing environment," *Automatic Control and Computer Sciences*, vol. 53, no. 4, pp. 342–350, 2019.

[14] M. Yoon, M. Jang, Y. S. Shin, J. W. Chang, "A bitmap based data encryption dcheme in cloud computing," *International Journal of Software Engineering & Its Applications*, vol. 9, no. 5, pp. 345-360, 2015.

Biography

Jinfeng Zhu,born in 1980, holds a bachelor's degree from Southeast University. Now he works in Nanjing Sanjiang University and is the director of the professional laboratory. Mainly engaged in the construction and management of professional laboratories, and participated in the teaching and scientific research of computer science and technology, software engineering, network engineering and other majors. Research interests: computer application, multimedia technology, cloud computing, etc.