Journal of Network Intelligence ©2024 ISSN 2414-8105
Taiwan Ubiquitous Information Volume **9**, Number **3**, August **2024**

Improved CP-ABE based Approach for IoT User Privacy Data Protection

Zhiqiang Peng

School of Information and Engineering
Sichuan Tourism University, Chengdu 610100, China
pengzhiqiang@mjc-edu.cn

*Corresponding author: Zhiqiang Peng

Received September 22, 2023, revised December 22, 2023, accepted March 6, 2024.

Abstract. *Data owners often encrypt their data to ensure its security during cloud data processing due to the complexity and sensitivity of user information. However, existing searchable encryption schemes face challenges in terms of cloud ciphertext search and low performance in document relevance and matching, resulting in inefficient search capabilities. To address this issue, this research proposes a search scheme that incorporates the layered attribute encryption algorithm CP-ABE. The scheme tests the homogeneity and adjusts the similarity of ciphertext, introduces the RAS algorithm to encrypt the weight vector, and designs the searchable scheme on the Hadoop platform. The effectiveness of this search scheme is analyzed and verified. It was discovered that the improved CP-ABE algorithm had a lower data processing error rate (12.36%) and a shorter delay (3.716ms) than the traditional CP-ABE algorithm, and its data confidentiality was far greater (82.36%¿58.13%). At the same time, the algorithm's memory consumption was low, and the score for collecting users' data confidentiality increased by 56.17%. The proposed improved encryption CP-ABE algorithm for encryption of weight vectors and support for the revocation function provides stronger control and management tools for the privacy protection of user data, which greatly improves the security of the algorithm and data confidentiality. This provides new means and tools for privacy protection of IoT user information, data sharing and collaboration, and security management of sensitive data.*

Keywords: improving CP-ABE; IoT users; private data; encryption; risk of leakage

1. **Introduction.** Internet of Things (IoT) can effectively realize the interconnection between various devices, and improve people's work efficiency and quality of life. It has been applied in a variety of fields such as smart cities, smart homes, industrial automation, health care, and so on [1]. Data collection and sharing is one of the core functions in the application of IoT, and its data storage location is mainly in the cloud database. Cloud storage generally relies on third-party data providers for data delivery and storage processing, which inevitably makes the delivery and storage process face a variety of security risks, seriously jeopardizing the privacy of IoT users and data security [2]. With the continuous popularization and application of IoT technology, users' personal information and sensitive data are also more and more likely to be threatened, which may include location information, health data, family life habits, and other sensitive content, and once leaked will bring serious privacy risks and security problems. Therefore, strengthening the protection of users' private data in the context of IoT plays an important role in the information security of individuals as well as the sustainability of the entire IoT

system. In the resource-constrained IoT, the most commonly used method to protect various types of privacy is encryption (usually homomorphic encryption), and the design of lightweight encryption schemes has become one of the most important methods in IoT privacy protection. One of the common data encryption methods is the ciphertext-policy attribute-based encryption (CP-ABE) scheme, which allows the data owner to define the access policy as well as the attribute rules, and control who can access the encrypted data accordingly, with flexible access control, and diversified access structure [3]. Many scholars have also conducted research on data encryption algorithms, such as Xue et al. proposed an improved CP-ABE encryption algorithm for the problem of greater security of data sharing in the cloud computing environment, which encrypted on the attributes of the ciphertext policy, combined with a fixed-length ciphertext for time control. Experimental results showed that the improved CP-ABE algorithm had better efficiency and higher reliability in processing data, indicating that the improved CP-ABE algorithm had better sharing of secure data [4]. Yu et al. proposed a crowdsourced privacy protection method based on ciphertext policy attribute encryption to improve data sharing privacy by combining multiple authorities to distribute keys separately, decentralizing and reducing the responsibility of each platform, lowering the computational cost of IoT users, and protecting the privacy of user data. Experiment results showed that the method can improve user privacy protection while also reducing computation time [5]. CP-ABE encryption technology may generate a large computational overhead during ciphertext generation and decryption, resulting in low efficiency, and it does not support data accessor and attribute revocation, which will cause a major hidden danger to system security once the user leaks the ciphertext. Moreover, this encryption technology is not compatible with the basic principle of hybrid encryption, and cannot fully utilize hybrid encryption to guarantee the security of information data. To create a better data protection mechanism system for users, the research introduces the support revocation function on the original CP-ABE algorithm and realizes the encryption processing and information management of user privacy protection data through authority authentication, homogeneity test, ciphertext similarity adjustment, and weight vector encryption.

This research first examines the current security-related algorithms of the CP-ABE algorithm, as well as its application fields and methods. It then analyzes the security of data storage and concludes that the CP-ABE algorithm can effectively enhance the protection of private data. Then, the traditional CP-ABE algorithm is improved by supporting the introduction of a revocation function and homogeneity test adjustment, so that it can better encrypt and protect the data. The improved CP-ABE attribute encryption algorithm proposed by the research has the following innovations. Firstly, the focus is on ensuring the privacy and non-traceability of cloud search data, as well as simplifying the process of ciphertext search operations. To achieve these objectives, improvements are proposed in various aspects of the searchable scheme, authority authentication, homomorphism test, and ciphertext similarity adjustment. These enhancements are based on the hierarchical attribute encryption algorithm CP-ABE. Secondly, the random asymmetric segmentation algorithm is used to encrypt the weight vectors and achieve the privacy protection of user data with the help of the cloud storage platform carrier form. In order to verify the high reliability of the encryption algorithm proposed in the research, performance tests and comparative experiments of related methods are carried out in terms of data search efficiency, data encryption processing security, and resource consumption, in order to create a better data protection mechanism system for users.

2. **Related Work.**

The CP-ABE algorithm can automatically control the scope of data sharing and assign a key to the encrypted data, making the data more secure. As a result, this algorithm is widely used in many fields, and many scholars have combined the CP-ABE algorithm to investigate privacy data protection, yielding numerous research results. Yadav et al. proposed an advanced encryption method based on attribute encryption to address the issue of inadequate user privacy protection in the cloud computing framework and experimentally demonstrated that the method had faster encryption and decryption times than traditional encryption methods [6]. Challagidad et al. proposed an improved CP-ABE control method for the problem of difficulty in protecting users' privacy in cloud storage by combining a role hierarchy algorithm and hierarchical access structure, and the experimental results showed that the method had a faster processing time than the original algorithm [7]. Feng et al. proposed a CP-ABE privacy protection method in order to solve the problem of high-security risks of user and data separation in cloud storage systems, which combined an explicit transfer algorithm, linking private key components and random identity of users, and experimental results demonstrated that the improved algorithm improved the degree of privacy protection [8]. Tian et al. proposed an attribute-based access management scheme for security and privacy protection in the cloud to promote the stability of IoT data storage security, combining the steps of key generation, access management, and content decryption to reduce the consumption expenses at the IoT user side, and the experimental results demonstrated that the proposed a management scheme had high computational efficiency and better security [9]. Xiong et al. developed an attribute-based encryption method combined with a partial hiding strategy to protect private information in order to address the issue of low data security and privacy in edge computing, and the experimental results demonstrated that the method can improve the privacy and security of data [10].

Zhang et al. developed an attribute-based encryption method to enhance the security and privacy of data in cloud computing. They proposed an improved encryption method (CP-ABE) that places greater emphasis on authorization verification and preventing privacy leakage. Experimental results showed that the improved method effectively enhances the security and privacy of data [11]. Han et al. in order to protect the sensitive information of IoT users designed an improved CP-ABE method to hide and revoke the parts of the method, and the experimental results demonstrated that the improved method can protect users' sensitive information more effectively [12]. For the communication system in IoT, ensuring network security and fine-grained access control is an important research inner tube. Li et al. proposed the ciphertext policy-based weighted attribute method to secure IoT data. With the help of new coding technology for access policy expression, ciphertext policy-weighted attribute-based encryption (CP-WABE) was constructed. The results showed that the method can effectively resist the security of selective plaintext attacks with high security [13]. Ma et al. proposed a CP-ABE-based data deletion method to address the issue of data privacy leakage in cloud databases. The experimental results showed that the method can securely and timely delete attribute data, thereby protecting IoT users' personal information [14]. Qaisar et al. designed a CP-ABE-based encryption system to protect the security and privacy of public cloud data by introducing cloud hosts and detecting the security environment of the cloud database. The experimental results showed that the encryption system was able to detect and automatically remove malware from the cloud database, effectively protecting the security and privacy of the cloud database [15]. Mohan et al. improved the bookkeeping and privacy protection of data owners in cloud storage degrees and designed a CP-ABE-based encryption method to enhance the protection of cloud storage. The experimental results showed that the encryption method can protect users' data and private information while also reducing cloud asset consumption [16].

Das and Namasudra believed that the growth of the IoT technology had a significant impact on data security, and CP-ABE had a high computing cost when using bilinear for internal operations. As a result, researchers proposed an encryption scheme based on elliptic curve cryptography to achieve data access control while significantly lowering the decryption cost of traditional CP-ABE systems [17]. To address the attribute revocation security issue in CP-ABE technology and ensure confidentiality and fine-grained access control, Das and Namasudra proposed outsourcing the decryption process to a decryption auxiliary entity. This approach reduced the time required for individual authorizations. Their demonstration showcased the effectiveness and security of this scheme [18]. Given the insecurity of data transmission, Pavithran et al. proposed a new cryptosystem based on a DNA password and finite state machine, which used a DNA character conversion table to increase the randomness of ciphertext. The test results showed that the system was more secure than a single DNA-based cryptosystem [19]. As the main carrier of data transmission, the IoT faces many security risks. Namasudra proposed a new cryptosystem for the basic design of the IoT in the cloud based on DNA cryptography and steganography which used long keys to achieve data encryption and resistance to related security attacks. The findings indicated that the scheme had high effectiveness and security [20]. Rechkoska et al. created a network development robot to ensure the quality of network services and mobile experience, as well as to improve the convenience and efficiency of user resource utilization through the design of Web services and user interfaces, thereby improving the quality of cloud connection experience [21].

The advancement of the Internet and wireless technology has led to significant development in the IoT. Wu et al. proposed enhanced authentication and key agreement protocols to address the security vulnerabilities and vulnerabilities of the intelligent healthcare service protocol, and improved the analysis using real or random models and informal security analysis. The results showed that the proposed protocol exhibited high security and performance stability [22]. Furthermore, Wu et al. proposed to improve the key agreement protocol in remote surgery by using random models and automatic verification tools. The results showed that the protocol passed the security analysis [23]. Wu et al. also conducted research on information authentication and key exchange in smart grids, based on the PAuth enhancement scheme to ensure the security of the improved scheme, and provided a detailed demonstration of the attack steps and implementation code [24]. Chen et al. proposed the lightweight authentication protocol LAP IoHT to address the risks of patient privacy and medical data leakage in the health IoT. They analyzed the security of this protocol using a random model and found that it had better environmental adaptability and performance advantages compared to other protocols [25]. For the problem of digital twin sensitive data in autonomous vehicles, Chen CM scholars and Miao et al. proposed an authentication mechanism to protect privacy, which was deployed in the automotive environment, and the real model test results showed that the method effectively reduced the calculation and transmission costs, with good applicability [26].

The research on private data protection, as described above, involved academics from diverse domains utilizing the CP-ABE algorithm. It is clear that CP-ABE can effectively promote the protection of personal privacy, and some researchers have further improved the security and application of the CP-ABE algorithm by introducing permission control design, IoT authorization, and DNA-based password system design. In view of the high risk of privacy data leakage of IoT users, this study will offer an improved CP-ABE algorithm to achieve attribute-based encryption, fix the shortcomings of its original method, and improve the protection of privacy data of IoT users in light of the increased danger of privacy data leakage among IoT users. With the rapid development of IoT and the explosive growth of user data information, the risk of user data leakage on third-party platforms is increasing. The original encryption algorithms for data protection are limited by the type of user data, so there is a need to propose a more efficient and flexible privacy data protection method. Currently, to ensure the confidentiality of cloud data, the data encryption of data owners makes it more challenging for their visitors to perform ciphertext searches, and the current searchable encryption schemes supported by a limited number of search keywords show lower search matching, and lower access strategy and search efficiency. The existing data protection encryption algorithms are often limited by the type of user data, providing insufficient support for large-scale and diverse IoT user data protection. In contrast, the improved CP-ABE proposed by the study improves the searchable scheme, permission authentication, homomorphism test, and ciphertext similarity adjustment. This is similar to Han et al.'s improved methods for hiding and revoking user-sensitive data [12] and Li et al. weighted ciphertext strategy improvement [13], which can effectively ensure data security to a certain extent. The improved encryption algorithm proposed in the study not only inherits the advantages of previous improved algorithms but also proposes improvements to cloud ciphertext search. Unlike previous scholars who have limited research content, the study has achieved confidentiality of keys and data from two aspects: weight vector encryption and data revocation function supplementation, which can meet the large-scale and diversified user data protection needs in the IoT environment. Weight vector encryption allows access policies to be defined based on vector form constraints, providing more flexibility for fine-grained access control of data. This allows algorithms to define more complex and specific access requirements based on the needs of data owners, adapt to various types of data protection needs, and improve the security protection requirements of algorithms in different application scenarios. This IoT user privacy data protection method fills and enriches the improvement gap of privacy methods in searchable encryption schemes, and greatly improves the processing error rate and delay amount of data.

3. IoT user privacy data protection mechanism design.

3.1. Improved CP-ABE control scheme design.

Attribute-Based-Encryption (ABE) is based on the encryption method of identity, and the data provider encrypts the data in conjunction with an access control scheme, and the cipher text can be decrypted only when the data visitor satisfies the access control policy used for encryption by the data provider, so the difficulty of sharing data is often accomplished using the ABE algorithm [27]. ABE can be divided into two classical directions according to the way the protection policy is deployed: CP-ABE and KP-ABE. In comparison, When encrypting data, the former can automatically set the access control strategy and dictate the size of the sharing. The KPABE scheme does not support both attribute revocation and user identity tracking. Among them, the issue of key abuse has always been a research hotspot of attribute-based encryption mechanisms. Users use attribute sets to identify their identities. If legitimate users share their private keys with other malicious users, they will disrupt the pre-defined access policies, thus causing access security risks. An initialization operation is performed on the access control mechanism in conjunction with a key distribution system (KDC) to facilitate the searchable encryption method to generate the parameters GP and master key MK for CP-ABE, which is then transmitted to the data provider via a secure channel [28]. To ensure data homomorphism, the generation of the key p, which is a naturally occurring positive prime number and $p \in \left[2^{\eta^2-1}, 2^{\eta^2}\right]$, is facilitated in conjunction with DGHV encryption, which has better integer homomorphism. The generated key p is used to encrypt the document set $Data$ and the weight vector set $VSM(Data)$ to obtain the encrypted ciphertexts $Data_{CT}$ and $VSM(Data)_{CT}$, where $VSM(Data)_{CT}$ can be abbreviated as VC. The homomorphic key ciphertexts CT, $Data_{CT}$ and $VSM(Data)_{CT}$ are transferred to the public cloud storage server HDFS, and the encrypted data will not be leaked during the transfer. In order to facilitate a flexible search of ciphertexts and improve the performance of ciphertext search, a searchable encryption scheme is introduced, and only when the visitor meets the legitimacy of the access, the user is able to implement the ciphertext search command and obtain the authentication of access rights. The user who meets the access requirements then needs to provide the search item st and preprocess it to obtain the weight vector of $VSM(st)$, so the weight vector cipher of st is $VSM(st)_{CT}$, which can be abbreviated as VC_{st}. The similarity between the documents d_j and st inside $Data$ is then calculated as shown in Equation (1).

$$sim\left(VC_{d_j}, VC_{st}\right) = \frac{\sum\limits_{i=1}^{m} w_{ij} \times w_i}{\sqrt{\sum\limits_{i=1}^{m} w_{ij}^2} \times \sqrt{\sum\limits_{i=1}^{m} w_i^2}} \tag{1}$$

In Equation (1), $sim\left(CV_{d_j}, CV_{st}\right)$ is used to describe the similarity between the weight vector ciphertext of the document VC_{d_j} and the weight vector ciphertext of the data visitor VC_{st}. w_i is used to describe the weight of the keyword on the retrieved item st. Because the efficiency of the DGHV homomorphic encryption method is inefficient, it is necessary to confirm whether the ciphertext search operation in the cloud satisfies homomorphism, as shown in Equation (2).

$$\sum_{i=1}^{m} w_{ij}^2, \sum_{i=1}^{m} w_i^2, \sum_{i=1}^{m} w_{ij} \times w_i \tag{2}$$

In Equation (2), w_{ij} denotes the TF-IDF weights of the keywords in each document in the data owner's document set. Through these steps, the data visitor is able to download the search results from the public cloud and decrypt them using a homomorphic key to obtain $sim\left(CV_{d_j}, CV_{st}\right)$. The similar results are sorted in descending order to obtain the top N documents with greater similarity, and then the top N documents with the greater similarity are downloaded from the public cloud, ensuring the privacy of the cloud search while improving the efficiency of the search.

However, in the whole control scheme, the DGHV algorithm must control the operation of the ciphertext within a limited number of times, or else the result obtained from the search lacks the homomorphic feature, and DGHV is less efficient in running division and opening square, and the ciphertext search operation process is more complicated, leading to a higher risk of eventually leaking private data [29]. The replacement of the RAS algorithm better ensures that the generated key and the data in the splitting and encryption process are all random in nature. This randomness increases the difficulty of the attacker to guess the key and data and improves the security of the algorithm. Random Asymmetric Splitting makes (RAS) it possible for the different parts of the key to be stored in multiple places or even processed by different entities, which provides a high level of security. Therefore, the superior RAS algorithm is utilized to replace the DGHV symmetric homomorphic encryption algorithm. Replacing the DGHV symmetric homomorphic encryption algorithm with the RAS algorithm requires the encryption operation of the weight vector to be completed. The randomness of RAS needs to be broken first, as shown in Equation (3).

$$x[i] = x_a[i] + x_b[i] \tag{3}$$

In Equation (3), x represents the m dimensional vector, which is split to give two randomly separated vectors x_a and x_b, and i represents the dimensionality. Simplifying equation (3), $x = x_a + x_b$ can be obtained. m which number product of the dimensional vector x and the dimensional vector m, y is calculated as shown in Equation (4).

$$x \cdot y = x_a \cdot y + x_b \cdot y \tag{4}$$

The RAS algorithm splits each vector dimension to obtain more randomly separated vectors. If the vector x and the vector y both have dimension 2, the vector x and the vector y can be expressed as:

$$x = (x_1, x_2), y = (y_1, y_2) \tag{5}$$

The dot product of the two vectors can be obtained by Equation (5) as shown in Equation (6).

$$x \cdot y = x_1 \cdot y_1 + x_2 \cdot y_2 \tag{6}$$

The two vectors are split as shown in Equation (7).

$$\begin{cases} X = \{x_a, x_b\} \\ Y = \{y_a, y_b\} \end{cases} \tag{7}$$

The vector x is split in the second dimension, and the vector y is split in the first dimension. The split vector relationship is shown in Equation (8).

$$\begin{cases} x_a = (x_1, x_{2a}), x_b = (x_1, x_{2b}) \\ y_a = (y_{1a}, y_2), y_b = (y_{1b}, y_2) \end{cases} \tag{8}$$

In Equation (8), y_{1a}, y_{1b} represent the random separation vectors on the first dimension, and x_{2a}, x_{2b} represent the random separation vectors on the second dimension. The split vector satisfies the condition that:

$$\begin{cases} x_{2a} + x_{2b} = x_2 \\ y_{1a} + y_{1b} = y_1 \end{cases} \tag{9}$$

Combining the RAS algorithm to split the two vectors x and y, two randomly separated vectors X and Y are then obtained, with equal dot products between them, as shown in Equation (10).

$$X \cdot Y = x_a \cdot y_a + x_b \cdot y_b = x_1 \cdot y_1 + x_2 \cdot y_2 = x \cdot y \tag{10}$$

The CP-ABE algorithm privacy data protection system model is shown in Figure 1.

As shown in Figure 1, use the key distribution center (KDC) as the third party, and use the public cloud to compute and store data. The data provider must create vectors and encrypt the uploaded documents in a homomorphic manner, as well as set corresponding access policies for the keys. The data visitor must first obtain a homomorphic encryption key, and enter search terms. Data can only be viewed after ciphertext search and clear text data (Data) acquisition. After obtaining a homomorphic encryption key, the data accessor can search the private data in the public cloud data sharing center, and, in conjunction with , the data visitor has access to the data and obtains the plaintext . In the traditional CP-ABE algorithm, the encryption keys for attributes are defined by the data owner in the access policy, which leads to an exponential growth in the number of keys. R-CP-ABE algorithm introduces randomness to generate encryption keys with the help of RAS, which effectively reduces the number of keys and improves the efficiency of

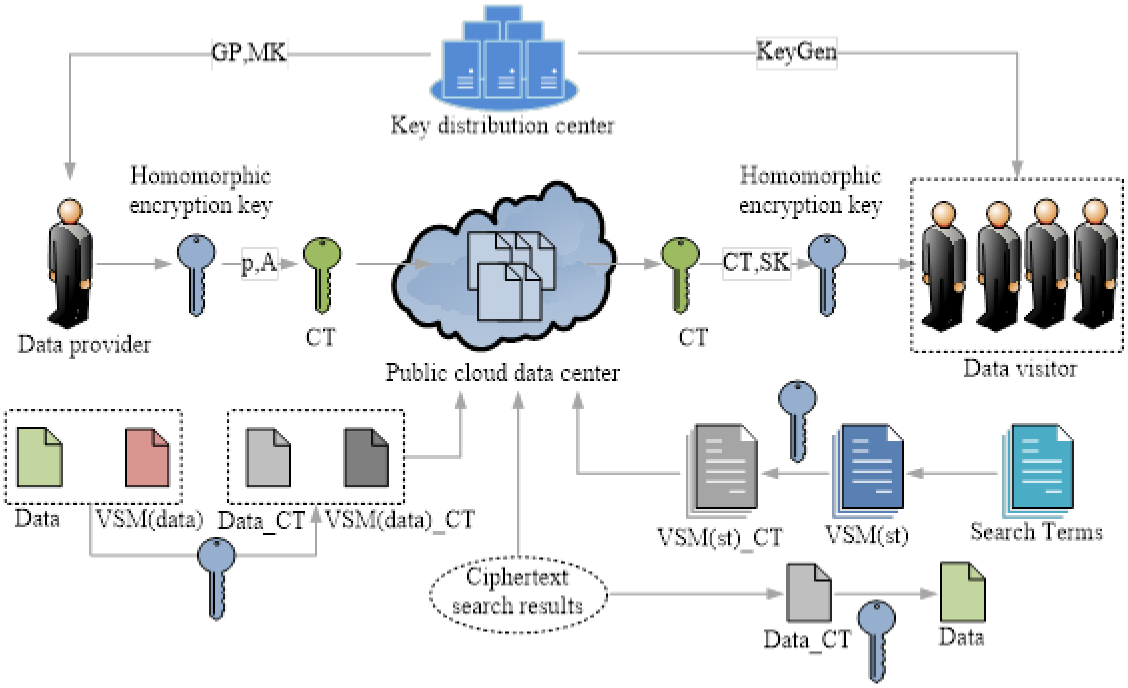


Figure 1. System model

the algorithm. The introduction of the splitting technique RAS can split the user's private key into multiple parts so that the user can decrypt the data by possessing only some of the attributes of the private key. The R-CP-ABE algorithm reduces the complexity of the key management and its generation is random in nature so it is difficult for the attacker to get the key information by analyzing it, thus increasing the efficiency of the algorithm. Simultaneously, because the traditional CP-ABE lacks the authority to revoke data visitors and attributes, and there is a risk of user privacy data leakage, the traditional CP-ABE will be improved so that the optimized control scheme has the function of double revocation to meet the needs in practical applications [30].

However, traditional protocols often face the risk of data interception during synchronization attacks, resulting in potential data leakage and loss. Therefore, research is being conducted on adding security protocols to resist desynchronization attacks in the process of data privacy protection. The specific content is that the server will use permutation functions to verify keys, random numbers, etc. before executing the agreement, and the client will display relevant identification data. When the protocol starts to execute, when the client sends the corresponding identification data to the server, the server will traverse and shrink it. If successful, the ciphertext retrieval will be synchronously initiated, and the server will generate a random number and calculate the checksum value based on the cross bit. The checksum calculation is shown in the Equation (11).

$$CS = CRC(Cor(K, \oplus Kd)), Cn \tag{11}$$

In Equation (11), $Cor()$ represents the cross-bit calculation function, $CRC()$ is the verification function, K, K_d is the public key, and C_n is the value of the server. According to the verification function, the server and client can update the server's keys and random numbers. The R-CP-ABE algorithm that supports the revocation function consists of two parts: $Setup(k)$ and $KeyGen(ID, S, MK, GP)$. Optimization of CP-ABE will modify the similarity calculation Equation (1) to obtain Equation (12).

$$sim\left(CV_{d_j}, CV_{st}\right) = \frac{CV_{dj} \cdot CV_{st}}{|CV_{dj}| \times |CV_{st}|} = \frac{\sum\limits_{i=1}^{m} w_{ij} \times w_i}{\sqrt{\sum\limits_{i=1}^{m} w_{ij}^2} \times \sqrt{\sum\limits_{i=1}^{m} w_i^2}} = \frac{a}{\sqrt{b}} \tag{12}$$

In Equation (12), a, b denote the constants, the value of b is large and will be computed several times, so to improve the efficiency of the searchable encryption mechanism, Equation (12) is simplified as follows.

$$sim'(CV_d, CV_{st}) = \sum_{i=1}^{m} w_{ij} \times w_i \tag{13}$$

The simplified formula for the calculation of the weight of $TF - IDF$ is still complex and reduces the efficiency of the ciphertext search to a certain extent, so a more suitable formula needs to be chosen instead of Equation (12).

$$Score(D, Q) = D \cdot Q = \sum_{t_i \in st} TF_{ij} \times IDF_i \tag{14}$$

In Equation (14), D is used to describe the search vector for a document in the data provider's document set $Data$, Q is meant to be a vector of keywords that the data visitor is waiting to search for, which is used to describe the TF_{ij}. TF value of the keyword t_i in the document d_j, IDF_i is used to describe the IDF value of the keyword t_i in the document set, and the length of the two vectors D and Q is equal to the total number of keywords m. TF_{ij} is calculated as shown in Equation (15).

$$TF_{ij} = \frac{TF'_{d_j,i}}{\sqrt{\sum_{i=1}^{m} (TF'_{d_j,i})^2}} \tag{15}$$

In Equation (15), $TF'_{d_j,i}$ is used to describe the TF value of t_i as it appears in d_j and IDF_i is computed as shown in Equation (16).

$$IDF_i = \frac{IDF'_i}{\sqrt{\sum_{i=1}^{m} (IDF'_i)^2}} \tag{16}$$

In Equation (16), IDF'_i is used to describe the IDF value of t_i as it appears in $Data$. Improvements to the CP-ABE algorithm enable an optimized system model to be obtained, as shown in Figure 2.

In Figure 2, the improved CP-ABE protection mechanism with revocation combines GP and MK to search for ciphertext data in the public cloud data-sharing center. Unlike the traditional model, the optimized model has two keys CT, allowing it to search for plaintexts in the cloud database faster and obtain plaintexts $Data$ faster.

3.2. **IoT user privacy data protection mechanisms combined with improved CP-ABE.** IoT data consists of data providers and data visitors, and the entire IoT system is made up of data providers, data visitors, and cloud storage platforms [31]. The data provider is the data owner of specific private data and can dictate access to the private data. The data visitor needs to obtain access consent from the data provider to access the private data, and when the data visitor obtains access permission, the visitor can download and decrypt the private data from the cloud storage platform. The cloud

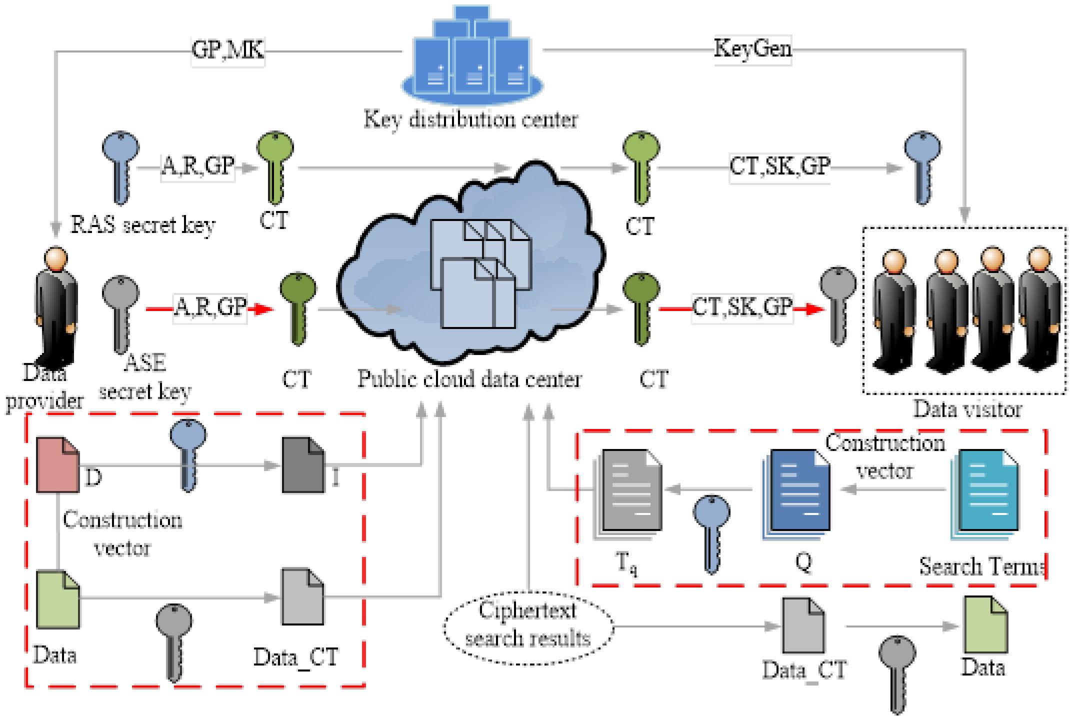


Figure 2. Optimized system model

storage platform stores the private data of IoT users, and generally, the private data are encrypted into highly private ciphertext packets [32]. In this regard, the flow of the data provider generating private data and the data visitor accessing a particular private data is shown in Figure 3.

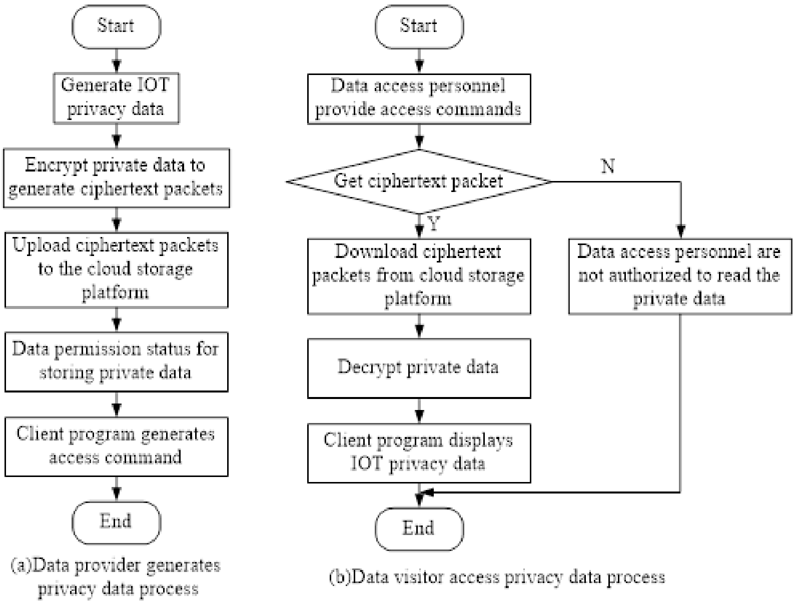

Start
Generate IOT privacy data
Encrypt private data to generate ciphertext packets
Upload ciphertext packets to the cloud storage platform
Data permission status for storing private data
Client program generates access command
End
(a)Data provider generates privacy data process
Start
Data access personnel provide access commands
Get ciphertext packet
N
Y
Download ciphertext packets from cloud storage platform
Data access personnel are not authorized to read the private data
Decrypt private data
Client program displays IOT privacy data
End
(b)Data visitor access privacy data process

Figure 3. SRCNN model structure

In Figure 3, after the data provider generates the IoT privacy data, it encrypts it using encryption methods to obtain the ciphertext data package, transmits it to the public cloud data platform, and generates an access command to that privacy data, i.e. successfully generates a specific privacy data. When a data visitor needs to access the privacy data, the visitor must first display the access command established by the data provider before retrieving the ciphertext data packet from the database. If the access is successful, the visitor can download the ciphertext data packet directly from the cloud storage platform and decrypt it to complete the access to the privacy data [33]. If the access to the ciphertext data fails, it means that the accessor does not have the right to access the private data. The workflow of the IoT user privacy data protection mechanism combined with the improved CP-ABE is shown in Figure 4.

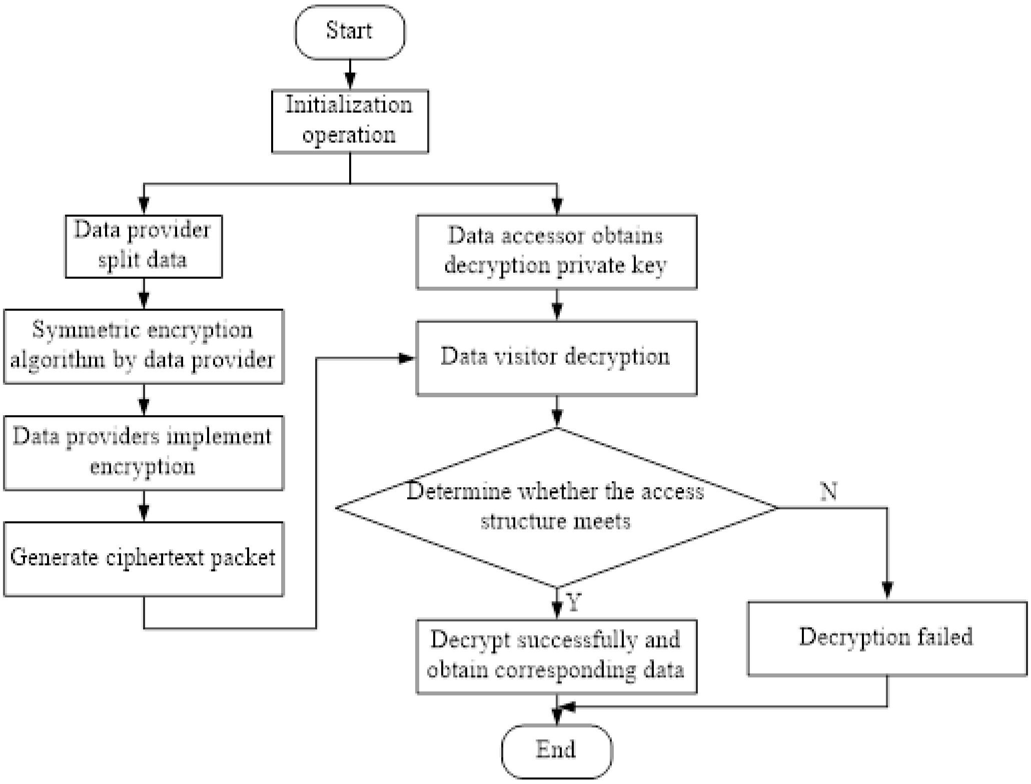

Start
Initialization operation
Data provider split data
Symmetric encryption algorithm by data provider
Data providers implement encryption
Generate ciphertext packet
Data accessor obtains decryption private key
Data visitor decryption
Determine whether the access structure meets
N
Y
Decrypt successfully and obtain corresponding data
Decryption failed
End

Figure 4. Flow chart of optimized IoT user privacy data protection mechanism

Figure 4 shows that the improved CP-ABE-based IoT user privacy data protection requires pre-processing first, and then separating and encrypting the privacy data to produce ciphertext data packets; the data accessor must have the decryption private key to decrypt the data if it meets the access requirements, the privacy data can be decrypted successfully and the privacy data can be obtained; if it does not meet the access requirements, that means the private data cannot be obtained. Therefore, the encryption step based on IoT user data can effectively authenticate the user visitor and encrypt the information in a hierarchical attribute during data processing, ensuring data accuracy and relevance under heavy data volume, and thus achieving privacy data protection.

4. Security analysis of user privacy data protection mechanisms.

4.1. **Performance analysis of the improved CP-ABE algorithm on user privacy data.** Due to the growing number of IoT users and the redundancy of multi-label data, current user information data exhibits characteristics such as diversity and privacy. Convenient network facilities bring great convenience to people's lives, but they also invariably increase the risk of leakage of user information data, making user data security and privacy better guaranteed [34]. The study added support for revocation to the original CP-ABE algorithm and created a searchable encryption scheme based on cloud storage, all while collecting IoT user data with the help of crawler technology, and after pre-processing this document data according to the privacy data type characteristics of the dataset, and then using 60% of this dataset as test data and 40% as application data. To verify the performance effectiveness analysis of this encryption algorithm, the latency condition and data processing error rate before and after the improvement of the CP-ABE algorithm were determined, and the results are shown in Figure 5.

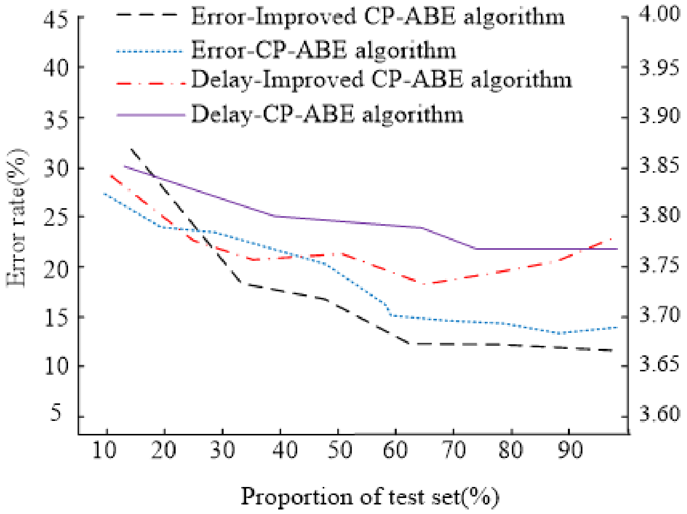


Figure 5. Comparison results of anti overfitting performance and delay of CP-ABE algorithm before and after improvement

Figure 5 shows that the CP-ABE algorithm had a significant difference in the anti-overfitting performance and latency before and after the improvement, with the improved data processing error rate. The anti-fitting performance of the algorithm better demonstrated how it was affected by the external interference environment when adjusting and updating parameters, and its delay reflected the efficiency of the algorithm. CP-ABE algorithm showed a decreasing trend as the percentage of the data set increased, and the slope of its decreasing curve was overall larger than that of the traditional CP-ABE algorithm, with the lowest value reaching 12.36%, which was much lower than the CP-ABE algorithm. The lowest value was 12.36%, which was significantly lower than the lowest value of the CP-ABE algorithm (14.12%). The maximum and minimum delays of the improved algorithm for information encryption were 3.824ms and 3.716ms, respectively, which were much lower than the maximum (3.851ms) and minimum (3.625ms) values of the CP-ABE algorithm, indicating that the algorithm was able to guarantee high performance when encrypting information data. Experimental data was also collected on the accuracy of the algorithm in privacy data feature screening, and the results are shown in Figure 6.

Figure 6(a) depicts the search accuracy results of the improved CP-ABE algorithm before and after different sample data amounts, in which the improved CP-ABE algorithm's search accuracy on data basically fluctuated above and below 80%, with its fluctuation range of no more than 10%, and the accuracy of its data sample points was much higher than the standard baseline value. The traditional CP-ABE algorithm, on the other hand, had a wide range of fluctuation in accuracy for data search, with most of the sample data points falling below the standard baseline value, and its average accuracy was 53.78%,

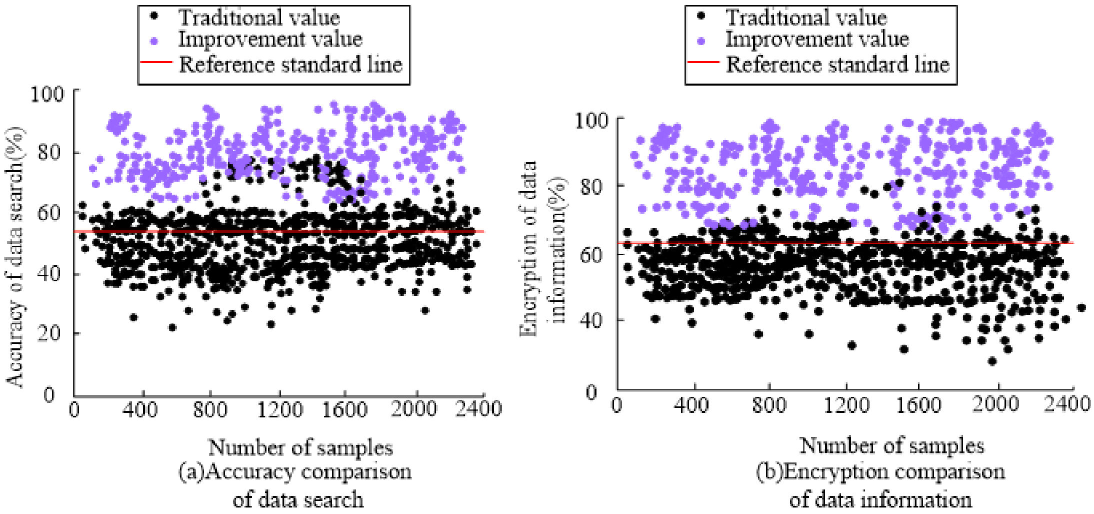


Figure 6. Comparison of accuracy and encryption performance of different algorithms in data processing

which was significantly lower than the effectiveness of the improved CP-ABE algorithm. Figure 6(b) depicts a comparison of the performance of different algorithms in terms of data information encryptability.

The results in the figure indicated that the improved CP-ABE algorithm performed more uniformly in terms of information data confidentiality, with a data confidentiality of 82.36%. However, the traditional CP-ABE algorithm had a basic encryption rate of 58.13% for data, with a maximum value of no more than 80%, and encrypted large sample data. The minimum value obtained was 18.33%. The aforementioned results demonstrated that the improved CP-ABE algorithm excelled in searching and encrypting data information, thereby providing enhanced protection for data with distinct personal characteristics. When using encryption algorithms to test information data, the time and resources consumed were first checked to ensure good practicality and scalability. Therefore, this study tested the application loadability of the algorithm on the basis of establishing a testing environment. The setup of the experimental testing environment is shown in Table 1.

Table 1. Experimental test environment

Name	Parameter
CPU	Intel(R) Core(TM) i7-2600K CPU @3.40GHz
Memory	12.0GB
Operating system	Ubuntu 16.04 64bit
System development tools	Eclipse Juno 23.0., Android Studio 2.2. , Sublime Text 3., MySQL WorkBench 6.3.8
Edition	Android 7.0

After the experimental parameter environment was determined, the data of the algorithm's resource consumption on the server was counted and the results are shown in Figure 7.

Figure 7 shows that the CP-ABE algorithm has a significant difference in time consumed and the percentage of resources consumed before and after improvement. Figure 7(a)

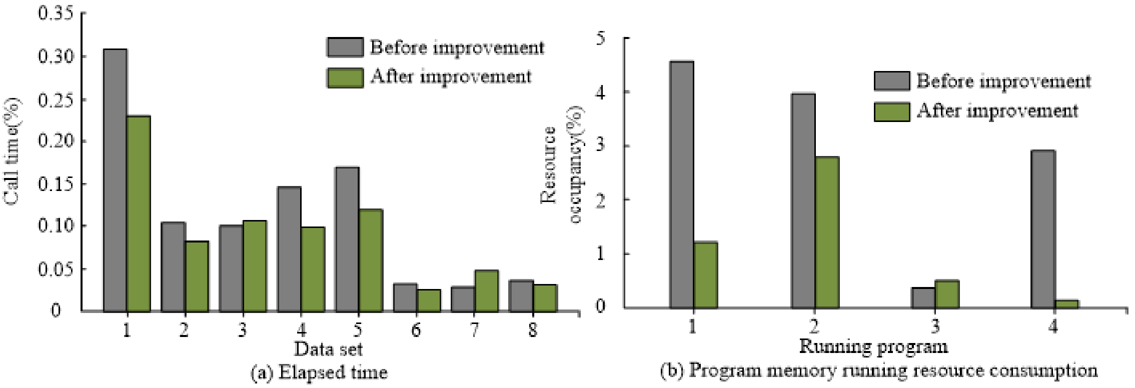


Figure 7. Comparison of time consumption and memory running resource consumption before and after algorithm improvement

shows that the improved algorithm took slightly less time under Dataset 4 and Dataset 7 than the algorithm before the improvement, while the time consumption in datasets with higher information content data was basically less than 0.10%. Figure 7(b) shows that the improved algorithm consumed an average of 1.05% of the memory resources, indicating its overall good adaptability.

4.2. **Analysis of the application of improved algorithms for protection mechanisms on IoT privacy data.** The rapid development of IoT technology makes it more convenient to obtain and distribute information, while also raising awareness of the importance of information encryption protection. The safety of encryption algorithms in information security protocols is linked to the safety of people's data. The key stream generated by a secure encryption algorithm must be random, and the size of its random independence is proportional to the security of the algorithm, that is, the greater the randomness, the higher the security. This idea was used in the research to check the security of the algorithm proposed in the research, and it was analyzed under the built privacy protection mechanism platform to reduce the probability of the occurrence of individual bytes, as well as to reduce the analysis cost while ensuring correctness. The randomness analysis results of the algorithm are shown in Table 2.

Table 2 shows that the improved CP-ABE algorithm and its key stream sequence of the CP-ABE algorithm passed the test results. However, the results of the improved algorithm in the frequency test, intra-block frequency test, and overlapping module matching test were higher than the data before the improvement, and the data difference between the two algorithms in the random walk state frequency test was -0.05255. The key stream sequence of the improved algorithm was better than the algorithm before the improvement, indicating that it had high randomness and security. To assess the effectiveness of the algorithm employed in the research on IoT privacy data, the study utilized a dataset consisting of privacy data. The application of the optimized encryption scheme was analyzed, and a performance analysis comparing the improved CP-ABE algorithm used in the research with TRSE, a similar searchable encryption scheme, was conducted. This analysis aimed to explore the search efficiency and classification accuracy of user information collection, the results are shown in Figure 8.

The results in Figure 8 show that the Institute's improved CP-ABE algorithm performed better in terms of indexing time consumption and search time than the traditional CP-ABE algorithm and TRSE search scheme in terms of indexing time construction. As the amount of information data increased, the improved CP-ABE algorithm maintained

Table 2. Randomness test results of improved CP-ABE algorithm and CP-ABE algorithm

Test Category	Improved CP-ABE algorithm		CP-ABE algorithm		Difference
	P value	Result	P value	Result	
Frequency inspection	0.42314	Pass	0.56243	Pass	0.13929
Intra block frequency test	0.44289	Pass	0.58241	Pass	0.13952
Run test	0.51542	Pass	0.64672	Pass	0.13130
Rank test of binary matrices	0.42177	Pass	0.40581	Pass	-0.01596
Discrete Fourier transform test	0.47258	Pass	0.44756	Pass	-0.02502
Overlapping module matching test	0.38452	Pass	0.42487	Pass	0.04035
Linear complexity test	0.42571	Pass	0.60376	Pass	0.17805
Sequence test	0.39673	Pass	0.37559	Pass	-0.02114
Approximate entropy test	0.45426	Pass	0.30372	Pass	-0.15054
Random walk test	0.51143	Pass	0.56539	Pass	0.05396
Random walk state frequency test	0.54013	Pass	0.48758	Pass	-0.05255

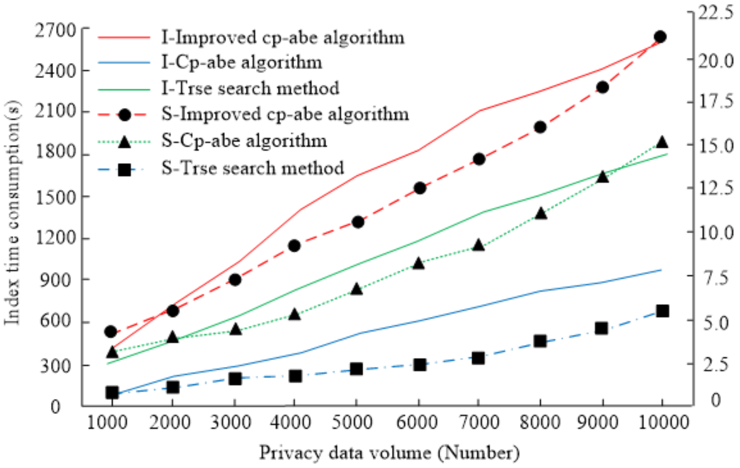
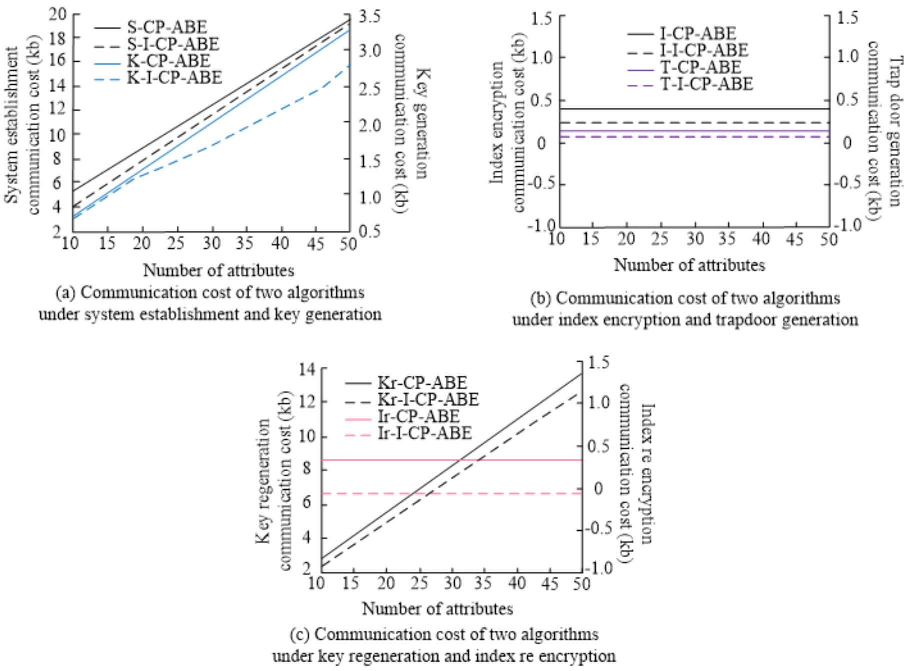


Figure 8. Comparison of search efficiency and classification accuracy of user information collection

a more stable time consumption trend, and the time consumption after the threshold of 6000 information data was less than 600 seconds. However, due to the subjectivity and uncertainty of its range information, it ranked second only to the improved CP-ABE algorithm in terms of search performance when performing private data [35]. The index encryption algorithms used in different access control structures consumed different amounts of computing and communication resources, as well as had a certain impact on the encryption burden of the data owner and the search efficiency of the server. The costs of various encryption algorithms were examined and compared within the context of the access control structure, and sorted out the data content. Figure 9 depicts the results.

In Figure 9, "I-CP-ABE" represents an improved algorithm. The results showed that the improved CP-ABE algorithm had a lower communication cost than the CP-ABE algorithm in system establishment key generation, index encryption, trap gate generation, key re-generation, and index re-encryption, with variation ranges of 2.4%, 3.7%, 2.6%, 1.3%, 2.8%, and 4.2%, respectively. The overall cost of the improved algorithm was



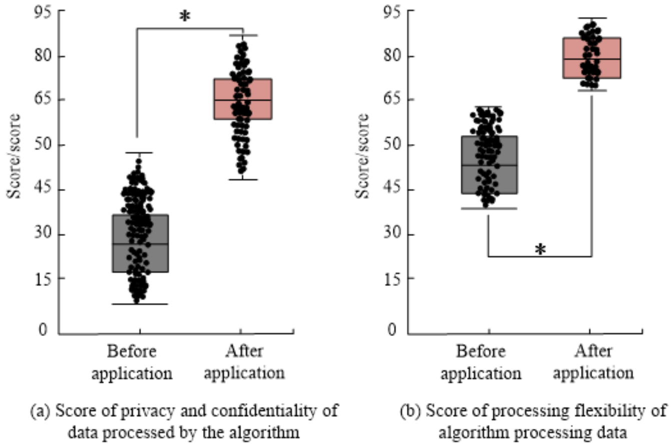
(a) Communication cost of two algorithms under system establishment and key generation

(b) Communication cost of two algorithms under index encryption and trapdoor generation

(c) Communication cost of two algorithms under key regeneration and index re encryption

Figure 9. Communication cost statistics of two algorithms under different operation steps

relatively low. Then, the application validity of the algorithm in privacy data protection was then statistically analyzed, and the results are shown in Figure 10.



(a) Score of privacy and confidentiality of data processed by the algorithm

(b) Score of processing flexibility of algorithm processing data

Figure 10. Score statistics of privacy, confidentiality and processing flexibility of data

The application algorithm was applied to the experimental data application process, and the data were statistically calculated before and after the experiment for the user's intended privacy confidentiality and processing flexibility scores. The results showed that the user's confidentiality score for the data increased from 18.24 to 65 after the application experiment, and their processing flexibility score for the search data increased from 45.12 to 80.13, and the data before and after the experiment had the same value. Before and after the experiment, the data were statistically significant ($P < 0.05$). The above results demonstrated that the algorithm can improve search rate while also enhancing data confidentiality when performing user data protection.

5. **Conclusion.** The IoT, as a network infrastructure that connects the information world and the physical world, has characteristics and requirements such as collaborative management and perceptual interaction, leading to severe security challenges in communication data. With the popularization of IoT devices and the large-scale collection of data, personal privacy faces greater risks of leakage. Protecting IoT data privacy helps prevent personal information leakage, illegal monitoring, and data abuse. This research focuses on improving CP-ABE and designing searchable solutions on the Hadoop platform, incorporates revocation support into the CP-ABE algorithm, and creates a protection mechanism in the form of a cloud platform carrier that meets user requirements and private data characteristics.

The improved algorithm is subjected to application performance testing, and the results showed that the data processing error rate of the improved CP-ABE algorithm showed a decreasing trend as the percentage of data sets increased, and the slope of its decreasing curve was overall larger than that of the traditional CP-ABE algorithm, with the lowest value reaching 12.36%, which was significantly lower than the lowest value of the CP-ABE algorithm (14.12%). The overall accuracy was much higher than the standard baseline value and the average accuracy of the traditional CP-ABE algorithm (53.78%). In terms of data information confidentiality performance, the improved CP-ABE algorithm achieved 82.36% confidentiality of information data, which was significantly higher than the CP-ABE algorithm. The application generalisability of the encryption algorithm was an important aspect of testing its performance, and in the results, it was discovered that the time consumption of the improved algorithm was essentially below 0.10% for data sets with higher information content data, and the memory resources were 1.05% on average. The effectiveness of the algorithm's user evaluation was investigated, and it was discovered that users' scores for data confidentiality improved from 18.24 to 65 after the algorithm was applied, and their scores for processing flexibility in searching data improved from 45.12 to 80.13, with statistically significant data before and after the experiment ($P < 0.05$). The improved algorithm used in the study can achieve hierarchical access to private data for each user role and has higher decryption efficiency, as well as better applicability and practicality. Future work should focus on addressing scalability and other issues related to the privacy data of IoT users. During the application process, attention must be given to the suitability of selected IoT device types, application scenarios, and methods, as well as the feasibility and cost-effectiveness of large-scale deployment. Moreover, it is crucial to consider the diverse needs of various IoT users. While the research content often involves basic theories, future research should concentrate on combining this encryption algorithm with practical application scenarios to provide better ideas for protecting user data privacy.

REFERENCES

[1] A. Beduschi, "Digital identity: contemporary challenges for data protection, privacy and non-discrimination rights," Big Data and Society, vol. 6, no. 2, pp. 1-6, 2019.

[2] E.-A. Ahmed, "An extended data protection model based on cipher-text-policy attribute based encryption model and an XACML framework in cloud computing," Journal of Advanced Science, vol. 28, no. 16, pp. 1021-1033, 2019.

[3] N. Momen, M. Hatamian, and L. Fritsch, "Did app privacy improve after the GDPR," IEEE Security and Privacy, vol. 17, no. 6, pp. 10-20, 2019.

[4] S. Xue, and 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.

[5] Y. Yu, L. Guo, S. Liu, J. Zheng, and H. Wang, "Privacy protection scheme based on CP-ABE in crowdsourcing-IoT for smart ocean," IEEE Internet of Things Journal, vol. 7, no. 10, pp. 10061-10071, 2020.

[6] C. Yadav, B. Patro, and V. Yadav, "Design engineering Aes-light weight CP-ABE based privacy protection framework with effective access control mechanism in cloud framework," Design Engineering (Toronto), vol. 2021, no. 6, pp. 2321-2336, 2021.

[7] P.-S. Challagidad, and M.-N. Birje, "Efficient multi-authority access control using attribute-based encryption in cloud storage," Procedia Computer Science, vol. 167, no. 1, pp. 840-849, 2020.

[8] T. Feng, X. Yin, Y. Lu, J. Fang, and F. Li, "A searchable CP-ABE privacy preserving scheme," International Journal of Network Security, vol. 21, no. 4, pp. 680-689, 2019.

[9] H. Tian, X. Li, H. Quan, C.-C. Chang, and T. Baker, "A lightweight attribute-based access control scheme for intelligent transportation system with full privacy protection," IEEE Sensors Journal, vol.32, no. 99, pp. 1-1, 2020.

[10] H. Xiong, Y. Zhao, L. Peng, and K.-H. Yeh, "Partially policy-hidden attribute-based broadcast encryption with secure delegation in edge computing," Future Generation Computer Systems, vol. 97, no. 8, pp. 453-461, 2019.

[11] L. Zhang, Y. Cui, and Y. Mu, "Improving security and privacy attribute based data sharing in cloud computing," IEEE Systems Journal, vol. 14, no. 1, pp. 387-397, 2020.

[12] D. Han, N. Pan, and K.-C. Li, "A traceable and revocable ciphertext-policy attribute-based encryption scheme based on privacy protection," IEEE Transactions on Dependable and Secure Computing, vol. 19, no. 1, pp. 316-327, 2020.

[13] H. Li, K. Yu, B. Liu, C. Feng, Z. Qin, and G. Srivastava, "An efficient ciphertext-policy weighted attribute-based encryption for the internet of health things," IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 5, pp. 1949-1960, 2021.

[14] J. Ma, M. Wang, J. Xiong, and Y. Hu, "CP-ABE-based secure and verifiable data deletion in cloud," Security and Communication Networks, vol. 2021, no. 3, pp. 1-14, 2021.

[15] Z.-H. Qaisar, R. Li, S.-H. Almotiri, M.-A. Ghamdi, A.-A. Nagra, and G. Ali, "A scalable and efficient multi-agent architecture for malware protection in data sharing over mobile cloud," IEEE Access, vol. 35, no. 99, pp. 1-12, 2021.

[16] A. Mohan, and P. Vamshikrishna, "Accounting and privacy preserving of data owner in cloud storage," International Journal of Innovative Technology and Exploring Engineering, vol. 10, no. 6, pp. 14-17, 2021.

[17] S. Das, and S. Namasudra, "Multi-authority CP-ABE-based access control model for IoT-enabled healthcare infrastructure," IEEE Transactions on Industrial Informatics, vol. 19, no. 1, pp. 821-829, 2022.

[18] S. Das, and S. Namasudra, "MACPABE: multi-authority-based CP-ABE with efficient attribute revocation for IoT-enabled healthcare infrastructure," International Journal of Network Management, vol. 33, no. 3, pp. e2200, 2022.

[19] P. Pavithran, S. Mathew, S. Namasudra, and A. Singh, "Enhancing randomness of the ciphertext generated by DNA-based cryptosystem and finite state machine," Cluster Computing, vol. 2023, no. 26, pp. 1035-1051, 2022.

[20] S. Namasudra, "A secure cryptosystem using DNA cryptography and DNA steganography for the cloud-based IoT infrastructure," Computers and Electrical Engineering, vol. 104, pp. 108416, 2022.

[21] U. Rechkoska, D. Djamtovski, D. Davcev, C. Ciulla, and J. Sikoski. "Design and development of an android accounting application using web services and quality of experience for mobile computing," AICT2014: The Tenth Advance International Conference on Telecommunications, pp. 181-186, 2014.

[22] T.-Y. Wu, L. Wang, and C.-M. Chen, "Enhancing the security: a lightweight authentication and key agreement protocol for smart medical services in the IoHT," Mathematics, vol. 11, no. 17, pp. 3701, 2023.

[23] T.-Y. Wu, Q. Meng, L. Yang, K. Saru, and P. Matin, "Amassing the security: an enhanced authentication and key agreement protocol for remote surgery in healthcare environment," CMES-Computer Modeling in Engineering & Sciences, vol. 134, no. 1, pp. 317-341, 2023.

[24] T.-Y. Wu, Y.-Q. Lee, C.-M. Chen, Y. Tian, and N.-A. Al-Nabhan, "An enhanced pairing-based authentication scheme for smart grid communications," Journal of Ambient Intelligence and Humanized Computing, 2021.

[25] C.-M. Chen, Z. Chen, S. Kumari, and M.-C. Lin, "LAP-IoHT: A lightweight authentication protocol for the internet of health things," Sensors, vol. 22, no. 14, pp. 5401, 2022.

[26] C.-M. Chen, Q. Miao, S. Kumar, and T.-Y. Wu, "Privacy-preserving authentication scheme for digital twin-enabled autonomous vehicle environments," Transactions on Emerging Telecommunications Technologies, vol. 34, no. 11, pp. e4751, 2023.

[27] M.-M. Abbassy, "The human brain signal detection of health information system in EDSAC: A novel cipher text attribute based encryption with EDSAC distributed storage access control," Journal of Advanced Research in Dynamical and Control Systems, vol. 12, no. SP7, pp. 858-868, 2020.

[28] S. Son, J. Lee, M. Kim, S. Yu, A.-K. Das, and Y. Park, "Design of secure authentication protocol for cloud-assisted telecare medical information system using blockchain," IEEE Access, vol. 8, no. 121, pp. 192177-192191, 2020.

[29] L. Yang, C. Li, Y. Cheng, S. Yu, and J. Ma, "Achieving privacy-preserving sensitive attributes for large universe based on private set intersection," Information Sciences, vol. 582, no. 230, pp. 529-546, 2022.

[30] K. Fan, T. Liu, K. Zhang, H. Li, and Y. Yang, "A secure and efficient outsourced computation on data sharing scheme for privacy computing," Journal of Parallel and Distributed Computing, vol. 135, no. 1, pp. 169-176, 2020.

[31] Z. Guan, X. Lu, W. Yang, L. Wu, N. Wang, and Z. Zhang, "Achieving efficient and privacy-preserving energy trading based on blockchain and ABE in smart grid," Journal of Parallel and Distributed Computing, vol. 147, no. 7, pp. 34-45, 2021.

[32] R. Katariya, and A. Dangi, "A review on CP-ABE for big data access control in cloud computing," International Journal of Computer Sciences and Engineering, vol. 7, no. 4, pp. 941-943, 2019.

[33] S. Zhou, G. Chen, G. Huang, J. Shi, and T. Kong, "Research on multi-authority CP-ABE access control model in multicloud," China Communications, vol. 17, no. 8, pp. 220-233, 2020.

[34] D. Kim, J.-R. Woo, J. Shin, J. Lee, and Y. Kim, "Can search engine data improve accuracy of demand forecasting for new products? evidence from automotive market," Industrial Management and Data Systems, vol. 119, no. 5, pp. 1089-1103, 2019.

[35] R. Xu, J. Joshi, and P. Krishnamurthy, "An Integrated privacy preserving attribute based access control framework supporting secure deduplication," IEEE Transactions on Dependable and Secure Computing, vol. 26, no. 99, pp. 1-1, 2019.