Hindawi
Journal of Function Spaces
Volume 2023, Article ID 9756926, 1 page
https://doi.org/10.1155/2023/9756926

Retraction

Retracted: Mathematical Modeling of Static Data Attribute Encryption Based on Big Data Technology

Journal of Function Spaces

Received 12 December 2023; Accepted 12 December 2023; Published 13 December 2023

This article has been retracted by Hindawi, as publisher, following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of systematic manipulation of the publication and peer-review process. We cannot, therefore, vouch for the reliability or integrity of this article.

Please note that this notice is intended solely to alert readers that the peer-review process of this article has been compromised.

Wiley and Hindawi regret that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] Y. Liu and Q. Zhang, "Mathematical Modeling of Static Data Attribute Encryption Based on Big Data Technology," *Journal of Function Spaces*, vol. 2022, Article ID 4292063, 10 pages, 2022.

Hindawi
Journal of Function Spaces
Volume 2022, Article ID 4292063, 10 pages
https://doi.org/10.1155/2022/4292063

Research Article

Mathematical Modeling of Static Data Attribute Encryption Based on Big Data Technology

Yutang Liu[1] **and Qin Zhang**[2]

[1]*School of Science, Henan Institute of Technology, Xinxiang, Henan 453003, China*
[2]*School of Mathematics and Statistics, Xinxiang University, Xinxiang, Henan 453003, China*

Correspondence should be addressed to Qin Zhang; zhangqin@xxu.edu.cn

Received 12 May 2022; Revised 8 July 2022; Accepted 15 July 2022; Published 2 August 2022

Academic Editor: Miaochao Chen

Attribute encryption is an effective one to many network communication technologies, which supports flexible access control strategies and is very suitable for fine-grained access control in large-scale information systems. In order to improve the attributes of static data, encryption technology can provide a reliable technical guarantee for network security. This paper presents a mathematical modeling method of static data attribute encryption based on big data technology. The big data redundancy elimination algorithm based on similarity calculation is analyzed. By using static data attribute encryption based on big data technology, the length of encrypted data packets will not increase, and partial redundancy of fragments can be eliminated, which can greatly improve the efficiency of the system. The attribute-based encryption mechanism uses attributes as public keys, and the decryption user is a group; so, the encryption efficiency is very high. It can realize efficient encryption and decryption, as well as flexible access control based on user attributes. This scheme can reflect the importance of attributes; so, it is more practical.

1. Introduction

Along with the continuous reform and opening up, China's economy has significantly improved, Internet technology has been widely popularized, and the increasing level of mobile device technology has promoted the further expansion and exponential growth of data volume [1]. At the same time, information technology attacks have emerged continuously, and users' private files cannot be protected by traditional password methods alone, which is a great security risk [2]. Lingfeng proposes a new big data encryption algorithm based on data deduplication technology. This method eliminates the redundancy in big data processing and fully reflects the advantages of elliptic curve encryption algorithm, such as high security and short key. At the same time, combined with CTR working mode, it fully shows the advantages of good parallelism and fast speed [3]. Big data are data sets with huge amount of data and many types of structures, which are difficult to be processed using existing relational databases or data processing tools [4]. Encrypting data on the basis of big data properties is one of the effective methods to control access to data users in a data outsourcing environment [5]. The characteristics of big data are mostly obvious, such as large volume, diversity, and complex sources, and some of them contain a large amount of information content with certain value [6]. Traditional encryption techniques convert private data information into unrecognizable ciphertext data, losing many semantic features of the data and rendering conventional data processing methods ineffective [7]. Thus, it severely hinders the data computing and processing services provided by cloud service providers to their tenants and is not applicable to the protection of private data in cloud platforms.

Therefore, the impact of factors such as security level and number of attributes on the algorithm should be analyzed in depth through the in-depth study of attribute

encryption, combined with the actual needs of engineering applications using programming simulation in terms of time overhead, communication overhead, CPU, and memory overhead [8]. Database system as the core component of information system and database files as the aggregation of information, obviously, its security will be an important indicator of information system security. This reinforces the need to protect not only economic information in business communications but also personal information in communications [9]. Databases must provide data protection measures and cannot leak information to unauthorized parties. For the present, there are various ways to encrypt network data attributes, including link encryption, node encryption, and end-to-end encryption. However, in the current distributed and network-based application environment, users access the system in various ways, and the database system faces various security threats [10]. So, these several network data encryption methods rely on traditional data encryption, but traditional data encryption and decryption techniques rely on application and individual operations. It is important to go through manual operations, which inevitably leads to leakage of important data through the network by human.

The explosive growth of data volume has brought us into the era of large-scale data processing, which is characterized by high computing intensity and the need for efficient concurrent computing and storage capacity. By using the attribute encryption system, the authentication and access rights of users in the system are no longer described by a single identity or certificate, but each user has a set of attributes and a set of keys corresponding to the group of attributes. However, the widely used encryption methods have different performance, and they need to be located according to the actual encrypted data object. If this method is constantly updated, it will affect the overall algorithm execution efficiency and increase the time overhead. As an extension of the identity-based cryptosystem, attribute-based cryptosystem has attracted more and more attention because of its special application significance and a wide range of usage scenarios.

The innovation points of this paper are as follows.

(1) The adopted redundant data detection technology can transparently encrypt and decrypt the network data in the passing intermediate layer, allowing users to customize filtering encryption rules and specify encryption algorithms

(2) This paper constructs a secure and scalable mathematical model of static data attribute encryption around attribute encryption technology, combining specific data types and realistic scenarios

(3) In In-depth study of database encryption technology and extended stored procedure technology, the paper proposes a redundancy elimination algorithm for big data based on similarity calculation, which enables application developers to have "transparent access" to ciphertext data, thus greatly reducing the difficulty of application development

2. Mathematical Modeling Idea of Static Data Attribute Encryption Based on Big Data Technology

2.1. Encryption Model Classification Method. Mathematical modeling is the process of describing actual phenomena with mathematical language. The actual phenomena here include both concrete natural phenomena such as free fall and abstract phenomena. For example, the value tendency of customers to a certain commodity. The description here includes not only the description of external forms and internal mechanisms but also prediction, experiment, and interpretation of actual phenomena. Through teaching, students can understand the whole process of using mathematical theories and methods to analyze and solve problems and improve their ability to analyze and solve problems. Improve their interest in learning mathematics and their awareness and ability to apply mathematics, so that they can often think of using mathematics to solve problems in their future work. Improve their awareness of making full use of computer software and contemporary high-tech achievements and be able to organically combine mathematics and computer to solve practical problems. Encryption systems are usually divided into three categories: symmetric encryption systems, asymmetric encryption systems, and hybrid cryptosystems. Static data attribute encryption is characterized by the fact that the sender of the message only needs to encrypt the message according to the attribute and does not need to know the number and identity of the user, which reduces the cost of encryption and protects the privacy of the user. The two subjects involved in data transmission can be summarized as follows: the transmitter (sender) and the receiver of the information. Network database encryption can be broadly divided into two ways: external encryption and internal encryption, as shown in Figure 1.

When both users need to communicate, the sender encrypts the plaintext packets with an encryption algorithm to form complex unreadable encrypted data and then sends it to the recipient through the medium. If A sends a message to B, the verification process of its signature is shown in Figure 2.

The first is the symmetric encryption system, in which the key used for encryption is exactly the same as the key used for decryption. The encryption algorithm is the process of generating an unreadable ciphertext from the original readable plaintext through a series of operations and transformations with the key. For the unsigned person j, a randomly selected value x_j on Z_p generates the corresponding signed private key for the attribute i as

$$T_{i,j} = g_1^{x_j/(\gamma+t_i)}. \tag{1}$$

If any of the commands in the transaction fails to execute, then the entire transaction needs to be undone, including the commands that were executed successfully. The state of the database then reverts back to the state before the transaction was executed. When used for general encryption functions, the public key is used to encrypt the plaintext and

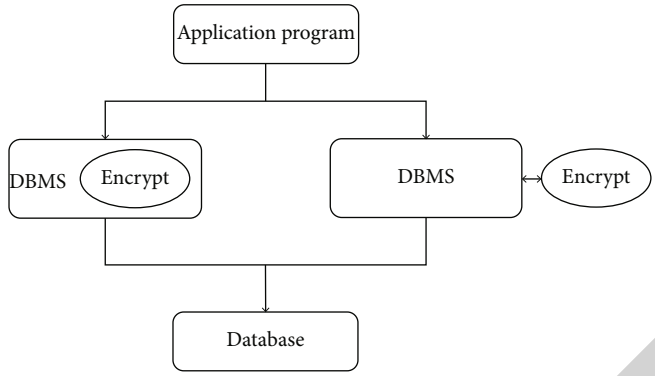

Figure 1: Database encryption mode.

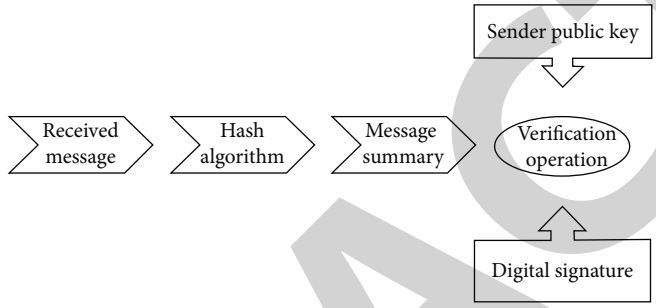

Figure 2: Verification process of signature.

the private key to decrypt the ciphertext. Authentication is the security portal of the network database system and the basis for access control. In general, the available measurement information will depend on the current moment. For example, if there is complete data from 0 to τ moment ,then the measurement information can be written as follows:

$$P_{\text{data}} = \left\{ (r, y, u) \in R \times U \times Y \left| P_r \begin{bmatrix} u - u_{\text{data}} \\ y - y_{\text{data}} \end{bmatrix} = 0 \right. \right\}, \quad (2)$$

where P_r is the time truncation operator, i.e.,

$$\mathrm{P}_\tau x(t) = \begin{cases} x(t), 0 \le t \le \tau \\ 0, \text{other} \end{cases}. \quad (3)$$

When the plaintext is encrypted with the private key and the ciphertext is decrypted with the public key, it can be used as a digital signature. The encryption step is called signature, and the decryption step is called verification. Therefore, the network database encryption system must provide a way to identify the user, and the user must provide a username, password, or other relevant security credentials, such as terminal key and user USB Key, in accordance with the system security requirements. If it is not determined behavior information system, it means that the environment has changed; otherwise, it means that the environment has not changed and has defined the environment trigger function:

$$h(s) = \begin{cases} 0, (a(x) = a(y)) \wedge (d(x) = d(y)), \\ 1, (a(x) = a(y)) \wedge (d(x) \neq d(y)). \end{cases} \quad (4)$$

$a(x)$ is the action selection of behavior x, and $\mathrm{d}(x)$ is the behavior x acceptable or unacceptable.

The next is the asymmetric encryption system, which is different from the symmetric encryption algorithm, and the asymmetric encryption algorithm needs two keys: public key and private key. At any point of time, the data in the database accessed by the user should be unique and up-to-date, and the integrity of the relational data and the consistency of the business logic cannot be destroyed by the transactions executed in the database. In addition to this, there is the escrow policy for passwords, and when the encryption algorithm is handed over to a third party for management, it must be strictly enforced in a controlled manner. The duplicate detection method is adopted to detect the relevant data that need to be encrypted, and all duplicate data among all data are deleted. That is, different records and different fields of each record in the database are encrypted with different keys. This is complemented by checksum measures to ensure the confidentiality and integrity of the database data storage and to prevent unauthorized access and modification of the data. Verification technology or password technology shall be adopted to ensure the integrity of important data during transmission, including but not limited to identification data, important business data, important audit data, important configuration data, important video data, and

important personal information. Then, in the hash table, the characteristic information of the repetitive data is stored. Let G_1 be a bilinear group of order prime p and g be its generating element, and the Lagrangian parameter is expressed as

$$\Delta_{i,s}(X) = \prod_{j \in S, j \neq i} \frac{x - j}{i - j} \tag{5}$$

SrepresentsZ_p.

Finally, there is the hybrid cryptosystem, which is a combination of symmetric and asymmetric encryption. When a transaction is successfully executed, the resultant changes to the database are permanently preserved, and even if the system crashes, the database is still able to recover to the state it was in after the transaction was successfully executed. The main point of the process is to use symmetric encryption to actually encrypt the plaintext and then encrypt the symmetric encryption key using the public key. Since the key length is shorter than the message, the defect that the processing speed of public key encryption is lower than that of symmetric encryption can be solved. In the database, not all data need to be encrypted, and only the sensitive data of users can be encrypted to improve the database access speed. Nonrepudiation means that in the process of information interaction between the two parties of communication, it must be confirmed that the information of both parties has the true unity. That is, it is impossible for all participants to deny or repudiate their true identity, as well as the originality of the information provided and the operations and commitments completed. In the attribute-based encryption mechanism, the sender controls the access policy, and the more complicated the access policy is designed, the more massive the public key is, and the more difficult it is to prove the security of the algorithm. Therefore, the network database encryption system must provide encryption setting function, so that users can set the database, table, field, etc. that need to be encrypted, which is conducive to the user's autonomy to balance between efficiency and security.

2.2. Mathematical Model Structure of Static Data Attribute Encryption. Big data refers to the technical architecture that extracts its value by capturing, computing, and analyzing a large amount, type, and source of complex data at high speed using economical methods. The database encryption system includes three functional parts: data reader, password manager, and encryption manager. Its structure is shown in Figure 3.

First is the data reader, receives the request and the data from the application, carries on the syntax analysis, and carries on the corresponding operation according to the user's request. After receiving the encrypted information, the data receiver can decrypt the encrypted information only by the same key and encryption algorithm and then read the data to make use of it. For a set of attributes, the t_i on Z_p is randomly selected as the private key of the attributes, and the public key corresponding to the attributes is published as

$$\left\{ T_1 = g^{t1}, \cdots, T_{|U|} = g^{t|U|} \right\}. \tag{6}$$

Using the full file detection algorithm, the hash table is initialized with an operation that sets a single file to granularity, detects all data, and finds duplicate data. The data must be encrypted before it is transmitted by the data transmission node and then decrypted after it is received by the data receiving node. Then, it needs to be encrypted again before it can be transmitted out, and here, it is important to note that the key used for reencryption is not for this link but for the next one. The physical NIC driver communicates directly with the transport layer driver, and the network packets are passed between them through the NDIS interface without any processing. When encryption is implemented at the application level, the encryption of data can be done by the application first, and then the encrypted data is transferred to the DBMS and stored in this encrypted form. The control law of the DBMS is

$$u(x) = k_p e(t) + k_i \int_0^t e(t)dt + k_d \frac{de(t)}{t} \tag{7}$$

or in the form of transfer function:

$$G(s) = k_p + \frac{k_i}{s} + k_d S. \tag{8}$$

k_p is the scale factor, k_i is the integral coefficient, and $k_d$is the differential coefficient.

Next is the password manager, which is the core part of the database encryption system and is responsible for completing the encryption and decryption of the data in the background, transparent to the user and the application. For securing the data, access control (i.e., authentication and authorization) has proved useful, as long as the data can be accessed using a predefined system interface. Therefore, it is required that both the decrypting party, i.e., the recipient and the sender, must have the encryption key and the encryption algorithm. Comparing the stored data in the hash table with the obtained hash function value, if an exact match is obtained, the file will be replaced by a pointer; if it is difficult to match correctly, the file can be stored. Byte substitution, row shifting, and column mixing operations are simple and reversible, and the corresponding inverse functions are used in the decryption algorithm. The round key plus step will reverse the operation on the same round key heterogeneous data group when encrypted. Under this model, the security of a cipher depends only on the security assumption on which the scheme is based, and unless that security assumption is breached, the security of the cipher can be proven to be unbreakable. If the variance of the cost function is negative, the original center of mass is replaced with a noncentral center of mass at the current location; otherwise, the center of mass remains unchanged. The fitness is suitably extended by the simulated annealing algorithm, and

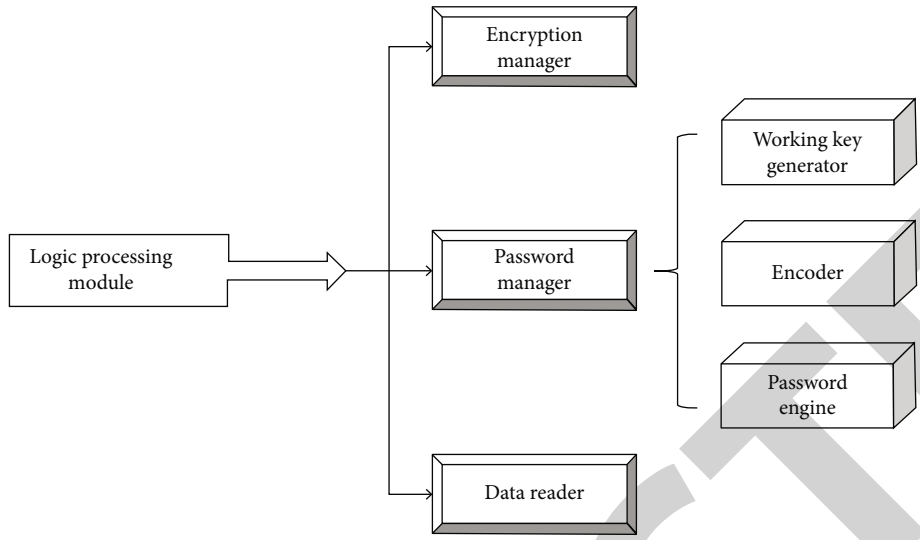

Figure 3: Structure diagram of the encryption system.

the fitness stretching is done as follows:

$$f_i = \frac{e^{f_i/T}}{\sum_{i=1}^{M} e^{f_i/T}}. \quad (9)$$

f_i is the fitness of the i individual.

Finally, the Encryption Manager is the module used by the system administrator to implement the encryption dictionary management functions and to interact with the encryption dictionary. It is responsible for the definition, modification, and deletion of the encryption dictionary. Both the user's private steel and cipher text are associated with attributes. The solution supports attribute-based threshold access policy, i.e., the user can only decrypt the cipher text when the number of attributes intersecting the set of attributes in the user's cipher text reaches the threshold value set by the system. In the case that the fingerprint matches with the matching condition, the block boundary can be set at the window, as well. For the set ω if $|\omega \cap \omega'| \geq d$, then select any d elements belonging to the intersection of the two sets and use Lagrange's difference theorem to get the following:

$$\frac{E'}{\sum_{i \in S} (e(D_i, E_i))^{\Delta i, S(0)}} = M \quad (10)$$

The provider of the data resource must use the public key of each user in the receiving group when encrypting the message. Then, the ciphertext is sent to the corresponding users separately, which leads to problems such as high processing overhead and high bandwidth consumption. Therefore, use mandatory access control, a mechanism that restricts access to an object based on the sensitivity of the information contained in the object and whether the subject has formal access authorization to the object of that sensitivity. At this point, you can first publish the public key encryption key on a reliable third-party website and let your friends obtain this public key. The receiver decrypts the digital signature with the sender's public key to obtain the decrypted information summary. And the hash algorithm used by the sender is used for the obtained message, so as to obtain the summary of the information. Compare the two information to determine whether the information has been modified. If the information has been modified, the two summaries must be different. Then, use this public key to encrypt the information and transmit it to you and then use the decryption key private key to recover the plaintext of the information for reading, in which the private key decryption key will not be transmitted in any form. When you write data, you do not need to maintain the relationship between data and data, and you do not need to fix the format of data in a table; in addition, you do not need to maintain the characteristics of ACID, which makes the performance greatly improved.

3. Analysis of Big Data Redundancy Algorithm Based on Similarity Calculation

3.1. Redundant Data Detection and Analysis. In the space of data structure, Bloom filter has the advantage of high data compression efficiency, and the feature values of this algorithm consist of Bloom filter data structure representation. Compared with the traditional data redundancy elimination algorithm, the Bloom filter algorithm has more advantages in query time and space efficiency and is more suitable for handling large data. The similarity calculation-based redundancy algorithm and CBC algorithm are shown in Figures 4 and 5 below to compare the time consumption of each encryption algorithm when the key length is 128 bit and 256 bit.

Firstly, the data in a data set with two or more duplicates are deleted to ensure that only the same data is retained in the final data set, so that the deleted redundant data are replaced by data pointers. The database files as a whole are encrypted with encryption keys and encryption algorithms to ensure the authenticity and integrity of all user data tables,

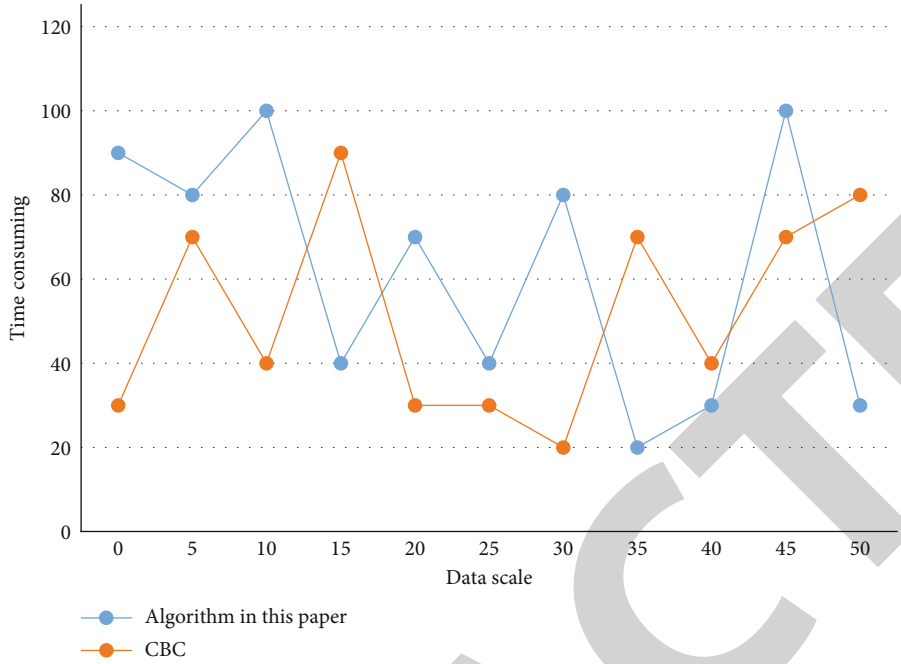

Figure 4: Comparison of time consumption of each algorithm when the key length is 128 bit.

system data tables, indexes, views, and stored procedures in the entire database. Data sharing is achieved by decrypting the entire database file with the decryption key. The information sent to others is encrypted with e, and as long as others can decrypt it with d, the information is proved to be sent by you, which constitutes a signature mechanism. It can be encrypted by AES advanced encryption standard using deduplication technology first. Since CTR mode encryption and decryption is fast and both the key and group length can be varied, it can be used as the working method of the group cipher algorithm. Moreover, with such an encryption method, the encryption process in database applications is transparent and can be used directly. Proxy reencryption is very suitable for use in scenarios where encrypted data is to be shared with multiple parties. There is no need to share private keys with recipients, and there is no need to encrypt the entire message for each recipient before proxy. This encryption allows the user to encrypt only once and then authorize the recipient according to his public key. In the proxy re encryption scheme, there is a semi trusted proxy with encryption. It is able to convert the ciphertext encrypted with the public language to the ciphertext encrypted with the public key, but the proxy does not obtain any information about the plaintext during the whole process. To achieve strong backward security requires updating the keys of the underlying encryption. A comparative analysis of the additional communication burden and computational burden required when the underlying encryption is stream cipher, group cipher, and public key cipher, respectively. The comparison of the extra burden required to update the underlying key for different underlying encryption schemes is shown in Table 1 below.

Secondly, the ratio of the number of bytes before redundant data deletion to the number of bytes processed is mainly used in determining the reduction rate of data to achieve. The group length and key length are both variable, and only to meet the requirements are the processed group size limited to 128 bits, while the key length is 128 bits or 256 bits various options. With table-level encryption granularity, the query performance of the system will be improved. Because the system performance will not be affected for the queries of unencrypted tables, and for the queries of encrypted tables, only the corresponding encrypted tables need to be decrypted instead of decrypting the whole database. Before the data reaches the target endpoint, the data has to be transmitted through a large number of link nodes. The serialization module completes the interconversion between byte stream and ABE data types, and it writes the structure variables output by other modules to byte arrays. When other modules need a specific type of variable, it reads the data from the byte stream to generate the variable of the target type. This encryption definition tool mainly defines the way to encrypt the data of each table in the database. After creating a certain database table, this tool defines the encryption method for the characteristics of that table.

Finally, all the possible attribute values are sorted according to some linear rule, and the corresponding statistics are then used to obtain a statistical histogram based on the attribute values. The entire record in the data table is encrypted and stored in the database file as a cipher text. The data block undergoes several data transformation operations, and each transformation operation produces an intermediate result, which is called state. Since the encryption and decryption keys are usually stored in the database, it also raises the risk of key management, which leads to the protection of the keys only through the access control of DBMS. The ciphertext data is decrypted, then encrypted using a different key, and then in transmission. In general,

Table 1: Comparison of the extra burden of different underlying encryption schemes when updating the underlying key.

Encryption mode	Communication burden	Client computing burden	Client computing burden
Stream cipher	17.26	$O(n)$	$O(1)$
DES	23.76	$O(n)$	$O(1)$
AES	31.74	$O(n)$	$O(1)$

the security of a cryptosystem can be divided into two forms: selective plaintext attack and selective ciphertext attack, depending on the attacker's target. In a cryptosystem, we can prove that the cryptosystem is secure as long as the probability of a successful attack by an attacker is calculated to be negligible. The linear hybrid layer ensures a high degree of diffusion over multiple rounds, and the nonlinear layer consists of 20 S boxes juxtaposed to act as an obfuscation and the key encryption layer heterogeneous subkeys to intermediate states. Since the polynomial running time is feasible according to the Turing machine model, the probability is considered to be nonnegligible when it is the inverse of the polynomial.

Complexity of the cryptographic algorithm is as follows: the complexity of the cryptographic algorithm is a basic condition to ensure password security. If the cryptographic algorithm used in a cryptographic system is not complex, or it seems to be complex, but there are institutional weaknesses, it is easy to be exploited by attackers. In addition to the complexity of the cryptographic algorithm, key length is also the basic factor to ensure the security of cryptographic system. The simplest way to crack the key is to try all kinds of possible keys to see which one is actually used. In this attack, the number of passwords to be attempted is closely related to the entire key space to be retrieved.

3.2. Analysis of Big Data Encryption Algorithm. ECC, known as "ellipse curve ctyptography", is a public key encryption algorithm based on elliptic curve mathematics. The use of elliptic curves in cryptography is proposed independently. Different from the traditional encryption algorithm based on the problem of large prime number decomposition, this encryption method is based on the mathematical problem of "discrete logarithm." The ECC algorithm is used to achieve encryption of big data, which has the advantages of low computational overhead and high encryption security performance in the encryption process, and is more suitable for research areas where computational power, space constraints, and bandwidth are limited. In describing the security proof, the simulator needs to simulate all realistic environments, including the use of realistic hash functions. It makes it possible to deceive the attacker to perform attack operations on the simulated environment. If the data is decrypted and then encrypted in the cloud, there is a risk of privacy leakage. The experimental analysis of the encryption algorithm and decryption algorithm on a laptop, encryption algorithm and decryption algorithm with number of attributes and number of keys, respectively, are shown in Figures 6 and 7 below.

First, when the receiver decrypts the encrypted data file, it needs to calculate the private key and the public key for decryption. The sender first encrypts the message using a symmetric cipher, then encrypts the key using the receiver's public key, and then sends both to the receiver. When the user's security level flag is higher than or equal to the security level flag of the data to be accessed, the data in the data table is decrypted using the three-level key, and the plaintext data is returned to the user. We can also use the same scheme to prove that if the attacker can recover the encrypted random number, then it can recover the information. The stream encryption algorithm theoretically encrypts increments of bits, but in reality, it usually encrypts every bit of the plaintext itself at a time. When decrypting the ciphertext, the user can only decrypt the message correctly if the set of attributes used for encryption matches the access control structure in the user's private key, thus allowing the user to access the message content. To access the data in HDFS, the user needs to request a block access token from the Name Node, and only with the correct token can the user access the corresponding data block in HDFS, in order to test the relationship between mixed encryption time, attribute encryption time, and file size. In the experiment, the length of symmetric encryption secret key is fixed to 128 bit, and the encryption time varies with the file size as shown in Table 2 below.

Secondly, in the process of detection calculation, this paper will take the data file that needs to be encrypted individually as the granularity and detect the duplicate data in the data file. When an attribute revocation event occurs in the system, the system must ensure that the original private key of the user who owns the attribute is invalid; that is, the private key cannot decrypt the original ciphertext again. In CP-ABE, each user is divided into a group of describable attributes, and the unified attribute authorization center distributes the attribute private key to each user according to the attribute set of each user. The receiver decrypts the symmetric key using its private key and then decrypts the message using the key. For the user to write data to the database, when the security level flag of the user is lower than or equal to the security level flag to be written, a random key is generated and used as the tertiary working key to encrypt the data to be written to the database, the secondary key is used to encrypt the tertiary key, and the encrypted key is written to the database. In the initialization process, the main function of the key is to churn the S-box. Therefore, the application goal of the system is to ensure that it can support specific types of queries on ciphertext data and at the same time can protect the privacy of user data. The runtime of the Decrypt algorithm for IAI-CP-ABE1 and IAI-CP-ABE2 on a resource-constrained physical device was simulated in the Intel Edison development board, Outsourced Decrypt algorithm, and a comparison of the algorithmic efficiency of the Decrypt algorithm, and the big data encryption algorithm is shown in Figure 8 below.

Finally, the data files that do not duplicate each other in the complete file detection method are rearchived, and the

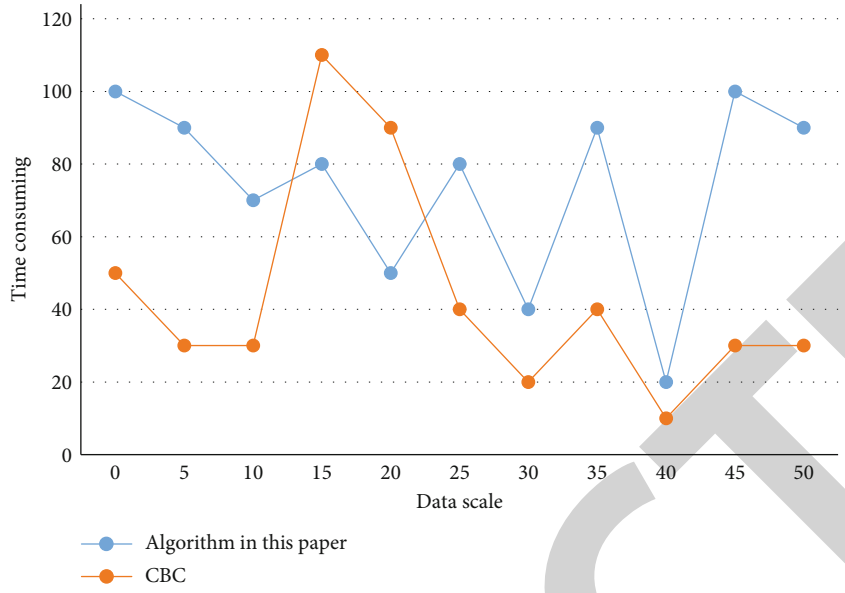

Figure 5: Comparison of time consumption of each algorithm when the key length is 256 bit.

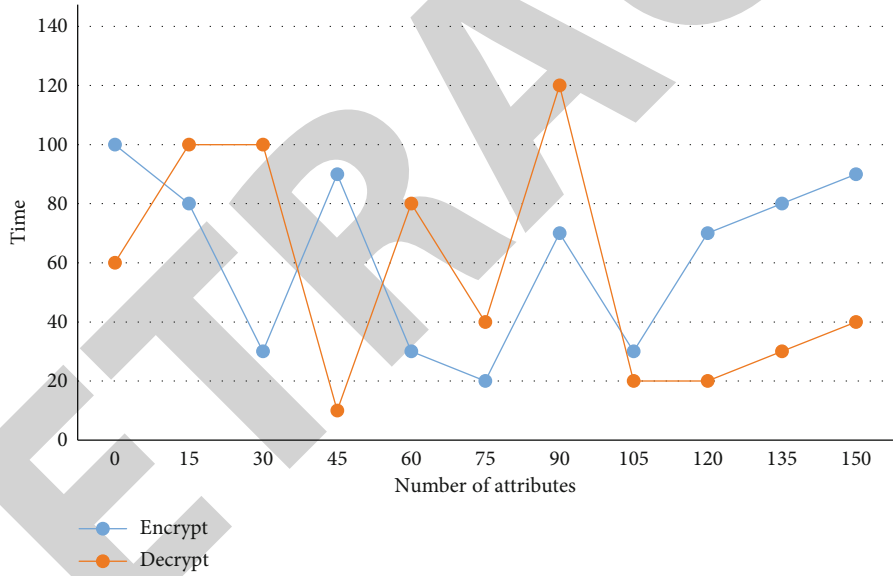

Figure 6: Relationship between Encrypt/Decrypt algorithm and the number of attributes.

Table 2: Experimental data of encryption time varying with file size.

File size (MB)	5	10	15	20	25	30
Encryption time (s)	17	28	37	42	58	63
Attribute encryption 128bit key time (s)			65			

process of this paper will use the CDC data block calculation method to archive them one by one from the file source. When the user accesses the data, the database system in the background must decrypt the data before returning it to the user. In order to respond to the user's request as soon as possible and reduce the user's waiting time, the speed of data encryption and decryption must be fast. The data owner sets the access control structure and uses it to encrypt the information and then share it in the network, and other users can decrypt the information when and only when the set of attributes in their attribute private key matches the access control structure in the cipher text. In the process of cryptosystem security proof, we always end up with a hypothetical hard problem, such as BDH and CDH problem. If field-level, record-level, or data-item-level encryption granularity is used, the key is used to encrypt a three-level key. Due to the existence of partial weak keys, it makes the subkey sequence to be completely duplicated in less than 10,000 bytes, and if it is partially duplicated, it may be able to be duplicated in less than 10,000 bytes. Therefore, the

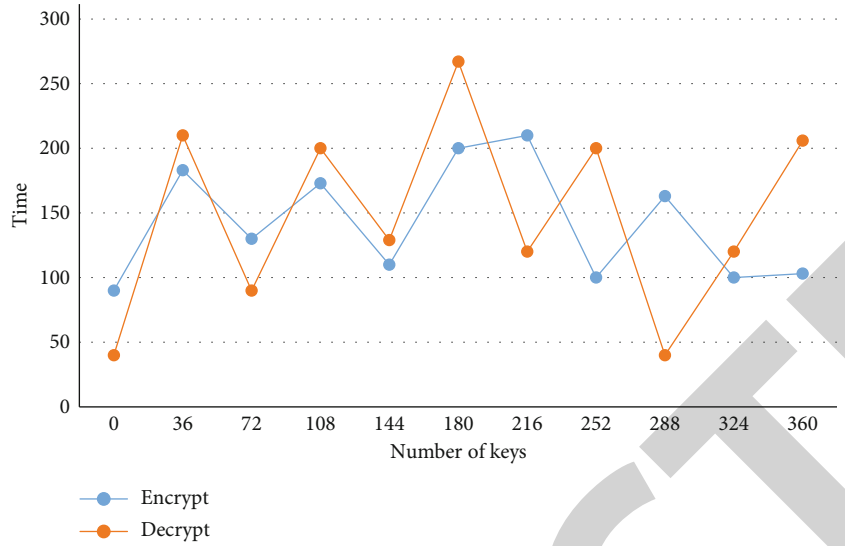

Figure 7: Relationship between Encrypt/Decrypt algorithm and the number of keys.

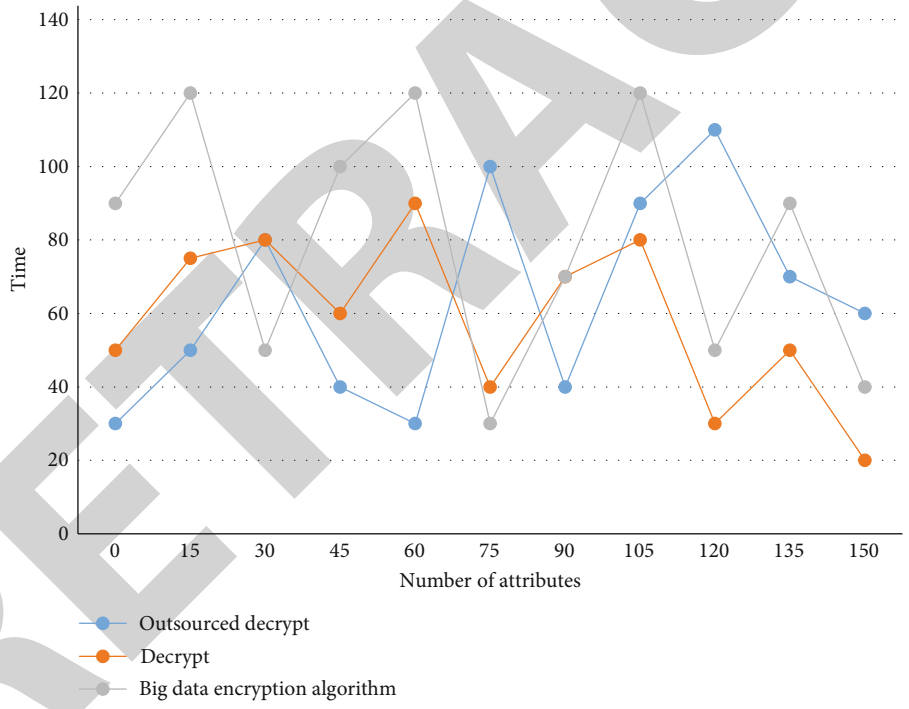

Figure 8: Comparison of algorithm efficiency between Outsourced Decrypt algorithm, Decrypt algorithm, and big data encryption algorithm.

encryption key must be tested to determine whether it is a weak key.

4. Conclusions

Network database is a new manifestation of database in the network era, which is a strong aggregate of network and database, and its application is very wide and has penetrated into various fields. Due to the huge amount of data in the big data environment, the traditional encryption scheme is inefficient in the big data application, while the traditional access control mode is also difficult to apply to the big data environment. The degree of information security in the big data environment directly affects the privacy protection of users, and the superior features provided by encryption algorithms can compensate for the drawbacks caused by relying solely on software security prevention strategies, thus overcoming the difficulties and challenges faced by information security in a more efficient and stable manner. In this paper, we propose a mathematical modeling idea of static data attribute

encryption based on big data technology and then analyze the big data redundancy elimination algorithm based on similarity calculation. By using static data attribute encryption based on big data technology, the length of encrypted data packets will not be increased, and the part of redundancy of fragmentation can be eliminated, which can greatly improve the efficiency of the system. The attribute-based encryption mechanism uses attributes as public keys, and the decryption users are a group; so, the encryption efficiency is very high. At the same time, attribute-based encryption can represent the access policy flexibly, which provides a good tool for access control of big data security applications. Therefore, attribute-based encryption has a good application prospect in big data security. However, there are still some deficiencies in the research. Attribute encryption applies a large number of bilinear mappings, and the calculation of bilinear pairs is time-consuming. This has become a difficult problem in the practical application of attribute encryption. I believe that with the continuous improvement of computer hardware, this problem will be solved.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the First-Class Undergraduate Course Programme of Henan Province (No. [2020]13136).

References

[1] Z. Yumei, "Mathematical modeling of big data attribute encryption based on data redundancy technology," *Computer Simulation*, vol. 38, no. 5, 2021.

[2] W. Songjing and X. Feng, "Research on "mathematical modeling" teaching reform driven by realistic data," *Journal of Ningbo University: Education Science Edition*, vol. 39, no. 5, 2017.

[3] Z. Lingfeng, "Research on new data encryption algorithm in big data environment," *Science and Technology Bulletin*, vol. 33, no. 6, 2017.

[4] D. Yujiao, "Research on data encryption method of big data platform," *Computer Knowledge and Technology: Academic Edition*, vol. 2X, 2017.

[5] G. Pengpeng, "Research on computer network information data encryption technology in big data environment," *Journal of Suzhou Institute of Education*, vol. 21, no. 3, 2018.

[6] A. Chehri, I. Fofana, X. Yang et al., "Security risk modeling in smart grid critical infrastructures in the era of big data and artificial intelligence," *Sustainability*, vol. 13, no. 6, p. 3196, 2021.

[7] M. Xiangzhi, "Discussion on information and communication data encryption technology under big data," *Communication World*, vol. 26, no. 10, 2019.

[8] Y. Zhang, "Application of information and communication data encryption technology in the era of big data," *Computer Paradise*, vol. 6, 2019.

[9] G. Baoping and M. Jianhong, "A secure storage method for unstructured big data based on revocable attribute encryption combined with fast density clustering algorithm," *Computer Applications and Software*, vol. 38, no. 5, p. 7, 2021.

[10] Y. Zuobin, M. Jianfeng, and C. Jiangtao, "Attribute encryption scheme supporting location verification and policy change," *Journal of Xidian University*, vol. 44, no. 2, 2017.