Dynamic encryption method of user privacy data based on association rules hiding

Zongrong Li1*, Nanfang Li1, Xuebin Yuan1, Lei Zhao1, Xuezhi Ma1 and Shengchun Li 1

1Electric Power Science Research Institute, State Grid Qinghai Electric Power Company, Xining, Qinghai, 810008, China

*E-mail: lzrong007x@qh.sgcc.com.cn

Abstract: Because the traditional data encryption method remains unchanged after the key is generated, the encryption scheme is single, the encryption cost is high, and it is easy to be cracked. Aiming at the above problems, the dynamic encryption method of user privacy data based on association rules hiding is studied. Based on the DGHV encryption scheme, a dynamic encryption scheme is designed to hide the mined user privacy data association rules. After the key is processed by confounding method, the dynamic encryption process of data is completed. By comparing with the traditional methods, it is verified that the encryption efficiency of the studied method is increased by at least 26.2%, which has the advantage of low encryption cost.

1. Introduction

User privacy information contains a large amount of personal information of network users, and once personal information is cracked and disclosed, it will not only cause trouble to personal work, but also cause serious interference to social stability and economic development due to serious personal privacy information data leakage [1]. Therefore, encrypting user privacy data becomes the main way to protect user data security. Tian et al. [2] proposed an anonymous privacy protection method based on dynamic data mining. According to the cipher scheme, the ciphertext of private data is solved and sent to the third party of data mining. Finally, the protection of anonymous privacy is realized. However, the data encryption method of this method is relatively simple and the encryption cost is increased, which can not effectively protect the security of user's private data.

Hiding the association rules in the user privacy data can further improve the encryption efficiency of the encryption method. According to the above analysis content, to protect the security of user privacy data in the network, this paper will study the dynamic encryption method of user privacy data hidden based on association rules, and verify the validity of the encryption method studied.

2. Dynamic encryption method of user privacy data based on association rules hiding

2.1. Mining association rules for private data

The increasing number of network users has caused great problems to protect the security of user privacy data, and there are a large number of association rules between user privacy data. Foreign attackers can collect all user privacy data by getting association rules in the data. Therefore, it is necessary to mine and hide association rules in user privacy data before encrypting them.
Set the user privacy data set as \( I = \{i_1, i_2, \cdots, i_m\} \), and any transaction \( T \) in the transaction database \( D \) corresponding to the data set is a set of user data items in the user privacy data set. For an item set \( X \) in the user privacy data set, the form of association rule between the item set and the user privacy data set is \( X \rightarrow Y, X \subseteq I, Y \subseteq I, X \cap Y = \emptyset \). This paper uses FP-tree frequency set algorithm to mine association rules of user privacy data \(^3\). Firstly, the user privacy database is scanned to record the frequent transaction items in the privacy data set and their corresponding support degree. The support calculation of frequent transaction items is shown in formula (1).

\[
\text{sup}(X) = \frac{\text{count}(X)}{|D|}
\]

In formula (1), \( \text{count}(X) \) is the number of transactions containing frequent transaction item \( X \) in the transaction database corresponding to the user privacy data set; \( |D| \) is the total number of frequent transaction items in the transaction database. The FP-Tree root node is created based on the frequent items scanned for the first time \(^4\). In FP-Tree, traversal is performed according to the path, and the minimum path whose support value is the sum of all nodes in the current path is obtained \(^5\). In order to protect the security of user privacy data, association rules hiding algorithm based on fuzzy logic is adopted to deal with user privacy data association rules.

2.2. User privacy data association rules are hidden

In this paper, association rules hiding algorithm based on fuzzy logic is used to hide association rules in user privacy data. The algorithm is divided into two stages. In the first stage, the association rule hiding algorithm traverses each transaction in the user privacy data set and collects the association rule information in the user privacy data set \(^6\). Each node is used to represent item \( i_k \) in transaction \( x_i \) in the form of \( \{R_k, |R_k|\} \), \( R_k = \{j \mid \text{SAR}_j \subseteq x_i, i_k \subseteq \text{SAR}_j\} \). The side effects of the hidden algorithm can be estimated by calculating the prior weight of transaction \( x_i \), and the prior weight of the transaction can be calculated according to formula (2).

\[
\begin{aligned}
    w_i &= \frac{\text{MIC}_i}{2^{|R_i| - 1}} \\
    \text{MIC}_i &= \max\left(|R_i|\right)
\end{aligned}
\]

In formula (2), \( w_i \) is the prior weight of transactions; \( \text{MIC}_i \) is the number of association rules in the association rule set. The order of modifying the transaction items in the user privacy data set is the priority weight of each transaction item, and the transaction items with a higher priority weight have priority to be modified \(^7\). Repeat the following steps until the association rule set SAR for user privacy data is empty \(^8\). Select a transaction from the storage table that stores transaction item weights of user’s ID and user privacy data to maximize the prior weight corresponding to the transaction; Select the transaction items to be fuzzed. For the given user privacy data association rules set, all the association rules in the rule set are hidden; Modify the privacy data association rule set. Repeat the above steps until the association rule set SAR for the user private data is empty. After hiding all data association rules based on fuzzy logic, the dynamic encryption scheme of user privacy data is designed.

2.3. Design data dynamic encryption scheme

This paper will improve the efficiency of dynamic data encryption and reduce the cost of encryption by means of public key compression. At the same time, the private key corresponding to the independent user factor data encryption is generated, and the digital key and public key are generated.
in the random number generator for xor operation to obtain the new public key after reconstruction\(^9\). According to the formula (3), the whole homomorphic encryption startup system is constructed.

\[
\begin{align*}
\{ (sk, pk) \leftarrow & \text{KeyGen}_e(\lambda) \\
\psi_i \leftarrow & \text{Encrypt}_e(pk, \pi_i) \\
& \text{Decrypt}_e(sk, \psi_i)
\end{align*}
\] (3)

In formula (3), \(\text{KeyGen}_e\) represents the encrypted public key \(pk\) and private key \(sk\) generated by the above process; \(\lambda\) is a safety parameter; \(\text{Encrypt}_e\) means to use encrypted public key and encrypted plaintext \(\pi_i\) to calculate the mapping of plaintext in ciphertext space under the random algorithm; \(\text{Decrypt}_e\) is the process of decrypting plaintext using private key and ciphertext \(\psi_i\) to decrypt output. In order to realize dynamic encryption of homomorphic data, an error-correcting code is constructed on the ring by using FHE construction method\(^10\). The design of data dynamic encryption scheme provides the basis for dynamic encryption of user privacy data.

In this paper, the key update strategy of dynamic confounding is adopted to encrypt user privacy data, and the randomly generated Numbers are inserted into the pseudo-position of ciphertext after preliminary processing according to bytes. This promiscuous encryption processing mode can prevent the attacker from obtaining both the ciphertext and the key to correctly crack the ciphertext.

3. Simulation experiment

This paper studies the dynamic encryption method of user privacy data based on association rule hiding. This section will design an experimental verification process to verify the effectiveness and feasibility of the encryption method.

3.1. Experiment content

This simulation experiment uses the dynamic encryption method based on association rules hiding as the experimental group, and the method in reference \(^2\) as the control group. Mainly from two aspects of security and encryption efficiency and compares the traditional way of user privacy data encryption, through contrast experiment on the previously proposed based on association rules hidden user privacy data dynamic encryption method is verified and analyzed. The running program of this simulation experiment takes visual studio 2014 as the development environment and uses C++ programming language to efficiently implement the experimental content. Process and analyze the experimental data to achieve the experimental preset verification objectives.

3.2. Experimental process

10 data sets with different data volumes were selected as the encryption processing objects of this experiment, and the data of one or several attributes in the experimental data set were divided and marked as private data. Respectively using two methods of privacy data encryption encryption processing of experimental data set in direct contrast two encryption method for experimental data set encryption effect, on the basis of comparing the two encryption method is time-consuming and encryption method of encryption tag number, to compare the encryption method of encryption and encryption efficiency overhead. MATLAB software was written to process and analyze the collected experimental data, and the corresponding conclusions of the two experimental indicator variables were combined to judge the performance of the two encryption methods.

3.3. Experimental results and analysis

After encryption method encrypts a data set. The results of time consuming comparison of the two encryption methods for encrypting data of different numbers of experiments are shown in the table 1.
Table 1 Encryption method encryption time

| Data set number | Encrypted data volume /GB | Experimental group method /s | Comparison group method /s |
|----------------|---------------------------|------------------------------|---------------------------|
| 1              | 11.4                      | 12.8                         | 26.3                      |
| 2              | 13.6                      | 14.6                         | 28.1                      |
| 3              | 10.7                      | 12.1                         | 27.9                      |
| 4              | 14.0                      | 14.2                         | 24.7                      |
| 5              | 21.2                      | 15.9                         | 25.3                      |
| 6              | 16.5                      | 13.7                         | 29.1                      |
| 7              | 17.3                      | 14.5                         | 28.8                      |
| 8              | 17.5                      | 15.9                         | 29.7                      |
| 9              | 16.8                      | 15.3                         | 30.3                      |
| 10             | 19.2                      | 14.4                         | 33.5                      |

Analysis of the data in table 1 shows that the encryption time of the comparison group method is greater than that of the experimental group method. The minimum encryption time difference between the two encryption methods was 8.8s, that is, the encryption efficiency of the experimental method was improved by at least 26.2%. Combined with the encryption effect analysis in figure 1, the ciphertext of the experimental group after encryption were all chaotic codes with no plaintext. However, some plaintext still appears in ciphertext after the contrast method is encrypted, which will leave a loophole for attack and intrusion, and the encryption effect is poor.

The encryption cost of the encryption method can be characterized by the relationship between the number of encryption tags, the amount of encrypted data and the encryption time. The figure 1 compares the encryption costs of two encryption methods when encrypting different amounts of data.

![Figure 1 Comparison of encryption overhead of the two encryption methods](image)

As can be seen from the figure 1, the interval of encrypted data is the same. With the increase of the number of encrypted tags, the encryption time of the experimental group first increases and then tends to be static, while the encryption time of the contrast group keeps increasing. All the above three aspects indicate that the experimental group has lower encryption cost when encrypting data. To sum up, compared with traditional encryption methods, the dynamic encryption method of user privacy data based on hidden association rules studied in this paper has the advantages of high encryption efficiency and low encryption cost.

4. Conclusion

Encryption of user privacy data has always been the focus of network security maintenance, and with the continuous upgrade of network attack technology, the challenges to user privacy data security protection are also increasing. Aiming at the problems existing in the traditional encryption methods
of user privacy data, this paper studies the dynamic encryption method of user privacy data based on association rules hiding. By comparing with the traditional data encryption methods, it is verified that the dynamic encryption method studied in this paper can effectively reduce the encryption cost, improve the encryption efficiency, and achieve better encryption processing effect. In the future research, we should start from the direction of encrypting data in different forms and formats, so as to study encryption methods with a wider range of application.

Reference

[1] HAN Q. (2020) Double Encryption Method of Network Privacy Information Based on Dynamic Key Selection [J]. Journal of Heilongjiang University of Technology (Comprehensive Edition), 20(03): 89-93.

[2] TIAN CR, LI ZX, LUO YX. (2019) Simulation of Anonymous Privacy Protection Method Based on Dynamic Data Mining [J]. Computer Simulation, 36(11): 171-174+233.

[3] WANG X. (2019) Computer Network Data Security Encryption Technology in the Internet Environment [J]. Mobile Communications, 43(03): 49-53.

[4] TONG W, HUANG QP, WANG Tangzhi. Research on Location Big Data Encryption Method Based on Privacy Protection [J]. Journal of Anhui Electrical Engineering Professional Technique College, 24(01): 118-122.

[5] LI YC, ZHANG P, ZHENG SQ. (2019) Privacy Protection of Power Consumption Big Data Based on Empirical Mode Decomposition and Homomorphic Encryption [J]. Power System Technology, 43(05): 1810-1818.

[6] GUI DW. (2018) Research on Privacy Protection Data Encryption Algorithm for Big Data in Mobile Cloud Computing [J]. Bulletin of Science and Technology, 34(08): 211-214.

[7] ZHOU YF. (2018) An Improved Data Encryption Algorithm Designed for Cloud Storage Platform [J]. Journal of Inner Mongolia Normal University (Natural Science Edition), 47(03): 237-240.

[8] WANG B, ZHANG L, ZHANG GY. (2018) Research on location information hiding method based on improved genetic algorithm [J]. China Sciencepaper, 13(08): 902-906.

[9] CHEN XY, LING BH, HU DH. (2018) Privacy Protection of Health Data for Network Based on Attribute-based Encryption [J]. Journal of Anqing normal university (Natural Science Edition), 24(01): 56-59.

[10] CAO LC, LIU YF, DONG XY, et al. (2018) User privacy-preserving cloud storage scheme on CP-ABE [J]. Journal of Tsinghua University (Science and Technology), 58(02): 150-156.