Computer Communications 163 (2020) 84–90

Contents lists available at ScienceDirect

Computer Communications

journal homepage: www.elsevier.com/locate/comcom

Application of data mining technology in alarm analysis of communication network

Qun Zheng [a,*], Yaofeng Li [b], Jie Cao [a]

[a] North East Electric Power University, School of Computer Science, Jilin 132000, Jilin, China
[b] Power Supply Company of Jilin, Electric Power Dispatching Control Center, Jilin 132000, Jilin, China

ARTICLE INFO

Keywords:
Communication network alarm
Data mining technology
Fuzzy association rule mining
Dynamic time window

ABSTRACT

Nowadays, with the continuous development of science and technology, the social demand for the network is more and more, and with the continuous expansion of the network scale, the complexity of the network is increasing day by day. How to locate the fault accurately and quickly from a large number of alarm information has become a problem. To solve this problem, we must use computer technology to study the automatic analysis technology of alarm correlation. Therefore, this paper applies data mining technology to communication network alarm analysis, and based on fuzzy theory, it can accurately describe the relationship between network alarm and fault reason, and combines data mining technology with fuzzy theory to form the network alarm correlation analysis method of fuzzy association rule mining. On this basis, in order to mine the fuzzy association rules of the fuzzy alarm database directly and avoid the interference in the process of converting the alarm database to the transaction database, this paper proposes a dynamic time window fuzzy association rule mining algorithm. Through the simulation analysis, compared with the traditional data mining technology, the fuzzy association rules mining method combined with the fuzzy theory has better performance, and the further proposed dynamic time window fuzzy association rules mining algorithm greatly improves the accuracy of the association rules confidence, and has good application performance in the network alarm correlation analysis.

1. Introduction

With the continuous development of science and technology, communication technology is constantly updated, and the competition of communication network market is more and more fierce. The traditional network operation and maintenance management mode is that the monitoring team composed of various specialties is responsible for the network management of the specialty, but in today's fiercely competitive communication network market, the mode has been unable to meet the market demand. The traditional network operation and maintenance management mode requires multi-disciplinary, multi-vendor and multi-platform decentralized management, which wastes human resources, increases the cost of network operation and maintenance, and lacks the ability of centralized processing of alarm data, which easily leads to the existence of missed alarm and false alarm. Therefore, it is imperative to unify the traditional decentralized network operation and maintenance management mode and realize the integrated application of the existing heterogeneous software and hardware systems, which is conducive to improving the operation and maintenance efficiency of the communication network and reducing the operation and maintenance cost. However, the centralized operation and maintenance mode of the communication network will lead to a large number of derivative alarm information generated simultaneously by multiple devices and links in the network when a component fails due to the centralized processing of all information, which brings a new challenge to the network monitoring center to quickly and accurately find and locate the fault. Therefore, in the face of new challenges, it is very necessary to use computer technology to study the correlation analysis mechanism of automatic positioning alarm fault, generate alarm tree through alarm correlation analysis, quickly and accurately analyze root alarm and derivative alarm, so as to locate the fault for maintenance personnel to repair the fault and ensure the stable operation of communication network system.

The correlation analysis of alarm data, in fact, is to shield the derivative alarm information irrelevant to root fault through correlation preprocessing, reduce the interference of these derivative alarm information, so as to quickly and accurately find the fault of communication network. In other words, the essence of alarm data correlation analysis is to select and filter the alarm information and then carry out pattern matching to realize fault identification and location. Data mining technology [1–3] is a scientific data analysis method, which is very suitable for the correlation analysis of alarm information in

* Corresponding author.
E-mail addresses: zhengqun@neepu.edu.cn (Q. Zheng), 45780338@qq.com (Y. Li), caojiell78@126.com (J. Cao).

https://doi.org/10.1016/j.comcom.2020.08.012
Received 24 April 2020; Received in revised form 24 July 2020; Accepted 17 August 2020
Available online 19 August 2020
0140-3664/

communication network. It can analyze the historical records of alarm and automatically find new alarm correlation rules, which is conducive to the rapid location of network alarm fault, saves the process of troubleshooting for network management staff, and improves the efficiency of network management staff in dealing with faults. After in-depth research by researchers, now most of them use data mining technology to process alarm information and mining association rules of alarm information [4,5], which effectively improves the efficiency of traditional troubleshooting methods. Apriori algorithm [6,7] was first proposed in association rule mining algorithm. Its disadvantages are that it will produce a large number of candidates sets and need to scan the database repeatedly. Therefore, the improvement of association rule mining algorithm is mainly from three aspects: reducing candidate set, compressing the general office of affairs and reducing the number of database scans. Some improved association rule mining methods such as Frequent Pattern tree (FP Tree) can reduce candidate set, total number of transactions and the number of database scans at the same time. Up to now, the research of association rules can be divided into two kinds: generating frequent set candidates and not generating frequent set candidates.

Although the research and application of association rule mining algorithm has developed rapidly, there are still some problems that need to be solved urgently, and the efficiency of association rule mining and the accuracy of troubleshooting are still lacking. This is mainly reflected in the fact that the traditional association rule mining algorithm only deals with the fault alarm information, but not the network fault directly. Although alarm information has a direct relationship with network fault, their essence is not the same, and the processing of alarm information will inevitably have a certain impact on network fault diagnosis. Specifically speaking, the alarm information is caused by the failure of a part of the communication network, which causes the equipment or link associated with it to send abnormal working signals, that is to say, a large number of alarm information may be caused by a failure, so the alarm information and failure are not a one object relationship, but a one to many fuzzy relationships. In many alarm information, what we need is not derivative alarm information, but to find the root information of the alarm or how much the alarm information solves the root alarm. So, we need a theory that can describe the alarm information intuitively. In the traditional alarm correlation analysis, Boolean logic is used to process quantitative alarm information, which is directly located in a certain language range, which makes the description of alarm information direct and hard. However, the fuzzy relationship between alarm information and fault performance is not definite, which leads to the deviation of network fault diagnosis caused by this kind of rough description, thus the accuracy of network fault location is affected. For this reason, fuzzy theory, which is used to solve the problem that traditional mathematics cannot accurately describe in scientific research, is used to solve this problem in this paper.

Fuzzy mathematics is an important branch of mathematical theory. Based on the fuzzy set theory [8,9], many theories, such as fuzzy association rules, fuzzy evaluation [10,11], are put forward. Combining data mining with fuzzy theory, the fuzzy association rule mining is different from the traditional association rule mining in that every alarm data absolutely belongs to a set of hard partition. Instead, the attribute is divided into fuzzy sets according to expert knowledge, and each alarm information is divided into several fuzzy sets. In this way, the alarm information does not belong to a fuzzy set, but may belong to each fuzzy set. The possibility of each alarm data belonging to each fuzzy set is calculated according to the set membership function. In this way, fuzzy association rules mining can avoid the deviation of hard partition of traditional Boolean logic association rules mining, and improve the accuracy of network fault diagnosis. Based on this, this paper introduces fuzzy theory on the basis of data mining technology, combines data mining technology and fuzzy theory, and applies the algorithm of mining fuzzy association rules to alarm correlation analysis, which improves the accuracy of network fault diagnosis. And on this basis, this paper further proposes a dynamic time window fuzzy association rules mining algorithm, which can directly mine the fuzzy association rules in the fuzzy alarm database, and improve the reliability of the association rules, thus further improving the accuracy of network fault diagnosis.

2. Theory of mining Fuzzy association rules

2.1. Data mining technology

Data mining is a combination of database technology [12], artificial intelligence [13], machine learning [14] and statistics [15]. The simple interpretation of data mining is to “mine” the knowledge needed in various fields from a large amount of data. Due to the wide application of data mining, different experts have different definitions. But generally speaking, the definition of data mining can be expressed as "data mining is a complex process of mining hidden or unknown, user interested and potentially valuable knowledge and rules for decision-making from a large number of data sets (which may be incomplete, noisy, uncertain, and various storage forms)". The mining rules represent the specific relationship between some objects in the database, show the useful information of these objects for decision-making, and can provide a strong basis for multiple decisions or planning.

The emergence of data mining lies in that people want to make full use of the stored data to provide decision support, and keep a good development prospect in the more than ten years since its birth. Data mining, which is used in many fields, usually can find four kinds of knowledge types: generalized knowledge that abstractly describes the category features, related knowledge that reflects the dependency between events, classification knowledge that reflects the common features of the same things and the differences between different things, and predictive knowledge that predicts future data through historical and current data. There are many kinds of data mining technologies. There are corresponding solutions to different problems, and different effects have their own unique advantages. Common technologies include clustering [16], association, neural network [17], decision tree [18], text analysis, web mining [19], etc. Nowadays, association and clustering technology is an important application technology in data mining.

Association rule mining was first studied by R. Agrawal and Jiawei Han [20], the main purpose is to get a set of all the rules that meet the user specified minsup and minconf, with a wide range of application scenarios, such as cross marketing. Its main basic definitions are as follows:

(1) Transaction database: in the correlation analysis of alarm data of communication network, transaction database represents a collection of all alarm information in a certain period of time.

(2) Item set support: the support of $I = \{i_1, i_2, \dots, i_n\}$ indicates the proportion of transactions including I in the total number of transactions. Expressed as:

$$\mathrm{support}(I) = \mathrm{support\_}count(I_i) / |D| \tag{1}$$

(3) Frequent item set: set the minimum support degree minsup. If the item set $\mathrm{support}(I) \geq \mathrm{minsup}$, the item set I_i is called frequent item set.

(4) Rule confidence: the form of expression is $A => B$ rule, A is the front and B is the back. $A => B$ confidence is:

$$\mathrm{confidence}(A => B) = \mathrm{support}(A \cup B)/\mathrm{support}(A) = P(B\,|A) \tag{2}$$

(5) Association rule: for a given minsup and minconf, if $\mathrm{support}(A \cup B) \geq \mathrm{minsup}$ and $\mathrm{confidence}(A => B) \geq \mathrm{minconf}$ are used, $A => B$ is called a strong association rule. In the mining of alarm data association rules in communication network, rule $A => B$ indicates that alarms in A may derive alarms in B.

According to the definition, association rules mining mainly includes two steps: finding frequent item sets and generating strong association rules from frequent item sets. In the two steps, the second step is easy to implement, so the overall performance of association rule mining depends on the implementation of the first step.

2.2. Fuzzy Theory

Set theory [21,22] put forward the principle of general situation for the first time, which shows that all things in the world are possible to have a certain attribute and not have a certain attribute, and there is no other possibility. That is to say, the attributes of all things are clear, not fuzzy. If this relationship is described mathematically, it means that there is an element w and element set X that belong to the domain W, then the relationship between w and X is shown as follows:

$$L(X, w, W) = \begin{cases} 1, w \in X \\ 0, w \notin X \end{cases} \tag{3}$$

However, with the development of the application of set theory, it is found that in many cases, it is not suitable, because the real world is not composed of clear things alone, 20 is composed of fuzzy things and clear things. So, the general situation of set theory does not apply to all things in the world. A new theory must be found to describe the fuzzy attribute. Therefore, in the middle of the 20th century, researchers put forward the fuzzy theory. Based on the classical set theory, the theory defines the specific meaning of "fuzzy", and puts forward the important fuzzy theory description tools such as fuzzy membership degree, which perfectly describes the fuzzy phenomenon of all things in the world. In the proposed fuzzy theory, the relationship between w and X is redefined as follows.

$$L(X, w, W) = f_X(w), f_X(w) \in [0, 1] \tag{4}$$

Where f is a fuzzy membership function that describes the attribution degree of an attribute to something, its value range is 0 to 1; $f_X(w)$ represents the fuzzy membership degree of w to X.

Since the fuzzy theory has been put forward and relied on, it has attracted extensive attention of the industry. Researchers all over the world have conducted in-depth research on it, and further put forward many more in-depth fuzzy technologies. At the same time, fuzzy technology is also applied to many research fields, such as the research of network fault diagnosis in this paper, which takes into account the fuzzy relationship between alarm information and network fault, introduces fuzzy theory on the basis of data mining, and combines data mining and fuzzy theory to diagnose the fault of network alarm.

2.3. Mining Fuzzy association rules

On the basis of data mining technology, fuzzy association rule mining combined with fuzzy theory is different from traditional association rule mining. In the traditional association rule mining, each alarm information is roughly divided into certain sets, while the fuzzy association rule mining is to divide the attributes into different fuzzy sets. Each alarm information may belong to different fuzzy sets, and the possibility of membership is calculated according to the specific membership function. In fuzzy association rule mining, many concepts are defined which are different from those in traditional association rule mining.

Fuzzy attribute item: suppose $I = \{i_1, i_2, \dots, i_n\}$ is a set of all attributes, which is called attribute item set. Each $i_k (k = , 2, \dots, n)$ is called an item. It is assumed that each attribute item i_k is fuzzy into attribute items of different evaluation intervals of f_k, that is, for any attribute item i_k, there is a corresponding fuzzy set $\left\{i_k^1, i_k^2, \dots, i_k^{f_k}\right\}$, where $i_k^{f_k}$ is called fuzzy attribute item. The I_f shape of the fuzzy attribute set consisting of all fuzzy attribute items is shown below.

$$I_f = \left\{i_1^1, i_1^1, \dots, i_1^{f_1}, i_2^1, i_2^2, \dots, i_2^{f_2}, \dots, i_n^1, i_n^2, \dots, i_n^{f_n}\right\} \tag{5}$$

The idea of fuzzy transaction T_f and fuzzy transaction database $D_f = \{t_{1f}, t_{2f}, \dots, t_{nf}\}(t_{kf}, (k = 1, 2, \dots, n)$ stands for the k transaction) is similar to that of traditional transaction database, but the transaction T is replaced by fuzzy transaction T_f, and the item in T becomes the fuzzy attribute item $i_k^{f_k}$ in T_f. The support of T_f of $i_k^{f_k}$ contained in it also becomes the membership degree with the value of $[0, 1]$.

Fuzzy support degree: the fuzzy support degree of fuzzy transaction database D_f to any fuzzy attribute item $i_k^{f_k}$ is expressed as:

$$\text{support}(i_k^{f_k}) = \frac{1}{n} \sum_{p=1}^{n} \text{support}(i_k^{f_k})_p \tag{6}$$

Where n is the total number of fuzzy transaction databases; support $(i_k^{f_k})_p$ is the membership degree of$i_k^{f_k}$, the fuzzy attribute item in the p transaction of D_f.

Fuzzy attribute item set $X_f = \{X_{1f}, X_{2f}, \dots, X_{mf}\}$, where $X_{kf}, k \in [1, m]$ is one of the fuzzy attribute set I_f, its fuzzy support expression is:

$$\text{support}(X_f) = \frac{1}{n} \sum_{p=1}^{n} \left[\prod_{k=1}^{m} \text{support}(X_{kf}) \right]_p \tag{7}$$

Fuzzy minimum support F_{minsup} : the minimum fuzzy support threshold set by the user to determine the frequency of any fuzzy attribute set X_f in the fuzzy transaction database D_f.

Fuzzy frequent k item set: the set of fuzzy attribute items is the set of fuzzy attribute items, which is called fuzzy item set for short. The number of fuzzy attribute items is the length of fuzzy item set. As the name implies, the fuzzy term set of k order is the fuzzy term set whose length is k, which is called fuzzy k term set for short. The fuzzy frequent k term set refers to the k term set of fuzzy support $F\_\text{support} \geq F_{\text{minsup}}$.

Fuzzy association rules: if there are two sets of fuzzy terms X_f and Y_f, the formula of $X_f => Y_f$ is called fuzzy association rules, $X_f \subseteq I_f, Y_f \subseteq I_f$ and $X_f \cap Y_f = \varnothing$. Then X_f and Y_f are the rule front part and the rule back part respectively. $X_f \cup Y_f$ does not contain the same attribute to different fuzzy evaluation interval. For example: use Low, $Middle$ and $High$ to describe the attribute of student's "score S", then the fuzzy attribute items $S.Low$, $S.Middle$ and $S.High$ cannot appear in a fuzzy association rule at the same time, for example: $S.Low => S.High$ cannot exist. Because such rules contradict with people's understanding, there is no meaning in real life.

Fuzzy confidence: if there is a fuzzy association rule $X_f => Y_f$, its fuzzy confidence can be expressed as the ratio of fuzzy item set $X_f \cup Y_f$ to fuzzy item set X_f support, which is recorded as:

$$confidence(X_f => Y_f) = \text{support}(X_f \cup Y_f)/\text{support}(X_f) = P(Y_f \,\middle|\, X_f) \tag{8}$$

Fuzzy minimum confidence $F_{\min conf}$: the minimum fuzzy confidence threshold set by the user to determine whether a fuzzy association rule is a strong rule.

Fuzzy strong association rule: if $F\_confidence(X_f => Y_f) \geq F_{\min conf}$, $X_f => Y_f$ is the fuzzy strong association rule.

3. Dynamic mining algorithm of time window Fuzzy association rules

In the past research, association rule mining is an object-oriented mining algorithm, and the alarm database we need to deal with is not a transaction database, so if we want to mine the fuzzy association rules of the alarm database, we must transform the alarm database into the representation of the transaction database. In the existing transformation methods, the sliding time window [23,24] is generally used to transform the data. But in this processing method, the alarm database is transformed into transaction database by using sliding time window. In the process of mining association rules, the support and confidence of association rules will be interfered by various factors, which can be roughly divided into two aspects: time window overlap and time window stage. As we all know, the confidence of association

Table 1
i_1 item set table.

Alarm	1	2	3
I_A	0 s	5 s	10 s
I_B	4 s	9 s	16 s
I_C	4 s	18 s	26 s
I_D	5 s	16 s	35 s
I_E	10 s	21 s	30 s
I_F	10 s	15 s	
I_G	14 s	14 s	37 s
I_H	1 s	5 s	
I_I	10 s		

rules has a great influence on the accuracy of network fault diagnosis, so a new technology is urgently needed to replace the sliding time window method. Based on this, this paper proposes a dynamic time window fuzzy association rules mining algorithm to mine the alarm information association rules, effectively avoiding the problems caused by using the sliding time window method.

The traditional sliding time window method is to transform the alarm database into a transaction database in the process of processing. When mining the fuzzy association rules, the transformed alarm transaction database is used to verify the number of times that the items contained in the candidate set appear in an alarm transaction and the fuzzy membership degree of each occurrence, so as to count the support degree of this item set. This shows that the traditional association rule mining is to mine the transformed transaction database, which will lead to the uncertainty of support and confidence in association rule mining. In order to avoid this problem, the algorithm in this paper wants to directly process the alarm database to determine whether some alarm data belong to the same alarm transaction. Therefore, the method designed in this paper is to use a fixed size dynamic time window to deal with the alarm database directly when counting the candidate sets, to make sure whether the candidate set items appear in the same time window and the number of times.

The main steps of dynamic time window fuzzy association rule mining algorithm are described as follows:

(1) First, scan the database composed of the selected alarm data, and store the alarm time and membership degree in the i_1 item set table at the same time. The specific storage information is shown in Table 1.

(2) The support of each item set is calculated by scanning the i_1 item set table, and the candidate i_1 item set is pruned to obtain the frequent i_1 item set table. $N = 2$.

(3) If $N \geq 2$ is used, the first $N - 1$ items of frequent i_n item sets are linked into candidate i_{N-1} item sets, and the support degree of candidate i_{N-1} item sets is calculated by querying frequent i_1 itemset table, then the candidate $N + 1$ item sets are pruned. $N++$.

(4) If the number of frequent i_N item sets obtained is greater than 1, go to step (3); otherwise, generate association rules.

The flow chart of dynamic time window fuzzy association rule mining algorithm is shown in Fig. 1.

From the flow chart of the dynamic time window fuzzy association rule mining algorithm in Fig. 1, it can be seen that the algorithm in this paper records the alarm time information in the i_1 itemset table, and the subsequent calculation of the i_N itemset ($N \geq 2$) support is based on the frequent i_1 itemset table, so the key of this algorithm is to determine the frequent i_1 itemset table. This algorithm is improved on the basis of traditional Apriori algorithm, so this algorithm also has the characteristics of Apriori algorithm forward pruning. Apriori algorithm shows that a candidate i_{N+1} item set is linked by two frequent i_N item sets with the same previous $(N - 1)$ items. The fuzzy support degree of candidate item set is calculated to verify whether the linked i_{N+1} item set is frequent. Based on this feature, the algorithm stores the time information of alarm occurrence in the frequent itemset table, and determines the alarm transaction in the support calculation stage of the candidate itemset. For example, this paper assumes that the size of

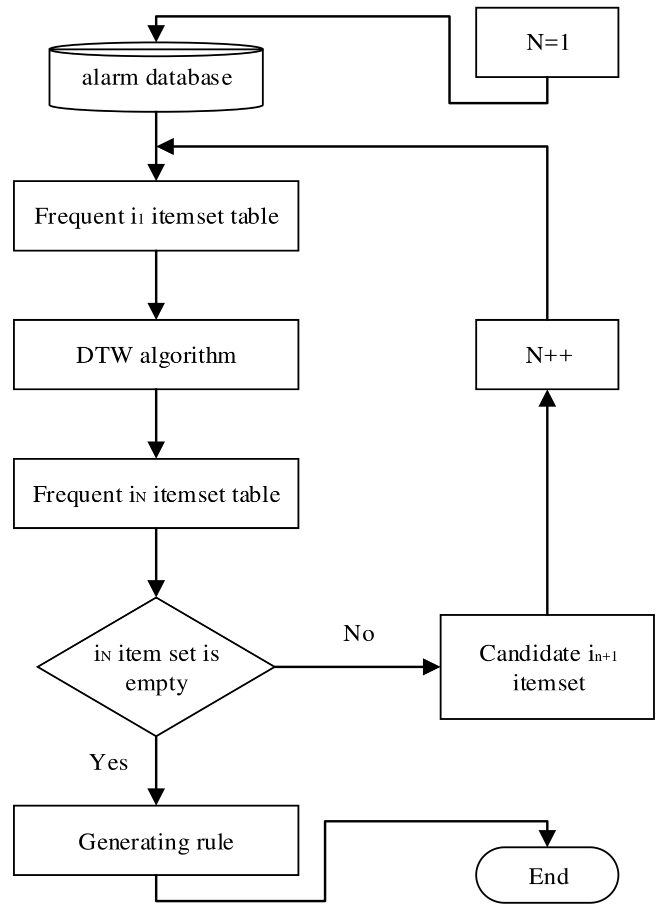

Fig. 1. Dynamic algorithm flow chart of time window fuzzy association rules mining.

the time window W is 5S, so we can get the candidate set $I_B I_C$. By querying the frequent i_1 item set table (assume that the alarm items in Table 1 given in this paper are frequent, that is to say, assume that Table 1 is the frequent i_1 item set table), we can know that the number of I_B and I_C occurrence is 3 times, among which the occurrence time interval less than W is the first occurrence of I_B and the first occurrence of I_C, the third occurrence of I_B and the second occurrence of I_C show that the support count of $I_B I_C$ is 2. After analysis, we can know that the method designed in this paper can accurately mine the support and confidence of alarm association rules in theory. It should be pointed out that the difference between fuzzy association rule mining and traditional Boolean association rule mining [25,26] is that the calculation formula of fuzzy support degree is different from Boolean data mining, and frequent i_1 item set table recording alarm time information and alarm fuzzy membership degree is set in fuzzy association rule mining. In addition, you can find that if you need to change the time window size during the mining process, you do not need to do any processing on the alarm database and frequent i_1 item set table, you only need to change the size of the time window for processing. In the traditional sliding time window method, if the time window needs to be changed in the mining process, the alarm transaction database that has been converted must be discarded, and a new alarm transaction database is generated according to the alarm database again, which will lead to the inefficiency of association rule mining.

4. Results analysis

In the dynamic time window fuzzy association rule mining algorithm, the selection of time window size is an important factor in the

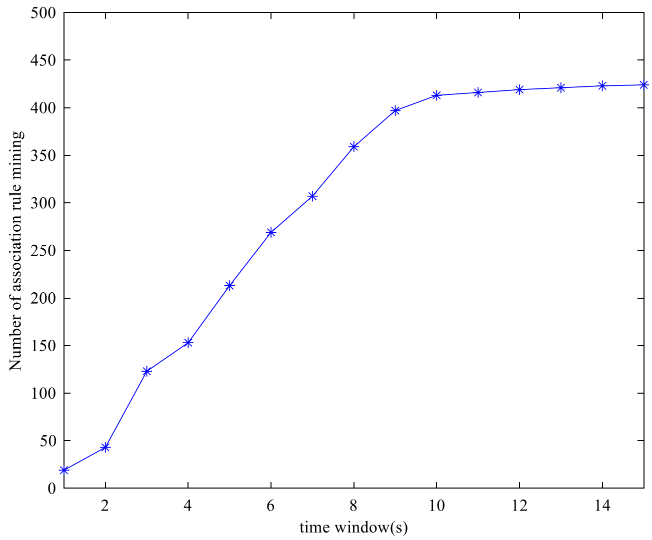

Fig. 2. The change of mining number of association rules with the size of time window.

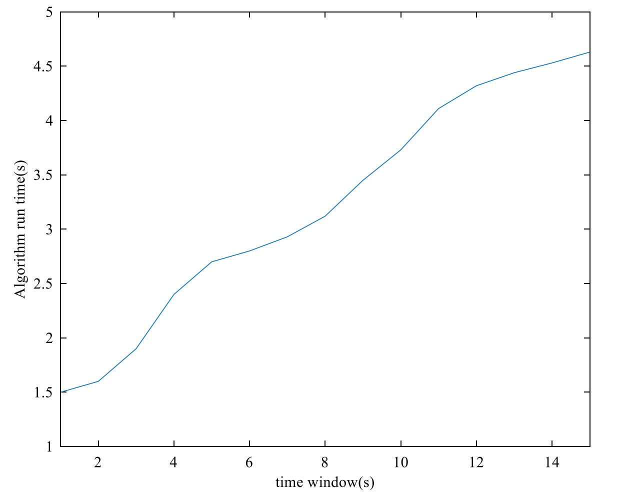

Fig. 3. The change of algorithm running time with the size of time window.

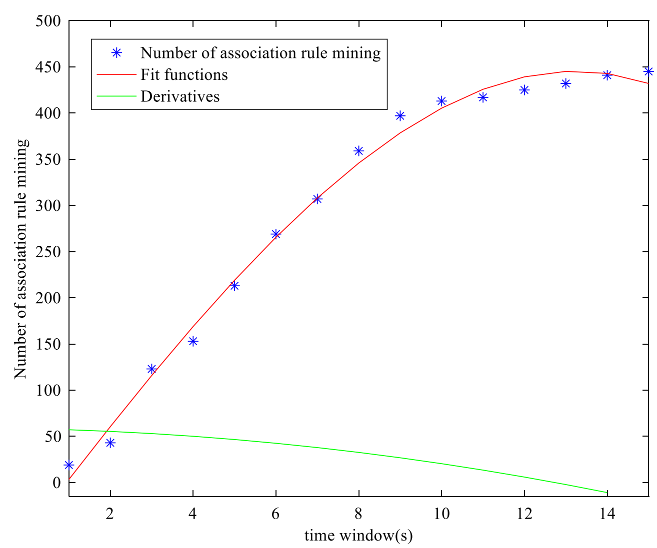

Fig. 4. Change of fitting function and derivative.

network alarm correlation analysis. The selection is directly related to the accuracy of association rule mining. This influence is embodied in the determination of alarm information association. Different time windows lead to different association judgments between different alarm information. 5S and 6S time windows lead to different association judgments between different alarm information. This shows that the larger the time window is, the more association rules will be mined. In this paper, the simulation analysis of this situation is carried out, and the results are shown in Fig. 2. As can be seen from Fig. 2, the number of mining alarm information association rules increases with the time window becoming larger. However, when the time window reaches 10 s, the number of association rule mining basically changes little. This is because although the time window becomes larger, there is not too much association between the alarm information mining with the time window becoming larger, so the number of rules does not change too much. It can also be seen from this that although the number of rule mining increases with the increase of time window, it is not that the larger the time window, the better.

The increase of time window will lead to the increase of the number of mining rules, but with the increase of time window, the number of alarms in a time window will increase, which will also lead to the increase of the time required for mining association rules, as shown in Fig. 3. As can be seen from Fig. 3, the processing time of the algorithm is proportional to the time window, and increases with the growth of the time window.

Combined with Figs. 2 and 3, we can know that although the number of association rule mining increases with the growth of time window, the running time of algorithm also increases with the growth of time window, which leads to the loss of more time while pursuing as many rules as possible, so for the efficiency of association rule mining, the larger the time window, the better. Therefore, in this paper, the number of association rules and the fitting function and derivative of the time window curve are used for analysis, and the results are shown in Fig. 4. It can be seen from Fig. 4 that the derivative curve shows a downward trend as a whole, and the curve of the fitting function shows a downward trend after the time window is 13, and the derivative at this time is less than 0. When the derivative is less than 0, it means that the number of association rules almost no longer increases with the increase of time window. In addition, considering the loss of algorithm running time, the selection of time window size should not be higher than 13 s. In this paper, the number of rules and the loss of time are integrated, and the time window is set as 10 s.

After choosing the length of time window, this paper analyzes the performance of the design algorithm in the accuracy of rule confidence, and compares it with the traditional Apriori algorithm and sliding time window method. In the simulation analysis, in order to facilitate the analysis of performance, this paper sets the support degree of X as $(1, 0, 0, 0)$, and inserts Y alarm after 80% data of alarm database is inserted into X alarm. According to the calculation of confidence degree of association rules, the confidence degree of $X => Y$ is 80%, and the minimum fuzzy support degree set at this time is 0.25%. Three methods are used to mine association rules of alarm database, and the confidence degree obtained by three methods under different number of alarm databases is shown in Fig. 5. As can be seen from the figure, after setting the length of time window, the confidence obtained in the process of mining association rules is stable at 80%. The other two methods are not accurate and vary with the increase of the number of alarms. The association rules mining method of sliding time window is slightly better than Apriori algorithm in accuracy and stability. Generally speaking, this method can accurately and stably solve the confidence degree in the process of mining association rules, and has better application effect in the mining of fuzzy association rules for network alarms.

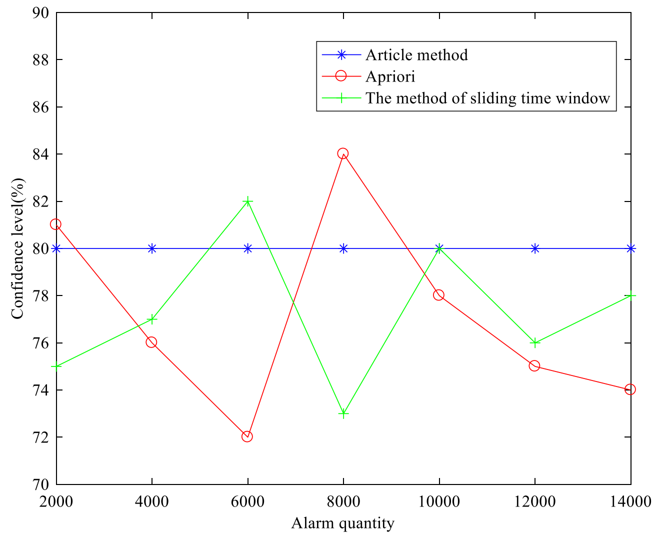

Fig. 5. Confidence obtained from association rules mining by three methods.

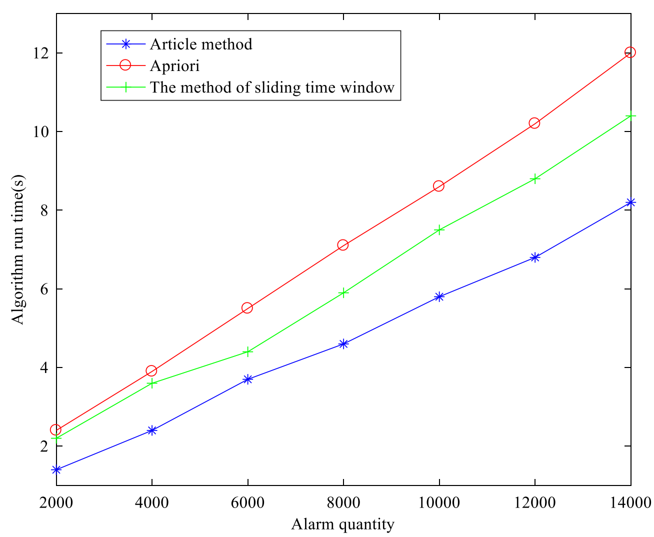

Fig. 6. Running time of association rule mining with three methods.

Then, this paper analyzes the lost time of the three methods in the process of mining association rules, and the running time of the three algorithms is shown in Fig. 6. It can be seen from Fig. 6 that the running time of the three algorithms in association rule mining increases with the increase of alarm database, among which the traditional Apriori algorithm has the largest growth trend, followed by the association rule mining method of sliding time window, and the growth trend of this method is the gentlest compared with the other two methods. This is because the method in this paper is to directly mine the alarm data in the alarm database, while the association rule mining method of sliding time window needs to convert the alarm database into the transaction database, and then mine the rules in the transaction database, which leads to the efficiency of association rule mining in this method is better than the other two mining methods.

5. Conclusions

Now, with the wide application and increasing complexity of communication network, the alarm data of communication network increases gradually, which brings great challenge to network managers to find and locate the exact location of fault in time. In order to locate the fault more conveniently and accurately, it is necessary to improve the traditional fault search method with the help of computer technology, and study a new automatic analysis technology of alarm correlation. In this paper, based on the shortcomings of traditional data mining technology, combined with data mining technology and fuzzy theory, a new mining method of fuzzy association rules is formed and applied to alarm correlation analysis. In addition, the traditional sliding time window mining method must transform the alarm database into a transaction database, which will bring deviation to the network alarm analysis. In order to solve this problem, this paper designs a dynamic time window fuzzy association rules mining algorithm based on the fuzzy association rules mining, which can directly mine the association rules of the alarm database and effectively avoid the interference in the process of converting the alarm database to the transaction database. In the simulation analysis, this paper analyzes the time window length selection of the design method. After choosing the reasonable length of time window, the performance of this method and the traditional mining method is analyzed, which shows that the method designed in this paper can obtain the confidence stably and accurately in the mining of fuzzy association rules, and needs less time in the mining process of fuzzy association rules, and has better efficiency of mining association rules.

CRediT authorship contribution statement

Qun Zheng: Conceptualization. **Yaofeng Li:** Writing - original draft. **Jie Cao:** Writing - original draft.

Declaration of competing 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.

Acknowledgment

Funding

This work was supported by the Science and Technology Project of the Jilin Provincial Education Department, China No. JJKH20200122KJ.

References

[1] R.P. Bonidia, J.D. Brancher, R.M. Busto, Data mining in sports: a systematic review. 16(1) (2018) 232-239.
[2] H.-L. Li, Y. Liang, S.-C. Wang, Review on dynamic time warping in time series data mining, Control Decis. 33 (8) (2018) 1345–1353.
[3] H. Chuang, C. Chen, S. Li, Incorporating monotonic domain knowledge in support vector learning for data mining regression problems, Neural Comput. Appl. 32 (2020) 11791–11805.
[4] Frederick Parente, John-Christopher Finley, Using association rules to measure subjective organization after acquired brain injury, Neurorehabilitation 42 (1) (2018) 9–15.
[5] H.S. Yang, Y.M. Xie, C. Chen, Y. Zhuang, Y. Zhang, [association rules analysis of fufang kushen injection in combination with modern medications in treating lung cancer: real-world study based on hospital information]. 43(8) (2018) 1708-1713.
[6] S. Cai, R. Sun, S. Hao, et al., Minimal weighted infrequent itemset mining-based outlier detection approach on uncertain data stream, Neural Comput. Appl. 32 (2020) 6619–6639.
[7] Mohamed Ali Kammoun, Nidhal Rezg, Toward the optimal selective maintenance for multi-component systems using observed failure: applied to the fms study case, Int. J. Adv. Manuf. Technol. 96 (2) (2018) 1–15.
[8] Xueling Ma, Jianming Zhan, Muhammad Irfan Ali, Nayyar Mehmood, A survey of decision making methods based on two classes of hybrid soft set models, Artif. Intell. Rev. 49 (4) (2018) 511–529.
[9] Mohamed Rafik N. Qureshi, Ram Karan Singh, Mohd Abul Hasan, Decision support model to select crop pattern for sustainable agricultural practices using fuzzy mcdm, Environ. Develop. Sustain. 20 (2) (2018) 1–19.

[10] Adam Khalifa Mohamed, Dan Liu, Mohamed A.A. Mohamed, Kai Song, Groundwater quality assessment of the quaternary unconsolidated sedimentary basin near the pi river using fuzzy evaluation technique, Appl. Water Sci. 8 (2) (2018) 1–12.

[11] Z. Cao, C.T. Lin, Inherent fuzzy entropy for the improvement of eeg complexity evaluation. 26(2)(2018) 1032-1035.

[12] Na Ding, Ye-Peng Liu, Lin-Wei Fan, Cai-Ming Zhang, Single image super-resolution via dynamic lightweight database with local-feature based interpolation, J. Comput. Sci. Tech. 34 (3) (2019) 537–549.

[13] Kipp W. Johnson, Jessica Torres Soto, Benjamin S. Glicksberg, Khader Shameer, Joel T. Dudley, Artificial intelligence in cardiology, J. Am. Coll. Cardiol. 71 (23) (2018) 2668–2679.

[14] Ian Goodfellow, Patrick McDaniel, Nicolas Papernot, Making machine learning robust against adversarial inputs, Commun. ACM 61 (7) (2018) 56–66.

[15] R.I. Smith, D. Fowler, J.N. Cape, The statistics of phytotoxic air pollutants, J. R. Stat. Soc. 152 (2) (2018) 183–198.

[16] Isuru Jayasinghe, Alexander H. Clowsley, Ruisheng Lin, Tobias Lutz, Carl Harrison, True molecular scale visualization of variable clustering properties of ryanodine receptors, Cell Rep. 22 (2) (2018) 557–567.

[17] Famao Ye, Yanfei Su, Hui Xiao, Xuqing Zhao, Weidong Min, Remote sensing image registration using convolutional neural network features, IEEE Geosci. Remote Sensing Lett. 15 (2) (2018) 232–236.

[18] S. Buschjäger, K. Morik, Decision tree and random forest implementations for fast filtering of sensor data, IEEE Trans. Circuits Syst. I. Regul. Pap. 65 (1) (2018) 209–222.

[19] Trang Van, Atsuo Yoshitaka, Bac Le, Mining web access patterns with super-pattern constraint, Appl. Intell. 48 (11) (2018) 3902–3914.

[20] Jian Pei, Jiawei Han, B. Mortazavi-Asl, Jianyong Wang, H. Pinto, Qiming Chen, et al., Mining sequential patterns by pattern-growth: the prefixspan approach, IEEE Trans. Knowl. Data Eng. 16 (11) (2004) 1424–1440.

[21] Mridu Sahu, N.K. Nagwani, ShrishVerma, Optimal channel selection on electroencephalography (eeg) device data using feature re-ranking and rough set theory on eye state classification problem, J. Med. Imaging Health Inform. 8 (2) (2018) 214–222.

[22] Ignagio Jane, Reflections on skolem's relativity of set-theoretical concepts, Philos. Math. 9 (2) (2018) 129–153.

[23] Xu Zheng, Liu Yunhuai, Neil Y. Yen, Lin Mei, Xiangfeng Luo, Xiao Wei, Chuanping Hu, Crowdsourcing based description of urban emergency events using social media big data, IEEE Trans. Cloud Comput. 8 (2) (2020) 387–397.

[24] M. Stojković, S. Kostić, S. Prohaska, J. Plavšić, V. Tripković, A new approach for trend assessment of annual streamflows: a case study of hydropower plants in serbia. 31(4) (2018) 1-15.

[25] Tutut Herawan, Mustafa Mat Deris, A soft set approach for association rules mining, Knowl.-Based Syst. 24 (1) (2011) 186–195.

[26] V. Bolon-Canedo, N. Sanchez-Marono, A. Alonso-Betanzos, Recent advances and emerging challenges of feature selection in the context of big data, Knowl.-Based Syst. 86 (sep) (2015) 33–45.