IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received June 25, 2020, accepted June 28, 2020, date of publication July 1, 2020, date of current version July 14, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3006345

Data Dependence Analysis for Defects Data of Relay Protection Devices Based on Apriori Algorithm

MINGWEI TIAN[1], LIE ZHANG[2], PENG GUO[2], (Member, IEEE), HANFANG ZHANG[2], QIAN CHEN[1], YANFEI LI[2], AND ANCHENG XUE[1], (Member, IEEE)

[1]State Key Laboratory of Alternate Electrical Power System with Renewable Energy Source, North China Electric Power University, Beijing 102206, China
[2]China Electric Power Research Institute, Beijing 100192, China

Corresponding author: Ancheng Xue (acxue@ncepu.edu.cn)

This work was supported by the project Research on the basic key technology of secondary equipment evaluation center under Grant WBS:5442JB190008.

ABSTRACT Currently, a large amount of defect data in relay protection devices (RPDs) is accumulated in operation. However, the defect data dependence analysis is absence and thus it could not meet the demand for further improving the management and operation RPDs. Based on 7-years defect data of RPDs in SGCC, this paper discovers the association rules (ARs) of defect data based on the Apriori algorithm. In detail, the ARs among different categories of PRDs, such as defect parts and defect causes are discovered and analyzed. Furthermore, the family characteristics of defects are illustrated, with the defect data of RPDs from different manufacturers. The analysis results show that the Apriori method can effectively reveal the hidden information in the defect data, such as the ARs between the vulnerable parts of RPDs, defect causes and other factors.

INDEX TERMS Relay protection devices (RPDs), defect analysis, data mining, association rules (ARs), Apriori algorithm.

NOMENCLATURES

Relay Protection Devices (RPDs), Relay Protection System (RPS), State Grid Corporation of China (SGCC), Association Rule (AR), Condition Based Maintenance (CBM), Merging Unit (MU).

I. INTRODUCTION

Relay protection is the first line of defense to ensure the safety of the power grid, thus its reliability is very important. However, statistical data show that various defects often result in failure of RPDs, which may threaten the safety of power grid, or even lead to blackouts [1].

Currently, with the expansion of power grid, the number of RPDs in operation is large. For example, there are over 180000 RPDs in SGCC. Thus, a lot of data, which contains a lot of defects information, is produced by the RPS in operation [2]. Thus, mining the above data, could recognize the regular defects patterns of RPDs, guide the operation, maintenance and management of RPDs, so as to improve the reliability of RPS.

There is some work on the mining detects data [3], however, they are focused on using traditional statistical methods to carry out simple classification, and this work is lack of deep mining, e.g., the dependence among data, the defects patterns. Thus, the hidden information lying in the defects data PRDs needs to be further discovered.

On the other hand, data mining could extract the patterns and rules concerned by users from seemingly unrelated massive data [4]–[6], thus it has been applied to the data analysis for the equipment in power grid [7]–[13]. For example, in Ref [7], the Apriori algorithm is used to mine the defects data of the automation equipment, which is part of secondary equipment, and reveals the weak link of the automation equipment and the causes of defect. In Ref [8], an improved algorithm called FP-growth is applied to analyze the defect data of secondary equipment the certain substation, discover the dependence among frequent, location, property and cause of the frequent defects in substation.

It is worth noting that, the applications of the defects data of PRDs are focusing on condition based maintenance (CBM)

The associate editor coordinating the review of this manuscript and approving it for publication was Mingjian Cui.

and risk assessment, but not the dependence of the data. For example, in Refs [9] and [10], based on the failure characteristics of RPD, the reliability evaluation model of RPD is constructed to determine the optimal maintenance period and replacement period of RPD. In Ref [11]–[13], based on the RPS reliability data, Markov state method is used to establish the system reliability evaluation model by considering the fixed inspection, self-inspection and protection configuration of RPD, and the optimal fixed inspection period of RPD is proposed. As state before, the above work has initially reflected the value of defects data of RPDs, but they are focusing on device reliability assessment, lacking of the dependence analysis to the defects data themself.

Recognizing the above problems, based on the 7-years defects data of PRDs in SGCC, this paper applies the Apriori algorithm to reveal the dependence among the defect data. The ARs of the defect parts and defect categories and defect causes of RPDs are discovered, and the vulnerable parts and defect causes of all kinds of RPD are determined. The similarities and differences of the obtained dependence rules according to the defect data of RPDs with different manufacturer discover the characteristics of family defect.

The contributions of the paper are as follows.

(1) The Apriori algorithm is applied to mine the hidden association information in defects data of RPDs, and the ARs in the device category, defect parts and defect causes for the common defects of RPDs are revealed.

(2) Based on the analysis of the ARs of different manufacturers, the characteristics of family defects of the RPDs are discovered, which could be conducive to the improvement of the production quality.

The remainders of the paper are as follows. Section II introduces the basic concept of ARs, including the support and confidence and mathematical calculation formula, after that, describes the implementation process of Apriori algorithm. Section III analyzes the structure of the data, and select appropriate feature items to participate in the generation of ARs according to the characteristics of the data. In Section IV, the defect data of RPDs in SGCC from 2012 to 2018 is mined as a whole, and the common ARs of defects are analyzed. Then, based on the data of three manufacturers, the family characteristics of devices are explored. Finally, section V gives out the conclusions and remarks.

II. ASSOCIATION RULES AND APRIORI ALGORITHM

This section introduces the basic concept of ARs, including the support and confidence and mathematical calculation formula, after that, describes the implementation process of Apriori algorithm.

A. ASSOCIATION RULES

The AR is correlation between different transactions, which is known as the rule of "beer-diaper" from the user's purchase record [14]. The AR mining, which is also popular known as "shopping basket analysis", could discover the hidden relationship between two or more things, provide the association mechanism, and even predict the occurrence of things. The mathematical description of AR is as follows [15].

Let I be a set of M different items, which is called itemset. A itemset with a length of K is called k-itemset. The sample set T used for association rule mining is a subset of itemset I. and all the samples form the sample database D. Then, there are two key indexes to evaluate the rule, one is the support and the other is the confidence. Support is the possibility of a rule, confidence is the degree of trust of a rule.

A and B are two subsets of itemset I, and their AR could be expressed as:

$$R : A => B$$

Count(A) is the number of samples in sample set T that contains A, and Count(B) is the number of samples in sample set T that contains B. the mathematical expression of the support of itemset A is as follows:

$$Support\ (A) = \frac{count\ (A)}{|D|} \tag{1}$$

The mathematical expression of support for rule R is as follows:

$$Support\ (A => B) = \frac{count\ (A \cup B)}{|D|} \tag{2}$$

Count (A∪B) indicates the number of samples in sample set T that contains A and B.

The mathematical expression of the confidence of rule R is as follows:

$$Confidence\ (A => B) = \frac{count(A \cup B)}{count(A)} \tag{3}$$

If one obtained AR is engaged with both the minimum support and the minimum confidence threshold, then it is called a strong AR. Furthermore, the AR mining is a process of finding frequent itemset from data sets and finally generating rules through filtering according to the needs of minimum support and minimum confidence.

B. APRIORI ALGORITHM

As one of the classical algorithms of association rule mining, Apriori algorithm extracts sub item sets that can represent the rules of data sets through iterating layer by layer. The basic implementation of the method is to discover the frequent itemset according to the combination of all different items and then gives out the ARs. The algorithm will end once the frequent itemset become empty. the corresponding flow chart is shown in Figure 1.

In detail, the general processes of Apriori algorithm are as follows:

Step 1: A minimum support G ($0 < G < 1$) is determined according to the total amount of data.

Step 2: In first iteration, all elements of the sample set T are called transactions, each of them are composed of many characteristics. These characteristics are members of candidate 1-term set C1. In this step, the algorithm scans all transactions, counts the number of occurrences of

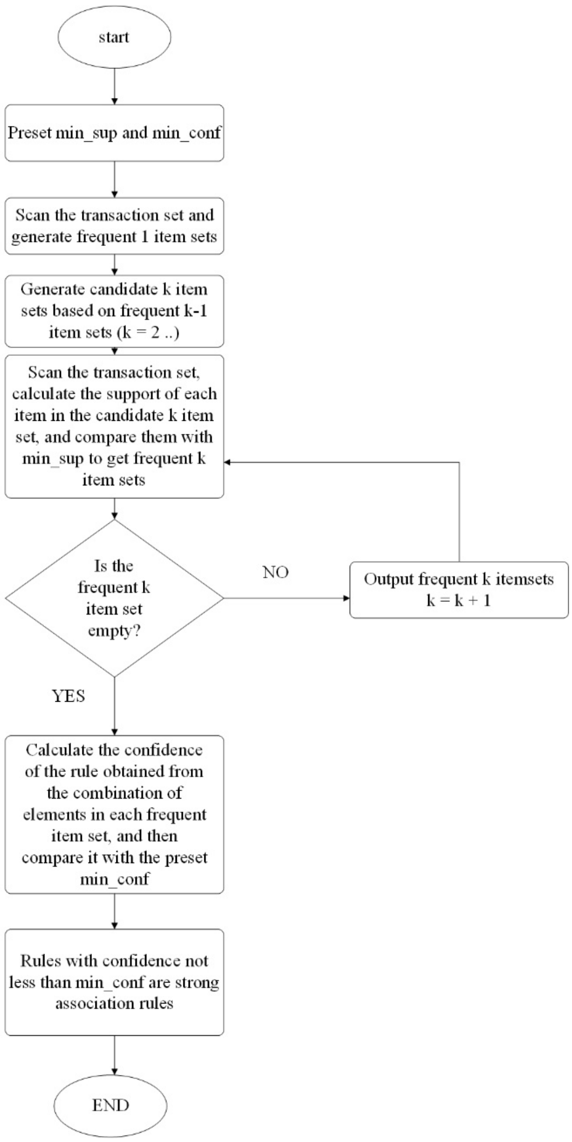

FIGURE 1. Flow chart of Apriori algorithm.

each characteristics, and then calculates the support according to Eq (1).

Step 3: If the support of a characteristic is less than the minimum support, it is abandoned; if the support of a characteristic is not less than the minimum support, it is retained. Furthermore, all characteristics not less than the minimum support G constitute frequent 1-itemset L1.

Step 4: The candidate 2-term called set C2 is generated by L1. C2 is composed of two pairs of each element in L1. Similarly, scan the transactions, calculate the support of elements in C2 according to Eq (2), and then retain the combinations that are not less than the minimum support G, which constitute frequent 2-itemset called set L2.

Step 5: When C3 is generated by L2, Apriori pruning is used: all subsets of frequent itemset must be frequent. Then take the same method as the above steps, scan the transactions, calculate the support of each element in C3 according to Eq (2), reserve combinations that are not less than the minimum support G, which constitute frequent 3-item set called L3.

Step 6: Follow steps (1) - (5) to search and iterate layer by layer until the frequent k-itemset cannot be found.

Step 7: A minimum confidence U ($0 < U < 1$) is determined according to the total amount of data.

Step 8: According to Eq (3), the confidence of the items included in each frequent itemset (L1, L2…Lk) are calculated

Step 9: Compare the confidence of the above items with the minimum confidence U. The items whose confidence are not less than the minimum confidence threshold are called ARs.

TABLE 1. Set of features included in device properties.

No.	feature	Category feature
1	Manufacturer	equipment attribute
2	Device model	
3	Protection category	
4	Specific classification of protection categories	
5	Software version	
6	Id of the protection device	
7	Factory category	
8	Defective equipment category	

TABLE 2. Set of features included in defect level.

No.	feature	Category features
1	Voltage level	defect description
2	Time of defect occurs	
3	Time of defect eliminated	
4	Duration of the defect	
5	Responsible department	
6	Defect part	
7	Specific defect cause	
8	The type of defect	
9	Defect cause	
10	Processing Record	

III. DEFECT DATA OF RPD AND FEATURE SELECTION

In this section, the structure of the defects data is analyzed, and furthermore, appropriate features are selected to participate in the generation of ARs, according to the characteristics of the data.

A. DEFECT DATA INTRODUCTION AND PROCESSING

The defect data of RPDs used in this paper is range from 2012 to 2018 years in SGCC. There are 18439 records of defect data. Every data contains 21 features, according to the different description property, the feature sets can be divided into three category features: equipment attribute; defect level; defect description. The specific features of each category are shown in tables 1, 2 and 3.

For the collected defect information of RPD, considering the quality problems such as the lack of data items and irregular description, it is necessary to clean the data. After data processing, there are 19 categories of defective equipment, including "RPD itself", "channel interface equipment", "communication transmission equipment", etc;

TABLE 3. Set of features included in defect description.

No.	feature	Category features
1	Voltage level	defect description
2	Time of defect occurs	
3	Time of defect eliminated	
4	Duration of the defect	
5	Responsible department	
6	Defect part	
7	Specific defect cause	
8	The type of defect	
9	Defect cause	
10	Processing Record	

the causes of defects are classified into 8 categories, including "poor manufacturing quality", "poor debugging quality", "poor operation and maintenance", "internal component damage"; the specific causes of defects are summarized as "plug-in damage", "device crash", "Damage of components and parts", "principle defects", etc.; The defective parts include 71 parts such as "CPU plug-in", "interface plug-in", "acquisition plug-in" and "MMI plug-in", etc; the degree of defects includes three categories, namely "general", "serious" and "critical".

B. DEFECT FEATURE SELECTION OF RPD

The purpose of this paper is to mine the value information hidden in the defect data of RPD, such as the correlation between the categories, the vulnerable parts and the causes. Therefore, in order to avoid the noise interference caused by irrelevant feature attributes, six features, which can describe the main characteristics of defects, including manufacturer, defect equipment category, defect reason, specific defect reason, defect position and defect degree, are selected from the attribute set as the data mining object. The selected features are expressed in the following composite forms:

$$Q = (F, N\,(a, b, c, d, e)) \tag{4}$$

In formula (4), F represents the dimension feature attribute, which is used to determine the dimension range of data mining, including the main manufacturers and the overall data; N represents the collection of mining feature attributes, which is used to participate in the generation of rules, including 5 specific features, where a represents the defect equipment category, b represents the defect reason, c represents the specific defect reason, d represents the defect part and e represents the degree of defect. Q is an abstract tuple structure composed of F and N, which is used to summarize the method for defect data mining.

It can be seen from N that every defect sample is a point in the 5-Dimensional space constructed by these 5 types of defect feature attributes. On the basis of the initial candidate set, based on the Apriori algorithm, the frequent item set is filtered through the threshold value, and finally the strong ARs are generated.

Based on the above methods and considering the dimensional characteristics, the common association characteristics of defects can be obtained from the perspective of the overall data, the differences of ARs of different manufacturers can be analyzed and compared to explore the characteristics of family defects of RPDs from the perspective of manufacturer data.

TABLE 4. Strong association rules in PRD itself.

No.	Association rules	Conf.
1	RPD itself => Power plug-in	17.6%
2	RPD itself => CPU plug-in	29.3%
3	RPD itself => Other parts	16.3%
4	RPD itself => Insert plug-in	6.4%
5	RPD itself => MMI plug-in	10.0%
6	RPD itself => Channel interface plug-in	6.1%
7	RPD itself => LCD screen	6.4%
8	RPD itself &Power plug-in => Poor manufacturing quality	70.2%
9	RPD itself & CPU plug-in => Poor manufacturing quality	74.4%
10	RPD itself &Insert plug-in => Poor manufacturing quality	76.3%
11	RPD itself & MMI plug-in => Poor manufacturing quality	73.8%
12	RPD itself & Channel interface plug-in => Poor manufacturing quality	62.6%
13	RPD itself & LCD screen =>Poor manufacturing quality	57.1%
14	RPD itself => Poor manufacturing quality	68.3%
15	RPD itself => Non-human factor	10.4%
16	CPU plug-in => Serious	38.6%
17	CPU plug-in => Critical	38.6%
18	Power plug-in => Critical	42.8%

*"Conf." represents the confidence of rule (%).

IV. CASE STUDIES

In this section, the defect data of RPDs from 2012 to 2018 year in SGCC is used in the case studies. Furthermore, due to large amount of data and low occupancy rate of different type of defect samples, the minimum support threshold and the minimum confidence threshold is set to 2% and 5%, respectively. Then, the Apriori algorithm is carried out to obtain the frequent item and ARs.

A. ANALYSIS OF OVERALL RESULTS

With the defect data from 2012 to 2018 as a whole, the strong ARs of defect equipment categories, defect parts and defect causes can be obtained. Among them, the defect equipment categories mainly involve the RPD itself, communication transmission equipment, channel interface equipment, merging unit and intelligent terminal. The specific rules are shown in Table 4~8, respectively:

(1) According to the ARs 1-7 in Table 4, the defects of RPD itself mainly focus on the power plug-in, CPU plug-in, input plug-in, MMI plug-in and channel interface plug-in, with confidence of 17.6%, 29.3%, 16.3%, 6.4%, 10.0% and 6.1% respectively, thus they are the vulnerable parts of RPD itself. Furthermore, the confidence of power plug-ins and CPU plug-ins (17.6%, 29.3%) is far greater than the others, so they are the weakest parts. Thus, in maintenances, attentions should be paid to the vulnerable parts, especially the power

TABLE 5. Strong association rules of communication transmission equipment.

No.	Association rules	Conf.
19	Communication transmission equipment =>optic cable	71.6%
20	optic cable =>Critical	47.3%
21	optic cable =>Serious	40.7%

TABLE 6. Strong association rules for channel interface devices.

No.	Association rules	Conf.
22	Channel interface equipment => Interface plug-in	23.3%
23	Channel interface equipment =>linker	22.3%
24	Channel interface equipment & Interface plug-in =>Poor manufacturing quality	69.1%
25	Channel interface equipment & linker => Poor operation and maintenance	22.8%
26	Channel interface equipment & linker => Poor manufacturing quality	21.6%
27	Interface plug-in =>Serious	42.9%
28	Interface plug-in =>Critical	33.1%
29	Linker => Serious	35.6%

plug-ins and CPU plug-ins. Furthermore, the rule 3 shows that when there are defects in RPD itself, most of the defect parts cannot be accurately located. Therefore, the data collection should be more specified, for the operation and maintenance personnel.

(2) According to rule 8-13 in Table 4, the main reason for the weakness parts of RPD itself is the poor manufacturing quality. Therefore, in order to improve the reliability of the RPDs, it is necessary to improve the design, use high quality parts, and strengthen the relative maintenance.

(3) The rule 14 in Table 4 shows that the confidence of defects of RPD itself caused by the poor manufacturing quality is 68.3%. The confidence of defects of RPD itself caused by the non-human factor is 10.0% according to rule 15. Thus, different measures should be taken. Devices with defects due to poor manufacturing quality should strengthen the management and control of the manufacturer. The non-human reason is related to the overdue service of RPDs, thus, special operation and maintenance strategies should be carried out for the RPDs with long service time. In detail, the overdue service functional plug-ins should be maintenance and replaced in time.

(4) Rule 19 in Table 5 shows that the defects of the communication transmission equipment mainly lies on optical cable, with the confidence 71.6%, so it is the vulnerable part. At the same time, rules 20 and 21 in Table 5 show that the defects related to optical cable are generally classified as serious or critical defects. Therefore, the manufacturing quality, maintenance and timely replacement of optical cables should be improved pertinently.

(5) Similarly, rules 22 and 23 in Table 6 show that interface plug-ins and linkers are vulnerable parts of channel interface equipment. Furthermore, according to rule 24, 25 and 26, the defects of interface plug-ins are mainly caused by poor manufacturing, and the defects of linkers are mainly caused by poor operation and maintenance or poor manufacturing. The quality of interface plug-ins and linkers in the channel interface equipment should be improved pertinently, and the level of operation, maintenance and timely replacement of linkers should be improved at the same time.

TABLE 7. Strong association rules of merging unit and intelligent terminal.

No.	Association rules	Conf.
30	Merging unit => Poor manufacturing quality	69.5%
31	Intelligent terminal => Poor manufacturing quality	68.3%
32	Intelligent terminal => Non-human factor	9.1%

TABLE 8. Other strong association rules.

No.	Association rules	Conf.
33	Poor operation and maintenance =>Critical	39.8%
34	Poor manufacturing quality => Plug-in damage	74.0%
35	Poor manufacturing quality =>Principle defect	6.5%
36	Poor manufacturing quality =>Internal communication interruption	10.4%
37	Poor manufacturing quality =>device crash	10.5%

(6) The rules 30, 31 and 32 in Table 7 show that the defects of the merging unit and the intelligent terminal are mainly caused by the poor manufacturing quality. Therefore, the quality of merging unit (MU) should be improved.

(7) The rule 33 in Table 8 shows that defects caused by poor operation and maintenance are generally defined as critical defects with a confidence of 39.8%. Therefore, the trained level of operation and maintenance personnel needs to be strengthened through training.

(8) According to rules 34, 35, 36 and 37, poor manufacturing quality is embodied in four aspects, i.e., plug-in damage, principle defect, internal communication interruption and device crash, thus, they should be paid more attention in operation.

B. ANALYSIS OF DIFFERENT MANUFACTURERS

By applying the proposed method to the data related to three different manufacturers, the corresponding comparison result can be obtained, as shown in Table 9.

Table 9 shows that the ARs of different manufacturers share the common characteristics. (1)The confidence of CPU plug-ins from three manufacturers is 40.5%, 25.0% and 34.1% respectively; the confidence of power plug-ins is 13.4%, 14.1% and 20% respectively, which are far greater than the preset confidence threshold, therefore CPU plug-ins and power plug-ins are the common defect prone parts of the three manufacturers' RPD itself. And the design and quality of CPU plug-in and power plug-ins should be improved. (2) Manufacturers A, B and C have 63%, 66.9% and 79.9%

TABLE 9. Comparison table of three manufacturers' partial association rules.

No.	Association rules	Conf.		
		Manufacturers		
		A	B	C
1	RPD itself =>CPU plug-in	40.5%	25.0%	34.1%
2	RPD itself =>Power plug-in	13.4%	14.1%	20.0%
3	RPD itself => Communication systems	\	\	12.3%
4	Communication transmission equipment =>Optic cable	63.0%	66.9%	79.9%
5	Channel interface equipment =>linker	\	\	43.1%

*"\" represents that the confidence of the rule is less than the preset minimum confidence threshold.

confidence in defect about optical cables respectively. Thus, the quality of optical cables also should be improved.

Furthermore, in case of defects in RPD itself, the manufacturer A and C engage with high confidence related to CPU plug-ins, while the manufacturer C engages with high confidence relative to the power plug-ins and communication systems.

In case of defects in communication transmission equipment, with respect to optical cable, the manufacturer C engages with the highest confidence.

In case of defects in the channel interface equipment, with respect to the linker, the manufacturer C engages with the highest confidence, while the manufacturer A and B are small than the preset minimum confidence threshold.

Therefore, from the viewpoint of manufacturers, the confidence of CPU plug-in defects of manufacturer A is higher than the average level of the three and special attentions should be paid. For manufacturer C, the confidence of power plug-ins, communication systems, optical cables and connectors is higher than that of other manufacturers, and the above family characteristics of defects need special attention.

V. CONCLUSIONS AND REMARKS

In this paper, based on the analysis of the characteristics of RPD defect data, a method using Apriori algorithm to analyze the dependence between data is established, which includes the selection of dimension features and attribute features. By analyzing the results of defect data mining, the following conclusions are obtained:

(1) When the mining dimension is focused on the whole data, this method can effectively explore the common characteristics of defects, including the potential dependence among different categories of PRDs, such as defect parts and defect causes which can be used to guide the operation and maintenance management of RPD.

(2) When the mining dimension is concentrated in different manufacturers, this method can effectively mine the family defect features of different manufacturers' devices. According to these features, the manufacturing quality of devices can be improved in the early stage of production.

It is worth noting that this paper only retains rules larger than the preset support and confidence threshold, and more general screening rules need further study.

REFERENCES

[1] A. Xue, R. Wang, W. Liu, B. Zhuang, N. Wang, and S. Huang, "Estimation methods for constant failure rate of protection equipment," *Automat. Electr. Power Syst.*, vol. 36, no. 4, pp. 11–15, 2012.

[2] H. Yuan, G. Qu, and W. Zhuang, "Discussion on the content of state monitoring of power grid secondary equipment," *Automat. Electr. Power Syst.*, vol. 38, no. 12, pp. 100–106, 2014, doi: 10.7500/AEPS201210277.

[3] L. Lei, S. Gao, and F. Xu, "Common methods of defect analysis and treatment in secondary equipment maintenance of hydropower plant," *Hydropower. New Energy*, no. 8, pp. 35–37, 2016.

[4] H. Shen, Y. Liu, Z. Xia, and M. Zhang, *An Efficient Aggregation Scheme Resisting on Malicious Data Mining Attacks for Smart Grid*. Amsterdam, The Netherlands: Elsevier, 2020.

[5] M. Cui, J. Wang, and B. Chen, "Flexible machine learning-based cyberattack detection using spatiotemporal patterns for distribution systems," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1805–1808, Mar. 2020.

[6] C. Chen, M. Cui, F. F. Li, S. Yin, and X. Wang, "Model-free emergency frequency control based on reinforcement learning," *IEEE Trans. Ind. Informat.*, early access, Jun. 9, 2020, doi: 10.1109/TII.2020.3001095.

[7] Y. Zhang, C. Hu, S. Huang, S. Feng, and G. Lin, "Apriori algorithm based data mining and analysis method for secondary device defects," *Automat. Electr. Power Syst.*, vol. 41, no. 19, pp. 147–151, 2017.

[8] Y. Xiao, S. Liu, W. Jian, and Y. Song, "A kind of defects analysis method for secondary device of substation based on FP-growth algorithm," *Elect. Meas. Instrum.*, pp. 1–8, 2019. [Online]. Available: https://kns.cnki.net/kcms/detail/23.1202.TH.20190815.1013.006.html

[9] R. Wang, A. Xue, T. Bi, and S. Huang, "Relay replacement strategy based on the least unit life cycle cost with minimum maintenance model," in *Proc. Int. Conf. Adv. Power Syst. Autom. Protection*, Oct. 2011, pp. 668–673.

[10] R. Wang, A. Xue, T. Bi, and S. Huang, "Time-varying failure rate estimation of relay protection devices and their regional differences analysis," *Automat. Electr. Power Syst.*, vol. 36, no. 5, pp. 11–15, 2012.

[11] R. Billinton and R. Allan, *Reliability Evaluation of Engineering Systems: Concepts and Techniques*, 2nd ed. New York, NY, USA: Plenum Press, 1992.

[12] R. Billinton, M. Fotuhi-Firuzabad, and T. S. Sidhu, "Determination of the optimum routine test and self-checking intervals in protective relaying using a reliability model," *IEEE Trans. Power Syst.*, vol. 17, no. 3, pp. 663–669, Aug. 2002.

[13] P. M. Anderson, G. M. Chintaluri, S. M. Magbuhat, and R. F. Ghajar, "An improved reliability model for redundant protective systems-Markov models," *IEEE Trans. Power Syst.*, vol. 12, no. 2, pp. 573–578, May 1997.

[14] R. Agrawal, "Fast algorithms for mining association rules," in *Proc. 20th Intl Conf. Very Large Database*, 1994, pp. 487–499.

[15] E. V. Altay and B. Alatas, *Intelligent Optimization Algorithms for the Problem of Mining Numerical Association Rules*, vol. 540. Amsterdam, The Netherlands: Elsevier, 2020.

MINGWEI TIAN received the B.S. degree in electrical engineering from the Hubei University of Technology, Wuhan, China, in 2018. He is currently pursuing the master's degree with North China Electric Power University. His research interest is the reliability analysis and application of artificial intelligence in relay protection system.

LIE ZHANG was born in Liaoning, China, in 1984. She received the master's degree in power system and its automation from the Huazhong University of Science and Technology, Wuhan, China, in 2008. She is currently a Senior Engineer with the China Electric Power Research Institute, where she joined, in March 2008. Her research interest includes relay protection operation evaluation.

PENG GUO (Member, IEEE) was born in Hebei, China, in 1988. He received the Ph.D. degree in power system and its automation from North China Electric Power University, Beijing, China, in 2017. He is currently an Engineer with the China Electric Power Research Institute, where he joined, in August 2017. His research interest includes relay protection state evaluation.

HANFANG ZHANG was born in Shanxi, China, in 1991. She received the master's degree from North China Electric Power University, Baoding, China, in 2017. She is currently an Engineer with the China Electric Power Research Institute, where she joined, in September 2017. Her research interest includes relay protection operation evaluation.

QIAN CHEN received the B.S. degree in electrical engineering from Jilin University, Changchun, China, in 2017. He is currently pursuing the master's degree with North China Electric Power University. His research interest is the reliability analysis and application of artificial intelligence in relay protection system.

YANFEI LI was born in Shandong, China, in 1988. She received the master's degree in computer science and technology from the Beijing University of Posts and Telecommunications, Beijing, China, in 2013. She is currently an Engineer with the China Electric Power Research Institute, where she joined, in March 2013. Her research interest includes relay protection state evaluation and analysis.

ANCHENG XUE (Member, IEEE) was born in Jiangsu, China, in 1979. He received the B.S. degree in applied mathematics and the Ph.D. degree in electrical engineering from Tsinghua University, Beijing, China, in 2001 and 2006, respectively. He held a postdoctoral position at the Institute of System Science, Chinese Academy of Sciences. He is currently a Professor with North China Electric Power University, where he joined, in March 2008. His research interest includes the model and data-driven power system stability analysis, control, and protection.