

## Research Article

# Research on Fault Diagnosis of Launch Vehicle's Power Transformation and Transmission System Based on Big Data

Yichi Zhang <sup>1</sup>, Tao Shu,<sup>1</sup> Xincheng Song,<sup>1</sup> Yan Xu,<sup>2</sup> and Pengxiang Zhang<sup>1</sup>

<sup>1</sup>Air Defense and Missile Defense College, Air Force Engineering University, Xi'an 710051, China

<sup>2</sup>Mathematics and Statistics Science College, Shaanxi Normal University, Xi'an 710119, China

Correspondence should be addressed to Yichi Zhang; [zyc412181588@ahlcti.com](mailto:zyc412181588@ahlcti.com)

Received 18 May 2021; Accepted 7 July 2021; Published 22 July 2021

Academic Editor: Jie Chen

Copyright © 2021 Yichi Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The on-board power supply system provides power for the launch vehicle. The power transmission and transformation system plays an irreplaceable role to ensure that the on-board power supply system receives the normal working voltage of the launch vehicle. There are many types of faults in power transmission and transformation systems. The traditional faulty diagnosis method of power transmission and transformation equipment has the disadvantages of being susceptible to experts' subjectivity and model's ossification. In this paper, a new method of equipment fault diagnosis based on big data is proposed. On the basis of big data, this paper introduces the failure mode clustering algorithm, the state parameter correlation analysis algorithm, the fault diagnosis method based on the correlation matrix, and other key fault diagnosis technologies. The fault record data of the 400 kV voltage grade oil-immersed transformer bushing in the past ten years by a Chinese combat unit is used as a case for demonstration. The results show that the accuracy rate of SC-LSTM-K-means clustering model exceeds 95%. And the fault classification mode can be accurately obtained. A priori correlation algorithm with TA coefficient can be used to evaluate the strong and weak relationship between the state parameters; the fault diagnosis matrix based on Pearson's correlation coefficient can accurately determine the fault mode consistent with the actual operation and maintenance test results. Therefore, the fault diagnosis method of power transmission and transformation system based on big data can both effectively obtain the inherent laws of historical data and realize more accurate fault diagnosis with data adaptability.

## 1. Introduction

The vehicle-mounted power supply system is the power source of the launch vehicle, and the safety of the power transmission and transformation system is the basis of reliable and stable operation of the launch vehicle's power grid, which is of great significance to mobile warfare. Effective and accurate evaluation, diagnosis, and prediction of equipment status can significantly improve the reliability of power supply and the intelligent level of power grid operation [1].

The research on condition monitoring, evaluation, and fault diagnosis technology of high-voltage power equipment was carried out earlier abroad [2, 3]. As early as 1951, engineers from Westinghouse Electric Corporation monitored and diagnosed the motor damage caused by electric discharge in normal operation [4, 5].

Before the 1970s, developed countries such as the former Soviet Union, Japan, the United States, Germany, and Canada made more explorations in live and online monitoring of power transmission and transformation systems. First, they opened up the research field of online monitoring technology and developed partial discharges of dissolved gases in transformer oil, transformers, and gas-insulated switchgear [6]. After the 1990s, equipment condition monitoring and diagnosis technology has developed rapidly, and measurement methods have been continuously improved with the development of sensors, computers, network communication, and other technologies. Monitoring objects have gradually expanded from substation equipment to transmission equipment, and condition information has become increasingly rich. There are also some other test instruments that reflect the equipment condition through

nonelectric quantity measurement, such as ultra-high-frequency partial discharge detection, gas chromatography sensor, optical fiber temperature online measurement, infrared equipment, ultrasonic equipment, etc.

The research on condition monitoring and evaluation of power system equipment in China began in 1970s–1980s [7]. Since the 1980s, the research of online monitoring technology has laid the foundation of development status assessment technology in China. In the past 10 years, live detection and online monitoring systems for primary equipment have been widely used in China. Particularly, with the construction and development of smart grid, online monitoring technology has been rapidly popularized and applied [8, 9]. In recent years, China's power grid companies have made a lot of explorations and attempts in the field of equipment operation and maintenance, gradually realizing the important value of accurately grasping equipment status information, and have begun to promote maintenance management strategies based on status evaluation. In recent years, with the rapid development of sensor technology, countries such as Europe, America, Australia, and Japan have significantly accelerated the research and application of intelligent diagnostic devices. Large foreign companies, especially European companies such as SIEMENS, ABB, Alston, AGE, etc., use online detection systems in high-voltage circuit breakers and GIS. In modern technology, there is still a hard connection between the compartment control cabinet and the primary components of the GIS [10]. In order to overcome this shortcoming, ABB has developed a serial fiber optic bus system, eliminating hard-wired cables and developing the third-generation secondary intelligent technology [11]. ABB adopts intelligent sensor technology and microprocessing technology in the equipment it develops and realizes online monitoring, diagnosis, process monitoring, and in-station computer monitoring of the equipment through digital communication [12].

At present, the widely used state evaluation methods of power transmission and transformation systems in Chinese power grid companies include equipment state scoring system method, expert system method, multidimensional equipment state evaluation method based on traditional machine learning, and sample training method introducing remote expert opinions. However, with the development of intelligent monitoring equipment in recent years, the amount of state parameter data of power transmission and transformation system has increased exponentially; the equipment status data comes from a number of different systems. Traditional state evaluation methods cannot deal with this kind of multisource heterogeneous massive data.

Firstly, this paper analyzes the shortcomings of traditional fault diagnosis methods for power transmission and transformation equipment, including the shortcomings of setting model parameters, being difficult to change after model training and forming, and some relationships being unable to be expressed by equations. A fault diagnosis method of power transmission and transformation system based on big data is proposed. The improved LSTM-K-means algorithm based on silhouette coefficient (SC) is used for fault classification, the a priori correlation algorithm is

combined with TA coefficient to obtain the strong-weak relationship between state parameters, and Pearson's coefficient is used to construct the fault diagnosis matrix. Finally, the feasibility and accuracy of the proposed method are verified by a fault example of 400 kV oil-immersed transformer bushing in combat unit launch vehicle in the recent ten years.

## 2. State Evaluation Methods of the Traditional Power Transmission and Transformation System

With the development of artificial intelligence algorithms such as neural networks, an equipment state evaluation method based on machine learning has been developed, as shown in Figure 1. This kind of method is usually based on limited sample training and adopts a certain mathematical modeling method to construct a predictable relationship between input and output. Compared with the traditional scoring system, the state evaluation method based on machine learning can use more time section data for sample training, and the prediction results obtained are more accurate than the traditional scoring system. In addition, this method can use complex physical and mathematical functions to model. Usually used mathematical methods include artificial neural network, Bayesian network, support vector machine, and Markov [13].

However, the data that can be used by machine learning method is still limited, and it is still difficult to consider the influence of external factors such as meteorological environment on equipment status in the modeling process of such functions. The model is solidified after setting and training. Unless it is modified or trained again, it cannot adapt to various changes in the process of equipment operation and maintenance, nor can it reflect the influence of differences between different equipment models and different operating environments on diagnosis results. Because the core problem of machine learning method is the selection of historical samples and the training of samples, there are limitations of training speed, training convergence, local minima, and other problems in practical application. Some improvement measures are usually adopted to solve the above problems, such as preimprovement algorithm, postimprovement algorithm, and so on.

On the basis of machine learning method and the concept of expert system, a diagnosis system based on remote expert intervention is developed. The system can train the system with expert opinions as new samples, which integrates the advantages of machine learning method and expert system method, and can improve the accuracy and reliability of subsequent diagnosis results. However, the introduction of remote expert opinions cannot solve the problem of model solidification.

The rise and development of big data mining analysis method has opened up a brand-new technical route for state evaluation and fault diagnosis of power equipment in weapons and equipment and put forward higher requirements for existing equipment state monitoring parameters.

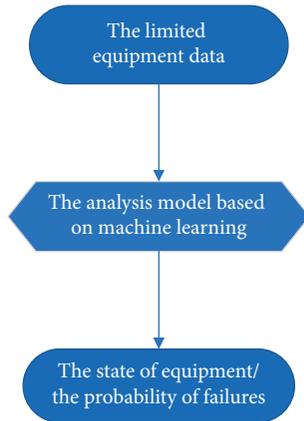


FIGURE 1: Equipment fault diagnosis method based on machine learning.

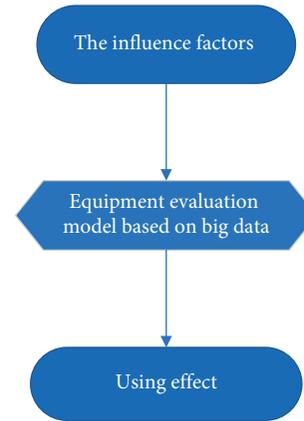


FIGURE 2: Diagram of equipment evaluating model based on big data.

This method introduces the theories and tools of mathematical statistics and pattern recognition. On the basis of large-scale data analysis, it focuses on mining the correlation between analyzed factors under uncertain model conditions. The equipment status evaluation model based on big data is shown in Figure 2. This method adopts the idea of big data mining, focusing on mining and investigating equipment defects and the correlation degree between fault state results and equipment state parameters.

Compared with traditional methods, the most fundamental difference between power equipment condition assessment and fault diagnosis methods based on big data mining analysis is as follows:

- (1) In the evaluation model of traditional methods, the equipment condition monitoring quantity is the input parameter, while the equipment defects and faults are the output parameters; in the analysis method of big data mining, equipment condition monitoring quantity and equipment defects and faults are all input parameters, while output quantity is the association rules, association degrees, and other elements among all input parameters. The traditional model formed by input and output training, once generated, cannot be changed unless retrained; the big data mining analysis method model uses dynamic correlation coefficient matrix to model the correlation between equipment condition index and equipment condition monitoring parameters, which can be continuously regressed and revised and flexibly changed according to the studied equipment objects, state parameters, fault types, etc., without reconstructing the model and without the problem of model solidification. The condition evaluation method of power transmission and transformation equipment based on big data mining is suitable for evaluating and predicting any parameter index of equipment, including equipment health and load capacity.
- (2) In the traditional methods, the evaluation model is the most critical; it is impossible to express the results

other than the preset logical relationship in the evaluation model between the input and output; it is difficult to properly reflect the personalized and differentiated elements such as equipment manufacturers, meteorological environment, habits of operation and maintenance personnel, etc.; however, the correlation mining analysis method based on big data analysis is different from the traditional method. On the basis of higher requirements for existing equipment condition monitoring parameters, it takes massive data as mining objects and uses data mining methods suitable for big data to mine the correlation between factors to be analyzed in uncertain models. In the data mining analysis method, the most important thing is the effective integration and fusion of massive data with multiple sources, multiple time scales, and multiple space-time dimensions, so as to find the inherent (known or hidden) correlation between various equipment condition monitoring quantities and equipment defects and faults, even if it is difficult to integrate the correlation of physical and logical models.

### 3. Key Technology of System Condition Evaluation Based on Big Data Analysis

#### 3.1. The Clustering Model

**3.1.1. K-Means Algorithm.** K-means algorithm is proposed by Ding and He [14], which can divide the data into  $k$  clusters that minimize the sum of squares of errors through continuous iterative calculation. The algorithm is widely used in all walks of life because of its simple and efficient operation, strong scalability, nearly linear time complexity, and suitability for processing large data sets [15]. The implementation steps of K-means clustering algorithm are as follows [16–18]:

- (1) Initializing the clustering center: randomly selecting  $K$  sample points from  $N$  sample data as the initial clustering center.

- (2) Cluster division: calculating the distance from the remaining sample points of  $K$  initial cluster centers to each initial cluster center and dividing the sample points into the cluster with the smallest distance.
- (3) Calculating a new clustering center: calculating the sum of the divided cluster sample points in step (2), completing the average value calculation, and taking the calculated average value as the new clustering center.
- (4) Convergence judgment:  $E$  function is usually used as judgment function, where  $E$  function is the sum of squares of errors between sample data and each clustering center. Steps (2) and (3) are circulated until the end of the division cycle that minimizes the  $E$  value, which is the best clustering result.

**3.1.2. SC-LSTM-K-Means Clustering Model.** Based on the long short-term memory (LSTM) network's strong non-linear deep learning capabilities [19] and the advantages of the  $K$ -means clustering algorithm, this paper proposes a hybrid clustering model combining LSTM and  $K$ -means to identify fault types in power transmission and transformation systems. The flowchart of the LSTM- $K$ -means hybrid model is shown in Figure 3.

In LSTM- $K$ -means clustering, the selection of  $K$  value of cluster number is very important. Only by finding out the appropriate  $K$  value can we get ideal clustering effect. Silhouette coefficient can solve this problem well. In this paper, by introducing SC, the two concepts of cohesion and separation are integrated, and it is more effective to evaluate the clustering effect by SC.

For the measurement of cohesion in a cluster, the way is to calculate the average distance between the  $f$ th element in the cluster and other elements in the cluster, which is denoted as  $a_i$ . For the quantification of the separation degree between clusters, the way is to select a cluster  $b$  other than the cluster of the above elements, calculate the average value of the distance between the element and all elements in  $b$ , then calculate the distance between the above elements and all elements in the other clusters, and find the minimum value of the distance between the element and other clusters, which is recorded as  $b_i$ . The equation for calculating the  $SC_i$  of the  $i$ th element is

$$SC_i = \frac{b_i - a_i}{\max(b_i, a_i)}. \quad (1)$$

Finally, the silhouette coefficients of all elements in all clusters are calculated, and the average value of the silhouette coefficients of each element is obtained as the overall silhouette coefficients of the current cluster.

**3.2. Correlation Analysis Algorithm of State Parameters.** As far as the current situation is concerned, there are many and complicated equipment state parameters, and there is a lack of mining and analysis of the relationship between equipment parameters, which leads to the lack of systematic

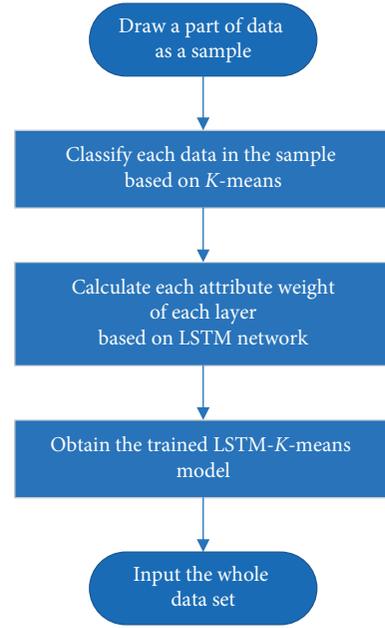


FIGURE 3: The flowchart of LSTM- $K$ -means model.

understanding of equipment parameters. Through the analysis of state parameter association rules, the effective combination of multiple state parameters of equipment, the extraction and merging of feature quantities, and the analysis of the mutual influence degree of state parameters can be realized. For association rules, the general form is the implication of  $X \geq Y$ , which can be understood as “if  $X$ , then  $Y$ ,” where  $X$  is the equipment state parameter in the preceding item, which can be a single state parameter or a set of multiple state parameters, and  $Y$  is the state parameter in the following item, which is generally a single state parameter.

Taking a priori algorithm used in this paper as an example, set the library to be mined of association rules be  $D$ , which is a collection of transactions  $T$ . If there are  $n$  transactions,  $D = \{T_1, T_2, \dots, T_j, \dots, T_n\}$ , for each transaction, consists of “ $m$  items,”  $T = \{I_1, I_2, \dots, I_l, \dots, I_m\}$ .

For item set  $X$ , the degree of support  $S$  is defined as

$$S(X) = \frac{N(X \subseteq T)}{|D|}. \quad (2)$$

For the association rules with  $X \geq Y$   $x > y$ , the degree of support is

$$S(X \longrightarrow Y) = \frac{N(X \cup Y)}{|D|}, \quad (3)$$

where  $S$  is the degree of support and  $N(X \subseteq T)$  is the number of sets.

The degree of support described in equation (3) reflects the probability of simultaneous occurrence of these two item sets. The support degree is equal to that of frequent sets.

Similarly, for association rules with  $X \geq Y$ , its credibility  $C$  is

$$C(X \longrightarrow Y) = \frac{S(X \longrightarrow Y)}{S(X)}. \quad (4)$$

The reliability described in equation (4) reflects the probability that if the item set contains  $X$ , it also contains  $Y$ . For users who use association rules, users can mine association rules with higher  $S$  and  $C$  by defining thresholds of minimum support and credibility.

**3.3. Fault Diagnosis Based on Correlation Matrix.** In order to diagnose the fault mode, it is necessary to consider the correlation between each state parameter and each fault mode, that is, the possibility of a certain fault mode when a certain state parameter is abnormal. After obtaining the correlation coefficient between each state parameter and each fault mode of the equipment, the equipment fault mode diagnosis matrix  $R$  can be obtained; that is,

$$R = \begin{bmatrix} R_{11} & \cdots & R_{1t} & \cdots & R_{1q} \\ \vdots & & \vdots & & \vdots \\ R_{r1} & \cdots & R_{rt} & \cdots & R_{rq} \\ \vdots & & \vdots & & \vdots \\ R_{p1} & \cdots & R_{pt} & \cdots & R_{pq} \end{bmatrix}, \quad (5)$$

where  $R_{rt}$  is the correlation coefficient of vector  $B_r$  of the  $r$ th failure mode under vector  $V_t$  of the  $t$ th state parameter; among them,  $r \in [1, p]$ , there are  $p$  fault modes;  $t \in [1, q]$ , and there are  $q$  state parameters. When calculating  $R_{kt}$ , there are many sets of data for each failure mode and state parameter, so  $B_r$  and  $V_t$  are vectors.

When calculating the correlation coefficient  $R_{rt}$ , Pearson's coefficient is used in this paper. The correlation coefficient is based on the deviation between the two variables and the average value of their respective variables, calculated by the product-difference method, multiplied by the two deviations, and the product is used to reflect the correlation degree between the two variables [20–22]. Pearson's coefficient ranges from  $-1$  to  $1$ . A value of  $0$  indicates that there is no significant linear relationship between the two variables.  $-1$  and  $1$  indicate that the two variables are completely negative or positive. The following equation shows the correlation coefficient of  $B_r$  and  $V_t$  applied to equipment fault diagnosis:

$$R_{rt}(B_r, V_t) = \frac{\text{Cov}(B_r, V_t)}{\sqrt{\text{Var}(B_r)} \cdot \sqrt{\text{Var}(V_t)}}, \quad (6)$$

where  $\text{Cov}(B_r, V_t)$  is the covariance of  $B_r$  and  $V_t$ ;  $\text{Var}(B_r)$  is the variance of  $B_r$ ; and  $\text{Var}(V_t)$  is the variance of  $V_t$ . After the diagnosis matrix  $R$  is obtained by the above method, the fault data can be diagnosed by the following equation:

$$F = R \cdot U. \quad (7)$$

In equation (7),

$$\begin{aligned} U &= [U_1, U_2, \dots, U_t, \dots, U_q]^T, \\ F &= [F_1, F_2, \dots, F_r, \dots, F_p]^T, \end{aligned} \quad (8)$$

where  $U$  is the data vector of the fault case to be diagnosed, including the state parameter level of each state parameter,

and  $F$  is a fault mode diagnosis result vector, and the value of each element in the vector can indicate the membership degree of the fault case under each fault mode. When the most likely failure mode is finally diagnosed, the failure mode with the largest membership degree (the largest value) can be selected as the final result.

## 4. Simulation Analysis of Fault Diagnosis Based on Big Data

In this paper, the fault cases of 400 kV oil-immersed transformer bushing in a combat unit launch vehicle in the recent ten years are taken as data mining objects, and the equipment fault diagnosis based on big data mining is studied.

**4.1. Preprocessing Data.** Firstly, the abnormal state data of the equipment to be mined are collected, with emphasis on the case data of faults and defects. The case code is represented by  $A_s$  ( $s = 1, 2, \dots, 34$ ) and the state parameter is represented by  $W_u$  ( $u = 1, 2, \dots, 26$ ). According to the representation of faults and defects, the state parameters are assigned. Since the construction of knowledge map is only to mine and analyze the state parameters or equipment abnormal cases themselves, only one state parameter needs to be known whether it is abnormal or not and does not involve the equipment state level or the deterioration degree of the state parameters, so only binary quantification is carried out. According to the severity, defects can be divided into emergency defects, major defects, and general defects.

When a certain state parameter is abnormal, the value of  $W_u$  is  $1$ , which means that the state parameter has faults or major or urgent defects. The value of  $W_u$  is  $0$ , which means that the state parameter is normal. The parameters of the state parameters considered in this paper are shown in Table 1. Through summary statistics, there are 34 groups of equipment failure and defect cases, of which 22 groups are failure data.

### 4.2. Clustering Analysis of Failure Cases

**4.2.1. Clustering Analysis of SC-LSTM-K-Means.** Firstly, 22 groups of fault cases in the original data are clustered hierarchically, and all faults can be divided into 3–8 categories by preliminary analysis.  $K$  is selected as 3, 4, 5, 6, 7, and 8, respectively, and the clustering effects are compared by calculating the silhouette coefficients of different cluster numbers  $K$ . The results are shown in Figure 4. As can be seen from Figure 4,  $k = 5$ , the clustering result is the most ideal when the failure modes are divided into 5 categories. Combined with the fault physical background and expert experience of transformer bushing, five common fault modes of transformer bushing can be summarized. Using LSTM-K-means to cluster fault cases, the results are shown in Table 2.

**4.2.2. Comparison of Three Models' Clustering Accuracy.** In recent years, some scholars have studied the combined clustering model of BP neural network and K-means

TABLE 1: Names of status parameters.

State parameter name	Symbol
Porcelain set of creepage	W1
External insulation configuration	W2
Porcelain insulation damage	W3
Oil level indicator	W4
Oil leakage check	W5
Casing grounding	W6
Lead wire at the end of the screen	W7
Infrared temperature measurement	W8
The main insulation dielectric loss and capacitance	W9
Dielectric loss at the end of screen and capacitance	W10
The main insulation resistance	W11
Insulation resistance at the end of screen	W12
Dissolved gas analysis in oil	W13
Outside the insulating oil	W14
Oil dielectric loss	W15
Insulating oil water-soluble acid	W16
Insulation oleic acid value	W17
Insulating oil flash point	W18
Insulating oil moisture	W19
Insulating oil interfacial tension (below 25 degrees Celsius)	W20
Insulating oil breakdown voltage	W21
Insulating oil volume resistance (below 90 degrees Celsius)	W22
Insulating oil sludge and sediment	W23
Insulating oil corrosive sulfur	W24
The carbamate insulating oil	W25
Insulating oil charged tendency	W26

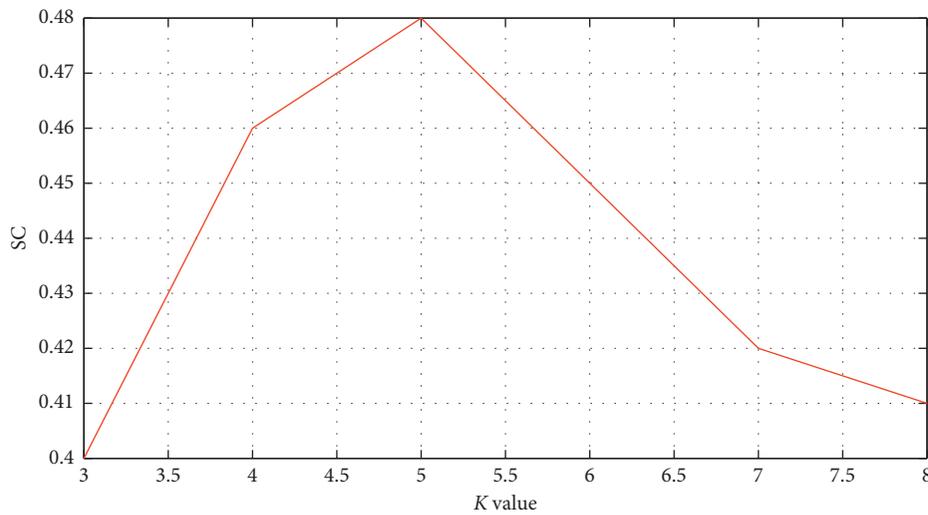


FIGURE 4: Silhouette coefficient under different K values.

TABLE 2: Clustering results of fault data ( $k = 5$ ).

The serial number	Case	Fault name
Failure 1	A4, A6, A11, A12, A19	External insulation affected with damp
Failure 2	A3, A10, A13, A22	Serious oil leakage
Failure 3	A7, A9, A15, A16, A18	Poor contact
Failure 4	A2, A14, A20	Insulating oil affected with damp
Failure 5	A1, A5, A8, A17, A21	Discharge at the end of screen



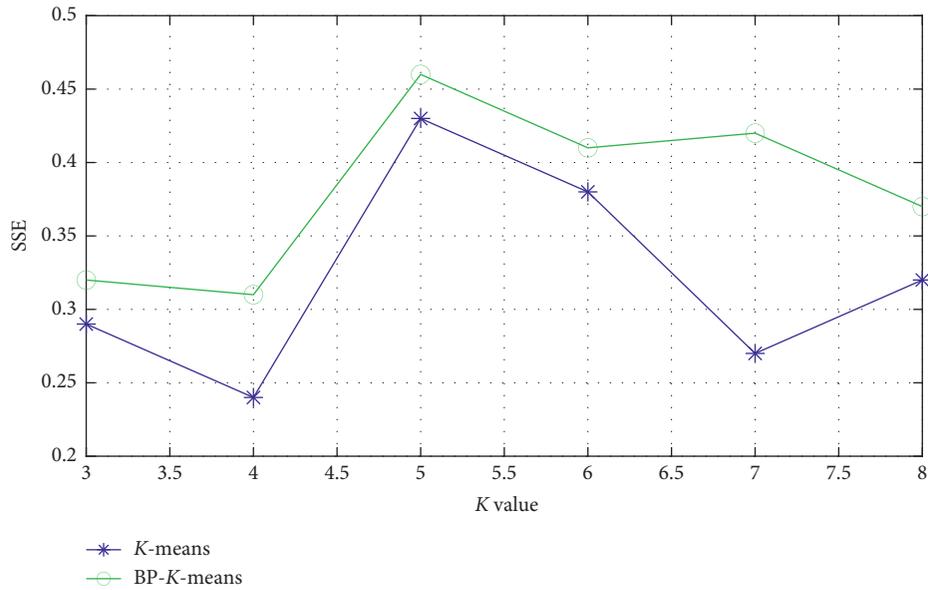


FIGURE 5: K values of K-means and BP-K-means.

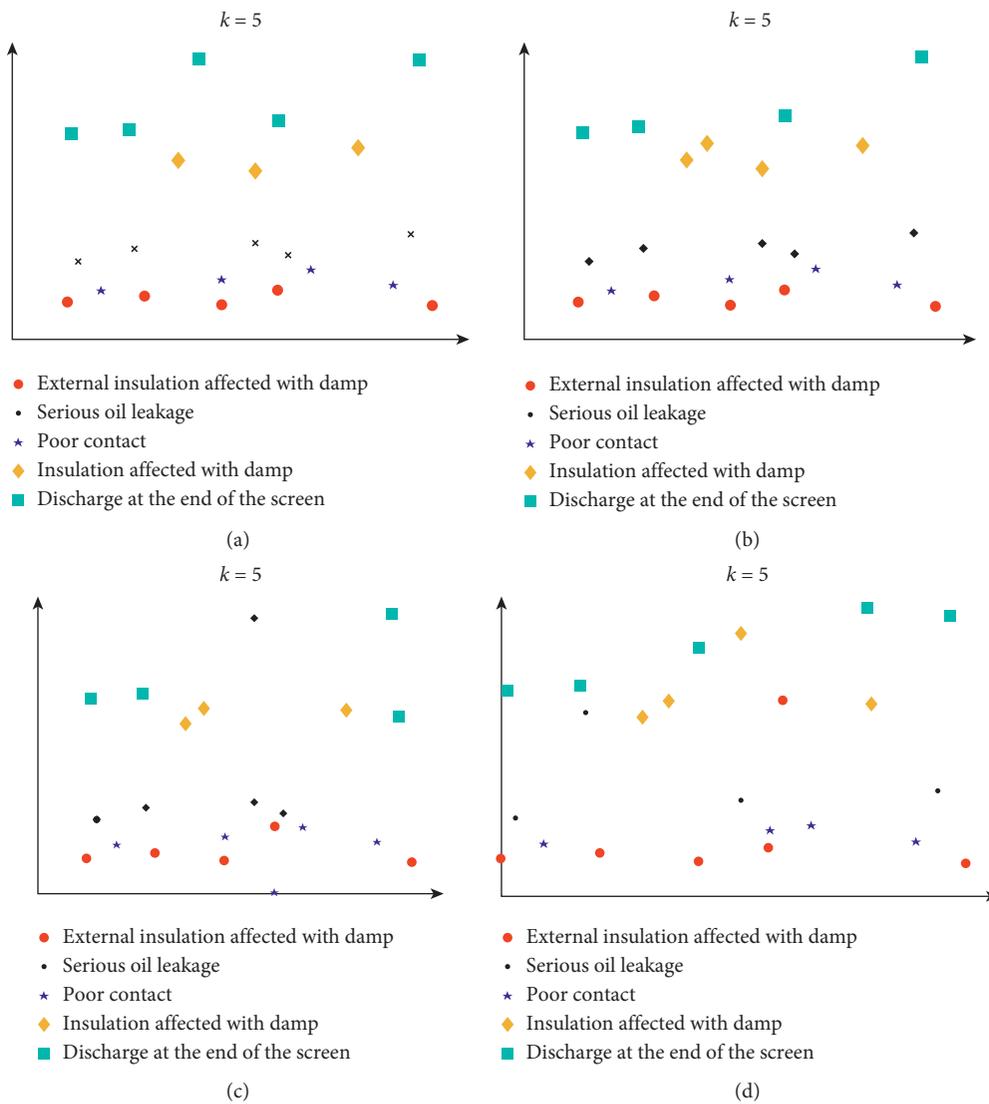


FIGURE 6: Contrast of the three models' clustering results. (a) SC-LSTM-K-means. (b) Real result. (c) K-means. (d) BP-K-means.

TABLE 3: Contrast of the three clustering models.

Models	Indicators		
	Number of correct cases	Number of wrong cases	Accuracy (%)
SC-LSTM-K-means	21	1	95.45
K-means	18	4	81.81
BP-K-means	19	3	86.37

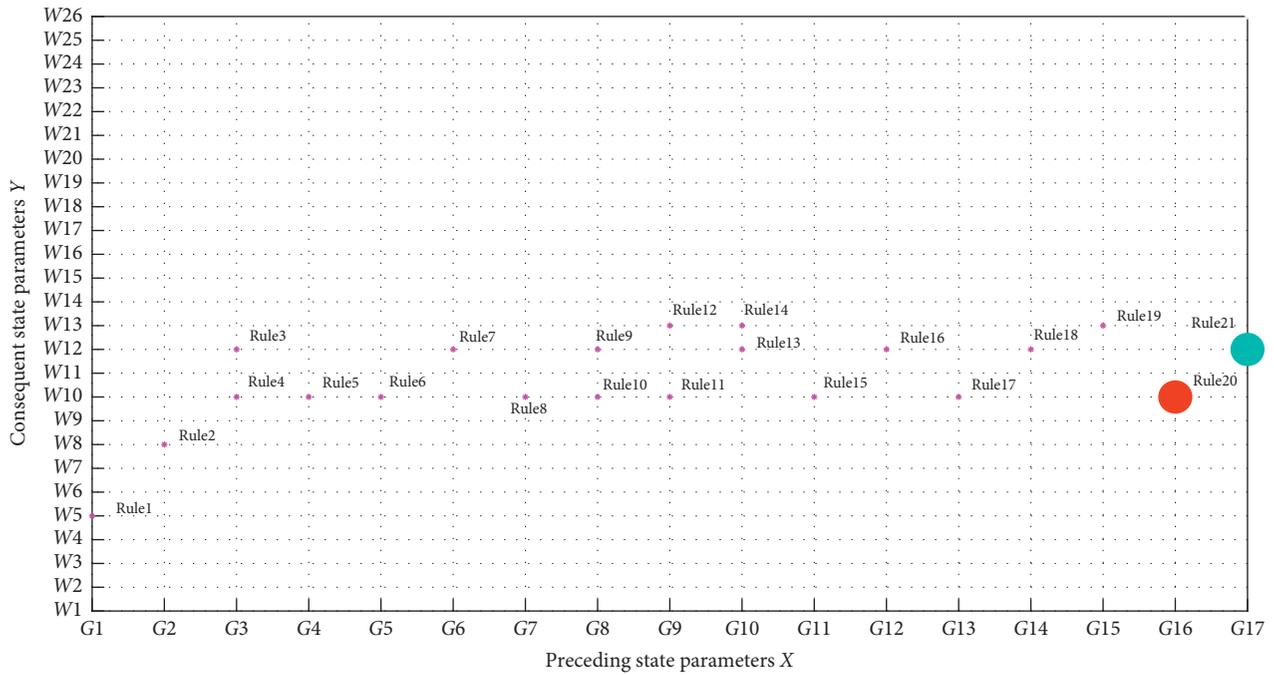


FIGURE 7: Diagram of fault relationship in bushing equipment of oil-immersed transformer.

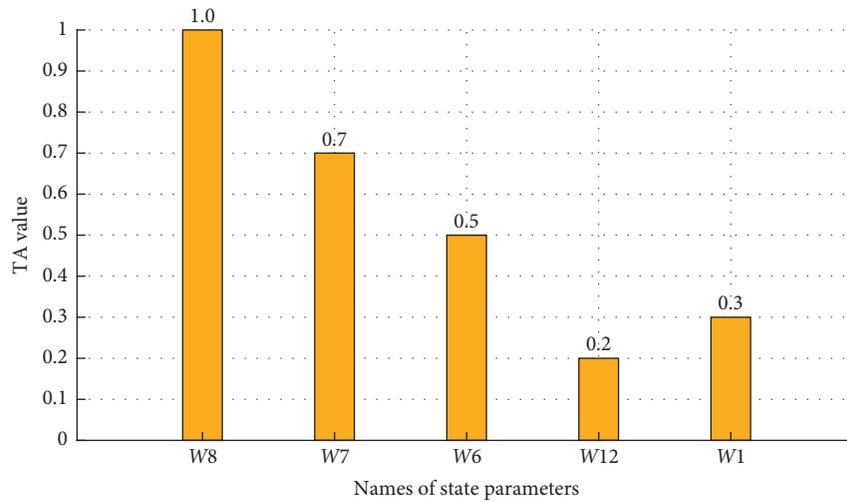


FIGURE 8: TA value of status under poor connection fault.

TABLE 4: Result of fault diagnosis.

Fault name	Fr (case 1)	Fr (case 2)
External insulation affected with damp	-3.1489	-3.0251
Serious oil leakage	2.3013	<b>4.8892</b>
Poor contact	<b>5.1630</b>	-0.7086
Insulating oil affected with damp	-1.3265	-2.2365
Discharge at the end of screen	-0.7208	-0.3216

Calculate the correlation coefficient of each state parameter in the terminal screen discharge fault mode; for example, the vector  $V_{13}$  of the state parameter for analyzing the dissolved gas in the oil is

$$V_{13} = [4, 0, 4, 0, 0, 0, 0, 2, 0, 0, 0, 0, 0, 4, 0, 2, 0, 0, 4, 0, 0]. \quad (10)$$

Therefore, the factor  $R_{5\&13}$  in the diagnosis matrix is calculated by

$$R_{5\&13} = \frac{\text{Cov}(S_5, V_{13})}{\sqrt{\text{Var}(S_5)} \cdot \sqrt{\text{Var}(V_{13})}}. \quad (11)$$

Similarly, the remaining elements in the diagnostic matrix  $R$  of equation (5) can be calculated. The condition evaluation results of oil-immersed transformer bushing equipment twice are taken as cases for example verification. In case 1, the parameters of casing wiring, terminal screen outgoing line, and infrared temperature measurement state are abnormal; in case 2, the porcelain insulation was damaged, and the parameters of oil level indication and oil leakage inspection state were abnormal. The evaluation of the state parameters is quantified by the deterioration level, and the vectors  $U$  of the two cases to be diagnosed are shown in the following equation:

$$U = \begin{cases} [0, 0, 0, 0, 0, 3, 4, 0, 0, 0, 0, 0, 0, 0], & \text{(For case 1),} \\ [0, 0, 2, 3, 3, 0, 0, 0, 0, 0, 0, 0, 0, 0], & \text{(For case 2).} \end{cases} \quad (12)$$

After substituting equation (7), the fault mode diagnosis result vector  $F$  of the two diagnosis cases is shown in Table 4.

If the fault mode with the highest diagnosis membership value is taken as the diagnosis result, the diagnosis result of the sample of diagnosis case 1 is poor contact; the diagnosis result of diagnostic case 2 is serious oil leakage. After on-site fault diagnosis by relevant operation and maintenance personnel, the former is an abnormal situation caused by the thread of the casing end screen not tightening, while the latter is an abnormal situation caused by the failure to observe the oil level through the oil level mirror. After inspection, it is caused by the failure to replenish oil for a long time and the normal aging and oil leakage of the casing.

## 5. Conclusion

Firstly, the SC-LSTM- $K$ -means clustering algorithm can be used to mine the fault modes of power transmission and transformation systems of launch vehicles, and the number of fault classifications can be determined through the SC. By

comparing the clustering effects with  $K$ -means and BP- $K$ -means algorithms, it can be concluded that the clustering accuracy of this model reaches 95.45%. Secondly, a priori algorithm based on Boolean association rules to mine frequent item sets can be used to mine the internal correlation relationship of characteristic parameters of power transmission and transformation system, so as to realize the effective combination of multiple state parameters of equipment, feature extraction and merging, and analysis of the mutual influence degree of state parameters; the obtained TA coefficient can characterize the strong-weak relationship of correlation. Finally, the equipment fault correlation matrix based on Pearson's coefficient can get that the maximum value of Fr for case 1 is 5.1630, and the maximum value of Fr for case 2 is 4.8892. The method accurately analyzes and diagnoses the equipment fault mode which is consistent with the actual operation and maintenance test results.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest regarding the publication of this paper.

## References

- [1] C. Li, Z. Tang, and Z. Zheng, "Recent progress and future perspective on condition monitoring and diagnostic techniques for power equipment in China," in *Proceedings of the International Conference on Condition Monitoring & Diagnosis*, no. 1, pp. 1-6, IEEE, Perth, Australia, September 2018.
- [2] Z. Liao, F. S. Wen, W. X. Guo et al., "An analytic model and optimization technique based methods for fault diagnosis in power systems," in *Proceedings of the International Conference on Deregulation and Restructuring and Power Technologies*, no. 13, pp. 1388-1393, IEEE, Weihai, China, 2018.
- [3] R. J. Song and X. D. Wang, "Research on assessment and analysis system of condition-based maintenance for power transmission and transformation equipment," *Power System Protection and Control*, vol. 36, no. 9, pp. 54-57, 2018.

- [4] A. Contin, A. Cavallini, G. C. Montanari et al., "Artificial intelligence methodology for separation and classification of partial discharge signals," in *Proceedings of the Conference on Electrical Insulation & Dielectric Phenomena*, vol. 2, no. 09, pp. 522–526, Toronto, Canada, October 2016.
- [5] Y. Jiang and Y. J. Zhou, "Application of distributed partial discharge online monitoring technology in Shanghai 500 kV XLPE power cable line," *High Voltage Engineering*, vol. 41, no. 04, pp. 1249–1256, 2015.
- [6] M. Ge, R. X. Du, G. C. Zhang et al., "Fault diagnosis using support vector machine with an application in sheet metal stamping operations," *Mechanical Systems and Signal Processing*, vol. 18, no. 1, pp. 143–159, 2019.
- [7] J. M. Babita, J. Amit, and B. Srinivasm, "A Web based expert system shell for fault diagnosis and control of power system equipment," in *Proceedings of 2008 International Conference on Condition Monitoring and Diagnosis*, no. 21, pp. 1310–1313, IEEE, Beijing, China, April 2008.
- [8] W. G. Chen, J. Liu, and M. Cao, "Diagnosis method of internal fault for transformers based on information fusion," *High Voltage Engineering*, vol. 41, no. 11, pp. 3797–3803, 2015.
- [9] J. Li, J. Zhang, C. Wang et al., "Research on integrated smart monitoring devices for internet of power transformation equipment," *High Voltage Engineering*, vol. 41, no. 12, pp. 3881–3887, 2015.
- [10] Y. He, "Research on the key technology of power transmission and transformation equipment condition monitoring based on the Internet of Things," Thesis, Hunan University, Hunan, China, 2014.
- [11] L. Qin, "Research and application of fault diagnosis method for power transmission and transformation equipment based on online monitoring," Thesis, North China Electric Power University, Beijing, China, 2014.
- [12] M. M. Lu, "Research on fault rate model and fault diagnosis of oil-immersed transformer," Thesis, Zhejiang University, Zhejiang, China, 2014.
- [13] Z. L. Wu, "Research oil fault diagnosing methods for power transformers," Thesis, North China Electric Power University, Beijing, China, 2018.
- [14] C. Ding and X. He, "K-means clustering via principal component analysis," in *Proceedings of the Twenty-First International Conference on Machine Learning ACM*, no. 3, pp. 15–21, Da Lat, Vietnam, January 2019.
- [15] L. Han, Y. Qiao, and H. T. Jing, "Wind power interval assessment based on error classification," *Automation of Electric Power Systems*, vol. 45, no. 1, pp. 97–104, 2021.
- [16] L. H. Zhang, "Research and application of K-means clustering algorithm," Thesis, Wuhan University of Technology, Wuhan China, 2017.
- [17] S. K. Sun, C. Dong, Z. Wang et al., "Wind power medium and long-term electricity forecasting methods considering different wind energy characteristics," *High Voltage Technology*, vol. 47, no. 4, pp. 1224–1233, 2021.
- [18] G. Q. Weng, Y. G. Gong, J. P. Shu et al., "Active distribution network power quality prediction based on clustering LSTM deep learning model," *High Technology News*, vol. 30, no. 7, pp. 687–697, 2020.
- [19] R. Sibson, "SLINK: an optimally efficient algorithm for the single link cluster method," *Computer Journal*, vol. 16, no. 1, pp. 30–34, 2010.
- [20] X. M. Li and W. H. Jiang, "Research on linear correlation and dependence measure," *Mathematics in Practice & Theory*, vol. 36, no. 2, pp. 188–192, 2016.
- [21] M. W. Lin, C. Huang, R. Q. Chen et al., "Directional correlation coefficient measures for Pythagorean fuzzy sets: their applications to medical diagnosis and cluster analysis," *Complex & Intelligent Systems*, preprint, 2021.
- [22] I. H. Hariyanto, D. G. Pratomo, and M. R. Darminto, "The correlation of tsunami heights and coastal conditions in Palu Bay using the contingency coefficient analysis," *IOP Conference Series: Materials Science and Engineering*, vol. 1052, no. 1, 2021.