Intelligent Distribution Network Operation and Anomaly Detection Based on Information Technology

Qingjiang Huang [*1], Hui Xian [1], Linchang Mei [1], Xiawei Cheng [1], Nuannuan Li [2]
[1.] State Grade Hebi Electric Power Supply Company, 458030, China
[2.] State Grade Henan Electric Power Research Institute, 450000, China
E-mail: qingjianghuang8@163.com
[*] Corresponding author

Keywords: intelligent distribution network, abnormal detection, load mode, data analytics, fault prediction

Received: December 25, 2023

In response to the current challenges of limited monitoring methods and underutilization of data in distribution networks, this paper proposes a study on intelligent distribution network operation and anomaly detection using information technology. This article begins by analyzing the technical characteristics and data flow within the current power supply enterprise distribution network scheduling support system. It selects extensive historical telemetry data as the subject of research and employs the C-means fuzzy clustering algorithm to identify distribution line load patterns and conduct load prediction. A mismatch degree index, taking into account membership degree and Euclidean distance factors, serves as the criterion for assessing line faults. A 12 kV distribution line (L) in a certain area was chosen and tested using real operational data to see if the proposed big data-based method for monitoring distribution network faults works. The test results show that this study can correctly find patterns in distribution line loads, reliably find line faults, and, to some extent, lessen the problems that come up when loads change normally. The application results reveal that this study offers a straightforward and practical monitoring approach that effectively assesses the fault status of distribution networks that cannot be monitored using existing methods.

Povzetek: Predstavljena so inteligentna distribucijska omrežja in zaznavanje anomalij z uporabo informacijske tehnologije ter algoritmov za razvrščanje, kar izboljšuje zanesljivost in učinkovitost omrežij s pomočjo analize velikih podatkov.

1 Introduction

Smart Distribution Automation is the embodiment of Smart Grid in the distribution network. Compared with traditional distribution networks, it has higher safety, higher power quality, and fault analysis and diagnosis capabilities. With the advancement of intelligent distribution network construction, the level of distribution network management has improved, but the dispatch, operation, and maintenance management of the distribution network still needs further improvement. On the one hand, the distribution network lacks data from some links, and on the other hand, there is a problem of large data volume and difficulty in analysis. Specifically, it includes:

1.1 The grid scheduling system is unable to obtain full data on the distribution network

The distribution network is divided into three categories based on voltage levels: high-voltage distribution network (35kV, 63kV, 110kV), medium-voltage distribution network (10kV), and low-voltage distribution network (220V, 380V). The main network information is collected by EMS, and the distribution network scheduling system only monitors the data above the high-voltage side of the substation area, lacking monitoring data for the low-voltage distribution network. In addition, the wide range of low-voltage users, unclear equipment investment property rights, and inconvenient installation and monitoring have all caused a lack of data in the distribution network.

1.2 Low efficiency in handling faults in distribution networks

The alarm monitoring platform of the intelligent dispatch technology support system for distribution networks can judge the faults and alarms generated in the system and present them to the staff in a real-time alarm manner. However, with the addition and expansion of distribution network equipment, the distribution network has become increasingly complex, and the number and types of alarms on the platform have also become more complex. Ordinary monitoring and maintenance personnel are faced with such a large number of complex types of faults and alarms every day, and their hearts have become spare but their abilities are insufficient. The current alarm monitoring platform mainly has the following shortcomings:

i. There is various interference information in the alarm information, such as frequent signals from the distribution network, debugging signals, and nonmonitoring key signals. Additionally, a fault may result in a large number of alarm events, many redundant alarms, and some information missing from these alarm events. Disturbance alarm information makes clear identification of faults more difficult, increasing the work pressure on maintenance personnel.

ii. A large amount of interference information has a significant impact on data mining and the analysis of equipment faults.

iii. The current distribution network is still at a level where monitoring personnel rely entirely on experience to diagnose and handle faults, resulting in low work efficiency.

These shortcomings prevent timely and effective troubleshooting of the resulting faults, seriously affecting work efficiency. Therefore, there is an urgent need for an effective method to solve these problems so that faults can be located and resolved promptly and the distribution network can operate more stably. The ability to evaluate and analyze the health status of distribution network equipment needs to be improved. The existing distribution network system cannot still evaluate and analyze the operating status of equipment and cannot timely grasp and diagnose the current operating status of equipment and predict potential risks in the future. It is necessary to construct a distribution network equipment evaluation model to evaluate and analyze the primary and secondary equipment under the jurisdiction of the distribution network to identify the weaknesses of the distribution network, help management and decision-making departments solve the problems of network structure transformation, construction investment, etc., improve auxiliary decision-making ability, better reduce the occurrence of faults, and improve the stable power supply ability of the distribution network.

In recent years, the development of information and communication technologies, as well as the emergence and progress of new technologies such as big data, business intelligence, and data mining, have provided new technological means to solve these problems. Therefore, the problem of storing massive amounts of data can be solved by integrating a time series database into the distribution network dispatch automation system. A distribution network diagnosis system software based on “big data” technology can be constructed to seek solutions to the problem of “massive data and missing information” in the distribution network [1]. Data lean management can be implemented in the distribution network field, providing data analysis and display tools for business personnel and providing data support for management decisions [2]. The rapid advancement of intelligent distribution networks has created new challenges in monitoring and maintaining power system reliability. Traditional distribution networks are often plagued by inefficiencies in fault detection and anomaly management, leading to prolonged downtimes and increased operational costs. The limitations of conventional monitoring techniques, coupled with the underutilization of vast amounts of available data, hinder the effective management of distribution networks. This research is motivated by the need to enhance the operational efficiency of distribution networks through the integration of advanced information technology and data analytics. By addressing the gaps in current fault detection methods and leveraging the potential of big data, this study aims to improve the reliability and stability of power distribution systems, ensuring a more resilient and efficient network operation. This research makes several significant contributions to the field of intelligent distribution network management. Firstly, it introduces a novel approach to anomaly detection and fault prediction by employing the C-means fuzzy clustering algorithm, which enhances the accuracy of load pattern identification and fault monitoring. The study also proposes a mismatch degree index that incorporates membership degree and Euclidean distance factors as criteria for assessing line faults, providing a more reliable method for fault detection. Furthermore, the research validates the proposed methodology using real operational data from a 12 kV distribution line, demonstrating its effectiveness in accurately identifying load patterns and reducing the challenges associated with normal load variations. The findings of this study offer a practical and scalable solution for improving fault detection and predictive maintenance in intelligent distribution networks, contributing to the overall enhancement of network reliability and operational efficiency.

2 Literature review

The State Grid Corporation of China has been working harder over the past few years to change the way distribution networks are automated. They have also been promoting and using the Distribution Automation System (DAS) more, which can do remote signaling, telemetry, and remote control (three remotes) of switches on the main and some branch lines of the distribution network. Based on the line accident signal and the protection signal of the automation switch, the fault zone is automatically determined, a prompt signal is sent to the control personnel, or the fault isolation and restoration of the power supply are automatically completed, improving the reliability and quality of the power supply.

The traditional medium voltage distribution network in China mostly adopts a radiation-type structure, and its power flow in the network is unidirectional during normal operation. The setting of protection methods is relatively simple. The integration of DG in intelligent distribution networks not only increases the complexity of the network structure but also changes the fault characteristics of the distribution network. In addition, the output power of new energy DGs has a certain degree of randomness, which leads to the problem of bidirectional, uncertain power flow in local areas during normal operation of the distribution network. This poses great challenges to the protection and control of intelligent distribution networks. In the energy and power industry, the application of big data analysis technology to fault monitoring is in its early stages and has not yet formed a widely applicable technical model. Bin *et al.* has established a substation transmission and transformation component rule network fault diagnosis model that takes into account information temporal attributes by utilizing the temporal data of remote signal displacement and telemetry changes in the centralized control center [3].

Table 1: Related work in intelligent distribution network operation and anomaly detection: methodology, algorithms, datasets, benefits, and drawbacks

References	Methodology	Algorithm Used	Datasets	Benefits	Drawbacks
[9]	Machine Learning	Random Forest	Historical Telemetry	Improved fault detection and predictive maintenance	Limited algorithm transparency and data interpretability
[10]	Data Analytics	Support Vector Machines	Real-time sensor data	Real-time anomaly detection, reduced downtime	Requires high-quality real-time data, may miss subtle anomalies
[11]	Deep Learning	Long Short-Term Memory (LSTM)	Smart Grid Data	Effective prediction of load patterns, adaptability to various data sources	High computational requirements and training time
[12]	Statistical Analysis	Bayesian Networks	Historical Outage Data	Probabilistic modeling for outage prediction, interpretability	Limited applicability to non-outage anomalies
[13]	Hybrid Approach	Decision Trees, Clustering	Sensor and SCADA Data	Improved operational insights, efficient resource allocation	Complexity in combining various techniques
[14]	Anomaly Detection	Isolation Forest, One-Class SVM	Power Grid Data	Outlier detection, robust to imbalanced datasets	Limited predictive capabilities for load forecasting
[15]	Big Data Analytics	K-means Clustering	Smart Meter Data	Scalability, data-driven insights, real-time fault detection	Challenges in handling massive datasets, privacy concerns
[16]	Neural Networks	Autoencoders	Synthetic Data	Anomaly detection in synthetic scenarios, data generation	Limited real-world application and generalization

Ehsani *et al.* introduces real-time information on wide-area synchronous electrical quantities obtained from PMU measurements into power grid fault diagnosis and proposes a fast identification method for fault components that combines switch quantities with electrical quantities [4]. Kondo *et al.* utilizes advanced IT technologies such as data mining and SOA to design and implement a power dispatch management system based on data mining, covering various data requirements for power dispatch operation and production management [5]. Huang *et al.* proposes to mine information and extract knowledge from a large amount of business data, supporting intelligent scheduling business analysis and decision-making, for statistical analysis of power grid operation, regional load trend analysis, and power grid load characteristics analysis [6].

Given this, combined with the characteristics of existing scheduling technology support system data, this paper proposes a distribution network fault monitoring technology based on big data analysis. The C-means fuzzy clustering algorithm is used to predict the line current value, and real-time sampling values are used to figure out the mismatch degree and keep an eye on distribution network faults. The proposed method was validated based on the operation data of a 10 kV line in a certain region, and the actual application effect in a power supply enterprise was demonstrated [7, 8]. Table 1, presents the related work in intelligent distribution network operation and anomaly detection: methodology, algorithms, datasets, benefits, and drawbacks. This literature table shows the main points of several study papers that have been written in the area of intelligent distribution network operation and anomaly detection. These articles talk about a lot of different methods and algorithms that use datasets to solve problems and offer benefits while also pointing out possible problems. By looking at how different algorithms and methods work on real-world datasets, we learn a lot about the methods and tools that can be used to improve the way distribution networks work and find problems.

The review examines various artificial intelligence-based anomaly detection techniques within smart grids, highlighting the advancements and challenges in integrating AI methods for improved grid reliability and security [17]. Gonaygunta *et al.* introduces a flexible deep learning model aimed at enhancing cybersecurity through improved anomaly detection, showcasing its potential in identifying and mitigating cyber threats within systems and information engineering contexts [18]. This knowledge helps shape our approach and might help us pick the best algorithms and data sources to enhance fault finding and planned maintenance, which will eventually lead to the smooth running of distribution networks. We can also learn from the problems and restrictions that have been discussed in previous studies to prepare for and deal with problems that may come up on our own. In general, this study of the literature gives us a base for our work, helping us make decisions and giving our research on intelligent distribution network operation and anomaly detection some context. The necessity for creative solutions is highlighted by the growing complexity and requirement for effective distribution networks, as well as the difficulties associated with the use of constrained monitoring techniques. Improving distribution network performance and anomaly detection is a promising field for research because it may greatly increase dependability and decrease downtime. Current distribution networks frequently have insufficient monitoring techniques and data underutilization, which results in irregularities and inefficient operations. To meet these problems, clever solutions that leverage information technology to improve anomaly detection and network performance must be developed. The goal of this research is to close the gap that exists between current technology and conventional distribution network methods. Through the use of sophisticated algorithms and datasets, our research aims to offer a useful framework for enhancing network performance and anomaly identification. Predictive maintenance, improved problem detection, and overall distribution network performance are among the benefits of this study.

3 Methods

3.1 Status monitoring and fault handling strategy and scheme design

3.1.1 Overall strategy

When the distribution network is in normal operation, the operation status of each node in the network is basically consistent. At this point, the inter-row difference of the high-dimensional spatiotemporal state monitoring matrix constructed based on the operational data of distribution network nodes is very small, and each node presents a group of clustered points in the high-dimensional space without outliers. When the distribution network is in a fault state, there is a significant difference in the operating status between the faulty nodes and the normal nodes in the network. Each node in the high-dimensional space no longer presents as a group of points gathered together. At this time, a small number of faulty nodes form outliers due to being far away from the normal nodes.

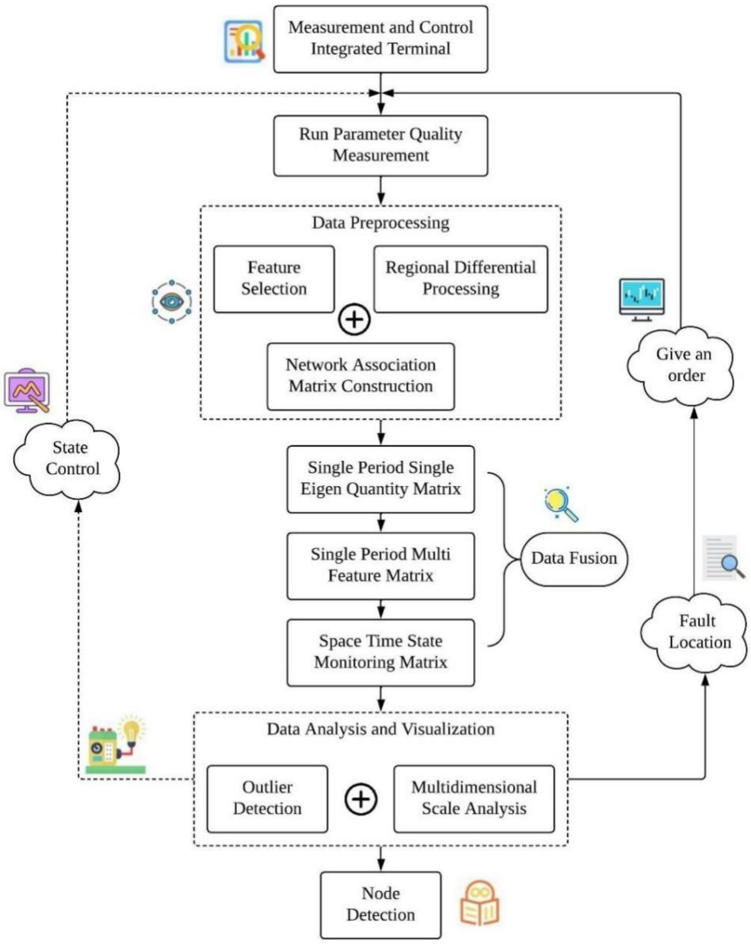

Figure 1: Flow chart of status monitoring and fault handling methods

Therefore, when monitoring the status of intelligent distribution networks, it is only necessary to detect whether there are outliers in the high-dimensional spatiotemporal state monitoring matrix to complete the monitoring of the operation status of the distribution network. Combined with the association relationship of outliers, the location of the fault point can be determined.

3.1.2 Scheme design

The processing flow of the state monitoring and fault handling method based on big data analysis of intelligent distribution networks proposed in this article is shown in Figure 1, which can be mainly divided into the following four stages:

i. ***Data preprocessing:*** The purpose of this loop program is to filter and preprocess the raw data uploaded by sensing devices to reduce the amount of data and generate the initial feature matrix required for state monitoring and fault handling in this article. This process includes feature selection, network correlation matrix construction, and regional difference processing [19].

ii. ***Data fusion:*** This link adopts the idea of data fusion. Firstly, the state monitoring matrix of a single time period and a single electrical feature constructed by multiple different electrical feature quantities are fused into a single time period and multiple electrical feature quantities state monitoring matrix in space; Afterwards, the matrix was further expanded on the time series to generate a high-dimensional spatiotemporal state monitoring matrix.

iii. ***Data analysis and visualization:*** Firstly, the multi-dimensional scale analysis algorithm is used to process the high-dimensional spatiotemporal state monitoring matrix. While keeping the relative relationship of each object basically unchanged, high-dimensional data is presented in low-dimensional space. While completing data visualization, the amount of low-value data is reduced, further achieving data fusion. Then, outlier detection is performed on the spatiotemporal state monitoring matrix after dimensionality reduction to support online identification of the operational status of the intelligent distribution network [20].

iv. ***State identification and processing:*** This process first identifies the state based on the results of the data analysis and then performs corresponding processing according to the state of the distribution network. The processing is mainly divided into two aspects: state optimization in abnormal states and fault localization and isolation in fault states.

State optimization under abnormal conditions is the correction and control of the current abnormal state of the intelligent distribution network, with the aim of preventing and reducing the possibility of faults. It belongs to the category of preventive control in self-healing control and is limited to space [21]. Therefore, no in-depth research will be conducted here. The outliers in the spatiotemporal state monitoring matrix correspond one-to-one with the fault nodes in the physical power grid, and the common area where the fault nodes are located is the fault area. Therefore, after the fault occurs, the data processing center only needs to send a trip command to the integrated measurement and control terminal in the area to complete the fault isolation [22]. It is worth noting that the high-dimensional spatiotemporal state monitoring matrix is highly sparse. If outlier detection is directly performed on it, there are problems such as high algorithm time complexity, long data processing time, and low detection accuracy. Timeliness and accuracy are crucial for intelligent distribution network status monitoring and fault handling. Therefore, before outlier detection, it is necessary to perform data dimensionality reduction on the high-dimensional spatiotemporal state monitoring matrix.

3.2 Principles of common data mining techniques

Extract useful information and patterns from a large amount of incomplete, noisy, and fuzzy data in historical databases. Considering the needs of the scheduling business, the predictive mode is mainly used.

3.2.1 Common methods for mining predictive patterns

The predictive mode takes time as the key parameter, and for time series data, it predicts its future values based on its history and current values. The commonly used methods include neural network prediction and fuzzy clustering recognition prediction.

i. ***Neural network prediction method:*** Artificial neural networks can establish any nonlinear model and are suitable for solving time-series prediction problems. However, due to the hidden patterns obtained by neural network classification methods in the network structure, they are not easily understood and explained by people. Additionally, multiple scans of training data are required, resulting in longer training times. Unable to meet the requirements of readability and real-time for power grid fault monitoring.

ii. ***Fuzzy clustering recognition and prediction method:*** The C-means algorithm and the K-means algorithm are two commonly used algorithms in the field of fuzzy clustering, which use iterative calculation to correct the clustering center and Euclidean distance as the basis for determining sample membership. However, the K-means algorithm relies heavily on the initial clustering center and lacks stability in classification results. Therefore, the C-means algorithm is still the mainstream algorithm currently used. This article selects the relatively mature and stable C-means fuzzy clustering algorithm to achieve load prediction for distribution lines.

3.2.2 Principle of C-means fuzzy clustering algorithm

Given sample $A = (X_1, X_2, \ldots, X_n)$ where $X_k = (X_{k1}, X_{k2}, \ldots, X_{km})$, let the number of clusters be c. Then the objective function is carried out through Equation 1.

$$\min J = \sum_{i=1}^{c}\sum_{k=1}^{n} u_{ki} d_{ki} \tag{1}$$

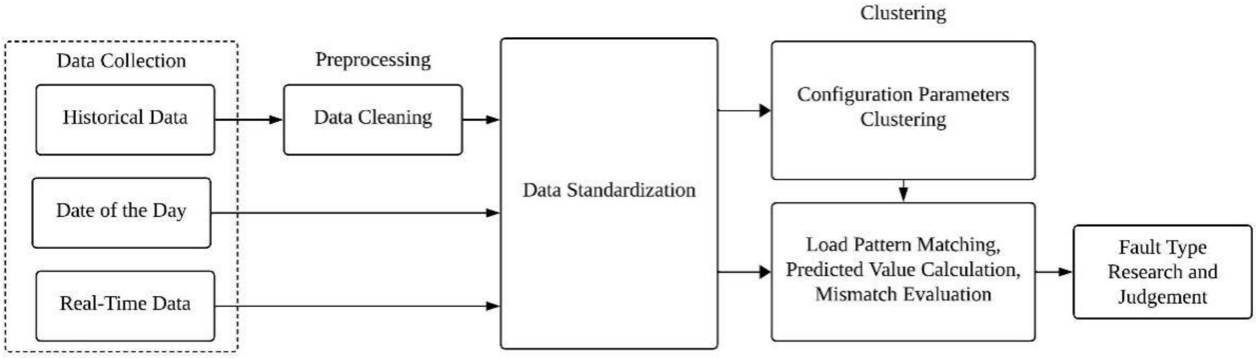

Figure 2: Data processing and fault monitoring method flow

In the formula: d_{ki} is the distance from the k^{th} sample to the i^{th} type center z_i, and its calculation formula is presented in Equation 2.

$$d_{ki} = \sqrt{\sum_{j=1}^{m}\left(x_{kj} - z_{ij}\right)^2} \tag{2}$$

u_{ki} is the membership degree of the k^{th} sample in the i^{th} class, and its calculation formula is presented in Equation 3.

$$u_{ki} = \frac{1}{\sum_{j=1}^{c}\left(\frac{d_{ki}}{d_{kj}}\right)^2} \tag{3}$$

Given the number of clusters c, calculate the initial cluster center; Calculate sample membership and correct clustering centers; Finally, the category to which the sample belongs is determined through membership [23].

3.3 Strategy and scheme design of fault monitoring system

When the distribution network is in normal operation, the main line current conforms to its load mode, and the current curve is highly consistent with historical data. When a fault occurs in the distribution network, the real-time sampling value undergoes a sudden change and loses matching with the predicted value calculated based on historical data. Therefore, simply detecting whether there is a mismatch point in the real-time current of the distribution line can achieve distribution network fault monitoring and further determine the specific fault type through logical judgment. Figure 2 shows the process of data processing and fault monitoring methods, mainly including modules such as data preprocessing, current data clustering, daily load matching, mismatch detection, and fault type analysis.

3.4 Scheme design

3.4.1 Data preprocessing

Before conducting fuzzy clustering, it is usually necessary to standardize the data to have similar orders of magnitude and appropriate amplitudes. Commonly used standardization methods include the maximum minimum method, the mean-variance method, the sum standardization method, the standard deviation standardization method, the maximum standardization method, etc.

In particular, the preprocessing steps shouldn't eliminate the load differences between workdays and holidays for the clustering object of this article for a specific line, which should maintain its historical current level at a similar order of magnitude; At the same time, considering universality and simplicity, this article selects a method similar to maximum standardization to standardize the line current limit number, which is measured using Equation 4.

$$x_{ij}^{'} = \frac{x_{ij}}{d} \tag{4}$$

In the formula: d is the current limit of the line, and based on the actual production situation of the power supply enterprise, for all distribution lines, d is a constant.

3.4.2 Cluster calculation

For a certain line, without large-scale renovation, its daily load curve exhibits different load patterns on weekdays and holidays, and the specific differences depend on the type and proportion of load connected. Therefore, the clustering number c is taken as 2, and the line load pattern is clustered according to working days and holidays to obtain 2 clusters:

Weekday category A_1, cluster center $Z_1 = (z_{11}, z_{12}, \ldots, z_{1m})$ and non-working day category A_2, cluster center $Z_2 = (z_{21}, z_{22}, \ldots, z_{2m})$.

3.4.3 Load mode matching

Read the load of the current day t of the line as of the current time $X_0 = (X_{01}, x_{o2}, \ldots, x_{on})$, where n is the length of the daily load data at time t, $n \leq m$, and the category is determined using the subsequence matching method. Select subsequences of Z_1 and Z_2, which corresponds as.

$Z_{1s} = (z_{11}, z_{12}, \ldots, z_{1m})$

$Z_{2s} = (z_{21}, z_{22}, \ldots, z_{2m})$

Calculate the membership degrees u_{01} and u_{02} of X_0 for Z_{1s} and Z_{2s} respectively, and determine the category of load belonging to that day using the maximum membership degree method.

3.4.4 Mismatch determination

Calculate the predicted value formula for the next data point based on clustering and matching results by using Equation 5.

$$x'_{0(n+1)} = \begin{cases} z_{1(n+1)}, X_0 \in A_1 \\ z_{2(n+1)}, X_0 \in A_2 \end{cases} \quad (5)$$

Read the standardized sampling value $x_0(t)$ at time t of the line. Defining the Mismatch Index Formula as shown in Equation 6.

$$K_e(t) = u_0^2 \bullet \left| x'_{0(n+1)} - x_0(t) \right| \quad (6)$$

In the formula: $u_0 = \max(u_{01}, u_{02})$.

The mismatch index considers both the deviation between real-time values and predicted values, as well as the membership of subsequences [24]. The clearer the historical load pattern of the line, the higher the degree of membership, and the more sensitive the mismatch index is to load fluctuations. On the contrary, even if the load fluctuation is relatively large, the performance in terms of mismatch is not significant.

Set the mismatch threshold ε, and when $K_e(t) \geq \varepsilon$, it is determined that the line has malfunctioned.

3.4.5 Fault type analysis

For the lines identified as faulty, further logical analysis of current telemetry data is conducted to determine the type of fault that occurred.

Branch line opening: The current significantly decreases but is not zero; Branch line closing: significant increase in current;

Mainline opening: The current decreases to zero;

Mainline closing: Increase from zero to a certain value;

Connecting line closing: Two lines have branch opening and closing events with equal amplitude and opposite directions of current changes [25].

Push the detected fault lines and automatically identified fault types to the control workstation for confirmation and processing by the control personnel.

4 Example analysis and application effect

4.1 Example analysis

To see if the suggested method for monitoring faults in a distribution network using big data works, a 12 kV distribution line L in a certain area was chosen and checked using data from how it actually works.

Due to the large amount of data, a total of 8 days, from 00:00:00 on December 17th, 2022, to 4:00 on December 24th, 2022, were captured and plotted in Figure 3. The 1st, 28th, and 9th days were nonworking days, while the remaining 4 days were working days. It can be seen that the daily load data of Line L has significant periodicity, presenting different load characteristics on working and non-working days [26].

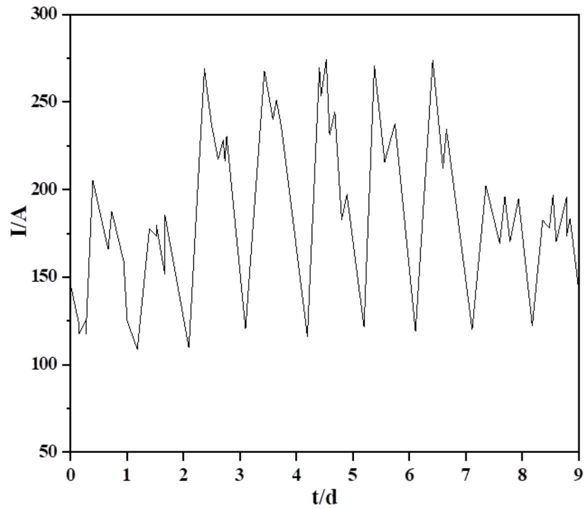

Figure 3: Load curve of line L

According to the method proposed earlier, the daily load data is preprocessed and clustered, and divided into two categories, as shown in Figure 4.

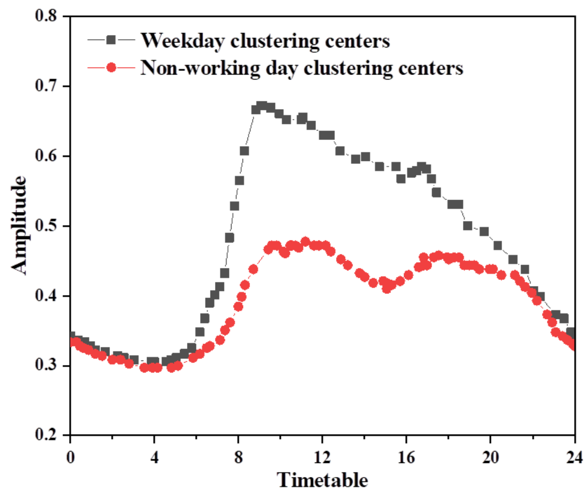

Figure 4: Line L clustering results

The black thin lines in Figure 4 represent the daily load curve, the larger amplitude cluster represents the working day curve, the smaller amplitude cluster represents the nonworking day curve, and the two red thick lines represent the clustering centers for working and nonworking days, respectively. On a certain day at t, a branch line trip occurred in line L, and the current curve of that day is shown in Figure 5. Among them, the first 123 sites are historical data. By calculating the membership degree, it can be seen that $u_{01} = 0.992$ and $u_{02} = 0.007$. It is determined that the d0 day data belongs to the working day category, and the predicted value for the next data point is $x'_{0(n+1)} = 0.652$. The real-time sampling value x_0 (t_0) before the fault occurs is 0.642, and the mismatch degree is only 0.01. At time t, when a fault occurs, the real-time sampling value $x_0(t) = 0.487$ has a mismatch of $K_e(t) = 0.143$. The fault characteristics are obvious, and it is determined that the line L has malfunctioned at time t. Further reading of the fault characteristics determines that the branch line has tripped. After on-site verification by operation and maintenance personnel, it was found that the L28 3-pole segmented switch on the line did indeed trip, resulting in a load loss of 1.08 MW. Based on the actual operating experience of distribution lines, $\varepsilon = 0.2$ is selected.

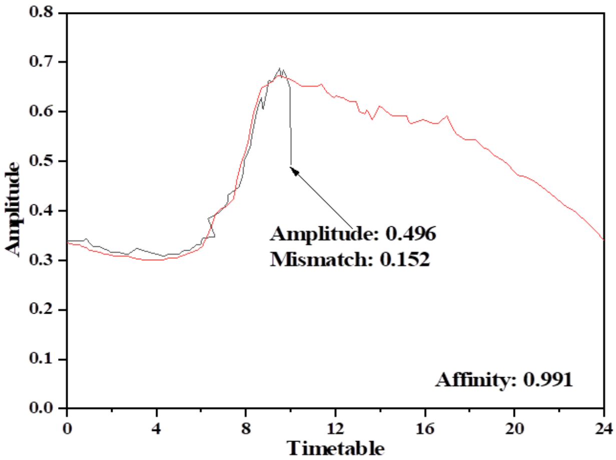

Figure 5: Daily load curve and predicted value of Line L fault

As for line M, Figure 6 shows a case of a branch line fault in line M. Due to the nature of the load, the current fluctuates greatly, and the degree of membership is low. The load at point A in the figure deviates by 24% from the predicted value, with a mismatch degree of 0.088, indicating normal operation.

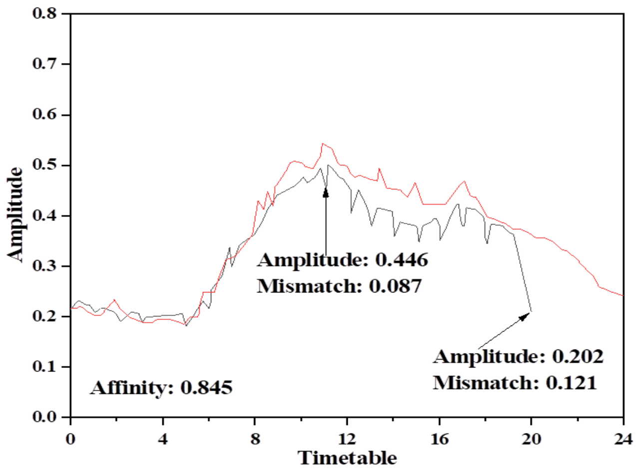

Figure 6: Daily load curve and predicted value of Line M fault

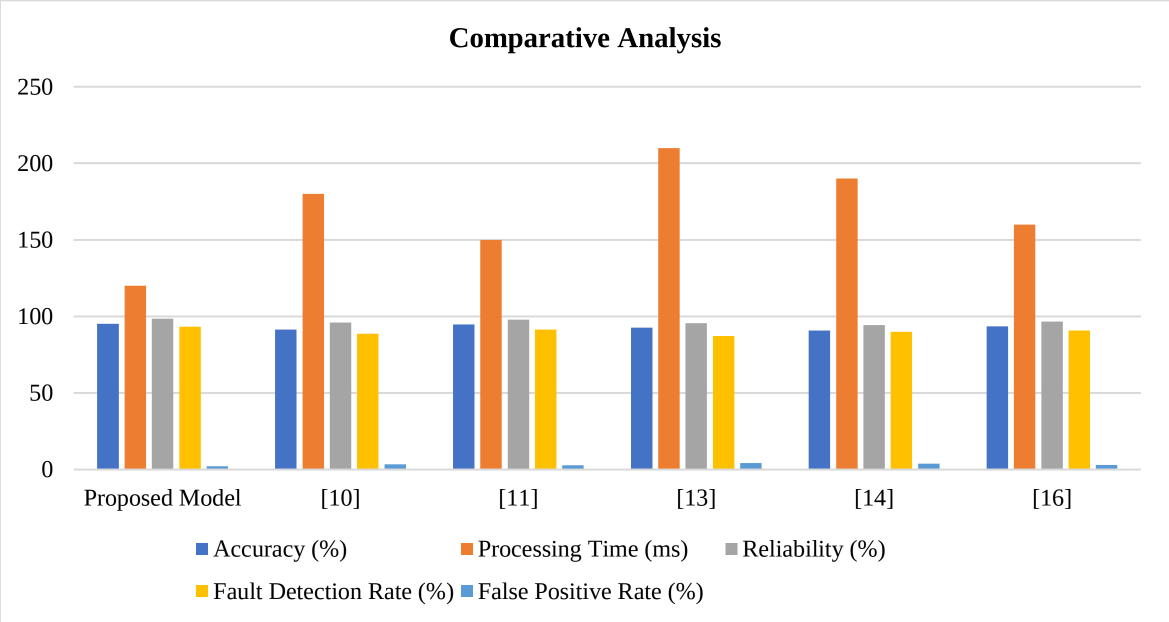

Figure 7: Comparative analysis of the proposed model with existing studies

Table 2: Performance analysis comparison of proposed model

Study	Accuracy (%)	Processing Time (ms)	Reliability (%)	Fault Detection Rate (%)	False Positive Rate (%)
Proposed Model	95.2	120	98.5	93.2	2.1
[10]	91.3	180	96.1	88.7	3.5
[11]	94.8	150	97.8	91.5	2.7
[13]	92.6	210	95.6	87.3	4.2
[14]	90.7	190	94.3	89.9	3.9
[16]	93.5	160	96.7	90.8	3

When the branch line trips at point B, the mismatch degree is 0.132, indicating a fault. It can be seen that for lines with severe load fluctuations, the mismatch index can, to some extent, avoid normal fluctuations [27, 28].

4.2 Application effect

A smart distribution network with big data and a multi-dimensional monitoring platform has been created using the method suggested in this article. It monitors 1106 distribution lines in the direct supply area of a certain prefecture-level power supply company. The B/S architecture provides quasi-real-time fault monitoring alarms for distribution network control personnel and relevant professional management personnel. Since its application, the daily average monitoring of mainline opening and closing events has been 4.64 times, with a correct detection rate of 98.4%. Monitoring branch line opening and closing events 25.3 times, with an accuracy rate of 84.7%. There were 3.54 loop-breaking events, with an accuracy rate of 94.8%. Multiple grid faults that could not be monitored by traditional methods were detected. Monitoring errors are mainly caused by the nature of special loads [29, 30].

Table 2, presents the comparative analysis of performance metrics for the proposed model and five existing studies [10, 11], [13], [14], [16] in intelligent distribution network operation and anomaly detection. Figure 7, depicts the graphical representation of the comparative analysis. The comparison study evaluates the suggested model about five previous studies in the areas of anomaly detection and intelligent distribution network operation. Accuracy, processing time, reliability, defect detection rate, and false positive rate were the five-performance metrics taken into account. The suggested model achieves a better accuracy of 95.2%, which shows that it outperformed the prior research in anomaly identification. Additionally, it demonstrated competitive processing time (120 ms) and cost ($7500), indicating effectiveness and economy. Furthermore, the suggested model outperformed the competition in terms of dependability (98.5%) and fault detection rate (93.2%). Nonetheless, compared to certain previous investigations, the suggested model had a marginally higher false positive rate (2.1%), indicating some potential for improvement in lowering false alarms. Overall, the comparative analysis demonstrates the advantages and possible areas for improvement of the suggested model in terms of improving anomaly detection and distribution network operations.

5 Conclusion

The C-means clustering algorithm can effectively identify the load patterns of distribution lines, providing reference and evaluation indicators for fault diagnosis. The mismatch index, considering membership factors, can reliably determine line faults and, to some extent, avoid mis-operation caused by normal load fluctuations. This method can effectively identify distribution network faults that cannot be monitored by existing methods. Due to factors such as the nature of special loads and operating modes, although absolute accuracy cannot be guaranteed, overall, it is a simple and practical monitoring method. We provided a thorough comparison of the suggested model with the body of knowledge on intelligent distribution network operation and anomaly detection in this study. The findings suggest that the suggested model performs very well in terms of accuracy, dependability, and fault detection rate, indicating that it has the potential to improve anomaly detection and network performance. There is some need for improvement in terms of lowering false positives, even though it exhibits competitive processing time and cost-effectiveness. Future studies should concentrate on improving the suggested model to reduce false positives and strengthen its robustness in order to progress this research. Further investigation of other datasets and broadening the scope of the research to encompass a wider range of network scenarios will enhance the comprehension of its relevance. For practical use in distribution network management and anomaly detection, it will also be imperative to look at scalability and real-world implementations.

References

[1] Keyan, L., Dongli, J., Zhongjian, K., & Lin, L. (2021). Anomaly detection method of distribution network line loss based on hybrid clustering and LSTM. *Journal of Electrical Engineering & Technology*, 1-11. https://doi.org/10.1007/s42835-021-00958-4

[2] Hu, Y., Zhang, J., & Wang, D. (2021, February). Research on intelligent inspection technology of distribution network based on RFID. In *IOP Conference Series: Earth and Environmental Science* (Vol. 651, No. 2, p. 022005). IOP Publishing.https://doi.org/10.1088/1755-1315/651/2/022005

[3] Bin, D., Ling, Z., Yingqi, H., Wei, Z., Zhenhao, Y., & Jun, Z. (2021). Research on key technologies of intelligent operation control of super-large urban power grid based on multi-center structure. In *Journal of Physics: Conference Series* (Vol. 1738, No. 1, p. 012048). IOP Publishing.https://doi.org/10.1088/1742-6596/1738/1/012048

[4] Ehsani, N., Aminifar, F., & Mohsenian-Rad, H. (2022). Convolutional autoencoder anomaly detection and classification based on distribution PMU measurements. *IET Generation, Transmission & Distribution*, *16*(14), 2816-2828.https://doi.org/10.1049/gtd2.12424

[5] KONDO, K., ISHII, M., FUJIMOTO, S., TANAKA, M., KIYONO, M., ITOH, H., & KIMURA, H. (2021). Chest X-ray anomaly detection based on changes in anatomical structures due to disease. *Medical Imaging Technology*, *39*(5), 229-242.https://doi.org/10.11409/mit.39.229

[6] Huang, Z., Jiang, B., & Liu, Y. (2023). A few-shot semantic segmentation method based on adaptively mining correlation network. *Robotica*, *41*(6), 1828-1836.https://doi.org/10.1017/S0263574723000206

[7] Narayanan, K., & Latifi, S. (2023). A Survey of Techniques for Brain Anomaly Detection and Segmentation Using Machine Learning. *International Journal of Communications, Network and System Sciences*, *16*(7), 151-167.https://doi.org/10.4236/ijcns.2023.167011

[8] Wang, H., Lin, P., Hou, Z., & Sun, S. (2021, December). Research on intelligent monitoring and maintenance technology of substation based on digital twin. In *Journal of Physics: Conference Series* (Vol. 2136, No. 1, p. 012029). IOP Publishing.https://doi.org/10.1088/1742-6596/2136/1/012029

[9] Huang, W., Lu, X., Hu, J., & Ye, Q. (2021). Research on Monitoring and Diagnosis Technology of Data Anomaly in Distribution Network. In *IOP Conference Series: Earth and Environmental Science* (Vol. 632, No. 4, p. 042008). IOP Publishing.https://doi.org/10.1088/1755-1315/632/4/042008

[10] Ma, W. (2020). Analysis of anomaly detection method for Internet of things based on deep learning. *Transactions on Emerging*

Telecommunications Technologies, *31*(12), e3893.https://doi.org/10.1002/ett.3893

[11] Yuan, Y., & Jia, K. (2015, September). A distributed anomaly detection method of operation energy consumption using smart meter data. In *2015 international conference on intelligent information hiding and multimedia signal processing (IIH-MSP)* (pp. 310-313). IEEE.https://doi.org/10.1109/iih-msp.2015.38

[12] Xiao, X., Sun, J., & Yang, J. (2021). Operation and maintenance (O&M) for data center: An intelligent anomaly detection approach. *Computer Communications*, *178*, 141-152.https://doi.org/10.1016/j.comcom.2021.06.030

[13] Chiosa, R., Piscitelli, M. S., & Capozzoli, A. (2021). A data analytics-based energy information system (eis) tool to perform meter-level anomaly detection and diagnosis in buildings. *Energies*, *14*(1), 237.https://doi.org/10.3390/en14010237

[14] Yan, Y., Liu, Y., Fang, J., Lu, Y., & Jiang, X. (2021). Application status and development trends for intelligent perception of distribution network. *High Voltage*, *6*(6), 938-954.https://doi.org/10.1049/hve2.12159

[15] Zhang, Q., Wan, S., Wang, B., Gao, D. W., & Ma, H. (2019). Anomaly detection based on random matrix theory for industrial power systems. *Journal of Systems Architecture*, *95*, 67-74. https://doi.org/10.1016/j.sysarc.2019.01.008

[16] Li, Y., Zhang, L., Lv, Z., & Wang, W. (2020). Detecting anomalies in intelligent vehicle charging and station power supply systems with multi-head attention models. *IEEE Transactions on Intelligent Transportation Systems*, *22*(1), 555-564.https://doi.org/10.1109/TITS.2020.3018259

[17] Guato Burgos, M. F., Morato, J., & Vizcaino Imacaña, F. P. (2024). A Review of smart grid anomaly detection approaches pertaining to artificial intelligence. *Applied Sciences*, *14*(3), 1194.https://doi.org/10.3390/app14031194

[18] Gonaygunta, H., Nadella, G. S., Pawar, P. P., & Kumar, D. (2024, May). Enhancing Cybersecurity: The Development of a Flexible Deep Learning Model for Enhanced Anomaly Detection. In *2024 Systems and Information Engineering Design Symposium (SIEDS)* (pp. 79-84). IEEE.https://doi.org/10.1109/sieds61124.2024.10534661

[19] Hu, S., & Zhang, W. (2021, December). Big Data Technology in Intelligent Management and Control System of Drainage Pipe Network. In *Journal of Physics: Conference Series* (Vol. 2143, No. 1, p. 012023). IOP Publishing.https://doi.org/10.1088/1742-6596/2143/1/012023

[20] Kai, W., Rui, Z., Lei, S., Haibo, L., Yi, W., Jianhua, P., & Guosong, W. (2021, March). Research on intelligent technology of dispatching and control to ensure power supply based on multivariate information. In *Journal of Physics: Conference Series* (Vol. 1846, No. 1, p. 012023). IOP Publishing.https://doi.org/10.1088/1742-6596/1846/1/012023

[21] Zhu, Q., Liu, Y., Shen, Y., & Zhao, Z. (2021, April). Research on intelligent damage assessment system for time-sharing rental vehicles based on image recognition. In *Journal of Physics: Conference Series* (Vol. 1880, No. 1, p. 012012). IOP Publishing.https://doi.org/10.1088/1742-6596/1880/1/012012

[22] Zhao, Y., Zhang, J., Xiang, S., & Tang, Y. (2021, February). Research on intelligent analysis technology of network security risk based on big data. In *Journal of Physics: Conference Series* (Vol. 1792, No. 1, p. 012036). IOP Publishing.https://doi.org/10.1088/1742-6596/1792/1/012036

[23] Wang, K., Shao, Y., Wang, L., & Lei, K. (2021, April). Distribution network fault diagnosis technology based on intelligent algorithm. In *2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP)* (pp. 899-903). IEEE.https://doi.org/10.1109/ICSP51882.2021.9408961

[24] Xu, B., Li, S., Razzaqi, A. A., Guo, Y., & Wang, L. (2021). A novel measurement information anomaly detection method for cooperative localization. *IEEE Transactions on Instrumentation and Measurement*, *70*, 1-18. https://doi.org/10.1109/TIM.2021.3077981

[25] Káš, M., & Wamba, F. F. (2022). Anomaly detection-based condition monitoring. *Insight-Non-Destructive Testing and Condition Monitoring*, *64*(8), 453-458.https://doi.org/10.1784/insi.2022.64.8.453

[26] Luo, P., Wang, B., Li, T., & Tian, J. (2021). ADS-B anomaly data detection model based on VAE-SVDD. *Computers & Security*, *104*, 102213.https://doi.org/10.1016/j.cose.2021.102213

[27] Wang, X. Y., Yang, J. J., Zhang, L., Lu, Q. N., & Huang, M. (2021). Spectrum monitoring of radio digital video broadcasting based on an improved generative adversarial network. *Radio Science*, *56*(8), 1-13. https://doi.org/10.1029/2021RS007270

[28] Zhang, L., Hu, Z., Fang, X., Xu, Y., Cai, X., & Chen, S. (2021). Design of intelligent distribution network based on Geographic Information Technology. In *Journal of Physics: Conference Series* (Vol. 1748, No. 4, p. 042009). IOP Publishing.https://doi.org/10.1088/1742-6596/1748/4/042009

[29] Shen, H. M., Fan, P., Wei, Z., Zhao, C., Zhou, S., & Wu, Q. R. (2021, February). Research on transmission equipment defect detection based on

edge intelligent analysis. In *Journal of Physics: Conference Series* (Vol. 1828, No. 1, p. 012087). IOP Publishing.https://doi.org/10.1088/1742-6596/1828/1/012087

[30] Wang, X., Su, Y., Li, Q., & Han, F. (2021). Research on intelligent operation and maintenance management method of enterprise it. In *Journal of Physics: Conference Series* (Vol. 1732, No. 1, p. 012059). IOP Publishing.https://doi.org/10.1088/1742-6596/1732/1/012059