Data Mining-based Fault Detection Method for Distributed Power Generation Systems

Huaqin Wu

Henan Technical Institute, Zhengzhou, Henan, 450042, China
E-mail: 13523405765@163.com

Received 14 May 2021; Accepted 27 May 2021;
Publication 28 July 2021

Abstract

Traditional fault detection methods for power generation systems use centralized fault processing analysis, which leads to long accuracy and response time of fault detection. To address these problems, a data mining-based distributed power generation system fault artificial intelligence detection method is studied. The depth-first search tree algorithm is used to divide the grid of distributed generation system. The network structure is modified to locate fault zones by processing anomaly mining of the system data after grid division. The combination of fuzzy logic and wavelet singular entropy is used to complete the detection and identification of system faults. Through simulation experiments, it is verified that the response time of the detection method is only 0.016 s, and its detection error rate and false negative rate are 1.23% and 1.25%, which are far lower than other methods.

Keywords: Data mining, distributed generation systems, grid faults, artificial intelligence, fault detection.

Introduction

Due to the continuous expansion of the scale and complexity of the structure of distributed power system, when the power system fails, a large number

Distributed Generation & Alternative Energy Journal, Vol. 36_4, 441–458.
doi: 10.13052/dgaej2156-3306.3646

of alarm information will flood into the dispatching center in a short period of time, which is far beyond the handling capacity of the operator, thus causing some difficulties for the dispatcher to identify the fault signal quickly and accurately. Therefore, it is very important to study how to analyze and judge these signals quickly and accurately and take corresponding effective measures for the normal operation of distributed generation system [1]. Domestic and foreign scholars have proposed methods such as expert system (ES), artificial neural network (ANN), Petri network, genetic algorithm and Bayesian network, etc. Most of these methods can achieve more satisfactory results for accurate and complete signals of the control center. These traditional fault detection methods are mostly centralized processing of data, assuming that all data are concentrated locally for anomaly detection [2]. Among them, Bayesian network fault detection method is a fault diagnosis model based on Bayesian network, which can deal with the uncertainty in power grid fault diagnosis, and has the characteristics of accurate semantics, fast reasoning and high learning efficiency. Petri net fault detection method is to obtain a steady Petri net model according to the occurrence of transition, and use graphical and analytical expressions to describe the power grid fault diagnosis process, so as to realize the functions of information preprocessing, startup detection, fault detection and so on. For distributed power systems, which are increasingly used in practice, their functions are becoming more and more complex, the size of the data sets are becoming larger, and the power supply data sets all need to be distributed in different subsystems, and the traditional centralized power system fault detection methods have limitations. At the same time, in the case of complex fault diagnosis modes such as incomplete signals, signal change or loss caused by the failure of communication devices, none of the above detection methods can achieve satisfactory results, and there are problems such as low detection rate and unstable detection efficiency [3].

Data mining technology can extract key information from the massive information, and in today’s era of informationization and networking, the application of this technology covers all aspects of finance, economy, and society. In the massive information system of smart grid, data mining technology can play an advantage well and combine with the specialized knowledge of power system to propose new solutions or countermeasures to many traditional problems [4]. Distributed generation system will significantly reduce energy consumption, save cost, and improve system flexibility if it can effectively cooperate with large power grids. Besides, data mining technology has the advantages of small investment, environmental friendliness, and high flexibility. As the role of distributed generation becomes more and more obvious and the level of load electricity consumption increases, the threat to the safe operation of the power system becomes more and more serious. Based on the characteristics of the novel problems in this field, the advantages of data mining technology can be brought into play. With respect to the above analysis, this paper will investigate the artificial intelligence detection method of distributed power generation system faults based on data mining.The depth-first search tree algorithm is used to search all nodes of distributed generation system, and grid division is carried out. The collected data is preliminarily cleaned, converted and integrated by data mining method, which avoids the large deviation caused by data redundancy and loss in data mining. Using wavelet singular entropy to analyze the original data of one cycle after the fault, the phase difference of the fault can be detected, and the wavelet singular entropy value of the fault line is larger than that of the normal line, thus realizing the fault detection.

1 Artificial Intelligence Detection Method for Distributed Power Generation System Faults Based on Data Mining

1.1 Distributed Power Generation System Grid Partitioning

The fault detection problem of distributed power systems can be effectively solved by artificial intelligence techniques. In order to improve the efficiency of artificial intelligence in detecting faults in distributed power generation systems, the network of power generation systems needs to be effectively partitioned. A weighted depth-first search tree algorithm is used to grid partition the distributed power generation system.

Depth-first search tree algorithm is a standard algorithm for searching all nodes of the power system node graph when performing fault detection in distributed generation systems. Assuming that G is a graph with n nodes, the steps to form a weighted depth-first search tree are as follows [5, 6].

(1) Select the node with the largest number n, label it as 1, which is the root node of the tree, and proceed to the next step with this node and the corresponding label number as the initial condition.
(2) For node i with annotation number k (annotation number k indicates the order of depth-first search), if all the nodes associated with i have been annotated, go to step (3); otherwise, select the node with the largest node number among the unannotated nodes associated with i and give it the smallest unused annotation number in the depth-first search sequence

$\{1, 2, \ldots, n\}$. Repeat step (2) with this freshly annotated node and the corresponding annotation number as the new starting point.

(3) If the labeling number k of node i satisfies $k > 1$ (i.e., node i is not the root node), then backtrack from node i along the search path to the previous node and repeat step (2) with this node and the corresponding labeling number as the new starting point; conversely, if k=1 (i.e., node i is the root node), the algorithm terminates. In order to balance the computational load of each subgrid network after partitioning, each node in the network is assigned a weight (the weight takes the value of an integer), which is used to indicate the computational load of the corresponding node. Related studies show that the computational load of a sub-network is mainly determined by the total number of possible faulty elements within that sub-network, so the weight of a node is defined as the number of possible faulty elements (busbars, transmission lines and transformers) associated with this node.

Assume that Y_n is the node derivative matrix of a given power network, where the nodes are arranged in descending order of the labeling number k. For node i the number of all non-zero elements in row i of the upper triangular array of Y_n (including the diagonal elements) is taken as its weight, denoted as $W_{node}(i) = w_i$. After considering the node weights, the tree generated by the DFS algorithm becomes a weighted DFS tree. Let T[i] denote the subtree with node i as the root node, then the total weight associated with this subtree is defined as [7]

$$W(T[\mathrm{i}]) = \sum_{j \in T[j]} W_{node}(j) \tag{1}$$

For a graph G of n nodes with node n as the root of the DFS spanning tree, $i = P(i)$ indicates that node i is the parent node of node j, i.e., i is an adjacent node of j in the path from j to the root node; also, j is called a child node of node i. A node without children is called a leaf node. The length of node i is defined as the number of nodes in the path from i to the root node. According to the depth-first search tree algorithm given above to partition the distributed power system network, if n_g is the desired number of sub-networks, the power system network is partitioned into n_g connected sub-networks by the algorithm, and the weights of each sub-network are made as close as possible to $W(T[\mathrm{i}])/n_g$. Then the specific process of partitioning the power system network is as follows [8].

(1) Let S denote the set of nodes of graph G, C_l denotes the set of nodes of the lth subnetwork, where $l = 1, 2, \ldots, n_g$. C_w is the temporary

working set that will be used in the algorithm. The initial conditions of the partitioning algorithm are set as $S = \{1, 2, \ldots, n\}$, $C_l = \Phi$, $C_w = \Phi =$, $l = 1$, Φ is the empty set.

(2) All leaf nodes with the maximum length from the root node are selected as the starting point of the search and added to the temporary set C_w. Also, assume that the maximum length of the leaf node is k and use it as a loop control pointer.

(3) Modify the loop control pointer $k = k - 1$.

(4) Find the parent nodes of all nodes in C_w, and do the following operations for each different parent node (denoted as node p) in turn. Find the children $\{j|p = P(j)\}$ of the parent node p, and arrange the weight values in descending order according to the weights $W(T[\mathrm{i}])$ of the corresponding subtree. Subsequently, judge these children nodes in turn if $T[j]$ satisfies the following relation [9].

$$\left| W(T[j]) - \frac{W(T[\mathrm{i}])}{n_g} \right| < \left| W(T[p]) - \frac{W(T[n])}{n_g} \right| \tag{2}$$

Then the subtree $T[j]$ constitutes a subnetwork, i.e., $C_l = T[j]$, and correct $l = l + 1$, while node j and the set of nodes $T[j]$ are removed from C_w and S, respectively; conversely, if a child node does not satisfy Equation (2), the child node is removed from C_w and its parent node p is added to C_w. (If the parent node p has more than one child node at the same time that does not satisfy Equation (2), then only the common parent node p is retained in C_w). When all the different parent nodes are tested go to step (5).

(5) Search all leaf nodes of length k and add them to the temporary set C_w. Then go to step (3) and repeat the above steps until $k = 0$, i.e., $S = \Phi$, all nodes are searched and the algorithm terminates.

After using the algorithm to divide the distributed power generation system grid, the data collected by the distributed power generation system state data collector is mined for anomalies using the data mining algorithm according to the results of the division.

1.2 Power Generation System Data Abnormal Mining

Preliminary data cleaning, data transformation and data integration processing are performed on the collected data to avoid large deviations in data mining caused by data redundancy and missing phenomena. Before data mining, it is necessary to first analyze the data features corresponding to different faults of distributed power generation systems, compose the data features into feature sets and evaluate them by feature subsets to obtain the final fault data features.

According to the measurement of sequential search for fault data feature subset search, the sequential forward search algorithm can be divided into the following steps [10].

(1) Generate feature subset. Starting from the empty set, in the kth round, each time l features are selected from $d-k$ candidate features and added to the selected optimal feature subset containing k features to form a new candidate feature subset.
(2) Compare feature subset evaluation values. While generating the candidate feature subset, compare it with the optimal subset evaluation value of the previous round, and update the optimal feature subset if it is greater than the optimal evaluation value of the previous round.
(3) Output feature results. When the optimal feature subset is no longer updated in a round, the search is stopped to output the optimal feature subset.

The sequential backward search algorithm starts from the full set and removes r features that contribute the least to the evaluation function each time until the evaluation function value no longer increases. The two-way search algorithm combines forward and backward search, increasing l relevant features and decreasing r irrelevant features in each round, and outputting the optimal feature subset when the optimal feature subset is no longer updated after n rounds. Bidirectional search is faster than backward search and has better performance than forward search. Therefore, this paper uses the two-way search algorithm as the algorithm for feature subset search.

The relevance-based feature subset evaluation algorithm is selected as the feature subset evaluation algorithm. The feature subset evaluation function used in this algorithm can select the optimal feature subset with low redundancy among features and high correlation between features and predictor variables (i.e., strongly correlated features), and the weighted correlation coefficient calculation method introduced in this algorithm can measure the correlation between various types of variables [11].

The value of the evaluation function for the subset of fault characteristics of the distributed generation system is calculated according to the following equation [12].

$$e(F) = \frac{d\bar{r}_{cf}}{\sqrt{d + d(d-1)\bar{r}_{ff}}} \tag{3}$$

In the above equation, $e(F)$ is the evaluation value of the candidate subset; d is the number of features contained in the candidate subset F; $\bar{r}_{cf}$ is the average correlation between all feature variables and predictor variables in the candidate subset F; $\bar{r}_{ff}$ is the average correlation between feature variables in the candidate subset F. If the feature redundancy within the feature subset is higher, i.e., the larger $\bar{r}_{ff}$ is, the smaller the value of $e(F)$ is; the more strongly correlated feature variables in the candidate feature subset F, i.e., the larger the value of $\bar{r}_{cf}$ is, the larger the value of $e(F)$ is. Therefore, by comparing the evaluation values of the candidate feature subsets calculated by Equation (3), redundant and non-strongly correlated variables can be eliminated and the optimal feature subset that is strongly correlated with the predictor variables can be selected.

Based on the optimal feature subset selected by the above process, the C4.5 decision tree algorithm is used for distributed generation system anomaly data mining. First, the information gain and gain rate of the candidate features (obtained by random sampling of candidate features without putting back from the feature set) are calculated. The information gain and gain rate of the candidate division features dividing the training sample set D are calculated by the following formula [13].

$$\begin{aligned} &\mathit{Gain}(D,a) = \mathit{Ent}(D) - \mathit{Ent}_a(D) \\ &\mathit{Gainr}(D,a) = \mathit{Gain}(D,a)/IV(a) \\ &IV(a) = -\sum_{v=1}^{V} \frac{|D^v|}{|D|} \log_2 \frac{|D^v|}{|D|} \end{aligned} \tag{4}$$

In the above equation, $\mathit{Gain}(D,a)$ is the information gain of the features of different data types; $\mathit{Gainr}(D,a)$ is the gain rate of the data features; $\mathit{Ent}(D)$ is the information entropy of the data set D; $\mathit{Ent}_a(D)$ is the information entropy after dividing the data set D according to the features; D^v is the information entropy of all the samples in the sample set D with the value a^v of feature a at the vth branch node of the decision tree. After calculating the information gain and gain rate of the sample set, the features with higher-than-average information gain are identified from the candidate division features, and then the feature with the highest gain rate is selected from them as the split node, and the initially selected feature is used as the root node of the decision tree. The current samples are divided according to the different values of the features on the split node to construct branches. Leaf nodes are generated when the samples in a branch belong to the same class. When all branches are generated leaf nodes, the tree construction is completed by pruning process using post-pruning method [14, 15]. The subset of fault features of the power generation system generated by the evaluation is used as the division criterion, and the constructed decision tree is used for data mining. After the decision tree algorithm processes the data collected by the data collector of the distributed power generation system, different types of fault data of the power generation system are obtained. After mining the abnormal data, in order to quickly locate the possible locations of faults in the power generation system, corrections for changes in the power generation system grid structure are performed in order to intelligently isolate the fault areas.

1.3 Power Generation System Grid Structure Change Correction

In this paper, the node-branch correlation matrix L is defined based on the idea of connected system. It can be adapted to the change of network topology only by correcting the fault information matrix Gp or tripping switch vector Ds which contains fewer elements.

To improve the reliability of power supply, a contact switch is usually connected between two main sources in the distribution network. When the branch line structure is changed, the contact switch will not close; while when the main line structure is changed, the contact switch may close and the backup power supply will continue to supply power to part of the line. This section investigates the structure correction algorithm based on the state of the contact switch after the network topology is changed. First, the switch state vector $K = \left[k_1\, k_2 \ldots k_m\right]$ is defined, where element k_i on the line is defined as [16]:

$$k_i = \begin{cases} 1, & \text{switch on} \\ 0, & \text{switch off} \\ -1, & \text{unknown} \end{cases} \tag{5}$$

When the removed line is a branch line, the contact switch does not operate, and the main line and the remaining branch lines are still powered by the original main power supply. If a fault occurs in the system, the fault detection and isolation algorithm does not need to be modified. Since the corresponding switch is disconnected when the line is removed, the elements of the trip switch vector Ds need to be modified based on the original algorithm in order to avoid issuing a trip command to a switch that is already disconnected when performing fault isolation as follows [17].

$$d_i' = d_i \,\&\, k_i \tag{6}$$

When the removed line is the main line, the contact switch is closed in order to improve the reliability of the power supply, and the power supply is continued from the backup power source to the non-faulted section, which affects the original fault detection algorithm because the power supply has been changed in some lines. In this case, the definition of the element g_i in the fault information vector Gp needs to be amended: g_i takes "0" when the fault direction detected by the measurement point on the main line is in the opposite direction, otherwise it takes "1"; "1" when the fault direction detected by the measurement point on the branch line is in the positive direction, otherwise it takes "0". After the correction of the grid structure, the system fault zone can be determined, and the theory of wavelet transform combined with fuzzy logic is used for fault detection and identification.

1.4 Distributed Generation System Fault Detection Implementation

Calculate the positive sequence components of the anomalous data of the power system after data mining processing, and use the wavelet coefficient matrix of the power generation system obtained by Haar wavelet analysis. The obtained wavelet coefficient matrix is subjected to singular value decomposition, and the SVD decomposition is defined as follows [18].

$$A = U^* Q^* W \tag{7}$$

In the above equation, A is the wavelet coefficient matrix of the anomalous data of the power system; W and U are the You matrices of the singular value decomposition; Q is the matrix composed of the diagonal elements of the matrix M; the matrix M is the diagonal matrix composed of non-zero singular values. The singular value matrix is analyzed, and if a larger entropy value is obtained, the greater the singularity of the signal. When a fault occurs in the line, there will be a high frequency signal, so that a large singularity will occur. The wavelet transform combined with Shannon entropy is defined as wavelet singular entropy to represent the degree of order of the time-frequency coefficient matrix after the wavelet transform of the data. By applying the wavelet singular entropy to analyze the raw data for 1 cycle after the occurrence of a fault, the phase difference of the fault can be detected and the wavelet singular entropy value of the line where the fault occurs is larger compared to the normal line. The wavelet singular entropy is defined as follows [19].

$$WSE(n) = \sum_{i=1}^{n} \Delta p_i \tag{8}$$

Where, $WSE(n)$ denotes the wavelet singular entropy of the nth order; Δp_i is the incremental wavelet singular entropy, and the calculation formula is shown in (9).

$$\Delta p_i = -\left(\frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j}\right) \log_2 \left(\frac{\lambda_i}{\sum_{j=1}^{n} \lambda_j}\right) \tag{9}$$

In the above equation, λ_i and λ_j denote the singular values on the diagonal matrix. Fault detection and identification is performed by combining fuzzy logic with wavelet singular entropy. The distributed generation system fault detection process is shown in Figure 1 below [20, 21].

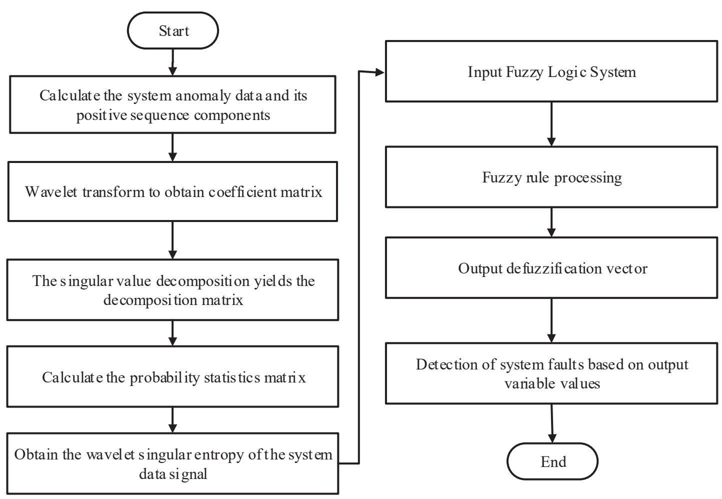

Figure 1 Fault detection process of distributed generation system.

In the fuzzy system, four inputs are set and two trapezoidal high and low affiliation functions are used for fault detection. two trapezoidal high and low affiliation functions correspond to the same input index, and the output of the fuzzy inference system editor selects the triangular affiliation function and sets the affiliation function values for normal and various fault cases. Then, according to the final calculated output, the specific fault location and fault type in the fault zone can be determined and detailed fault detection results can be obtained. Thus, the research on the artificial intelligence detection method of distributed power generation system faults based on data mining is completed.

2 Simulation Experiment

The artificial intelligence detection method of distributed power generation system fault based on data mining is proposed above, and this section will simulate and study the fault detection method.

2.1 Simulation System Modeling

Figure 2 below shows a simple distributed power system simulation model consisting of one distributed power supply unit connected to the load through the PCC common point. The transformer is placed at the end of the distributed power supply unit, which ensures that current and voltage signals are available in different situations. The simulation is performed in MATLAB 2015 and the fault occurs at the PCC.f in the figure is the fault occurrence point.

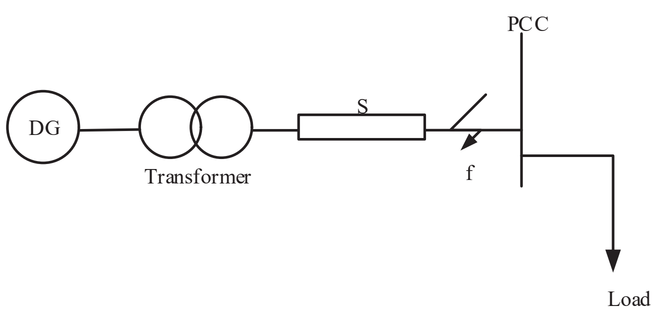

Figure 2 Distributed power system simulation model.

Distributed power supply DG: capacity 9 MVA, voltage 400 V, frequency 50 Hz. transformer: capacity 10 MVA, frequency 50 Hz, rated ratio 400 V/20 kV, transformer one winding resistance R1 = 0.00375 p.u., impedance X1 = 0.1 p.u., magnetoresistance Rm = 500 p.u., electromagnetic impedance Xm = 500 p.u. Distribution line: PI-type model, length is 5 km , rated voltage 20 skV, rated capacity 20 MVA, line resistance R0 = 0.1948 Ω/km, R1 = 0.027 Ω/km, C0 = 9e-9F/km C1 = 12.7e-9 F/km, L0 = 2.067e-3H/km, L1 = 0.88586e-3H/km. load: capacity of 3 MW.

2.2 Experimental Procedure

This simulation experiment uses a comparative format to compare the fault detection method proposed above with the Bayesian network-based fault detection method and the Petri network-based fault detection method in two stages: the response time of fault detection and the detection of false misses. The experiments are done in MATLAB software. After establishing the power system simulation model shown in Figure 2 above, different faults are set and the simulation experimental study is completed after comparing the data corresponding to the experimental indexes.

2.3 Experimental Results

The fault is set to occur at 1s, and the fault occurs at 2.5 km. Figure 3 shows the time taken by each method to detect the fault under various faults. If the

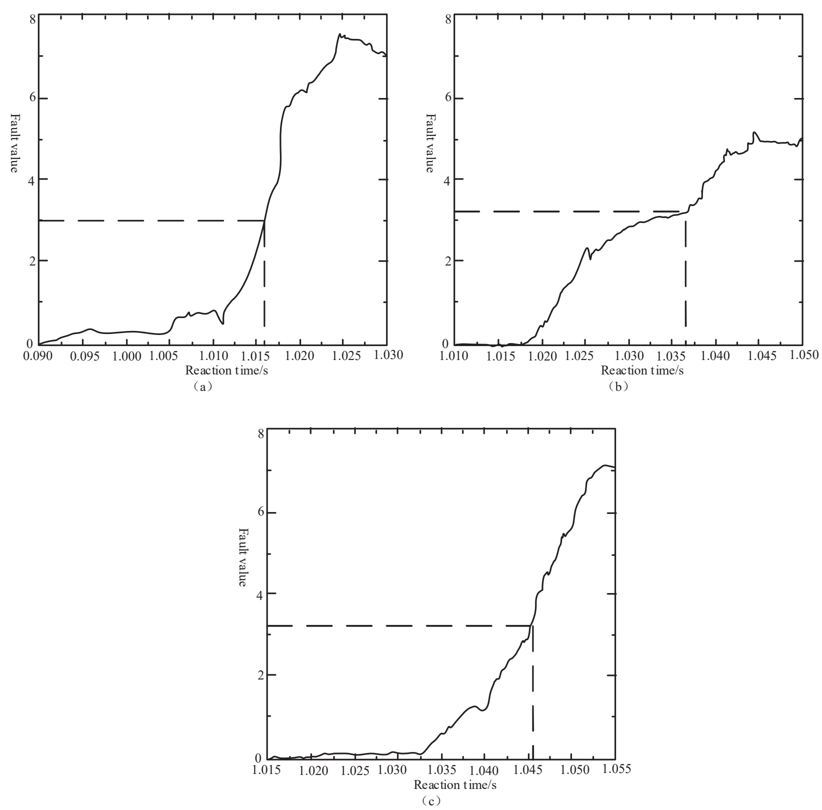

Figure 3 Detection response time of fault detection methods.

fault time period is refined, the response time of each detection method can be seen. In Figure 3 below, Figure 3(a) shows the fault detection response time of the method in this paper; Figure 3(b) shows the fault detection response time of the method based on Bayesian network; and Figure 3(c) shows the fault detection response time of the method based on Petri network.

Analysis of the curves in Figure 3 above shows that the fault is detected in 1.016 s for the method in this paper starting at 1s, in 1.037 s for the Bayesian network-based method, and in 1.045 s for the Petri network-based method, showing the rapidity of the method in fault detection in this paper.

Table 1 below shows the comparison of the false and missed detection rates of the three fault detection methods for different fault types of faults.

Table 1 Comparison of fault detection methods in terms of misjudgment rate and false missing rate

	Method in this Article		Bayesian Network-based Approach		Petri Network-based Approach	
Fault Type	Misjudgment Rate/%	Missing Rate/%	Misjudgment Rate/%	Missing Rate/%	Misjudgment Rate/%	Missing Rate/%
AG	1.14	1.24	3.84	2.79	3.77	3.09
BG	1.20	1.26	3.53	2.87	4.01	3.08
CG	1.38	1.24	3.51	2.84	4.24	3.12
ABG	1.19	1.24	3.58	2.75	4.13	3.14
ACG	1.31	1.24	3.86	2.93	3.76	3.16
BCG	1.13	1.26	3.77	2.88	3.83	3.13
AB	1.13	1.25	3.69	2.72	4.21	3.19
AC	1.37	1.26	3.84	2.76	4.18	3.03
BC	1.22	1.26	3.83	2.75	3.74	2.95
ABC	1.21	1.24	3.74	2.83	3.96	3.01
ABCG	1.19	1.25	3.54	2.81	3.85	2.94

Analyzing the data in the above table, it can be seen that the false positive rate and the leakage rate of the method in this paper are much lower than the other two fault detection methods when detecting different types of faults. The average false alarm rate and leakage rate of this method are 1.23% and 1.25%, respectively; the average false alarm rate and leakage rate of the Bayesian network-based method are 3.70% and 2.81%, respectively; the average false alarm rate and leakage rate of the Petri network-based method are 3.97% and 3.08%, respectively. In summary, the artificial intelligence detection method of distributed power generation system based on data mining proposed in this paper has a superior detection performance.

3 Conclusion

The joint development of technology, public environmental policies and the expansion of the electricity market have made distributed generation an important energy option in the new century, and the application of distributed generation systems has effectively improved the reliability of electricity services and power quality. It is difficult to ensure stable and high correct diagnosis rate when using traditional fault detection methods for distributed generation systems. Therefore, in view of the problems of traditional fault detection methods, this paper investigates the artificial intelligence detection method of distributed power generation system faults based on data mining , the grid data of distributed generation system is divided by depth-first search tree algorithm, and the abnormal data is analyzed by fuzzy logic and wavelet singular value method, and the fault detection is determined., and verifies the effectiveness of the method through simulation experiments.However, the application of this method in fault detection of power generation system still has some limitations. For different types of power plants and power stations, it needs to be adjusted according to the needs in actual use, and its adaptive performance should be further improved in future research.

References

[1] AbdLkader AG, Saleh SM. Zero noncetection zone assessment for anti-islanding protection in rotating machines based distributed generation system[J]. International Journal of Energy Research, 2020, 45(10):521–540.

[2] Murthy G, Sivanagaraju S, Rao BH. Reliability improvement of radial distribution system with distributed generation[J]. International Journal of Engineering Science and Technology, 2020, 4(9):4003.

[3] F. Dewi Cahya, Daniar Fahmi, I Made Yulistya Negara, et al. Loading Effect with Distributed Generation (DG) Injection on Distribution System on Temperature and Breakdown Voltage of Oil Transformer in Distributed System[J]. Journal of Physics: Conference Series, 2019, 1201(1):012011.

[4] Maged N.F. Nashed, Mona N. Eskander. Novel circuit to compensate the effect of source open circuit fault in distributed generation system[J]. Energy Reports, 2020, 6(Supl.2):76–80.
[5] Hannoon, Naeem M.S., Vijayakumar, V., Vengatesan, K., et al. Small Signal Fault Analysis for Renewable Energy (Wind) Power System Distributed Generation by Using MATLAB Software (Simulink)[J]. Journal of Computational and Theoretical Nanoscience, 2019, 16(2):537–543(7).
[6] Donghan Feng, Fan Wu, Yun Zhou, et al. Multi-agent-based Rolling Optimization Method for Restoration Scheduling of Distribution Systems with Distributed Generation[J]. Journal of Modern Power Systems and Clean Energy, 2020, 8(04):737–749.
[7] Galiveeti Hemakumar Reddy, Arup Kumar Goswami, Nalin B. Dev Choudhury. Impact of Distributed Generation Integration on Distribution System Reliability[J]. Indian Journal of Science and Technology, 2018, 11(34):1–13.
[8] Huadong Mo, Giovanni Sansavini. Impact of aging and performance degradation on the operational costs of distributed generation systems[J]. Renewable Energy, 2019, 143:426–439.
[9] Faiz MT, Khan MM, Jiang H, et al. H ∞ robust control with improved harmonics suppression for inverter-based distributed generation systems[J]. IET Power Electronics, 2020, 13(13):2742–2755.
[10] Emilio Barocio, Petr Korba, Walter Sattinger, et al. Online coherency identification and stability condition for large interconnected power systems using an unsupervised data mining technique[J]. IET Generation, Transmission & Distribution, 2019, 13(15):3323–3333.
[11] Yingzhe Kan, Dongye Sun, Yong Luo, et al. Optimal design of the gear ratio of a power reflux hydraulic transmission system based on data mining[J]. Mechanism and Machine Theory, 2019, 142:103600.
[12] Shan Liu, Li Sun, Senlin Zhu, et al. Operation strategy optimization of desulfurization system based on data mining[J]. Applied Mathematical Modelling, 2020, 81:144–158.
[13] Oujdi Sihem, Belbachir Hafida, Boufares Faouzi. C4.5 Decision Tree Algorithm for Spatial Data, Alternatives and Performances[J]. Journal of computing and information technology, 2019, 27(3):29–43.
[14] Sandoval Betsy, Barocio Emilio, Korba Petr, et al. Three-way unsupervised data mining for power system applications based on tensor decomposition[J]. Electric Power Systems Research, 2020, 187:106431.

[15] Michał Jasiński, Tomasz Sikorski, Zbigniew Leonowicz, et al. The Application of Hierarchical Clustering to Power Quality Measurements in an Electrical Power Network with Distributed Generation[J]. Energies, 2020, 13(9):1–19.
[16] Beatrice Marchi, Simone Zanoni, Marco Pasetti. Multi-Period Newsvendor Problem for the Management of Battery Energy Storage Systems in Support of Distributed Generation[J]. Energies, 2019, 12(23):1–13.
[17] Lucas F, Costa P, Batalha R, et al. Fault detection in smart grids with time-varying distributed generation using wavelet energy and evolving neural networks[J]. Evolving Systems, 2020, 11(2):165–180.
[18] Sharma Nikhil Kumar, Samantaray Subhransu Ranjan. Assessment of PMU-based wide-area angle criterion for fault detection in microgrid[J]. IET Generation, Transmission & Distribution, 2019, 13(19):4301–4310.
[19] Roy Subhamita, Debnath Sudipta. PSD based high impedance fault detection and classification in distribution system[J]. Measurement, 2021, 169(1):108366.
[20] Du Wenhua, Guo Xiaoming, Wang Zhijian, et al. A New Fuzzy Logic Classifier Based on Multiscale Permutation Entropy and Its Application in Bearing Fault Diagnosis.[J]. Entropy (Basel, Switzerland), 2019, 22(1):27.
[21] Ghahderijani MM, Camacho A, Moreira C, et al. Imbalance-Voltage Mitigation in an Inverter-Based Distributed Generation System Using a Minimum Current-Based Control Strategy[J]. IEEE Transactions on Power Delivery, 2019, PP(99):1–1.

Biography

Huaqin Wu, female, majored in computer and application, Zhengzhou University, with a Bachelor's degree in engineering in 2002; In 2008, he graduated from Huazhong University of science and technology with a master's degree in computer science and technology. He is an associate professor of Henan Technical Institute. His main research direction is software architecture and cloud computing.