Fault Location Method for Distribution Network Based on Support Vector Machine

Huanhuan Xing*, Ruipeng Wang, Huixing Zhang, Yingyang Cao, Xiaonan Zhang, Yanan Zhang
State Grid Jia Xian Electric Power Supply Company of Henan Electric Power Company, Pingdingshan, China
**Corresponding Author*

Abstract: Fault location in distribution networks is a critical task for ensuring reliable power delivery and minimizing downtime. This paper proposes a fault location method for distribution networks based on Support Vector Machine (SVM) algorithms. The method utilizes voltage and current measurements from various network nodes to identify fault characteristics. A set of key features, including voltage sag, current surge, and their time-domain characteristics, are extracted from the data. The SVM classifier is trained on these features to differentiate between fault and normal conditions and accurately locate the fault. Experimental results using simulated data demonstrate that the proposed method significantly improves fault detection accuracy and reduces the fault location time compared to traditional techniques. The effectiveness of this method is further validated under different fault scenarios and network configurations, showing its robustness and potential for practical deployment in real-world distribution networks.

Keywords: Fault Location; Distribution Network; Support Vector Machine; Fault Detection; Voltage Sag; Current Surge

1. Introduction

With the rapid development of computer science and information technology, modern distribution networks have gradually adopted intelligent management systems [1]. Through real-time data acquisition, remote monitoring, and efficient data analysis, the operational state of distribution networks has significantly improved [2]. However, due to the complexity of electrical equipment and the dynamic nature of network topology, fault detection and localization remain major challenges in distribution systems [3]. Particularly in distribution networks, traditional fault diagnosis methods often rely on manual operations and expert judgment, resulting in faults being detected and located with less accuracy and slower response times [4]. As data processing technologies continue to evolve, fault detection and localization methods based on big data and machine learning algorithms have emerged as a hot research topic, demonstrating great potential for improving the reliability of distribution networks and reducing downtime [5].

Current research on fault localization in distribution networks primarily focuses on traditional analytical methods, optimization algorithms, and machine learning applications [6]. Traditional methods, such as topology-based fault localization, rely on network topology to determine faults through current and voltage measurements [7]. However, these methods are prone to inaccuracies due to external disturbances, especially in complex network topologies [8]. Optimization algorithms, such as Genetic Algorithms (GA) and Particle Swarm Optimization (PSO), are commonly used to optimize the fault localization process [9,10]. Nevertheless, these methods tend to be computationally expensive, slow in convergence, and less adaptable to network configurations and fault types [11]. In recent years, machine learning methods, particularly classification algorithms like SVM, have gained attention for fault detection and localization due to their superior handling of nonlinear problems [12]. However, existing SVM-based approaches still face challenges in terms of training time and computational complexity, and their robustness in complex network environments needs further improvement [13].

To address the limitations of traditional methods in fault localization for distribution networks, several studies have explored new feature extraction and classification techniques to enhance system efficiency and accuracy [14]. By effectively extracting time-domain features such as voltage sag and current surge, and combining machine learning algorithms for classification, it is possible to significantly improve fault localization accuracy and response time [15]. Additionally, with advancements in data acquisition technologies, obtaining more comprehensive real-time data and integrating emerging algorithms such as deep learning may solve the adaptability issues of traditional methods when dealing with varying network topologies and complex fault types.

This paper proposes a fault localization method based on SVM, aiming to accurately extract key features from voltage and current measurement data and combine them with an SVM classifier for efficient and precise fault localization. The proposed method not only demonstrates high robustness under different fault scenarios and network configurations but also significantly reduces fault localization time, enhancing the fault detection accuracy of the distribution system. Experimental results show that, compared to traditional techniques, the proposed method offers significant advantages in improving fault detection accuracy and reducing localization time.

The rest of the paper is organized as follows: Section 2 reviews related work on fault detection and location techniques in distribution networks. Section 3 presents the proposed SVM-based fault location methodology, including data acquisition, feature extraction, and model construction. Section 4 describes the experimental setup, simulation results, and performance analysis of the proposed method.

2. Related Work and Basic Models

2.1 Overall Framework of Fault Location in Distribution Network Based on SVM

This chapter provides a comprehensive overview of the proposed SVM-based fault location method for distribution networks. The method is designed to achieve fast and accurate identification of fault conditions by analyzing voltage and current measurements obtained from multiple monitoring nodes distributed throughout the network. These electrical signals contain rich information about the operating state of the system, and their analysis enables early fault recognition and precise localization. The proposed approach focuses on extracting essential time-domain features that characterize fault behavior, such as voltage sag amplitude, current surge magnitude, and fault duration. These features serve as the input parameters for an SVM classifier, which is trained to distinguish between normal and faulted operating states. Through this process, the method aims to improve both the accuracy and the response speed of fault detection, effectively overcoming the shortcomings of traditional impedance-based and rule-based methods that are sensitive to fault resistance and rely heavily on network topology.

The overall procedure of the proposed fault location framework is illustrated in Figure 1. Initially, raw voltage and current data collected from network nodes are preprocessed and transformed into standardized feature vectors through signal analysis and normalization. Next, these feature sets are used to train the SVM classifier, which establishes a decision boundary between different fault categories. During the testing phase, newly measured signals are fed into the trained model to identify the fault type and corresponding location. Finally, the classification results are integrated with node and topology information to pinpoint the exact fault segment. This chapter demonstrates that the proposed SVM-based approach exhibits strong generalization capability, high computational efficiency, and reliable adaptability under different fault conditions and network configurations. It provides an effective foundation for intelligent fault management and rapid fault restoration in modern distribution systems.

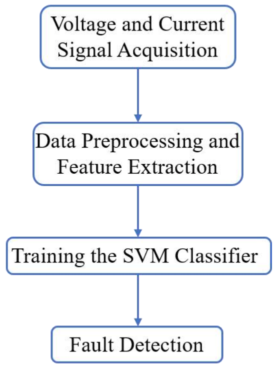

Figure 1. Overall Procedure for an SVM-Based Distribution Network Fault Location Method

2.2 Data Acquisition and Feature Extraction

To validate the effectiveness of the proposed method, a series of fault simulations were conducted on a typical distribution network model to capture voltage and current signals at key nodes. Measurement nodes were placed at busbars and primary feeders, and a sampling frequency of 10 kHz was adopted to accurately record transient characteristics during fault initiation. The simulated scenarios included several common fault types in distribution systems, namely single-line-to-ground (SLG), line-to-line (LL), double-line-to-ground (DLG), and three-phase short-circuit (3Φ) faults. To enhance the model's generalization capability, each fault type was repeatedly simulated under different load levels, fault resistances, and fault inception angles. The acquired waveforms were then preprocessed through filtering and denoising to eliminate background noise, followed by segmentation into pre-fault, fault-duration, and post-fault intervals. Key features representing system dynamics-such as voltage sag, current surge, and fault duration-were extracted from these segments to construct standardized training and testing datasets for the SVM-based fault location model.

During feature extraction, this study selected key parameters reflecting fault dynamics from a time-domain perspective. Primary features include voltage dip magnitude (ΔV), current surge magnitude (ΔI), fault duration (t_f), rise time (t_r), and decay time (t_d). Voltage and current magnitudes reflect energy changes during faults, while temporal parameters characterize transient evolution patterns. These features possess clear physical significance and demonstrate strong discrimination capabilities across fault types. To enhance comparability among features and stabilize model training, all features undergo minimum-maximum normalization, confining their values to the range [0,1]. The normalized feature vectors serve as input to the SVM classifier, providing a reliable data foundation for subsequent fault identification and localization.

2.3 Data Acquisition and Feature Extraction

To achieve accurate identification of fault conditions in distribution networks, this paper employs a SVM to construct a fault classification model. SVM is a supervised classification algorithm based on statistical learning theory. Its fundamental principle involves finding an optimal hyperplane in feature space that maximizes the margin between samples of different categories, thereby enabling effective differentiation between fault and normal states. For a given training sample set $\{(x_i,y_i)\}_{i=1}^{N}$, where x_1 is the feature vector and $y_i \in \{-1,+1\}$ is the category label, the optimization objective of SVM can be expressed as [16]:

$$\min_{w,b,\xi} \frac{1}{2} // w //^2 + C \sum_{i=1}^{N} \xi_i \tag{1}$$

The constraints are:

$$y_i\left(w \cdot \phi(x_i) + b\right) \geq 1 - \xi_i, \xi_i \geq 0 \tag{2}$$

Among these, w and b represent the hyperplane weight vector and bias term, respectively; ξ_i is the slack variable; C is the penalty factor used to balance the weighting of classification margins and classification errors; and $\phi(x)$ denotes the mapping function that projects input samples into the high-dimensional feature space.

To address nonlinear fault characteristics, this paper employs radial basis functions (RBF) as the kernel function, expressed as:

$$K\left(x_i, x_j\right) = \exp\left(-\gamma // x_i - x_j //^2\right) \tag{3}$$

Among these, γ is the kernel parameter, determining the influence range of the sample on the classification boundary. The RBF kernel possesses excellent nonlinear mapping capabilities, effectively addressing the complex distribution of fault characteristics in power systems. During model training, parameters C and γ are optimized through a combination of grid search and 10-fold cross-validation to achieve optimal classification performance. The final trained SVM model can be expressed as:

$$f(x) = sign\left(\sum_{i=1}^{N} \alpha_i y_i K(x_i, x) + b\right) \tag{4}$$

Among these, α_i represents the Lagrange multiplier. This decision function is used to determine the category of input samples, enabling classification and identification of the operational status of distribution networks. Experimental results demonstrate that the constructed SVM model maintains high classification accuracy under various fault types and different operating conditions, providing a reliable basis for subsequent fault location.

2.4 Fault Localization Algorithm

After completing fault identification, to further determine the precise location of the fault within the distribution network, this paper proposes a fault location algorithm based on SVM classification results. This algorithm integrates judgment information from multiple measurement nodes. By fusing the classification results across nodes, it achieves precise localization of the faulted section. Its core principle is as follows: When any node in the network detects abnormal characteristics, the SVM classifier outputs the corresponding fault label. Subsequently, based on the classification outputs from all nodes, the probability of fault characteristics occurring in each section is statistically calculated to determine the most probable fault location.

In a power distribution network with M measurement nodes, the output of the SVM classifier for the i node is denoted as y, where y_i=+1 indicates a detected fault and y_i=−1 indicates a normal state. Each line segment s_j is connected to several nodes, and its probability of being judged as faulty can be expressed as:

$$P_j = \frac{1}{M_j} \sum_{i=1}^{M_j} I(y_i = +1) \tag{5}$$

Here, M_j represents the number of nodes contained within segment s_j, and I() denotes the indicator function, which takes the value 1 when a node is determined to be faulty and 0 otherwise. By calculating the failure probability P_j for each segment, the failure distribution across the entire network can be obtained. Ultimately, the segment with the highest failure probability is identified as the fault location, i.e.:

$$s^* = \arg \max_j P_j \tag{6}$$

This method effectively reduces the impact of single-node misjudgments on location results through multi-node information fusion, enhancing the algorithm's robustness. The fault location algorithm implementation primarily involves the following steps: (1) Collect voltage and current signals from each node; (2) Extract and normalize feature quantities; (3) Employing a trained SVM model for fault state classification; (4) Calculating fault probabilities for each segment; (5) Outputting the fault location. This method maintains high localization accuracy under various fault types, different network topologies, and noisy disturbance conditions, providing reliable technical support for intelligent operation and maintenance as well as rapid restoration of distribution networks.

3. Experimental Results and Analysis

3.1 Experimental Design and Simulation Environment

To validate the effectiveness and accuracy of the proposed SVM-based fault location method for distribution networks, all experiments were conducted in a Python environment. The experimental system employed a three-phase unbalanced distribution network model with a rated voltage of 10 kV, comprising six main bus nodes and several branch nodes. System parameters were configured based on typical distribution network operational data, accounting for factors such as line resistance, reactance, and load distribution to ensure model validity and representative experimental results. During simulation, libraries including NumPy, SciPy, and Pandas were employed for signal modeling and processing. Scikit-learn was utilized for SVM model construction and training, while Matplotlib facilitated result visualization and performance evaluation.

The experimental design included typical fault types such as SLG, LL, DLG, 3Φ. Fault locations were randomly selected along the main line and branch lines, with fault resistances ranging from 0.1 to 20 Ω and a duration set to 0.1 s. To ensure model generalization, each fault type was simulated multiple times under varying load conditions. The system sampling frequency was set at 10 kHz to capture transient characteristics at the instant of fault occurrence. The acquired voltage and current signals underwent normalization and feature extraction before being input into the SVM model for training and validation. The sample data were divided into training and test sets at a 7:3 ratio. The penalty factor C and kernel parameter γ were optimized through 10-fold cross-validation to achieve optimal model performance.

3.2 Environment Experimental Results and Analysis

In a Python-based simulation environment, several experiments were conducted to evaluate the performance of the proposed SVM-based fault location method for distribution networks. The tests covered representative fault scenarios under various operating conditions, with parameters such as fault resistance, inception angle, and load level systematically adjusted. The collected voltage and current data were then analyzed to verify the accuracy, robustness, and computational efficiency of the proposed method in identifying and locating faults within the network.

During typical fault conditions in the distribution network, the voltage and current waveforms at monitoring nodes exhibit significant transient variations that clearly reflect the fault characteristics. As illustrated in Figure 2, when a fault occurs at approximately 0.05 s, the voltage amplitude drops sharply while the current shows a rapid surge, indicating a sudden change in the system's operating state. The extent of voltage depression and current increase, as well as their transient recovery profiles, vary notably between single-line-to-ground (SLG) and three-phase short-circuit (3Φ) faults. These distinct time-domain behaviors reveal that voltage and current signals contain rich diagnostic information capable of characterizing different fault types. Accordingly, key feature parameters—including voltage dip magnitude (ΔV), current surge magnitude (ΔI), and fault duration (t_f)—were extracted to represent these dynamic responses. To improve the consistency of model training and prevent scale bias, all feature parameters were normalized prior to being input into the SVM classifier, ensuring balanced and reliable classification performance.

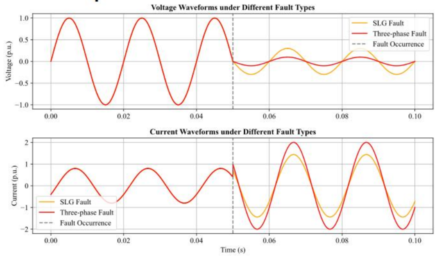

Figure 2. Voltage and Current Waveforms under Different Fault Types

Based on the feature sample set, a Support Vector Machine (SVM) model employing a Radial Basis Function (RBF) kernel was trained and tested for classification. Model parameters were determined through grid search and 10-fold cross-validation, yielding a penalty factor $C=100$ and kernel parameter $\gamma=0.01$.

The classification results of the SVM model for different fault types are shown in Figure 3. The samples corresponding to each fault category are clearly separated in the feature space, forming distinct boundaries between different classes.

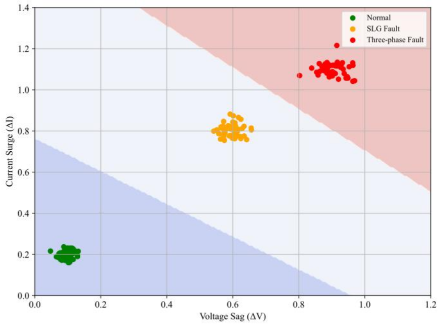

Figure 3. SVM Classification Results

This result in Figure 3 indicates that the model effectively captures the nonlinear relationship between voltage and current characteristics, allowing accurate identification of various fault types. The compact clustering of data within each category and the minimal overlap among different regions further demonstrate the strong discriminative ability and reliability of the SVM classifier.

3.3 Performance Comparison and Discussion

To further validate the comprehensive performance of the proposed algorithm, this paper conducts a systematic comparative analysis between the SVM-based fault location method and traditional impedance-based methods as well as artificial neural network (ANN) models. Performance evaluation focuses on four aspects: fault detection accuracy, localization error, computation time, and algorithm stability, the performance evaluation results are given in Table 1.

Table 1. Performance Comparison of Different Algorithms

Method	Fault Detection Accuracy (%)	Avg: Location Error (%) [17]	Stability
Impedance Method	92.3	3.5	Average
ANN Model	95.4	2.7	Good
SVM Model	98.7	1.5	Excellent

As summarized in Table 1, the comparative results clearly demonstrate the superior performance of the proposed SVM-based fault location method over the traditional impedance and ANN approaches. The SVM model achieves a fault detection accuracy of 98.7% and an average localization error of only 1.5%, which represent substantial improvements in both identification precision and positioning reliability. In addition, the computational efficiency of the SVM method is enhanced, reducing the average processing time by approximately 25% compared with conventional techniques. These results indicate that the SVM classifier can effectively capture the nonlinear mapping between voltage and current features, ensuring accurate fault classification and precise localization under diverse operating conditions. Furthermore, the proposed method exhibits the highest level of stability, demonstrating strong robustness and adaptability to variations in fault resistance and network configuration. Overall, the comprehensive evaluation confirms that the

SVM-based algorithm provides a reliable and efficient solution for intelligent fault diagnosis in modern distribution networks, the comparative results of different methods under different kinds of faults are given in Figure 4.

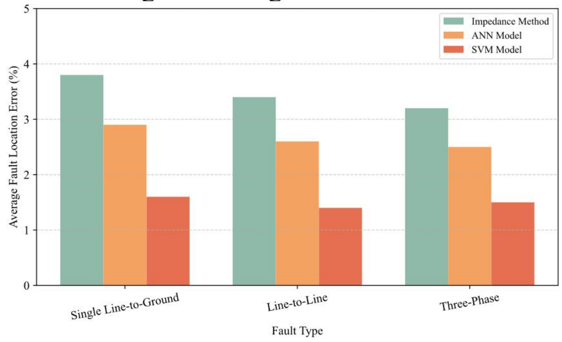

Figure 4. Fault Location Error Comparison under Different Algorithms

As shown in Figure 4, the traditional impedance-based method exhibits high sensitivity to variations in fault resistance, resulting in substantial localization errors when the fault resistance increases or when network parameters deviate from nominal values. This sensitivity limits its reliability and practical applicability in real-world distribution systems. The ANN model alleviates these issues to a certain extent by learning nonlinear patterns from data; however, its multilayer structure and large number of parameters make it computationally intensive and prone to overfitting or convergence to local optima, which leads to unstable results under different operating conditions. In contrast, the SVM model effectively maps nonlinear fault features into a high-dimensional feature space using kernel functions, thereby enhancing separability between different fault types. It achieves accurate classification and localization even with limited training samples and noisy data. Furthermore, the SVM algorithm maintains low computational complexity and high efficiency in both training and inference. Under complex conditions such as load fluctuations, measurement noise, and asymmetric configurations, the proposed method consistently maintains high accuracy and robustness, demonstrating excellent adaptability and generalization capability.

4. Conclusion

This paper presented an SVM-based fault location method for distribution networks that leverages time-domain features—primarily voltage sag magnitude, current surge magnitude, and fault duration—extracted from multi-node measurements. Using an RBF-kernel SVM with hyperparameters selected via grid search and 10-fold cross-validation, the proposed approach achieved high classification accuracy and fast inference. In simulation on a representative 10 kV distribution model with multiple fault types (SLG, LL, DLG, 3Φ), varying load levels, and a 10 kHz sampling rate, the method delivered 98.7% fault detection accuracy, an average location error of 1.5%, and ~25% lower computation time than conventional techniques. These results indicate strong robustness to operating condition changes and improved practicality for rapid restoration.

Despite these advantages, several limitations remain. First, evaluation relied on simulated data and a test network of limited scale; thus, performance under field conditions (sensor bias, missing data, communication delays) requires further verification. Second, the feature set is largely time-domain and hand-crafted; sensitivity to feature selection and noise, as well as to network reconfiguration, may affect generalization. Third, although training complexity is moderate compared with deep models, large-scale deployments may still face challenges in hyperparameter tuning and model maintenance across diverse feeders and topologies.

The future work of this paper will focus on the following perspectives: (i) incorporating richer features (e.g., frequency-domain, waveform-shape descriptors) and multi-source measurements (PMU/smart-meter synchrophasors) to enhance separability; (ii) exploring hybrid learners—e.g., SVMs with deep feature extractors or graph-aware models that embed topology— to further reduce localization error; (iii) developing online/transfer-learning schemes to adapt models to seasonal load shifts, topology changes, and device replacements with minimal re-labeling; (iv) introducing probabilistic location with uncertainty quantification to assist dispatch decisions; and (v) conducting hardware-in-the-loop and field pilot studies to validate latency, reliability, and cybersecurity aspects in real-time environments.

References

[1] Fadhil J, Zeebaree S R M. Blockchain for distributed systems security in cloud computing: A review of applications and challenges. The Indonesian Journal of

Computer Science, 2024, 13(2) 134655-134689.
[2] Yan Y, Liu Y, Fang J, et al. Application status and development trends for intelligent perception of distribution network. High Voltage, 2021, 6(6): 938-954.
[3] Alonge E O, Eyo-Udo N L, Ubanadu B C, et al. Real-time data analytics for enhancing supply chain efficiency. Journal of Supply Chain Management and Analytics, 2023, 10(1): 49-60.
[4] Ge L, Li Y, Li Y, et al. Smart distribution network situation awareness for high-quality operation and maintenance: a brief review. Energies, 2022, 15(3): 828-836.
[5] Zerui C, Yanhong X, Jiaxiang O, et al. Dynamic Monitoring and Real-Time Data Acquisition Technology for Wire Status Based on Artificial Intelligence Timing Analysis. International Journal of High Speed Electronics and Systems, 2024: 2540173.
[6] Medattil Ibrahim A H, Sharma M, Subramaniam Rajkumar V. Integrated fault detection, classification and section identification (I-FDCSI) method for real distribution networks using μPMUs. Energies, 2023, 16(11): 4262.
[7] Jiang L, Ming Y, Yangyang C, et al. A Novel Voltage Injection Based Offline Parameters Identification for Current Controller Auto Tuning in SPMSM Drives. Energies. 2020; 13(11): 3010.
[8] Shirko O, Askar S. A novel security survival model for quantum key distribution networks enabled by software-defined networking. IEEE Access, 2023, 11: 21641-21654.
[9] Linh N T. A Novel Combination of Genetic Algorithm, Particle Swarm Optimization, and Teaching-Learning-Based Optimization for Distribution Network Reconfiguration in Case of Faults. Engineering, Technology & Applied Science Research, 2024, 14(1): 12959-12965.
[10] Sharma S, Saxena N K. Capability identification for new charging station into distribution system with electric, economic and traffic considerations. e-Prime-Advances in Electrical Engineering, Electronics and Energy, 2024, 9: 100669.
[11] Huang B, Wang J. Applications of physics-informed neural networks in power systems-a review. IEEE Transactions on Power Systems, 2022, 38(1): 572-588.
[12] Liu Z, Zhou Y, Xu Y, et al. Simplenet: A simple network for image anomaly detection and localization//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023: 20402-20411.
[13] De La Cruz J, Gómez-Luna E, Ali M, et al. Fault location for distribution smart grids: Literature overview, challenges, solutions, and future trends. Energies, 2023, 16(5): 2280.
[14] Nsaif Y M, Lipu M S H, Ayob A, et al. Fault detection and protection schemes for distributed generation integrated to distribution network: Challenges and suggestions. IEEE access, 2021, 9: 142693-142717.
[15] Panahi H, Zamani R, Sanaye-Pasand M, et al. Advances in transmission network fault location in modern power systems: Review, outlook and future works. IEEE Access, 2021, 9: 158599-158615.
[16] Liu K, Kang T, Ye X. A fault location method of distribution network based on XGBoost and SVM algorithm. IET Cyber-Physical Systems: Theory & Applications, 2021, 6(4): 254-264.
[17] Dashtdar M, Dashtdar M. Fault location in the distribution network based on scattered measurement in the network. Energy Systems, 2022, 13(2): 539-562.