A PLC communication characteristics-based fault location method in MV meshed distribution networks

Mingang Tan¹, Yi Tang², Chaohai Zhang¹, Bin Chen¹, and Junliang Qian¹

¹Nanjing University of Aeronautics and Astronautics
²Southeast University

April 4, 2023

Abstract

A power line carrier (PLC) communication characteristics-based method is proposed for single-phase-to-ground fault location in neutral isolated medium voltage (MV) meshed distribution networks in this paper. The carrier signals with a time-varying frequency and constant amplitude are processed by a set of PLC transmitters and receivers, whose placement is optimized by regarding the power network as an undirected graph. Two signal encoding and decoding algorithms for the PLC terminals are proposed to avoid using expensive timing systems between the terminals. The fault location technique is implemented by comparing the cosine similarity of amplitude attenuation (AA) and phase offset (PO) between the fault and a library. The node corresponding to the maximum cosine similarity of the characteristics between the present fault and the library is selected as the location of the current fault. Only one set of low-cost PLC communication terminals and the widely available power lines are needed in the fault location system, making this approach highly practical. Numerical simulations using MATLAB/Simulink have been performed to verify the technique’s feasibility. The results show that the method can accurately locate faults in neutral isolated MV meshed distribution networks. Besides, the presented approach achieves high accuracy in estimating transition resistance values.
A PLC communication characteristics-based fault location method in MV meshed distribution networks

Tan Mingang¹*, Tang Yi², Zhang Chaohai³, Chen Bin¹, Qian Junliang⁴

¹ Dept. of Electrical Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China
² School of Electrical Engineering, Southeast University, Nanjing, China
* E-mail: tannmingang1223@163.com

Abstract: A power line carrier (PLC) communication characteristics-based method is proposed for single-phase-to-ground fault location in neutral isolated medium voltage (MV) meshed distribution networks in this paper. The carrier signals with a time-varying frequency and constant amplitude are processed by a set of PLC transmitters and receivers, whose placement is optimized by regarding the power network as an undirected graph. Two signal encoding and decoding algorithms for the PLC terminals are proposed to avoid using expensive timing systems between the terminals. The fault location technique is implemented by comparing the cosine similarity of amplitude attenuation (AA) and phase offset (PO) between the fault and a library. The node corresponding to the maximum cosine similarity of the characteristics between the present fault and the library is selected as the location of the current fault. Only one set of low-cost PLC communication terminals and the widely available power lines are needed in the fault location system, making this approach highly practical. Numerical simulations using MATLAB/Simulink have been performed to verify the technique’s feasibility. The results show that the method can accurately locate faults in neutral isolated MV meshed distribution networks. Besides, the presented approach achieves a high level of accuracy in estimating transition resistance values.

1 Introduction

With more Distributed Energy Resources (DERs) being connected to medium voltage (MV) distribution networks through Inverter-Based Power Sources (IBPS), the meshed networks are expected to become more prevalent. However, this shift from traditional radial systems to meshed networks presents novel challenges for fault location [1]. The single-phase-to-ground location in neutral isolated distribution networks has been a topic of research for a long time and there are numerous available methods [2, 3], which can be categorized into four types as follows: steady-state signal-based (SSBM) [4, 5], transient signal-based (TSBM) [6, 7], signal or disturbance injection-based (SIBM) [8], and artificial intelligence-based methods (AIMB) [9].

The SSBM locates faults by calculating the impedance between the fault point and the measurement point. It uses voltage and current measurements from the faulty circuit and determines fault distance by comparing the impedance with the unit length impedance. Different information acquisition positions result in various methods, such as single-end, double-end, and multi-end methods [10]. A classified analysis of various scenarios and fault currents for distributed power access is presented in [11], and a fault location strategy for overhead distribution networks based on current information is also proposed. Furthermore, [12] considers the characteristics of phase-to-phase short-circuit faults in active distribution networks and proposes a diagnostic method based on fault steady-state components. The SSBM is based on a simple principle and requires only basic calculations. In general application scenarios where there are enough meters in the power grid, the fault location accuracy of SSBM is sufficient.

The TSBM commonly uses the temporal characteristics of the fault waveform to calibrate the feature quantities before and after the fault on the time axis for fault location. Techniques such as calculating the time-frequency matrix of the transient waveform after the fault, measuring the transient energy within the characteristic frequency band [13], or quantifying the transient current mutation quantity can be employed to achieve fault location. It is commonly implemented through methods such as the calculation of the time-frequency matrix of the transient waveform after the fault [14], the measurement of the transient energy within characteristic frequency bands, and the determination of the abrupt change of the transient current. The method possesses high accuracy while also placing a high demand on the performance of the monitoring device.

The SIBM involves injecting a specific frequency signal into a system through a signal injection device. A faulty line will form a signal loop at the point of the fault, and the location of the fault can be determined by detecting the path that the signal flows through using a fault indicator installed on the line. A cable fault location method is proposed in [15], which injects high-frequency distributed zero-sequence current carrier signals through current transformers and has high independence and anti-interference capabilities. Various methods for analyzing reflection waveform and determining fault locations using different types of carrier reflectometry, such as time domain reflectometry (TDR), noise domain reflectometry (NDR), frequency domain reflectometry (FDR), and others are summarized in [16]. The accuracy of this method for locating stable grounding faults is high, but it is difficult to accurately detect some intermittent and transient faults. At the same time, additional signal generation and reception devices need to be installed in the power grid, which results in high investment costs.

The development of AI has led to the emergence of data-driven and AI-based methods for fault location in active distribution networks [17, 17]. These methods use historical fault knowledge models that consider multiple factors and integrate data collected from supervisory control and data acquisition (SCADA) systems for more accurate results. Examples of such methods include expert systems [18], fuzzy nets [9], Petri nets [19], and Bayesian networks [20], among others. AIBM are highly regarded because they do not require precise mathematical models and are not influenced by fault type, transition resistance, or system operating mode. However, obtaining data samples and accurately training the models pose significant challenges in practical research.

High localization accuracy is demonstrated by SSBM and TSBM in radial distribution networks, but they perform poorly in meshed
ones. Power outages are required by SIBM, which should disrupt normal user production and incur high implementation costs. Significant time costs are imposed by AIBM and the final performance is difficult to ensure. The power line carrier (PLC) communication technology utilizes the widely available power lines in the power grid as a physical channel, eliminating the need for additional investment [21, 22]. Twisted-pair cables are typically used in the distribution network, which possesses excellent anti-interference capability and transmission stability, ensuring the quality of communication. The above advantages make PLC an economical, practical, and reliable communication technology. In order to quickly and accurately locate faults in a meshed power grid with a minimum impact range, a low-cost fault location method based on PLC communication characteristics is proposed in this article. This method requires only one set of PLC terminals and is effective under both live and power-off conditions. It addresses the cost and several other shortcomings of the mentioned method, and its main contributions are highlighted below.

1) The low-cost PLC terminals and the widely available power lines are used for fault location.

2) The PLC communication indexes, amplitude attenuation (AA), and phase offset (PO), are chosen as fault characteristics.

3) Examining the AA and PO of the PLC communication system by utilizing the signals with regular changing of frequency and duration, and further locating faults by cosine similarity.

4) Determining the placement nodes of PLC communication terminals using path search methods in graph theory.

The rest of the article is sectioned as follows: System modeling and problem formulation for PLC communication characteristics-based fault location method are introduced in Section 2. Section 3 demonstrates the simulation results and discussions about the effectiveness of the proposed method. The conclusion is conducted in the last section.

2 Methodology

2.1 Theoretical Framework

As shown in Figure 1, a typical PLC communication process involves three parts: a sending modem (SM), a receiving modem (RM), and the power lines (PLs). SM modulates the data to be transmitted into a high-frequency signal and loads it onto the PL through a coupling capacitor in conjunction with a filter. RM extracts the corresponding frequency band carrier by using a combined filter and demodulates it to obtain the elemental data. PLs provide the communication channel for high-frequency signals. The main idea of this article is to locate the bus where a fault has occurred by varying the signal frequency at the SM and then detecting the AA and PO at the RM.

According to the information nodes-based [21] technique, a PLC communication channel modeling method for distribution networks to address the challenge of decomposing loop or mesh networks into multiple sub-networks in the frequency domain model is proposed. In this method, each PL is considered a sub-network, and the transmission equation is established using the transmission line model. Equations are also established for the remaining branches using Kirchhoff’s voltage and current laws.

Assuming a power grid with \( m \) PLs and \( n \) loads in the power grid, a total of \( 2m + n + 2 \) information nodes are created, each with two unknowns: node voltage and branch current. Therefore, the total number of unknowns in the network is \( 4m + 2n + 4 \), and Kirchhoff’s voltage and current laws can be used to establish \( 4m + 2n + 3 \) linearly independent equations as shown in Eq. 1 and 2. In Eq. 1, \( E \) represents the information node matrix of the topology, where the element \( e_{ij} \) demonstrates the connection relationship between information nodes \( i \) and \( j \). \( U_s \) and \( I_x \) (\( x = S, 1, 2, \ldots, 2m + n + 1 \)) in Eq.2 refer to the node voltage and the current flowing out of node \( x \).

Therefore, the ratio of any two physical quantities can be obtained, such as \( \frac{U_{y,x}}{x} \) (\( x = 1, 2, \ldots, 2m + n + 1 \)). In other words, the AA(f) and PO(f) in Eq.3 can be characterized as expressions that are dependent on both the communication frequency and the topological structure parameters (\( E \) in Eq.1).

\[
\begin{align*}
\text{AA}(f) &= 20 \log \frac{|U_{RM}(f)|}{|I_{RM}(f)|} \\
\text{PO}(f) &= \arctan \left( \frac{\text{Im}(f)}{\text{Re}(f)} \right)
\end{align*}
\]

\( f \) represents the communication frequency of the PLC, while AA(f) and PO(f) respectively indicate the amplitude attenuation and phase offset of RM relative to SM in the frequency domain. \( \text{Im}[-] \) and \( \text{Re}[-] \) denote the operators for obtaining the imaginary and real parts of a complex number separately. \( U_{SM} \) and \( U_{RM} \) individually represent the carrier signals transmitted from SM and received at RM under a specific \( f \).

When a fault occurs at a specific node within the power grid, it can be modeled as the introduction of a load at the faulted node. As a result, shown in Eq.4 and 5, the number of equations and unknowns in the system increases by two each. However, despite these additional complexities, the values for AA(f) and PO(f) in Eq.3 can still be determined.

\[
E' = \left[ \begin{array}{c|c}
0 & \cdots & -1 & \cdots & 0 & 1 & 0 \\
0 & \cdots & 1 & \cdots & 0 & 1 & r
\end{array} \right]
\]

\[
E' \cdot [P] \cdot U_{2m+n+2} = 0 \quad (5)
\]

where, \( P = \{ U_s, U_1, \ldots, U_{2m+n+1}, I_s, I_1, \ldots, I_{2m+n+1} \} \). \( U_{2m+n+2} \) and \( I_{2m+n+2} \) refer to the voltage and current at the faulted node, and \( r \) represents the transition resistance of the fault. There is an element in the \( km + 2n + 4 \)-th row of the matrix \( E' \) that is -1, the column in which the element is located corresponds to the column number of the fault voltage.

In Eq.5, the relationship between AA(f) and \( E' \), and PO(f) and \( E' \) are established. When the fault location remains constant, the only variable in matrix \( E' \) is \( r \) (several elements in matrix X'). Therefore, the values of AA(f) and PO(f) in the frequency domain is solely influenced by \( r \). Therefore, this paper posits that faults occurring at the same location in a certain power network may exhibit similarities in their AA(f) and PO(f) characteristics. Accordingly, a fault location method based on cosine similarity is proposed in this paper. The cosine similarity between matrix \( X \) and \( Y \) is defined as follows:

\[
\cos(X, Y) = \frac{X \cdot Y}{|X||Y|} = \sum_{i=p+1}^{n} \sum_{j=1}^{q} x_{ij} y_{ij} = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij} y_{ij}}{\sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij}^2} \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} y_{ij}^2}}
\]

\( X \) and \( Y \) are two matrices of identical dimensions, with \( p \) rows and \( q \) columns. The elements of \( X \) and \( Y \) are represented by \( x_{ij} \) and \( y_{ij} \), respectively. The cosine similarity of any two non-zero matrices with the same dimensions numerically falls within the range of -1 to 1. A value of \( \cos(X, Y) = 1 \) indicates that the elements of the two matrices are linearly correlated, whereas a value of \( \cos(X, Y) = 0 \) suggests that the two matrices are completely independent. On the other hand,
cos (X, Y) = -1 reflects that the elements of the two matrices vary in opposite directions, while other values indicate some degree of similarity between X and Y.

In this paper, we propose a fault location method based on establishing a fault feature library in the frequency domain for each node position. When a fault occurs in the system, the features of the fault are compared with those in the library through similarity calculations. The node corresponding to the maximum similarity with the current fault in each library is selected as the location of the current fault. This method provides a reliable and efficient approach to identifying the location of faults in power systems.

### 2.2 Placements of SM and RM

The method described in [23, 24] requires monitoring devices to be installed on each circuit, and it is also described to install multiple monitoring devices in [25]. The method proposed in this article only requires one set of PLC communication equipment, namely one SM and one TM. To achieve better fault location results, this article proposes a method for placing the SM and TM of the PLC communication system.

The method presented in this paper can be considered a variation of SIBM. Specifically, the feature data becomes more sensitive to the fault as the signal passes through more nodes in the channel from SM to RM in a specific network topology. This increased sensitivity to faults can aid in locating the spot accurately. Additionally, since signals tend to flow along branches with smaller impedance, this paper maps the circuit network to a weighted graph and employs the breadth-first search (BFS) method from graph theory to solve the single-source shortest path problem of the graph.

1) **Mapping the electric power network to a weighted graph:**

The elements in power networks can be classified into three categories: bus, transmission line, and load. The buses correspond to nodes in graph theory, while the transmission lines and loads can be mapped to weighted edges connecting two nodes due to their impedance. The weight of each edge is equal to the impedance between the two nodes. Therefore, a weighted undirected graph can be constructed to represent the power network:

\[ G = (V, E, W) \]  

where \( V \) and \( E \) are sets of all nodes and edges, respectively, and \( W \) is a set of edge weights. Figure 2 shows an example of a power grid topology with six buses. The impedance of each transmission line is indicated after the “-” on the label. The corresponding sets \( V, E, \) and \( W \) [26] are as follows:

\[ V = \{1, 2, 3, 4, 5, 6\} \]  

\[ E = \{L_1 : e_{1,2}, L_2 : e_{2,4}, L_3 : e_{2,3}, L_4 : e_{3,4}, L_5 : e_{3,5}, L_6 : e_{4,5}, L_7 : e_{5,6}\} \]  

\[ W = \{w_{1,2} : 1, w_{2,4} : 2, w_{2,3} : 3, w_{3,4} : 1, w_{3,5} : 5, w_{4,5} : 4, w_{5,6} : 1\} \]

2) **Traversing the path with the smallest impedance between any two nodes using BFS:**

**Fig. 2:** The meshed network diagram with six nodes

In graph theory, the shortest path is the path with the minimum total weight among all possible paths. For example, suppose someone is driving from City A to City B and there are multiple routes to take. The shortest path refers to the route with the smallest total weight, which can be interpreted as the total distance traveled. Many BFS methods can be used to solve the aforementioned problem, and Dijkstra’s algorithm [26] is adopted in this article. Assuming that node 1 is chosen as the origin (s), the shortest paths from \( s \) to all vertices in \( V \setminus \{1\} \) as are as follows:

\[ \text{short}[2] : 1 \rightarrow 2, \Rightarrow \text{smallest}[1 \rightarrow 2] = w_{1,2} = 1 \]  

\[ \text{short}[3] : 1 \rightarrow 2 \rightarrow 3, \Rightarrow \text{smallest}[1 \rightarrow 3] = w_{1,2} + w_{2,3} = 4 \]  

\[ \text{short}[4] : 1 \rightarrow 2 \rightarrow 4, \Rightarrow \text{smallest}[1 \rightarrow 4] = w_{1,2} + w_{2,4} = 3 \]  

\[ \text{short}[5] : 1 \rightarrow 2 \rightarrow 4 \rightarrow 5, \Rightarrow \text{smallest}[1 \rightarrow 5] = w_{1,2} + w_{2,4} + w_{4,5} = 7 \]  

\[ \text{short}[6] : 1 \rightarrow 2 \rightarrow 4 \rightarrow 5 \rightarrow 6, \Rightarrow \text{smallest}[1 \rightarrow 6] = w_{1,2} + w_{2,4} + w_{4,5} + w_{5,6} = 7 \]

where \( \text{short}[a] \) represents the shortest distance from the origin \( s \) to vertex \( s \) in set \( V \setminus a \), which corresponds to the smallest value of impedance between bus \( A \) and \( B \) in the circuit diagram, denoted as \( \text{smallest}[A-B] \).

3) **Determining the placement nodes of SM and TM in a power network:**

For each element in \( V \), consider it as the source node \( s \) and apply the above method to find the shortest path from \( s \) to the remaining ones. For each \( s \), count the number of nodes that each shortest path traverses. Identify the origin and target vertices that cover the most number of nodes among all paths as the positions of TM and SM. The corresponding maximum coverage node numbers are shown in Table 1 when each node in \( V \) is taken as the origin point.

It should be noted that when there are multiple shortest paths with the same total distance, the path that covers the most nodes is selected as the shortest path. For example, in the two paths between nodes 1 and 3 in Figure 3, \( \{1 \rightarrow 2 \rightarrow 3\} \) and \( \{1 \rightarrow 2 \rightarrow 4 \rightarrow 3\} \), the latter covers more nodes with a total of 4 compared to the former’s
Table 1: The maximum coverage node numbers correspond to nodes.

<table>
<thead>
<tr>
<th>Source node</th>
<th>Target node</th>
<th>Coverage nodes</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6</td>
<td>[1, 2, 4, 5, 6]</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>[2, 4, 5, 6]</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>[1, 2, 3, 4]</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>[1, 2, 4]</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>[1, 2, 4, 5]</td>
<td>4</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>[1, 2, 4, 5, 6]</td>
<td>5</td>
</tr>
</tbody>
</table>

Fig. 3: The meshed network diagram with six nodes

3 although their distances are the same. Therefore, in Table 1, the latter is chosen as the shortest path.

The shortest path between nodes 1 and 6 in Table 1 covers the most nodes and will be selected as the location for placing SM and RM, with the option of interchanging their positions.

### 2.3 Proposed Method

Figure 4 shows the flowchart for the implementation stages of the proposed PLC channel-based fault location algorithm. It starts with stage 1 where the topology data of the power system is obtained, and then the model is completed in the simulation software. The placement of PLC terminals is optimized with the highest coverage of nodes using the graph theory-based minimum path search algorithm. In order to generate a fault library, the communication frequency and fault nodes are first set, and then random values of the fault resistances are sampled. After that, all faults are simulated, and the AA and PO at the PLC receiver are extracted. In stage 2, fault x is set with random transition resistance at arbitrary nodes on the basis of the power system model, and the AA and PO are simulated at various frequencies. In stage 3, the location of the fault is determined by computing the cosine similarity between its characteristics and those of faults in the library. The fault features in the library with the same nodes are considered as the same region. The cosine similarity between the features of the current fault and each feature set in each region of the library is calculated. The cosine similarity is further averaged within each region. The maximum average cosine similarity among all the regions is identified, and the corresponding node in the region is determined as the location of the current malfunction.

In stages 1 and 2, calculations involving AA and PO are required. The implementation process is as follows: when a drop in the grid frequency voltage is detected by the SM, a fixed-length handshake protocol is first sent out to agree with the RM on the starting time of frequency scanning. Then, based on the frequency points to be scanned, a set of time-domain signals with a constant amplitude of \( A_m \) and frequency is encoded into a frequency-varying complex signal by the SM. The arranged signal is transmitted by the SM end and received by the RM end. Features are then decoded and calculated from the received complex signal by the RM based on the corresponding parameters and signals. The encoding and decoding processes are further detailed in Algorithm 1 and Algorithm 2, respectively.

As shown in Figure 5, the waveform \( u_{SM}(t) \) transmitted by the SM end located at node 1 and the waveform \( u_{RM}(t) \) received by the RM end located at node 6 are presented. Figure 5 shows the voltage waveform \( u_{SM}(t) \) at SM and \( u_{RM}(t) \) at RM when faults are arranged at nodes 1 to 6, respectively. The scanning frequency ranges from 1 MHz to 50 MHz, and the fault transition resistance is set to 1 ohm. Indeed, there are 50 complete sine waves in Figure 5, and their frequency increases in steps of 1 MHz. The difference of \( u_{RM}(t) \) at RM is caused by the amplitude attenuation and phase offset induced by the change of fault node.

Algorithm 1 Algorithm for encoding composite signals at the SM end

**Require:** \( N_f \): The number of scanned frequencies; \( N_c \): The number of cycles for each frequency; \( A_m \): The number of samples per cycle at maximum frequency; \( f_c \): The initial frequency and its interval; \( L(m) \): The amplitude of all single-frequency signals

**Ensure:** \( u_c \): Composite signal

1: \( f_s = (np.arange(N_f)+1) \times f_c \)
2: \( t_s = 0 \)
3: for \( f = f_s[0] \rightarrow f_s[-1] \) do
4: \( \begin{align*} & t_s = t_s + N_c/f \\ & t_p = f_s/[0] \rightarrow f_s[-1] \end{align*} \)
5: end for
6: \( t = np.linspace(0, t_s, int(np.ceil(N_f \times N_c \times t_s))) \)
7: \( t_p = f_s/[0] \rightarrow f_s[-1] \)
8: \( t_u = unit(t) \cdot unit(t - t_p) \)
9: \( u_c = A_m \times np.cos(2\pi \times f_s[0] \times t) \times t_u \)
10: for \( f = f_s[0] \rightarrow f_s[-1] \) do
11: \( \begin{align*} & t_s = t_s + t_p \\ & t_p = f_s/[0] \rightarrow f_s[-1] \end{align*} \)
12: \( t_u = unit(t - t_s) \cdot unit(t - t_p) \)
13: \( u_c = A_m \times np.cos(2\pi \times f_s[0] \times t) \times t_u \)
14: end for
15: return \( u_c \)

where np.arange(\( N \)) generates a sequence of integers from 0 to \( N-1 \) with an interval of 1; np.linspace(start, stop, N) is used to generate a set of \( N \) equally spaced numbers from start to stop; unit(\( t \)) is a step function; np.cos() is a cosine function. And np.fft(data, frequency) is used to calculate the amplitude and initial phase of a signal with a frequency in the data. Hence, the main technical contributions and advantages of the proposed fault location method are highlighted below.
1) The approach presented in this article can be applied to any type of meshed or radial power network by introducing the PLC information node model.

2) It is only required one set of PLC communication systems, which includes one SM and one RM, to implement the fault location device easily. The placement strategy for these components is also provided based on the graph theory principles.

3) Once the signal parameters for localization are determined, the SM only needs to transmit signals, while fault location can be accomplished by performing PO and AA calculations at the RM end.

4) The need for a timing system is eliminated in both the SM and RM, as relative timestamps are used as the time reference in frequency point signal scanning.

5) The similarity-based fault location method performs better for complex networks because the proportion of fixed elements in the matrices representing AA and PO features of the RM terminal increases with complex topology.

3 Case study

3.1 Feasibility analysis

In this section, the proposed method is applied in two systems: meshed and radial test feeders mentioned in [21], and all line parameters of the test systems remain constant. Firstly, the feasibility of the proposed method in this paper was verified using the meshed test feeder.

Based on the statistics presented in Table 1, nodes 1 and 6 were selected as the placement points for PLC terminals. The input parameters for algorithm 1 were presented in Table 2. A sequence of fault transition resistances is generated by creating 50 numbers ranging from 0.01 to 1, calculated using the function logspace(-3,0) in MATLAB. The fault feature library, which is based on the above transitional resistance, is shown in Figure 6. This figure illustrates the changing process of AA and PO under various faults with their respective transition resistances.

### Table 2 The input parameters for algorithm 1.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$A_m$</th>
<th>$N_f$</th>
<th>$N_c$</th>
<th>$N_s$</th>
<th>$f_b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Values</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>64</td>
<td>1e6</td>
</tr>
</tbody>
</table>

150 random transition resistors ranging from 0.001 to 1 ohm were selected and evenly distributed among 6 nodes (nodes 1-6) to trigger faults. Subsequently, 25 fault occurrences were recorded in each node and their cosine similarity with 6 sets of fault libraries was illustrated in Figure 7. It is evident from the study case that the fault that occurred at node 6 ($x = 1, 2, 3, 4, 5, 6$) has the highest cosine similarity with the corresponding fault features in the fault library. Out of the 150 reported faults, only one was erroneously localized at node 1 (0.93664182) instead of its actual location at node 6 (similarity score of 0.93213024). However, fault localization for all other nodes was entirely accurate, resulting in a remarkable fault localization accuracy of 99.33%.

Based on the above method, when the tree and ring network described in reference [21] are adopted, the corresponding PLC terminal configurations and localization results are obtained as shown in Table 3, with the algorithm input parameters presented in Table 2. Therefore, the method proposed in this article, which involves establishing a fault feature library and utilizing the cosine similarity between faults as the criterion for fault localization, has proven to be effective.
Fig. 6: The fault feature database when SM and RM are placed at nodes 1 and 6 respectively

(a) Feature library of AA

(b) Feature library of PO

Fig. 7: The average similarity between the fault characteristics occurring at each node and the features in the library section.

Table 3 The fault location accuracy under different network topologies.

<table>
<thead>
<tr>
<th>Network</th>
<th>TM node</th>
<th>RM node</th>
<th>Size of library</th>
<th>Number of faults</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tree</td>
<td>1</td>
<td>2</td>
<td>400</td>
<td>200</td>
<td>1.00</td>
</tr>
<tr>
<td>Ring</td>
<td>1</td>
<td>4</td>
<td>400</td>
<td>200</td>
<td>0.995</td>
</tr>
<tr>
<td>Mesh</td>
<td>1</td>
<td>6</td>
<td>300</td>
<td>150</td>
<td>0.993</td>
</tr>
</tbody>
</table>

3.2 Impact of PLC terminal placement

In order to examine the influence of SM and RM positioning on fault localization, this paper takes the network illustrated in Figure 2 as an example. After randomly selecting 10 deployment schemes, the resulting cosine similarity, and localization result statistics are presented in Figure 8 and Table 4, respectively.

Table 4 The fault location accuracy under different SM and RM positions.

<table>
<thead>
<tr>
<th>Category</th>
<th>SM</th>
<th>TM</th>
<th>Covered nodes</th>
<th>Number</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1, 2</td>
<td>2</td>
<td>0.9933</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>3, 4</td>
<td>2</td>
<td>0.9933</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>5</td>
<td>2, 4, 5</td>
<td>3</td>
<td>0.9967</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>5</td>
<td>3, 4, 5</td>
<td>3</td>
<td>0.9967</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>4</td>
<td>1, 2, 4</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>6</td>
<td>3, 4, 5, 6</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>5</td>
<td>1, 2, 4, 5</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>6</td>
<td>2, 4, 5, 6</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>6</td>
<td>1</td>
<td>1, 2, 4, 5, 6</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>6</td>
<td>1, 2, 4, 5, 6</td>
<td>5</td>
<td>1</td>
</tr>
</tbody>
</table>

Based on the statistical results in Table 4, it can be observed that the accuracy of the proposed method for fault localization can reach 0.99 when the paths of SM and RM cover only two nodes. Moreover, as the number of covered nodes increases, the accuracy of fault localization quickly reaches 1. Although the accuracies of schemes 5-10 are all 1, it can be seen from Figure 8 that the cosine similarity of scheme 10 (with the most covered nodes) is more capable of distinguishing the specific location of the fault. Therefore, the placement schemes of SM and RM have a significant impact on the discriminability of cosine similarity and the accuracy of fault localization. The results indicate that the placement scheme that covers the maximum number of nodes performs the best in terms of fault localization accuracy.
3.3 Comparison of characteristics

The fault location method mentioned above utilizes two physical quantities, AA and PO, which are combined to form a characteristic list. Based on this characteristic list, the cosine correlation degree between the fault and the library is calculated. To analyze the rationality of the proposed method, this paper separately utilizes AA and PO as characteristic parameters. It calculates the cosine similarity of the characteristics between 150 random faults and the library as shown in Figure 9. The corresponding fault location accuracy under the three different feature conditions mentioned above is shown in Table 5.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>AA</th>
<th>PO</th>
<th>PO and AA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accuracy</td>
<td>0.86</td>
<td>0.9733</td>
<td>0.9933</td>
</tr>
</tbody>
</table>

Table 5: The fault location accuracy under different characteristics.

Figure 9 clearly indicates that when only AA is used as the fault characteristic, the cosine similarity of features between the faults at nodes 1, 2, 5, and 6 and the library are close, leading to low distinguishability. However, when PO is used alone, it performs significantly better than AA. Moreover, using both AA and PO as fault characteristic parameters displays the most distinguishable cosine similarity, as demonstrated by the fault location accuracy in Table 5 for all three. Hence, it is essential to consider both AA and PO as fault characteristic parameters in this paper.

3.4 Estimation of transition resistance

Based on the equation, it can be deduced that once the fault node is identified, the only parameter that undergoes variation in matrix $E'$ is $r$. Following the aforementioned localization principle, this study endeavors to leverage the cosine similarity of AA and PO to estimate the transition resistance value of the fault point. In this study, the deployment scheme of the 10th PLC terminal in Table 4 is selected. The fault feature library contains transition resistance values distributed in logspace(-3, 1, N). The test group’s transition resistance values are obtained by selecting a list of 20 numbers from logspace(-3,1,22), excluding the first and last values. The relative estimation errors of each transition resistance value in the test group are presented in Figure 10 for varying $N$ values of 50, 100, 200, and 500.

Based on the results presented in Figure 10, it can be observed that the relative errors associated with all values of $N$ are below 5%. The estimation error of the transition resistance gradually decreases as $N$ increases. When $N$ is less than 400, the decreasing trend of the relative error is more obvious. And it is maintained at about 0.1% after $N$ reached 400. In summary, the method proposed in this paper, which is based on the cosine similarity of PLC characteristics, achieves a high level of accuracy in estimating transition resistance values when there are a sufficient number of transition resistance library features, provided that the faulty node location can be accurately determined.

4 Conclusion

In this paper, a novel PLC characteristics-based fault location technique for meshed networks is proposed based on the cosine similarity of amplitude attenuation and phase offset between the fault and library. The PLC signals with time-varying frequency and constant amplitude are injected into the power lines along which the signals travel to the receiving terminal where they are picked up and treated by the filter to get the information for fault location. Only one set of PLC communication terminals is needed in the system, which greatly simplifies the feasibility of the proposed approach. The following is a summary of the results of this paper:

By the information node-based PLC channel model, a theoretical framework is constructed to incorporate faults as a distinct type of branch in the mathematical analysis. The conjecture is put forth that

![Fig. 8: The comparison of cosine similarity scores for SM and RM at varying positions](image)
**Fig. 9**: The comparison of cosine similarity scores under different characteristics

**Fig. 10**: The relative estimation error for different transition resistance values.

Fault localization could be achieved through the use of amplitude attenuation and phase shift as identifying characteristics.

The topology of the power grid was mapped into a weighted undirected graph by regarding bus categories as nodes, power lines as edges, and the impedance of each line as the weight of its corresponding edge. The placement nodes of PLC communication terminals are determined using path search methods in graph theory.

An algorithm flow for fault location using PLC signal characteristics is proposed, with signal encoding and decoding processes described for both the PLC transmitter and receiver.

Numerical simulations based on MATLAB/Simulink have been performed. The results verify the correctness and feasibility of the novel technique. In future works, more detailed simulations and experimental model tests considering every situation and real application will be performed. A full-scale hardware prototype will be constructed and evaluated online.

**Author contributions**

Tan Mingang: Conceptualization, Formal analysis, Investigation, Methodology, Validation, Writing - original draft, Writing - review & editing. Tang Yi: Conceptualization, Methodology, Data curation, Supervision, Writing - review & editing. Zhang Chaohai: Supervision, Funding acquisition, Validation, Writing - review & editing. Chen Bin: Investigation, Software, Writing - review & editing. Qian Junliang: Investigation, Writing - original draft.

**Conflict of interest statement**

The authors declare no conflicts of interest.

**Funding information**

The work was supported by the National Key R&D Program of China (No. 2021YFF0603100)

**Data availability statement**

The data that support the findings of this study are available from the corresponding author upon reasonable request.
References