Electrical Power and Energy Systems 81 (2016) 459–468

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes

A new fault location method for distribution networks using sparse measurements

S. Jamali *, A. Bahmanyar

Center of Excellence for Power System Automation and Operation, School of Electrical Engineering, Iran University of Science and Technology, Tehran 16846-13114, Iran

ARTICLE INFO

Article history:
Received 28 August 2015
Received in revised form 8 December 2015
Accepted 25 February 2016

Keywords:
Distribution networks
Fault location
Measurement uncertainty
Outage management
State estimation

ABSTRACT

In response to the growing demand to improve reliability and quality of power supply, distributed monitoring devices are gradually being implemented in distribution networks. On the other hand, utilities are demanding more accurate and reliable fault location systems to reduce the economic impact of power outages. This paper presents a novel method that takes full advantage of all available measurements to provide an accurate fault location. The developed method first uses an iterative state estimation based algorithm to find the nearest node to the fault location. It then examines all lines connected to the selected node and locates the fault. The performance of the proposed method is studied by simulation tests on a real 13.8 kV, 134-node distribution system under different fault scenarios. The results verify the accuracy of the algorithm and its robustness even under uncertain measured data. The method robustly handles measurement errors, and is applicable to any distribution network with laterals, load taps and heterogeneous lines.

Introduction

Power distribution networks, because of their geographical dispersion in urban and rural areas, can be significantly affected by faults arising from different sources such as adverse weather conditions, bird contacts, vegetation growth and equipment failure. Over 80% of customer service interruptions are owing to faults on distribution networks; thus, in order to minimize outage times and improve the continuity of supply, distribution network automation has been applied to enhance the reliability, efficiency, and quality of power supply. In this context, fault management can be stated as the core of distribution network automation. As one of the main building blocks of the fault management, fault location enables fast service restoration and narrows down the search area to find the fault point.

Considerable studies have been devoted to the development of practical methods of fault location, thereby reducing the average outage time and hence improving the reliability. These studies can be categorized into two main groups. The first group, also known as outage mapping, are a combination of techniques applied to narrow down the search area by using various available data sources like customer outage calls, weather data or fault indicator signals to estimate the most likely affected area [1–3]. The second group utilize field measurements to locate the fault and can be classified into impedance based algorithms [4–7], methods based on traveling waves [8–11], artificial intelligence based methods [12,13] and algorithms based on sparse voltage measurements [14–18].

Because of their simplicity and practical feasibility, impedance based algorithms are the most widely used type of fault location methods. However, due to the branched nature of distribution networks, these algorithms are prone to multiple location estimation problem. Installation of fault indicators can solve the multiple estimation problem [19], but increases the implementation cost. Other solutions such as injection of two sinusoidal signals with different frequencies [20] or fault diagnosis based on the fault current patterns [6] would be more cost effective. Traveling wave based methods produce accurate results for networks with small configurations and very limited number of laterals and branches [10,11]. However, in distribution systems with a large number of laterals and load taps, these methods often require very high frequency sampling rate to identify the exact fault location. Methods based on artificial intelligent systems, such as neural networks, despite their accuracy and simplicity, require large training data and retraining following a change in the distribution system topology [21].

* Corresponding author.
E-mail addresses: sjamali@iust.ac.ir (S. Jamali), bahmanyar@iust.ac.ir (A. Bahmanyar).

http://dx.doi.org/10.1016/j.ijepes.2016.02.046
0142-0615/

Recent advances in metering and communication systems, and the advent of Intelligent Electronic Devices (IEDs) such as power quality meters, digital protective relays and digital fault recorders have greatly improved monitoring and protection of modern distribution networks and have provided new opportunities to enhance fault location methods. Accordingly, another class of fault location methods has been proposed trying to benefit from the sparse voltage measurements, provided by IEDs, in addition to the voltage and current measured at the upstream substation in order to overcome some of the aforementioned problems in the previous algorithms [14–18].

The fault location method proposed in [16], first estimates fault current by summing fault current contributions from all sources. It then injects the calculated current at all system nodes to calculate the change in three phase voltages at all measurement nodes. Finally, comparing the measured values with calculated values, the method identifies the faulted node. This work has a simple procedure; however, its approximate estimate of fault current affects the accuracy of the results. Authors in [18], propose a method with the same principles, but instead of estimating the change in voltages, they estimate the fault currents and identify the faulted node by comparing the estimated currents. The proposed method has acceptable results with or without synchronization of the measured values, though, it requires a large number of meters.

The fault location methods proposed in [14,17] are based on the fact that each fault causes voltage sags with different characteristics at different nodes. Therefore, knowing the voltage sag magnitudes at certain measurement nodes, it would be possible to locate the faulted node. The algorithm assumes the fault at each node throughout the network and calculates voltage sags using a load flow program. It then determines the faulted node by comparing how well the calculated values for each node match the measured values. The proposed methods successfully find the nearest node to the fault without synchronization of the measured values, but they cannot identify the actual location of the fault.

The method presented in [15] follows the same principles. For each node, the method first employs a set of short circuit analysis to estimate fault resistance. It then applies the estimated resistance and explores the similarity between the measured and calculated voltage sags to find the faulted node. In the next step, the algorithm considers the lines connected to the selected node, moves the fault along these lines and locates the fault. The method is accurate and can identify the exact fault location; however, using three iterative stages for estimation of fault resistance, identification of faulted node and exact fault location increases the computational burden, especially for large networks.

Despite the shortcomings mentioned, all of the previously mentioned works provide acceptable results. However, since all of these methods rely on indices defined based on measured voltages, they would be very sensitive to measurement inaccuracies. Therefore, whilst all of the previously proposed algorithms use measurements only to define their indices, the main idea of this work is to incorporate all measured values into calculations through a state estimation based algorithm to improve both the accuracy and reliability of the results even under uncertain conditions. Each time a fault is detected, the developed method identifies the nearest node to fault location using an iterative state estimation based algorithm. It then examines all lines connected to the selected node and finds the exact fault location. The method has been tested by simulation studies on a 134-node, real, radial distribution network. The results verify accuracy and robustness of the method for different fault types, positions, and resistances even when there are measurement and load estimation errors.

The rest of the paper is organized as follows: the concept of Branch-current-based State Estimation (BSE) is described in Secti on 'Branch-current-based state estimation'. Section 'The algorithm' presents the outline and a detailed description of the proposed fault location method. Case study is given in Section 'Case study' where the performance of the proposed method is evaluated and Section 'Conclusion' concludes the paper.

Branch-current-based state estimation

State estimation is the process of determining the most likely state of a system by using its mathematical model and a set of measurements, which may include any combination of current, voltage and power measurements.

The measurements relate to system states by a set of nonlinear functions:

$$z = h(x) + e \tag{1}$$

where $h(x)$ is the vector of nonlinear functions, x is the vector of system states, and e is the vector of measurement errors.

Branch-current-based state estimation [22,23], commonly used in distribution networks, is developed based on the weighted least squares method and uses branch currents as state variables:

$$x = [I_r, I_x] \tag{2}$$

where I_r and I_x are the vectors of real and imaginary parts of the branch currents.

BSE estimates the state variables (i.e. branch currents) by minimizing the following objective function:

$$j(x) = \sum_{i=1}^{m} w_i (z_i - h_i(x))^2 = [z - h(x)]^T W [z - h(x)] \tag{3}$$

where W is a weighting diagonal matrix with w_i elements.

Minimizing this objective function leads to the best possible values of state variables using all available measuring resources. In order to minimize $j(x)$, the first-order optimality conditions have to be satisfied:

$$g(x) = \frac{\partial j(x)}{\partial x} = -H^T(x) W [z - h(x)] = 0 \tag{4}$$

where $H(x) = \left[\frac{\partial h(x)}{\partial x}\right]$.

Expanding the nonlinear function $g(x)$ into its Taylor series and ignoring the higher order terms leads to an iterative solution scheme known as Gauss–Newton method:

$$x^{k+1} = x^k - \left[G(x^k)\right]^{-1} g(x^k) \tag{5}$$

where:

$G(x) = \frac{\partial g(x)}{\partial x} = H^T(x) W H(x)$ is the gain matrix [24].

Starting from an initial guess, the state estimation algorithm iteratively updates the state variables until the objective function is minimized. The details of BSE are described in [22,23].

The algorithm

The proposed fault location algorithm is developed based on the concept of the branch-current-based state estimation method. As shown in the flowchart of Fig. 1, the input data is first checked for identification and elimination of bad data from the measurement set. Then, the faulted zone is determined as a group of neighboring nodes suspected of being the fault location. The main fault location process begins in the next step where all the suspected nodes are ranked by applying the fault at each of them and calculating a predefined index. The node with the largest value of index is identified as one end of the affected line. During this process, the fault is modeled as a special load temporarily connected to each analyzed node, one at a time, and its current is estimated using an iterative state estimation algorithm. Finally, the list of ranked

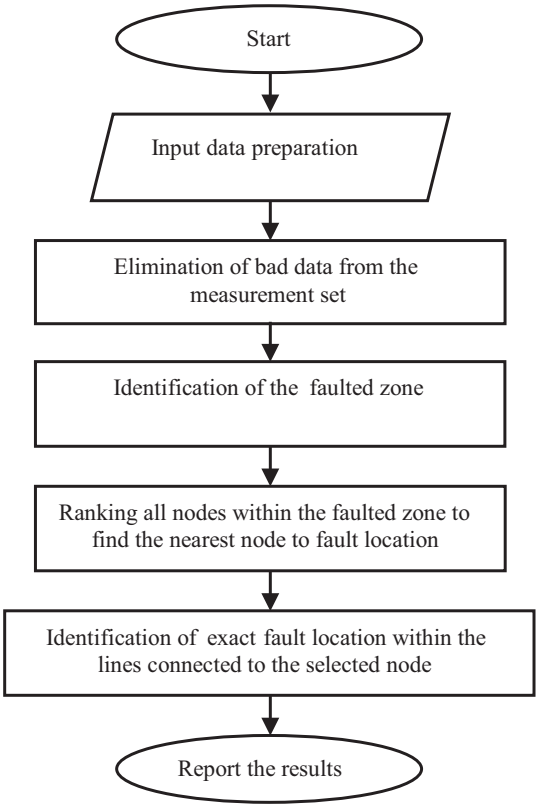

Fig. 1. Outline of the proposed method.

nodes is passed on to the next block where the faulted line is identified among the lines connected to the selected node and the fault is exactly located. Detailed description of each step is described in the following.

Input data preparation

The required input data for the algorithm are as follows:

(1) Fault type: A fault classification algorithm, such as that described in [25], is required to identify fault type and faulted phase or phases. However, since the development of such algorithms is beyond the scope of this paper, it is assumed that this information is available.
(2) System information: These data are extracted from the distribution system database and include network topological information, such as line sections length and impedance, and location of loads.
(3) Estimated load demand: Forecasted load demand is required to provide pseudo measurements to ensure the observability of the distribution system.
(4) Real measurements: In addition to the voltage and current measurement at the head of main feeder, synchronized voltages at sparse measurement nodes are used to perform the proposed method. There are two main methods for synchronization of the measurements: global positioning system (GPS) and computer network.
(5) Other available measurements such as non-synchronized voltages, line currents or powers can also be employed to enhance the method performance. The details of integration of different measurements into state estimation are described in [22,23].
(6) Field data: All accessible field data such as smart meter data or fault indicator signals, if available, could provide the required information for identification of the faulted zone [1].

Bad data detection and identification

Measuring devices are imperfect and may contain errors from various sources. The errors caused by measurement inaccuracies or noise are usually small but the errors caused by communication failures and improper connection or calibration are normally large and called bad data, which should be eliminated from the measurement set.

Several methods are reported in the literature for detection and elimination of bad data from the measurement set. In this paper, the normalized residual test is used which is composed of the following steps [24]:

(1) Solve the state estimation problem and obtain the measurement residuals:

$$r_i = z_i - h_i(\hat{x}), \qquad i = 1, \ldots, m \tag{6}$$

where $\hat{x}$ is the estimated value of the states;
(2) Calculate the normalized residuals:

$$r_i^N = \frac{|r_i|}{\sqrt{\Omega_{ii}}}, \qquad i = 1, \ldots, m \tag{7}$$

where Ω_{ii} is the i-th diagonal element of the residual covariance matrix:

$$\Omega = W^{-1} - HG^{-1}H^T \tag{8}$$

where W, H and G are previously defined in Section 'Branch-current-based state estimation';
(3) Find the largest element in the vector of normalized residuals ($r_i^{N^{max}}$).
(4) Compare $r_i^{N^{max}}$ against a threshold. If $r_i^{N^{max}} > c$, then the i-th measurement will be suspected as bad data. Else, stop, no bad data is detected. The threshold (c) is a constant previously determined based on the maximum value of the normalized residuals when no bad data is detected.
(5) Eliminate the i-th measurement if it is a bad data and go to step 1.

Identification of the faulted zone

Power companies normally depend on customer trouble calls for outage mapping and hence reducing the search space for fault location. However, in modern distribution networks, Advanced Measuring Infrastructure (AMI) for processing smart meter data has become an available option for determining the affected part of the network [1,26]. Furthermore, the use of IEDs with communication capabilities, that are able to detect and report a fault condition, could provide the required information for identification of

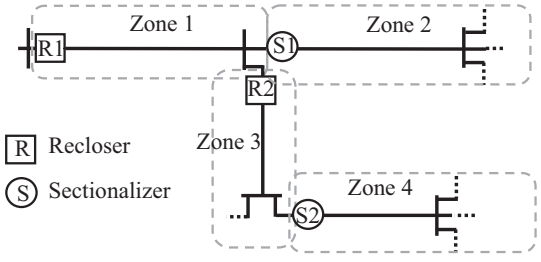

Fig. 2. Identification of the faulted zone in a sample distribution system.

the faulted zone [1,27]. As an example, consider the system shown in Fig. 2. If a fault occurs in each zone, all devices upstream of the faulted zone will indicate a fault condition; therefore, the affected zone would be the one with all the devices from the path from the substation reporting the fault. For example, if R1, R2 and S2 indicate a fault condition, zone 4 would be the affected part of the network.

Regardless of the method employed, identification of faulted zone narrows down the search space and significantly reduces the required time and computational burden.

Ranking the nodes

Considering the fault as a special load temporarily connected at the faulted node, the presence of this new load changes the actual load demand of the node. Thus, performing state estimation using the previously estimated load demand and during fault real time measurements, results in highly erroneous estimates. In other words, the occurrence of a fault leads to creation of large normalized residual values in state estimation results (see Eq. (7)).

The proposed fault location algorithm is based on the fact that the correction of the bad data (i.e. load demand of the faulted node) significantly reduces the normalized value of residuals. Therefore, the node which modification of its load demand results in the least values of normalized residuals would be the faulted node. This modification is performed by correction of the equivalent current of the estimated load demand of the faulted node:

$$I_{Li}^{ECM} = \left(\frac{(P_{Li} + jQ_{Li})}{V_i}\right)^* \qquad I_{Li} = I_{Li}^{ECM} + I_F \tag{9}$$

where I_{Li}^{ECM} and I_{Li} are the equivalent current of the estimated load demand and the modified value of the load current for node i, respectively, and I_F is the fault current.

In the proposed method, all suspect nodes are examined by applying the fault at each of them and performing state estimation. At each iteration of the state estimation, the fault current is calculated by Eq. (10), and the load current of the node under investigation is modified by Eq. (9).

$$I_F^{(k)} = I_F^{(k-1)} + \left(I_S^{mea} - I_S^{cal}\right) \tag{10}$$

where I_S^{mea} and I_S^{cal} are measured and calculated current at the feeder head and $I_F^{(k)}$ is the fault current at k-th iteration.

The node at which applying the fault (i.e. modifying its load current) produces the least value of normalized residuals would be the nearest node to the actual location of the fault. Therefore, for each node, the following index is calculated and the node with the largest value of the index is identified as one end of the faulted line:

$$R = \sum_{i=1}^{n}\sum_{j=1}^{3} r_L^N(i,j) \qquad Index = \frac{1}{R+\varepsilon} \tag{11}$$

where $r_L^N(i,j)$ is the value of normalized residual of the load pseudo measurement for j-th phase of i-th node, and ε is a small number to avoid division by zero.

The flowchart of the proposed ranking algorithm is illustrated in Fig. 3. Once the list of the suspect faulted nodes is prepared, the ranking algorithm starts from the first node. For each node, considering the fault as a load temporarily connected at the node and performing state estimation, the *Index* value is calculated. Applying the fault at the nearest node to the fault location would result in the least differences between the calculated and measured values (i.e. residuals). Thus, in the final step all nodes are ranked based on the *Index* which is a cumulative representation of normalized values of residuals.

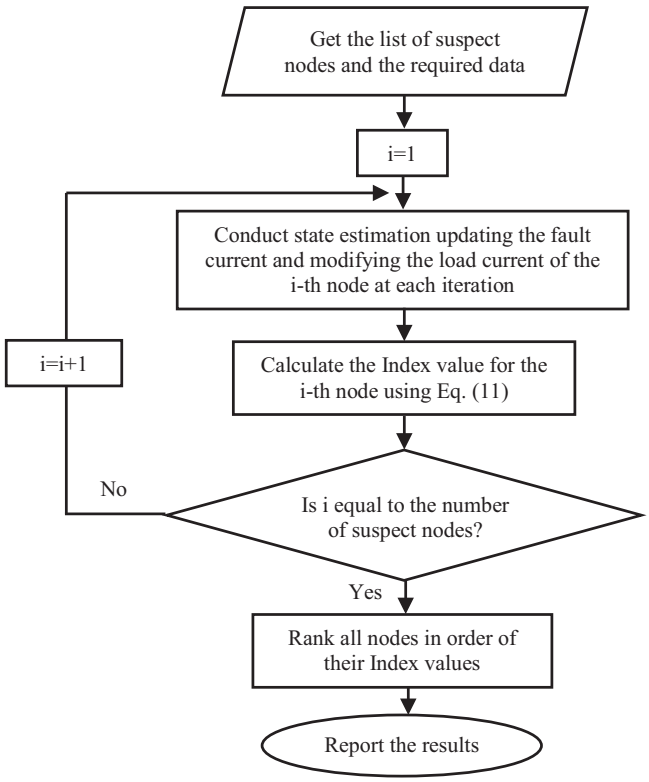

Fig. 3. Flowchart of the proposed node ranking algorithm.

Identification of the faulted line and fault location

During the process of ranking, a fault current value is estimated for each node. In [17], this fault current is used to define an angle index $\angle Zx$ for all analyzed nodes and the angle index is then employed to better rank the nodes.

$$\angle Zx^i = \angle V_F^{calc,i} - \angle I_F^{calc,i} \tag{12}$$

where $\angle V_F^{calc,i}$ and $\angle I_F^{calc,i}$ are the calculated angle of node voltage and fault current at node i.

In this work, the $\angle Zx$ value helps to identify the faulted line among the lines connected to the previously selected node and to locate the fault. Applying a fault at the beginning of the faulted line and moving it along the line, the estimated value of $\angle Zx$ will change as shown in Fig. 4. This change is to compensate the reactive part of line impedance and since line reactance has a linear relationship with its length, the trend will be linearly upward. This

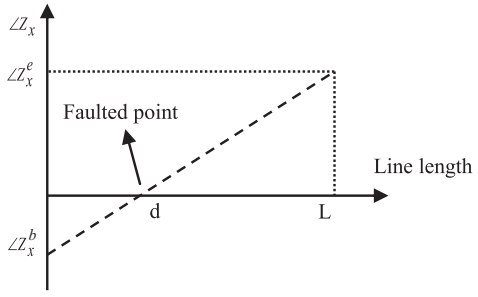

Fig. 4. Change of the $\angle Zx$ by moving the fault along the line.

means that the sign of $\angle Zx$ is always negative at the beginning of the faulted line, and positive at its end. Therefore, among the lines connected to the selected node, fault will be on the line with opposite sign of $\angle Zx$ at its ends. Furthermore, according to the trend shown in Fig. 4, the fault location can be simply calculated by solving the linear equation between the angle index and the distance:

$$d = \frac{L \times \angle Z_x^b}{\angle Z_x^b - \angle Z_x^e} \tag{13}$$

where d is the distance to the fault from the beginning of the line, L is the length of the line, and $\angle Z_x^b$ and $\angle Z_x^e$ are the estimated angle index values at the beginning and at the end of the line.

During the process of fault location, once the nodes are ranked based on the procedure explained, the algorithm will be applied to all top ranked nodes, one by one, to identify the exact fault location around them. Starting from the first top ranked node, the steps of the algorithm are as follows:

(1) Use Eq. (12) to calculate an angle index $\angle Zx$ for all analyzed nodes.
(2) Find all lines connected to the i-th top ranked node.
(3) If any of the selected lines has the opposite sign of $\angle Zx$ at its ends, use Eq. (13) to find the exact fault location in that line.
(4) Examine the next top ranked node.

The algorithm continues by considering a defined number of nodes and finds all possible fault locations around them.

Case study

The overhead, three phase, 13.8 kV, 134-node distribution system from [18] is used to test the proposed fault location algorithm. Five voltage measurements are arbitrarily placed in the system, at nodes 20, 51, 87, 118 and 127. The test feeder is simulated in the Alternative Transient Program (ATP) with loads modeled as constant impedances. The phasor values of voltage and current at the substation and the phasor value of voltage at each meter position are obtained by applying a full cycle Fourier algorithm on the corresponding waveforms [28].

Two sets of simulation tests are carried out. The first test set is performed under ideal conditions, whereas in the second test set the ability of the proposed method is evaluated in the presence of errors in measurements and load forecasts. In each test set, several cases are investigated considering different fault scenarios (see Fig. 5).

Performance under ideal condition

In this part, it is assumed that the field recorded data are free of errors and the loads are accurately estimated. In order to assess the accuracy of the proposed method, faults are applied at different lines considering various fault resistances. In each case, first, nodes are ranked and then the fault location algorithm is performed for the first two top ranked nodes.

Fig. 6 shows the Index values for a single line to ground (AG) fault at line 107–112, 85 m from node 107 with $R_f = 10\ \Omega$. Node 112 has the largest Index value and would be one end of the faulted line. The other possible ends are nodes 107 and 113 that are the nodes on the other side of lines initiating from node 112. Fig. 7 shows the angle index values ($\angle Zx$). As explained in Sectio n 'Identification of the faulted line and fault location', the line between nodes 112 and 113 cannot be a possible location of the fault because they have the same sign of $\angle Zx$, but line 107–112 can be a candidate. Hence, using Eq. (13) the fault is located at 80.4 m from node 107 at line 107–112. Node 107 is the second node with the highest value of the Index. Considering this node as a possible end of the faulted line, nodes 106, 108 and 112 are the other possible ends. Investigation of the $\angle Zx$ values in Fig. 7, reveals that the line between nodes 106 and 107 cannot be a possible location of the fault. Line 107–112 is already investigated and thus the fault is located at line 107–108, 23.55 m from node 107 as the second candidate.

Several other cases are also investigated considering different fault types, positions and resistances. Table 1 shows the deviation between the estimated and actual fault points (D) in meters for single line to ground faults (AG). The results are obtained for different fault scenarios and in each case, the selected lines are listed in order. In the upper part of Table 1, the estimations are made assuming that the affected zone is previously identified, whilst in the lower part the results are obtained by applying the algorithm to all nodes of the entire network. For example, for an AG fault occurred at line 90–119 with $R_f = 5\ \Omega$, either the faulted zone is previously identified or not, the fault location software informs the maintenance crew that the fault is located at line 90–199 and estimates its location by an error of 8.64 m. But if the affected zone is not identified, the software also selects line 90–91 as the second candidate. Comparison of results obtained for different fault resistance values reveals that the fault location error increases with the increase of fault resistance. Nevertheless, in all cases the faulted line is correctly identified and the estimations are always accurate.

Reviewing the results presented in the upper and lower parts of Table 1 for different fault scenarios, it can be concluded that the proposed algorithm is able to produce the same results with or without identification of the faulted zone. However, by determining the affected zone, instead of applying the algorithm to each node throughout the network, it should only be applied to the nodes within the identified zone, hence significantly reducing the computational burden.

Table 2 shows the difference between the estimated and actual fault locations in meters for the same fault scenarios but different fault types. Comparison of the results shows that the proposed fault location algorithm reproduces almost the same results for different fault types. Table 3 shows the results of the same tests for the fault location method proposed in [14]. This method finds the nearest node to the fault location; therefore, the errors reported in Table 3 for this algorithm are the distance between the selected node and actual fault location and the reported lines are the lines connected to the selected node. Compared to Table 2, almost in all cases the proposed algorithm produces more accurate results. For example, for a fault at line 74–75, the method proposed in [14] selects node 74 that is 55 m away from the fault location, whereas the proposed method gives more accurate results. As another example, for a fault on line 38–40, the method proposed in [14] selects node 40 as the nearest node with an estimation error of 109 m. For the same fault, the proposed algorithm selects node 38 and introduces line 38–40 and line 38–39 as the candidate faulted lines. The main advantage of the method proposed in [14] is its low computational time. For the test cases presented in Tables 2 and 3, the mean computational times for the proposed method and the algorithm proposed in [14] was 7 s and 22 s, respectively. Compared to required time for the consequence activities such as repair crew dispatching and line patrolling, both computational times are negligible.

Performance under non-ideal condition

In this part, the accuracy and robustness of the proposed algorithm are evaluated under load estimation uncertainties, measurement inaccuracies and presence of bad data. Different fault scenarios are simulated and since the single line to ground faults occur most frequently, results are only presented for this type of fault.

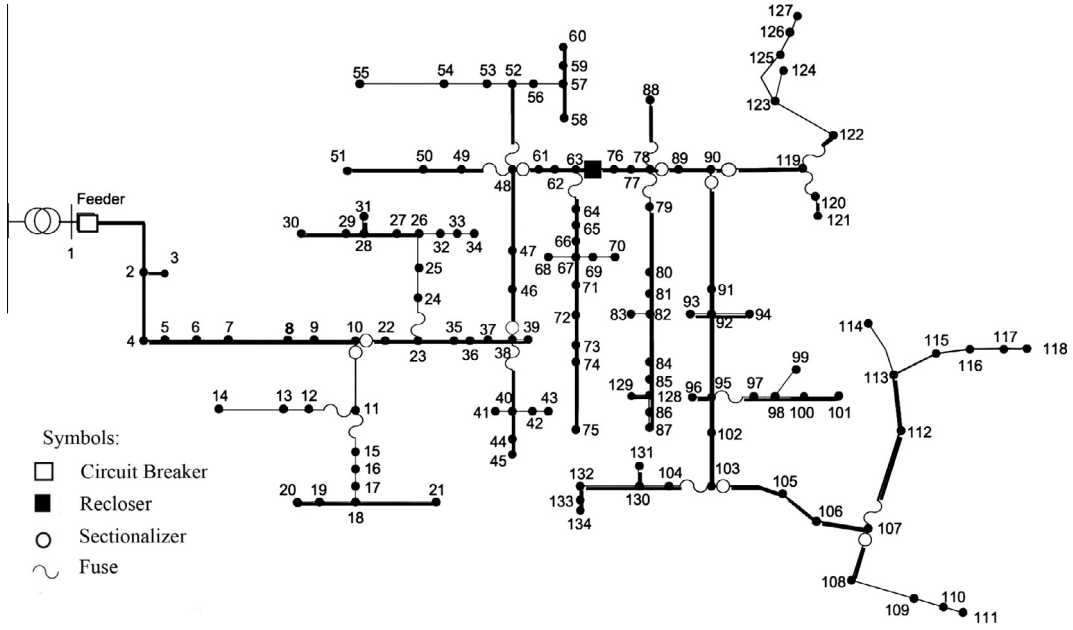

Fig. 5. Topology of the 134-node distribution network.

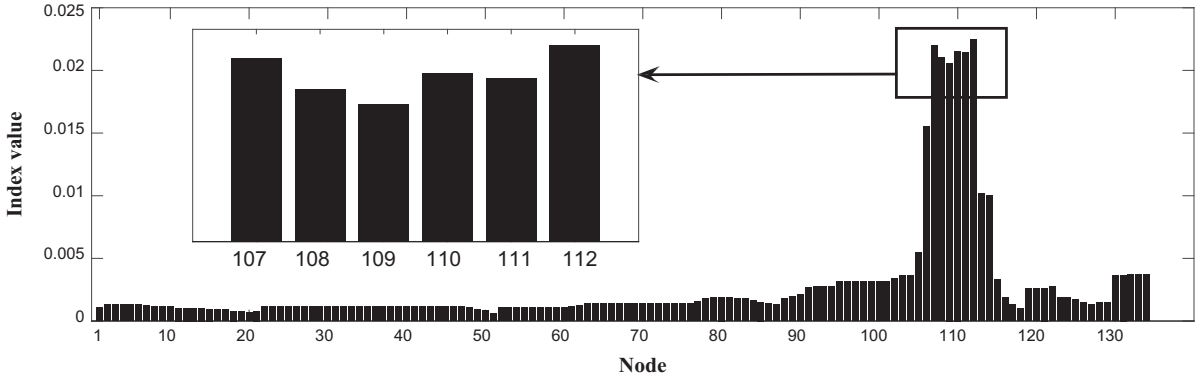

Fig. 6. Index values obtained for different nodes for a single line to ground (AG) fault at line 107–112.

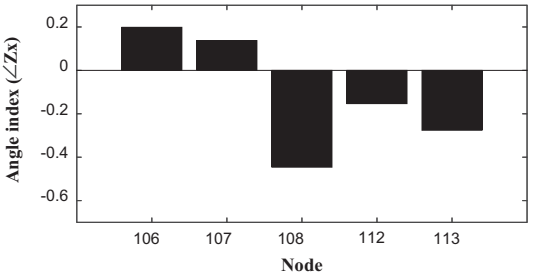

Fig. 7. Angle index values ($\angle Zx$) for nodes under investigation.

Effect of voltage measurement and load estimation inaccuracies

In distribution systems, loading data are often obtained by processing the historical costumer load data and due to costumer behavior uncertainties, the load profiles cannot be accurately estimated. In order to test the performance of the proposed method under uncertain loading conditions, the load data employed in simulations is taken as the benchmark and the forecasted load data is created by the random variation of them within specified ranges:

$$\begin{aligned} P_{i,j} &= P_{i,j}^{act}\left(1+\varepsilon_{i,j}^{p}\sigma_L\right) \\ Q_{i,j} &= Q_{i,j}^{act}\left(1+\varepsilon_{i,j}^{q}\sigma_L\right) \end{aligned} \tag{14}$$

where $P_{i,j}$, $Q_{i,j}$, $P_{i,j}^{act}$, $Q_{i,j}^{act}$ and $\varepsilon_{i,j}$ are the forecasted and actual load data and a random number between −1 and 1 for i-th node and j-th phase, respectively, and σ_L is the range of deviation considered.

Voltage measurements are also prone to errors due to bad calibration, unreliable communications, noises and meter inaccuracies. In order to study the effect of meter inaccuracies and noises, random errors are added to the data generated during simulation:

$$V_{Mi} = V_{Mi}^{act}(1+\varepsilon_i\sigma_V) \tag{15}$$

where V_{Mi}, V_{Mi}^{act} and ε_i are the measured voltage, the actual value of voltage and a random number between −1 and 1 for the i-th voltage meter, respectively, and σ_V is the range of deviation considered.

Besides the mentioned sources of measurement inaccuracies, signal phase angle deviations due to synchronization error would also cause measurement errors. There are two main methods for synchronization of the measurements: global positioning system (GPS) and computer network. Whilst the IEEE Standard for Synchrophasor Measurements specifies a theoretical accuracy of

Table 1
Distance between the estimated and actual fault points (D) in meters for single line to ground faults (AG).

	R_f	Fault at line 7–8 160 m from 7		Fault at line 38–40 39 m from 38		Fault at line 74–75 55 m from 74		Fault at line 90–119 44 m from 90		Fault at line 107–112 85 m from 107	
		Selected lines in order	D (m)	Selected lines in order	D (m)	Selected lines in order	D (m)	Selected lines in order	D (m)	Selected lines in order	D (m)
With identification of the faulted zone	1 Ω	7–8	18.99	38–39	98.83	74–75	0.34	90–119	10.34	107–112	9.07
				38–40	0.27					107–108	114.59
	5 Ω	7–8	12.9	38–39	99.3	74–75	0.07	90–119	8.64	107–112	7.24
				38–40	1.21					107–108	113.59
	10 Ω	7–8	2.6	38–40	4.85	74–75	0.25	90–119	1.08	107–112	4.59
				38–39	94.95					107–108	108.55
	20 Ω	7–8	43.3	38–40	20.07	74–75	20.28	90–119	30.89	107–112	59.22
				38–39	80.1					107–108	92.2
Without identification of the faulted zone	1 Ω	7–8	18.99	38–39	98.83	74–75	0.34	90–119	10.34	107–112	9.07
				38–40	0.27			90–91	71.06	107–108	114.59
	5Ω	7–8	12.9	38–39	99.3	74–75	0.07	90–119	8.64	107–112	7.24
				38–40	1.21			90–91	77.48	107–108	113.59
	10Ω	7–8	2.6	38–40	4.85	74–75	0.25	90–119	1.08	107–112	4.59
				38–39	94.95			90–91	71.88	107–108	108.55
	20Ω	7–8	43.3	38–40	20.07	74–75	20.28	90–119	30.89	107–112	59.22
				38–39	80.1			90–91	51.53	107–108	92.2

Table 2
Distance between the estimated and actual fault points (D) in meters for different fault types.

Fault type	Fault at line 7–8 160 m from 7		Fault at line 38–40 39 m from 38		Fault at line 74–75 55 m from 74		Fault at line 90–119 44 m from 90		Fault at line 107–112 85 m from 107	
	Selected lines in order	D (m)	Selected lines in order	D (m)	Selected lines in order	D (m)	Selected lines in order	D (m)	Selected lines in order	D (m)
Single line to ground fault (AG) (R_{fa} = 2 Ω)	7–8	18.45	38–39	99.76	74–75	0.112	90–119	10.48	107–112	10.03
			38–40	0.27					107–108	114.8
Line to line fault (AC) (R_{fa}, R_{fc} = 2 Ω)	7–8	17.18	38–39	101.12	74–75	0.43	90–119	11.74	107–108	111.21
			38–40	0.91					107–112	7.06
Line to line to ground fault (BCG) (R_{fb}, R_{fc} = 2 Ω, R_f = 5 Ω)	7–8	15.43	38–39	100.81	74–75	4.08	90–119	12.61	107–112	68.83
			38–40	0.65					107–108	90.39
Three phase to ground fault (ABCG) (R_{fa}, R_{fb}, R_{fc} = 2 Ω, R_f = 5 Ω)	7–8	13.69	38–39	70.09	74–75	0.08	90–119	9.086	107–112	4.54
			38–40	0.21					107–108	115.71

Table 3
Distance between the estimated and actual fault points (D) in meters for different fault types obtained by the method proposed in [14].

Fault type	Fault at line 7–8 160 m from 7		Fault at line 38–40 39 m from 38		Fault at line 74–75 55 m from 74		Fault at line 90–119 44 m from 90		Fault at line 107–112 85 m from 107	
	Selected lines in order	D (m)	Selected lines in order	D (m)	Selected lines in order	D (m)	Selected lines in order	D (m)	Selected lines in order	D (m)
Single line to ground fault (AG) (R_{fa} = 2 Ω)	8–9	40	39–38	109	74–75	55	90–119	44	108–107	85
	8–7								108–109	
Line to line fault (AC) (R_{fa}, R_{fc} = 2 Ω)	8–9	40	39–38	109	74–75	55	90–119	44	108–107	85
	8–7								108–109	
Line to line to ground fault (BCG) (R_{fb}, R_{fc} = 2 Ω, R_f = 5 Ω)	8–9	40	39–38	109	74–75	55	90–119	44	108–107	85
	8–7								108–109	
Three phase to ground fault (ABCG) (R_{fa}, R_{fb}, R_{fc} = 2 Ω, R_f = 5 Ω)	8–9	40	39–38	109	74–75	55	90–119	44	108–107	85
	8–7								108–109	

synchronization better than 1 μs (0.02 degree in 60 Hz) [29], the real world tests on phasor measurement units have shown the phase angle errors are in the range of 0.1 degree [30]. On the other hand, in [31] it is shown that network synchronization can achieve synchronization accuracy of better than 4 μs (0.08 degree in 60 Hz).

As an example to illustrate the accuracy of the proposed state estimation based method in fault analysis, Fig. 8 shows the differences between the calculated and actual values of fault voltage magnitudes under uncertain loading conditions. The deviations in voltage magnitudes calculated by the proposed state estimation based algorithm are remarkably low and about one-tenth of the during fault load flow method proposed in [14]. It should be noted that such small differences would have considerable impacts on the performance of the fault location algorithms which are based on the deviation between the measured and calculated values. Fig. 8 shows that the proposed state estimation based approach, by using all available measuring sources for during fault calculations, reduces the effect of the inaccuracies and ensures its effectiveness even under uncertain conditions.

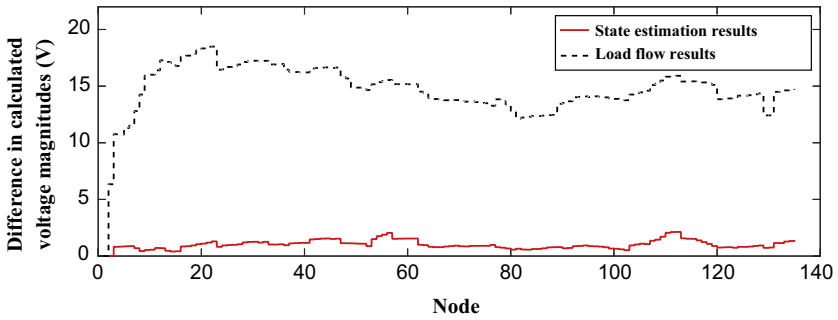

Fig. 8. Difference between the calculated and actual voltage magnitudes for a single line to ground fault at node 80 with $R_f = 10\ \Omega$.

Seven cases are considered to assess the accuracy and reliability of the algorithm under non-ideal conditions:

Case (1) Random variation of all loads within 20% of deviation ($\sigma_L = 0.2$).
Case (2) Random variation of all loads within 50% of deviation ($\sigma_L = 0.5$).
Case (3) Random variation of all measured voltages within 0.5% of deviation ($\sigma_V = 0.005$).
Case (4) Random variation of all measured voltages within 1% of deviation ($\sigma_V = 0.01$).
Case (5) Random variation of phase angle of all measured voltages within 0.1 degrees.
Case (6) Random variation of all loads and measured voltage magnitudes and phase angles within 50%, 1% and 0.1 degree respectively.

In all the considered scenarios, it is assumed that the faulted zone is previously identified. In each case, first, the algorithm described in Section 'Ranking the nodes' is applied to rank all the nodes within the affected zone. Then the fault location algorithm is performed for the first three top ranked nodes.

The results are summarized in Table 4, where almost in all cases, the affected line is correctly identified. By comparing Tables 4 and 1, it can be concluded that load estimation uncertainties do not have a significant impact on the method performance. Variation of all loads within 20% and 50% of deviation has led to almost the same results. The comparison also indicates that random errors within 0.5% of measured voltages do not affect the fault location results, but higher rate of errors may slightly decrease the method performance. The results for Case 5 shows that the proposed method can tolerate phase angle errors within 0.1 degrees but compared to ideal condition (i.e. Table 1) its accuracy decreases.

Case 6 is the worst situation considered where loads, voltage magnitude and phase angle measurements are randomly changed. Compared to the results obtained for Case 2, Case 4 and Case 5, the estimation errors are increased but all the results are still satisfactory.

Effect of bad data and the quantity of measurements

In the previous part, the effect of meter inaccuracies and noise in voltage measurements is studied. But, sometimes problems such as bad calibration or improper connection of measurements can cause higher errors in some meters. In such a case, bad data detection and identification should be performed to filter the erroneous data.

In the proposed method, as shown in Fig. 1, before performing the process of fault location, the prefault data and measurements are first employed to identify and eliminate the bad data from the measurement set.

In order to evaluate the ability of the algorithm in identifying the bad data and hence working with a reduced number of measurements, voltages measured at nodes 87 and 118 are multiplied by factors of 0.9 and 1.1, respectively. Using the prefault measurements, state estimation is performed and the normalized residuals for measured voltages are calculated. As shown in Fig. 9, the calculated values clearly indicate the presence of bad data. Hence, as described in Section 'Bad data detection and identification', the voltage at node 118 is eliminated from the measurement set. State estimation is subsequently repeated using the remaining measurements. As can be seen in Fig. 9, after elimination of the

Table 4
Distance between the estimated and actual fault points (D) in meters for single line to ground faults (AG) with $R_f = 10\ \Omega$ (the performance of the proposed method under non ideal conditions).

Cases under study	Fault at line 7–8 160 m from 7		Fault at line 38–40 39 m from 38		Fault at line 74–75 55 m from 74		Fault at line 90–119 44 m from 90		Fault at line 107–112 85 m from 107	
	Selected lines in order	D (m)	Selected lines in order	D (m)	Selected lines in order	D (m)	Selected lines in order	D (m)	Selected lines in order	D (m)
Case 1	7–8	9.28	38–39	91.19	74–75	6.46	90–119	5.54	107–112	6.15
			38–40	9.42					107–108	109.54
Case 2	7–8	18.54	38–39	91.24	74–75	6.8	90–119	6.01	107–112	9.3
			38–40	9.32					107–108	108.31
Case 3	7–8	13.07	38–39	93.75	74–75	3.06	90–119	5.6	107–112	5.86
			38–40	6.61					107–108	112.52
Case 4	7–8	21.85	38–39	88.11	74–75	7.69	90–119	5.77	107–112	7.85
			38–40	12.19					107–108	113.85
Case 5	7–8	20.823	38–39	88.75	74–75	8.8	90–119	12.71	107–112	10.16
			38–40	15.19					107–108	115.73
Case 6	9–10	48.17	38–39	89.5	74–75	11.04	90–119	16.69	107–108	121.3
			38–40	16.01					107–112	11.32

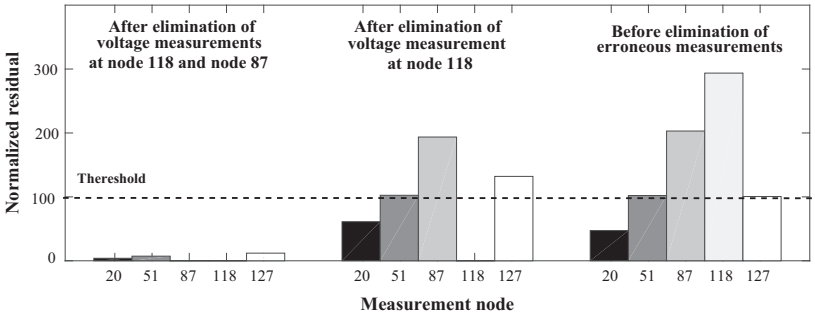

Fig. 9. Normalized residuals of measured voltages in presence of bad data at nodes 87 and 118.

Table 5
Distance between the estimated and actual fault points (D) in meters using a reduced number of measurements.

R_f	Fault at line 7–8 160 m from 7		Fault at line 38–40 39 m from 38		Fault at line 74–75 55 m from 74		Fault at line 90–119 44 m from 90		Fault at line 107–112 85 m from 107	
	Selected lines in order	D (m)	Selected lines in order	D (m)	Selected lines in order	D (m)	Selected lines in order	D (m)	Selected lines in order	D (m)
1 Ω	7–8	19.05	38–39	70.09	74–75	0.41	90–119	10.41	107–108	166
			38–40	0.2					107–112	1.2
5 Ω	7–8	13	38–39	99.33	74–75	1.85	90–119	7.79	107–108	165.53
			38–40	1.46					107–112	0.64
10 Ω	7–8	2	38–39	95.67	74–75	2.68	90–119	3.94	107–108	161.85
			38–40	5.07					107–112	6.83
20 Ω	7–8	50.12	38–39	20.26	74–75	21.94	90–119	44	107–108	148.75
			38–40	80.51					107–112	18.48

detected bad data, the calculated values of normalized residuals are still over the detection threshold. Therefore, voltage measurement at node 87 is also identified as a bad data and eliminated from the measurement set. State estimation is repeated using the modified measurement set. As shown in Fig. 9, after elimination of the bad measurements, the largest value of normalized residuals is below the threshold, ensuring the quality of the remaining measurements.

In order to evaluate the ability of the proposed algorithm using the remaining 3 voltage measurements, different fault scenarios are considered and Table 5 summarizes the results. In all cases the affected line is correctly identified and the results are still satisfactory. Therefore, the proposed method is robust enough to work with a reduced number of measurements, but its accuracy and error tolerance may be decreased. On the other hand, increasing the number of meters improves both the accuracy and robustness of the method.

In this paper, it is assumed that the employed measurement devices are already installed for other purposes like power quality monitoring. The proposed method aims to use the data from the available measurements to enhance the fault location accuracy. The minimum number of the measurement devices required for fault location depends on their location. However, if a distribution company decides to dedicate a number of measurement devices for the fault location purpose, a method similar to the optimal meter placement algorithm proposed in [17] can be employed.

Conclusion

This paper presents a new algorithm that uses sparse measurements for fault location in distribution networks. The proposed method is tested by simulation studies for different fault types, positions, and resistances. The simulation results validate the accuracy of the proposed method. The major contributions of the paper are summarized as follows:

(1) A new fault location algorithm is proposed that incorporates all of the available measurements such as voltage, current or power into calculation to provide the best estimation.
(2) Simulation studies have shown that compared to the previously proposed methods, the presented algorithm is more accurate.
(3) The integrity of the data collected from IEDs in distribution networks is not guaranteed. Measuring devices may contain small errors due to measurement inaccuracies or noises, and large errors caused by communication failures and improper connection or calibration. Compared to the previously proposed methods that mostly require accuracy in measurements and load profile estimation, the proposed method is able to handle measurement errors. The method first assesses the input data and eliminate the measurements having large errors. It then uses a state estimation based method incorporating all remaining measurements which significantly reduces the impact of their random small errors and load estimation uncertainties.
(4) The proposed method requires only a few number of meters such as digital fault recorders or power quality meters to enhance its accuracy and robustness.

References

[1] Chen P-C, Dokic T, Kezunovic M. The use of big data for outage management in distribution systems. In: International conference on electricity distribution (CIRED) workshop, Rome; 2014.
[2] Sridharan K, Schulz NN. Outage management through AMR systems using an intelligent data filter. IEEE Trans Power Del 2001;16:669–75.
[3] Teng J-H, Huang W-H, Luan S-W. Automatic and fast faulted line-section location method for distribution systems based on fault Indicators. IEEE Trans Power Syst 2014;29:1653–62.
[4] Mora-Flòrez J, Meléndez J, Carrillo-Caicedo G. Comparison of impedance based fault location methods for power distribution systems. Electr Power Syst Res 2008;78:657–66.

[5] Jamali S, Talavat V. Accurate fault location method in distribution networks containing distributed generations. Iranian J Elect Comput Eng 2011;10:27–33.
[6] Lee Seung-Jae, Choi Myeon-Song, Kang Sang-Hee, Jin Bo-Gun, Lee Duck-Su, Ahn Bok-Shin, et al. An intelligent and efficient fault location and diagnosis scheme for radial distribution systems. IEEE Trans Power Del 2004;19:524–32.
[7] Nouri H, Alamuti MM. Comprehensive distribution network fault location using the distributed parameter model. IEEE Trans Power Del 2011;26:2154–62.
[8] Borghetti A, Corsi S, Nucci CA, Paolone M, Peretto L, Tinarelli R. On the use of continuous-wavelet transform for fault location in distribution power systems. Int J Electr Power Energy Syst 2006;28:608–17.
[9] Borghetti A, Bosetti M, Nucci CA, Paolone M, Abur A. Integrated use of time-frequency wavelet decompositions for fault location in distribution networks: theory and experimental validation. IEEE Trans Power Del 2010;25:3139–46.
[10] Hizman H, Crossley PA, Gale PF, Bryson G. Fault section identification and location on a distribution feeder using travelling waves. In: IEEE power engineering society summer meeting, vol. 1103; 2002. p. 1107–112.
[11] Bellis EJ, Gale PF. Apparatus for locating faults in electric cables. In: U.S. Patent No. 4,491,782; 1985.
[12] Thukaram D, Khincha HP, Vijaynarasimha HP. Artificial neural network and support vector machine approach for locating faults in radial distribution systems. IEEE Trans Power Del 2005;20:710–21.
[13] Al-shaher MA, Sabry MM, Saleh AS. Fault location in multi-ring distribution network using artificial neural network. Electr Power Syst Res 2003;64:87–92.
[14] Pereira RAF, da Silva LGW, Kezunovic M, Mantovani JRS. Improved fault location on distribution feeders based on matching during-fault voltage sags. IEEE Trans Power Del 2009;24:852–62.
[15] Lotfifard S, Kezunovic M, Mousavi MJ. Voltage sag data utilization for distribution fault location. IEEE Trans Power Del 2011;26:1239–46.
[16] Brahma SM. Fault location in power distribution system with penetration of distributed generation. IEEE Trans Power Del 2011;26:1545–53.
[17] Dong Y, Zheng C, Kezunovic M. Enhancing accuracy while reducing computation complexity for voltage-sag-based distribution fault location. IEEE Trans Power Del 2013;28:1202–12.
[18] Trindade FCL, Freitas W, Vieira JCM. Fault location in distribution systems based on smart feeder meters. IEEE Trans Power Del 2014;29:251–60.
[19] Almeida MCD, Costa FF, Xavier-de-Souza S, Santana F. Optimal placement of faulted circuit indicators in power distribution systems. Electr Power Syst Res 2011;81:699–706.
[20] Han Fengling, Yu Xinghuo, Al-Dabbagh M, Wang Yi. Locating phase-to-ground short-circuit faults on radial distribution lines. IEEE Trans Ind Electron 2007;54:1581–90.
[21] Lotfifard S, Kezunovic M, Mousavi MJ. A systematic approach for ranking distribution systems fault location algorithms and eliminating false estimates. IEEE Trans Power Del 2013;28:285–93.
[22] Whei-Min L, Jen-Hao T, Shi-Jaw C. A highly efficient algorithm in treating current measurements for the branch-current-based distribution state estimation. IEEE Trans Power Del 2001;16:433–9.
[23] Teng JH. Using voltage measurements to improve the results of branch-current-based state estimators for distribution systems. IEE Proc Generat Transmiss Distrib 2002;149:667–72.
[24] Abur A, Exposito AG. Power system state estimation: theory and implementation. Mercel Dekker; 2004.
[25] Zhang J, He ZY, Lin S, Zhang YB, Qian QQ. An ANFIS-based fault classification approach in power distribution system. Int J Electr Power Energy Syst 2013;49:243–52.
[26] Tram H. Technical and operation considerations in using Smart Metering for outage management. In: IEEE/PES trans distrib conf expo Chicago; 2008. p. 1–3.
[27] Yanfeng G, Guzman A. Integrated fault location system for power distribution feeders. IEEE Trans Ind Appl 2013;49:1071–8.
[28] Yong G, Kezunovic M, Deshu C. Simplified algorithms for removal of the effect of exponentially decaying DC-offset on the Fourier algorithm. IEEE Trans Power Del 2003;18:711–7.
[29] IEEE Standard C37.118.1.2011. IEEE standard for synchrophasor measurements for power system; 2011.
[30] Depablos J, Centeno V, Phadke AG, Ingram M. Comparative testing of synchronized phasor measurement units. In: IEEE power engineering society general meeting. IEEE; 2004. p. 948–54.
[31] Johannessen S. Time synchronization in a local area network. IEEE Control Syst Mag 2004;24:61–9.