Fault Location Algorithm for Radial Distribution Systems Capable of Handling Insufficient and Inaccurate Field Data

Yimai Dong
Department of Electrical and Computer Engineering
Texas A&M University
College Station, USA
dongyimai@neo.tamu.edu

Mladen Kezunovic
Department of Electrical and Computer Engineering
Texas A&M University
College Station, USA
kezunov@ece.tamu.edu

***Abstract*— A fault location algorithm for radial distribution systems is proposed in this paper. The probability model of data error is developed and the standard deviation of the error is selected as one of the inputs to the fault location algorithm. Data processing technology is introduced to evaluate data condition and detect bad data before fault location calculation is obtained. The contribution of inaccurate field-recorded data is recognized in the stage of faulted node selection. Performance of the proposed algorithm is tested and compared with a similar algorithm, and the result shows that the proposed algorithm is more reliable when field-recorded data is insufficient or inaccurate.**

I. INTRODUCTION

Fault location in distribution system is gaining attention in recent years for the reason of system reliability. On one hand, customer requirement on the quality of power supply is higher and on the other, electric utilities wish to reduce the revenue loss caused by outage to the minimum. Efforts to reduce outage time in distribution system may be achieved through introducing condition-based maintenance in asset management, and developing new algorithms for fault location. Having an accurate fault location program helps improving system reliability because the average time for field crew for finding the faults in the system is reduced so isolating the faults can be done more efficiently.. By reducing duration of unexpected outages (outage caused by faults) and confining the search for faults within a relatively small area, the accuracy of fault location algorithms may have a huge impact in system reliability.

Distribution fault location algorithms developed in recent years can be classified into 4 different groups: Intelligent processing of trouble calls [1], [2]; fault location based on fault distance calculation [3]-[6]; Fault location using transient signal features [7]-[10]; model-based fault location [11], [12]. In general, trouble call-based methods have the least capability of pinpointing faults accurately, followed by the distance-based methods. The accuracy of distance-based methods relies highly on the estimation of fault resistance, and the results become unreliable with the existence of laterals between measurement points. Fault location methods using transient signal features require high-frequency sampling and time synchronization, and their accuracy relies highly on the algorithm used for feature extraction.

The implementation of the algorithms mentioned above in distribution systems is hampered by the availability of field data. In distribution systems the field collected data has several issues: a) Data comes mostly as phasors or magnitudes; b). Very few sensors are installed in the system exclusively for collecting voltage and current measurements, except at the root of feeders; c). A limited data is available from IEDs (power quality meter, etc.); d). The quality of data is not guaranteed as the distribution level is not considered as critical as the transmission level.

Constrained by the availability and quality of data, algorithms that perform perfectly under ideal condition may not be even acceptable under the practical condition of insufficient or inaccurate field data. The model-based algorithms based on comparison between simulated and filed captured data are capable of handling sparse measurements, butt such algorithms rely on pretty good information about the distribution models being readily available, which is not a very practical requirement.. Assuming the model data is readily available, the problem raised by the scarcity of data may be solved, but the problem of the results being affect by the quality of data still exists.

This paper proposes a model-based distribution fault location method that is capable of dealing with data with poor quality. A discussion of source of data error and stochastic method to reduce the impact of data error are introduced in section II. Description of the algorithm is given in section III. Section IV presents the evaluation results of the proposed

This work was supported by Power Systems Engineering Research Center (PSerc) under the project T-36 titled “Integration of asset and outage management tasks for distribution application”.

algorithm when compared with the algorithm introduced in [12]. Conclusions are given in section V.

II. Error in Field-Recorded Data

A. Data condition

As mentioned in the introduction, the availability of filed data at the distribution level is not as good as at transmission level. The imperfection of field data has two aspects:

1) Insufficiency: sensors placed in distribution systems for protection and monitoring purpose are very few because of the lack of instrument transformers and communication facilities along feeders . In addition, data from available sensors are mostly phasors or magnitudes that are not time-synchronized.

2) Inaccuracy: data recorded in the field is prone to erros due to unreliable communications and potential calibration problems with the sensors.

Based on the above,, a fault location algorithm implemented in distribution systems must be able to deal with the poor data condition. A model-based algorithm may be selected to deal with the insufficiency of data, but data processing technology is needed for dealing with the inaccuracy of data. Data error needs to be analysis carefully before any method is proposed to reduce the impact of error.

B. Modeling of data error

The data required for the fault location algorithm proposed in this paper are phasors from feeder root and scalars from some nodes in the distribution system. Data acquired may be "contaminated" in two ways: from the sensor and during transmission. A/D conversion, phasor calculation and electro-magnetic interference (EMI) are all possible sources of error. The model of acquired data may be represented as:

$$\widehat{\mathrm{X}} = \mathrm{X} + \mathrm{e}(\mathrm{X}) = \mathrm{X} + [(\mathrm{G} - 1) \cdot \mathrm{X} + \mathrm{D}(\mathrm{X}) + \mathrm{x}\,] \tag{1}$$

where

$\widehat{\mathrm{X}}$ is the contaminated data;

G is the gain ratio;

$\mathrm{e}(\mathrm{X})$ is the total error inserted;

X is the true value of the electric quantity;

$\mathrm{D}(\mathrm{X})$ is the offset associated with X;

x is the random error (white noise).

The error consists of three parts: gain factor G, offset D and random error x. The first part is proportional to the true value of data, which comes from differences in the calibration of measured value, caused by the. ratio of instrument transformer, voltage reference in A/D conversion, etc. Offset is a constant value introduced mostly by the difference in the ground voltage and random error x may come from various sources such as instrument transformer saturation or EMI. Although it is hard to predict the random error, it is reasonable to assume that it has a normal distribution:

$$\frac{\mathrm{x}}{\mathrm{X_N}} \sim \mathrm{N}(0, \sigma^2) \tag{2}$$

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot exp(-\frac{x^2}{2\sigma^2}) \tag{3}$$

where

X_N is the rated value of X;

$p(x)$ is the density function of x;

σ^2 is the variance of x.

C. Reduction of the error impact

The approaches for reducing the impact of data error are:

1) Cancel out the gain and offsset parts of data error by doing simpleprocessing operations such as subtraction or division;

2) Rely more on accurate data and less on inaccurate data;

3) Detect and eliminate bad data when data error exceeds the threshold.

Methodology for implementing such approaches is described in the following section.

III. Fault Location Algorithm

A. Data requirement

The proposed fault location algorithm is based on the one proposed in [12]. However, some improvements are made for handling insufficient and inaccurate data. Following is a detailed description of data requirements.

1) Field-recorded data: pre-fault and during-fault phasors of voltage and currents recorded from roots of feeders; pre-fault and during-fault magnitudes of voltages recorded at some nodes in the system.

2) System information: topology information from feeder database, such as length and per-mile impedance of line sections, location and parameters of transformers and other devices.

3) Load information: location of loads and connected KVA.

4) Measurement information: location of voltage measurements and the standard deviation σ of measurement error.

B. Algorithm flow chart

The flow chart of proposed fault location algorithm is shown in Fig.1. The algorithm consists of four steps: Pre-fault load flow calculation, estimation of applicability, fault simulation and faulted node selection. The four steps will be described separately in the following sections.

Processing of data takes place in the step of estimation of applicability where the data condition is estimated using $\mathrm{J}(\widehat{\mathrm{X}})$detection test [13]. If the number of recorded data points and accuracy cannot satisfy the requirements for implementing the algorithm, bad data is removed from input values. The procesure is repeated until the data is good enough for the algorithm to be executed or no more data can be removed and the program is terminated.

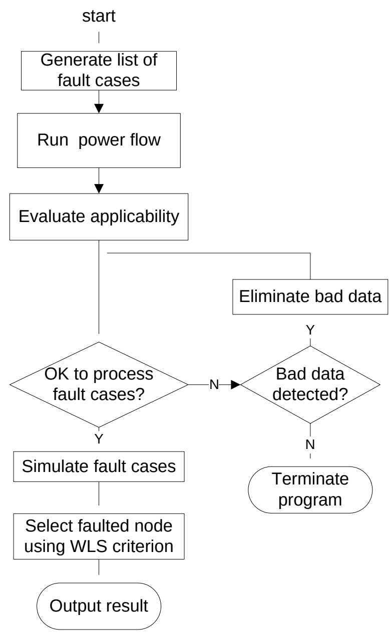

Figure 1. Flow chart

C. *Power flow solution*

The load flow algorithm for radial system described in [14] is used to calculate pre-fault voltage magnitudes. Fixed-impedance model is used for load modeling. In the initial stage, all node voltages are assigned with voltage recorded at the root of feeders. Back-sweeping to update branch currents using (4) and (5) and forward-sweeping to update node voltages using (6) is done in every iteration. The stopping criterion for iterations is defined by eg. (7).

$$I_{j\_n}^{(k)} = {Z_{L\_n}}^{-1} \cdot V_n^{(k-1)} \tag{4}$$

$$I_{b\_i}^{k} = \sum I_{b\_p}^{(k)} + I_{j\_n}^{(k)} \tag{5}$$

$$V_n^{(k)} = V_m^{(k)} - Z_{b\_i} \cdot I_{b\_i}^{(k)} \tag{6}$$

$$\max\left\{\left|V_n^{(k)} - V_n^{(k-1)}\right|\right\} < \varepsilon,\ \text{n=1,…,N} \tag{7}$$

where

k is the number of iteration;

$I_{j\_n}^{(k)}$ is the injection current at node n;

Z_{L_n} is the three phase load impedance matrix at node n;

$V_n^{(k)}$ is the node voltage of the down-stream node of branch i;

$I_{b\_i}^{k}$ is the branch current of branch i, which flows from node m to node n;

$I_{b\_p}^{(k)}$ is the branch current of branch p, which flows out from node n;

Z_{b_i} is the three phase line impedance matrix for branch i;

ε is the threshold for change in node voltage.

N is the total number of nodes.

D. *Estimation of applicability*

The $J(\hat{X})$detection test from [13] is applied to estimate the condition of data, i.e. if the number and accuracy of voltage measurements are good enough for a reliable output.

Calculated value of voltage magnitude at node i from pre-fault load flow calculation is designated as $\left|V_{i,pre}^{cal}\right|$, while field-recorded value is designated as $\left|V_{i,pre}^{meas}\right|$. Weighted difference J_i is defined as (V_N *is the rated voltage* is the rated voltage):

$$J_i = \left(\frac{\left|V_{i,pre}^{cal}\right| - \left|V_{i,pre}^{meas}\right|}{V_N \cdot \sigma_i}\right)^2 \tag{8}$$

The J index is the summation of J_i:

$$J = \sum J_i \tag{9}$$

Reliability index of field-recorded data is defined as:

$$RI = \frac{\sqrt{J-m}}{\sqrt{2m}} \tag{10}$$

where m is the number of redundant measurements. For the proposed algorithm, load flow calculation relies only on voltage and current phasors at feeder roots, m is the total number of voltage measurements.

The value of J_is and RI reveal the condition of data. Large value of individual J_i indicates that data from measurement i is very likely to be bad data and should be eliminated; Large value of RI indicates that either the number of measurements are not enough for a reliable output, or bad data exists, or sever error exists in system model, such as wrong topology or load information.

J_i and RI are used as double criteria. If $J_i \leq 25$ stands for all J_i s, and if $RI < 3$, the data condition is considered as acceptable, and the program will proceed to fault simulation. If for one or two $J_i > 25$, data from the corresponding measurements will be eliminated and RI will be recalculated. If the criteria can not be met by eliminating bad data, the program is be considered not applicable under the current data condition.

E. *Fault-case simulation*

A list of fault cases is generated according to the affected area. All nodes within the affected area are considered as a suspect faulted node. Fault-case simulation is executed for each cases, and the calculated value of node voltage magnitudes at nodes with voltage measurements are recorded.

The algorithm for fault case simulation is similar to pre-fault load flow algorithm. Fault is considered as a special load connected to the faulted node, as is shown is Fig.2. The total injection current is the summation of fault current and load current.

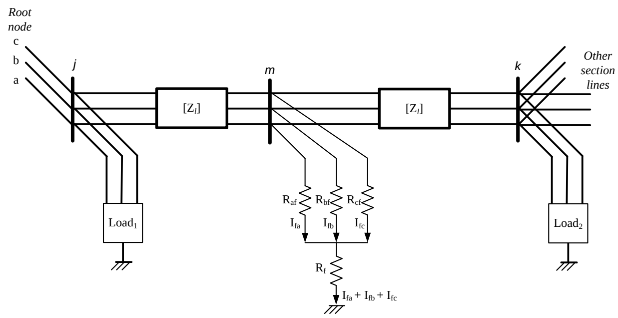

Figure 2. Current injection from a fault at node m [12]

The equivalent impedance of the fault is not of interest. The fault current is calculated at the end of each iteration and added as current injection caused by fault at the faulted node using (11) and (12):

$$I_f^{(k)} = I_f^{(k-1)} + (I_{rn}^{df,meas} - I_{rn}^{df,cal}) \quad (11)$$

$$I_{j\_n}^{df,(k)} = I_{j\_nl}^{df,(k)} + I_f^{(k)} \quad (12)$$

where

$I_f^{(k)}$ is fault current;

$I_{rn}^{df,meas}$ is the current measured at feeder root;

$I_{rn}^{df,cal}$ is the calculated current at feeder root;

$I_{j\_n}^{df,(k)}$ is the injection current at faulted node n;

$I_{j\_nl}^{df,(k)}$ is the injection current from load connected to n.

The flow chart for fault case simulation is shown in Fig.3.

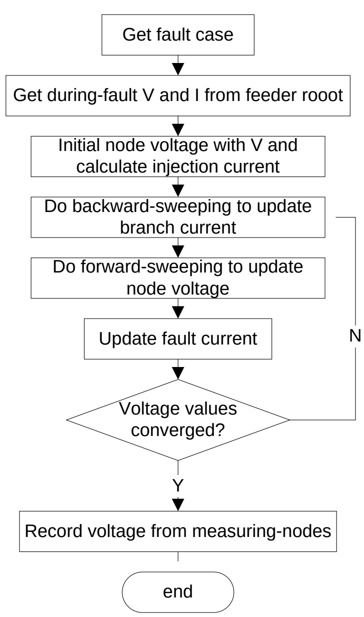

Figure 3. Flow chart

F. *Faulted node selection*

The likely fault location is selected taking into account all analyzed nodes during the fault location process. Weighted-deviation is used for locating the fault.

For each analyzed node, the during-fault magnitude deviation between measured and calculated voltage sags is computed:

$$\delta_k^j = \left\| \Delta V_k^{j,cal} - \Delta V_k^{meas} \right\|, \mathrm{k} = 1, \ldots \mathrm{m}, \mathrm{j} = 1, \ldots \mathrm{np} \quad (13)$$

where

$\Delta V_k^{j,cal}$ is the difference in three-phase pre-fault and during-fault voltage magnitudes (voltage sags) calculated at node k considering node j as the faulted node;

ΔV_k^{meas} is the three-phase voltage sags measured at node k;

m is the total number of voltage measurements;

np is the total number of fault cases simulated.

The weighted-deviation is calculated as

$$\gamma_j = \sum_{k=1}^{m} (\delta_k^j / \sigma_k)^2 \quad (14)$$

The faulted node is the one with the smallest value of γ_j.

$$n_f = j \mid \gamma_j = \min\{\gamma_s\}, s = 1, \ldots np \quad (15)$$

G. *Description of error-impact reduction*

The algorithm is capable of minimizing the impact of offset error and random error.

The offset error is removed by the calculation of voltage sags—offset from pre-fault and during fault data cancels out in subtraction.

As the selection of faulted node relies on the weighted-deviations, the contribution of data from less accurate measurements is reduced in proportion with the variance of the random error, which means that data more likely to have high random error has a lower impact on the result.

The proportional error is not considered in the proposed algorithm.

IV. SIMULATION RESULTS

A. *Description of test system*

A 13.8 kV, 134-node, overhead three-phase primary distribution feeder is used as the test system. Fig.4 shows the topology of the feeder.

Root voltage and current are recorded at node 1.

Four voltage measurements are placed in the system, at node 30, 48, 103 and 118 respectively. They are marked as measurement 1- 4 respectively.

Both the algorithm reported in [12] and the algorithm proposed in this paper are implemented and the results are compared.

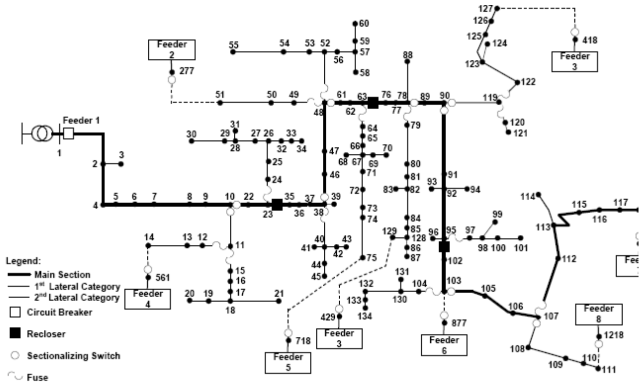

Figure 4. Test system

B. Case study

Case 1: perfect condition

In this case, the field-recorded data are not contaminated by errors. Fault scenarios are listed in Table 1.

TABLE I. FAULT SCENARIOS FOR CASE 1

Faulted node	***Fault type***	***Fault resistance (Ω)***
17, 36, 42, 107	A-G	1
63, 90	A-G	10
5, 77	A-B-C	5
86	A-B	1

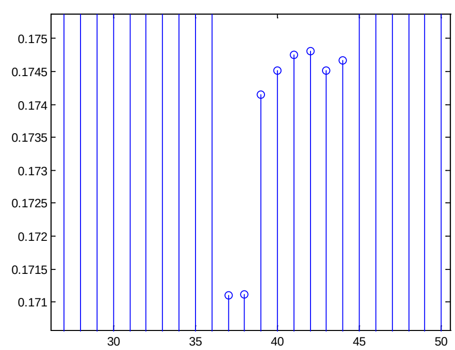

Figure 5. Nodes with smallest γ_j

Both the algorithm reported in [12] and the proposed algorithm give correct result for all scenarios. Fig.5 shows the smallest γ_j calculated for fault occurring at node 36.

Case 2: Bad data

A-G fault at node 36 is simulated, but pre-fault and during-fault voltage magnitude recorded by measurement 2 are added with errors of 20% and 15%. Variances of random error σ_i for all voltage measurements are 0.01.

The faulted node selected by the algorithm proposed in [12] is node 48, which is incorrect.

J_i and RI calculated by the proposed algorithm are listed in Table 2. J_2 is very large, indicating that data from measurement 2 are bad data and should be eliminated. RI after bad data elimination is less than 3, and the program continues with data from measurement 1, 3 and 4.

TABLE II. J_i AND RI

J_i	***RI***
0.1213 398.7 0.1374 0.1151	Before data removal: 7.027 After data removal: 0.6638

The faulted node selected by the proposed algorithm after bad data elimination is node 36. The proposed algorithm selected the right node again.

Case 3: Data with errors

TABLE III. LIST OF DATA CONDITION AND RESULT

Measurements	***σ_i***	***Correct times***
1, 2, 3, 4	0.01, 0.01, 0.01, 0.01	10
1, 2, 3, 4	0.01, 0.1, 0.01, 0.01	10
1, 2, 3	0.01, 0.01, 0.01	10
1, 3, 4	0.01, 0.1, 0.01	8

A-G fault at node 36 is simulated. Errors with density function of (3) is generated according to σ_i of the measurement and added to the true value of the measured pre-fault and during-fault quantities. Four data conditions are designed and ten sets of data are generated and fed to fault location program for each data condition. Data condition and the times that the fault location program provides correct output are listed in Table 3.

The outputs for the two cases where the selected node is wrong are 39 and 33.

It appears that with four voltage measurements, the program can tolerate much more inaccurate data. When one measurement is removed, and especially when the one close to faulted node is removed, the performance is not as good as before.

V. CONCLUSION

A model-based fault location algorithm for radial distribution systems is proposed in this paper. The proposed algorithm use voltage magnitudes from the sparse voltage measurements installed in distribution systems. The following are major contributions of this paper:

- Stochastic analysis is introduced in the algorithm to reduce the impact of data error on the output.
- The algorithm is capable of assessing the measurement applicability, detecting bad data and adjusting the

contribution of field-recorded data from different measurements according to the accuracy of measurements.

- Simulation results show that the performance of the proposed algorithm is capable of handling is insufficient and/ or inaccurate data, which is common problem in distribution systems.

ACKNOWLEDGMENT

The authors of this paper gratefully acknowledge the contributions of Rodrigo A.F. Pereira to development of the fault location technique mentioned in this paper.

REFERENCES

[1] Liu Yan, N.N. Schulz, "Knowledge-based system for distribution system outage locating using comprehensive information", IEEE Trans. on Power Systems, vol. 17, issue 2, pp. 451 – 456, May 2002.

[2] Z.L. Gaing, C.N. Lu, Y.T. Lin, "Object-oriented design of trouble call analysis system on personal computer", IEE PRroceedings on Generation, Transmission and Distribution, vol. 142, issue 2, pp. 167 – 172, March 1995.

[3] Zhu Jun, D.L. Lubkeman, A.A. Girigis, "Automated fault location and diagnosis on electric power distribution feeders", IEEE Trans. on Power Delivery, vol. 12, issue 2, pp. 801 – 809, April 1997.

[4] Myeon-Song Choi, Seung-Jae Lee, Duck-Su Lee, Bo-Gun Jin, "A new fault location algorithm using direct circuit analysis for distribution systems", IEEE Trans. on Power Delivery, vol. 19, issue 1, pp. 35 – 41, Jan. 2004 .

[5] Seung-Jae Lee, Myeon-Song Choi, Sang-Hee Kang, Bo-Gun Jin, Duck-Su Lee, Bok-Shin Ahn, Nam-Seon Yoon, Ho-Yong Kim, Sang-Bong Wee, "An intelligent and efficient fault location and diagnosis scheme for radial distribution systems", IEEE Trans. on Power Delivery, vol. 19, issue 2, pp. 524 – 532, April 2004.

[6] R.H. Salim, M. Rensener, A.D. Filomena, K. Rezende Caino de Oliveira, A.S. Bretas, "Extended fault-location formulation for power distribution systems", IEEE Trans. on Power Delivery, vol. 24, issue.2, pp. 508-516, April 2009.

[7] E.C. Senger, G., Jr. Manassero, C. Goldemberg, E.L.Pellini, " Automated fault location system for primary distribution networks", IEEE Trans. on Power Delivery, vol. 20, issue 2, part 2, pp. 1332 – 1340, April 2005.

[8] F.H. Magnago, A. Abur, "A new fault location technique for radial distribution systems based on high frequency signals", IEEE Power Engineering Society Summer Meeting 1999, vol. 1, pp. 426 – 431, July 1999.

[9] A.Borghetti, M.Bosetti, M.Di Silvestro, C.A.Nucci, M.Paolone, "Continuous-Wavelet Transform for Fault Location in Distribution Power Networks: Definition of Mother Wavelets Inferred From Fault Originated Transients", IEEE Trans. on Power Systems, vol. 23, issue 2, pp. 380 – 388, May 2008.

[10] El-Hami, M., L. L. Lai, D. J. Daruvala, and A. T. Johns, "A new traveling-wave based scheme for fault detection on overhead power distribution feeders", *IEEE Transactions on Power Delivery*, vol. 7, nº 4, pp. 1825 – 1833, 1992.

[11] Z.Galijasevic, A.Abur, "Fault Location Using Voltage Measurements", IEEE Trans. on Power Delivery, vol. 17, issue 2, pp. 441-445, April 2002 .

[12] R.A.F. Pereira. L.G.W. da Siva, M. Kezunovic, J.R.S. Mantovani, "Improved fault location on distribution nfeeders based on matching during-fault voltage sags". IEEE Trans. on Power Delivery, Vol. 24, Issue 2, pp. 852-862, April 2009,

[13] A. Monticelli, "state estimation in electric power systems: a generalized approach". Kluwer academic publishers, Boston, Dordrecht, London, 1999.