# A Systematic Approach for Ranking Distribution Systems Fault Location Algorithms and Eliminating False Estimates

Saeed Lotfifard, *Member, IEEE*, Mladen Kezunovic, *Fellow, IEEE*, and Mirrasoul J. Mousavi, *Member, IEEE*

***Abstract*—The need for distribution reliability enhancement in the age of smart grids requires reliable methods for locating faults on distribution systems leading to a faster service restoration and maintenance cost optimization. Given the numerous fault location methods, one faces the challenge of objectively evaluating and selecting the most proper method. In this paper, a two-step approach is proposed and discussed for ranking available fault location methods that takes into account application requirements and modeling limitations and uncertainties. The ranking method formulated as uncertainty analysis utilizes $2n + 1$ point estimation to calculate the statistical moments of the fault location estimation error. These moments plugged into the Chebyshev's inequality provide a basis for ranking the fault location method. The selected method may still suffer from multiple fault location estimations. To address this caveat, voltage sag characteristics reported by few intelligent electronic devices (IEDs) along the feeder are utilized. The number and location of these IEDs are determined through an optimal approach specifically formulated for this problem. The proposed two-step ranking methodology and the IED placement optimization approach were implemented on a simulated distribution system and their effectiveness was demonstrated through a few select scenarios and case studies.**

***Index Terms*—Distribution systems, fault location, intelligent electronic devices (IEDs), uncertainty analysis.**

## I. Introduction

FAULT LOCATION in distribution systems is an important function for outage management and service restoration, which directly impacts feeder reliability and quality of the electricity supply. Improving fault location methods supports the Department of Energy (DOE) "Grid 2030" [1] initiatives for grid modernization. A number of methods have been proposed by various researchers specifically for fault location in distribution systems. In [2]–[4], the apparent impedance, defined as the ratio of selected voltage to selected current based on the fault type and faulted phases, is utilized for locating faults by comparing apparent impedance with the preknown line data. A common caveat for impedance-based methods is multiple location estimations due to the reliance solely on the substation voltage and current signals[5]. In [4], data collected from fault indicators along the network, which determine the direction of the fault current, are utilized to solve the problem of multiple fault location estimation.

In [6] and [7], a method based on direct circuit analysis is proposed where the fault location equation was derived by applying the matrix inverse lemma. In [8] and [9], a method using the superimposed components of the voltages and currents is proposed. In this method, the assumed fault point is varied systematically until the actual fault point is found. The fault is located based on the fact that the amount of the superimposed current in the healthy lines should be minimal.

Methods based on intelligent systems, such as neural networks and fuzzy logics, are proposed in [10]–[12]. In [10], the faulted area is detected by training an adaptive neurofuzzy inference system (ANFIS) net with extracted features based on the knowledge of protective device settings. In [12], the problem of multiple fault location estimation is obviated using the learning algorithm for multivariable data analysis (LAMDA) classification technique. In [13]–[15], methods based on traveling waves generated by the fault have been proposed. The time difference between successively captured traveling waves is used for fault location.

The availability of various fault location methods provides a wide range of options for integration into distribution-management systems. However, adopting a suitable method is an outstanding challenge that requires an objective evaluation and ranking based on a set of predefined criteria. The hardware requirements of the host-processing platform and lack of extensive instrumentation in distribution systems, as well as the requirements and underlying assumptions of the algorithm itself, such as the heterogeneity of the lines, presence of laterals, and load taps are among the limitations. Taking a systematic approach to select the best fault location method would ensure that the predefined user criteria related to data availability, robustness against uncertainties in the measurements and/or network parameters modeling, and feasibility of implementation are satisfied. This is the focus of this paper.

Moreover, the availability of feeder data from IEDs in modern distribution systems provides new opportunities to enhance fault location methods. Examples of such measuring devices include, but are not limited to, digital protective relays,

Manuscript received February 08, 2012; revised June 03, 2012; accepted July 16, 2012. Date of publication November 19, 2012; date of current version December 19, 2012. This work was supported in part by ABB Inc. under the contract entitled "Optimized fault location in distribution systems." Paper no. TPWRD-00134-2012.

S. Lotfifard is with the Department of Electrical Engineering and Computer Science, University of Central Florida, Orlando, FL 32816 USA (e-mail: s.lotfifard@ucf.edu).

M. Kezunovic is with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843-3128 USA (e-mail: kezunov@ece.tamu.edu).

M. J. Mousavi is with ABB Corporate Research, Raleigh, NC 27606 USA (e-mail: mirrasoul.j.mousavi@us.abb.com).

Digital Object Identifier 10.1109/TPWRD.2012.2213616

capacitor bank and recloser controllers, residential meters, power-quality (PQ) meters and various types of low-cost current and voltage sensors that may be selectively installed outside the substation fence. A systematic approach is needed to take advantage of the data for fault location purposes.

In this paper, an optimized approach for selecting a proper fault location method given a set of user requirements is proposed. First, the potential candidate methods from the pool of available fault location methods are qualitatively ranked. Uncertainty analysis is used as a tool to quantify the robustness of the fault location methods against uncertainties in the measurements and/or network parameters and modeling. The method less affected by the uncertainties would be more desirable. To eliminate false fault location estimation by the selected fault location method, the voltage sag data gathered from few IEDs installed along the feeder are utilized. An optimal IED placement approach is proposed to determine the number and location of these IEDs.

This paper is organized as follows. In Section II, the proposed approach for optimized fault location selection is explained. In Section III, the method for eliminating false fault location estimation using IED data is detailed. Section IV describes an optimal IED placement that determines the number and location of the required IEDs to be installed. In Section V, the performance of the proposed fault location approach is demonstrated on a simulated distribution system. Finally, Section VI provides conclusions for this paper.

## II. Fault Location Method Selection

The proposed optimized fault location selection procedure is conducted in two steps presented as follows.

### A. Qualitative Analysis

For the qualitative analysis, the candidate fault location methods are classified into five groups based on data requirements as well as topological/modeling requirements. Table I summarizes these categories and their characteristics.

- apparent impedance measurement;
- direct three-phase circuit analysis;
- superimposed components;
- traveling waves;
- artificial intelligence.

The traveling wave-based methods require high-frequency sampling in the megahertz range depending on the tower configuration (propagation velocity) and the desired accuracy and, hence, costly for distribution applications. The presence of laterals and load taps that reflect the traveling waves creates multiple reflection difficulties and confusion when applied to distribution systems. The artificial-intelligent methods require training data and retraining subsequent to a change in the circuit topology. The superimposed components and direct three-phase circuit analysis require an accurate data-acquisition process[6], [8].

Simplicity and practical feasibility of the impedance-based methods are important advantages of this category. Some of the impedance-based fault location methods do not consider heterogeneity of distribution lines and presence of laterals which ultimately undermine their accuracy. In [3] and [4], all peculiarities of distribution systems have been taken into account including heterogeneity of the lines, and the presence of laterals. Two methods have been short listed and further evaluated in the next step.

### B. Quantitative Analysis

To select the most suitable fault location method quantitatively, the uncertainty analysis is conducted to rank the select fault location methods. Uncertainty analysis focuses on quantifying uncertainty in model output due to uncertainties in the inputs [16]. The method less affected by the uncertainties in the inputs would be more desirable.

In fault location applications, the location estimation error is a function of the algorithm inputs and underlying modeling assumptions. The uncertainty problem could be formulated as follows:

$$e = f(x_1, x_2, \ldots x_n) \tag{1}$$

where

| | |
|---|---|
| $e$ | fault location estimation error; |
| $f(.)$ | function that relates the inputs ($x_i$'s) to the output (estimated fault location); |
| $x_i$ | factor that affects. For instance: |
| $x_1$ | measured voltage signal at the substation; |
| $x_2$ | measured current signal at the substation; |
| $x_3$ | line parameter; |
| $x_4$ | estimated load at each node; |
| $x_5$ | fault resistance |

There are uncertainties related to each of these inputs. For example, line parameters do change as the resistivity of the soil and conductor temperature change; each measuring device has its own uncertainties and there are inaccuracies associated with the load estimation.

To estimate the error due to the said uncertainties, the exhaustive approach enumerating every possible combination of loads and circuit parameters is impractical and poses unmanageable computational burden. Alternatively, a probabilistic approach may be carried out in which a probabilistic distribution function is assigned to each load and/or parameter of the network. Instead of calculating a deterministic value for the error $(e)$, its statistical moments are calculated.

Having the mean $(\mu)$ and standard deviation $(\sigma)$ of $e$, we define the following function to estimate the uncertainty:

$$X_{\text{threshold}} = \mu + c\sigma \tag{2}$$

where $C$ is a real number and $X_{\text{threshold}}$ is an index that quantifies the uncertainty.

TABLE I
SUMMARY OF THE MAIN CHARACTERISTICS AND REQUIREMENTS OF SELECT FAULT LOCATION METHODS FOR DISTRIBUTION SYSTEMS

| Analyzed aspects | Fault Location Method | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Impedance Measurement | | | Direct Three-phase Circuit Analysis | Superimposed Components | Traveling Wave | | | Artificial Intelligence | | |
| | [2] | [3] | [4] | [6] | [8] | [13] | [14] | [15] | [10] | [11] | [12] |
| Data Requirements | | | | | | | | | | | |
| • Substation data | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| • Feeder data (outside the fence) | - | - | √ | - | - | - | - | √ | - | - | - |
| • Voltage signals (abc) | √ | √ | √ | √ | √ | √ | √ | √ | - | - | √ |
| ○ Pre-fault data | √ | √ | √ | - | √ | - | - | - | - | - | √ |
| ○ During-fault data | √ | √ | √ | √ | √ | √ | √ | √ | - | - | √ |
| ○ Post-fault data | - | - | - | - | - | - | - | - | - | - | √ |
| • Current signals (abc) | √ | √ | √ | √ | √ | - | √ | - | √ | √ | √ |
| ○ Pre-fault data | √ | √ | √ | - | √ | - | - | - | √ | √ | √ |
| ○ During-fault data | √ | √ | √ | √ | √ | - | √ | - | √ | √ | √ |
| ○ Post-fault data | - | - | - | - | - | - | - | - | √ | - | - |
| • Symmetrical Components | √ | - | √ | - | - | - | - | - | - | - | - |
| • Phasor components | - | √ | - | √ | √ | - | - | - | - | - | √ |
| • Time domain | - | - | - | - | - | √ | √ | √ | √ | √ | - |
| • High frequency sampling | - | - | - | - | - | √ | √ | √ | - | - | - |
| Topological and Modeling Requirements | | | | | | | | | | | |
| • Line model | Short | Short | Long | Short | Short | Long | Long | Long | Long | Short | Long |
| • Load model | Z = constant | Z = f(V) | Z = f(V) | Z = constant | Z = constant | Z = constant | Z = constant | Z constant | Z = f(V) | Z = constant | Z = f(V) |
| • Non-homogeneity | √ | √ | √ | √ | √ | - | - | - | √ | - | √ |
| • Laterals (3-phase or 1-phase branches) | √ | √ | √ | - | - | - | √ | - | √ | - | √ |
| • Load taps | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| • Communication | - | - | √ | - | - | - | - | √ | - | - | - |

According to the Chebyshev's inequality [17], for any random variable $X$ with mean $\mu$ and standard deviation $(\sigma)$, regardless of its distribution, the following holds:

$$p[\mu - C\sigma \leq X \leq \mu + C\sigma] \geq 1 - \frac{1}{C^2}. \tag{3}$$

For instance, if two standard deviations are used $(C = 2)$, the probability that $X$ is bounded by $\pm C\sigma$ is greater than or equal to 75%. For a 95% confidence interval criterion on $\sigma^2$ (which is an accepted and commonly used value for classification purpose), $C$ would be 4. Therefore, (2) becomes

$$X_{\text{threshold}} = \mu + 4\sigma. \tag{4}$$

The method with the smallest $X_{\text{threshold}}$ wins the draw as this means it is less affected by the input uncertainties. The statistical moments of $e$ ($\mu$ and $\sigma$) can be calculated using a computationally efficient $2n + 1$ point estimation method discussed in the next section.

### C. $2n + 1$ *Point Estimation Method*

According to the previous section, to rank the fault location methods, $X_{\text{threshold}}$ in (4) should be calculated, which means statistical moments of $e$ ($\mu$ and $\sigma$) should be estimated. These may be computed via the $2n + 1$ point estimation method that is explained in [18]–[22]. Here, a brief summary of the method is presented.

For $z = h(X) = h(x_1, x_2, \ldots x_n)$, where $X$ is a set of random variables $x_k$ $k = 1, 2 \ldots, n$, $\mu_k$, $\sigma_k$ $\mu_k$ denote the mean, and standard deviation of $x_k$, respectively. $\rho_{ij}$ is defined as correlation coefficient between variable $x_i$ and $x_j$ $i \neq j$. It is assumed $\rho_{ij} = 0$, $i \neq j$. When $\rho_{ij} \neq 0$, the rotational transformation based on the eigenvector of the

covariance matrix can be utilized to transform the set of correlated random variables $X$ into an uncorrelated set of random variables $X'$[18]. The $p_{k,i}$ is concentrations/weights located at $(\mu_1, \mu_2, \ldots, x_{k,i}, \ldots, \mu_{n-1}, \mu_n)$ where

$$x_{k,i} = \mu_k + \xi_{k,i}\sigma_k \; i = 1, 2, \ldots, , m \; k = 1, 2, \ldots, n \tag{5}$$

where $m$ is the number of concentrations per input variable. The standard location $\xi_{k,i}$ and the weight $p_{k,i}$ are obtained by solving the following equations [16]:

$$\sum_{i=1}^{m} p_{k,i} = \frac{1}{n} \quad k = 1, 2, \ldots, n \tag{6}$$

$$\sum_{i=1}^{m} p_{k,i}(\xi_{k,i})^j = \lambda_{k,i} \quad i = 1, 2, \ldots, m \; k = 1, 2, \ldots, n \tag{7}$$

where $\lambda_{k,j}$ is the ratio of the $j$th moments about the mean of $x_k$ to $(\sigma_k)^j$ i.e.,

$$\lambda_{k,j} = \frac{M'_j(X)}{(\sigma_k)^j} \quad j = 1, 2, 3, \ldots \tag{8}$$

$$M'_j(X) = \int (x - \mu_k)^j f_k(x) dx \tag{9}$$

where $\lambda_{k,1}$ equals 0, $\lambda_{k,2}$ equals 1, and $\lambda_{k,3}$ and $\lambda_{k,4}$ are the coefficient of skewness and coefficient of kurtosis of $x_k$, respectively. The $j$th raw moment of the output random variables is estimated as follows:

$$E(z^j) \cong \sum_{k=1}^{n} \sum_{i=1}^{m} p_{k,i} \times (h(\mu_1, \mu_2, \ldots, x_{k,i}, \ldots \mu_{n-1}, \mu_n))^j. \tag{10}$$

The $j$th central moment of the output random variable can be estimated as soon as the $j$th raw moment is estimated. For instance, the mean and standard deviation of the random variable can be calculated as follows [18]:

$$\mu_z = E(z) \cong \sum_{k=1}^{n} \sum_{i=1}^{m} p_{k,i} \times (h(\mu_1, \mu_2, \ldots, x_{k,i}, \ldots, \mu_{n-1}, \mu_n)) \tag{11}$$

$$\sigma_z = \sqrt{E(z^2) - E(z)^2}. \tag{12}$$

For $m = 2$, the following holds:

$$\xi_{k,i} = \frac{\lambda_{k,3}}{2} + (-1)^{3-i}\sqrt{n + \left(\frac{\lambda_{k,3}}{2}\right)^2} \quad i = 1, 2 \quad k = 1, 2, \ldots, n. \tag{13}$$

And

$$p_{k,i} = \frac{1}{n}(-1)^i \frac{\xi_{k,3-i}}{\xi_k} \tag{14}$$

where

$$\xi_k = 2\sqrt{n + \left(\frac{\lambda_{k,3}}{2}\right)^2}. \tag{15}$$

In (13), the standard location $\xi_{k,i}$ depends on the number of input random variables. When $n$ becomes large, inaccuracies occur as indicated in [19]. In the fault location methods, uncertain parameters could be the measured voltage and current signals, line parameters, and the estimated load at each node. Therefore, for power systems with realistic size, the number of uncertain parameters becomes large. To overcome this problem, the $2n + 1$ scheme has been proposed in [20] and [21]. Only one additional evaluation of the function compared to the 2n scheme is needed. This scheme is derived by solving (6) and (7) for $m = 3$ with one of the three standard location $\xi_{k,i}$ set to zero.

Let $\xi_{k,3}$. Then, the standard locations and weights are

$$\xi_{k,i} = \frac{\lambda_{k,3}}{2} + (-1)^{3-i}\sqrt{\lambda_{k,4} - \frac{3}{4}{\lambda_{k,3}}^2} \quad i = 1, 2 \quad \xi_{k,3} = 0 \tag{16}$$

$$p_{k,i} = \frac{(-1)^{3-i}}{\xi_{k,i}(\xi_{k,1} - \xi_{k,2})} \quad i = 1, 2 \tag{17}$$

$$p_{k,3} = \frac{1}{n} - \frac{1}{\lambda_{k,4} - {\lambda_{k,3}}^2}. \tag{18}$$

It is noted that in (5), setting $\xi_{k,3} = 0$ yields $\xi_{k,3} = \mu_k$. So $n$ of $3n$ locations are the same point $(\mu_1, \mu_2, \ldots, \mu_k, \ldots, \mu_{n-1}, \mu_n)$. It is enough to run only one evaluation of the function at this location, provided that the corresponding weight is updated as follows:

$$p_0 = \sum_{k=1}^{m} p_{k,3} = 1 - \sum_{k=1}^{m} \frac{1}{\lambda_{k,4} - {\lambda_{k,3}}^2}. \tag{19}$$

From (16), it can be seen that the standard location values of the scheme $2n + 1$ do not depend on the number of input random variables $n$ as do the $m \times n$-type schemes. This is a common feature of all $m \times n + 1$ concentration schemes [21].

Now by using (11) and (12), statistical moments ($\mu$ and $\sigma$) of $e$ in (1) could be estimated. Having $\mu$ and $\sigma$, according to (4), $X_{\text{threshold}}$ could be calculated. At first, parameters that have impacts on $e$ in (1), such as measured voltage and current signal at the substation, line parameter, the estimated load at each node, and the fault resistance are identified. To model uncertainty of these parameters, probability distribution functions are assigned to each of them. (In Section V, this will be explained through a case study.) In the next step, $\mu$ and $\sigma$ are estimated according to (11) and (12). Finally $X_{\text{threshold}}$ is calculated according to (4).

Based on this discussion, the procedure to select the most proper fault location method for each application involves a qualitative step that reduces the number of methods for numerical ranking in the subsequent step. The ranking is accomplished according to (4). The method that has the least $X_{\text{threshold}}$ value is selected as the most appropriate fault location method given the uncertainties in the measurements and modeling errors. The selected method may still suffer from multiple fault location estimation. In the next section, a method to eliminate this caveat is proposed.

## III. False Fault Location Estimation Elimination

One of the main caveats of the apparent impedance-based fault location method is the multiple fault location estimation due to the presence of laterals and branched nature of the circuit.

Where additional feeder data outside the substation are available, it is possible to eliminate the false fault location estimations. The fact that a fault causes voltage sags with different characteristics at different nodes throughout the network is utilized for this purpose.

First, the potential locations are determined by the optimized apparent impedance-based fault location method selected in Section II. Voltage sags at IED locations are calculated using a load-flow program for the faults at each estimated location of the fault by the apparent impedance method. The fault location is determined by comparing how well each calculated case matches up to what was actually observed by the IEDs in the network. The case that shows the highest similarity is considered to be the location of the fault.

In [23] and [24], a method for voltage sag characterization has been discussed. The characteristic voltage ($\bar{V}$), which is a phasor that quantifies the severity of the voltage sag, is calculated. It is defined as

$$\bar{V} = \bar{V}_{\text{pos}} - \bar{V}'_{\text{neg}} \tag{20}$$

where

$$\bar{V}'_{\text{neg}} = \bar{V}_{\text{neg}} e^{-jT60^\circ}$$
$$T = \frac{1}{60^\circ} \arctan\left\{\frac{\bar{V}_{\text{neg}}}{1 - \bar{V}_{\text{pos}}}\right\} \tag{21}$$

where $T$ is rounded to the nearest integer number and $\bar{V}_{\text{pos}}$ and $\bar{V}_{\text{neg}}$ are the positive and negative sequence voltages, respectively. The absolute value and the argument of the characteristic voltage $\bar{V}$ are referred to as "magnitude" and "phase-angle jump" of the voltage dip, respectively. Therefore, $\bar{V}$ has enough information about amplitude and phase-angle jump of the voltage sag and is the only required data that should be transferred to the distribution control center.

The following equation quantifies the similarity between the calculated and the recorded voltage sag [25]:

$$\text{Error} = \varepsilon^2_{\text{amplitude}(V)} + \varepsilon^2_{\text{phase}(V)} + \varepsilon^2_{\text{phase}(I)} \tag{22}$$

where

$\varepsilon_{\text{amplitude}(V)}$ difference between the amplitude of the characteristic voltage of the recorded ($\bar{V}_{\text{recorded}}$) and calculated voltage sags ($\bar{V}_{\text{calculated}}$);

$\varepsilon_{\text{phase}(V)}$ difference between the phase angle of the characteristic voltage of the recorded ($\varepsilon_{\text{recorded}}$) and calculated voltage sags ($\bar{V}_{\text{calculated}}$);

$\varepsilon_{\text{phase}(I)}$ difference between the phase angle of the recorded and calculated current at root node.

The following criterion is defined for the faulted node detection

$$\text{Flag} = \frac{1}{\text{Error} + \Delta} \tag{23}$$

where Error is defined in (22) and $\Delta$ is a small number just to prevent the division by zero. The location that has the largest Flag value (or smallest Error) is the faulted location.

To estimate the fault resistance, a binary search is conducted. At first, a fault with a typical value $R_f = 10$ is applied at an assumed faulted location and the steady state fault current at the root node (substation) is calculated using short circuit analysis methods. This calculated fault current is compared with the recorded current at the substation

$$\varepsilon = |I_{\text{calculated}}| - |I_{\text{recorded}}| \tag{24}$$

where

$|I_{\text{calculated}}|$ amplitude of the calculated current phasor at the substation;

$|I_{\text{recorded}}|$ amplitude of the recorded current phasor at the substation.

If $\varepsilon$ does not fall within the convergence tolerance, $R_f$ must be increased/decreased until $\varepsilon$ falls within the convergence tolerance. If $\varepsilon$ is positive (negative), $R_f$ should be increased (decreased) according to the binary search rule (half-interval search). For example, if the actual value of $R_f$ is 8 $\Omega$, the values of $R_f$ at each iteration would be 10, 5, and 7.5 $\Omega$.

## IV. Optimal IED Placement

To eliminate false fault location estimations, the voltage sag data are gathered from few IEDs installed at select nodes along the feeder. A mobile telephone network or other available communication architecture is used to transmit the recorded data to the control center. The objective of the optimal IED placement problem is to find the minimum number of IEDs and the locations at which these IEDs should be installed to provide adequate information for fault location.

First, all pairs of nodes that are at the same electrical distances from the substation are determined. The nodes that are at both sides of each line (sections) should be checked when the fault location falls between two nodes. For each pair of $i$ and $j$, faults with different impedance are applied at $i$ and $j$, one at a time, and the voltage sag characteristic at the other nodes (candidate location for IED placement) is calculated. All nodes that capture different voltage sag characteristics due to the fault at node $i$ and $j$ are potential nodes for IED placement. All fault types should be checked. The fault type that causes worst case (i.e., generates voltage sags with least difference) is considered for further study. If the method holds for the worst case it holds for other types of fault as well.

The goal of the optimal IED placement is to provide the required information with the minimum number of IED installations to minimize the solution cost. Therefore, the IED placement is formulated as an optimization problem with the objective of minimizing the number of required IEDs so that for each pair of nodes $i$ and $j$ (that are at the same distance from the substation), at least one IED is installed at nodes that capture different voltage sag characteristics.

To formulate the IED placement problem as an optimization problem, first the measurement status vector $X$ is defined where

the $i$th element $x_i$ shows whether an IED is required to be installed at node $i$

$$x_i = \begin{cases} 1, & \text{if an IED is needed at node } i \\ 0, & \text{if an IED is not needed at node } i \end{cases}. \quad (25)$$

The optimization problem is formulated as follows:

$$\text{Min} \sum_i x_i. \quad (26)$$

Subject to

$$(x_i + \cdots + x_n) + (x_o x_p + \cdots x_q x_r) + (x_c x_d x_e + \cdots + x_f x_g x_h) + (x_i x_j \cdots x_n) \geq 1. \quad (27)$$

Constraint (27) should hold for each pairs of nodes that are at the same electrical distance from the substation. This constraint ensures that for each pair of nodes that are at the same distance from the substation ($j$th pair), at least one set of IEDs is installed to determine the actual faulted node.

If voltage sag characteristics captured by single IEDs for faults at nodes of the $j$th pair are not disparate (at least the difference should be more than the IED measurement errors), a set of IEDs should be installed. In (27), the term $X_o X_p$ means to discriminate the fault at the $j$th pair, two IEDs should be installed at nodes $o$ and $p$. The term $X_c X_d X_e$ means to detect the actual fault location the fault at $j$th pair, three IEDs should be installed at nodes $c$, $d$, and $e$. The last term in (27) means the IEDs should be installed at every node.

The optimization problem can be solved using CPLEX solver [26]. The constraint (27) is nonlinear. To reduce the computational effort of the solver, constraint (27) is linearized and the optimization problem is reformulated as follows:

$$Min \sum_i x_i \quad (28)$$

$$\sum_{k_j} Y_{jk_j} \geq 1 \quad \forall_j \quad (29)$$

$$\sum_{i \in k_i} x_i \geq |k_j| \times Y_{jk_j} \quad \forall j,\ k_j$$

$$x_j, Y_{jk_i} \in \{0, 1\} \quad (30)$$

where

$j$ $j$th pair of nodes that are at the same distance from the substation. For instance, if $\{(K, L),\ (M, N),\ (O, P)\}$ are pairs of nodes that are at the same distance from the substation, $j = 1$ refers to $(K, L)$ and $j = 2$ refers to $(M, N)$ and $j = 3$ refers to $(O, P)$;

$Y_{jk_j}$ is the $k$th set of required IEDs to discriminate faults at the $j$th pair of nodes that are at the same distance from the substation;

$Y_{jk_j}$ 1 if the $k_j$ set of IEDs is used to discriminate faults at the $j$th pair. For instance, if the possible sets of IEDs for discriminating fault at nodes $(M, N)$, (which means $j = 2$) are $x_1$, $x_2$, $x_3$, $x_1x_2$, $x_1x_3$, $x_1x_2x_3$, then $Y_{21} = x_1$, $Y_{22} = x_2$, $Y_{23} = x_3$, $Y_{24} = x_1x_2$, $Y_{25} = x_1x_3$, $Y_{26} = x_1x_2x_3$ ;

$k_j$ pointer to the $k$th set of IEDs to discriminate faults at the $j$th pair;

$|k_j|$ number of IEDs at the $k$th set of IEDs to discriminate faults at the $j$th pair. For instance, in the above example $|k_1| = 1$, $|k_2| = 1$, $|k_3| = 1$, $|k_4| = 2$, $|k_5| = 2$, $|k_6| = 3$.

The constraint (29) ensures that for each pair of the nodes that are at the same distance from the substation ($j$th pair), at least one set of IEDs is installed to determine the actual faulted node. The constraint (30) ensures that a minimum number of IEDs placements are selected. The larger sets have larger $|k_j|$, so (30) tries to select smaller sets.

In addition to the size of the sets, constraint (30) tries to select the sets with minimum new IED installment requirement. In some cases, the larger sets that do not require new IED installment are preferred. For instance, assume $(K, L)$, $(M, N)$, $(O, P)$ are pairs of nodes that are at the same distance from the substation and one IED is installed at node 1 to discriminate $K$ from $L$ and one IED is installed at node 2 to discriminate $M$ from $N$. If possible options for IED placement to discriminate $O$ from $P$ are installing either one IED at node 3 or two IEDs at nodes 1 and 2, the latter $(x_1x_2)$ is preferred because it does not require new IED placement as two IEDs are already installed at node 1 and 2 to discriminate pairs $(K, L)$ and $(M, N)$.

Solving the aforementioned binary linear programming problem yields the number and location of required IEDs. To use the already installed IEDs, which means an IED is already installed at node $i$, $X_i$ is set to 1.

It is also quite possible that the topology of the network changes as a result of circuit reconfiguration. This limits the candidate nodes for IED installment. It may happen that due to the topology change and switching operation, an IED that is responsible for providing information for the proposed fault location method becomes unavailable. To address this potential challenge, the circuit is divided into several zones. Each zone is defined by the installed switches along the feeder as illustrated in Fig. 1. The rule requires that the nodes that are at the same distance from the substation within a given zone should be differentiated from each other by using only the IEDs that are installed at the same zone. For instance and in reference to Fig. 1, to discriminate faults at nodes $A$ and $B$, the candidate locations of IED placement are only the nodes inside the Zone 1. If the nodes that are at the same distance from the substation are at two different zones, all nodes within both zones are possible candidates for IED placement. If only one of the zones is in service, there is no confusion about the actual location of the fault and no information is required from the IEDs. For instance and in reference to Fig. 1, to discriminate faults at nodes $C$ and $D$, the candidate locations for IED placement are

TABLE II
FAULT LOCATION ERRORS (%) USING THE PROPOSED METHOD IN [3]

| Location of the fault | | Node# 1 | | | | Middle of line 7-8 | | | | Node# 13 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Incidence angle | | 0 deg. | | 90 deg. | | 0 deg. | | 90 deg. | | 0 deg. | | 90 deg. | |
| Measurement error | | Without error | With error | Without error | With error | Without error | With error | Without error | With error | Without error | With error | Without error | With error |
| $R_f$=0 Ω | a-g | 0.79 | 6.21 | 0.81 | 6.5 | 9.88 | 14.90 | 9.81 | 14.1 | 5.91 | 8.01 | 6.04 | 8.52 |
| | b-c | 0.87 | 5.92 | 0.85 | 6.02 | 11.16 | 16.05 | 11.361 | 16.59 | 0.96 | 4.98 | 092 | 4.75 |
| | a-b-c | 0.74 | 6.21 | 0.79 | 6.53 | 9.92 | 14.90 | 9.68 | 15.1 | 3.99 | 6.98 | 4.01 | 7.02 |
| $R_f$=10 Ω | a-g | 0.7 | 4.39 | 0.7 | 4.52 | 9.44 | 14.38 | 9.44 | 14.42 | 4.69 | 7.41 | 4.72 | 7.83 |
| | b-c | 1 | 4.14 | 1 | 4.83 | 10.20 | 15.10 | 10.37 | 15.7 | 0.05 | 5.02 | 0.05 | 5.18 |
| | a-b-c | 6.63 | 11.6 | 6.92 | 17.3 | 8.72 | 13.58 | 8.81 | 13.73 | 3.12 | 6.35 | 3.91 | 6.84 |

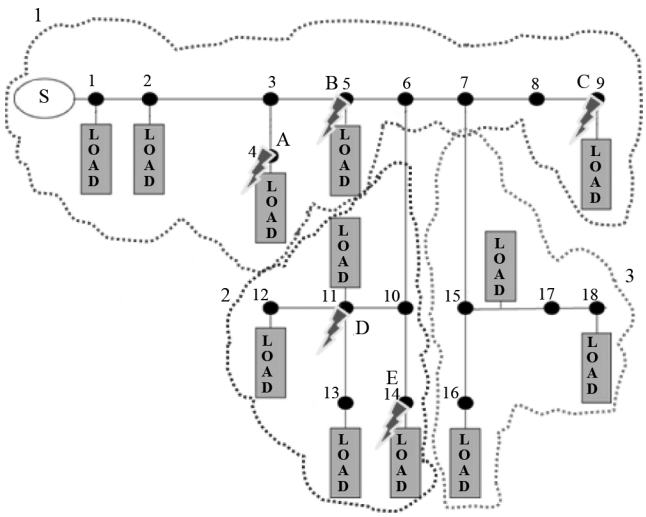


Fig. 1. Different zones of the network defined by the installed switches along the feeder.

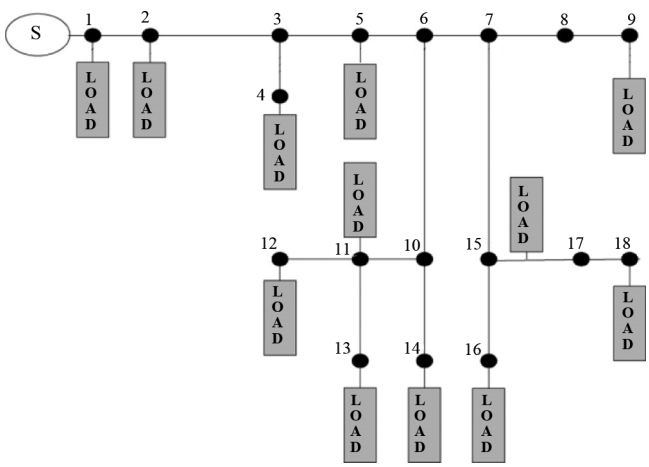


Fig. 2. Power distribution system, Saskpower, Canada [5].

nodes inside Zone 1 and Zone 2. Moreover, some zones are connected to the substation through other zones. For example, in Fig. 1, if Zone 1 goes out of service, zone 2 will also go out of service. Therefore, to discriminate faults at nodes $D$ and $E$, the candidate locations for IED placement are nodes inside Zone 1 and Zone 2. This fact reduces the number of required IEDs. It means that the IED that is installed at Zone 1 can cover faults at both zones. Therefore, if an IED in Zone 1 can discriminate faults in zone 2, the number of required IEDs is reduced. The same situation applied to Zones 1 and Zone 3. However, Zones 2 and Zone 3 are independent from each other and to discriminate fault at nodes $E$ and $D$, nodes at Zone 3 are not candidates for IED placement. To take into account the above in the optimization problem, if node $k$ is not a possible location for discriminating faults at node $i$ from node $j$, $X_k$ is set to zero. It means node $k$ is eliminated from possible location for IED placement because according to (25), $X_k = 0$ means an IED should not be installed at node $k$

It is notable that the optimal IED placement is done for the present network. If the network is expanded in the future and new lines and nodes are added to the network, the optimization procedure should be done again. It may happen that the new optimal locations for IED placements are different from the existing ones. To take into account this issue, if an IED is located at node $I$ in the measurement status vector $X$, $x_i$ is set to 1 which means an IED already exists at node $I$. This will force the optimization to take into account the present location of already installed IEDs

## V. CASE STUDY

In this section, the proposed two-step procedure for ranking fault location methods is applied to a select distribution feeder for illustration purposes, which also includes the process to eliminate false fault location estimations.

According to the qualitative step presented in Section II, two algorithms henceforth referred to as algorithm I [3] and algorithm II [4] were short listed for the subsequent quantitative step. These methods were implemented in MATLAB software. The SaskPowerfeeder model, shown in Fig. 1[5], was simulated in Alternate Transients Program (ATP) software; the simulated voltage and current signals were exported to MATLAB for fault location estimation. The results for a few select scenarios are tabulated and presented in Tables II and III. The error is defined as follows:

$$\text{error} = \frac{|\text{actual fault location - estimated fault location}|}{\text{total line length}} \times 100. \quad (31)$$

The ranking is done via the $2n + 1$ point estimation method. The line parameters and loads are assumed to follow a normal distribution which is typically the case. The mean values are set

TABLE III
FAULT LOCATION ERRORS (%) USING THE PROPOSED METHOD IN [4]

| Location of the fault | | Node# 1 | | | | Middle of line 7-8 | | | | Node# 13 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Incidence angle | | 0 deg. | | 90 deg. | | 0 deg. | | 90 deg. | | 0 deg. | | 90 deg. | |
| Measurement error | | Without error | With error | Without error | With error | Without error | With error | Without error | With error | Without error | With error | Without error | With error |
| $R_f$=0 Ω | a-g | 0.8 | 5.73 | 0.8 | 5.84 | 0.41 | 21.34 | 0.5 | 24.5 | 0.02 | 1.24 | 0.02 | 2.03 |
| | b-c | 0.8 | 6 | 0.8 | 6 | 0.31 | 4.72 | 0.35 | 4.63 | 0.22 | 1.20 | 0.19 | 1.42 |
| | a-b-c | 0.8 | 5.87 | 0.8 | 5.6 | 0.8 | 3.9 | 0.8 | 4.1 | 0.453 | 3.05 | 0.462 | 3.41 |
| $R_f$=10 Ω | a-g | 10.67 | 21.33 | 10.8 | 12.1 | 0.45 | 2.87 | 0.52 | 2.63 | 3.05 | 3.61 | 3.1 | 3.61 |
| | b-c | 10.67 | 16.66 | 11.03 | 16.58 | 1.64 | 2.87 | 1.86 | 2.87 | 2.17 | 2.65 | 2.17 | 2.77 |
| | a-b-c | 2 | 7.33 | 1.8 | 7.42 | 0.92 | 5.74 | 0.92 | 5.92 | 0.44 | 3.13 | 0.44 | 3.34 |

TABLE IV
FALSE FAULT LOCATION ESTIMATION ELIMINATION RESULTS

| Scenario | Fault type | Actual location | Possible locations using [3] | Flag value | Final estimation |
|---|---|---|---|---|---|
| 1 | b-g (5 Ω) | node# 10 (30.093km) | node #10 (29.96 km) | $1.3\times10^{-2}$ | node# 10 |
| | | | node#7 (29.94 km) | $2.02\times10^{-3}$ | |
| 2 | a-b (10 Ω) | node# 4 (20.92 km) | node #4 (20.68 km) | $1.0\times10^{-2}$ | node# 4 |
| | | | Line between node 3 and 5 (20.73km) | $3.25\times10^{-3}$ | |
| 3 | a-c-g (5 Ω) | node# 4 (20.92 km) | node #4 (20.70 km) | $1.1\times10^{-2}$ | node# 4 |
| | | | line between node 3 and 5(20.75 km) | $3.46\times10^{-3}$ | |
| 4 | c-g (2 Ω) | line between node 3 and 5(19.5 km) | line between node 3 and 5 (19.38 km) | $0.9\times10^{-2}$ | line between node 3 and 5 (19.38 km) |
| | | | line between node 3 and 4 (19.13 km) | $2.94\times10^{-3}$ | |

at the nominal (initial) values and the standard deviations are assumed to be 10% of the mean values. It is also assumed that the measurement error can be modeled by a normal distribution with standard deviations of 2% of the mean values. The ranges of errors in voltage phasors were selected per IEC standards for voltage transformers [27].

The fault resistance can vary from 0 Ω to that of high impedance faults. If historical data are available, it is possible to derive the associated probability distribution function. In view of the lack of historical data, it was assumed that the fault resistance follows a uniform distribution function. The lower limit is assumed to be 0 Ω and the upper limit is set to the value ensuring that the fault could still produce short circuit currents large enough to cause the protective devices to operate, typically 1.5 times load current. Using the $2n+1$ point estimation method, the $X_{\text{threshold}}$ in (4) for method II is 0.1326 and for method I is 0.1138, which means that particular to this network and assumptions, the first algorithm ranks higher than the alternative. To eliminate the false fault location estimation by method I, the voltage sag data gathered from select IEDs along the feeder were utilized. According to the proposed optimal IED placement method presented in Section V, nodes 8, 11, and 17 are the optimal locations for IED placement. The defined zones are similar to Fig. 1 which means it is assumed there are two switches at nodes 6 and 7

Three different scenarios are included to demonstrate the performance of the proposed method to eliminate false fault location estimations. In each scenario, the location that has the largest value of the Flag is the selected location of the fault. Table IV shows the results where the false fault location estimations are eliminated, leaving the actual location of fault as the only answer

## VI. CONCLUSIONS

A systematic approach for fault location method evaluation and selection that involves a qualitative analysis and a quantitative ranking is presented as follows.

- The qualitative analysis is first conducted to short list the most appropriate fault location methods meeting the limitations and requirements of the network under study.
- A quantitative procedure for fault location selection is developed based on the $2n+1$ point estimation method and the outputs of this analysis are ranked following a rule derived from the Chebyshev's inequality.
- In cases where the highest ranking method still suffers from multiple fault location estimations, a method based on voltage sag data gathered from few IEDs along the feeder is proposed to eliminate false fault estimations.
- An optimization problem for IED placement was formulated and solved to minimize costs and determine the most appropriate number and location of IEDs.

The proposed approach was demonstrated in the simulation environment where the performance of the proposed methodology steps was examined and discussed through a few scenarios and case studies.

## REFERENCES


[1] "GRID 2030" a national vision for electricity's second 100 years. 2003. [Online]. Available: http://www.climatevision.gov/sectors/electricpower/pdfs/electric_vision.pdf

[2] A. Girgis, C. M. Fallon, and D. L. Lubkeman, "A fault location technique for rural distribution feeders," *IEEE Trans. Ind. Appl.*, vol. 29, no. 6, pp. 1170–1175, Nov./Dec. 1993.

[3] J. Zhu, D. L. Lubkeman, and A. A. Girgis, "Automated fault location and diagnosis on electric power distribution feeders," *IEEE Trans. Power Del.*, vol. 12, no. 2, pp. 801–809, Apr. 1997.

[4] R. Das, "Determining the locations of faults in distribution systems," Ph.D. dissertation, Univ. Saskatchewan, Saskatoon, SK, Canada, 1998.

[5] J. Mora-Florez, J. Melendez, and G. Carrillo-Caicedo, "Comparison of impedance based fault location methods for power distribution systems," *Elect. Power Syst. Res.*, vol. 78, pp. 657–666, 2008.

[6] M. S. Choi, S. J. Lee, D. S. Lee, and B. G. Jin, "A new fault location algorithm using direct circuit analysis for distribution systems," *IEEE Trans. Power Del.*, vol. 19, no. 1, pp. 35–41, Jan. 2004.

[7] M. S. Choi, S. J. Lee, S. I. Lim, D. S. Lee, and X. Yang, "A direct three-phase circuit analysis-based fault location for line-to-line fault," *IEEE Trans. Power Del.*, vol. 22, no. 4, pp. 2541–547, Oct. 2007.

[8] R. K. Aggarwal, Y. Aslan, and A. T. Johns, "New concept in fault location for overhead distribution systems using superimposed components," *Proc. Inst. Elect. Eng., Gen., Transm. Distrib.*, vol. 144, no. 3, pp. 309–316, May 1997.

[9] R. K. Aggarwal, Y. Aslan, and A. T. Johns, "An interactive approach to fault location on overhead distribution lines with load taps," *Develop. Power Syst. Protect.*, pp. 184–187, Mar. 1997.

[10] J. J. Mora, G. Carrillo, and L. Perez, "Fault location in power distribution systems using ANFIS nets and current patterns," in *Proc. IEEE Power Eng. Soc. Transm. Distrib. Conf. Expo. Latin America*, Venezuela, 2006, pp. 1–7.

[11] L. S. Martins, J. F. Martins, C. M. Alegria, and V. Femilo Pires, "A network distribution power system fault location based on neural eigenvalue algorithm," presented at the IEEE Bologna Power Tech Conf., Bologna, Italy, Jun. 23–26, 2003.

[12] J. Mora-Florez, V. Barrera-Nuez, and G. Carrillo-Caicedo, "Fault location in power distribution systems using a learning algorithm for multivariable data analysis," *IEEE Trans. Power Del.*, vol. 22, no. 3, pp. 1715–1721, Jul. 2007.

[13] Z. Q. Bo, G. Weller, and M. A. Redfern, "Accurate fault location technique for distribution system using fault-generated high-frequency transient voltage signals," *Proc. Inst. Elect. Eng., Gen. Transm. Distrib.*, vol. 146, no. 1, Jan. 1999.

[14] D. W. P. Thomas, R. J. O. Carvalho, and E. T. Pereira, "Fault location in distribution systems based on traveling waves," presented at the IEEE Bologna Power Tech Conf., Bologna, Italy, Jun. 23–26, 2003.

[15] H. Nouri, C. Wang, and T. Davies, "An accurate fault location technique for distribution lines with tapped loads using wavelet transform," *Proc. IEEE Porto Power Tech.*, vol. 3, pp. 1–4, Sep. 2001.

[16] A. Saltelli, M. Ratto, T. Andres, F. Campolongo, J. Cariboni, D. Gatelli, M. Saisana, and S. Tarantola, *Global Sensitivity Analysis: The Primer*. Hoboken, NJ: Wiley, 2008.

[17] D. C. Montgomery and G. C. Runger, *Applied statistics and probability for engineers*. Hoboken, NJ: Wiley, 2003.

[18] H. P. Hong, "An efficient point estimate method for probabilistic analysis," *Rel. Eng. Syst. Safety*, vol. 59, pp. 261–267, 1998.

[19] J. T. Christian and G. B. Baecher, "The point-estimate method with large numbers of variables," *Int. J. Numer. Anal. Meth. Geomech.*, vol. 26, pp. 1515–1529, 2002.

[20] P. Caramiaa, G. Carpinelli, and P. Varilonec, "Point estimate schemes for probabilistic three-phase load flow," *Elect. Power Syst. Res.*, vol. 80, pp. 168–175, 2010.

[21] J. M. Morales and J. Perez-Ruiz, "Point estimate schemes to solve the probabilistic power flow," *IEEE Trans. Power Syst.*, vol. 22, no. 4, pp. 1594–1601, Oct. 2007.

[22] S. Lotfifard, L. Xie, and M. Kezunovic, "Quantifying the impact of unscheduled line outages on locational marginal prices," presented at the 42nd North Amer. Power Symp., Arlington, TX, Sep. 26–28, 2010.

[23] M. H. J. Bollen, *Understanding Power Quality Problems: Voltage Sags and Interruptions*. New York: IEEE, 1999.

[24] M. H. J. Bollen, "Algorithm for characterizing measured three-phase unbalanced voltage dips," *IEEE Trans. Power Del.*, vol. 18, no. 3, pp. 937–944, Jul. 2003.

[25] S. Lotfifard, M. Kezunovic, and M. J. Mousavi, "Distribution fault location using voltage sag data," *IEEE Trans. Power Del.*, vol. 26, no. 2, pp. 1239–1246, Apr. 2011.

[26] CPLEX. [Online]. Available: http://www.ilog.com/products/cplex/

[27] *Voltage Transformers*, IEC 60186-am2, Ed 2.0, Int. Elect. Tech. Comm, 1995.


**Saeed Lotfifard** (S'08–M'11) received the Ph.D. degree in electrical engineering from Texas A&M University, College Station, TX, in 2011.

Currently, he is an Assistant Professor at the University of Central Florida, Orlando. His research interests include power systems protection and control, intelligent monitoring and outage management-self healing grids, distribution automation, cyber-physical energy system/smart grid technologies, as well as applications of statistical methods and signal processing in power systems.

**Mladen Kezunovic** (S'77–M'80–SM'85–F'99) received the Dipl. Ing. degree from the University of Sarajevo in 1974, and the M.S. and Ph.D. degrees in electrical engineering from the University of Kansas, Lawrence, in 1977 and 1980, respectively.

Currently, he is the Eugene E. Webb Professor of Electrical and Computer Engineering at Texas A&M University, College Station, TX. He also serves as the Site Director of Power Engineering Research Center (PSerc) and the Deputy Director of Electrical Vehicles Transportation and Electricity Convergence (EV-TEC) Center. He is the Principal of Test Laboratories International, a consulting firm specializing in automated fault analysis and intelligent electronic devices (IED) testing since 1992. His main research interests are digital simulators and simulation methods for relay testing, as well as application of intelligent methods to power system monitoring, control, and protection.

Dr. Kezunovic is a member of CIGRE and a Registered Professional Engineer in Texas.

**Mirrasoul J. Mousavi** (S'01–M'05) received the Ph.D. degree in electrical engineering from Texas A&M University, College Station, TX, in 2005.

Currently, he is a Principal Scientist Engineer with ABB US Corporate Research, Raleigh, NC, responsible for leading smart-grid projects related to power system automation, intelligent monitoring, control, and protection applications. He was a Researcher in the Power System Automation Laboratory and a Graduate Lecturer at Texas A&M University prior to joining ABB. From 1999 to 2001, he was an R&D Engineer with Niroo Research Institute (NRI). His current professional interests are related to power system automation, fault diagnosis, asset management, as well as power system modeling and simulation.

Dr. Mousavi is a member of the IEEE Power and Energy Society (PES) and IEEE Dielectrics and Electrical Insulation Society (DEIS).