Novel system model-based fault location approach using dynamic search technique

Sundaravaradan Navalpakkam Ananthan | Alvaro Furlani Bastos | Surya Santoso

Abstract
Impedance-based fault location (IBFL) approaches are the most commonly used fault location methods in digital relays. However, each IBFL approach is designed specific to a line or network configuration and thus is not universal. For example, complex line configurations such as lines with mutual coupling between them and three-terminal lines have to employ individual IBFL algorithms derived specifically for them. Furthermore, they suffer from several sources of errors such as non-homogeneous system, and CT saturation. Hence, this paper presents a novel fault location approach that utilises a system model to overcome these limitations. The proposed model-based fault location (MBFL) approach estimates the fault location by identifying the closest match among various anticipated fault scenarios obtained using the system model and the actual fault scenario. It uses a dynamic search technique to implement the MBFL efficiently. A key highlight of the proposed approach is identifying the location of a fault on a neighbouring line using limited measurements, as few as only the through fault current flowing in a neighbouring line. The advantages of the approach and its practical applicability have been demonstrated by implementing it in complex network configurations as well as field data.

1 | INTRODUCTION
Transmission and distribution lines are crucial components of any power system. They form the backbone of the present-day interconnected power system and have enabled the transfer power to loads spread across vast regions. However, for a wide variety of reasons such as natural events, equipment failure, physical accidents, and misoperation, short-circuit faults on overhead lines are unavoidable [1, 2]. Transmission lines can span for long distances ranging from a few tens of miles to hundreds of miles. On the other hand, a distribution system can be a complex network consisting of several taps along its downstream path. Hence, it is essential to clear the fault and restore the line to normal operation. Dispatching the maintenance crew directly to the fault location saves a tremendous amount of resources in the restoration process, which has led to the development of several fault location approaches. Though fault location has been a subject of research interest for decades, only a few algorithms have been used due to practical considerations and limitations of these approaches.

For several decades, impedance-based fault location (IBFL) methods have been used in commercial industrial relays [3–6]. Each IBFL approach has its own input data requirements and is designed specific to a line configuration. For example, most of the commonly used IBFL algorithms are valid only for simple two-terminal lines. Each complex line configuration, such as three-terminal line, mutually coupled line, series compensated line, and parallel transmission lines, has to employ a separate IBFL algorithm derived specifically for it [6–11]. Furthermore, they assume the line impedances to be homogeneous. Hence, these algorithms provide accurate fault location estimates only when their specific requirements and assumptions have been met. Some of the common factors which affect IBFL methods and their sources of errors are discussed in [12, 13].

Recently, apart from IBFL approaches, traveling-wave (TW) fault location methods are being used in relays as well [3, 14, 15]. These methods require highly accurate timing information and waveforms to be recorded at a very high sampling frequency (in the rate of 100s of kHz or MHz) [9]. As a result,
they need appropriate sensors, devices with very high sampling frequency, large storage and processing capability, which makes the approach very expensive to implement. In addition, [6] discusses some factors which affect the accuracy of traveling-wave based approaches. Though there are other methods proposed based on frequency-domain analysis or wide-area fault locating techniques based on synchrophasor measurements, they are not yet commonly commercially used. Hence, there is a need for a new fault location approach. The new approach would be desirable to suit any power system configuration, and have fewer restrictions and measured data requirements. Furthermore, an approach that does not require the installation of additional hardware such as high-frequency sampling devices or phasor measurement units will make it more economical to implement.

The power system model and information regarding the system are usually underutilised when performing fault location. IBFL and many other fault location approaches commonly use only the line parameters of the faulted line and source impedances connected to the faulted line in their algorithm. However, valuable information surrounding the faulted line, though they are known, are not exploited. For example, through current is the fault current flowing through a line feeding an external fault (a fault present elsewhere on the system and not on that line). These through fault current measurements can be used to identify the location of a fault on a neighbouring line using the system model. Loss of accurate measurements caused by a variety of factors and events hamper the fault location process. A common cause is current transformer (CT) saturation and there are several works in the literature citing the observation of CT saturation in event reports [16, 17]. Improper or missing connections to measurement devices also lead to loss of input data [18, 19]. Besides, some relays may not have the ability to record measurements or may not be triggered by the fault [20]. Wide-area fault location algorithms have been proposed to tackle scenarios where the fault recording devices at the terminals of a line fail to record the fault waveform [21]. However, these algorithms usually require measurements from multiple locations, which severely limit their practical application to only systems with wide-area monitoring capabilities. Hence, identifying the location of a fault present on neighbouring lines using minimal number of measurements (as less as currents from only one terminal of a neighbouring line) is a unique and valuable benefit. Furthermore, this capability can be used as a second step verification process to the fault location estimated using other methods.

Fault resistance is a critical factor that affects the fault currents and voltages measured by a relay. Though most approaches aim to negate the effect of fault resistance when estimating the fault location, it is rather beneficial to identify the fault resistance value as it presents an insight into the root cause of the fault [22, 23]. Furthermore, knowledge of the fault resistance value enables the usage of several other fault analysis algorithms that require the fault resistance value in its calculations as well as to replicate the fault scenario in a power system model simulation.

Current and voltage measurements extracted from event reports recorded by relays or monitoring devices are used to identify the location of a fault. The power system parameters and network configuration affect these measured quantities and the relationship between them is used to identify the fault location. Engineers use the power system model for a variety of purposes such as to perform protection coordination, system power flow, and stability analysis. Hence, electric utilities commonly have their power system modeled in a power system simulation software such as those presented in [24, 25]. Furthermore, several approaches have been developed to extract various power system parameters to form, update, and verify the power system model [26–28]. Thus, system model based fault location techniques are promising fault location approaches that can overcome the limitations of IBFL and TW based methods.

A form of model-based fault location (MBFL) approach matches the actual fault currents recorded by a relay or fault recorder with fault currents of simulated fault scenarios. Though similar ideas have been presented in [29, 30], they need to be further explored with rigorous analysis to uncover the full potential of the approach. Furthermore, an effective implementation method as well as a study on critical factors that affect these approaches are required.

Based on the above background and motivation, the objective of this paper is to present a fault location method that overcomes the limitations of IBFL approaches as well as provides the novel benefit of identifying the location of a fault on a neighbouring line using limited number of measurements. The challenges in identifying the fault location in multi-terminal lines and on neighbouring lines are analysed and their input data requirements are deduced. Furthermore, critical factors such as the effect of remote source voltage as well as remote source impedance on the fault location estimates obtained using the proposed method are studied. The paper also presents the successful application of the proposed method in complex network configurations as well as in real-world fault event to demonstrate its practical applicability and superior performance over conventionally used fault location methods.

The major contribution is the novel fault location method that utilises the system model to identify the location and fault resistance of a fault on the line monitored by the relay as well as on neighbouring lines that are outside the primary protection zone of the relay. The paper is organised as follows. Section 2 introduces MBFL and Section 3 presents the proposed approach and a demonstration of the implementation and application of the approach. Section 4 discusses the challenges involved in identifying the location of a fault on multi-terminal lines and neighbouring lines, and extends the application of the proposed method to these cases. Section 5 studies the effect of remote voltage source and system parameters on the proposed approach. Section 6 presents comparisons with IBFL methods and highlights the benefits of the approach.
2 MODEL-BASED FAULT LOCATION

The underlying concept of a model-based fault location approach is to simulate a variety of fault scenarios using the power system model and identify fault scenarios that closely match the measured fault currents recorded by the relay to estimate the fault location. An overview of the proposed fault location approach is presented in Figure 1. When a fault occurs on the system, the electric utility operators receive fault event reports recorded by relays or fault recorders. The first step is the data preprocessing step where fault current phasors and other information such as fault type are extracted from the fault event reports. The next step is to determine whether the fault is on the line monitored by the relay or on a neighbouring line. This process is discussed in Section 3.1. Once the fault is determined to either be on the line monitored by the relay or on a neighbouring line, the fault location and fault resistance are estimated using the steps presented in Section 3.2.

The approach is implemented using two software programs. A control and processing software (CPS) is used to preprocess the event report and perform analysis and calculations. A power system simulation software (PS) is used to simulate fault scenarios and feed the results to the CPS. Though electric utilities contain large circuit models of a vast region of the transmission network, the fault location analysis focuses only on a small region surrounding the relay recording the measurements. Hence, fault scenario simulations in this approach are carried out using a reduced equivalent circuit (REC) rather than the large full circuit model to speed up the fault analysis process and reduce the memory and computation required. The REC will typically include the line monitored by the relay and neighbouring lines up to one or two buses away from the monitored line.

The proposed method uses fundamental frequency fault current phasors extracted from the steady-state fault portion of the event report as its input. The phasors represented in rectangular coordinates are used for all the calculations and analysis. Using phasors provide additional information regarding the underlying system impedance during the fault rather than using only the fault current magnitude. The process of extracting phasors for post-fault analysis is a well-known process [31, 32]. The fault type is usually available from the event report. It can also be determined using the fault current and/or voltage measurements [4, 21, 33]. Though fault type identification is a common process, it needs to be detected accurately as only fault scenarios of the detected fault type will be simulated in the proposed dynamic search algorithm. The event report will also contain information regarding which relay recorded the event report. This information can be extracted in the data preprocessing step to identify the appropriate REC to be used for analysis from a library of RECs for various relay locations.

3 PROPOSED FAULT LOCATION APPROACH

The proposed approach is presented along with a demonstration. The test circuit used for this purpose is the same transmission network as that used in [30]. The circuit contains complex network configurations such as three-terminal lines and parallel lines with mutual coupling. Figure 2 presents the REC formed from a large transmission network. It shows the equivalent sources and actual lines present but does not show the equivalent line impedances that were introduced into the system during the circuit reduction process. The circuit models used in this paper are classical short-circuit models and do not consider load currents. This is a reasonable assumption because the loads’ current levels in a transmission system are negligible in comparison to the fault current levels. As a result, the fault currents in the non-faulted phases are negligible or zero in the examples presented in this paper.

In this paper, OpenDSS was used as the PS to simulate fault scenarios on the REC. MATLAB was used as the CPS to drive OpenDSS to perform data preprocessing as well as post-fault analysis. To demonstrate the approach, a fault was created in the large transmission network model and the observed fault current phasors were used as inputs to the proposed approach as shown in Figure 3. In real-world applications, the fault current phasors can be extracted from event reports as illustrated in [31] and be used in the proposed algorithm implemented inside the protection device such as a relay or in a computer at the substation.

Consider a single line-to-ground (SLG) fault in phase A on a two-terminal line (Line 4) as shown in Figure 2. The measurements from relay R installed near Bus 153 are used to identify the location of this fault. The fault currents in phases A, B, and C were 2562.01∠−70.4° A, 17.17∠−69.2° A, and 17.17∠−69.2° A, respectively. The proposed approach first identifies whether the fault is on the monitored line or a neighbouring line and then estimates the fault location.
3.1 Stage 1: Determining the relative location of the fault

The process to determine whether the fault lies on the monitored line or a neighbouring line is outlined in Figure 4. For this purpose, a set of boundary equations are created using four sets of simulated fault scenarios by varying the fault location and fault resistance values. The four sets of fault scenarios are:

1. Fault with various fault resistance values within the fault resistance range at the local end of the line (fault at 0.01 pu of the line length)
2. Fault with various fault resistance values within the fault resistance range at the remote end of the line (fault at 0.99 pu of the line length)
3. Faults at various locations along the monitored line with the least fault resistance value (0 Ω)
4. Faults at various locations along the monitored line with the maximum fault resistance value (50 Ω in this analysis)

The fault current in one of the faulted phases for each set of boundary fault scenarios is fitted to a cubic polynomial shown in (1). A cubic polynomial was identified to be sufficient to capture the fault current trend in each of the four cases. Any curve fitting approach, such as the method of least squares, can be used to form these equations.

\[ b = \alpha_0 + \alpha_1 a + \alpha_2 a^2 + \alpha_3 a^3. \] (1)

The \( b \) component is the real part of the fault current and the \( a \) component is the imaginary part of the fault current.

Each set of the above listed fault scenarios consists of the extreme or boundary fault cases: faults at the beginning and far end of the line, and faults with minimum and maximum fault resistance values. Hence, faults anywhere on the line with any fault resistance value will lie within the region delimited by these boundary equations and is proved in the appendix section. As a result, when the measured fault current lies within the region delimited by these boundary equations, it is determined to be on the monitored line. Otherwise, it is located on a neighbouring line.
This process is demonstrated using Figure 5(a) for the example fault scenario. A-phase currents were used to plot the boundary equations as the fault was an SLG fault on phase A. In this paper, faults with 0, 5, 12.5, 25, and 50 Ω were used as the fault resistances in sets 1 and 2 and faults at 0.01, 0.33, 0.66, and 0.99 pu of line length were used as the fault locations in sets 3 and 4. The maximum fault resistance considered for analysis is 50 Ω as faults with higher fault resistance values have very low fault current and require special high-impedance fault location techniques. The four cubic polynomial boundary equations were formed using the curve fitting toolbox in MATLAB and plotted as four curves in the figure. Curves 1–4 in the figure correspond to the above enumerated list, that is, curve 1 shows faults with various fault resistance values at 0.01 pu of line length on the monitored line, curve 2 shows faults with various fault resistance values at 0.99 pu of line length on the monitored line, curve 3 shows faults at various locations along monitored line with a fault resistance value of 0 Ω, and curve 4 shows faults at various location along the monitored line with a fault resistance value of 50 Ω. The fault current in the faulted phase, indicated by the ‘x’ marker, falling within the boundary conditions identifies the fault to be on the monitored line, as expected. Figure 5(b) presents an example plot where the fault is not present on the monitored line, but instead on a neighbouring line. The fault scenario in this plot is an SLG fault on phase A present on Line 5 at 0.4 pu from Bus 153 with a fault resistance of 2 Ω. The fault current measured by relay R, indicated by the ‘x’ marker, lies outside the region delimited by the boundary equations, thus indicating that the fault lies outside the monitored line (Line 4).

Once the fault is determined to be on the monitored line, the fault location and fault resistance are estimated by performing fault analysis on the monitored line. If the fault is determined to lie outside the monitored line, there are two ways to proceed as discussed in Section 4.3. Fault analysis is either performed on each neighbouring line and the result from the appropriate faulted line is chosen with the assistance of a protection engineer, or available additional input measurement is used to estimate the fault location independently.

3.2 Stage 2: Estimating the fault location and fault resistance

The fault location and fault resistance estimation process is presented in Figure 6. It involves three steps. The first step is determining the fault simulation scenarios and is presented in Section 3.2.1. The next step is to simulate the fault cases and identify the closest simulation scenarios to the actual fault scenario, as explained in Section 3.2.2. This process is repeated until the end of simulations search criteria is reached. Finally, the fault
location and fault resistance are estimated using the weighted mean method discussed in Section 3.2.3.

3.2.1 Determining simulation scenarios

The simulation scenario set defines all fault scenarios to be simulated by the PS for each iteration of the search algorithm. The fault location (FL) and fault resistance (FR) of each simulated fault scenario are determined in this step. The set of simulation scenarios is an interpolation between the range of values to be explored for fault location and fault resistance, respectively. For each fault location value in the fault location set (L\(_{\text{set}}\)), all the fault resistance values in the fault resistance set (R\(_{\text{set}}\)) are simulated. There are two parameters for both FL and FR that determine their simulation set: a) the incremental step size, and b) the range of values.

First, a default initial set of fault scenarios are simulated to be used in the first iteration of the k-nearest neighbour (kNN) search (discussed in 3.2.2). This simulation set explores the entire range of possible fault location (between 0.01 pu and 0.99 pu of the line length) and fault resistance (between 0 \(\Omega\) and 50 \(\Omega\)) values with a very large step size. The subsequent iterations will narrow down to the actual fault location and will employ smaller step sizes as discussed later in this subsection. The initial fault location set is L\(_{\text{set}}\) = \{0.01, 0.33, 0.66, 0.99\} pu and the fault resistance set is R\(_{\text{set}}\) = \{0, 5, 12.5, 25, 50\} \(\Omega\). The variable quantity such as fault location or fault resistance is denoted using capital letters, the subscript following the variable provides details on the variable, and the superscript refers to the iteration number. This initial set of simulated fault current data is also used to determine whether the fault lies on the monitored line discussed in Section 3.1.

The initial simulation set is used to run the first search iteration. The results from each iteration of the kNN search is used to determine the step size and range of FL and FR values for the next set of simulation scenarios.

Step size

\[
L_{\text{step}}^{n+1} = \begin{cases} 
\frac{L_{\text{max}}^{n} - L_{\text{min}}^{n}}{n - n} & \text{if } L_{\text{max}}^{n} - L_{\text{min}}^{n} > \delta_L, \\
\frac{L_{\text{step}}^{n}}{n - n} & \text{for all other cases}
\end{cases}
\]

(2)

\[
R_{\text{step}}^{n+1} = \begin{cases} 
\frac{R_{\text{max}}^{n} - R_{\text{min}}^{n}}{n - n} & \text{if } \frac{R_{\text{max}}^{n}}{R_{\text{min}}^{n}} > \delta_R, \\
\frac{R_{\text{step}}^{n}}{n - n} & \text{for all other cases}
\end{cases}
\]

(3)

In these equations, \(\text{max}\) refers to the largest value and \(\text{min}\) refers to the smallest value among the results from the search algorithm discussed in Section 3.2.2. For example, \(L_{\text{max}}^{n}\) refers to the fault scenario with the largest FL value among the closest fault scenarios chosen by the search algorithm in the \(n^{th}\) iteration. \(N\) is the maximum number of iterations the approach is programmed to run. In the various fault scenarios tested by the authors, five iterations were sufficient to accurately locate the fault. Hence, \(N\) was set to five in this paper. The initial step sizes before the first iteration are taken as \(L_{\text{step}} = 0.33\) pu and \(R_{\text{step}} = 5\ \Omega\). On analysing a variety of circuits and fault types, a suitable value for both \(\delta_L\) and \(\delta_R\) was identified as 0.05 pu. The step size gradually decreases with the number of iterations in (2) and (3) to obtain a more local search within the identified range of FL and FR values in the subsequent iterations.

After calculating \(L_{\text{step}}^{n}\) and \(R_{\text{step}}^{n}\), if \(L_{\text{step}}^{n} = L_{\text{start}}^{n}\) and \(R_{\text{step}}^{n} = R_{\text{start}}^{n}\), then

\[
L_{\text{step}}^{n+1} = \frac{L_{\text{step}}^{n}}{n} \quad \text{and} \quad R_{\text{step}}^{n+1} = \frac{R_{\text{step}}^{n}}{n}.
\]

(4)

This is to ensure that the search is progressive and prevent the approach from getting stuck with the same set of scenarios in a few rare unusual cases.

Range

The range of values to be simulated is defined using the start and the end values obtained using the following equations.

\[
L_{\text{start}}^{n+1} = \begin{cases} 
L_{\text{min}}^{n} & \text{if } L_{\text{min}}^{n} - L_{\text{step}}^{n} < 0, \\
L_{\text{min}}^{n} - L_{\text{step}}^{n} & \text{for all other cases}
\end{cases}
\]

(5)

\[
L_{\text{end}}^{n+1} = \begin{cases} 
L_{\text{max}}^{n} & \text{if } L_{\text{max}}^{n} + L_{\text{step}}^{n} > 1, \\
L_{\text{max}}^{n} + L_{\text{step}}^{n} & \text{for all other cases}
\end{cases}
\]

(6)

\[
R_{\text{start}}^{n+1} = \begin{cases} 
R_{\text{min}}^{n} & \text{if } R_{\text{min}}^{n} - R_{\text{step}}^{n} < 0, \\
R_{\text{min}}^{n} - R_{\text{step}}^{n} & \text{for all other cases}
\end{cases}
\]

(7)

\[
R_{\text{end}}^{n+1} = \begin{cases} 
R_{\text{max}}^{n} & \text{if } R_{\text{max}}^{n} + R_{\text{step}}^{n} > 50, \\
R_{\text{max}}^{n} + R_{\text{step}}^{n} & \text{for all other cases}
\end{cases}
\]

(8)

The start and end values of each successive iteration are a step value beyond the minimum and maximum values of their respective quantities among the search results for the \(n^{th}\)
iteration unless they exceed the boundary values. However, in cases where \( L_{\text{max}}^{e} = L_{\text{min}}^{e} \), then two steps beyond the lower and upper limits are chosen for the start and end values, respectively.

The fault location set \((L_{\text{sf}})\) and fault resistance set \((R_{\text{sf}})\) are created as follows:

\[
L^{e+1}_{\text{sf}} = \left[ L^{e+1}_{\text{start}} : L^{e+1}_{\text{end}} \right] \text{ with step size of } L^{e+1}_{\text{step}},
\]

\[
R^{e+1}_{\text{sf}} = \left[ R^{e+1}_{\text{start}} : R^{e+1}_{\text{end}} \right] \text{ with step size of } R^{e+1}_{\text{step}}.
\]

The CPS software drives the PS to simulate the specific fault scenarios identified in the fault location and fault resistance sets. The total number of scenarios simulated for each iteration is: number of elements \((L_{\text{sf}}) \times \text{number of elements} \,(R_{\text{sf}})\).

### 3.2.2 Identifying closest simulation scenarios

The data obtained from simulating the scenarios in the fault location and fault resistance sets are now used to narrow down the search space and obtain fault scenarios that closely represent the actual fault scenario. This step uses the k-nearest neighbour (kNN) search to determine the region to zoom into and focus the search in the next iteration. Given a set of points \( S \) in a metric space \( P \) and a query point \( q \in P \), the kNN search finds the \( k \) nearest or closest points in \( S \) to \( q \). Here, the set of points \( S \) are the simulated fault currents and the query point \( q \) is the actual measured fault current. In this paper, Euclidean distance is used as the distance metric for applying kNN search. The approach uses fault currents observed at the relay location of all three phases represented in terms of rectangular coordinates. Hence, the query point and the metric space are 6-dimensional \((\mathbb{R}^6)\). In short, the kNN algorithm identifies \( k \) number of scenarios in the search space that are closest to the actual fault scenario using Euclidean distance as the distance metric.

The smallest and the largest fault location \((L_{\text{min}}^{e} \text{ and } L_{\text{max}}^{e})\) and fault resistance \((R_{\text{min}}^{e} \text{ and } R_{\text{max}}^{e})\) values among the \( k \) simulation scenarios identified by the kNN search algorithm are used to determine the step size and range of values to be simulated in the next iteration (discussed in Section 3.2.1). A very small \( k \) aggressively narrows down the search space, thus rapidly speeding up the fault location process. However, it can also mislead the approach’s search space and not converge to the actual fault location. On the other hand, a larger \( k \) value gradually narrows down on the search space. Though it slows down the process, it is less likely to mislead the approach’s search space in the subsequent iterations. Hence, an optimal value of \( k \) needs to be chosen. A \( k \) value of four worked well for all the scenarios tested by the authors and is used in all the demonstrations presented in this paper.

As the proposed search algorithm progresses, it moves closer to the actual fault location. This can be observed with the decrease in the Euclidean distance \((D)\) between the simulated fault current vector and the measured fault current vector in the simulation scenarios identified by the kNN search algorithm. The search simulations are stopped when either of the following two criteria is reached. The first is when \( D \) of at least one of the simulation scenarios identified by the kNN search algorithm is less than a set value \( \gamma \). The value of \( \gamma \) was determined by observational analysis and set at 50 in this paper. The other criterion is where the maximum number of allowed iterations \((N)\) is reached.

### 3.2.3 Determining fault location and fault resistance

A weighted mean approach is finally used to estimate the fault location and fault resistance using the closely representative simulation scenarios determined by the kNN search performed in the previous steps. The weights \((W)\) are assigned using (11).

\[
W'(p) = \frac{100}{D(p)},
\]

where \( p \) represents the index of the \( k \) nearest neighbours identified at the last iteration of the dynamic simulation search. In (11), a larger weight is assigned when the simulated fault scenario is closer to the actual fault location (when the distance \( D \) is lesser).

Let \( FL_{\text{kNN}} \) and \( FR_{\text{kNN}} \) represent the fault locations and the fault resistances of the simulated fault scenarios identified by the kNN search in its final iteration when the search is stopped. Then, the final estimated fault location and fault resistance value of the given fault are calculated using (12) and (13), respectively.

\[
\text{Estimated Fault Location} = \frac{\sum_{p=1}^{k} W'(p) \ast FL_{\text{kNN}}(p)}{\sum W'}, \quad \text{pu}, \quad (12)
\]

\[
\text{Estimated Fault Resistance} = \frac{\sum_{p=1}^{k} W'(p) \ast FR_{\text{kNN}}(p)}{\sum W'} - \Omega, \quad (13)
\]

As only the fault scenarios that are relevant to the actual fault scenario are simulated rather than sweeping through the entire line with various fault resistance values, this search method is called as dynamic search (DS) technique in this paper.

### 3.2.4 Demonstrating stage 2—The fault location and fault resistance estimation process

Table 1 shows the outcome of each iteration of the fault location process described in Stage 2 for the example fault scenario. The kNN search was implemented using the \texttt{knnsearch} function of the statistics and machine learning toolbox in MATLAB. In
Table 1, the first simulation set (iteration 0) consists of default initial simulation scenarios discussed in Section 3.2.1. The first iteration of the kNN search identifies the four closest simulation scenarios or neighbours (as k value was chosen as four) to the actual fault scenario from the above set of simulated fault scenarios based on the fault currents. The fault distance (FL), fault resistance (FR), and the Euclidean distance (D) (between actual fault currents and simulated fault currents) of the four closest fault simulation scenarios from every iteration of kNN are presented in the second column of the table. From this set, the smallest and the largest fault location and fault resistance values are extracted. At the end of the first iteration, the smallest fault location value ($L_{\text{min}}^1$) was 0.33 pu and the largest fault location value ($L_{\text{max}}^1$) was 0.99 pu. Similarly, the smallest fault resistance value ($R_{\text{min}}^1$) was 0 Ω and the largest fault resistance value ($R_{\text{max}}^1$) was 5 Ω. For the next iterations, fault scenario simulations explore fault location and fault resistance values within the ranges defined by these minimum and maximum values. The fault location and fault resistance step sizes ($L_{\text{step}}^1$ and $R_{\text{step}}^1$, respectively) are calculated using (2) and (3). The next set of fault simulation scenarios are determined from these values as shown in (9) and (10) in Section 3.2.1.

Observe that fault location and fault resistance values simulated and analysed are narrowed down in the second iteration. The $L_{\text{step}}^2$ is between 0.165 pu and 0.99 pu compared to 0.01 pu and 0.99 pu in the $L_{\text{step}}^1$. Similarly, the $R_{\text{step}}^2$ is between 0 Ω and 6.25 Ω compared to 0 Ω and 50 Ω in the $R_{\text{step}}^1$. Hence, the simulation scenarios preformed in successive iterations are dynamically determined based on the kNN search such that they converge to the actual fault location rather than simulating a fixed set of fault scenarios along the entire line length.

This above process is repeated until the simulation scenarios are close to the actual fault location, that is, either D of any simulation scenario is less than 50 or the maximum number of iterations ($N$) of 5 is reached as discussed in Section 3.2.2. It can be observed that with every iteration of the kNN, D decreases showing that the simulation scenarios are converging to the actual fault scenario. At the end of the 4th iteration, D of one of the simulation scenarios from the kNN search was less than 50. Then (12) and (13) were applied to estimate the fault location and fault resistance values, respectively.

Table 2 presents the observed fault currents in each phase and the estimated fault location and fault resistance using the proposed method. The error in fault location and fault resistance are calculated using (14) and (15), respectively.

Fault Location Error (pu) = $|FL_{\text{Actual}} - FL_{\text{Estimated}}|$.  \( (14) \)

Fault Resistance Error (Ω) = $|FR_{\text{Actual}} - FR_{\text{Estimated}}|$.  \( (15) \)

The proposed approach was able to accurately identify the fault location and fault resistance as shown in the Table 2.

### 4 | IMPLEMENTING THE APPROACH TO ESTIMATE FAULT LOCATION ON NEIGHBOURING LINES

This section will expand the application of the approach beyond faults on the monitored line. The proposed method can independently identify the fault location accurately using only measurements from one terminal of a line when the measured fault current is unique for a given fault location. This section first shows that the fault current measured in a one-terminal line is unique for a given source voltage for every fault location and fault resistance value. Then, it explains the challenges in obtaining the accurate location of a fault in a two-terminal line and a fault on a neighbouring line using data from only one relay location. Finally, the additional data requirements and incorporation of this available supplemental data into the approach proposed in Section 3 to obtain fault location are discussed.
TABLE 2 Estimating fault location and fault resistance using currents observed by relay near Bus 153 monitoring Line 4 for a fault on Line 4

| Data Type        | Fault Current Input (A) | Fault Type |
|------------------|-------------------------|------------|
| Phasor           | $I_A$, $I_B$, $I_C$     | AG         |
| (Real, Imag.)    | $2562.01\angle-70.4^\circ$, $17.17\angle-69.2^\circ$, $17.17\angle-69.2^\circ$ |           |
| Actual FL on Line 4 (pu) | Estimated FL by DS on Line 4 (pu) | FL Error by DS (pu) | Actual FR by DS (Ω) | Estimated FR by DS (Ω) | FR Error by DS (Ω) |
| 0.528            | 0.530                   | 0.002      | 1.61         | 1.55         | 0.06        |

FIGURE 7 Simplified circuit representation of a one-terminal line

4.1 Analysing fault current in one-terminal transmission line

This subsection demonstrates that when the Thevenin impedance to the fault point from the relay location is equal for different fault scenarios, the uniqueness in the relation between fault current and fault location is lost. The notations used for analysis equations and figures in this paper are as follows. The variable quantities are denoted using capital letters, that is, current, voltage, and impedance are represented by $V$, $I$, and $Z$, respectively. The subscript refers to the component being measured, such as Bus 1 and Line y. The superscript corresponds to the sequence component or other additional information. For example, $Z_{ij}^1$ refers to the positive-sequence impedance of Line y.

Consider the one-terminal transmission line shown in Figure 7. The reduced system has an equivalent source voltage of $V_S$, source impedance of $Z_S$, line impedance of $Z_y$, and a relay monitoring Line y near Bus 1. The system experiences a fault at F located at a distance $d$ pu from Bus 1 with a fault resistance $R_F$. Currents and voltages in the following equations are calculated according to the fault type as demonstrated in [12].

\[
V_F = dZ_{ij}^1 I_{1,F} + R_F I_F, \tag{16}
\]

\[
\frac{V_F}{I_{1,F}} = Z_{Relay} = dZ_{ij}^1 + R_F \frac{I_F}{I_{1,F}}, \tag{17}
\]

where $Z_{Relay}$ is the impedance seen by the relay. IBFL methods are based on (17). In (17), $V_F$ and $I_{1,F}$ are measured by the relay and there are two unknowns: $R_F$ and $d$.

Symmetrical components are commonly used for fault analysis because they decouple a complex three-phase network into three balanced sequence networks, enabling each sequence network to be analysed independently. The different sequence circuits are connected together based on the fault type as explained in [34].

The fault current is a function of total Thevenin impedance to the fault. For example, in a single line-to-ground fault,

\[
I_{F}^0 = I_{F}^1 = I_{F}^2 = \frac{V_{S}^{1-pre}}{Z_{Th}^0 + Z_{Th}^1 + Z_{Th}^2 + 3R_F}, \tag{18}
\]

where $Z_{Th}$ is the Thevenin impedance to the fault from the equivalent source (whose prefault voltage is $V_{S}^{1-pre}$) and comprises of the source impedance ($Z_S$) and the line impedance ($Z_y$).

\[
I_{F}^0 = I_{F}^1 = I_{F}^2 = \frac{V_{S}^{1-pre}}{Z_{S}^0 + Z_{S}^1 + Z_{S}^2 + d(Z_{Th}^0 + Z_{Th}^1 + Z_{Th}^2) + 3R_F}. \tag{19}
\]

Fault impedance is typically resistive. Changing the fault resistance affects only the real part of the denominator of (19), whereas varying the fault location ($d$) changes the real and imaginary parts of the denominator of (19). Hence, for a given prefault voltage, the fault current is unique for each fault location and fault resistance value. Conversely, the prefault voltage and fault current are sufficient to determine the fault location and fault resistance.

The above characteristic is visualised in Figure 8 using Line 6 from the demonstration circuit (shown in Figure 2) for various fault resistance values.

FIGURE 8 Plot of fault currents represented in rectangular coordinates for faults at various locations along Line 6 (of the system shown in Figure 2) for various fault resistance values.
example. Line 6 is a one-terminal line with only one source for fault current (from Bus 156). The figure presents an example plot of fault current (observed by relay at Bus 156 monitoring Line 6 in this case) in the faulted phase for a one-terminal line. Each curve is a plot of phase A fault currents observed on phase A simulated along the length of the line for a particular fault resistance value. The currents are plotted in rectangular coordinates. The location of fault on the line is shown using the colorbar and the fault resistance is mentioned adjacent to the curve. It can be observed that the curves in the figure are relatively parallel to each other and do not overlap, showing that the fault current measured by the relay is, indeed, unique for each fault location and fault resistance value.

Furthermore, when the fault currents and prefault voltages are known, the fault voltage measurements at the bus do not provide any additional information regarding the fault, as shown by (20)–(22). This is because the voltage at the bus during a fault can be calculated using the prefault bus voltage, fault current, and system impedance parameters. For example, the voltage at Bus 1 measured by relay R in Figure 7 can be calculated using the following equations [35].

\[
V_1^0 = -Z_s^0 I_{1,F}, \quad (20)
\]
\[
V_1^1 = V_{s1}^{1,pre} - Z_s^1 I_{1,F}, \quad (21)
\]
\[
V_1^2 = -Z_s^2 I_{1,F}^2. \quad (22)
\]

4.2 Challenges in identifying the location of a fault on a two-terminal line or a fault on a neighbouring line

Identifying the location of a fault on neighbouring lines is very challenging using IBFL methods because new sets of equations have to be derived for every different network configuration. In addition, there are more unknowns than available equations, thus requiring assumptions to discount the effect of some variables (resulting in more sources of errors) or seek additional measurements. These are demonstrated in this subsection using fault scenarios in three different circuits.

Consider a two-terminal line shown in Figure 9 as Case 1. Though (16) and (17) remain the same, an additional unknown apart from \( R_F \) and \( d \) is introduced into the equation. \( I_F \) is now an unknown because of the remote source contribution \( (I_{2,F}) \) as shown in (23). As a result, there are more unknowns than available equations. Hence, different IBFL methods make different assumptions or collect additional data to overcome this challenge to estimate the fault location and fault resistance.

\[
I_F = I_{1,F} + I_{2,F}. \quad (23)
\]

The fault current \( I_{2,F} \) can be calculated when the remote source \( (V_{s2}) \) parameters are known. Hence, when MBFL approach is used with an accurate system model, the remote source fault current contribution is a known parameter and the fault location can be estimated precisely. Further analysis of the remote source and their effect on the fault location estimate are discussed in Section 5.

Now, consider Case 2 where a reduced circuit contains two lines connected to the same terminal of a two-terminal line as shown in Figure 10. As discussed in Section 1, there can be a variety of reasons such as CT saturation that result in erroneous measurements from relay R2. As a result, the measurements from relay R2 cannot be used to identify the fault location accurately. The through fault current flowing in Line y monitored by relay R1 will be used to identify the location of the fault on neighbouring line (Line \( s_x \)). Identifying the location of a fault present after the tap point on a three-terminal line is similar to this circuit.

The following equations derive the impedance observed by relay R1 during a fault to illustrate the loss of uniqueness in the relation between through fault current and fault location.

\[
V_2 = dZ_{s1} I_{2,F} + R_F I_F, \quad (24)
\]
\[
I_{2,F} = I_{4,2} + I_{1,2}, \quad (25)
\]
\[
V_2 = V_1 + Z_s^1 I_{1,2}. \quad (26)
\]

Substituting (25) and (26) in (24):

\[
V_1 + Z_s^1 I_{1,2} = dZ_{s1}(I_{4,2} + I_{1,2}) + R_F I_F. \quad (27)
\]

Finally, the impedance seen by relay R1 can be written as:

\[
\frac{V_1}{I_{1,2}} = dZ_{s1} \left( 1 + \frac{I_{4,2}}{I_{1,2}} \right) + R_F \frac{I_F}{I_{1,2}} - Z_s^1. \quad (28)
\]
FIGURE 11 Case 3: Simplified circuit representation to demonstrate the identification of the location of a fault located two buses away

Comparing (28) to that obtained for a two-terminal line (17), an additional unknown variable, $I_{4-2}$, is introduced into the equation. Furthermore, a very similar set of equations can be derived for identifying the location of a fault on Line $x_2$ using measurements from relay $R_1$ as shown in (29).

$$\frac{V_1}{I_{1-2}} = dZ_{x_3}^1 \left(1 + \frac{I_{3-2}}{I_{1-2}}\right) + R_F \frac{I_F}{I_{1-2}} - Z_{x_1}^1. \quad (29)$$

As a result, fault current observed at relay $R_1$ for a fault on Line $x_2$ can be similar to the fault current observed for a fault on Line $x_1$. Hence, it is challenging to identify the location of a fault on Line $x_1$ using measurements from $R_1$ instead of $R_2$. Additional information (discussed in Section 4.3) is required to identify the faulted line and location of the fault on a neighbouring line independently.

The circuit shown in Figure 11 (Case 3) presents a more complicated fault location scenario where fault current measurements from relay $R_1$ are used to identify the location of a fault on a line two buses away (Line $x_3$). Equation (30) presents the impedance observed by the relay $R_1$ for a fault on Line $x_3$ derived in the same way as (28).

$$\frac{V_1}{I_{1-2}} = dZ_{x_3}^1 \left(1 + \frac{I_{6-4}}{I_{1-2}} + \frac{I_{3-2}}{I_{1-2}}\right) + R_F \frac{I_F}{I_{1-2}} - Z_{x_1}^1 - Z_{x_4}^1 \left(1 + \frac{I_{3-2}}{I_{1-2}}\right). \quad (30)$$

Equation (30) has two more new unknown variables ($I_{6-4}$ and $I_{3-2}$) in comparison to (17) or one additional unknown variable compared to (28). The presence of additional fault current contributors between the relay and the fault introduces further unknowns in the impedance equation. Besides, faults at different locations on the REC can cause the same through fault current to be measured at relay $R_1$ location similar to the previous case.

4.3 Additional data requirements for complex fault location scenarios

Section 4.2 showed that the fault location of a fault might not be unique for a set of fault currents observed by a relay. When there are only two lines connected at a bus, then by Kirchhoff’s current law (KCL), no new unknown variable is introduced in the impedance seen by the relay equation. In each of the cases in Section 4.2, one extra unknown is introduced in the relay impedance equation for every extra line or source greater than two connected to a bus. In other words, for every extra connection greater than two to a bus, an extra variable is unknown.

In Case 2, to uniquely identify the fault location using the proposed method, some additional information from Line $x_1$ or Line $x_2$ is required to identify the fault location. A current or voltage measurement from either of the lines (Line $x_1$ or Line $x_2$) is sufficient to determine the faulted line. Furthermore, even a simple indication from a device like fault circuit indicator (FCI) is sufficient instead of current or voltage measurement to determine the faulted line. FCIs are relatively inexpensive devices that provide an indication of the fault path [36]. There are other methods to determine the faulted line. For example, it can be identified using the circuit breaker status communicated to the relay. If the faulted neighbouring line is known beforehand, then the dynamic search can be executed directly on the known faulted line to obtain the fault location. In situations where neither additional fault current or voltage measurement or FCI device is available, the approach can estimate the possible fault location in each neighbouring line. The electric utility protection engineer can use his local knowledge about the fault such as customer reports or phone calls about a power outage or relay trip reports to pick out the fault location result from the appropriate line.

To independently identify the fault location without requiring any external assistance such as from a protection engineer, for every $n$ connections to a bus ($n > 2$), a minimum of $n - 2$
unique measurements are required. Each of the \( n - 2 \) measurements must provide unique information, that is, each additional measurement must be from different lines apart from the monitored line.

The kNN search described in Section 3.2.2 is modified such that the feature space includes additional current or voltage measurements. Hence, the metric space \( q \) is considered as the local source and the effect of the remote source voltage on the measured local current in cases where the system model is not updated to reflect the exact remote source voltages. For example, a decrease in the local fault current as a result of an increase in remote source voltage that was not reflected in the circuit model can result in overestimating the fault location, that is, estimating the fault location to be farther than the actual fault location.

Consider a line-to-line (LL) fault involving phases B and C on Line 6 in the circuit shown in Figure 2. Assume accurate measurements from relays monitoring Line 6 are unavailable; hence, through fault current measurements from a relay monitoring Line 5 near Bus 153 are used to identify the fault location. Table 3 shows possible fault locations on Line 6 using through fault current seen by relay monitoring Line 5 near Bus 153 in rectangular coordinates. The lower half of Table 3 shows the fault location identified by the approach using the additional available data. It is able to accurately identify the faulted line and the fault location with the help of the additional current measurement. Similarly, if an FCI device mounted on Line 1 indicates no fault present on this line, the approach can be programmed to look for possible fault locations only on Line 6.

### 5 SENSITIVITY ANALYSIS ON REMOTE VOLTAGE SOURCE

Variations in the remote source voltage and impedance affect the current contribution from the local end, and this, in turn, affects the fault location estimate. This analysis is vital to understand the impact of the variation of remote source voltage on the measured local current in cases where the system model is not updated to reflect the exact remote source voltages. For example, a decrease in the local fault current as a result of an increase in remote source voltage that was not reflected in the circuit model can result in overestimating the fault location, that is, estimating the fault location to be farther than the actual fault location.

Consider a generic two-terminal line shown in Figure 9. \( V_{S1} \) is considered as the local source and the effect of the remote source voltage \( (V_{S2}) \) on the local measured current \( I_{F1} \) will be analysed. The local measured current in the faulted phase is expressed in terms of symmetrical components in (31).

\[
I_{F1} = I_{F1}^1 + I_{F1}^0 + I_{F1}^2.
\]  

A system having balanced voltage sources results in sources only in the positive-sequence network [37]. The various sequence circuits are connected with each other at the fault point. Hence, the voltage at fault point is a key parameter that affects the local current measured.

Partially differentiating (31) with respect to the positive-sequence voltage at the fault point \( (V_F^1) \):

\[
\frac{\partial I_{F1}^1}{\partial V_F^1} = \frac{\partial I_{F1}^0}{\partial V_F^1} + \frac{\partial I_{F1}^2}{\partial V_F^1},
\]

The circuit impedances in the negative- and zero-sequence circuits are fixed values and are connected either in series or parallel to the positive-sequence circuit depending on the fault type. As the negative- and zero-sequence circuits do not have voltage sources, \( \frac{\partial I_{F1}^0}{\partial V_F^1} \) and \( \frac{\partial I_{F1}^2}{\partial V_F^1} \) in (32) are constants. Hence, we focus the sensitivity analysis on the positive-sequence circuit.

### TABLE 3 Estimating the fault location and fault resistance of a fault on Line 6 using through fault current seen by relay monitoring Line 5 near Bus 153

| Fault Type | Actual FL on Line 6 (pu) | Actual FR (Ω) | Possible FL and FR Estimated by DS |
|------------|--------------------------|---------------|-----------------------------------|
| BC         | 0.131                    | 3.11          | 0.132 pu on Line 6 from Bus 156 with FR of 3.18 Ω |
|            |                          |               | 0.428 pu on Line 1 from Bus 156 with FR of 2.43 Ω |

### Additional Current Input from Relay Monitoring Line 1 Near Bus 40 (A)

| Fault Type | Actual FL on Line 6 (pu) | Actual FR (Ω) | Possible FL and FR Estimated by DS |
|------------|--------------------------|---------------|-----------------------------------|
| BC         | 0.132 pu on Line 6 from Bus 156 with FR of 3.18 Ω |
|            |                          |               | 0.428 pu on Line 1 from Bus 156 with FR of 2.43 Ω |

\[
I_F = I_{F1} + I_{F1}^0 + I_{F1}^2,
\]

\[
\frac{\partial I_{F1}^1}{\partial V_F^1} = \frac{\partial I_{F1}^0}{\partial V_F^1} + \frac{\partial I_{F1}^2}{\partial V_F^1},
\]

\[
I_{F1} = I_{F1}^1 + I_{F1}^0 + I_{F1}^2.
\]
FIGURE 12 Positive-sequence circuit of the two-terminal line shown in Figure 9

Analysing the positive-sequence circuit of the two-terminal line shown in Figure 12:

\[ V_{F1}^{\prime} - V_{F2}^{\prime} = I_{S1}^{\prime} (Z_{S1}^{\prime} + dZ_{F}^{\prime}), \quad (33) \]

\[ V_{F2}^{\prime} - V_{F1}^{\prime} = I_{S2}^{\prime} (Z_{S2}^{\prime} + (1-d)Z_{F}^{\prime}). \quad (34) \]

In the scenario where the remote voltage source \((V_{F2}^{\prime})\) is varied, an increase in \(V_{F1}^{\prime}\) causes an increase in \(I_{S1}^{\prime}\) and \(V_{F1}^{\prime}\). Partially differentiating (33) with respect to \(V_{F1}^{\prime}\):

\[ \frac{\partial I_{S1}^{\prime}}{\partial V_{F1}^{\prime}} = \frac{-1}{Z_{S1}^{\prime} + dZ_{F}^{\prime}}. \quad (35) \]

From (35), an increase in \(V_{F1}^{\prime}\) causes a decrease in \(I_{S1}^{\prime}\). Furthermore, when the system has a larger local source impedance \((Z_{S1}^{\prime})\) value (weaker local source), the impact of variation of \(V_{F1}^{\prime}\) on \(I_{S1}^{\prime}\) is lesser, that is, the same \(V_{F1}^{\prime}\) variation results in a smaller change in \(I_{S1}^{\prime}\).

To study the impact of change in \(V_{F1}^{\prime}\) on \(V_{F1}^{\prime}\),

\[ \frac{\partial V_{F1}^{\prime}}{\partial V_{F1}^{\prime}} = 1 - \tau, \quad (36) \]

where \(\tau = \frac{\partial I_{S2}^{\prime}}{\partial V_{F2}^{\prime}} \times (Z_{S2}^{\prime} + (1-d)Z_{F}^{\prime}). \quad (37)\]

By Ohm’s law, \(\partial I_{S2}^{\prime}/\partial V_{F2}^{\prime}\) is inversely proportional to the net total circuit impedance or proportional to the net total circuit admittance viewed from \(V_{F2}^{\prime}\). Though the net total circuit impedance depends on how the sequence networks are connected, any increase in the total circuit impedance leads to a decrease in \(\partial I_{S2}^{\prime}/\partial V_{F2}^{\prime}\). As a result, in fault scenarios with higher fault resistance values, \(\tau\) becomes smaller resulting in \(\partial V_{F1}^{\prime}/\partial V_{F2}^{\prime}\) to increase. Hence, \(I_{S1}^{\prime}\) becomes more sensitive to variations in \(V_{F2}^{\prime}\) in fault scenarios with higher fault resistance.

An increase in remote source impedance \((Z_{S2}^{\prime})\), that is, a weaker remote source, results in \(\tau\) closer to one as \(Z_{S2}^{\prime}\) forms a larger portion of the net total circuit impedance as viewed from \(V_{F2}^{\prime}\). Consequently, \(\partial V_{F1}^{\prime}/\partial V_{F2}^{\prime}\) becomes smaller, implying that \(V_{F2}^{\prime}\) varies less with changes in \(V_{F2}^{\prime}\). Hence, \(I_{S1}^{\prime}\) becomes less sensitive to variations in \(V_{F2}^{\prime}\). This is expected because a weaker remote source is expected to have a lesser impact on the measured local current.

When a fault is farther down the line from the relay monitoring the line, the fault distance in per-unit \((d)\) is larger. In (36), a decrease in \(1 - d\) value results in a decrease in \(\tau\), this causing \(\partial V_{F1}^{\prime}/\partial V_{F2}^{\prime}\) to increase. This implies that variation on \(V_{F1}^{\prime}\) has a larger impact on \(V_{F2}^{\prime}\) as the fault moves closer to the remote end of the line. However, according to (35), \(\partial I_{S1}^{\prime}/\partial V_{F1}^{\prime}\) decreases with an increase in \(d\). This means that \(I_{S1}^{\prime}\) is less sensitive to variations in \(V_{F1}^{\prime}\) as expected because the fault is much farther away from the local source. Therefore, the impact of variation of \(d\) is influenced by both (35) and (36) depending on the fault scenario.

The impact of each circuit parameter on \(I_{S1}^{\prime}\) discussed above has been summarised in Table 4. The first line indicates the parameter varied keeping the rest of the system and fault conditions constant. The following lines present the gist of the analysis. The last line presents whether the variation of \(I_{S1}^{\prime}\) increases or decreases with the change in the parameter indicated on the first line. Similar analysis can be performed for the negative- and zero-sequence circuits as well.

For all shunt fault scenarios, change in \(V_{F2}^{\prime}\) affects \(I_{S1}^{\prime}\). However, a peculiar fault scenario is a bolted three-phase to ground fault (LLLG fault). A LLLG fault involves only the positive-sequence circuit and a bolted fault results in \(V_{F1}^{\prime}\) being fixed at zero. In this case, any change in \(V_{F2}^{\prime}\) does not affect \(V_{F1}^{\prime}\) and in turn \(I_{S1}^{\prime}\) according to (33). Hence, any change in the remote source voltage \((V_{F2}^{\prime})\) does not affect the fault location estimate obtained using local source current measurement \((I_{S1}^{\prime})\).

The above characteristic behaviour of the remote voltage source on the local measured fault current can be extended to three-terminal lines as well as for cases where the fault is located on a neighbouring line. The effect of remote voltage sources on the fault current observed at the local end diminishes because of multiple impedances present between both the local source and fault, and remote source and local measurement. These fault scenarios are similar to having additional impedances in the local source branch as well as the remote source branch in Figure 12. As seen in Table 4, an increase in \(Z_{S1}^{\prime}\) or \(Z_{S2}^{\prime}\) reduces the influence of \(V_{F2}^{\prime}\) on \(I_{S1}^{\prime}\).

$\text{5.1 Tackling variations in remote voltage source}$

An essential requirement for the proposed approach to work successfully is having an accurate system model. This includes the equivalent source voltages in the RECs. The currents and voltages in the system during the fault depend on the equivalent source voltages as seen earlier in this section. Each of the new
TABLE 4  Impact of circuit parameters on $I_{11}^1$

| Parameter: ↑ $Z_{11}^1$ | Parameter: ↑ $Z_{12}^2$ | Parameter: ↑ $R$ | Parameter: ↑ $d$
|----------------------|----------------------|----------------------|----------------------|
| ⇒ Local source is weaker | ⇒ Remote source is weaker | ⇒ Fault resistance is higher | ⇒ Fault is farther down the line from the relay |
| $\frac{\partial I_{11}^1}{\partial V_{1}^{\text{pre}}}$ in (35) ↓ | $\frac{\partial I_{12}^1}{\partial V_{2}^{\text{pre}}}$ in (36) ↓ | $\frac{\partial I_{11}^1}{\partial V_{1}^{\text{pre}}}$ in (36) ↓ | $\frac{\partial I_{12}^1}{\partial V_{2}^{\text{pre}}}$ in (36) ↓ |
| ⇒ Less impact on $I_{11}^1$ | ⇒ $\text{in (36)}$ is closer to 1 | ⇒ $\text{in (36)}$ is closer to 1 | ⇒ $\text{in (36)}$ is closer to 1 |
| $\frac{\partial I_{11}^1}{\partial V_{1}^{\text{pre}}}$ in (36) ↓ | $\frac{\partial I_{12}^1}{\partial V_{2}^{\text{pre}}}$ in (36) ↓ | $\frac{\partial I_{11}^1}{\partial V_{1}^{\text{pre}}}$ in (36) ↓ | $\frac{\partial I_{12}^1}{\partial V_{2}^{\text{pre}}}$ in (36) ↓ |
| ⇒ Less impact on $I_{11}^1$ | ⇒ $\text{in (36)}$ is closer to 1 | ⇒ $\text{in (36)}$ is closer to 1 | ⇒ $\text{in (36)}$ is closer to 1 |
| ⇒ More impact on $I_{11}^1$ | ⇒ $\text{in (36)}$ is closer to 1 | ⇒ $\text{in (36)}$ is closer to 1 | ⇒ $\text{in (36)}$ is closer to 1 |
| ⇒ Impact on $I_{11}^1$ depends on both (35) and (36) | ⇒ Impact on $I_{11}^1$ depends on both (35) and (36) | ⇒ Impact on $I_{11}^1$ depends on both (35) and (36) | ⇒ Impact on $I_{11}^1$ depends on both (35) and (36) |

unknowns introduced by the presence of additional fault current contributors between the relay and the fault in the complex circuits presented in Section 4.2 will no longer be unknowns in the proposed approach as they are directly related to the prefault voltage of the remote buses.

A crucial yet reasonable assumption made in the circuit models is that equivalent source voltages are taken as 1 pu if their exact value is unknown. This is a rational assumption because transmission line voltages are strictly required to operate within their ratings as stipulated in national electric grid codes [38]. Distribution systems too are required to operate within specific voltage limits (such as between 0.95 pu to 1.05 pu [39]). Hence, the voltages typically do not deviate more than 5–10% of their rated voltage levels.

Fault recording devices are usually set up to record a few cycles of prefault data [40]. In a two-terminal line, the prefault voltage and prefault current information can be used to calculate remote source voltage according to (38).

$$V_{12}^{\text{pre}} = V_{1}^{\text{pre}} + (Z_{1}^{n} + Z_{2}^{n})_{1-2}^{n}, \tag{38}$$

where superscript $n$ refers to the $n^{th}$ symmetrical component. The system model can be updated after the data preprocessing step in Figure 1 to reflect the latest remote source voltage. As source impedance and line impedance parameters of the two-terminal line are known, the method can exactly identify the fault location without making any assumptions.

An approach to circumvent assumptions regarding prefault voltages is modifying the implementation of the proposed framework. The electric utility engineer can apply the approach for a range of remote source voltage values and obtain the span of possible fault locations. For example, in the two-terminal line case shown in Figure 9, the electric utility engineer can run the approach for $V_{12}^{\text{pre}} = \{0.98, 0.99, 1.0, 1.01, 1.02\}$ pu and obtain the five possible fault location values for each case. This can provide further information to the maintenance crew to swiftly locate the fault.

With the rapid increase in communication capability in recent times, it is no longer a difficult task to know the system voltage at various locations at regular intervals. An ideal implementation of this approach would involve regularly updating the system model to reflect real-time conditions using state-estimation techniques on frequently measured system information [20].

6 | COMPARISON WITH TRADITIONAL FAULT LOCATION METHODS

IBFL methods have been the benchmark for fault location standards and used in commercial relays for decades [3–5]. Hence, this paper compares the fault location estimates from the proposed approach with commonly used IBFL approaches such as Simple Reactance, Takagi, and Modified Takagi algorithms explained in [12] for complex network configurations.

6.1 | Complex network configurations

The first complex network configuration presented is lines with mutual coupling between them. Lines 2 and 3 in the test circuit shown in Figure 2 are mutually coupled. Consider a single-line-to-ground (SLG) fault on phase C at a distance of 0.35 pu from Bus 153 on Line 3 with a fault resistance of 5.8 $\Omega$. Table 5 presents the fault currents observed from a relay monitoring Line 3 at Bus 153 and the fault location estimates obtained using the proposed approach (dynamic search) and various IBFL methods.

The presence of mutual coupling between the lines affects the zero-sequence fault circuit [41]. As a result, the fault location estimates obtained using Simple Reactance, Takagi, or Modified Takagi methods are less accurate as they do not account this factor into their calculation. Hence, special IBFL algorithms such as those presented in [9] derived exclusively for mutually coupled lines have to be used to get accurate fault location results. The estimated fault location using the proposed method is very close to the actual fault location. It also identified the fault resistance value accurately, else additional impedance-based approaches are required to estimate the fault resistance.

The next complex network configuration is a three-terminal line. The lines between Bus 153, 40, and 2876 in the test circuit shown in Figure 2 can be visualised as a three-terminal line with Bus 156 as the tap point. In this scenario, measurements from relay monitoring one end of the line are used to identify the location of a fault after the tap point from the relay’s perspective. This scenario is similar to identifying the location of a fault on a neighbouring line. The fault is a line-to-line (LL) fault involving phases B and C at a location of 0.485 pu on Line 6 from Bus 156 with a fault resistance of 1.1 $\Omega$. Table 6 presents
the fault currents measured from relay monitoring Line 5 at Bus 153 and the fault location estimates obtained using the proposed approach and various IBFL methods.

The proposed approach was able to accurately identify the fault location, whereas the IBFL methods estimated the fault location to be beyond the line length or out of the line. This is because of the current contribution from the remote terminal at the tap point, which affects the fault current measured by the local relay and hence the fault location. Similarly, in this scenario as well, separate IBFL algorithms developed for three-terminal lines such as those presented in [6, 7, 9] have used to get accurate fault location results.

6.2 Application to field data

The fault event analysed is a line-to-line (LL) fault involving phases A and B on a sub-transmission line caused by lightning and failed arrester. The line connects Station A to Station B and is 9.1 miles long. It is protected with a recloser. Figure 13 shows the current and voltage waveforms recorded at Station A during the recloser action while the fault was still present in the system. The fault was at a distance of 5.33 miles (0.586 pu) from Station A.

The fundamental frequency fault current phasors were extracted using fast Fourier transform (FFT) after the DC offset had decayed. A one cycle window starting at 5.44 cycles was used for this purpose. The fault current phasors and the estimated fault location using the proposed approach as well as the IBFL methods are presented in Table 7. The fault location estimate using the proposed approach (Dynamic Search) is highly accurate and closer to the actual fault location than the IBFL approaches.

The authors extensively tested the proposed approach for several fault scenarios including various fault types, fault locations, and fault resistance values. Selected few cases were presented in this section to demonstrate the benefits of the approach. The fault location scenarios shown in Section 6.1 illustrate the need for special IBFL algorithms for each complex network configuration such as three-terminal lines and mutually coupled lines. On the other hand, the proposed approach can handle these complex network configurations without any

FIGURE 13  Current and voltage waveforms recorded at station A
advantages over various fault location

6.3 Advantages over various fault location methods

Impedance-based and traveling-wave based fault location approaches are the two most common fault location approaches used in relays as discussed in Section 1. A valuable unique highlight of this method is identifying the fault location on neighboring lines using limited data. The benefits of the proposed model-based fault location approach over impedance-based fault location methods have already been shown in [30, 42] as well as in Sections 4 and 6.1.

The primary benefit of this approach lies in being able to use the system model available to the electric utility. Because of the sensitivity of the proposed MBFL approach to system parameters, it can take into account all the factors that affect the fault current and overcome some of the limitations of IBFL approaches. Hence, it does not require separate algorithms for each complex network configuration and the same method can be used without much modifications.

Traveling-wave fault location methods require waveforms to be recorded at very high sampling frequency as discussed in Section 1. Hence, they cannot be used on reports from existing relays in the system that do not record at such high sampling frequency. For example, the event report in the fault scenario presented in Section 6.2 was recorded at just 16 samples per cycle and traveling-wave fault location approach cannot be applied. The proposed approach does not impose any new hardware requirements (such as high sampling frequency device, or new installation of phasor measurement units (PMU) across the transmission network). Hence, it provides additional benefits with no significant extra cost.

There are some model-based fault location approaches in the literature such as the search and find (S&F) approach in [30] and machine learning-based approaches such as artificial neural networks (ANN) [43, 44] that use the system model but they have their issues and limitations as well. The S&F approach suffers from a loss of precision based on the number of cases simulated as it is not possible to simulate fault scenarios at all possible fault locations with all possible fault resistance values. In addition, the number of cases that can be simulated at a given time depends on the computational capability and available memory of the computing device.

A common issue while using ANNs is that they have to be trained again when there is a change in the system network configuration. Furthermore, the fault location estimate will be affected by local and remote voltage levels if they are not taken into consideration while developing the ANNs. A key limitation of using most ANN based techniques is the black-box nature of the approach that limits its credibility and reliability. Furthermore, it is relatively hard and complex, both conceptually as well as computationally, to train ANNs for common electrical engineers who are less proficient in advanced mathematics. These are the few key reasons why they are not yet commonly used by electric utilities for fault location.

On the other hand, the proposed approach is transparent, enabling electric utility engineers to understand the underlying calculations and the convergence of the simulated fault scenarios towards the actual fault location. Furthermore, any change in the system network configuration only requires the REC to be updated to reflect the modification. There are no hassles involved such as simulating a large number of scenarios again and training new ANNs. The dynamic search approach further reduces the setup and run time as it simulates much fewer fault scenarios than the S&F MBFL technique. The approach is successful even when the fault event lies outside the simulated set of fault scenarios. It also offers the flexibility to incorporate as much additional data as and when available, as discussed in Section 4.3, to increase the fault location accuracy as well as to function independently without any aid from an engineer. The proposed approach can also be extended to distribution systems similar to [42] using fault circuit indicators.

7 Conclusion

A novel fault location algorithm that overcomes several limitations of conventional fault location methods is presented in this paper. A dynamic search method is proposed to perform the search process of a model-based fault location (MBFL) approach more effectively and efficiently. A key highlight of the proposed approach is identifying the location and

| Data Type | Fault Current (A) | Estimated Fault Location From Station A |
|-----------|------------------|----------------------------------------|
| Phasor    | $1574.5 \angle -49.53^\circ$ | $1522.2 \angle 130.78^\circ$ | $1.1 \angle -145.714^\circ$ | AB |
| Per-unit (pu) | 0.586 | 0.584 | 0.579 | 0.563 |
| Miles (mi) | 5.33 | 5.31 | 5.27 | 5.13 |
resistance of a fault on a neighbouring line with limited measurements. The approach is universal and advantageous to implement in the presence of complex network configurations such as three-terminal lines and lines with mutual coupling between them. The approach derives its benefits by utilising information regarding the surrounding system by using the system circuit model. Furthermore, the practical applicability of the approach was demonstrated by implementing it on field data.

This paper also analyses a critical factor that affects the accuracy of the proposed method, the remote source voltage. The approach is flexible and allows it to incorporate additional measurements as and when available. This can be used to identify the faulted line in scenarios where the fault is on a neighbouring line and increase the accuracy of the fault location estimate.

**ORCID**

Sundararavanan Navelpakham Ananthan 🐘 https://orcid.org/0000-0002-5588-9199

Alvaro Furlani Bastos 🐘 https://orcid.org/0000-0001-8786-6197

Surya Santoso 🐘 https://orcid.org/0000-0002-8675-9610

**REFERENCES**

1. Ram, B., Vishwakarma, D.N.: Power System Protection and Switchgear. Tata McGraw Hill Publication Company Limited, New Delhi (2005)
2. Blackburn, J.L., Domin, T.J.: Protective Relaying Principles and Applications, 4th edn. CRC Press, Boca Raton (2015)
3. Fischer, N., et al.: Protective relay traveling wave fault location. In: 11th IET International Conference on Developments in Power Systems Protection (DPSP 2012), pp. 1–3. IET, Stevenage (2012)
4. SEL-411L Relay: Protection and Automation System - Instruction Manual. Pullman, Washington DC (2017)
5. SEL-221F/121F: Distance Relay - Instruction Manual. Pullman, Washington DC (1992)
6. IEEE guide for determining fault location on AC transmission and distribution lines. IEEE Std C37.114-2014, pp. 1–76. IEEE, Piscataway (2015)
7. Tziovouaras, D.A., et al.: New multi-ended fault location design for two- or three-terminal lines. In: 2001 Seventh International Conference on Developments in Power System Protection (IEE), pp. 395–398. IEE, London (2001)
8. Girgis, A.A., et al.: A new fault location technique for two- and three-terminal lines. IEEE Trans. Power Delivery 7(1), 98–107 (1992)
9. Saha, M.M., et al.: Fault Location on Power Networks, 1st edn. Springer-Verlag, London (2010)
10. Saha, M.M., et al.: A new accurate fault locating algorithm for series compensated lines. IEEE Trans. Power Delivery 14(3), 787–799 (1999)
11. Saha, M.M., et al.: Accurate location of faults on series-compensated lines with use of two-end unsynchronised measurements. In: 2008 IET 9th International Conference on Developments in Power System Protection (DPSP 2008), pp. 338–343. IET, London (2008)
12. Das, S., et al.: Impedance-based fault location in transmission networks: theory and application. IEEE Access 2, 537–557 (2014)
13. Dalecastagné, A.L., Zimath, S.L.: A study about the sources of error of impedance-based fault location methods. In: Proceedings of IEEE/PES Transmission and Distribution Conference and Exposition, pp. 1–6. IEEE, Piscataway (2008)
14. Schweitzer, E., et al.: Accurate single-end fault location and line-length estimation using traveling waves. In: 13th International Conference on Developments in Power System Protection. IET, Stevenage (2016)
15. Mars, S., et al.: Traveling wave fault location in protective relays: Design, testing, and results. Paper presented at 16th Annual Georgia Tech Fault and Disturbance Analysis Conference, Atlanta, 2013
16. Hangrave, A., et al.: Beyond the knee point: A practical guide to CT saturation. In: 2018 71st Annual Conference for Protective Relay Engineers (CPRE), pp. 1–23. IEEE, Piscataway (2018)
17. Costello, D.: Lessons learned through commissioning and analyzing data from transformer differental installations. IEEE Trans. Ind. Appl. 44(4), 979–1003 (2008)
18. Costello, D.: Lessons learned analyzing transmission faults. In: 2008 61st Annual Conference for Protective Relay Engineers, pp. 410–422. IEEE, Piscataway (2008)
19. Das, S., et al.: Relay performance verification using fault event records. Protection Control Mod. Power Sys. 3(1), 22 (2018)
20. Kezunovic, M.: Smart fault location for smart grids. IEEE Trans. Smart Grid 2(1), 11–22 (2011)
21. Chen, K., et al.: Fault detection, classification and location for transmission lines and distribution systems: A review on the methods. High Voltage 1(1), 25–33 (2016)
22. Xia, B., et al.: Estimation of fault resistance using fault record data. IEEE Trans. Power Delivery 30(1), 153–160 (2015)
23. Das, S., et al.: Estimating zero-sequence line impedance and fault resistance using relay data. IEEE Trans. Smart Grid 10(2), 1637–1645 (2019)
24. Bam, L., Jewell, W.: Review: Power system analysis software tools. In: IEEE Power Engineering Society General Meeting, vol. 1, pp. 139–144. IEEE, Piscataway (2005)
25. Hay, S., Ferguson, A.: A review of power system modelling platforms and capabilities. Paper 3, Institution of Engineering and Technology (IET), Stevenage (2015)
26. Amberg, A., et al.: Validating transmission line impedances using known event data. In: 2012 65th Annual Conference for Protective Relay Engineers, pp. 209–280. IEEE, Piscataway (2012)
27. Ritzmann, D., et al.: A method for accurate transmission line impedance parameter estimation. IEEE Trans. Instrumentation and Measurement 65(10), 2204–2213 (2016)
28. Das, S., et al.: Estimating zero-sequence impedance of three-terminal transmission line and Thevenin impedance using relay measurement data. Protection Control Mod. Power Sys. 3(1), 36 (2018)
29. Baran, M., et al.: A model based fault locating method for distribution systems. In: 2014 IEEE PES General Meeting | Conference Exposition, pp. 1–5. IEEE, Piscataway (2014)
30. Ananthan, S.N., Santoso, S.: Universal model-based fault location for improved system integrity. IET Gener. Transm. Distrib. 13(8), 1212–1219 (2019)
31. Ananthan, S.N., et al.: An automated technique for extracting phasors from protective relay’s event reports. Inventions 3(4), 81 (2018)
32. Min, K.W., Santoso, S.: DC offset removal algorithm for improving location estimates of momentary faults. IEEE Trans. Smart Grid 9(6), 5503–5511 (2018)
33. Ananthan, S.N., et al.: Wavelet based transmission line fault analysis: A literature survey. In: 2014 14th International Conference on Environment and Electrical Engineering, pp. 254–259. IEEE, Piscataway (2014)
34. Grainger, J.J., Stevenson, W.D.: Power System Analysis. McGraw-Hill, New York (1994)
35. Santoso, S.: Fundamentals of Electric Power Quality. CreateSpace, Scotts Valley (2009)
36. Ferrer, H.J.A., Schweitzer, E.O.: Modern Solutions for Protection, Control, and Monitoring of Electric Power Systems. Schweitzer Engineering Laboratories, Pullman (2010)
37. Chase, S., et al.: Analyzing faulted transmission lines: Phase components as an alternative to symmetrical components. In: 2018 71st Annual Conference for Protective Relay Engineers (CPRE), pp. 1–19. IEEE, Piscataway (2018)
38. Central Electricity Regulatory Commission and Others: Indian Electricity Insulation Code, 2010
39. American National Standard for Electric Power Systems and Equipment-Voltage Ratings (60 Hertz). National Electrical Manufacturers Association, Rosslyn (2016)
40. Perez, J.: A guide to digital fault recording event analysis. In: 2010 63rd Annual Conference for Protective Relay Engineers, pp. 1–17. IEEE, Piscataway (2010)
APPENDIX A

This section provides proof that the measured fault current of a fault anywhere on the line with any fault resistance value will lie within the region delimited by these boundary equations presented in Section 3.1. Consider an SLG fault on phase A of a two-terminal line shown in Figure 7. The current on the faulted phase ($I_{F}^{A}$) is

$$I_{F}^{A} = \frac{3 \times I_{1,pre}^{A}}{Z_0 + Z_1 + Z_2 + d(Z_0^2 + Z_1^2 + Z_2^2) + 3R_F}. \quad (A.1)$$

(A.1) is obtained by converting the sequence fault currents given in (19) to phase components. The only varying parameters in (A.1) for a given prefault voltage that affect the fault current are the location of fault on the monitored line ($d$) and fault resistance ($R_F$). Curves 1 and 2 capture the boundaries of $d$ by fixing parameter $R_F$ and varying $R_F$. Curves 1 and 2 are formed from fault scenarios simulated with various fault resistance values at 0.01 pu (beginning of the line) and 0.99 pu of the line length (end of the line). Similarly, curves 3 and 4 capture the boundaries of $R_F$ by fixing parameter $R_F$ and varying $d$. Curves 3 and 4 are formed from fault scenarios simulated at various locations along the line at 0 Ω and 50 Ω (the minimum and maximum considered fault resistance values). Hence, the measured fault current for any standard fault scenario on the monitored line will fall within the four boundary equations formed using faults scenarios at each of the extreme values (min and max) of ($d$) and ($R_F$). Similar equations can be derived for all other fault types as well.