Fault Location in Power Networks Using a Sparse Set of Digital Fault Recorders

Cesar Galvez, *Student Member, IEEE*, and Ali Abur, *Fellow, IEEE*

***Abstract*—This paper presents a robust fault location method that can be used for radial or meshed power systems. It addresses several challenges encountered in previous work such as signal attenuation by power transformers, measurement errors, the impact of Inverter Based Power Sources (IBPSs), differentiating faults from other events. This is accomplished via a KMeans clustering of DFRs and using a weighted directed tree model enabling accurate fault location. The proposed algorithm is tested on both transmission and distribution systems under different fault scenarios.**

***Index Terms*—Fault location, power system faults, electromagnetic transients, wavelet transform, traveling waves.**

I. Introduction

RECENT increase in the number of Distributed Energy Resources (DERs) connected to transmission and distribution systems via Inverter-Based Power Sources (IBPS) and the changing configuration of traditional radial systems to meshed networks introduce new challenges in detecting and locating fault. Fault location has been a topic of research for a long time and proposed methods so far can be grouped as those based on apparent impedance measurements [1], [2], traveling wave (TW) methods [3], and methods using artificial intelligence (AI) [4]. AI-based algorithms are mostly used for fault classification and faulted section identification and cannot determine the precise location of the fault [4]. Assuming that the faulted line is known, phasor measurements can be used to estimate the fault location as shown in [5], [6]. Use of sparse estimation approach for fault location was initially shown in [7] and later used in [8]–[11]. However, phasor based methods have limitations caused by unknown fault resistances [12], asymmetry of lines, current transformer capacitive effects, and the fault ride-through grid codes [13] that dictate responses of IBPSs under faults affecting the fault location accuracy. On the other hand, traveling waves (TW) based methods do not have such limitations [14]. Their accuracy depends on the sampling rates of Digital Fault Recorders (DFRs). Given the availability of 1MHz sampling rate fault recorders [15] and the integrated optical sensors for voltage and current measurements [16], [17], the TW-based methods developed for transmission systems using wide area synchronized voltage measurements [18], [19] can be readily used in distribution networks as well. The computational burden of these methods based on a two-step linear optimization problem is removed in [20] in particular for networks with ring type topology. However, it requires a fault recorder at each bus and does not address previously stated challenges brought up by DERs. In [21], a method based on time differences of recorded faults in hybrid networks is presented, but its precision decreases in the presence of different cable parameters. In [22], a minimum detector placement strategy and fault location based on two invariant graphs are proposed, but details of detector placement for large grids and impact of real disturbances in power grids are not provided. In [23], arrival time differences are minimized to find the fault location but the impact power transformers or topology changes on TW are not addressed.

In our earlier work [14] we have shown that faults could be detected and located in radial distribution networks in the presence of IBPS. This was accomplished by placing DFRs at terminal nodes of all laterals. Despite its high accuracy, the technique is limited to strictly radial distribution networks. More recently, this work is extended to locate faults in meshed distribution networks [24], [25]. However, attenuation of TW due to power transformers in meshed networks limits application of this approach. The method presented in this paper addresses this and several other shortcomings of the method in [24], [25]. The main contributions of the proposed algorithm are highlighted below.

1) For cases where the faulted bus has many incident lines, the method of [24] is improved via the use of virtual buses as will be explained in the sequel.
2) The proposed method remains robust against bad data and synchronization-time errors.
3) Instead of using all DFRs, small clusters are created via KMeans clustering. Only those DFRs closest to the fault are identified and used avoiding ToA errors due to the attenuation of TWs passing through power transformers.
4) An optimal DFR placement strategy is developed for implementation of the method for any type of network.

Implementation of the method will be facilitated by an optimal DFR placement (ODP) strategy based on the approach in [26].

Manuscript received 3 September 2021; revised 8 March 2022; accepted 16 April 2022. Date of publication 20 April 2022; date of current version 23 August 2022. This work was supported in part by the Engineering Research Center Program of the National Science Foundation and the Department of Energy through NSF under Award EEC-1041877, and in part by the Center for Ultra-Wide-Area Resilient Electric Energy Transmission Networks (CURENT) Industry Partnership Program. Paper no. TSG-01419-2021. *(Corresponding author: Ali Abur.)*

The authors are with the Electrical and Computer Engineering Department, Northeastern University, Boston, MA 02151 USA (e-mail: cgalvezn@ece.neu.edu; abur@ece.neu.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSG.2022.3168904.

Digital Object Identifier 10.1109/TSG.2022.3168904

TABLE I
PROPOSED METHOD [⋆] VS LITERATURE

	[21]	[18]	[20]	[23]	[22]	[⋆]
Large Grids FL	x	✓	x	x	x	✓
Grids include power transformers	x	x	x	x	x	✓
FL using a DFR subset	x	x	x	x	x	✓
FL using a sub-network	x	x	x	x	x	✓
Topology change Robust	x	x	x	x	x	✓
FL computation redundancy	✓	x	x	x	x	✓
Line-cables Hybrid Network FL	✓	x	x	x	x	✓
Radial Networks FL	✓	x	x	✓	✓	✓
Meshed Networks FL	x	✓	✓	✓	✓	✓
FL in both Transmission and Distribution systems	x	x	x	x	x	✓
Harmonics, capacitor and load switching Robust	x	x	x	x	x	✓
Time-synchronization errors and bad data Robust	✓	x	✓	x	x	✓
DFR Placement Strategy	✓	✓	x	x	✓	✓
DFR placement reduction in radial or meshed distribution system	x	x	x	x	x	✓

The optimally placed DFRs are synchronized via GPS to register the first time of arrivals (ToAs) of faulted-generated voltage signals using Discrete Wavelet Transform (DWT). The method can either use the entire set or a subset of DFRs to compute the fault distance. For large networks with power transformers, the algorithm clusters DFRs closest to the fault to avoid bad data and arrival time errors. Otherwise, the entire DFR set will be used to determine the fault distance. The subset is determined using the K-Means clustering, which clusters DFRs based on feature vector formed by arrival times and shortest traveling times taking as initial reference the DFR closest to the fault. Using a DFR subset will only use a small sub-network from the entire network to compute the fault distance, making the proposed algorithm robust against topology changes.

Having the ToAs of selected DFRs, the proposed algorithm proceeds in two stages. In the first stage, the entire network is modeled as a weighted undirected graph and using the ToAs of all or just clustered DFRs and placing virtual buses the faulted section is identified as the one yielding the minimum ℓ_2 norm. The identified virtual bus will belong to the faulted line section. The second stage models the shortest travel time routes from the fault to selected DFRs using a weighted directed tree network to determine the fault location in a redundant and robust fashion. Detailed formulation and implementation of these stages will be described below. A comparison of the key features of the proposed fault location (FL) algorithm and several relevant references is shown in Table I.

II. PROBLEM FORMULATION

A small 8-bus transmission system shown in Fig. 1 will be used to describe the proposed fault location method. The model contains wind turbines connected at bus 1, and there are 138KV and 345KV voltage networks interconnected via power transformers. The meshed system contains four breakers at the terminals of overhead line and power transformers between buses 2-6 and 5-8. The system is meshed when the breakers

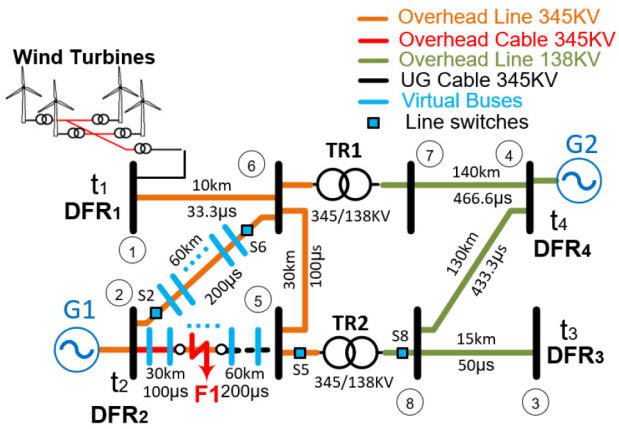

Fig. 1. 8-Bus Example System.

are closed, but becomes strictly radial when they are open. Four DFRs are placed in the network at buses 1, 2, 3 and 4. Consider the fault F1 occurring along the line section formed by 345KV overhead lines and cables as shown in Fig. 1. While the proposed method is illustrated for a meshed transmission network, it can also be used in distribution networks, as will be shown in the simulation section.

A. Placement of DFRs

1) DFR Placement in Radial Distribution Networks: The DFR placement in radial networks uses the procedure proposed in [14], where a DFR needs to be placed at the terminal bus of every lateral. Doing this, the shortest routes formed by DFR pair combinations will cover all network edges. Assuming that the breakers are open, the 8-bus system of Fig. 1 will be strictly radial. Thus, one DFR will be placed at the terminal buses 1, 2 and 3 to ensure that any fault in the system will be identifiable.

2) DFR Placement in Meshed Transmission Networks: In meshed transmission networks, DFRs are placed in two stages. First, one DFR will be placed at the terminal bus of each lateral. Next, the optimal DFR placement (ODP) method of [26] is used to install the rest of the DFRs to make any fault in the system identifiable. The ODP problem can be formulated as a linear integer programming problem:

$$\text{minimize} \quad \sum_{i=1}^{N} w_i \cdot x_i \tag{1}$$

$$s.t. \quad [A][X] \geq b \tag{2}$$

where,

$$A = [a_{ij}]_{N\times N}$$
$$X = [x_1,\ x_2,\ x_3, \ldots, x_N]^T_{N\times 1}$$
$$b = [1,\ 1,\ 1, \ldots, 1]^T_{N\times 1}$$
$$a_{ij} = \begin{cases} 1 \text{ if } i = j \\ 1 \text{ if } i \text{ and } j \text{ are connected} \\ 0 \text{ otherwise} \end{cases}.$$

X is a binary vector whose entries x_i will be 1 or 0 if a DFR is placed at bus i or not, and w_i is the installation cost of the

DFR at bus i. In the above problem, the corresponding binary variables will be initially set equal to 1 for those buses with already installed DFRs. Four switches in the 8-bus system of Fig. 1 will be closed to create a meshed network. Hence, in the first stage, two DFRs will be installed at lateral terminals 1, and 3. Therefore, $x_1 = 1$ and $x_3 = 1$ in vector X. Then, ODP method will be applied to place the remaining DFRs, which yields two additional DFRs at buses 2 and 4. Note that each installed DFR will only measure voltages. Unlike PMUs which measure both voltage and current phasors, in this approach voltage measurements made by DFRs will be sufficient to estimate the precise fault location.

3) DFR Placement in Meshed Distribution Networks: DFR placement in meshed distribution networks uses the same procedure as meshed transmission system. However, meshed distribution systems typically have many laterals connected to the main meshed network. Hence, total number of DFRs can be reduced by placing a DFR at each lateral line terminal as done in radial networks. As an illustration, consider that switches S2 and S6 in Fig. 1 are opened. This will yield a network with three lateral lines connected to a ring formed by buses 5, 6, 4, and 8. Here, installing DFRs at the terminals of each lateral will be sufficient to locate any fault in the system.

B. Initial Considerations

The transmission and distribution conductors used in this work are modeled using Frequency Dependent (FD) distributed line model [27]. Three-phase voltage signals captured by DFRs are transformed from phase to modal domain using Clarke Transformation [28]. Voltage signals in modal domain are decomposed into five wavelet transform coefficients (WTCs) using the Discrete Wavelet Transform (DWT) and the mother wavelet Daubechies2 (DB2) [29]. The WTCs at level 1 (D1) are then processed to determine the first arrival times due to their rich high-frequency content. Finally, to reduce the noise in WTCs, these signals are squared to obtain WTC^2 as also done in [3].

C. Selecting an Optimal Subset of DFRs

Fault location in large meshed grids with a limited number of DFRs is challenging due to the long line lengths as well as the ground connection of power transformer neutrals, which attenuate the fault-generated TWs. For instance, the TW generated by a fault at F1 in Fig. 1 will be correctly captured by DFRs 1 and 2 due to their proximity to F1. However, it will be attenuated by TR2 transformer before reaching DFRs 3 and 4. In addition, long line lengths further increase attenuation of the TWs captured by DFR4. Therefore, using ToAs recorded by all DFRs may yield incorrect fault distance estimates. Instead of using the ToAs recorded by all available DFRs, a subset of DFRs can be strategically selected. This is accomplished by using K-Means clustering on the 2D feature vector $x^{(i)}$ given by:

$$x^{(i)} = (t_i, \tau_{(R,j)}) \tag{3}$$

where, $x^{(i)}$ is the 2D feature vector for p DFRs, $i = 1, 2, \ldots, p$. t_i is the ToA recorded by DFR_i and $\tau_{(R,j)}$ is the shortest travel time, where R is the reference bus with the earliest ToA and j refers to other DFR buses. For a fault at F1 in 8-bus example, $R = 2$ being the reference bus, the shortest travel times are computed from bus 2 to buses 1, 3 and 4. Alternatively, magnitudes of squared wavelet transform coefficient (WTC^2) can also be added to $x^{(i)}$ to create a 3D feature vector as depicted in (4). The 3D feature vector $x^{(i)}$ will yield the same cluster as 2D, but it will illustrate the attenuation of TWs as they pass by a power transformer. Logarithm is used for WTC^2 to better visualize clusters.

$$x^{(i)} = \Big(t_i,\ \tau_{(R,j)},\ log_2(WTC^2 + 1)\Big) \tag{4}$$

The K-Means algorithm [30] is then used to determine suitable assignments of clusters $(C_1, \ldots, C_k)$. It first selects the cluster number k using the well-known elbow method [30]. Next, the representatives of each cluster $(z_1, \ldots, z_k)$ are randomly selected and then the process is repeated to find the clusters using:

$$Cost(C_1, \ldots, C_k) = \sum_{i=1}^{N} \min_{j=1,\ldots,K} \left\| x^{(i)} - z_j \right\| \tag{5}$$

Given $C_1, \ldots, C_k$ find the best representatives $z_1, .., z_k$ so that:

$$z_j = \text{argmin}_z \sum_{i \in C_j} \left\| x^{(i)} - z \right\|^2 \tag{6}$$

For the fault at F1 in Fig. 1, K-Means will yield the cluster $DFR_i, i = 1, 2$ and 3 due to their proximity to the fault. Using the three DFRs will yield an accurate fault location. K-Means is relevant for fault location in large networks containing thousands of buses since the proposed method uses only a small tree-like sub-network around clustered DFRs. Hence, any topological change outside of the sub-network will have no impact on algorithm performance.

D. Proposed Fault Location Method

The proposed method has two stages. First, the network is modeled as a weighted undirected graph denoted by $G = (\upsilon, e, \tau)$, where υ is the set of buses or nodes with size n, e is the set of edges or lines with size L in which $L \geq n - 1$, and τ is the set of weights with size L represented by the travel times τ_i for each edge i. The travel times τ_i (μs) are computed by ($\tau_i = l_i/v_i = l_i\sqrt{L_i.C_i}$), where l_i is the line length, and v_i is the wave speed for the edge i. Also, it can be computed using the conductor capacitance C_i and inductance L_i per unit length. Depending on the presence of power transformers in large grids, all or a subset of available DFRs will be selected. Then, the weighted undirected graph will be used to identify the bus closest to the fault based on the minimum ℓ_2-norm of ToAs captured by DFRs. This bus will be referred as the fault-closest bus in the sequel. The faulted section is then identified by placing virtual buses along each edge incident to the fault-closest bus and re-calculating the "fault-closest bus" among the virtual buses. This will yield the faulted edge containing the fault-closest virtual bus. The second stage determines the shortest routes from the fault to all DFRs or clustered DFRs using a weighted directed tree graph to estimate the ToA vectors at each terminal bus of faulted section. In addition, the

tree graph enables the use of all possible ToA pairs to compute the fault distance. Hence, it will increase the estimation redundancy and robustness against errors in ToA. Finally, using the two-terminal fault location method presented in [14], fault will be accurately located.

1) Identifying the Faulted Section Using Virtual Buses: Depending on the network type, the algorithm first places strategically DFRs based on the approach presented in Section II-A. In the 8-bus example system, the algorithm uses the optimal DFR placement (ODP) to install four DFRs, DFR_i, $i = 1, \ldots, 4$ in the network. The weighted undirected graph $G = (\upsilon, e, \tau)$ is then used to identify the fault-closest bus and faulted section. Considering the fault at F1 in Fig. 1, DFR2 will be the reference and ToAs registered by DFRs at buses 1, 2, 3, and 4 will be stored as:

$$ToA_1{}^T = \begin{bmatrix} t_1 & t_2 & t_3 & t_4 \end{bmatrix} \tag{7}$$

The relative ToA_2 vector with respect to $DFR2$ can be formed as:

$$ToA_2{}^T = \begin{bmatrix} t_1 - t_2 & t_2 - t_2 & t_3 - t_2 & t_4 - t_2 \end{bmatrix} \tag{8}$$

The following $p \times n$ matrix can be formed showing the shortest travel times between buses:

$$\tau_A = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ . \\ . \\ p \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & \cdots & n-1 & n \end{array} \\ \begin{bmatrix} 0 & \tau_{1.2} & \cdots & \tau_{1.(n-1)} & \tau_{1.n} \\ \tau_{2.1} & 0 & \cdots & \tau_{2.(n-1)} & \tau_{2.n} \\ \tau_{3.1} & \tau_{3.1} & \cdots & \tau_{3.(n-1)} & \tau_{3.n} \\ . & . & \cdots & . & . \\ . & . & \cdots & . & . \\ \tau_{p.1} & \tau_{p.2} & \cdots & \tau_{p.(n-1)} & \tau_{p.n} \end{bmatrix} \end{array} \tag{9}$$

where p is the number of buses with DFRs and n is the number of buses in the network. Now, define a $p \times p$ matrix D as:

$$D = \begin{array}{c} \\ 1 \\ k \\ 3 \\ . \\ . \\ p \end{array} \begin{array}{c} \begin{array}{ccccccc} 1 & k & 3 & \cdots & p-2 & p-1 & p \end{array} \\ \begin{bmatrix} 1 & -1 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & -1 & 1 & \cdots & 0 & 0 & 0 \\ . & . & . & \cdots & . & . & . \\ . & . & . & \cdots & . & . & . \\ 0 & -1 & 0 & \cdots & 0 & 0 & 1 \end{bmatrix} \end{array} \tag{10}$$

where k is the reference DFR bus and column k contains all "-1"s except for the k row which is 0. Taking the product of D and τ_A, a reduced relative ToA matrix τ_{red} can be obtained:

$$\underset{p\times n}{\tau_{\text{red}}} = \underset{p\times p}{D} \times \underset{p\times n}{\tau_A} \tag{11}$$

Subtracting ToA_2 of (8) from every column of τ_{red}, a travel time mismatch matrix $\Delta\tau$ is obtained as:

$$\underset{p\times n}{\Delta\tau} = \underset{p\times n}{\tau_{red}} - \underset{p\times 1}{[ToA_2]}.\underset{1\times n}{[1\ 1\ 1\ \cdots\ 1]} \tag{12}$$

The fault-closest bus j will then be given by the bus that minimizes the ℓ_2 norm given by:

$$\ell(j) = \min_j \left(\|\Delta\tau_j\|_2;\ j \in \{1, 2, \ldots, n\} \right) \tag{13}$$

For the fault at F1 in Fig. 1, $j = 2$ and $n = 8$. Once the fault-closest bus (bus 2) is identified, the faulted section can be identified by placing virtual buses along the incident edges connected to bus 2. In long line lengths, the number of virtual buses for each incident edge will be estimated by rounding their incident edge weights τ_{e_u} to their nearest integer, where τ_{e_u} are the travel times (μs) of incident edges $u = 1, 2, \ldots, w$. Therefore, the total number ε of virtual buses for the entire network can be estimated as follows:

$$\varepsilon = \tau_{e_1} + \tau_{e_2} + \cdots + \tau_{e_w} \tag{14}$$

In the case of short line lengths (e.g., smaller than 1 km), installing a small number of virtual buses (e.g., 5 or 10 fictitious buses for each incident edge) will always lead to the identification of faulted section. Each incident edge must have equally spaced fictitious buses. The τ_A matrix will then be modified to add these virtual buses as shown below:

$$\tau_A^* = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ . \\ . \\ p \end{array} \begin{array}{c} \begin{array}{ccccc|ccc} 1 & 2 & \cdots & n & n+1 & \cdots & n+\varepsilon \end{array} \\ \left[\begin{array}{ccccc|ccc} 0 & \tau_{1.2} & \cdots & \tau_{1.n} & \tau_{1.(n+1)} & \cdots & \tau_{1.(n+\varepsilon)} \\ \tau_{2.1} & 0 & \cdots & \tau_{2.n} & \tau_{2.(n+1)} & \cdots & \tau_{2.(n+\varepsilon)} \\ \tau_{3.1} & \tau_{3.1} & \cdots & \tau_{3.n} & \tau_{3.(n+1)} & \cdots & \tau_{3.(n+\varepsilon)} \\ . & . & \cdots & . & . & \cdots & . \\ . & . & \cdots & . & . & \cdots & . \\ \tau_{p.1} & \tau_{p.2} & \cdots & \tau_{p.n} & \tau_{p.(n+1)} & \cdots & \tau_{p.(n+\varepsilon)} \end{array}\right] \end{array} \tag{15}$$

Likewise, the τ_{red} and $\Delta\tau$ matrices are modified to include the fictitious buses as below:

$$\underset{p\times(n+\varepsilon)}{\tau^*_{\text{red}}} = \underset{p\times p}{D} \times \underset{p\times(n+\varepsilon)}{\tau_A^*} \tag{16}$$

$$\underset{p\times(n+\varepsilon)}{\Delta\tau^*} = \underset{p\times(n+\varepsilon)}{\tau^*_{red}} - \underset{p\times 1}{[ToA_2]}.\underset{1\times(n+\varepsilon)}{[1\ 1\ 1\ \cdots\ 1]} \tag{17}$$

The fault-closest virtual bus y will then be given by the bus that minimizes the ℓ_2 norm given by:

$$\ell(y) = \min_y \left(\|\Delta\tau_y^*\|_2;\ y \in \{1, 2, \ldots, (n+\varepsilon)\} \right) \tag{18}$$

where y represents the fault-closest virtual bus, which will reveal the faulted line section. In the example system of Fig. 1, the incident edges (2-5, and 2-6) to bus 2 are split into equally-spaced line sections by placing 300 and 200 virtual buses to their incident edges and the faulted line is thus identified as the edge 2-5. Knowing the fault-closest fictitious bus, a preliminary estimation of fault distance X_v can be made by:

$$X_v = \tau_f . V_s \tag{19}$$

where τ_f is the travel time from the fault-closest virtual bus to the fault-closest bus, and V_s is the travel time modal speed of the faulted section. This first computation is straight-forward, but not precise and it doesn't create redundancy in the fault distance computation compared to the following approach using a weighted directed tree network. Note that if the K-Means clustering is utilized only a subset (DFRs 1, 2, and 3) from all existing DFRs will be used to identify the fault-closest bus and faulted section.

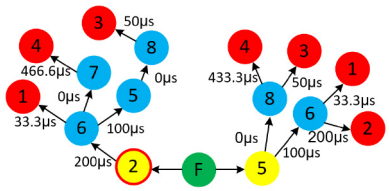

Fig. 2. Directed Tree from the Fault to the DFRs.

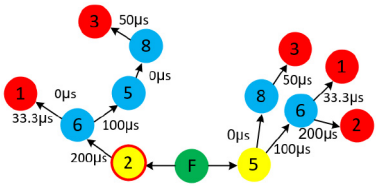

Fig. 3. Directed Tree from the Fault to the DFR 1, 2, and 3.

2) Finding the Shortest Path Directed Tree: Once the faulted line section is identified, the shortest travel time routes from fault to the DFRs can be visualized by a "weighted directed tree" as shown in Fig. 2. This facilitates the estimation of ToAs at each terminal bus of the faulted line and provides redundancy in the fault distance computation, making it more robust against bad data and ToA errors.

In Figure 2, the fault location F1 is indicated by green, terminal buses of the faulted line by yellow, DFR locations by red and remaining buses by blue. Note that bus 2 is circled in red to indicate the presence of DFR 2. The shortest paths taken by the fault traveling wave from F1 to DFRs are equivalent to the shortest paths from the terminal buses 2 and 5 to DFRs. Hence, the shortest travel time τ_i from each terminal bus to the DFRs can be calculated by:

$$\tau_i(i,k) = \sum_{j=1}^{N} \tau_j \tag{20}$$

where: $i = a, b$, correspond to terminal buses of faulted section, and $k = 1, 2, \ldots, p$, are the DFR indices. τ_j is the travel time of line section j along the shortest path from each terminal bus to the DFRs, and N is the number of line sections in the shortest path. The shortest path travel times for each terminal bus computed by (20) are then stored in the vectors τ_a and τ_b as shown below:

$$\underset{1\times p}{\tau_a} = \begin{bmatrix} \tau_{a_1} & \tau_{a_2} & \tau_{a_3} & \cdots & \tau_{a_p} \end{bmatrix} \tag{21}$$

$$\underset{1\times p}{\tau_b} = \begin{bmatrix} \tau_{b_1} & \tau_{b_2} & \tau_{b_3} & \cdots & \tau_{b_p} \end{bmatrix} \tag{22}$$

Moreover, the t_a and t_b arrival time vectors at each terminal bus of the faulted line can also be computed by (23) in (24), where, ToA_1 is the vector containing the ToAs of p DFRs.

$$\underset{1\times p}{t_a} = \underset{1\times p}{ToA_1{}^T} - \underset{1\times p}{\tau_a} \tag{23}$$

$$\underset{1\times p}{t_b} = \underset{1\times p}{ToA_1{}^T} - \underset{1\times p}{\tau_b} \tag{24}$$

Once the vectors t_a and t_b are computed, distance to the fault can be estimated by forming arrival time pairs and using the two-terminal fault location method presented in [14]. However, only one of the two arrival times stored in these two vectors will be the true value since the shortest path may be through one (bus 2) or the other (bus 5) terminal bus. Thus, the vectors t_a and t_b are evaluated as below in order to find the correct ToAs corresponding to the shorter path between those from the two terminal of incident edges to the DFRs.

$$t_a(j) = \begin{cases} 0 & \text{if } \tau_a(j) \geq \tau_b(j) \\ t_a(j) & \text{otherwise} \end{cases} \tag{25}$$

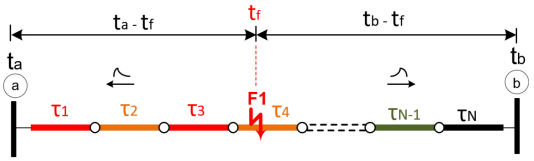

Fig. 4. Faulted Line Section.

$$t_b(j) = \begin{cases} 0 & \text{if } \tau_b(j) \geq \tau_a(j) \\ t_b(j) & \text{otherwise} \end{cases} \tag{26}$$

where j are the indices at each vector. For the above example of Fig. 2, τ_a and τ_b vectors will be as given below:

$$\tau_2 = \begin{bmatrix} 233.33\mu s & 0\mu s & 350\mu s & 666.67\mu s \end{bmatrix} \tag{27}$$

$$\tau_5 = \begin{bmatrix} 133.33\mu s & 300\mu s & 50\mu s & 433.33\mu s \end{bmatrix} \tag{28}$$

Using (25) and (26), the vectors t_a and t_b will be given as below:

$$t_2 = \begin{bmatrix} 0 & t_{a_2} & 0 & 0 \end{bmatrix} \tag{29}$$

$$t_5 = \begin{bmatrix} t_{b_1} & 0 & t_{b_3} & t_{b_4} \end{bmatrix} \tag{30}$$

Note that three arrival time pairs can be formed using the vectors t_2 and t_5. Hence, there are three possible ToA pairs that can be used to compute the fault distance. In the example where a subset of DFRs (DFRs 1, 2, and 3) clustered by K-Means, is used for fault identification, the tree graph will be as shown in Fig. 3. This configuration also will yield a precise estimation of fault distance but only using a pair of ToAs as below:

$$t_2 = \begin{bmatrix} 0 & t_{a_2} & 0 \end{bmatrix} \tag{31}$$

$$t_5 = \begin{bmatrix} t_{b_1} & 0 & t_{b_3} \end{bmatrix}. \tag{32}$$

3) Determining the Fault Occurrence Time t_{fD}, Faulted Section S_D, Fault Distance X_D Arrays: Having identified the faulted line section $(a-b)$, and the arrival time vectors t_a and t_b computed at each terminal bus of faulted section, the fault location problem can be solved by the two-terminal method described earlier in [14]. Consider the faulted line section of Fig. 4 to derive the vectors of the fault occurrence time t_{fD}, faulted section S_D and the fault distance X_D. Note that the different colors along the faulted line section in Fig. 4 represent different conductors such as overhead lines, overhead and underground cables. This line-cable hybrid configuration have traveling wave modal speed differences which impact in the precision of fault location. However, identification of the faulted incremental line section along this line section will enable accurate estimation of the fault distance. Detailed derivation of the method, and demonstration of its validity are already presented in [14]. Therefore, only a brief overview of

the approach will be given here. The number of arrival time pairs n formed by t_a and t_b to estimate the fault distance is given by $n = i \times j$, for $p = 1, 2, \ldots, n$, where i and j are the number of arrival times stored in t_a and t_b respectively. Each arrival time t_{a_i} and t_{b_j} and traveling times τ_k for each section k from buses a and b are used in (33) to estimate the fault occurrence time $t_{f_{(p)}}$ for each arrival time pair and whose values are then stored in vector t_{fD} as given in (34).

$$t_{f_{(p)}} = \frac{1}{2}\left(t_{a_i} + t_{b_j} - \sum_{k=1}^{N} \tau_k\right) \tag{33}$$

$$t_{fD} = \begin{bmatrix} t_{f_1} & t_{f_2} & t_{f_3} & \cdots\cdots\cdots & t_{f_{n-1}} & t_{f_n} \end{bmatrix} \tag{34}$$

The fault occurrence times $t_{f_{(p)}}$ stored in t_{fD} are then used to find the faulted sections $s_{(p)}$ for n pairs, which must satisfy the condition:

$$t_{a_i} - t_{f_{(p)}} \leq \sum_{k=1}^{s_{(p)}} \tau_k \tag{35}$$

These computed sections are then stored in an vector S_D:

$$S_D = \begin{bmatrix} s_1 & s_2 & s_3 & \cdots\cdots\cdots\cdots & s_{n-1} & s_n \end{bmatrix} \tag{36}$$

For $s_{(p)} = 1$, the fault distances $X_{(p)}$ can be computed by:

$$X_{(p)} = V_{s_{(p)}}\left(t_{a_i} - t_{f_{(p)}}\right) \tag{37}$$

Otherwise, for $s_{(p)} \geq 2$ the following expression must be used:

$$X_{(p)} = V_{s(p)}\left(t_{a_i} - t_{f_{(p)}} - \sum_{k=1}^{s_{(p)}-1} \tau_k\right) + \sum_{k=1}^{s_{(p)}-1} V_k \tau_k \tag{38}$$

where: $V_{s(p)}$ is the faulted section aerial modal speed (mode 1), τ_k and V_k are the travel time and aerial mode velocity for each line section k. The computed fault distances $X_{(p)}$ are stored in the fault distance vector X_D. These distances in X_D must be approximately equal. Therefore, their **"median"** is used to detect and remove outliers from X_D. Finally taking the **"mean"** of remaining entries in X_D will yield the accurate fault distance.

$$X_D = \begin{bmatrix} X_1 & X_2 & X_3 & \cdots\cdots & X_{n-1} & X_n \end{bmatrix}. \tag{39}$$

E. Differentiating Between Faults and Other Disturbances

A practical concern is differentiation between faults and other disturbances such as switching capacitors, sudden load changes, and high-frequency harmonics. Short-circuits mostly occur along lines, while capacitor switching and sudden load variations always occur at substation buses. Therefore, the identified location can be used as one discriminating factor. Furthermore, the recorded transient voltage at the closest DFR to the disturbance will easily confirm if it is a switching event or a fault. Additionally, the proposed algorithm is robust against high-frequency harmonics. Transients generated by short-circuits are higher than the harmonics whose WTC^2 values remain insignificant compared to faults.

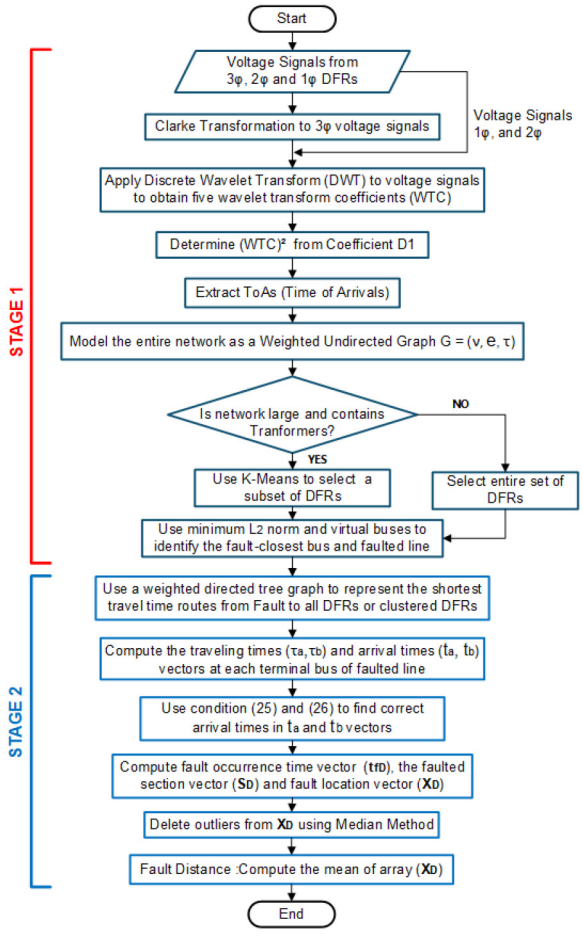

Fig. 5. Flowchart of Proposed Fault Location Method.

F. Flowchart of Fault Location Method

Figure 5 shows the flowchart for the implementation stages of the proposed fault location algorithm. It starts with stage 1 where the ToAs are extracted from recorded fault transients and network is modeled as a weighted undirected graph. For large networks with power transformers, DFRs are clustered by K-Means method. Otherwise, all available DFRs are used. ToAs of all DFRs or selected DFRs along with the undirected graph are used to identify the "fault-closest bus". Then virtual buses are placed along its incident edges to identify the faulted line. In stage 2, a weighted directed tree network is used to form arrays of travel times (τ_a and τ_b) and arrival times (t_a and t_b) for the faulted line terminals a-b. Using (25) and (26) the correct ToAs are determined for t_a and t_b. These are then used by the two-terminal fault location method to compute the arrays t_{fD}, S_D, and X_D. Finally, outliers are removed from X_D by using the median and the precise fault distance is found by calculating the mean of remaining X_D entries. Hence, the main technical contributions and advantages of the proposed fault location method are highlighted below.

1) The presented approach can be used for any type of meshed or radial power network. Optimal DFR placement strategy for both types of systems are also described.
2) ToA errors due to TW attenuation of fault-generated travel waves passing through power transformers are

avoided by selecting a subset of DFRs using K-Means clustering method.

3) Modeling of shortest paths from fault to DFRs using the weighted directed tree graph enables the use of all possible DFR pairs to estimate the fault distance. Hence, the proposed fault distance estimation has significant built-in redundancy.
4) The proposed algorithm is robust against bad data, time-synchronization errors, and network topology changes. It can accurately estimate the fault distance using either a subset or the entire set of DFRs. Since this leads to including only the shortest routes between the fault and selected DFRs, topology changes will not impact the method's performance unless they occur along any of these shortest routes. Finally, use of the median in X_D facilitates detection and removal of outliers from X_D, making the proposed method even more robust against erroneous data and synchronization-time errors.
5) Use of the travel times of each line section eliminates the TW modal speed differences in hybrid line-cable sections. Since the two-terminal fault location method of [14] is used for the identified faulted line, the issue of TW modal speed differences is eliminated since the specific faulted line section along the faulted line will be determined before estimating the fault distance. Thus, the non-homogeneity of hybrid line-cable sections will not impact the accuracy of the proposed algorithm.
6) Load or capacitor switching events and faults are differentiated by their location (switchings cannot occur along lines but at substation buses) as well as by inspecting the voltage transients recorded at the closest DFR, characteristics of switching events are easily identified and differentiated from faults.
7) High-frequency harmonics have no impact on the proposed algorithm's precision. Plot of WTC^2 magnitudes shows that harmonic distortion remains imperceptible in the presence high-frequency transients generated by short-circuits.

III. Practical Implementation

A. Simulation Model for Transmission Networks

The IEEE 118 bus system shown in Fig. 6 whose model data for EMTP-type simulation studies developed by CIGRE [31] is used to validate the fault location algorithm for meshed transmission networks. The model includes 138KV and 345KV voltage networks interconnected via nine power transformers. Seven lateral lines connected to the meshed network are identified in Fig. 6. Therefore, seven DFRs are manually placed in these line terminals. Twenty-eight additional DFRs are installed in the meshed network following the approach of the Optimal DFR Placements(ODP). In total, 35 DFRs are required to make any fault identifiable in the system. The transient simulations are carried out by using a sampling frequency of 1MHz in order to maintain high accuracy. Besides, the voltage signals captured by DFRs are exposed to White Gaussian Noise of 70db. Four wind parks of 443.42MVA capacity are installed at buses 10, 26, 65 and 116 as shown in Fig. 6. Each

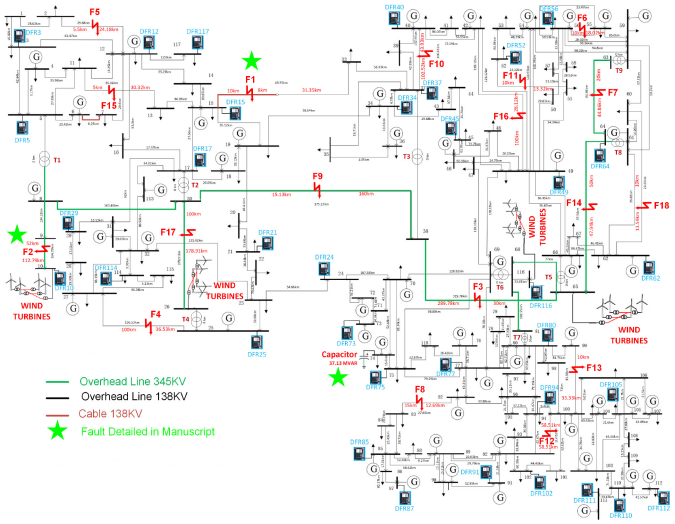

Fig. 6. IEEE 118 Bus System-CIGRE Model.

wind farm includes Full-Size Converter(FSC) Wind Turbines 266 connected at buses via a park power transformer of rating 500 MVA, 34.5/345KV and whose Inverter-Based Power Sources(IBPSs) is set to operate in voltage-controlled mode at 1 p.u. Each turbine is rated to deliver 1.667 MVA. The FSC WT uses a permanent magnet synchronous generator(PMSG), and the ac-dc-ac converter system consists of two voltage source converters(VSCs): machine side converter (MSC) and grid side converter (GSC). The MSC converts the 0.575KV provided by PMSG to 1150 VDC link voltage, and the GSC converts the DC link voltage to 34500 VAC. The grid collector estimated for the interconnection of each wind farm into grid are $R = 0.02\Omega$, $L = 0.06mH$, and $C = 47\mu F$. The model also includes a wind park controller (WPC), which controls reactive power at the point of interconnection (POI). The converters are modeled using the Average Value Model(AVM) type converter, whose detailed model can be found in [32]. Furthermore, in order to illustrate the performance of the proposed fault location approach in the presence of renewables with fault ride-through (FRT) features, the controls of four wind farms are accordingly simulated following the recommendation of ISO-NEW England [33].

B. Fault Simulations in Transmission Networks

1) DLG Fault at 10 km From Bus 15: A double-line-to-ground (DLG) fault (F1) shown in Fig. 6 occurs along the 49.35 km hybrid line-cable branch connecting buses 15-33, 10 km from bus 15. The fault is simulated assuming a 30 Ω fault resistance and a 30° fault inception angle, which is set to $30.69ms$. Due to the presence of power transformers, only a subset of DFRs will be selected by the K-Means method based on their feature vector $x^{(i)}$ formed by ToAs (t_i) and shortest travel times ($\tau_{(R,i)}$). Here, DFR15 at bus 15 is used as the reference being closest to the fault. These shortest travel times used in $x^{(i)}$ are computed using the graph of Fig. 7. For this case, selecting the blue cluster of DFRs 15, 17, 37, 34, 29, 21, 12, 117, 114, 5 and 3 as shown in Table II and Fig. 8 are sufficient to estimate the precise fault location. These 11

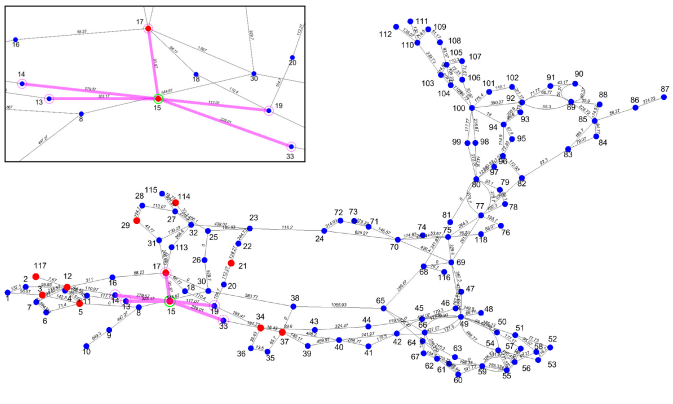

Fig. 7. Weighted Undirected Graph For IEEE 118 Bus System.

TABLE II
TOA(MS) FOR FAULTS F1, F2, AND CAPACITOR

DFR	ToA$_{F1}$	DFR	ToA$_{F2}$	DFR	ToA$_{Cap}$
15	30.758	10	32.44	75	33.514
17	30.849	5	32.69	77	33.779
37	31.037	12	32.82	73	33.896
34	31.053	117	32.83	116	33.921
29	31.101	3	32.83	80	34.055
21	31.141	-	-	-	-
12	31.148	-	-	-	-
117	31.156	-	-	-	-
114	31.220	-	-	-	-
5	31.284	-	-	-	-
3	31.360	-	-	-	-

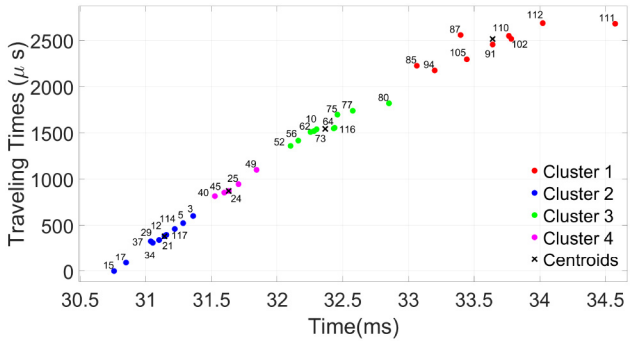

Fig. 8. K-Means Clustering for F1 in IEEE 118 Bus System.

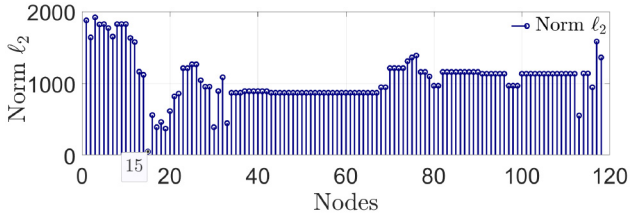

Fig. 9. ℓ_2-Norm Estimation For F1 in IEEE 118 Bus System.

ToAs of clustered DFRs circled in red in Fig. 7 are then used by the minimum ℓ_2-norm method to identify the closest bus to the fault. Bus 15 in Fig. 9 is identified as the fault-closest bus, which is also indicated by the green circled node in Fig. 7.

The algorithm then selects five incident edges connected to the fault-closest bus, which are the magenta edges 15-13, 15-14, 15-17, 15-19, and 15-33, as shown in Fig. 7. The next

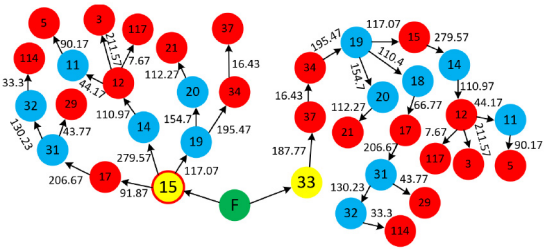

Fig. 10. Directed Tree Model For Fault F1.

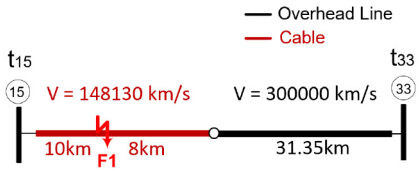

Fig. 11. Faulted Section for the F1 Fault.

step is to place virtual buses along the selected edges to identify the faulted line. The virtual bus number in transmission systems is selected by rounding the traveling time of each incident edge into their nearest integer. Hence, 323 (15-13), 280 (15-14), 92 (15-17), 117 (15-19), and 226 (15-33) virtual buses are installed equally spaced in each incident edge respectively. The virtual bus with minimum ℓ_2-norm between buses 15 and 33 is identified.

The next step is to compute the travel times (τ_a and τ_b) and arrival times (t_a and t_b) vectors at each terminal bus 15 and 33 using weighted directed tree graph and shortest routes from each terminal bus to DFRs as shown in Fig. 10. Note that the incident bus 15 is circled in red to indicate the presence of a DFR at bus 15. The travel times (μs) at each terminal bus are then computed using (20) and stored in the following two vectors:

$$\tau_{15} = \begin{bmatrix} 0 & 91.87 & 328.97 & 312.54 & 342.31 & 384.04 & 390.54 \\ & 398.21 & 462.07 & 524.88 & 602.11 \end{bmatrix}$$

$$\tau_{33} = \begin{bmatrix} 516.74 & 576.84 & 187.77 & 204.2 & 827.28 & 666.64 & 907.28 \\ & 914.95 & 947.04 & 1041.6 & 1118.8 \end{bmatrix}$$

The t_{15} and t_{33} vectors at each incident bus are computed using (23) and (24), and whose values are evaluated through the conditions (25) and (26) to find the correct ToAs to be stored as below.

$$t_{15} = \begin{bmatrix} 30.758 & 30.758 & 0 & 0 & 30.759 & 30.758 & 30.758 & 30.758 \\ & & & & & 30.758 & 30.758 & 30.758 \end{bmatrix}$$

$$t_{33} = \begin{bmatrix} 0 & 0 & 30.85 & 30.85 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Finally, having the arrival time vectors t_{15} and t_{33} on each side of faulted line enables the computation of fault location vectors 34, 36, and 39. Each vector will have $n = 18$ entries, which is the possible number of arrival time pairs that can be formed between elements of t_{15} and t_{33}. Equations (33), (35), and (38) are evaluated for all 18 cases. Here, (38) is used instead of (37) since the faulted section contains a hybrid line-cable conductor as shown in Fig. 11. The different modal

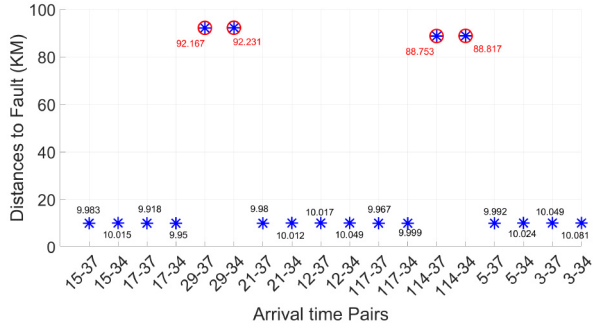

Fig. 12. Extreme Outliers generated by Topology Changes.

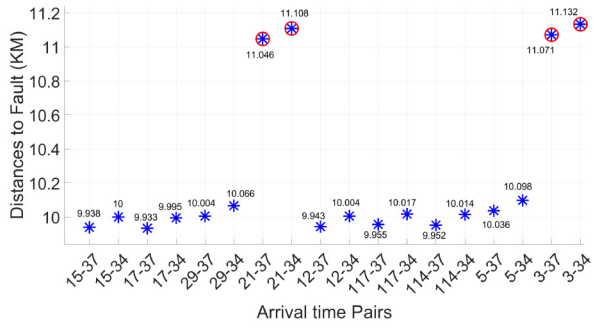

Fig. 13. Gross Errors Detected by Median Method.

velocities of different line sections are properly accounted for before estimating the fault distance.

The proposed algorithm chooses the fault-closest bus 15 as the reference to estimate the fault location vectors t_{fD}, S_D and X_D as shown below. Taking the median of the entries in X_D, the influence of bad data and synchronization-time errors are eliminated. Once the detected outliers are removed, the mean of remaining entries in X_D will yield the best estimate for fault distance, i.e., 9.9928 km. Comparing it with the true distance of 10 km yields a absolute error of 7.2 m.

$$t_{fD} = \begin{bmatrix} 30.69 & 30.69 & 30.69 & 30.69 & 30.69 & 30.69 & 30.69 & 30.69 & 30.69 \\ 30.69 & 30.69 & 30.69 & 30.69 & 30.69 & 30.69 & 30.69 & 30.69 & 30.69 \end{bmatrix}$$

$$S_D = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

$$X_D = \begin{bmatrix} 9.938 & 9.99 & 9.933 & 9.994 & 10 & 10.06 & 9.935 & 9.996 & 9.942 & 10 & 9.955 \\ 10.01 & 9.952 & 10.014 & 10.09 & 10.036 & 9.959 & 10.02 \end{bmatrix}.$$

2) Robustness Against Topology Changes: As shown previously, only 31% of DFRs are required to identify the fault-closest bus and faulted line section for fault F1. Then, the directed tree network is used to represent the shortest travel times from fault to clustered DFRs as shown in Fig. 10. Hence, topology changes outside the tree network will have no impact on fault location. Moreover, the redundancy created by the use of directed tree will facilitate detection of outliers via the median estimation in case of a topology change in the directed tree. Consider the same fault F1 and five topology changes in the IEEE 118 bus system of Fig. 6: three line sections (35-36, 39-40, and 22-23) outside and two line sections (29-31 and 32-114) inside the directed tree will be open. The last two changes only impact the array X_D. However, using median as the estimator will detect them as extreme outliers as shown in Fig. 12. These outliers are then removed and taking the mean of X_D will yield a fault distance of 10.002 km, with an absolute error of 2 m. illustrating the method's robustness against topology changes.

3) Performance in the Presence of Gross Errors in ToAs: Consider the same fault F1 but assume that ToAs measured by DFRs 3, 21 and 24 have errors of $20\mu s$, $15\mu s$, and $30\mu s$ respectively. As shown in Fig. 8, KMeans clustering excludes DFR 24 from the selected cluster of 11 DFRs (marked in blue) to estimate the fault distance. Therefore, its error will not impact fault location, leaving only the ToAs

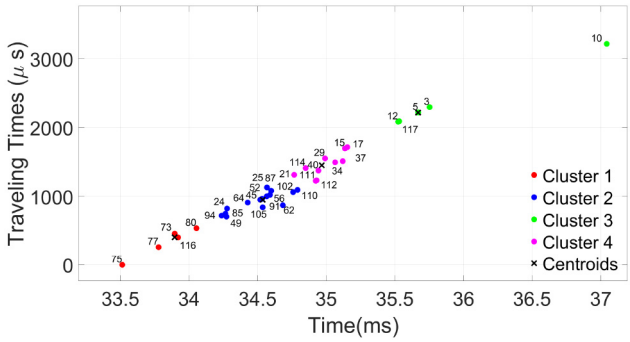

Fig. 14. K-Means For Switching of Capacitor at bus 74.

from DFR 21 and 3 as the two whose errors could affect fault location. However, taking the median of X_D, fault distances 11.046 km, 11.107 km, 11.07 km, and 11.132 km are identified as outliers which are marked in red in the graph of Fig. 13 where the differences between these distances and the rest of the estimated distances can be more clearly seen. Note that the error in ToA of DFR 21 generates two outliers due to the arrival time pairs created with DFRs 37 and 34. Likewise, DFR 3 yields two outliers with the same DFRs 37 and 34. The median removes the outliers in X_D and subsequent calculation of the mean of X_D yields a fault distance of 9.9969 km, with an absolute error of 3.1 m. This illustrates the robustness of the method against ToA errors thanks to the use of KMeans clustering and median estimator to detect and remove outliers from X_D.

4) Differentiating Between Faults and Switching Events: Transients caused by switching of capacitors or load/generation need to be differentiated from faults. Consider a transient caused by switching of a 138KV, 37.13 MVAR capacitor at bus 74 as shown in Fig. 6. The switch closing angle is set to 90°, which is equal to 33.46 ms. The K-Means algorithm yields the cluster of DFRs 75, 77, 73, 116, and 80 as shown in red in Fig. 14. The algorithm uses the arrival times of clustered DFRs shown in Table II to identity bus 74 and line 74-75 as the fault-closest bus and faulted line section respectively. Then the fault distance vector X_D will be calculated as:

$$X_D = \begin{bmatrix} 0.0055 & 0.0 & 0.0 & 0.0 \end{bmatrix}$$

The mean of X_D yields the fault distance of 0.001375 km. from bus 74, with an absolute error of 1.375 m. Being so close

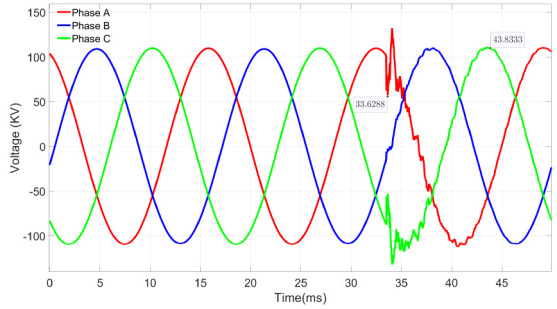

Fig. 15. Voltage Transients of Capacitor Switching at bus 74.

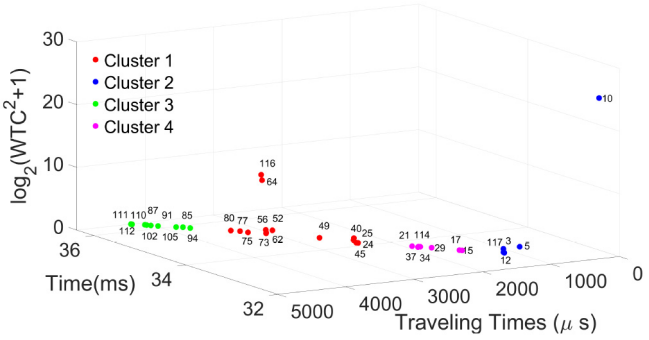

Fig. 16. K-Means For F2 in IEEE 118 System.

to the bus, this result clearly implies that there is a chance of this disturbance may be caused by a switching event at the bus. It can be easily verified by inspecting the recorded voltage transient by DFR 75 the closest to the identified disturbance location. As shown in Fig. 15, the 3-phase transient voltage lasts about 10.2 *ms* before it returns to its previous steady state value, identifying the cause as a possible switching event.

5) Impact of IBPSs on Fault Location: It is widely known that the asymmetry of lines, fault resistance, and fault ride-through (FRT) requirements implemented in IBPSs adversely impact fault detection and may lead to malfunctioning of distance relays. To illustrate the robustness of the proposed approach to these three features consider a single line to ground (SLG) fault at (F2) in Fig. 6 along 164.79km long line connecting buses 9-10, 112.79 km from the 443.422 MVA wind park connected to bus 10, and 52 km from bus 9. The fault is simulated with a 100 Ω fault resistance and a fault inception angle of 60°, equivalent to 32.07 ms. Note that fault F2 occurs at 345KV line and its TW signal passes by power transformer T1 345/138KV to reach and spread to the 138KV side. Hence, the attenuation generated by T1 will be better visualized in a 3D cluster using the 3D feature vector given by (4). Applying the K-Means algorithm, the blue cluster of DFRs 10, 5, 12, 117, and 3 in Fig. 16 are chosen. It is clearly shown in Fig. 16, DFR10 captures the largest WTC^2 due to its proximity to the fault. The other blue clustered DFRs registered lower magnitudes due to their location on the 138KV side. The proposed technique uses the arrival times captured by clustered DFRs and shown in Table II to estimate the fault-closest bus and faulted line section as 9 and 9-10 respectively. The fault distance array X_D will be

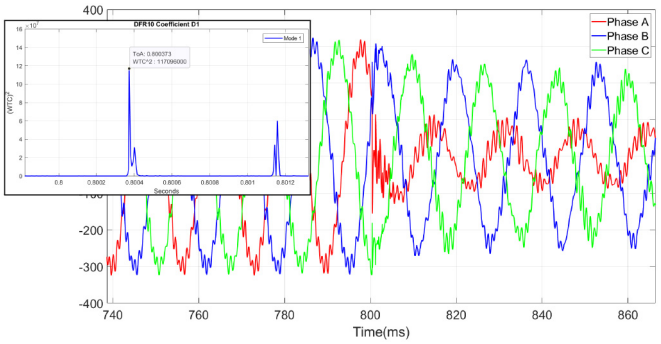

Fig. 17. Harmonic Distortion by Wind Turbine Entrance.

obtained as below:

$$X_D = \begin{bmatrix} 52.025 & 52.03 & 52.046 & 52.04 \end{bmatrix}$$

Extracting the mean of X_D will yield a fault distance 52.035 km and comparing it with the true distance of 52 km yields an absolute error of 35.25 m. As evident from these results, the Fault Ride-Through protection of IBPs has no effect on the accuracy of proposed method.

6) Robustness Against High-Frequency Harmonics: Switching of IBPSs generates harmonics, however these harmonics occur in lower frequencies than those generated by short-circuits. Let us consider a drastic case of a wind farm being connected at bus 10 as shown in Fig. 6. The wind farm is connected at 0.3s and a SLG fault occurs at 0.8s. Assuming that the wind farm is not equipped with a proper harmonic filter resulting in harmonic distortion as shown in Fig. 17. The WTC^2 for the voltage signal recorded by the fault-closest DFR10 is shown in the left upper corner of Fig. 17. It is evident that the high-frequency harmonics are imperceptible in the presence of high-frequency transients generated by the short-circuit. Therefore, the fault distance of 52.035 km computed for this case will be similar to the previous case, validating that high-frequency harmonics will not have any effect on the accuracy of proposed algorithm.

C. Summary of Simulated Cases for the IEEE 118-Bus System

Fault scenarios as depicted in Fig. 6 are simulated for the IEEE 118-bus system by using different fault parameters for the fault resistance and fault inception angle, in the presence of non-transposed lines and IBPSs in the network. Simulation results give maximum and average absolute errors of 27.46m and 10.19 m respectively as shown in Table III. These results provide experimental evidence that the configuration and fault types do not have significant impact on the performance of the proposed algorithm.

D. Fault Location in Distribution Networks

IEEE 123-node test feeder [34] shown in Fig. 18 is used to test the proposed method for radial and meshed distribution networks.The radial system is created by closing the green-colored switches and opening the two blue-colored switches. Otherwise, meshed configuration is created in the network.

TABLE III
FAULT LOCATION ERRORS AT IEEE 118 SYSTEM

Fault	Location	Rf	Fault Angle	Fault Type	Err (m)
F3	9.38% L_{65-38}	40Ω	135°	A-g	22.5
F4	26.75% L_{25-27}	20Ω	90°	C-g	5.55
F5	18.53% L_{2-12}	40Ω	60°	BC-g	4.8
F6	35.68% L_{56-55}	30Ω	45°	ABC-g	5
F7	30.83% L_{63-64}	50Ω	90°	ABC	25
F8	45.82% L_{82-83}	60Ω	10°	B-g	2.27
F9	8.63% L_{30-38}	100Ω	30°	AC-g	15.7
F10	30% L_{40-37}	120Ω	75°	AB	8.75
F11	42.88% L_{52-51}	140Ω	90°	ABC	6.35
F12	50% L_{100-92}	80Ω	135°	A-g	3.25
F13	28.3% L_{99-100}	90Ω	60°	AB-g	4.45
F14	48.94% L_{65-64}	70Ω	45°	B-g	18.25
F15	14.15% L_{11-13}	100Ω	60°	AC-g	2.45
F16	21.9% L_{42-49}	75Ω	150°	ABC-g	7.56
F17	35.85% L_{30-26}	100Ω	30°	AC-g	27.46
F18	42.39% L_{60-62}	90Ω	90°	AC	3.75

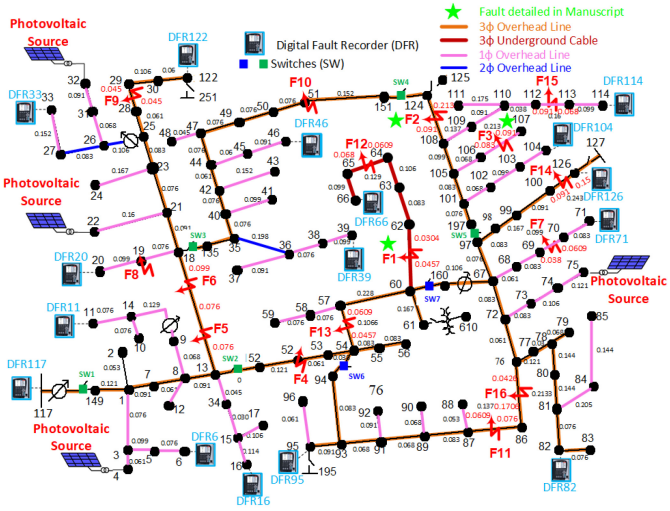

Fig. 18. IEEE 123 Test Node Feeder.

Four single-phase PV units are connected to the system via Inverter Based Power Sources (IBPSs). Each IBPS provides 20 KVA to the grid in "power factor control" mode with p.f set at 0.95. Description of the "Average Value Model" (AVM) type converter used in this study can be found in [32]. A DFR is placed in each lateral terminal bus to identify any fault network, as was described in Section II-A. Hence, sixteen DFRs are placed strategically at these lateral terminals, six of which are 3-phase, and the rest are single-phase. Due to the short line lengths, a sampling frequency of 10 MHz is required to obtain precise fault distances. While no DFRs are placed for very short laterals, in case of a fault occurring along these short laterals, the proposed algorithm will estimate the fault location as the junction where the faulted lateral connects to the main feeder or through identification of buses with minimum ℓ_2-norm as demonstrated in [14], [25].

E. Fault Simulations in Radial Distribution Network

Consider the green switches closed and blue switches open in the network of Fig. 18 which forms a radial configuration. A double-line to ground (DLG) fault F1 shown in Fig. 18 occurs along a 0.0761 km long underground cable between buses 60 - 62, at 0.03047 km from bus 62. The existing underground cable in the network goes all the way from bus 60 to 66, marked in brown in Fig. 18. The fault resistance and inception angle are set to 140 Ω and 120° (34.81ms) respectively. This radial system doesn't contain any power transformer in its main feeder. Therefore, the entire DFR set will be used to estimate the fault distance. The arrival times captured by sixteen DFRs shown in Table IV are used by the minimum ℓ_2-norm method to identify the fault-closest bus and placing five virtual buses in each incident edge connected to bus 62, the faulted section is identified as 60-62. The shortest path from fault to DFRs are modeled by the weighted directed tree graph to compute the fault distance vector X_D. Taking the median in X_D leads to identification and removal of outliers to obtain the array as shown below. Finally, calculating the mean in X_D will yield a fault distance 0.03064 km, having 0.17 m absolute error with respect to the true distance of 0.03047 km.

$$X_D = \big[0.0293\ \ 0.0307\ \ 0.033\ \ 0.033\ \ 0.0314\ \ 0.0306\ \ 0.00276\ \ 0.0306\ \ 0.0294\ \ 0.0318\ \ 0.0326\ \ 0.0304\ \ 0.0268\big].$$

TABLE IV
TOA(MS) FOR FAULTS F1, F2, AND F3

DFR	ToA$_{F1}$	DFR	ToA$_{F2}$	DFR	ToA$_{F3}$
16	34.8129	114	34.8120	114	30.6926
117	34.8135	104	34.8123	126	30.6930
11	34.8135	71	34.8125	95	30.6933
20	34.8136	46	34.8127	117	30.6947
66	34.8139	126	34.8131	66	30.6960
6	34.8140	95	34.8134	122	30.6963
39	34.8147	82	34.8136	-	-
46	34.8147	20	34.8139	-	-
122	34.8150	39	34.8141	-	-
33	34.8151	16	34.8142	-	-
114	34.8184	117	34.8148	-	-
104	34.8186	11	34.8149	-	-
71	34.8189	122	34.8152	-	-
126	34.8195	6	34.8154	-	-
82	34.8200	33	34.8154	-	-
95	34.8204	66	34.8161	-	-

F. Fault Simulations in Meshed Distribution Network

Closing the blue and green switches will create a meshed distribution system of Fig. 18. The meshed configuration has many laterals connected to its main feeder, hence placing DFRs in lateral terminal buses similar to the radial configuration will be sufficient to identify any fault. Consider a single-line to ground (SLG) fault at F2, 0.0914 km from bus 108 as shown in Fig. 18. Fault resistance and incidence angle are set to 30 Ω and 120° (34.81ms). The sixteen DFRs capture the ToAs shown in Table IV to form the vector X_D as shown below. Note that using the entire DFR set forms 55 ToA pairs, a highly redundant set. Calculating the median of X_D removes all outliers and mean of remaining entires will yield a fault distance 0.0916 km, having an absolute error of 0.2m.

$$X_D = \Big[0.0886\ \ 0.0855\ \ 0.0931\ \ 0.08965\ \ \cdots\cdots\cdots\ \ 0.0919\ \ 0.08794\Big].$$

G. Use of Reduced Number of DFRs

When working with a limited budget, it may be preferable to use a reduced number of DFRs in meshed and radial distribution networks by placing DFRs in terminal buses that can

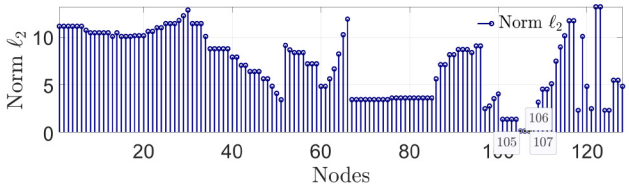

Fig. 19. ℓ_2-Norm Estimation For Faulted Lateral.

TABLE V
FAULT LOCATION ERRORS AT IEEE 123 SYSTEM

Fault	Location	Rf	Fault Angle	Fault Type	Err (m)	Network Type
F4	Bus52	120Ω	135°	ABC-g	3.62	Radial
F5	30.27% L_{13-18}	10Ω	60°	BC	1.5	Meshed
F6	39.44% L_{18-13}	150Ω	10°	AB-g	1.85	Radial
F7	38.46% L_{69-70}	180Ω	45°	A-g	2.6	Meshed
F8	Bus19	200Ω	90°	A-g	3.5	Radial
F9	50% L_{28-29}	160Ω	135°	ABC	9.5	Meshed
F10	Bus51	170Ω	90°	B-g	2.82	Radial
F11	44.44% L_{87-86}	100Ω	30°	AC-g	6.29	Radial
F12	47.05% L_{64-65}	250Ω	45°	C-g	2.18	Meshed
F13	42.85% L_{54-57}	5Ω	45°	C-g	1.94	Radial
F14	37.5% $L_{100-126}$	260Ω	45°	AC-g	1.86	Meshed
F15	42.5% $L_{113-112}$	3Ω	150°	A-g	1.56	Meshed
F16	20% L_{76-86}	8Ω	90°	AB-g	2.34	Radial

form shortest routes to cover only the main feeders of the network. The method will still locate accurately any fault in these main feeders while only identifying faulted laterals if that lateral does not have a DFR at its terminal bus. To illustrate this for the meshed configuration, let us place three-phase DFRs only at (117, 122, 126, 95 and 66) and one single-phase DFR at 114. Consider a fault along the single-phase lateral at F3, 0.0914 km from bus 107 in Fig. 18, and whose ToAs are shown in Table IV. There is no DFR at this lateral's terminal bus. Fault resistance and angle are set to 100 Ω and 30° (30.69ms). As shown in Fig. 19, the minimum ℓ_2 norm is the same for buses 105, 106, and 107. Hence, the algorithm identifies the faulted lateral as 105-107, but not the exact fault point due to the missing terminal bus DFR.

H. Summary of Fault Simulations in IEEE 123-Node Feeder

This section summarizes the results for faults in meshed and radial networks shown in Fig. 18, which are obtained using different fault parameters in order to test the proposed method under different types of faults . Table V shows the results with maximum and average absolute errors of 9.5m and 3.19 m, respectively. These results experimentally validate insensitivity of the method to changes in network configuration and fault type in radial and meshed distribution systems.

IV. CONCLUSION

This paper presents and validates a robust fault location method that can be used for radial or meshed power systems. The method avoids the ToA errors due to the attenuation by power transformers by selecting a subset of DFRs closest to the fault using the K-Means clustering method. The method is robust against gross errors in measurements, fault-ride-through controls of IBPSs, topology changes and can differentiate between faults and other switching and harmonics events. Meshed and radial networks are used to validate the proposed method performance under different fault types and conditions.

REFERENCES

[1] M. S. Sachdev and R. Agarwal, "A technique for estimating transmission line fault locations from digital impedance relay measurements," *IEEE Trans. Power Del.*, vol. 3, no. 1, pp. 121–129, Jan. 1988.

[2] E. Styvaktakis, M. H. J. Bollen, and I. Y. H. Gu, "A fault location technique for two and three terminal lines using high frequency fault clearing transients," in *Proc. PowerTech Budapest Abstract Rec.*, Aug. 1999, p. 255.

[3] F. H. Magnago and A. Abur, "Fault location using wavelets," *IEEE Trans. Power Del.*, vol. 13, no. 4, pp. 1475–1480, Oct. 1998.

[4] S.-J. Huang, X.-Z. Liu, W.-F. Su, and T.-C. Ou, "Application of enhanced honey-bee mating optimization algorithm to fault section estimation in power systems," *IEEE Trans. Power Del.*, vol. 28, no. 3, pp. 1944–1951, Jul. 2013.

[5] S. Azizi, M. Sanaye-Pasand, and M. Paolone, "Locating faults on untransposed, meshed transmission networks using a limited number of synchrophasor measurements," *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4462–4472, Nov. 2016.

[6] M. Gholami, A. Abbaspour, M. Moeini-Aghtaie, M. Fotuhi-Firuzabad, and M. Lehtonen, "Detecting the location of short-circuit faults in active distribution network using PMU-based state estimation," *IEEE Trans. Smart Grid*, vol. 11, no. 2, pp. 1396–1406, Mar. 2020.

[7] G. Feng and A. Abur, "Identification of faults using sparse optimization," in *Proc. 52nd Annu. Allerton Conf. Commun. Control Comput. (Allerton)*, 2014, pp. 1040–1045.

[8] G. Feng and A. Abur, "Fault location using wide-area measurements and sparse estimation," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 2938–2945, Jul. 2016.

[9] I. Rozenberg, Y. Beck, Y. C. Eldar, and Y. Levron, "Sparse estimation of faults by compressed sensing with structural constraints," *IEEE Trans. Power Syst.*, vol. 33, no. 6, pp. 5935–5944, Nov. 2018.

[10] K. Jia, B. Yang, X. Dong, T. Feng, T. Bi, and D. W. P. Thomas, "Sparse voltage measurement-based fault location using intelligent electronic devices," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 48–60, Jan. 2020.

[11] K. Jia, B. Yang, T. Bi, and L. Zheng, "An improved sparse-measurement-based fault location technology for distribution networks," *IEEE Trans. Ind. Informat.*, vol. 17, no. 3, pp. 1712–1720, Mar. 2021.

[12] A. D. Filomena, R. H. Salim, M. Resener, and A. S. Bretas, "Ground distance relaying with fault-resistance compensation for unbalanced systems," *IEEE Trans. Power Del.*, vol. 23, no. 3, pp. 1319–1326, Jul. 2008.

[13] S. Paladhi and A. K. Pradhan, "Adaptive distance protection for lines connecting converter-interfaced renewable plants," *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 9, no. 6, pp. 7088–7098, Dec. 2021.

[14] C. Galvez and A. Abur, "Fault location in active distribution networks containing distributed energy resources (DERs)," *IEEE Trans. Power Del.*, vol. 36, no. 5, pp. 3128–3139, Oct. 2021.

[15] "SEL-T400L Time-Domain Line Protection." [Online]. Available: https://selinc.com/products/T400L/ (Accessed: 2022).

[16] J. Ding, X. Wang, Y. Zheng, and L. Li, "Distributed traveling-wave-based fault-location algorithm embedded in multiterminal transmission lines," *IEEE Trans. Power Del.*, vol. 33, no. 6, pp. 3045–3054, Dec. 2018.

[17] S. Xie *et al.*, "Application of integrated optical electric-field sensor on the measurements of transient voltages in AC high-voltage power grids," *Appl. Sci.*, vol. 9, no. 9, p. 1951, 2019.

[18] M. Korkalı, H. Lev-Ari, and A. Abur, "Traveling-wave-based fault-location technique for transmission grids via wide-area synchronized voltage measurements," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 1003–1011, May 2012.

[19] M. Korkali and A. Abur, "Optimal deployment of wide-area synchronized measurements for fault-location observability," *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 482–489, Feb. 2013.

[20] K. Yu, J. Zeng, X. Zeng, F. Xu, Y. Ye, and Y. Ni, "A novel traveling wave fault location method for transmission network based on directed tree model and linear fitting," *IEEE Access*, vol. 8, pp. 122610–122625, 2020.

[21] R. Chen, X. Yin, X. Yin, Y. Li, and J. Lin, "Computational fault time difference-based fault location method for branched power distribution networks," *IEEE Access*, vol. 7, pp. 181972–181982, 2019.
[22] E. C. M. Maritz, J. M. Maritz, and M. Salehi, "A travelling wave-based fault location strategy using the concepts of metric dimension and vertex covers in a graph," *IEEE Access*, vol. 9, pp. 155815–155825, 2021.
[23] R. Liang *et al.*, "Fault location method in power network by applying accurate information of arrival time differences of modal traveling waves," *IEEE Trans. Ind. Informat.*, vol. 16, no. 5, pp. 3124–3132, May 2020.
[24] C. Galvez and A. Abur, "Fault location in meshed and active power distribution networks," in *Proc. IEEE Madrid PowerTech*, 2021, pp. 1–6.
[25] C. Galvez and A. Abur, "Fault location in meshed distribution systems using a minimum number of digital fault recorders," in *Proc. North Amer. Power Symp. (NAPS)*, 2021, pp. 01–06.
[26] B. Xu and A. Abur, "Observability analysis and measurement placement for systems with PMUs," in *Proc. IEEE PES Power Syst. Conf. Exposit.*, vol. 2, 2004, pp. 943–946.
[27] J. R. Marti, "Accurate modelling of frequency-dependent transmission lines in electromagnetic transient simulations," *IEEE Trans. Power App. Syst.*, vol. PAS-101, no. 1, pp. 147–157, Jan. 1982.
[28] H. W. Dommel, *EMTP Theory Book*. Vancouver, BC, Canada: Microtran Power Syst. Anal. Corp., 1996.
[29] I. Daubechies, *Ten Lectures on Wavelets*, vol. 61. Philadelphia, PA, USA: Siam, 1992.
[30] K. P. Murphy, *Machine Learning: A Probabilistic Perspective*. Cambridge, MA, USA: MIT Press, 2012.
[31] "Power System Test Cases for EMT-Type Simulation Studies." CIGRE. [Online]. Available: https://e-cigre.org/publication/736-power-system-test-cases-for-emt-type-simulation-studies (Accessed: Feb. 3, 2018).
[32] U. Karaağaç *et al.*, "A generic EMT-type model for wind parks with permanent magnet synchronous generator full size converter wind turbines," *IEEE Power Energy Technol. Syst. J.*, vol. 6, no. 3, pp. 131–141, Sep. 2019.
[33] "Inverter Source Requirement Document of ISO New England (ISO-NE)." ISO-NE. 2018. [Online]. Available: https://www9.nationalgridus.com/
[34] "IEEE 34 Test Node Feeder." IEEE. Sep. 2010. [Online]. Available: http://sites.ieee.org/pes-testfeeders/resources/

Cesar Galvez (Student Member, IEEE) received the B.S. degree from Callao University, Peru, in 2005, and the M.S. degree from Northeastern University, Boston, in 2016, where he is currently pursuing the Ph.D. degree in power system with specialization in control and power system modeling.

Ali Abur (Fellow, IEEE) received the B.S. degree in EE from Orta Doğu Teknik Üniversitesi, Turkey, and the M.S. and Ph.D. degrees from The Ohio State University. He is currently a Professor with the Electrical and Computer Engineering Department, Northeastern University, Boston, MA, USA.