Jamali et al. Protection and Control of Modern Power Systems (2020) 5:17
https://doi.org/10.1186/s41601-020-00162-y

Protection and Control of Modern Power Systems

ORIGINAL RESEARCH Open Access

Hybrid classifier for fault location in active distribution networks

Sadegh Jamali*, Alireza Bahmanyar and Siavash Ranjbar

Abstract

This paper presents a fast hybrid fault location method for active distribution networks with distributed generation (DG) and microgrids. The method uses the voltage and current data from the measurement points at the main substation, and the connection points of DG and microgrids. The data is used in a single feedforward artificial neural network (ANN) to estimate the distances to fault from all the measuring points. A k-nearest neighbors (KNN) classifier then interprets the ANN outputs and estimates a single fault location. Simulation results validate the accuracy of the fault location method under different fault conditions including fault types, fault points, and fault resistances. The performance is also validated for non-synchronized measurements and measurement errors.

Keywords: Artificial neural networks, Distributed generation, Distribution networks, Fault location, K-nearest neighbors

1 Introduction

Following the occurrence of a short circuit fault in a distribution network, the restoration process may take from tens of minutes to hours to complete. During fault management, fault location (FL) can serve as an effective tool to narrow down the search area, thereby considerably reducing the inspection and service restoration time. There are a wide variety of FL methods proposed in the literature [1]. These can be classified into impedance-based, travelling wave-based, sparse measurements-based, artificial intelligence-based, and hybrid methods. Although many of the proposed methods provide satisfactory results, artificial intelligence techniques such as the artificial neural network (ANN) and support vector machine (SVM) have gained increased attention because of their capability of solving nonlinear problems and very short execution time. If these methods provide the required accuracy, they can be applied in modern distribution networks with advanced communication, measurement and switching infrastructure for fast fault location and service restoration to reduce the duration of a power outage [2].

Reference [3] employs an SVM to identify the fault type and a feedforward ANN is then applied to estimate the line reactance to the fault for each fault type using the three-phase steady-state voltage and current. In [4], fault voltage waveforms are processed using a wavelet transform to extract wavelet coefficients to feed to a feedforward ANN to estimate the fault distance. In [5], a similar method employing a three-layer feedforward ANN and a fuzzy logic system is used to classify the fault type and estimate its distance.

In [6], high and low frequency components of voltage and current transients are extracted using a wavelet transform, and the components are fed into a fuzzy ANN to locate the single-phase to ground faults. In [7], the current entropy and energy of wavelet detail coefficients are employed as input features of an ANN for faulted section identification and fault location. The method in [8] uses a fuzzy ANN to map the information extracted from the fault voltage to the location of single-phase to ground faults. In [9], frequency components of voltage and current signals are used as inputs to an ANN classifier to locate faults in a small-scale distribution feeder. In [10], the current signals are analyzed to

* Correspondence: sjamali@iust.ac.ir
Centre of Excellence for Power System Automation and Operation, School of Electrical Engineering, Iran University of Science and Technology, P.O. Box: 16846-13114, Tehran, Iran

extract effective features to feed to an adaptive neuro-fuzzy inference system (ANFIS) to estimate the fault zone. The method proposed in [11] uses neural networks for fault location in distribution networks with distributed generation (DG). For each type of fault, a separate multi-layer perceptron ANN is trained to estimate the fault distance from the main substation and all DGs. In [12, 13] the authors propose data mining methods for the protection of microgrids. These methods are effective for online identification and isolation of a faulted line section, but they cannot locate the fault. Similar methods are proposed in [14, 15] to locate the fault, but they require extensive measurement.

Table 1 compares the aforementioned ANN-based fault location methods. Electrical power distribution networks mostly have tree branched structures, and therefore, the methods proposed in [3–5, 8, 9], which estimate the distance to the fault, may result in misidentification of multiple candidate locations having the same distance from the main substation. In [11], the authors propose a method that estimates the distance to the fault from all DGs and main substation to overcome the multiple solution problem. However, considering the error in the estimated distance, it is difficult to correctly interpret the results to find the fault location. The methods in [4–9] use the information of high frequency transients and hence require measurements with high sampling rates. Such equipment is costly and is not available in most distribution systems. In addition, the methods proposed in [4, 6, 8] are designed only for earth faults and most of the discussed FL methods are designed for conventional radial networks [3, 4, 7–10] and are not applicable to active distribution networks.

In order to overcome the limitations discussed above and listed in Table 1, this paper presents a new hybrid (ANN and KNN) FL method for active distribution grids. It is assumed that meters are installed at DG terminals, the microgrid point of connections and the main substation to measure the three-phase voltage and current (synchronized or non-synchronized). These measurements are used as inputs to a single feedforward ANN to estimate the distance to the fault from all measurement points. At the next stage of the proposed method, a k-nearest neighbor classifier is employed to interpret the ANN outputs and estimate the faulted line section and fault location. As shown in Table 1, compared to existing similar methods, the proposed method solves the multiple estimation problem in DG penetrated networks for all types of faults.

The rest of the paper is organized as follows: Section 2 presents the details of the proposed method. The simulation results under ideal and non-ideal conditions, with or without synchronised measurements, are presented in Section 3. Finally, Section 4 concludes the paper and highlights the contributions of the study.

2 Proposed method

This paper presents a fast hybrid FL method that uses measurements as inputs to a feedforward ANN to estimate distances to a fault from all measuring points. Then, a KNN classifier interprets the ANN outputs and estimates the fault location.

Multi-layered feedforward neural networks, also known as multi-layered perceptron (MLP) have one input layer, one output layer and one or more hidden layers. As shown in Fig. 1, the number of neurons in the input and output layers depends on the number of selected inputs and the number of outputs, while the number of neurons in the hidden layers is usually determined by trial and error depending on the complexity of the problem. The neurons of each two concessive layers are interconnected by weighted communication links. The weights represent information being used to solve the problem and are determined by a training algorithm. In this work, the well-known Levenberg-Marquardt (LM) training algorithm is employed.

A KNN classifier is a simple algorithm that classifies new cases based on their similarity to previous cases using a measure (e.g., distance). The new case will be assigned to the most common class among its k nearest neighbors. The appropriate value of k can be chosen using cross-validation to test several values of k in order to determine which one works best.

2.1 The method architecture

Figure 2 shows the architecture of the proposed hybrid method. First, the fault voltage and current signals of all sources are collected and processed by a full cycle discrete Fourier transform (DFT). The magnitudes and phase angles of all the three-phase voltages and currents are then extracted and the three-phase apparent impedances of all sources calculated. The three features are passed to the ANN, which outputs the distances to the fault from all sources. The distances are then fed into the KNN classifier to interpret the ANN outputs and identify a correct class, which has the information of the faulted line-section, fault location and distances to the fault from all sources.

2.2 Selection of the input and output variables

Selection of input variables influences the range of applicability and its success in estimating the fault location. Many of the previously proposed methods, such as those in [4–9], are based on the information extracted from high frequency fault transients using high sampling rates, which are costly and not available in most distribution systems. In this paper, the fundamental frequency

component of three-phase voltage and current from all sources (i.e. main substation, DGs and microgrids) are the only required measurements.

These measurements can be retrieved from the available smart meters, digital relays, or digital fault recorders, as the accessibility of such measurements at source terminals is a primary requirement of DG connection. A communication infrastructure such as the one described in [16] can collect the required measurements.

For measurement scenarios, two cases are considered. In the first case, it is assumed that the measurements are synchronised. Therefore, for n sources, magnitudes and phase angles of three-phase voltage, current and apparent impedances of all sources are selected as ANN input features (i.e. $2 \times 3 \times 3 \times n$ inputs). In the second case, the measurements are not synchronized, and therefore, the magnitudes of three-phase voltage, current and apparent impedances of all sources are employed as ANN input features (i.e. $3 \times 3 \times n$ inputs). The apparent impedance of each source is calculated as:

$$Z_i = \frac{V_i}{I_i} \tag{1}$$

where V_i and I_i are the voltage and current of the ith source, respectively.

The proposed method employs an ANN to estimate the fault distance. Since the ANN-based fault locators mostly select the distance to fault from the main substation as the output [3–5, 8, 9], they may result in misidentification of multiple locations having the same distance from the main substation. In the proposed method, the ANN estimates distances to fault from all sources (i.e. n outputs) to overcome the multiple solution problem. Nevertheless, the ANN outputs are not accurate enough and it is difficult to match the obtained distance to find a specific fault location.

For example, considering the fault in the network of Fig. 3, if the ANN underestimates the distance to fault from the first source (D_1), there are three different points with the same distance, as possible locations.

If the ANN overestimates the distance to the fault from the second source (D_2), there are three other different points with the same distance, as other possible locations. In ideal conditions, one of the points with the estimated distance D_1, matches one of the points with the estimated distance D_2 exactly on the fault location. However, in real-world conditions, as shown in Fig. 3, there will be estimation errors and the estimated distance does not match. Consequently, this will lead to confusion.

Therefore, at the final stage of the proposed method, the use of a KNN classifier is proposed to interpret the

Table 1 Comparison of the ANN-based fault location methods

Method	Type	Inputs	Outputs	Considered fault types	Considering the effect of DG
Thukaram et al. [3]	Feedforward ANN	Three-phase steady-state voltage and current	Line reactance to the fault from the main substation	All types	–
Pourahmadi-Nakhli and Safavi [4]	Feedforward ANN	Energy content of voltage transients around the characteristic frequencies	Fault distance from the main substation	Single-line to ground	–
Rafinia and Moshtagh [5]	Feedforward ANN	Information extracted using DWT and analysis of the recorded transients	Fault distance from the main substation	All types	✓
Chen et al. [6]	Feedforward ANN	Information extracted from voltage signals recorded at different locations	Distance to the fault from all voltage sensors and the faulted line- section	Earth faults	✓
Dashtdar et al. [7]	Feedforward ANN	High frequency components of current signals	Line impedance to the main substation, faulted line-section	All types	–
Chunju et al. [8]	WFNN	High frequency and low-frequency components of voltage and current transients	Fault distance from the main substation	Single-line to ground	–
Aslan and Yagan [9]	Feedforward ANN	High frequency and low-frequency components of voltage and current transients	Fault distance from the main substation	All types	–
Mora et al. [10]	ANFIS	Patterns obtained from measured substation current	Faulted zone	All types	–
Javadian et al. [11]	Feedforward ANN	Ratio of injected fault currents of DGs to substation current	Distance to the fault from all DGs and the main substation	All types	✓
The proposed method	Feedforward ANN	Three-phase steady-state voltages and currents of substation and DGs	Fault location, faulted line-section and distance to the fault from all DGs and the main substation	All types	✓

ANN output to find the fault location and the faulted line-section. In the example of Fig. 3, the KNN classifier receives the estimated distances to the fault from all sources and selects the nearest class with the most similar distances. The KNN outputs are the fault location, faulted line-section, and distances to the fault from all sources in the selected class.

2.3 Data generation

All the ANN-based methods have an offline training phase using the already available data (real fault cases) or generated data (simulated cases).

In this work, for generating training and/or testing data, active and reactive power of all network loads are varied randomly within specified ranges of their corresponding base values, according to the following relations:

$$\begin{aligned} P_i(k) &= P_i^a(1+\delta_{Li}(k)) \\ Q_i(k) &= Q_i^a(1+\delta_{Li}(k)) \end{aligned} \tag{2}$$

where $P_i(k)$ and $Q_i(k)$ are the respective active and reactive power of the ith load for the kth training pattern, whereas P_i^a and Q_i^a denote active and reactive power of the base case load, respectively. δ_{Li} is a randomly generated number from a normal distribution with zero mean and standard deviations of 20%.

For each fault type, faults are simulated at every 200 m of all line-sections, with four different fault resistances (i.e. 1 Ω, 5 Ω, 20 Ω, 50 Ω). As shown in Fig. 4, under random load data, for each of the training or testing data, the considered fault type, fault resistance and fault location are fed to the simulated system.

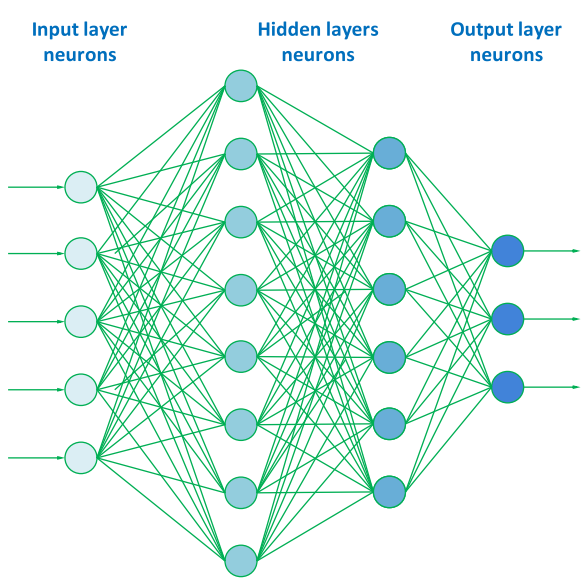

Fig. 1 A typical multi-layered perceptron (MLP) neural network with two hidden layers

For each simulated case, the voltage and current waveforms at all source terminals are recorded and the full-cycle discrete Fourier transform is used to calculate the fundamental phasors. Having the voltage and current phasors, the apparent impedance is calculated using (1) and these features are stored as inputs of each training or testing data pattern. The target output is calculated as the distance to the considered fault location from all sources. This procedure is repeated to generate sufficient numbers of training and/or testing patterns for building the proposed neural network model.

3 Results and discussion

A simplified three-phase symmetrical model of the IEEE 34-node radial test feeder [17] is used to demonstrate the performance of the proposed FL method. As shown in Fig. 5, the test feeder is modified by adding two microgrids and one DG unit at nodes 822, 838 and 848, respectively. The lengths of the line-sections are also illustrated in Fig. 5. The microgrids are modelled according to [18] and the DG units are modelled as a source behind an impedance with unity power factor representing inverter-based generation. The lines are modelled by three-phase pi-equivalents and loads to be of a constant impedance type, whereas the DG units serve 50% of the total system load.

The considered fault scenarios for different fault types, fault resistances and fault locations are simulated on the

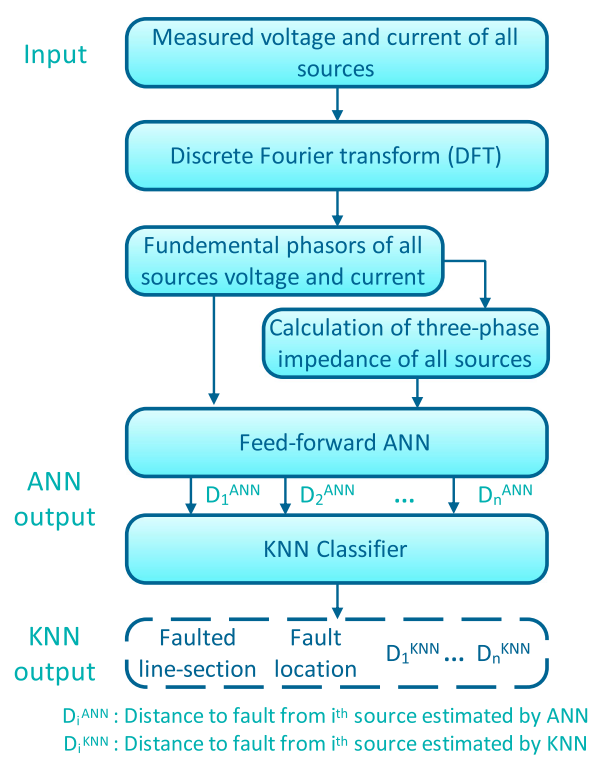

Fig. 2 Architecture of the proposed method

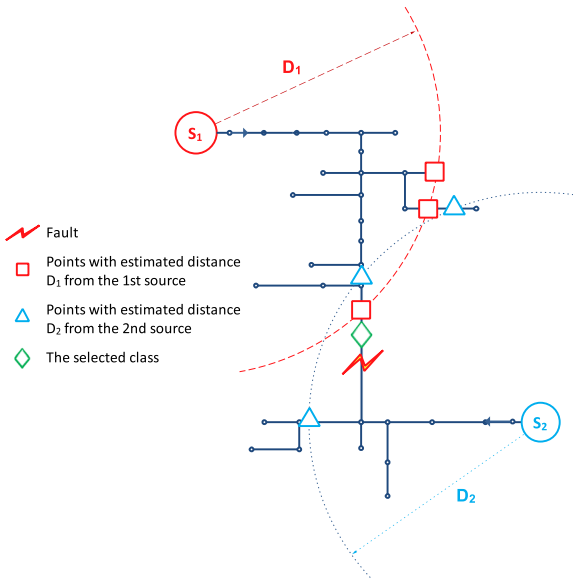

Fig. 3 The difficulty of interpretation of ANN outputs in finding a specific fault location

test system to generate data. Using the procedure described in Section 2.3, 1056 cases are generated for each fault type, simulated at different locations (i.e. every 200 m of all line-sections), with four different fault resistances (i.e. 1 Ω, 5 Ω, 20 Ω, 50 Ω). From the generated cases, 90% are employed as training data and the remaining 10% are employed as testing data. The input features for all the different fault types are normalized within the range of [− 1, 1], and are fed to a single ANN

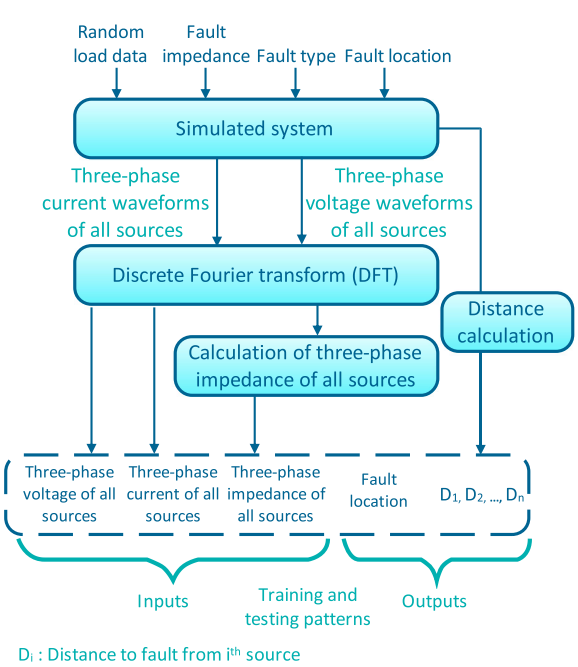

Fig. 4 Generation of training and testing data

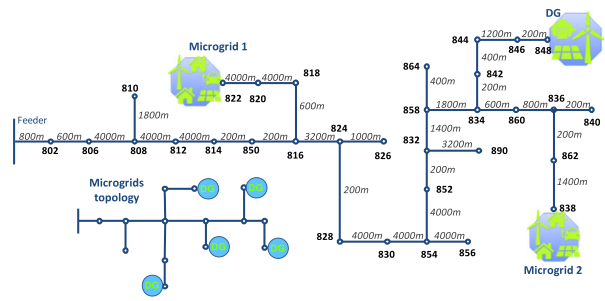

Fig. 5 Topology of the test feeder

to estimate the FL. Table 2 lists the details of the generated data.

In order to test the generalization capability of the proposed hybrid method, the root mean-squared error (*RMSE*) and the maximum absolute error (*MAE*) between the actual FL and the estimated FL are calculated. These indices can be defined as:

$$\begin{aligned} RMSE &= \sqrt{\frac{1}{NP}\sum_{p=1}^{NP}(error(p))^2} \\ MAE &= \max(error) \end{aligned} \tag{3}$$

where p is the pattern number, NP is the total number of patterns, and *error* is the fault distance estimation error in meters calculated as:

$$error(p) = |actual(FL(p)) - estimated(FL(p))| \tag{4}$$

The structure of the neural network is determined by trial and error. It has one input layer, two hidden layers consisting of 15 and 10 neurons each, and one output layer. The input features consist of the magnitudes and phase angles of three-phase voltage, current and apparent impedances of all sources, while the outputs are the distances to the fault from all sources. A hyperbolic tangent transfer function and linear transfer function are used for the hidden layers and the output layer neurons, respectively. The proposed MLP ANN is trained using the Levenberg–Marquardt method. On a personal computer with 2-GHz Intel Core 2 Duo processor and 2 GB of RAM, it takes around 265 s with 21 epochs to train the network. Once the MLP ANN model is trained, the testing data is employed to evaluate the model

Table 2 Training/testing data details

Simulation parameters	Details	Count
Fault type	Phase to ground, phase to phase, phase to phase to ground, three-phase to ground	4
Fault resistance (Ω)	1 Ω, 5 Ω, 20 Ω, 50 Ω	4
Fault location	Each 200 m of all line-sections	264
Total conditions	3800 cases for training and the 424 cases for testing	4224

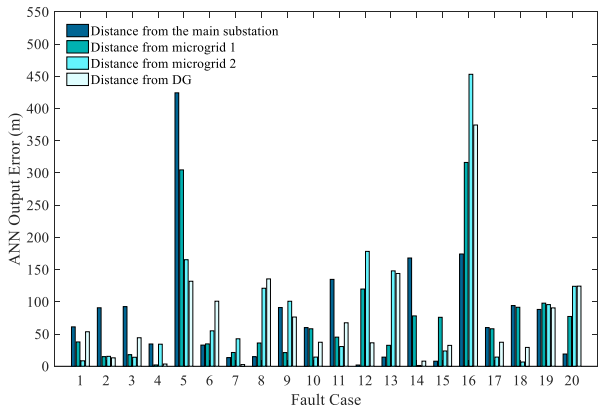

Fig. 6 ANN outputs error of estimated fault distances from all sources, for 20 randomly selected testing data

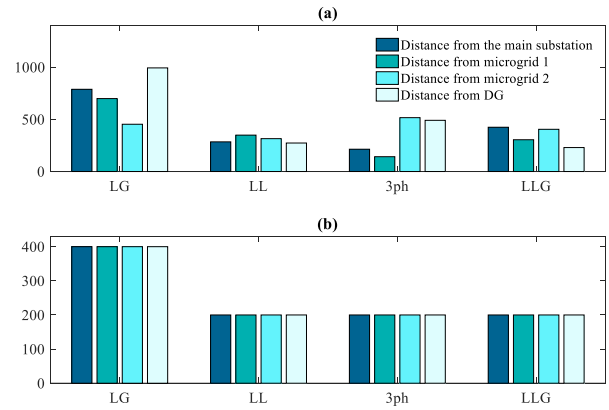

Fig. 8 Output *MAE* values for different fault types for 424 test cases. **a** for ANN outputs **b** for KNN outputs

performance. For KNN, after several trials with different values of k and different distance metrics, k is selected to be equal to 1 and the Euclidean distance metric is selected. When k = 1, a new case is simply assigned to the class of the single nearest neighbor.

3.1 Tests with synchronized measurements

In this section, it is considered that the measurements are synchronised, and the voltage and current signals are provided with magnitudes and phase angles. In this case, the input feature vectors have the dimensions of 3800×72 for the training set and 424×72 for the testing set, respectively. Figure 6 shows the ANN output error between the estimated and the corresponding actual fault distances from all sources, for 20 randomly selected testing data. The accuracy of the results is acceptable, though as discussed in Section 2.2, it is difficult to interpret the output distances of the ANN to find the actual fault location. Therefore, the ANN outputs are passed to a KNN classifier to interpret the results to find the faulted line-section and the fault location. Figure 7 shows the errors of the KNN outputs for the same fault scenarios. It can be seen that the use of the KNN as an interpreter not only helps find the faulted line-section and the fault location, but also refines the accuracy of the results and for most of the selected data, the output error is zero.

3.1.1 Results for different fault types

Figures 8 and 9 show the *MAE* and *RMSE* of ANN and KNN outputs for all testing data. The results indicate that the proposed single MLP ANN is able to estimate the fault distance of all different fault types with acceptable accuracy. However, in some cases, the obtained *MAE* is considerably large. The KNN refines the ANN results and for all testing data, the maximum output error is less than 400 m. The small *RMSE* values of less than 100 m for all the considered test scenarios shown in Fig. 9b, clearly indicate the good performance and acceptable generalization accuracy of the proposed method for different fault

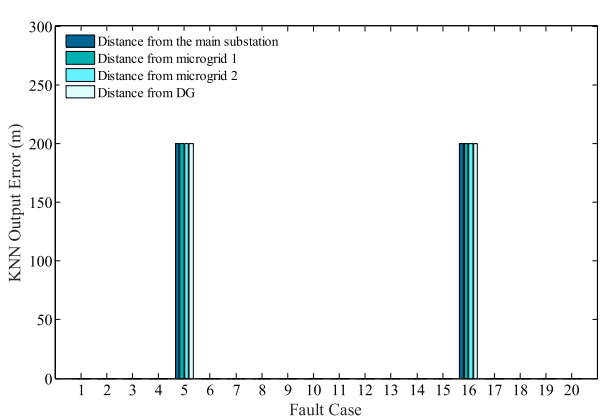

Fig. 7 KNN outputs error of estimated fault distances from all sources, for 20 randomly selected testing data

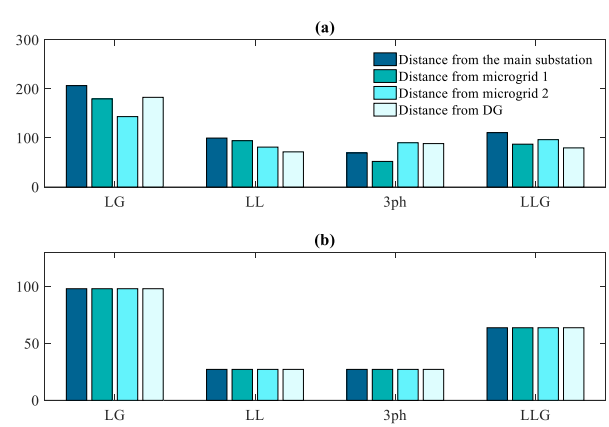

Fig. 9 Output *RMSE* values for different fault types for 424 test cases. **a** for ANN outputs **b** for KNN outputs

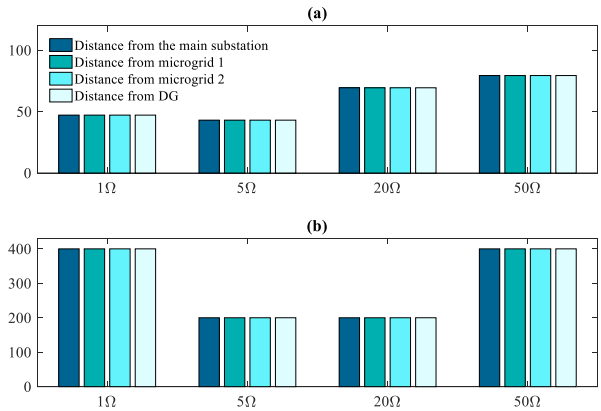

Fig. 10 Output errors of the KNN for different fault resistances for 424 test cases. **a** *RMSE* **b** *MAE*

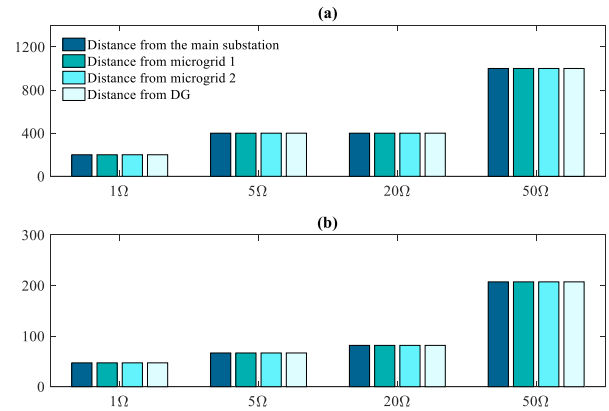

Fig. 12 Output errors of the KNN for different fault resistances using non-synchronized measurements for 424 test cases. **a** *MAE* **b** *RMSE*

types. Also, the proposed method estimates the FL almost instantaneously.

3.1.2 Results for different fault resistances

In addition to fault type, fault resistance is another factor and can change the voltage and current of the sources during the fault, and hence complicates the fault location problem. Figure 10 shows the *RMSE* and *MAE* values of the KNN outputs for different fault resistances. It can be seen that the *RMSE* increases along with the increase of the fault resistance. However, the proposed method has an *RMSE* value of less than 100 m and an *MAE* value of less than 400 m for all the testing cases. In 99.3% of the considered test scenarios, the faulted line-section is correctly identified. In the other 0.7%, the distance between the estimated and the actual fault locations is less than 400 m.

3.2 Tests with non-synchronized measurements

In the previous sections, it was assumed that the measurements are fully synchronised and that they are able to provide the magnitudes and phase angles of three-phase voltage and current of all sources. However, even in some of the recent smart meter deployments, synchronized measurements are not available [16]. In this section, we consider the situation where only the measured magnitudes of three-phase voltage and current are available. For this condition, the input feature vectors have the dimensions of 3800×36 for the training set and 424×36 for the testing set. The structure of the neural network is similar to the previous one, but with 36 inputs in the input layer (i.e. magnitudes of three-phase voltage, current and apparent impedances of all sources). In this case, it took about 385 s with 43 epochs to train the network.

Figures 11 and 12 show the *MAE* and *RMSE* of the KNN outputs for different fault types and fault resistances. It can be seen that even with non-synchronised measurements, the proposed method is able to estimate the fault distance of all different fault types and fault resistances with a single MLP ANN. Compared to the case with synchronised data, the estimation errors are increased, but in 98.84% of the considered test scenarios the faulted line-section is correctly identified.

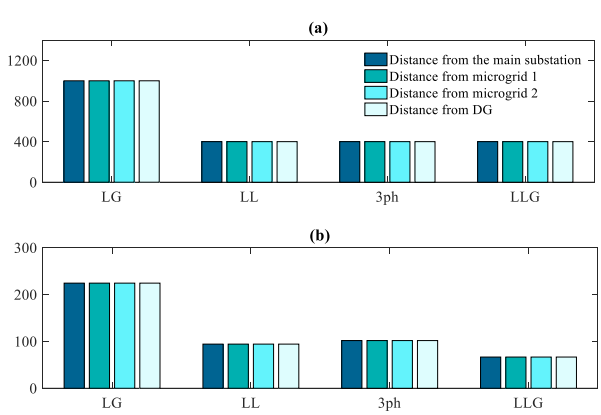

Fig. 11 Output errors of the KNN for different fault types using non-synchronized measurements for 424 test cases. **a** *MAE* **b** *RMSE*

3.3 Tests with measurement errors

Measurements may be influenced by errors due to noise, meter inaccuracy, etc. In order to test the influence of measurement errors on the performance of the proposed method, the magnitudes of measured voltage and current are varied randomly within specified ranges:

$$\begin{aligned} V_{Mi} &= V^{a}_{Mi}(1+\delta_{mi}) \\ I_{Mi} &= I^{a}_{Mi}(1+\delta_{mi}) \end{aligned} \tag{5}$$

where $V^{a}{}_{Mi}$ and $I^{a}{}_{Mi}$ are the actual values of voltage and current of the ith source, respectively, whereas

V_{Mi} and I_{Mi} are the erroneous inputs of the fault locator. δ_{mi} is a randomly generated number from a normal distribution with zero mean and standard deviations of σ_{mi}.

Measurements can be synchronised using the global positioning system (GPS) or a computer network. Network synchronization can provide an accuracy of 4 μs (0.08 degree in 60 Hz) [19], while GPS synchronization can achieve a theoretical accuracy of 1 μs (0.02 degree in 60 Hz) [20]. If the phase angles are employed as the feature to the ANN, the effect of such measurement synchronization errors may also affect the results.

Four cases are considered to assess the accuracy of the proposed method under measurement errors:

Case 1) Random variation of the magnitudes of the measured voltage and current within 1%;
Case 2) Random variation of the magnitudes of the measured voltage and current within 2%;
Case 3) Random variation of the phase angles of the measured voltage and current within 0.5 degree;
Case 4) Random variation of the magnitudes and phase angles of the measured voltage and current within 2% and 0.5 degree, respectively.

The test results for all cases are compared against the ideal condition without measurement errors, as summarised in Table. 3. When measurements are synchronised, the proposed method can well handle the measurement errors and in the worst case (Case 4), more than 96.99% of the faulted line-sections are correctly identified. When non-synchronized measurements are employed, random variation of measured magnitudes has a more considerable impact on the estimated distance, though synchronization errors do not affect the fault locator performance. Comparing Case 3 with synchronized measurements and the ideal condition with non-synchronised measurements reveals that even poorly synchronized measurements can help to provide more accurate results. Overall, the proposed method shows a high generalization potential, which helps to correctly identify the faulted line section in the presence of measurement errors.

4 Conclusions

As a practical function in distribution system automation, accurate fault location can lead to fast service restoration following a fault. However, the presence of DGs and microgrids in active distribution grids makes fault location a challenging problem. In this paper, the use of a single feedforward ANN in combination with a KNN classifier (a hybrid method) is proposed to estimate the fault location. The simulation results show that the method provides accurate results for different fault types, fault locations and fault resistances, with or without measurement synchronization. Compared to previously proposed methods, the main contributions are as follows:

1. The methods in [3–5, 8, 9] which estimate the distance to fault may result in misidentification of multiple locations with the same distance, whereas the proposed ANN estimates distances to the fault from all the sources and the KNN classifier interprets the ANN outputs to provide a single fault location candidate;
2. Many of the previously proposed methods, such as [4–9], require measurements with high sampling

Table 3 RMSE and MAE values and the Percentage of the correctly identified line-sections for all test scenarios

Performance measures	*RMSE* (m)				*MAE* (m)				Percentage of the correctly identified line-sections
Outputs	D1[1]	D2[2]	D3[3]	D4[4]	D1	D2	D3	D4	
Cases	**With synchronised measurements**								
Ideal	77	77	77	77	400	400	400	400	99.31%
Case 1	111	111	102	104	1000	400	400	600	98.84%
Case 2	163	109	108	109	2000	400	400	400	98.38%
Case 3	81	81	79	81	400	400	400	400	99.31%
Case 4	174	178	159	178	1800	1800	1800	1800	96.99%
	With non-synchronised measurements								
Ideal	115	115	113	115	1800	1800	1800	1800	98.84%
Case 1	178	178	166	166	2200	2200	2200	2200	98.61%
Case 2	222	222	253	253	2200	2200	2600	2600	97.91%
Case 3	115	115	113	115	1800	1800	1800	1800	98.84%
Case 4	347	334	291	292	4000	4000	3600	3600	95.6%

1. Distance from the main substation, 2. Distance from microgrid 1, 3. Distance from microgrid 2, 4. Distance from DG

rates. In contrast, the proposed method uses lower sampling rates;

3. The proposed method employs a single ANN for all different fault types and thus does not require the fault type identification process, which is a source of error;
4. Distribution level meters are prone to errors caused by meter inaccuracy, noise etc., while their effect on fault location has not been reported in previous artificial intelligence-based distribution FL methods. The simulation results show that the proposed hybrid method provides acceptable results even with less accurate measurements;
5. In distribution systems, even with smart meter installation, the available load data is inaccurate because of their low reporting rate. In this paper, in contrast to previous work, the load data uncertainty is considered with 20% random error in the training phase and thus the ANN can provide acceptable results despite such inaccuracy.

Nevertheless, the main limitation of all learning-based approaches is the need for retraining following changes in distribution network topology or DG connection status. Moreover, while the proposed method is designed to provide acceptable results with 20% load variation, it may not provide satisfactory results under larger load variations. These limitations can be to some extent overcome by training an ANN for each planned topology, DG status and load level.

Abbreviations

ANFIS: Adaptive neuro-fuzzy inference system; ANN: Artificial neural network; DFT: Discrete Fourier transform; DG: Distributed generation; FL: Fault location; GPS: Global positioning system; KNN: K-nearest neighbors; LM: Levenberg-Marquardt; MAE: Maximum absolute error; MLP: Multi-layered perceptron; RMSE: Root mean-squared error; SVM: Support vector machine

Acknowledgements

Not applicable.

Authors' contributions

S. Jamali supervised the research study, the method and the manuscript preparation. A. Bahmanyar developed the proposed method and prepared the manuscript. S. Ranjbar contributed in method development, data analytics, data interpretation and manuscript preparation. All authors read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

The datasets generated and/or analysed during the current study are not publicly available but can be requested from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Received: 6 January 2020 Accepted: 2 July 2020
Published online: 26 August 2020

References

1. Bahmanyar, A., Jamali, S., Estebsari, A., & Bompard, E. (2017). A comparison framework for distribution system outage and fault location methods. *Electric Power Systems Research*, *145*, 19–34.
2. Bahmanyar, A., Estebsari, A., Pons, E., Patti, E., Jamali, S., Bompard, E., & Acquaviva, A. (2016). *Fast fault location for fast restoration of smart electrical distribution grids*, *IEEE international smart cities conference (ISC2)* (pp. 1–6).
3. Thukaram, D., Khincha, H., & Vijaynarasimha, H. (2005). Artificial neural network and support vector machine approach for locating faults in radial distribution systems. *IEEE Transactions on Power Delivery*, *20*, 710–721.
4. Pourahmadi-Nakhli, M., & Safavi, A. A. (2011). Path characteristic frequency-based fault locating in radial distribution systems using wavelets and neural networks. *IEEE Transactions on Power Delivery*, *26*, 772–781.
5. Rafinia, A., & Moshtagh, J. (2014). A new approach to fault location in three-phase underground distribution system using combination of wavelet analysis with ANN and FLS. *International Journal of Electrical Power & Energy Systems*, *55*, 261–274.
6. Chen, X., Yin, X., & Deng, S. (2017). A novel method for SLG fault location in power distribution system using time lag of travelling wave components. *IEEJ Transactions on Electrical and Electronic Engineering*, *12*, 45–54.
7. Dashtdar, M., Dashti, R., & Shaker, H. (2018). *Distribution network fault section identification and fault location using artificial neural network*, (pp. 273–278). Istanbul: International Conference on Electrical and Electronic Engineering (ICEEE).
8. Chunju, F., Li, K., Chan, W., Weiyong, Y., & Zhaoning, Z. (2007). Application of wavelet fuzzy neural network in locating single line to ground fault (SLG) in distribution lines. *International Journal of Electrical Power & Energy Systems*, *29*, 497–503.
9. Aslan, Y., & Yagan, Y. (2017). Artificial neural-network-based fault location for power distribution lines using the frequency spectra of fault data. *Electrical Engineering*, *99*, 301–311.
10. Mora, J., Carrillo, G., & Perez, L. (2006). *Fault location in power distribution systems using ANFIS nets and current patterns*, (pp. 1–6). Latin America: IEEE/PES Transmission & Distribution Conference and Exposition.
11. Javadian, S., Nasrabadi, A., Haghifam, M.-R., & Rezvantalab, J. (2009). *Determining fault's type and accurate location in distribution systems with DG using MLP neural networks*, *IEEE international conference on clean electrical power* (pp. 284–289).
12. Manohar, M., Koley, E., & Ghosh, S. (2018). Reliable protection scheme for PV integrated microgrid using an ensemble classifier approach with realtime validation. *IET Science, Measurement and Technology*, *12*, 200–208.
13. Jamali, S., Ranjbar, S., & Bahmanyar, A. (2020). *Identification of faulted line section in microgrids using data mining method based on feature discretisation*, *International transactions on electrical energy systems* (pp. 1–16).
14. Dharmapandit, O., Patnaik, R. K., & Dash, P. K. (2017). A fast time-frequency response based differential spectral energy protection of AC microgrids including fault location. *Protection and Control of Modern Power Systems*, *2*, 1–28.
15. Zhang, F., & Mu, L. (2019). New protection scheme for internal fault of multi-microgrid. *Protection and Control of Modern Power Systems*, *4*, 1–12.
16. Bahmanyar, A., Estebsari, A., Pons, E., Jamali, S., Patti, E., Bompard, E., & Acquaviva, A. (2016). *Emerging smart meters in electrical distribution systems: opportunities and challenges*, *24th Iranian conference on electrical engineering (ICEE)* (pp. 1082–1087).
17. IEEE PES Radial distribution test feeders. 2017. Available: http://www.site.ieee.org/pes-testfeeders/resources/.
18. Papathanassiou, S., Hatziargyriou, N., & Strunz, K. (2005). *A benchmark low voltage microgrid network*, *Proceedings of the CIGRE symposium: power systems with dispersed generation* (pp. 1–8).
19. Johannessen, S. (2004). Time synchronization in a local area network. *IEEE Control Systems*, *24*, 61–69.
20. IEEE Standard for Synchrophasor Measurements for Power Systems, (2011). IEEE Std C37.118.1-2011 (Revision of IEEE Std C37.118-2005), pp.1-61, 28 Dec. 2011, https://doi.org/10.1109/IEEESTD.2011.6111219.