SAND2022-5080C

Asynchronous Traveling Wave-based Distribution System Protection with Graph Neural Networks

Miguel Jiménez-Aparicio
Electric Power Systems Research
Sandia National Laboratories
Albuquerque, USA
mjimene@sandia.gov

Matthew J. Reno
Electric Power Systems Research
Sandia National Laboratories
Albuquerque, USA
mjreno@sandia.gov

Felipe Wilches-Bernal
Electric Power Systems Research
Sandia National Laboratories
Albuquerque, USA
fwilche@sandia.gov

***Abstract*—The paper proposes an implementation of Graph Neural Networks (GNNs) for distribution power system Traveling Wave (TW) - based protection schemes. Simulated faults on the IEEE 34 system are processed by using the Karrenbauer Transform and the Stationary Wavelet Transform (SWT), and the energy of the resulting signals is calculated using the Parseval's Energy Theorem. This data is used to train Graph Convolutional Networks (GCNs) to perform fault zone location. Several levels of measurement noise are considered for comparison. The results show outstanding performance, more than 90% for the most developed models, and outline a fast, reliable, asynchronous and distributed protection scheme for distribution level networks.**

***Index Terms*—Power System Protection, Traveling Waves, Distribution Systems, Graph Neural Networks, Stationary Wavelet Transform**

I. Introduction

Promising changes are occurring in power system protection. Numerous studies that apply Traveling Wave (TW) - based methods to distribution systems will make ultra-fast protection a reality at the distribution level [1]. These approaches are usually based on Machine Learning/Deep Learning methods due to the complexity of the task: the TW propagation in distribution systems is heavily affected by the propagation path characteristics, such as line lengths and elements present in the system [2].

The application of Graph Neural Networks (GNNs) is a relatively new topic in power systems protection, but it is already showing impressive results in comparison to previously presented Machine Learning (ML) or Deep Learning (DL) methods in many different tasks [3]. These advancements are only the beginning of a new generation of GNNs-based methods, which will continue to improve practical implementations of ML/DL protection schemes.

One of the reasons for this success is that GNNs not only use node measurements, but they also give a spatial relationship between nodes. Besides a potential superior performance, GNNs provide other desirable properties to protection schemes: GNNs assume a distributed communication system between the nodes. In the most common type of GNNs, the Graph Convolutional Networks (GCNs), the nodes update their information with data coming from their immediate neighbors. Distributed systems are resilient to system failures and cyber-attacks, as a significant number of nodes would have to be incapacitated in order to pose a risk to the system [4].

The remaining of the paper is organized as follows. Section II introduces recent fault location methods based on GNNs and some references that use TWs to perform fault location on power distribution systems. Section III explains the theory of GCNs. Next, Section IV describes the fault simulation procedure, the signal-processing stage, and the proposed GCN models. Section V gathers the accuracy results for the considered models with and without measurement noise. Section VI includes a brief discussion about the results shown by other methods on a similar task and the proposed method. Section VII provides some insights about how the work in this paper could be expanded. Finally, the conclusions of this paper are gathered in Section VIII.

II. Background

The usage of GNNs in power systems is a very recent topic. Reference [3] provides a recent review of recent developments in that area. For power system protection only a handful of approaches have been proposed. In [5], GCNs are used to predict fault location on the IEEE 123 bus test case, taking into account measurement noise and changes in topology. The input features are the voltage and current phasors. The GCNs achieve an accuracy close to 90%, with a one-hop accuracy of more than 95%. This level of accuracy cannot be reached by other ML methods that are included for comparison. Similarly, in [6], voltage and current phasors are used as the node features, while the admittance matrix is used to define the edge features. In [7], a fault diagnosis method for the shipboard power system is presented. The employed data time window

This material is based upon work supported by the Laboratory Directed Research and Development program at Sandia National Laboratories and the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE) under Solar Energy Technologies Office (SETO) Agreement Number 36533.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

978-1-6654-6591-5/22/$31.00 ©2022 IEEE

is 5 seconds. The reported accuracy for fault type and location prediction is greater than 99%. The work in [8] proposes an alternative representation of the adjacency matrix to show the correlation between historical data and current samples. In addition, it gives some practical insights about the most suitable optimization algorithm for training, and about the length of the GCNs. Unsurprisingly, the GCNs outperforms all the other methods in terms of accuracy.

In summary, GNNs offer new modeling capabilities that include spatial relationships, which have been reported as beneficial in terms of model accuracy and generalization in many areas of power systems. To the best of the authors' knowledge, there aren't any GNNs approaches on Traveling Wave (TW)-based power system protection. The TWs phenomena occurs in the first microseconds after the fault inception. These wide-band waves propagate through the system at nearly the speed of light and are attenuated and distorted due to the line impedance and discontinuities in the system [2]. There is extensive research on how to use the information contained in those waves for fault location and classification purposes. For example, the works in [9], [10] aim to find the fault location using time differences between TW arrivals. On the other hand, the works in [11]–[13] develop advanced ML-based method for fault location using Random Forest. A similar task is performed in [14]–[16] employing Convolutional Neural Networks (CNNs) instead. These papers report high accuracy, around 90%, for asynchronous DL-based protection on the IEEE 34 nodes system.

III. The Method

A. The Fault Location Task

The test case is the IEEE 34 nodes system, which is divided into 11 protection zones, considering laterals as individual protection zones. The main backbone is divided into three smaller zones using the voltage regulator as a boundary, as shown in Fig. 1. The goal of the approach is to be able to identify which protection zone must be isolated after a fault occurs, which is treated as a supervised classification problem.

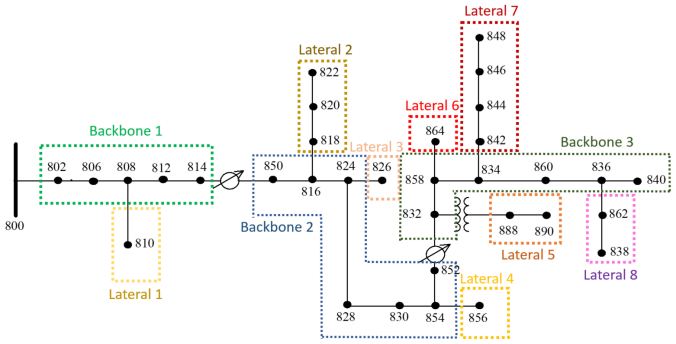

Fig. 1. Proposed protection zone distribution

Note that a power system is, at the most basic level, a set of nodes interconnected by lines (in graph theory, they correspond to the "edges"). As it was mentioned before, GCNs are able to provide spatial relationships to the measurements. In this application, the node features are the 3-Phase (3P)/ 1-Phase (1P) voltage and current measurements. The edge feature is just the line length, as the distance is one of the most critical factors in TW propagation.

B. Graph Neural Networks

The proposed method is based on GCNs, which are able to provide a spatial relationship to the voltage and current measurements. The features will form a matrix X of size $n \times d$, where n is the number of nodes in the system and d is the length of the time series per node. The physical connections between the nodes is given by the adjacency matrix A, which has a size of $n \times n$ [17]. This matrix can be either binary or weighted. If binary, as it is in this paper, the element A_{ij} is 1 if the nodes i and j are connected, and it is 0 otherwise. A graph-based neural network model $f(X, A)$ takes both node features matrix X and adjacency A as a minimal input. State-of-the-art spatial GCNs are a sequence of the following layer-wise propagation rule, as explained in [17]:

$$H^{(l+1)} = \sigma(b^{(l)} + D^{-\frac{1}{2}} A' D^{-\frac{1}{2}} H^{(l)} W^{(l)}), \tag{1}$$

where $A' = A + I$ is the modified adjacency matrix A considering the self-connections adding the identity matrix I, $b^{(l)}$ is the optional layer-specific trainable bias term, $W^{(l)}$ is the layer-specific trainable weight matrix, σ is the activation function, D is the diagonal node degree matrix of the adjacency matrix A', which is defined as:

$$D_{ii} = \sum j A'_{ij}, \tag{2}$$

and $H^{(l)}$ is the output of the l^{th} layer. Note that in the first layer $H^{(0)} = X$. A GCN layer updates the node data, but not the graph connectivity. The input and output graphs are defined by the same adjacency matrix [18]. The multiplication between the feature matrix $H^{(l)}$ and the modified adjacency matrix A' is an efficient way of message-passing between the different nodes of the system: The state of a node is only updated with the self-state and the state of the immediately connected nodes. This propagation rule effectively maps the information contained in a structured graph into a fixed-length matrix that can be multiplied by a set of trainable weights during the forward propagation stage.

If a weight is assigned for each edge, the formula is modified to:

$$H_i^{(l+1)} = \sigma\left(b^{(l)} + \sum_{j \in \mathcal{N}(i)} \frac{e_{ji}}{c_{ji}} H_j^{(l)} W^{(l)}\right), \tag{3}$$

where $H_i^{(l)}$ is the output of the l^{th} layer for node i, c_{ji} is the product of the square root of node degrees and e_{ji} is the scalar weight on the edge from node j to node i. Note that both $b^{(l)}$

and $W^{(l)}$ are the same for all the nodes. Finally, the edge weights can be normalized applying the following expression:

$$c_{ji} = \sqrt{\sum_{k \in \mathcal{N}(j)} e_{jk}} \sqrt{\sum_{k \in \mathcal{N}(i)} e_{ki}}. \quad (4)$$

For training purposes, Eq. (3) is implemented as Eq. (1) with a weighted multiplication to incorporate the edge weights.

IV. The Use Case

A. Fault Simulation

The dataset is composed of 2034 fault cases. The fault conditions depend on whether the node is 3P or 1P. There is only one fault in the system for each case. The system is in a steady state when the fault occurs. For 3P nodes, the conditions are as follows:

- Two fault inception angles: 1 millisecond apart.
- Three fault resistance values: 0.01, 1, and 10 ohms.
- Eleven fault types: 3 Single-Line-to-Ground (SLG) faults, 3 Line-to-Line (LL) faults, 3 Line-to-Line-to-Ground (LLG) faults, 3-Phase fault (3P) and 3-Phase-to-Ground (3PG) fault.

This sums up a total of 66 faults per 3P node. For 1P nodes:

- Nine inception angles: Same as for the 3P case, plus extra seven inception angles equally distributed in a period of 13 milliseconds.
- Six fault resistance values: 0.01, 0.1, 1, 2, 5 and 10 ohms.
- One fault type: From the corresponding phase to ground (SLG).

This yields a total of 48 faults cases for each 1P node. Splitting on an 80/20% training/testing dataset, the total number of faults per protection zone, as defined in Fig. 1, is shown in Table I. A learning rate scheduler is employed to modify the learning during the learning process. Part of the training data acts as a validation set in which the model learning is tested in every epoch. If the validation loss has not improved in 4 consecutive epochs, the current learning rate is multiplied by 0.9. This procedure achieves smoother training. The Adam optimizer is selected, and the initial learning rate is 0.003. In order to get a sense of the model variance to training data, a 10-fold cross validation is applied to the training set. The portion of data that is not used in each iteration is the validation set.

TABLE I
Training, Validation and Testing Sets

Protection Zone	Training Set	Validation Set	Testing Set
Backbone 1	238	26	66
Backbone 2	333	36	93
Backbone 3	285	31	80
Lateral 1	35	3	10
Lateral 2	104	11	29
Lateral 3	35	3	10
Lateral 4	35	3	10
Lateral 5	95	10	27
Lateral 6	35	3	10
Lateral 7	190	21	53
Lateral 8	82	9	23

B. Pre-processing Stage

Similar to [11], the measured current and voltages undergo a pre-processing stage that transforms the raw measurements into more useful data that can be fed into the models. The following steps are performed independently in each node and in an asynchronous way.

1) Given the raw voltage and current measurements (either 3P or 1P depending on the node), the first step is to detect the TW arrival timestamp. Once this is achieved, the signal is cropped $\pm$ 50 μs. Note that the 3P measurements are transformed into the ground mode by the Karrenbauer Transformation (KT). For both 3P and 1P nodes, the voltage and current measurements are reduced to arrays of 1,000 samples each, which correspond to 100 μs at a sampling rate of 10 MHz.
2) Voltage and current arrays are decomposed into 6 frequency bands using the SWT. Note that the resulting coefficients of the SWT are reconstructed to return voltage and current magnitude again. The frequency bands can be observed in Table II. The output of this step is a total of 12 arrays of length 1,000 samples.
3) The energy of each of the signals along the 100 μs window is computed using Parseval's Energy Theorem. Therefore, after this step all the signals are positive, monotonically increasing and magnitude jumps are amplified.
4) Finally, the 12 energy arrays of length 1,000 samples are concatenated and downsampled by 10 (for the sake of computation speed) to form a final array of length 1,200 samples.

TABLE II
SWT Boundaries for Frequency Bands

Decomposition Level	Lower Frequency	Upper Frequency
1	2.5 MHz	5 MHz
2	1.25 MHz	2.5 MHz
3	625 kHz	1.25 MHz
4	312.5 kHz	625 kHz
5	156.25 kHz	312.5 kHz
6	78.125 kHz	156.25 kHz

C. The Graph Neural Networks Models

A total of four GCN model versions are considered, varying the number of GCN layers they contain. Hereafter, models with 1 to 4 GCN layers will be labeled with the letters "A" to "D", respectively. The layer's size can be observed in Table III. All models end with a LogSoftmax layer that returns the log-probabilities that the fault is in each protection zone. As an example, the structure of model "C" can be observed in Fig. 2.

Note that, by computing Eq. (1), the output of the model is a set of 34 predictions, one per node, each one of size equal to the number of protection zones. In GCNs, each node is able to give a predicted fault location based on its own measurements and some messages coming from neighboring

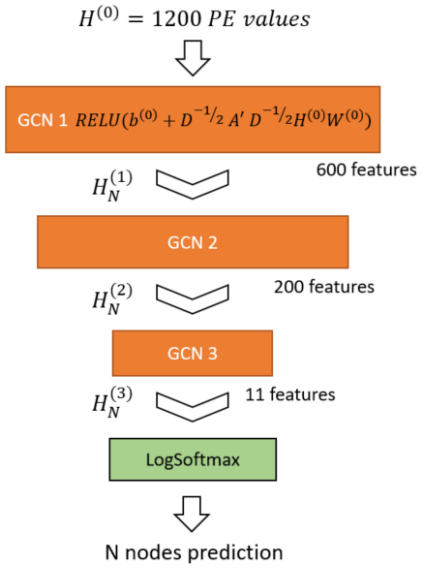

Fig. 2. Model with 3 GCN layers

TABLE III
GCN LAYERS' PARAMETERS SIZE

Model	GCN layer 1	GCN layer 2	GCN layer 3	GCN layer 4
A	1200-by-11	-	-	-
B	1200-by-300	300-by-11	-	-
C	1200-by-600	600-by-200	200-by-11	-
D	1200-by-1000	1000-by-500	500-by-200	200-by-11

nodes. For training purposes, all the nodes' predictions are considered at the same time (i.e., for a specific fault case, all the nodes are expected to predict the correct fault location independently of their position in the system). The loss metric to evaluate the model performance is cross-entropy.

V. RESULTS

This section studies the performance considering nodes as individual protection devices (a "distributed" protection scheme), and considering the system as a quorum in which the output is the average of all the individual node predictions (a "centralized" protection scheme). This last approach gives a single prediction for the model, which is similar to what traditional ML/DL approaches return. Results are shown for cases with and without measurements noise. Two Signal-to-Noise-Ratios (SNRs) of 60 and 45 dB have been used.

Table IV shows the accuracy results without measurement noise of nodes that individually give a correct prediction for the 10-fold cross-validation trained models on the training, validation and testing sets. In general, for models "C" and "D", more than 90% of all the nodes are accurate considering all the fault cases. The maximum accuracy on the testing set is the criteria for selecting the best model, which is model "D". For this model, one of the 10-fold models achieves an accuracy of 94.38%, which means that in average that percentage of nodes correctly identify the faulty protection zone for each fault case. This outstanding performance can be explained by the communication between nodes, that is able to pass information through several layers of neighboring nodes. In the case of models "A" and "B", the lower amount of communication brings down the accuracy.

TABLE IV
MEAN (STANDARD DEVIATION) ACCURACIES FOR DISTRIBUTED APPROACH WITHOUT MEASUREMENT NOISE

Model	Training (%)	Validation (%)	Testing (%)	Testing Max. (%)
A	33.24 (0.19)	32.94 (1.21)	32.99 (0.247)	33.41
B	57.75 (1.69)	56.61 (2.14)	54.59 (1.57)	58.42
C	88.39 (2.23)	85.58 (2.92)	84.14 (2.34)	88.01
D	**95.65 (2.65)**	**93.09 (2.96)**	**91.51 (3.18)**	**94.38**

Tables V and VI display the same results as Table IV but for SNRs of 60 and 45 dB, respectively. As it can be seen, the best GCN model is model "D" in all the cases. Its maximum accuracy decreases slightly down to 93.40% in a small measurement noise scenario of SNR equal to 60 dB. For a more significant measurement noise, the maximum accuracy is 84.42%. This is a good result taking into account that just ± 50 μs of data are used. Models "A" to "C" exhibit similar results.

TABLE V
MEAN (STANDARD DEVIATION) ACCURACIES FOR DISTRIBUTED APPROACH WITH SNR = 60 DB

Model	Training (%)	Validation (%)	Testing (%)	Testing Max. (%)
A	33.42 (0.38)	32.98 (1.08)	33.05 (0.38)	33.53
B	61.71 (1.82)	60.60 (2.09)	58.96 (1.82)	62.38
C	89.73 (2.29)	85.06 (3.07)	83.63 (2.54)	87.84
D	**95.92 (2.25)**	**90.68 (2.67)**	**89.50 (2.54)**	**93.40**

TABLE VI
MEAN (STANDARD DEVIATION) ACCURACIES FOR DISTRIBUTED APPROACH WITH SNR = 45 DB

Model	Training (%)	Validation (%)	Testing (%)	Testing Max. (%)
A	32.99 (0.19)	32.42 (0.97)	32.31 (0.22)	32.63
B	48.52 (1.93)	47.30 (1.63)	45.98 (1.59)	48.05
C	82.22 (1.31)	76.15 (2.20)	75.35 (1.53)	78.47
D	**94.21 (0.56)**	**84.33 (1.25)**	**83.71 (0.39)**	**84.42**

When averaging the nodes' prediction to return a single prediction for each fault case, similar insights can be extracted from Table VII. Note that accuracies are higher because it is easier for a group of nodes to give the correct prediction, rather than just a single node. It is especially significant in the case of model "D", which has a maximum accuracy of 99.51% on the testing set. In addition, it can be observed in the averaged outcome that the standard deviation of the 10-fold cross validation models is significantly lower than when taking into account individual nodes.

As could be expected, the centralized protection scheme that averages all the nodes prediction is also affected negatively by measurement noise. The results are presented in Tables VIII and IX. However, the effect is much less noticeable. For

TABLE VII
MEAN (STANDARD DEVIATION) ACCURACIES FOR CENTRALIZED APPROACH WITHOUT MEASUREMENT NOISE

Model	Training (%)	Validation (%)	Testing (%)	Testing Max. (%)
A	72.94 (1.68)	72.43 (3.67)	71.65 (1.51)	73.47
B	84.78 (0.94)	84.61 (1.76)	82.31 (1.49)	85.82
C	99.91 (0.08)	99.07 (0.99)	98.51 (0.29)	99.02
D	**100.00 (0.00)**	**99.39 (0.94)**	**98.88 (0.54)**	**99.51**

TABLE VIII
MEAN (STANDARD DEVIATION) ACCURACIES FOR CENTRALIZED APPROACH WITH SNR = 60 DB

Model	Training (%)	Validation (%)	Testing (%)	Testing Max. (%)
A	74.02 (1.42)	73.46 (3.95)	71.75 (1.94)	74.20
B	91.18 (2.36)	90.02 (2.82)	88.80 (2.90)	91.97
C	99.94 (0.03)	98.90 (1.04)	98.63 (0.47)	99.27
D	**100.00 (0.00)**	**99.38 (0.55)**	**98.97 (0.28)**	**99.27**

TABLE IX
MEAN (STANDARD DEVIATION) ACCURACIES FOR CENTRALIZED APPROACH WITH SNR = 45 DB

Model	Training (%)	Validation (%)	Testing (%)	Testing Max. (%)
A	72.80 (1.63)	71.98 (2.72)	70.90 (1.37)	73.23
B	81.52 (3.12)	78.95 (3.84)	78.85 (2.90)	82.23
C	99.91 (0.04)	98.52 (1.00)	98.41 (0.39)	99.02
D	**100.00 (0.00)**	**98.84 (1.14)**	**98.61 (0.53)**	**99.51**

model "D", this is reduced to an accuracy drop of about 1% in training, validation, and testing accuracies. However, the maximum testing accuracy remains above 99% in all the cases. Similar insights can be extracted from the other models. This shows the strong performance of GNNs and the benefits of quorum.

VI. DISCUSSION

Other previous works have already studied the problem of fault location in the IEEE 34 nodes system. Even if the task is not identical (those works aim to find the node instead of the protection zone), some insights can be taken for comparison. First, in [16], accuracies for no-noise measurements are around 90% (between 80 to 95% depending on the node), which is just slightly below the 94.38% achieved by the GCNs. However, when noise is present, the accuracy drops significantly down to 70% in average (some nodes even report a 40 or a 60% accuracy). Using GCNs boosts the performance of all the nodes in such a way that, even in scenarios with noise, up to 84% of the individual nodes are able to give an accurate prediction. The work in [15] reports a similar performance, but it uses ± 0.5 ms of data instead.

Generally speaking, most of the DL approaches in the literature don't require a distributed implementation, and using data from all nodes simultaneously is a common practice. The reported results (in this case, averaging the output in order to give a single prediction) are superior to those shown in [14] for a similar task. Finally, the results of this work are aligned with those of other GCN implementations, such as those in [5] and [7], with the difference that the amount of information used in our method is hundreds of times smaller.

VII. FUTURE WORK

The implementation of GNNs in power system protection is a growing topic. Some papers already showed the superiority of this method in comparison to previously developed approaches. Future work in TW protection with GNNs could include a complete accuracy comparison with other methods. Second, in this work it is supposed that the required communication works without any problem. A reliability study could be conducted studying the accuracy under tampered communications.

In addition, in this work only up to four GCN layers have been considered. An optimization study considering a wider number of layers could be performed. Finally, in the considered distribution system, there are many nodes that are located close in distance to their neighbors. A certain aggregation analysis could be carried out to study the drop in performance when several close-by nodes are summarized in just a single node.

VIII. CONCLUSION

The method proposed in this paper, based on the usage of Graph Convolutional Networks (GCNs), achieves an outstanding performance using Traveling Waves (TWs) to locate where faults happened on the distribution system. Results show that notably high accuracies can be achieved both on a node-wise or system-wide basis ("distributed" and "centralized" protection schemes, respectively). The models are tested on fault cases with and without measurement noise.

The model with 4 GCN layers achieves a 94.38% and 99.51% accuracy without measurement noise on the testing set for the "distributed" and "centralized" approaches, respectively. Models with fewer GCN layers achieve lower accuracies. In the case of Signal-to-Noise-Ratio (SNR) of 60 dB, these accuracy values remain close: 93.40% and 99.27%. For a SNR equal to 45 dB, the "distributed" approach performance decreases down to 84.42%, while the "centralized" approach reports an accuracy of 99.51%. Note that fault detection is handled separately with 100% accuracy, so these accuracies for determining the location of the fault are very high.

REFERENCES

[1] F. Wilches-Bernal, A. Bidram, M. J. Reno, J. Hernandez-Alvidrez, P. Barba, B. Reimer, R. Montoya, C. Carr, and O. Lavrova, "A Survey of Traveling Wave Protection Schemes in Electric Power Systems," *IEEE Access*, vol. 9, pp. 72949–72969, 2021.

[2] M. Jiménez-Aparicio, M. J. Reno, and F. Wilches-Bernal, "Traveling Wave Energy Analysis of Faults on Power Distribution Systems," *Energies*, vol. 15, no. 8, 2022.

[3] W. Liao, B. Bak-Jensen, J. R. Pillai, Y. Wang, and Y. Wang, "A Review of Graph Neural Networks and Their Applications in Power Systems," 2021.

[4] A. Helmrich, S. Markolf, R. Li, T. Carvalhaes, Y. Kim, E. Bondank, M. Natarajan, N. Ahmad, and M. Chester, "Centralization and decentralization for resilient infrastructure and complexity," *Environmental Research: Infrastructure and Sustainability*, vol. 1, p. 021001, sep 2021.

[5] K. Chen, J. Hu, Y. Zhang, Z. Yu, and J. He, "Fault Location in Power Distribution Systems via Deep Graph Convolutional Networks," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 1, pp. 119–131, 2020.

[6] H. Sun, S. Kawano, D. N. Nikovski, T. Takano, and K. Mori, "Distribution system fault location analysis using graph neural network with node and link attributes," in *IEEE PES Innovative Smart Grid Technologies Conference - Europe (ISGT Europe)*, oct 2021.

[7] R. A. Jacob, S. Senemmar, and J. Zhang, "Fault Diagnostics in Shipboard Power Systems using Graph Neural Networks," in *2021 IEEE 13th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED)*, vol. 1, pp. 316–321, 2021.

[8] W. Liao, D. Yang, Y. Wang, and X. Ren, "Fault diagnosis of power transformers using graph convolutional network," *CSEE Journal of Power and Energy Systems*, vol. 7, no. 2, pp. 241–249, 2021.

[9] X. Chen, X. Yin, and S. Deng, "A novel method for SLG fault location in power distribution system using time lag of travelling wave components," *International Transactions on Electrical Energy Systems*, vol. 12, no. 1, pp. 45–54, 2017.

[10] F. V. Lopes, K. M. Dantas, K. M. Silva, and F. B. Costa, "Accurate Two-Terminal Transmission Line Fault Location Using Traveling Waves," *IEEE Transactions on Power Delivery*, vol. 33, no. 2, pp. 873–880, 2018.

[11] M. Jiménez Aparicio, M. J. Reno, P. Barba, and A. Bidram, "Multi-Resolution Analysis Algorithm for Fast Fault Classification and Location in Distribution Systems," *9th International Conference on Smart Energy Grid Engineering (SEGE)*, 2021.

[12] F. Wilches-Bernal, M. Jiménez Aparicio, and M. J. Reno, "An Algorithm for Fast Fault Location and Classification Based on Mathematical Morphology and Machine Learning," in *2022 IEEE Innovative Smart Grid Technologies North America (ISGT NA)*, 2022.

[13] F. Wilches-Bernal, M. Jimenez-Aparicio, and M. J. Reno, "A Machine Learning-based Method using the Dynamic Mode Decomposition for Fault Location and Classification," in *Thirteenth Conference on Innovative Smart Grid Technologies (ISGT)*, 2022.

[14] M. Jimenez Aparicio, S. Grijalva, and M. J. Reno, "Fast Fault Location Method for a Distribution System with High Penetration of PV," in *54th Hawaii International Conference on System Sciences*, 2021.

[15] S. Paul, S. Grijalva, M. Jimenez Aparicio, and M. J. Reno, "Knowledge-Based Fault Diagnosis for a Distribution System with High PV Penetration," in *Thirteenth Conference on Innovative Smart Grid Technologies (ISGT)*, 2022.

[16] M. Jimenez Aparicio, *Fast Fault Location Method For a Distribution System With High Penetration of PV*. School of Electrical and Computer Engineering Theses and Dissertations, Georgia Institute of Technology, 2020.

[17] T. N. Kipf and M. Welling, "Semi-Supervised Classification with Graph Convolutional Networks," *CoRR*, vol. abs/1609.0, 2016.

[18] B. Sanchez-Lengeling, E. Reif, A. Pearce, and A. Wiltschko, "A Gentle Introduction to Graph Neural Networks," *Distill*, vol. 6, aug 2021.