SAND2021-15234C

A Machine Learning-based Method using the Dynamic Mode Decomposition for Fault Location and Classification

Felipe Wilches-Bernal
Sandia National Laboratories
Albuquerque, NM
fwilche@sandia.gov

Miguel Jiménez-Aparicio
Sandia National Laboratories
Albuquerque, NM
mjimene@sandia.gov

Matthew J. Reno
Sandia National Laboratories
Albuquerque, NM
mjreno@sandia.gov

***Abstract*—A novel method for fault classification and location is presented in this paper. This method is divided into an initial signal processing stage that is followed by a machine learning stage. The initial stage analyzes voltages and currents with a window-based approach based on the dynamic mode decomposition (DMD) and then applies signal norms to the resulting DMD data. The outputs for the signal norms are used as features for a random-forests for classifying the type of fault in the system as well as for fault location purposes. The method was tested on a small distribution system where it showed an accuracy of 100% in fault classification and a mean error of ∼30 m when predicting the fault location.**

***Index Terms*—Dynamic Mode Decomposition, Fault Location, Fault Classification, Random Forests, Machine Learning**

I. Introduction

The main objective of power system protection techniques is to keep system operations safe with a minimal reduction in performance. To do so, protection techniques try to isolate only the parts of the system that are affected by the fault. For this task to be effective, protection devices must be able to detect, classify and locate faults with high accuracy. The time it takes for protection devices to engage their corrective actions has decreased significantly in the past decades. This has been partly possible by introducing time-domain protection schemes like the theory of traveling waves, high sampling resolution of the monitored signals, and advanced techniques to extract information from those signals. The most popular of these techniques are the discrete wavelet transform (DWT) [1] and mathematical morphology [2]. Research on power system protection, has also shown the benefits of using machine learning (ML) and deep learning (DL) algorithms in fault classification and location tasks [3].

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.

Combining DWT approaches with machine learning for protection applications such as fault detection, classification, and location has been reported previously [4]–[8]. For MM techniques, this has also been investigated [9]. Other than [8], these techniques typically requires data of at least a couple of milliseconds to be able to perform their protective task. However, the work in [8] demonstrated that even with a window of 0.1 ms a combined ML and DWT approach is able to classify and detect faults in a power system.

It has been demonstrated that the dynamic mode decomposition is a useful tool for analyzing power system signals with distortions due to power system faults [10]. The approach proposed in [10] shows that the real part of the DMD eigenvalues varies in cases where faults in the system are present. This work proposes a novel approach that combines the DMD technique in [10] with ML for classifying and localizing faults in power systems. The proposed method initially obtains time signals related to the DMD eigenvalues of the voltages and current signals. Then ℓ_p-norms are used to obtain metrics from the DMD signals which are then used as the features of an ML algorithm, which was selected to be a random forest (RF). The paper shows the importance of the features for the classification and location tasks. The method is tested in the same system as [8]. The results of the proposed approach show that the mean predicted error for the fault location task was halved with respect to that previous work. Because the results in [8] show an accuracy of 100%, those results cannot be improved but were matched by the proposed method.

The rest of the paper is organized as follows. The power system used in this work is presented in Section II. The proposed method that uses DMD to process the data and RFs for fault location and classification is presented in Section III. The evaluation of the proposed method in the test system is presented in Section IV. The conclusions and potential continuations of this work are presented in Section V.

II. Test System

The test power system used in this paper is shown in Fig. 1. The system consists of a *variable-length* distribution line at 12.47kV that is connected to a 300 kVA per phase load. The generator is connected at the primary side of the transformer at a 115 kV voltage level. The faults of study in this system are included at varying locations from 25 m until 4000 m in steps of 25 m. These faults are always included at the fault bus. This work considers three different types of faults: single-line-to-ground (SLG), line-to-line (LL), and line-to-line-to-line (LLL). It also considers seven different values of fault resistances: 0.01, 0.05, 0.1, 0.5, 1, 5, and 10 Ω. In total, 3360 simulations were used in this work. The system was implemented in PSCAD and the simulations were performed with a resolution of 10 MHz (for a time step of $0.1\mu s$).

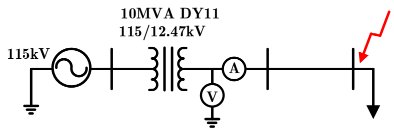

Fig. 1: Schematic of the system used in this work.

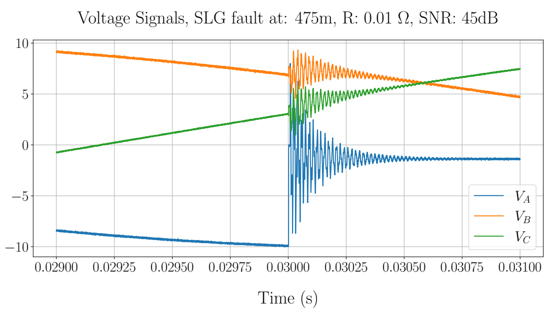

Fig. 2: Voltages for the system in Fig. 1 for the case of a SLG fault applied at 475 meters with a resistance of 0.1 Ω.

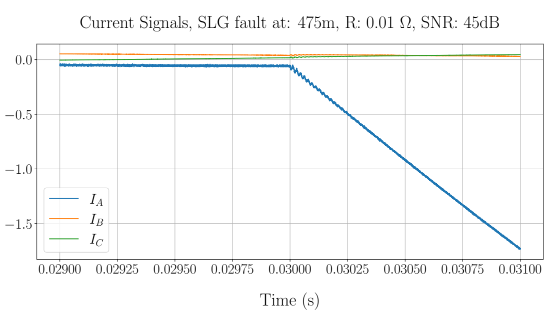

Fig. 3: Currents for the system in Fig. 1 for the case of a SLG fault applied at 475 meters with a resistance of 0.1 Ω.

Figs. 2 and 3 show, respectively, the voltages and currents of the system for a SLG fault with a resistor of 0.1 Ω. These results are obtained from the measurement point (the secondary of the transformer) and correspond to a time window of 2 ms, which extends 1ms before and after the fault inception time. The signals shown in these figures are those obtained from the simulations of the system in Fig. 1 but contaminated with noise. These signals have a signal-to-noise ratio (SNR) of 45 dB, which is the level of noise used in this work.

III. Method for Fault Location and Classification

The proposed method for fault location and classification is presented in this section. The method uses a window of 100 μs of voltage and current measurements from the power system with 50 μs coming before and 50 μs coming after the fault detection. These data are analyzed with the DMD and then with norms-based metrics. The data that is generated in the metrics stage is then used to train a random forest algorithm for fault location and another for fault classification.

A. The Dynamic Mode Decomposition

DMD is a method aimed at estimating system dynamics from measurement data [11], [12]. This method computes the best linear dynamics of the system for which the data is captured even if that system is nonlinear. DMD approximates the Koopman (or composition) operator which is a linear, infinite-dimensional operator able to represent nonlinear systems on a (Hilbert) space of measurement functions of its state. Even though it is linear, the Koopman operator is able to represent nonlinear dynamics by being an infinite-dimensional operator [13]. DMD determines the eigenvalues and eigenvectors of a finite-dimensional linear system that can be interpreted as approximating, the infinite-dimensional Koopman operator. DMD is also interpreted as a dimensionality reduction approach that, at its core, uses singular value decomposition.

Given a sequential set of measurements, $\{\mathbf{x}_1, \mathbf{x}_2 \dots, \mathbf{x}_m\}$ where $\mathbf{x}_i \in \mathbb{R}^n \; \forall i = 1 \dots m$, taken from a system at regular intervals, DMD estimates

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k \tag{1}$$

which is a linear system representation that captures the dynamics present in the measurement set. Formally, the input of DMD is a measurement set and its outputs are the following two matrices:

$$\mathbf{\Phi} = \begin{bmatrix} | & | & & | \\ \phi_1 & \phi_2 & \dots & \phi_r \\ | & | & & | \end{bmatrix} \tag{2}$$

$$\mathbf{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_r \end{bmatrix} \tag{3}$$

where $\mathbf{\Lambda}$ and $\mathbf{\Phi}$ are matrices that, respectively, contain the eigenvalues of the system and the corresponding right eigenvectors. It is important to note that the index r represents the relevant dimensions of the DMD algorithm and is user-defined parameter. Note that $r \leq n$ and for the case $r = n$ no dimensionality reduction was performed by DMD.

It has been demonstrated that DMD is a tool capable of analyzing power system signals with disturbance information. Particularly, the work in [10] proposes a window-based approach that uses DMD to detect when power system signals are distorted due to an event in the system. That work uses the real part of the DMD eigenvalue for its fault detection task. In this work, the DMD method of [10] is used as an initial processing stage to compute the features for a machine learning algorithm. The output dimensions of the DMD method are set to 2 (i.e. $r = 2$), and the three cases of input signals are considered: voltages, currents, and a case with both voltages and currents are analyzed at once. Note that for this latter case both types of signals, voltages and currents, were normalized before being used as the input of the DMD method. Fig. 4 shows the real part of the first eigenvalue, α_1, obtained from DMD as a function of time. The data used to obtain these results is a window 100 μs around the fault inception from the signals in Fig. 2. This figure shows that for all DMD window sizes considered, α_1 fluctuates after the fault occurs.

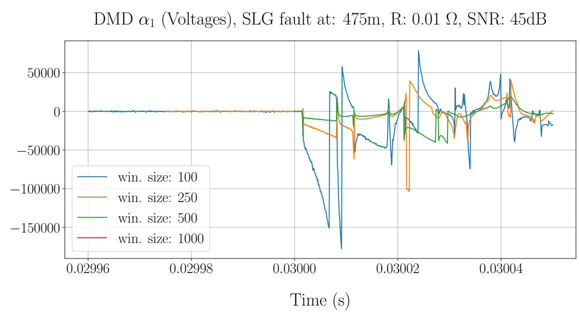

Fig. 4: Real part of the first DMD eigenvalue (α_1) as a function of time for different window sizes.

B. Signal Norms

The ℓ_p-norm of a one-dimensional discrete signal in $\mathbb{R}^n$ is defined by

$$||x||_p = \left(\sum_{k=1}^{n} |x_k|^p \right)^{1/p} \qquad \forall 1 \leq p \leq \infty \tag{4}$$

This function performs $\mathbb{R}^n \mapsto \mathbb{R}_{\geq 0}$, where $\mathbb{R}_{\geq 0}$ represents the set of non-negative real numbers. For $p = 1$, (4) represents the taxicab norm, for $p = 2$, (4) represents the Euclidean norm (which is related to the energy of the signal), and when $p = \infty$, (4) describes the infinity norm. These norms have the following property

$$||x||_\infty \leq ||x||_2 \leq ||x||_1. \tag{5}$$

This work uses the 1 and ∞ norms which are defined by

$$||x||_1 = \sum_{k=1}^{n} |x_k| \tag{6}$$

$$||x||_\infty = \max_{1 \leq k \leq n} |x_k|. \tag{7}$$

The norms in this work are used on the signals obtained from the DMD approach such as those presented in Fig. 4 as a way to *encapsulate* their intrinsic information. Fig. 5 shows an example of this approach for the 1-norm of α_1 as a function of the fault distance. The results in this figure are for a SLG fault with the shades highlighting the differences due to the different fault resistance values.

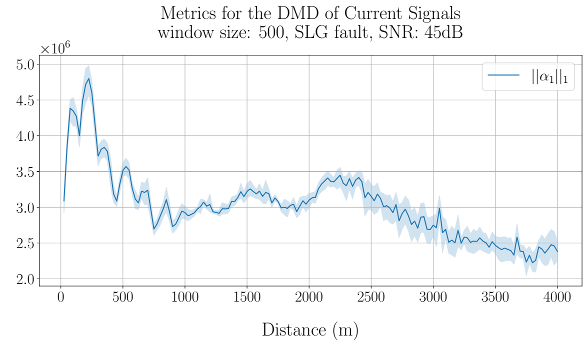

Fig. 5: 1-norm of α_1 as a function of the distance between the sensor and a SLG fault. The shades indicate differences due to the fault resistance values. In this figure, each point corresponds to a different simulation with a fault applied at a different location.

C. Random Forests for Fault Location and Classification

The simulation setup explained in Section II for the system in Fig. 1 produces 3360 simulations. For each of these simulations, the voltage, current, and a combination of voltage and current signals are analyzed with DMD and metrics are obtained based on the aforementioned norms in Section III-B. There are 48 quantities that can be obtained for each simulation, as follows

- Three measurements (voltages, currents, and voltage and currents together).
- Real part of the two eigenvalues of DMD.
- Four window sizes (100, 250, 500, and 1000 points).
- Two norms (1 and ∞-norm).

These quantities are used as features for a RF as the selected ML algorithm in this work for the tasks of fault classification and location. RFs are a popular ML technique because of its ability to work with a wide variety of data that does not require any particular scaling, and because of the interpretability of its results. RFs are supervised ML algorithms that are built from decision trees. Using a technique called ensemble learning, RFs aggregate the output of the decision trees that form part of it. This technique is powerful and by using it the predictive power of RFs is enhanced [14]. At its core, ensemble learning is based on the law of large numbers where good prediction can be built even from weak learners (that is ML algorithms that are barely better than random guessing). However, this only happens when the methods (or algorithms) that make up the ensemble set are completely independent. For this reason, when using ensemble learning, there is a need for the predictors to be as independent as possible. One way to create an ensemble of different predictors is to use different algorithms. Another way is to use the same algorithms but to train them with different subsets of the training set selected at random. *Bagging* and *pasting* are the names of the sampling techniques, when the sampling occurs respectively with and without replacement [14]. In RFs the algorithms that are aggregated are the same, decision trees as mentioned above, usually trained with the bagging approach.

IV. Results in the Test System

This section presents the results of using the method for fault location and classification described in Section III to a set of simulated data obtained from the system in Fig. 1. In this work two separate RFs are developed, one for the classification task and the other for the fault-location task. The 3360 simulations were divided into 2730 for the training set and 630 for the test set.

A. Fault Classification Task

The objective of this task is to determine the type of fault that occurred, either SLG, or LL, or LLL. The RF obtained in this work was trained with the 48 features for each one of the 2730 simulations that were labeled accordingly. The metrics of performance on the testing data for this RF classifier are:

- Accuracy of 100%.
- Precision of 100%.
- Recall of 100%.
- F-score of 100%.

These results are the same as those in [8] where the initial processing stage was based on the DWT and the signals considered did not have any noise.

This work also analyzed the feature importance of the RF for the classification task. Fig. 6 shows the 15 most important features for the RF classifier. The results in this figure show these features use both the infinity norm coupled with α_1. In this figure E, I, and C stand respectively for voltages, currents, and the combination of both types of signals. These results also show that all signals and window sizes are used in this task. That is, the RF algorithm finds useful information at different DMD resolutions.

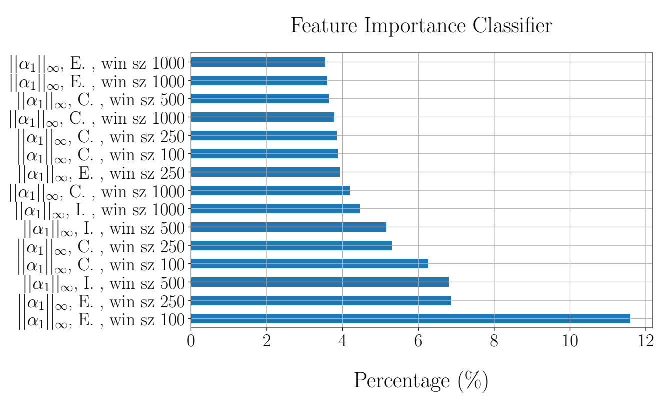

Fig. 6: Feature importance for the RF classifier for the fault classification task.

B. Fault Location Task

The objective of this task is to determine how far away the fault occurred. This is a regression task that was also approached with a RF trained with the same type of data as the classifier in Section IV-A. The main difference is that the labels in this case are the distance at which the fault occurred. The structure of RFs is defined by a set of variables known as hyperparameters. These variables define things such as the number of decision trees in the RF and the minimum number of leaves that they can have. These parameters ultimately determine the performance of the RF. This research performs a randomized search cross validation to determine the best set of parameters for the fault-location task among a large set of candidates. Table I shows the hyperparameters that were part of the randomized cross-validation search. This table also shows the search space and the optimal value of those hyper-parameters. Further information about the hyperparameters of RF can be found at [15]. The predictions with the RF for fault location trained in this work have the mean error and standard deviation presented in Table II. This table also presents the same metrics for the work in [8]. The results in this table show that the proposed approach has around half of the mean error and around 40% of the standard deviation than the work in [8].

TABLE I: RF parameters for randomized search (for regressor)

Hyperparameter	Range	Optimal Value
Number of estimators	40 to 400 in steps of 10	220
Min. samples split	[2, 5, 10]	2
Min. samples leaf	[1,2,4]	1
Max. features	`[auto, sqrt]`	`sqrt`
Max. depth	10 to 110 in steps of 10	40
Bootstrap	`[True, False]`	`False`

TABLE II: Comparison between the proposed method and the method in [8]

Method	Mean Error (m)	Standard Deviation (m)
Proposed approach	31.88	38.19
Approach in [8]	62.95	94.77

Fig. 7 shows the distribution of prediction errors per fault-type. These results show that all these distributions are fat tailed but that the distribution for LLL faults has a much shorter tail than those for the other two fault types. These results are different from those in [8] where the distribution for LL faults has a much larger tail than the other two.

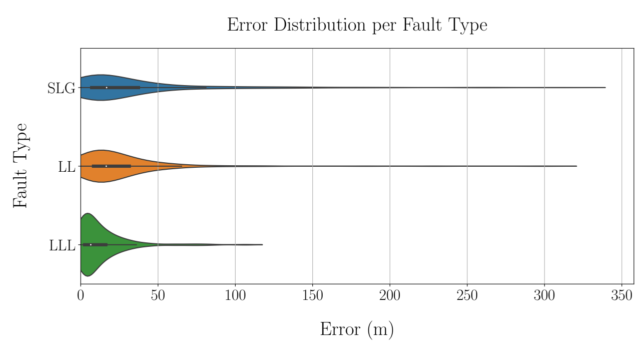

Fig. 7: Error distributions per fault-type for the RF regressor.

Fig. 8 shows the 15 most important features for the RF regressor for the fault location task. Similar to the results for the classifier in Fig. 6, the results in this figure show these features use both the infinity norm coupled and real part of the first DMD eigevalue (α_1). In addition, these features are based on the voltages and currents but not in the case that combines them both.

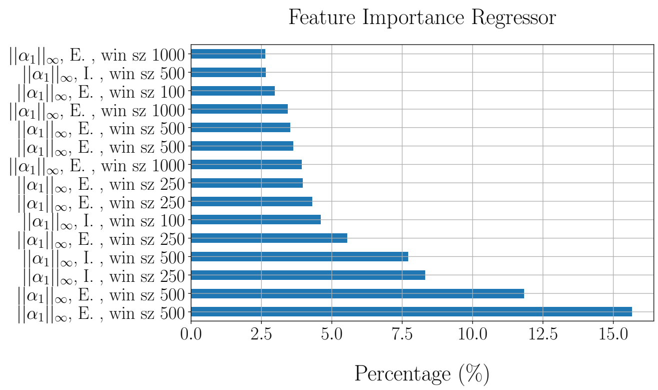

Fig. 8: Feature importance for the RF regressor for the fault location task.

V. Conclusions and Future Work

This paper presents a new method to classify and locate faults in power systems. The proposed method initially parses voltage and current signals with the dynamic mode decomposition. Then metrics based on the ℓ_p-norms are used on the signals resulting from the DMD stage. These metrics become the features for a machine learning approach that is based on random forests. The proposed method is tested in signals for a test system in noisy conditions, 45 dBs of SNR. The data required in the proposed method is just a window of 100 μs which is significantly less than most existing approaches. The method is shown to have an accuracy of 100% for the classification task and a prediction error of around 30 m for the fault location task. The average fault location error is around half of the one reported for the same system in an earlier work.

This work can be expanded in three ways: (i) the parsing of the voltage and current signals can be extended from DMD to other methods such as the Wavelet Transform or Mathematical Morphology, (ii) different machine learning and deep learning algorithms can be used instead of (or along with) RFs, and (iii) perform the test in larger and more complex power systems.

References

[1] A. Osman and O. Malik, "Transmission line distance protection based on wavelet transform," *IEEE Transactions on Power Delivery*, vol. 19, no. 2, pp. 515–523, 2004.

[2] F. Namdari and M. Salehi, "High-speed protection scheme based on initial current traveling wave for transmission lines employing mathematical morphology," *IEEE Transactions on Power Delivery*, vol. 32, no. 1, pp. 246–253, 2016.

[3] 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. 72 949–72 969, 2021.

[4] V. Veerasamy, N. I. A. Wahab, M. L. Othman, S. Padmanaban, K. Sekar, R. Ramachandran, H. Hizam, A. Vinayagam, and M. Z. Islam, "Lstm recurrent neural network classifier for high impedance fault detection in solar pv integrated power system," *IEEE Access*, vol. 9, pp. 32 672–32 687, 2021.

[5] N. Zhang and M. Kezunovic, "Transmission line boundary protection using wavelet transform and neural network," *IEEE Transactions on Power Delivery*, vol. 22, no. 2, pp. 859–869, 2007.

[6] X. Magagula, Y. Hamam, J. Jordaan, and A. Yusuff, "Fault detection and classification method using dwt and svm in a power distribution network," in *2017 IEEE PES PowerAfrica*. IEEE, 2017, pp. 1–6.

[7] A. Yadav and A. Swetapadma, "Combined dwt and naive bayes based fault classifier for protection of double circuit transmission line," in *International Conference on Recent Advances and Innovations in Engineering (ICRAIE-2014)*. IEEE, 2014, pp. 1–6.

[8] M. Jiménez Aparicio, M. J. Reno, P. Barba, and A. Bidram, "Multi-resolution analysis algorithm for fast fault classification and location in distribution systems," in *2021 IEEE International Conference on Smart Energy Grid Engineering (SEGE)*. IEEE, 2021, pp. 1–7.

[9] M. Mishra, R. R. Panigrahi, and P. K. Rout, "A combined mathematical morphology and extreme learning machine techniques based approach to micro-grid protection," *Ain Shams Engineering Journal*, vol. 10, no. 2, pp. 307–318, 2019.

[10] F. Wilches-Bernal, M. J. Reno, and J. Hernandez-Alvidrez, "A dynamic mode decomposition scheme to analyze power quality events," *IEEE Access*, vol. 9, pp. 70 775 – 70 788, 2021.

[11] P. J. Schmid, "Dynamic mode decomposition of numerical and experimental data," *Journal of fluid mechanics*, vol. 656, pp. 5–28, 2010.

[12] J. H. Tu, C. W. Rowley, D. M. Luchtenburg, S. L. Brunton, and J. N. Kutz, "On dynamic mode decomposition: Theory and applications," *Journal of Computational Dynamics*, vol. 1, no. 2, pp. 391–421, 2014.

[13] J. N. Kutz, S. L. Brunton, B. W. Brunton, and J. L. Proctor, *Dynamic mode decomposition: data-driven modeling of complex systems*. SIAM, 2016.

[14] A. Géron, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O'Reilly Media, 2019.

[15] "sklearn RandomForestClassifier," https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html, accessed: 2021-08-04.