A Fault Classification Method for Medium Voltage Networks with a high Penetration of Photovoltaic Systems using Artificial Neural Networks

Tran The Hoang, Quoc Bao Duong, Tuan Tran Quoc, Yvon Besanger, Tung Lam Nguyen

► To cite this version:

Tran The Hoang, Quoc Bao Duong, Tuan Tran Quoc, Yvon Besanger, Tung Lam Nguyen. A Fault Classification Method for Medium Voltage Networks with a high Penetration of Photovoltaic Systems using Artificial Neural Networks. 2020 IEEE PES General Meeting: “Are Big Data, Machine Learning & Electric Transportation Transforming the Grid?”, Aug 2020, Montréal, Canada. hal-03145523

HAL Id: hal-03145523

<https://hal.science/hal-03145523>

Submitted on 18 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Fault Classification Method for Medium Voltage Networks with a high Penetration of Photovoltaic Systems using Artificial Neural Networks

Tran The Hoang*, Quoc Bao Duong[†], Quoc Tuan Tran[‡], Yvon Besanger* and Tung Lam Nguyen*

*G2Elab, UGA, CNRS, Grenoble INP, 38000 Grenoble, France

Email: tran-the.hoang@g2elab.grenoble-inp.fr, Yvon.Besanger@g2elab.grenoble-inp.fr,
tung-lam.nguyen@g2elab.grenoble-inp.fr

[†]AIP PRIMECA Dauphiné Savoie, Grenoble INP, 38031 Grenoble, France

Email: quoc-bao.duong@grenoble-inp.fr

[‡]CEA/INES, 73375 Le Bourget-du-Lac, France

Email: QuocTuan.Tran@cea.fr

Abstract—With the rapid advancement of power electronic technologies and the reduction of photovoltaic cell price, the share of solar energy in the total power production has been booming recently. On the one hand, the increase in the amount of power delivered by solar energy can be beneficial in many economic and environmental aspects. On the other hand, this can cause various technical challenges to network operators. One of these issues is related to classifying faults located in distribution networks with high penetration of photovoltaic systems. Although many studies have paid significant attention to developing new algorithms applicable for a more active today distribution networks, there is still space for other improvements. Hence, after reviewing state-of-the-art researches, this paper was intended to develop a fault classification that is based on artificial neural networks. In particular, a technique so-called Multiplayer Perceptron Classifier was selected for the proposed algorithm. First, the authors generated a data set for the study by modeling and simulating a real distribution network with practical parameters provided by a local utility in the environment software PowerFactory/DigSILENT. Multiple fault scenarios were simulated. Second, a part of the generated data collection was used for network learning. Finally, the performance of the proposed methodology was demonstrated via testing on the remaining number of generated data.

Index Terms—Artificial neural network, Multilayer Perceptron Classifier, fault classification, distribution network, photovoltaic system;

I. INTRODUCTION

Faults on distribution feeders are one of the main factors causing disturbances or sometimes interruption in power supply to customers, and therefore need to be quickly and accurately detected and cleared by protection systems. These faults then should be isolated by fault management systems, and the supply is recovered to as many affected customers as possible. For these systems to realize fast, accurate, and effective performance, the role of the fault classification scheme is of vital importance.

The development of fault classification algorithms for distribution networks has attracted lots of research over the last decades. According to [1], fault classification methods may be

approximately divided into the following categories, including logic flow-based, decision tree, support vector machine, fuzzy inference systems, and artificial neural network. Prior to the era of smart grid technologies, most researches for distribution networks had used the same principles developed for transmission systems. Authors in [2] used a negative sequence factor derived from pre-fault and post-fault measurements of negative-sequence and zero-sequence currents as inputs for their neural network-based fault classification algorithm. The scheme in [3] also proposed to use μ PMUs located at different locations across the feeder for obtaining the input data for the classification system. A method in [4] developed an ANFIS-based method that can adapt to different grounding systems, fault resistance, load levels, and the change in network configuration, achieving a rather high accuracy of about 99.4%. A method published in [5] was based on Hilbert-Huang Transform and Convolution Neural Network that employed transient and stationary fault signals as inputs. Research in [6] made use of fuzzy-logic, avoiding the requirement for training stages. The main disadvantage of these methods in [4]–[6] is the requirement for a high sampling rate (10 kHz) that would impose huge computation burdens for the fault recording systems. Hence, a wavelet-based scheme was described in [7] that utilized Daubechies fourth-order wavelet for reducing the computational burden. Nonetheless, none of these above-mentioned researches took the impacts of distributed generators and their detailed fault characteristics into considerations.

In the event of a fault, new inverter-interfaced generating units, such as photovoltaic (PV) systems, Battery Energy Storage System (BESS), converter-based wind turbine (WT) ... and so on, respond to faults in a way differing from that of conventional synchronous generators. As a result, as the integration of such distributed generations (DGs) has increased, fault characteristics of the distribution systems have changed significantly. Moreover, many countries have upgraded their grid codes to require these DGs, especially those connected

to medium- and low-voltage grids, to actively be involved in supporting grid stability in the event of faults, resulting in their unneglectable impacts on transient processes of the overall network. Thus, those methods developed for transmission networks cannot be applied. Moreover, the distribution networks are no longer radial after the installation of DGs downstream of the outgoing feeders. Nondirectional overcurrent based protection units are not sufficient for protecting the whole grid in different operating scenarios. Other methods, such as directional or distance protections, should be applied [8]–[10]. Since these protection principles require information on fault types before moving to the next operating stage, the development of new effective fault classification algorithms dedicated to distribution networks with high penetration of inverter-interface DGs has become more and more essential. However, the number of fault classification studies for a new kind of active distribution network is still limited. In [11], the wavelet energy spectrum entropy-based method was utilized. Different fault locations, fault resistances, fault inception angles, load angle variations, network topology changes as well as the connection of distributed generators, were taken into account. A similar methodology was also presented in [12]. The processing time of this method is less than 10ms, which is quick enough for protection purposes. Nonetheless, the effects of fault resistances and variations of DG generation were ignored. Reference [13] suggested wavelet transform for extracting distinct features for identifying faulted phases. However, the computation time and sampling rates of input data that are very important for practical implementation were not mentioned.

For the application of new kinds of protection principles rather than nondirectional overcurrent, new fault type classification that can cope with the high integration of PV systems and their intermittent nature should be developed. In the context of smart grids, especially the potential application of big data analytics for utilizing historical data accumulated from practical power system operation, machine learning is an inevitable trend of research, and of course, it would be the most suitable application tool in this case. Therefore, in this paper, a fault-type classification scheme based on Artificial Neural Network that can deal with these shortages mentioned above was proposed. The Artificial Neural Network was theoretically explained in subsection III-A. The performance of the proposed system was investigated in different fault scenarios.

II. THE STUDIED NETWORK

The studied network is a realistic medium-voltage distribution network whose parameters are provided by a local utility. The single-line diagram of the network is shown in Fig. 1. The network consists of 7 diesel generators feeding the loads via 7 transformers whose connection groups are $\Delta/Y-11$. The rating of each generator is 1.125 MVA. The total installed capacity of PV and wind systems are 8 MW and 3 MW, respectively. The rating of a PV system is 500 kW. There are two PV plants located in different locations with PV-1 at the generation center and PV-2 near the end of feeder 1. The

number of PV systems installed at PV-1 and PV-2 are 6 and 10, respectively.

Fig. 1. Single-line diagram of the studied network

From the 22 kV busbar systems located in the generation center, three radial feeders are originated from the 22 kV busbar systems of the generation center, supplying various single-phase and three-phase LV loads. Moreover, feeder 1 and feeder 2 mesh at several points by normally-open switches. The earthing system used for the MV network is the solidly and undistributed earthed method. Regarding the earthing LV grid, the TT system was selected since it appears as the most suitable method, allowing the PV systems to be coupled with the grid without grounding their neutral point [14].

Since the total maximum load demand, which is 8 MW, is less than the overall installed capacity of the PV and wind energy, a battery energy storage system (BESS) was installed at the production center to store the surplus amount of power generation during peak hours of PV systems. On the other hand, the BESS will be discharged to feed the loads during some periods over a day when required, for instance, during no-generation hours of PV systems.

Note that the PV and wind systems in the study are modeled

to fulfill the new grid code requirements. In other words, PV and wind systems should be capable of staying connected during the fault and feeding reactive current to support the network voltage recovery. The control scheme of PV systems is illustrated in Fig. 2 [15]. During the fault event, the reference signal generator block generates the reference currents that follow the dynamic voltage support requirements imposed by the grid code indicated in [16]. Also, the output current should be confined to the inverter thermal limit, which in the case is equal to 1.2 times its rated current [16].

Fig. 2. The control diagram of the PV systems

Additionally, the three-phase loads were partially modeled as dynamic loads, i.e., frequency- and voltage-dependent, to obtain more realistic load effects on fault data. In contrast, single-phase loads were represented as static load, i.e., constant impedance.

III. THE PROPOSED FAULT CLASSIFICATION ALGORITHM

A. Multi-layer perceptron classifier-based artificial neural network

Artificial Neural Networks (ANNs) are statistical learning algorithms that are widely used recently in numerous applications, including classification schemes. Many libraries and frameworks are available to develop neural networks. Among these other, Multi-layer Perceptron (MLP) Classifier appears to be one of the most suitable candidates for developing classification purposes. Opposed to other classification methods, such as Support Vector Machines or Naive Bayes Classifier, MLP Classifier applies an underlying Neural Network for conducting classification tasks [17]. Therefore, the authors proposed to use the MLP classifier in this paper thanks to their ability to solve problems stochastically, which often allows approximate solutions for extremely complex problems like fitness approximation. The layer-based architecture of the applied MLP classifier can be observed in Fig. 3.

Each layer contains N neurons connected to the neurons of the previous layer. Each neuron, in turn, receives many inputs from other nodes and computes a single output based on the inputs and the connection weights. The layers which are between these two are named hidden layers. In our proposal, we applied a three-layer neural network. The first layer consists of a set of 50 neurons, whereas the two followings comprise sets

Fig. 3. The layer-based architecture of the applied MLP classifier

of 30 and 4 neurons, respectively. The output of the neural network should be of 4 since we were intended to classify the fault into four categories. The number of the inputs was justified in subsection III-B. For properly triggering each layer, a suitable activation function should be chosen according to the application goals. In our case, an activation function, namely sigmoid, was chosen and described as:

$$\sigma(z) = \frac{1}{1 + e^{-z}} = \frac{e^z}{e^z + 1} \quad (1)$$

The computation that is carried out in the neuron of the first layer would be as:

$$z^{(1)} = \left[W^{(1)} \right] \cdot X + b^{(1)} \quad (2)$$

where X is the input vector, $b^{(1)}$ is the bias vector of all units whose dimension is $[50 \times 1]$, and $W^{(1)}$ is the weight matrix of dimension $[50 \times N]$ described by (3):

$$W^{(1)} = \begin{bmatrix} \omega_{1-1}^{(1)} & \cdots & \omega_{1-N}^{(1)} \\ \vdots & \ddots & \vdots \\ \omega_{50-1}^{(1)} & \cdots & \omega_{50-N}^{(1)} \end{bmatrix} \quad (3)$$

The calculations in the two hidden layers are defined as:

$$z^{(2)} = \left[W^{(2)}\right]^T \cdot \sigma\left(z^{(1)}\right) + b^{(2)} \quad (4)$$

$$z^{(3)} = \left[W^{(2)} \right]^T \cdot \sigma \left(z^{(2)} \right) + b^{(3)} \quad (5)$$

where $W^{(2)}$ and $W^{(3)}$ are weight matrix whose dimensions can be obtained from Fig. 3. The sizes of $b^{(2)}$ and $b^{(3)}$ are $[30 \times 1]$ and $[4 \times 1]$ respectively. The output neuron of dimension $[4 \times 1]$ is calculated as:

$$\hat{y} = \sigma \left(z^{(3)} \right) \quad (6)$$

The learning of the neural network is accomplished by altering the weight of the connection. By updating the weight iteratively, the performance of the network is improved. In this developed algorithm, we utilized the Back-propagation method presented in [18]. Back-propagation is a standard method whose goal is to train the neural network for calculating the weights and bias. In this case, we had the known data output (supervised learning), by a method to lessen the error function E , we can obtain weights $\omega_{ij}^{(l)}$. We can estimate the error for each output neuron by adopting the squared error function and sum them to find the total error:

$$E = \frac{1}{2} \cdot \sum_{n=1}^N \sum_{i=1}^m [y_i^l(x_n) - t_{n,i}]^2 \quad (7)$$

where x_n is the n -th element of the input vector X , $t_{n,i}$ is the n -th element of the output vector of the i -th class, $y_i^l(x_n)$.

The back-propagation step computes the gradient of E with respect to this input as follows:

$$\omega_{ij}^{(l)} \leftarrow \omega_{ij}^{(l)} - n \cdot \frac{\partial E}{\partial \omega_{ij}^{(l)}} \quad (8)$$

where n represents the learning rate.

The learning process of the proposed algorithm is described in Algorithm 1 [18].

Algorithm 1 The learning process of the proposed algorithm

- 1: Feed-forward computation: use equation (2), (3), (4), (5) and (6) for calculate for each layer. The vectors $z^{(1)}$, $z^{(2)}$ and $z^{(3)}$ are computed and stored. The evaluated derivatives of the activation functions are also stored at each unit
- 2: Back-propagation to the output layer
- 3: Back-propagation to the hidden layer
- 4: Weight updates

B. Data measurement and feature extraction

The faults in power systems can be classified into four types, including single-phase-to-ground (Type-1), two-phase (Type-2), two-phase-to-ground (Type-3), and three-phase (Type-4). The last type of fault may be to earth or clear of the earth that has the same nature [19]. As justified in the introduction, with high penetration of PV systems and the introduction of new grid codes, the fault currents have become bidirectional rather than unidirectional. Moreover, the PV systems respond to faults in a way differing from that of conventional well-established synchronous generators. Therefore, one-point measurement appears impossibly enough for fault type identification within such an active grid.

In this research, after having considered the number of measurement data set, the author proposed to install three measuring units (MU) on the concerned feeder, i.e., feeder 1, as shown in Fig. 1 including MU1, MU2, and MU3. The measuring units may be μ PMUs or smart meters that

can provide timely and time-synchronized current and voltage phasors effectively for data analysis in our case, [3]. Each MU samples the measured three-phase current and voltage values at the sampling rate of 800Hz. The derived currents and voltages were filtered by using a cosine-filtering algorithm that is quite popular in relay technologies, as presented in Fig. 4 [20].

Fig. 4. Computation of current phasor from sampled waveform

Application of machine learning-based approaches for fault type identification, especially for large networks, necessitates a large data set of faults, requiring a great number of fault scenarios. In the study, the authors proposed to carry out fault simulation at different hours during a day, from 0 to 20 hours by a step of 4 hours. The varying operation times at which the fault happened also allowed us to cover the impacts of intermittent nature of PV generation. Moreover, we took the effects of fault resistance into account by increasing its value from 0 to 60 Ohm by a step of 10 Ohm. Faults were assumed to occur at 14 different sections of three feeders. By combining these above fault scenarios with 4 fault types, a totally of 2352 faults scenarios were simulated. In the study, faults were assumed to occur at $t = 1.5s$ and be cleared at $t = 1.65s$, i.e., lasted for 7.5 cycles; however, the authors only made use of 4 samples per cycle per each quantity for a duration of 2 first cycles from the instant of fault inception. So, for each fault scenario, there were 8 data windows taken into consideration.

Based on these sampled values, the data processing unit carried out the calculation of magnitudes and angles of current and voltage symmetrical sequence components. For each data window received from each MU, 12 quantities would be derived. As a result, the authors obtained 36 quantities from three MUs plus one quantity standing for the time, a totally of 37 quantities for each data window, as dictated in Fig. 5, corresponding to vector input X that was provided to the proposed ANN (Fig. 3).

C. Data set for training and testing the designed ANN

The data set was split into two smaller sets, one for training and another for testing with a ratio of 75:25 correspondingly, as shown in Fig. 5. Firstly, the first part of the data set was loaded into the designed ANN for learning purposes following the Algorithm 1 described in Subsection III-A. Afterward, the remaining part of the data set was used for testing the proposed approach. Fig. 5 presents the flowchart of the proposed method.

IV. EVALUATION AND DISCUSSION

The result of testing the developed network presented in Table I indicates that the proposed method identified fault

Fig. 5. The flowchart of the proposed method

TABLE I
FAULT TYPE CLASSIFICATION RESULTS PER HOUR

Hour	Average computation time, ms	Accuracy, %
0	1.889	100
4	2.016	100
8	2.259	99.1
12	2.432	98.6
16	3.391	98.7
20	1.950	99.3
Average	2.323	99.3

types with high accuracy, the overall average error is lower than 1%. One more notable feature of the proposed approach is that its average computation time for all cases in the testing data set is only 2.323 millisecond, rather fast compared with 10 ms of a method presented in [12]. It can also be observed that the computation time and the error increase at 8, 12, and 16 hours compared with three remaining testing hours. This is due to the participation of PV systems in power production. The type of faults that occurred during a high generation of PV systems was more difficult to be identified. It took the developed network more time and iteration steps to reach the final results.

Although the developed method was tested with the different grid configurations, it needs to be checked and upgraded in case of a large modification of the concerned network, e.g., disconnection of an outgoing feeder or a primary transformer.

V. CONCLUSION AND PERSPECTIVE

An MLP classifier based artificial neural network was proposed in this paper for fault classification within medium voltage distribution networks with a high share of PV systems. The proposed approach can deal with different fault scenarios with varying network operational times and scenarios. Detailed models of PV systems and loads were also employed for making the simulation data as close to practice as possible. High accuracy and low computation time of the method were demonstrated. Moreover, the proposed method can be employed for more complicated electrical networks since its performance just depends on the data set provided.

In the further step, the authors are determined to develop the presented methodology for fault location since this topic is rather challenging for distribution network operators in the context of high penetration of DGs.

REFERENCES

- [1] K. Chen, C. Huang, and J. He, "Fault detection, classification and location for transmission lines and distribution systems: a review on the methods," *High voltage*, vol. 1, no. 1, pp. 25–33, 2016.
- [2] A. Oliveira, P. Garcia, L. Oliveira, H. Silva, J. Pereira, and M. Tomim, "New algorithms for fault classification in electrical distribution systems," in *IEEE PES General Meeting*. IEEE, 2010, pp. 1–3.
- [3] M. U. Usman, J. Ospina, and M. O. Faruque, "Fault classification and location identification in a smart distribution network using ann," in *2018 IEEE Power & Energy Society General Meeting (PESGM)*. IEEE, 2018, pp. 1–6.
- [4] J. Zhang, Z. He, S. Lin, Y. Zhang, and Q. Qian, "An anfis-based fault classification approach in power distribution system," *International Journal of Electrical Power & Energy Systems*, vol. 49, pp. 243–252, 2013.
- [5] M.-F. Guo, N.-C. Yang, and W.-F. Chen, "Deep-learning-based fault classification using hilbert-huang transform and convolutional neural network in power distribution systems," *IEEE Sensors Journal*, 2019.
- [6] B. Das, "Fuzzy logic-based fault-type identification in unbalanced radial power distribution system," *IEEE Transactions on Power Delivery*, vol. 21, no. 1, pp. 278–285, 2005.
- [7] U. Dwivedi, S. Singh, and S. Srivastava, "A wavelet based approach for classification and location of faults in distribution systems," in *2008 Annual IEEE India Conference*, vol. 2. IEEE, 2008, pp. 488–493.
- [8] H. Laaksonen, D. Ishchenko, and A. Oudalov, "Adaptive protection and microgrid control design for hailuoto island," *IEEE Transactions on Smart Grid*, vol. 5, no. 3, pp. 1486–1493, 2014.
- [9] N. K. Choudhary, S. R. Mohanty, and R. K. Singh, "A review on microgrid protection," in *2014 International Electrical Engineering Congress (iEECON)*. IEEE, 2014, pp. 1–4.
- [10] A. A. Memon and K. Kauhaniemi, "A critical review of ac microgrid protection issues and available solutions," *Electric Power Systems Research*, vol. 129, pp. 23–31, 2015.
- [11] A. C. Adewole and R. Tzoneva, "Fault detection and classification in a distribution network integrated with distributed generators," in *IEEE Power and Energy Society Conference and Exposition in Africa: Intelligent Grid Integration of Renewable Energy Resources (PowerAfrica)*. IEEE, 2012, pp. 1–8.
- [12] M. Dehghani, M. H. Khooban, and T. Niknam, "Fast fault detection and classification based on a combination of wavelet singular entropy theory and fuzzy logic in distribution lines in the presence of distributed generations," *International Journal of Electrical Power & Energy Systems*, vol. 78, pp. 455–462, 2016.
- [13] K. Lout and R. K. Aggarwal, "Current transients based phase selection and fault location in active distribution networks with spurs using artificial intelligence," in *2013 IEEE Power & Energy Society General Meeting*. IEEE, 2013, pp. 1–5.
- [14] N. Jayawarna, N. Jenkins, M. Barnes, M. Lorentzou, S. Papathanassiou, and N. Hatziahyriou, "Safety analysis of a microgrid," in *2005 International Conference on Future Power Systems*. IEEE, 2005, pp. 7–pp.
- [15] M. Castilla, J. Miret, J. L. Sosa, J. Matas, and L. G. de Vicuña, "Grid-fault control scheme for three-phase photovoltaic inverters with adjustable power quality characteristics," *IEEE transactions on power electronics*, vol. 25, no. 12, pp. 2930–2940, 2010.
- [16] W. Bartels, F. Ehlers, K. Heidenreich, R. Huttner, H. Kuhn, T. Meyer, T. Kumm, J. Salzmänn, H. Schäfer, and K. Weck, "Generating plants connected to the medium-voltage network," *Technical Guideline of BDEW*, 2008.
- [17] A. Music and S. Gagula-Palalic, "Classification of leaf type using multilayer perceptron, naive bayes and support vector machine classifiers," *Southeast Europe Journal of Soft Computing*, vol. 5, no. 2, 2016.
- [18] R. Rojas, *Neural networks: a systematic introduction*. Springer Science & Business Media, 2013.
- [19] P. M. Anderson and P. M. Anderson, *Analysis of faulted power systems*. IEEE press New York, 1995, vol. 445.
- [20] W. Rebizant, J. Szafran, and A. Wiszniewski, *Digital signal processing in power system protection and control*. Springer Science & Business Media, 2011.