



INTELLIGENT FAULT DETECTION SCHEME FOR MICROGRIDS USING WAVELET-BASED NEURAL NETWORK

¹Gita Ghule, ²Dr. V.N.Ghate, ³Prof. V.M.Harne

¹PG Scholar, ²Associate Professor, ³Assistant Professor

Electrical Engineering

Government College of Engineering, Amravati, Maharashtra, India

Abstract-The protection of microgrids (MGs) is one of the most important and dangerous operational challenges with the gradual implementation of renewable energy sources in recent power systems. MGs are generally combined with photovoltaic (PV) arrays, wind turbines, fuel cells. Fault detection in MG is very complicated due to complex structures and so many bus bars available in MG, so fault detection and classification is necessary for MG operation and control, as it allows the system to perform fast fault isolation and recovery. Otherwise, MG component like transformer, loads, generators, and insulators may get damage due to long-duration faults presents in the system. In this paper, an intelligent fault detection method for MG based on wavelet transform (WT) and neural network (NN) is used. The main objective of the proposed scheme is to provide fast fault type information for MG protection and recovery. In this scheme, branch currents are pre-processed by discrete WT to extract statistical features. Then all the available data is given as input to NNs to developed fault information. All the tests are conducted on the IEEE 14 bus system.

Key words-Microgrid protection, fault detection, discrete wavelet transform (DWT), Multi-resolution analysis (MRA), neural network (NN)

1. INTRODUCTION

MG is defined as a network that includes distributed energy resources (DERs) such as distributed generators (DGs), storage devices, and small load clusters. MG can operate in a grid-connected mode or it can run in an islanded (non-grid connected) mode as a standalone System [1]. The main aim of the MG is to increase system efficiency, power quality, and reliability [2]. The use of different types of distributed generation (DG) for power generation presents challenges for operating, controlling, and protecting the MG [3]. With the slow implementation of renewable energy sources in recent power systems, MG are normally combined with inverter interfaced DGs (IIDG), such as photovoltaic DGs (PVDG) and battery energy storage systems (BESS). Old protective relays that are used for distribution system fault detection depend on huge fault currents. But, IIDGs can only offer unimportant fault currents such that the protection schemes are not initiated. So these relays fail to protect MG [5].

Fault detection in MG has three main objectives [6].

1. If there is any fault in the system, a fault detection scheme should detect the fault type (e.g. line to ground (LG), line to line (LL), etc.)
2. A fault detection scheme should detect the fault phase in unbalanced faults.
3. A fault detection scheme should determine the fault location where fault occurs.

In present years, a growing body of research uses data driven and digital signal processing methods for MG fault type, phase and location detection. Now a days, decision tree (DT) and random forest are widely used to detect faults in MGs (i.e. grid-connected and islanded MGs) [7]. Other machine learning techniques are support vector machine (SVM) and k-nearest neighbors (KNN) method, have also used for fault detection ([3], [8]). The computation speed of these approaches is very high and computation time requires is very less. Also fault classification can be done in real-time. Discrete Fourier transform (DFT) and discrete wavelet transform (DWT) are widely used to 'pre-process' the input signals to extract statistical features for analysis [3], [4], [8].

This paper presents an intelligent fault detection method based on DWT and NN, which is a set of data-driven machine learning techniques. DWT is used to pre-process the input signals to extract features. After that, the data is given to the NN for fault detection and classification. Fig. 1 shows the schematic block diagram of the proposed fault detection method. The proposed scheme takes branch

current magnitude of three phases as input data. The measurements are preprocessed by DWT to extract statistical features (time-frequency domain features). After that, all available data is given as input to NN for fault classification.

2. MICROGRID SYSTEM

In this paper, all the study has been done on IEEE 14 bus system [9]. Single line diagram of the IEEE 14 bus system is given in following figure 1. IEEE 14 bus system consists of 5 generators, 11 loads and 14 buses. The data used for design of this system is given in Appendix A. base values are as follows.

Base Power = 100MVA, Base Voltage = 11KV. This MG system is modeled in MATLAB software.

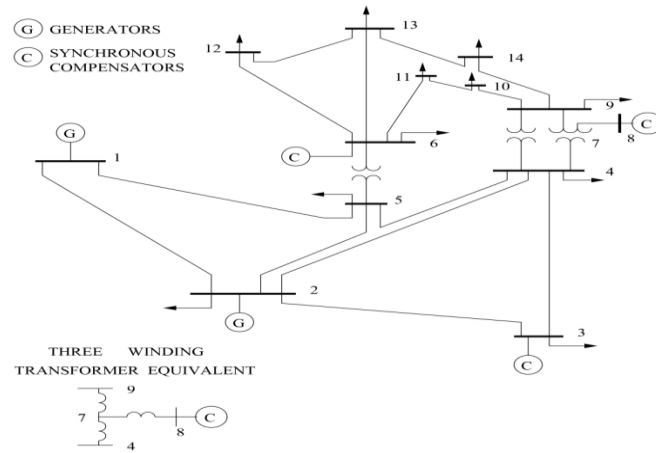


Fig 1:IEEE 14 bus system diagram

3. DISCRETE WAVELET TRANSFORM ANALYSIS

The word 'wavelet' means an oscillatory vanishing wave with time-limited extend, which has the power to explain the time-frequency plane with atoms of various time supports. DWT is a digital signal processing technique which transforms a time-series into a mutually orthogonal set of data. It can extract the hidden time-frequency domain characteristics of the fault current. In the proposed fault detection scheme, DWT serves an important role in preprocessing the input data for NNs in the scheme [5].

A. Continuous and discrete wavelet transform

Continuous wavelet transform (CWT) in [10] is one of the signal processing tools employed in analysis of various signals. The term multi resolution analysis (MRA) is closely associated with this signal processing tool. MRA analyses the signal at different frequencies with different resolutions. This makes this technique different from Short Time Fourier Transforms (STFT). STFT analyses the entire signal at same window length. This produces better resolution at some frequencies but poor at some other. MRA is meant to supply excellent time resolution and poor frequency resolution at high frequencies, and good frequency resolution and poor time resolution at low frequencies. Fortunately, most of the real signals analyzed in this work have high frequency components for short durations and low frequency components for long durations (fig 2). This is the reason why an MRA had been used in this signal analysis [11]-[12].

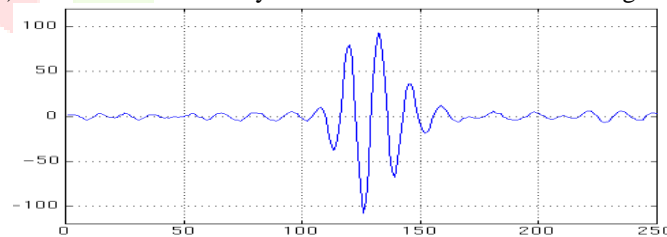


Fig 2:Real signal from vibration sensor

The continuous wavelet transform is computed by the following formula:

$$CWT^\psi(\tau, s) = \int_{-\infty}^{\infty} x(t) \psi^* \left(\frac{t-\tau}{s} \right) dt$$

The transformed signal is a function of two variables, τ and s , the translation and scale parameters respectively. $\Psi(t)$ is the transforming function, and it is called the mother wavelet. The mother wavelet implies to a small but finite length oscillatory function which is used as a prototype for generating the other window functions [11].

Multi-resolution analysis (MRA) was formulated based on the study of Bio-orthogonal, compactly supported wavelet bases. Wavelets can present multiple resolutions in both time and frequency. The signal is decomposed at distinctive resolution levels by the use of wavelet and scaling functions in multiple resolution analysis. The detail form of the decomposed signal will be generated by the wavelet function and the approximated form of the decomposed signal will be generated by the scaling function. It means that the wavelet function consists of the high pass filter and low pass filter is contained in the scaling function [13]-[14].

Let $V(t)$ is the original signal obtained from a measuring device. $V(t)$ is decomposed into detail and approximation. $cA1$ and $cD1$ are the decomposed signals of level 1 in the multi resolution technique as shown in figure 3. Where, $cA1$ is the approximation of the original signal and the $cD1$ is the detail version of the original signal. $cA1$ and $cD1$ are defined as given in equation.

$$cA1(t) = \sum_k V(t) \cdot L_d(k - 2t)[1]$$

$$cD1(t) = \sum_k V(t) \cdot Hd(k - 2t)[2]$$

Where,

Ld is the low-pass filter and Hd is the high-pass filter.

These filters are related to mother wavelet ψ . High frequency components of the signal are contained in cD1 whereas cA1 contains the low frequency components of the signal. When the original signal $V(t)$ is passed through the low pass and high pass filters, it gets decomposed into cA1 and cD1 coefficients of the signal. Here the cA1 and cD1 are level1 coefficients. Further approximation cA1 is decomposed into cA2 and cD2 coefficients of the signal which are the level 2 coefficients. And this procedure is repeated again and again until the required level is obtained for a particular application.

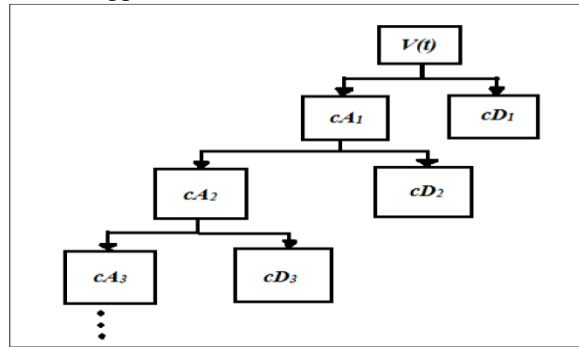


Fig 3:Wavelet decomposition tree

B. Mother Wavelet and Decomposition Level

Many wavelet families have adopted in previous work for DWT in MG fault detection, e.g. coiflets, daubechies, discrete meyer, haar, biorthogonal and symlets. While there must exist an optimal set of wavelet members in there families that will lead to the optimal performance for fault detection. This selection is based on the characteristics of analyzed data. When the data contain sufficient samples, the db and sym families are generally preferred. In this work we used db as mother wavelet to transform the input signal [5]-[15]. Besides the mother wavelets, decomposition level is another important parameter that affects the signal decomposition performance. A larger level may provide a more detailed description of the input signal. So here four decomposition levels are used.

4. FAULT DETECTION USING ANN

A. General Procedure for Fault Detection

A general flowchart of the NN-based classifier is shown in Fig 4. In a general, the three-phase input current is captured. Features that are essential statistical parameters are calculated, and these are given as an input to NN. Weights are initialized to little random values, using the various learning rules and cross-validation (CV) process, an error is calculated through the error back propagation algorithm. Precaution is taken so that network should not be stuck up in the local minima, and hence, according to the calculated error, weights are updated and the process is repeated until the global minimum is attained. After that NN is trained and tested carefully on different data sets, namely, testing, CV, and training data sets for the various performance measures. So in this way ANN is used here to classify different faults.

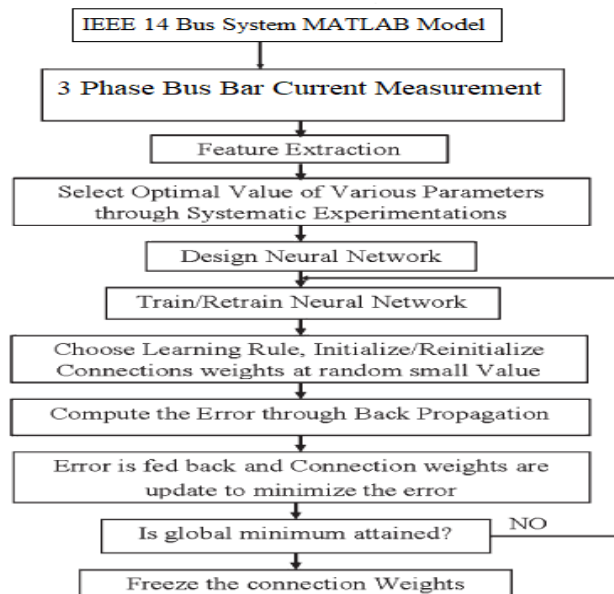


Fig 4: Schematic block diagram of the proposed scheme

B. Feature Selection and Extraction

This paper investigates four of the most common types of faults in the IEEE 14 bus system. The fault which we are classifying here is line to ground (LG), line to line (LL), double line to ground fault (LLG), three-phase fault (LLL). For experimentation and data collection, the IEEE 14 bus system is selected, whose details are given in [9].

Here using db4 as a mother wavelet DWT be able to decompose a sequence of the input signal into a series of coefficients $a_{M,k}$, and $d_{j,k}$. Choosing suitable features to represents the characteristics of the input signal is critical for fault detection [5]. There are various features available from which we can classify the faults which are given below:

- The max value of the coefficient: $\max\{s\}$
- The min value of the coefficient: $\min\{s\}$
- The mean value of the coefficient: $\mu_s = E[s]$
- The energy of the coefficient: $\sum s^2$

Here we use 's' to represent the coefficient $a_{M,k}$ or $d_{j,k}$.

So here in this paper feature selected for fault classification is energy. The energy of the signal is calculated using DWT. This feature vector is later inputted into the neural network to develop the fault detection results [5].

The wavelet energy is the sum of the square of detailed wavelet transform coefficients. The energy of the wavelet coefficient is varying over different scales counting on the input signals.

5. DESIGN OF NN-BASED CLASSIFIER

For the fault classification, RBF and MLP-based classifiers are designed, and their performances are listed in Table I. It is observed that the performances of these classifiers are not satisfactory. Thus, a cascade connection of RBF and MLP NN is proposed in order to design the classifier [16]. Where CV data means cross validation data

TABLE I
COMPARATIVE PERFORMANCES OF RBF AND MLP CLASSIFIERS

Classifier	Performance	Testing on Test Data	Testing on CV Data
RBF	MSE	0.113920085	0.097053
	Accuracy	55.20833333	59.52381
MLP	MSE	0.74507132	0.06345
	Accuracy	60.23456756	65.34281

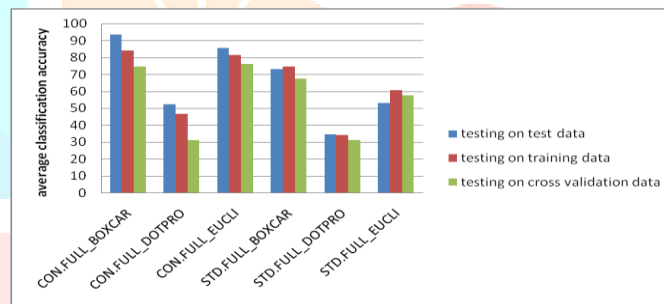


Fig 5: Variations of Average Classification Accuracy of RBF-MLP with Different competitive rule and metrics

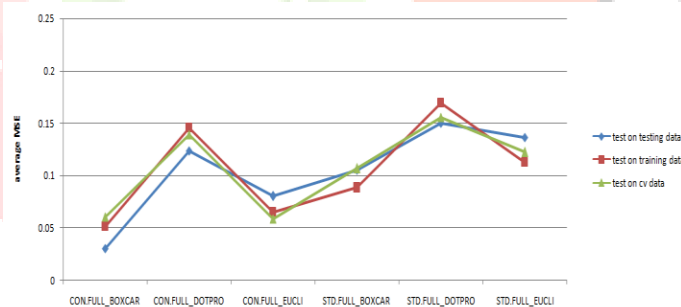


Fig 6: Variations of Average MSE of RBF-MLP with Different competitive rule and metrics

In the first layer, a RBF with a Gaussian function is designed with appropriate competitive rule and metric combination. Convergence to the global minimum depends upon the competitive rule and metric, and hence, the different competitive rules and metrics are verified for the convergence. From Fig 5 and Fig 6, it is observed that the ConscienceFull competitive rule and boxcar metric provide the optimum performance.

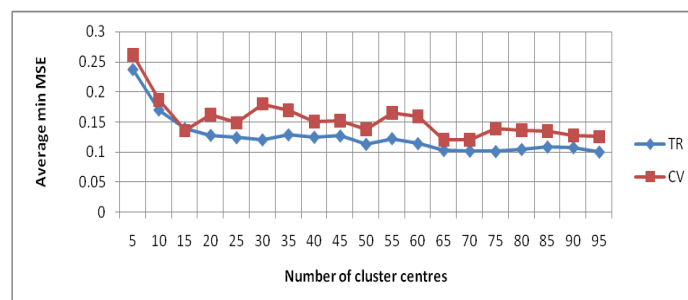


Fig 7: Variations of Training and CV MSE with Cluster Centers for RBF-MLP NN

It is shown in Fig 7 that there is no major change in the MSE beyond 65 cluster centers. Hence, 65 cluster centers are chosen because the more the number of cluster centers, the more will be the number of connection weights, and hence, the network will be more complex; thus, an optimal number of cluster centers should be chosen to keep the network simple.

The selection of PEs in the hidden layer, step size, and momentum rate of hidden and output layers is responsible for the global minimum, and it will decide the convergence rate to the minimum. Hence, these are the main important parameters to be selected. For the selection of optimal step size and momentum rate of hidden and output layers, several experimentations are performed, and results are shown in Figs 8 and 9.

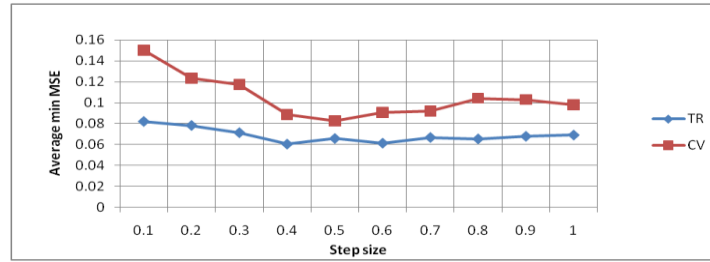


Fig 8: Variations of Training and CV MSE with step size for RBF-MLP NN

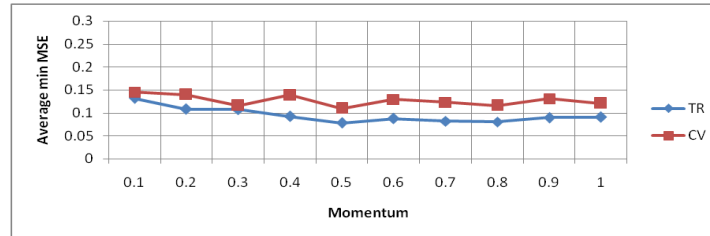


Fig 9: Variations of Training and CV MSE with momentum for RBF-MLP NN

Fig 8 shows the variations of average minimum training and CV MSE with a step size. For a small number of stepsize, we saw the variations and large values of MSE and as the step size increase the MSE goes on decreasing. It is seen that there is no significant change in the MSE beyond 0.4 step size. Hence 0.4 step size is chosen.

Fig 9 shows the variations of average minimum training and CV MSE with momentum. For a small number of momentum, we observed the variations and large values of MSE and as the momentum increase the MSE goes on decreasing. It is seen that there is no significant change in the MSE beyond 0.5 momentum. Hence 0.5 momentum is chosen.

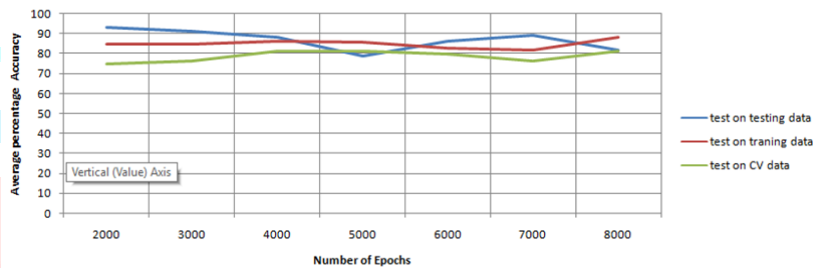


Fig 10: Variations of percentage accuracy on Training, testing and CV with Number of Epochs

From fig 10 it is observed that there is no significant change in training, testing and CV accuracy after 3500 epochs. It clear from the experimentation that, 4000 epochs must be selected, as predefined value for the training of network.

Finally, the required network is designed with following specifications:

- 1) Number of inputs: 15;
- 2) Stopping condition: 4000 epochs;
- 3) Competitive rule: ConscienceFull;
- 4) Number of hidden layers: 01;
- 5) Number of PEs in the hidden layer: 17;
- 6) Error criterion: L2 norm;
- 7) Metric: Boxcar;

Hidden Layer:

Transfer function: Tanh Learning Rule: Momentum
Step size: 0.4 Momentum: 0.5

Output Layer:

Transfer function: Tanh Learning Rule: Momentum
Step size: 0.1 Momentum: 0.7

- 8) Number of cluster centers: 65;
- 9) Training time required per epoch per exemplar: 0.054 ms;
- 10) Number of connection weights: 1304.

A. Training and Testing of NN

The next difficulty is to train the network and test it on the hidden data, i.e., testing data. As discussed earlier, the designed network must be more generalized. Different data sets are produced using variable split ratios and Leave-N-Out CV technique. The proposed NN is trained on various data sets and later validated carefully, to ensure that its performance does not depend on the specific-data-partitioning scheme. The performance of the NN must be every time-optimal over all the data sets for MSE and classification accuracy. Finally, the designed network is retrained five times with different random weight initializations and tested on testing, CV, and training data sets. Initially, small data sets are given for the training, and large data are kept for testing. Gradually, the size of training data is increased, and the performance of the classifier is observed, as shown in Figs 11 and 12. Finally, it is determined that about 70% of data must be used for training and the other 30% can be reserved for testing.

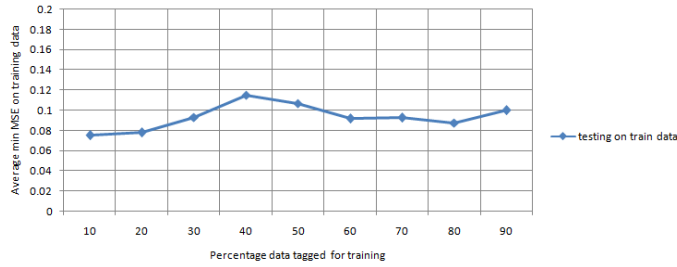


Fig 11: Variation of MSE with percent data tagged for training.

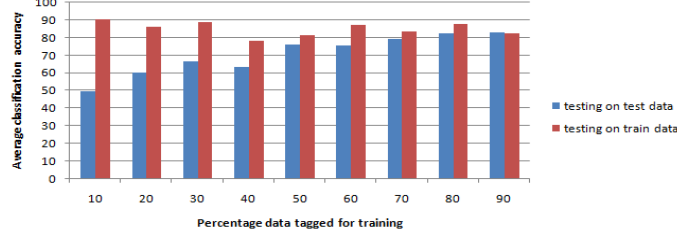


Fig 12: Variation of average classification accuracy with percent data tagged for training.

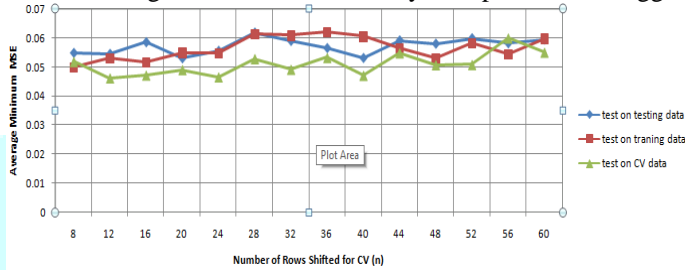


Fig 13: Variation of average minimum MSE with number of rows shifted for CV (n).

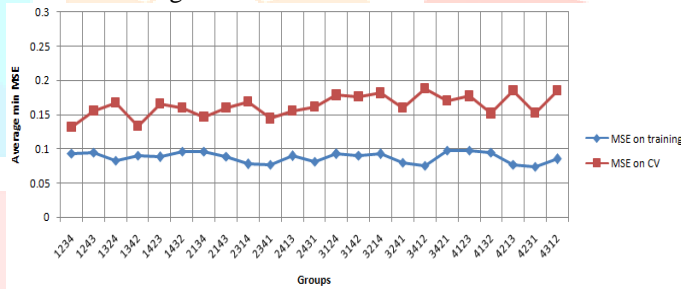


Fig14: Variation of average minimum MSE, when different data groups are used for training, CV, and testing.

For training and testing, the Leave-N-Out method and data tagging by various groups are used. Leave-N-Out training is a technique that allows one to estimate how well the model generalizes. It is also very useful for small data sets since it allows one to use the entire data set for training and testing. The algorithm trains the network multiple times, each time omitting a different subset of the data and using that subset for testing. To check the learning ability and classification accuracy, an additional 23 data sets are prepared. The network is carefully trained and tested by different methods, and results are as shown in Figs. 13–15.

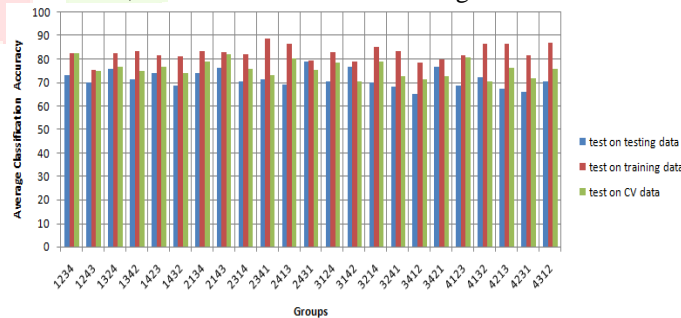


Fig 15: Variation of average classification accuracy, when different data groups are used for training, CV, and testing.

B. Sensitivity Analysis

The total numbers of inputs given to NN are 15. From fig. 15 we can see that out of 15 inputs only AD3, CD1, BD2, AD2, and BD3 these features or inputs decide the type of faults. These five parameters are most sensitive. While other parameters or features are not affecting the classification that much. So if we reduce the number of inputs then the time required for classification of fault is reduced. The complexity of the network is also reduced.

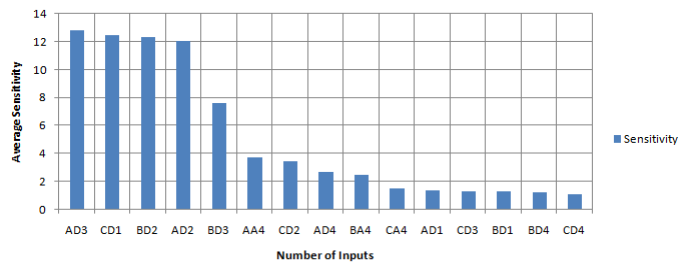


Fig 16: Sensitivity analysis

So 5 features i.e. AD3, CD1, BD2, AD2, and BD3 which are most sensitive are given to the network as an input as shown in fig. 16 and similar experimentations which are performed above are performed. By performing similar experimentation we have to design a new network. So step size, momentum, cluster center, number of PE's all are changed.

Finally, the second network is designed with following specifications:

- 1) Number of inputs: 5;
- 2) Stopping condition: 4000 epochs;
- 3) Competitive rule: ConscienceFull;
- 4) Number of hidden layers: 01;
- 5) Number of PEs in the hidden layer: 7;
- 6) Error criterion: L2 norm;
- 7) Metric: Boxcar;

Hidden Layer:

Transfer function: Tanh Learning Rule: Momentum

Step size: 0.7 Momentum: 0.8

Output Layer:

Transfer function: Tanh Learning Rule: Momentum

Step size: 0.1 Momentum: 0.7

- 8) Number of cluster centers: 45;
- 9) Training time required per epoch per exemplar: 0.02050 ms;
- 10) Number of connection weights: 454.
- 11) Percent reduction in weight: $(1304-454)/1304 = 65.18\%$
- 12) Percent reduction in time: $(0.054-0.02050)/0.054 = 62.32\%$

The above network is trained and tested in a similar way as the first network. All the results obtained indicate that the network's output closely follows the desired output. The advantage of reducing the network is the training time per epoch of the network is reduced by 62.32% and reduction in connection weight by 65.18%.

TABLE II
PERFORMANCE OF RBF-MLP CASCADE NN-BASED FAULT CLASSIFIER

Performance	Testing on test data	Testing on CV data
Classification Accuracy	91.6369048	90.3831169
MSE	0.05250344	0.0618758

6. CONCLUSION

A new approach to intelligent fault detection and classification of IEEE 14 bus MG based on RBF-MLP cascade NN has been proposed in this paper. Simple statistical characteristics, such as energy is extracted to derive rich faulty information from 3 phase bus current. In the design, the first layer of cascade NN, which is the RBF with ConscienceFull competitive rule and Boxcar metric with 65 cluster centers, is found to be the best. For the second layer of the network, the Momentum learning rule and Tanh transfer function give the optimal results. It has been found that the network is able to detect the faults in MG with average classification accuracies of 90% and 91% when tested on testing data and CV data, respectively. Performance measures are shown in table II. The training time required per epoch per exemplar is 0.054 ms, which indicates that the network is fast enough. From the results, it is observed that a minimum 70% of data is required for training so that the results are optimum. With 70 % data given for training and 30 % data reserved for testing. By reducing the number of inputs that are given to NN it is observed that the training time required per epoch is reduced by 62.32% and reduction in connection weight by 65.18%. This indicates that network complexity is also reduced.

REFERENCES

- [1] E. Casagrande, W. L. Woon, H. H. Zeineldin, and D. Svetinovic, "A differential sequence component protection scheme for microgrids with inverter-based distributed generators," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 29–37, Jan. 2014.
- [2] R. F. Arritt and R. C. Dugan, "Distribution system analysis and the future smart grid," *IEEE Trans. Ind. Appl.*, vol. 47, no. 6, pp. 2343–2350, Nov./Dec. 2011.
- [3] T. Abdelgayed, W. Morsi, and T. Sidhu, "A new approach for fault classification in microgrids using optimal wavelet functions matching pursuit," *IEEE Trans. Smart Grid*, to be published. [Online].
- [4] R. Mohanty, U. S. M. Balaji, and A. K. Pradhan, "An accurate noniterative fault-location technique for lowvoltage DC microgrid," *IEEE Trans. Power Del.*, vol. 31, no. 2, pp. 475–481, Apr. 2016.
- [5] James J. Q. Yu, Albert Y. S. Lam, "Intelligent Fault Detection Scheme for Microgrids with Wavelet Based Deep Neural Network," *IEEE Trans. Smart Grid*, vol. 10, no. 2, March 2019.
- [6] R. M. Cheney, J. T. Thorne, and G. Hataway, "Distribution single phase tripping and reclosing: Overcoming obstacles with programmable recloser controls," in *Proc. 62nd Annu. Conf. Protect. Relay Eng.*, Austin, TX, USA, Mar. 2009, pp. 214–223.
- [7] Guillén, Daniel, et al. "Algorithm for transformer differential protection based on wavelet correlation modes." *IET Generation, Transmission & Distribution* 10.12 (2016): 28712879.
- [8] S. Dhar, R. K. Patnaik, and P. K. Dash, "Fault detection and location of photovoltaic based DC microgrid using differential protection strategy," *IEEE Trans. Smart Grid*, to be published. [Online]. Available: <http://ieeexplore.ieee.org/document/7820172/>
- [9] D. K. Ray, S. Chattopadhyay, K. D. Sharma and S. Sengupta, "Steady state harmonic stability analysis in IEEE 14 bus system for fault at generator bus," *Michael Faraday IET International Summit 2015, Kolkata, 2015*, pp. 261-266.
- [10] A. Graps, "An Introduction to Wavelets", in *Proc. IEEE Computational Science and Engineering, Summer Meeting, 1995*, vol. 2, no. 2, pp. 50-61.
- [11] The Wavelet tutorial, Robi Polikar, <http://users.rowan.edu/~polikar/wavelets/wtpart1.html>
- [12] Saravanababu, K., Balakrishnan, P. and Sathiyasekar, K., "Transmission Line Faults Detection, Classification, and Location Using Discrete Wavelet Transform," *International Conference on Power, Energy and Control (ICPEC)*, pp. 233-238, 2013
- [13] S. G. Mallat, "A theory for multi resolution signal decomposition: The wavelet representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 7, pp. 674–693, Jul. 1989.
- [14] Youssef, O.A.S., "Fault classification based on wavelet transform," *Transmission and Distribution Conference and Exposition IEEE/PES*, vol. 1, pp. 531-536, 2001.
- [15] Reddy, R.B., Kumar, M.V., Suryakalavathi, M. and Prasanth Babu, C., "Fault detection, classification and location on transmission lines using wavelet transform," *Electrical Insulation and Dielectric Phenomena*, pp.409-411, Oct., 2009.
- [16] Ghate, V. N., & Dudul, S. V. (2009). "Fault Diagnosis of Three Phase Induction Motor Using Neural Network Techniques," *2009 Second International Conference on Emerging Trends in Engineering & Technology*. doi: 10.1109/icetet.2009.100