Fault classification on transmission line using LSTM network

Abdul Malek Saidina Omar, Muhammad Khusairi Osman, Mohammad Nizam Ibrahim,
Zakaria Hussain, Ahmad Farid Abidin
Faculty of Electrical Engineering, Universiti Teknologi MARA, Malaysia

Article Info

Article history:
Received Feb 10, 2020
Revised Apr 15, 2020
Accepted Apr 19, 2020

Keywords:
Artificial neural network
Deep learning
Long short-term memory
Signal processing
White gaussian noise

ABSTRACT

Deep Learning has ignited great international attention in modern artificial intelligence techniques. The method has been widely applied in many power system applications and produced promising results. A few attempts have been made to classify fault on transmission lines using various deep learning methods. However, a type of deep learning called Long Short-Term Memory (LSTM) has not been reported in literature. Therefore, this paper presents fault classification on transmission line using LSTM network as a tool to classify different types of faults. In this study, a transmission line model with 400 kV and 100 km distance was modelled. Fault free and 10 types of fault signals are generated from the transmission line model. Fault signals are pre-processed by extracting post-fault current signals. Then, these signals are fed as input to the LSTM network and trained to classify 10 types of faults. The white Gaussian noise of level 20 dB and 30 dB signal to noise ratio (SNR) is also added to the fault current signals to evaluate the immunity of the proposed model. Simulation results show promising classification accuracy of 100%, 99.77% and 99.55% for ideal, 30 dB and 20 dB noise respectively. Results has been compared to four different methods which can be seen that the LSTM leading with the highest classification accuracy. In line with the purpose of the LSTM functions, it can be concluded that the method has a capability to classify fault signals with high accuracy.

1. INTRODUCTION

Electrical power system faults lead to electrical supply interruptions, which increase maintenance costs including device failure and recovery costs. Fault in transmission lines accounts for more than 80% of all power system faults [1, 2]. Fault in transmission lines consists of four different categories known as Single Line to Ground Fault (SLGF), Double Line to Ground Fault (DLGF), Line to Line Fault (LLF) and Three Line Fault or Three Line to Ground Fault (LLL or LLGF) [3, 4]. All the fault categories extended to 10 different types of fault on transmission line for three phase electrical power system. The fault diagnosis cycle will start with fault detection and proceed with the fault classification process, which refers to the ability to detect and classify types of fault. Then, the process will lead to the location estimation, known as fault location, in order to locate where the fault has occurred. A transmission lines fault diagnosis can be performed through three different methods that have been used which are known as prominent, hybrid and modern methods. The prominent method consists of three well-known approaches which are the wavelet approach, the artificial neural network (ANN) approach and the fuzzy logic approach. Hybrid methods are a combination of more than one of the prominent methods to produce new four techniques. Modern methods use others techniques such as support vector machines, genetic algorithms, decision trees etc as a tool to classify fault on transmission lines [5, 6]. The wavelet approach is a numerical tool used for signal processing...
to select wavelet functions known as “mother wavelet”. A recent study has been made involving discrete wavelet transform for fault location identification on a double circuit transmission line. The purpose of the study is to analyse fault location algorithm detection capability and accuracy using MATLAB-Simulink with several types of fault. The simulation result showed small percentage error by using discrete wavelet transform which starts with a 0.025% margin of error for a 20km fault location and increased up to a 0.25% margin of error for a 200km fault location [7]. Another study that used the same approach has been performed to classify fault type with time delay value as a main feature. The algorithm has been successfully tested with several types of fault, fault resistances, fault location and inception times to see the system reliability [8]. However, the method suffers in the mathematical calculation analysis to produce a signification result.

These days, the ANN approach [9-11] and fuzzy logic approach [12-14] also known as artificial intelligent (AI) techniques, have led to an improved output in terms of system robustness and a highly accurate network. Typically, the AI methods required feature extraction to extract important feature information for diagnostic purposes. Since the transmission line fault is easily influenced by several parameters such as fault type, fault resistance and inception angle, the best features selection is a must in order to improve the overall system performance in terms of robustness and accuracy. Another advance approach known as hybrid technique that combines more than one of the previous methods is also used to perform the fault analysis. For instance, a neuro-fuzzy technique called Adaptive based Fuzzy Inference System (ANFIS) is used to perform fault classification with excellent classification results [15-17]. Another hybrid technique combination used are wavelet and ANN approach (DWT-ANN) which was applied to the 765kV transmission line fault classification [18]. While modern techniques such as support vector machine are becoming another option of diagnostic tool for fault classification. The algorithm has been tested and performed on IEEE 34 bus and IEEE 123 bus system which is able to classify types of fault with high accuracy [19]. Although the methods show good performance, the techniques suffer from features extraction and complicated signal processing.

Recently, Deep Learning (DL) also knowns as Deep Neural Network (DNN) has become popular for fault diagnosis due to a good accomplishment record and has successfully improved the performance for most of the research fields. For example, the application of the autoencoder with unsupervised features learning and automatic features learning technique for transmission line fault detection and classification [20, 21]. The application of Adaptive Deep Belief Network (ADBN) that was used to identify overhead transmission line fault cause [22]. Finally, the application of Convolutional Neural Network (CNN) for fault location on 500kV HVDC transmission line system [23], voltage sag estimation approach on the IEEE 68 bus network [24] and fault location estimation on AC transmission line for back to back MMC-HVDC system [25]. However, over the application of DNN in power system protection schemes, the Long Short-Term Memory (LSTM) has not been tested yet, compared to the other tools on DNN methods. The LSTM network has a capability to remember a uniqueness of the fault signals and very suitable for time-series data as transmission line fault signals. Therefore, this paper presents fault classification on transmission line using LSTM network techniques to classify 10 different types of fault. The method will be tested with various types of fault condition such as fault resistance, fault location and inception angle. It also will be tested with 3 different sets of data which consists of ideal dataset and noise added dataset at 30 dB and 20 dB white Gaussian Noise to measure the robustness of the network. The simulation will use MATLAB-Simulink as a software to perform the analysis through transmission line models as a fault signal generation. This paper has been organized into four main sections. Section 2 presents the proposed method for fault classification using LSTM network. Section 3 gives the results and discussion. Finally, Section 4 concludes the outcome of the study.

2. METHODOLOGY

Fault classification on transmission lines using DL method consists of five main process flows which are transmission line model development, fault signal generation and data collection, data pre-processing, fault classification using LSTM technique and performance assessment. A block diagram of the workflow for transmission line fault classification can be shown in the block diagram as Figure 1.

![Figure 1. Fault classification with DL method](image-url)
2.1. Model development and data collection

In this study, a transmission line model has been developed by using MATLAB-Simulink. The model is used as a simulation platform to simulate the transmission line using various types of fault conditions which is influenced by fault location, fault resistance and inception angle. Figure 2 shows the single ended transmission line model used in this study. The model consists of two three phase sources and two three phase section lines with a 400-kV voltage supply and 100 km distance between the two different sources. Simulation data is generated through the transmission line model for fault and fault free current signals collection. The simulation data consists of faulty datasets which are assigned with fault numbers from one to ten to represent 10 different types of fault. While the fault free dataset has been assigned with zero to represent non-fault signal. The simulation run has been performed up to 20 times for training and 40 times for testing purposes which replicate every single type of fault including non-fault signal. Every single sample will indicate with different types of fault location, fault resistance and fault angle for both faulty and non-fault signal.

![Figure 2. Transmission line model on MATLAB-simulink](image)

2.2. Data pre-processing

Firstly, preprocessing data involved data collection and data being stored on the data frame with specific variable names as fault type, fault number, fault parameters, simulation run, sample number and three phase current amplitude measurement. Fault types indicated 10 different fault classes listed as Ag Fault, Bg Fault, Cg Fault, AB Fault, BC Fault, CA Fault, ABg Fault, BCg Fault, CAg Fault and ABC/ABCg Fault. While fault number indicates the number that varies from zero to ten. A fault number zero represents non-fault signals whereas fault numbers one to ten represent 10 different fault types on the transmission line. The simulation parameters have been set at various fault conditions which reflect fault location, fault resistance and inception angle which were purposely performed to measure system reliability. The simulation run indicates the number of times that has been obtained for every training and test dataset. The number of simulations has been simulated up to 20 times for non-fault data and 40 times for faulty data. The sample number indicates the number that has been recorded for every single simulation with up to 50 samples each. Then, three phase current amplitude measurement contained measured value through simulation. Figure 3 shows sample fault signals for Single Line to Ground Fault (SLGF), Line to Line Fault (LLF) and Three Line Fault (LLL) that occurred on the transmission line.

The final dataset which consists of training, validation and testing data will be used to classify the correct fault number. Long Short-Term Memory (LSTM) will be used in this study as a deep learning tool to classify all the faulty and non-faulty signals. Basically, LSTM provides a high level of performance to time series data which able to remember the uniqueness of the previous signals as a reference to classify the new signals that feed to the network. The training dataset that is collected will be divided into 1:4 ratio for validation and training purposes whereas the testing data remains at 100 percent. Mean and standard deviation is applied to all simulation signals in the training and testing dataset for data normalization. This condition will make sure all the data value in the dataset at the standard scale to ensure the variables with a small value is not dominated by variables with larger value.
2.3. Network architecture

The classification process will be the network architecture design stage. In this stage, the input layer has been set to the same value of the input signal as the sequence input layer. The hidden layer of the network consists of three LSTM hidden layers which are set to the same number of input signals for the first hidden layer followed by 40 and 25 for the next two layers. In order to prevent the network from overfitting, a dropout layer has been set between each of the LSTM layers that produce another three dropout layers. Then, a fully connected layer has been fitted at the final part of the network for classification purposes which has been set with the same size of the number of faulty and non-faulty conditions equal to 11 for this research. A SoftMax layer followed by a classification layer have been set after the fully connected layer for possibility prediction for each fault class and fault type as output prediction purpose respectively. Finally, network training stage will be proceeded after network setting completed together with validation process and network testing by using testing dataset to classify the fault type on the transmission line.

3. RESULTS AND ANALYSIS

In this section, the performance of the LSTM network for transmission line fault classification will be discussed. There are three different fault conditions that have been used to measure the LSTM network through ideal dataset and noise added dataset which was set at 30 dB and 20 dB white Gaussian noise. Figure 4 shows faulty signals without noise and with noise at 30 dB and 20 dB for sample Ag Fault. It can be seen that the lower value of the white Gaussian noise will show more distortion to the fault signals. This condition will increase difficulties of the classification task due to the signal distortion. Figure 5 shows a sample of the network training together with confusion chart as predicted output classes compared to true fault classes. From the figure, it shows the training achieved the 100% of the classification accuracy at iteration number of 50. The confusion chart provides a useful information of the classification performance results for each of the fault class. The overall classification accuracy of the network then calculated through the confusion chart output to visualize the percentage of the accuracy of the testing data predictions. Table 1 shows the average accuracy of the classification result by using the LSTM network performance for three different sets of datasets. It can be seen that the accuracy of the LSTM network shows a decrease from 100% to 99.77% and 99.55% for ideal to 30 dB and 20 dB added noise data respectively. However, the accuracy of the network shows high accuracy value for the overall performance.
Fault classification on transmission line using LSTM network

(Faithful Malek Saidina Omar)

Figure 4. Faulty signals without and with 30 dB and 20 dB white gaussian noise

Figure 5. Confusion matrix of the LSTM classification network

Fault classification on transmission line using LSTM network… (Abdul Malek Saidina Omar)
Table 1. Overall testing accuracy of LSTM method

| Fault Number | Fault Types | Ideal | 30 dB | 20 dB |
|--------------|-------------|-------|-------|-------|
| 0            | Non-Fault   | 100.00| 100.00| 100.00|
| 1            | Ag Fault    | 100.00| 100.00| 100.00|
| 2            | Bg Fault    | 100.00| 100.00| 100.00|
| 3            | Cg Fault    | 100.00| 100.00| 100.00|
| 4            | AB Fault    | 100.00| 100.00| 100.00|
| 5            | BC Fault    | 100.00| 100.00| 100.00|
| 6            | CA Fault    | 100.00| 100.00| 100.00|
| 7            | ABg Fault   | 100.00| 97.50 | 100.00|
| 8            | BCg Fault   | 100.00| 100.00| 97.50 |
| 9            | CAg Fault   | 100.00| 100.00| 97.50 |
| 10           | ABC/ABCg Fault | 100.00| 100.00| 100.00|

Average Accuracy 100.00 99.77 99.55

The performance of the LSTM network has been compared with four different proposed methods for fault classification on transmission line which are [26-29]. Table 2 shows the accuracy of fault classification for three different datasets on every single method that has been compared. From the table it can be seen that the LSTM leading with the highest accuracy at 100%. 99.77% and 99.55% for ideal, 30 dB and 20 dB datasets. The result shows that the LSTM network achieved the highest classification accuracy compared to the other four different methods that were used. From these results, it can be concluded that the LSTM network has a capability to perform the fault classification on transmission line. The results also shown better performance compared to other techniques and is more reliable to noise conditions.

Table 2. Performance comparison with LSTM network

| Method Reference | Testing Accuracy (%) | Ideal | 30 dB | 20 dB | Average |
|------------------|----------------------|-------|-------|-------|---------|
| [26]             | -                    | 97.45 | 97.45 |       |
| [27]             | 99.20                | -     | 98.30 | 98.75 |
| [28]             | -                    | 99.37 | -     | 99.37 |
| [29]             | 99.52                | 99.37 | 99.19 | 99.36 |
| LSTM Method      | 100.00               | 99.77 | 99.55 | 99.77 |

4. CONCLUSION

This research has presented fault classification on transmission line using LSTM network. A power transmission line model has been simulated on MATLAB-Simulink to generate fault signals as a training and testing data to the LSTM network. The simulated data consists of various types of fault conditions that influence the fault signals named as fault location, fault resistance and inception angle. Overall network performance has been evaluated through three different types of dataset which consists of ideal dataset and noise added dataset with 30 dB and 20 dB white Gaussian noise. Through the simulation, the result showed that the accuracy of fault classification was decreased with the increasing of the noise value. However, the accuracy results still at the highest level of classification compared to the other methods as explained in the previous section. Therefore, it can be concluded that the LSTM network has the ability to classify 10 different types of fault on transmission lines with high accuracy for three types of datasets that have been fed into the network. For future work, LSTM network can be used to perform on transmission line fault detection since the system required fault detection prior classification to detect fault occurrence. Other than that, it also can be applied to the fault location estimation in order to complete the cycle of transmission line fault protection scheme.

ACKNOWLEDGEMENTS

The financial support of this research was provided by Universiti Teknologi MARA, Cawangan Pulau Pinang, Malaysia.

REFERENCES

[1] S. A. Aleem, N. Shahid, and I. H. Naqvi, “Methodologies in power systems fault detection and diagnosis,” Energy Syst., vol. 6, no. 1, pp. 85–108, 2014, doi: 10.1007/s12667-014-0129-1.
Fault classification on transmission line using LSTM network... (Abdul Malek Saidina Omar)
BIOGRAPHIES OF AUTHORS

Ir. Ts. Abdul Malek Saidina Omar was born in Malaysia on January 14th, 1983. He obtained Bachelor Degree of Electrical Engineering (Hons) from Universiti Tun Hussein Onn Malaysia (UTHM) in 2006 and hold a Master of Science in Electrical Power Systems from University of Birmingham United Kingdom in 2014. His main research interest includes Power System Protection and Artificial Intelligent. He is currently a Senior Lecturer at the Faculty of Electrical Engineering, Universiti Teknologi Mara (UiTM) Malaysia since November 2015.

Dr. Muhammad Khusairi Osman obtained his B.Eng Degree in Electrical and Electronic Engineering in 2000 and MSc in Electrical and Electronic Engineering in 2004 from Universiti Sains Malaysia (USM). In 2014, he obtained his Ph.D. in Medical Electronic Engineering from Universiti Malaysia Perlis (UniMAP), Malaysia. He is currently a Senior Lecturer at Faculty of Electrical Engineering, Universiti Teknologi MARA (UiTM), Malaysia. His research interest is in Image Processing, Pattern Recognition and Artificial Intelligence.

Ts. Dr. Mohammad Nizam Ibrahim obtained a Degree of Electrical Engineering (Hons) from University of Portsmouth in 1996 and Master of Science in Engineering Electronic System Design from Universiti Sains Malaysia in 2005. In 2012, he completed his PhD study in Power Protection from University of Adelaide, South Australia. Currently, he is a Senior Lecturer at the Faculty of Electrical Engineering, Universiti Teknologi Mara (UiTM) Malaysia. His main research includes Power System Transmission Line Protection, Uncertainty and Sensitivity Analysis and Transmission Lines Fault Simulation.

Assoc. Prof. Ir. Dr. Zakaria Hussain was born in Malaysia on April 19th, 1975. He obtained his Bachelor Degree in Electrical and Electronics Engineering from The University of Huddersfield, United Kingdom in 1997 and hold a Doctor of Philosophy in Automatic Control and System Engineering from The University of Sheffield, United Kingdom in 2010. His main research interest includes Intelligent Control and Artificial Intelligent. He is currently a Senior Lecturer at the Faculty of Electrical Engineering, Universiti Teknologi Mara (UiTM) Malaysia since September 1999.

Assoc. Prof. Ir. Dr. Ahmad Farid Abidin was born in Malaysia on Dec 25, 1978. He obtained Bachelor of Engineering in Electrical, Electronic, and System Engineering in 2000 from Universiti Kebangsaan Malaysia, MSc in Electrical Engineering in 2005 from Universiti Teknologi MARA (UiTM) Malaysia, and PhD in Electrical Engineering from Universiti Kebangsaan Malaysia in 2011. His main research interests are in power system stability, power quality and power system protection. He is currently a Senior Lecturer at the Faculty of Electrical Engineering, UiTM and the Director, Community of Research (CoRe), Advanced Computing and Communication (ACC), Institute of Research Management and Innovation, UiTM since Feb 2018.