DETECTION, CLASSIFICATION AND IDENTIFYING SOURCE LOCATION OF SHORT CIRCUIT FAULTS IN MODERN POWER SYSTEM USING MACHINE LEARNING

A Project Report Submitted

by

ADITYA KUMAR

(2001EE85)

In Partial Fulfilment

of the Requirements for the award of the degree

BACHELOR OF TECHNOLOGY

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY PATNA

6TH OCTOBER 2023

THESIS CERTIFICATE

This is to certify that the thesis titled **DETECTION, CLASSIFICATION AND IDENTIFYING SOURCE LOCATION OF SHORT CIRCUIT FAULTS IN MODERN POWER SYSTEM USING MACHINE LEARNING**, submitted by **Aditya Kumar,** to the Indian Institute of Technology, Patna, for the award of the degree of **Bachelor of Technology**, is a bonafide record of the research work done by him under our supervision. The contents of this thesis, in full or in parts, have not been submitted to any other Institute or University for the award of any degree or diploma.

Dr. S.K Parida
Supervisor
Associate Professor
Dept. of Electrical Engineering
IIT-Patna, 800 013

Place: Patna

Date: 6th October 2023

ACKNOWLEDGEMENTS

I want to express my gratitude to Dr.S.K.Parida, my BTP supervisor, for his invaluable guidance throughout the project. His support in conceptualizing the project and providing essential resources to tackle the problem statement was highly appreciated.

ABSTRACT

Transmission lines play a vital role in power systems by transmitting power from the source to the load end. However, faults in these overhead transmission lines are frequent events that can disrupt the power system's smooth operation. It is crucial to swiftly and accurately classify faults and pinpoint the exact fault location to isolate the affected area. This is essential for maintaining the reliability of the power system. Till now, in this report, an 11-bus system has been simulated through MATLAB Simulink to create the fault in different regions. After carrying out the fault simulation, the results of faulty voltages and currents at each phase at different locations have been taken.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

The electrical power system comprises three main subsystems: generation, transmission, and distribution. Transmission and distribution lines, being exposed to the environment, are susceptible to a range of faults. In normal circumstances, the power system carries standard voltage and current. However, faults, which are abnormal conditions, can occur between phases or between phases and ground in transmission and distribution lines. These faults cause a drop in voltage at the point of fault and result in excessive current flow, leading to overheating and damage to the power system's equipment. Analyzing faults is crucial due to the maintenance challenges, inconvenience to users, and financial losses that often result from electrical faults. Rectifying these faults can be a time-consuming process, sometimes taking hours or even days.

Transmission line faults are usually divided into two categories: symmetrical and unsymmetrical faults. The single line-to-ground (S-L-G), double line-to-ground (L-L-G), and line-to-line (L-L) faults are considered unsymmetrical faults, while the three-phase faults (L-L-L-G and L-L-L) are considered as the symmetrical faults in power system.

After going through several literature we found that either their analysis was done on a small sub-system or they took data from a particular fault location only instead of taking it from multiple non-fault locations as well. In this report, we discussed the way we generated data from a 2-area 11-bus system considering all the parameters and cases.

In the future, we will be developing a robust and scalable machine learning model that can detect whether the fault has happened or not, detect the fault location if happened, and the type of fault that has happened.

1.1 Types of Fault

A fault within the power system refers to a deviation from the intended pathway, leading to a disruption in the flow of current. The fault in the power system is mainly categorized into two types, they are: -

1. Open circuit fault (Series Fault)
2. Short circuit fault (Shunt Fault)

1.1.1 Open circuit fault

The open circuit fault mainly occurs because of the failure of one or two conductors. The open circuit fault takes place in series with the line, and because of this, it is also called the series fault. The open circuit fault is categorized as: -

- Open Conductor Fault
- Two conductors Open Fault
- Three conductors Open Fault

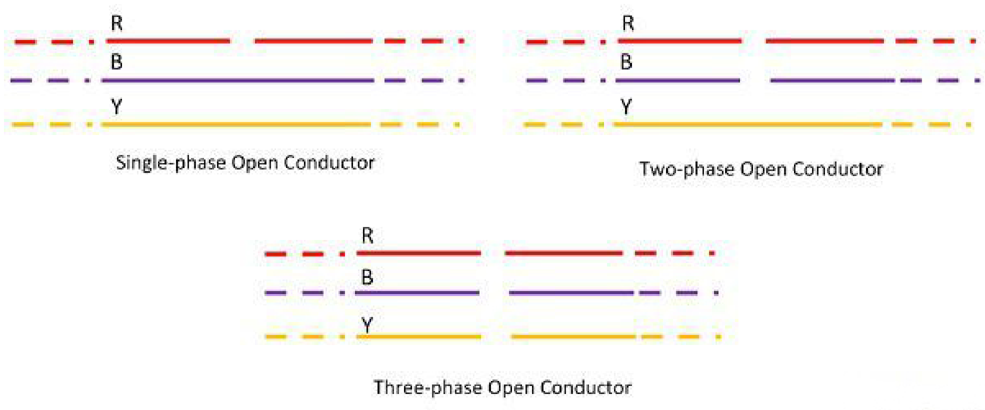

Figure 1.1: Open Circuit Faults

1.1.2 Short circuit fault

In this fault scenario, conductors from distinct phases make unintended contact, typically involving a power line, power transformer, or another circuit element. This contact results in a significant surge of current within one or two phases of the system. Short-circuit faults are categorized into symmetrical and asymmetrical faults based on their characteristics.

Symmetrical Faults: -

The faults which involve all three phases are known as the symmetrical fault. Such types of fault remain balanced even after the fault. The symmetrical faults mainly occur at the terminal of the generators.

1. **Line-Line-Line (L-L-L) Fault:** The L–L–L fault occurs rarely, but it is the most severe type of fault that involves the largest current. This large current is used for determining the rating of the circuit breaker.

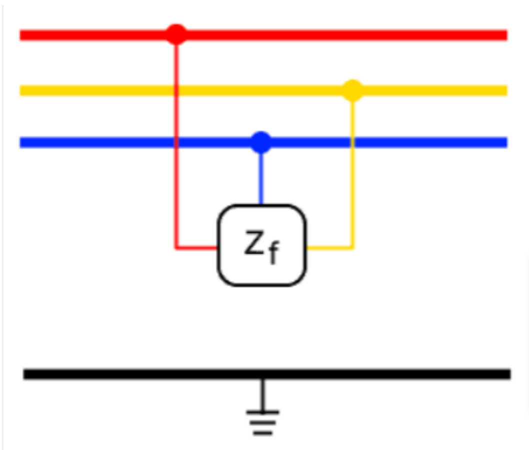

Figure 1.2: L-L-L Fault

2. **Line-Line-Line-Ground (L-L-L-G) Fault:** The L–L–L–G fault occurs between the three phases and the ground of the system. The probability of occurrence of such type of fault is nearly 2 to 3 percent.

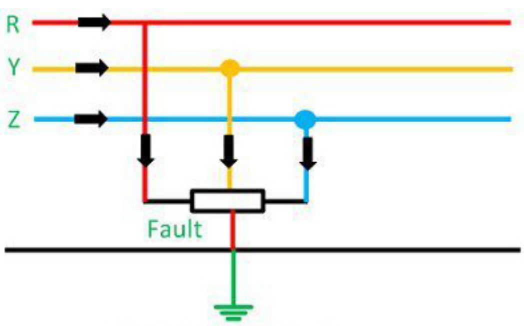

Figure 1.3: L-L-L-G Fault

Unsymmetrical Faults: -

The fault gives rise to unsymmetrical current, i.e., current differing in magnitude and phases in the three phases of the power system are known as the unsymmetrical fault. The unsymmetrical makes the system unbalanced.

1. **Single Line to Ground (S-L-G) Fault:** The single line of ground fault occurs when one conductor falls to the ground or contacts the neutral conductor. 70 – 80 percent of the fault in the power system is the single line-to-ground fault.

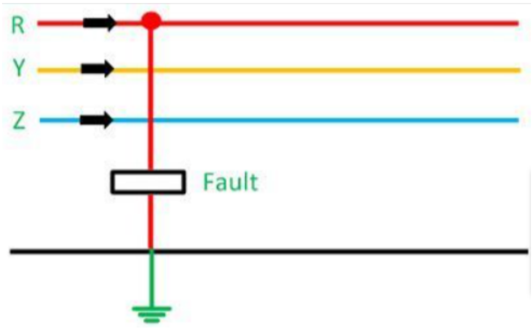

Figure 1.4: S-L-G Fault

2. **Line-Line (L-L) Fault:** A line-to-line fault occurs when two conductors are short circuited. The major cause of this type of fault is the heavy wind. The heavy wind swings the line conductors which may touch together and hence cause a short-circuit. The percentage of such types of faults is approximately 15 – 20 percent.

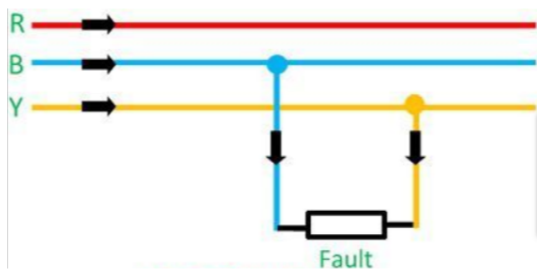

Figure 1.5: L-L Fault

3. **Line-Line-Ground (L-L-G) Fault:**In a double line-to-ground fault, the two lines come in contact with each other along with the ground. The probability of such types of faults is nearly 10 percent.

CHAPTER 2

LITERATURE REVIEW

1. In [1], the use of ANN for fault classification and detection of faults in power systems has been thoroughly investigated. The phase currents, phase voltages, and zero sequence components of a system model created with Simulink/MATLAB were used to apply the ANN algorithm. The data obtained from the different fault models was fed to the Artificial Neural Network for both training and testing purposes. The results were impressive. They can also reach better results by increasing the number of buses, creating more fault locations, and finding exact locations of fault.

2. In [2], a three-phase medium power transmission line system under study was converted into a Pi-model and simulation using the MATLAB/ Simulink® environment. Simulated values of transmission systems are used as training data for neural networks. Feed-forward BPNN algorithms were used for fault classification and detection. Performance analysis of three BPNN algorithms was done and results are presented.

Paper Name	**Publication**	**Methodology**	**Research Gap**
Fault Detection and Classification in Power System using ANN.	2nd International Conference on Intelligent Technologies (CONIT),2022	Uses 2 ANN models, one for detection and one for classification.	They can reach better results by increasing the number of buses, creating more fault locations, and finding exact locations of fault.
Fault Detection and Classification in Power Transmission Lines using BPNN.	2020 International Conference on Smart Electronics and Communication (ICOSEC)	Uses BPNN model with Levenberg-Marquardt (LM) algorithm for fault detection and classification.	The study is done on a small 3-phase transmission line for only A-G and A-B-G type faults.

Table 2.1: Summary of Literature review.

CHAPTER 3

SIMULATION WITH TWO AREA SYSTEM

In this project, we have used the famous 2 area kundur system for fault analysis at various regions. The system consists of 2 areas where each area consists of 4 buses, 2 generators, and 2 transformers. The two areas have been bridged through 3 additional buses, making it total of 11 buses in use.

The whole system is divided into 6 regions. Region 5-6, Region 6-7, Region 7-8, Region 8-9, Region 9-10, Region 10-11 of lengths 25 km, 10 km, 110 km, 110 km, 10 km, 25 km respectively.

After simulation, various faults have been observed through below waveforms: -

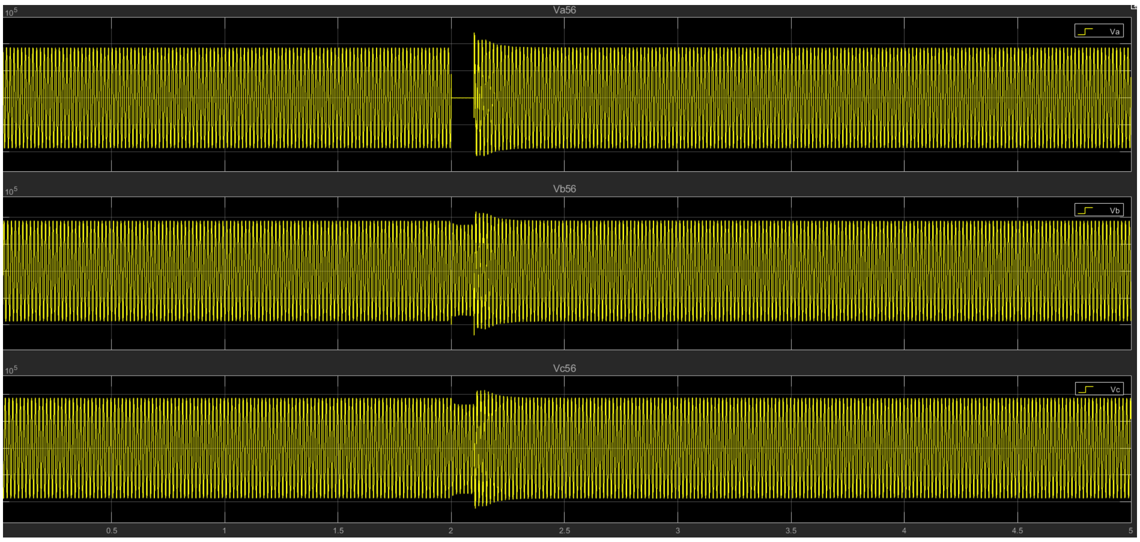

Figure 3.1: Voltage waveform after SLG Fault on line A.

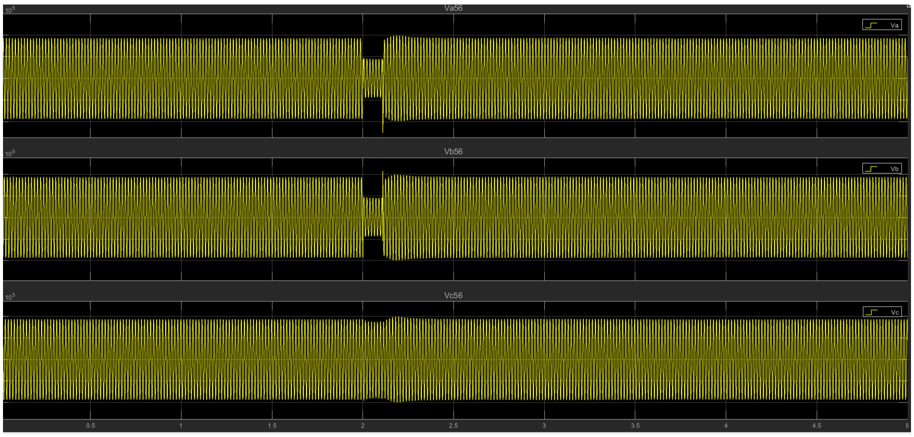

Figure 3.2: Voltage waveform after LL Fault on line A and B.

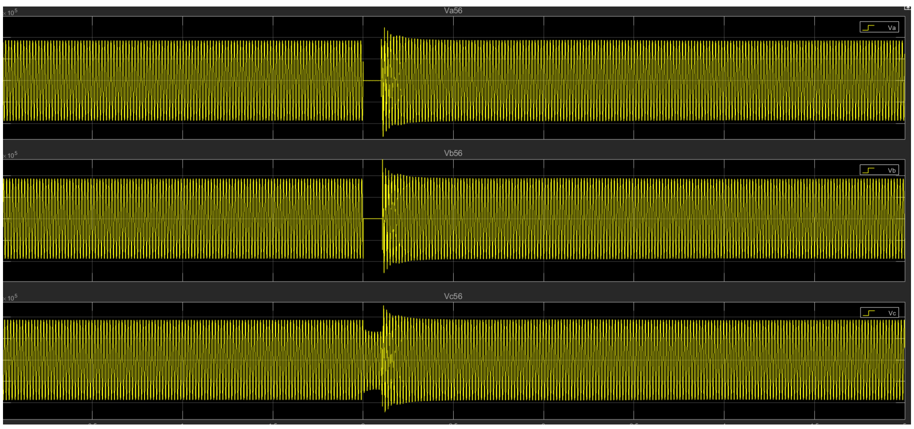

Figure 3.3: Voltage waveform after LLG Fault on line A and B.

CHAPTER 4

METHODOLOGY

Our approach will involve a methodical progression through a set of clearly outlined stages, depicted in the accompanying flow chart [Fig 4.1]. The process initiates with gathering data from diverse sources such as sensors and monitoring devices. Subsequently, we undergo thorough data pre-processing to maintain data quality and uniformity. Next, we employ feature extraction methods to capture crucial insights from the data, and feature selection is used to streamline the dataset.

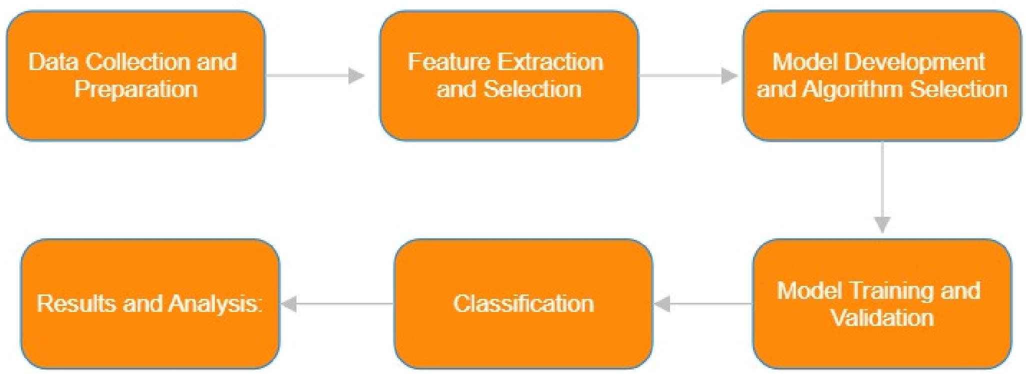

Figure 4.1: Flowchart of Methodology for Shunt Fault Detection and Classification Model.

4.1 Data Collection

We utilized an IEEE 2 area system (referenced as Fig 4.2) with a total of 11 buses to generate our primary dataset. This system serves as a model representing an electrical power network, providing an effective means to simulate real-world conditions and fault scenarios. Our data generation process within this system is thorough, encompassing all common types of electrical faults encountered in power systems.

In order to replicate the dynamic behavior of the system during fault events, we specifically focused on collecting voltage and current values. These measurements were obtained from a carefully selected subset of 5 buses within the network. This subset was strategically chosen to include buses from diverse locations, allowing us to gather data that mirrors a broad spectrum of system responses to faults. Notably, our data collection approach was designed to accommodate fault occurrences from 4 distinct locations situated between these 5 selected buses.

By taking into account both the variety of fault types and the precision of fault locations, our data collection strategy ensures that our dataset is both extensive and indicative of genuine power system conditions. This comprehensive dataset serves as the foundation for our research on robust fault detection and classification.

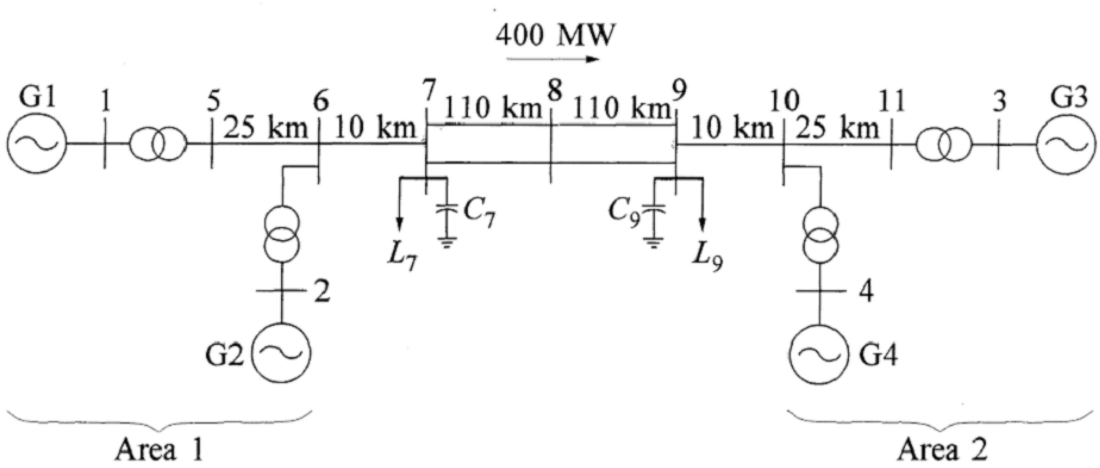

Figure 4.2: 2 Area Kundur System.

Creating a high-quality dataset is a fundamental pillar of our research. We derived our dataset from a complex 2-area electrical power system, intricately structured with 11 buses, 4 generators, and 4 transformers. This system's extensive configuration mirrors the intricate dynamics encountered in actual power networks. To comprehensively simulate fault scenarios, we meticulously adopted a post-fault data collection approach. Voltage and current data were diligently recorded following each fault event, enabling a detailed analysis of the system's responses.

A significant feature of our dataset generation is the deliberate consideration of diverse fault locations. We intentionally induced faults at four specific locations within the system, each chosen strategically to enhance the model's resilience and efficacy in handling various fault scenarios. Additionally, our dataset is meticulously time-stamped, with measurements precisely taken at intervals of 2-2.1 seconds, 3-3.1 seconds, and 4-

4.1 seconds after each fault. This temporal precision facilitates time-dependent analysis and model training.

The dataset encompasses a broad range of fault types, including single-line-to-ground (S-L-G), line-to-line (L-L), line-to-line with ground (L-L-G), triple line-to-line (L-L-L), and triple line-to-line with ground (L-L-L-G) faults. Alongside these fault scenarios, we also accounted for a no-fault scenario, ensuring a well-rounded dataset for comprehensive research. These carefully planned data collection strategies guarantee that our dataset is not only diverse but also a faithful representation of real-world power system conditions. This robust dataset forms the cornerstone of our research on fault detection and classification.

CHAPTER 5

FUTURE WORK

The succeeding phases in our research process are equally important in preparing the data for analysis and model building after the data are generated from Simulink. After obtaining the raw data, we start a thorough data preprocessing phase. To remove any noise, outliers, or inconsistencies that may have been introduced during the simulation, careful data cleaning is required. In order to standardise the data and ensure that it falls within a consistent range for successful analysis, data normalization procedures are used. Additionally, techniques for data augmentation may be used to increase dataset diversity and replicate real-world differences.

The crucial phase that comes after data preprocessing is feature extraction and selection. In order to accurately depict the underlying patterns and properties of the data, we extract pertinent features from the pre-processed data in this phase. These characteristics act as inputs to our machine-learning model. The next step is feature selection, which aims to find the most useful and discriminating characteristics while lowering dimensionality. The model's effectiveness and efficiency are maximized by careful feature selection.

To assign each data point with the appropriate fault class, fault location, and specific time of occurrence, data labelling is then done. Labeling is a crucial part of supervised learning since it makes the training and evaluation of the model easier. A key source for precise labeling is ground truth data, which offers details about actual fault occurrences and locations.

The dataset is prepared for usage in the creation of our machine learning model after the data preprocessing and labelling processes are finished. The main input for training, verifying, and testing our model is the created dataset, which has been enhanced with significant features and labeled data. The effectiveness of our research in defect detection and classification within electrical power systems is supported by these post-data production stages, which are fundamental in translating raw simulation data into a refined and structured dataset.

REFERENCES

[1] S. M. Chopdar and A. Koshti, “Fault detection and classification in power system using artificial neural network,” in *2022 2nd International Conference on Intelligent Technologies (CONIT)*, 2022, pp. 1–6.

[2] O. N. Teja, M. S. Ramakrishna, G. Bhavana, and K. Sireesha, “Fault detection and classification in power transmission lines using back propagation neural networks,” in *2020 International Conference on Smart Electronics and Communication (ICOSEC)*, 2020, pp. 1150–1156.