Virginia Commonwealth University
VCU Scholars Compass

Theses and Dissertations

Graduate School

2018

Fault Classification and Location Identification on Electrical Transmission Network Based on Machine Learning Methods

Vidya Venkatesh

Fault Classification and Location Identification on Electrical Transmission Network Based on Machine Learning Methods

A thesis submitted in partial fulfillment of the requirements for the

degree of Master of Science

at Virginia Commonwealth University.

by

Vidya Venkatesh

Master of Science, Electrical and Computer Engineering

Director: Dr. Umit Ozgur, Professor and Graduate Program Director, Department of Electrical and Computer Engineering

Virginia Commonwealth University

Richmond, Virginia

July 2018

Copyright Page

Acknowledgements

I would like to express my gratitude to those who helped me throughout completion of this work. First, I would like to thank my advisor, Dr. Carl Elks, for giving me the opportunity to work at VCU's Cyber Physical Systems Lab and for his expert advice and encouragement throughout the past two years.

This work would not have been possible without the support of Dominion Energy, VA. I am especially indebted to Brian Starling, whose constant encouragement and support throughout has made this research work possible. Special thanks to Jeremy Crider for his great support and willingness to help me. It has been wonderful working with you both.

Above all, I would like to thank my parents; whose love and guidance are with me in whatever I pursue. I would also thank my loving and supportive husband, Adarsh, for his constant devotion to my success and happiness. I dedicate this to my parents and Adarsh.

Abstract

FAULT CLASSIFICATION AND LOCATION IDENTIFICATION ON ELECTRICAL TRANSMISSION NETWORK BASED ON MACHINE LEARNING METHODS

By: Vidya Venkatesh, MS.

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science at Virginia Commonwealth University.

Virginia Commonwealth University, 2018

Major Director: Dr. Umit Ozgur, Professor and Graduate Program Director, Department of Electrical and Computer Engineering

Power transmission network is the most important link in the country's energy system as they carry large amounts of power at high voltages from generators to substations. Modern power system is a complex network and requires high-speed, precise, and reliable protective system. Faults in power system are unavoidable and overhead transmission line faults are generally higher compare to other major components. They not only affect the reliability of the system but also cause widespread impact on the end users. Additionally, the complexity of protecting transmission line configurations increases with as the configurations get more complex. Therefore, prediction of faults (type and location) with high accuracy increases the operational stability and reliability of the power system and helps to avoid huge power failure. Furthermore, proper operation of the protective relays requires the correct determination of the fault type as quickly as possible (e.g., reclosing relays).

With advent of smart grid, digital technology is implemented allowing deployment of sensors along the transmission lines which can collect live fault data as they contain useful information which can be used for analyzing disturbances that occur in transmission lines. In this thesis, application of machine learning algorithms for fault classification and location identification on the transmission line has been explored. They have ability to "learn" from the data without explicitly programmed and can independently adapt when exposed to new data. The work presented makes following contributions:

1) Two different architectures are proposed which adapts to any *N*-terminal in the transmission line.
2) The models proposed do not require large dataset or high sampling frequency. Additionally, they can be trained quickly and generalize well to the problem.
3) The first architecture is based off decision trees for its simplicity, easy visualization which have not been used earlier. Fault location method uses traveling wave-based approach for location of faults. The method is tested with performance better than expected accuracy and fault location error is less than ±1%.
4) The second architecture uses single support vector machine to classify ten types of shunt faults and Regression model for fault location which eliminates manual work. The architecture was tested on real data and has proven to be better than first architecture. The regression model has fault location error less than ±1% for both three and two terminals.
5) Both the architectures are tested on real fault data which gives a substantial evidence of its application.

Table of Contents

List of Figures

List of Tables

Chapter 1 : Purpose and Significance of the Research

1.1 Introduction

Modern society relies heavily upon complex and widespread electric grids for critical service capabilities such as healthcare, transportation, household heating and cooling, and industrial manufacturing to name a few. As our energy delivery systems (electric and other) age, natural disasters and man-made perturbations are expected to threaten grid integrity more often. Furthermore, urban infrastructure energy delivery networks are highly reliant on the electric grid and consequently, the vulnerability of infrastructure networks to electric grid outages is becoming a major national concern. Electric power transmission is the bulk movement of electrical energy from a generating site, such as a power plant, to an electrical substation. Essentially an electrical grid is an interconnected network for delivering electricity from producers to consumers. It consists of generating stations that produce electrical power, high voltage transmission lines that carry power from distant sources to demand centers and distribution lines that connect individual customers or businesses. Transmission lines are a vital part of the electrical distribution system, as they provide the path to transfer power between generation and load. Transmission lines operate at voltage levels from 100kV to 1000kV and are ideally tightly interconnected for reliable operation. In recent years, advanced sensors, intelligent automation, hierarchical control, communication networks, and operations technologies (OT) have been integrated into the electric grid to enhance its performance and efficiency. These new OT devices allow for large amounts of information from numerous grid systems and transmitting needed information to operations personnel in a timely manner that could not be envisioned when previous generation and transmission systems were designed and built decades ago.

In recent years, power quality has become a main concern in power system engineering – with 85-87% of power system faults occur on distribution lines [1]. However, the faults that occur on the transmission lines (the transmission grid) though fewer have a more significant and widespread impact on the consumers. The performance of a power system is affected by faults on transmission lines, which results in interruption of power flow. As the power transmission configurations (networks) become more complex quick detection of faults and accurate estimation of fault location is critical. The rapid dispatch of repair and restoration of supply voltage is essential for minimizing local and regional economic impacts, reducing overall power outages and improving customer satisfaction.

When a fault occurs in transmission line, it initiates a transition condition. Transients produce over currents in the power system, which can damage the power system depending upon its severity of occurrence. To avoid fault recurrences and the high cost associated with finding line faults, utilities endeavor for developing more accurate fault-locating methods. Transmission protection systems are designed to identify the location of faults and isolate only the faulted section of the network. The key challenge to the transmission line protection lies in reliably detecting and isolating faults compromising the security of the system – with significant accuracy. With the advent of OT devices, new measurement devices like phasor measurement unit (PMU), Digital Fault Recorders (DFR) are often used to provide detailed information on the health the grid. These OT advances in power system has led to massive volumes of data from the continuous monitoring of transmission lines. The massive volumes of data is both a blessing and curse- large amounts of data easily can overwhelm storage facilities, but with the advent of machine learning algorithms this opens potential to implement smart and robust fault location algorithms [2].

Section 1.2 discusses the fundamental terms and concepts used in today's electric power system. The basics and types of transmission network is presented in section 1.3. Section 1.4 describes the issue addressed in the thesis and section 1.5 lays the outline of the thesis.

1.2 Basics of Power System

Electric power systems are real-time energy delivery systems. Real-time meaning power is generated, transported, and supplied the moment light switch is turned on. Electric power systems are not storage systems like water systems and gas systems. Instead, generators produce the energy as the demand calls for it. Figure 1 shows the basic building blocks of an electric power system. Starting with generation, where electrical energy is produced in the power plant and then transformed in the power station to high-voltage electrical energy that is more suitable for efficient long-distance transportation. The power plants transform other sources of energy as well in the process of producing electrical energy. For example, heat, mechanical, hydraulic, chemical, solar, wind, geothermal, nuclear, and other energy sources are used in the production of electrical energy. High-voltage (HV) power lines in the transmission portion of the electric power system efficiently transport electrical energy long distances to the consumption locations. Finally, the remote substations are responsible for transforming this HV electrical energy for delivery on lower high voltage power lines called "Feeders" that are more suitable for the distribution of electrical energy. This electrical energy is again transformed to even lower voltage services for residential, commercial, and industrial consumption.

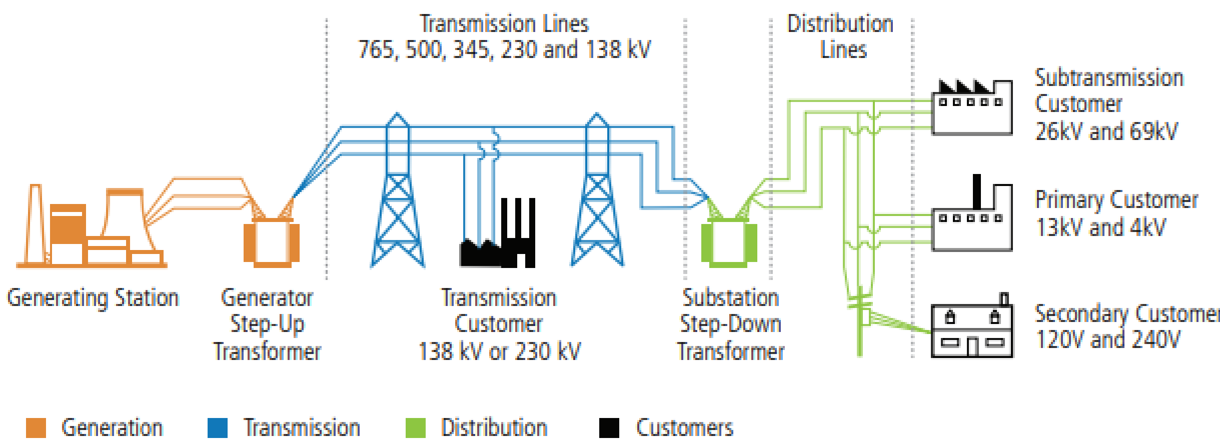

Figure 1: Building Blocks of Electric Power System [3]

The Power Generation and Distributions has four stages:

1) *Generation*: Power generation plants produce the electrical energy that is ultimately delivered to consumers through transmission lines, substations, and distribution lines. Electrical energy must be generated at the same rate at which it is consumed. A sophisticated control system is required to ensure that the power generation very closely matches the demand.
2) *Transmission*: Transmission lines are necessary to carry high-voltage electricity over long distances and connect electricity generators with electricity consumers. Transmission-level voltages are typically at or above 110,000 volts or 110 kV, with some transmission lines carrying voltages as high as 765 kV[3]. Power generators, however, produce electricity at low voltages and the generation voltage is stepped-up to transmission voltages. To make

high-voltage electricity transport possible, the electricity must first be converted to higher voltages with a step-up transformer.

3) *Distribution*: Distribution systems are responsible for delivering electrical energy from the distribution substation. Most distribution systems in the United States operate at primary voltages between 12.5 kV to 34.5 kV and some operate at lower distribution voltages such as 4kV. These low-voltage distribution systems are being phased out because of their relatively excessive cost for losses (low voltage requires high currents, which means high losses). These networks carry the power to consumer units like businesses or residential entities.
4) *Load*: This stage accounts for electrical energy used by various loads on the power system. Electricity is consumed and measured several ways depending on whether the load is residential, commercial, or industrial and whether the load is resistive, inductive, and capacitive.

The electrical network's or the grid's ability to supply a clean and stable power supply is very critical on day-to-day. High power quality ideally creates a perfect power supply that is always available, has a pure noise-free, sinusoidal wave shape and is always within voltage and frequency tolerances. A well-functioning power transmission network enables:

1) *Economies of scale*: The behavior of the electricity sector is directly related to economic factors such as Gross Domestic Product (GDP). In this manner, the demand for electricity be a "thermometer" of the market. As such, growth of the economy as well as increases in purchasing power and quality of life must be accompanied by improvements in the power system, with the objective being compliance with current and future situations.
2) *Rural electrification*: Extending electrical grids into countryside will not only help cater to residential houses for lighting and household purposes but also allows for mechanization of many farming operations especially in areas facing labor shortages.
3) *Increased transmission reliability*: Reliability refers to the extent to which customers have a continuous supply of electricity. As electricity cannot be easily stored, a reliable supply of electricity requires generators to produce electricity and the transmission and distribution networks to transport the electricity to customers in real time. Therefore, a good transmission system will ensure affordable, high-quality electric service is essential for modern life.

4) *Decreased costs*: Transmission network carries the high-voltage power from the generating sites to the distribution stations. The development and improvement of algorithms that allow the analysis and diagnosis of failures in transmission lines can have an important economic impact, for power utilities by reducing operation costs, as they enable the continuity and reliability of the electric sector.
5) *Increases potential for power pools, markets and bulk power transactions:* A reliable transmission network will enable more advanced methods of power transfers like power pool, bulk power transfers etc. It primarily helps to balance electrical load over large network than a single utility by providing mechanism for interchange of power between two or more utilities.

Section 1.3 describes more about the transmission network and its configurations used by utilities.

1.3 Power Transmission Networks

The United States' bulk electric system consists of more than 360,000 miles of transmission lines, including approximately 180,000 miles of high voltage lines, connecting to about 7,000 power plants [4]. High-voltage (up to 765 KV) transmission lines transport power long distances much more efficiently than lower voltage (12 - 34.5 KV) distribution lines for two main reasons. First, high-voltage power transmission allows for lesser resistive losses in transit which is about 6% on average in the United States [5] . This efficiency of high voltage transmission allows for the transmission of a larger proportion of the generated power to the substations and in turn to the loads, translating to operational cost savings. Second, raising the voltage to lower the current allows one to use smaller conductor sizes, or have more conductor capacity available for growth. Transmission line systems relay the power from production sites to the users. Failure of these structures can lead to power cuts and therefore disrupt the day to day life of people as well as the industries dependent on electricity.

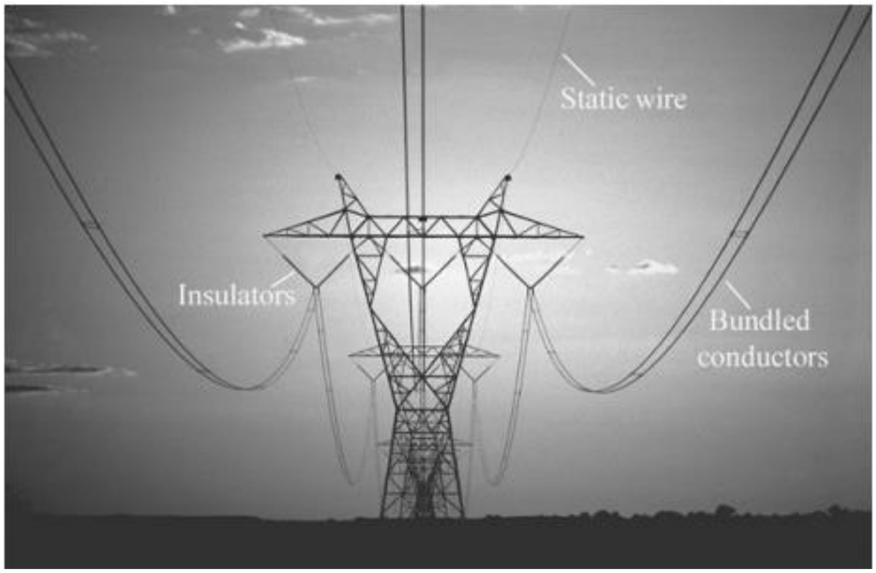

Figure 2: Transmission Network

A transmission grid is a network of power stations, transmission lines, and substations. Energy is usually transmitted within a grid with three-phase AC. Transmission lines are either overhead power lines or underground power cables. Overhead cables are not insulated and are vulnerable to the weather but can be less expensive to install than underground power cables. Overhead and underground transmission lines are made of aluminum alloy and reinforced with steel; underground lines are typically insulated. Figure 2 shows a three-phase 500 kV transmission line with two conductors per phase. The two conductors per phase option is called bundling. Multiple conductors are bundled together per phase to double, triple, or greater to increase the power transport capability of a power line, lower losses and improve other operating characteristics of the line such as electromagnetic fields and audible noise.

Typically, there are three types of line configurations used in the transmission network. These line configurations include (a) radial (one-terminal), (b) two-terminal, and (c) multi-terminal of which three-terminal is possibly the most prominent multi-terminal type. It should be noted that "terminals" in this context, refers to source terminals and not-tapped transformer terminals or stations. The two-terminal line configuration is the most dominant type followed by radial, and the three-terminal lines are the exceptions.

Three-terminal systems are used in power transmission networks to connect three power sources, A, B, and C. The power sources are either generators or Thevenin equivalent of a connected network. As shown as in Figure 3, the three terminals are connected through a Tap-point T which does not contain any measuring devices. Protection systems are like that of two-ended lines except with more sophisticated techniques. In many cases, an existing two-terminal line is converted to three-terminal line as part of program to reinforce the power system. At least one (generally two) communications-based protection groups are normally used with three-terminal line applications.

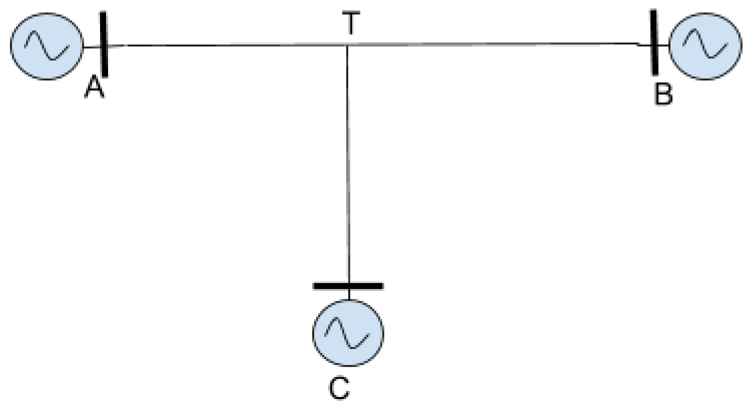

Figure 3: Three-terminal Transmission Lines

Two-terminal line systems are used for bulk power transfer and to supply loads from two power sources- Terminals A, B and are very common. Figure 4 shows the two-terminal transmission line. To obtain proper selectivity and coordination, directional distance relays [6] for phase and ground fault detection are used normally. Directional ground overcurrent relaying is sometimes applied in addition to, or in place of, directional ground distance relay functions. One or two communications-based protection groups are normally used with two-terminal line applications at the transmission voltages greater than 200KV.

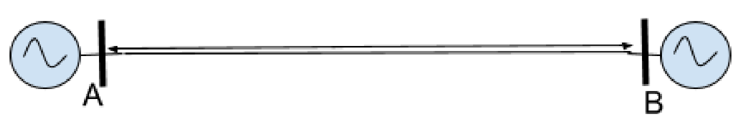

Figure 4: Two-terminal Transmission Line

Radial lines are lines that supply loads from single power source- Terminal A as shown in Figure 5. Nondirectional overcurrent or distance relays are normally used to protect these types of lines. Communications based tripping is not generally necessary.

Figure 5: Radial Configuration

1.4 Problem Statement

Transmission lines or transmission network is a crucial part of the electric grid as it carries high voltage power from generating site to the substations where the voltage stepped-down for end-use consumption transported via distribution lines. Though the frequency of faults is much higher in distribution lines, faults on transmission lines have more widespread impact and faults in buried transmission lines take longer to locate and repair. Additionally, since the transmission lines carry high voltages, faults on these lines might lead to unsafe conditions. Therefore, safeguarding against exposed fault is the most critical task in the protection of power system. The protection schemes or mechanisms for the transmission lines become challenging as configurations of the transmission lines become increasingly complex.

Three-terminal and other multi-terminal line construction are generally a trade-off of planning economics and protection complexities. Two-terminal lines with long tap(s) supplying remote load from the main line may display many of the same protection and load ability issues as three-terminal lines. The complexity of protecting these line configurations increases from the relatively simple radial, to the more difficult two-terminal, and to the still more difficult three-terminal. Relaying three-terminal lines has been and continues to be a challenge for protection engineers [7].

Primary and biggest challenge with protecting three terminal circuits is “Infeed”. During a fault on the transmission line, distance relay measures impedance which is equal to the positive sequence (A balanced three-phase system with the same phase sequence as the original sequence), if there are no sources of fault current on the transmission line between the line terminal where the relay is located and the fault.

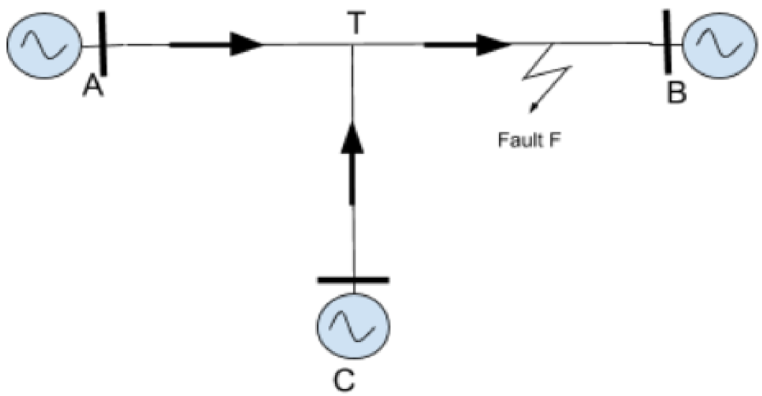

Figure 6: Infeed Effect at Three-terminal

From the Figure 6, the actual line impedance from the relay terminal (Terminal A) to the fault is not always the impedance measured by the relay. This is because the third line terminal (Terminal C) tapped (Tee point) to a line is an additional source of current for a line fault. Current will be supplied to a fault that occurs on the line section beyond the tap of Terminal C through both Terminal A and Terminal C. The voltage drops resulting from the input of fault current from each of these sources into the common section of the line will be measured by the distance relay at the Terminal A. Since the current input from Terminal C is not applied to the relay at Terminal A, the impedance measured by this relay is higher than the actual impedance from the Terminal A to the fault. The relay will under reach; that is, for a given relay setting the relay does not cover the same length of line it would if the additional current source were not present. Due to infeed, most of the impedance based and traveling wave-based methods are not successful in identifying faults and often give erroneous results [8].

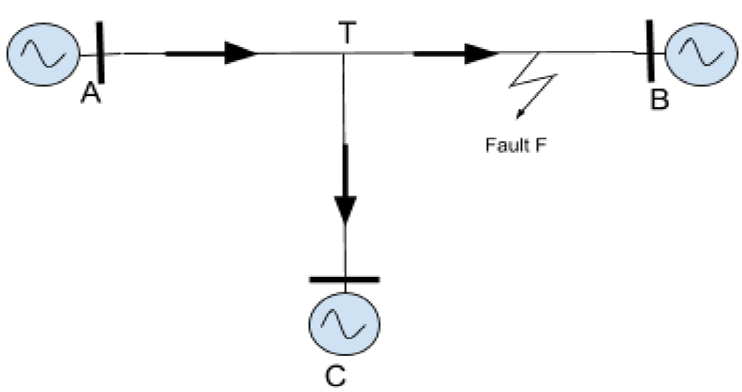

Figure 7: Outfeed Effect at the three-terminal

It is also possible to experience an “Outfeed” at the T location, in which case there will be tendency to overreach as shown in

Figure 7. This phenomenon is not too common but can cause delayed or sequential tripping at the terminals.

Thirdly, transmission lines could traverse long distances, in which case the line A-B ends up being in one region and line C-T in different region with separate set of environmental conditions which directly influences the impedance of the respective lines. This causes line non-homogeneity and since impedance on line C-T is different that of line A-B which makes fault location on line C-T trickier. Therefore, there is a need of an adaptable, resilient method for fault classification and location on transmission lines which could learn from system behavior and detect unknown faults rather than hardcoded methods (algorithms)which follow specific set of rules.

Taken all together, faults on transmission lines and the varying environmental conditions present a complex classification and detection problem. With the advent of new machine learning methods and supervised learning methods, these challenges may be more effectively addressed. Machine learning methods are based on the idea that systems can learn from data, identify patterns and make decisions with minimal human intervention. The ability to automatically apply complex mathematical calculations to big data – over and over, faster and faster give these algorithms potential to identify insights in the data which would be otherwise an impossible task for humans. The availability of high-resolution/high-volume data, due to the proliferation of intelligent electronic devices in smart grids, paves ground to implement more accurate and intelligent machine learning methods for fault classification and location identification on the transmission lines.

1.5 Contributions and Thesis Outline

Majority of the faults on the transmission lines are shunt faults and with around 5% of them being symmetric (all three phases are equally affected). This research presents supervised-learning method for classification and location of shunt faults on three-terminal and two terminal transmission lines. The main contributions of the research are:

- Two different architectures are proposed which adapts to any N-terminal in the transmission line (dimensional scaling).
- The models proposed do not require large dataset or high sampling frequency. Additionally, they can be trained quickly and generalize well to the problem.
- The first architecture is based off decision trees for its simplicity, easy visualization which have not been used earlier. In this instance, fault location method uses traveling wave-based approach for location of faults. The method is tested with performance better than expected accuracy and fault location error is less than ±1%.
- The second architecture uses single Support Vector Machine to classify ten types of shunt faults and Regression model for fault location which eliminates manual work. The architecture was tested on real data and has proven to be better than first architecture. The regression model has fault location error less than ±1% for both three and two terminals.
- Both the architectures are tested on real fault data which gives a substantial evidence of its application.

The thesis is organized into 5 chapters. Chapter 2 discusses relevant literature and presents a survey of different machine learning methods proposed over past years. The two new architectures for fault classification and location is proposed in Chapter 3. Chapter 4 details the experimental setup, implementation and results are presented. Chapter 5 summarizes the conclusions and proposes the possibilities of future work.

Chapter 2 : Background and Related Work

2.1 Introduction

Given the electrical power grid is a complex power system consisting of power generating stations, high voltage transmission lines and distribution lines, fault classification and location identification is necessary to improve protection mechanisms and have reliable, high-speed protection devices. Most often, electrical faults result in mechanical or material damage to the lines or structures, which must be repaired before returning the line to service. As it is noted earlier, repair and restoration is extremely important for maintaining critical and societal services. The restoration process is hampered if the location of the fault cannot be estimated with accuracy or confidence (less than a mile). Various methods have been proposed over the years, and each method have their own merits and disadvantages.

Section 2.2 presents an overview on several types of electric faults occurring on transmission lines. In Section 2.3, to encapsulate the current state of art methods, a survey review is presented on popular machine learning algorithms used for fault classification and location on transmission lines and summaries are given in Section 2.4.

2.2 Review of the Faults on the Transmission Line

In an electric power system, a fault or fault current is any abnormal electric current. For example, a short circuit is a fault in which current bypasses the normal load and an open-circuit fault occur if a circuit is interrupted by some failure. Transmission line carry 3-phase AC [9]. Under ideal state, all phase voltages have same maximum value but differ in phase from each other at angle 120 degrees. Shunt faults are caused by short circuit between lines. For example, a line to ground fault occurs when one conductor drops to the ground or comes in contact with the neutral conductor (ground). This causes a rapid decrease in respective phase voltage of the line involved in the fault and increase in phase current. Figure 8 shows the phase C-to-ground fault in time (μs). V_a, V_b, V_c are the phase voltages of phase a, b, c and I_a, I_b, I_c are the phase currents of phase a, b, c respectively. From the figure we can see that, magnitude of phase C voltage decreases and phase

C current increases during the fault.

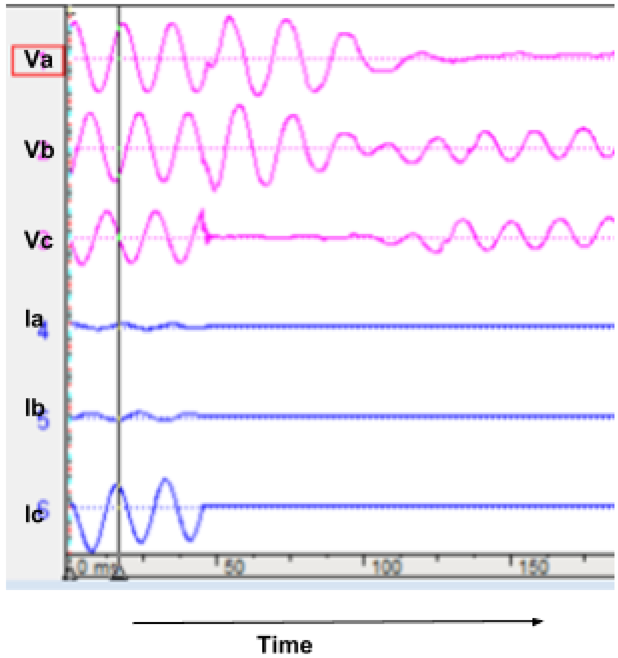

Figure 8: Phase to ground fault in time

Faults can be categorized as the shunt faults and series faults [10] described below:

2.2.1 Series Faults

Series faults represent open conductor and take place when unbalanced series impedance conditions of the lines are present. These faults disturb the symmetry in one or two phases and are therefore unbalanced faults. Two examples of series fault are when the system holds one or two broken lines, or impedance inserted in one or two lines. In the real world a series faults takes place, for example, when circuit breakers control the lines and do not open all three phases, in this case, one or two phases of the line may be open while the other/s is closed [10]. Series faults are characterized by increase of voltage and frequency and fall in current in the faulted phases.

2.2.2 Shunt Faults

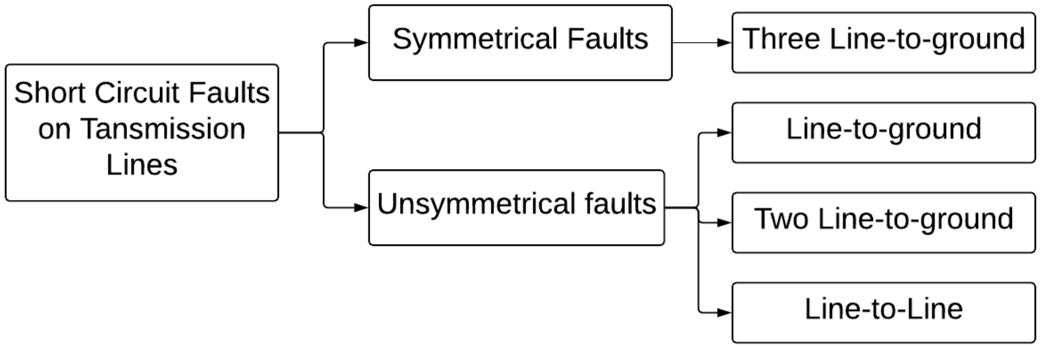

Figure 9: Classification of Short Circuit faults

There are two types of short circuit faults or shunt faults which can occur on transmission lines; balanced faults and unbalanced faults also known as symmetrical and unsymmetrical faults respectively as shown in Figure 9. In symmetrical faults, also called three phase short circuits, all the three phases are short circuited to each other and often to earth also. Such faults are balanced and symmetrical as the system remains balanced even after the occurrence of the fault. Though the symmetrical faults are rare, they generally lead to the most severe fault current flow. Most faults that occur in a power system are unsymmetrical faults involving only one or two phases. The most common type of unsymmetrical fault is a short circuit between a phase and the earth.

The shunt faults are the most common type of fault taking place in the field. They involve power conductors or conductor-to-ground or short circuits between conductors. One of the most important characteristics of shunt faults is the increment the current suffers and fall in voltage and increase frequency. Shunt faults can be classified into four categories [11].

1. *Line-to-ground fault*: This type of fault exists when one phase of any transmission lines establishes a connection with the ground either by ice, wind, falling tree or any other incident. About 70% of all transmission lines faults are classified under this category [12].
1. *Line-to-line fault*: Because of high winds, one phase could touch anther phase & line-to-line fault takes place. Approximately 15% of all transmission lines faults are considered line-to-line faults [12].
2. *Double line-to-ground*: Falling tree where two phases become in contact with the ground could lead to this type of fault. Two phases will be involved instead of one

in the line-to-ground faults scenarios. Ten percent of all transmission lines faults are under this type of faults [12].

3. *Three phase faults*: In this case, falling tower, failure of equipment or even a line breaking and touching the remaining phases can cause three phase faults. In reality, this type of fault not often exists which can be seen from its share of 5% of all transmission lines faults [12]. The first three of these faults are known as asymmetrical faults.

2.3 Survey of the methods

This section presents a survey on different fault classification and location identification techniques in transmission lines highlighting the implementations machine learning methods in the past. In this review, only short circuit faults are considered as they are more common.

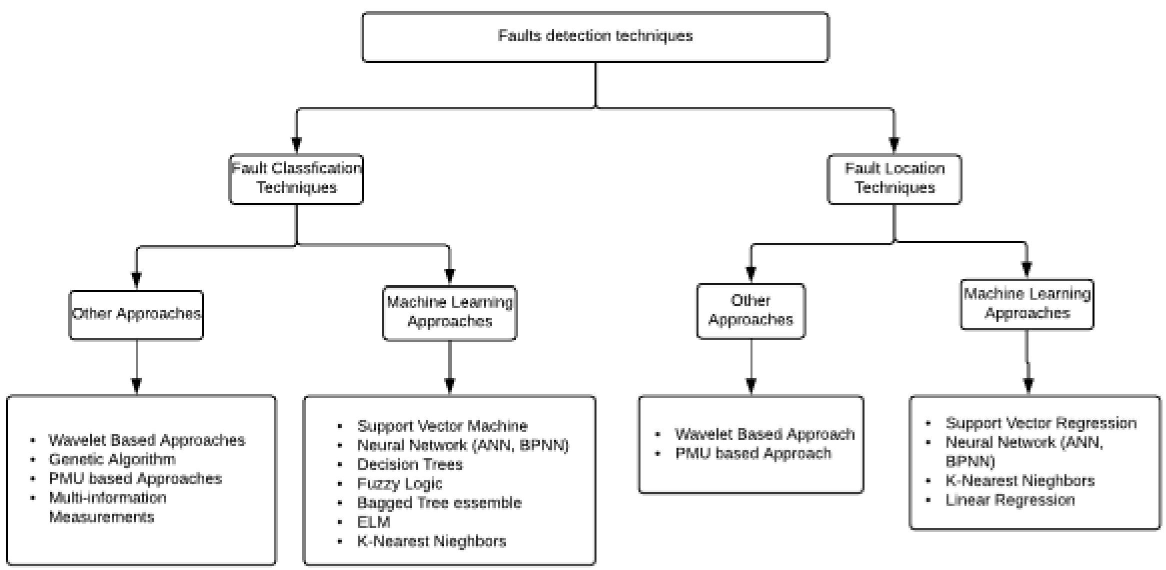

Figure 10: Fault detection techniques

The survey is mainly divided into two parts:

1) Fault classification techniques - Methods that determine the fault type
2) Fault Location Techniques - Methods that calculate the distance of the fault

Both techniques play a vital role in development of protection mechanisms for a given power system model. There have been various approaches used to develop a fast speed and reliable method to deal with faults as shown in Figure 10.

Wavelet based approaches primarily use the time difference between the traveling wave reflections which assume higher sampling rate and synchronized measurements at the terminals for fault identification making it difficult for a practical application, especially for three-terminal circuits due to infeed problem. Though this method has lower error of estimation, it has a higher computational burden [13]. Phasor measurement unit (PMU) based approaches require synchronized phasor quantities from all the terminals of the transmission lines. Genetic Algorithms are the heuristic search and optimization techniques that mimic the process of natural evolution. When applied to classify and location faults on transmission lines, they are often slow and complex to be implemented [14].

Machine learning is a subset of artificial intelligence in the field of computer science that often uses statistical techniques to give computers the ability to "learn" (i.e., progressively improve performance on a specific task) with data, without being explicitly programmed [15]. Machine learning algorithms can learn and improve themselves by studying high volumes of available data. They are very helpful in fields where traditional programming rules do-not operate or rules keep evolving. Since the faults occurring on power grid are very unlikely to be similar and power system can change depending on the future demand, use of machine learning algorithms for solving such problems might. They can benefit from learning correlation between events and can give insights helping human uncover factors causing the faults and to find the solution of complex multiobjective nonlinear systems, the above-said methods are used to get faster solution and less error.

2.3.1 Fault Classification Techniques- Machine Learning

Classification of power system faults is the first stage for improving power quality and ensuring the system protection. For this purpose, a robust classifier is necessary.

Most prominent machine learning approaches for fault classification are explained below:

A. Support Vector Machines (SVM)

Support Vector Machines are supervised learning models which maps the high dimensional input space to target space [16]. The main advantage is its regularization parameter, it tells the SVM to avoid misclassifying each training sample [17]. Secondly, SVM works well with continuous data and can learn more from less number of samples. Section 3.3.1 describes the SVM in more detail.

One promising indicator is that researchers in the past have used SVM for as a classifier to carry out fault classification in the transmission lines [13]. Babu et al. [18] proposed fault classification using Empirical Mode Decomposition (EMD) and SVM's. EMD was used to decompose the voltages of transmission line into Intrinsic Mode Function (IMF's). The characteristic features from the IMF's were extracted by Hilbert Huang Transform which was given as input to three SVM's trained for three phases respectively to predict their involvement in the fault. The method was tested on the simulated data with acceptable levels of accuracy.

K. Li et al. [19] presented fault detection and classification method based on Principal Component Analysis (PCA) and SVM. In the first step, PCA is used to reduce the dimensionality as well as find violating point of the signals according to the confidential limit. In the second step, extracted features are used to build SVM networks to use the pattern recognition to identify faulty phase. However, the PCA cannot identify non-linear relationships which might not work well with three-terminal circuits.

Malathi et. al [20] proposed an approach for fault classification in transmission line using multi-class Support Vector Machine (SVM). Wavelet decomposition using Discrete Wavelet Transform (DWT) of post fault phase current signals are used as a feature set for the SVM which predicts the fault class. This method has been tested extensively on various simulated conditions on the transmission line with different network conditions.

Dubey et. al. [21] used proposed fault classification method using Least Square SVM (LS-SVM). The main advantage of this method is that it requires lesser training sets. Post fault one-fourth cycle current signal is used as input to four LS-SVM, for prediction of three phases and ground respectively. The results have been validated on the simulated data under various fault conditions.

An approach of combining Discrete Wavelet Transform (DWT) with SVM for fault classification was proposed by Hanif et. al. [22]. DWT is used to extract the post fault voltage energies and normalized energies are used as input to the four SVM classifiers used to predict the involvement of each phase in the fault. The method was evaluated using simulated data for two different networks; an overhead line combined with an underground cable and a 6-bus distribution network.

Youssef et. al [23] proposed approach to detect and classify faults in real-time using SVMs. The main idea is to use fault inception angle as input to SVM to recognize the patterns. Phase angle after the fault were recorded for nine types of fault types and used to train the two SVMs.

There have been researches implementing SVM for detecting and classifying faults on series compensated circuit [24]–[26] which is out of boundaries for this research.

B. Neural Network

Artificial Neural Networks are computing systems vaguely inspired by the biological neural networks that constitute animal brains. An ANN is based on a collection of connected units or nodes called artificial neurons which loosely model the neurons in a biological brain. Each connection, like the synapses in a biological brain, can transmit a signal from one artificial neuron to another. An artificial neuron that receives a signal can process it and then signal additional artificial neurons connected to it. The original goal of the ANN approach was to solve problems in the same way that a human brain would. Figure 11 shows an artificial neural network which is an interconnected group of nodes, akin to the vast network of neurons in a brain. Here, each circular node represents an artificial neuron and an arrow represents a connection from the output of one artificial neuron to the input of another. However, over time, attention moved to performing specific tasks, leading to deviations from biology. Many researches have used neural network for fault classification and detection explained in [27].

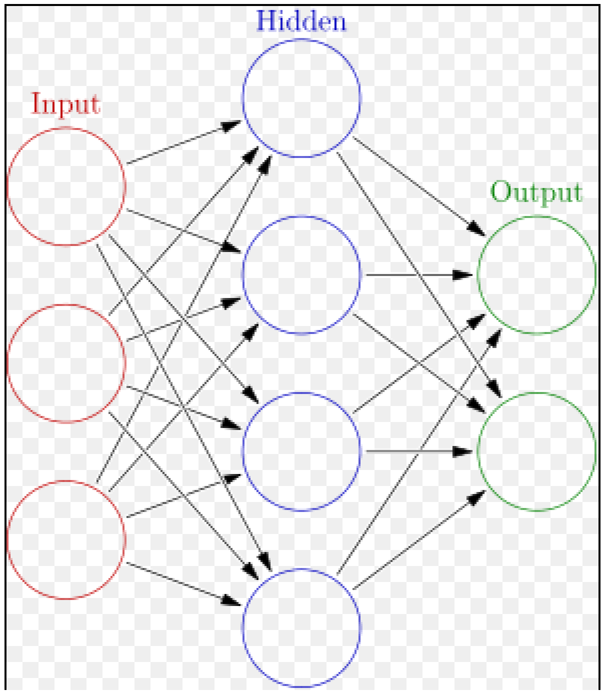

Figure 11: Artificial Neural Network [28]

Koley et. al. [29] proposed a hybrid wavelet transform and modular artificial neural network based fault detector, classifier and locator for six phase lines using single end (single terminal) data. The standard deviation of the approximate coefficients of voltage and current signals obtained using discrete wavelet transform are applied as input to the modular artificial neural network for fault classification and location.

Ahmad et. al. [30] proposed wavelet based artificial neural networks for fault classification. Discrete wavelet transforms (DWT) is used to extract high-frequency components of the aerial modal currents. A feature vector is built using the wavelets details coefficients of one level of the aerial modes and is used to train an ANN. The proposed method is tested on the simulated data with acceptable accuracies.

Rao et. al. [31] has presented a fault classification and detection method using discrete wavelet transform and artificial neural networks. Discrete wavelet Transform (DWT) is applied to the fault phase currents to obtain energy values which is used as input to train the neural network. The proposed method was tested on a simulated network values using MATLAB.

Few other researchers have used Back-Propagation Neural Network (BPNN) for identifying and classifying faults on transmission lines. Backpropagation is a method used in artificial neural networks to calculate a gradient that is needed in the calculation of the weights to be used in the network. It is commonly used to train deep neural networks, a term referring to neural networks with more than one hidden layer.

Saini et. al. [32] has proposed new algorithm for fault detection and classification on parallel transmission lines using Discrete Wavelet Transform (DWT) and Back-Propagation Neural Network (BPNN). Wavelet energies coefficients of alpha and beta mode currents obtained by clark's transformation are used as input to train BPNN with two hidden layers. The proposed method is tested on different networks and fault scenarios.

C. Fuzzy Logic

Fuzzy logic is a form of many-valued logic in which the truth values of variables may be any real number between 0 and 1. It is employed to handle the concept of partial truth, where the truth value may range between completely true and completely false. By contrast, in boolean logic, the truth values of variables may only be the integer values 0 or 1. They do not need detailed knowledge of the system as the decisions are based on rules determined by humans. Figure 12 shows the basic building block of a fuzzy scheme consisting of 3 stages. In the fuzzification, the inputs (for fault classification voltage/current transients) fuzzified into fuzzy membership functions. In the Fuzzy Inference System, all the rules in the rule base are used to compute the fuzzy output functions. De-fuzzification stage, maps the fuzzy output functions to get predicted fault type.

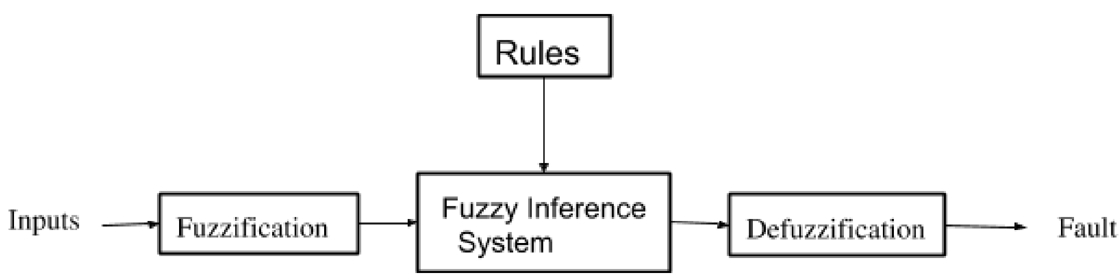

Figure 12: Block Diagram of Fuzzy Logic System

Saradarzadeh et. al. [33] proposed an algorithm for fault type recognition of shunt faults that occur on the transmission line. The proposed method uses the phase sequence components of three phase voltages and currents that are available in most of the power system protection relays.

A fuzzy method is used to identify the type of fault from the current and voltage signals separately and then combines the results to provide more accurate fault-type recognition.

Prasad et. al. [34] proposed a method for fault classification. Post fault currents from three phases of one terminal is used as input to the Fuzzy Inference System (FIS) to classify faults. The proposed technique using two classifiers one is for ground faults (Fuzzy classifier-I) and second one is for phase faults (Fuzzy classifier-II). The method is tested on the simulated network.

Adhikari et. al. [35] used three phase currents data obtained from Compact Reconfigurable i/o (CRIO) devices as input to their method using Fuzzy logic. Once the rule base is prepared for classification, compiled fuzzy logic is dumped to FPGA (field-programmable gate array) to get a real-time performance.

D. Other Techniques

The other machine learning techniques used in fault classification and identification.

Jamehbozorg et. al. [36] proposed Decision Tree based method for fault classification in Double-Circuit Transmission Lines. The proposed method needs voltages and currents of only one terminal of the protected line. After detecting the exact time of fault inception and calculating the odd harmonics of the measured signals, up to the nineteenth, a decision tree algorithm is employed for recognition of the intercarrier fault type. Also, the proposed method is extended for classification of crossover faults in these transmission lines.

Mishra et. al [37] have used bagged tree ensemble technique. Bagging stands for bootstrap aggregation; whereby random samples are drawn through replacing the training datasets. Bagging is a simple method that can be employed to reduce the variance for those machine learning techniques with high variance. Post fault current is decomposed by Fast Discrete Orthonormal S-Transform (FDOST) and bagged tree ensemble technique is used to classify faults. The proposed method is then tested under different fault scenarios on the simulated data.

K-Nearest Neighbor (k-NN) based classification method was proposed by Majd et. al. [38]. Distances of each sample and its fifth nearest neighbor in pre-determined in the default window which determines the fault occurrence time and phase. Therefore, k-NN is applied to the instantaneous values of normalized three phase currents.

Ray et. al. [39] proposed an Extreme Learning Machine (ELM) based fault classification technique in a series compensated transmission line. Extreme learning machines are feedforward neural networks for classification, regression and feature learning with a single layer or multiple layers of hidden nodes, where the parameters of hidden nodes (not just the weights connecting inputs to hidden nodes) need not be tuned. Discrete Wavelet Transform (DWT) is used to decompose the instantaneous current signals used as input to ELM for fault classification. The proposed method is tested with simulated data.

Dasgupta et. al. [40] proposed a method for detecting and classifying transmission line faults using cross-correlation and k-Nearest Neighbor (k-NN). This method computes the cross correlation between pure and faulty current signals. Extracted features are used as input the k-NN algorithm which then computes distance of a given sample to all other samples in the set and class of the sample with least distance is predicted.

Linear Regression Index-Based Method for fault classification and detection is presented by Musa et. al. [41]. The proposed algorithm has constructed a rule as follows: when the system is running under healthy condition, the Linear Regression Coefficient Indices (LRICs) will be equal to zero; when the system is subjected to the fault condition, the LRICs of faulted phases will be greater than zero. For each possible scenario of faults, the proposed algorithm required only the three-phase current measurement of the local measurement.

2.3.2 Fault Location Techniques – Machine Learning

Accurate fault location that occur on transmission lines is highly important from aspect of quick identification of weak points on the transmission line and taking respective counter measures to decrease the probability of those faults. Various approaches have been developed over the time to address these issues and in those few are hard coded. On the other hand, power of algorithms which can learn from real world pattern are explored. Figure 10 shows few fault location techniques that have been developed by the researches in the past. In this section, machine learning techniques employed in past have been summarized.

A. Neural Networks

Neural Networks becomes first choice when the goal is to model non-linear and complex relationships which follow real-life pattern. They generalize well which results in a model that predicts well on unseen data.

Twafik et. al. [42] proposed Artificial Neural Network (ANN) for estimating fault location on transmission lines. Prony method is used to extract the modal information from voltage or current signal. ANN are then used to estimate the fault distance based on the modal information. The model is trained and tested using the simulated data.

Fathabadi et. al. [43] proposed an hybrid framework consisting of a proposed two stage Finite Impulse Response (FIR) filter, four Support Vector Machines (SVMs), and eleven Support Vector Regressions (SVRs). The proposed two-stage FIR filter together with the SVMs are used to detect and classify short-circuit faults while the SVRs are utilized to locate short-circuit faults and predict distances.

Yadav et. al. [44] have written a comprehensive and exhaustive survey will reduce the difficulty of new researchers to evaluate different ANN based techniques with a set of references of all concerned contributions. From the survey they concluded that ANN is found to be robust, accurate, and efficient approach for transmission line fault detection, classification, localization, direction discrimination, and faulty phase selection.

B. Support Vector Regression (SVR)

The Support Vector Regression (SVR) uses the same principles as the SVM for classification. Because output is a real number it becomes very difficult to predict the information at hand, which has infinite possibilities. In the case of regression, a margin of tolerance (epsilon) is set in approximation to the SVM which would have already requested from the problem.

Ray et. al. [45] proposed Support Vector Machine (SVM) for fault classification and Support Vector Regression (SVR) for fault location. Fault classification consists of four SVMs each predicting the involvement of each phase and fault location consisting of SVR. Both the models have been trained by best features out of decomposed post fault current signals using Wavelet packet transform (WPT). The proposed method has been trained and tested on the simulated data and a comparison study has been carried out with methods published by other researchers.

Hosseini et. al. [46] presented hybrid method for fault classification and location. Post fault voltage samples are decomposed by discrete wavelet transform which are used with post fault current samples as input to SVM for the fault classification model. Four SVMs are used to predict the fault in each phase. Depending on the fault type in the first stage, the second stage one out of four SVRs is selected for fault location.

C. Other Techniques

Farshad et. al. [47] proposed a method to classify and locate single-to-ground faults using k-Nearest Neighbors (k-NN) algorithm. Various features are extracted from the voltage signals measure from a single terminal. Decomposed signals from the discrete Fourier Transform is used as input to k-NN for fault type classification. k-NN in regression mode is used for fault location. Since, current signals are not used, the proposed approach is immune against current-transformer saturation and its related errors.

Ray et. al. [48] in proposed fault location technique using extreme learning machine (ELM) in series compensated transmission line. The proposed method uses one cycle of post fault current and voltage signals which are decomposed by wavelet transform. Best features, selected by genetic algorithm are then used as input to the ELM. The method has been tested on variety of simulated data.

2.4 Summary

Variety of approaches have been used to increase reliability and robustness of fault classification and location methods.

Neural Network follows a black box model making the explanation for the result typically difficult to understand. Neural network-based fault classification methods show good accuracy, however the training time is quite large due to which the task becomes more complex. The artificial neural network techniques suffer from the requirement of large training data. However, these methods are hard to implement practically. If the fault can't be identified quickly, it will produce many ill-effects such as line outages during the period of peak load leading to severe economic losses. There may be a chance for the entire grid to collapse which is called as blackout and the reliability of the system would be affected.

Few of the fault classification and location methods use fuzzy logic based architectures [13]. Fuzzy logic uses rule-based relationship for making decisions. Though they have lot less computation burden, it is tedious to develop fuzzy rules and membership functions and fuzzy outputs can be interpreted in many ways making analysis difficult. In addition, it requires lot of data and expertise to develop a fuzzy system.

Other impedance measurement-based methods for fault location depend on fundamental concept of calculating line impedances pre- and post-faults to determine the distance of the fault. However, in three terminal circuits, due to infeed, the impedance values are measured are much larger than actual line impedance which gives rise to erroneous results. Secondly, the lines A-B and C-T may be in different terrains which results in different environmental conditions. In this scenario, synchronized impedance measurements might provide erroneous results.

It appears, Current state of the art Machine Learning methods presented in above sections have tested models on simulated data which have same distribution, pattern/trend as training data. Therefore, the robustness of the methods is not completely known and the question of whether these methods are applicable to 3 terminal networks is yet to be answered.

Chapter 3 : Design

3.1 Introduction

In the recent past, many researchers have proposed approaches for fault classification and location identification. However, they were not applicable to a variety of transmission network configurations, particularity 3 terminal networks and were not evaluated with real data to ensure the effectiveness of the methods in locating and classifying faults on the transmission line.

Based on the literature review and preliminary designs, the design criteria that emerged is as follows:

1. *Scalability*. Transmission network expand as the demand for the power increases every year. Therefore, fault location methods should be scalable.
2. *Confidence*. Should produce estimates that are timely, sound and reliable, otherwise the confidence in the methods would be weakened and longer lengths of transmission line would have to be examined to find the exact fault location.
3. *Network* Topology. Adaptive to different configurations, applicable to changing fault data.
4. *Relevant*. Make use of existing power line health data
5. *Extensible*. Interface with existing OT power line monitoring devices.

Based on recent advances in machine learning, it was decided to explore the utility and applicability of machine learning to fault classification and location on transmission lines. Machine learning is a form of data analysis that automates analytical model building [49]. Using algorithms that continuously assess and learn from data, machine learning algorithms enables hidden insights into complex behavior and relationships. It can handle multi-dimensional and multi-variable data in dynamic environments. A key aspect to machine learning algorithms is that they learn the behavior of the system representative and synthetic datasets to produce reliable, repeatable decisions and results. For fault classification and location on transmission lines, these attributes are desirable. In addition, machine learning methods do not require synchronized measurements at the terminals. Additionally, they can be employed in real time to monitor the grid as well.

Input data is very crucial aspect for the machine learning algorithms and the correctness of prediction is based on the quality of data that the model was trained with. Post- fault voltage or current transients from the live grid are prone to have small amount of noise (may be negligible). Therefore, to ensure the quality of the data used to train the machine learning models and to extract required features from the post-fault signal, Discrete Wavelet Transform (DWT) is used. The advantage of DWT lies in determination of the key components in the signal like energy and entropy. The training data which consists of these components is used to train the predictive models so that they can learn from the information given by these components.

In predictive modeling, the idea is to create a function which is isomorphic to original function/process which was used to generate the training data. Therefore, this predictive model can predict new data points using the "new" function. The fault type and faulty line identification on the transmission line is a classification problem – in which we want to build a classification model to classify ten fault types into the target classes. Based on the preliminary research, we focused on two machine learning based architectures. The first architecture employs decision trees and second architecture use multi-class Support Vector Machines (SVM) for fault classification and faulty line identification. Both SVM and Decision trees follow white box (subsystem whose internals can be viewed but usually not altered) approach and are capable of handling continuous data (floating point).

Section 3.2 gives the details about the overall process for fault identification in the transmission lines.

3.2 Overview

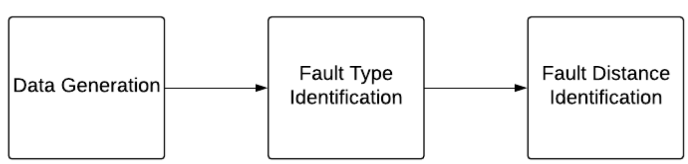

Figure 13: High Level Overview of the Process

This section provides an overview of the proposed method as shown in Figure 13. The goal is to employ machine learning methods to identify faults and fault location on a given transmission line or circuit rather than hard coding the values.

The process has mainly three major steps below:

1) *Data Generation*: The machine learning algorithms need to be trained before deployment in real time to detect faults and identify locations. In this step, simulated data of fault phase voltages is collected from an emulated Transmission model resembling the live transmission model. The collected data then is normalized and massaged before it can be used as training data to train the models.
2) *Fault Type Identification*: The goal is to correctly classify ten types of Short circuit faults as described Chapter 2. Two different architectures are presented in the later sections consisting machine learning algorithms to predict the fault type.
3) *Fault Location Identification*: The methodology differs for three terminals and two terminal circuits. Fault location identification in first architecture is using wavelet-based traveling wave method to calculate the distance. The second method is to employ regression models to predict the distance of the fault.

Section 3.2 describe the data generation process in detail. Section 3.3 and 3.4 present two different architecture for fault type classification and location methods.

3.3 Data Generation

Data Generation is a crucial step to any method employing machine learning algorithms. The algorithm or model needs to be trained beforehand to predict the outcome for the test sample. The model is initially fit on a training dataset that is a set of examples used to fit the parameters of the model and tested on the test dataset which is used to provide an unbiased evaluation of a final model fit on the training dataset. Machine learning algorithms learn from data. Therefore, it is critical that you feed them the appropriate data for the problem being addressed. Additionally, the type of the data collected governs the choice of machine learning algorithm to be applied to obtain best results. For the fault classification and location problem, the data generation process is presented in Figure 14.

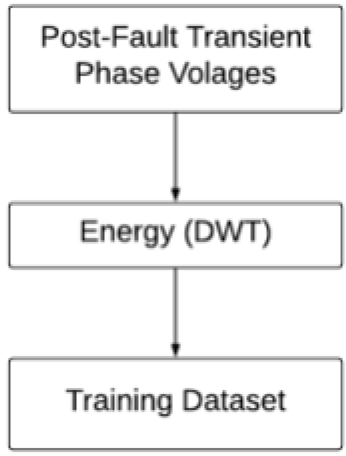

Figure 14:Data Generation Process

It is a three-step process provided as follows:

1) Post fault transient three phases (V_a, V_b, V_c) and ground mode voltages are recorded for one cycle on each terminal of the transmission model under study. Simulated post fault transient voltages from all the terminals is obtained from simulating the faults on the network using Aspen One-liner [50]. Post fault transient phase current value also could be used in place of phase voltages which is to be tried in future work.
2) Discrete Wavelet Transform is applied to the transient phase and ground mode voltages to get the wavelet transform approximation coefficients (WTAC's). To minimize noise effect wavelet coefficients are squared [51]. Energy of the wavelet is obtained by summation of all the $WTAC^2$ over one cycle after the fault has occurred.

$$E_m = \sum_{k=0}^{K-1} WTAC^2(k) \; \forall \, m \in \{a, b, c \; and \; g\} \; and \; k \in \{0, K-1\}$$

Where K is the number of cycles

The obtained wavelet energies are normalized as

$$E_{Nk} = \frac{E_m}{E_{ma} + E_{mb} + E_{mc} + E_{mg}} \; \forall \, m \in \{a, b, c \; and \; g\}$$

3) The input features are calculated wavelet energies at all the terminals. For three terminal models, the data set consists of wavelet energies from all the three terminals A, B and C. Classification of fault is done from the obtained energy of the approximation coefficients.

4) For architecture I, the number of target labels to be predicted is 4, phase A, B, C and Ground. Therefore, three-terminal training set would be Nx12 matrix features and Nx4 target labels, where N is the number of samples in the dataset. Two terminal has Nx8 feature matrix and same target labels. For architecture II, the feature space for two and three terminals does not change, the target matrix reduces to Nx1.

All the machine learning algorithms are trained with above training simulated dataset. Once the training is done, the algorithms will have optimal decision boundaries which will be used to predict the outcome of the real-time fault during the testing phase.

The next section briefly describes Discrete Wavelet Transform (DWT) and then the later sections present the proposed architectures.

3.3.1 Discrete Wavelet Transform

A discrete wavelet transform (DWT) is any wavelet transform for which the wavelets are discretely sampled. For fault classification and location technique, DWT is a method of preparing the data. Traveling wave theory is utilized in capturing the travel time of the transients along the monitored lines between the fault point and the relay. Time resolution for the high frequency components of the fault transients, is provided by the wavelet transform. Using wavelets for fault location was first proposed in [52].

Traveling wave or ultra-high-speed fault location method utilizes the higher frequency contents of the transient fault signals due to its use of traveling wave theory and shorter sampling windows.

Wavelet transform possesses some unique features that make it very suitable for this application. It maps a given function from the time domain into time-scaling domain. The wavelet, the basis function used in the wavelet transform, has bandpass characteristics which makes this mapping like a mapping to the time-frequency plane. The wavelet transform is often compared with the Fourier transform, in which signals are represented as a sum of sinusoids. In fact, the Fourier transform can be viewed as a special case of the continuous wavelet transform with the choice of the mother wavelet $(t) = e^{-2\pi i t}$. The main difference in general is that wavelets are localized in both time and frequency whereas the standard Fourier transform is only localized in frequency. This localization allows the detection of the time of occurrence of abrupt disturbances, such as fault transients. Fault generated traveling waves appear as disturbances superposed on the power frequency signals recorded by the relays. Processing these signals using the wavelet transform reveals their travel times between the fault and the relay locations.

Implementation of the discrete wavelet transform, involves successive pairs of high-pass and low-pass filters at each scaling stage of the wavelet transform. This can be thought of as successive approximations of the same function, each approximation providing the incremental information related to a scale. The first scale will cover a broad frequency range at the high frequency end of the spectrum, and the higher scales will cover the lower end of the frequency spectrum with progressively shorter bandwidths. Conversely, the first scale will have the highest time resolution and lowest frequency resolution and higher scales will cover longer

time intervals and shorter frequency ranges.

The DWT of a signal x is calculated by passing it through a series of filters. First the samples are passed through a low pass filter with impulse response g resulting in a convolution of the two:

$$y[n] = (x * g)[n] = \sum_{k=-\infty}^{\infty} x[k]g[n-k]$$

The signal is also decomposed simultaneously using a high-pass filter h as shown in Figure 15. The outputs giving the low frequency detail coefficients (from the high-pass filter) and high frequency approximation coefficients (from the low-pass). It is important that the two filters are related to each other and they are known as a quadrature mirror filter.

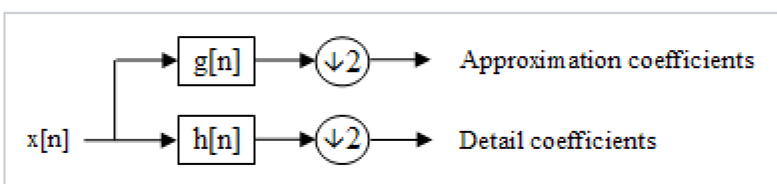

Figure 15: Block diagram of one level DWT [53]

However, since half the frequencies of the signal have now been removed, half the samples can be discarded according to Nyquist's rule. The filter output of the low-pass filter g in the diagram above is then subsampled by 2 and further processed by passing it again through a new low- pass filter g and a high- pass filter h with half the cut-off frequency of the previous one, i.e.:

$$y_{high}[n] = \sum_{k=-\infty}^{\infty} x[k]h[2n-k]$$

$$y_{low}[n] = \sum_{k=-\infty}^{\infty} x[k]g[2n-k]$$

With summation operator ↓, the above equation can be written more concisely as:

$$y_{low} = (x * g) \downarrow 2$$

$$y_{high} = (x * h) \downarrow 2$$

DWT differs from Continuous Wavelet Transform (CWT) in time-frequency plane that is considered as multi-resolution wavelet analysis. The purpose is to decompose signal in multiple frequency bands, to process the signal in multiple frequency bands differently and independently.

Santoso et. al. [51] studied the power quality via wavelet analysis. In their study, they concluded that, for short and fast transient disturbances, Daub4 and Daub6 wavelets are better, while for slow transient disturbances, Daub8 and Daub10 are particularly good. Rioul et al. in their seminal paper [54] have studied the computational complexity of wavelet transforms in detail. In general, the computations are periodic in 2^m for an m-level wavelet. The selection of an appropriate mother wavelet without knowing types of transient disturbances (which is always the case) is a formidable task. Therefore, instead of creating algorithms to select appropriate wavelets (which surely adds complexity to the main problem), we utilize one type of mother wavelet in the whole course of detection and localization for all types of disturbances. In this study, Db4 at scale 2 is used for decomposition of the post fault transient voltage signals.

3.4 Architecture I

Locating faults on Three Terminal transmission models are complex [Section 1.4] due to interactions between all the three terminals. A reasonable approach to address complex interactions is a "divide and conquer" approach where the problem is broken down into sub-problems. This approach is not only resilient but also helps towards optimization of sub-problems. In architecture I as shown in Figure 16, for the fault classification, decision is made at each phase. Four decision tree classifiers, one for each phase (A, B, C and G) for its involvement in the fault. The location identification, is done by using travelling wave-based method. However, for three-terminal circuits, the faulty-branch is identified first and then the distance is calculated.

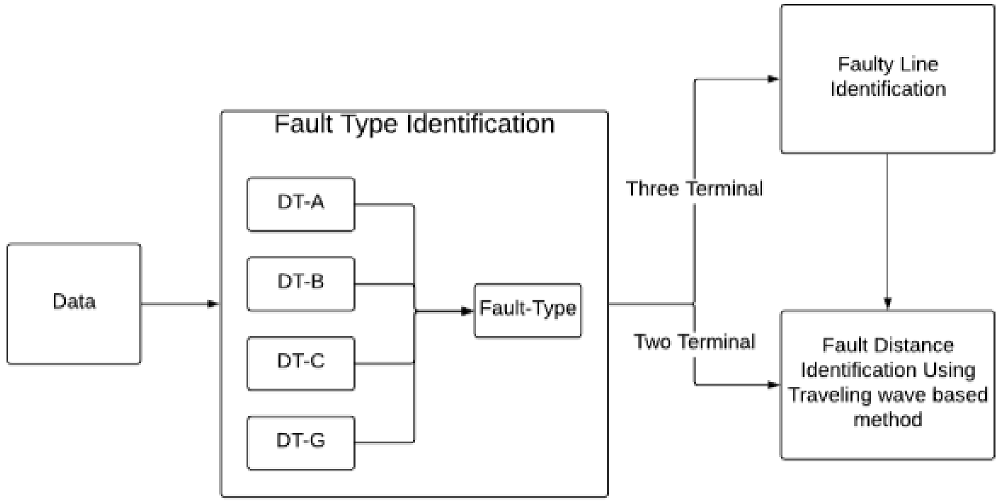

Figure 16: Overview of Architecture I

The architecture can be applied to simple system consisting of two-terminal transmission line to a complex system consisting of N lines, which makes it highly scalable.

Section 3.3.1 has an overview of Decision Tree model used in fault classification followed by fault classification and location identification processes in Section 3.3.2 and 3.3.3 respectively.

3.4.1 Review of Decision Trees

Decision Trees are non-parametric supervised learning method used in classification and regression. Classification and Regression Trees (CART) were first introduced by [55]. The model predicts the value of the target variable by learning decision rules from the feature set also called input variables. For Fault Classification, since the target value to be predicted is discrete, the trees are called classification trees. Whereas, for the regression the target value is a real number.

Decision Trees are constructed top-down approach, by choosing a variable at each step that best splits the target into homogeneous sets. It follows greedy search for one feature value which can "best" split the target space which is a decision value at a given node. This process repeats for the next child node. There are various metrics used by different algorithms constructing the decision trees to find the "best" feature.

Below are few metrics used to provide a measure of the quality of split:

1) *Gini Impurity*: It is a measure of how often a randomly chosen element from the set would be incorrectly labeled if it was randomly labeled according to the distribution of

labels in the subset. The Gini impurity can be computed by summing the probability p_i of an item with label i to being chosen times the probability:

$$\sum_{k \neq i} p_k = 1 - p_i$$

To compute Gini impurity for a set of items with J classes, suppose i belongs to {1, 2…J}, and let p_i be the fraction of items labeled with class i in the set.

$$I_G(p) = \sum_{i=1}^{J} p_i \sum_{k \neq i} p_k = \sum_{i=1}^{J} p_i \ (1 - p_i) = 1 - \sum_{i=1}^{J} p_i^2$$

2) *Information Gain*: It is based out of concept of Entropy from Information Theory.

Entropy is defined as:

$$H(T) = -\sum_{i=1}^{J} p_i \log_2 p_i$$

where p_1, p_2, ...are fractions that add up to 1 and represent the percentage of each class present in the child node that results from a split in the tree

$$\overbrace{IG(T,a)}^{\text{Information Gain}} = \overbrace{H(T)}^{\text{Entropy(parent)}} - \overbrace{H(T|a)}^{\text{Weighted Sum of Entropy(Children)}}$$

A good example of decision tree is mentioned in [56]. The iris data set has three classes namely - Setosa, Versicolour, Virginica. At each node, an observation traverses to the left child node only if the condition at that node is true. The pair of numbers below each terminal node gives the number of misclassified over the node sample size as shown in Figure 17.

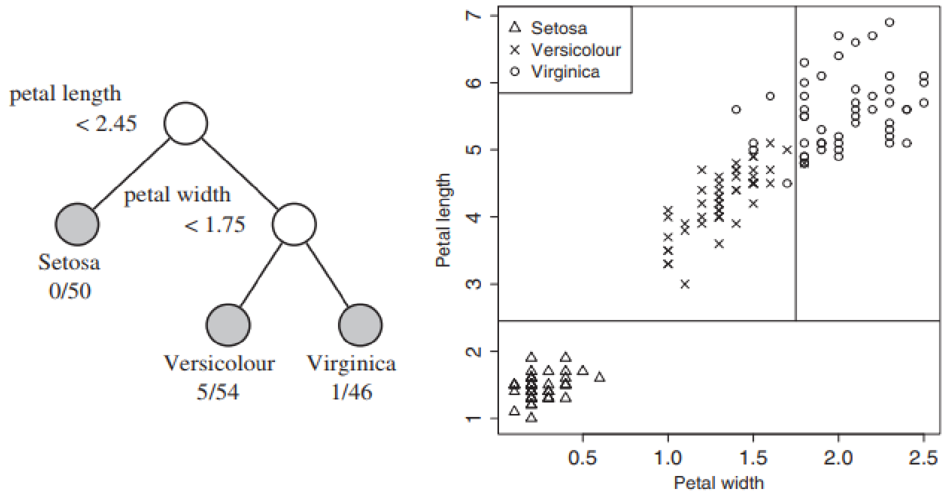

Figure 17:Classification tree model for iris dataset [42]

Decision Trees are simple to understand, interpret and visualize. They work well with both continuous and discrete data which makes them suitable to work with the generated dataset. They require little data preparation compared to other models and mirrors human decision making closely which can greatly help in reasoning about a situation. Decision Trees work well with Binary/Multiclass Labels. A class is the category for a classifier which is given by the target. The number of class to be predicted define the classification problem. A class is also known as a label.

On the Other hand, Decision trees can be non-robust i.e. get over-fitted to the training data. A slight change in the training set can cause substantial changes in the predictions. Various statistical metrics should be used to generalize the model. Since the transmission model doesn't change frequently, Decision Tree model once made could be sustained for long time.

3.4.2 Fault Classification

The goal is to predict ten types of Short Circuit fault which occur on the transmission line. Figure 18 shows the process of the fault classification for two and three terminal circuits. The idea is to use divide and conquer approach on the system and collectively interpret the result.

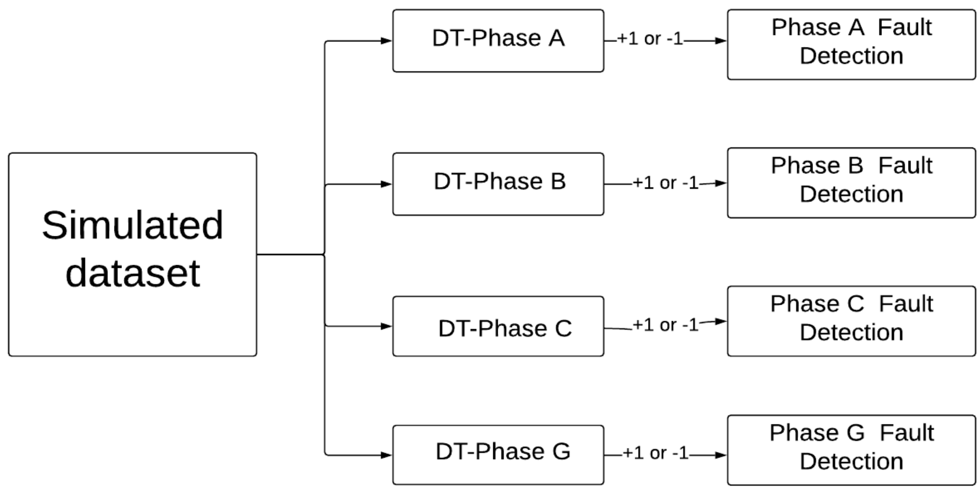

Figure 18: Architecture I Fault Classification for Two and Three Terminal Circuit

The simulated data has phase voltage wavelet energies at each Terminal as features and four labels (A, B, C and G) each taking value of either +1 or -1. The value of +1 denotes that the phase is involved in the fault that occurred. For example, for a phase A to Ground fault, the values of A, B, C and G would be +1, -1, -1 and +1 respectively.

Simulated data set described earlier is used as input to train the decision trees and tested on the real dataset. Each of four decisions have two target classes (binary). Decision Trees corresponding to each phase are used to predict the involvement of each phase in the fault. For example, DT-Phase A is a decision tree for phase A predicts +1 if phase A is involved in the fault and -1 if it's not. The result is interpreted as combination of prediction of 1 bit from 4 Trees. The Target space for the fault classification is $\sim 2^{12}$.

Note that in this architecture, trees are assumed to be independent of each other, meaning prediction of one tree is an independent event. This may not hold true as Short Circuit faults are a combination of any two phases.

3.4.3 Fault Location Identification

Once the fault type is classified using the decision tree classifiers, the next step is to identify the line where the fault has occurred. In case of two terminal transmission line, since the system studied has just a single line, we skip this step and do the wave-distance method to get the fault location distance directly. For Three-Terminal (TEED) Circuits, the branch with the fault first needs to be identified. Figure 19 shows the Fault Location Identification process for TEED

Circuits.

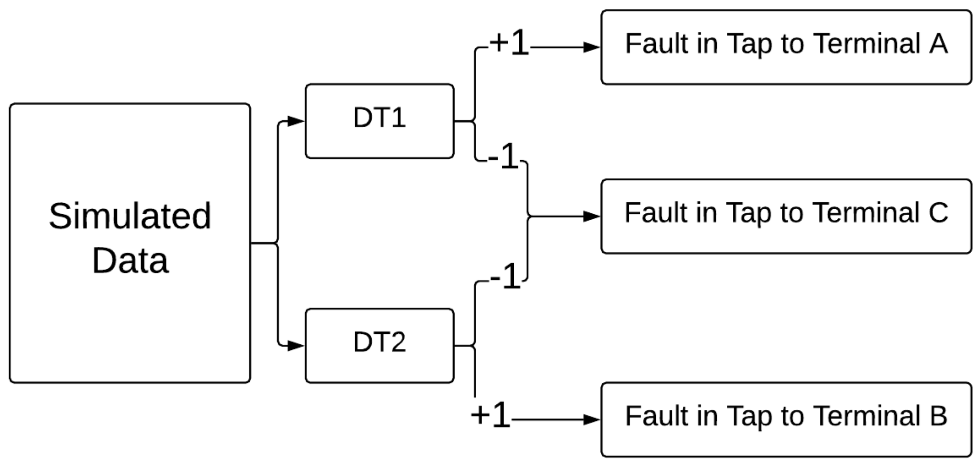

Figure 19: Fault Location Identification Process for Three-Terminal Circuits

The fault location identification process mainly consists of two Decision Tree models which predict whether the fault occurred was in line A-Tap or B-Tap. First Decision Tree, DT1, predicts +1 if the fault has occurred on line A-Tap and value -1 if not. Second Decision Tree, DT2, makes decision for the line B-Tap. If both the trees have predicted the value of -1, then the fault has occurred in line C-tap. The simulated data is used to train the Decision Tree models. Max-depth parameter has been used to restrict the tree from over-fitting the training data. In case of a complex system with *N* lines, *N-1* Decision Trees would be used to make predictions.

The fault location for the two terminal circuits do not require previous step as they have only one line and therefore fault distance can be calculated as described below.

Once the line is identified, the fault location equations in [52] is used to calculate the fault distance. The aerial mode voltages are obtained by,

$$S_{mode} = T\, S_{phase}$$

where, S_{mode} and S_{phase} are the modal and phase signals (voltages) vectors respectively, and T is Clarke’s constant. All transmission line models are assumed to be fully transposed, and therefore the well-known Clarke's constant and real transformation matrix given by:

$$T = \frac{1}{3}\begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & -1 \\ 1 & \sqrt{3} & -\sqrt{3} \end{bmatrix}$$

is used. Clarke's transformation is real and can be used with any transposed line. If the studied line is untransposed, then an eigenvector-based transformation matrix, which is frequency dependent, will have to be used. This matrix should be computed at a frequency equal or close to the frequency of the initial fault transients.

If the fault is in A-Tap or B-Tap, fault distance from Tap point is calculated using:

$$S = L_{AT} - \frac{v^{line}\,\Delta t}{2}$$

Where LAT - length of the line A-Tap. LBT for line B-Tap

v^{line} - velocity of the traveling wave in mi/s = ~ speed of light in aerial mode [57]

$$v^{line} = 1.85 * 10^5\ mi/s$$

Δt - is the time difference between the first and second peaks of DWT coefficients squared of aerial mode voltages at Terminal A or B respectively. [Clarks transformation]. DWT gives the higher frequency transients generated by the fault.

For fault in C-Tap, the fault distance is calculated using:

$$S = L_{AT} + \frac{v^{line}\,\Delta t}{2}$$

Where L_{AT} - length of the line A-Tap.

v^{line} - velocity of the traveling wave in mi/s = ~ speed of light in aerial mode [57]

$$v^{line} = 1.85 * 10^5\ mi/s$$

Δt - is the time difference between the first and second peaks of DWT coefficients squared of aerial mode voltages at Terminal A.

The Two-Terminal transmission model consists of a single line. Therefore, line identification step could be skipped as shown in Figure 20.

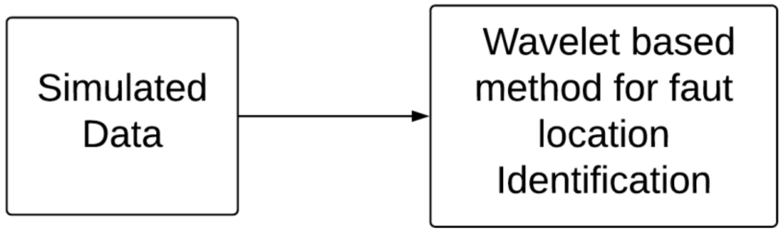

Figure 20: Fault location Identification Process in two terminal circuit

The fault distance is calculated from terminal A (reference terminal) using:

$$S = \frac{v^{line}\,\Delta t}{2}$$

Where v^{line} - velocity of the traveling wave in mi/s

Δt - is the time difference between the first and second peaks of DWT coefficients squared of aerial mode voltages at Terminal A.

3.5 Architecture II

In Architecture I, the fault type classification done by using four decision tree classifiers that did not generalize well and couldn't predict the test set well. Whereas, the fault location method used travelling wave method, which though accurate requires higher sampling rate for a reasonable accuracy. This section describes the fault classification and location identification method using a different approach. The simulated dataset for this process consists of wavelet energies of post-fault transient phase voltages as features described in previous sections and multi-class categorical target labels for fault classification. Each short circuit fault (e.g. "AB") is considered as one class, with total 10 classes. In this architecture as shown in Figure 21, the fault type classification uses multi-class SVM classifier and the fault location identification is done by regression model which predicts the distance of the fault. However, for three-terminal circuits the faulty line is identified before calculating the fault distance.

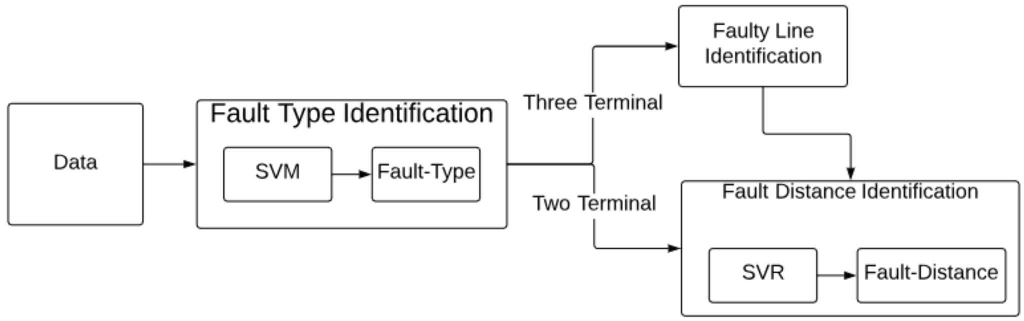

Figure 21: Overview of Architecture II

Sections 3.3.1 provides a brief overview of the Support Vector Machine algorithm from the point of view of the current problem and 3.3.2 has a brief description of Regression model.

3.5.1 Review on Support Vector Machine

The dataset for the fault classification has multiclass Label. Multiclass classification (*not to be confused with multilabel classification)* means a classification task with more than two classes; e.g., classify a set of images of fruits which may be oranges, apples, or pears. Multiclass classification assumes that each sample is assigned to one and only one label: a fruit can be either an apple or a pear but not both at the same time.

Support Vector Machine (SVM) algorithm was first introduced by Vapnik, Vladimir Naumovich [58].It is a supervised learning model which uses maximum margin hyperplane to create decision boundary between two different classes. A support vector machine constructs a hyperplane or set of hyperplanes in a high- or infinite-dimensional space, which can be used for classification, regression, or other tasks like outliers detection. Intuitively, a good separation is achieved by the hyperplane that has the largest distance to the nearest training-data point of any class (so-called functional margin), since in general the larger the margin the lower the generalization error of the classifier.

There are two approaches to use SVM in multiclass mode:

1) "*One vs one*" mode where a classifier is created for each pair of classes. Therefore, number of classifiers = n (n-1)/2 where n is number of classes. For 10 classes, it would create 450 classifiers. This method is resource consuming and slow.
2) "*One vs rest*" mode creates one classifier per class. Linear SVM uses one vs the rest scheme

to build classifiers. This method would give 10 classifiers. This mode has been chosen to build fault type classifiers for this problem.

Linear SVM creates a maximum-margin hyperplane separating data points from two different classes. In case of multiclass, "one vs rest" when building a hyperplane for each class, other classes are considered negative class as shown in Figure 22.

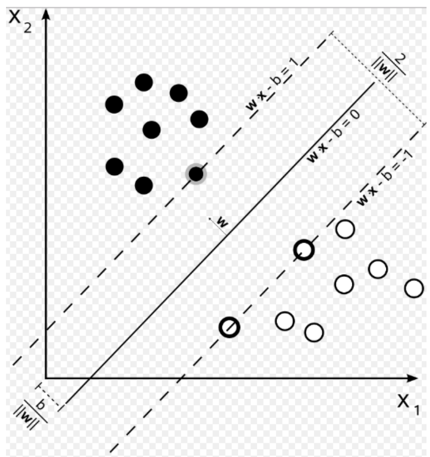

Figure 22: Maximum-margin hyperplane and margins for an SVM trained with samples from two classes. Samples on the margin are called the support vectors [59].

If we are given a set of N data points form of (vector x, linear y) (x_1, y_1) …. (x_n, y_n) where y_n is either +1 (positive class), -1 (negative class) and x is m-dimensional real vector. The goal is to find the "maximum-margin hyperplane" that divides group of points x_i for which $y_i = +1$ from the other points of x_i for which $y_i = -1$. Any hyperplane can be written as:

$$\vec{\omega}^T \vec{x} = -b$$

Where $\vec{\omega}$ is a normal vector to the hyperplane. The parameter $\frac{b}{||\vec{\omega}||}$ determines the offset of the hyperplane from the origin along the normal vector $\vec{\omega}$.

Linear classifier is then given as:

$$f(\vec{x}) = sign(\vec{\omega}^T \vec{x} + b)$$

3.5.2 Review on SVM Regression

Fault distance identification is done by Regression model. Regression analysis is used to estimate the dependent variable (continuous) given a set of independent variables. Since there is a non-linear relation between variables, for fault location, SVR (support vector Regression) is used to predict the distance, which is the dependent variable and the post-fault wavelet energies are the independent variables.

A version of SVM for regression was proposed by Vapnik et al. [58] .This method is called Support Vector Regression (SVR). The model produced by support vector classification (as described above) depends only on a subset of the training data, because the cost function for building the model does not care about training points that lie beyond the margin. Analogously, the model produced by SVR depends only on a subset of the training data, because the cost function for building the model ignores any training data close to the model prediction.

Training the original SVR is to:

$$minimise\ \frac{1}{2}\ ||\omega||^2$$

$$subject\ to \begin{cases} y_i - (\omega, x_i) - b\ \le\ \varepsilon \\ (\omega, x_i) + b -\ y_i \le \varepsilon \end{cases}$$

where x_i is the training sample with target value y_i. ε is a free parameter that serves as a threshold. The prediction of a given sample is $(w, x_i) + b$.

3.5.3 Fault Classification

In this section, the fault classification method for three and two terminal transmission models is presented. Fault type classification is performed using a single multiclass SVM as shown in the Figure 23.

Figure 23: Fault Classification Process

The simulated dataset consisting of the wavelet energies of the post-fault transient phase voltages as features and fault type (‘AG’, ‘ABG’ …) as labels is used to train the SVM model. SVM works well with continuous and nonlinear data and thus proves as a good fit for the fault classification problem.

3.5.4 Fault Location Identification

After the fault type has been identified, the next step is to locate the fault. For Three-Terminal (TEED) Circuits, the faulty line (branch) is identified first. Figure 24 shows the process of fault location identification specifically for TEED circuits.

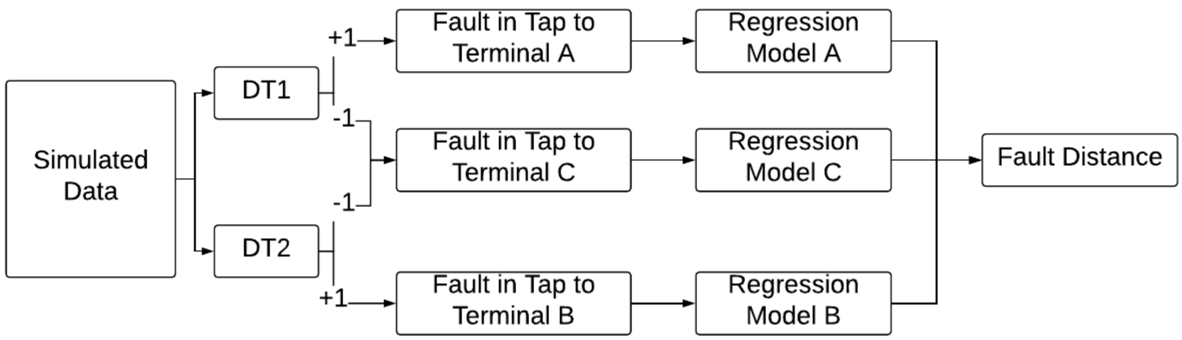

Figure 24: Fault Location Identification Process in Three-terminal Circuits

Two Decision Trees DT1 and DT2 is used to predict whether the fault has occurred in A-Tap or B-Tap respectively. When the predictions of both the trees is -1 simultaneously, then the fault has occurred in C-Tap line. This step is important to provide a reference point for the regression model as the distances and the input feature space are non-linear in nature.

Once the faulty-line has been identified, respective regression models are used to predict the fault distance. The fault distance obtained is the distance of the fault from the Tap point. The regression models are trained for each line of the transmission network and it remains consistent for all types of short circuit faults.

Figure 25: Fault Location Identification process in Two-Terminal circuits

For Two- Terminal circuits, the simulated data is used to train the regression model as Shown in Figure 25. The regression model predicts the fault distance from Terminal A. Since the features are linearly related to the distance, Linear SVR is used for prediction of the fault distances. Next chapter describes the implementation of both the architectures followed by results.

Chapter 4: Implementation and Evaluation

4.1 Experimental Setup

The three terminals and two terminal networks from Dominion Energy, Virginia, USA Transmission network in North Eastern region was used to experiment the proposed architecture. The testbed contains two-terminal and three-terminal 230KV transmission lines shown below in Figure 26 and Figure 27 with their configuration details shown tabulated in Table 1 and Table 2 respectively.

Figure 26: Two-terminal Transmission Line Details

A-B details:
Length- 24.64 mi
R=0.003
X = 0.0338
R_0 = 0.0214
X_0 = 0.0887

Table 1: Line details of two terminal Circuit

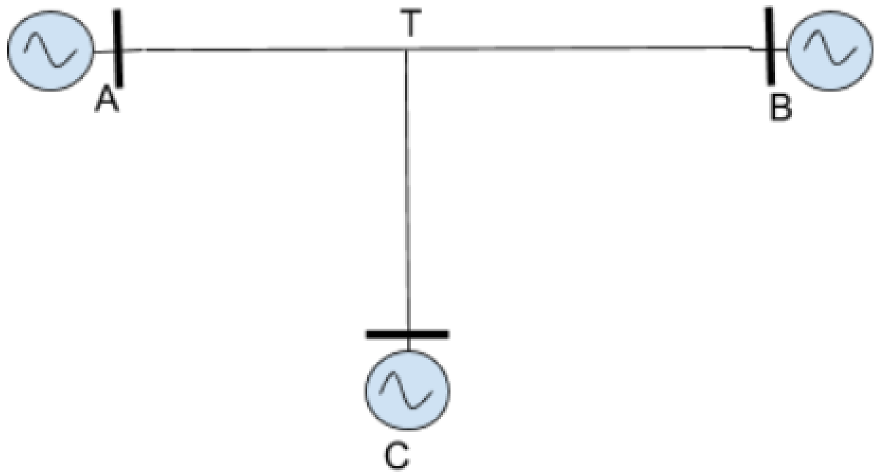

Figure 27: Three-terminal Transmission Lines Details

A-T details:	B-T details:	C-T details:
Length- 25.55 mi	Length- 25.33 mi	Length- 36.91 mi
R=0.004	R=0.0045	R=0.00659
X = 0.0277	X = 0.0275	X = 0.433
R_0 = 0.0229	R_0 = 0.0321	R_0 = 0.03992
X_0 = 0.0742	X_0 = 0.0867	X_0 = 0.12675

Table 2: Line Details of the three-terminal transmission model

An evaluation framework was created used to deploy the two architectures. The both transmission models were emulated in Aspen One Liner [58] to collect the simulated data. *Aspen OneLiner* is a PC-based short circuit and relay coordination program. It is a productivity tool which is useful for simulations. In aspen one-liner, any type of fault could be simulated under different conditions and fault data for relay testing can be exported. All 10 types of shunt faults as mentioned in *Table 3* were simulated at every 1%, 2%,3%....99% of each line and instantaneous post-fault phase voltages from each terminal was collected over one cycle. Aspen OneLiner Script was written to simulate and collect the data which takes about 60 seconds.

Types of Shunt faults	Fault Type
	AG (phase A to Ground)

Line to ground faults	BG (phase B to Ground)
	CG (phase C to Ground)
Double Line to ground faults	ABG (phase A-phase B to Ground)
	BCG (phase B-phase C to Ground)
	ACG (phase A- phase C to Ground)
Three Line to ground faults	ABCG (phase A- phase B-phase C to Ground)
Line to Line faults	AB (phase A to phase B)
	BC (phase B to phase C)
	AC (phase A to phase C)

Table 3 : Types of Shunt Faults

Python was used in the framework for data preparing which consists of computing wavelet energies for the transient post fault phase voltages collected, normalization and then assigning target classes to each sample. The simulated data was used to train and validate the models. Real fault data were collected from Digital Fault Recorders placed at the terminal of the Dominion Energy's transmission line under study. The test set consists of real fault data which is independent of the training dataset (simulated) and used only to assess the performance (i.e. generalization) of a fully specified classifier. For three terminal line configuration, the test set consists of 2 samples (two faults) whereas two terminal test set consists of one sample (one fault).

Classifiers were built with scikit-learn: machine learning library in Python [60]. Training and testing time for the classifiers were within few microseconds/milliseconds. The experiment was carried on a machine with 8GB RAM, intel i5 processor and 1TB memory.

4.2 Methods to Evaluate the model

This section briefly describes various metrics used in evaluating the machine learning models. From the training dataset presented to the algorithm, it tries to tweak its internal parameters to better understand the data. If the model is over trained meaning the model tries to mimic the exact behavior of the training data, then it will be able to identify all the relevant information in the training data but will fail miserably when presented with the new data which is independent of the training set. Then it is said that the model is not generalizing well or that it is overfitting the training data. Below are few metrics which help in estimating the generalization capabilities of a given model:

1) *Cross validation*: One of the model validation techniques for assessing how the results of a statistical analysis will generalize to an independent data set. It is mainly used in a supervised setting. The goal of cross-validation is to test the model's ability to predict new data that were not used in estimating it, to flag problems like overfitting.

   To evaluate the models of the proposed model k-fold cross validation is used where k = 5. It divides the complete dataset into 5 sets, then trains model on 4 sets and evaluates performance on the remaining set. This is repeated 5 times and the 5 results can then be averaged to produce a single estimation which is used to present the result.

2) *Absolute Error (AE):* This measure is used to evaluate regression models. It is the magnitude of the difference between the exact value and the approximation. If $Y_{predicted}$ is a value of single prediction generated from a sample on all variables, and Y_{true} is the observed value of the variable being predicted, then AE of the predictor is computed as:

$$AE = \left|\left|Y_{true} - Y_{predicted}\right|\right|$$

4.3 Architecture I Results

In this section, testing is carried out for the fault classification and location identification process for architecture I. The evaluation results presented are the accuracy measures for both two and three terminal transmission models.

4.3.1 Two Terminal Transmission Line Configuration

In Architecture I, Fault Classification process for two terminal circuits have four decision trees which predict the involvement of three phases and ground in a fault respectively. Table 4 shows the fault classification results for two terminals. All the numbers in the table represent accuracy of the model. DT- A is the decision tree for phase A. Similarly, DT-B, DT-C and DT-G is classification trees for B, C and Ground phase. These columns present accuracy measure for the individual tree. Since the result at the end is combination of 4-bits predicted by 4 trees, it is crucial to study the accuracy of this combined result as well. Column 4/4-bit accuracy represents the accuracy of correct prediction where a prediction is considered correct only if all the four bits predicted by each of the four decision trees match the actual value. All the incorrect predictions are discarded in calculation of accuracy. This metrics also gives the threshold value or tolerance for a prediction to be considered as incorrect. Similarly, ¾ bit accuracy is the measure of accuracy when a prediction is considered correct only if at least 3 bits out of 4 are correct.

Metrics	DT-A	DT-B	DT-C	DT-G	4/4-bit accuracy	3/4-bit accuracy
Baseline	90%	90%	90%	90%	90%	100%
5-fold cross validation on 100% training set	87.22%	87.22%	87.23%	89.99%	90%	100%
5-fold cross validation on 80% training set	100%	94.23%	100%	100%	94.56%	100%
20% validation set	100%	94%	100%	100%	100%	100%
Test set	100%	100%	100%	100%	100%	100%

Table 4: Architecture I Two-terminal Fault Classification Results

Baseline is the minimum accuracy a model should achieve for it to be considered good. It gives a measure which can be used to compare the model's performance. In this case, baseline is 90% meaning that the currently used technology to classify and locate faults gives right outcome 90% of the times. 5-fold cross validation on 100% of the training data gives us a measure of generalization of the model over whole training set. To test the model, the training set is divided into 80% training set and 20% validation set. The results of fault classifiers show that they perform well on validation set.

	Actual value				Predicted value			
Classifiers	Phase A	Phase B	Phase C	Phase G	Phase A	Phase B	Phase C	Phase G
Fault 1	-1	-1	+1	+1	-1	-1	+1	+1

Table 5: Architecture I Two-terminal Fault Classification Model Prediction Results

Lastly, the models are tested on the real-dataset and all the faults are predicted exactly right. Predictions from the classifying trees are shown in Table 5. We see that all the predicted values from the classifying trees match the actual or True value.

The real fault data has been collected from Digital Fault Recorders (DFR) placed at the terminals and for two-terminal transmission line the DFR's used have a sampling frequency of 5.7 KHz. Therefore, the time between the samples of the time series fault data collected from DFR is given by:

$$T_s = \frac{1}{F_s} = \frac{1}{5700} = 0.175\ ms$$

Discrete wavelet transform is applied at scale 1 to the fault data to obtain the wavelet coefficients. The approximation coefficients and Detailed coefficients have half the samples that of original signal. Approximation coefficients are the high frequency components of the signal, whereas detail coefficients capture the low frequency components of the signal. The difference between the arrival times of waveform reflections at one terminal is used to compute fault distance. The transmission model under study is 24.64 mi long and it takes about 0.133µs for the voltage signal in aerial mode to travel from one end of the line to other. Therefore, to capture the reflection waves the sampling time should be less than 0.133 µs or sampling frequency should be greater than at least 7.5MHz. Since the frequency of the DFR is much lower, the wavelet coefficients have overlapping peaks which makes it impossible to calculate the time difference between peaks which is required to calculate the Δt to be imputed in the below equation to calculate the distance of the fault.

$$S = \frac{v^{line}\,\Delta t}{2}$$

Therefore, travelling wave method could only be successfully applied if sampling time is much lower than the time taken for voltage signal in aerial mode to travel the through the total length of the line. That is, sampling frequency of the DFR's should be higher for shorter lines or transmission line length should be greater to use DFR's with lower sampling frequency. Secondly, while calculating the time difference between the two arrival peaks manually, since these quantities are in μs, there might be human precision error which could lead to difference in the final distance calculation.

4.3.2 Three Terminal Transmission Line Configuration

For three terminal transmission line configuration, the fault classification has same process as that of two terminals with four decision trees, but the fault location process has two additional decision trees DT1 and DT2 to identify the faulty line. Table 6 shows the resulting accuracies of all the 6 decision trees for various metrics. All the numbers in the table represent accuracy of the model. Accuracy here means percentage of correct prediction over total prediction.

Metrics	DT-A	DT-B	DT-C	DT-G	DT1	DT2	6/6 bits accuracy	5/6 bits accuracy
Baseline	80%	80%	80%	80%	80%	80%	80%	90%
5-fold cross validation on 100% training set	98.5 %	88.5 %	96.77 %	98.2 %	91.5%	92.1%	92%	100%
5-fold cross validation on 80% training set	100%	94.23 %	100%	100%	96.12 %	95.6%	94%	100%
validation set 20%	100%	94%	100%	100%	95.2%	90.7%	84.84%	100%
Test set	100%	100%	100%	50%	100%	100%	50%	100%

Table 6: Architecture I Three-terminal Fault Classification and Location Results

All the individual classifying trees perform better than the baseline. Column 6/6 bits accuracy, is the accuracy of correct prediction where a prediction is considered correct only if all the 6 bits predicted from the 6 decision trees classifiers is correct. For the validation set, approximately 85% of the validation set was predicted correctly when 6/6-bit accuracy is considered and 100% of the validation set was predicted correctly when ⅚ bit accuracy was the threshold. Therefore, 15% incorrect prediction were considered incorrect as one of the bits out of 6 were predicted wrong. This is due to errors of each tree cascading for the final error.

The test dataset consists of the real data for two faults i.e. two samples. From the table it is evident that the accuracy of the decision tree for phase A, B and C are 100% meaning both the samples in the test set is classified correctly. Whereas, decision tree for ground phase is 50%. This is because decision tree for ground phase predicts the second fault incorrectly as shown in Table 7. Therefore, when we consider a strict measure of 6/6-bit accuracy, if the prediction has even one of the bits incorrect the prediction is considered incorrect. Therefore, since only one sample out of two was predicted correctly, accuracy is 50%. On the other hand, the threshold for error is increased in 5/6-bit accuracy, both the test samples were correctly predicted.

	Actual value						Predicted value					
Classifiers	DT-A	DT-B	DT-C	DT-G	DT1	DT2	DT-A	DT-B	DT-C	DT-G	DT1	DT2
Fault 1	+1	+1	-1	-1	-1	+1	+1	+1	-1	-1	-1	+1
Fault 2	+1	-1	-1	+1	-1	+1	+1	+1	-1	-1	-1	+1

Table 7: Architecture I Three-terminal Model Prediction Results

The individual bit predictions from 6 decision trees has been illustrated in Table 7. On a closer look its seen that the Decision Tree for Ground phase has predicted -1 instead of actual value +1, whereas the fault 1 is predicted correct.

Once the type of fault is identified, the next step is location identification. The results of few scenarios have been presented below. The scenario assumes that the fault type identification is done prior. Scenarios 1 to 3 are simulated faults data to calculate the fault distance.

Scenario 1: The first scenario assumes the fault is phase A-to-Ground that has occurred in line A-Tap at 12.775 miles from Bus A.

The fault type was accurately classified since DT-A (corresponding to phase A) and DT-G (corresponding to ground phase) had predicted +1 while another DT's predicted -1. The two faulty-line identifiers DT1 and DT2 predicted +1 and -1 respectively indicating that the fault is the line A-Tap.

Discrete wavelet transform is applied to aerial mode voltage at Bus A and the WTC^2 are plotted with respect to time shown in Figure 28.

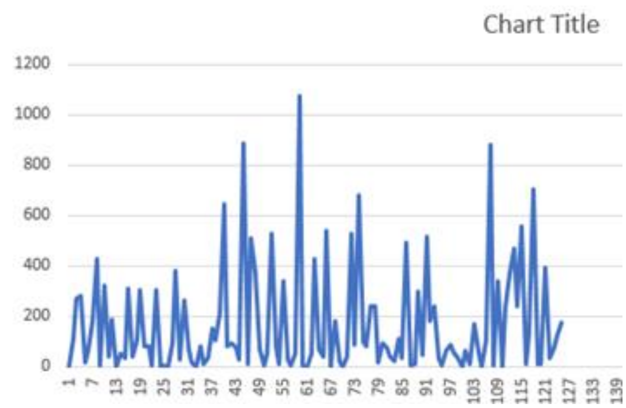

Figure 28: WTC^2 vs time at Terminal A for fault on line A-Tap

Δt (time difference between first and second peaks) = 59 - 45 = 14 x10^{-5} s

Fault location can be calculated using:

$$S = \frac{(v_i^{line} * \ \Delta t)}{2}$$

$$= \frac{(1.85*10^5)(14* \ 10^{-5})}{2} = 12.95 \text{ mi}$$

Absolute Error in calculation is ~ 0.175 mi

Scenario 2: The second scenario assumes the fault is line to line i.e. phase B-to-phase C that has occurred in line B-Tap at 2.533 miles from Tap point towards bus B.

The fault type was accurately classified since DT-B (corresponding to phase B) and DT-C (corresponding to phase C) had predicted +1 while another DT's predicted -1. The two faulty-line identifiers DT1 and DT2 predicted -1 and +1 respectively indicating that the fault is the line B-Tap.

Discrete wavelet transform is applied to aerial mode voltage at Bus B and the WTC^2 are plotted with respect to time shown in Figure 29: WTC2 vs time at line B-Tap for fault on line B-Tap.

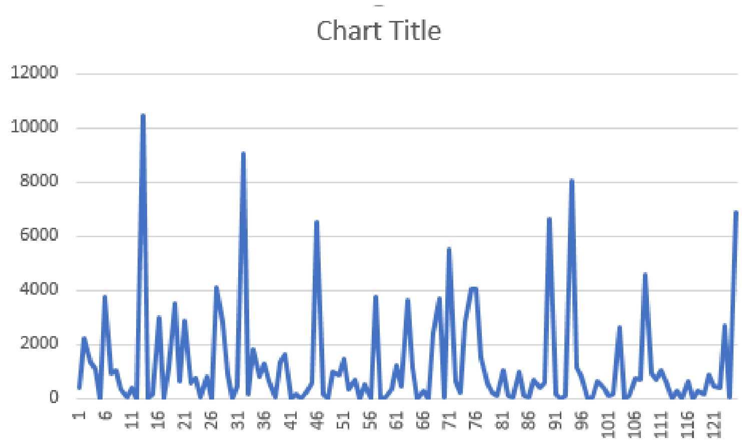

Figure 29: WTC^2 vs time at line B-Tap for fault on line B-Tap

Δt (time difference between first and second peaks) = 16 -13 = 3 $\times 10^{-5}$s

Fault location can be calculated using:

$$S = \frac{(v_i^{line} * \Delta t)}{2}$$

$$= \frac{(1.85 * 10^5)(3 * 10^{-5})}{2} \quad = \mathbf{2.775\ mi}$$

Absolute Error in calculation is ~ 0.2 mi

Scenario 3: The third scenario assumes the fault is Three-phase-to-ground that has occurred in line C-Tap at 34.55 miles from bus A.

The fault type was accurately classified since all the four DT's predicted +1. The two faulty-line identifiers DT1 and DT2 predicted -1 and -1 respectively indicating that the fault is the line C-Tap.

Discrete wavelet transform is applied to aerial mode voltage at Bus A and the WTC^2 are plotted with respect to time shown in Figure 30.

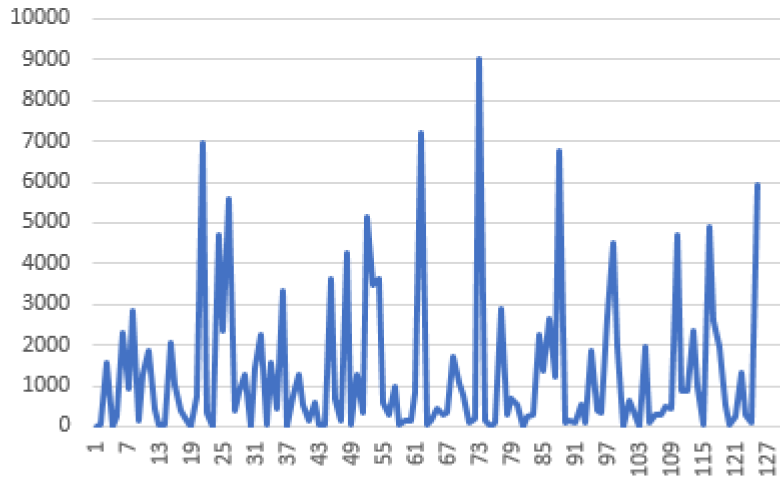

Figure 30: WTC^2 vs time at terminal A for fault on line C-Tap

Δt (time difference between first and second peaks) = 73-62 = 10 $x10^{-5}$ s

Fault location can be calculated using:

$$S = L_{AT} + \frac{v^{line}\,\Delta t}{2}$$

$$= 25.33 + (0.51x10^{6}) \text{ x } (10x10^{-5})/2 = 34.58 \text{ mi from A}$$

Absolute Error in calculation is ~ 0.03 mi

The error for faults on all the three terminal is within 1 mile which is within the acceptable range of ± 1% to ± 2% of the line length. To calculate the fault distance of the real fault (test data), the data from the DFR's at terminal A (4.8 KHz), terminal B (5.7KHz) and terminal C (9.6 KHz) is used. As we have concluded in previous section, the sampling frequency of DFR's need to be higher to capture two peaks in the wavelet transform, which is the case in three-terminal as well. This makes it impossible to locate the fault with travelling wave method in these real-world scenarios.

In architecture I, for three terminal line configuration, the test samples were not classified correctly and therefore had lower accuracy. Judging by its performance based on current results the accuracy might or might not improve when tested on more bigger set. Furthermore, the travelling wave-based method needs higher sampling frequency and has an associated human precision error which make it more challenging for it to be applied in real world. Therefore, architecture II results are investigated in Section 4.4 for fault classification and location identification on transmission line which uses multiclass support vector machine for predicting fault type and regression models for calculating the distance.

4.4 Architecture II Results

In this section, a different technique of fault type classification and location identification process is presented. The method is evaluated on two-terminal as well three-terminal transmission model. Following Section 4.4.1 presents the results for Two Terminal and Section 4.4.2 for Three Terminal model.

4.4.1 Two Terminal Transmission Line Configuration

In architecture II, the fault classification process consists of one Support Vector Machine (SVM) which could predict all 10 types (10 classes) of Short Circuit faults. Table 8 presents the accuracy results of fault classification for two terminal transmission lines. All the numbers in the table represent accuracy of the model. Accuracy here means percentage of correct prediction over total prediction.

Metrics	Fault Type
Baseline	90%
5-fold cross validation 80%	99.5%
5-fold cross validation 100%	95.37%
20% validation set	100%
Test set	100%

Table 8: Architecture II Two-terminal Fault Classification Results

The accuracies of the SVM based on various metrics is evaluated. Baseline prediction is 90% meaning the model should have at least 90% accuracy to have an acceptable result. Model validation technique of 5-fold cross validation on the entire training set had ~95% accuracy meaning on an average 95% of the samples were always classified correctly. Validation set is predicted with 100% accuracy. Test set consisting of real fault is correctly classified as shown in Table 9.

Faults	Actual Label	Predicted Label
Fault 1	CG	CG

Table 9: Architecture II Two-terminal Fault Classification Model Prediction Results

Once the fault type is identified, the next step is to locate the fault. In architecture II, the fault distance is predicted by the regression model. Table 10 shows the prediction results of regression model.

Faults	Expected distance from Terminal C	Model Prediction	Absolute Error
Fault 1	7.42 mi	7.5 mi	~0.08mi

Table 10: Architecture II Two-terminal Fault Location Results

The model has predicted 7.5 miles for a fault with expected distance of 7.42 mi from Bus C. The absolute error for the prediction is about ~0.08 mi which is well within acceptable ±1% of the line length (±0.2464 mi).

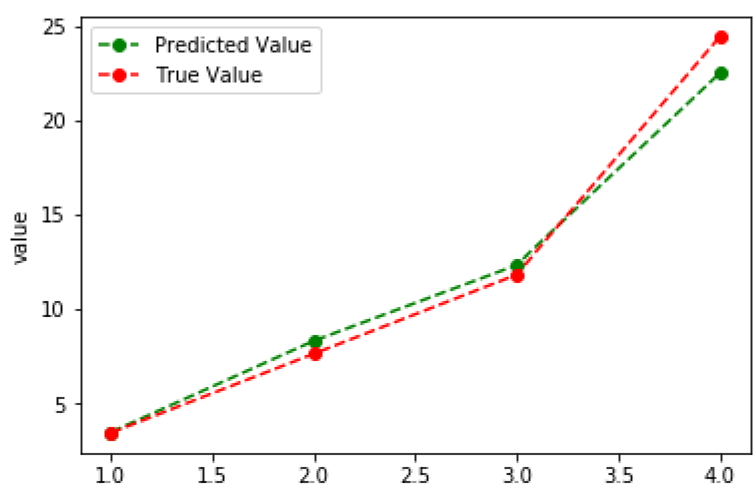

Figure 31: Actual vs Predicted distance plot for phase C-to-Ground fault (4 samples)

The test data (DFR fault data at the terminal) for the two-terminal transmission is available for only single fault which provides very less evidence to believe that the model has achieved good generalization. To validate the correctness of the model prediction, 4 samples were picked randomly from the simulated data which constituted of the validation set. The model was trained on rest of the samples. Figure 31 shows the actual distance vs the predicted distance plot for phase C- to-Ground fault. From the graph, its seen that the error in the prediction of the fault distances is well below the acceptable absolute error of ±1% to ±2% of the line length.

Absolute Error is given by $:= \| true\ fault\ distance - predicted\ fault\ distance \|$

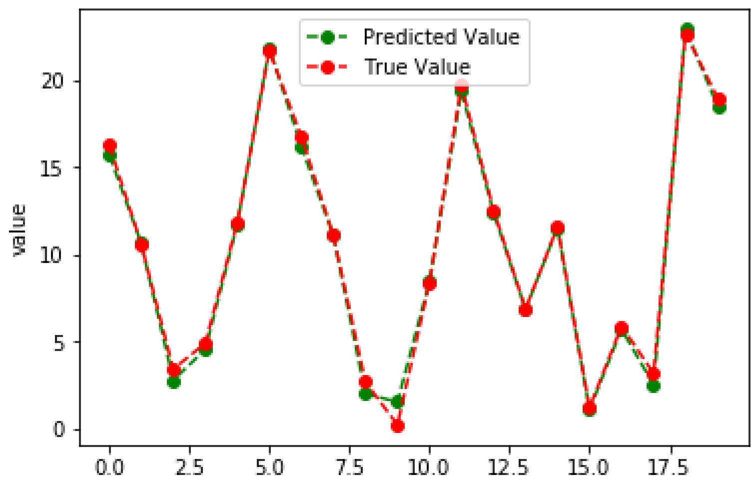

Figure 32: Actual vs predicted fault distances for phase C-to-Ground fault

To validate the robustness of the model, the training data was randomly sampled without replacement into 20% validation set and remaining 80% as training set. Figure 32 shows the graph of prediction vs actual true fault distance value. From the graph its seen that, the regression model has predicted fault distances with error less than ±1% of the length of the line (~ ±2.5 mile) which is a better than precision of other conventional methods.

4.4.2 Three Terminal Transmission Line Configuration

In architecture II, the fault location process is same for two and three terminal transmission models. Table 11 illustrates the Fault classification results for three terminal circuit. All the numbers in the table represent accuracy of the model. Accuracy here means percentage of correct prediction over total prediction.

Metrics	Fault-Type
Baseline	90%
5-fold cross validation 100%	99.2%
5-fold cross validation 80%	89.5%
20% validation set	100%
Test set	100%

Table 11: Architecture II Three-terminal Fault Classification Results

The model is validated with complete training set by 5-fold cross validation which has an accuracy of 89.5% very close to the baseline prediction. Model predictions are tested on validation set which is 20% of the samples in the training set. Finally, the model is tested on the test dataset which is independent of the training dataset achieving 100% accuracy on the test set meaning all the samples in the test set were predicted correctly. Table 12 shows the predicted value by the model.

Faults	Actual Label	Predicted Label
Fault 1	AB	AB
Fault 2	AG	AG

Table 12: Architecture II Three-terminal Fault Classification Model Prediction Results

Faults	Results from DT1 and DT2	Expected distance from TAP Point	Predicted distance from TAP Point	Absolute Error
Fault 1	BT	21.46 mi	20.77 mi	~0.69 mi
Fault 2	BT	14.308 mi	14.72 mi	~0.4 mi

Table 13: Architecture II Three-terminal Fault Location Results

Once the fault type is identified, the faulty-line identification is done by the two decision trees. Table 13 shows the fault location identification results. The first column has the results from two decision trees, DT1 and DT2. For both the faults in the test dataset, the output of DT1 and DT2 were -1 and +1 respectively indicating that the fault occurred is in line B-Tap. Then the respective regression models are used to predict the fault distance from the tap point. Column 2 and 3 give the expected and predicted distances. We observe that the error is less than 1 mile in both the cases, which is well within the acceptable error range of ±1 to ±2% of the length of the line.

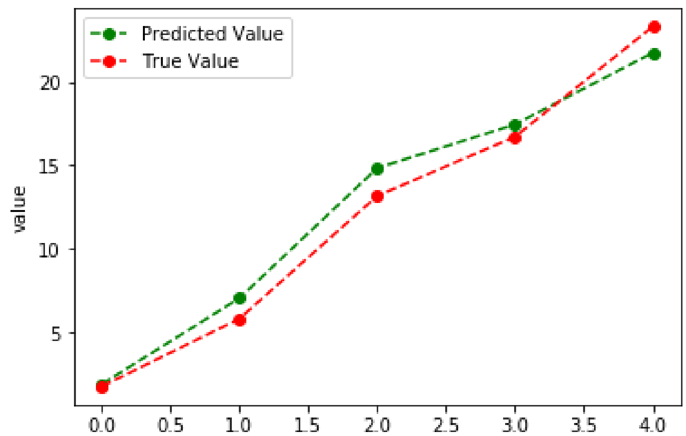

Figure 33: AG Fault on B-Tap line

For the three-terminal model, the test data (DFR fault data at the terminal) has only two samples of the real-world faults. Let's take a scenario where phase A-to-ground has occurred on line B-Tap. To validate the correctness of the model prediction,4 samples were picked randomly from the simulated data which constituted of the validation set which have extreme ranges to test the corner cases. The model was trained on rest of the samples. Figure 33 shows the actual distance vs the predicted distance plot for phase A- to-Ground fault. From the graph, its seen that the error in the prediction of the fault distances is well below the acceptable absolute error of ±1% to ±2% of the line length where Absolute Error is given by:

$$err = \|true\ fault\ distance - predicted\ fault\ distance\|$$

Figure 34 shows the comparison of true and predicted value on 20% validation set formed by randomly sampling the training data without replacements. Even for three-terminal circuit, which is complex, the regression model predictions are well within the acceptable absolute error.

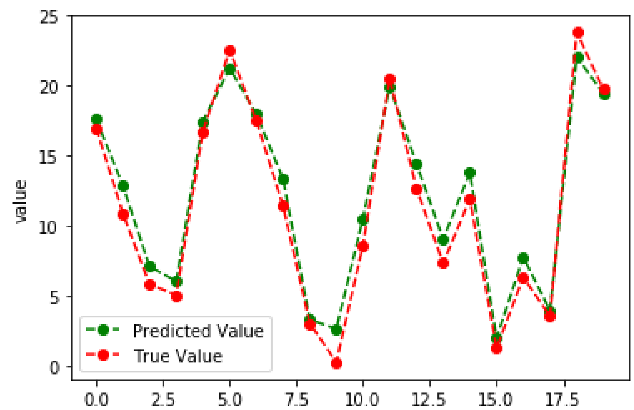

Figure 34: Actual vs predicted fault distances for phase A-to-Ground fault on validation set

4.5 No Fault Scenario

Under no fault scenario the power system remains at a steady state with specified current and voltage values. Therefore, the fault classification and location method could be deployed if the voltage or current values differ from the steady state values.

4.6 High Level Comparison of the Two Architectures

Both the architectures presented in this thesis implement machine learning algorithm to predict the fault type and distance on the transmission line and they can be applied to any configuration of the transmission line. However, the design criteria and the choice of algorithm vary between the architectures. The differences have been outlined below:

- Each machine learning model has some amount of prediction error associated with it. In architecture I, the fault classification approach uses four decision tree classifiers to predict the fault type. Unlike architecture I, in architecture II uses one multiclass- SVM classifier to predict the fault type. The prediction error associated with the model in architecture II is lower than that of architecture I.
- For fault location identification, architecture I uses traveling wave-based method, though accurate it requires higher sampling frequency and synchronized measurements at the terminal. Additionally, the travelling wave-based method has human error associated with it while calculating the difference between the arrival time of the peaks. Currently, there are no software's that can automate this process.

Whereas architecture II, uses regression model to calculate the fault distance which provides faster and reasonable accurate results.

Chapter 5: Conclusion

5.1 Summary and Findings

Transmission lines safeguard against exposed fault is the most critical task in the protection of power system. The purpose of a protective relaying is to identify the abnormal signals representing faults on a power transmission system. So, fault classification and location is necessary for reliable and high speed protective relaying.

The research work presented in this thesis provides two promising architectures for fault classification and location on transmission line using machine learning techniques. The first architecture deploys four decision tree models one for each phase for fault identification in each phase. "Divide and conquer" strategy has been used to classify faults on each phase with greatest accuracy. Each tree classifier has two possible target classes [-1, +1]. Class of '+1' indicates that a phase is involved in the fault. Once the type of fault is known, travelling wave method is used to locate the fault on the transmission line. The wavelet transformation coefficients (WTCs) of the post fault transient phase voltage at a terminal is computed. The time difference between first two peaks of WTC^2 is used to calculate the distance from a reference terminal. While travelling wave-based methods have few demerits, they are very reliable and can give accurate results. The method was tested on real data for both two and three terminal transmission models. The fault classification did better than baseline accuracy and the fault distance error calculated was within ±1% to ±2% of the line length.

Decision trees give better prediction with more data. Then we come to a question of how much training data is sufficient for getting a good prediction. SVM is an attractive classifier when classification process needs to be made useful and economical with smaller number of samples. SVM models have similar functional form to neural networks and radial basis functions, both popular data mining techniques. However, neither of these algorithms has the well-founded theoretical approach to regularization that forms the basis of SVM. The quality of generalization and ease of training of SVM is far beyond the capacities of these more traditional methods which makes it a better pick than decision trees.

In the second architecture, multi-class support vector machine model is used to classify 10 types of short circuit faults. When tested SVM fault classification model on real data, it outperformed decision trees with 100% accuracy on two and three terminal transmission lines. For fault location, Support Vector Regression was used for three terminal transmission line and Linear Regression for two terminal transmission lines. The error of the fault distance was same as the travelling wave-based method used in first architecture. However, regression models don't require any manual calculations or measurements.

Since both the architecture do not use current signals, the presented approach is immune against current-transformer saturation and its related errors. Additionally, the presented method uses simulated data for training which means it can be deployed even when a new transmission line is added to the transmission network which does not have any historical records of fault. The presented architectures in this thesis are tested and validated against real data of faults. In this thesis, the test set sizes for both two and three terminal line configuration had only three samples in total. Therefore, the method needs to be tested more robustly. Overall, preliminary findings are promising but more data is needed to gain confidences in the reliability of the methods.

5.2 Challenges

In the initial stages, collecting the simulated data had two stages. First, the instantaneous post fault transients simulated in ASPEN were collected in .CFG as ASPEN only had one format option to save the data. After collecting the simulated data, the CFG was saved to a CSV file with help of tool called WAVEWIN (use to analyze the signals). This process creating each file manually was very tedious even to have a reasonable number of samples. To overcome this, a ASPEN script was written which could write the instantaneous values directly to the CSV file which reduced the data preparation time to 30 secs which otherwise would take several hours manually.

Secondly, since transmission line faults are rare compared to the faults on distribution lines, there was availability of limited real data to test the presented architectures. However, the preliminary results show strong potential in the architectures.

5.3 Future Work

In this thesis, problem of a having an adaptive, effective and resilient method for fault classification and location is addressed. There are several lines of research arising from this work which could be pursued as mentioned below:

- Since the test set considered in the thesis had just three samples in total (small size), the architectures need to be tested and evaluated with more test samples robustly in future. Additionally, architecture I had low accuracy of about 50% on the test set which may be due to low sample size. Therefore, with more test set samples, this result needs to be re-evaluated. However, the preliminary results are promising to lead to further investigation.
- Four Terminal transmission lines aren't very far from implementing. The proposed method may be possible to be extended N-terminal transmission line in the future work which would give different insights.
- The described method assumes single fault on the transmission line. However, chances of multiple faults, double circuit faults etc. though are low cannot be ignored from protection perspective.
- "Evolving Faults" are faults beginning in one phase and spreading to another phase after a few cycles. Thus, evolving fault consists of two faults: primary fault and secondary fault according to their fault inception times. Extending proposed method to classify and locate these faults would be an interesting topic for future work.
- The method proposed could be used for real-time monitoring of grid using spark-python. Once the fault has occurred, an alert could be set up from the method which would give the fault type and distance instantly.
- Complex fault sequence elements could be incorporated in data which will increase the robustness of the machine learning technique and results could be compared.

Bibliography

[1] M. Singh, B. K. Panigrahi, and R. P. Maheshwari, “Transmission line fault detection and classification,” in *2011 International Conference on Emerging Trends in Electrical and Computer Technology*, 2011, pp. 15–22.

[2] “SPCTF-3TerminalLines091906.pdf.” .

[3] “Electric power transmission,” *Wikipedia*. 12-Jul-2018.

[4] M. Gilstrap, “United States Electricity Industry Primer,” p. 94.

[5] “How much electricity is lost in transmission and distribution in the United States? - FAQ - U.S. Energy Information Administration (EIA).” [Online]. Available: https://www.eia.gov/tools/faqs/faq.php?id=105&t=3. [Accessed: 18-Jul-2018].

[6] “Protective relay,” *Wikipedia*. 27-Jun-2018.

[7] A. Guzman, B. Kasztenny, Y. Tong, and M. V Mynam, “Accurate and Economical Traveling-Wave Fault Locating Without Communications,” 2017.

[8] “SPCTF-3TerminalLines091906.pdf.” .

[9] “Three-phase Y and Delta Configurations | Polyphase AC Circuits | Electronics Textbook.” [Online]. Available: https://www.allaboutcircuits.com/textbook/alternating-current/chpt-10/three-phase-y-delta-configurations/. [Accessed: 19-Jul-2018].

[10] T. Gönen, *Electric power transmission system engineering analysis and design*, 2nd ed. Boca Raton: CRC Press, 2009.

[11] “Electrical Power Systems, Fifth Edition by Wadhwa, C.L.: New Age International, New Delhi 9788122424683 Soft cover - Books in my Basket.” [Online]. Available: https://www.abebooks.com/servlet/BookDetailsPL?bi=22644546648&searchurl=tn%3Delectrical%2Bpower%2Bsystems%26sortby%3D17%26an%3Dwadhwa&cm_sp=snippet-_-srp1-_-title1. [Accessed: 13-Jul-2018].

[12] H. Saadat, *Power System Analysis*. PSA Publishing, 2010.

[13] D. P. Mishra and P. Ray, “Fault detection, location and classification of a transmission line,” *Neural Comput. Appl.*, Dec. 2017.

[14] A. Prasad, J. Belwin Edward, and K. Ravi, “A review on fault classification methodologies in power transmission systems: Part-II,” *J. Electr. Syst. Inf. Technol.*, vol. 5, no. 1, pp. 61–67, May 2018.

[15] "Machine learning," *Wikipedia*. 11-Jul-2018.

[16] "Support vector machine," *Wikipedia*. 15-Jul-2018.

[17] "Overfitting," *Wikipedia*. 30-Jun-2018.

[18] N. Ramesh Babu and B. Jagan Mohan, "Fault classification in power systems using EMD and SVM," *Ain Shams Eng. J.*, vol. 8, no. 2, pp. 103–111, Jun. 2017.

[19] Y. Guo, K. Li, and X. Liu, "Fault Diagnosis for Power System Transmission Line Based on PCA and SVMs," in *Intelligent Computing for Sustainable Energy and Environment*, vol. 355, K. Li, S. Li, D. Li, and Q. Niu, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 524–532.

[20] V. Malathi and N. S. Marimuthu, "Multi-class Support Vector Machine approach for fault classification in power transmission line," 2008, pp. 67–71.

[21] H. C. Dubey, A. K. Tiwari, Nandita, P. K. Ray, S. R. Mohanty, and N. Kishor, "A novel fault classification scheme based on least square SVM," 2012, pp. 1–5.

[22] H. Livani, "A Fault Classification Method in Power Systems Using DWT and SVM Classifier," p. 5.

[23] O. A. S. Youssef, "An optimised fault classification technique based on Support-Vector-Machines," in *2009 IEEE/PES Power Systems Conference and Exposition*, 2009, pp. 1–8.

[24] U. B. Parikh, B. Das, and R. Maheshwari, "Fault classification technique for series compensated transmission line using support vector machine," *Int. J. Electr. Power Energy Syst.*, vol. 32, no. 6, pp. 629–636, Jul. 2010.

[25] P. K. Dash, S. R. Samantaray, and G. Panda, "Fault Classification and Section Identification of an Advanced Series-Compensated Transmission Line Using Support Vector Machine," *IEEE Trans. Power Deliv.*, vol. 22, no. 1, pp. 67–73, Jan. 2007.

[26] P. Tripathi, A. Sharma, G. N. Pillai, and I. Gupta, "Accurate Fault Classification and Section Identification Scheme in TCSC Compensated Transmission Line using SVM," *Int. J. Electr. Comput. Eng.*, vol. 5, no. 12, p. 7, 2011.

[27] A. Prasad, J. Belwin Edward, and K. Ravi, "A review on fault classification methodologies in power transmission systems: Part—I," *J. Electr. Syst. Inf. Technol.*, vol. 5, no. 1, pp. 48–60, May 2018.

[28] "Artificial neural network," *Wikipedia*. 31-Jul-2018.

[29] E. Koley, K. Verma, and S. Ghosh, "An improved fault detection classification and location scheme based on wavelet transform and artificial neural network for six phase transmission line using single end data only," *SpringerPlus*, vol. 4, no. 1, Dec. 2015.

[30] A. Abdullah, "Ultrafast Transmission Line Fault Detection Using a DWT-Based ANN," *IEEE Trans. Ind. Appl.*, vol. 54, no. 2, pp. 1182–1193, Mar. 2018.

[31] Y. S. Rao, G. R. kumar, and G. K. Rao, "A New Approach for Classification of Fault in Transmission Line with Combination of Wavelet Multi Resolution Analysis and Neural Networks," vol. 8, no. 1, p. 9, 2017.

[32] M. Saini, A. A. bin Mohd Zin, M. W. Bin Mustafa, A. R. Sultan, and Rahimuddin, "Transmission Line Using Discrete Wavelet Transform and Back-Propagation Neural Network Based on Clarke's Transformation," *Appl. Mech. Mater.*, vol. 818, pp. 156–165, Jan. 2016.

[33] M. Saradarzadeh and M. Sanaye-Pasand, "An accurate fuzzy logic-based fault classification algorithm using voltage and current phase sequence components: FUZZY LOGIC-BASED FAULT CLASSIFICATION ALGORITHM," *Int. Trans. Electr. Energy Syst.*, vol. 25, no. 10, pp. 2275–2288, Oct. 2015.

[34] A. Prasad, J. Belwin Edwar, C. Shashank Roy, G. Divyansh, and A. Kumar, "Classification of Faults in Power Transmission Lines using Fuzzy-Logic Technique," *Indian J. Sci. Technol.*, vol. 8, no. 30, Nov. 2015.

[35] S. Adhikari, N. Sinha, and T. Dorendrajit, "Fuzzy logic based on-line fault detection and classification in transmission line," *SpringerPlus*, vol. 5, no. 1, Dec. 2016.

[36] A. Jamehbozorg and S. M. Shahrtash, "A Decision Tree-Based Method for Fault Classification in Double-Circuit Transmission Lines," *IEEE Trans. Power Deliv.*, vol. 25, no. 4, pp. 2184–2189, Oct. 2010.

[37] P. K. Mishra, A. Yadav, and M. Pazoki, "A Novel Fault Classification Scheme for Series Capacitor Compensated Transmission Line Based on Bagged Tree Ensemble Classifier," *IEEE Access*, vol. 6, pp. 27373–27382, 2018.

[38] A. Asadi Majd, H. Samet, and T. Ghanbari, "k-NN based fault detection and classification methods for power transmission systems," *Prot. Control Mod. Power Syst.*, vol. 2, no. 1, Dec. 2017.

[39] P. Ray, B. K. Panigrahi, and N. Senroy, "Extreme Learning Machine based Fault Classification in a Series Compensated Transmission Line," p. 6.

[40] A. Dasgupta, S. Debnath, and A. Das, "Transmission line fault detection and classification using cross-correlation and k-nearest neighbor," *Int. J. Knowl.-Based Intell. Eng. Syst.*, vol. 19, no. 3, pp. 183–189, Oct. 2015.

[41] M. H. H. Musa, Z. He, L. Fu, and Y. Deng, "Linear regression index-based method for fault detection and classification in power transmission line: Linear Regression Index-Based Method for Fault Detection and Classification in Power Transmission Line," *IEEJ Trans. Electr. Electron. Eng.*, vol. 13, no. 7, pp. 979–987, Jul. 2018.

[42] M. M. Tawfik and M. M. Morcos, "ANN-Based Techniques for Estimating Fault Location on Transmission Lines Using Prony Method," *IEEE Trans. POWER Deliv.*, vol. 16, no. 2, p. 6, 2001.

[43] H. Fathabadi, "Novel filter based ANN approach for short-circuit faults detection, classification and location in power transmission lines," *Int. J. Electr. Power Energy Syst.*, vol. 74, pp. 374–383, Jan. 2016.

[44] A. Yadav and Y. Dash, "An Overview of Transmission Line Protection by Artificial Neural Network: Fault Detection, Fault Classification, Fault Location, and Fault Direction Discrimination," *Adv. Artif. Neural Syst.*, vol. 2014, pp. 1–20, 2014.

[45] P. Ray and D. P. Mishra, "Support vector machine based fault classification and location of a long transmission line," *Eng. Sci. Technol. Int. J.*, vol. 19, no. 3, pp. 1368–1380, Sep. 2016.

[46] Department of Electrical Engine ering, Islamic Azad University, South Tehran Branch, Tehran, Iran and K. Hosseini, "Short Circuit Fault Classification and Lo cation in Transmission Lines Using A Combination of Wavelet Transform and Support Vector Machines," *Int. J. Electr. Eng. Inform.*, vol. 7, no. 2, pp. 353–365, Jun. 2015.

[47] M. Farshad and J. Sadeh, "Accurate Single-Phase Fault-Location Method for Transmission Lines Based on K-Nearest Neighbor Algorithm Using One-End Voltage," *IEEE Trans. Power Deliv.*, vol. 27, no. 4, pp. 2360–2367, Oct. 2012.

[48] P. Ray, "Fast and accurate fault location by extreme learning machine in a series compensated transmission line," 2014, pp. 1–6.

[49] "Machine Learning: What it is and why it matters | SAS." [Online]. Available: https://www.sas.com/en_us/insights/analytics/machine-learning.html. [Accessed: 27-Jul-2018].

[50] "Aspen OneLiner." .

[51] S. Santoso, E. J. Powers, W. M. Grady, and P. Hofmann, "Power quality assessment via wavelet transform analysis," *IEEE Trans. Power Deliv.*, vol. 11, no. 2, pp. 924–930, Apr. 1996.

[52] F. H. Magnago and A. Abur, "Fault location using wavelets," *IEEE Trans. Power Deliv.*, vol. 13, no. 4, pp. 1475–1480, Oct. 1998.

[53] "Discrete wavelet transform," *Wikipedia*. 20-Jul-2018.

[54] O. Rioul and P. Duhamel, "Fast algorithms for discrete and continuous wavelet transforms," *IEEE Trans. Inf. Theory*, vol. 38, no. 2, pp. 569–586, Mar. 1992.

[55] J. Van Ryzin, "Review of Classification and Regression Trees," *J. Am. Stat. Assoc.*, vol. 81, no. 393, pp. 253–253, 1986.

[56] W.-Y. Loh, "Fifty Years of Classification and Regression Trees," *Int. Stat. Rev.*, vol. 82, no. 3, pp. 329–348.

[57] C. Christopoulos and A. Wright, *Electrical Power System Protection*. Springer Science & Business Media, 1999.

[58] "Vapnik, V. (1998). Statistical learning theory (Adaptive and learning systems for signal processing, communications, and control). New York: Wiley." .

[59] "Support vector machine," *Wikipedia*. 23-Jul-2018.

[60] "Documentation scikit-learn: machine learning in Python — scikit-learn 0.19.2 documentation." [Online]. Available: http://scikit-learn.org/stable/documentation.html. [Accessed: 27-Jul-2018].