12-2012

Power Transmission Line Fault Classification Using Support Vector Machines

Zhuokang Jia
Western Michigan University

POWER TRANSMISSION LINE FAULT CLASSIFICATION
USING SUPPORT VECTOR MACHIINES

Zhuokang Jia, M.S.E.

Western Michigan University, 2012

Electrical transmission lines are prone to faults and failures. When a fault occurs, it is impossible most of the times to fix it manually. Many methods have been adopted in the past in order to serve the purpose as fault diagnosing application. In this thesis, I discuss the method of Support Vector Machine (SVM) for fault diagnosis. SVM has the edge of good generalization over other fault diagnosing applications because it is based on pattern recognize algorithms. The aim is to classify the type of fault in the lines. Furthermore, in this work, the current and voltage of each phase are sampled, calculated and then utilized as an optimal learning pattern. Using this method, experimental simulations will show that SVM can identify each class accurately in comparison to previously used methods such as; Expert System, Artificial Neural Network, Petri Net, Fuzzy Theory. The results of simulation tests demonstrate the effectiveness of the proposed method to automatic fault diagnosis.

POWER TRANSMISSION LINE FAULT CLASSIFICATION USING SUPPORT VECTOR MACHIINES

by

Zhuokang Jia

A Thesis
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Degree of Master of Science in Engineering (Electrical)
Department of Electrical and Computer Engineering
Advisor: Dr. Ikhlas Abdel-Qader, P.E.

Western Michigan University
Kalamazoo, Michigan
December 2012

ACKNOWLEDGMENTS

I would like to take this chance for thanking my advisor Dr. Ikhlas Abdel-Qader, P.E., my family and friends for the support they provided to me while working on this thesis and during my graduate studies, and for their belief in me as well as for the guidance they provided without which, I would have never been able to do this research.

Zhuokang Jia

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: BACKGROUND AND RESEARCH OBJECTIVES

1.1 Introduction

With the rapid development of power electronics technology, fault detection and localization are the focus of research efforts in the area of transmission and distribution system. Because faults in electrical power systems cannot be avoided, enough information provided from the fault analysis is needed to recognize the cause, and interpret the broken down system. It is also needed to restore as soon as possible the transfer of power, in addition to learning more about the system and aim to reduce the occurrence of future faults if possible. Circuit breakers and other control components are designed to help protect the relay and to take appropriate action and thus minimizing power loss and length of power disruption.

In electrical power system control centers, a great amount of information and data are collected from the transmission lines. Fault signals must be detected and classified in real time and accurately in order to protect the whole system. The related signals have to be received and processed by the operators in the control center. An effective and accurate mathematics based program and process is usually in place to help the operators to detect, classify, locate, and isolate faults in the transmission system once one occurs.

The general procedure followed is based on preset threshold values according to the fault signal values of current and voltage. When a fault occurs, the transient DC offset and high-frequency transient components will be extracted along with the power frequency components from the fault current and voltage signals. The fault current and voltage vary with fault type, location, size of the fault, and the fault inception angle and other system conditions. These variations cause space to be non-linearly separable and no threshold values can be found that satisfy the various system and fault conditions (Wake & Sakaguchi, 1984). Furthermore, when faults take place, the faulted phases have an effect on the healthy phases due to mutual coupling between these phases. This problem is compounded by the fact that this coupling is highly non-linear in nature and is dependent on a complex interplay amongst a number of variables (Wake & Sakaguchi, 1984).

Digital signal processing methods have become more and more in use to analyze fault signals and to allow for the fault detection, localization and classification in power systems. Recent research includes Principal Component Analysis (Qais, 2010), Artificial Neural Network (ANN) (Cardoso & Rolim, 2004), Wavelets transform incorporated with Probabilistic Neural Network (PNN) (KocSavas & Aydogmus, 2009), and Adaptive Resonance Theory (ART) neural fuzzy system (Monsef et al., 1997) to name few. However, each has its own disadvantages and so more research is needed. Among the intelligent systems to be investigated is Support Vector Machine (SVM) which has been looked at for fault diagnosis in power transmission line systems (LeeByung et al., 2006). The SVM superiority over conventional schemes was reported and that prompted this thesis work.

As pattern classifiers, SVM has been adopted in the field of power transmission line classification and localization (Ravikumar et al., 2008). By using SVM, classification is done of the unknown fault pattern once the learning process is accomplished using training patterns. Several key purposes and applications of SVM are proposed in the literature that will help to understand the concept of adopting it as a tool for fault detection and classification in power transmission line systems.

1.2 Research Motivation

Transmission line is the most likely element in the power system to be the location and reason of faults especially when their physical dimension is taken into consideration (Qais 2010). This thesis aims to classify faults associated with transmission lines using their phase voltage and current signals once a fault has occurred using Support Vector Machine. There are four types of faults that might occur in the three phase transmission line. They include; single phase to ground fault, double phases to ground fault, phase to phase fault, and three phases fault. Therefore, in order to prevent lengthy periods of power loss and any of its negative impacts on the power system itself because of a faulted line. We need to identify as soon as possible the type of fault and the line where fault occurred.

SVM is a classifier that has a high learning capacity proved to be efficient in many applications. The learning capacity can be judged from its ability to model complex and real world issues that may include image and text classification, bioinformatics, hand-writing recognition and analysis of bio-sequence. SVM possesses the capacity to perform explicitly well on sets of data having many attributes, even there is a shortage of cases for the training of the model. SVM has faster and more accurate computational ability than traditional techniques, because of the inherent property of SVM as it requires less training data than other methods (Milenova et al., 2005). Without the need for large amount of training data and thus long training time, employing this accurate classifier in transmission line fault diagnosis does validate its usefulness and encourages engineers to use this technique in many other problems in power systems.

1.3 Research Objectives

In view of the above discussion, the objectives of this research are to utilize support vector machine as tool for fault detection and classification in power transmission line systems. Specifically, I intend to identify the type of fault in the lines and to determine which segment of the line has experienced a fault. In order to achieve these objectives, the learning patterns that are sampled from the three-phase transmission line are optimized. Here, optimization illustrates the use of technique to design a mathematical model that can best use the parameters of the transmission lines. Optimization generally uses a mathematical model for resolving any type of problem where the decision variables are problem's parameters.

However, in this work, the current and voltage of each phase are sampled, calculated and then utilized as an optimal learning pattern. Using this method, experimental simulations will show that SVM can classify each class more accurately in comparison to previously used methods such as; Expert System, Artificial Neural Network, Petri Net, Fuzzy Theory. It shows that the advantage of SVM method is due to its learning process using quadratic optimization and thus the number of operations is substantially reduced during the training process. Our results show that our SVM approach is accurate and effective for power systems fault detection and analysis.

1.4 Research Questions

RQ1: How can support vector machine be utilized as a tool for fault detection and classification in power transmission line systems?

RQ2: What is the method to identify the type of fault in the lines and to determine which segment of the line has experienced a fault?

RQ3: How can the current and voltage of each phase are sampled, calculated and then utilized as an optimal learning pattern?

RQ4: Has SVM proved to be more accurate in classification and diagnosing of faults in comparison to previously used methods such as; Expert System, Artificial Neural Network, Petri Net and Fuzzy Theory?

1.5 Pertinent Literature

Electrical Systems have so far, been improved very well, and now we are almost completely electrified except in some tribal, mountainous and remote places, where, though Human habitat is possible, Technical and Engineering advancements is still a nightmare. Assuming that it is a part of future development, we Engineers are bound to restrict ourselves to protecting the present systems intact and ensure its safe, steady, and balanced operation. Being Electrical Engineers, our objective is to ensure continuous and intermittent power supply to already electrified places as far as possible. But, just like all the other Engineering systems, Electrical Networks are also prone to fail, go out of Synchronism, suffer from a fault, thereby removing some connected loads suddenly. Faults, though very uncommon now, need to be serviced and reported or signaled to the concerned operator responsible.

Support vector machine (SVM) is an intelligent classification algorithm. It was presented by Vapnik and Cortes originally and became a novel machine for data analysis (Cortes & Vapnik, 1995). Especially, it has advantages for solving problems with small quantity of nonlinear samples even in high dimensions. It has strong ability to solve ill-posed problems as a regularized method. Nowadays, many applications have taken advantage of SVM such as faults classification in power transmission lines, wind turbines and induction motors (Huang, 2002), face recognition (Guo, Li, & Chan, 2000), time series prediction (Wang, 2007), and non linear systems modeling (Lee, Kim, Seok, & Won, 2006). SVM is found effective for power system applications due to its ability to work with existing pattern of data.

The Support Vector Machine is an effective and efficient tool in fault detection and classification in power transmission line. A wide variety of published work in the fault diagnosis of power systems has been reported in the literature. For instance, Ravikumar, Dhadbanjan, and Khincha (2007) employed support vector machine as an intelligent method for fault diagnosis in power transmission systems. Ganyun, Haozhong, Haibao and Lixin proposed a SVM classifier with multiple layers for fault diagnosis of power transformer (Ganyun, Cheng, Zhai, & Dong, 2005). Jin and Renmu presented a new algorithm of improving fault location based on SVM (Jin & Renmu). Also, Samantaray, Dash and Panda used support vector machine in fault classification and ground detection (Dash, Samantaray, & Panda, 2007).

Rongjie and Haifeng presented an application of SVM in the fault diagnosis of power electronics rectifier (Rongjie & Haifeng, 2008). It should be noted that most of the implementations in the above references are used in other power systems. The main differnece between this study and the literature of past studies is that the present study solely aims at fault classification of power transmission lines. In addition, the current and voltage of each phase are sampled, calculated and then utilized as an optimal learning pattern. More discussion about the applications used in the area of fault diagnosis is presented in Chapter 3.

1.6 Key Definitions

1.6.1 Protection Systems

A protection system must have the function to disconnect the part with fault from the system as fast as possible. The faulty system component, such as a line, a bus, or a transformer, should be separated from the whole system in a timely manner to prevent black out through the action of protective devices.

1.6.2 SVM

Support Vector Machine (SVM) is a novel intelligent machine self-learning method based on the statistical learning theory. According to this method, the original input is mapped into a higher-dimensional space in order to make it easier to find the optimal hyper-plane that can classify the input, which is a set of non-linear samples in this case.

1.6.3 SVM Regression

SVM has the ability to solve and simplify regression based problems by adopting epsilon-insensitive loss function (Janik & Lobos, 2006). SVM regression tries to find a continuous function such that the maximum number of data points lie within the epsilon-wide insensitivity tube. The predictions that fall within the distance of epsilon for the original target value are not considered as errors.

1.6.4 One-Class SVM

SVM is used as a one-class classifier by Oracle Data Mining for the purpose of anomaly detection. When using SVM as an anomaly function, there remains a function for classification mining but the target becomes absent (Janik & Lobos, 2006). In the adoption of one-class SVM model, there exists a probability and a prediction in the scoring data for every individual case.

1.6.5 Class Weights

In SVM classification, the concept of class weights signifies the biasing mechanism for the purpose of identifying the pertinent importance of the targeted values i.e. classes (Janik & Lobos, 2006). The initializing of SVM models is done in manner to get the prediction in the best possible way for all the present classes.

1.6.5 Optimization in SVM

Optimization illustrates the use of technique to design a mathematical model that can best use the parameters of the transmission lines. Optimization generally uses a mathematical model for resolving any type of problem where the decision variables are problem’s parameters.

1.7 Ethical Considerations

Any research process is required to conform to ethical, legal and moral standards (Plooy, 1997). Basically, the research should be consistent with requirements for the protection of human rights if any of the individual is contacted (Plooy, 1997). As a security measure and among one year following completion and acceptance of the research study, hard copies of transcribed notes will be shredded. At this time, soft-copies of transcriptions will be deleted from all electronic media storage and portable forms of electronic media storage containing this data such as: CDs and memory sticks will be destroyed. While retrieval of electronically deleted documents from hard drives is possible, the cost and effort required to retrieve erased data makes this possibility unlikely.

Researcher is fully aware of the ethical issues involved in this work. The responsibility of all procedures and ethical issues related to the project lies with the principal investigators. The research will be conducted so that the integrity of the research enterprise will remain and negative side effects that may decrease the potential for future research were avoided. The choice of research topics is based on the best scientific approach and an evaluation of the potential benefits. This study is related to a major intellectual problem.

1.8 Thesis Outline

Apart from the introduction, this thesis is spread over five further chapters. Chapter 2 provides a general idea about the research methodology that has been selected for this research. The methodology mainly includes; the research design, data collection method, reliability and validity for the research. In Chapter 3, the fault diagnosis in power system, issues and problems related to the protection of typical power transmission line systems, and the fault types and protection schemes available in the literature are briefly presented. The transmission line protection and relaying are also discussed in this chapter.

A brief introduction to the previous approaches and a detailed one to support vector machine is proposed. Chapter 4 presents the mathematical modeling and simulation of the proposed SVM for detection and classification of the faults as applied to a typical three-phase transmission line system. In Chapter 5, simulation results and discussion are presented. Chapter 6 provides conclusions and further work suggestions for the use of SVM in the field of transmission line fault detection.

CHAPTER 2: METHODOLOGY

2.1 Research Design

The research is based on secondary data accumulation. The data is pressed out from various journals, articles and books. Secondary research depicts information assembled by literature, broadcast media, publications, and other non-human origins. In this research, we are also employing the descriptive method of research. The research accession used is qualitative. Qualitative research is practically more immanent than quantitative one. This type of study is often less costly than quantitative studies and is exceedingly effectual in acquiring information. It is a frequently used method in instances where valued measurement is not commanded (Silverman 2001, 45-89).

2.2 Plan of Work

This study has used the sampling the voltage and current signals in the three-phase transmission lines instead of the traditional method that employs time-current sampling and time-voltage sampling. Sampling the voltage and current signals of the three-phase transmission lines will allow us to detect a fault in the system and to specify the fault type as well. In three-phase line system, we will deal with three SVMs one per phase. The inputs to each SVM for each single phase are the three-phase currents and voltages of the systems, these are;

$I = [I_a\ I_b\ I_c]^T$ And

$V = [V_a\ V_b\ V_c]^T$

The output of each SVM will be the declaration of a fault if it occurs. The results of these three SVM is combined to obtain the four signals to distinguish the four types of faults (Saadat, 2005), namely,

- Single phase to ground fault,
- Double phases to ground fault,
- Phase to phase fault, and
- Three phases fault

In order to perform the above two tasks, mathematical models of the transmission line system are constructed. Simulation of the model is performed using MATLAB and its accompanying Support Vector Machine Toolbox.

2.3 Research Method

The research method applied in this study was the secondary research involving the undertaking of study on a qualitative basis. Professional education practice is complex and a difficult social interaction. I believe, qualitative method (non-positivist approach) of analysis is better suited than the quantitative (positivist approach) in this type of the setting. Various online libraries, journals and databases were researched and reviewed in order to gain an extensive insight and comprehension of the issue being studied. Some of the examples of databases used include Google Scholar, Academia Search Premier, EBSCO, Pro-Quest and other relevant websites were explored. The major keywords searched include Curriculum change, self learning and self assessment, passive and active learners, etc.

2.4 Data Analysis Method

The literature was reviewed while considering the analysis of all the major facts and figures. An in-depth analysis of all the data collected was done and the real life examples and evidences were also taken into account. However, a literature too old to appropriate the search and the topic was excluded (to some extent) from this research and analysis. Only the latest figures which were sufficient enough to assist the exploration and illustration were utilized.

2.5 Secondary Data

The analysis of secondary data plays a vital role in many fields of study, including the social sciences. The definition of secondary versus primary data is not based on specific qualities of the data itself but on its history and relationship to a specific analysis. A simple definition is that primary data are collected by a research group for the specific analysis in question, whereas secondary data are collected by someone else for some other purpose. So if a researcher conducts a survey and analyzes the results for his or her analysis, the data from the survey are primary data.

If the researcher deposits the data in an archive and someone else unrelated to the original research team analyzes it 20 years later, then the results for that analysis the data are now secondary data (Bulmer, 2006). One reason analysis of secondary data is becoming more popular in the social sciences is the availability of large data sets collected and processed by the government and made available for researchers to analyze. It would be beyond the capability of most if not all research teams to collect data on this scale, but the data from these projects are available for anyone with a connection to the Internet to download for free.

2.6 Qualitative Research

Qualitative research can be defined as a process of inquiry that builds a complex and holistic picture of a particular phenomenon of interest by using a natural setting (Thermalite, 2005). Thus, qualitative research involves the analysis of words, pictures, videos, or objects in the context in which they occur. It is critical to understand that qualitative inquiry starts with a very different relationship to social reality than that of most quantitative researchers. Constructivists challenge positivist views about the nature of reality, how to study it, who can know, what knowledge is, and how it is produced. This influences every aspect of the research design.

2.7 Assumptions of Qualitative Research

The goal of qualitative research is to understand social issues from multiple perspectives to have a comprehensive understanding of a particular event, person, or group (Thermalite, 2005). As with quantitative research, there are several key assumptions underlying qualitative research methods:

i. Reality is socially constructed, and there are multiple realities.
ii. The researcher interacts and often works closely with the individuals or groups under study and serves as the primary instrument for data collection and analysis.

iii. The research is value laden, and the researchers become a part of the research, attempting to understand the lives and experiences of the people they study.

iv. Research is context bound and based on inductive forms of logic that emerge as a study progresses.

v. The purpose of research is to find theories that help explain a particular phenomenon.

2.8 Literature Search

The measurements of choice for literature were relevancy to research topic and year of publishing. Both public and individual libraries as well as online libraries were chaffered to approach the data. Facts were accumulated to support the study that has been conveyed. To gather proper information about the study, scholarly as well as non-scholarly articles were searched and considered. Some of online databases that were accessed are SAGE, Questia, Emerald, Pro-quest, EBSCO and so on.

2.9 Validity

Whilst reliability is known for the credibility and repeatability from the instrument, validity accounts the power of this instrument in order to correctly characterize the characteristics of a phenomenon. Validity may be measured via three points of views: “validity of the content, validity of the construct and validity that is related to criterion” (Maykut and Morehouse 2004, p.39). Validity of the content is actually concerned about trials and instruments found in the research and defines the level that helps to make certain that the phenomenon is investigated in satisfactory form.

Validity of the construct is often a task at the proximity of the implementation towards construct that has being researched. Validity that is related to criterion identifies contrasting of the method, discussions and findings from the research in opposition to an established standard. Participant verification or validation is a phenomenon that indulges reassessing the study's results along with the participants of the study.

Even though the validation of the data and the participants can be done through this method, but Long and Johnson warns that that such strategy can prove to be time consuming and a transition stage may occur between the data collection and in the analysis of the results (Maykut and Morehouse 2004, p.39). Although this strategy might be used as some sort of validation strategy, Long & Johnson cautioned we now have problems such as a time lag between your collection associated with data as well as the processing associated with results and data that that are inherent to this technique that may make this ill-suited in order to gauge this validity from the study. This method was not found in this examine. Thus, this method has not been employed in this research.

CHAPTER 3: PROTECTION OF POWER TRANSMISSION LINE SYSTEMS

3.1 Power System Faults

Power systems are prone to faults at times. When faults occur, consumer supply is disrupted thereby leading to loss of economy too (Lee et al., 2006). The objective of any Electrical Power system is to maintain continuity of supply so faults are undesirable. Though many fault identification methods exist, use of SVM is almost a novel approach in this concern. A common question arises asking what the need of a system to identify faults is, when there are various signaling systems already installed (Lee et al., 2006). It is just another system for the same purpose, which uses Learning Techniques to learn Fault characteristics and apply the learnt math to say whether a new input, Bus here, is faulted or healthy.

For all these undesirable reasons, faults should immediately be identified and isolated. Use of SVM permits better classification of faulted and healthy buses using its indigenous learning algorithm and the resulting model can be used for fault identification in other systems too and the model trained and tested to 100% performance in most cases except a few (Lee et al., 2006).

3.2 Background to Fault Diagnosis Technologies

Transmission lines are one of such power system components that hold the highest level of probability of fault incidence, since they are exposed to the environment. Line faults due to lightning, storms, vegetation fall, fog and salt spray on dirty insulators are beyond the control of man. The balanced faults in a transmission line are three phase shunt and three phases to ground circuits. Single line-to-ground, line-to-line and double line-to-ground faults are unbalanced in nature. On a transmission system, protective relaying systems are placed for the detection of unusual signals that indicates faults and works electronically to separate the faulted part from the remaining system with minimal disturbance and equipment damage. This survey attempts to cover the various developments in digital relays for transmission line protection reported in the literature up to October 2010 and point to some of the references showing promising directions.

Rockefeller first presented the implementation of digital relaying in 1969 (Rockefeller, 1969). The advances in the very large scale integrated (VLSI) technology and software techniques led to the development of microprocessor based relays that were first offered as commercial devices in 1979 (Sachdev, 1979).

Distance relaying principle, due to their high speed fault clearance compared with the over current relays is a widely used protective scheme for the protection of high and extra high voltage (EHV) transmission and sub-transmission lines. A distance relay estimates the electrical distance to the fault and compares the result with a given threshold, which determines the protection zone. The hardware has changed with the passage of time from distance relays to electromechanical relays, and now digital relays that operate through microprocessor. The first installation of digital computer for relaying began in 1960's which made it possible to store information so that the relay engineer can control the reach characteristics of a distance relay to suit the application and develop fault location algorithms (Gilcrest et al., 1972). Such digital fault locators calculate the reactance of a faulty line anticipated through the computation of current and voltage phasors on line terminals (Adu, 2004). But these fault location methods require an easy kind of hypothesis for the calculation of fault distance that affects the reliability and validity of the outcomes. The one terminal approach is simple and easy to implement although the algorithms for two-end which performs processing of signals from both terminals of the line are superior as compared to the one-end advancements (Girgis et al., 1992).

In the 70's research was concentrated on "ultra high speed protection" based on the travelling wave. The post fault wave forms in the first one or two cycles after the occurrence of a fault contain high frequency transient wave fronts. Different algorithms proposed for implementation of travelling wave distance protection are reported by Desikachar et al., (1984), Shehab-Eldin et al., (1988), Ancell et al., (1994), Ernesto Vazquez-Martinez, (2003), Dong et al., (2009). The "positional protection" utilizes the transit times of the high frequency fault generated transients to identify the faulted line section (Bo et al., 2000). It has been noted by most of the researchers that the travelling wave based method does not perform well for faults close to the relaying point and for faults with small fault inception angle, besides they require a very high sampling rate and their implementations have much higher cost in comparison to the impedance techniques implementation. As the difficulty of the power network increases, the transmission line protection and control must be based on real time power system changes and it must be at high speeds for making sure that there will be no transient stability problems in the power system. Several papers have considered the real time power system changes and have reported about accurate, fast faulty phase selection and fault location.

In the late 80's synchronized measurement technology emerged as a promising prospect in achieving real time protection. With global positioning system (GPS), digital measurement can be executed synchronously at different line terminals (Crossley et al., 1998; Bo et al., 2000). They are more accurate than distance relaying algorithms which are affected with inadequate modeling of parameter uncertainty and transmission lines due to aging of the line, line asymmetry and environmental factors. The Phasor Measurements Units (PMU) are the most commonly used devices for synchronized measurement, whose measurements are synchronized with respect to a GPS clock and PMU-based fault locators are more accurate than the method based on unsynchronized phasors. Such impedance based fault location methods have been presented by Izykowski et al., (2006), Dascastagne et al., (2008), having very little fault location error if parameters for phasor and transmission line are accurate. Although the use of GPS, phasor measurement units (PMUs), digital communication technologies, high precision signal transducers have facilitated accurate protection of power system over a wide area, they are subjected to software insecurity and communications latency.

There is a need for the measuring algorithms that posses the quality to adapt with dynamism to the system operating conditions that include changes in the system configuration, source impedances and fault resistance. Keeping this in view the trends since 90’s, intelligent techniques are under investigation to increase consistency, rate and accuracy of present digital relays that have the basis on the Expert System, Artificial Neural Network (ANN), Fuzzy Logic (FL), Fuzzy-Neuro and Fuzzy Logic-Wavelet based systems and Support Vector Machine.

3.3 Fault Diagnosis Applications

Fault Diagnosis in power system is considered a system level fault diagnosis in the dispatch center. It is to analyze the action information of various types and levels of protection devices and circuit breakers, the voltage and current measurement of electrical quantities. According to the logic and operation of the protection action experience, fault diagnosis infers possible fault type and fault location, and provides the relevant criterion for the dispatcher's decision-making. When the grid fails, accurate, rapid, automatic fault diagnosis has great significance to quick recovery of the power grid.

Because artificial intelligence techniques are good at simulating human thinking process, easily taking into account the experience and learning ability, it has been widely investigated in the transmission line fault diagnosis, such as the Expert System, Artificial Neural Network, Petri Net, Fuzzy Theory, Optimization method, and Support Vector Machine. This chapter briefly introduces the basic concepts of these methods and reviews the reported adaptations in literatures.

3.3.1 Expert System

An expert system is the most mature branch in the field of artificial intelligence. It uses computer technology to integrate the theoretical knowledge of the underlying field with the practical knowledge of the experts. Through the effective connection between databases, knowledge base, inference engine, man-machine interface, clear procedures and knowledge access program (Park, Kim & Sohn, 1997), it has the capacity to resolve complex problems in the given expertise area. The expert system has a variety of knowledge representations, such as predicate logic notation, the production rule representation, frame-based knowledge representation, semantic network representation, object-oriented technology and procedural representation.

Park, Kim and Sohn presented a logical method to deal with fault in the diagnosis phase. The relay with the relationship is represented as a logical relationship and fault reasoning is translated into Boolean algebra. Due to the lengthy process of logical reasoning, there is a need to pre-complete the major implications offline in order to achieve the targeted operation while online (Park, Kim, & Sohn, 1997).

Jadid, Jeyasurya, and Khaparde (1989) presented a typical rule-based expert system for fault diagnosis in which protection expert knowledge was translated into IF-THEN type of fault diagnosis rules using PROLOG language. During the time of diagnosis, fault information is matched with the diagnostic rules to achieve the diagnostic results. As a precondition of a completed failure information, and assuming that there is up to only one protection or circuit breaker malfunction, fault tolerance and robustness of the diagnostic system were not strong enough, and it was difficult to meet the requirements of practical implementation. It is indeed a very difficult and tedious work to establish a rule-based expert system that is complete and self-consistent for fault analysis. Huang proposed a method of man-machine interface interacted with the dispatcher to automatically create diagnostic rules. (Huang, 2002) It can greatly improve the efficiency of knowledge acquisition and transformation. Even so, rules having conflicting resolution in the rule-based expert system are still often occurring. In addition, the slow reasoning speed is a disadvantage of rule-based systems.

Wake and Sakaguchi (1984) introduced a model-based fault diagnosis expert system. This method uses logic circuit to model the structure and function of the relay protection system, simulates the fault conditions in accordance with the model, and compares the simulation results and the real alarm information in order to judge the authenticity of the hypothetical fault. The model-based expert system needs an effective algorithm that can correct the hypothetical accident set according to the results of the simulation, until it finds the correct fault set (Lee et al., 2006). To enhance their results, they used some of the ideas in the optimization algorithms. Model-based fault diagnosis method can be adapted to the loss of the fault information, but the reliability of the protection system and circuit breakers need to be addressed during the simulation.

3.3.2 Fuzzy Theory

Fuzzy theory uses the concept of fuzzy membership to describe the imprecise and uncertain objects and it adopts approximate reasoning rules to express expert knowledge effectively. The fuzzy set theory is used for fault type identification on a transmission line by Ferrero et al., (1995), Das et al., (2005), without any computationally expensive training of ANN or expert domain knowledge. These algorithms are fairly accurate only under certain assumptions of fault distance, pre-fault power flow, fault resistance and line length. When there is a requirement for the representation of uncertain knowledge, fuzzy sets serves the purpose in a better way. On the other hand, neural networks have the efficient ability to learn from examples.

In a fuzzy neural network (FNN), a neural network is used to implement a fuzzy rule-based system from input/output data to enhance the learning capabilities, plus knowledge illustration of fuzzy logic system. Wang et al., (1998), proposed three different neuro-fuzzy networks in series to classify the fault in transmission line protection using both designer's experiences and sample data sets. A distance relaying scheme based on FNN is proposed by Dash et al., (2000) in which the fuzzy view point is utilized to simply the model, but the FNN's calculate the fault distance within 80% of the line. It has also been found that Fuzzy theory has a strong fault tolerance capacity (Cho & Park, 1997).

Therefore, Fuzzy theory can be more suitable to handle the uncertainty and missing information of fault diagnosis in grid protective actions. Cho & Park pointed out the reliability of protection action and the uncertainty of protection scope in the existed relay protection system. The range line of main protection is about 85% of line length; the scope of the back-up protection extends to 120% -150% of the line length (Cho & Park, 1997). On this basis, with fuzzy directed graph, the uncertain relationship was established between the faulty components and the protection and circuit breakers. In their work they assumed that the failure information to be reliable and complete, and used approximate reasoning to achieve the extent possible of each failure component. For this, they build sagittal diagrams for the representation of fuzzy relations in the power system.

These sigittal diagrams were then used for the diagnosis of fault sections. The next study that Cho & Park (1997) did, was to verify relay logic with relations of suspicious components to identify the protection and circuit breakers with incorrect action, and find the false alarm information and circuit breakers. Differently, Monsef divided credibility into five levels (VL, L, M, H, and VH) in accordance with the classical theory of fuzzy. Fuzzy input signal and approximate reasoning of failure are based on these five fuzzy levels (Monsef, Ranjbar, & Jadid, 1997). In their work, they found that the use of prior probability information for the grid is not sufficient but the robustness was better.

3.3.3 Artificial Neural Network

Artificial neural network (ANN) has characteristics of parallel processing, nonlinear mapping, and associative memory capacity and online learning abilities. It is widely used in electrical power systems and other areas with successful results. Different modes of failure in the grid will produce different combinations of fault information. Fault diagnosis problem can be regarded as the problem of pattern recognition using ANN for processing but with a more complete set of training samples needed to be established. By using preselected accident samples as input, and fault information set as an output, the neural network (NN) is trained.

Chan presented an earlier BP (error back propagation) neural network applied to fault diagnosis in power system (Chan, 1989). But the training speed is slow, and it possibly falls into the shortcomings of local optima. In the BP neural network, the introduction of additional momentum factor can solve this problem. Also, Radial basis function (RBF) NN has the ability of arbitrary function approximation and has a faster learning speed. Bi, Ni, and Wu presented a new RBF NN to solve the problem of fault diagnosis. The simulation results of 4-bus test system show that the ability of RBF NN in grid fault diagnosis is superior to traditional BP NN (Bi, Ni, & Wu, 2002). In order to solve unsuitable problem for the common NN in case of network topology change (need to re-train), Cardoso and Rolim comprehensive utilization of the accurate capability of multilayer perception (MLP) NN.

They also utilized the characteristics of generalized regression neural network (GRNN), such as quick learning, global optimum and low requirements of comprehensive sample (Cardoso & Rolim, 2004). Using trained MLP NN to describe the behavior of the rules of relay protection, the failure information is fed first into the MLP, the reasoning results then outputted to the GRNN for the pattern classification, and eventually give the fault diagnostic conclusions, or declare that the fault information is inadequate. (Cardoso & Rolim, 2004) Compared with the expert system diagnostic methods, the neural network fault diagnosis methods can avoid the formation of expertise and expert heuristic knowledge and expression and thus saving a real amount of tedious work. ANNs can solve the overreach and the under reach problems which are very common in the conventional distance relay design.

ANN utilizes samples of currents and voltages directly as inputs without computation of phasors and related symmetrical components. Various kinds of neural network such as multi-layer perception (MLP), recurrent, radial basis function (RBF), probabilistic neural network etc. are being applied for fault classification and fault location. Different algorithms have been used for the designing purpose such as; orthogonal least square, back propagation and extended Kalman filter. For selecting the appropriate network configurations, the performance criteria are fault tolerance, minimal response time and generalization capabilities. The approach of ANN has been positively utilized for the improvement of many of the standard functions that are operated in transmission lines. They have been related to fault direction discrimination (Sidhu et al., 2004), fault detection and classification, distance protection, improvements in fault distance computation, protection of series compensated lines, adaptive distance protection and adaptive reclosing (Aggarwal et al., 1994).

To make the ANN responsive to time varying voltage and current waveforms different types of recurrent networks were considered that allow the hidden units of the network to see their own previous output, so that the subsequent behavior can be shaped by previous response. Inside these ANNs the operations that take place are not clearly defined and hence they are not considered highly reliable. Further development is the concept of supervised clustering to reduce the number of iterations in the learning process of multi layer feed forward networks (Kezunovic et al., 1995). A neural network simulator is developed by Venkatesan et al., (2001), to identify the optimum ANN structure required for training the data and to implement the ANN in hardware.

Still the problem with ANN's is that no exact rule exists for the choice to the number of hidden layers and neurons per hidden layer. So it is uncertain whether the ANN based relay gives the optimum output, to maintain the integrity of the boundaries of the relay characteristics. A high speed distance relaying scheme based on radial basis function neural network (RBFNN) is proposed by Pradhan et al., (2001), due to its ability to distinguish faults with data falling outside the training pattern. The use of separate ANNs, for faults involving earth and not involving earth has proved to be convenient way of classification of transmission faults based on RBF neural networks by Mahanty et al., (2004).

The existing ANN based solutions easily converge on local minima whenever input patterns with large dimensionality are present and when designed for specific applications, are prohibitively expensive or infeasible for real time implementations. It is observed that the ANN based distance relays need much larger training sets and hence the training of these networks is time consuming and further research work shall produce a hardware realization with proper modification in the learning methodology and preprocessing of input data that would improve the learning rate performance, efficiency and the reliability many folds. Presently research efforts are in the direction of evolutionary computational techniques such as genetic algorithms (GA) for determining the neural network weights and thereby avoid training of ANN.

3.3.4 Wavelet Approach

The fundamental frequency components of the post fault voltages and currents need to be extracted as quickly and accurately as possible for the quick response of a digital distance relay. Wavelet approach is one of the new tools in this direction which is useful for power system transient analysis, since the conventional signal processing techniques have the inherent disadvantages of long discrimination time, errors in impedance calculations and misclassifications (during CT saturation and in presence of fault resistance) (Liang et al., 2004; Bhowmik et al., 2009; Magnago et al., 1998). Wavelet transform (WT) has the ability to perform local analysis of relaying signals without losing the time-frequency information.

WT in conjunction with AI/Fuzzy/Expert system/SVM based techniques have the advantage of fast response and increased accuracy in fault type and location identification. A preprocessing module based on discrete wavelet transform (DWTs) considerably simplifies the input signal, reducing the volume of input data fed into an

ANN that classifies the fault events (Martin et al., 2003). A solution for protection of parallel transmission lines by decomposing fault current signals using WT and by comparing the magnitudes of line currents in the corresponding phases is presented by Osman et al., (2004).

The ability of wavelets to decompose the signal into different frequency bands using multi resolution analysis (MRA) allows detecting and classifying faults as well as extracting the voltage and current fundamental phasors needed to calculate the impedance to the fault point in distance protection (Osman et al., 2004) and with filtering algorithms proposed by Kleber M.Silva et al., (2010), fast relay operating times are obtained. Discrete wavelet transform based MRA is used for feature extraction by Samantaray et al., (2007) and the features extracted from fault current signals are used to train and test the support vector machine (SVM) for fault classification and the fault location from the relaying point is computed by RBFNN. A solution to the complexities of the protection of series compensated transmission lines is proposed by Parikh et al., (2008), which is a combination of wavelet-SVM technique for fault zone identification.

In a Fuzzy-logic-wavelet based technique the wavelet transform of current signal provides hidden information of a fault situation to FLS, to classify the fault and these are reported in (Youssef et al., 2004; Reddy et al., 2006; Reddy et al., 2007; Pradhan et al., 2004). These fuzzy procedures solve the problem with simple computational procedures rather than using more complex algorithms in the deterministic way. Some more improved solutions to detect the faults precisely with wavelet transform based digital protection for transmission lines are proposed by Zhang et al., (2007), Valsan et al., (2009), Da silva et al., (2010).

Wavelet singular entropy (WSE) technique which indicates the uncertainty of the energy distribution in the time-frequency domain is used to extract features from fault transients for the fault diagnosis in EHV transmission lines (Zhengyou He et al., 2010).The capabilities of wavelets are affected owing to the existence of noises riding high on the signal and the problem lies in identification of the most suitable wavelet family that is more approximate for use in estimating fault location. Most of the wavelet based techniques employ multi-level wavelet decomposition, which requires multi-level filtering followed by complex computations. Wavelet transform will emerge as a powerful tool in transmission line protection provided further work is done in reducing the algorithm complexity, computational burden and response time.

3.3.5 Support Vector Machine

Support Vector Machine (SVM) is a novel intelligent machine self-learning method based on the statistical learning theory. According to this method, the original input is mapped into a higher-dimensional space in order to make it easier to find the optimal hyper-plane that can classify the input, which is a set of non-linear samples in this case. Based on the statistical learning theory, the optimal hyper-plane is found by adopting the optimization theory (Welling, 2004). Dash and others (2007) introduced a new approach for fault classification and section identification of the thyristor-controlled series compensator (TCSC) using a support vector machine (SVM) (Dash, Samantaray, & Panda, 2007).

SVM has capacity to optimize the classification in non-linear plane between two classes, even when the different classes' samples are very close to each other. In addition, SVM regression has strong ability for multiclass classification. Electrical Systems have so far, been improved very well, and now we are almost completely electrified except in some tribal, mountainous and remote places, where, though Human habitat is possible, Technical and Engineering advancements is still a nightmare. Assuming that it is a part of future development, we Engineers are bound to restrict ourselves to protecting the present systems intact and ensure its safe, steady, and balanced operation. Being Electrical Engineers, our objective is to ensure continuous and intermittent power supply to already electrified places as far as possible. But, just like all the other Engineering systems, Electrical Networks are also prone to fail, go out of Synchronism, suffer from a fault, thereby removing some connected loads suddenly.

Faults, though very uncommon now, need to be serviced as soon as possible, and resume supply. For this, faults need to be identified, Fault Identification and Location in Distribution Systems using Support Vector Machines and reported or signaled to the concerned operator responsible. Many applications have taken advantage of SVM such as faults classification in power transmission lines, wind turbines and induction motors (Huang, 2002), face recognition (Guo, Li, & Chan, 2000), time series prediction (Wang, 2007), and non linear systems modeling (Lee, Kim, Seok, & Won, 2006). SVM is found effective for power system applications due to its ability to work with existing pattern of data. The Support Vector Machine is an effective and efficient tool in fault detection and classification in power transmission line.

A wide variety of published work in the fault diagnosis of power systems has been reported in the literature. For instance, Ravikumar, Dhadbanjan, and Khincha (2007) employed support vector machine as an intelligent method for fault diagnosis in power transmission systems. Ganyun, Haozhong, Haibao and Lixin proposed a SVM classifier with multiple layers for fault diagnosis of power transformer (Ganyun, Cheng, Zhai, & Dong, 2005). Jin and Renmu presented a new algorithm of improving fault location based on SVM (Jin & Renmu). Also, Samantaray, Dash and Panda used support vector machine in fault classification and ground detection (Dash, Samantaray, & Panda, 2007). Rongjie and Haifeng presented an application of SVM in the fault diagnosis of power electronics rectifier (Rongjie & Haifeng, 2008).

3.3.5.1 Basic Operation of SVM

Starting point for building a support vector machine is a set of training objects for which each is known, what class they belong. Each object is represented by a vector in a space represented (Janik & Lobos, 2006). The object of the support vector machine is in this space a hyper plane fit, which acts as a release surface and divides the training objects into two classes (Ravikumar et al., 2007). The distance between those vectors that are closest to the hyper plane is maximized thereby. This wide, empty edge will later ensure that objects that do not correspond exactly to the training objects are classified as reliably as possible. When inserting the hyper plane, it is not necessary to consider all training vectors (Janik & Lobos, 2006).

Vectors, the further away from the hyper plane and "hidden" behind a front sense other vectors are, do not affect the location and position of the separating plane. The hyper plane depends only on the closest vectors and only these are needed to describe the level of mathematical accuracy (Ravikumar et al., 2007). These vectors are closest to their function support vectors and helped the Support Vector Machines to their name. A hyper-plane cannot be "twisted", so that a clean separation of a hyper plane is possible only if the objects are linearly separable. This is in real applications, in general not the case. Support Vector Machines in the case of using non-linearly separable data to the kernel trick to feed a class of nonlinear boundary.

The idea behind the kernel trick is to convert the vector space, and thus the training vectors therein in a higher-dimensional space. In this higher-dimensional space now separating hyper plane is determined. In the transformation back to the lower dimensional space, the hyper plane is linear to a non-linear, may not even contiguous hyper surface, which separates the clean training vectors into two classes. In this process, two problems present (Ravikumar et al., 2007). The transformation is extremely high computational-heavy and the presentation of the division in the low dimensional space is generally unlikely complex and thus practically useless. At this point, the kernel trick is handy.

One uses to describe the interface suitable kernel functions describing the hyper plane in the high-dimensional and still remain in the low-dimensional "benign", so it is possible to implement the return transformation, without them actually have to execute computationally. Here also satisfies one of the vectors, namely the support vectors in turn to describe the class boundary completely. Both linear and non-linear support vector machines can include additional slack variables make it more flexible. The slack variables allow the classifier to classify individual objects wrong to "punish" but at the same time any such erroneous classification. In this way, on the one hand is over fitting avoided, and secondly, to reduce the required number of support vectors.

3.3.5.2 Benefits of SVM as a Classification Technique

All classification techniques have got benefits and drawbacks that happen to be pretty much essential according to the information that happens to be reviewed, and therefore have a general importance (Ravikumar et al., 2007). SVM generally is a useful application regarding insolvency analysis and research, when it comes to non-regularity from the information, by way of example once the information is not regularly sent out or produce a mysterious submitting (Zhang et al., 2009). It will also help to evaluate facts, i.e. financial ratios that ought to possibly be transformed just before going into the rating regarding time-honored distinction tactics. The advantages of the SVM process might be summarized as follows:

1. Through presenting the kernel, SVMs acquire mobility from the choice of the shape on the threshold separating solvent coming from financially troubled corporations, which in turn desires not possibly be linear and also desires not have access to identical well-designed type for those info, since it's operate is non-parametric along with performs in your area. For that reason they are able to use fiscal quotients, which in turn show the non-monotone regards to the rating and the probability regarding default, or that happen to be non-linearly dependent and this also without wanting any distinct focus on each non-monotone variable.
2. Because the kernel implicitly possesses a non-linear alteration, zero assumptions regarding the well-designed way of the alteration, which are info linearly separable, is essential. The particular alteration occurs implicitly using a sturdy theoretical groundwork along with human being skills judgment before man expertise ruling in advance is not necessary.

3. SVM gives a beneficial out-of-sample generalization, when the guidelines C along with r (in the situation of the Gaussian kernel) tend to be adequately selected. Because of this, by picking out the right generalization level, SVMs might be sturdy, even if working out sample possesses a few prejudices. Through picking out diverse r values regarding diverse input values, it gets possible to rescale outliers.

4. SVMs produce an exceptional remedy, since the optimality problem is convex. That is an advantage in comparison to Neural Networks, which have many answers related to local minima along with because of this is probably not sturdy above diverse samples.

5. Using the choice of the right kernel, including the Gaussian kernel, anybody can place more stress around the likeness involving corporations, for the reason that more similar the financial structure regarding a pair of corporations is, the higher could be the benefit on the kernel.

3.4 Protection Systems

A protection system must have the function to disconnect the part with fault from the system as fast as possible. The faulty system component, such as a line, a bus, or a transformer, should be separated from the whole system in a timely manner to prevent black out through the action of protective devices (Zhang et al., 2009). Transmission lines are one of such power system components that hold the highest level of probability of fault incidence, since they are exposed to the environment. Line faults due to lightning, storms, vegetation fall, fog and salt spray on dirty insulators are beyond the control of man. The balanced faults in a transmission line are three phase shunt and three phases to ground circuits. Single line-to-ground, line-to-line and double line-to-ground faults are unbalanced in nature.

On a transmission system, protective relaying systems are placed for the detection of unusual signals that indicates faults and works electronically to separate the faulted part from the remaining system with minimal disturbance and equipment damage. This survey attempts to cover the various developments in digital relays for transmission line protection reported in the literature up to October 2008 (Zhang et al., 2009) and point to some of the references showing promising directions. The protection system includes many other subsystems such as the relays, transducers (voltage transformers (CVT) and current transformers (CT)), and the circuit breakers. The function of the relay is to provide contacts between the battery and the trip coil closely. The circuit breaker is controlled through the triggering of its trip coil from the station battery which interrupts current and isolates the flow.

The relay is the most important component of the protection system. The response of relay depends on the status of its inputs namely the current and voltage values. Once the status of its input is a fault, it generates a corresponding output to breakers. Reliability of the relay is the degree of its certainty to perform as intended (Zhang et al., 2009). A relay is designed such that it produces a trip output only for faults it is responsible for and not for faults generated by other equipments or systems which have no relation with the relay. There two ways, in which relays can be unreliable, the first is they may fail to operate when they are expected to and the second is they may operate when they are not supposed to such as when it responds to disturbance signal in power system.

3.5 Classification of Relays

Generally, there are several relays that are used in power systems (Gonen, Mordern Power System Analysis, 1988) as follows:

- Magnitude relays, which respond to the magnitude of the input quantity.
- Directional relays, which respond to the phase angel between two inputs.
- Ratio relays, which respond to the ratio of two input signals expressed as phasors, such as distance relays.
- Differential relays, which respond to the magnitude of the algebraic sum of two or more inputs.
- Pilot relays, which employ communicated information from remote locations as input signals.

3.6 Zones of Protection

Protection zone is a part of the power system. The responsibility of a protection system is formalizing by assigning zones of protection to various protection systems. A one-line diagram of a portion of a power system containing a transformer and four buses, illustrates the concept of zone protection.

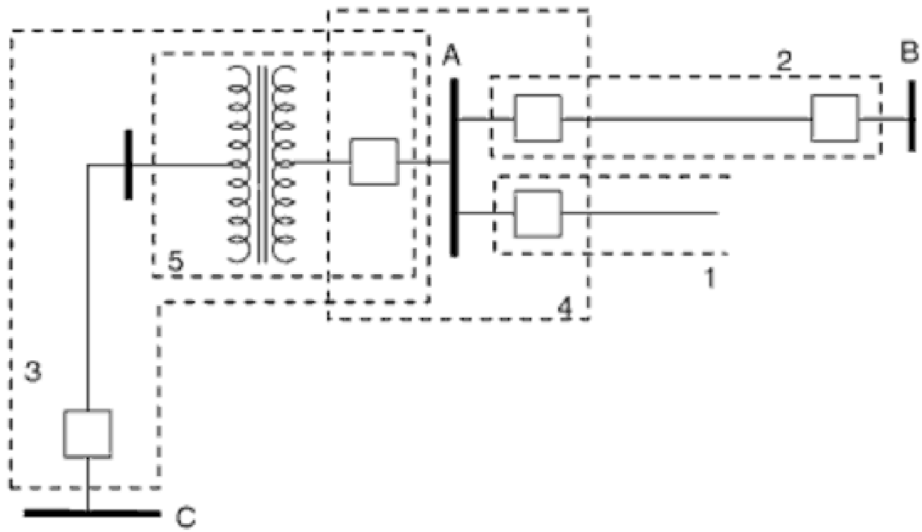

Figure 1. Zones of Protection (Stevenson, 1982)

Zone 2 defines the boundary for protection of transmission line A-B. Zone 4 defines bus-A protection. Zone 5 defines transformer protection (Stevenson, 1982). The closed dashed lines show that five zones of protection by which the various power system components are covered. Each zone contains one or more of the system components in addition to two or more circuit breakers. The boundary of each zone defines a part of the power system when a fault occurs in any part of that zone, and the protection system would take action to isolate everything within that zone from the rest of the system. The relays in that zone activate trip coils of certain circuit breakers in order to isolate the faulty zone. A circuit breaker should be inserted at each point where the equipment inside a zone and the rest of the system, for example, the circuit breaker defines the boundaries of the zones of protection.

3.7 Protection of Transmission Lines

Transmission line faults occur more frequently among the faults of any power system. About two thirds of the faults in power systems occur in the transmission line network (Sanaye-Pasand & Malik, 2001). Faults in transmission line can be:

1. Signal phase-to-ground faults
2. Double phase-to-ground faults
3. Phase-to-phase faults
4. Three phase fault

3.7.1 Signal Phase-To-Ground Faults

A single line to ground fault is a type of fault that occurs between a single phase A and the ground of the power system.

3.7.2 Double Phase-To-Ground Faults

A double phase to ground fault is a type of fault that occurs between two phases and the ground of the power system with the exception of the ground and the remaining third phase. The two participant phases in this fault include Phase B and Phase C.

3.7.3 Phase-To-Phase Faults

A phase to phase fault is a type of fault that occurs between two phases with the exception of the ground and the remaining third phase. The two participant phases in this fault include Phase B and Phase C of the power system.

3.7.4 Three Phase Fault

A 3 phase to ground fault is a type of fault that occurs between three phases and the ground of the power system. This means that all the four lines are involved in this type of fault. The two participant phases in this fault include Phase B and Phase C.

Three-phase faults comprise 5% of faults while 70%-80% of faults are single-line-to-ground faults. Traditional protective relaying methods are based on preset values of the parameters and conditions of the system under normal conditions. However, parameters values and conditions vary widely when a fault occurs. This influences the relays that are set to perform well only during predetermined fault conditions.

CHAPTER 4: SUPPORT VECTOR MACHINE MODELING

Based on the statistical learning theory, Support Vector Machine (SVM), as a new machine learning method, was developed by Vapnik and Corinna in 1995. SVM is designed to find the best compromise between the ability of recognizing random samples without mistake and the accuracy of learning the feature samples. It implements the minimum empirical risk and confidence interval, and thus obtains better generalization ability and accuracy (Welling, 2004).

4.1 SVM for Classification

SVM is comparatively a new method that is based on computational learning. It is actually designed on the basis of statistical learning theory. In SVM, the input space is made into a dot product of high dimension that is called a feature space. The optimal hyper-plane is considered to make the generalization ability higher of the classifier in the feature space. The optimal hyper-plane is determined by exploring the theory of optimization, and by respecting the critical information provided by the statistical learning theory. SVM has got the ability to counter a large featured space. This is because SVM passes through such training which maintains the fact that the classified vectors' dimension does not create a distinctive impact on the performance of conventional classifiers. Because of this feature, it is considered as highly efficient in extended classification problems.

This has also got the ability to assist in fault classification as there is no restriction on the number of features in SVM that are required as the basis for fault diagnosis.

Furthermore, the classifiers, that are SVM based, possess better generalization properties in comparison to the conventional classifiers. The difference has generated due to the fact that, the training of the SVM classifier aims at minimizing the risk of structural misclassification. On the other hand, traditional classifiers are trained for the minimization of the empirical risk. SVM method is actually matched up with the radial basis function (RBF) neural network in the task for industrial fault classification. SVM may possess some problems with wide sets of data, but it is not even available in the progress of fault classification routines.

4.2 Classification of SVM

The classification of SVM relies on the concept of decision planes that is associated with decision boundaries (Janik & Lobos, 2006). A decision plane is actually a concept that gets separated between the given set of elements possessing variable class memberships. The support vectors in the SVM are the separators that provide the widest and the broadest class separation. Both binary and multiclass targets are supported by SVM classification.

4.2.1 Class Weights

In SVM classification, the concept of class weights signifies the biasing mechanism for the purpose of identifying the pertinent importance of the targeted values i.e. classes (Janik & Lobos, 2006). The initializing of SVM models is done in manner to get the prediction in the best possible way for all the present classes. Though, the model can be made biased for the purpose of compensation if the data of training does not signify a realistic distribution for class values that are under the representation. The increase in the weight of a class results in the increase in the percentage of correct prediction, which means both are directly related to each other.

4.2.2 One-Class SVM

SVM is used as a one-class classifier by Oracle Data Mining for the purpose of anomaly detection. When using SVM as an anomaly function, there remains a function for classification mining but the target becomes absent (Janik & Lobos, 2006). In the adoption of one-class SVM model, there exists a probability and a prediction in the scoring data for every individual case. The case has to be considered as typical if the prediction is 1. While, the case has to be considered as anomalous if the prediction is 0. The behavior shows that the model is completed with normal data.

4.2.3 SVM Regression

SVM has the ability to solve and simplify regression based problems by adopting epsilon-insensitive loss function (Janik & Lobos, 2006). SVM regression tries to find a continuous function such that the maximum number of data points lie within the epsilon-wide insensitivity tube. The predictions that fall within the distance of epsilon for the original target value are not considered as errors.

4.3 Linear Classification Machine

Linear classifier is a simple and effective classifier and it describes the core concept of support vector machines. To clarify the concept, an example of classification with only two types of samples in a two-dimensional space is shown in Figure 3.1.

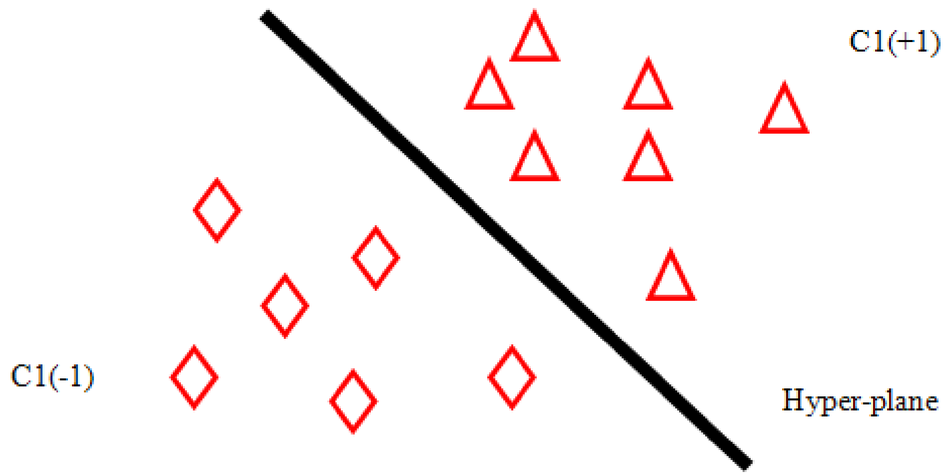

Figure 2. Linear Classifier 1 (Welling, 2004)

In this figure, C1 and C2 are two categories that need to be distinguished in the two-dimensional plane. The line between the two classes, Hyper Plane, is a classification function that can completely separate the two classes of samples. Generally, if two classes of samples can be separated completely and correctly by a linear function, then it is linearly separable, otherwise the data is non-linearly separable and some modifications are needed to allow for the separation.

Suppose the training samples are(x_i, y_i), $i = \{1, \dots n\}$, $x \in R^d$ which are input training vectors and $y \in \{+1, -1\}$ which are class symbols. In the case of linear separable, there will be a hyper-plane $f(x)$ that can make a complete separation of two classes of samples, such that all samples with $y_i = +1$ fall on the right top side and there is

$f(x) > 0$, on the other hand samples with $y_i = -1$ fall on the other side and there is $f(x) < 0$.

Here;

- y_i is either 1 or −1, that represents the class that the x_i point belongs
- x_i is a real vector of p-dimension
- n is the dimensionality of vector space
- $x \in R^d$ represents any data vector of n-dimension in which every sample is associated with any of the two classes under the label $y \in \{+1, -1\}$.

According to the method of $y_{test} = sign(x_{test})$, we could determine which class the test sample belongs to. Because both w and b are vectors, equation 1 shown below represents a plane but not a single line.

$$f(x) = w^T \cdot x + b = 0 \qquad \text{……………………1}$$

The interval between a sample point and the hyper-plane is defined as shown below：

$$\delta_i = y_i(w^T \cdot x_i + b) \qquad 2$$

4.4 Optimal Hyper-Plane

Generally there are thousands of different such hyper-planes obtained by tiny perturbations of such a given solution. Therefore, how to find optimal one among these planes is the challenge. For instance, if the hyper-plane is close enough to samples of one particular class with $y_i = +1$, obviously, when we test the cases, the error on cases with $y_i = -1$ will be minimum. Oppositely, there will be much more misclassifications on cases that should be classified with $y_i = +1$. The solution of choosing optimal hyper-plane can be viewed as solving the quadratic programming problem. The interval between a sample points and the hyper-plane is defined as:

$$\delta_i = y_i(w^T \cdot x_i + b) \qquad 3$$

If some particular samples belong to the class with $y_i = +1$, then $w^T \cdot x_i + b > 0$ and $y_i > 0$. Otherwise, $w^T \cdot x_i + b < 0$ and $y_i < 0$. This means the value of $y_i(w^T \cdot x_i + b)$ is always positive, and it is intuitively equal to $|f(x_i)| = |w^T \cdot x_i + b|$.

We make normalization to w and b using $\frac{w}{\|w\|}$ and $\frac{b}{\|b\|}$ instead of the original w and b respectively. Therefore, the equation of interval can be represented as below:

$$\delta_i = \frac{1}{\|w\|}|f(x_i)| \qquad 4$$

Geometrically, the equation is just similar with the distance from a single point x_i to a line. But in here, it represents the interval from a vector point to the hyper-plane $f(x_i) = 0$. Two more hyper-plane parallel to the separating plane are defined to cut through the closet training samples on each side. These two hyper-plane are called Support Hyper-Planes, because the sample-vectors the contain support the plane and these vectors are called Support Vectors (Welling, 2004).

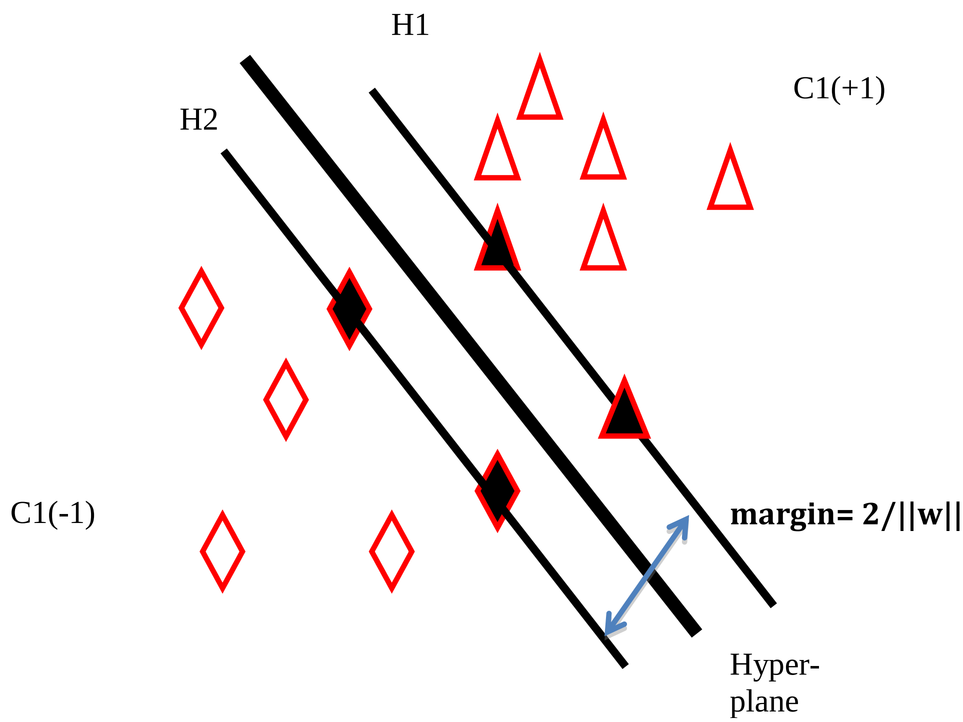

Figure 3. Optimal Hyper-Plane (Welling, 2004)

In Figure 3.2, H is the separating plane, H1 and H2 are parallel with H and cut through the vectors of each class that are most closed to H. The distance between H1 and H, H2 and H are the geometry interval. The geometry interval is related with the mistake rate of classification.

$$Mistake\ Rate\ \leq \left(\frac{2R}{\delta}\right)^2 \quad 5$$

Respectively, δ is the interval between set of samples and classification hyper-plane, $R = max\|x_i\|$, that R is the longest length value among all vectors of samples. Obviously, the upper bound of the mistake rate is determined by the geometric interval. The solution with bigger geometry interval has smaller upper bound of the error. Therefore, maximize geometry interval becomes the goal of training stage.

4.5 The Solution of Linear Classifier

Geometry interval:

$$\delta_i = \frac{1}{\|w\|}|f(x_i)| \qquad 6$$

Pay attention to the geometry interval | is inversely proportional to $\|w\|$. Hence, maximize geometry interval and minimize $\|w\|$ are completely same thing. And the method that we commonly is not seeking the largest geometry interval with fixed $\|w\|$, but looking for the smallest $\|w\|$ with the fixed interval (such as fixed for 1).

Those methods of seeking a function in the minimum value (or maximum) can be called optimization problem (also called a programming problem). For example, when we are looking for the smallest $\|w\|$, it can be expressed as below:

$$min\ \|w\| \qquad 7$$

For a giving training sample, find the optimal value of weights w and offset b, in order to minimum cost function of weights.

$$min\emptyset(x) = \frac{1}{2}\|w\|^2 \qquad 8$$

Simply, when $\|w\| = 0$, the objective function gets the minimum value. But obviously, the solutions keep same whatever the data are. From figure 3.2, that is the distance between H1 and H2 is infinite. In this case, all the sample points included both positive and negative samples are dropped into the area between H1 and H2. But our original intent is that, the samples on the right side of the H1 are divided into positive class, the samples on the left side of the H2 are divided into negative class, and the samples located between H1 and H2 are declined to classify.

The reason caused the results is only considering the target when describing the problem, without adding constraint conditions. Constraint conditions are the sample points must be in a side of the H1 or H2 (or at least on H1 and H2), not in the middle between H1 and H2. Because the geometry interval was fixed as 1, it means the smallest interval among all the sample points is 1.

$$y_i(w \cdot x_i + b) \geq 1 \qquad i = \{1 \dots l\} \qquad 9$$

$$y_i(w \cdot x_i + b) \geq 0 \qquad i = \{1 \dots l\} \qquad 10$$

Hence, our classification is also transformed into its form of mathematics, the minimum value problem with constraint:

$$\min \frac{1}{2} \|w\|^2 \qquad 11$$

$$subject\ to \quad y_i(w \cdot x_i + b) - 1 \geq 0 \quad i = \{1 \dots l\} \qquad 12$$

4.6 The Solution of Non-Linear Classifier - Kernel Functions

Here is a well-known simple non-linear classifier example in 2-D:

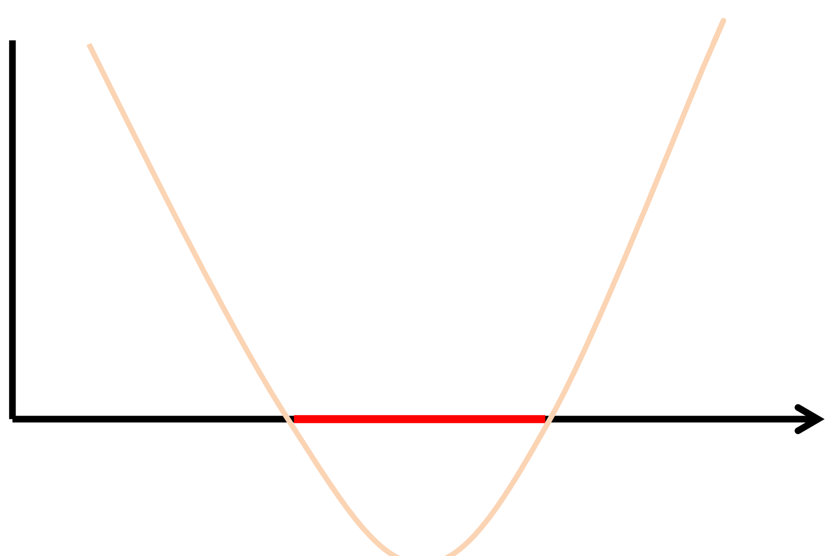

Figure 4. Graph for Non-Linear Classifier 1

As shown in Figure 3.3, the points in red on the horizontal axis belong to a positive class, on the other hand, the points on both sides of the black part are negative. Obviously, there is not a linear function that can classify these two. But we can find a curve function to classify them. Through the location of point related to the curve, the type of points can be determined. And this curve function is obviously a quadratic curve. Its function expression can be written as below:

$$g(x) = c_0 + c_1x + c_2x^2 \qquad 13$$

This function is a non-linear function. Creating two new vectors y and a:

$$y = \begin{matrix} y_1 \\ y_2 \\ y_3 \end{matrix} = \begin{matrix} 1 \\ x \\ x^2 \end{matrix}, \qquad a = \begin{matrix} a_1 \\ a_2 \\ a_3 \end{matrix} = \begin{matrix} c_0 \\ c_1 \\ c_2 \end{matrix} \qquad 14$$

In this way, $g(x)$ can be transformed into $f(y) =< a, y >$. In any dimensions of space, this kind form of function is a linear function. Only a and y are multidimensional vectors. The original problem in 2-D space in a linear inseparable problem, but after mapping to 4-D space, it becomes a linear separable problem. Hence, the core principle is transforming to high dimensional space in order to make it linear separable.

The most critical part of transformation is that looking for the method of x to y mapping. Unfortunately, there is not such a systematic way to find this mapping. But there is a kind of function $K(w, x)$ called kernel function. It accepts input value in low dimensional space, but can work out an inner product value $< w', x' >$ in high dimensional space.

When an input x in low dimensional space is given, there is:

$$g(x) = K(w, x) + b \quad 15$$

$$f(x) =< w', x' > +\text{b} \quad 16$$

These two functions present the same results. Thus we can use the input x in low dimensional directly into the function $g(x)$. As long as meeting the conditions of Mercer, the function can be used as kernel function.

Recently, there are three commonly used kernel functions:

D-order polynomial:

$$K(x_i \cdot x) = [(x_i \cdot x) + 1]^d \quad 17$$

Radial Basis Function (RBF):

$$K(x_i \cdot x) = exp\{-v\|x - x_i\|^2\} \quad 18$$

Sigmoid Function:

$$K(x_i \cdot x) = tanh[-v(x_i \cdot x) + c] \quad 19$$

Considering the linear classifier we talked about in last section, its form should be as shown below:

$$f(x') = \sum_{i=1}^{n} a_i y_i < x_i', x > +b \quad 20$$

Now this is the linear function in high dimension space. We can also use a function in low dimensional space instead.

$$g(x) = \sum_{i=1}^{n} a_i y_i K(x_i, x) + b \quad 21$$

From these two equations above, we can find that the values of a, y, b of these two equations are exactly same. Although the given problem is non-linear, we can use the same way to solve it using the kernel function.

By solving the quadratic programming using Lagrange multipliers method, it obtains the non-zero optimal α^*, the relative samples with α^* is defined as support vector. When SVM tests the input test samples, the class of x is determined by the symbol of the optimal decision function.

$$f(x) = sgn(\sum_{i \in SV} a_i^* y_i (x_i \cdot x) + b^*) \qquad 22$$

When the samples are non-linear, the sample x can be mapped to a high dimensional feature space H, and using linear classification machine in H.

$$\emptyset: R^d \rightarrow H \qquad 23$$

$$x \rightarrow \emptyset(x) = (\emptyset_1(x), \emptyset_2(x), \dots, \emptyset_i(x), \dots)^T \qquad 24$$

Where, $\emptyset_i(x)$ is a real function. According to Mercer condition, the inner product function adopted in optimal classification plane $K(x_i, x_j)$ is called Kernel function. The corresponding optimal decision function becomes as shown below:

$$f(x) = sgn(a_i^* y_i K(x_i \cdot x) + b^*) \qquad 25$$

4.7 Slack Variables

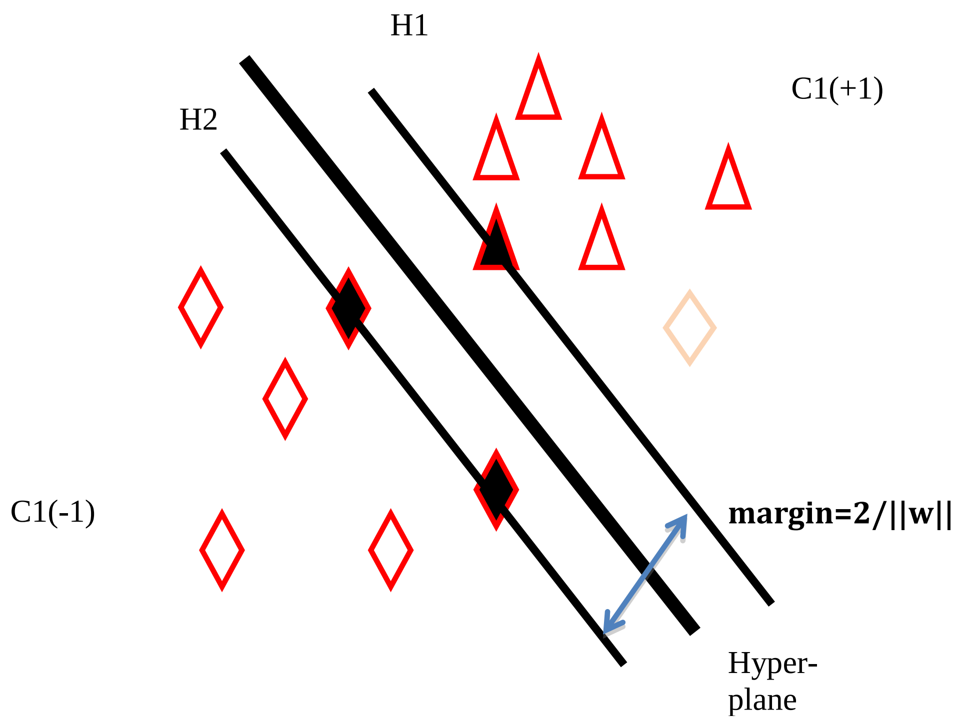

Figure 5. Non-Linear Classifier 2 (Welling, 2004)

Now we have transformed a non-linear inseparable classification problem into a linear one through the mapping to high dimension space as shown in Figure. There are thousands of diamond and triangle points in all the samples. Assume that there is another set of samples take part in. After mapping to the high dimension space, we can find additional point in the samples as shown in the Figure. It is a diamond one but it is also a negative one dropped in the positive area. This a single sample point, make originally the problem of linear separable into a linear inseparable. It is possible that this sample point is an error, a kind of noise. We could ignore this noise and adopt the original classifier. But the program of original classifier cannot ignore this point. The process of classification will consider all the samples including this noise. As a result, the original classifier cannot get any solutions due to this single noise.

In order to solve this problem, imitating the thinking of people, allowing the distance between some points and the hyper-plane cannot meet the original requirements of the classification.

The original requirements are:

$$y_i = [wx_i + b] \geq 1 \quad i = \{1 \ldots . l\} \qquad 26$$

If we want to bring in fault tolerance, adding a relaxation variables to 1 this hard that allowed:

$$y_i = [wx_i + b] \geq 1 - \xi_i \quad i = \{1 \ldots . l\} \qquad 27$$

$$\xi_i \geq 0$$

The relaxation variable is negative, therefore the final result is the requirement of interval can be smaller than 1. But when some points appear in this kind of situation (these points also called outliers), it means that we gave up the precise classification for these points and this is also a kind of loss for our classifier. But giving up these points also has brought benefits, it is to make the classification hyper-plane don't have to move face these points, thus geometry interval can be larger. In the low dimensional space view, classification plane has much more smooth boundary.

When the training sample is non linear, it needs to use non-negative slack variables. The problem of optimal hyper-plane is transferred to as shown below:

$$min\emptyset(x) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}\xi_i \qquad 28$$

$$y_i(w \cdot x_i + b) \geq 1 - \xi_i \quad i = \{1 \ldots l\} \qquad 29$$

$$\xi_i \geq 0 \quad i = \{1 \ldots l\} \qquad 30$$

Where, C is penalty parameter, the greater C means that the punishment of error classification is the greater.

4.8 Modeling and Framework for Fault Analysis

4.8.1 Symmetrical Components

Through the technique of symmetrical components, the unbalanced systems could be solved using balanced methods (Fortescue, 1918). This theory presents that an unbalanced system can be transformed into a balanced system in symmetrical components. According to Fortescue's theory, the balanced sets of symmetrical components can be generated from three unbalanced vectors of a three phase system, as shown below (Yu & Chen, 1995):

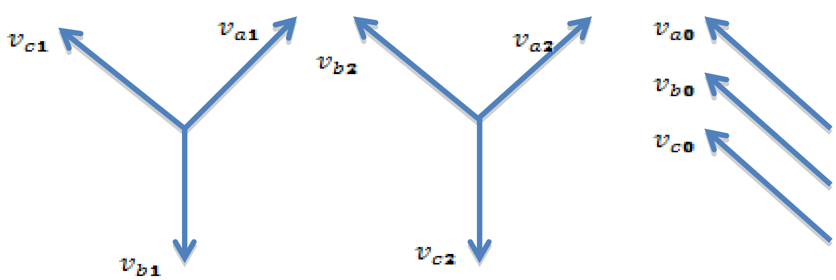

Figure 6. Symmetrical Components 1 (Koc & Aydogmus, 2009)

1. Each component of the three positive sequences is equal in magnitude, rotated 120 degrees from each other as shown in Figure 4.1. The phase sequence is same as the original ones.
2. Each component of the three negative sequences is equal in magnitude, rotated 120 degrees from each other as shown in Figure 4.1 as well. But the phase sequence is opposite to the original ones.

3. Each component of the three zero sequences is equal in magnitude with no rotation as shown in Figure 4.1

V_a, V_b, and V_c represent the original three phase voltages and subscripts 1, 2 and 0 represent positive, negative and zero sequence components respectively. The sum of each phase's symmetrical components is equal to the original unbalanced phase voltage (Fortescue, 1918).

$$V_a = V_{a1} + V_{a2} + V_{a0} \quad (1)$$

$$V_b = V_{b1} + V_{b2} + V_{b0} \quad (2)$$

$$V_c = V_{c1} + V_{c2} + V_{c0} \quad (3)$$

In electrical power system, symmetrical components are rotated 120 degrees from each other. "a" is an operator that present the difference of phase, and it is represented by complex numbers. (Hawary, 2002)

$$a = 1\angle 120^O = 1. e^{\frac{j2\pi}{3}} = -0.5 + j0.866 \quad (4)$$

With this operator a, symmetrical components of all phases could be represented by components of single phase. For example, the components of both V_b and V_c could be expressed by components of V_a with operator a. (Fortescue, 1918)

$$V_{b1} = a^2 V_{a1} \quad (5)$$

$$V_{b2} = aV_{a2} \quad (6)$$

$$V_{b0} = V_{a0} \quad (7)$$

$$V_{c1} = aV_{a1} \quad (8)$$

$$V_{c2} = a^2 V_{a2} \quad (9)$$

$$V_{c0} = V_{a0} \quad (10)$$

We could obtain the phase voltages by substituting equations (5) – (10) into (1) – (3) respectively.

$$V_a = V_{a0} + V_{a1} + V_{a2} \tag{11}$$

$$V_b = V_{a0} + a^2 V_{a1} + a V_{a2} \tag{12}$$

$$V_c = V_{a0} + a V_{a1} + a^2 V_{a2} \tag{13}$$

A matrix equation can express the symmetrical components of three phase voltages.

$$\begin{matrix} V_a \\ V_b \\ V_c \end{matrix} = \begin{matrix} 1 & 1 & 1 \\ 1 & a^2 & a \\ 1 & a & a^2 \end{matrix} \cdot \begin{matrix} V_{a0} \\ V_{a1} \\ V_{a2} \end{matrix} \tag{14}$$

A matrix is represented as shown below:

$$A = \begin{matrix} 1 & 1 & 1 \\ 1 & a^2 & a \\ 1 & a & a^2 \end{matrix} \tag{15}$$

CHAPTER 5: SIMULATIONS AND EXPERIMENTAL RESULTS

After studying the support vector machine and power system fault modeling with complex mathematical work, this chapter presents the MATLAB simulation of various fault types in order to generate beneficial experimental results.

5.1 Data Gathering

For data gathering, a 3 phase transmission line model has been presented. This transmission line model is also linked with a 3 phase generator (Koc & Aydogmus, 2009). All the presented values are given in the per unit form. In this circuit, base voltage is selected as 22KV, whereas base current is selected as 400MVA. Parameters of the circuit are presented below:

(All values are in per-unit, Vbase=22KV, Sbase=400MVA)

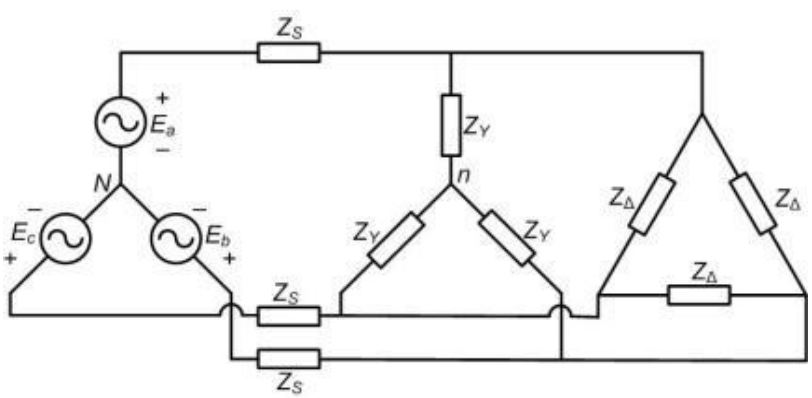

Figure 7. Three Phase Transmission Systems. (Saadat, 2005)

Parameters:

Frequency of generator is f=50Hz

Ea=1 pu

Impedance sequences of transmission line are:

Positive sequence: Z1=0.603j pu

Negative sequence: Z2=0.603j pu

Zero sequence: Z0=0.089j pu

Fault impedance: Zf=0.00~0.19j pu

There are four types of Fault that will be simulated as shown below:

1 Single Line to ground Fault (a-g)

2 Line to Line Fault (b-c)

3 Double Line to ground Fault (bc-g)

4 3 Phase to ground Fault (abc-g)

Sampling: Sampling starts from the first ¼ circle after fault occurs at the frequency of Sampling: fs=1000. Therefore, there are 5 times of sampling from 0.00s to 0.05s. Simulation Results of fault data:

- Phase A is represented by the color Red
- Phase B is represented by the color Blue
- Phase C is represented by the color Green

For single line to ground fault (a-g), the data is as shown in Figure 8. A single line to ground fault is a type of fault that occurs between a single phase A and the ground of the power system. The result can be seen in the graph at sampling frequency fs=1000. Therefore, there are 5 times of sampling from 0.00s to 0.05s. Phase A is represented by the color Red in this diagram.

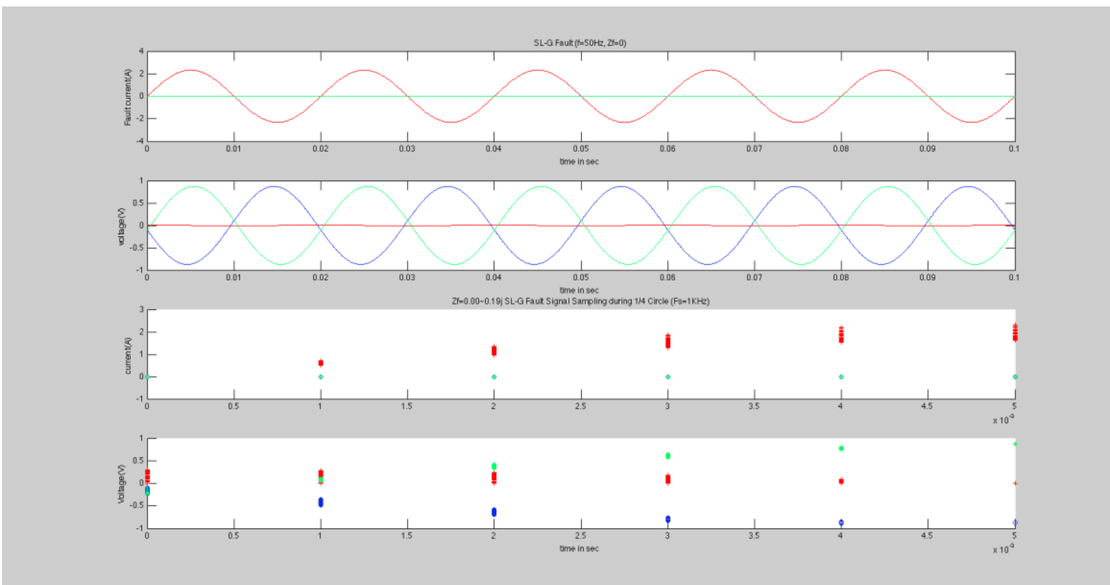

Figure 8. Single Line to Ground Fault

Figure 9 shows the line to line fault (b-c) samples. A line to line fault is a type of fault that occurs between two phases with the exception of the ground and the remaining third phase. The two participant phases in this fault include Phase B and Phase C of the power system. The result can be seen in the graph at sampling frequency fs=1000. Therefore, there are 5 times of sampling from 0.00s to 0.05s. Phase B is represented by the color Blue and phase C is represented by the color Green in this diagram.

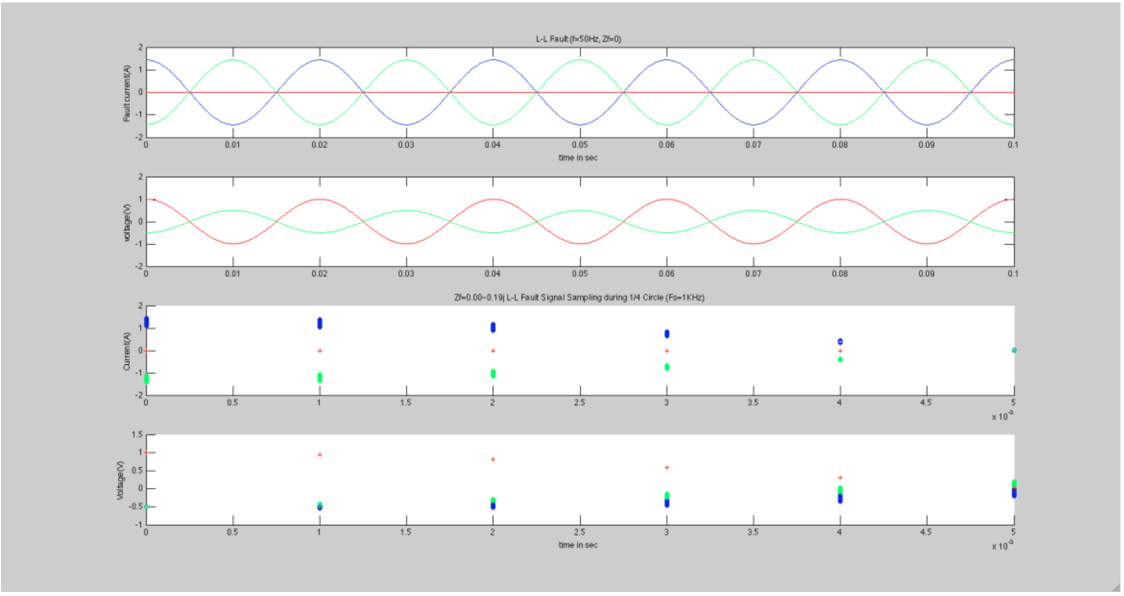

Figure 9. Line to Line Fault

The double line to ground fault (bc-g) data is shown in Figure 10. A double line to ground fault is a type of fault that occurs between two phases and the ground of the power system with the exception of the ground and the remaining third phase. The two participant phases in this fault include Phase B and Phase C. The result can be seen in the graph at sampling frequency fs=1000. Therefore, there are 5 times of sampling from 0.00s to 0.05s. Phase B is represented by the color Blue and phase C is represented by the color Green in this diagram.

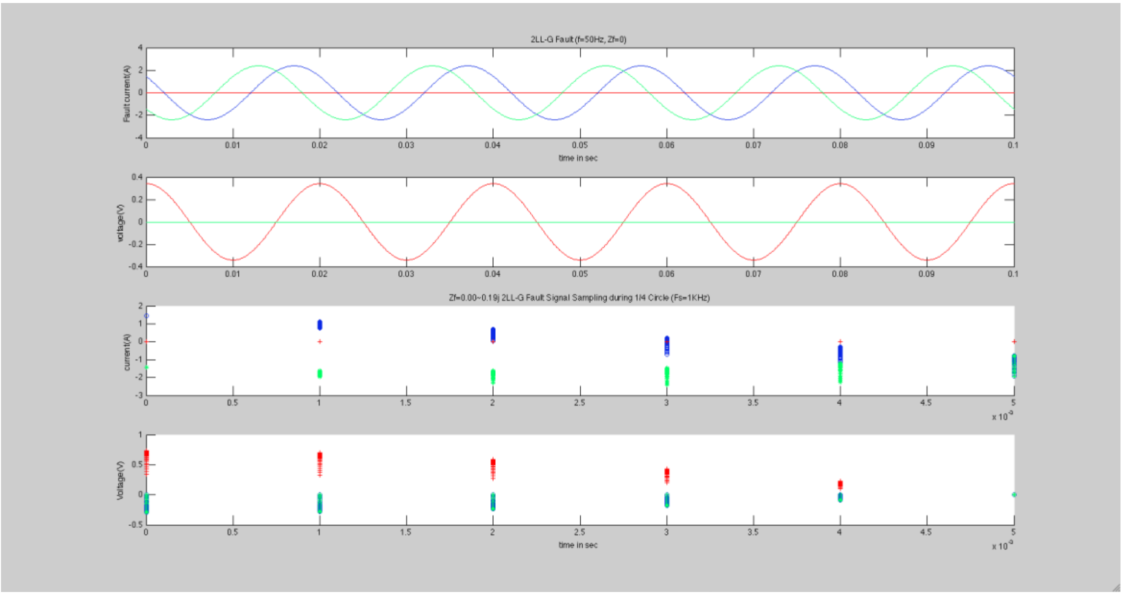

Figure 10. Double Line to Ground Fault

Figure 11present data for phase to ground fault (abc-g). A phase to ground fault is a type of fault that occurs between three phases and the ground of the power system. This means that all the four lines are involved in this type of fault. The two participant phases in this fault include Phase B and Phase C. The result can be seen in the graph at sampling frequency fs=1000. Therefore, there are 5 times of sampling from 0.00s to 0.05s. Phase A is represented by the color Red, Phase B is represented by the color Blue and phase C is represented by the color Green in this diagram.

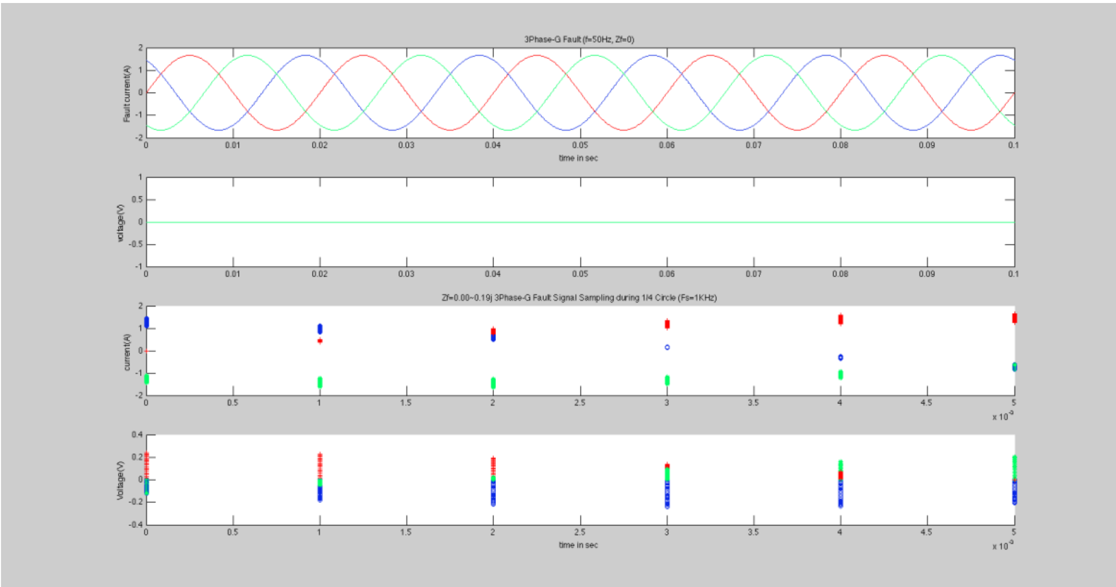

Figure 11. 3 Phase to Ground Fault

5.2 Development of Pattern Data

Consider about each Phase separately, define a condition parameter Y which results “1” for fault and “-1” for no fault.

	Phase A	**Phase B**	**Phase C**
SL-G (A-g)	1	-1	-1
L-L (B-C)	-1	1	1
2L-G (BC-g)	-1	1	1
3P-G (ABC-g)	1	1	1

Table 1. Development of Pattern Data 1

Therefore, target is changed from 4 classes' classification to 2 classes' classification that is operated 3 times. There are SVM-A for Phase A, SVM-B for Phase B, and SVM-C for Phase C.

The input data for SVM are the Voltage and Current during ¼ circles after the fault occurs. Simulation results for the case of input of:

Red: SL-G Fault Blue: L-L Fault Green: 2L-G Fault Yellow: 3P-G Fault

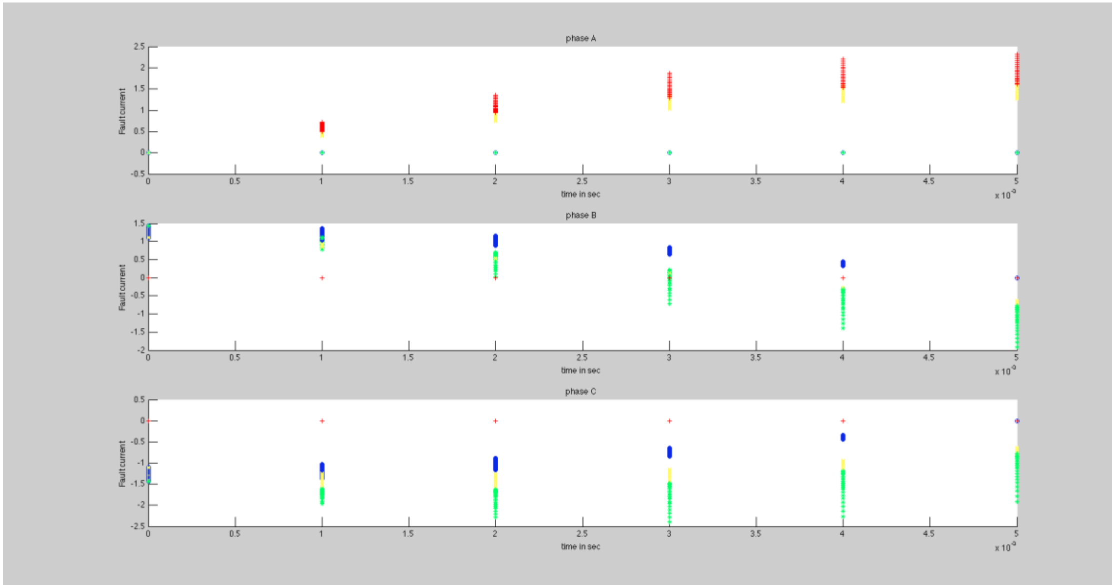

Figure 12. Fault Current Samples on Time-Current Axis

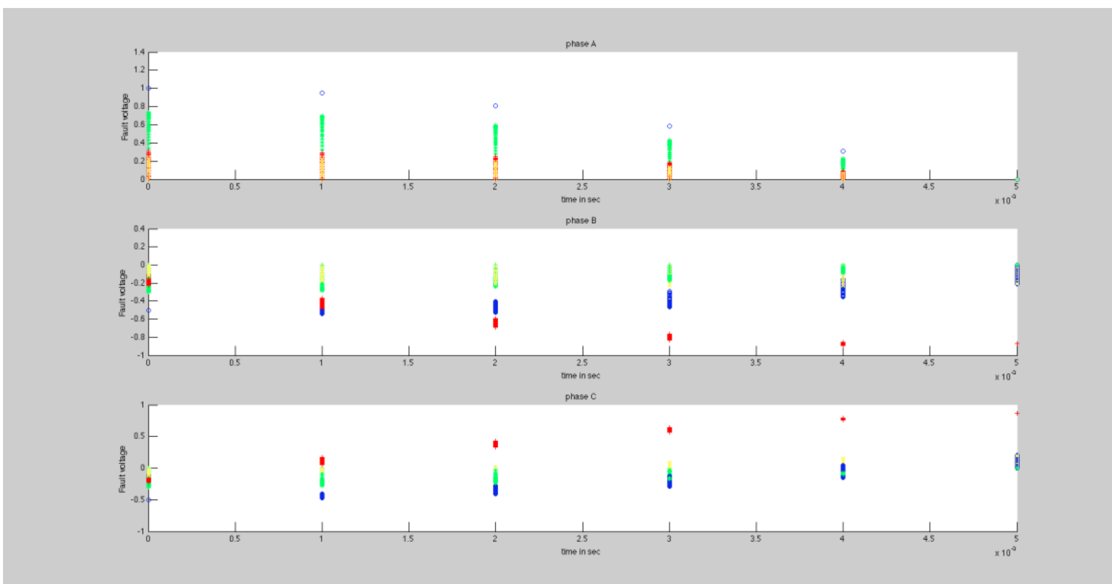

Figure 13. Fault Voltage Samples on Time-Voltage Axis

In order to increase the accuracy of classification, it is necessary to make sure that the patterns have no superposition part from each other. I changed the start sampling time from 0.00s to 0.02s, and sampling still ends at 0.05s. In addition, I adopt voltage-current axis instead of time-voltage or time-current as input learning pattern shown below:

Red: SL-G Fault Blue: L-L Fault Green: 2L-G Fault Yellow: 3P-G Fault

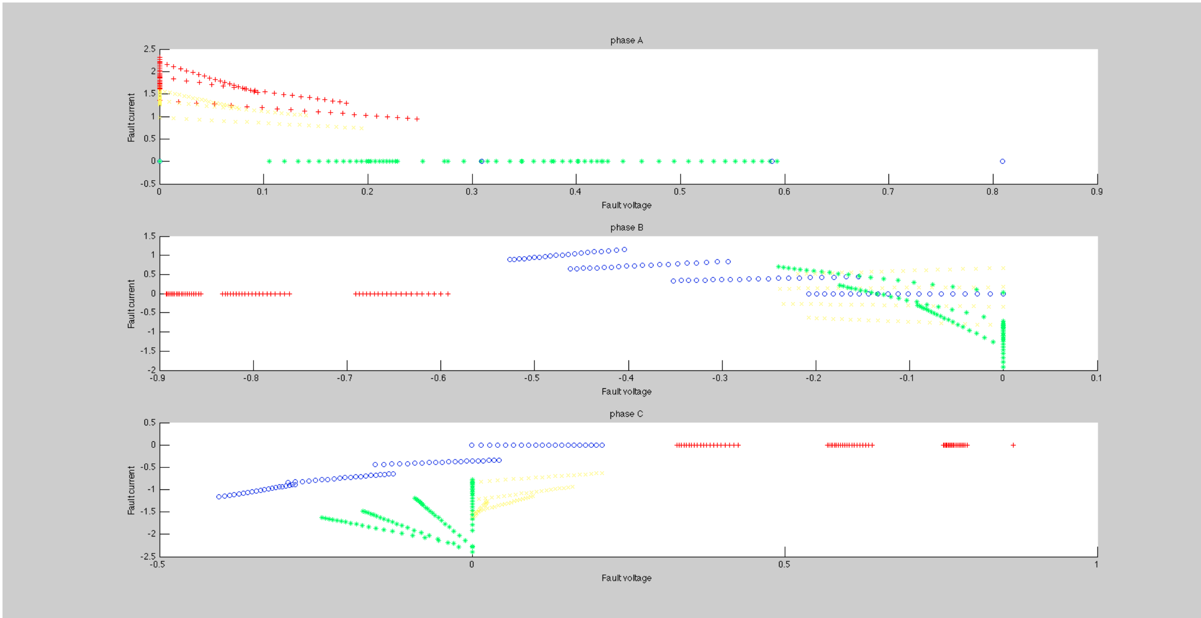

Figure 14. Fault Samples on Voltage-Current Axis

	Phase A	**Phase B**	**Phase C**
SL-G (A-g)	**1**	**-1**	**-1**
L-L (B-C)	**-1**	**1**	**1**
2L-G (BC-g)	**-1**	**1**	**1**
3P-G (ABC-g)	**1**	**1**	**1**

Table 2. Development of Pattern Data 2

5.3 Learning and Testing Results

Using the patterns in voltage-current domain as the learning data, a classification SVM is created for each phase. In the results shown below, the red area presents that the pattern of voltage-current is error, and the blue area presents that the pattern of voltage-current is correct. The plane of "0" is the hyper-plane $f(x)$ that can make a complete separation of two classes of samples. In addition, the black circle points are the support vector.

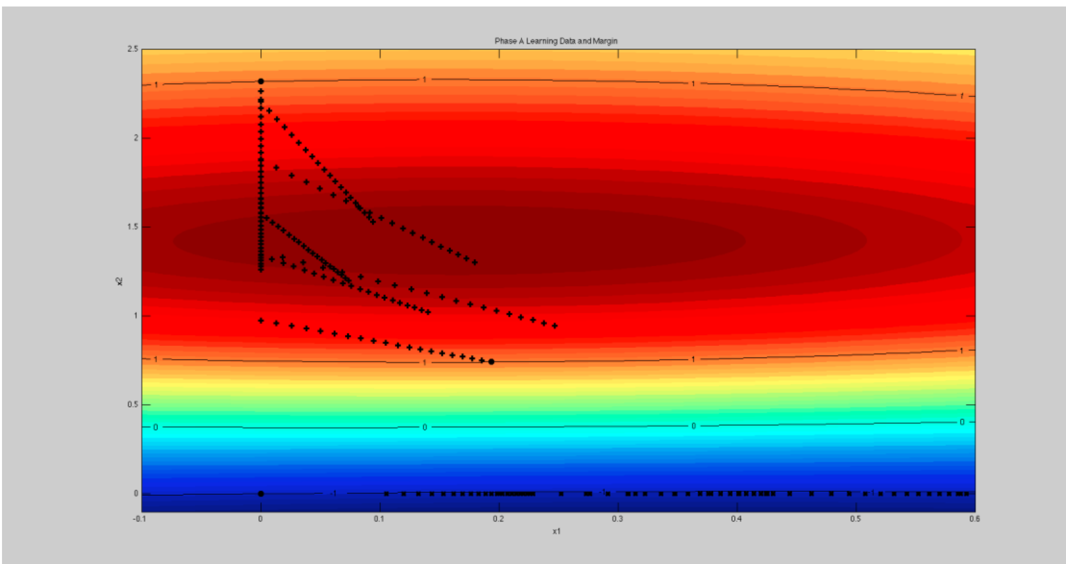

Figure 15. Showing Error, Correct and Support Vector Patterns in Phase A

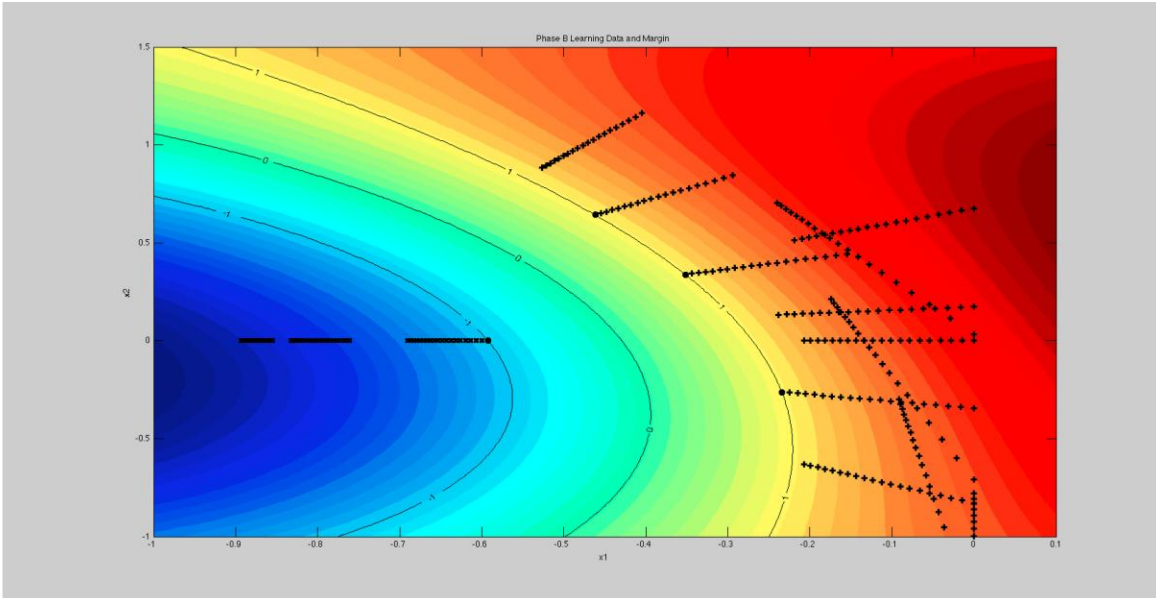

Figure 16. Showing Error, Correct and Support Vector Patterns in Phase B

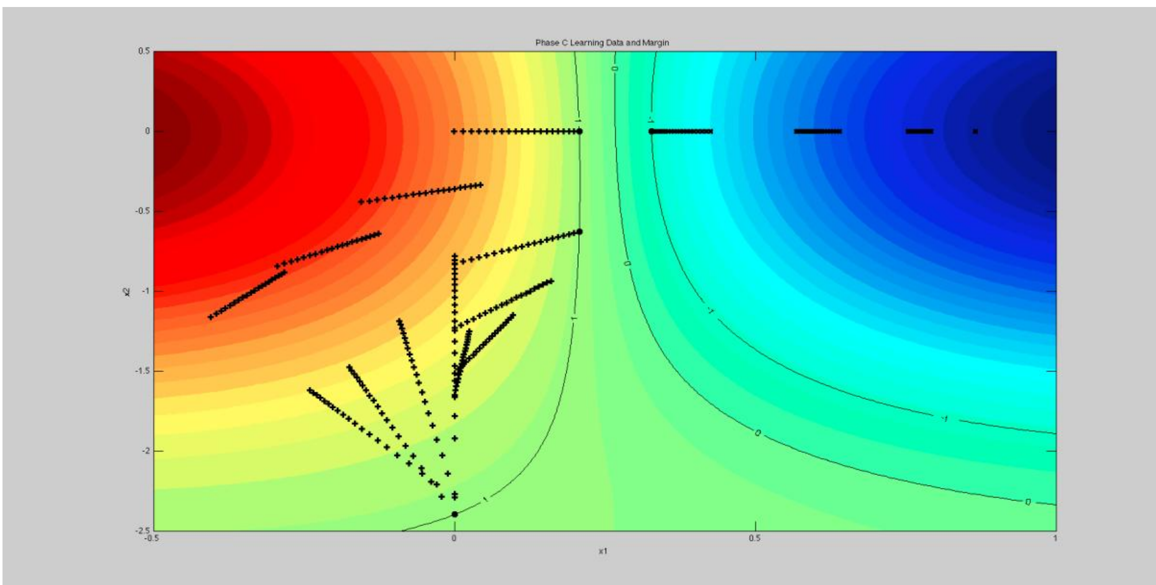

Figure 17. Showing Error, Correct and Support Vector Patterns in Phase C

It is worth to notice that in order to increase the accuracy of classification, it was necessary to make sure that the patterns have no superposition part from each other. The starting sampling time was also changed from 0.00s to 0.02s, and sampling still ends at 0.05s. This was done to achieve higher accuracy. In addition to this, voltage-current axis was adopted instead of time-voltage or time-current as input learning pattern which also enhanced the results.

CHAPTER 6: CONCLUSIONS

For fault feature extraction and fault diagnosis, this thesis presents a classification method based on SVM as a solution for the diagnosis of faults in power transmission lines. The simulation results show that this algorithm has high accuracy in the classification performance and wider generalization ability even in small group of samples using the learning and testing patterns in the voltage-current domain. In future work, validation of the work for classification in multiple busses is necessary. Furthermore, in this work, only faults occurring in designated phases were considered and so future validation can be confirmed by verifying the fault analysis results in every possible point.

The presented framework can always be updated to yield better results, as new data becomes available from the same system. Despite the good performance of SVM, it suffers from many difficulties in finding the optimum solution when the overlaps occur between the different classes patterns due to the similar sample input data. In order to enhance the performance, we can extend the sampling period from a quarter cycle to a half cycle or even longer. It can slightly enhance the accuracy due to there are more learning and testing samples that can represent significant features of different classes, it can even effectively avoid the classification errors due to other conditions like weathers other than fault. However, with the longer sampling time, the classifier cannot make the decision instantly due to requires more time to sample, learn and classify. To make the best choice between accuracy and speed, the quarter cycle was selected. In future work, there are more investigations that can be performed among the different sampling times needed to ensure best performance.

The ANN, fuzzy logic, genetic algorithm, SVM and wavelet based techniques have been quite successful but are not adequate for the present time varying network configurations, power system operating conditions and events. Therefore, it seems that there is a significant scope of research in AI techniques which can simplify the complex nonlinear systems, realize the cost effective hardware with proper modification in the learning methodology and preprocessing of input data and which are computationally much simpler. Also development of reliable software and communication system will pave the way for better relaying and fault location performance using multi terminal synchronized phasor measurement based on GPS.

BIBLIOGRAPHY

Bhowmik, K. and D. G. Flinn (2009) A Windows Based Graphical Package for Symmetrical Components Analysis. *IEEE Transactions on Power System.*

Bi, T., Ni, Y., & Wu, F. (2002). A novel neural network approach for fault section estimation. *Proceddings of the CSEE , 22* (2), pp. 73-78.

Cardoso, G., & Rolim, J. (2004). Application of neural-network modules to electric power system fault section estimation. *IEEE Transactions on Power Delivery , 19* (3), pp. 1034-1041.

Chan, E. (1989). A application of nerual network computing in intelligent alarm processing. *Power Industry Computer Application Conference* (pp. 246-251). Seattle, USA: IEEE.

Cho, H., & Park, J. (1997). An expert system for fault section diagnosis of power systems using fuzzy relations. *IEEE Transactions on Power Systems. , 12* (1), pp. 342-348.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning , 20* (3), pp. 273-297.

Dash, P., Samantaray, S., & Panda, G. (2007). Fault classification and section identification of an advanced series-Compensated Transmission Line Using Support Vector Machine. *IEEE Transactions on power delivery , 22* (1), pp. 67-73.

FarhatIbrahim. (2003). Fault Detection, Classification and Location in Transminssion Line Systems Using Neural Networks.

Fortescue, C. (1918). *Method of Symmetrical Coordinates Applied to the Solution of Polyphase Networks* (Vol. 37). Trans. AIEE, pt. II.

Ganyun, L., Cheng, H., Zhai, H., & Dong, L. (2005). Fault Diagnosis of Power Transformer Based on Multi-layer SVM Classifier. *Electric Power Systems Research , 74*, pp. 1-7.

Gonen, T. (1986). *Electric Power Distribution System Engineering.* McGraw Hill Book Comp.

Gonen, T. (1988). *Mordern Power System Analysis.* John Willey and Sons Inc.

Guo, G., Li, S., & Chan, K. (2000). Face recognition by support vector machines. *Automatic Face and Gesture Recognition* (pp. 196-201). Fourth IEEE International Conference.

Guo, L., Chen, J., & Li, X. (2009). Rolling Bearing Fault Classification Based on Envelope Spectrum and Support Vector Machine. *Journal of Vibration and Control , 15*, 1349.

Hawary, M. E. (2002). Electrical Power Systems. *Power Systems Engineering Series* .

Huang, Y. (2002). Fault section estimation in power system using a novel decision suport system. *IEEE Transactions on Power Systems. , 17*, pp. 439-444.

Jadid, S., Jeyasurya, B., & Khaparde, S. (1989). Power system fault diagnosis expert system using PROLOG[A]. *Fourth IEEE Region 10 International Conference* (pp. 778-781). Bombay(India): IEEE.

Janik, P., & Lobos, T. (2006). Automated Classification of Power Quality Disturbances using SVM and RBF Networks. *IEEE TRANSACTIONS ON POWER DELIVERY , 21*, 3.

Koc, S., & Aydogmus, Z. (2009). A Matlab/GUI Based Fault Simulation Tool For Power System Education. *Mathematical and Computational Application , 14* (3), 207.

Lee, B.-h., Kim, S.-u., Seok, J.-w., & Won, S. (2006). Nonlinear System Identification based on Support Vector Machine using Particle Swarm Optimization. *SICE-ICASE, 2006* (pp. 5614-5618). International Joint Conference.

Liang, G. and Sakaguchi, T. (2004). Method to determine he fault components of power system based on description of structure and function of relay system. *IEEE Trans. Japan , 104-B(10)*, pp. 248-256.

Milenova, B.L., Yarmus, J.S., Campos, M.M., (2005) SVM in Oracle Database 10g: Removing the Barriers to Widespread Adoption of Support Vector Machines, Proceedings of the 31st VLDB Conference, Trondheim, Norway.

Monsef, H., Ranjbar, A., & Jadid, S. (1997). Fuzzy rule-based expert system for power system fault diagnosis. *IEE Proceedings: Generation, Transmission and Distribution , 144* (2), pp. 186-192.

Park, Y. M., Kim, G. W., & Sohn, J. M. (1997). A logic based expert system (LBES) for fault diagnosis of power system[J]. IEEE Transaction on Power System, pp.363-369.

Qais, H. A. (2010). Pattern Recognition For Fault Detection, Classification, and Localization in Electrical Power Systems.

Rabbani, M., Nasiruzzaman, A., Sheikh, R., & Shamim Anower, M. (2006). Matlab Based Fault Analysis Toolbox for Electrical Power System. *4th International Conference on Electrical and Computer Engineering.*

Ravikumar, B., Thukaram, D., & Khincha, H. (2007). Intelligent Approach for Fault Diagnosis in Power Transmission Systems Using Support Vector Machine. *International Journal of Emerging Electric Power Systems, 8* (4).

Rongjie, W., & Haifeng, Z. (2008). Application of SVM in Fault Diagnosis of Power Electronics Rectifier. *7th World Congress on Intelligent Control and Automation.* Chongqing, China.

Saadat, H. (2005). *Power System Analysis* (2nd ed.). McGraw-Hill Primis Custom Publishing.

Sanaye-Pasand, M., & Malik, O. (2001). Neural Network-Based Fault Direction Discrimination for High-Speed Transminssion Line Protection. *Electric Power Components and Systems, 29* (8), pp. 757-770.

Stevenson, W. (1982). *Elements of Power System Analysis.* McGraw-Hill.

Valsan, K. and Campos, M.M. (2009) Removing the Barriers to Widespread Adoption of Support Vector Machines, Proceedings of the 31st VLDB Conference, Trondheim, Norway.

Wake, A., & Sakaguchi, T. (1984). Method to determine he fault components of power system based on description of structure and function of relay system. *IEEE Trans. Japan , 104-B(10),* pp. 248-256.

Wang, W. (2007). Time Series Prediction Based on SVM and GA. *Electronic Measurement and Instruments, 2007. ICEMI '07* (pp. 2-307-2-310). 8th International Conference.

Welling, M. (2004). *Support Vector Machines.* Toronto: University of Toronto.

Yu, D., & Chen, D. (1995). S. Ramsamy and D. G. Flinn, A Windows Based Graphical Package for Symmetrical Components Analysis. *IEEE Transactions on Power Systems*, (pp. 1742-1749).

Zhang, Y., Liu, Y., Wang, X., & Wang, Z. (2009). Fault Pattern Recognition in Power System Engineering. *International Conference on Industrial Mechatronics and Automation.*