Mitigating Future Blackouts via Smart Relays:  
A Machine Learning Approach

Submitted in partial fulfillment of the requirements for  
the degree of  
Doctor of Philosophy  
in  
Department of Electrical and Computer Engineering

Yi Zhang

B.S., Engineering Physics, Tsinghua University  
M.S., Electrical and Computer Engineering, Carnegie Mellon University

Carnegie Mellon University  
Pittsburgh, PA

July, 2010
**Keywords:** Power systems, Protective relays, Machine learning, Support vector machines, Hypothesis testing, Decision making
Abstract

As the electric power systems in the United States become increasingly large, complex, and interconnected, the conventional relays and protection systems are proving to be inadequate during some abnormal conditions. In particular, there exists a significant history of relay protection schemes malfunctioning and, ultimately, leading to the infamous system-wide failures, known as electric power blackouts. The malfunctioning ranges from: (1) disconnecting a functional equipment component because of ‘false alarms’ which are caused by abnormal conditions elsewhere in the system, and triggering cascading failures of other components; (2) not clearly differentiating the equipment failures from unusually large load demand deviations; and, (3) not providing sufficient coordination of the affected components to disconnect service only to the minimal number of customers and to isolate the rest of the system from the effects of the triggering events. Considering the possibility of carefully planned malicious attacks on the electric power system, today’s protection systems would be inadequate during such conditions, as well. More intelligent relays are, therefore, needed to meet both security and reliability requirements of the current and future electric power grids. In this dissertation, we investigate the existing logic of protection relays in electric power systems and their roles in preventing or mitigating large-scale blackouts. We review several proposed solutions to this problem which employ communications and intelligent algorithms. After reviewing such solutions, we propose a new machine learning based approach for the design of smart protective relays. The goal of smart relays is to classify and discriminate normal conditions from fault conditions based on local measurements. It is shown that the proposed SVM-based smart relays can detect the location of an initial fault (in terms of which zone it belongs to) using local current, voltage, real power and reactive power measurements; and by continuing to monitor these metrics they can make a correct decision even when the state of the system changes after some equipment failure. By making an intelligent decision on whether and when to trip, and communicating the changes observed to SCADA for quick and intelligent decision making, SVM-based smart relays have the potential to mitigate large-scale blackouts and confine them to much smaller areas. Notably, we show that by using SVM-based smart relays only at relatively few critical locations where they have the highest probability to be tripped incorrectly, the probability of cascade of failures and a blackout can be substantially reduced.
For my parents
Acknowledgments

I would like to thank my advisors, Professor Marija Ilic and Professor Ozan Tonguz, for their guidance and continuous support during my six years of study at Carnegie Mellon University (CMU). Prof. Ilic was there whenever I needed her help. She offered generous help on numerous occasions. Prof. Tonguz provided me with innovative ideas and great insights. He helped me to identify the most impactful goals and encouraged me to move towards them. Besides providing me with invaluable insights and direction, he also taught me how to do research by introducing me to the scientific method, which I did not know before starting my PhD studies at CMU.

I would also like to thank other members in my thesis committee. Professor Vijayakumar Bhagavatula has given me valuable suggestions after my proposal, which help me to greatly improve the quality of my work. Doctor Damir Novosel at Quanta Technology was instrumental in providing useful feedback from both technical and industrial perspectives, adding an important dimension to my research work.

I am very grateful to Eric Udren, the mentor of my summer internship with Quanta Technology. He is one of the best teachers I have had the pleasure to meet. With discussions and insightful comments he has greatly contributed to my research and understanding of protective relays. A considerable portion of my understanding of the protective relays is due to Eric Udren, during my summer internship with Quanta Technology. My understanding of the protective systems would not have been the same without him.

My deepest thanks also go to other members in my research groups: Marija Prica, who has been my research partner, gave me invaluable advice, and guided me through my first year at CMU in the M.S program; Le Xie, who is one of my best friends and colleagues, helped me to clarify definitions and solve problems via hundreds of discussions; Wantanee Viriyasitavat and Hsin-Mu Tsai, although they were working on different research topics, they offered me generous help and encouragement; Paisarn Sonthikorn, Mate Boban, and Xindian Long tremendously helped me in every ways in my research. I would also like to thank my fellow
graduate students in the ECE and CS department; Qi Wu and Xiaoqian Jiang, who have provided generous help in understanding, applying and modifying the machine learning mechanism I used in my research.

I would also like to greatly acknowledge the financial support from the Army Research Office through grant number DAAD19-02-1-0389 ("Perpetually Available and Secure Information Systems") and W911NF-09-1-0273 to Carnegie Mellon University’s CyLab, and the U.S. National Science Foundation under award CNS-0428404.

Last but not least, I would like to thank my parents, Kaixian Zhang and Yiping Wang, for their love and encouragement. In spite of the distance, they are always with me in all my endeavors.
# Contents

1 Introduction .......................................................... 1

2 Background .......................................................... 4
   2.1 Blackouts ......................................................... 4
   2.2 Current Practice of Protective Relays ...................... 6
   2.3 Modern Relays in Practice ................................... 11
   2.4 Approaches to Smart Protective Relays .................... 16

3 Machine Learning based Smart Protective Relays .......... 27
   3.1 Problem Formulation .......................................... 27
   3.2 Introduction to Machine Learning ......................... 28
   3.3 Hypothesis Testing Based Smart Protective Relays .... 29
      3.3.1 Introduction of Hypothesis Testing .................. 29
      3.3.2 Hypothesis Testing Based Relays .................... 30
   3.4 Support Vector Machine Based Protective Relays ....... 32
      3.4.1 Introduction of Support Vector Machines (SVM) .. 32
      3.4.2 Proposed SVM-based Relays .......................... 37
      3.4.3 Model Selection by Cross-Validation and Grid Search . 41
      3.4.4 Dimensionality Reduction: Feature Extraction and Feature Selection ... 41
   3.5 Simulation results ............................................ 47
      3.5.1 Simulation results on IEEE test networks: SVM Training and Testing . 47
      3.5.2 Scalability ................................................ 51
   3.6 Summary ......................................................... 56
## 4 Comparison of SVM-based Smart Relays and Conventional Relays

4.1 Case Study .......................................................... 58

4.2 Benchmarking Performance of SVM-based Smart Relays and Conventional Relays in Cascading Failures .......................................................... 61

4.2.1 Benchmarking Performance ..................................... 61

4.2.2 Critical Locations .................................................. 67

4.3 Benchmarking Performance of SVM-based Smart Relays and Modern Distance Relays in Cascading Failures ......................................................... 70

4.4 Summary .............................................................. 77

## 5 Handling of Missing Measurements and Updating

5.1 Handling of Missing Measurements ............................. 78

5.1.1 Introduction of K-Nearest Neighboring ..................... 78

5.1.2 Estimation of Missing Measurements ..................... 81

5.2 Dynamic updating on SVM-based smart relays ............. 84

5.2.1 Proposed Approach ............................................. 85

5.2.2 Simulation Results ............................................. 90

5.3 Summary .............................................................. 93

## 6 Transient state SVM-based Smart Relays

6.1 Proposed Approach .................................................. 94

6.2 Simulation Results .................................................. 96

6.3 Obstacles to implementing Transient Measurement Based SVM relays ....................... 98

6.4 Summary .............................................................. 99

## 7 Conclusion

7.1 Main Contributions of the Thesis ............................... 100

7.2 Practical Implications ............................................. 102

7.3 Future Work ......................................................... 106

## 8 Appendix

.......................................................... 108
List of Figures

2.1 Basic Structure of Power Systems and their Protection Systems. Observe that protection systems exist in all three components (generation, transmission, and distribution) of the electrical power system. .......................................................... 7

2.2 (a)Protection Zones and (b) Characteristic Curve for Over-current Relays. .............................. 8

2.3 (a)RX diagram and (b)Protection Zones for Distance Relays. R and X correspond to the resistive and reactive components respectively of the impedance measured by a distance relay. .......................... 9

2.4 Two bus system with distance relay. .......................................................................................... 10

2.5 Typical R-X diagram in digital distance relays with Load-encroachment characteristic. [11]. 14

2.6 Phasor measurement units that function with the aid of GPS satellite ................................. 15

2.7 Voltage profile and gradient with and without contribution of distributed generators after [15]. Here T denotes the transformer, G denotes the distributed generator, U denotes voltage, and x denotes the distance from bus 1. The voltage along this transmission line is a monotonically decreasing function of the distance x if there is no DG. However, the connection of DG can change the voltage, and hence the current direction on the line. ............ 19

2.8 Typical LAN structure in substation automation after [28]. The IEDs in this figure are primary and backup protection, and the HMI stands for Human Machine Interface. .............. 23

2.9 Power system states and transactions between them [29] .............................................................. 26

3.1 Conditional probability density functions (pdf) of normal current and fault current in an over-current relay. ........................................................................................................ 30
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2</td>
<td>Multi-threshold setting for smart relays. Observe that there are several thresholds defining the normal region and faulty region. Accordingly, a current magnitude larger or smaller than the normal current range can indicate a fault.</td>
</tr>
<tr>
<td>3.3</td>
<td>The separating hyperplane that maximizes the margin. (‘o’ is a positive data point, <em>i.e.</em>, ( f('o') &gt; 0 ), and ‘•’ is a negative data point, <em>i.e.</em>, ( f('•') &lt; 0 )).</td>
</tr>
<tr>
<td>3.4</td>
<td>The separating hyperplane that maximizes the margin in non-linear separable case. (‘o’ is a positive data point, <em>i.e.</em>, ( f('o') &gt; 0 ), and ‘•’ is a negative data point, <em>i.e.</em>, ( f('•') &lt; 0 )).</td>
</tr>
<tr>
<td>3.5</td>
<td>Support Vector Machine classification with 2 features: magnitude of voltage and current.</td>
</tr>
<tr>
<td>3.6</td>
<td>Grid Search for Model Selection in SVMBC Relays.</td>
</tr>
<tr>
<td>3.7</td>
<td>The histogram of the distributions of six features, which are the magnitude and phasor of current and voltage, real and reactive power. Dark blue is the normal class, red indicates zone 1 fault and light blue indicates relax class.</td>
</tr>
<tr>
<td>3.8</td>
<td>The plot matrix of the three most important features, which are reactive power (feature 6), real power (feature 5), and phasor of voltage (feature 4). Blue is the normal class, red indicates zone 1 fault and green indicates relax class.</td>
</tr>
<tr>
<td>3.9</td>
<td>The Accuracy of SVM-based Relays with Various Fault Impedances.</td>
</tr>
<tr>
<td>3.10</td>
<td>The Accuracy of SVM-based Relays Which Are Trained with Different Inaccurate Parameters.</td>
</tr>
<tr>
<td>3.12</td>
<td>Testing and Training Accuracy in Different Size Networks.</td>
</tr>
<tr>
<td>4.1</td>
<td>7-Bus System in an European Power System [46].</td>
</tr>
<tr>
<td>4.2</td>
<td>The current in 7bus network (a)before and (b)during a fault in transmission line 1-2. This fault current can trigger the relay on line 1-7 to false trip.</td>
</tr>
<tr>
<td>4.3</td>
<td>The ratio of possible initial disturbances which will lead to incorrect tripping in different load levels.</td>
</tr>
<tr>
<td>4.4</td>
<td>The total number of incorrect tripping after possible initial disturbances in different load levels.</td>
</tr>
<tr>
<td>4.5</td>
<td>Typical 3 zones design of the conventional distance relays.</td>
</tr>
<tr>
<td>4.6</td>
<td>Typical 3 zones design of the conventional distance relays.</td>
</tr>
</tbody>
</table>
4.7 Required number of simulations as a function of the number of circuit elements in N-1 (left) and N-2 (right) studies ................................................................. 68
4.8 The histogram of false tripping events on the relay in different load level (relay 23-25) .... 69
4.9 The histogram of false tripping events on the relay in different load level (relay 89-92) ... 70
4.10 The R-X diagram for critical relays ........................................................................... 71
4.11 The R-X diagram for non-critical relays .................................................................... 71
4.12 The histogram for node degree in the IEEE 118 bus system ..................................... 72
4.13 The R-X diagram with Load Encroachment Element in Distance Relays .................. 73
4.14 The functional logic and the apparent impedance seen by line 63-59 in different system conditions ......................................................................................... 74

5.1 Estimate the missing measurement via K-NN .......................................................... 80
5.2 Parameter selection on K-NN .................................................................................... 81
5.3 The estimation and the nearest neighbors when the reactive power measurement is missing in the target data ................................................................. 82
5.4 The estimation and the nearest neighbors when the phase of current measurement is missing in the target data ................................................................. 83
5.5 Comparison of $\epsilon\alpha$-estimators as a bound against leave one out error. (a) Without concept drift. (b) with concept drift. ................................................................. 91
5.6 Locating concept drifting point with dynamic step functions. Simulation results indicate that a fitted step function locates a concept drifting which occurs at data point 413 whereas its real value is 403. ................................................................. 92
5.7 Simulation results indicate that the training time with bunch learning method in [50] increases sharply as the bunch size decreases while our algorithm requires constant time. Our incremental method has a great advantage in terms of time consumption. ......... 92

6.1 Typical Transient Measurements of Current Magnitude ......................................... 96
6.2 Typical transient current measurements and feature extraction .............................. 96
6.3 Processed Transient current measurements and feature extraction ......................... 97

7.1 Current Topology and Structure of Power Grid (modified from [59]) ..................... 103
7.2 Representative oversimplified power grid as an abstraction. ............................ 105

8.1 General Structure of a Neural Network ............................................................. 108
# List of Tables

2.1 Some notable cascading failures in North America and Europe .......................... 5
2.2 Communication Media Used in Pilot Protective Scheme by Different Utilities ........... 12

3.1 Error rate of hypothesis testing-based smart relays in power systems of different sizes .......................... 32
3.2 Accuracy of SVM-based Smart Relays with Two Features .................................. 47
3.3 Accuracy of SVM-based Smart Relays with Three Features .................................. 48
3.4 Accuracy of SVM-based Smart Relays with Six Features .................................. 48
3.5 Accuracy of SVM-based Smart Relays in (N-1) Conditions .................................. 48

4.1 Accuracy of SVM-based Smart Relays on Line 1-7 with Six Features ................. 60
4.2 Accuracy of SVM-based Smart Relays on Line 6-7 with Six Features ................. 60
4.3 Accuracy of SVM-based Smart Relays on Line 1-7 with Three Features ............... 61
4.4 Accuracy of SVM-based Smart Relays on Line 6-7 with Three Features ............... 61
4.5 Probability of Cascading Failure after Initial Disturbances in Systems with Different Protective Relays ................................................................. 76

5.1 The best value of parameter $k$ on K-NN with different missing features ............... 81
5.2 The testing accuracy with missing measurement in 6 features ............................. 83
5.3 The testing accuracy with missing measurement in 6 features ............................. 84
5.4 The testing accuracy with missing measurement in 4 features ............................. 84
Chapter 1

Introduction

Today’s power systems are very large, complex, and interconnected systems. Because of the increasing dependence on electricity, ensuring its delivery in a secure and reliable manner is of paramount importance to both customers and suppliers. To meet these requirements, the systems have to be monitored, controlled, and protected. This has led to an increase in the number of meters and sensors within the electric power system, such as Phasor Measurement Units (PMUs) and Advanced Metering Infrastructure (AMI). With the aid of communications, such devices can monitor and obtain the information needed about the entire system continuously.

Of particular interest in this dissertation is the role of existing protective relays in preventing or mitigating large-scale blackouts. Their key function is to quickly identify equipment exposed to conditions outside the acceptable operating range. These abnormal conditions are caused by hard-to-predict events, such as short-circuiting caused by trees touching transmission lines, and the like. At the same time, the protection systems are expected to reconnect the affected equipment as soon as the conditions return to normal. Consequently, they are essential for ensuring as little interruption as possible to customers’ electricity service. At the system level, we observe that the protective relays, which are installed to protect every piece of equipment in power system, can function as distributed sensors. Protective relays, which are already installed on every piece of equipment in contemporary power systems, continuously measure magnitude of voltage and current, and sometimes, voltage phasor. They obtain information from these local measurements and determine the system’s condition independently. Subsequently, with or without the aid of some communications, they can isolate the abnormal parts from the remainder of the system, keep the rest of the system...
intact, and send information and decisions to SCADA. Therefore, in fact, modern relays are functioning as sensors and protectors at the same time. To achieve this goal, a protective relay has to have an accurate and smart functional logic.

The conventional relays are not sophisticated enough to satisfy today’s needs. In some situations, they are not adaptive enough to discriminate between fault and normal conditions, or to react correctly to faults. Malfunctioning of relays is among the most common modes of failure that accelerates the geographic spread (or the cascade) of faults. Every four months, the United States experiences a blackout large enough to leave half a million homes in dark [1]. According to the historical data, relay malfunctioning is one of the major contributing factors to 70% of the major disturbances in the United States [2][3].

At the same time, the trend in power system planning that utilizes tight operating margins with less redundancy, addition of distributed generators, and independent power producers, makes the power system more complex to operate and to control and, therefore, more vulnerable to disturbances. Current control strategies are sometimes inadequate to stop the spreading of disturbances. In such cases, one could only rely on protective relays to protect the system from the wide-spread effects of fast disturbances. This suggests that the protection systems should be more reliable, secure, and robust. Therefore, more intelligent and sophisticated protective relays are needed.

In this dissertation, we propose a novel protective logic for digital relays based on a machine learning technique known as Support Vector Machines (SVM). By training a classifier, SVM can provide a non-linear decision boundary based on multiple measurements [4]. The conventional relays which respond to preset, non-adaptive tripping thresholds are based only on local voltage and current measurements. For example, a distance relay has pre-determined tripping thresholds based on the ratio of the magnitude of local voltage and current on a pre-defined setting of the system. These thresholds might not be valid when the state of the power system changes, for example after equipment failure or other disturbances. Compared to conventional relays, SVM-based smart relays have a decision boundary based on the statistical information obtained from several local measurements including the power. They can therefore accurately detect and locate the initial disturbance in the system, as well as the system situation or state after the isolation of this disturbance. Based on these decisions, the SVM-based smart relays can decide whether and when to trip a transmission line. This can stop the propagation (or cascading) of failures and/or confine it to a limited small area. It is important to understand that the protective algorithm we propose recognizes the fact that the current power grid is a hierarchical and centralized network and attempts to make the current power grid a self-organizing
network in spite of its inherent topological constraints. Thus, the SVM-based smart relays employ a stand-alone algorithm to make intelligent decisions without major changes in the current topology and physical structure of the power grid.

The remainder of this dissertation is organized as follows. In Chapter II, we review the principle of operation of protective relays that are currently used in electric power grids. Both conventional relays and modern digital relays are considered. We also review several proposed solutions for enhancing protection in the power grid. These solutions are not widely utilized yet, but they suggest possible options one can take to improve the protection systems. In Chapter III, a novel approach using hypothesis testing and support vector machines based smart relays is introduced. Simulations are reported, illustrating the performance of SVM-based smart relays and the scalability issues associated with them. As a comparison between SVM-based smart relays and conventional distance relays, Chapter IV presents the results of a real case study and the benchmarking performance during cascading failures. The case study shows that SVM-based smart relays can deal with complicated situations which conventional relays cannot. The benchmarking performance demonstrates that the implementation of SVM-based smart relays can mitigate the cascading failures. Chapter V proposes the methodologies to deal with missing data and updating problems in implementing SVM-based smart relays. An attempt to utilize SVMs with transient state measures to build system-wide protection is showed on Chapter VI. Finally, Chapter VIII provides a detailed discussion on the principle of operation of the proposed solution and the implications of using the proposed SVM-based relays. The main contributions and future works of this dissertation are also presented in this Chapter.
Chapter 2

Background

2.1 Blackouts

A blackout refers to the total loss of power to an area and is the most severe form of power outage that can occur. Blackouts which result from or result in power stations tripping are particularly difficult to recover quickly. Outages may last from a few minutes to a few weeks depending on the nature of the blackout and the configuration of the electrical network [5]. Despite efforts to mitigate blackout risk, the data available from the North American Electric Reliability Council (NERC) indicate that the frequency of large blackouts in the United States is not decreasing. Blackouts come with substantial direct economic and non-financial losses. In the August 2003 blackout, for example, it is estimated that there were $3 billion in insurance claims [3]. It also resulted in significant other problems, such as subway passengers stranded underground and emergency vehicles stalled in traffic due to failed traffic lights. The social cost of a blackout are a function of many factors including the size of the blackout, the duration of the blackout, its location and the time of day. Clearly, blackout costs increase with both the geographic extent of the event and the amount of energy that is left unserved as a result of the grid failure.

Table 2.1 summarizes the most notable blackouts in North America and Europe. The initial disturbances in the event series of blackouts range from variety of triggering events including natural disasters, human error, and mechanical failure. However, the trigger of hidden failures is always a critical event during each blackout. In most of the blackouts, the critical events caused by hidden failures are considered as the real reasons for the cascading failures or the turning points on which the propagation of cascading failures are
<table>
<thead>
<tr>
<th>Date</th>
<th>Location</th>
<th>Size</th>
<th>Initial Reason</th>
<th>Critical Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>11-9-1965</td>
<td>Northeast US</td>
<td>30,000,000 people</td>
<td>Maintenance person incorrectly set a protective relay to trip too low on one of the transmission lines between the Niagara generating station Sir Adam Beck Station No. 2 in Queenston, Ontario.</td>
<td>Two generators, with no outlet for their power (the transmission line connected with them have been tripped), were automatically shut down to prevent damage. Within five minutes the power distribution system in the northeast was in chaos as the effects of overloads and loss of generating capacity cascaded through the network.</td>
</tr>
<tr>
<td>07-02-1996</td>
<td>Western North-America</td>
<td>55,000,000 people</td>
<td>3:42pm, a 500kV line sagged into a tree. 3:47pm, another line shorted out. 3:48pm, the 13 turbines at McNary Dam, tripped off line.</td>
<td></td>
</tr>
<tr>
<td>03-11-1999</td>
<td>Southern Brazil blackout</td>
<td>75,000,000-97,000,000 people</td>
<td>Lightning strike occurred at an electricity substation in Bauru, causing most of the 440kV circuits at the substation to trip.</td>
<td>A lot of generators automatically shut down because they did not have any load (the 440kV circuits tripped are the routes connected these generators to loads). The transmission lines connected to other plants to the rest of the system could not take the load and tripped too.</td>
</tr>
<tr>
<td>08-14-2003</td>
<td>Northeast US, Canada</td>
<td>50,000,000 people/57669 MW</td>
<td>2:02 p.m. The first 345kV transmission line sagged into a tree and initiated the blackouts.</td>
<td>256 power plants are off-line, 85% of which went off-line after 4pm, most due to the action of automatic protective controls.</td>
</tr>
<tr>
<td>09-28-2003</td>
<td>Italy</td>
<td>56,000,000 people</td>
<td>A transmission line was damaged by storm.</td>
<td>2 400kV power lines tripped due to sudden increased demand from those two power lines.</td>
</tr>
<tr>
<td>08-18-2005</td>
<td>Java, Bali</td>
<td>100,000,000 people</td>
<td>A 500kV transmission line failed at 10:23 am local time; this led to a cascading failure that shut down two units of the Paitan plant in East Java and six units of the Suralaya plant in West Java.</td>
<td></td>
</tr>
<tr>
<td>11-4-2006</td>
<td>German, France</td>
<td>15,000,000 households</td>
<td>Operators disconnected a double circuit line over the river Ems to allow a cruise ship to pass through safely. This caused other parts to be overload triggered.</td>
<td></td>
</tr>
<tr>
<td>11-10-2009</td>
<td>Brazil and Paraguay</td>
<td>60,000,000 people</td>
<td>Heavy rains and strong winds caused three transformers on a key high-voltage transmission line to short circuit, cutting the line and automatically causing the complete loss of 14 GW of power and the shutdown of the Itaipu Dam for the first time in its 25-year history.</td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1: Some notable cascading failures in North America and Europe
Hidden failures in protection systems are defined as a permanent defect that will cause a relay or a relay system to incorrectly and inappropriately remove a circuit element(s) as a direct consequence of another switching event [6]. As conveyed by the definition, hidden failures remain dormant until a particular event causes its manifestation and associated relay malfunction. The modes of hidden failures are a function of the relay type and it is closely associated with the relay’s functional logic. Although the overall probability of a protective relay having a hidden failure is relatively small, a hidden failure is always associated with intrinsic high risk. Especially, hidden failures are prone to manifest themselves under stress conditions and therefore their consequences tend to be rather noteworthy. Prevailing system conditions like overloaded lines, voltage dips, and overloaded generators, which are likely to happen during or after initial disturbances, boost the probability of hidden failures.

NERC outages report indicates that hidden failure are involved in over 70% of the cascading failures. The great blackout in 1965 was initiated by a hidden failure in a distance relay, whose setting was outdated. The setting of the relay was based on a typical load in 1963; however, the line loading increased during the following two years. In 1965, it finally reached the tripping setting of the relay and initiated the cascading event which left 30 million people in dark. The report in [3] shows that there are a total of 14 inappropriate relay trippings during the major blackout in August 2003.

In the last 10 years, significant research effort has been reported in developing technology to detect and prevent hidden failures. However, non-decreasing frequency of blackouts has proven that the effects of the proposed solutions in today’s power protective systems are not sufficient. The major blackouts in North America and Europe provide evidence that fundamental weaknesses exist in today’s power transmission infrastructures, especially the transmission protection systems. Novel functional logic of protective relays will therefore be needed to prevent such hidden failures and protect the system from cascading failures.

2.2 **Current Practice of Protective Relays**

With the increasing dependence of human population on a consistent supply of electricity, the need for an acceptable level of reliability and security of service has become crucial to both the suppliers and the customers. Hence, the protection system has become a critical part of the entire electric power system. Its shortcomings and concrete ways to enhance them must therefore be carefully considered.
Figure 2.1: Basic Structure of Power Systems and their Protection Systems. Observe that protection systems exist in all three components (generation, transmission, and distribution) of the electrical power system.

Fig. 2.1 shows a generic picture of an electric power system with its key components which include generation, transmission, and distribution. It is important to note here that power plants, high voltage transmission lines, transformers, distribution lines, etc., all have their own protection systems.

To start with, protective relays are the basic safeguard against faults in power systems. Their objective is to protect hardware from immediate damage by disconnecting it temporarily from the rest of the system. A protective relay must be reliable, fast, selective, and cost-efficient. Reliability has two main functions: security and dependability. The protective relay design requires intelligent decisions concerning the tradeoff between the security and the dependability of the relay. Dependability is a measure of the relay’s ability to correctly clear a fault. Security is a measure of the relay’s ability not to trip incorrectly. In the design of relays high dependability is typically considered to be more important for protection systems than security. There are some tradeoffs between these two elements. One element is often ensured at the expense of the other, and this leads to inherent biases in relay design. This design bias is one of the major reasons which cause the malfunctioning of relays in modern power systems, which may lead to a cascading failure of a larger section of the network, as observed in the August 2003 blackout in the USA.
CHAPTER 2. BACKGROUND

**Conventional Relays**

In transmission systems (see Fig. 2.1), distance relays and over-current relays are the commonly used conventional relays [7]. Their principle of operation is straightforward and easy to understand. They worked well with the electro-mechanical relays that were initially deployed and more recently have been merged with digital relays that are currently used.

**Over-current relays**

Over-current relays are widely used in the protection of power systems. Typically, three protection zones around a fault are used in order to protect a section of a line and to provide back-up protection to remote sections (see Fig. 2.2(a)). To achieve coordination between over-current relays in different zones, a characteristic curve is used in over-current relay to decide how long a relay should wait before tripping a line (see Fig. 2.2(b)). Typically, a over-current relay trips its breaker within 8 cycles (around 100ms) when a fault is detected in its zone 1. It waits another half to one second to trip break when a fault is detected in zone 2. The smaller the current, the longer the tripping time as this is intended to ensure that the line under protection does not trip in response to far away faults.

The over-current relay has only one threshold based on the magnitude of current. There are several assumptions in deciding the threshold, which include the assumptions that: (i) the normal current magnitude is always smaller than short-circuit current; and (ii) fault which is closer to the relay causes higher fault current. These assumptions work well and are adequate in situations where a fault is followed with a higher current. However, they can be grossly invalid in today’s complex power transmission systems and may lead
to malfunctioning of a relay. We will examine this issue in the next sections.

**Distance relays**

Distance relays are another type of commonly used protective relays on transmission lines. Distance relay responds to the impedance seen by the relay when the fault occurs. The R-X diagram is an indispensable tool for describing and analyzing a distance relay characteristic (see Fig. 2.3(a)). R and X in the diagram are the real and imaginary part of the apparent impedance, which is equal to the ratio between voltage and current. The zones of distance relays are shown in Fig. 2.3(b), the desired zone of protection is shown with a dotted line. However, to prevent a relay from tripping in response to a fault beyond its zone (over-reaching), the real zones are shown in solid lines [7]. Usually, for a distance relay, zone 1 covers 75% - 85% of the length of the local line from the location the relay is installed; zone 2 is defined as 120% of the length of the local line; and zone 3 is defined as the total length of the local line plus 150% of the length of the longest adjacent line. Therefore, a fault beyond 85% of the local line will be treated as zone 2 even if it is located on the protected line. When this happens, the fault will be isolated at different times by the relays on different terminals of the line.

Similar assumptions as for over-current relays are made for distance relays, which are not always valid and can cause the malfunctioning of relays.

Consider the 2 bus system with the distance relay shown in Fig. 2.4. Let us assume that bus 1 is the generator bus, the voltage on which has a magnitude of 1 unit and angle of 0 degree. As is well known, bus
2 which is connected to a load can be treated as the PQ bus. Therefore, the equation describing the power flow in this 2 bus system are:

\[
V_1 = 1; \\
I = \frac{V_1 - V_2}{Z_{\text{line}}}; \\
S_2 = V_2I^* = V_2\left(\frac{V_1^* - V_2^*}{Z^*}\right) = \frac{V_1V_2 - |V_2|^2}{Z^*} = P + jQ; \\
Z_{\text{apparent}} = \frac{V_1}{I}
\]

Obviously, the apparent impedance seen by the distance relay on the transmission line on bus 1 is a function of the load and generation level. Therefore, if the distance relay is designed based on the apparent impedance of a pre-determined load and generation level, it may malfunction when the system is in fact operating at a different load and generation level.

While conventional relays have been working well for the last 80 years, they have failed to meet the current requirements of electric power systems since the power systems are becoming increasingly complex. Because of these, the protection systems have experienced major changes over the past 30 years, mainly via the application of microprocessor relays. Microprocessors facilitate the use of communication, Internet technology, and artificial intelligence (AI) techniques in relays. Most recently, many utilities in the world have begun using or considering using digital relays for implementing substation monitoring, protection, and control functions. The major new features of digital relays are more accurate measurement and the
ability to communicate. We will introduce modern relays in the next subsection, such as differential relays and pilot relays, which require communications. From an unpublished survey of 20 major utilities in the US; out of 17 respondents, 7 have reported between 75% and 100% of relays are microprocessor-based, whereby 4 of these users have relatively new fleets comprising all or almost all microprocessor-based relays 6 or less years old. Six respondents have reported less than 30% microprocessor relays. Of these same 17 utilities, 5 have reported 50% or more electromechanical relays, 3 of the respondents have reported the use of large or predominant populations of analog solid state relays that are 30 years old. Generally speaking, about one third of the relays in the power systems are digital relays, and utilities are considering to replace some of the old relays with digital ones [8].

2.3 Modern Relays in Practice

Initially, protective relays were built using electro-mechanical technology. About 30 years ago, with the help of cheap and reliable microprocessors, the power industry started switching to digital technology. Digital relays can sample and process input signals at a sampling rate of 1 kHz or higher; hence, they can react to internal faults in a protected element within few tens of milliseconds or faster [9]. In contrast to electro-mechanical relays, they employ analog-to-digital (A/D) conversion of the incoming voltages and currents and Fourier transform concepts to analyze and extract the A/D converter outputs. They are capable of implementing advanced relay logic which is user-configurable. The modern versions of digital relays contain advanced metering and communication protocol ports, allowing the relays to become a focal point in the SCADA system.

Differential Relays

As described in the previous section, it is impossible to isolate a fault on a transmission line instantaneously from both ends by distance relays if the fault is located close to one terminal of the line. This is because of the time difference in isolation after detecting the fault, and the reason that there is a time difference is because the logic of the distance relay is designed to prevent overheating and underheating problems. In this situation, using a differential relay is one of the ideal techniques for fault detection and isolation. Differential relays are first used for protecting power transformers, generators, and busbars. The communication capability that exists in digital relays makes it possible to protect long transmission lines. In the differential
relays, the instantaneous values of currents or powers are compared at each terminal of a protected element. The differential relays are very sensitive to zone 1 faults because the differential signal indicates an internal fault. However, they cannot function as backup relays since they are not designed to detect the faults outside their primary protective areas. On the other hand, the differential relays are robust to the topology changes in the system because their logic is not dependent on the structure information, therefore they can function well during the change of topology of the power system (such as maintenance and blackouts) without modifying the settings. However, they can be affected by the saturation of current transformers, inrush and over-excitation phenomena in power transformers and a number of other phenomena.

**Pilot Relays**

Pilot relays are widely used in transmission line protection as a variant of the differential relays. Although differential relays have several strengths, such as they are sensitive to the detection of internal faults and they are not dependent on topology information, they are not very practical because they are more expensive than other relays. Moreover, to detect a fault, a differential relay depends on the current difference between two terminals of a line. The current difference can be caused by the inaccurate measurements of currents in the transformers and line capacitances. Such measurement errors may cause malfunction of differential relays in some cases [7]. Therefore, differential relays are always used to protect short lines (in fact, with the increasing use of digital relays and cables, differential relays are becoming popular in long transmission lines), while pilot relays are used in long transmission lines.

As previously mentioned, a pilot relay is a variant of a differential relay that was designed for circumventing the shortcomings of a differential relay. The term “pilot” refers to the communication channel between two ends of the transmission line. The communication media used for pilot relays are generally power line carrier, microwave, fiber optics, or cables. The communication media used by pilot relays in practice are shown in Table 2.2, based on a survey of 20 utilities [8].

<table>
<thead>
<tr>
<th>Channel type</th>
<th>PLC</th>
<th>Digital microwave</th>
<th>Analog microwave</th>
<th>Multiplexed fiber</th>
<th>Dedicated fiber</th>
<th>leased copper analog</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of utilities</td>
<td>13</td>
<td>15</td>
<td>3</td>
<td>15</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Used as dominant channel</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
There are several kinds of pilot schemes. Until recently, the most common pilot relaying scheme in the USA has been the directional comparison blocking scheme, using power line carrier. In this scheme, a distance relay can be used as a fault detector and it will transmit a blocking signal when it locates a fault outside its protected range. If a blocking signal is received, it will prevent the circuit breaker from tripping. Similarly, there are other schemes such as directional comparison unblocking scheme, direct transfer tripping (DTT), and permissive overreaching transfer trip (POTT). The latter two schemes are more viable protection systems since the communication channel is independent of the power lines [7].

Pilot relays are becoming common since they employ a new generation of digital relays and communication channels. A blocking mode is usually used when the communication channel is integrated with power lines and a tripping mode is used when they are independent from each other. In both of these modes, to make decisions, the relays utilize the information from the other terminal of the same transmission line. With the help of communications, the accuracy and performance of relays can be improved significantly.

Adaptive Relays

Adaptive relays are not a new concept in power protective relays and they were first proposed by S.H. Horowitz, A.G. Phadke, and J.S. Thorp as the ability of relay to change their settings, operation, or logic to adapt to prevailing system conditions[10]. Although the concept of adaptive relaying has been arisen for decades, there are only few schemes have been implemented. The most important reason for this is that the conventional relays react in a predetermined setting and cannot be updated when the system’s conditions change. All protective relaying decisions involve a fundamental compromise. One of the most difficult compromises to cope with is that of reliability. As explained in the first section, reliability has two aspects, dependability and security. The conventional relay design with a bias toward dependability is one of the factors which caused malfunctioning of conventional relays. The dependability of the conventional relays make the protective relays react without considering the condition of the whole system after a fault is detected. Their fault detection capability is based on their local voltage and current and therefore may become unreliable under abnormal stress conditions. In other words, the assumptions hidden in the relay settings do not hold in extreme cases, which lead to hidden failures and inappropriate trippings. Therefore, an adaptive relay is needed to achieve the compromise required and optimize system performance.

Although the concept of adaptive relays has been proposed long ago, it has not been widely implemented until the introduction of the new generation of digital relays. Most of the modern digital relays can keep up to
eight different settings corresponding to different conditions. Even in a single set of settings, modifications have been made to make the relays adaptive. For example, Fig. 2.5 is a typical R-X diagram of a digital distance relay. Compared with the conventional R-X diagram, it has a shaded area which indicates the heavy load condition. Although multiple settings help protective relays to be more adaptive and robust, these setting are still pre-determined and recorded in the relays, which can only be triggered and changed by supervisory commands. The relay itself cannot automatically update the setting during the rapidly changing conditions of the power grid during a blackout. There will always be cases which are not considered during the setting of the relays. Therefore, if the relay cannot “learn” and update its settings by itself, it can clearly fail under some unusual circumstances.

**Phasor Measurement Units (PMUs) for Modern Digital Relays**

While most of relays still only use magnitudes of voltage and current measurements, a new technology is available for accurately measuring voltage phase angles (phasors). These measurements could offer new information that can be used to improve the functional logic of protective relays.

The idea of phasor measurement was introduced after the blackout of 1965 in North-East US. The
first prototype phasor measurement unit (PMU) is developed by a Virginia Tech research team in 1988 [12]. However, the PMUs were not widely used in power systems until recently. This has changed after the introduction of the new generation of digital relays. PMUs measure the positive sequence voltage at two substations separated by hundreds of miles which are synchronized precisely with the aid of a GPS satellite system (see Fig. 2.6). The precise time-tags are attached with samples, and this information is exchanged over communication channels and collected by control centers and/or substations. By extracting the relevant information from these measurements, phasor information can be obtained at any node where PMUs are installed in the power grid. This can be used to do more accurate state estimation, control, and protection. Most modern digital relays have PMU integrated as their standard component. In these relays, besides current and voltage, phasor information has become an important measurement in decision making.

Problems with Current Protective Relays

Conventional relays, which are typically over-current relays and distance relays, have been known to malfunction in today’s complex power grids. Digital relays are recently being used more widely in power systems as they are capable of utilizing more complex logic with microprocessors and most of them are integrated with communication ports. Some modern digital relays even have a PMU module. All of these improve the decision making capability and performance of protective relays and help them to form a reliable and robust protection system. However, they certainly do not solve all the aforementioned problems in the protection systems; in fact, they introduce new ones. First, as mentioned before, most of the new protec-
tive schemes are based on a reliable communication channel, which cannot always be guaranteed in practice.
For example, when there is a communication failure, most differential relays and pilot relays cannot function well during the abnormal conditions. However, during blackouts or cascading failures in the power grid, as system conditions change significantly and rapidly, more information exchanges may be required by the control centers and substations. In other words, the communication channels are operating with high load and therefore become more vulnerable when the power grid is in contingent conditions. Thus, relying on the communication channel for decision making may not be the optimal solution for protective relays, although it might be beneficial to have information exchange. Moreover, as explained in the section on digital relays, these new protective schemes are built based on the conventional logic employed by protective relaying, which includes over-current and distance relay principles. This means that the digital relays are also based on the assumptions made for conventional relays, which are clearly invalid sometimes. Therefore, without changing the basic decision making logic of protective relays, the malfunctioning of protective relays cannot be avoided. At the same time, most of the solutions given in the previous parts of this section attempt to enhance the performance of the relays when a zone 1 or internal fault occurs. Compared to this improvement reported for primary relays, the performance of the relay’s backup function has not improved much. For example, the highly accurate differential relays cannot function as backup relays because they are not capable of detecting outside faults beyond zone 1. The pilot relays can function as backup relays because they are basically directional distance or over-current relays. However, all the schemes which are designed for pilot relays (e.g., DTT and POTT) cannot improve the capability of relays as backup relays. Thus, a new and more comprehensive logic is needed in protective relays, and the application of digital relays makes it possible to utilize new logic which was not possible before.

2.4 Approaches to Smart Protective Relays

The development of relays has a long history from just introducing the digital technology in the late seventies and early eighties, to progressing further with the development of a number of digital solutions for relaying in late eighties, and finally coming up with a variety of mature products and systems during the nineties [13]. This trend offers the opportunity to apply new technologies and mechanisms to extend and enhance the functions of relays. In current systems, the protective relays are functioning as monitoring and control devices in addition to performing protection.
Since the late 1990s, power systems have been pushed closer to their limits, resulting in a growing risk for a local failure to propagate and develop into a cascade of failures which may result in a large-scale catastrophic blackout. Under these conditions, power system as a whole needs to be well monitored, controlled, and protected. To this end, communication and information exchange have been used in the power systems and in this setting protective relays can function as distributed sensors in the system with no installation cost and low communication cost. To accomplish this goal, a protective relay must be accurate, adaptive, and fast. Therefore, many new technologies, such as artificial neural networks (ANN), decision tree, and fuzzy logic (FZL), are proposed to be embedded into digital relays. The recent developments in substation automation, which has started around the early seventies, might facilitate such cost-effective systems.

**Challenges of Protective Relay Design**

The current trends in power system planning utilize tight operating margins with less redundancy, because of new constraints placed by economical and environmental factors. At the same time, addition of non-utility generators and independent power producers, the increase in interchanges, an increasingly competitive environment and introduction of the flexible alternating current transmission system (FACTS) devices make the power system more complex to operate and to control and, thus, more vulnerable to a sudden major disturbance [14]. Therefore, there are several challenges to the design of accurate and robust protection systems. Below, we briefly discuss some of these challenges.

- **High-impedance faults**

  High-impedance faults (HIFs) usually occur when a conductor touches the branches of a tree having a high impedance or when a broken conductor touches the ground. This type of fault, in general, is difficult to detect through conventional protective relays such as distance or over-current relays. This is mainly due to relay insensitivity to the very low level fault currents and/or limitations on other relay settings imposed by high-impedance faults (HIFs). In the case of an over-current relay, the low levels of current associated with HIF are not high enough to trigger the relay. In the case of a distance relay, which relies on the estimation of impedance to faults based on the measured voltages and currents, the accuracy of the estimation can be significantly affected by the high-impedance fault.

- **Meshed network structure/parallel lines**
As modern power grids are getting increasingly more complex, especially power delivery systems are becoming more interconnected. To meet the reliability requirements, parallel lines are connected between substations so that there is a backup path to deliver power when one line goes out. Parallel lines and meshed network structure bring circuit loops into the system. A fault located within a loop may change the direction of the current, instead of only changing the magnitude of the current. This may cause a big problem for relays which have direction sensitive components, such as directional distance relays, directional over-current relays, and the pilot relays which utilize these basic relays as a fault detector. They may malfunction because of the change of current directions.

- Change of topology
  Different network topologies require different protection schemes. The settings of most relays are decided based on the current topology of systems and a conventional relay’s setting cannot be updated automatically. Therefore, a relay may malfunction during the changing of topology, when equipment goes off-line or re-connects on-line again. For example, an isolation of a certain equipment in the power grid can cause a significant change in the magnitude of current or even in the direction of current on other transmission lines. If the protective relays on those lines do not change their settings accordingly, they may misclassify between (N-1) condition as a fault condition. Relays must be adaptive and they should be updated automatically to handle this problem.

- Distributed generators
  Most recently there has been a growing trend toward deploying distributed generators (DGs) in the power grid. Distributed generators are usually small rating, privately owned generators in distribution systems. They are usually not centrally dispatched. Moreover, they were not considered when the local grid was planned. Therefore, effect of distributed generation was not considered in the settings of the relays. The overall problem when integrating DGs in existing networks is that their connection to network can change the node voltage’s magnitude, line current’s magnitude, or even the direction of line current. All of these can and will further confuse the relays when they are making decisions. An example is given in Fig. 2.7.

- Coordination between relays
  Conventional relay settings are based on network topology and coordination between primary and backup relays, which means a correct setting of a relay has to ensure that it trips for local faults but
Figure 2.7: Voltage profile and gradient with and without contribution of distributed generators after [15]. Here $T$ denotes the transformer, $G$ denotes the distributed generator, $U$ denotes voltage, and $x$ denotes the distance from bus 1. The voltage along this transmission line is a monotonically decreasing function of the distance $x$ if there is no DG. However, the connection of DG can change the voltage, and hence the current direction on the line.
stays closed for far away faults. In some cases, especially in meshed networks, it is hard to find a suitable setting for a protective relay which can make it selective, dependable, and secure. In some extreme cases, there is not even a way to find a set of settings to meet all these requirements. The main reason for this is the fact that today’s protective relays are mainly based on the magnitude of current and voltage. Their boundaries are decided based on the comparison of currents and voltages, and sometimes phasor measurements. We will give a detailed example in the next section to show that a non-linear decision boundary based on more measurements can overcome this problem.

Although the challenges which are caused by the meshed networks and high-impedance faults have been better addressed by modern relays, the problems associated with distributed generators and changing of topologies remain unsolved.

**Improvement of Relay Logic**

The most straightforward approach to building accurate and robust protective relaying systems is to make individual relays more intelligent and flexible. This translates into the introduction of new designs that can accommodate more functions that are capable of not only accommodating new capabilities, but also making the implementation flexible [13]. The implementation of accurate algorithms using the data from both terminals of the transmission lines to locate faults, the application of intelligent systems such as ANN and FZL are all solutions proposed in this category.

- Fault location estimation by two terminal measurements

  Following the occurrence of a fault, a power utility strives to restore power as quickly as possible. Therefore, a good algorithm that can provide an accurate estimation of the fault location is very important. Novosel et al. proposed a novel unsynchronized two-terminal fault location estimation algorithm in [16][17], for which synchronized measurements are not required, instead of using only one terminal measurements or synchronized two-terminal measurements [18][19][20]. This new method estimates fault location on a transmission line with the voltage and current phasors from both terminals of the protected line, but they do not have to be synchronized. The unknown synchronization angle can be solved by an iterative Newton-Raphson method, after which the fault location can be estimated. This algorithm does not make assumptions on the source impedance and the loads. Thus, generally speaking, this is an accurate algorithm for estimating the fault location with unsynchronized measurements
CHAPTER 2. BACKGROUND

from both ends of the transmission line. However, this mechanism relies on the communication between two terminals of the line. In addition, the convergence time of the iterative Newton-Raphson method can be as long as minutes in extreme cases. Therefore, this algorithm is suitable for estimating the fault location in the post-fault period to help the engineers restore the service. However, it may not work well in protecting a power system during cascading failures, when the communication channel and the substation computers could be compromised because of the heavy load they are carrying.

• Application of intelligent systems

It has been recognized by several researchers that the conventional relays are getting worse when handling some special cases in today's complex power grids. Therefore, new techniques are proposed to overcome the shortcomings of conventional relays. Although these techniques, including ANN, FLZ, and decision tree, have not been proven to be reliable and robust enough to be widely used or replace the current relays, they are pointing to a new direction which can help to improve the performance of protection systems.

Neural networks have been used to solve fault analysis problems since 1989. The major advantage of neural networks approach is that it can take into account several features of the input signals simultaneously and compare the patterns according to their mutual similarity instead of hard thresholds. It works well in terms of fault detection, classification, location and zone determination [21], most of which are applications related to improvements in the distance relaying principle. For fault detection, most neural network studies report a higher than 90% accuracy. Besides these real-time applications, at the same time, neural networks also work well in off-line applications such as post-mortem study of fault events recorded with digital fault recorders (DFRs). More details can be found in [21]-[24]. Although NN works successfully in protective relays in most cases, they have inherent shortcomings. The output of NN always falls into the local optimum, instead of the global optimum. This means that the accuracy of NN is highly dependent on the initial setting when tuning the parameters. Second, without data pruning, NN easily overfits the input data. In addition, the initialization of the NN classifier can be very slow and time consuming.

In addition to improving the decision making capability of individual relays, E. Bernabeu and J. Thorp proposed a new voting mechanism based on the application of decision tree among several relays [25]. In this mechanism, the logic of individual relays is not changed at all; however, a voting mechanism
will be held among three relays after they make individual decisions and share measurements with each other. This decision making mechanism has been experimented in a power grid in California and was shown to work well with excellent performance. However, this mechanism needs information exchange which depends on communication channels. The decision tree voting process is intuitive but for some occasions it may be too simple to handle complex situations.

Other proposed solutions include FLZ [26], expert systems [27] and new AI methods in protection systems, but most of these are used to improve the coordination between relays instead of improving the individual relays.

Substation Automation

The development of substation automation systems dates back to the early seventies when the first all-digital substation monitoring control and protection systems were proposed. However, it was not widely utilized for developing cost-effective systems until late eighties when the communication and processing technology became mature enough to support power systems [13]. The advent of industry deregulation has placed greater emphasis on the availability of information, the analysis of this information, and the subsequent decision-making to optimize system operation in a competitive environment. Intelligent electronic devices (IEDs) being implemented in substations today contain valuable information, both operational and non-operational, needed by many user groups within the utility. Fig. 2.8 is a typical Local Area Network (LAN) structure in a substation. Most IEDs are dedicated to a given function or a set of functions, which include fault locating, digital fault recording, and digital protective relays. The future trend will be to make the individual IEDs more compatible and even universal. These IEDs are metering and monitoring the system and report their measurements and records to control centers. The supervisory control and data acquisition (SCADA) system can estimate the condition of the system, do fault analysis, and plan maintenance by utilizing these different pieces of information. The automation of substation also helps the protective relays to be robust and more reliable.

In summary, new mechanisms and technologies have been proposed to enhance the capability of individual relays and protection systems as well. Although these mechanisms are not ready to be widely implemented in today’s power grid, they offer valuable information and options for improving the protective relays. Moreover, the mature nature of the communication systems and substation automation in power grids offer a good opportunity for integrating new algorithms with digital relays.
Figure 2.8: Typical LAN structure in substation automation after [28]. The IEDs in this figure are primary and backup protection, and the HMI stands for Human Machine Interface.

It is important to note, however, that all of the proposed solutions mentioned in this section have their own limitations. Most of them are add-ons to conventional protective principles, which do not change the protective logic of conventional relays. They can therefore enhance the performance of relays but cannot solve the blackout problem completely. Since we have the ability to obtain more physical measurements from the system, it is possible to build a completely new decision logic by utilizing all of these different pieces of information. To date, little research has been done or reported on improving the inherent logic of individual relays. The following section attempts to bridge this gap by introducing new methods for improving the relays’ logic for making decisions and for providing rapid reactions. In particular, we propose hypothesis testing and support vector machine (SVM) techniques for improving the performance of individual protective relays in the following sections.

**Special Protection Schemes and Wide Area System**

To achieve adaptivity in protection systems, besides the improvement on single relays, there are also proposed adaptive schemes based on group of relays. The most famous among these schemes is Special Protection Schemes (SPS) based on Wide Area Measurement (WAMs).

Power systems have originally arisen as individual self-sufficient units, where the power production need
CHAPTER 2. BACKGROUND

to be matched with the consumption all the time. Today’s power systems’ size and complexity have grown significantly to satisfy increasingly larger power demand. Some new phenomena have thus emerged which endanger the normal operation of power systems. They are [29]:

- **Frequency instability**: A power system cannot maintain the steady frequency within the operation limits.

- **Voltage instability**: A power system cannot maintain the steady voltage level at each bus in abnormal operating conditions during and/or after disturbances. A systems is voltage unstable if the voltage magnitude for at least one bus in the system decreases as the reactive power injection in the same bus is increased [30].

- **Transient angular instability**: A power system cannot maintain synchronism when subjected to a severe transient disturbance.

- **Power oscillations**: A power system cannot maintain synchronism under small disturbances. Such disturbances can occur continually on the system because of small variations in loads and generations.

These phenomena always happen in a global scale instead of happening only in local area, therefore, it is important to obtain and use information from remote locations. These data, combined with the local data, can be utilized for wide area monitoring, protection, and control. These schemes are often referred to as Special Protection Schemes (SPS). According to [31] they are defined as “a protection scheme that is designed to detect a particular system condition that is known to cause unusual stress to the power system and to take some type of predetermined action to counteract the observed condition in a controlled manner. In some cases, SPSs are designed to detect a system condition that is known to cause instability, overload, or voltage collapse. The action prescribed may require the opening of one or more lines, tripping of generators, ramping of HVDC power transfers, intentional shedding of load, or other actions that will alleviate the problem of concern. Common types of line or apparatus protection are not included in the scope of interest here.”

According to [30], there are five states of operating conditions, which are normal, alert, emergency, in extremis and restorative. Fig. 2.9 shows the basic schematic to provide the framework in which control strategies and operator actions are determined to deal effectively with each state. In case of highly reliable
SPS with a good performance, a normal power system operation could be shifted from the Normal state to Alert state. This confidence in SPS would allow much better utilization of existing assets.

Similar to the design criteria for conventional relays, there are four main design criteria for SPS [32]:

- **Dependability**: The certainty that the SPS operates as designed in all cases where emergency controls are required to avoid a system collapse.

- **Security**: The certainty that the SPS will not operate in the cases that controls are not required. This means the SPS will not take emergency controls unless they are necessary to avoid a system collapse.

- **Selectivity**: The ability to select the correct and minimum amount of action to perform the intended function, that is, to avoid using disruptive controls such as load shedding if they are not necessary to avoid a system collapse.

- **Robustness**: The ability of the SPS to provide dependability, security and selectivity over all possible dynamic and steady state operating conditions that may encounter.

Although there is a trend to commission more SPSs in the power system, there are problems which stay unsolved with SPSs. With the installation of SPSs in power systems, the degree of complexity is rapidly increasing, therefore, the impact of SPSs on daily operation of a power grid seems unclear. Moreover, all installed SPSs are dedicated solutions for particular power systems, there is no scheme that could be applied to another power system with minimal modifications. Therefore, to cope with huge changes in the power system, the SPSs need to be re-designed to avoid false or undesired reaction.

On the other hand, SPS is only one kind of functional applications of the wide area systems, which is becoming a popular option in today’s power grid. Wide area system may be a platform serving various purposes. It acquires data from both local and remote locations with Remote Terminal Units (RTUs), then communicates them into one central location - Supervisory Control and Data Acquisition (SCADA), where they can be processed and decisions can be made. The wide area systems include [29]:

- **Wide area monitoring**: the system can continuously offer accurate information (synchronized measurements with a high sampling rate) about the states which would not be observable by other ways, such as oscillations, load dynamics etc. (the displayed quantities may range from power flows, magnitudes and phase angles of voltages and currents to stability indicators)
Figure 2.9: Power system states and transactions between them [29]

- **Wide area protection:** SPS in the traditional, or conventional sense. In situations endangering the power system (detected incipient instability), SPS executes a single action.

- **Wide area control:** the system continuously, after the recognition of a state prone to instability, influences the behavior of power system to follow a certain trajectory to avoid instability and keep the power system within safe boundaries. A feedback control loop is employed to do so.

- **Wide area optimization:** It refers to the basically economic in nature and aims at the operation of the network in the most profitable way, such as minimization of losses and similar tasks which are usually done by Energy Management Systems.

Many schemes have been proposed based on wide area systems, while only a few of them have been deployed so far. In today’s systems, the deployment and advances in PMUs have greatly improved the performance of these wide area system functions and therefore make the utilization of them possible.
Chapter 3

Machine Learning based Smart Protective Relays

In this dissertation, we proposed smart protective relays which are based on machine learning mechanisms from computer science. First, we give a problem statement.

3.1 Problem Formulation

The North American electric power system is one of the great engineering achievements of the 20th century. Our modern society has come to depend on reliable electricity service as an essential resource for national security, health and welfare, communications, finance, transportation, food, water supply, heating, cooling, lighting, computers and electronics, commercial enterprise, and even entertainment and leisure; in short, electricity is essential in nearly all aspects of modern life [3]. The electricity has to be always available to all types of customers. Meeting this expectation is an enormously complex technical challenge even during normal operating conditions when the equipment status is as planned and the demand is at expected levels.

In addition to the overall design challenge for transmission systems and reliable equipment, correctly functioning protection schemes and relay settings are critical to the high quality electric power system operations and service. To start with, relays are installed in power systems to protect individual pieces of equipment, such as generators, transmission lines, transformers, and buses. Moreover, the security and stability of the electric power system as a whole are affected by each of the subsystems which is installed or
connected to it. Protection systems, as many other artificial systems, are not impeccable. While relying on protection systems for basic safety, the integrity of the electric power systems as a whole may also be affected in undesired ways by protection systems. According to historical data, protection relay malfunctioning, which is associated with hidden failures in protection systems, is one of the major contributing factors to about 70% of the major disturbances in the United States [2]. Similar findings were reported by an investigation team of experts who studied the blackout of August 2003. The experts pointed out that the relay malfunctioning was one of the key reasons for this serious event. The basic logic of a conventional protective relay can be invalid in today’s complex power transmission systems and trigger cascade of failures. As demonstrated by the blackout of 2003, some relays responded to overloads and N-1 conditions even though there were no faults on the lines they were protecting. This type of relay malfunctioning is the common mode of failure that accelerates a geographic spread of the cascading faults [3]. To reduce the damage on power systems hardware and the system as a whole, more intelligent relays are therefore needed.

3.2 Introduction to Machine Learning

Machine Learning is defined as the process of designing and developing of algorithms that allow computers to evolve behaviors based on empirical data [33]. In this dissertation, the major focus of machine learning is to enable the relays to automatically uncover the complex rules or patterns which are carried out on samples of measurements, and ultimately to learn to make intelligent decisions based on their physical measurements. The tasks of machine learning commonly involve four different classes [34]:

- **Classification**: Arranges the data into predefined groups. For example, a face detection program might attempt to detect human’s faces from background in pictures. Common algorithms include decision tree learning, nearest neighbor, naive Bayesian classification and neural networks.

- **Clustering**: Is similar to classification but the classes are not predefined, so the algorithm will try to group similar items together.

- **Regression**: Attempts to find a function which models the data with the least error.

- **Association rule learning**: Searches for relationships between variables.
CHAPTER 3. MACHINE LEARNING BASED SMART PROTECTIVE RELAYS

Most of the problems involved in power protection systems are in the category of Classification and Regression.

The field of machine learning is not associated with any specific algorithm or problems, but it is a conglomerate of methods that include decision trees, parameter associating, k-nearest neighbor, neural networks, cluster analysis, regression trees, bayesian networks, graphic models, etc. Some of these methods are already in use or have been proposed for use in power systems. For example, bayesian networks concepts have been applied for network validation in the power systems on the shipboard; decision trees have been proposed to establish the voting schemes between relays in contingency conditions [25].

The machine learning methods used in this dissertation are Hypothesis Testing and Support Vector Machines (SVM). The motivation is provided by the natural simplicity of these two mechanisms, their intuitive representation of knowledge discovery, and their outstanding classification accuracy.

3.3 Hypothesis Testing Based Smart Protective Relays

3.3.1 Introduction of Hypothesis Testing

Hypothesis testing is widely used in several disciplines such as statistical communication theory [35][36]. It is a general method for making decisions about accepting or rejecting a hypothesis. The hypothesis being tested is referred to as the null hypothesis and denoted by \( H_0 \). Rejection of the null hypothesis implies acceptance of its complement, which is referred to as the alternative hypothesis and is denoted by \( H_1 \) [37].

In the power protection system, we take the viewpoint that the normal condition of the power system can be presented by hypothesis \( H_0 \), and condition with fault by hypothesis \( H_1 \).

The conditional distributions of normal and fault current that will be used in the hypothesis testing conducted in this dissertation is shown in Fig. 3.1.

Usually, a threshold is used to discriminate \( H_0 \) and \( H_1 \). Where to place the threshold is an important decision and is usually decided by the performance criteria used for the discrimination. It can be seen that the main challenge in determining a threshold is in a situation where distribution of normal current overlaps the distribution of fault current. In such a case, the total probability of error is

\[
P(\text{error}) = P(H_0|H_1) \cdot P(H_1) + P(H_1|H_0) \cdot P(H_0)
\]  

(3.1)
Our goal is to minimize the probability of error (malfunctions), so the threshold should be chosen to minimize the result of equation (3.1).

### 3.3.2 Hypothesis Testing Based Relays

As introduced in Section 3.3.1, hypothesis testing is a general and well-established theory for deciding when to accept or reject a hypothesis. When applied to the logic of smarter over-current relays, hypothesis testing helps to discriminate the abnormal conditions from normal conditions based on the different current characteristics. In this thesis, the current characteristics used are the distributions of normal and fault current. Simulations were performed using the IEEE 14 bus systems [38]. To generate the distributions of normal and fault currents, the following assumptions were made:

1. The loads are random variables which have a uniform distribution in the range from 80% to 120% of peak value;

2. Loads are modeled as constant power sinks; and,

3. The generators always meet the load requirements.
These uncertainties were captured using Monte Carlo simulations, resulting in typical current distributions as shown in Fig. 3.2. Obviously, the distributions of the current magnitude in normal and abnormal conditions have several important properties that can be observed from Fig. 3.2:

- Normal current has a Gaussian distribution;
- There are several peaks in the distribution of fault current, which are similar in size and very scattered;
- Different peaks indicate different fault locations;
- The peak which is closest to normal current is caused by faults that happen in the most distant buses;
- The highest current is caused by the fault in zone 1;
- The lowest current is caused by the fault in generators’ buses.

By comparing normal and fault current distributions, thresholds are determined and stored in relays. Dual thresholds are used to differentiate fault conditions from normal conditions. Multiple thresholds can provide even more accurate differentiation, including the occurrence and the location of the fault.

Hypothesis testing is a viable and promising method for making relays more intelligent and adaptive when the faults are located close to buses. In complex situations, which involve faults in various locations including the faults in the middle of transmission lines, the accuracy of smart relays deteriorates. This is because the probability distributions of current tend to be flatter and have fatter tails in complex fault conditions and this causes larger errors in decision making.

Table 3.1 shows the accuracy of smart relays in systems with different sizes.
### 3.4 Support Vector Machine Based Protective Relays

Our preliminary research has shown that hypothesis testing is a viable and promising method for making relays more intelligent and adaptive when the faults are located in buses. In complex conditions which might involve several faults in various locations, however, more features are needed to be taken into consideration. To determine the hyperplane based on these features, we propose to use support vector machine classification technique in the next subsection.

#### 3.4.1 Introduction of Support Vector Machines (SVM)

In this section, we first overview of the SVM technique.

**Linear Support Vector Machines**

To start with the simplest case, we introduce linear SVM first.

**The separable case**

A classification task usually involves with training and testing data which consist of some data instances. Each instance in the training set contains one “target value” (class labels) and several “attributes” (features). The goal of SVM is to produce a model which predicts target value of data instances in the test set when one is only given the values of the attributes [39].

To get a basic understanding of SVM, first consider a linear binary classification task, as depicted in Figure 3.3. For this problem, SVM finds the separating hyperplane \((\bar{w} \cdot \bar{x} = 0)\) that maximizes the margin, denoting the distance between the hyperplane and closest data points (i.e., support vectors).

Suppose we have \(l\) observations. Each observation consists of a pair: a vector \(\bar{x}_i \in \mathbb{R}^n, i = 1, ..., l\) and the associated “truth” \(y_i\) which is given as trusted. Consider the two classes case, where \(y_i \in \{-1, 1\}\). Suppose we have some hyperplane which separates the positive from the negative examples. The points \(\bar{x}\) which lie on the hyperplane satisfy \(\bar{w} \cdot \bar{x} + b = 0\), where \(\bar{w}\) is perpendicular to the hyperplane and \(||\bar{w}|||\) is the Euclidean norm of \(\bar{w}\). \(|b|/||\bar{w}|||\) is the perpendicular distance from the hyperplane to the origin. Let

<table>
<thead>
<tr>
<th>Systems</th>
<th>3 Bus System</th>
<th>14 Bus System</th>
<th>30 Bus System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probability of Single Fault</td>
<td>0.01</td>
<td>0.08</td>
<td>0.01</td>
</tr>
<tr>
<td>Minimum Error Rate</td>
<td>0.0017</td>
<td>0.0106</td>
<td>0.0018</td>
</tr>
</tbody>
</table>
$d_+$ be the shortest distance from the separating hyperplane to the closest positive example, and $d_-$ be the shortest distance to the closest negative example. In the linearly separable case, the support vector algorithm simply looks for the separating hyperplane with the largest margin. This can be formulated as a quadratic optimization problem as follows:

Suppose that all the training data satisfy the following constraints:

\[
\vec{w}\vec{x}_i + b \geq +1 \text{ for } y_i = +1
\]
\[
\vec{w}\vec{x}_i + b \leq -1 \text{ for } y_i = -1
\]

which is equal to

\[
y_i(\vec{w}\vec{x}_i + b) - 1 \geq 0 \quad \forall i
\]

The points which satisfy the equality in the equation lie on the hyperplane $H_1 : \vec{w}\vec{x}_i + b = +1$ or $H_2 : \vec{w}\vec{x}_i + b = -1$. The points on the hyperplane $H_1$ has the perpendicular distance from the origin $|1-b|/||\vec{w}||$. The points on the hyperplane $H_2$ has the perpendicular distance from the origin $|-1-b|/||\vec{w}||$. 

Figure 3.3: The separating hyperplane that maximizes the margin. (‘o’ is a positive data point, i.e., $f(o') > 0$, and ‘•’ is a negative data point, i.e., $f(•') < 0$)
Hence \( d_+ = d_- = 1/||\vec{w}|| \) and the margin is \( 2/||\vec{w}|| \). Thus, one can find the hyperplane which gives the maximum margin by minimizing \( ||\vec{w}|| \), subject to the constraints in equation 3.2.

**Lagrangian formulation of SVM**

The linear-separable case has been proposed as a quadratic optimization problem in the previous section. However, it is not always easy to solve the optimization problem. Moreover, in practice, one might not always have linear-separable cases. To solve these two problems, one needs the Lagrangian formulation of SVMs. Thus, we introduce positive Lagrange multipliers \( a_i \), where \( i = 1, 2, ..., l \), one for each of the inequality constraints. This gives the Lagrangian equation as

\[
L_p = \frac{1}{2} ||\vec{w}||^2 - \sum_{i=1}^{l} a_i y_i (\vec{w} \cdot \vec{x}_i + b) + \sum_{i=1}^{l} a_i
\]  

(3.4)

Then we can minimize \( L_p \) with respect to \( \vec{w} \), \( b \) and require the derivatives of \( L_p \) with respect to all \( a_i \) vanish simultaneously. This is a convex quadratic problem, which means we can solve the dual problem for the primal problem. This requires that the gradient of \( L_p \) with respect to \( \vec{w} \) and \( b \) vanish; therefore,

\[
\vec{w} = \sum_i a_i y_i \vec{x}_i
\]

\[
\sum_i a_i y_i = 0
\]

(3.5)

We subject these to equation (3.4) to get

\[
L_D = \sum_i a_i - \frac{1}{2} \sum_{i,j} a_i a_j y_i y_j \vec{x}_i \cdot \vec{x}_j
\]

(3.6)

SVM therefore amounts to maximizing \( L_D \) with respect to \( a_i \) subject to the constraints in equation (3.5) and all the \( a_i \) are positive. As mentioned before, there is a \( a_i \) for the training point \( \vec{x}_i \). In the solution, those points whose \( a_i \) is non-zero are called “support vectors”, and they lie on one of the margins. All the other training points whose \( a_i \) equals to zero do not contribute to the location of the separating hyperplane and they lie outside the margins.

**The non-separable case**

The algorithm and equations stated in previous section are for separable data, when the training data are not linear separable, there will be no feasible solutions. To extend these ideas to handle non-separable data
in Fig. 3.4, we have to relax the constraints. To do this, one can introduce positive slack variables $\xi_i$, where $i = 1, ..., l$ into each constraint, which yields:

$$\begin{align*}
\vec{w} \vec{x}_i & \geq 1 - \xi_i \\
\vec{w} \vec{x}_i & \leq -1 + \xi_i \\
\xi_i & \geq 0 \quad \forall i
\end{align*}$$

Hence $\sum_i \xi_i$ is the upper bound of training errors, which could be plugged into the target function with a weight parameter as

$$||\vec{w}||^2 + \nu (\sum_i \xi_i)$$

where $\nu$ is a parameter to be chosen arbitrarily. Larger $\nu$ corresponding to assigning higher penalty to errors, while smaller $\nu$ gives more tolerance to errors. The dual problem at the same time becomes:

$$Max: \quad L_D = \sum_i a_i - \frac{1}{2} \sum_{i,j} a_i a_j y_i y_j \vec{x}_i \vec{x}_j$$
\[ s.t.: 0 \leq a_i \leq \nu \]
\[ \sum_i a_i y_i = 0 \] (3.10)

**Nonlinear Support Vector Machines**

To deal with the linear non-separable case, besides relaxing the constraints with a slack variable, we can also generalize to the case where the separating hyperplane is a non-linear function. This can be done by using a “trick” on the dual problem of the SVM classification. To generate a non-linear hyperplane, intuitively, one can map the training data to some other higher dimensional Euclidean space \( H \) by using a mapping function \( \Phi \):

\[ \Phi : \mathbb{R}^d \rightarrow H \] (3.11)

In the dual form of the SVM problem, the training data only appear in the form of inner products. Therefore, the training algorithm will depend also on the inner product in the space \( H \), in the form of \( \Phi(x_i)\Phi(x_j) \). If we replace all the \( x_i x_j \) by \( \Phi(x_i)\Phi(x_j) \), this implies that we project all the training data to higher dimensional space and solve the optimization problem there. Hence, to generate the non-linear separating hyperplane to deal with nonlinear separable problem, one can use the same optimization function in equation (3.10) with all the \( x_i x_j \) replaced by \( \Phi(x_i)\Phi(x_j) \). We call \( K(x_i, x_j) = \Phi(x_i)\Phi(x_j) \) a kernel function. There are four commonly used kernels in SVM, which are

- **Linear**: \( K(x_i, x_j) = (x_i)^T (x_j) \)
- **Polynomial**: \( K(x_i, x_j) = (\gamma(x_i)^T (x_j) + r)^d, \gamma > 0 \)
- **Radial basis function (RBF)**: \( K(x_i, x_j) = exp(-\gamma ||(x_i) - (x_j)||^2), \gamma > 0 \)
- **Sigmoid**: \( K(x_i, x_j) = tanh(\gamma(x_i)^T (x_j) + r) \)

where \( T \) stands for the transpose operator.

We use RBF kernel in this dissertation, since it non-linearly maps samples into a higher dimensional space. Therefore, compared with a linear kernel, it can handle the case when the class labels and features are coupled in a non-linear manner. Moreover, the linear kernel is a special case of RBF kernel, thus the RBF kernel can handle a classification task at least as well as a linear kernel. On the other hand, the RBF kernel has less hyper parameters than the polynomial and sigmoid kernels. The number of hyper parameters
influence the complexity of model selection. To summarize, RBF kernel is the one with the least numerical difficulties but with the highest probability to determine an optimal hyperplane. Therefore, RBF kernel will be used in this dissertation.

Besides the kernel selection, the SVM model selection, which means the chosen of parameters $\gamma$ and $\nu$, is very important to the accuracy of the SVM classifier. In any predictive learning task, such as the classification task in this dissertation, a well chosen parameter is the first step for obtaining a high level of performance of the learning machine. The physical meanings of the parameters $\gamma$ and $\nu$ are:

- the regularization parameter $\nu$ determines the trade-off between minimizing the training error and maximizing the margin;
- parameter $\gamma$ in the RBF kernel function implicitly defines the nonlinear mapping from input space to some high-dimensional feature space

These two parameters are determined by using grid search, which will be stated in detail in the simulation subsection of this Chapter.

Theoretically, the mechanism which is described above is designed for binary two-class problems. However, it can be extended to multi-class problems. Basically, there are two ways of extending this binary two-class classification mechanism to multi-class classification: 1) modify the model of SVM and find the optimal boundary under modified model to fit the multi-class problem; 2) combine several binary classifiers. We use the second mechanism in this dissertation. In this mechanism, a set of binary classifiers are determined in a way that each of them is trained by the data from two classes. Then majority vote is taken among these binary classifiers to establish an optimal classifier to predict the multi-class problems.

3.4.2 Proposed SVM-based Relays

To improve the accuracy of smart relays in complex conditions, more than one feature is taken into consideration at a given time. Besides the magnitude of current, which can be selected to be the representative feature for hypothesis testing based smart relays, phase of current, magnitude of voltage, phase of voltage, real power, and reactive power are all good candidate features for SVM-based smart relays. To get the sample values of these features during normal and abnormal conditions, simulations were performed in Matlab.
Recall from Chapter II, example in Fig. 2.4, that the conventional relay settings for zone1 through 3 are based on assuming a given network topology and generation and load level at each node. Based on this assumption a deterministic circuit analysis is done to compute the short-circuit current, voltage, and the corresponding apparent impedance. The knowledge about $R$ and $X$ is key to compute and set the tripping logic of each relay, as shown earlier in Fig. 2.3. However, as the network topology and generation and load levels change, in real operations the apparent impedance will change. Since the relay logic does not take into consideration these changes, relay malfunctioning can occur. In particular, this could result in false alarms by distance relays, which in turn could lead to cascading failures [3].

To circumvent this problem, we have proposed in [40] an SVM approach for defining the boundaries between normal, zone 1, 2, and 3 fault conditions in higher than two dimensional spaces (for example, $R$ and $X$ in distance relays and magnitude of current in over-current relays). The boundaries are obtained by extensive off-line training using many network topologies and load and generation levels. A simple illustration is given in Fig. 3.5 in two dimensional space with the magnitude of voltage and current. The solid dots are the measurements obtained from faulty conditions and the empty circles are the measurements.
obtained from the normal conditions. The measurements are called features in SVM training, and their corresponding situations in the system are known as classes or class labels. By training the classifier with these features, \(|V| and |I|\), we want to determine the parameters of the boundary described by the equation 
\[
\omega \cdot \vec{x} - b = 0
\]
with the positive margin of 
\[
\omega \cdot \vec{x} - b = 1
\]
and negative margin of 
\[
\omega \cdot \vec{x} - b = -1
\]
where \(\omega\) and \(\vec{x}\) are vectors, \(b\) is a scalar, and \(\vec{x} = [|V|, |I|]\), so that all the data from the normal condition class are located on the right hand side of the negative margin and the data from the fault condition class are located on the left hand side of the positive margin. The parameters \(\omega\) and \(b\) can be determined by solving the quadratic optimization problem:

\[
\max \frac{2}{||\omega||} \\
\text{s.t.} \\
\omega \cdot \vec{x}_i + b \geq 1 \text{ for all } \vec{x}_i \text{ in fault condition} \\
\omega \cdot \vec{x}_i + b \leq -1 \text{ for all } \vec{x}_i \text{ in normal condition} \\
\text{where } \cdot \vec{x}_i = [|V_i|, |I_i|]^T;
\]

(3.12)

After these two parameters, \(\omega\) and \(b\), are obtained, the decision boundary and the two margins can be determined. Whenever a new data point with a specific voltage and current value is measured (for example, the square point in Fig. 3.5), its distance to the boundary will indicate which class it belongs to.

**Sampling in Normal and Fault Conditions**

The normal condition refers to the condition when no faults happen and no equipment is disconnected from the system. When there is a short circuit somewhere in the system, this is known as the fault condition. To obtain the samples of features (i.e., the magnitude and phase of current or voltage, real or reactive power) during normal and fault currents, power flow and short-circuit analyses should be run for different conditions in the power system. During these simulations, the following assumptions are made in this dissertation:\(^1\):

1. Loads are random variables that have a uniform distribution in the range from 50% \((P_{\text{min load}})\) to 150% \((P_{\text{max load}})\) of the given load level. The loads are independent and identically distributed (i.i.d.) random variables.

---

\(^1\)In the SVM-based smart relays proposed in this dissertation, we only focus on improving the functional logic for detecting and locating faults. We have not considered making any changes on the measures that are taken after the faults are detected. Since the distance relays are the most widely used relays in power transmission systems, we utilized the concepts of zones and time delays common to distance relays.
variables. There is no correlation between any two loads.

2. Loads are modeled as constant power sinks;

3. There is sufficient power generation to meet demand;

4. 3-phase-to-ground faults\(^2\), whose fault impedance is varied from 0 to as high as 30\% of the line impedance, can happen at any location of the system, including the middle of transmission lines and buses, with the same probability.

The most widely used measurements, which can be obtained easily by today’s equipment in power systems are: 1) magnitude of current; 2) phase of current; 3) magnitude of voltage; 4) phase of voltage; 5) real power; and 6) reactive power. Although these six measurements can be used as features separately or together in power protection systems, they don’t have the same information content (or “energy”) as the candidate features. To select the most powerful features, principal component analysis (PCA) is performed before the SVM training.

**Smart Relays with SVM Classifier**

SVM classifier is built based on the training over sample features, in both noisy and noise-free conditions. Equipped with this classifier, smart relays can make a decision with high accuracy in complex conditions. Unlike the hypothesis testing based smart relays which work well when faults are located on buses, SVM-based smart relays can also discriminate, with high accuracy, the normal and fault conditions when faults are located in the middle of transmission lines.

SVM-based smart relays differ from hypothesis testing based relays and traditional relays in three important aspects:

- They employ classifiers which are determined via SVM training;

- They sample the features they need online, and make quick decisions using these online data; and

\(^2\)Although 3-phase-to-ground faults are not likely to happen in power systems, they happen during blackouts. In practice, 80\% of the faults on transmission systems are single phase-to-ground faults and this particular type of fault with high impedance is most likely to trigger a false trip in relays. However, this type of miss-operation of relays is not the major factor which leads to cascading failures. While the proposed SVM-based relay functional logic and simulation results shown in this dissertation are based on the study of 3-phase-to-ground faults, they can be easily extended to other types of faults.
They can update the SVM classifier online. This maintains a high accuracy when the system conditions change.

### 3.4.3 Model Selection by Cross-Validation and Grid Search

The non-linear kernel used in this dissertation is RBF kernel, which is the most widely used kernel function in SVM. There are two parameters when one uses RBF kernels: $\gamma$ and $\nu$. It is not known beforehand which $\gamma$ and $\nu$ are the best for one problem; consequently, model selection, which indeed is the parameter search, must be done before the training. The goal is to identify good parameters so that the classifier can accurately predict the unknown target value for each instance. It may not be useful to achieve high training accuracy in a single set of training data, since the training and test data may not be consistent all the time. Therefore, a common way is to separate training data into two parts where one part is considered unknown in training the classifier. Then the prediction accuracy on this set can more precisely react to the performance on classifying unknown data. This is known as cross-validation. In this dissertation, we use 10-fold cross-validation. We first divide the training set into 10 subsets of equal size. Sequentially one subset is tested using the classifier trained on the remaining 9 subsets. Thus, each instance of the whole training set is predicted once so the cross-validation accuracy is the percentage of data which are correctly classified [41].

Since parameters $\gamma$ and $\nu$ are coupled with each other, they cannot be determined separately. We use grid-search on $\gamma$ and $\nu$ with using cross-validation. Basically, different pairs of $\gamma$ and $\nu$ are tried and the one with the best cross-validation accuracy is picked. We observed that when one increases the set of values of $\gamma$ and $\nu$ exponentially, the grid search returns optimal results. Although grid-search seems to be a time-consuming method, it is a practical and safe method to identify good parameters.

The result of grid-search of 10-cross validation is shown in Fig. 3.6. As stated in this figure, the best accuracy we can obtain is around 94% where the parameters $\gamma$ and $\nu$ are equal to 32 and 128, respectively.

### 3.4.4 Dimensionality Reduction: Feature Extraction and Feature Selection

In statistics, dimensionality reduction is the process of reducing the number of random variables under consideration, and can be divided into feature selection and feature extraction [42]. Feature selection approaches try to find a subset of the original features. Two strategies are to use filter (e.g., information gain) or wrapper (e.g., search guided by the accuracy) approaches. Feature extraction transforms the data in the high-dimensional space to a space with fewer dimensions. The data transformation may be linear, as in
principal component analysis (PCA), but many non-linear techniques also exist. In this dissertation, feature selection is the main mechanism used before the SVM classifier is trained. The motivation is, compared with feature extraction, feature selection is advantageous for reasons that include the cost reduction in making, storing, and processing measurements. However, feature extraction is also considered as a supplementary option. The feature selection mechanism used in this dissertation is RELIEF, and the feature extraction mechanism is Principle Component Analysis (PCA).

RELIEF is developed by Kira and Rendell in [43]. It is shown to be very efficient in estimating features. The key idea of RELIEF is to weigh features according to how well their values differ for instances that are close to each other. To do this, RELIEF searches the two nearest neighbor for a given training data, one of which is from the same class (nearest hit) and the other from the different class (nearest miss). RELIEF in fact use $W(A)$ of the feature $A$, which is an approximation of the difference of probabilities to weigh each
\[ W(A) = P(\text{different value of } A - \text{nearest instance from different class}) \]
\[ - P(\text{different value of } A - \text{nearest instance from same class}) \]  
(3.13)

The rationale is that good feature should differentiate between training data from different classes and should have the same value for data from the same class. In a simplified version of RELIEF in Support Vector Machine technique, one uses the value of \( \vec{w} \) to approximate the probability in equation (3.13). As stated in previous section, \( \vec{w} \) is the vector normal orthogonal to the separating hyperplane and the hyperplane with the maximum margins is the one which can maximize \( ||\vec{w}||^2 \). Moreover, \( \vec{w} \) is a vector which has the same length as the dimension of features. For example, in a four feature SVM classifier, \( \vec{w} \) is a \( 1 \times 4 \) vector with each of the components corresponding to a single feature. Therefore, the feature corresponding to the largest component in \( \vec{w} \) is the most important feature, because it contributes most in deciding the location of the decision hyperplane and the width of margins.

In statistics, principal component analysis (PCA) is a technique that is used for simplifying a data set, by reducing multidimensional data sets to lower dimensions for analysis. Formally speaking, PCA is an orthogonal linear transformation that transforms the data to a new orthogonal (or orthonormal) coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on. PCA can be used for dimensionality reduction in a data set while retaining those characteristics of the data set that contribute most to its variance, by keeping lower-order principal components and ignoring higher-order ones. Such low-order components often contain the most important aspects of the data and the relevant information. The PCA algorithm comprises the following steps:

- Subtract the mean: The mean subtracted is the average across each dimension. This produces a data set with zero-mean;
- Calculate the covariance matrix: The covariance matrix should be a semi-positive \( n \times n \) matrix, where \( n \) is the dimension of features;
- Calculate the eigenvalues and the eigenvectors of the covariance matrix: The eigenvalues and the eigenvector for the covariance matrix can tell us the patterns in the data. The largest eigenvalue indicates the most important new features which can be constructed from the original features, and
the eigenvectors tell us how to construct the new features from the original features. The eigenvector with the highest eigenvalue is the principle component of the data set. It is important to point out that since eigenvectors are norm-orthogonal to each other, the new features will be orthogonal to each other.

- Choosing components and forming a feature vector: This is where the notion of data compression and reduced dimensionality comes into the picture. Once the eigenvectors and eigenvalues are found in the previous step, the eigenvectors are ordered by their corresponding eigenvalues, highest to lowest. This gives the components in the order of significance. In this step one can also decide to ignore the components with lesser significance. By doing this, one does lose information, but if the eigenvalues are small, one does not lose much.

- Deriving the new data set: Once we have chosen the components that we want from the previous step, we simply take the transpose of the vector and multiply it on the left of the original data set, then transpose back. This will give us the original data solely in terms of the vectors we chose.

After doing all these steps, basically, we have transformed our data so that the data are expressed in terms of the patterns between them, where the patterns are the lines that most closely describe the relationships between the data.

Both RELIEF and PCA can reduce the feature dimensionality. The difference is PCA will generate new features which may not have any physical meaning, while RELIEF will pick up the most important subset of original features which may not be as discriminative as the features generated by PCA. In this dissertation, both PCA and RELIEF are used to analyze the patterns and relationships between features, but dimension reduction is done by applying RELIEF to pick up the subset from original features with physical meaning.

RELIEF is applied to the feature matrix to select the best features with highest “energy”. The RELIEF analysis we conducted shows that real and reactive power are the two features with most “energy” (i.e., the largest amount of information). By using the combination of these two features, we can obtain about 75% of the complete information. If the phase of voltage is also taken into consideration, then 85% of the total information is recovered from these three features. On the other hand, we have observed that the magnitude of current and voltage, which are widely used in conventional relays, do not have strong discrimination properties in most of the relay locations. What needs to be pointed out here is the importance of reactive power. Reactive power is not commonly used in current protection systems, because it is not
Figure 3.7: The histogram of the distributions of six features, which are the magnitude and phasor of current and voltage, real and reactive power. Dark blue is the normal class, red indicates zone 1 fault and light blue indicates relax class.

readily measurable compared to other features. However, the reactive power has either the highest or the second highest discrimination capability among all these six features in most of the relay locations. It contains more than 60% of the total information in some of the relay locations. If this important feature, which is not used in conventional relays, is taken into account in SVM-based smart relays, then the decision accuracy will be increased significantly.

Fig. 3.7 shows the histogram of the 6 features in three classes, which is consistent with the RELIEF analysis results. Obviously, when using a single feature, reactive power delivers the most discriminative among three classes. This can also be demonstrated by Fig. 3.8.

PCA also delivers consistent results with RELIEF. The strongest two new features which is generated by PCA lie very close to the direction of reactive and real power. 80% of the strongest new feature is constructed by the reactive power and the rest of the 20% are from other 5 features. This, once again, demonstrates that the reactive power and real power, which is not widely used in today’s protective relays, are in fact very discriminative.
Figure 3.8: The plot matrix of the three most important features, which are reactive power (feature 6), real power (feature 5), and phasor of voltage (feature 4). Blue is the normal class, red indicates zone 1 fault and green indicates relax class.
### 3.5 Simulation results

#### 3.5.1 Simulation results on IEEE test networks: SVM Training and Testing

To generate samples of the these six features (magnitude and phase of current, magnitude and phase of voltage, real power and reactive power), simulations were performed using the IEEE 118 bus network [38].

In the Monte Carlo simulations performed on the IEEE 118 bus network, load level was varied from 50% to 150% of the normal design values and various locations for faults were considered. As a result of these extensive simulations, we obtained nearly 10,000 samples for each feature. In other words, the feature space we have is presented as a matrix which contains 10000 rows (approximately) and 6 columns. Besides the feature matrix, we have another important vector which contains all the class labels corresponding to each feature. In this dissertation, we have three classes, which are relaxed condition, stand-by condition, and emergency condition. The emergency condition corresponds to the condition when fault is located in zone 1. Under this condition, a protective relay should open the transmission line immediately. The stand-by condition occurs when the fault happens within zone 2 and zone 3. Under such conditions, a protective relay should stand by as the back-up relay, waiting for the primary relay to work first. The relaxed condition, in fact, includes the normal condition and the far-away fault condition. In normal condition, there is no fault anywhere in the system; while in the far-away fault condition, a fault is located beyond zone 3. Under both of these two conditions, a protective relay should have no reaction. In our experiment, we label the emergency, stand-by, and relax conditions as $-1$, $0$, and $1$, respectively.

In this section, we present the training and testing results from SVM on IEEE 118 bus test networks. The software used in this part are matpower [45] and LIBSVM [41]. Tables 3.2, 3.3, and 3.4 show the SVM training and testing results in both noisy and noise-free conditions.

Tables 3.2, 3.3, and 3.4 show the SVM training and testing results under both noisy and noise-free conditions. As expected, the training and testing accuracy in noise-free conditions increases when more features are taken into consideration. It was observed that the effect of the noise is not significant for

---

<table>
<thead>
<tr>
<th>Table 3.2: Accuracy of SVM-based Smart Relays with Two Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Features: Real Power and Reactive Power</td>
</tr>
<tr>
<td>Training (Noise Free)</td>
</tr>
<tr>
<td>Test Accu.</td>
</tr>
<tr>
<td>Train Accu.</td>
</tr>
<tr>
<td>95.1071%</td>
</tr>
<tr>
<td>95.1044%</td>
</tr>
<tr>
<td>95.0254%</td>
</tr>
</tbody>
</table>
training and testing in this experiment. This means that the gaps between the sampling data in different classes are large enough. However, there may be some cases in which these gaps are relatively small. In these cases, some of the features may be sensitive to noise.

Moreover, similar simulations were set up assuming the (N-1) condition. (N-1) condition is defined as the system condition in which one piece of equipment has already been disconnected from the system. In this dissertation, it refers to one of the transmission lines being disconnected. This is because several switches stay open after previous faults or because that transmission line is under maintenance. Although (N-1) conditions are considered as part of the normal conditions, they represent more vulnerable situations in which a disturbance or the isolation of a disturbance has a higher probability to lead to false trigger of relays. The malfunctioning of traditional over-current relays after (N-1) condition was one of the critical factors which caused the blackout in the USA in August, 2003 [3]. We will prove that the SVM-based smart relays can achieve high accuracy even in this critical condition. Table 3.5 shows the simulation results of SVM-based smart relays in this condition. In Table 3.5, we still use real and reactive power as the two main features, and add the phase of voltage to build the three features combination. Table 3.5 shows the training and testing accuracy when we take the sample data resulting from response

<table>
<thead>
<tr>
<th>Table 3.3: Accuracy of SVM-based Smart Relays with Three Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>3 Features: Real Power, Reactive Power and phase of Voltage</td>
</tr>
<tr>
<td>Training (Noise Free)</td>
</tr>
<tr>
<td>Testing (Noise Free)</td>
</tr>
<tr>
<td>96.4824%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3.4: Accuracy of SVM-based Smart Relays with Six Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>6 Features</td>
</tr>
<tr>
<td>Training (Noise Free)</td>
</tr>
<tr>
<td>Testing (Noise Free)</td>
</tr>
<tr>
<td>96.8351%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3.5: Accuracy of SVM-based Smart Relays in (N-1) Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>2 Features</td>
</tr>
<tr>
<td>84.2147%</td>
</tr>
</tbody>
</table>
to (N-1) condition as a separate class from the previous defined three classes. This means, the protective relays have an 86% chance of making correct decisions when discriminating the zone 1 fault, zone 3 fault, normal condition, and (N-1) condition from each other, by using all the aforementioned six features. When considering (N-1) conditions as part of the class 3 (normal and far-away fault), in which the relay will not trip a transmission line, we can achieve a training and testing accuracy as high as 94% which looks very promising.

The correct interpretation of the accuracies of SVM-based smart relays reported in this Section is paramount. The accuracies shown in Tables 3.2, 3.3, 3.4 and 3.5 are the testing accuracies generated by an SVM classifier. Hence, they do not represent the correct tripping rate of SVM-based relays in practice. In other words, a 95% accuracy in these tables does not indicate that a relay will make 5 incorrect decisions out of every 100 estimations. To illustrate this statement, assume 9000 instances are used to test a relay in the power system. These 9000 instances are generated from normal and abnormal system states. To make sure the test is unbiased, same numbers of instances from each class are used: e.g., 3000 instances are generated from each of the three classes; class 1, class 0, and class -1, respectively. Therefore, when the testing accuracy is calculated in the simulations, the same prior probabilities (1/3) are used for each class. Moreover, to ensure that the number of instances from different classes is the same, we replicate the instances from the classes that contain fewer scenarios. For example, there are various scenarios in the class 1 which correspond to normal situations, while there are fewer scenarios in the class -1 and class 0 which correspond to abnormal situations. To balance the instance numbers from each class, some of the scenarios in class -1 and 0 have to be replicated. Therefore, when a scenario, which is misclassified by a relay, is replicated n times, it is counted as n instead of 1 when we calculate the accuracy. Because of these two reasons, the testing accuracy after simulation is not equal to the real accuracy of decisions in practical relays. For example, if the practical probabilities of normal, zone 1 fault and zone 2 fault conditions are 94%, 1%, and 4% respectively, the correct decision rate of the SVM-based smart relays with N-1 conditions considered is 99.1%, instead of 94% accuracy in Table 3.5. In fact, further analysis on the testing process shows that most of the incorrect decisions made by an SVM-based relay are the zone 3 faults in class 0. The effects of these several misclassified scenarios are magnified because of the scenario replication that was performed in our simulations.\textsuperscript{3}

\textsuperscript{3}In practice, the general number of trippings of HV protection systems (500 kV and 230 kV) in a well maintained utility is approximately 15 to 20 per year, out of which about 20-25% are miss-operations of some kind.
The SVM-based smart relays are robust to various fault impedances if they are trained with various fault impedances. In other words, they can function well with high accuracy to detect high-impedance faults. The accuracy of the SVM relays with various fault resistances, which is from 0 to 0.1 p.u., is shown in Fig. 3.9. However, if an SVM relay is trained with only the data which are generated with zero-impedance fault, it cannot function accurately when dealing with high-impedance faults. Therefore, it is required to train a relay with all possible fault resistances to ensure that it can detect a fault accurately in practice. Of course, if one has the pre-knowledge of the range of fault resistances, one can use a larger weighing factor for the training cases which are likely to happen in practice to emphasize their importance in achieving higher accuracy.

We have to note that the accuracy of an SVM-based smart relay is highly dependent on the accuracy of the model’s parameters. Inaccurate parameters can greatly decrease the accuracy of SVM relays, even trigger the relays to false trip their breakers. To illustrate this statement, simulations were performed on the IEEE 14 bus systems. In these simulations, the effects on accuracy caused by inaccurate models parameters are estimated. The simulations are designed as follows:

- Loads are random variables that have a uniform distribution in the range from 50% to 150% of the given load level. The loads are independent and identically distributed (i.i.d.) random variables. There is no correlation between any two loads.

- Loads are modeled as constant power sinks;

- There is sufficient power generation to meet demand;

- 3-phase-to-ground faults can happen at any location of the system, including the middle of transmis-
sion lines and buses, with the same probability.

- A relay is trained in a scenario where the parameters of a transmission line are inaccurate.
- A relay is tested in a system where the parameters of every component are as designed.
- To simulate the case of inaccurate transmission line parameters, the impedance of a transmission line is varied from 50% to 150% of its designed value.

Three scenarios which were tested in the simulations are:

- Local strong lines scenario: The parameters of a strong transmission line are inaccurate and the relay on this line was in question.
- Local weak lines scenario: The parameters of a weak transmission line are inaccurate and the relay on this line was in question.
- Inaccurate impedance on neighboring lines scenario: A relay is trained with a case where the parameters of a neighboring transmission line are inaccurate.
- Inaccurate impedance on far away lines scenario: A relay is trained with a case where the parameters of a far-away transmission line are inaccurate.

The simulation results are shown in Fig. 3.10.

Obviously, the testing accuracy is related to the accuracy of the model’s parameters. Inaccurate parameters significantly affect the accuracy of local and neighboring relays, while they have less effect on far away relays. Inaccurate parameters of a strong line significantly decrease the decision accuracy of the relay which is located on the line. Therefore, accurate parameters, especially accurate parameters of major components in the system, are required to ensure that an SVM-based smart relay can function correctly as designed.

### 3.5.2 Scalability

Scalability is a crucial requirement for the proposed SVM-based smart relays. Simply put, the questions one needs to address are the following: 1) Given that the training and testing of the SVM-based smart relays is done on the relatively small ‘toy networks’ (such as the IEEE 30 bus system or the IEEE 118 bus system), how can one guarantee that when one deploys these relays in a huge network like the U.S. power grid, their
Figure 3.10: The Accuracy of SVM-based Relays Which Are Trained with Different Inaccurate Parameters

performance will be satisfactory and robust; 2) To ensure such excellent and robust performance, what are the necessary and sufficient conditions (e.g., how many tiers are needed for training and testing); 3) Which approach or methodology can be used to establish the scalability of SVM-based smart relays; 4) How many features will be needed for the training and testing of SVM-based smart relays for acceptable performance.

Scalability is very important for managing combinatorial off-line simulations, especially the ones using statistical methods. In this context, scalability critically depends on how much and which data can be either disregarded or approximated as unchanged when Monte Carlo simulations are carried out for very large electric power networks. By performing simulations on the IEEE 118 bus system, it was found that training an SVM-based smart relay in a 5 tier network is sufficient because the settings of the relay is highly related to the components and states within the 5 tier network surrounding its location, but not to tiers beyond the 5th tier. In other words, if an SVM-based smart relay is only trained with an accurate model of its neighboring 5 tier network, it can be implemented in a large scale network and it will function well in terms of detecting and locating faults. This also indicates that any inaccurate components’ parameters or changes outside the 5 tier network do not have a strong bearing on the settings of the relays. Therefore, when training and testing an SVM-based relay, one can treat the impact of the power system which is outside the 5 tier network as an equivalent power injection and ignore the changes that occur beyond the 5th tier. This is the key finding that makes the implementation of SVM-based smart relays in practice possible. The setting of a practical
SVM-based relay, which is installed and works in the U.S. power grid containing hundreds and thousands of buses and lines, depends only on the neighboring 5 tier network which is as large as the IEEE 30 bus system. Thus, it is sufficient to train a practical SVM relay with the information of a sub-network, which can be obtained from a single utility, and implement it in the U.S. power grid. Moreover, because the SVM-based smart relay is scalable, its accuracy which was obtained by performing simulations on the IEEE 118 bus system represents its performance when it is implemented in the large-scale system. On the other hand, the scalability also ensures that the setting of the SVM relays does not need to be updated frequently because of the changes that may occur in the system. Since the setting of an SVM relay is only highly related to the 5 tier network surrounding its location, any long term or short term changes incurred in the network which is outside this 5 tier network will not trigger the updating of the relay’s setting. In other words, any change in the system only triggers several nearby relays to update their settings without affecting the relays in the rest of the system. Therefore, when a long term change takes place in the system, for example a connection of a new transmission line, only a very small portion of the total number of relays which are close to the change (i.e., the new transmission line) need to be updated.

Moreover, since the SVM-based smart relays are scalable, which stems from the fact that their settings depend only on their neighboring 5 tier network, the simulation time and data volume needed for training the protection relays before installing them on the transmission lines can be significantly reduced. Thus, within a reasonable training period and an attainable volume of simulation data, the protection systems may become more adaptive during the failures and help avoid widespread blackouts. This implies that SVM classification can be used effectively in novel protection relays. Ultimately, the most important challenge is to develop methods for adjusting the method’s complexity as a function of the degree of scalability, as the system conditions change. This is the subject of our ongoing work.

In this dissertation, we introduce a concentric relaxation-like approach for testing the scalability property of smart relays. While electric power systems exhibit under most conditions a localized response, which means the effect of change dies out with the electrical distance, the extent to which the rest of the system is affected greatly depends on the actual electrical characteristics of the transmission system and on the level of power transferred by the system. It is for this reason that one needs to determine how far the effect of a fault, e.g., a line outage, spreads.

We propose an algorithm to test the scalability of smart relays by performing simulations in a similar manner as gradually growing the test areas in a concentric relaxation-like way, tier by tier. A tier $I$ is defined
as a set of network nodes directly connected to the set of nodes in tier \((I - 1)\) surrounding the fault location [40].

The algorithm for scalability testing has four basic steps:

**Step 1. Defining the initial area**

Because scalability is a characteristic which is relevant to enlarging the systems under investigation, the starting test area is a single transmission line whose relay is being tested. For example, if the relay in the transmission line between bus 1 and bus 2 is being tested, this transmission line would be chosen as the first tier testing area.

**Step 2. Growing the testing area tier by tier**

In our algorithm, the testing area is enlarged tier by tier. In Step 2, the testing area is divided into three parts: inner networks, boundary, and outer networks. The inner network consists of the transmission lines and buses which are completely located inside the testing area. The outer part includes the transmission lines and buses which are not covered by the testing area. The boundary part includes the buses which are connected to the transmission line from the inside and the outside of the testing area.

Fig. 3.11 shows the process of enlarging a testing area from a single transmission line. To enlarge the testing network from \(N\) tiers to \((N + 1)\) tiers network, the buses which are one hop away from the boundary buses are counted in, as well as the transmission lines connected in between. Subsequently, these buses become the new boundary buses in the \((N + 1)\) tier network.

**Step 3. Equivalencing the Testing Network**

After a testing area is defined, the outer network is equivalenced as injections on the boundary buses; the injections have the same value as the tie-line power flows. The equivalenced injection into bus \(i\) is

\[
P_i = \sum_{j \in i} V_i \cdot I_{ij}^*\]

where \(P_i\) is the equivalent complex valued power injection into bus \(i\). \(j \in i\) represents the set of buses directly connected to bus \(i\). \(V_i\) is the complex-valued voltage on bus \(i\), \(I_{ij}\) is the complex-valued current in the transmission line connecting buses \(i\) and \(j\).

**Step 4. SVM Classification and Comparison of the Results**

After the equivalenced network is created, SVM classification method is applied to the obtained network, starting with the Tier 1 network to the Tier \(M\) network. If the testing accuracy in the \(M\)th and the \((M + 1)\)st network are not significantly different, then the algorithm reports that relay located in Tier 1 is scalable to the degree of \(M\)-tier network. We refer to this relay as being scalable to the degree \(M\).

Fig. 3.12 shows the simulation results for scalability testing of SVM classification based smart protective
Figure 3.11: The Process of Enlarging the Testing Area Tier by Tier Starting with a Single Transmission Line

Figure 3.12: Testing and Training Accuracy in Different Size Networks
relays in the IEEE 118bus systems [38]. In Fig. 3.12, the blue curve stands for the testing accuracy in each network when training is done in the same network with six features; while the red curve shows the testing accuracy when training is done in the same network with only three features. The black and green curves show the testing results on a different scale network when the classifier is training only on a five tier network, by using 6 and 3 features, respectively. By comparing these curves in pairs, one can draw several important conclusions:

1. Training and testing with six features and/or a larger network can always lead to better results, because including more far away areas or more features can increase the gap between the clusters of normal and faulty conditions.

2. Larger networks can lead to better testing accuracy, because including more far away areas in the training and testing area can increase the gap between the clusters of normal and faulty conditions. This can be concluded from each of the four curves shown in Fig. 3.12.

3. SVM-based smart protective relays are scalable, in the sense that they can deliver a near-optimum testing accuracy when the classifier is trained using a limited area. Obviously, one can always have a better testing accuracy when one uses a classifier which is trained in the same network; the difference between these two cases, however, is minor: less than 2% with six features and less than 4% with three features. Considering the decrease in training time and computational complexity, this implies that in practical scenarios using three features could be adequate both in terms of accuracy and computational complexity. Moreover, to apply this in practice, another important conclusion should be drawn:

4. When comparing the classifiers which are obtained from a smaller network with an optimal classifier (which is obtained from the complete system), the one which is trained with six features is closer to the optimal classifier than the one which is trained only with three features.

3.6 Summary

In this Chapter, we have proposed a novel functional logic for smart relays based on machine learning techniques. We first examined the possibility of apply binary hypothesis testing to protective relays, and then investigated the application of Support Vector Machines to smart relays. The simulation results indicated that the use of these machine learning techniques in protective relays has the potential to improve the decision making accuracy of protective relays. Moreover, we have observed that some rarely used measure-
ments, such as real and reactive power, are in fact very strong features for discriminating different system conditions. We also have demonstrated the scalability of the proposed SVM-based smart relays approach, which makes the practical implementation of smart relays possible.
Chapter 4

Comparison of SVM-based Smart Relays and Conventional Relays

4.1 Case Study

Fig. 4.1 shows a 7-bus network, which is a sub-system with a distributed generator (DG). This 7-bus network can be simplified to the one-line diagram in Fig. 4.2. This 7-bus network is from a DMS ([46]) demo system, which is representative of a real power system in Europe. Because of the distributed generator which is connected to bus 7, it is hard to set up suitable thresholds for relays. As shown in the figure, the setting of the relay on bus 7, which is designed to protect from a local fault on line 6-7, will malfunction and trip when the fault happens on line 1-2 (see Fig. 4.2).

To obtain the values of settings for the protective relays in substation 7 so that they can clear the local fault accurately without causing malfunction due to faraway faults, support vector machines (SVM) can be used. By using SVM, multiple features need to be considered at the same time, with a non-linear decision boundary. The decision making accuracy demonstrated that SVM-based smart relays can handle this situation much better than conventional relays.

In Table 4.1 and Table 4.2 we present the training and testing results. In this simulation, we test the protections in substation 7, on both line 1-7 and line 6-7. The first two columns of the table shows the training and testing accuracy with respect to all possible faults in the network. The third column shows the testing accuracy when making decisions only on Zone1 faults; and the last column is the decision accuracy.
Figure 4.1: 7-Bus System in an European Power System [46]
CHAPTER 4. COMPARISON OF SVM-BASED SMART RELAYS AND CONVENTIONAL RELAYS

![Diagram of a network with labels](image)

**Figure 4.2:** The current in 7bus network (a) before and (b) during a fault in transmission line 1-2. This fault current can trigger the relay on line 1-7 to false trip.

Table 4.1: Accuracy of SVM-based Smart Relays on Line 1-7 with Six Features

<table>
<thead>
<tr>
<th></th>
<th>Training Accuracy for All</th>
<th>Testing Accuracy for All</th>
<th>Testing Accuracy for Zone1 Fault</th>
<th>Testing Accuracy for Fault in Line 1-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear Kernel</td>
<td>92.5%</td>
<td>92.43%</td>
<td>100%</td>
<td>90.37%</td>
</tr>
<tr>
<td>Non-linear Kernel</td>
<td>98.67%</td>
<td>98.68%</td>
<td>100%</td>
<td>92.98%</td>
</tr>
</tbody>
</table>

...when dealing with the faults between substation 1 and substation 2. As discussed in the previous section, a conventional relay cannot identify the proper setting to discriminate the faults on line 1-2 from its local fault.

Even a subset of 6 features (e.g., 4 features) in this case can deliver a high accuracy in decision making. Table 4.3 and Table 4.4 show the decision accuracy of the protective relays in substation 7 with three features. Obviously, the decision accuracy obtained by using three features is lower than that by using six features, but it is at an acceptable level. Hence, four features are probably good enough for most of the cases.

Table 4.2: Accuracy of SVM-based Smart Relays on Line 6-7 with Six Features

<table>
<thead>
<tr>
<th></th>
<th>Training Accuracy for All</th>
<th>Testing Accuracy for All</th>
<th>Testing Accuracy for Zone1 Fault</th>
<th>Testing Accuracy for Fault in Line 1-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear Kernel</td>
<td>91.56%</td>
<td>91.65%</td>
<td>100%</td>
<td>86.23%</td>
</tr>
<tr>
<td>Non-linear Kernel</td>
<td>96.66%</td>
<td>96.47%</td>
<td>100%</td>
<td>98.42%</td>
</tr>
</tbody>
</table>
Table 4.3: Accuracy of SVM-based Smart Relays on Line 1-7 with Three Features

<table>
<thead>
<tr>
<th></th>
<th>Training Accuracy for All</th>
<th>Testing Accuracy for All</th>
<th>Testing Accuracy for Zone1 Fault</th>
<th>Testing Accuracy for Fault in Line 1-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear Kernel</td>
<td>90.22%</td>
<td>90.38%</td>
<td>100%</td>
<td>88.98%</td>
</tr>
<tr>
<td>Non-linear Kernel</td>
<td>94.7%</td>
<td>95.23%</td>
<td>100%</td>
<td>93.28%</td>
</tr>
</tbody>
</table>

Table 4.4: Accuracy of SVM-based Smart Relays on Line 6-7 with Three Features

<table>
<thead>
<tr>
<th></th>
<th>Training Accuracy for All</th>
<th>Testing Accuracy for All</th>
<th>Testing Accuracy for Zone1 Fault</th>
<th>Testing Accuracy for Fault in Line 1-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear Kernel</td>
<td>87.93%</td>
<td>87.65%</td>
<td>100%</td>
<td>80.42%</td>
</tr>
<tr>
<td>Non-linear Kernel</td>
<td>95.21%</td>
<td>94.88%</td>
<td>100%</td>
<td>95.76%</td>
</tr>
</tbody>
</table>

Thus, our results show that by using SVM-based smart protective relays, compared to conventional relays, high accuracy in decision making by single relays and easier coordination among relays can be achieved.

4.2 Benchmarking Performance of SVM-based Smart Relays and Conventional Relays in Cascading Failures

4.2.1 Benchmarking Performance

In most of the cascading failures in electric power systems, the malfunctioning of protective relays is the key factor that accelerates the propagation of a fault. As explained in previous sections, the key reason for relay malfunctioning during cascading failures is the fact that the conventional logic of protective relays is not adaptive to the changes in the system, so they cannot distinguish between the (N-1) condition and the faulty condition. Therefore, erroneous tripping may take place after the correct isolation of the initial fault. To solve this problem and mitigate the cascading failures in electric power systems, the conventional distance relays in the power system can be fully or partially replaced by the SVM-based smart relays. The simulations were performed using the IEEE 118 bus systems and the benchmark case is shown in Fig. 4.3 and Fig. 4.4.

The assumptions used in the simulation are as follows:

- Conventional distance relays make decisions based on the R-X diagram shown in Fig. 2.3.
CHAPTER 4. COMPARISON OF SVM-BASED SMART RELAYS AND CONVENTIONAL RELAYS

- A conventional distance relay will trip a transmission line in 0.5 second\(^4\) when a zone 1 fault is detected, in 1 second when a zone 2 fault is detected, and in 2.5 seconds when a zone 3 fault is detected.

- SVM classification based smart relays are trained by up to 6 features, which include magnitude of current, phase of current, magnitude of voltage, phase of voltage, real, and reactive power.

- SVM classification based smart relays are trained to classify the system into 6 classes, which are normal, zone 1 fault, zone 2 fault, zone 3 fault, beyond zone 3 fault, and (N-1) conditions.

- SVM-based smart relays trip a transmission line in 0.5 second when zone 1 fault is detected, in 1 second when zone 2 fault is detected, and in 2.5 seconds when zone 3 fault is detected. They generate warning messages but stay closed when a beyond-zone 3 fault and/or (N-1) conditions are detected.

In the IEEE 118 bus system, there are 118 buses, 54 generators, and 188 transmission lines; therefore, 376 transmission line relays are installed on each terminal of the transmission lines. Simulations are performed by placing the initial fault on any place in each transmission line which is directly connected to a generator, from 100% to 200% of designed load level. The results for benchmarking are shown in Fig. 4.3 and Fig. 4.4, respectively.

The four simulation scenarios are defined as:

- All distance relays scenario: all the 376 relays installed on each terminal of the 188 transmission lines are distance relays with pilot channel. Each distance relay is equipped with 3 zones which is typically as shown in Fig. 4.5. Each relay in the system has pilot channels to the relay on the other terminal of the same transmission line in POIT scheme.

- All SVM relays scenario: all the 376 relay relays in the 118bus system are smart relays equipped with SVM classifiers. They make decisions on all 6 features from their local measurements, which include magnitude of current, phase of current, magnitude of voltage, phase of voltage, real and reactive power. There is no information exchange between any two relays.

- All SVM relays with pilot scenario: all the 376 relays in the system are smart SVM relays and there are pilot channels between each pair of relays on the same transmission line. However, the pilot

\(^4\)In practice, a protective relay usually trips the breaker within 8 cycles (around 100 ms) when a zone 1 fault is detected. However, in this section, a tripping delay of 0.5 s is used for better visualization and for explaining the concept.
scheme utilized in this scenario is different from the all distance relays scenario. Through pilot communication, a relay, besides its local measurements, will know the decision made by the other relay on the remote terminal of the same transmission line. For example, on the transmission line between bus 1 and bus 2, there are two relays which are located close to bus 1 and bus 2 respectively. Each of them can make their own decision via SVM classification based on 6 local measurements, which we call the first time SVM decision making. After this, they exchange their independent decisions through the pilot channel between them, through which the relay close to bus 1 knows the decision by the relay on bus 2, while the relay on bus 2 also knows the decision made by the relay on bus 1. Through this pilot communication, both of these two relays know 7 different pieces of information, which include the 6 local measurements and the decision made by the relay on the remote terminal of the same transmission line. Then, the smart relays will perform another SVM classification using all these 7 features and take action according to their final decisions, which we call the second time SVM decision making. Compared with the conventional distance relays with pilot channels which use the decision from the remote relay to trip or block breakers, smart SVM relay utilize this information as a supplementary feature for the second time decision making.

- Distance relays with several SVM relays on critical locations: The conventional distance relays in the critical locations which have the highest probability of incorrect tripping after disturbances have been identified via extensive simulations and replaced by SVM-based smart relays. The relays which is replaced by smart SVM relays are only a very small portion of all the relays in the IEEE 118 bus systems. In this dissertation, only 20 out of 376 conventional distance relays have been replaced by smart SVM relays. In this scheme, there is no pilot communication between any pair of relays, and all the relays only execute a one time decision making.

In Fig. 4.3, the ratio of possible initial disturbances which leads to incorrect tripping is computed as:

\[
\text{Ratio} = \frac{\text{# of disturbances that will lead to incorrect tripping}}{\text{# of total possible disturbances}} \quad (4.1)
\]
Figure 4.3: The ratio of possible initial disturbances which will lead to incorrect tripping in different load levels.

Figure 4.4: The total number of incorrect tripping after possible initial disturbances in different load levels.
In Fig. 4.4, the total number of incorrect tripping is computed as:

\[
\text{# of incorrect tripping} = \sum_i (\text{# of relays which are false tripped after disturbance } i),
\]

where \( i = \{1, 2, \ldots N\} \) and \( N \) is the total number of possible initial disturbances.

(4.2)

Obviously, the optimal solution is the scenario where most of the relays in the system are conventional distance relays and smart SVM relays are only installed at critical locations. By replacing the conventional relays at critical locations where a false tripping is likely to happen, the probability of a disturbance propagating to a larger area can be decreased significantly. Therefore, by installing and using SVM-based smart relays judiciously at the most critical locations, which comprise a very small portion of the total number of protective relays in the entire system, cascading failures can be effectively mitigated.

There are several important observations that can be made from Fig. 4.3 and Fig. 4.4:

* Load level is a very important factor involved in cascading failures. When the load level is low; for example, less than 20% overloading of the designed load level, a disturbance in the power system hardly leads to a false trip of other relays in the network. When the load level is less than 140% of designed level, even though there is a probability that an initial fault will trip other relays incorrectly, it is very unlikely that a false trip will lead to cascading failures. In other words, when the load level is less than 40% overloading, there is a near-to-zero probability that a disturbance will false trip relays

Figure 4.5: Typical 3 zones design of the conventional distance relays
CHAPTER 4. COMPARISON OF SVM-BASED SMART RELAYS AND CONVENTIONAL RELAYS

in the power system to go to more than two steps. However, when the load level increases to more than 40% overloading, the probability of incorrect tripping increases significantly. In a system where there are only conventional distance relays or no pilot communications, almost every initial fault on strong lines will trip other relays incorrectly. Therefore, the ratio of disturbances which trigger hidden failures in the Fig. 4.3 increases from 10% to more than 95% in the scenario with all distance relays and the scenario with non-pilot SVM relays.

- Compared to conventional distance relays with pilot channel, non-pilot SVM relays cannot decrease the ratio of disturbances which will trigger hidden failures. This means that even if all the relays in the system are equipped with SVM classifiers, without pilot communication, in heavily loaded condition, initial faults on strong transmission lines will still trip at least one other relays in the system incorrectly. However, the application of SVM classification mechanism in relays can decrease the total number of false tripping events, as shown in Fig. 4.4. For example, when the load level is 160% of the designed value, the ratio of disturbances which will trigger hidden failures are close to 100% in both the distance relays and SVM relays scenarios in Fig. 4.3, however, the total number of false tripping events in the SVM relays scenario is only one third of the number in the distance relay scenario. In other words, on average, a disturbance which triggers 3 other distance relays to malfunction will only trigger 1 SVM relay to open incorrectly. Moreover, the simulation results also indicate that the locations of the relays which are malfunctioning are much less important when one uses SVM relays. Therefore, although in heavily loaded conditions, a disturbance in the SVM relay equipped system has a high chance to trigger a few incorrect trippings, it has a low probability of reaching to a cascade of failures. In fact, in most cases the false tripping only continues one more step and then stops. Fig. 4.6 shows a typical case in an extremely heavily loaded condition.

- The pilot communication can help to improve the performance of SVM relays significantly. It is shown in Fig. 4.3 and Fig. 4.4 that the SVM relays perform much better with pilot communications than without, in terms of both the ratio of disturbances that trigger hidden failures and the total number of false tripping events. This is most obvious when the load level is between 140% to 180% of the designed value. When the load is higher than 180% of the designed level, even with the pilot communications, almost every disturbance on strong lines will trigger incorrect trippings on other SVM relays. However, at this load level, the pilot communication, although it cannot decrease the
ratio of disturbances which trigger hidden failures, it can help to decrease the total number of false tripping events and therefore it can decrease the probability of cascading failures.

- The optimal scenario is the hybrid scheme where most of the relays in the system are conventional distance relays and smart SVM relays are only installed at critical locations. This is because the locations where the distance relays tend to undergo false trip are different from the locations where the SVM relays tend to malfunction. Simply out, in heavy load condition, the false trippings always happen on the relays of strong transmission lines in distance relays scenario, while in the SVM relays scenario they tend to happen on relatively weak lines. This will be further explained in the next section.

### 4.2.2 Critical Locations

Traditionally, protection systems have an intrinsic bias towards dependability at the expense of security. However, it was argued that due to the manner in which power systems have evolved and the increasing complexity and connectivity of the power system, this philosophy needs to be changed. Under stressed system conditions, a bias towards security is beneficial, otherwise the false alarm may destroy the power system. This false alarm, which is caused by the design defects and make relay to trip incorrectly under certain stressed condition, is one kind of hidden failure. An analysis of NERC outages report indicates that

![Typical 3 zones design of the conventional distance relays](image-url)
hidden failures are the underlying cause of over 70% of all cascading outages [2]. In this dissertation, the location in the power grid where a false trip is caused by a hidden failure is defined as critical location. Generally speaking, although significant research effort has been reported in developing technology to detect hidden failures, there is no clearcut way to identify the critical locations except exhaustive search. A methodology to index the severity for identifying critical locations based on dynamic simulations was proposed in [47] and a systematic procedure to identify and rank the critical locations was presented in [48]. However, these methodologies turn out to be either infeasible or too expensive when the system is complex and the number of circuit components is large. The number of simulations required for an exhaustive study can be computed as:

\[ \text{Simulations} = \frac{N!}{(N-k)!k!} \]  

(4.3)

where \( N \) is the total number of circuit elements in the systems and \( k \) is the number of elements being removed [48]. Therefore, the simulations required for an exhaustive N-1 and N-2 search are shown in Fig. 4.7

Although there is no better way than exhaustive search to identify the critical locations, one can try to identify the properties of the critical locations and therefore decrease the complexity of simulations. To achieve this, we need to analyze the critical locations which were identified by the exhaustive search in the IEEE 118 bus system in this dissertation.
By analyzing the relays which are incorrectly tripped by the initial disturbances, they can be classified into two different groups as shown in Fig. 4.8 and Fig. 4.9.

The malfunctioning pattern of the first group of relays is shown in Fig. 4.8. It can be observed that these relays begin to trip incorrectly at a relatively low load level, say only at 20% overload. However, the total number of false tripping events keeps at an extremely low level (10 false tripping events out of 900 simulation scenarios) without increasing as the load level increases. On the other hand, the other group of relays malfunctions in a completely different pattern which is shown in Fig. 4.9. They begin to false trip at a relatively high load level, say above 50% overloading. However, once the load level is higher than 150% of the designed value, they malfunction with a high probability which is around 90% of all the simulation scenarios. In other words, the first kind of malfunctioning seems more related to a certain disturbance, while the second kind of malfunctioning seems more related to the load level. Additional simulation results indicate that the first kind of relays tends to malfunction to the specific disturbance which happens close to their locations, while the second kind of relays can be tripped incorrectly by far away disturbances in heavily loaded cases. Therefore, in practice, the second kind of malfunctioning is more critical and more attention needs to be paid. In this dissertation, the relays which are associated with the second kind of malfunctionings are called critical relays. They are the relays which need to be replaced by the smart SVM relays.

In the IEEE 118bus system, in which there are 376 relays on total, there are 20 such critical relays. They
Figure 4.9: The histogram of false tripping events on the relay in different load level (relay 89-92)

have the properties of:

- They are all located on strong transmission lines, which has lower impedance.

- In their R-X diagram, their circles for zone 3 are usually very close to normal circle as shown in Fig. 4.10, compared with other relays’ as shown in Fig. 4.11. Therefore, when the system is overloaded, it is easy for the relays to misclassify the normal condition or N-1 condition as a zone 3 fault.

- They are all located close to the buses with higher connectivity. Network connectivity is often measured using the degree of nodes in the network, which is the number of edges connected to a given node. The histogram of the connectivity of the IEEE 118bus system is shown in Fig. 4.12. Based on the simulation results, the critical relays are all located close to the buses with 4 or more connections, which is only a small portion of all the buses.

4.3 Benchmarking Performance of SVM-based Smart Relays and Modern Distance Relays in Cascading Failures

During the 2003 blackout, there were a total of 14 inappropriate trippings of the protective system. Therefore, after a comprehensive study of this blackout, improvements have been made to enhance the functional logic of distance relays. One of the major changes on the distance relays is the implementation of load
CHAPTER 4. COMPARISON OF SVM-BASED SMART RELAYS AND CONVENTIONAL RELAYS

Figure 4.10: The R-X diagram for critical relays

Figure 4.11: The R-X diagram for non-critical relays
encroachment element. As per NERC Task Force requirements [49], Phase distance settings and other applicable phase and ground distance zone settings must permit loading of the line without trip to 150% of emergency line ampere rating, with 0.85 per unit bus voltage and a load angle of 30 degrees. The Load Encroachment element/function is set to prevent tripping of Distance Protection Elements on load. With the load encroachment element, the functional logic of a distance relay is shown in Fig. 4.13.

The settings of distance zones were done according to the following criteria:

- The first zone is an under-reaching fast tripping zone, set to 80% of the line-length. The time delay was set to trip in 100 ms.

- The second zone is a time delayed overreaching zone, set to 120% of the line length and it is applied to a permissive pilot scheme. The time delay was set to 300 ms.

- The third zone is a time delayed overreaching zone that reaches the next remote busbar and is used as back-up protection, set to 100% of the line length and 150% of the next longest line length. The delay was set to 1 s.

- Load encroachment zone is set as the equivalent impedance with 150% of emergency line ampere rating, 0.85 per unit bus voltage, and a load angle of 30 degrees.

The zone 1, zone 2, and zone 3 of distance relays are set up similarly as conventional distance relays,
the major difference being in the setting of load encroachment element. In conventional distance relays, any conditions whose apparent impedances are within zone 3 were considered as zone 3 fault. However, in today’s distance relays, only the conditions whose apparent impedances are within zone 3 but outside the load encroachment area are considered as fault, otherwise, they are considered as loads. To demonstrate the improvement of today’s distance relays and compare them with the proposed SVM-based smart relays, simulations were performed on the IEEE 118 bus system to mimic the disturbances which happened in 2003 blackout.

In the 2003 blackout, the first several disturbances/events happened in the systems were as follows [3]:

- 1:31 p.m: The Eastlake, Ohio generating plant shuts down.
- 2:02 p.m: The first of several 345 kV overhead transmission lines in northeast Ohio fails due to contact with a tree in Walton Hills, Ohio.
- 3:05 p.m: A 345 kV transmission line known as the Chamberlain-Harding line fails in Parma, south of Cleveland, due to a tree.

We mimicked these three events in the simulations conducted on the IEEE 118 bus system in the following manner:

- Shut down the generator on bus Tidd (bus 59);
Figure 4.14: The functional logic and the apparent impedance seen by line 63-59 in different system conditions

- Increase the load level to 130% of the designed level;

- Place a three-phase-to-ground fault on the transmission line between bus W. Kammer (bus 61) and bus Kammer (bus 64), then isolate this line in 100 ms by tripping its primary protection;

- Place a three-phase-to-ground fault on transmission line Muskingum (bus 66) and bus Summerfld (bus 67), then isolate this line in 100 ms by tripping its primary protection;

Fig. 4.14 shows the impedance of the protective relay on the transmission line between two buses on Tidd (bus 63 and bus 59):

As shown in Fig. 4.14, the green circles are apparent impedances which are seen by line 63-59 in normal condition, after the generator has been disconnected, under overloading condition, after the line 61-64 has been isolated, and after the line 66-67 has been isolated. As the system changed its conditions, the apparent impedance seen by the relay on line 63-59 was increasingly smaller and finally fell into zone 3 area. In other words, after a generator and two transmission lines have been isolated in the overloading condition, the relay on line 63-59 determined a zone 3 fault happened in the system and it tripped its breaker after 1 second delay. After the isolation of line 63-59, the whole system, which lacked a generator and three lines, was in an extremely overloaded condition. Such overloading conditions then triggered many other relays to trip their breakers. This is very similar to what happened in the last phase of 2003 blackout. Of course,
this is not the only case which can mimic the cascading failures in 2003. In fact, when a generator has been shut down and the load level is more than 130% of the designed level, there are quite a few combinations of two transmission lines which, if isolated, could cause a cascading failure. This clearly demonstrates that the distance relays, which are equipped with load encroachment, cannot survive when similar events which occurred in the 2003 blackout happen again in the system.

On the other hand, if the SVM-based smart relays are implemented in the system, things will be very different. We set up the SVM-based smart relays as follows:

- The SVM-based smart relays were trained with 6 features, which include magnitude of current, phase of current, magnitude of voltage, phase of voltage, real, and reactive power.

- The system conditions were classified into three different classes when training SVM relays, which are relaxed condition, primary protection, and backup protection. Relaxed condition included normal, far-away fault, and N-1 conditions. An SVM relay stays closed when it determines the system is in the relaxed condition. Primary protection class is defined when a fault is located in zone 1 and backup protection class is defined when a fault is located in zone 2 or in zone 3.

- An SVM-based smart relay trips the breaker in 100 ms when primary protection class is determined; it trips the breaker in 0.3 s to 1 s when backup protection class is determined; and it stays closed in the relaxed condition.

Simulations have been performed on the IEEE 118 bus system to mimic the first several disturbances which happened during the 2003 blackout. The generator on bus 59, line 61-64, and line 66-67 have been isolated from the system, respectively, and the load level has been increased to 130% of the designed level at the same time. Under these conditions, all the SVM relays determined that the system is in relaxed condition and were staying closed. Therefore, the whole system, in which three important components have been isolated, stayed connected and stable. In fact, one does not need to replace every distance relay with an SVM relay, since most of the distance relays function correctly. It is required to install the SVM relays in locations where the distance relays have a high probability to miss-trip. Such distance relays form only a small portion of the total number of relays used in the system.

To summarize, the simulation results are shown in Table 4.5. In these simulations, we randomly isolated a generator (column 1) and two strong transmission lines (there are 95 strong lines in total, 93 of which
were used in this simulation) in the 30% overloading conditions. The probability of cascading failures is calculated as follows:

\[
P(\text{cascading failures}) = \frac{\text{Number of cases which lead to cascading failures}}{\text{Total number of possible cases}}
\]

\[
= \frac{C}{\binom{N}{k}}
\]

where,

- \( C \) = Number of series of event which leads to cascading failures;
- \( N \) = Total number of transmission lines which are tested in this simulation; and
- \( k \) = Number of transmission lines which have be tripped.

For example, when the generator on bus 59 has been shut down and two transmission lines have been tripped, there are a total of \( \binom{93}{2} = 4278 \) possible series of events, 91 of which leads to cascading failures. Therefore, the probability of cascading failures is \( \frac{91}{4278} \times 100\% = 2.127\% \), as shown in the first entry of column 3 of Table 4.5. Comparing this number with the first entry of column 2 verifies that the implementation of load encroachment element in distance relays have indeed greatly decreased the probability that initial disturbances lead to cascading failures. When one compares the results in the 4th column of Table 4.5 with those in the 3rd column, the substantial improvement that can be achieved with SVM Relays is quite impressive. Table 4.5 thus provides compelling evidence that by replacing critical distance relays with SVM relays at few critical locations, one can significantly reduce the probability of cascading failures.

Table 4.5: Probability of Cascading Failure after Initial Disturbances in Systems with Different Protective Relays

<table>
<thead>
<tr>
<th>Generator</th>
<th>Probability of cascading failures (basic)</th>
<th>Probability of cascading failures (load encroachment)</th>
<th>Probability of cascading failures (SVM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bus 59</td>
<td>97.71%</td>
<td>2.127%</td>
<td>0</td>
</tr>
<tr>
<td>Bus 12</td>
<td>0.164%</td>
<td>0.095%</td>
<td>0.024%</td>
</tr>
<tr>
<td>Bus 54</td>
<td>0.187%</td>
<td>0.095%</td>
<td>0.024%</td>
</tr>
<tr>
<td>Bus 103</td>
<td>0.164%</td>
<td>0.117%</td>
<td>0</td>
</tr>
<tr>
<td>Bus 111</td>
<td>0.164%</td>
<td>0.117%</td>
<td>0</td>
</tr>
<tr>
<td>Bus 46</td>
<td>0.117%</td>
<td>0.070%</td>
<td>0</td>
</tr>
</tbody>
</table>
4.4 Summary

In this Chapter, we have compared the performance of the proposed SVM-based smart relays with the conventional distance relays. The case study on the 7-bus system from an Europe network described the performance of relays in the network with distributed generators. In this case, a conventional distance relay cannot discriminate the system conditions clearly and find a proper setting for executing the correct and selective tripping, while an SVM-based smart relay can. We have performed a benchmarking performance between SVM-based smart relays and conventional distance relays during cascading failures. We have demonstrated that the implementation of smart relays at critical locations can decrease the chance of cascading failures significantly. Although the critical relays have to be identified via an exhaustive search over the whole network, important properties of these critical relays have been presented in this Chapter which can help to reduce the calculation complexity during such an exhaustive search. We have also performed simulations to mimic the event series which happened in 2003 blackout to compare the performance of SVM-based smart relays, conventional distance relays and today’s distance relays with load encroachment. We have demonstrated that the implementation of load encroachment element, which is required by NERC after 2003 blackout, can greatly decrease the probability of cascading failure. However, by implementing SVM-based smart relays in critical locations, such probability can be further decreased.
Chapter 5

Handling of Missing Measurements and Updating

5.1 Handling of Missing Measurements

As stated in the previous section, in the proposed SVM-based smart relays, there are six features. These features are magnitude of current, phase of current, magnitude of voltage, phase of voltage, real, and reactive power. The best accuracy can be achieved when all of these six features are available. However, in practice, real data are not perfect and there are cases where some of the measurements are missing because of the unavailability of communication channels or malfunction of meters. Under these circumstances, we need to propose a mechanism which can handle the missing measurements. In this dissertation, we apply the K-nearest Neighbor (K-NN) concept to estimate the missing measurements before SVM classification, which is usually known as imputation of missing data with K-NN.

5.1.1 Introduction of K-Nearest Neighboring

In pattern recognition, the K-nearest neighbors algorithm (K-NN) is a common and intuitive classification mechanism based on closest training examples in the feature space. k-NN is a type of instance-based learning where the function is only approximated locally and all computation is deferred until classification. The k-nearest neighbor algorithm is amongst the simplest of all machine learning algorithms: an object is classified by a majority vote of its neighbors, with the object being assigned to the class most common amongst
its k nearest neighbors (k is a positive integer, typically small). If \( k = 1 \), then the object is simply assigned to the class of its nearest neighbor.

The same method can be used for regression, by simply assigning the property value for the object to be the weighted average of the values of its \( k \) nearest neighbors. A weight factor can be assigned to each neighbor, so that the nearer neighbors contribute more to the average than the more distant ones. The weighting scheme used in this dissertation is to give each neighbor a weight of \( 1/d \), where \( d \) is the distance to the neighbor. This scheme is a simple generalization of linear interpolation.

The neighbors are taken from a set of objects for which the correct classification (or, in the case of regression, the value of the property) is known. In this dissertation, we use the training set of SVM classifier. We use the example in Fig. 5.1 to explain the process of estimating missing measurements via K-NN. Assume we want to use three features to do SVM classification, namely, reactive power, current, and voltage, but the measurements for reactive power are unavailable. Suppose we are in a situation where we cannot simply abandon this incomplete set of data and wait for the next set of data. We therefore have to use K-NN to estimate the value of reactive power before we use this set of data to do SVM classification. The steps for estimating the value of reactive power via K-NN are:

- We have an incomplete set of measurements whereby the value of the reactive power is missing. Our goal in applying the K-NN is to estimate the value of reactive power. We refer to this set of measurements as target data. These target data are in two dimensions which are voltage and current.

- At the same time, we have all the training data which are complete in three dimensions which include voltage, current, and reactive power.

- Compute the Euclidean distance from the target data to those training data in the 2D space, on which the x-axis is voltage and the y-axis is current, as shown in Fig. 5.1. For each training data \( i \), the distance to the target data is \( d_i \).

- Pick up \( k \) training data with the shortest distance from the target data.

- Estimate the value of reactive power with the target data as

\[
Q_{est} = \sum_{i=1}^{k} w_i Q_i
\]  

(5.1)
CHAPTER 5. HANDLING OF MISSING MEASUREMENTS AND UPDATING

Figure 5.1: Estimate the missing measurement via K-NN

Where \( w_i = \frac{1/d_i}{\sum_{1}^{k} 1/d_i} \) is the weight factor for training data \( i \), and \( Q_i \) is the value for reactive power of data \( i \).

The factor \( k \) in K-NN mechanism is very important. The accuracy of the K-NN algorithm can be severely degraded by the chosen value of \( k \). The best choice of \( k \) depends upon the data. Generally, larger values of \( k \) reduce the effect of noise on the regression, but make boundaries between classes less distinct. For example, in Fig. 5.1, green dot is the target data and other dots are the training datas. Blue and red dots are from different classes. After applying K-NN to locate the nearest neighbors of the target data, the solid circle is the boundary when \( k = 2 \), and the dotted circle is generated when \( k = 4 \). Larger \( k \) is more tolerant to noise but may introduce more data from different classes and therefore decease the accuracy of decision making. A good \( k \) value can be selected by various heuristic techniques. We use cross-validation in this dissertation. Detailed results will be shown in the simulation section.
5.1.2 Estimation of Missing Measurements

As introduced in the previous section, K-NN is applied to estimate the value of missing measurements.

Parameter Selection

$k$ is the most important factor in the K-NN mechanism. The accuracy of the K-NN algorithm can be critically dependent on the chosen $k$ value. Generally speaking, a too large $k$ value can count too many data from different classes and therefore decease the accuracy of decision making, while a too small $k$ can make the regression sensitive to noise. A good $k$ value can be selected by various heuristic techniques. We use cross-validation in this dissertation. Fig. 5.2 shows the estimation error corresponding to different $k$ values. The error rate is calculated by

$$ Error = \frac{||Measurement_{est} - Measurement_{true}||}{||Measurement_{true}||} $$

(5.2)

![Figure 5.2: Parameter selection on K-NN](image)

As shown in Fig. 5.2 and Table 5.1, the best $k$ to estimate different features are different.

| Missing Measurement | $|I|$ | $\angle I$ | $|V|$ | $\angle V$ | $P$ | $Q$ |
|---------------------|-----|----------|-----|----------|----|----|
| Best $k$            | 8   | 4        | 3   | 3        | 8  | 4  |

Table 5.1: The best value of parameter $k$ on K-NN with different missing features

Obviously, the estimation of real and reactive power are relatively worse than the estimation of other features. This is because real and reactive power are the strongest two features, when one of them is missing,
it is hard to estimate the value by finding the nearest neighbors based on weaker features. This can be observed in Fig. 5.3 and Fig. 5.4. In Fig. 5.3 and Fig. 5.4 the red and green cross stand for the true and estimation value of the target data, accordingly; the circles stand for the nearest neighbors of the target data when $k$ varies from 1 to 10. The red circles are the instance from the same class with the target data, while the blue circles are the neighbors from different classes. Obviously, when $k$ is carefully selected, the missing measurements, whether it is the reactive power or the phase of current, can be well estimated. However, the real and reactive power are strong features which are very discriminative between different classes. In other words, when considering the real and reactive power features, the instances from the same class tend to cluster closely. Therefore, when the value of real and reactive power are known and are used to estimate other measurements, the nearest neighbors are very likely to be found from the same classes. Therefore, the estimation value is close to the true value and the target instance is very likely to be correctly classified. On the other hand, the magnitude, phase of current and magnitude of voltage are weak features, which means the instances from the same class are scattered in these dimensions. Therefore, when estimating other measurements based on these three features, the target instance can be easily mis-classified.

Testing Accuracy

The comparison of the accuracy with and without missing measurement is shown in Table 5.2. The simulation was performed on the IEEE 118 bus system and the testing accuracy with complete data set is 97.7%
CHAPTER 5. HANDLING OF MISSING MEASUREMENTS AND UPDATING

Figure 5.4: The estimation and the nearest neighbors when the phase of current measurement is missing in the target data

Table 5.2: The testing accuracy with missing measurement in 6 features

| Missing Measurements | $|I|$ | $\angle I$ | $|V|$ | $\angle V$ | $P$ | $Q$ |
|----------------------|-----|----------|-----|----------|-----|-----|
| Accuracy             | 96.85% | 96.20% | 92.3% | 77.75% & | 68.3% | 68.2% |

Obviously, when the missing measurement is one of phase of voltage, real and reactive power, the accuracy is very low and we may need to abandon the whole data set. This is because the phase of voltage, real and reactive power are strong features, while the magnitude of voltage and phase of current are weak features. In the case that we need to estimate the strong measurements based on the weak measurements, the classification accuracy is very low. Therefore, we may need to abandon the incomplete data set. However, we could modify the scheme a little so that we can have a better estimation and better classification mechanism. Instead of using all the known measurements to locate the nearest neighbor, we only use the measurements from relatively strong features. For example, in a complete data set, we have 6 measurements which are magnitude of current, phase of current, magnitude of voltage, phase of voltage, real and reactive power. After doing feature selection, it is known that these 6 features can be ordered as reactive power, real
power, phase of voltage, magnitude of current, magnitude of voltage, phase of current, from the best features to the worst. Therefore, when facing with missing measurements and a imputing is needed, we can use the best three available features to locate the nearest neighbors. For example, when the real power measurement is missing, we use reactive power, phase of voltage and magnitude of current to locate the neighbors; when the magnitude of voltage measurement is missing, we use reactive power, real power and phase of voltage to find the neighbors. By doing these, the effect of weak features can be decreased. The accuracy is shown in Table 5.3.

| Missing Measurements | $|I|$ | $\angle I$ | $|V|$ | $\angle V$ | $P$ | $Q$ |
|----------------------|-----|----------|-----|---------|-----|-----|
| Accuracy             | 96.86% | 96.22% | 92.28% | 84.35% | 88.2% | 86.4% |

When only four features, which are magnitude of current, phase of voltage, real and reactive power, are considered, the testing accuracy is shown on Table 5.4.

| Missing Measurements | $|I|$ | $\angle I$ | $|V|$ | $\angle V$ | $P$ | $Q$ |
|----------------------|-----|----------|-----|---------|-----|-----|
| Accuracy             | 97.2% | 95.7% | 94.06% | 94.66% |     |     |

5.2 Dynamic updating on SVM-based smart relays

Many machine learning applications employ algorithms for learning concept descriptions. Usually, as time changes, new instances obtained are added to the training dataset. Then the concept description should be properly updated to take the new instances into consideration. These applications often face the challenges that real world concepts tend to change over time, i.e., a concept description learned from all previous examples may not hold for future data. Hence, some of the old observations that are out-of-date have to be forgotten or omitted. This problem is known as concept drift [51], formally defined as the probability that two subsequent concepts disagree on a randomly drawn example [50].

Two kinds of concept drift that may occur in the real power systems are normally distinguished in the literature: 1) sudden (abrupt, instantaneous), for example, the concept drifting caused by sudden disturbances
in systems, and 2) gradual concept drift, for example, the concept drifting caused by the gradually change of loads and generations during day. In this section, we aim to develop a system that accurately detect the first case and be able to adapt to the second type of concept changes quickly.

In the implementation of SVM-based smart relays, the concept drift can caused by the topology changes of the power systems. This is a important problem which need to be solved. The two kinds of concept drift which are described in last paragraph can both happen in power systems. During cascading failures, the system topology changes suddenly and some time in an unexpected matter. This can be treated as the sudden concept drift. On the other hand, the long-term system construction, which include the connecting of a transmission line or a disconnecting with a generator, can be considered as a gradual concept drift.

In machine learning, concept drifting is usually handled by the following two categories of ideas: 1) detecting concepts change by time windows of fixed or adaptive size on the training data, e.g., Klinkenberg et al. developed a support vector machine based window size adjustment approach that results in a low predictive error [52]; 2) Weigh data or parts of the hypothesis according to their age and/or utility for the classification task. For example, Koychev and his colleagues designed a time-base forgetting function, which makes the last observations more significant for learning algorithms than old ones [52]. The second category, known as information filtering, use many heuristic approaches but it involves parameter tuning and often not generalizable across data. On the other hand, the methods in the first category are usually limited by computational complexity and their application has not been fully appreciated. In this dissertation, the concept drift is detected by a $\epsilon\alpha$-estimator, while the data are weighed by their age.

5.2.1 Proposed Approach

Different from offline single-source learning tasks, which treat arriving instances as equally important contributors to the final concept, we assign higher weights to most last observations that are more significant for learning algorithms than old ones. Thus, we seek learners that learn and unlearn examples to update model quickly. Here, we present a three phase algorithm outlined as incremental learning, $\epsilon\alpha$-estimator and out-of-date data prune.

- Incremental Learning

Support Vector Machine (SVM), for being theoretically well founded in statistical machine learning is used as our core algorithm.
CHAPTER 5. HANDLING OF MISSING MEASUREMENTS AND UPDATING

SVM Classification Mechanism, which can deliver a maximum marginal classifier, in fact, is a quadratic optimization problem:

\[
\begin{align*}
\min_{w,y} & \quad \frac{1}{2}w'w + \nu e'y \\
\text{s.t.} & \quad D(Aw - e\gamma) + y \geq \epsilon \quad \text{and} \quad y \geq 0
\end{align*}
\]

which minimizes the reciprocal of the margin (i.e., \(w'w\)) and the error (i.e., \(e'y\)). The slack variable \(y\) is larger than zero when the point is on the wrong side or within the margin area. The soft margin parameter \(\nu\) is tuned to balance the margin size and the error. The weight vector \(w\) and the bias \(\gamma\) will be computed by this optimization problem. The class of a new data \(x\) will be determined by \(f(x) = w'x - \gamma\), where the class is positive if \(f(x) > 0\), otherwise negative.

Applying the Lagrange multipliers on the primal problem, we usually solve the following dual problem by a QP solver [4]:

\[
\begin{align*}
\min_{\alpha} & \quad \frac{1}{2}\alpha'Q\alpha - e'\alpha \\
\text{s.t.} & \quad 0 \leq \alpha_i \leq \nu \quad \text{and} \quad \sum_i d_i \alpha_i = 0, \quad i = 0, \ldots, m
\end{align*}
\]

However, this original form of SVM needs a complete training data set, which makes SVM to be limited in many applications. Some of the applications deal with huge amount of data, which may make the SVM training impossible. Some of the applications require a decision making hyperplane which can be updated online, quickly and accurately, to be robust with the change of the conditions. A sophisticated robust online SVM mechanism which can learn the target concept incrementally is therefore needed. The former case is a purely incremental learning to keep low memory and low time consumption, while the latter case is dealing with a changed target concept, which is called concept drifting. In this case, old examples may be misleading. Some incremental SVM machinists have been reported in [53] and [54]. [53] reported a naive incremental learning which only keep the support vectors from previous iteration. [54] introduced an exact incremental and decremental SVM (Exact
I & D SVM) learning mechanism which ensure the accuracy and complete the hyperplane updating quickly.

- **Naive Incremental Support Vector Machine**

  An SVM does not depend on all the training data but only its Support Vectors. As the number of Support Vectors typically is very small compared to the number of training examples, SVMs promise to be an effective tool for incremental learning by compressing the data of the previous batches to their Support Vectors. This approach is shown in [53] with a comparable accuracy to non-incrementally trained SVM.

  The way the experiment was designed as follows:

  **Algorithm 1** Naive Incremental Support Vector Machine

  **Input:** \( n \) (Training data \( TR \)) (Testing data \( TE \))

  1: split \( TR \) into \( n \) parts: \( TR_i \), where \( i = 1, 2 \ldots n \)
  2: initialize \( SV_0 \) as an empty set;
  3: for \( i = 1 \) : \( n \) do
  4: train \( SVM_i \) on data \( [SV_{i-1}, TR_i] \);
  5: \( SV_i \) = Support vectors which are gotten in the training;
  6: test \( SVM_i \) in \( TE \);
  7: end for

  **Output:** number of SVs and testing accuracy in each iteration;

- **Exact Incremental and Decremented Support Vector Machine**

  The dual problem with equation (5.3) can be translated into the following new format by introducing a Lagrangian multiplier \( b \).

  \[
  W = \min_{0 \leq \alpha_i \leq C} \frac{1}{2} \alpha^T Q \alpha - e^T \alpha + b * y * \alpha 
  \]

  (5.7)

  Here \( Q_{i,j} = \sum y_i y_j K(x_i, x_j) \), according the KKT condition
CHAPTER 5. HANDLING OF MISSING MEASUREMENTS AND UPDATING

\[ g_i = W_{\alpha_i}^T = \sum_j Q_{ij}a_j + y_i b - 1 = \begin{cases} > 0 & a_i = 0 \\ = 0 & 0 \leq a_i \leq C \\ < 0 & a_i = C \end{cases} \]

(5.8)

\[ h = W_b^T = \sum_j y_j \alpha_j = 0 \]

(5.9)

Equation (5.8) divided the array into three different groups, margin support vectors set \( S (g_i = 0) \), error support vectors set \( E \) violating the margin \( (g_i < 0) \) and the remaining is considered as reserved set \( R \) as potential candidates for error and margin vectors. As new data \( U_i \) came into the system, we wish to come up with a solution to update the kernel matrix with least perturbation on the trained kernel matrix, which is known as adiabatic increments in [54].

We have already calculated the partial derivatives wrt \( \{\alpha_i, b\} \) in the entire \( S \) set using equation (5.7).

The perturbation on both \( g_i \) and \( h \) can be expressed in the following form:

\[
\Delta g_i = \sum_{k \in S} Q_{ik} \Delta \alpha_k + \sum_{l \in U} Q_{il} \Delta \alpha_l + y_i \Delta b, \text{ for } i \in S
\]

(5.10)

\[
\Delta h = \sum_{k \in S} y_k \alpha_k + \sum_{l \in U} y_l \alpha_l = 0
\]

(5.11)

If we replace \( \Delta \alpha_k = \beta_k \Delta p (k \in S) \), \( \Delta \alpha_l = \lambda_k \Delta p (l \in U) \) and \( \Delta b = \beta \Delta p \) in the above equation, we obtain the following formulation:

\[
\gamma_i = \frac{\Delta g_i}{\Delta p} = \sum_{k \in S} Q_{ik} \beta_k + \sum_{l \in U} Q_{il} \lambda_l + y_i \beta, \text{ for } i \in S
\]

(5.12)

\[
\frac{\Delta h}{\Delta p} = \sum_{k \in S} y_k \beta_k + \sum_{l \in U} y_l \lambda_l = 0
\]

(5.13)

Putting them into a matrix format, we obtain the following form \( Q \beta = -\sum_{l \in U} \lambda_l V_l \). Here, \( \beta = [\beta, \beta_{s1}, \ldots, \beta_{sn}]^T \) and \( V_l = [y_l, Q_{s1l}, \ldots, Q_{snl}] \) and
CHAPTER 5. HANDLING OF MISSING MEASUREMENTS AND UPDATING

\[
Q = \begin{pmatrix}
0 & y_{s1} & \ldots & y_{sn} \\
y_{s1} & Q_{s1, s1} & \ldots & Q_{s1, sn} \\
\vdots & \vdots & \ddots & \vdots \\
y_{sn} & Q_{sn, s1} & \ldots & Q_{sn, sn}
\end{pmatrix}
\] (5.14)

What we really need here is the inverse of \(Q\), namely \(R = Q^{-1}\) to compute the sensitivities \(\beta = -\sum_{l \in U} \lambda_l R V_l\). The trick here is that we don’t really need to invert the entire matrix to incorporate the incoming support vectors. Given the previous inverse matrix \(R_n\) and candidate \(k\) to be added to the working margin set \(S\), \(R_{n+1}\) is extended by applying Sherman- Morrison-Woodbury formula for block matrix inversion. We obtain:

\[
R_{n+1} = \begin{pmatrix}
R_n & 0 \\
0 & 0
\end{pmatrix} + \frac{1}{K} \begin{pmatrix}
\beta_k \\
1
\end{pmatrix} \begin{pmatrix}
\beta_k^T & 1
\end{pmatrix}
\] (5.15)

where, \(K = Q_{kk} - \eta_k^T R \eta_k, \eta_k = [y_k, Q_{k:s}]\), note that \(s\) refers to the elements in set \(S\). The update of the inverse matrix thus only involves expansion with zero row and column and addition of a rank-one matrix obtained via a matrix-vector multiplication \([52]\). The running time needed for such an update is quadratic in the size of \(calQ\) and much faster than explicit form inversion which make online SVM training possible.

The following is a very high level description of the incremental procedure for SVM training.

**Algorithm 2 Incremental Support Vector Machine, adding a new vector \(c\)**

1: Initialize \(\alpha_c\) to zero.
2: Calculate \(g_c\), if \(g_c \leq 0\), terminate because \(c\) is not a margin or error vector.
3: Otherwise, apply the largest possible increase on \(\alpha_c\) so that one of the following three conditions occurs.
   (a) \(g_c = 0\): Add \(c\) to the margin set \(S\), update \(R\) by removing \(c\) and terminate
   (b) \(\alpha_c = C\): Add \(C\) to the error set \(E\), and terminate.
   (c) Elements of previous data migrate across \(S, E\) and \(R\). Update membership of elements accordingly.

- \(\epsilon\alpha\)-estimator

\(\epsilon\alpha\)-estimators are theoretical bounds of leave-one-out estimation. While the leave-one-out estimate is
CHAPTER 5. HANDLING OF MISSING MEASUREMENTS AND UPDATING

statistically almost unbiased, it is very expensive to compute. $\epsilon\alpha$-estimators overcome this problem by calculating an upper bound on the number of leave-one-out errors instead of calculating them brute force [55]. For stable soft-margin SVMs, the $\epsilon\alpha$-estimators of the error rate is

$$Err_{\epsilon\alpha}(h_L) = \frac{d}{n} \text{ with } d = |i : (\mu\alpha_iR^i + \epsilon_i) \geq 1|$$

(5.16)

with $\mu = 2$ and $\epsilon, \alpha$ are the solution of the primal and dual optimization problem of SVMs. In our application, we will incrementally learn the training data backwards, and calculate the $\epsilon\alpha$-estimators for each additional data.

In the last phrase, when the $\epsilon\alpha$-estimators have been obtained for each previous training data during incremental learning, we conduct

- Out-of-Date Data Prune

As explained, we are expecting a jump up in the $\epsilon\alpha$-estimators curve when concept drifting happens. All the data after this jumping point should be pruned, because those can only decrease the accuracy. To locate the jumping point, we use a step function to fit the $\epsilon\alpha$-estimators curve. Thus the problem can be stated as:

$$\min_{p,b,x} |px + b - \epsilon\alpha Estimator|^2$$

s.t.

$$p > 0;$$

$$b \geq 0;$$

and $x_i \in \{-1, 1\}$

This is a convex optimization problem and can be solved quickly. The point where $x_i$ jumps from -1 to 1 indicates the concept drifting point.

5.2.2 Simulation Results

A simulation was done with mushrooms dataset [56], which is a public test dataset. Fig. 5.2.2 demonstrates the performance of the $\epsilon\alpha$-estimator. In both the cases when the training data contains a concept drift and
is without concept drift, the \( \epsilon \alpha \)-estimator bounds the leave-one-out error completely. Moreover, the \( \epsilon \alpha \)-estimator follows the same pattern with the leave-on-out error, which is very critical for this project. This makes it possible to detect the concept drift and prune the out-of-date data by only taking a look at the \( \epsilon \alpha \)-estimator possible.

Fig. 5.6 shows how a step function is used for fitting the \( \epsilon \alpha \)-estimation curve and locate the concept drifting point. Compared with the real drifting point which is data 403, the step function locate the drifting at data 413, which is very close to its real value.

To explain what’s happening in Fig. 5.7, we calculate time complexity bound on both algorithms [50] and ours. In [50], suspicious data, \( \text{size} = |M| \) is divide into \( N \) piece and it trains a combination of one piece of incoming data \( T \) with 1, 2, ..., \( N \) previous pieces of data for \( N \) iterations. The concept drift happens at the point where the minimum Leave-One-Out Error is reported. The complexity of each training process can be roughly bounded by \( O(M^3) \), thus the whole algorithm has a complexity of \( O(NM^3) \).

Our method uses incremental SVM and only goes through previous data in a single pass and each kernel matrix expansion costs \( O(K^2) \) if the incoming data turns out to be a support vector, where \( K \) is the size of current support vectors. So the algorithm is bounded by \( O(MK^2) \), where \( K << M \) in general. That explains why our algorithm has a great advantage in time and makes online concept drifting possible.

Figure 5.5: Comparison of \( \epsilon \alpha \)-estimators as a bound against leave one out error. (a) Without concept drift. (b) with concept drift.

Note that data enters the system once a time. Decisions of whether an incoming vector should belong to error, marginal and reserved sets are made on the fly. In the meantime of classifying the new data, previous samples get evaluated whether they were transferred across from different sets.

To summarize, \( \epsilon \alpha \)-estimators based incremental learning is a potential way to update the SVM classifi-
Figure 5.6: Locating concept drifting point with dynamic step functions. Simulation results indicate that a fitted step function locates a concept drifting which occurs at data point 413 whereas its real value is 403.

Figure 5.7: Simulation results indicate that the training time with bunch learning method in [50] increases sharply as the bunch size decreases while our algorithm requires constant time. Our incremental method has a great advantage in terms of time consumption.
cation based smart relays. In practice, after a relay has been installed, it was not left alone but was checked and updated automatically. The decision making results of a relay along time will be stored in memory and be used to decide whether a updating is needed. Once misclassifications happened continuously in a single relay or when a maintenance cycle is due, a $\epsilon\alpha$-estimators based incremental learning will take place to update the decision making boundary of relays. However, since it takes several minutes to accomplish a single updating, the $\epsilon\alpha$-estimators based incremental learning is only suitable for the boundary updating caused by gradual changes in the power system. It cannot match the speed of cascading failures in or before blackouts. Moreover, The updating process requires the aid of the communications with SCADA. This is because a single relay, which can only obtain the measurements from its location, cannot see the real situation of the system when the system topology is changing or changed. Only the SCADA, which collects information from the whole system can identify the system conditions correctly. Therefore, although a relay can obtain the attributes for each instance for updating from its local measurements, it needs to be given the class label by SCADA. This is another reason that the automatic updating of SVM-relays can be achieved corresponding to the long term changes in the power system, but cannot be performed during the cascading failures.

5.3 Summary

In this Chapter, we have proposed mechanisms for handling the missing data and updating problems that might occur with SVM-based smart relays. The missing data problem can be effectively handled by applying K-NN imputation, while the dynamic updating of SVM-based relays can be achieved via concept drifting detection and incremental learning. However, the updating of SVM relays requires communications with SCADA. This is because only the information from SCADA can offer the ground truth of class label for each set of measurements. Further research is needed on how to implement the dynamic updating of SVM-based smart relays.
Chapter 6

Transient state SVM-based Smart Relays

Static state measurement based SVM smart relays was stated in previous chapters. These SVM smart relays were designed for detecting and locating faults. They were not designed to prevent system-wide effects. However, there are some system-wide effects before and/or during blackouts which need to be carefully considered. In this Chapter, to prevent system-wide effects, a transient measurements based SVM smart relays will be proposed, as an extension to static state measurements based SVM smart relays.

6.1 Proposed Approach

The power grid works as a whole system, therefore a disturbance will not only cause local response but also system-wide effects. Thus, besides local equipment protection relays, special protection schemes (SPS) are needed which should be designed for system-wide protection.

Typically, major power system blackouts have been initiated by local disturbances that subsequently cascaded across the transmission networks. Significant studies were performed to understand their causes. It can be inferred that major power system contingencies typically comprise three phases depending on their duration: the initial phase where temporary system faults occur, which is rapidly cleared in 100 milliseconds; the intermediate phase where the system gets partitioned into undesirable parts in seconds; and the final phase where load and generation imbalance causes a blackout in minutes [57]. Since the late 1990s, power systems have been pushed closer to their limits, resulting in a growing risk for a local failure to cascade into a large-scale catastrophic blackout. The most common triggering fault of such an event is a short-circuit that occurs on high voltage (HV) or extra-high voltage (EHV) transmission lines of the system.
At the inception of a disturbance, the relays located on faulty transmission lines attempt to clear the fault. This induces variations of the electrical power generator outputs while the generator mechanical inputs remain almost constant. The resulting effect of this power imbalance is the formation of groups of coherent generators operating at different speeds, swinging one against the others. Eventually this may lead to a loss of synchronism and the splitting of the network. However, the islands formed in this manner may not have a balanced operation in terms of generation and load, which makes the failure to propagate further until a complete collapse of the system. To prevent such events, utilities have installed special protection schemes based on under-frequency and under-voltage relays that perform load and generation shedding and line tripping. To prevent cascading failures from propagating further throughout a power system, various controlled separation schemes have been proposed and installed by many utilities. A good scheme reduces the impact of an outage on the customers and the economy of the affected area while reducing the possibility of damage to equipment. Currently, controlled separation schemes have been implemented mainly in elongated and isolated power systems to split the system along pre-determined boundaries.

However, the special protection schemes are using conventional relays, while redefining their coordination principle, to cope with the cascading failure. Moreover, the islanding algorithm used by SPS is pre-defined and may not be the optimal solution at all times. Therefore, we proposed Support Vector Machine Classification based relays with transient measurements (TS-SVMCB relays) for system wide stability control. TS-SVMCB relay can be used as binary fast controller in the system for fast stabilization after disturbances. It uses the features extracted from the time-continuous transient measurements to make decisions. Fig. 6.1 shows a typical transient measurements of current magnitude in normal, during fault and after fault conditions. Now suppose a case where a fault happened at time 166 ms and been cleared at time 332 ms\(^5\). Features will be extracted from these measurements to describe the shape of this time-continuous function, including the peak values and time constant. Classes are defined by the impact after disturbances as

- Class 1: Serious impact;
- Class 2: Moderate impact;
- Class 3: Minor impact; Extra relays will be tripped if serious impact is detected and warnings will be sent to critical relays when moderate impact is detected.

\(^5\)We simulated to sample the current magnitude at the rate of 6000Hz. Therefore, 500 in the Y-axis equals to 83 ms, 1000 equals to 166 ms, and 2000 equals to 332 ms, respectively.
6.2 Simulation Results

Simulation has been performed on the IEEE 14 bus system [38]. The typical transient current measurement is shown in Fig. 6.2. This time-continuous measurements consist of three parts: the measurements during the normal condition, during fault, and after fault condition. The normal condition measurements are made when there is no short-circuit fault in the system. The other measurements are made when there is a short-circuit fault in the system and after its clearance. The short-circuit fault mentioned here is not necessarily a local fault.

As explained in last sub-section, we focus on the shape of this time-continuous transient measurements to determine whether extra relays should be tripped. Therefore, 9 features are extracted from these transient measurements to describe the shape of this time-continuous function and to make decisions. The first 5 fea-
features are extracted directly from these measurements, they are: maximum current during faulty condition, the amplitude of vibration during faulty condition, the minimum current during faulty condition, the maximum current after fault clearance, and the amplitude of vibration after fault clearance. These 5 features indicate the basic shape of the measurements, including the peak values and the amplitude of vibration in different conditions. The other 4 features are extracted from processed measurements. Since in most of the cases, the current in the systems dies out as an exponential function, we take logarithm of the measurements of current magnitude which are during and after the fault for the other 4 features. We do straight line curve fitting on the value after taking logarithm to get two lines, which are shown in Fig. 6.3. Then we use the two parameters of each line, which are four in total, as the other 4 features. They are the slope and intersection of fitting curve 1 and 2. These 4 features describe the speed of decay of the current after fault clearance. Combined with the first 5 features, we can describe the shape of this time-continuous function completely.

Besides 9 features, we define 3 classes as:

- **Class 1**: Serious impact; in this class, we have to trip more extra relays to prevent the system from collapse since the after fault current will trigger instability in the rest of the system.

- **Class 2**: Moderate impact; in this class, we have to warn critical relays and/or may need to trip more relays, since the after fault current may trigger hidden faults if the critical relays cannot handle the condition precisely.

- **Class 3**: Minor impact; in this class, we don’t need to take extra action other than only tripping the local protections.
CHAPTER 6. TRANSIENT STATE SVM-BASED SMART RELAYS

Simulation results indicated that the accuracy is above 94% when three classes are correctly defined.

6.3 Obstacles to implementing Transient Measurement Based SVM relays

The obstacle to implementing SVM mechanism in a transient measurement based smart relay is the assignment of classes. As is well known, SVM is a statistical decision making mechanism which is first trained by the training data. Each training data is combined with several features and a class label. The performance of the smart relays with SVM classifiers, after their installation on line, is highly dependent on the off-line training process. Therefore, it is critical to assign each training instance with a correct and proper class label. The class label assigned on training data represents the system conditions corresponding to the measurements for each instance and indicates the reaction a relay needs to take at that time. The class assignment for system-wide protection is not as simple as the class assignment for equipment protection. As described in previous chapters, the smart SVM protective relays based on static state measurements is designed for detecting and locating local faults. The class labels of their training data are assigned based on the location of the faults, in a one-to-one mapping manner. The class assignment for system-wide protective relays are much more complex. In the simulation which is described in previous section, the class labels were assigned by an exhaustive study of the IEEE 14 bus system. In the study, any tripping of breaker which leads to instability at any bus was defined as class 1; any tripping of breaker which causes the system to take a long time to achieve a new equilibrium point was considered as class 2; and others were considered as class 3. However, in practical system-wide protection, the class labels which represent the system conditions cannot always be determined by a single or current snapshot of the measurements. The situation of the system on time \( t \) is related to 1) the system structure at time \( t \); 2) the control or protection purpose at time \( t \) or after; and 3) even the condition of the system on time \( t - \) or before. Therefore, the same condition of the system can be treated differently in different cases. Hence, the assignment of class labels to each instance is difficult, which needs to take the control purpose and the system structure at that time into consideration. This is the main obstacle in implementing SVM in transient measurement based relays for system wide protection.
6.4 Summary

In this Chapter, we have proposed another possible application of SVMs in power protection relays, based on the transient state measurements. This might be a potential way to do system-wide protection, and therefore significantly decrease the chance of cascading failures and improve the system’s stability after disturbances. However, the complexity of the implementation of these transient measurements based smart relays is unclear, because the training process of these relays can be extremely complex. Further research is needed for finding ways to simplify the training process. This, in turn, could make the use of SVM-based protective relays a viable option for system-wide protection.
Chapter 7

Conclusion

In this dissertation, we have shown that a plausible way for mitigating the cascade of failures that lead to the infamous blackout problem is to design and deploy smart relays based on machine learning techniques. In particular, it is shown that smart relays using support vector machines could be instrumental in monitoring, detecting, and locating the initial faults on transmission lines. Based on measurements of current, voltage, real and reactive power at the local level, SVM-based smart relays can make intelligent decisions on whether and when to trip a transmission line. By avoiding unnecessary trips, SVM-relays can help SCADA to dispatch and re-distribute the excess power to several different transmission lines as opposed to overloading a single transmission line. This, in turn, can provide the much needed “breathing time” for SCADA to maintain the stable (N-1) mode of operations, thus avoiding the cascade of failures and a large-scale blackout.

In the remainder of this chapter we first summarize the main contributions of this PhD thesis. Then, we discuss the practical implications of these contributions and the key findings of our study. Finally, we point out some possible future directions for extending the work reported in this thesis.

7.1 Main Contributions of the Thesis

In this dissertation, we have shown that a plausible way to mitigate the cascading of failures that lead to the infamous blackout problem is to design and deploy smart relays based on machine learning techniques. In particular, it is shown that smart relays using support vector machines could be instrumental in monitoring, detecting, and locating the initial faults on distribution and transmission lines. Based on measurements of
current, voltage, real and reactive power at the local level, SVM-based smart relays can make intelligent
decisions on whether and when to trip a transmission line. By avoiding unnecessary trips, SVM-relays
can help SCADA to dispatch and re-distribute the excess power to several different transmission lines as
opposed to overloading a single transmission line. This in turn can provide the much needed “breathing
time” for SCADA to maintain the stable (N-1) mode of operations. Simulations demonstrated some rarely
used measurements, such as real and reactive power, should be considered during fault detection. In spite of
their rare utilization, they are strongly capable in discriminating the system conditions. We have also shown
in this thesis that the missing data and updating problems associated with SVM-based smart relays can be
solved, with the aid of communications with SCADA. Benchmarking performance between SVM-based
smart relays and conventional relays demonstrated that, by installing SVM-based smart relays on critical
locations, cascade of failures and a large-scale blackout can be avoided or mitigated.

More specifically, the main contributions of this dissertation can be summarized as follows:

• We have proposed, for the first time, the use of hypothesis testing and SVM technique to improve the
decision making capability of a single protective relay. This approach radically changes the functional
logic of protective relays, making them much more intelligent devices which, in turn, can help make
the power systems much more intelligent.

• We have shown the importance of real and reactive power in the protection systems when they make
decisions. This is a fairly counter-intuitive result as these two measurements are rarely used in today’s
protective relays for fault detection.

• We have demonstrated the scalability of SVM-based smart relay. We have also proposed a methodol-
dogy for handling missing data and for updating the SVM-based smart relays with the aid of SCADA.
All of these make the practical implementation of SVM-based smart relays possible.

• We have demonstrated via extensive simulations that by replacing the critical conventional relays by
SVM-based smart relays, the probability of cascading failures can be significantly decreased. This
result is quite significant in and of itself, as it implied that a massive replacement of existing conven-
tional relays is not needed for the proposed scheme’s success. Instead, it would be sufficient to use
SVM-based smart relays at critical locations which form only a very small portion of the total number
of relays that exist in the current power grid.
• We have proposed to utilize the SVM classification technique with transient state measurements for system-wide protection. Although the class assignment can cause problems to the practical implementation of these smart relays, this approach may improve the system stability after disturbances.

### 7.2 Practical Implications

The SVM-based smart protective relays use a non-linear decision boundary, which is computed from multiple physical measurements made locally, to detect the faults in the power grid. Besides this, they can also detect the rough location (in terms of tiers) of the initial fault with a single decision. By collecting measurements and monitoring continuously, a smart protective relay estimates the distance between itself and the fault location. Based on this monitoring which could last for several minutes, SVM-based relays can make intelligent decisions on whether and when to trip and/or reclose. At the same time, they can send messages about their status and/or conclusion on the system-wide conditions to the SCADA to alert the system. The SCADA can also utilize the information which is collected from smart protective relays for state estimation and re-adjustment of the power grid [58].

The current topology and structure of power grid is shown in Fig. 7.1. Obviously, the power system has a hierarchical structure and it is not completely distributed. Also observe from Fig. 7.1 that the connectivity of the power grid, in a graph-theoretic sense, is not the same as the connectivity of the communication network overlaid on top of the power grid. If one takes the viewpoint that protective relays can be considered as sensors in the power grid, then the current configuration of power grid can be considered as a sensor network with a centralized control center (SCADA). While SVM-based relays can substantially increase the efficiency of decision process, it is not clear whether this approach can completely eliminate the blackout problem. In fact, it is well known that in hierarchical networks, the higher-up is the subsystem affected by a fault or unexpected situation, the more difficult things could be. Part of the difficulty is precisely this legacy hierarchical architecture with centralized control.

Many people and research groups in the past have identified the root cause of blackout problem and showed its connection to the structural or topological organization of power grid. Some of the important works in this area include [60], [61], and [62]. The stochastic nature of the initial fault/failure makes it very hard to predict whether it would lead to a cascade of failures. It has been shown before that, depending on the subsystem hit by the initial failure, the propagation of failures could be very fast or relatively slow.
While these excellent studies shed light on the underlying reasons and dynamics of the cascade of failures, how to eliminate the blackout problem completely without changing the current hierarchical and centralized topology of the current power grid is a daunting task. In this sense, it is important to understand that the SVM-based smart relays can ameliorate the situation after an initial failure occurs by not making an ill-informed decision and tripping. In other words, our work shows that, based on the local information they collect (i.e., magnitude and phase of current and voltage, real power, reactive power, etc.), the SVM-based relays can accurately predict the location of the fault and make an intelligent decision on whether a relay should trip or not. This intelligent behavior when combined with fast communications with SCADA could substantially mitigate (see the difference between the third and fourth column of Table 4.5) the propagation (or cascading) of failures and thus prevent a large-scale blackout.

To see this more clearly, consider Fig. 7.2. To explain the principle of operation of the proposed SVM-based smart relays, this oversimplified representative power grid could be sufficient. It is worth recalling that in this dissertation, due to cost considerations, we propose to deploy a relatively small number of SVM-based smart relays on transmission lines. Suppose each load uses, as an example, 2 kW of power. If the

**Figure 7.1:** Current Topology and Structure of Power Grid (modified from [59])
initial fault, for instance, hits the distribution line\(^6\) D2 in Fig. 7.2, then the power supplied by D1 and D3 will increase as well as the power supplied by the back-up distribution line D4 coming from DS2. This, in turn, will cause D1 and D3 to trip, thus overloading the back-up line D4. In this example, to continue serving the loads L1, L2, and L3 of group 1, the power supplied by the back-up line will increase by 6 kW which eventually could increase the power carried by T2 significantly. In other words, if T2 is the main transmission line delivering power from the generator G1 to DS2, the power carried by T2 will almost double because of the cascade of failures of the loads in Group 1. Given that current transmission lines currently work with 80% capacity, this could be sufficient to trip T2. If T2 is out, then the power has to be supplied by transmission line T3 from the generator G1 and transmission lines T6 and T7 from the generator G2, which, with high probability, will trip the transmission line T3. Then because of the fact that G1 lost all possible transmission lines to support the demand of the loads connected with DS1, DS2 and DS3, G2 will have to supply the loads originally served by these three distribution substations. Since G2 will not be able to supply such a large power, its protective relay will trip taking G2 out of operation. This, in turn, will imply that all the remaining groups of loads will attempt to get their power supply from G1. Since G1 will not be able to provide such a high power, it will also trip causing the whole representative power grid to go to blackout, thus cutting off the electricity service to all loads.

When one deploys an SVM-based relay on T2, however, after the distribution lines D1, D2, and D3 trip, by monitoring the increase in power carried over a period of time (say several minutes), *an SVM-based smart relay will keep T2 closed and not allow it to trip*. In the mean time, a message will be sent to SCADA alerting it to the alarming rate of increase in the power carried by T2. Subsequently, SCADA will attempt to redistribute the excess power to other transmission lines (in Fig. 7.2, for instance, T3, T6, and T7) as well, trying to prevent the trip of T2. This way, the (N-1) condition can be maintained via the dispatching of SCADA and the aforementioned “load balancing” operation. In other words, the SVM-based intelligent relays can provide the necessary “breathing time” for SCADA to balance the power distribution over several transmission lines without tripping any major transmission lines (in Fig. 7.2, the extra power will be distributed over 4 transmission lines T2, T3, T6, and T7. In practice, there will be many more lines that are connected between generators and distribution substations which can be used to share the excess

\(^6\)It is worth mentioning that the example given in this section and depicted in Fig. 7.2 is for illustration purpose only. Thus, the fact that the initial fault starts at a distribution line in the provided example is not important in an absolute sense. The underlying mechanism which governs the propagation of a fault would not be different if the initial fault starts at a major transmission line instead. Similarly, the principle of operation of SVM-based relays and the benefit in mitigating blackouts will remain even if the initial fault starts at a major transmission line.
power), thus maintaining the (N-1) condition which is guaranteed to provide a stable mode of operation. This, in turn, prevents the cascading of failures and confines the problem to a much smaller area as opposed to a wide-spread blackout.

Figure 7.2: Representative oversimplified power grid as an abstraction.

The blackout problem has several parallels in other areas, such as infectious disease propagation in the society (malaria, HIV, other viruses), as well as malicious attacks on certain internet routers, etc. Although the physical, social, and economical mechanisms responsible for the occurrence of cascades are complex and may vary significantly across different systems (e.g., cascading failures in infrastructure and organizational networks, cultural fads, economic systems, etc.), some generic features of cascades and cascading failure can be explained in terms of the connectivity of the complex network through which influence is transmitted to individuals or individual components. During cascading failures, individual elements of a population exhibit herd-like behavior because they are making decisions based on the actions of other individual elements rather than relying on their own information about the problem [62]. Although they are generated by quite different mechanisms, cascades in social and economic systems are similar to cascading failures in physical infrastructure networks and complex organizations in that initial failures increase the
likelihood of subsequent failures, leading to eventual outcomes that, like the August 2003 cascading failure in the US, are extremely difficult to predict, even when the properties of the individual components are well understood.

It is important to emphasize that in this dissertation we view the connectivity and the topology of the current power grid as a given and explore how to mitigate or eliminate the cascading failure problem that eventually may lead to a large-scale blackout. In general, it is much harder, if not impossible, to have a self-organized network behavior when the network is centralized as opposed to distributed. For example, a similar problem exists with cellular wireless networks when one has the “hot spot” problem [63][64]. Researchers have proposed to alleviate the “hot spot” problem by deploying relays around base stations which can serve the incoming calls or handover requests by relaying the incoming requests to other base stations in the network that may have available channels [65]. A synergistic approach could also be possible for the hierarchical and centralized power grid whereby the excess power after the loss of a transmission line (or other subsystems) is somehow “rerouted”.

Finally, it is worth mentioning that the herein advocated machine learning based smart relays at critical locations for mitigating future blackouts is in stark contrast to an approach that uses less reliable relays to build a reliable power protection system. One can trace the origins of the latter approach to John von Neumann and his seminal paper [66]. As future work, it would be interesting to pursue such an approach, quantify its requirements and results and compare with the results presented in this dissertation.

### 7.3 Future Work

The future research directions which could be pursued as an extension to this thesis are:

- **Transient Measurements Based SVM Relays**

  Transient Measurements based SVM Relays can be used as system-wide protection relays, instead of a local equipment protective relays. They can serve independently with the static state measurement based SVM relays. By extracting information from the time-continuous function of local current and voltage, as well as the information from remote terminals, a transient measurements base SVM relay can be used to 1) estimate the best time for relays to re-close the breaker so that the system after disturbances is stable; 2) locate the most proper relays to trip after disturbances to make the system
reach the new equilibrium point quickly; 3) monitor the propagation of the cascading failures and warn the control center in a timely manner.

- **Locating Critical Relays via Power System Topology Analysis**

Critical relays are the ones which have the highest probability to experience a false trip in the protection systems. Our research has shown that by only replacing the critical conventional relays by SVM-based smart relays, a cascading failures after an initial disturbance can be mitigated. However, there is no efficient algorithm to locate the critical relays without going through an exhaustive search over the whole system. Obviously, the location of the critical relays are closely related to the structural properties of the power system. Therefore, a detailed study on the power system topology should be conducted to identify the critical relays. For example, if the power systems have the properties of a “scale free” network, the critical relays are likely to be located on the hub buses.

- **Security Issues in Smart Protection Systems**

Security is always an important issue in protection systems. Although several features are used in smart relays and the parameters of the non-linear decision boundary are not available for an outsider, it is still possible to attack the power systems by feeding a relay with false measurements to trigger cascading failures. Therefore, one needs to study the functional logic of the smart relays and propose a self-validation algorithm. This self-validation algorithm should utilize the properties of the functional logic of the smart relays and be embedded into the decision making process. By applying such a self-validation algorithm, if it cannot correct the false data, a smart relay should be able to identify and abandon the false data when making decisions. This could be achieved either with a software based or hardware based solution, and a software based solution might be preferable considering the cost of the smart relays.
Chapter 8

Appendix

Comparison between NNs and SVMs

A Neural Network (NN) is a massively parallel computing system consisting of a large number of processors (nodes) with many interconnections. NN models have nodes (neurons) and directed edges (with weights) between neuron outputs and neuron inputs. NN can learn complex nonlinear input-output relationships. A general structure of NN is shown in Fig. 8.1

![General Structure of a Neural Network](image)

**Figure 8.1:** General Structure of a Neural Network
Each node is represented by a logistic function:

\[ output = \frac{1}{1 + \exp(w_0 + \sum_i w_i x_i)} \]  

(8.1)

where \( w \)'s are weights of all inputs \( x \). The training of NN involves determining the weights of all nodes to minimize the sum of the squared errors at network outputs. This optimization is performed based on the gradient descent search.

SVMs and NNs are both statistical decision-making mechanism that can generate non-linear decision boundaries. Significant advantages of SVMs compared with ANNs are:

- **ANNs can suffer from multiple local minima; the solution to an SVM is global and unique.**

  SVM is a perfect quadratic convex problem with one and only one optimal solution, while the ANNs utilize gradient descent search to determine the optimal solutions. A major defect of the gradient descent search is that depending on the initial value, it may converge to a local minimum that is closest to the initial value.

- **Unlike ANNs, the model complexity of SVMs is automatically determined.**

  The model complexity of SVMs is a function of the number of support vectors (SVs) that are automatically determined during the SVM training. Therefore, the model complexity of SVMs is self-determined. On the other hand, the model complexity of ANN is related to the number of layers and nodes in the network, which can be arbitrarily chosen by users. Hence, part of the training process of ANN is an extensive exploration on how many layers and how many nodes should be used in order to achieve the highest decision-making accuracy. Therefore, the training of ANN can be extremely time-consuming.

- **In practice, SVMs are less prone to overfitting than ANNs.**

  Overfitting, which can be a problem in any kind of machine learning methodology, can increase the training accuracy as well as the model complexity. In SVM, the target function contains two terms. One is the model complexity and the other is the training accuracy. By minimizing the target function as a whole, overfitting can be avoided automatically. However, the target function of ANN does not contain any terms related to the model complexity. Therefore, without an extra step on data pruning, it always overfits the training data.
Because of these three major advantages, we have chosen SVM as our classification scheme. In addition, the advantage of SVM algorithm becomes more apparent in the context of protective relays. In other words, the proper network structures of ANNs are highly related to the power system and the protective relays in question. Therefore, only if one is very familiar with the power system and protective relays in question, the shortcomings of ANNs can be avoided (e.g., by choosing the proper number of the layers and nodes, as well as the initial value of the weights). Under such an ideal circumstance, the trained-ANN will be as accurate as a well-trained SVM. However, since one typically does not have sufficient knowledge about the system in general, the implementation of ANNs on protective relays can be difficult and the accuracy cannot be guaranteed.
Bibliography


[8] Private conversation with Eric Udren, Quanta Technology, Raleigh, NC, 2009


[34] U. Fayyad; P. S. Gregory, and P Smyth, “From Data Mining to Knowledge Discovery in Databases”, 2008.


[46] DMS Group, http://www.dmsgroup.rs/contact.htm,


[53] N. Syed, H. Liu, and K. Sung, “incremental learning with support vector machines.”, In Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (IJCAI-99), San Diego, CA, USA,1999


[56] Mushrooms Data Set, in http://archive.ics.uci.edu/ml/datasets/Mushroom


