AN INTELLIGENT DECISION MAKING SYSTEM FOR DETECTING HIGH IMPEDANCE FAULTS

A Dissertation

by

CHARLES JONG KIM

Submitted to the Office of Graduate Studies of

Texas A&M University

in partial fulfillment of the requirement for the degree of

DOCTOR OF PHILOSOPHY

December 1989

Major Subject: Electrical Engineering

ABSTRACT

An Intelligent Decision Making System

for Detecting High Impedance Faults. (December 1989)

Chair of Advisory Committee: Dr. B. Don Russell

The clearing of distribution line faults is usuallys accomplished by devices which can sense the overcurrent produced by a fault and react to disconnect the faulted section of the feeder from the healthy section. However, high impedance faults do not draw sufficient fault current to be detected by such a conventional protection scheme. Such faults may be caused by a conductor on the ground. Arcing is usually associated with these faults, which may result in a fire hazard.

The harmonic currents characterized by an arc are variable, transitory, and random in their behavior. The relative amplitude increase of harmonic currents is very large in some high impedance faults, but sometimes it is very low, and other times very similar to the level of the normal state.

While a few techniques to detect high impedance faults have been proposed, and some progress has been made, a complete solution has not been found. This research concentrates on designing an intelligent decision making system which uses multiple detection techniques incorporated with an appropriate detection reasoning method and a learning ability to provide a more effective solution for high impedance fault detection.

Major parts of this system are a technique selection, a technique combination, and an induction process. The method of decision making under

incomplete knowledge is used to select appropriate techniques because the information on the performance of techniques are available but not complete. With these selected techniques, the Dempster-Shafer theory is adopted to find a final belief about the system status by combining the belief from each technique. Inductive reasoning with minimum entropy is applied to find decision rules and thus to adjust the technique selection process.

A learning detection system which combines all three major parts is proposed to realize this intelligent decision making system. The learning detection system synthesizes the final belief of the combined techniques, the status output of a decision tree from the inductive reasoning process, and an event detector output to detect and identify the system status. The intelligent decision making system makes a smart decision on an example execution with a complex test set of sample data.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

There is a problem which currently exists in the electric utility industry with the detection of high impedance faults. High impedance faults can be described as those distribution feeder faults that do not draw sufficient fault current to be detected by conventional protective devices such as overcurrent relays. Such faults may be caused by a downed conductor on the ground or in contact with a grounded object. Arcing is often associated with these faults, which may result in a fire hazard.

The harmonic current characterized by an arc is variable, transitory, and random in its behavior. The relative amplitude increase of harmonic currents is very large in some high impedance faults, but sometimes it is very low, and other times very similar to the level of the normal state.

The behavior of high impedance faults can be characterized by a large relative amplitude increase in harmonic current through the duration of the arcing burst. However, the behavior of various high impedance faults is not consistent. Under similar conditions, faults show different characteristics in behavior and may draw less harmonic current than other faults on the same soil surfaces. On surface types such as grass, asphalt, and concrete, the behavior of faults can be very similar to the phenomenon of switching events. Hence, the detection and classification of high impedance faults, switching events, and the normal state is a complex problem.

While a few techniques to detect high impedance faults have been proposed, and some progress has been made, a complete solution has not been found. This research concentrates on an intelligent detection system which uses multiple

Journal Model is *IEEE Transactions on Power Delivery.*

detection techniques. It is expected that appropriate decision making methods incorporated with a learning ability will provide a more effective solution for high impedance fault detection.

The primary research objective is to find an intelligent detection system which combines multiple detection techniques. This system employs a decision making method for technique selection, a technique combination, and a learning ability, to detect/classify high impedance faults and discriminating them from switching events and normal states. Adaptability to outside factors will also be considered.

Existing fault detection techniques have considered and focused only on one parameter for the detection of a high impedance fault under given conditions. However, the behavior of high impedance faults is affected by many outside factors such as feeder configuration, surface conditions, weather conditions, and so on. Therefore, a single technique cannot provide effective fault detection. It is proposed that the use of a combination of detection techniques can improve overall detection and discrimination under a wide variety of conditions. If adaptability is added to this scheme, it is felt that most high impedance fault cases can be detected and classified.

An intelligent decision making system for high impedance fault detection will be delivered with logical structures, methodologies, implementation methods, programs, and application examples. Major parts of this system are a decision making method for a technique selection, a technique combination method which handles multiple pieces of information with uncertainty, and an automatic knowledge acquisition (or learning) method. In a time when utility companies are wanting to implement a high impedance fault detection system and to add intelligence to existing conventional techniques, this research appears to have the potential of achieving that very goal.

The following examines the characteristics of high impedance fault and switching events and how they are affected by environmental parameters. The statistical analysis was performed with staged fault data. Chapter III examines the existing detection techniques. Even though the purpose of this work is not to compare the performance of each technique, the advantages and weaknesses of each technique are investigated. Chapter IV explains the intelligent system in general and the system structure of an intelligent decision making system which consists of a technique selection, a technique combination, and an induction process for adaptability. Chapter V presents the details of the technique selection process which include an initialization stage, a technique selection stage, and a technique selection update stage. In Chapter VI, the technique combination process using the Dempster-Shafer theory with selected detection techniques is discussed. The combination of basic beliefs is illustrated with many examples. The data classification with inductive reasoning is discussed in Chapter VII. Pattern classification and threshold value calculation are also discussed. In Chapter VIII, the total and final intelligent decision making system which is realized as a learning detection system, is discussed. Its realization, implementation, and an example execution are discussed. In Chapter IX, the summary and conclusions about the final results are discussed.

CHAPTER II

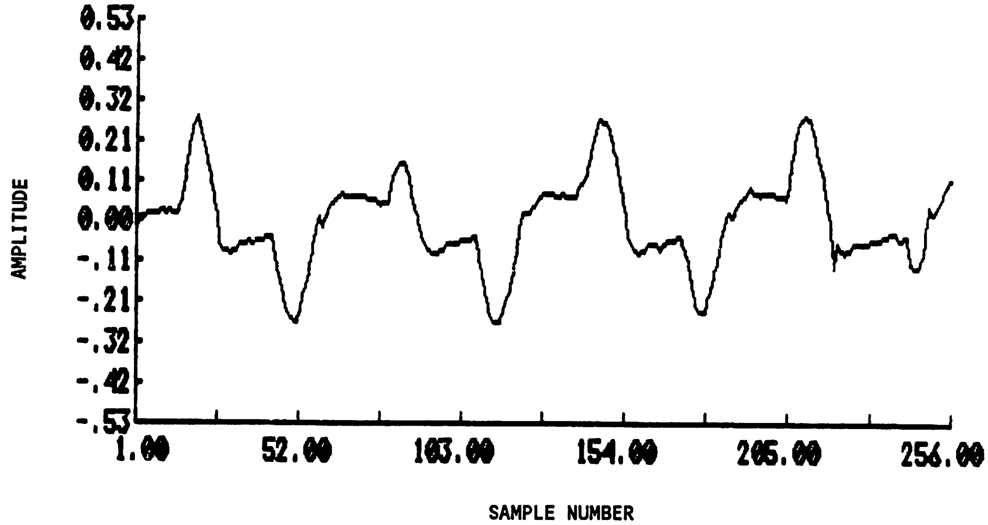
HIGH IMPEDANCE FAULTS

A. Introduction

High impedance faults typically occur on distribution circuits at voltage levels below 15 KV. Faults on a higher voltage circuit usually exhibit sufficient fault current to be detected by overcurrent protective devices. High impedance faults are usually caused by broken conductors on the ground. Arcing is often associated with these faults and they can cause fires and personnel hazard. The behavior of an arc has been studied by many scientists and, for a power system, can be summarized as follows. If two conductors are separated by a small gap and have a small potential difference between them, the air acts as an insulator. As the potential difference is increased, the resistance of the air gap decreases and the current flows between the conductors. Rapid ionization accounts for the sudden ability of the air to conduct current. Typically, a stable arc can be formed after several strikes and restrikes occur within a short period of time. In general, for the case of arcing in AC systems, stable arcs will form and be extinguished every half cycle[1].

Current harmonics generated by arcing have their origin in the nonlinear voltage-current characteristics of the arc voltage. The effect results in typical wave shapes and thus harmonics, sub-harmonics, and almost all frequency components. The harmonic current characterized by an arc is variable and although odd harmonic frequencies dominate, all harmonic and non-harmonic frequency components are produced at times. Figure 1(a) shows a differenced waveform, which comes from the fault current waveform after subtracting the normal current waveform. This differenced waveform starts at every quarter

(a)



(b)

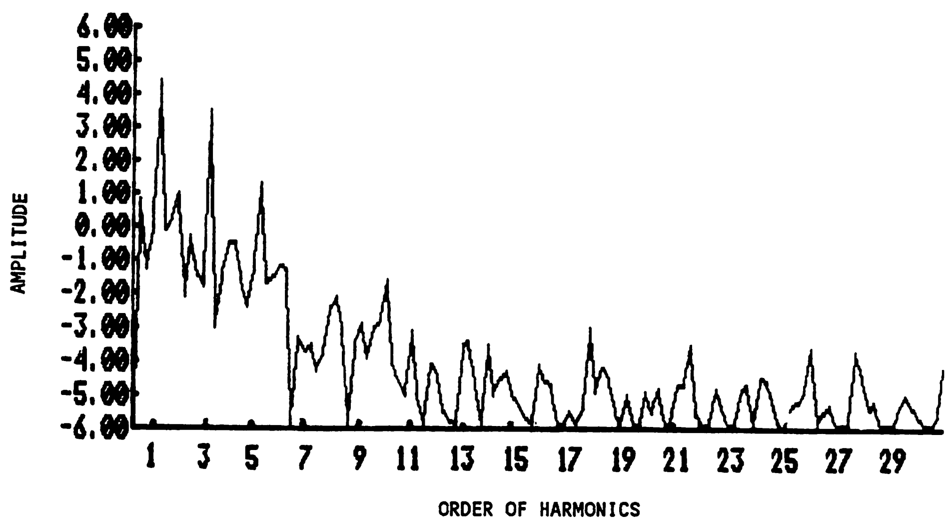


Figure 1. (a) A Differenced Waveform and (b) its Periodogram.

point of a cycle and remains for a quarter cycle; it is zero at the other portion of the cycle. Figure 1(b) shows the periodogram (squared amplitude of each frequency) of the differenced waveform. For this particular waveform analysis, the TIMESLAB software package was used to see which frequency components are present[2]. All orders of harmonics are present in the differenced waveform. Some harmonics show large magnitude increase, others small. The characteristics of the behavior of high impedance faults were investigated in the frequency domain.

B. Statistical Analysis

Utility statistics for broken and fallen conductor faults are usually poor and in many instances no formal records maintained. In order to understand the nature of high impedance faults and characterize their behavior, Texas A&M conducted a series of staged fault tests in collaboration with several utility companies in the country. These tests were staged under a variety of conditions of loading, soil condition, surface condition, and feeder configuration. In addition, measurements were taken during various switching operations like capacitor bank and air switch operation. Normal feeder conditions were monitored to determine feeder background characteristics. The data was recorded in analog form on a high-quality instrumentation recorder so that fault waveforms could be reproduced easily in the laboratory.

Texas A&M began research on the high impedance fault problem in early 1977. Initial research concentrated on understanding high impedance faults, developing models to describe them and performing laboratory simulation of faults. This research was based primarily on the high frequency (above 2.0 KHz) changes in current associated with arcing high impedance faults[3, 4]. Later, behavior of low frequency spectra during high impedance faults and normal events were investigated for odd, even, and sub-harmonics up to 360 Hz on different surface conditions[5]. The harmonic behavior of high impedance faults with and without a capacitor bank on the feeder line was also investigated while concentrating on even harmonics and the ratio of odd to even harmonics up to 1200 Hz (medium frequency components)[6].

To measure each frequency component variation during high impedance fault, a method for determining the "energy" contained in the normal system signal was needed. Individual spikes of noise were not of interest, but rather an indicator of the cumulative effect of many spikes over a short period of time was desired. "Energy" is defined as the summation of squared sample values of current over a cycle of 60 Hz.

High Frequency Components

For the purpose of data recording for the investigation of the behavior of high impedance faults, several utilities staged fault tests. A total of 86 separate faults were staged in a variety of locations with different soil conditions (including faults on concrete and asphalt), varying degrees of soil moisture, and a wide range of available fault currents. From preliminary research, the key characteristic identified for high impedance faults were the arcing normally associated with these faults. It was believed that arcing produced high frequency transients which could be used for fault detection.

The field tests showed that high impedance faults do exhibit a marked increase in high frequency current components over normal system conditions, and they persist for the duration of the arcing. The exceptions are due to the presence of grounded wye capacitor banks on the feeder. In this case most of the high frequency signal is shunted to ground by the capacitor bank. Figure 2 shows

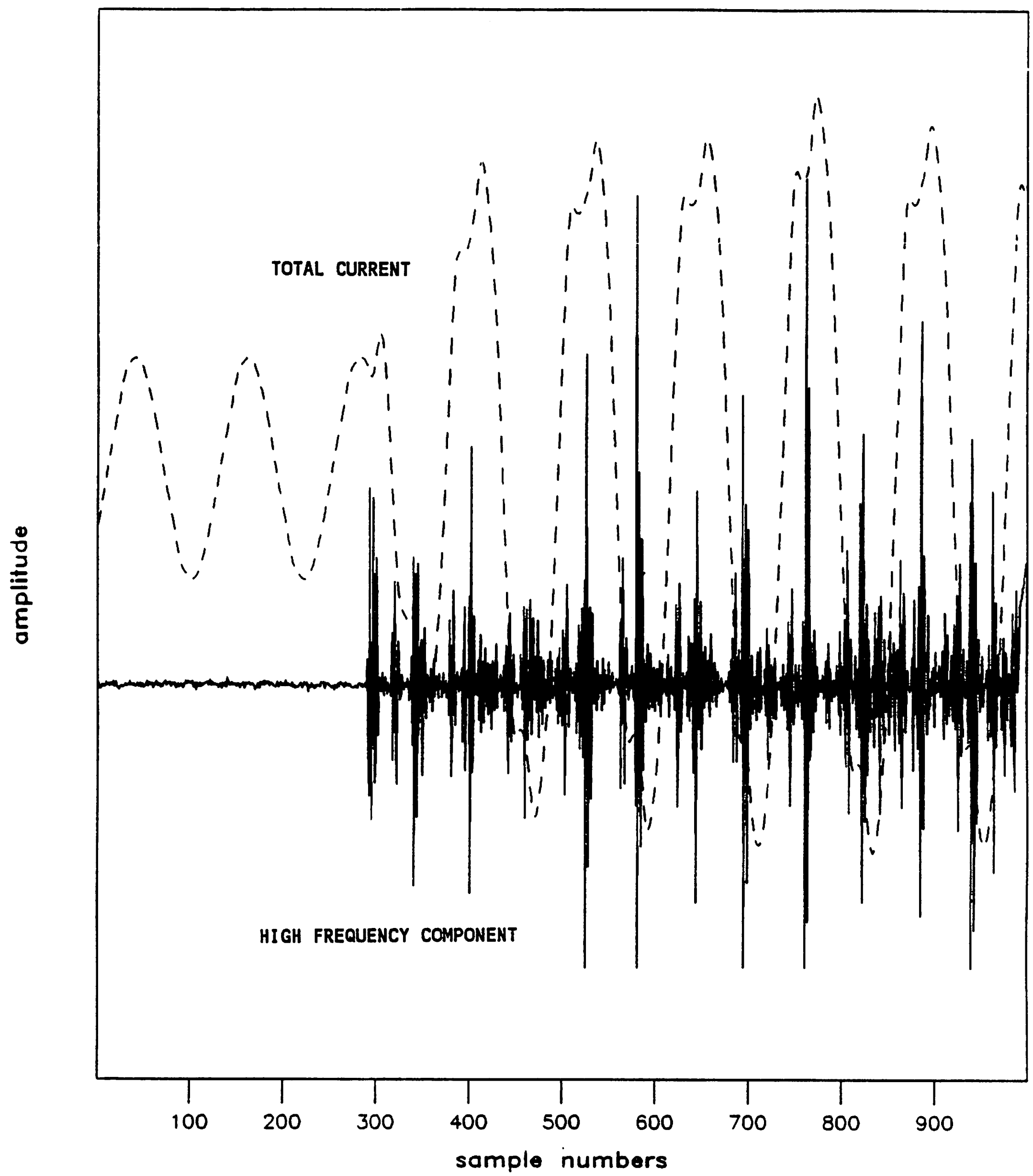


Figure 2. High Frequency Component Increase during a High Impedance Fault.

the increase of the high frequency current components during high impedance fault. Where total current has "notches," the high frequency current shows a dynamic transient of spikes. To neglect the harmonics, which vary widely under different loads, signals above 2 KHz are selected. Due to deteriorating signal levels at higher frequencies, 10 KHz was chosen as the upper bandlimit of interest[7].

Medium Frequency Components

With the recorded staged fault data, a frequency domain investigation has been done for the fundamental to 20th harmonic (1200 Hz) range concentrating especially on high impedance faults with and without a capacitor bank on the feeder. The analysis is focused on a differenced waveform, which is obtained from the fault waveform after subtracting the normal waveform. Odd harmonics, such as the 3rd and 5th, are predominant, but some even harmonics, such as 2nd, 4th, 6th, 8th, and 10th, have significant amplitudes; some even harmonics are greater than some odd harmonics in magnitude.

When a capacitor bank is present on a distribution feeder, some harmonics are increased in some cases and eliminated in other cases. A normal current waveform of a distribution feeder, when a capacitor bank is present, shows no significant changes in phase current in the amplitude or in the wave shape. Odd harmonics (especially 3rd, 5th, and 7th) dominate and the overall shapes of the frequency components are quite similar to that of the normal current waveform without a capacitor bank.

Figure 3 shows the relative magnitude increase of the medium frequency spectra during a high impedance fault with and without a capacitor bank. When a capacitor bank is present, the odd harmonics are predominant, but even harmonics can be seen to some extent. However, the relative increase (compared

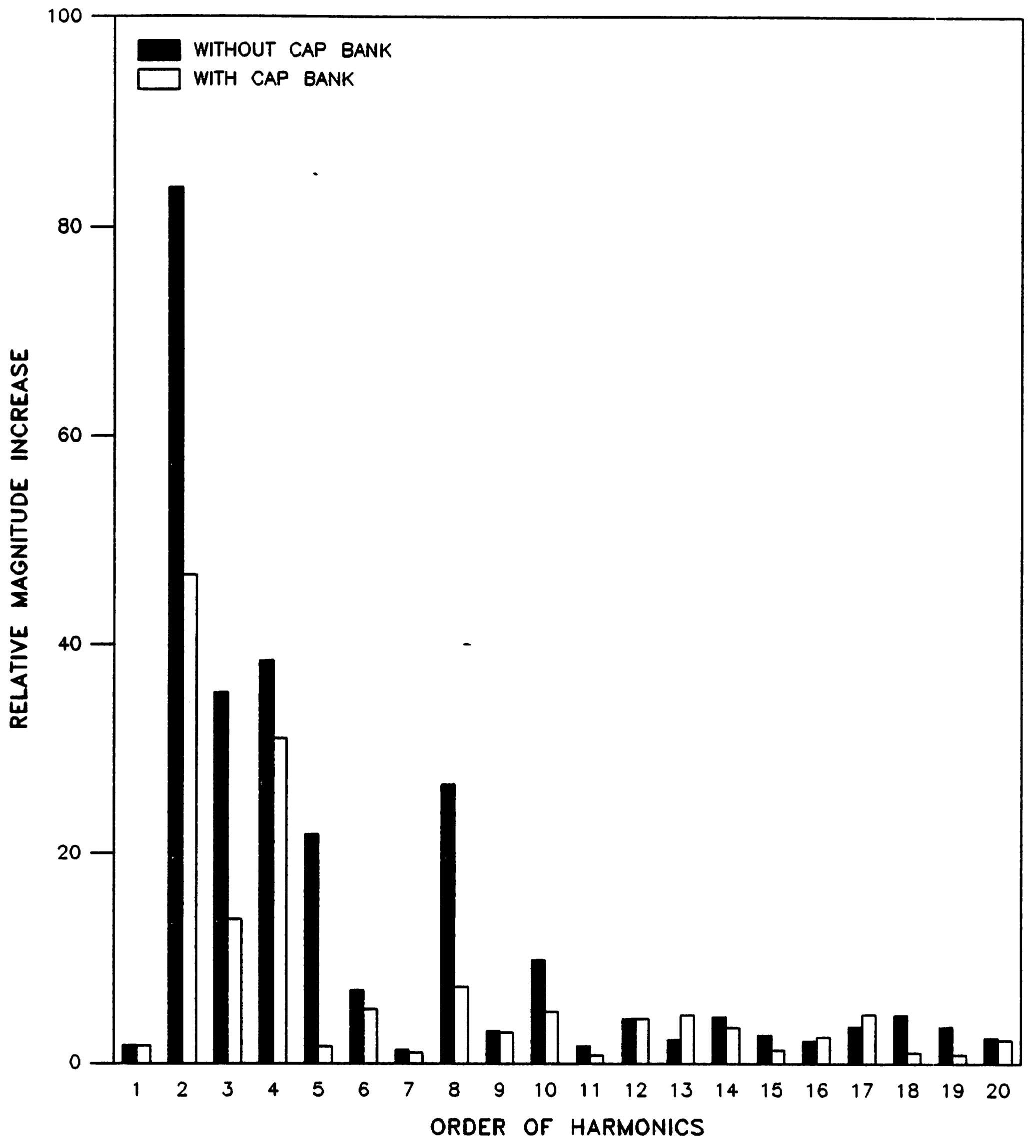


Figure 3. Frequency Response of High Impedance Fault Current with and without a Capacitor Bank.

to the normal status) of the even harmonics, from the 8th to the 16th, is much greater than that of the odd harmonics. Without a capacitor bank, the odd harmonics are still predominant, but the relative increase of the even harmonics is much greater than that of the odd harmonics. In addition, the increases in magnitude of all the harmonics without a capacitor bank are larger than those with a capacitor bank.

It is concluded that, up to the 8th harmonic, a comparison of the relative increase of the even harmonics with those of the odd harmonics, or simply the relative increase of the even harmonics, can discriminate a high impedance fault condition from the normal state.

Low Frequency Components

It has been observed that the arcing phenomenon associated with high impedance faults causes certain low frequency spectra to change in magnitude and phase from pre-fault conditions. The frequency spectra selected for investigation were:

1. Odd harmonics (60, 180, 300 Hz)
2. Even Harmonics (120, 240, 360 Hz)
3. Sub-Harmonics (30, 90, 150, 210, 270, 330 Hz).

This investigation was done on the recorded staged fault data. All the frequencies studied show an increase of at least 3 times in magnitude under high impedance fault conditions. For the statistical analysis, a frequency domain analysis was performed on the individual events in a cluster to determine the relative change in magnitude of the various spectra from pre-fault to fault conditions. Next, mean and standard deviation of the magnitude change was determined for each frequency spectrum by averaging all the 30 different events in the cluster and determining variation about the mean. Figure 4 shows the

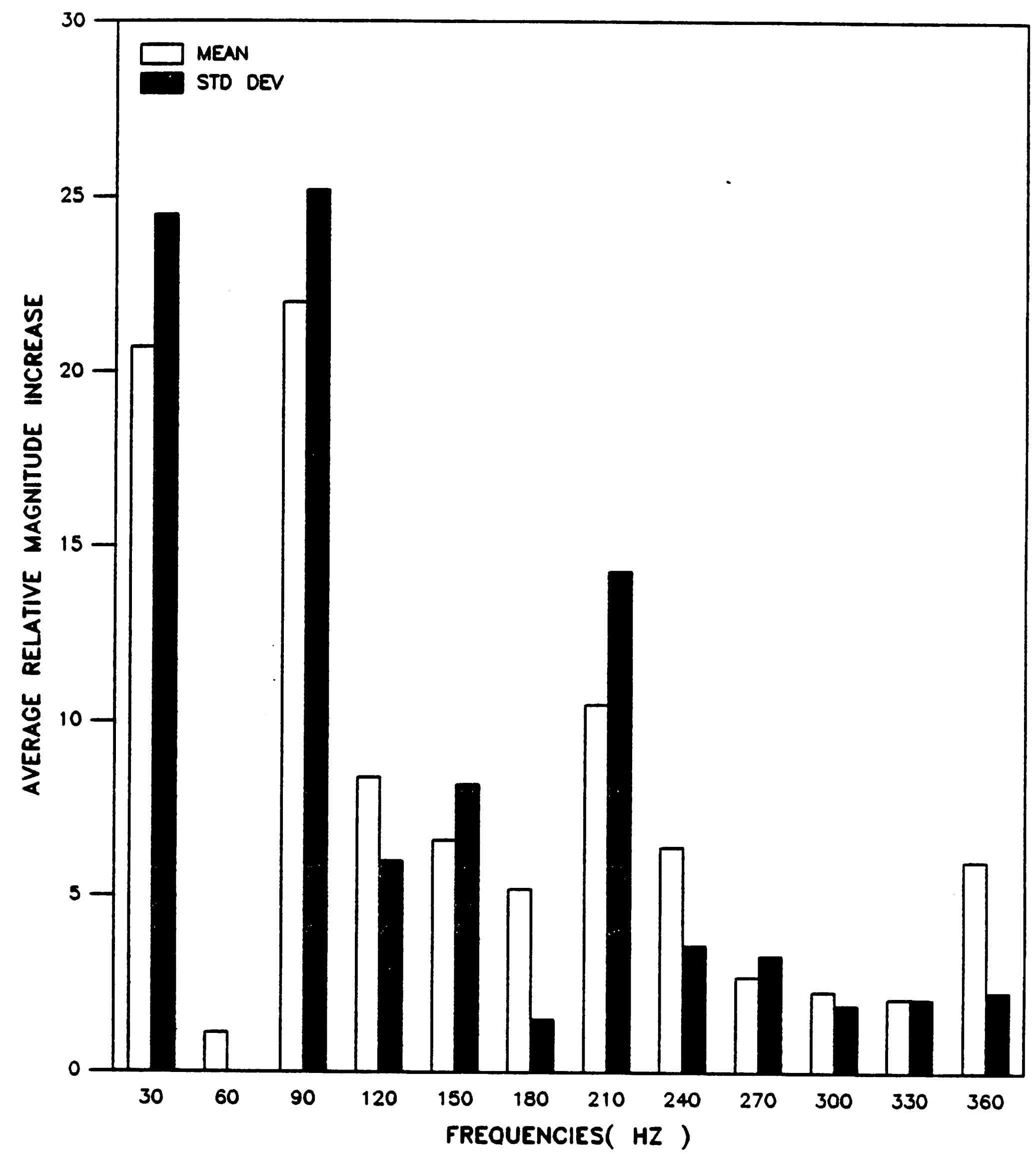


Figure 4. Average Magnitude Increase of Low Frequency Spectra during a High Impedance Fault.

average relative magnitude increase of low frequency spectra during a high impedance fault on dry soil. As is seen from the figure, sub-harmonics indicate a large dynamic change in magnitude (5 - 20 times the pre-fault level) under fault conditions. However, this change is very random as indicated by a large standard deviation. This means that even though the sub-harmonics show a large percentage increase in magnitude, the dispersion about the mean is also large, thereby indicating the precise amount of increase is unpredictable. The harmonic frequencies show a reliable change in average magnitude, however. This suggests that the relative change in magnitude is consistent for the different events comprising the cluster. The distributions of the magnitude increase of the several frequencies with different soil conditions are shown in Figure 5. As this graph shows, the relative magnitude increase becomes larger, as the soil moisture level becomes higher.

The dependency of the burst duration was also investigated with respect to soil conditions. It was seen that a large percentage of the events were comprised of short bursts, typically 2 or 3 cycles in duration, on wet soil. On dry soil, the arcing usually lasted for 4 to 20 cycles. On sandy soil, it was observed that the arcing persists longer, usually greater than 20 cycles in duration. The distributions of the arcing duration on three different soil conditions show similar exponential distribution with different mean values.

C. Switching Events

Switching events are investigated mainly in the low frequency range and in the high frequency range. Capacitor bank operations, air switch operations, and LTC (load tap changer) operations were performed. These events were also investigated in the frequency domain.

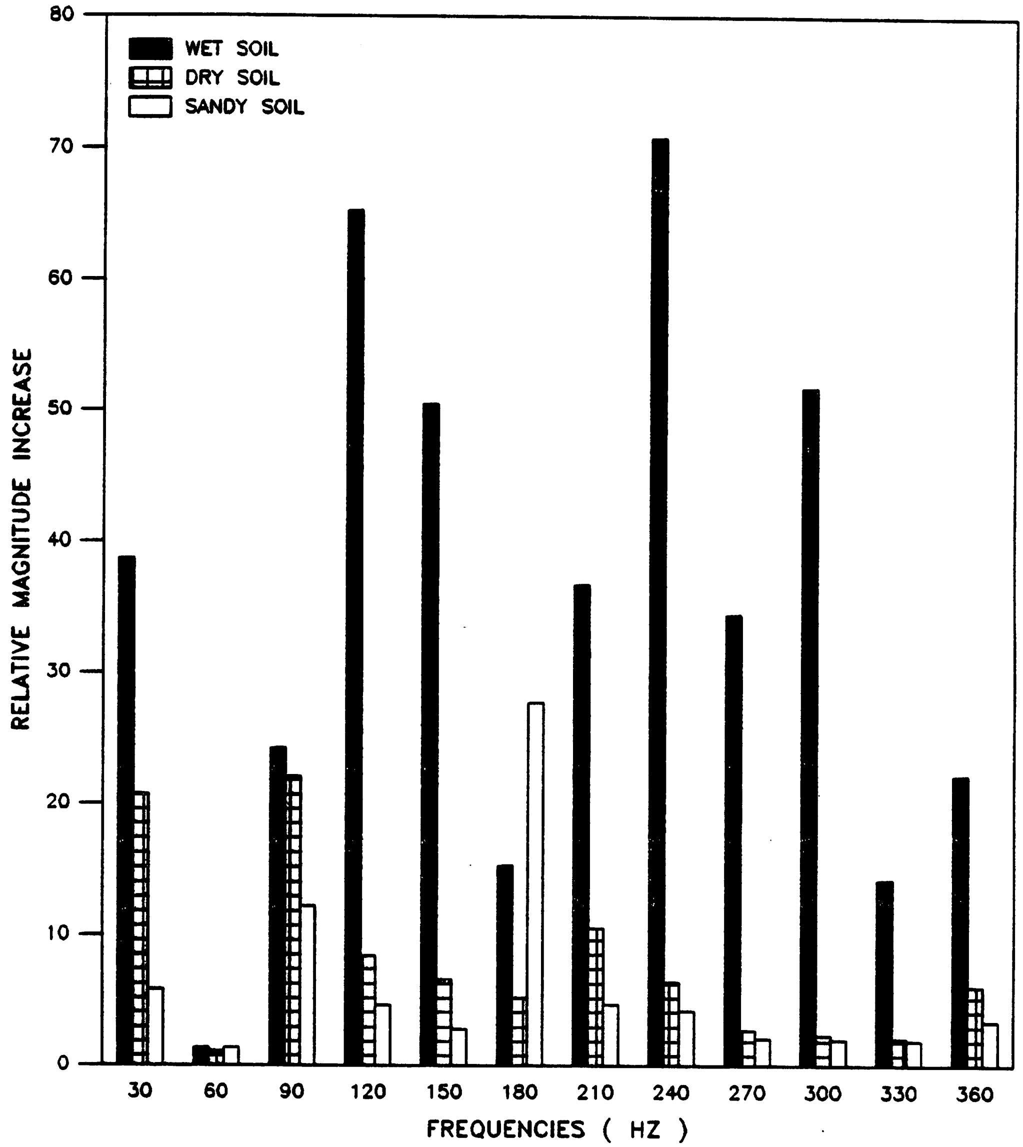


Figure 5. Distribution of the Magnitude Increase of Several Frequency Components on Different Soil Conditions.

High Frequency Components

Switching operations generate high frequency components which are definitely time-limited. Capacitor bank operations show a large magnitude of high frequency components. An increase in the high frequency current component which lasts much longer than the duration of a switching event provides an indicator of a fault. Figure 6 shows the energy variation of high frequency current component during a capacitor bank operation. When a switching event occurs, there is an impulse-like spike with large energy increase.

Low Frequency Components

With the same frequency spectra selected for the high impedance fault investigation, air switch operations and capacitor operations were investigated. For the behavior of the spectra during a capacitor bank operation, it is observed that the sub-harmonic frequencies remain fairly constant during the switching operation whereas the harmonic frequencies show a substantial change in magnitude. The odd harmonic energy shows a step increase in magnitude which persists the entire time the capacitor bank is energized. On the other hand, subharmonics magnitude change is more of an impulse lasting for a short duration of time due to the transitions of the switching operation itself. Figures 7(a) and 7(b) show odd harmonic energy variation and sub-harmonic energy variation during a capacitor bank operation. In the case of an air switch operation, all the harmonics show an impulse lasting for a short duration with a significant magnitude increase. Figures 8(a) and 8(b) show the odd and even harmonic energy variation during an air switch operation.

D. Environmental Parameters

From several investigations to see the behavior of high impedance fault, it

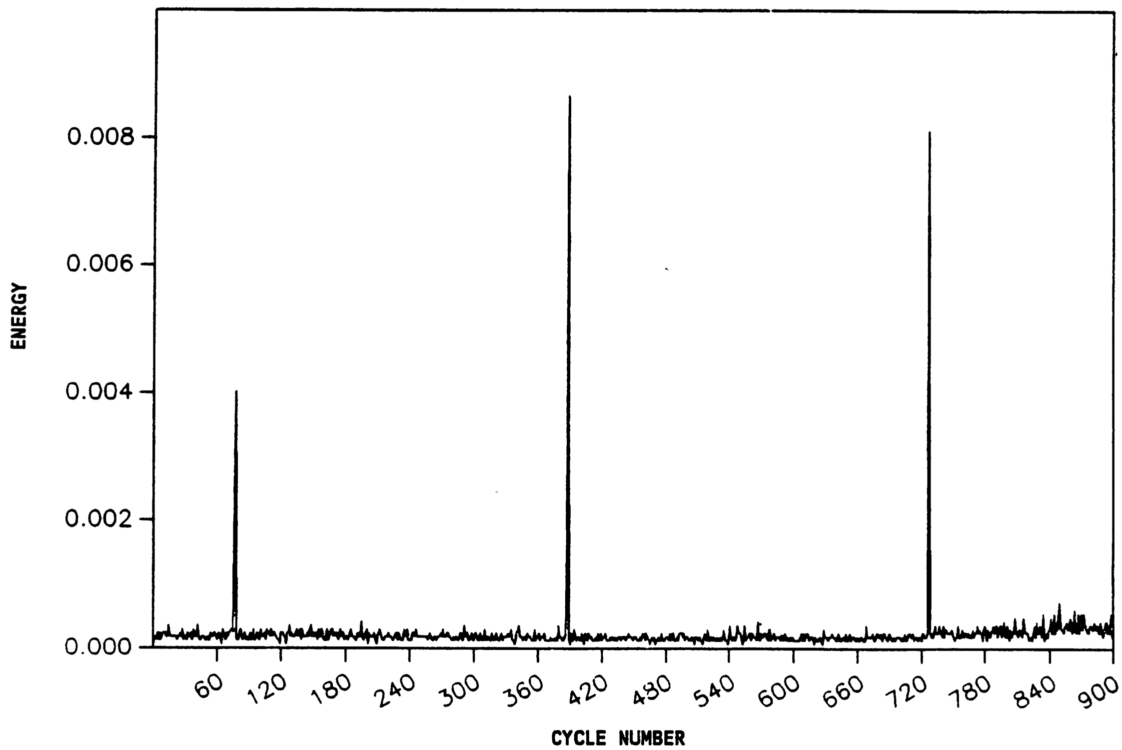


Figure 6. Energy Variation of High Frequency Spectra during a Capacitor Bank Operation.

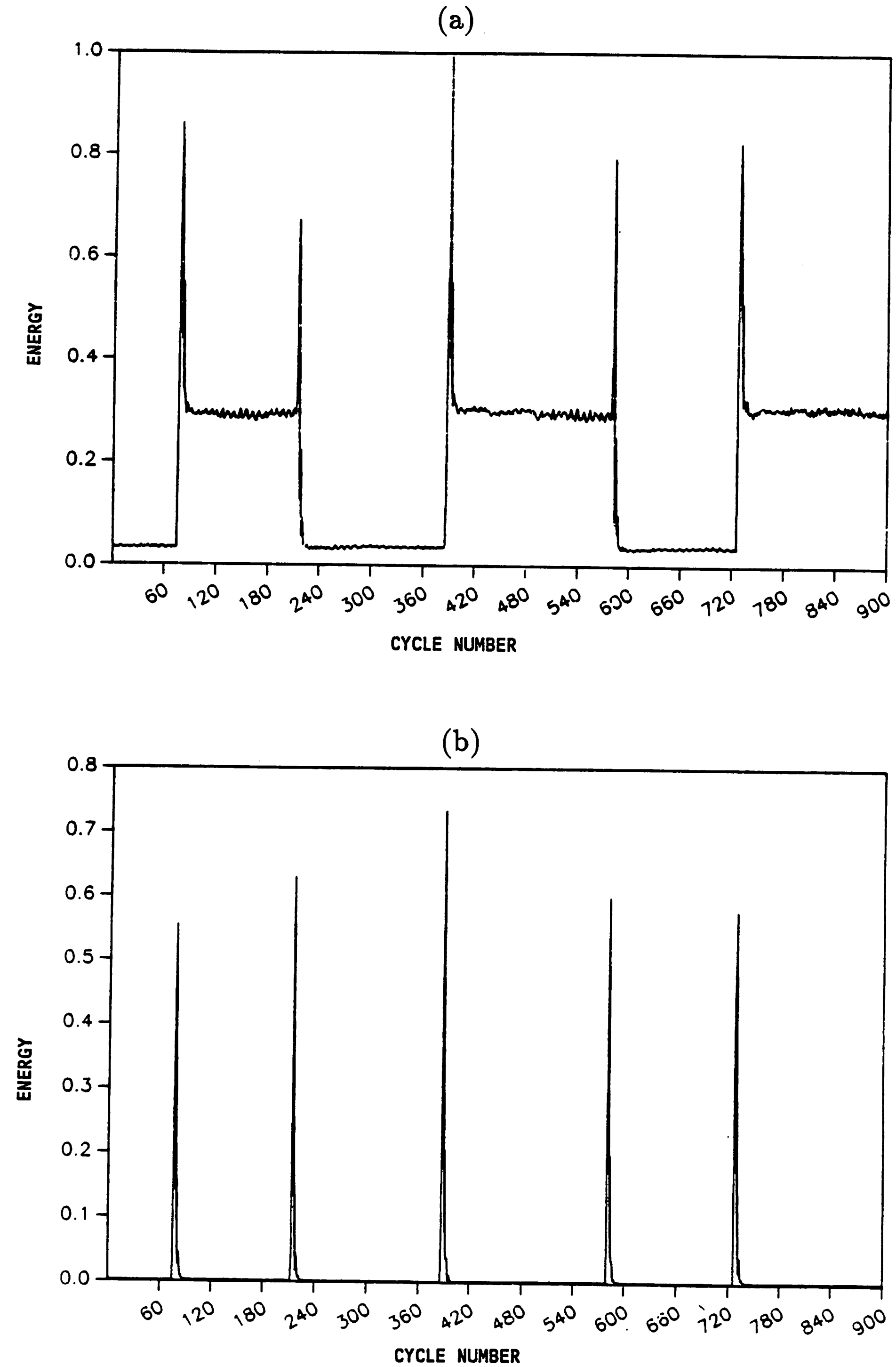
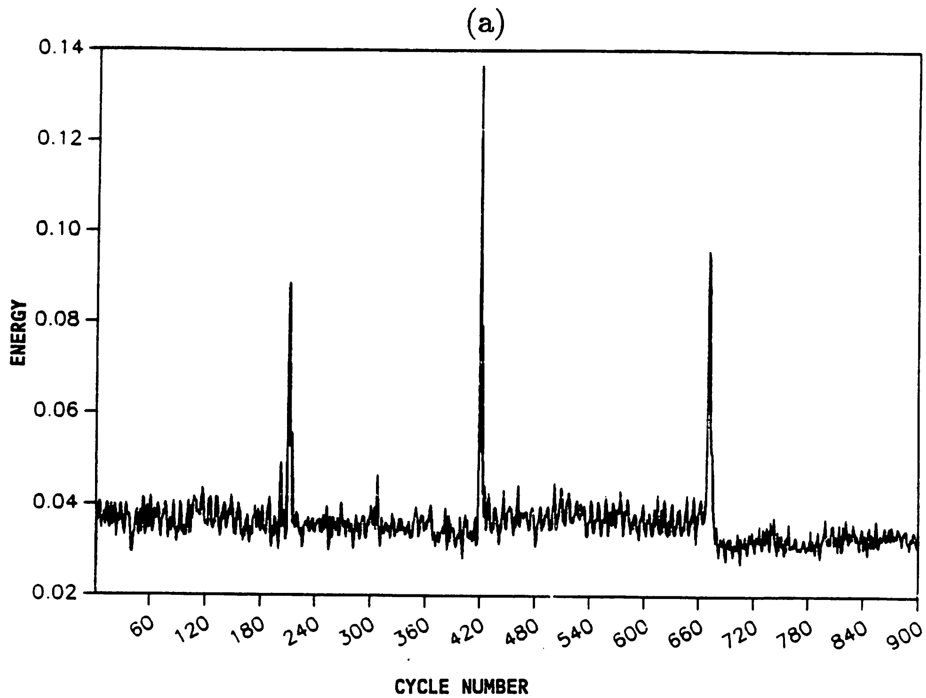


Figure 7. Energy Variation of (a) Odd Harmonic Current and (b) Sub-harmonic Current during a Capacitor Bank Operation.



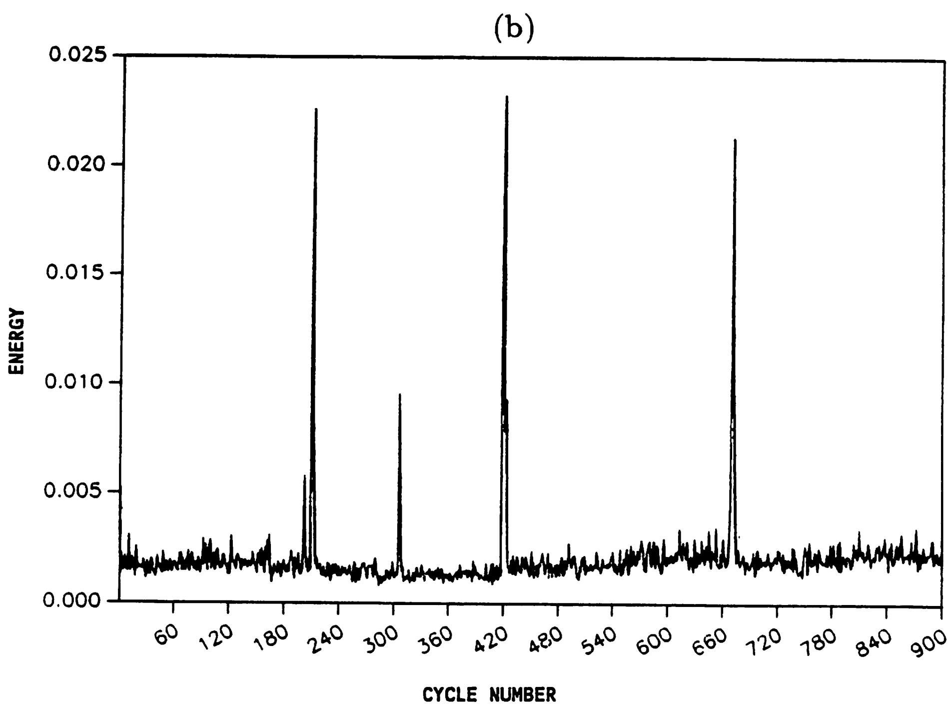


Figure 8. Energy Variation of (a) Odd and (b) Even Harmonic Current during an Air Switch Operation.

is found that there are some factors or parameters which change the behavior of the magnitude increase and burst duration. Here are examples of the changed behavior of the magnitude increase and burst duration.

Several investigations were performed to study the behavior of high impedance fault on unusual surface such as asphalt, grass, concrete, and so on. The behavior of these faults were examined with several harmonic currents at the frequency range up to 1000 Hz. Figure 9(a) shows an odd harmonic energy variation during high impedance fault on wet grass. The energy change during the fault shows a large energy increase with a short burst duration. Figure 9(b) shows an odd harmonic energy variation during a high impedance fault on concrete. The energy change also shows a large energy increase which lasts only a few cycles. There are some similarities between these faults on unusual surfaces and the switching events with respect to the magnitude increase and the burst duration.

In addition, the presence of a capacitor bank will depress the activity of high frequency components during fault conditions. The behavior of several sub-harmonics during a high impedance fault will show very different behavior in magnitude (or energy) and burst duration in the same situation. The surface condition will change the characteristics of the harmonics in the burst duration. The factors which are called "environmental parameters" are parameters or environment which changes the behavior of high impedance faults and thus will affect the detection techniques which use those frequency components vulnerable to the outside environments. Four environmental parameters are chosen and the reasons of their choice are mentioned as follows:

1. System Unbalance: load level is changing slowly and rapidly, which will cause misoperation of certain techniques which monitor the magnitude

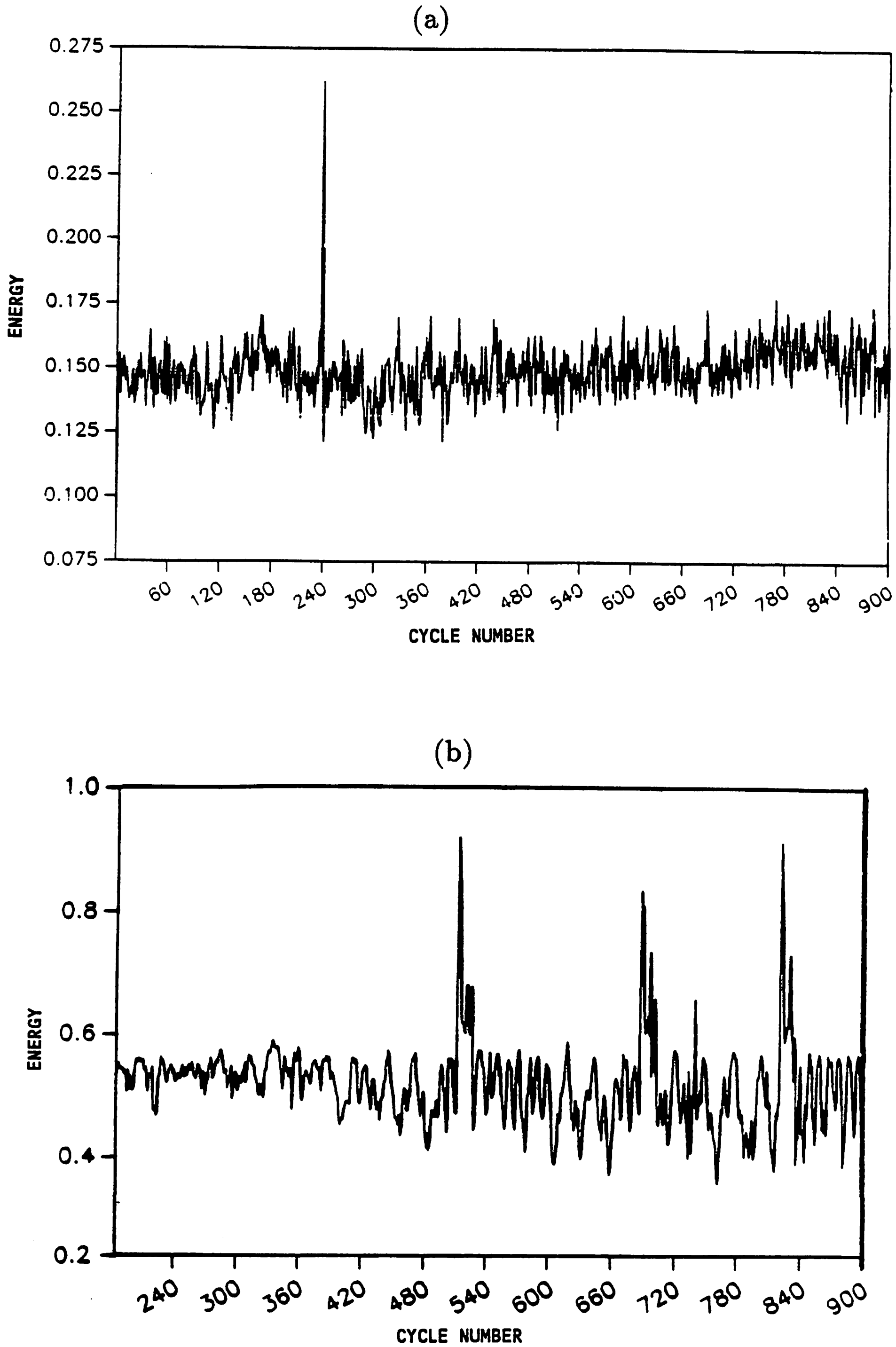


Figure 9. Energy Variation of Odd Harmonic Current during a High Impedance Fault (a) on Wet Grass and (b) on Concrete.

variation for detecting high impedance faults.

2. Feeder Configuration: the presence of capacitor banks attenuates high frequency components.
3. Load Types: certain loads produce some harmonics which have somewhat similar behavior to high impedance faults, which might lead certain techniques to make wrong decisions.
4. Surface Conditions: the surface conditions affect the behavior of the high impedance faults in magnitude and burst duration of harmonic and sub-harmonic fault currents.

E. Summary

Statistical analysis was performed in the frequency domain to investigate the behavior of high impedance faults in several frequency ranges. It was observed that the magnitude or energy of high frequency components shows a large increase. The magnitude increase and burst duration of harmonics and sub-harmonics in the low frequency range depends on the soil condition. Certain switching events show very distinctive behavior, but other switching events, such as an air switch operations, shows very similar behavior to that of high impedance faults on grass or concrete. Environmental parameters which affect the behavior of high impedance faults were suggested to measure the performance of various detection techniques under different situations.

CHAPTER III

EXISTING DETECTION TECHNIQUES

A. Introduction

The clearing of distribution line faults is usually accomplished by devices that can sense the overcurrent produced by a fault and quickly react to disconnect the faulted section of the feeder from the healthy section. Such devices usually include overcurrent relay and circuit breaker combinations, reclosers, and fuses. In addition to providing overcurrent protection, these devices must be capable of carrying an acceptable increase in the level of load current and transient overcurrents. They must not trip on inrush and surge currents resulting from transformer energization and feeder switching activity. In order to maintain reliability of service, the threshold of operation must therefore be set at a relatively high current level.

The current which flows during a fault depends on a number of variables, one of which is the fault impedance. A high impedance at the fault point can limit the fault current to levels that are well below the threshold of operation of conventional fault clearing devices. In such circumstances, the existence of a fault cannot be detected by overcurrent devices and will persist until some external event clears it. In this chapter, a brief summary of the proposed detection techniques, and their limitations and weaknesses with respect to environmental parameters are examined.

B. Energy Technique

This technique was proposed by Texas A&M researchers. The term "energy" refers to the summation of the squared sample values of current over one 60 Hz cycle. An early energy technique used the amplitude increase of the high frequency (2 - 10 KHz) current to detect high impedance faults[4]. Later, a revised energy technique was developed, which used a low frequency component as a parameter along with some fixed amount of cycles as a test window in a hierarchical detection scheme[8].

The energy technique determines the energy contained at any time in the current signal. A typical distribution system exhibits a periodic load cycle. Hence the detection technique must be adaptive to these changes in the load and at the same time avoid a complicated procedure for calibrating arbitrary pickup levels for fault detection.

A hierarchical nature is bred into the technique by making use of three levels in the detection process before signaling a fault. The system starts by recognizing a "disturbance" and on the occurrence of a disturbance, the system devotes its attention to trying to verify if the disturbance qualifies as an "event." An event recognition is followed by an attempt to classify the episode as either a "fault" or a normal occurrence.

Disturbance: A cycle of data showing a certain percent increase of energy over the average energy per cycle, the average being calculated over some previous period of time, constitutes a disturbance. Thus, if a cycle shows a certain percentage increase in energy over the previous average, a disturbance is said to have occurred. If the energy present in the present cycle is reasonably equal to the previous average, then a new average is calculated and disturbance detection is begun again.

Event: Once a disturbance is detected, a preselected series of cycles of data are tested. If a set percentage of these cycles show a certain percentage increase in energy per cycle over the average energy per cycle (the average being calculated over some previous period of time), then an event is said to have occurred. Typically, five cycles could be tested where 3 of 5 showing the requisite increase might dictate a necessity to proceed on the next hierarchical level of fault detection. By varying the number of test cycles, the sensitivity of detection can be varied.

Fault: Once an event is recognized, control is transferred to the fault identification routine in the detection technique. The average is frozen in the fault identification routine. This action is necessary because percentage (relative) change in the signal magnitude is used to detect faults as opposed to absolute changes from predefined fixed thresholds. One choice involved is the number of cycles after an event that must show an increase in activity before a fault is specified. The compromise is between correct identification of intermittent or very low grade faults and the possibility of identifying a normal event as a fault. Thus, an important parameter involved is the choice of the length of time after an event that is allowed for evaluating the low frequency current to make the trip/notrip decision.

The energy technique utilizes a hierarchical detection strategy in which a disturbance is classified as either a normal feeder event or a fault based on a discriminatory time window. A drawback in using such arbitrarily chosen windows is that a sporadic fault with short burst duration and regular intervals of inactivity between successive bursts may go undetected, again due to an inflation of the threshold. This technique's logic incorporates a threshold based on a set percentage of the average energy calculated from a moving time window. It is perceived that this percentage may have to be retuned depending on the load, surface condition, and the system configuration. Another factor to prevent use of this technique with high frequency components is its inability to detect high impedance fault when grounded capacitor banks intervene between the fault location and the substation.

C. Randomness Technique

The sporadic nature of high impedance faults causes problems for many proposed techniques which rely on an increase in magnitude over some predetermined period of time. Texas A&M researchers developed a technique which would not be adversely affected by this randomness, and would in fact use this randomness as an indicator of an high impedance fault.

The difference between an increase in harmonic fault current activity caused by high impedance fault and an increase caused by the energization of a "noisy" load lies in the cycle-to-cycle variations in the increase. The noise generated by a noisy load will be repetitive from one cycle to the next, while the noise generated by a high impedance fault will be quite sporadic, varying greatly from one cycle to the next (or even from one half-cycle to the next). During a fault, violent bursts will be present during some cycles, while no fault current will be drawn during other cycles; still other cycles will display a level of burst between the two extremes.

Thus, a technique which looks for several arcing bursts (periods of arcing followed by normal cycles) can discriminate step increases in noise levels from the sporadic increase caused by a high impedance fault. Also, by requiring several of these bursts in some short period of time, switching events can be recognized as such. For larger, more persistent faults, a fault can be distinguished from a step increase in noise by looking for large variations from one cycle to the next. For these reasons, a technique was developed to look for several bursts in a short period of time or for wild variations in cycle-to-cycle noise, and signal a fault if either or both are present[9].

Two separate indicators are used to detect the random, sporadic nature of arcing during a high impedance fault. The first of these is called "jumps." Jumps are meant to count the number of times that the high frequency current makes a step increase or a step decrease in a pre-determined amount of time (30 cycles). By keeping an average energy value, an abnormal increase can be detected, an event flagged, the average frozen, and the next 30 cycles checked for "jumps." It was determined that an increase in energy of ten times the calculated average indicates an event. Following this event identification, the technique checks the next 30 cycles for "jumps." A jump counter is incremented when the energy level of a cycle makes a transition from greater than ten times the average to less than three times the average or a transition from less three times the average to greater than ten times the average. An energy less than three times the average indicates an "off-cycle" in which there is no arcing. At the end of the 30 cycle decision time, if the jump counter has reached a value of five or more, a fault is indicated. If three or four jumps have been counted, the technique checks the next 30 cycles. If, however, only one or two jumps are counted, it is assumed that the event was a switching transient of some sort and the technique returns to normal system monitoring.

Some faults do not display "off-cycles." However, there always is a large difference between the fault current energy levels of consecutive cycles. To take advantage of this, the absolute difference in energy levels is calculated while the number of "jumps" are being counted. If this difference is greater than 20 times (for high frequency components of fault current) the average unfaulted energy value, a difference counter is incremented. At the end of the 30 cycle decision time, if the difference counter has reached 15, a fault is signaled. This runs in parallel to the "jump" detection portion of the technique, with a fault being indicated if the jump counter reaches 5 and/or the difference counter reaches 15.

This technique is mainly based on the behavior of the high frequency components of the fault current; therefore, the presence of a large capacitor bank may mislead this technique. Extremely noisy feeders, that is, feeders which display a high level of normal energy may prevent the detection of high impedance faults by the technique in its present form. In other words, this large normal noise level will tend to produce a large static average value which might make the parameter multiples inappropriate.

D. Phase Relationship Technique

High impedance fault current has a "notch" on the leading edge of each half-cycle when no fault current flows, due to the finite voltage necessary to break down the airgap between the conductor and ground at the fault point. This "notch" ensures that the current will be rich in odd harmonics. With this in mind, a fault detector based on third harmonic current was developed. This technique monitors the relative third harmonic phase angles between each of the three phases with respect to either the fundamental frequency current[10] or the fundamental frequency voltage[11].

If the phase of one line changes by at least 15°, the phase value before the change is held and compared against each succeeding sample for the next 5 seconds. This technique also monitors the 60 Hz line current for the rapid change indicative of normal system switching, and if a rapid change is detected, an "inhibit" signal is produced. If a relative phase change of at least 15° persists for at least 5 seconds concurrent with an increase in 60 Hz current of at least 15 amperes and no inhibit signal has been generated, a fault is indicated.

This technique monitors the relative phase changes in the third harmonic fault signal; however, third harmonic current can exist due to certain load imbalances. This technique requires at least 15 amperes of 60 Hz fault current for a duration of 5 seconds for its operation. Low grade faults that do not cause a substantial change in the fundamental component can therefore go undetected. Unbalanced capacitor bank operations result in an increase in the level of odd harmonic frequencies and may cause false operation.

E. Harmonic Sequence Component Technique

Under normal operating conditions on a balanced three-phase system, the fundamental 60 Hz current component theoretically consists of only a positive sequence component . Similarly, the third harmonic current has only a zero sequence component and the fifth harmonic current consists of only a negative sequence component. It is assumed that the degree of normal unbalance is small, and that this unbalance will increase due to any single phase fault. Thus, the non-characteristic sequence current components of the fundamental, third, and fifth harmonics can be monitored and a high impedance fault can be recognized by noting a sudden change in these components. Also, since it was seen from test data that the phase angle of each of these components often changes considerably during a high impedance fault, each component was divided into its real and imaginary parts. This technique monitors all of these quantities collectively[12].

To detect a sudden increase in the non-characteristic components, a test statistic was formed. It was assumed that each of the sequence components follows a normal distribution, and that the test statistic has a Chi-square distribution with 12 degrees of freedom and its value will increase under high impedance fault conditions.

Initial work showed that occasional "spikes" could occur in the test statistic due to such events as the energization of a large single phase load. Because of these spikes, the threshold above which a fault was detected would have to be so high that many low-grade faults would be undetectable. Thus, a modified test statistic was developed which was a smoothed version of the original test statistic. By using this modified statistic, the threshold of detection could be set lower without increasing the probability of a false detection.

The problems seen in using this detection scheme can be divided into two main categories. The first of these categories consists of true high impedance faults which are not detected. The second category, which is seen as the more serious of the two problems, consists of false detections under normal system conditions.

One of the basic assumptions made by this technique is that the detection parameters (*i.e.* the non-characteristic components of the first, third, and fifth harmonic current components) will be generated continuously by the fault. However, it has been shown that a grounded high impedance fault will draw fault current intermittently; a high impedance fault typically will arc for a period of time (on the order of a few cycles to a few tens of cycles); then, there will be no activity generated by the fault for a similar period of time[5]. Because of the snapshot delay time built into the detection scheme, it seems quite unlikely that the detector will happen to sample a "faulted" cycle for four consecutive snapshots.

Due to the design of the detection scheme and to the nature of high impedance faults, a common scenario which would produces false trips is seen. It has been observed that the activation of a capacitor bank is accompanied by a large increase in the level of odd harmonics[5]. This increased level persists until the bank is removed from the line. Since the addition of capacitor to one or two of the three phases will greatly increase the unbalance of the third and fifth harmonics, the activation of such a bank would look like a high impedance fault to a detection scheme such as this one.

F. Amplitude Ratio Technique

The even harmonic currents are very small in a system under normal condition. During a high impedance fault, with an activated arc, even harmonic currents increases due to the unsymmetrical characteristics of the arc current. Therefore a detection technique based on sensing a ratio of the phase even harmonic current to an odd phase harmonic current can be used for high impedance fault detection. A technique which monitors the ratio of the even to odd harmonic currents at the range of the first to seventh harmonic has been proposed[6]. Another technique senses the ratio of the second harmonic to fundamental current[13].

The limitation of this techniques is that the stability of this technique may not be good if the amplitude of the chosen even harmonic current is very unstable.

G. Other Techniques

The Ratio Ground Relay which is intended to detect broken conductors at the power distribution level was developed[14]. This relay is an electromechanical device whose operating characteristic is based solely on system unbalance. The operating torque is derived from the $3I_0$ summation of currents wound on an E-type magnetic element. The basic idea behind the proper operation of the relay is that under balanced conditions the $3I_0$ summation of current will be approximately zero. Furthermore, the positive sequence current will dominate. The great weakness of this technique is its fundamental dependency on system balance for proper operation. Most power distribution systems are operated unbalanced.

A sequence Component Phase Selection Technique was developed as a broken conductor detector based on changes in the symmetrical components as measured at the substation [15]. It is fairly simple to show that

$$\frac{\Delta I_2}{\Delta I_1} = (1, a, a^2)$$

for a broken conductor on phase a, b, or c, respectively. While this correctly will identify the faulted phase, it is felt that the phase must be identifiable by simple looking at which phase experiences a loss of current, without calculating the symmetrical components. This is necessarily true, since the symmetrical components are calculated from the raw phase currents; thus, for the change to be significant enough to be seen in the symmetrical components, it must also be large enough to be seen by other methods. If, however, the symmetrical components are already available because of some other technique being executed, this formulation does provide an easy way to make the phase selection decision.

H.. Summary

Several detection technique are examined. Their principles and their weaknesses are also discussed. The comparison of each technique is shown in Table I. Considering the environmental parameters, which were discussed in the Chapter II, the weakness of each technique is summarized. It is noted that while each technique may offer a solution to some degree, none to date has proven to be a powerful technique. This chapter is summarized as follows:

Table I. Comparison of Detection Techniques.

Techniques	Parameter	Attribute	Weakness
Energy Technique	Harmonics Sub-Harmonics	Energy Burst Duration	Surface Condition
Randomness Technique	High-Freq. Component	Jump Difference	Feeder Configuration
Phase Relation Technique	3rd Harmonics	Phase Angle	System Unbalance Load Type
Harmonic Seq. Comp. Technique	1st, 3rd and 5th Harmonics	Amplitude	Sys. Unbal. Load Type Feeder Config.
Amplitude Ratio Technique	Even and odd Harmonics	Amplitude Energy	Surface Condition Feeder Config.

1. No single technique covers all the cases of high impedance faults.
2. Each technique is very well formulated and based on theory and experiments, but it might work only under specific and well-defined conditions.
3. Each technique uses only a parameter to sense its attributes. The behavior of parameters of high impedance faults is very random and, in some cases, unpredictable; therefore the result of each technique is not always certain.
4. Each technique has at least one weakness with respect to the environmental parameters.
5. Each technique, even though it has at least one weakness with respect to environmental parameters, does not give any means to provide adaptability to the environment.

CHAPTER IV

AN INTELLIGENT DECISION MAKING SYSTEM

A. Introduction

A system that possesses and acquires knowledge and reasons with that knowledge, like a human, seems the most desirable intelligent system. An intelligent system is characterized by an effective analysis and synthesis of information, and a learning ability which makes the system's reactions appropriate to each situation [16]. Existing techniques have considered and focused only on one parameter for the detection of a high impedance fault under given conditions. However, the behavior of a high impedance fault is affected by many environmental parameters such as feeder configuration, surface condition, system unbalance, and load type. Therefore, a single technique cannot provide effective fault detection. Moreover, as was discussed in Chapter III, each technique has at least one weakness. To improve the analysis and synthesis of information from a distribution system, and thus to improve overall detection and discrimination under a wide variety of conditions, a combination of detection techniques is proposed. If adaptability is added to this scheme, it is felt that most high impedance fault cases can be detected and classified. Conclusively, an intelligent decision making system consists of multiple detection techniques and a learning system for adaptability.

B. Intelligent Systems

Considering an intelligent system, the term "intelligent" gives rise to confusions in its interpretation. There are many interpretations of "Someone is intelligent." However, the most satisfactory interpretation is "Someone's actions are appropriate to each situation."

Two of the ways that a person demonstrates his intelligence are by effective analysis and synthesis of a knowledge from outside, and by acquiring new knowledge through experience using learning ability. With respect to the detection of high impedance faults, an intelligent system means a system which can receive information effectively and synthesize it to detect reliably using various detection techniques. It also acquire new knowledge about each status of the power distribution system through experience, and then makes the detection adaptive by the learned knowledge.

A typical learning system consists of four major components: Critic, Learner, Rules, and Performer[17]. The Critic compares the actual with the desired output. In practice this may be a human expert, or teacher. The job of the critic is known as 'credit assignment' or alternatively 'blame assignment'. The Learner is the heart of the system. This is the portion that has responsibility for amending the knowledge base to correct erroneous performance. The Rules are the data structure that encode the system's current level of expertise. They guide the activity of the performance module. The Performer is the part of the system that carries out the task. This uses the rules in some way to guide its activity. Thus when the rules are updated, the behavior of the system as a whole changes.

An important part of learning is finding out what the basic descriptive units are in a situation. For example, in the detection of high impedance fault, a system must learn the information of each status with which descriptions of experience can be formed. In learning to classify fault and non-fault status, the system gradually becomes acquainted with such distinctive features as harmonic current amplitude variation (such as odd harmonic current increase and high frequency current variation). Amplitude, mean value, standard deviation, and distribution shape features of a power system current signal are gradually acquired as tools with which to describe the classification experience.

However, the knowledge in the high impedance fault detection environment is incomplete or inexact. In cases like these, the knowledge is inadequate to support the desired decision making. However, humans have ways of drawing inferences from incomplete, inexact or uncertain knowledge and information. Some decision making methods allow a system to use incomplete or uncertain knowledge in ways that take the incompleteness and uncertainty into account.

C. Structure of an Intelligent Decision Making System

For an intelligent detection with multiple techniques, there are three stages to consider. The first stage is to collect all the possible implementable techniques. This stage was investigated and summarized in Chapter III. Five major techniques were chosen for this stage. Another necessary step here is to change each technique slightly so that each has a final output with a basic belief instead of a decisive system status output. Because the behavior of high impedance fault is transient, random, and uncertain, it is appropriate for the output of each technique to have a certain belief level of the system status. In addition, each detection technique relies on a threshold value for detection which is arbitrarily chosen. This threshold value must be tuned depending on the surrounding circumstances. It seems to be very difficult to tune the threshold value appropriately. However, by using the threshold values of the variables in an induction process, some techniques could tune their threshold values.

The second stage is to select some techniques which prove very effective in a given situation. This is called the technique selection process. For the technique selection process, the environmental parameters are taken to be states of nature in the method of decision making under incomplete knowledge to make a technique selection database. This technique selection database consists of various important parameters and corresponding techniques. For a parameter, a best and a second best techniques are chosen using a decision making under incomplete knowledge approach. When the most important parameters in a given situation are chosen by a decision tree of an induction process, those corresponding techniques are activated to be run.

The third stage is to combine multiple pieces of information (basic beliefs) from selected techniques. This is called a technique combination process. For the technique combination process, it is necessary to find a reasoning method which can handle multiple pieces of information with uncertainty. The Dempster-Shafer's evidence theory was adopted for this reasoning process, which receives multiple basic beliefs from techniques and combines them into a final belief.

Classifying and learning a concept by induction for each class (i.e., fault and non-fault) by receiving knowledge, and using this knowledge is very important to improve the technique selection process. In addition, combined with other information such as a final belief output from combined techniques, an output from event detector, and an operator interaction, learning a new concept for each class of system status under surrounding environments and updating its detection rule are most desirable for an intelligent and adaptive detection. The technique selection process is advised by the decision tree which indicates the active and indicative parameters for a high impedance fault. To classify system status and add adaptability to its environments, an inductive reasoning with minimum entropy technique is adopted for this induction process.

As was mentioned above, an intelligent decision making system for detecting high impedance faults consists of three major parts: a technique selection process, a technique combination process, and an induction process. These all three parts are interconnected to make a learning detection system. Figure 10 shows the block diagram of an intelligent decision making system which is realized as a learning detection system.

The technique selection process examines all the possible implementable techniques and selects some of them. The reason that a method of decision under incomplete knowledge is used is that the information and the performance outcome of each technique is available but incomplete. It is very difficult or impossible to numerate exactly each performance outcome under a given condition. The environmental parameters are used as decision criteria to cover all the surrounding environments. The electrical parameters are chosen by experience. The chosen electrical parameters must positively identify high impedance faults and at the same time, possess an ability to discriminate transients associated with normal system events. Harmonic current components are the major electrical parameters. For each electrical parameter, with the method of decision under incomplete knowledge, the best and second best techniques are obtained. From the technique selection database of a parameter, the technique selection process starts. This will be explained thoroughly in Chapter V.

The technique combination process starts with the selected multiple techniques. It is assumed that each technique has its own basic belief about a system status. With these basic beliefs, the combination is performed to obtain a final belief. This will be fully investigated in Chapter VI.

The induction process will be accomplished using inductive reasoning with

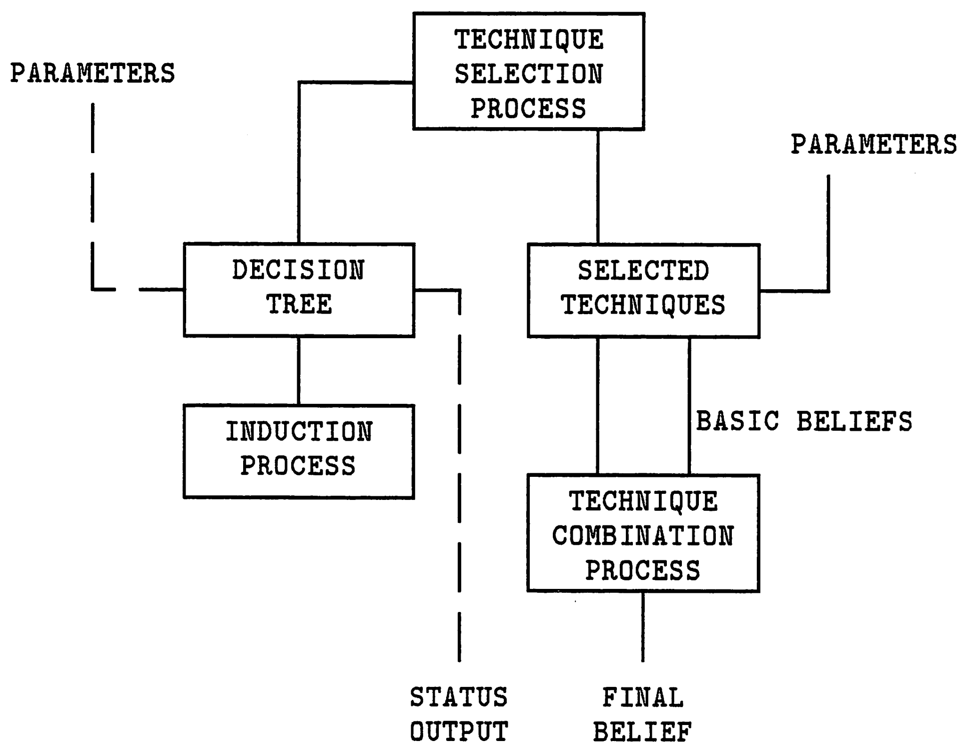


Figure 10. Block Diagram of an Intelligent Decision Making System.

minimum entropy. From the induction process a decision tree is obtained. This decision tree not only makes another detector but also indicates the most active and indicative parameters. This information will adjust the technique selection process. The induction process will be explained in Chapter VII.

All these three processes are combined into a decision control block. The final status output from the decision tree from an induction process will be combined to decide a system status, with the final belief of the combined techniques, an operator interaction, and an event detector output. The detailed structure of a learning detection system will be discussed in Chapter VIII.

D. Summary

Artificial Intelligence is a field of study that encompasses computational techniques for performing tasks that apparently require intelligence when performed by human. For an intelligent system for detecting high impedance faults, effective analysis and synthesis, and learning concepts of each class through experience is required. For learning in high impedance fault detection, the ability to be acquainted with such distinctive features as harmonic current energy variation is very important to describe the basic unit in each situation.

The structure of an intelligent decision making system is studied. It combines the idea of an intelligent system and decision making methods. An intelligent decision making system consists mainly of a technique selection process, a technique combination process, and an induction process for adaptability. All these are interconnected into a learning detection system for an intelligent and adaptive detector/learner. More detailed examination of each process will be given in following chapters.

CHAPTER V

TECHNIQUE SELECTION PROCESS

A. Introduction

One of the principal concerns of decision making is the development of analytical method for guiding the choice of a single course of action from among a series of alternatives. The classical decision model assumes that a decision maker can select one of a number of strategies open to it. As the performance of each technique under various situations is uncertain, the selected technique must operate under one of a number of mutually exclusive and exhaustive states of nature or environmental parameters. The eventual outcome of the selected technique will depend on the state of nature which happens to arise. The classical decision model with 3 techniques and 3 states of nature, for example, is shown below[18]:

$$\begin{array}{cc} & \begin{array}{ccc} E_1 & E_2 & E_3 \end{array} \\ \begin{array}{c} T_1 \\ T_2 \\ T_3 \end{array} & \begin{pmatrix} O_{11} & O_{12} & O_{13} \\ O_{21} & O_{22} & O_{23} \\ O_{31} & O_{32} & O_{33} \end{pmatrix} \end{array}$$

where

T_i: Technique

E_j: Mutually exclusive and exhaustive states of nature.

O_{ij}: Outcome of technique T_i given state of nature E_j.

Adopting the above classical decision making model and a decision tree of an induction process, a technique selection process consists of three stages: an initialization stage to provide a technique selection database, a technique selection stage, and a selection update stage. In the initialization stage, with environmental parameters as states of nature for each parameter, a best and a second techniques are obtained and stored in the database. The technique selection stage is very simple: select techniques which are stored in the technique selection database as best for given parameters which are indicated in the decision tree of an induction process. The selection update stage begins when the decision tree is changed and thus the important parameters are changed. After a specific decision model is discussed, an investigation of each stage follows in the following sections.

B. Decision Under Incomplete Knowledge

The framework of the decision problem forms a common starting point for the three approaches to decision making, which are distinguished by the amount of information they assure to be available about the probabilities with which the states of nature are likely to occur.

The first approach assumes that the decision maker is working in conditions of complete uncertainty about the future, i.e., that no information about the probabilities is available to it. When it is unable to make any statement about the vector of probabilities of the states of nature, this situation is referred to as decision making under uncertainty.

The second approach takes the view that probabilities of the states of nature can be specified uniquely, either by repeated experimentation or by eliciting unique subjective probabilities from the decision maker. When it is able to specify exactly values for all the vectors of probabilities of the states of nature, it is referred to as decision making under conditions of risk.

The third approach attempts to strike a balance between two approaches.

It assumes that in many decision problems some information is available about the probabilities of the states of nature, but that it is not comprehensive enough to enable exact specification of the probabilities. This is called decision making under conditions of incomplete knowledge.

In decision under incomplete knowledge, it is assumed that the decision maker was able to rank states of nature in terms of their importance such that $E_1 \geq E_2 \geq ... \geq E_n$. With two given techniques, T_1, T_2, each technique may have to operate in one of $j = 1, ..., n$ environmental parameters and a subjective *a priori* ranking of the importance of the environmental parameters exists, $E_1 \geq E_2 \geq ... \geq E_j \geq ... \geq E_n$. Then this approach seeks to determine under what circumstances the expected value (or performance) of T_1 will exceed or equal that of T_2, i.e., $E[T_1] \geq E[T_2]$, so that the first technique could be said to dominate the second, and vice versa. The conditions for statistical dominance are determined indirectly. Exploiting summation identity, the below theorem was derived[18].

If $E_1 \geq E_2 \geq ... \geq E_n$, then $E[T_1] \geq E[T_2]$, and
if $\left(\sum_{k=1}^{j} O_{1k}\right) \geq \left(\sum_{k=1}^{j} O_{2k}\right)$ for all $j = 1, ..., n$

The proof of this theorem is a useful step towards decision making in a situation of incomplete knowledge about the probabilities of states of nature. However, its principal practical drawback is clear. Only on a very small number of occasions are the $[O_{ij}]$ in a decision problem likely to obey the very stringent requirements of this theorem. It can easily happen that T_1 performs better than T_2 under the most likely environment, i.e., $O_{11} > O_{21}$, but T_2 may be greatly preference to T_1 under the second mostly likely environment, i.e., $O_{22} > O_{12}$. If some practical decision making aid is to be developed, a more general result than this is required, even if it may not be so unambiguous in its interpretation.

Two practical methods are listed[18]:

Derivation of maximum and minimum expected values given weak ranking of importance: The extreme expected outcomes of any technique, given an *a priori* ranking of the importance of the environmental parameters, may be found by computing n partial averages.

$$\bar{O}_j = \frac{1}{j}\sum_{k=1}^{j} O_k, \qquad j = 1, ..., n \tag{1}$$

The largest such partial average will be the maximum expected outcome and the smallest will be the minimum expected outcome.

Derivation of maximum and minimum expected values given strict ranking of importance: This case has more information about ranking of importance, i.e.,

$E_j - E_{j+1} \geq K_j, j = 1, 2, ..., n,$

where $E_{n+1} = 0$, and K_j are positive constants.

The optimum values of $E(T)$ may be found by evaluating the below equation for $j = 1, 2, ..., n$. The largest of these gives the maximum $E(T)$ and the smallest the minimum.

$$E(T) = \frac{1}{j}Y_j(1 - \sum_{j=1}^{n} jK_j) + \sum_{j=1}^{n} K_j Y_j \tag{2}$$

where

$$Y_j = \sum_{k=1}^{j} O_k$$

Here is an example. As it is mentioned in Chapter III, even though each technique has its own original parameter to monitor, many other parameters can be used without any performance degradation. Therefore, for each parameter, it is necessary to find which technique is best and which is second. Below is an example of how to calculate the outcome matrix and thus find the best technique for a given parameter. The outcome matrix is shown below:

$$\begin{array}{c} \\ T_1 \\ T_2 \\ T_3 \\ T_4 \end{array} \begin{array}{c} \begin{array}{cccc} E_1 & E_2 & E_3 & E_4 \end{array} \\ \begin{pmatrix} O_{11} & O_{12} & O_{13} & O_{14} \\ O_{21} & O_{22} & O_{23} & O_{24} \\ O_{31} & O_{32} & O_{33} & O_{34} \\ O_{31} & O_{32} & O_{33} & O_{44} \end{pmatrix} \end{array}$$

T_i indicates the i^{th} technique, E_i indicates the i^{th} environmental parameter as the state of nature, and O_{ij} indicates the performance outcome of the technique i and the environmental parameter j. This O_{ij} value in the matrix indicates the subjective or *a priori* detection performance of the technique. From the performance outcome of every situation, the average performance outcome of each technique is obtained. After obtaining the performance outcome, a technique is selected for a parameter with a certain reliability value without regard to any outside environment conditions, ignoring whatever environment is outside.

Quantifying the performance outcome is a subtle problem. While it is possible to predict the performance variation in each environmental parameter, it is a different matter to quantify it. Therefore the only way is to rely on the previous experiments and the experts' subjective or *a priori* ratings. Then, the outcome matrix can be rated as below:

$$\begin{array}{c} \\ T_1 \\ T_2 \\ T_3 \\ T_4 \end{array} \begin{array}{c} \begin{array}{cccc} E_1 & E_2 & E_3 & E_4 \end{array} \\ \begin{pmatrix} 80 & 10 & 10 & 80 \\ 80 & 20 & 80 & 70 \\ 80 & 50 & 40 & 80 \\ 80 & 80 & 80 & 50 \end{pmatrix} \end{array}$$

This table shows that T_1 shows 80 percent performance when it is exposed to the E_1 environmental condition, 10 percent for E_2, and so on. It is assumed that the order of the importance of the environmental parameters is $E_1 \geq E_2 \geq E_3 \geq E_4$, and this rating is weakly ordered. Then, the partial average of each technique can be calculated as below in terms of environmental parameters:

For T_1:

$O_1 = 80/1 = 80, O_2 = (80 + 10)/2 = 45$

$O_3 = (80 + 10 + 10)/3 = 33, O_4 = (80 + 10 + 10 + 80)/4 = 45$

Therefore the maximum is 80 (for E_1) and minimum is 33 (for E_3).

For T_2:

$O_1 = 80/1 = 80, O_2 = (80 + 20)/2 = 50$

$O_3 = (80 + 20 + 80)/3 = 60, O_4 = (80 + 20 + 80 + 70)/4 = 62$

Therefore the maximum is 80 (for E_1) and minimum is 50 (for E_2).

For T_3:

$O_1 = 80/1 = 80, O_2 = (80 + 50)/2 = 65$

$O_3 = (80 + 50 + 40)/3 = 56, O_4 = (80 + 50 + 40 + 80)/4 = 62$

Therefore the maximum is 80 (for E_1) and minimum is 56 (for E_3).

For T_4:

$O_1 = 80/1 = 80, O_2 = (80 + 80)/2 = 80$

$O_3 = (80 + 80 + 80)/3 = 80, O_4 = (80 + 80 + 80 + 50)/4 = 72$

Therefore the maximum is 80 (for E_1) and minimum is 72 (for E_4).

Therefore the maximum is 80 and the maximin is 72 at T_4, so the 4th technique is chosen as best one for a given parameter.

C. Initialization

Initialization starts with the existing techniques. The existing techniques are:

T1: Energy Technique.

T2: Randomness Technique.

T3: Phase Relationship Technique.

T4: Harmonic Sequence Component Technique.

T5: Amplitude Ratio Technique.

The electrical parameters are not restricted, but the behavior of high impedance faults is characterized and well indicated with harmonic currents. The electrical parameters are:

P1: Odd Harmonic Current.

P2: Even Harmonic Current.

P3: Sub-Harmonic Current.

P4: High Frequency Current (2 - 10 KHz).

P5: Third Harmonic Current.

P6: Third and Fifth Harmonic Current.

P7: Second Harmonic Current.

The activities of these electrical parameters are monitored and indicated in terms of the mean, standard deviation, and the mean absolute value of each parameter to be used as variables of an induction process.

The next step is to find a best technique for each parameter. To find a technique, a specific decision making method is used. While the information and the performance of each technique is available but incomplete, it is not impossible to rank the importance of the environmental parameters with experts' knowledge and subjective ratings. Therefore, the decision making method under incomplete knowledge is adopted. For the states of nature, environmental parameters are used, which are the factors affecting the performance of the techniques and the behavior of high impedance faults. The environmental parameters as the states of nature are:

E1: System Unbalance.

E2: Surface Condition.

E3: Feeder Configuration.

E4: Load Type.

Each environmental parameter is important, but some are more important than others. It is assumed that the environmental parameters above are listed by the order of level of importance. From all these the outcome matrix is formed.

The next step is to find the outcome of each technique-environmental parameter pair. This is done with respect to the advantages/disadvantages of each technique which are investigated in Chapter III, using the decision making method under incomplete knowledge.

Here is an example outcome calculation. For parameter P2, even harmonic current, the energy technique is not substantially affected in its detection performance by the system unbalance because the energy technique is based on the arcing phenomenon characteristic of a high impedance fault. However, because of this, the energy technique is affected by the surface condition, since this surface condition determines the arcing behavior in terms of the amplitude increase and the length of arc burst duration. In general, even harmonic current do not vary under changed feeder configurations such as presence/absence of a capacitor bank. Usual load and some solid state devices generate only odd harmonics, thus the energy technique using even harmonics is not affected by the load type except in the case of arc welding process.

With these discussions and assumptions, and inevitably also by subjective ratings, the outcomes of the energy technique for environmental parameters are:

$$\begin{array}{cc} & \begin{array}{cccc} E1 & E2 & E3 & E4 \end{array} \\ T1 & \begin{pmatrix} 90 & 60 & 90 & 70 \end{pmatrix} \end{array}$$

From this matrix, the partial average of the energy technique is derived using Equation (1).

$O11 = 90/1 = 90, \qquad O12 = (90 + 60)/2 = 75$

$O13 = (90 + 60 + 90)/3 = 80, \ O14 = (90 + 60 + 90 + 70)/4 = 78$

Therefor the maximum is 90 and the minimum is 75 for the energy technique. The outcomes of other techniques for the parameter P2 can be obtained similarly. The largest partial average gives the maximum performance and the smallest the minimum. Selecting a technique with their own maximums and minimums uses maximin partial average to indicate the best technique and maximax for the second best in case of extreme optimism.

In its original form, each technique uses only one parameter to indicate the system status. Because the behavior of parameters vary situation to situation it is desirable to use the most active parameter to indicate high impedance faults. Therefore, one parameter can be replaced by another; however, every technique cannot use all the parameters for detection. For example, the phase relationship technique uses only a single odd harmonic current to detect high impedance fault. Table II indicates the parameters and the usable techniques

Table II. Electrical Parameters and Usable Techniques.

Harmonics	Techniques				
	Energy	Random -ness	Phase Relation	Harmonic Sequence	Amplitude Ratio
Odd Harmonics	x	x	x	x	
Even Harmonics	x	x			x
Sub-Harmonics	x	x			
High Freq. Component	x	x			
3rd Harmonic	x	x	x		
3rd and 5th Harmonics	x	x	x	x	
2nd Harmonic	x	x			x

for each parameter.

The partial averages of the technique for other parameters are obtained similarly. Tables A(1) through A(7) of Appendix A show the outcome matrices for parameters. Even though the matrices can be obtained, the matrix for each parameter is not complete. It seems that there is no possible way to get a complete outcome matrix. This outcome matrix needs many experiments and tests for its refined form. This matrix will be changed and refined by several real operation experiences. The changing and refining of the outcome matrix is left unsolved in this research.

From Tables A(1) - A(7) of Appendix A, the technique selection database is formed as shown in Table III.

D. Technique Selection

Technique selection is performed using the technique selection database and a decision tree. The decision tree has the form of a tree whose branches have important parameters which are very active in a given situation. Using the branch information, technique selection will pick the best technique for each parameter on the branch. If the number of selected techniques to be run is less than two, the second best techniques may be picked. Therefore, at least two different techniques are run at the same time. Figure 11 shows an example of a decision tree, a technique selection database, and selected techniques to be run.

When field data comes into a data treatment block, the whole parameters which are listed at the initialization set-up are generated through A/D converters and various filters. With all the parameters arrived at the statistical measure calculation block, statistical measures such as mean, absolute mean, and standard deviation of each parameter are obtained. The statistical measures of chosen

Table III. A Technique Selection Database.

Parameters	Best Technique	Second Technique
Odd Harmonics	Randomness Technique	Energy Technique
Even Harmonics	Energy Technique	Randomness Technique
Sub-Harmnics	Randomness Technique	Energy Technique
High Frequency Components	Energy Technique	Randomness Technique
3rd Harmonic	Randomness Technique	Phase Relation Technique
3rd and 5th Harmonic	Harmonic Seq. Comp. Technique	Energy Technique
2nd Harmonic	Randomness Technique	Amplitude Ratio Technique

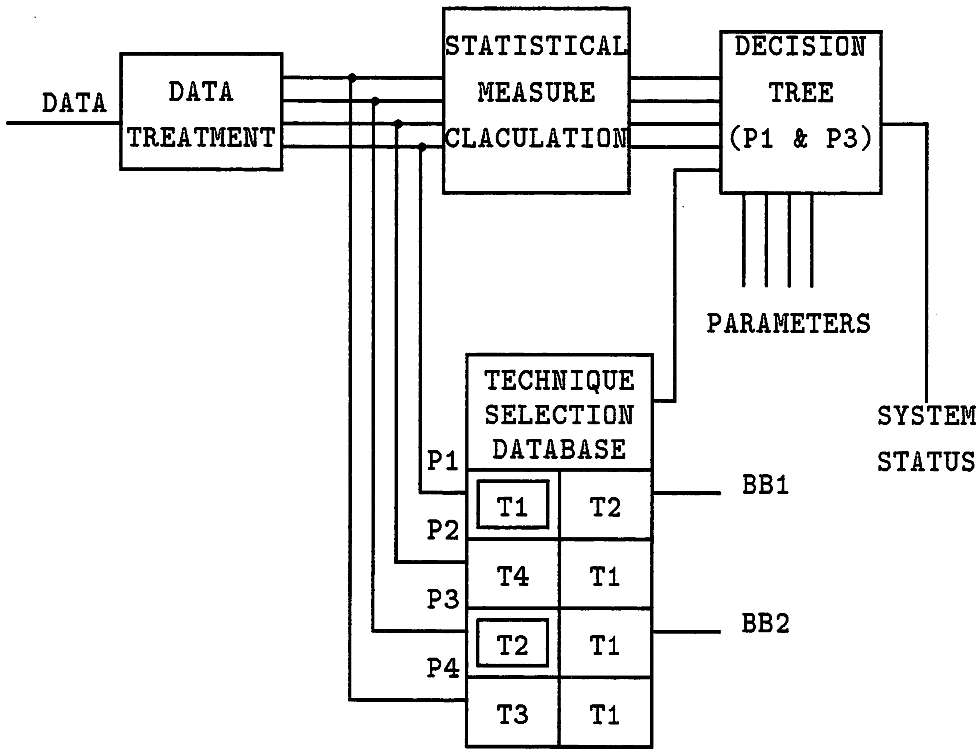


Figure 11. An Example of Technique Selection Process.

parameters are examined by a decision tree to identify a system status. The statistical measures of all the parameters are forwarded to a learning detection system to decide the class of the data. At the same time, all the parameters arrive at the technique selection block. At this block, the best and second best techniques are connected to a corresponding parameter; for example, for even harmonic current signal, a randomness technique and an energy technique are directly connected.

Then, by a decision tree which has rules indicating important indicative parameters for detection purpose, the parameter list will trigger and activate techniques to be run with the parameters. In this example, the chosen parameters are P1 and P3. The decision tree activates the techniques T1 and T2. T1 will be run with the parameter P1 and T2 with P3. Then, two basic beliefs come out of these two techniques about the system status.

Selection is updated whenever a decision tree structure is changed and thus the active parameter list is updated. Thereafter the list of techniques to be run is changed. The other case in which the selection update occurs is when the outcome matrices are changed by any means. This will force the selection process to be updated.

E. Summary

Technique Selection is performed by an initialization, a technique selection, and a selection update stage. When a decision tree is updated, technique selection is also updated. With the assumptions and discussion made, the total structure of technique selection was illustrated and explained.

CHAPTER VI

TECHNIQUE COMBINATION PROCESS

A. Introduction

One of the many difficulties associated with the development of an intelligent detection system is the fact that the information from each detection technique contains a considerable amount of uncertainty. The behavior of various high impedance faults is not consistent. Under similar conditions, faults show different characteristics in behavior and may draw less harmonic current than other faults under the same conditions. Sometimes the behavior of faults can be very similar to the phenomenon of the switching events. Moreover, the behavior of high impedance faults under various environmental conditions is so random that it is nearly impossible to characterize it by probabilistic qualities.

To accommodate uncertainty, an intelligent detection system must have some way of calculating its confidence in a conclusion in proportion to the level of the evidence. The most widely used representation of uncertainty has been in terms of a probability distribution which has a profound mathematical basis. However, many situations do not lend themselves to such convenient treatment and special approaches have to be utilized[19]. One approach widely used to tackle uncertain outcome, termed the *engineering approach*, is to avoid explicit representation of uncertainty and substitute an uncertain system with an equivalent *certain* system. The engineering approach to uncertainty is simply to formulate one's problem to avoid explicit representation of and reasoning with uncertainty. Most AI programs engineer uncertainty out of the task domain to some extent. The potential problem with this is that one may lose a valuable source of constraints. This is especially the case when it is uncertain how to use available information, and so it is simply ignored[20].

The evidence-gathering process in high impedance fault detection with multiple techniques requires a method of combining the support for a hypothesis, or for its negation, based on multiple, accumulated observations. Dempster-Shafer's evidence theory provides a formal proposal for its management[21]. Given two belief functions, based on two observations, Dempster's combination rule computes a new belief function that represents the impact of the combined evidence.

B. Evidence and Basic Belief

The evidence theory (or Dempster-Shafer theory), like Bayesian theory, relies on degrees of belief to represent uncertainty. Unlike Bayesian theory, however, it permits one to assign degrees of belief to subsets of hypotheses. In Bayesian theory, one constructs a probability distribution over all individual singleton hypotheses, but in the evidence theory, a distribution is constructed over all subsets of hypotheses. This is a great advantage.

For example, in a high impedance fault detection problem, there are three different system status to be identified: $Fault$, $Event$, and $Normal$. Imagine information from a power distribution system in a given time, and the hypothesis about the system status. In addition to the three singleton hypotheses there are two others: that the system is under $Disturbance$ status which indicates that system is under $Event$ or $Fault$ status, and that the system is under $Non-Fault$ status which indicates either $Normal$ or $Event$ status. Suppose the evidence that the probability of $Disturbance$ is 0.6 is obtained. In the evidence theory it is possible to assign the probability 0.6 to the set $\{Event, Fault\}$ without committing any of that 0.6 to either member of the set. In the Bayesian approach it would be natural to artificially divide the 0.6 among *Event* and *Fault*. Furthermore, in the Bayesian scheme, it seems to be forced to assign the remaining 1 - 0.6 = 0.4 of the probability distribution to *Normal*. In contrast, the evidence theory approach assigns the remaining 0.4 to the set $\{Fault,\ Event,\ Normal\}$, that is, to the set that reflects ignorance about system status.

The detection techniques introduced in Chapter III ignore all the uncertainty in the high impedance fault environment. Therefore it is necessary to change them slightly to have a probabilistic output such that the output shows the form of certainty level on the system status. One example output is that, based on the information for the high impedance fault detection monitoring data, the output result of the technique A is $Fault(0.6)$, or $Event(0.7)$.

The Dempster-Shafer theory uses a number in the range[0,1] to indicate belief in a hypothesis given a piece of evidence[21,22]. This number is the degree to which the evidence supports the hypothesis. The impact of each distinct piece of evidence on the subsets of *frame of discernment*, denoted Θ, which are all the cases of cause or status, is represented by a function called a *basic belief*. A basic belief is a generalization of the traditional probability density function; the latter assigns a number in the range [0,1] to every singleton of Θ such that the numbers sum to 1. The quantity $m(A)$ is a measure of that portion of the total belief committed exactly to A, where A is an element of 2^{Θ} and the total belief is 1. This portion of belief cannot be further subdivided among the subsets of A and does not include portions of belief committed to subsets of A. It would be useful to define a function that computes a total amount of belief in A. This quantity would include not only belief committed exactly to A but belief committed to all subsets of A. Such a function is called a *belief function*.

A belief function, denoted Bel, corresponding to a specific basic belief, m, assigns to every subset A of Θ the sum of the beliefs committed exactly to every subset of A by m. For example,

$$\begin{aligned}Bel(fault, event, normal) = {} & m(fault, event, normal) + m(fault, event) \\ & + m(event, normal) + m(fault, normal) \\ & + m(fault) + m(event) + m(normal).\end{aligned}$$

Thus, $Bel(A)$ is a measure of the total amount of belief in A and not of the amount committed precisely to A by the evidence giving rise to m. $Bel(\Theta)$ is always equal to 1 since $Bel(\Theta)$ is the sum of the values of m for every subset of Θ.

C. Combination of Belief Functions

As discussed earlier, the evidence-gathering process or reasoning process for detection with multiple techniques requires a method for combining the support for a hypothesis, or for its negation. The Dempster-Shafer model also recognizes these requirements and provides a formal proposal for their management. Given two belief functions with the same frame of discernment, Dempster's combination rule computes a new belief function that represents the impact of the combined evidence.

Let Bel_1 and Bel_2 and m_1 and m_2 denote two belief functions and their respective basic beliefs. Dempster's rule computes a new basic belief, denoted $m_1 \oplus m_2$, which represents the combined effect of m_1 and m_2. The corresponding belief function, denoted $Bel_1 \oplus Bel_2$, is then easily computed from $m_1 \oplus m_2$ by the definition of a belief function.

As an example, suppose that for a given system one technique supports a disturbance or not-normal status to degree $0.6(m_1)$ whereas another unconfirms the fault (i.e, confirms $(event,\ normal)$) to degree $0.7(m_2)$. Then the net belief based on both observations is given by $m_1 \oplus m_2$. For computational purposes, an "intersection table" with values of m_1 and m_2 along the rows and columns, respectively, is a helpful device. Table A(8) of Appendix A shows the intersection table with values of m_1 and m_2.

In this example, a subset appears only once in the table and $m_1 \oplus m_2$ is easily computed for the status:

$m_1 \oplus m_2(event) = 0.42$

$m_1 \oplus m_2(fault, event) = 0.18$

$m_1 \oplus m_2(event, normal) = 0.28$

$m_1 \oplus m_2(\Theta) = 0.12$

$m_1 \oplus m_2$ is 0 for all other subsets of Θ.

Since $Bel_1 \oplus Bel_2$ is fairly complex, only a few sample values are given:

$$\begin{aligned} Bel_1 \oplus Bel_2(fault, event) &= m_1 \oplus m_2(fault, event) \\ &+ m_1 \oplus m_2(fault) + m_1 \oplus m_2(event) \\ &= 0.18 + 0.0 + 0.42 = 0.60 \end{aligned}$$

$$\begin{aligned} Bel_1 \oplus Bel_2(event, normal) &= m_1 \oplus m_2(event, normal) \\ &+ m_1 \oplus m_2(event) + m_1 \oplus m_2(normal) \\ &= 0.28 + 0.42 = 0.70 \end{aligned}$$

$$Bel_1 \oplus Bel_2(fault) = m_1 \oplus m_2(fault) = 0.0$$

In this example, the table does not contain null entries because every two sets with nonzero basic belief values always have at least one element in common.

In general, nonzero products of the form $m_1(X)m_2(Y)$ may be assigned when X and Y have an empty intersection. The Dempster-Shafer theory deals with this problem by normalizing the assigned values so that $m_1 \oplus m_2(\emptyset) = 0$ and all values of the new basic belief lie between 0 and 1. This is accomplished by defining κ as the sum of all nonzero values assign to $\emptyset$ in a given case. This theory then assigns 0 to $m_1 \oplus m_2(\emptyset)$ and divides all other values of $m_1 \oplus m_2$ by $1 - \kappa$.

As an another example to the above, suppose now that for the same system, a third technique(m_3) confirms the status of fault to the degree of 0.8. Now it is necessary to compute $m_3 \oplus m_4$, where $m_4 = m_1 \oplus m_2$ of the above example. The intersection table of 3 values is shown in Table A(9) of Appendix A.

In this example, there are two null entries in the table, one assigned the value 0.336 and the other 0.224. Thus

κ=0.336+0.224=0.56 and $1 - \kappa$=0.44

$m_3 \oplus m_4(fault) = (0.144 + 0.096)/0.44 = 0.545$

$m_3 \oplus m_4(event) = 0.084/0.44 = 0.191$

$m_3 \oplus m_4(fault, event) = 0.036/0.44 = 0.082$

$m_3 \oplus m_4(event, normal) = 0.056/0.44 = 0.127$

$m_3 \oplus m_4(\Theta) = 0.024/0.44 = 0.055$

$m_3 \oplus m_4$ is 0 for all other subsets of Θ.

Therefore the final belief function for fault status in this example is 0.545.

D. Technique Combination with Dempster-Shafer Theory

High impedance fault detection with multiple techniques is well suited for implementation with the Dempster-Shafer theory. First, mutual exclusivity of singletons in a frame of discernment is satisfied by the sets of hypotheses in the detection scheme constituting the frames of discernment (Fault, Normal, Event). Second, the belief functions that represent the evidence in the detection scheme are of a particularly simple form. Moreover, the way the techniques indicate the system status fits very well into the theory.

The conceptual structure in the detection scheme is very simple, which has the form of (status, basic belief). This double gives rise to a single hypothesis of the form "the technique shows the status with a confidence of basic belief." A frame of discernment would then consist of all doubles with the same technique. Thus the number of doubles, or hypotheses in Θ, will equal the number of possible status that the technique may assume for the system in question. In other words, the frame of discernment is either (Fault, Event, Normal) or (Fault, Normal). The system status is fault, event, and normal; thus the representation scheme is particularly well suited to the mutual exclusivity demand of the theory. The simplified belief function, several types of evidence combination, and an example of the evidence combination will be discussed in this section.

In the most general situation, a given piece of evidence supports many of the subsets of Θ, each to a varying degree. The simplest situation is that in which the evidence supports only one subset to a certain degree and the remaining belief is assigned to Θ. Because of the characteristics of the detection technique, this applies in the combination of the conclusion of detection techniques. If the premise confirms the conclusion of a technique with degree s, where s is above threshold value, then the technique's effect on belief in the subsets of Θ can be represented by a basic belief. This basic belief assigns s to the singleton corresponding to the hypothesis in the conclusion of the technique, call it A, and assigns $1 - s$ to Θ. If the premise unconfirms the conclusion with degree s, then the basic belief assigns to the subset corresponding to the negation of the conclusion, $\neg A$, and assigns $1 - s$ to Θ[22].

There are three types of evidence combination in the high impedance fault detection environment. They are explained below.

Type 1: Two techniques are both confirming or both unconfirming of the same conclusion. For example, both techniques confirm $(fault)$, one to degree 0.4 and the other to degree 0.7. The effect of triggering these two is represented by basic beliefs m_1 and m_2, where $m_1(fault) = 0.4$, $m_1(\Theta)$=0.6, and $m_2(fault) = 0.7$, $m_2(\Theta)$=0.3. The combined effect on the belief is given by $m_1 \oplus m_2$.

Type 2: One technique is confirming and the other unconfirming of the same singleton hypothesis. For example, one technique confirms $(fault)$ to degree 0.4, and the other unconfirms $(fault)$ to degree 0.8. The effect of triggering these two techniques is represented by basic beliefs m_1 and m_3, where m_1 is defined in the example from Type 1 and $m_3(\neg fault)$=0.8, $m_3(\Theta)$=0.2. The combined effect on belief is given by $m_1 \oplus m_3$.

Type 3: Two techniques involve different hypotheses in the same frame of discernment. For example, one technique confirms $(fault)$ to degree 0.4, and the other unconfirms $(normal)$ to degree 0.7. The triggering of the second technique gives rise to m_4 defined by $m_4(\neg normal)$=0.7, $m_4(\Theta)$=0.3. The combined effect on belief is given by $m_1 \oplus m_4$.

An implementation in the detection scheme of the Dempster-Shafer methods which minimizes computational complexity is discussed here. Combining the functions in a simplifying order to reduce the computations[23] is adopted as is well shown in the reference[22]. The steps for simplified evidence combination are explained below.

First, for each singleton hypothesis, combine all basic beliefs representing techniques confirming that hypothesis. If $s_1, s_2, ..., s_k$ represent different degrees of support derived from k techniques confirming a given singleton, then the combined support is

$$1 - (1 - s_1)(1 - s_2)...(1 - s_k)$$

Similarly, for each singleton, combine all basic beliefs representing techniques unconfirming that singleton. Thus all evidence confirming a singleton is pooled and represented by a basic belief, and all evidence unconfirming the singleton is pooled and represented by another basic belief. Then there are $2n$ basic beliefs, where n is the size of Θ. In the high impedance fault detection case, $n = 3$: fault, event, and normal. Secondly, combine the two basic beliefs computed in the first step. Then there are n combined basic beliefs, which are denoted $CBB_1, CBB_2, ..., CBB_n$. Thirdly, Combine the combined basic beliefs computed in the second step in one computation, using simplified formulae, to obtain a final belief function Bel.

The form of the required computation is shown below[23]. Let $[i]$ represent the ith of n singleton hypotheses in Θ, that is, $i = 1$ for fault, $i = 2$ for event, and $i = 3$ for normal. Then p_i indicates the ith confirming basic belief of n singleton hypotheses, c_i the ith unconfirming basic belief, and r_i the rest of the basic belief as shown below.

$CBB_i[i] = p_i$

$CBB_i[\neg i] = c_i$

$CBB_i[\Theta] = r_i$

Since $p_i + c_i + r_i = 1$, $r_i = 1 - p_i - c_i$. Let $d_i = c_i + r_i$. Then it can be shown that the function Bel resulting from combination of $CBB_1, ..., CBB_n$ is given by

$$Bel([i]) = K[p_i \prod_{j \neq i} d_j + r_i \prod_{j \neq i} c_j]$$

For a subset $\neg A$ of Θ,

$$Bel(\neg A) = K([\prod_{all} d_j][\sum_{j \notin A} p_j / d_j] + [\prod_{j \in A} c_j][\prod_{j \notin A} d_j] - \prod_{all j} c_j)$$

where

$$K^{-1} = [\prod_{all j} d_j][1 + \sum_{all j} p_j/d_j] - \prod_{all j} d_j$$

as long as $P_j \neq 1$ for all j.

As an example, consider the net effect of the following set of basic beliefs from various techniques regarding the signal information of the system. Assume that the final conclusion about the beliefs in the competing hypotheses from detection techniques will be based on the following successful information:

T1 unconfirms $(fault)$ to the degree 0.6

T2 unconfirms $(fault)$ to the degree 0.2

T3 confirms $(normal)$ to the degree 0.4

T4 unconfirms $(event)$ to the degree 0.8

T5 confirms $(normal)$ to the degree 0.3

T6 unconfirms $(fault)$ to the degree 0.5

T7 confirms $(fault)$ to the degree 0.3

T8 confirms $(event)$ to the degree 0.7

Each step of belief combination follows.

Step 1:

Considering first confirming and then unconfirming basic belief for each status, following is obtained:

a. $(fault)$ confirmed to the degree s_1=0.3, unconfirmed to the degree s_1'=1-(1-0.6)(1-0.2)(1-0.5)=0.84.

b. $(event)$ confirmed to the degree s_2=0.7, unconfirmed to the degree s_2'=0.8.

c. $(normal)$ confirmed to the degree s_3=1-(1-0.4)(1-0.3)=0.58, unconfirmed to the degree s_3'=0.

Step 2:

Combining the conforming and unconfirming basic belief for each system status:

$$CBB_1(fault) = \frac{0.3(1-0.84)}{1-(0.3)(0.84)} = 0.064 = p_1$$

$$CBB_1(\neg fault) = \frac{0.84(1-0.3)}{1-(0.3)(0.84)} = 0.786 = c_1$$

Thus r_1=0.15 and d_1=0.786+0.15=0.936.

$$CBB_2(event) = \frac{0.7(1-0.8)}{1-(0.7)(0.8)} = 0.318 = p_2$$

$$CBB_2(\neg event) = \frac{0.8(1-0.7)}{1-(0.7)(0.8)} = 0.545 = c_2$$

Thus r_2=0.137 and d_2=0.545+0.137=0.682.

$$CBB_3(normal) = 0.58 = p_3$$

$$CBB_3(\neg normal) = 0 = c_3$$

Thus r_3=0.42 and d_2=0.42.

Step 3:

Assessing the effects of belief in the various system status on each other:

$K^{-1} = d_1 d_2 d_3(1 + p_1/d_1 + p_2/d_2 + p_3/d_3) - c_1 c_2 c_3$

=(0.936)(0.682)(0.42)(1+0.064/0.936 +0.318/0.682 +0.58/0.42)

-(0.786)(0.545)(0)

=0.268(1+0.068+0.466+1.38)

=0.781

Therefore, K=1.28

The final combined belief function of each status is calculated:

$Bel(fault) = K(p_1 d_2 d_3 + r_1 c_2 c_3)$=0.023

$Bel(event) = K(p_2 d_2 d_3 + r_2 c_1 c_3)$=0.160

$Bel(normal) = K(p_3 d_1 d_2 + r_3 c_1 c_2)$=0.704

From above final belief functions, the only important belief is a belief of fault. Therefore, if the belief of fault is larger than a preset threshold value, the final system status will be output as a single form of (fault). In the above case, the final status is (non-fault). Dempster-Shafer theory is particularly appealing in its potential for handling evidence bearing on types of the status. However, the representation of the system status needs to be studied more. Some techniques make an output result of fault or event, but others fault but not event.

E. Summary

Technique combination using the evidence theory is investigated in this chapter. The basic belief and the combination of basic beliefs are illustrated with examples of high impedance fault detection case. A detailed calculation process for a final combined basic belief of a status is explained with a simplified formula.

CHAPTER VII

CLASSIFICATION WITH INDUCTIVE REASONING

A. Introduction

An intelligent device, including man, can be thought of as something capable of adjusting to its environment. To make such adjustments the device must continually be classifying slightly different status of the environment as equivalent or not equivalent[24]. It has been shown that inductive reasoning can improve the classification of faults and switching events and, combined with other information such as a final belief from combined techniques and operator interaction, can be used as a learning system to adapt a detection system to its environments with induced rules[25].

For a learning process in an intelligent decision making system, inductive reasoning for classification/recognition is adopted. Inductive reasoning for pattern classification is based on entropy minimization. Entropy is a measure of the disorder in classifying data. Entropy minimization is an ordering principle by which it is determined which past events are more like future events in ways that are sufficient for predicting the outcome. If one wishes to classify a set of samples into a more ordered state, he attempts to lower the entropy. Therefore, by minimizing entropy, rules for classification can be induced. By this rule, the concept of each class, fault and non-fault, is learned from sample event. This rule and decision tree indicate the knowledge of the classes of a given situation.

B. Induction Theory

The end of induction is to discover a law having objective validity and universal application. However, the conclusions of a process of inductive reasoning need be no more general than merely annexation of one new particular to those particulars taken in the beginning. Thus beginning with the particular "all high impedance faults observed to date show high frequency current bursts," it might be concluded with the universal "all high impedance faults show high frequency current bursts" or with the less general "all high impedance faults observed to date show high frequency current bursts and it is highly probable that the next high impedance fault to be observed shows high frequency current bursts."

The essential principles of induction have been known for centuries. Three laws of induction are summarized below[26]:

1) Given a set of irreducible outcomes of a trial, the induced probabilities are those probabilities consistent with all available information which maximize the entropy of the set.
2) The induced probability of a set of independent observations is proportional to the probability density of the induced probability of a single observation.
3) The induced rule is that rule consistent with all available information for which the entropy is minimum.

It is not necessary to assign probability factors if the separation of two classes can be done easily and formally. The third law is appropriate for classification and the second law for calculating the reliability of the rule and mean probability of each step of separation. From the second law, the following is derived[26].

$$<p>=\frac{x+t}{n+t+f}$$

where

$<p>$ is a mean probability,

t is the number of distinguishable True states,

f is the number of distinguishable False states,

n is the number of observations,

and x is the number of observations classified as True.

When there are only two classes, the mean probability becomes,

$$< p >= \frac{x+1}{n+2}.$$

The classical problem of the third law of induction is the problem of pattern recognition, that is, classification. In classification the probability aspects of the problem is completely disregarded; it is simply asked whether it is True or False. Using this, a description of two classes are obtained.

C. Pattern Classification

Information consists of statements as to the values of probabilities. A key goal of entropy minimax analysis is to determine the quantity of information in a given data set. A quantity of information is defined as proportional to the negative of the logarithm of probability.

$I = -k \log P.$

where $P = 1/m$ in the equi-probability case.

This information measure compares the contents of data one receives to prior state of expectation. The higher was one's prior estimate of the probability for an outcome to occur, the lower will be the information one gains by observing it to occur.

The entropy on a set of possible outcomes of a trial where one and only one outcome is True is defined as

$$S = -k \sum_{i=1}^{N} p_i \ln p_i$$

In other words, the entropy is the expected value of the information. On the other hand, the entropy of a rule is defined as below:

$$S = -k \sum_{i=1}^{m} x_i p_i \ln p_i$$

where,

m is the number of total steps,

i is the step,

x_i is the number of samples of either True or False at step i,

p_i is a mean probability of the i^{th} step for either True or False,

and k is a constant.

The third law of induction which is typical in pattern recognition says that the entropy of a rule should be minimized to have a simple and reliable rule. To find the minimum entropy and its corresponding rule, all the possible combinations of steps and of rules must be investigated. The number of possible n digit numbers is $N = 2^n$. A rule which separates these N numbers into two classes is desired. There are N^2 ways of separating N numbers into two classes. If there are only m samples, then the number of available patterns are 2^m. This means, if there are 7-digit numbers and only 31 sample patterns, then there are 2^{31} ways of separation.

For simplifying the steps in rule induction, some easily derivable relations and tips for entropy minimization are derived.

1) The closer p_i to 1 or 0 is, the smaller S is.

2) The bigger x_i is, the higher p_i is.

3) The bigger the digit index of digit n is, the bigger x_n is.

In addition, the *digit index* is defined as the rating of separation into two classes in each digit for both True and False classes. For example, suppose there are two classes of True and False and each class has four, 3-digit numbers as below:

TRUE	FALSE
101	001
010	110
011	111
000	100

If the first digit index is considered, and choose 0xx as True (x means don't care), then there is one wrong separation in True and another one in False. For the second digit, if either x0x as True or x1x as True was chosen, there are two wrong separations at both classes. So, apparently using the first digit is better than using the second digit.

The following are steps to find a digit index to determine which digit is most important to separate numbers into two classes. First, count the number of 1's under True and count the number of 0's under False, and divide each number by the number of samples in each class. Then the digit counts, d_n's, are as given below:

TRUE			FALSE		
d_1	d_2	d_3	d_1	d_2	d_3
0.25	0.5	0.5	0.25	0.5	0.5

Then, adding together digit by digit, the result is

d_1	d_2	d_3
0.5	1.0	1.0

If the value of d_n is closer to 1, that digit is not important to separate. That is, 1's and 0's have the same weight (number, frequency, or importance) on both sides. If the value of d_n is away from 1, there are less 1's or 0's in one class, so the digit index of digit n, I_n, is defined here using d_n. The maximum value of I_n is 1.

$$I_n = |d_n - 1|$$

Therefore, from the above example,

$I_1 = 0.5$, I_2=0.0, and I_3 =0.0, so the first digit has the maximum digit index. This digit can be used to separate two classes if only one digit is used to classify. The following is the simplified version of entropy minimization.

1) Find the maximum digit index. Then separate two classes using the digit index by either 1 for True or 0 for False.
2) Eliminate those samples which are subsets of the above step.
3) Find the maximum digit index from remainders.
4) Do steps 1) - 3) until the two classes are empty.

D. Threshold Value

In a pattern classification problem, a rule derivation assumes the values of interests are 1's and 0's. In a discrete system, it is easy to assign 1's and 0's, but in a continuous system, certain values must be set to divide the sample events into 1's and 0's. This is called a threshold value. If the threshold value is changed, the 1/0 table, which is a converted sample list of binary form, is changed and so are the entropy values and rules. How to set a threshold value to separate two classes efficiently and simply is the point of the next discussion.

The idea is simple: find a threshold value for each variable which results in minimum entropy for that variable. In this case, the problem is still a classification system. The only difference is that this case deals with only a variable, but the previous one (rule induction) is classification with all the variables. Therefore, the idea of minimum entropy of the induced rule is similarly used here in threshold value calculation. Thus, finding a threshold value is finding a rule for a variable. Assume that a threshold value for a variable in the range of $X1$ to $X2$ is being sought. Then for this variable only, the entropy equations are written as below[26]:

$$S(x) = p(x)S_p(x) + q(x)S_q(x)$$

$$S_p(x) = -\sum_{k=1}^{m} p_k(x) \ln p_k(x)$$

$$S_q(x) = -\sum_{k=1}^{m} q_k(x) \ln q_k(x)$$

where,

$S(x)$ is the entropy of a variable value x,

$S_p(x)$ is the entropy of a variable in the region $X1$ to $X1 + x$,

$S_q(x)$ is the entropy of a variable in the region $X1 + x$ to $X2$,

$p_k(x)$ is the probability of the k^{th} outcome class given that the variable value is in the region $X1$ to $X1 + x$,

and $q_k(x)$ is the probability of the k^{th} outcome class given that the variable value is in the region $X1 + x$ to $X2$.

m is the number of outcome classes,

In addition, relatively unbiased estimates for $p_k(x)$ and $p(x)$ are defined below.

$$p_k(x) = \frac{n_k(x)}{n(x)}$$

$$p(x) = \frac{n(x)}{n}$$

where,

$n_k(x)$ is the number of samples located between $X1$ and $X1 + x$ in the k^{th} outcome class,

$n(x)$ is the total number in this region in all the classes,

and n is the total number of samples.

Now x can be chosen by various manipulations. This x can be any value between $X1$ and $X2$. Here steps are made in the value of x, i.e., x_i's. If $S(x_i)$ is minimum, then $X1 + x_i$ is the threshold value. So if each digit has its minimum entropy at the threshold value, then the total entropy for rule is also minimized.

The following is a simplified and coarse example for finding a threshold value. Here there are only two classes, so k in the above equations has values 1 and 2 or T (for Non-Fault) and F (for Fault). Therefore m equals 2. The sample data for each class is shown below.

T: 10 20 16 18 24 33 47 74

F: 76 83 45 90 66 33 72 84

From the above list, $X1$=10 and $X2$=90. By following the procedure of calculating a threshold value, it is easily found that the minimum occurs at x_2 , i.e. at 26. A detailed calculation example will be shown with staged fault data in Chapter IX. Thus, 36 (this value comes from $X1 + x_2$) can be used as a good threshold value. Now, 36 is a border line for 1 and 0 logic, so the 1/0 table is:

T: 0 0 0 0 0 0 1 1

F: 1 1 1 1 1 0 1 1

Assuming that only this digit is used for separation, the rule is:

$0 \rightarrow$ T, or $1 \rightarrow$ F.

Here there are two wrong decisions in T and one wrong decision in F. This result can be compared with that of the median. If the border line is chosen as the median, i.e., 47, there are two wrong decisions in T and another two wrong decisions in F. The 1/0 table is omitted for the case of the median. Here it is shown that the threshold value is better than any arbitrarily chosen value in this simple example. However, the more important point is that it has a basis for finding a border line for 1 and 0 that leads into a better position in rule induction. Thus the induced rule will be simpler and more reliable with this threshold value. These threshold values which minimize entropy can be adopted by an induction process in its threshold value calculation for the variable. In the technique selection process, by using these threshold values, each technique adjusts and tunes its threshold value appropriately. A variable value, in T (Non-Fault) class, which is greater than a threshold value may become a detection threshold for a technique.

E. Rule Induction and Decision Tree

Induction with entropy minimization is basically for pattern recognition, that is to classify two classes. However, this tool for detection can be applied to a learning process. How this induction is combined with other processes to become a learning system is explained below.

1) Initialize the induction process with the variables' sample data. The variables are given by historical data as an initial setting. The class of the sample is also given by a human expert or historical data with experience of actual operation or staged fault tests.
2) Start the induction process calculating the threshold values of each variable. With threshold values the variables of each sample data are converted into

binary values. From these binary values, the simplified minimum entropy process follows. The final decision rule will be derived with an induction process that indicates which parameters are most active in a given situation and, if this rule is applied as a detector, it is combined with the final belief of the combined techniques, an output from event detector, and an operator interaction to find the final status about a distribution feeder.

3) A decision rule or decision tree is updated when experimental sample data are provided as operational experiences. This process is realized by a learning detection system. The same induction process resumes to obtain a new decision rule and decision tree.

For inductive reasoning, sample data with a corresponding class is essential. Historical staged fault data or actual operation data can be used for initial set-up. With induction as many variables as possible are used to better describe each class effectively. As discussed in Chapter V, harmonic currents are used as parameters. These are listed below:

Odd Harmonic Current

Even Harmonic Current

Sub-Harmonic Current

High Frequency Current

Third Harmonic Current

Third and Fifth Harmonic Current

Second Harmonic Current

With these parameters, it is necessary to find any statistical measure which can indicate the activities of each parameter for each class, and thus indicate the high impedance fault. Mean, standard deviation, and mean absolute value are chosen for this purpose. Then, the variables which can be used for the induction process are the combination of parameters and statistical measures as shown in Table IV. Therefore, there are 21 variables to be used in the induction process. The length of data for measuring these variable values are limited to, for convenience, 30 cycles.

The induction process is used to draw a decision rule or tree from the sample data. With the sample data, threshold values are calculated as the first step. With calculated threshold value for each variable, the 1/0 binary table is formed. With this 1/0 binary table, a simplified formula for deriving a rule applies. The final rule and decision tree result. The rule is based on no more than the available information, and on all of the available information. Therefore, the detection rule might be changed in different cases. An investigation was done to find how the detection rule changes as the high impedance fault case changes.

As was mentioned in Chapter II, the behavior of high impedance faults is random and transitory. However, it is possible to categorize its behavior in terms of the fault current amplitude increase and the arc burst duration. Five different cases were studied. The first and second cases are high impedance faults and non-faults on wet soil. The third, fourth, and fifth cases are on dry soil. The result of this study is summarized in Table V. Detailed rule derivation is explained in Chapter VIII and Appendix B. From Table V, the following are found.

1. No one general rule is found, therefore, different rules are necessary at different sites.
2. Even harmonics and sub-harmonics are generally important parameters to classify the activities of high impedance faults and non-faults.
3. If amplitude increase is small and arc burst duration is short, then the sub-harmonic is the best parameter for classification.

Table IV. The Variables for an Induction Process.

MEAN OF ODD HARMONIC CURRENT (mnodd)

MEAN ABSOLUTE OF ODD HARMONIC CURRENT (amodd)

STANDARD DEVIATION OF ODD HARMONIC CURRENT (sdodd)

MEAN OF EVEN HARMONIC CURRENT (mneve)

MEAN ABSOLUTE OF EVEN HARMONIC CURRENT (ameve)

STANDARD DEVIATION OF EVEN HARMONIC CURRENT (sdeve)

MEAN OF SUB-HARMONIC CURRENT (mnsub)

MEAN ABSOLUTE OF SUB-HARMONIC CURRENT (amsub)

STANDARD DEVIATION OF SUB-HARMONIC CURRENT (sdsub)

MEAN OF HIGH FREQUENCY CURRENT (mnhgh)

MEAN ABSOLUTE OF HIGH FREQUENCY CURRENT (amhgh)

STANDARD DEVIATION OF HIGH FREQUENCY CURRENT (sdhgh)

MEAN OF 3RD HARMONIC CURRENT (mn3rd)

MEAN ABSOLUTE OF 3RD HARMONIC CURRENT (am3rd)

STANDARD DEVIATION OF 3RD HARMONIC CURRENT (sd3rd)

MEAN OF 3RD AND 5TH HARMONIC CURRENT (mn35)

MEAN ABSOLUTE OF 3RD AND 5TH HARMONIC CURRENT (am35)

STANDARD DEVIATION OF 3RD AND 5TH HARMONIC CURRENT (sd35)

MEAN OF 2ND HARMONIC CURRENT (mn2nd)

MEAN ABSOLUTE OF 2ND HARMONIC CURRENT (am2nd)

STANDARD DEVIATION OF 2ND HARMONIC CURRENT (sd2nd)

Table V. Rules for Different Fault Cases.

Examples	Faulted Signal Description		Harmonic Parameters				Remarks
	Mag.	Burst	Rule 1	Rule 2	Rule 3	Rule 4	
1 (site 1)	Large Medium	Long Medium Short	Even	ANY			Even
2 (site 1)	Medium Small	Medium Short	Sub	ANY EXCEPT EVEN			Sub
3 (site 2)	Medium	Long Medium	Sub	Sub 3rd			Sub
4 (site 2)	Large	Long	ANY	ANY			
5 (site 2)	Medium Small	Medium Short	Sub	Odd 3rd			Sub Odd
1 + 2 (site 1)	Large Medium Small	Long Medium Short	Even	Even			Even
3 + 4+ 5 (site 2)	Large Medium Small	Long Medium Short	Sub	Odd 3rd			Sub Odd
1+2+3+4+5 (site 1+2)	Large Medium Small	Long Medium Short	Even	Even	2nd	Sub	Even Sub

4. If amplitude increase is large and arc burst duration is long, then any harmonic parameters can be selected for classification.

When there is any discrepancy between the confirmed status and the actual status, the decision tree needs to be changed and updated. For the induction process, data with known classification is necessary. However, when the system is operated, there is not any way to confirm that the final system status decided upon by the intelligent system is correct. In this circumstance, it is impossible to save the data and class as a correct sample data to be used in the induction process. An intelligent system always can be ready to ask and to get replies from an operator about whether there is any discrepancy. However, the way the operator replies and provides the data and corresponding class to the memory has a time frame problem. To solve this problem, a learning detection system which combines a technique selection, a technique combination, and an induction process is proposed. A decision control block inside the learning detection system combines and synthesizes the status output from an induction process, the final belief from the combined techniques, an operator input, and an output from an event detector. Here, it is assumed that when the final combined decision from a learning detection system is a "Fault", then a "TRIP" signal is issued and the distribution feeder is disconnected. Then, an operator provides a confirmation or unconfirmation input to the learning detection system. The event detector indicates any disturbance or event with high frequency current or sub-harmonic current. Combining all these inputs, a learning detection system successfully learns about the system and parameters, identifies the data which makes the controller decide such results, and places the identified data into a fixed-size memory. This memory is accessed by an induction process later to update rules. The newly induced tree will replace the old one, and thus update the technique selection process, too. The learning detection system will be discussed thoroughly in Chapter VIII.

F. Summary

Inductive reasoning is explained. The initial set-up, the induction process, and the update process are examined. A learning detection system is briefly discussed. A detailed example, with a learning detection system implementation, will be shown in Chapter VIII with a training set and a test set of staged fault data.

CHAPTER VIII

IMPLEMENTATION OF AN INTELLIGENT DECISION MAKING SYSTEM

A. Introduction

An intelligent decision making system is useful for effective analysis of information and thus for effective and reliable detection of high impedance faults. In addition, it adjusts to its environment making the system adaptive by using a learning concept. The intelligent decision making system consists of a technique selection, a technique combination, and an induction process. The realization and implementation of an intelligent system with these three major parts will be discussed in this chapter. A learning detection system which combines all three parts is proposed to realize an adaptive detector and learner. An example execution with this implementation method is shown with sample training data and testing data.

B. Learning Detection System

As was mentioned, an intelligent decision making system consists of mainly a technique selection, a technique combination, and an induction process. However, the interconnection among them and the basic scheme of learning and detection has not been discussed. A learning detection system which combines all these processes to make an adaptive detector/learner is the subject of this discussion.

As important indicative parameters are selected with an existing rule, a best and second best technique are selected with respect to the chosen parameters. Then these techniques receive the corresponding input parameters from a power distribution system to detect a fault or normal status. At the same time, a decision tree which also was formed from an existent rule accepts parameters to decide the system status. The outputs from both a decision tree and combined techniques can be mixed or compared in a certain manner and have a final decision on the system status. This final decision activates a relay to trip or not to trip.

In this case, the problem is the quality of the final decision. There is a chance that this final decision is not correct. Then with this incorrect classification, the corresponding data as an incorrect class will be stored into memory from which an induction process updates its rule. Because using incorrectly identified stored data is very dangerous and unreliable, the problem of the reliability of the final decision, which is also known as the problem of an unreliable teacher, should be minimized and ideally eliminated. The problems of an unreliable teacher are summarized below.

1. Type I error: The system is under a Fault status, but the final decision is "Non-Fault."
2. Type II error: The system is under a Non-Fault status, but the final decision is "Fault."
3. There is no way to indicate the actual system status with only the three major parts.

To meet these problems of an unreliable teacher, operator interaction and an event detector are assumed. Operator interaction provides confirmation or unconfirmation information on the final decision of fault. When a final decision of fault is made and the "TRIP" signal is issued, with a power line disconnected, then an operator checks and replies if it is actually a fault or not. This information provides the actual system status (i.e., reliable teacher) and is very essential to eliminate Type II errors. By this confirmation/unconfirmation, the data class to be stored is correctly identified. However, this operator interaction still does not provide any useful information to eliminate Type I errors, because an operator interaction is activated only after a final decision of "Fault."

To minimize Type I errors, an almost perfect event detector is assumed. This event detector is a part of an existent fault detector which has a sensitivity problem. It detects any disturbance or event on the distribution line. High frequency current or sub-harmonic current, which is very sensitive to any kinds of transients, is the best candidate for an event detector. If a disturbance or event is detected by this detector, a '1' output will be issued, otherwise, a '0' output will be issued.

If the final decision is not "Fault," but there are some numbers of consecutive events, then it is suspected that a fault condition might be present and remain undetected. In this particular situation, a "WARNING" signal is issued to make an operator pay attention to the power system. Until there is a response from an operator, all the data which are involved in this particular decision will be remained unclassified and thus not stored as valid data. When an operator response is entered, with confirmation or unconfirmation, then a correct system status is obtained and thereafter data classification can be achieved.

To realize the above functions, a learning detection system is proposed which controls an event detector, decision rule, technique selection, and technique combination, in a decision control block. The structure of a learning detection system is shown in Figure 12. Data storage is memory where classified data are stored. This data storage, which has a fault data section and a non-fault data section, is a fixed-size FIFO (First In First Out) memory. Therefore, when a new classified data comes in, the first data goes out.

The signals inside a learning detection system are E, R, T, O, and D. E

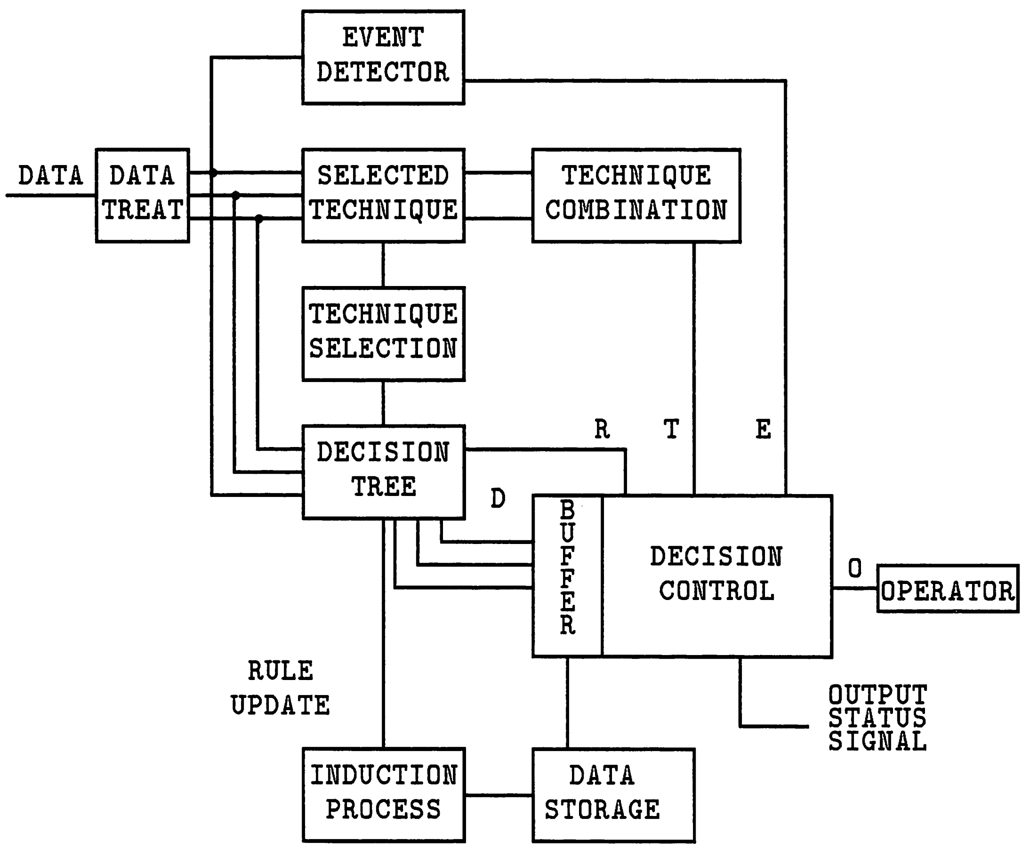


Figure 12. Structure of a Learning Detection System.

indicates an event output (1 for event, 0 for otherwise), R indicates a status output from a decision tree of an induction process (1 for fault, 0 for otherwise), T indicates a final belief from combined techniques (1 for fault, 0 for otherwise), O indicates an operator input (1 for confirmation, 0 for unconfirmation), and D indicates a data class to be stored as a valid data (1 for fault class, 0 for otherwise). These signals are combined and synthesized in the decision control block. The reasons that all the signals are used to get one final output are summarized below.

1. A decision rule from an induction is based on no more than the available information, and on all of the available information. Therefore, it might miss some unexperienced and very unusual data.
2. Techniques are usually based on a general theory and a few experiments. Therefore they are very sensitive in some cases and insensitive in other cases.
3. Technique selection is *indirectly* adjusted by a decision rule. Technique selection is performed by experimental and subjective ratings with respect to the parameters which are chosen by a rule. This technique selection database is seldom complete.
4. An ideal match between the status output from a decision tree and a final belief from combined techniques is always expected. However, in case, higher security against false decision making can be highly guaranteed by combining all the processes.

A decision control block controls and decides a sequence to a final decision: "TRIP," "WARNING," or nothing. If assuming 3 digit numbers and each number indicate the status of E, R ,and T from first to last. Then any uninterrupted 3 counts of {011, 111, 101, 110} will cause a "TRIP" signal, any uninterrupted 6 counts of {010, 001, 100} with less than 2 counts of any {011, 111, 101, 110} will cause a "WARNING" signal, and any counts of {000} will cause no output and the corresponding data is classified and stored as non-fault data (i.e., D=0). In addition, less than 6 counts of any combination of {010, 001, 100} will cause no output and the involved data will be classified and stored as non-fault data. The numbers 3 and 6 are adjustable. The detailed functions of a decision control block and some examples using this function are summarized in Table VI.

The time frame of the learning detection system is shown in Figure 13. When a "TRIP" signal is issued, a relay is tripped and the line is disconnected, then the operator input makes a learning detection system classify the data which are involved with this particular final decision and remained unclassified. The classified data are stored in the memory as valid experience data. With these new memory contents, a new rule is induced. This new rule will substitute an existent rule, and adjust a technique selection.

The decision comes out of a learning detection system every half a second. Sometimes, the amount of data which are pending unclassified should be considered. Therefore, a data buffer is necessary in the decision control block. The structure of the data at the decision control block consists of the statistical measures of each parameter and the flags of the D, E, R, and T signals. The data flow between a decision tree and the technique selection block was already discussed in Chapter V. The data storage contains such data that was classified. The data will be used in the induction process to update rules. This data flow in a learning detection system is already shown in Figure 12. High frequency current is directly connected to an event detector without any control output.

Table VI. Control Functions of a Decision Control Block.

3 $\{E \times (R+T)\} \longrightarrow$ ``TRIP", D is unclassified.
3 $\{E \times \cdot (R+T)\} \times O \longrightarrow D=1$ for all 3.
3 $\{E \times (R+T)\} \times \bar{O} \longrightarrow D=0$ for all 3.

3 $(R \times T) \longrightarrow$ ``TRIP", D is unclassified.
3 $(R \times T) \times O \longrightarrow D=1$ for all 3.
3 $(R \times T) \times \bar{O} \longrightarrow D=0$ for all 3.

6 $\{\bar{E} \times (R \oplus T)\} \longrightarrow$ ``WARNING", D is unclassified.
6 $\{\bar{E} \times (R \oplus T)\} \times O \longrightarrow D=1$ for all 6.
6 $\{\bar{E} \times (R \oplus T)\} \times \bar{O} \longrightarrow D=0$ for all 6.

$\bar{E} \times \bar{R} \times \bar{T} \longrightarrow D=0$.

6 $E \longrightarrow$ ``WARNING", D is unclassified.
6 $E \times O \longrightarrow D=1$ for all 6.
6 $E \times \bar{O} \longrightarrow D=0$ for all 6.

EXAMPLES:

E	R	T		
1	1	0		
0	1	1		
1	0	1	$\longrightarrow$	``TRIP"

E	R	T		
1	0	1		
1	0	0		
1	1	0		
1	0	0		
0	1	0		
1	0	0	$\longrightarrow$	``WARNING"

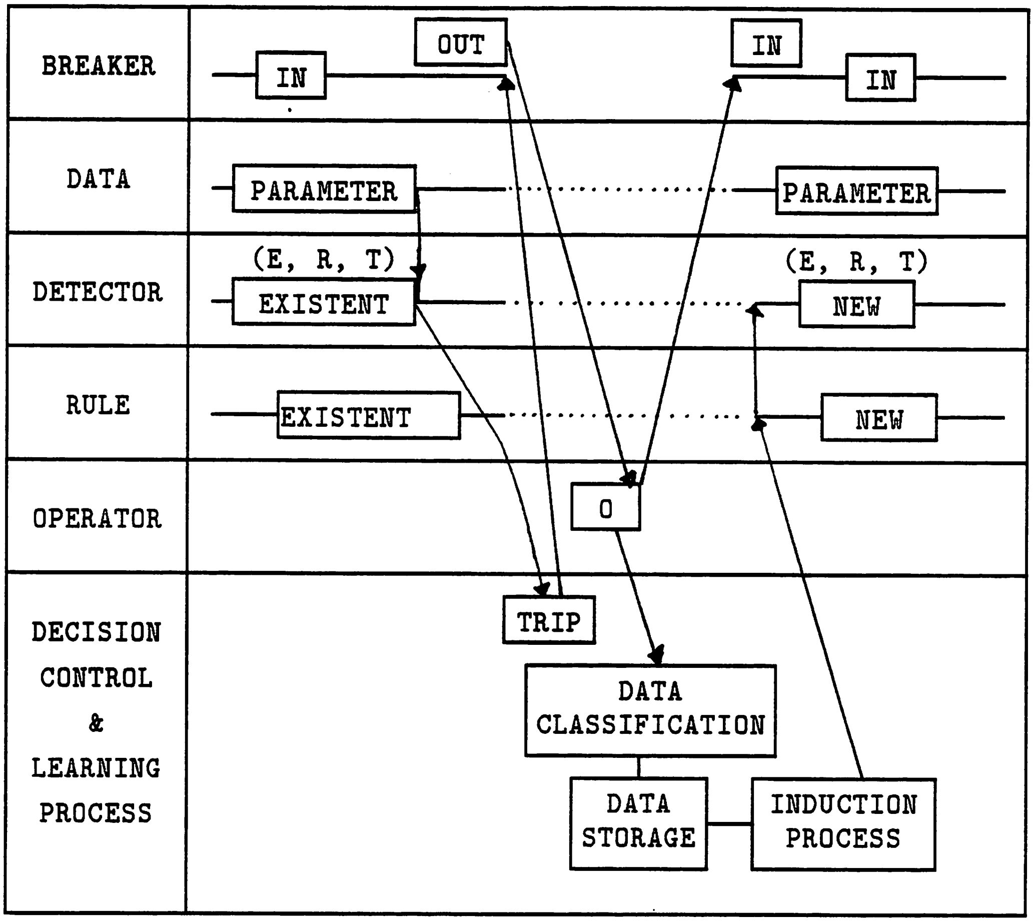


Figure 13. Time Frame of a Learning Detection System.

C. An Example Execution

With sample data, an example execution with this intelligent decision making system is performed. The training sample data consists of high impedance faults, capacitor bank operations, and load tap changer operations. In the real world, the high impedance fault case is not frequent; therefore, for the number of data in the training case, it seems appropriate to provide more non-fault data than fault data. A total of 58 sample data were chosen and 17 of them were fault data and the rest of them were non-fault data. Figure 14 shows part of a high impedance fault waveform which is used for the knowledge acquisition purpose in a learning process. With these training sample data, an induction process can describe each class in terms of statistical measures. To calculate the statistical measures for each parameter, the waveforms were passed through a notch-filter to minimize the fundamental component. From this notched waveform, each electrical parameter is obtained through a corresponding filter. Figure 15 shows waveforms of two electrical parameters. Odd harmonics and even harmonics are shown. With these electrical parameters, means, standard deviations, and mean absolute values of parameters are obtained over a fixed window of 30 cycles. The Section B(1) of Appendix B shows the variable list for this training waveform.

Then, the threshold value calculation follows. A program was written to calculate each threshold value with the statistical measures. This program just follows the threshold value calculation process explained in Chapter VIII. This threshold value calculation program is shown in Appendix C. The threshold value for each variable is shown in the Section B(2) of Appendix B. The threshold value calculation for the first variable of odd harmonic currents is also shown in Section B(2) in detail. The 1/0 binary table shown in the Section B(3) of Appendix B is formed using these threshold values. From this 1/0 table the induction process

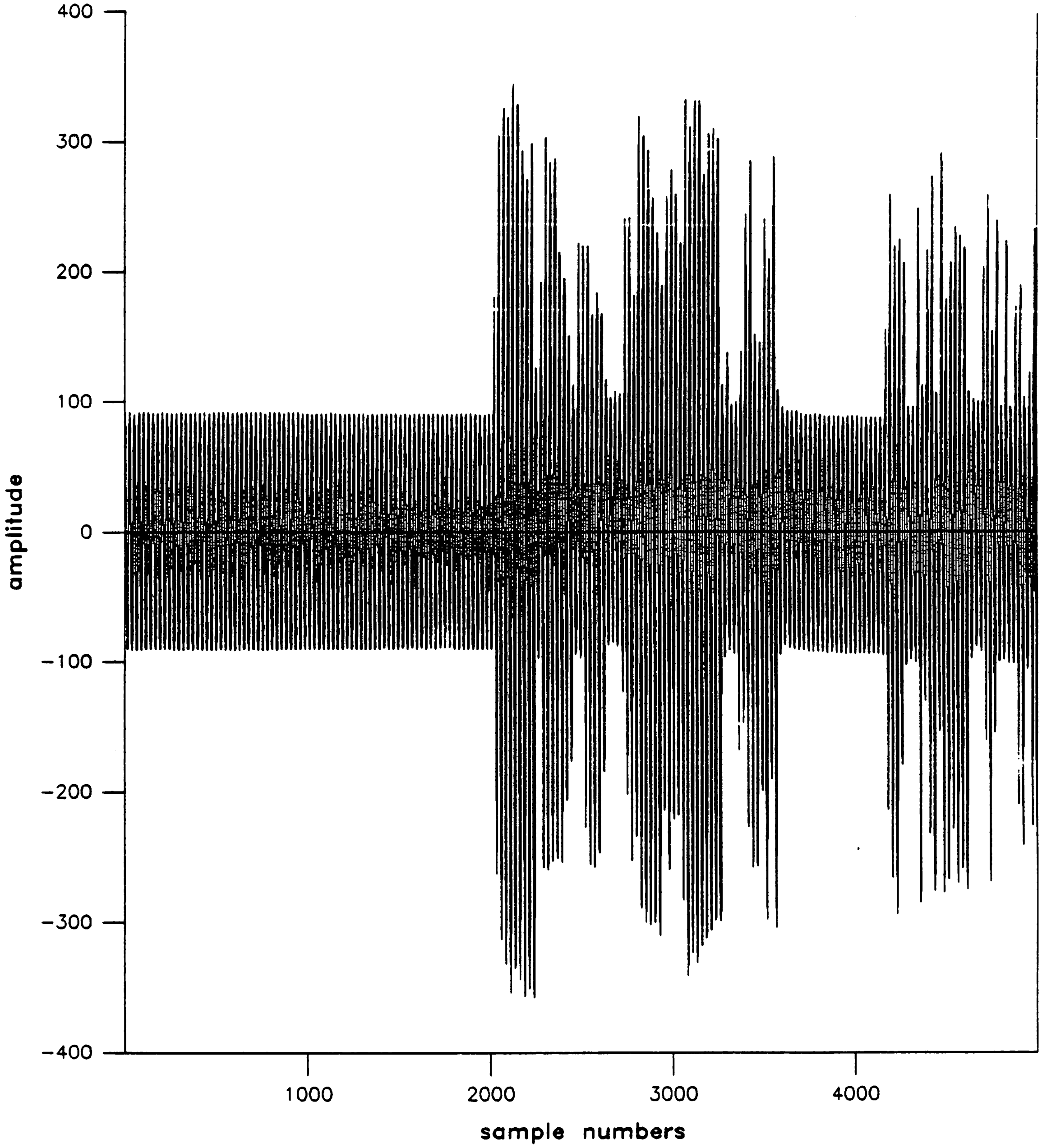


Figure 14. Training Sample Waveform.

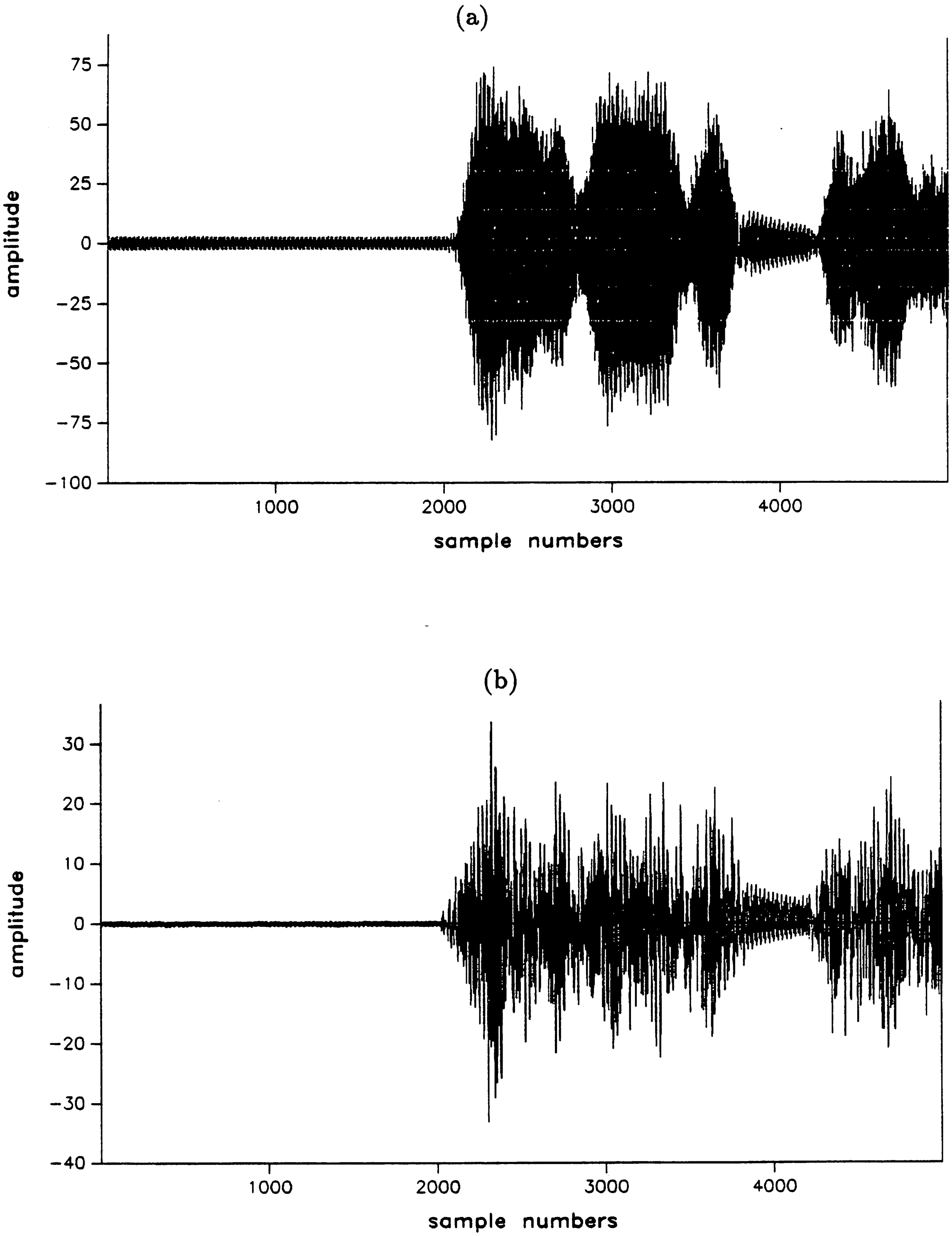


Figure 15. (a) Odd and (b) Even Harmonic Components of Training Waveform.

starts to derive rules. Along with the simplified minimum entropy formula explained earlier, the maximum digit index to classify two classes is found. From the maximum digit index, the classification starts. With this procedure a first rule is derived. The next step is to eliminate all the samples which have the same digit formation as the first rule. From the remaining samples, the same procedure repeats until the two classes are empty. This process is performed by writing a rule induction program. This simplified minimum entropy formula program is shown in Appendix C. In the Section B(4) of Appendix B, with the 1/0 table of Section B(3), the rule derivation process is illustrated. The 6th digit has a maximum digit index, so the first rule is on this 6th digit: 1 for Fault. After this first rule, all the samples which have 1 in their 6th digit are eliminated. From the remaining samples, using a similar procedure, Rule 2 is derived. From these rules, a decision tree is finally derived.

Testing sample data is somewhat complicated in the high impedance fault case. It has various kinds of bursts in terms of amplitude increase and of length. Some parts of this sample data have a few normal or event-looking data, but the conductor is still on the ground. These testing sample data with calculated statistical measures of each variable are listed in the Section B(5) of Appendix B. Figure 16 shows a test sample waveform. The first four samples are apparently normal. At the fifth sample, a fault starts with a large amplitude increase, and long and medium length burst durations. Sample numbers 10 and 11 look normal, or at least like an event by observing the total current, but, the conductor is on the ground; they are actually under a fault. Sample numbers 12 through 16 show apparent fault only by the total current. Sample numbers 17 and 18 show normal or event by observing total current; however, they are under a fault. Sample 19 is apparently fault data. By the decision rule of the

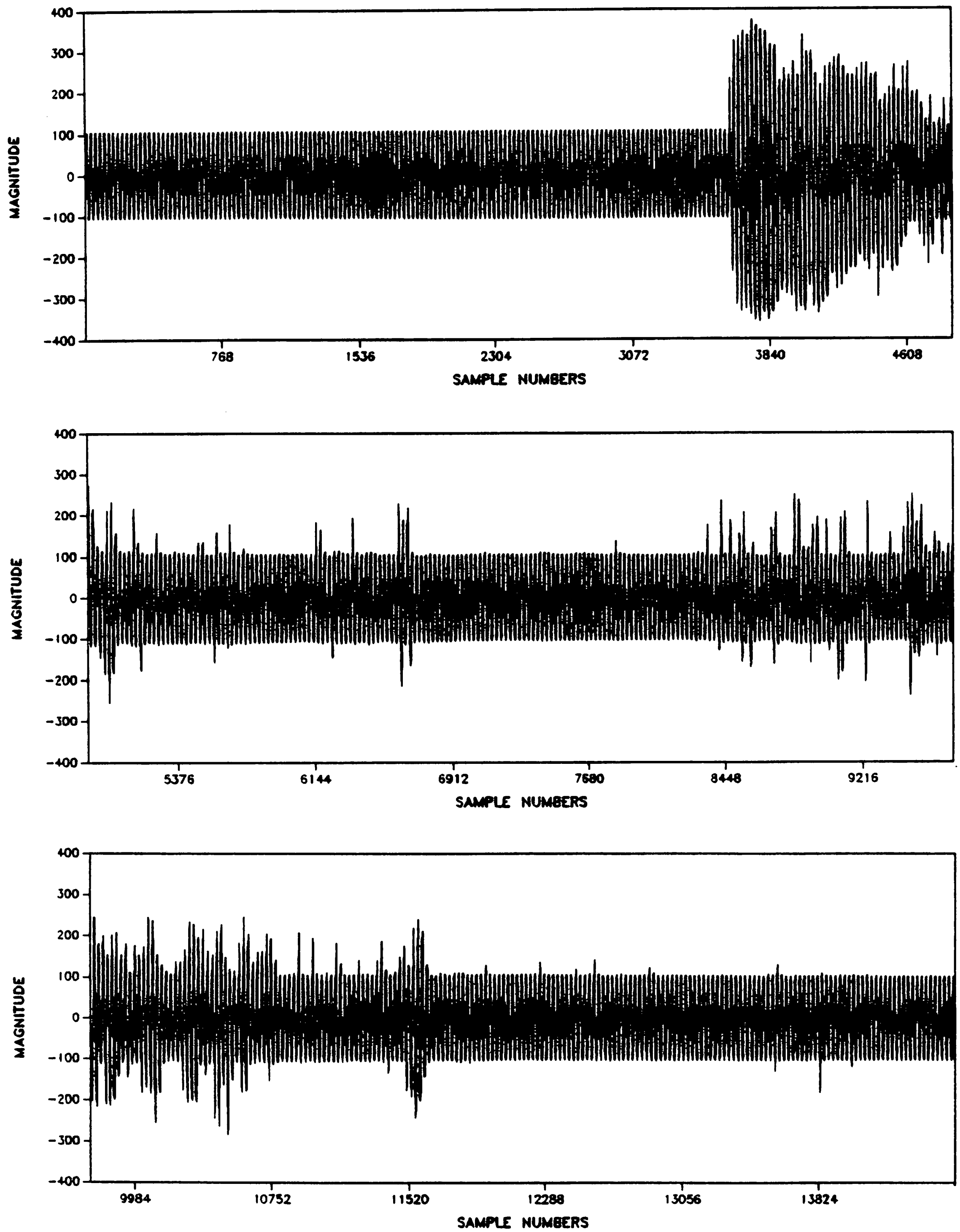


Figure 16. Test Sample Waveform.

Section B(3) of Appendix B, from the technique selection database of Table III, the techniques to be run are selected; energy technique with even harmonics and the randomness technique with odd harmonics. For the energy technique, a burst is defined as a cycle of energy which is greater than 10 times the normal static average energy. This threshold value for defining a burst is obtained roughly by using the threshold value of a variable of the induction process, AMEVE (mean absolute value of even harmonic current). In the Non-Fault class, a value which is maximum or above the threshold value should be the threshold for detecting a fault. The average of the all the values which are below the threshold value for AMEVE, which is 1.936, is obtained as 0.728. There is only one sample which is above the threshold value, and the value is 1.977. Because the definition of a burst is expressed in terms of a cycle of not sample values but energy increase, the squared ratio of above two calculated values results: This value is 7.37. However, to secure a strict safety against false trips, the threshold value of 10 is chosen. For the basic belief, for convenience, the following relationship is derived by the technique and the experiments:

1. If the number of bursts in the 30 cycle window is less than 3, the basic belief is (Normal, 0.8). This means, in this case, the technique is 80 percent confident that it met a normal status.
2. If the number of bursts in the 30 cycle window is between 4 and 6, the basic belief is (Event, 0.8).
3. If the number of bursts in the 30 cycle window is between 7 and 9, the basic belief is (Fault, 0.6).
4. If the number of bursts in the 30 cycle window is between 10 and 14, the basic belief is (Fault, 0.7)
5. If the number of bursts in the 30 cycle window is more than 14, the basic

belief is (Fault, 0.8).

In Table VII, for each test sample, the number of bursts and the corresponding basic beliefs from the energy technique using even harmonics are shown.

For the randomness technique, instead of "jumps," the "differences" are chosen because "jumps" are not as sensitive to long duration bursts. The "difference" is the absolute energy difference between the current cycle energy and the previous cycle energy. If this difference is 5 times greater than the normal static average, a difference is counted. For this threshold value, a variable AMODD (mean absolute of odd harmonic current) was used. For basic beliefs, the following relationship is set by the technique itself and by experiments:

1. If the number of differences in the 30 cycle window is less than 2, the basic belief is (Normal, 0.8).
2. If the number of differences in the 30 cycle window is between 3 and 4, the basic belief is (Event, 0.8).
3. If the number of differences in the 30 cycle window is between 5 and 6, the basic belief is (Fault, 0.7).
4. If the number of differences in the 30 cycle window is more than 7, the basic belief is (Fault, 0.8).

For each test sample, the number of differences and the corresponding basic beliefs from the randomness technique using odd harmonics are shown in Table VII.

The purpose of this research and the example execution is not to compare the performance of each technique in a given situation. Therefore detailed performance variation sample by sample is not reviewed.

From these two sets of basic beliefs, a combined final belief is derived from a simplified formula. This formula assumes that each technique has a dominant confirming basic belief. For example, if a technique has the following basic beliefs,

$\{Fault\}(0.2)$, $\{Event\}(0.7)$, and $\{Normal\}(0.1)$,

then a confirming hypothesis for this technique is $\{Event\}(0.7)$.

Because there is no unconfirming basic belief, the simplified formula is simpler. If

p_1 is a supporting degree of fault,

p_2 is a supporting degree of event, and

p_3 is a supporting degree of normal, and if

$r_1 = 1 - p_1$,

$r_2 = 1 - p_2$, and

$r_3 = 1 - p_3$,

then, with two basic beliefs as in this example execution, p's are derived as

$$p_1 = 1 - (1 - BB1(fault))(1 - BB2(fault))$$

$$p_2 = 1 - (1 - BB1(event))(1 - BB2(event))$$

$$p_3 = 1 - (1 - BB1(normal))(1 - BB2(normal))$$

where BB indicates a basic belief.

The constant K is derived as follows.

$$K' = r_1 r_2 r_3 (1 + p_1/r_1 + p_2/r_2 + p_3/r_3)$$

$$K = 1/K'$$

Then, final beliefs of each status is derived as below.

$$Bel(fault) = K(p_1 r_2 r_3)$$

$$Bel(event) = K(p_2 r_1 r_3)$$

$$Bel(normal) = K(p_3 r_1 r_2)$$

From three different final beliefs, the only belief of interest is the belief of fault; the other two are non-fault cases. To give a message to operators

Table VII. Test Results.

TEST SAMPLE NO.	ACTUAL STATUS	TECH1 BB	TECH2 BB	FINAL BELIEF	DECISION TREE STATUS	EVENT DETECTOR OUTPUT	DECISION CONTROL SIGNAL
1	NF	(N,0.8)	(N,0.8)	NF	NF	0	N/A
2	NF	(N,0.8)	(N,0.8)	NF	NF	0	N/A
3	NF	(N,0.8)	(N,0.8)	NF	NF	0	N/A
4	NF	(N,0.8)	(N,0.8)	NF	NF	0	N/A
5	F	(F,0.6)	(F,0.8)	F	F	1	N/A
6	F	(F,0.8)	(F,0.8)	F	F	1	N/A
7	F	(F,0.8)	(F,0.8)	F	F	1	TRIP
8	F	(F,0.8)	(F,0.8)	F	F	1	
9	F	(F,0.8)	(F,0.8)	F	F	1	
10	F	(F,0.8)	(N,0.8)	NF	F	1	
11	F	(F,0.7)	(N,0.8)	NF	F	1	
12	F	(F,0.8)	(F,0.8)	F	F	1	
13	F	(F,0.8)	(F,0.8)	F	F	1	
14	F	(F,0.8)	(F,0.8)	F	F	1	
15	F	(F,0.8)	(F,0.8)	F	F	1	
16	F	(F,0.8)	(F,0.8)	F	F	1	
17	F	(F,0.8)	(N,0.8)	NF	F	1	
18	F	(F,0.8)	(E,0.8)	NF	F	1	
19	F	(F,0.8)	(F,0.8)	F	F	1	

or directly trip a relay, this numerated belief of fault should be delivered in terms of a definite status, that is, fault or not. To insure strict safety against false trips, about which most utility persons are most concerned, if the final combined belief of fault is greater than 0.9, that is, 90 percent sure that it is under a fault status, the definite output status is a 'fault.' This combination, final belief, and a definite output are obtained using a program which applied a simplified formula. This program is shown in Appendix C. In Table VII, these outputs are shown in the column of 'Final Belief.'

The status from the decision tree is also obtained. The variables passing through the branches are the standard deviation of the even harmonic current and mean absolute odd harmonic current. Of course, sub-harmonic current or even harmonic current can replace the odd harmonic current as shown in Rule 2, but in this example, odd harmonic current is arbitrarily chosen. The first branch is the standard deviation of even harmonic current and the second branch is the mean absolute value odd harmonic current. The output status is shown in the column 'Decision Tree Status' in Table VII. While the techniques cannot classify some normal-like faults, the decision tree classifies the testing samples perfectly because it has experience with similar faulted data at the same test site.

The output from an event detector was obtained using sub-harmonic currents. If 3 of 5 cycles have 1.75 times greater energy than the average, then an event flag '1' will be issued. This result is also shown in Table VII.

The combination or synthesis of all the results are summarized at each step. The data of samples 1 through 4 has been automatically classified and stored as non-fault data. These data already replaced the first 4 data of the non-fault data section of an existing memory. The data of sample 5 and 6 are pending unclassified. At the 7th sample, a "TRIP" signal is issued, and if a confirmation input from an operator arrives at the learning detection system, each data of sample 4 through 7 is classified as fault data. These data go to the memory and replace the first 3 existing fault data. A new rule will be induced with this new memory structure. A new rule is derived with the new memory structure. The new memory structure and a rule derivation is shown in the Section B(6) of Appendix B. The new rules say that:

Rule 1: If the 6th digit 1 (that is, if the standard deviation is higher than 0.609), then it is Fault.

Rule 2: If the 7th digit is 1 (that is, if the mean value of sub-harmonic currents is greater than 0.407), then it is Non-Fault.

Rule 3: If the second digit is 0 (that is, if the mean absolute of odd harmonic currents is less than 4.762), then it is Non-Fault.

The new rule is slightly different from the existent rule.

It is not appropriate to draw any general and convincing conclusions by the execution of a learning detection system on this particular set of test samples. However, it is found that there is no false trip even though both techniques miss some low-level faults. The decision tree from an induction process shows very good performance as a detector. Both reliable detection and security against false trip can be acquired by combining multiple techniques and a decision tree. By adding an event detector, which shows a perfect performance in indicating the occurrence of events, the learning detection system is getting experience and adapting itself to a variety of conditions.

D. Constraints for On-line, Real-time Operation

An intelligent decision making system performs quite well; however, it has some constraints when it is used as an on-line, real-time system.

"On-line" is a term which implies automated, autonomous detection of high impedance faults. In the sense of this detection, this means the intelligent decision making system is continuously in operation receiving parameters, running techniques, combining basic beliefs, and confirming and controlling the trip/notrip actions. To meet the on-line requirements of an intelligent system, all the processes are done continuously and automatically except one. During usual operation, the system does not need any operator interaction. However, when there is any discrepancy between the final status and the actual status, the system has to learn and update the induction process and thus the whole detection reasoning process.

Real-time systems are systems which are guaranteed to respond within a fixed period of time. The value of the fixed period of time can vary greatly, depending on the specific system in question[27]. For the operation of an intelligent decision making system in a real time base, the "real-time" refers a few seconds or more by the degree of the fault itself, which is considered in terms of the importance to the interests of the utility companies. The decision time is fixed to 30 cycles, that is, a half second. Considering the execution speed of today's microcomputer or minicomputer, this amount of time is long enough to handle all the necessary processes. One of the unsatisfactory cases is a decision tree update process. When it is necessary to update a decision tree, the whole induction process should begin. It involves a data-class replacement by correct actual class and corresponding data, a threshold value calculation, a rule derivation, and a technique selection update. This might takes less than a quarter second, however, providing appropriate data takes much time. Even though this will not hurt the detection itself, lose any data, or interrupt the on-going detection process, it might lose some normal data when a distribution feeder is connected again after a trip.

E. Summary

Several implementation methods are examined. A comparative implementation is chosen for the structure of the intelligent decision making system. An example execution is shown with staged fault sample data. The detailed test of decision tree output, final belief of techniques, and a learning detection system is examined. The intelligent decision making system works quite well; however, it has some constraints. Constraints of the system as an on-line, real-time system are also investigated.

CHAPTER IX

SUMMARY AND CONCLUSIONS

The objective of this research is to find an intelligent decision making system which combines multiple detection techniques. This intelligent system is to have three specific properties. First, it is to be capable of taking advantage of multiple techniques. Secondly, it is to make a decision under uncertainty by combining multiple basic beliefs from various techniques. Thirdly, it is to provide an adaptability using inductive reasoning realized in a learning detection system. The research work to achieve this goal is summarized and concluded in this chapter.

To understand the behavior of high impedance faults, statistical analysis was performed in wideband spectra. It was found that the behavior of various high impedance faults was not consistent. Faults, under similar conditions, showed very different characteristics in behavior so that predicting and classifying the behavior of faults and switching events were very difficult. The environmental parameters which affect the behavior of high impedance faults were investigated and suggested to find a best technique in a given condition. Existing detection techniques were investigated. The principles and weaknesses with respect to the environmental parameters were examined. With this examination, the basic structure of a technique selection database was formed.

Several decision making methods which can be applied to general decision making problems were discussed. To find a best technique for parameters which indicate the current outside condition effectively, the method of decision making under incomplete knowledge was applied. For the technique combination which should handle multiple pieces of information with uncertainty, a evidence theory and and a simplified belief combination formula were adopted. Inductive reasoning which minimize the entropy was applied as an induction process to acquire the knowledge of fault class and non-fault class and apply this knowledge to make an intelligent system adaptive to surrounding environments.

A learning detection system was implemented to fulfill a learning process by combining all the processes. An example execution was shown with an initial technique selection database and a decision tree which was derived using training sample data consisting of high impedance faults, switching events, and normal status. A complicated testing sample data was used to test the performance of the intelligent system. It was found that the intelligent decision making system made smart decisions even when it met very complicated situations. The constraints of this system in on-line, real-time operation was investigated. If an induction update takes much time, even though this will not hurt any detection process, some amount of normal data might be lost when a distribution feeder is connected again after a trip.

REFERENCES

[1] W. L. Beasley, "An Investigation of the Radial Signals Produced by Small Sparks on Power Lines," Ph.D. Dissertation, Texas A&M University, January 1970.

[2] H. J. Newton, *TIMESLAB:A Time Series Analysis Laboratory*. Pacific Grove, CA: Wardsworth &Brooks/Cole, 1987.

[3] Texas A&M University, "Detection of Arcing Faults on Distribution Feeders," *Research Report*, EPRI EL-2757, EPRI, December 1982.

[4] B. M. Aucoin, B. D. Russell, "Distribution High Impedance Fault Detection Utilizing High Frequency Current Component," *IEEE Transactions on Power Apparatus and Systems*, vol. 101, no. 6, pp. 1596 - 1606, June 1982.

[5] B. Don Russell, R. P. Chinchali, C. J. Kim, "Behavior of Low Frequency Spectra during Arcing Fault and Switching Events," *IEEE Transactions on Power Delivery*, vol. 3, no. 4, pp. 1485 - 1492, October 1988.

[6] C. J. Kim, B. Don Russell, "Harmonic Behavior during Arcing Faults on Power Distribution Feeders," *Electric Power System Research*, vol. 14, no. 3, pp. 219 - 225, June 1988.

[7] Mike Aucoin, "Status of High Impedance Fault Detection," *IEEE transactions on Power Apparatus and Systems*, vol. 104, no. 3, pp. 638 - 644, March 1985.

[8] B. Don Russell, K. Mehta, R. P. Chinchali, "An Arcing Fault Detection Technique using Low Frequency Current Components-Performance Evaluation using Recorded Field Data," *IEEE Transactions on Power Delivery*, vol. 3, no. 4, pp. 1493 - 1500, October 1988.

[9] C. L. Benner, P. W. Carswell, B. D. Russell, "Improved Algorithm for Detecting Arcing Faults using Random Fault Behavior," presented at the Southern Electric Industry Application Symposium, New Orleans, LA, November 15-16, 1988.

[10] Huges Aircraft Co., "High Impedance Fault Detection Using Third Harmonic Current," *Research Report*, EPRI EL-2430, EPRI, June 1982.

[11] D. I. Jeerings, J. R. Linders, "Unique Aspects of Distribution System Harmonics Due to High Impedance Ground Faults," presented at the *IEEE/PES Conference and Exposition on Transmission and Distribution*, New Orleans, LA, April 1989.

[12] Power Technologies, Inc., "Detection of High Impedance Faults," *Research Report*, EPRI EL-2413, EPRI, June 1982.

[13] C. L. Huang, H. Y. Chu, M. T. Chen, "Algorithm Comparison for High Impedance Fault Detection based on Staged Fault Tests," *IEEE Transactions on Power Delivery*, vol. 3, no. 4, pp. 1427 - 1435 , October 1988.

[14] H. Calhoun, M. T. Bishop, C. H. Eichler, "Development and Testing of an Electro-Mechanical Relay to detect Fallen Distribution Conductors," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-101, no. 6, pp. 1643 - 1650, June 1982.

[15] A. G. Phadke, He Hankun, "Detection of Broken Distribution Conductors," *Proceedings of IEEE Southeast Relay Conference*, Raleigh, NC, Paper No. CH2161-8/85/0000-0074, August 1985

[16] S. L. Tanimoto, *The Elements of Artificial Intelligence.* Rockville, MD: Computer Science Press, 1987.

[17] R. Forsyth, R. Rada, *Machine Learning: Applications in Expert Systems and Information Retrieval.* Chichester, UK: Ellis Horwood Limited, 1986.

[18] Z. W. Kmietowicz, A. D. Pearman, *Decision Theory and Incomplete Knowledge.* Hampshire,UK: Gower Publishing Company Limited, 1981.

[19] B. G. Buchanan, E. H. Shortliffe(eds.), *Rule-Based Expert Systems.* Reading, MA: Addison-Wesley, 1984.

[20] Paul R. Cohen, *Heuristic Reasoning about Uncertainty: An Artificial Intelligence Approach.* Marshfield, MA: Pitman Publishing Inc., 1985.

[21] J. Gordon, E. H. Shortliffe, "The Dempster-Shafer Theory of Evidence," in *Rule-Based Expert System.* B. G. Buchanan, E. H. Shortliffe, ED., Reading, MA: Addison-Wesley, 1984, chap. 13, pp. 272 - 292.

[22] G. Shafer, *The Mathematical Theory of Evidence.* Princeton, NJ: Princeton University Press, 1976.

[23] J. A. Barnett, "Computation methods for a Mathematical theory of evidence ," in *Proceedings of the International Joint Conference on Artificial Intelligence*, Vancouver, B.C., 1981.

[24] E. B. Hunt, J. Marin, P. J. Stone, *Experiments in Induction.* New York, NY: Academic Press, 1966.

[25] C. J. Kim, B. D. Russell, "Classification of Faults and Switching Events by Inductive Reasoning and Expert System Methodology," *IEEE Transactions on Power Delivery*, vol. 4, no. 3, pp. 1631 - 1637, July 1989.

[26] Ronald Christensen, *Entropy Minimax Sourcebook.* Lincoln, MA: Entropy Limited, 1980.

[27] D. Leinweber, J. Perry, "Controlling Real-Time Processes on the Space Station with Expert Systems," in *Proceedings of SPIE 729 - Space Station Automation II*, Cambridge, MA, October 1986.

SUPLEMENTARY SOURCES CONSULTED

[1] B. D. Russell, R. P. Chinchali, "A Digital Signal Processing Algorithm for Detecting Arcing Faults on Power Distribution Feeder," *IEEE Transactions on Power Delivery*, vol. 4, no. 1, pp. 132 - 140, January 1989.

[2] S. J. Balser, K. A. Clements, D. J. Lawrence, "A Microprocessor-Based Technique for Detection of High Impedance Faults," Presented at *IEEE Power Engineering Society* Winter Meeting, 86 WM 155-6,1986

[3] J. Carr, G. L. Hood, "High Impedance Fault Detection on Primary Distribution Systems," Canadian Electrical Association Final Report, Project No. 78-75, Montreal, 1979.

[4] Frank H. Knight, *Risk, Uncertainty, and Profit.* Chicago, IL: The University of Chicago Press, 1971.

[5] B. Don Russell, Karan Watson, "Power Substation Automation using a Knowledge Based System-Justification and Preliminary Field Experiments," *IEEE Transactions on Power delivery*, vol. 2, no. 4, pp. 1090 - 1097, October 1987.

[6] B. Don Russell, K. Watson, "Knowledge Base Expert Systems for Improved Power System Protection and Diagnostics," in *Symposium on Expert Systems Application to Power Systems*, Finland, August 1988.

[7] Donald A. Waterman, *A Guide to Expert System.* Reading, MA: Addison-Wesley, 1986.

[8] D. I. Jeerings, J. R. Linders, "Ground Resistance—Revisited," *IEEE Transactions on Power Delivery*, vol. 4, no. 2, pp. 949 - 956, April 1989.