University of Kentucky Master's Theses   Graduate School

2007

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION BY WAVELET TRANSFORMS AND ADAPTIVE RESONANCE THEORY NEURAL NETWORKS

Karthikeyan Kasinathan
University of Kentucky, kkasinathan@gmail.com

ABSTRACT OF THESIS

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION BY WAVELET
TRANSFORMS AND ADAPTIVE RESONANCE THEORY NEURAL NETWORKS

This thesis aims at detecting and classifying the power system transmission line faults. To deal with the problem of an extremely large data set with different fault situations, a three step optimized Neural Network approach has been proposed. The approach utilizes Discrete Wavelet Transform for detection and two different types of self-organized, unsupervised Adaptive Resonance Theory Neural Networks for classification. The fault scenarios are simulated using Alternate Transients Program and the performance of this highly improved scheme is compared with the existing techniques. The simulation results prove that the proposed technique handles large data more efficiently and time of operation is considerably less when compared to the existing methods.

Keywords: Adaptive Resonance Theory Neural Network, Wavelet Transform, Alternate Transient Program, Transmission line fault detection, fault type classification

_____
(Author's Signature)

_____
(Date)

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION USING WAVELET
TRANSFORMS AND ADAPTIVE RESONANCE THEORY NEURAL NETWORKS

By

Karthikeyan Kasinathan

_____

Director of Thesis

_____

Director of Graduate Studies

_____

Date

RULES FOR THE USE OF THESIS

Unpublished thesis submitted for the Master's degree and deposited in the University of Kentucky Library are as a rule open for inspection, but are to be used only with due regard to the rights of the authors. Bibliographical references may be noted, but quotations or summaries of parts may be published only with the permission of the author with the usual scholarly acknowledgements.

Extensive copying or publication of the thesis in whole or in part also requires the consent of the Dean of the Graduate School at the University of Kentucky.

A library that borrows this thesis for use by its patrons is expected to secure the signature of each user.

<u>Name</u>   <u>Date</u>

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

_____

THESIS

Karthikeyan Kasinathan

The Graduate School

University of Kentucky

2006

POWER SYSTEM FAULT DETECTION AND CLASSIFICATION USING WAVELET
TRANSFORMS AND ADAPTIVE RESONANCE THEORY NEURAL NETWORKS

_____

THESIS

_____

A thesis submitted in partial fulfillment of the
requirements for the degree of Masters of Science in
the College of Engineering
at the University of Kentucky

By
Karthikeyan Kasinathan
Lexington, Kentucky
Director: Dr. Yuan Liao, Assistant Professor of Electrical Engineering
Lexington, Kentucky
2006

Dedicated to my parents and friends

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude for this thesis, to the following people, as they are the people who made me contribute towards the completion of this thesis.

First, my thesis chair, Dr. Yuan Liao, I would like to thank him for being the strongest mainstay of this thesis. He gave me valuable guidance throughout this thesis. Association with him has been an enlightening experience not only in the field of research but also in terms of personal growth. I extend the opportunity to thank him for guiding me through all the phases during my thesis.

In addition to the technical assistance above, I received equally important assistance from family and friends. I desire to give my deepest thanks to my parents for being the strongest support and stand beside me in all the odds and still encouraging me. I would also like to appreciate my friends and roommates, who gave me strongest support and kept the confidence alive in me till the completion of this thesis.

Lastly, I would like to thank the entire department of Electrical Engineering, University of Kentucky for being so supportive and facilitating a healthy environment for the students.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FILES

Filename: Karthik_Thesis.pdf

Type   : PDF

Size   : 1.71 MB

Chapter One: Introduction

An important objective of all the power systems is to maintain a very high level of continuity of service, and when abnormal conditions occur, to minimize the outage times. It is practically impossible to avoid consequences of natural events, physical accidents, equipment failure or misoperation which results in the loss of power, voltage dips on the power system. Natural events can cause short circuits i.e. faults which can either be single phase to ground or phase to phase or phase to phase to ground or a three phase fault. Most faults in an electrical system occur with a network of overhead lines are single-phase to ground faults caused due to lightning induced transient high voltage and from falling trees. In the overhead lines, tree contact caused by wind is a major cause for faults. The appropriate percentages of occurrences various faults are listed below [1]:

Single line to ground fault – 70-80%
Line-Line to ground fault - 10-17%
Line-Line fault   – 8-10%
Three phase   – 2-3%

When faults occur in the power system, they usually provide significant changes in the system quantities like over-current, over or under-power, power factor, impedance, frequency and power or current direction. The most common and also the one used in this thesis is the over-current and so over-current protection is widely used.

Power system protection is the art of applying and setting up relays or fuses or both, to provide maximum sensitivity to faults and abnormal conditions and also to avoid false alarms during normal operating conditions. So it is desirable that a correct decision be made by the protective device as to whether the trouble is an abnormal condition or just a transient which the system can absorb and return to normal working condition. The protective relays are more of a preventive device which operates after a fault has occurred and it helps in minimizing the duration of trouble and limiting the damage, outage time and related

1

problems. For the system to operate properly, it is necessary to isolate the trouble area quickly with a minimum number of system disturbances. Both failure to operate and incorrect operation can result in major system upsets involving increased equipment damage, increased personnel hazards and possibly long interruption of service.

A typical logic representation of a relay is shown in the *figure 1.1* [1]. The components can be electromechanical, solid-state or both. The features and designs vary with different manufacturers. Solid state relays are more common, because they provide higher accuracy, reduced space, lower equipment and installation costs, give wider application and setting capabilities. They include control logic, data acquisition, event recording, fault location, remote sensing and self-monitoring and setting.
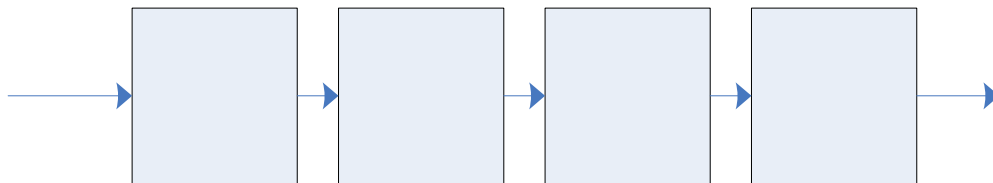
Figure 1.1: Logic representation of a relay

Protective relays are connected to the power system through the current or voltage transformers. These instrument transformers are used to provide insulation from the high power system voltages and hence reduce the magnitudes to practical system. The other end of the protection system is connected to a circuit breaker. Since the protective devices are low energy devices, they do not possess the energy to open and isolate the problem area of the power system. Thus protective devices and circuit breakers work together. The protective relay instructs the circuit breaker to isolate the fault. Thus both are necessary for prompt isolation of the trouble area or the damaged equipment. A diagram of the basic protection relay is shown in *figure 1 2*.
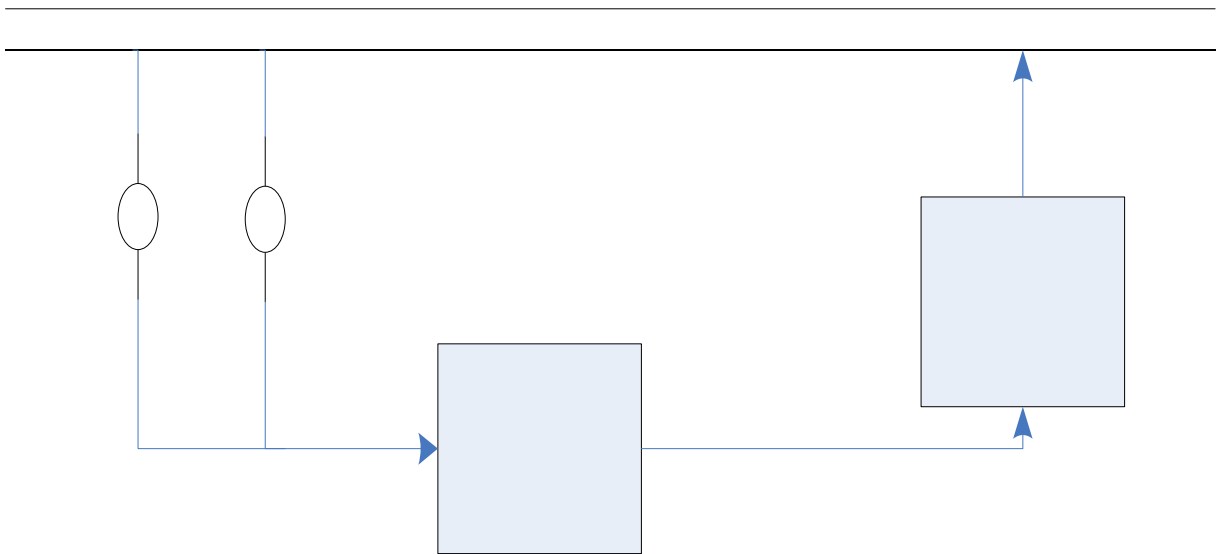
Figure 1.2: Basic Protection Relay system

Considering the fact that there could be a problem with the protection relays as well, a portion of the transmission line is usually protected by more than one relay. There is a primary relay which is assigned to operate at the first sign if trouble in their assigned protective zone. If they fail, they are backed up by various backup systems which operate to clear the trouble. The circuit breakers are the high power device which helps in physically isolating the faulted section of the power system. The fault detection and classification system which is discussed in this report is a part of the protective device. The five basic qualities of a protective relays [1] are

1. Reliability
2. Selectivity
3. Speed of operation
4. Simplicity
5. Economics

Various factors like availability of circuit breakers and fault indicators affect the protection system. The abnormalities, faults and intolerable conditions must provide a distinguishable difference between from the normal operating or tolerable conditions. In this thesis, voltage and current waveforms at both the sending and receiving end are used for detection and classification of the problem. Any significant changes in these quantities provide a means to detect abnormal conditions and hence employed for relay operation. So the important step is to first determine the quantity and the associated difference value that would separate the faulted and non-faulted waveforms. There could be two kinds of problems; a faulted waveform could be detected as a normal wave. This results in delivering the power exceeding the currents and frequency limits to the customers leading to equipment failures. The other error could arise if a normal waveform is detected as a faulted wave. In this case, a certain portion of the power system is unnecessarily isolated from the system.

In general, the first step in the power system relaying algorithms is the detection of fault and the next step is classification. This thesis uses a combination of Wavelet Transforms and 2 different Adaptive Resonance theory (ART) neural networks for detecting and classifying power system faults. The Objective of this thesis is to classify the faults according to the following parameters:

1. Fault type
2. Fault Location
3. Fault Resistance

The fault cases are classified as one of ten different types of faults viz. three Single line to ground faults, three Line to Line faults, three Line to Line to ground faults and a three phase fault. The fault location is an important parameter especially in high voltage power systems. The knowledge of fault location leads to high speed fault clearance as well as improved transient stability. For simulation purposes, a 200 mile transmission line is discussed in this thesis. Resistance is an important parameter for clearing the fault faster and to resume normal operation of the power system. The experimental setup and the simulation studies are explained in detail under Chapter 4.

The wavelets are developed as an alternative for Short time Fourier transform. The wavelet transforms is capable of providing the time and frequency information at the same time hence giving a time frequency representation of the signal. It is used to capture the dynamic characteristics of a non-stationary window using a short data window. The Current and the voltage waveforms are convoluted with the Daubechies mother wavelet to obtain the wavelet transform of the signal. The procedure is discussed in detail in Chapter 3.

The continuing increase in the size and complexity of power systems has led to the need for sophisticated computer based tools for solving the difficult problems that arise in the planning, operation, diagnosis and design of these systems. Neural Networks have been used in a lot of applications like pattern classification, pattern recognition, prediction and

automatic control [2]. Most commonly, neural networks have been used in the areas for fault classification/location, security assessment and load forecasting [3].

The function of a power system is to deliver the power requested by the customers, without exceeding voltage and frequency limits. This has to be accomplished in real time and in safe, economic manner. If the system is considered unsafe due to outages, then a control action has to be taken to restore it back to the normal working conditions. In general there are three operational states; the normal state is the one in which the customer demands are met and operating limit is within presented limits; alert state in which the demands are met and the operating limit is within the specified limits, but the system will become unstable with little disturbance; emergency state in which the limits are exceeded and the system becomes unstable [4]. This shows that the system has more than one dimension. To overcome this problem, a Multilayered Perceptron (MLP) with back propagation training algorithm neural network [5] is used. It helps is reducing the number of contingences, characterization of security boundaries, the dimension of the operating vector and also quantifies the operating point into a reduced number of classes. Since forecasting is usually done offline, neural networks with a lot of input factors like weather conditions, holidays, weekends etc. help in short term and long term forecasting. Owing to its flexibility to noisy data, neural networks are used in fault diagnosis [6]. The application of neural networks in this area is discussed in detail in the next two Chapters.

The basic Adaptive Resonance Theory (ART) system is an unsupervised learning model. It consists of a comparison field and a recognition field composed of neurons, a vigilance parameter, and a reset module. The vigilance parameter has considerable influence on the system: higher vigilance produces more categories, while lower vigilance results in more general memories i.e. fewer categories [7]. The comparison field takes an input vector and transfers it to its best match in the recognition field. Its best match is the single neuron whose set of weights i.e. the weight vector most closely matches the input vector. Each recognition field neuron outputs a negative signal which is proportional to that neuron's quality of match to each of the other recognition field neurons and inhibits their output accordingly. In this way the recognition field exhibits lateral inhibition, allowing each neuron in it to represent a

category to which input vectors are classified. After the input vector is classified, the reset module compares the strength of the recognition match to the vigilance parameter. If the vigilance threshold is met, training commences. Otherwise, if the match level does not meet the vigilance parameter, the firing recognition neuron is inhibited until a new input vector is applied; training commences only upon completion of a search procedure. In the search procedure, recognition neurons are disabled one by one by the reset function until the vigilance parameter is satisfied by a recognition match. If no committed recognition neuron's match meets the vigilance threshold, then an uncommitted neuron is committed and adjusted towards matching the input vector.

The Wavelet Transform and the Adaptive Resonance theory training will be discussed in detail in Chapter 3.

Chapter Two: Literature survey

The problem of detecting and classifying faults in a transmission line has been going for a very long time. It has been one of the major concerns of the power industry. Normally, protective relays, recording devices and special control and protection software systems are responsible for detecting the fault occurrences and isolating the faulted portion from the system. Thus it is necessary for the faults to be detected quickly and precisely. It is also equally important to know the details about the fault that has occurred so that it can be corrected soon.

2.1 Power system fault detection techniques:

The Discrete Fourier Transform of the signal gives the frequency components that exist in the signal. But since power system fault waveforms are non-stationary by nature and since we are also interested in determining the place where the fault has occurred, Short Time Fourier Transform (STFT) was used for detecting the power system transmission line faults. In STFT, the signal is divided into small enough segments, where these segments (portions) of the signal can be assumed to be stationary [8]. For this purpose, a window function "w" is chosen. The width of this window must be equal to the segment of the signal where its stationarity is valid.

This window function is first located to the very beginning of the signal. That is, the window function is located at t=0. Let's suppose that the width of the window is "T" seconds. At this time instant (t=0), the window function will overlap with the first T/2 seconds. The window function and the signal are then multiplied. By doing this, only the first T/2 seconds of the signal is being chosen, with the appropriate weighting of the window. Then this product is assumed to be just another signal, whose FT is to be taken. In other words, FT of this product is taken, just as taking the FT of any signal.

The result of this transformation is the FT of the first T/2 seconds of the signal. If this portion of the signal is stationary, as it is assumed, then there will be no problem and the obtained

result will be a true frequency representation of the first T/2 seconds of the signal. The next step would be shifting this window to a new location, multiplying with the signal, and taking the FT of the product. This procedure is followed; until the end of the signal is reached by shifting the window with "t1" seconds intervals.

Thus Short Time Fourier Transform of the signal can be defined as

$$STFT_{x(t,\ f)} = \int [x(t)*w*(t-t')]*e^{-i2\pi ft} dt \qquad (1)$$

Where $x(t)$ is the signal, $w(t)$ is the window function and $w*$ is the complex conjugate. Thus we have a time frequency representation of the signal. So it gives the frequency response as well the occurrence of that frequency in time. It is computationally fast and detects faulted waveforms effectively. But this technique depends on the selection of a good window. This is called the Resolution problem, in other words choosing the appropriate size of the window plays a vital part in the fault detection process. Narrower windows have good time resolution but poor frequency resolution, but on the other hand broader windows have good frequency resolution but poor time resolution [8]. In the next Chapter, we will see how the wavelet transform overcomes the resolution problem.

2.2 Power system fault classification techniques:

Different techniques have been used for power system fault classification in the past. In the next few sections, a few of these approaches are discussed.

2.2a Back propagation neural network:

Various applications of Neural Networks have been used in the past to improve the protection scheme in the transmission lines [8]. They have been used in fault classification, fault section estimation, adaptive relaying and fault diagnosis. Many of these methods are based on back propagation, Radial basis function and Finite Impulse response neural networks. A few of these approaches are discussed here.

A typical Back Propagation Neural Network is a non-linear regression technique which attempts to minimize the global error. Its training includes both forward and backward propagation, with the desired output used to generate the error values for back propagation to iteratively improve the output. The back propagation neural network must have at least one input layer and one output layer. The hidden layers are optional.

A Typical back propagation neural network is shown in the *Figure 2.1* [10]:
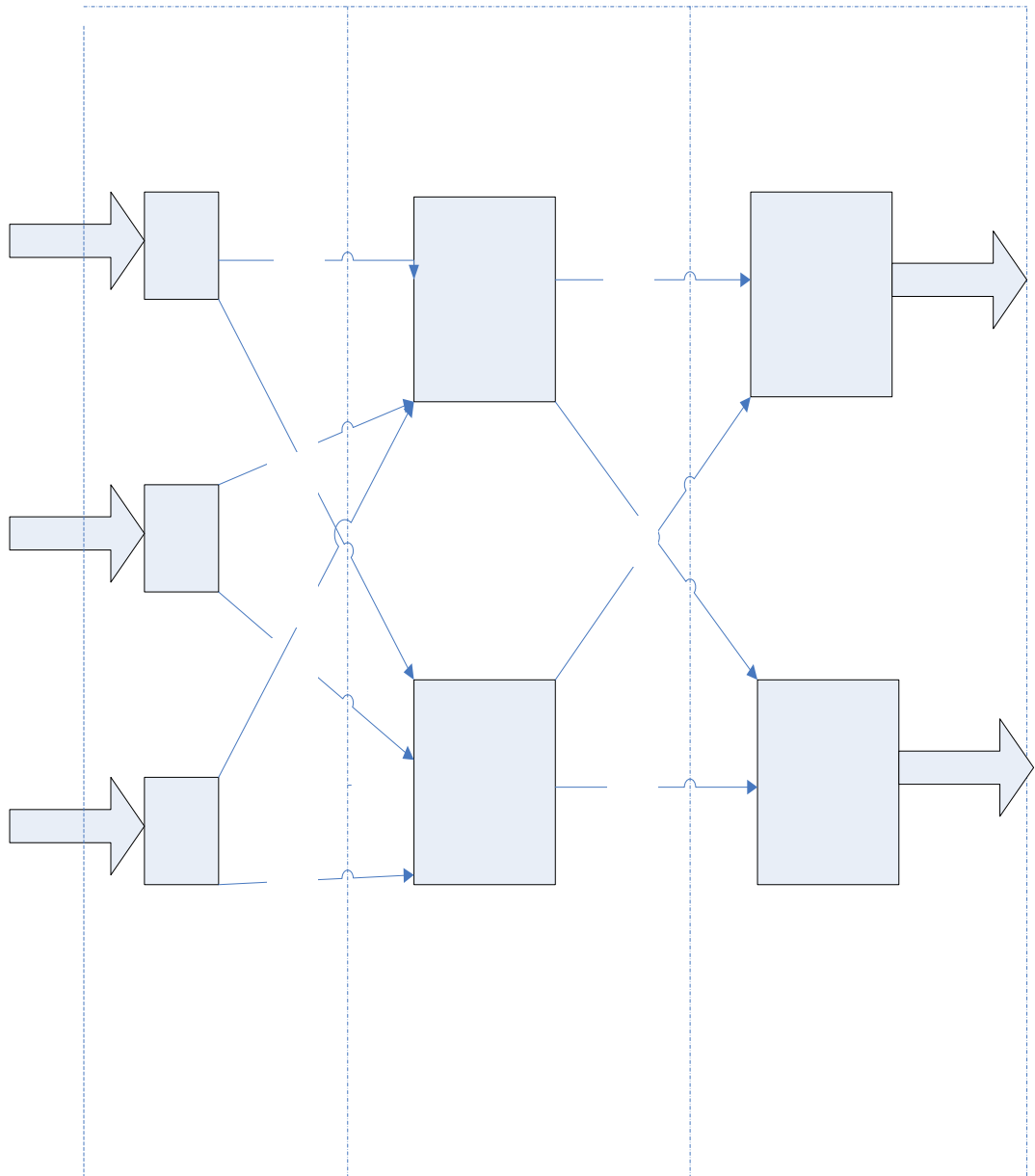
Figure 2.1: Back propagation neural network

x1

N

The Back propagation neural network consists of three layers: an input layer with three neurons, a hidden layer with two neurons and an output layer with two neurons. In the figure, we see that the output of a neuron in a layer goes to all neurons in the following layer and each neuron has its own weights. Initially the weights of the input layer are assumed to be 1 for each input. The output of the back propagation neural network is reached by applying input values to the input layer, passing the output of each neuron to the following layer as input.

The number of neurons in the input layer depends on the number of possible inputs we have, while the number of neurons in the output layer depends on the number of desired outputs. The number of hidden layers and how many neurons in each hidden layer cannot be well defined in advance, and could change per network configuration and type of data. In general the addition of a hidden layer could allow the network to learn more complex patterns, but at the same time decreases its performance [10]. Ideally, we could start a network configuration using a single hidden layer, and add more hidden layers if we notice that the network is not learning as well as we like.

The Back propagation training algorithm [11] could be summarized as follows: The Input data sample is first presented to the network and then the network's output taken from the output layer is compared with the desired output and the error is calculated in each output neuron. And now for each neuron, a scaling factor called the local error is calculated which indicates how much higher or lower the output must be adjusted to match the desired output. The weights are modified to lower this local error. This process gets repeated until the error falls within the acceptable value (pre-defined threshold) which would indicate that the neural network has been trained successfully. On the other side, if the maximum number of iterations is reached, then it indicates that the training was not successful.

2.2b Fuzzy based classification:

Fuzzy logic (FL) can be defined as a problem-solving control system methodology that lends itself to implementation in systems ranging from simple, small, embedded micro-controllers to large, networked, multi-channel PC or workstation-based data acquisition and control systems. Fuzzy based classification technique employs a simple, rule-based IF X AND Y THEN Z approach to a solving control problem rather than attempting to model a system mathematically.

In the case of Neural Networks, considerable amount of training time under different operating conditions is required for good performance of the system. The fuzzy logic based fault classification scheme does not require training and so it is computationally much faster when compared to fault classification by Artificial Neural Network (ANN) methods. This fuzzy logic based scheme is capable of accurately predicting the exact type of fault under wide range of operating conditions [12].

The Fault Classification is based on Angular differences among the sequence components of the fundamental during fault current as well as on their relative magnitudes. The phasor diagram of a phase "a" to ground fault is shown in the *Figure 2.2*. The zero, positive and negative sequence components of the post fault currents relative to phase "a" are denoted as $I_{aof,}$ I_{a1f} and I_{a2f} respectively. The angles between the positive and negative sequence components of phase a, b and c are given as

$$\text{arg\_} A = \mid Arg(I_{a1f}) - Arg(I_{a2f}) \mid = 0^0$$
$$\text{arg\_} B = \mid Arg(I_{b1f}) - Arg(I_{b2f}) \mid = 120^0 \qquad\qquad (2)$$
$$\text{arg\_} C = \mid Arg(I_{c1f}) - Arg(I_{c2f}) \mid = 120^0$$

The magnitudes of I_{aof}, I_{a1f} and I_{a2f} are related by

$$R_{of} = | I_{aof}/ I_{a1f}| = 1 \text{ and } R_{2f} = | I_{a2f}/ I_{a1f}| = 1 \tag{3}$$

Similarly, the magnitudes and angle between the positive and negative sequence components are obtained for other types of asymmetric faults.

For every type of fault, there exists a unique set of these five parameters. So it is possible to formulate simple logic base for determining the fault type from the values of the five inputs. The different inputs are represented by a corresponding fuzzy variable. Now a fuzzy rule was developed using these five variables to detect the type of fault.
For example:

If arg_A is "approximately 30^{0}" and arg_B is "approximately 150^{0}" and arg_C is "approximately 150^{0}" and R_{of} is "high" and R_{sf} is "high" then fault type is "a-g"

In this method, only 3 parameters are sufficient and it identifies 10 types of short-circuit faults accurately. But the main disadvantage with this method is that it is applicable to only asymmetric faults and it is not very effective if you are looking to classify not just by the type of fault.
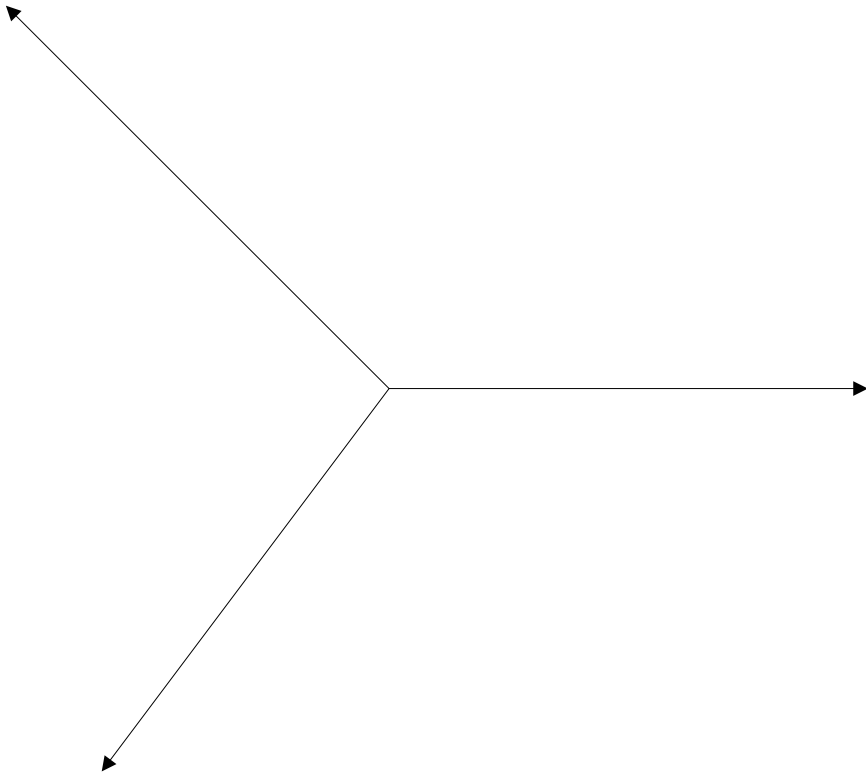
Figure 2.2: Phasor diagram for a-g fault

$$I_{b2f} = a\text{...}$$
$$I_{c1f} = a\text{...}$$

Chapter Three: The Proposed Approach for Fault Detection and Classification

3.1 Objective of the system:

The Objective of the system is to detect and classify fault cases based on three parameters; fault type, fault location and fault resistance. The fault is detected using the coefficients of the discrete wavelet transform. In the classification stage, the fault cases are classified as one of either 20 or 60 clusters obtained during training. To achieve this, two different experiments are conducted separately. The 20 clusters are obtained from 10 different types of faults (3 single line to ground, 3 line to line, 3 line to line to ground and 1 three phase fault) and 2 fault locations (0-90% of the length of the transmission line and 90-100% of the length of the transmission line from the sending end). The 60 clusters are obtained from 10 different types of faults, 2 fault locations and 3 fault resistances (10, 50 and 100 Ohms). The aim of this thesis is to classify a fault case into one of either 60 or 20 clusters in separate tests. The proposed approach is explained in the next few sections.

3.2 Overview of the system:

In this Chapter, the proposed optimized solution to the power system fault classification problem is discussed. A few of the Neural Network techniques discussed in the last section were not suitable for real-time applications since the testing process is too slow. The new concept proposed in this thesis is based on Wavelet Transform and Adaptive Resonance Theory Neural Networks. The Wavelet Transform serves to isolate the Faulted and Non-Faulted waveforms before training and the adaptive resonance theory neural networks are used for pattern classification. This optimized training algorithm targets the speed of the training process and improves it considerably. The first step in the training process is to isolate the faulted and non-faulted waveforms. The technique is based on the absolute sum of coefficients in multi resolution signal decomposition based on discrete wavelet transform. A fault criterion is defined based on a number of experiments with different types of faults.

After isolation of non-faulted waveforms, the faulted waveforms are sent to the neural networks for fault classification. The proposed approach uses two different neural networks for classification. The first neural network uses the sending end voltage waveforms for training and testing. The second neural network separates further refines the classification by using the sending and receiving end current waveforms. It was seen from the simulations that at any instant after the occurrence of the fault, the current waveforms at the sending and receiving end will be approximately in phase for non grounded faults. It was seen that for some cases, as shown in the *figure 3.1*, the two-phase grounded and the non-grounded (aerial) faults are classified as the same type of fault. This is also illustrated in [13]. The proposed approach uses a fuzzy based classification in addition to the two neural networks to make sure that the fault type is been classified accurately. The figure 3.1 shows the sending end voltage waveforms for grounded and non-grounded fault. It has been seen after repeated tests, that the absolute sum of the modulus of all the 3 phase currents for both sending end and receiving end can overcome this problem. The absolute sum of 3 phase currents for aerial of non-grounded faults are seen to be in the order of the magnitudes of 10 exp -7 and for non-grounded or aerial fault it is more than 1. So the absolute sum of current waveforms at sending end and receiving end are calculated and this approach simply employs a threshold detector which would detect a magnitude of more than 10 exp -6. If the magnitude is less than 10 exp -6, then it is an aerial fault and if is not then it is classified as a grounded fault. This procedure is repeated for every test case, after the two neural networks to make sure that the type of fault; especially for double-line faults classified is reliable. A three step method for fault classification has been discussed in [13]. But it employs three different neural networks instead of two being used in the proposed technique and so the training and testing time is considerably slower when compared to this approach which has been demonstrated in the results section of Chapter 4.

The proposed technique is shown in *figure 3.3*. The first block describes obtaining the three-phase current and voltage waveforms from the Alternate Transients Program (ATP) software [14]. Different fault scenarios are simulated by varying the fault type, fault distance and resistance. A total of 3230 fault scenarios have been generated to train the proposed neural network and another 3230 different fault cases have been generated to test the performance

of the two neural networks. To demonstrate the working of this block, faulted as well as non-faulted waveforms are given as the input. The next block differentiates the faulted and non-faulted waveforms. This is done with the help of discrete wavelet transform. The procedure is explained in detail in the next section. It performs the operation of isolating the faulted and non-faulted waveforms. So, the output of this block to the neural networks will only include the faulted waveforms.

The faulted waveforms obtained from the wavelet transform are used for training the neural networks. In the first neural network, the sending end voltage waveforms are used for training. It was found that the first neural network by just employing the sending end voltages does not accurately classify different fault scenarios based on fault type, resistance and distance. Since, the proposed approach is targeted at not only classifying the type of fault but also the location as well as the resistance, it is required to have more number of clusters and also a well defined boundary separating them for improving the accuracy of the cases which fall close to the boundaries. So it is required to have more knowledge of the fault case. For this purpose, a second neural network with the current waveforms is used.

The training matrix for the first neural network is given as

$$A = \begin{bmatrix} Vsa_1, Vsa_2, \ldots, Vsa_N \\ Vsb_1, Vsb_2, \ldots, Vsb_N \\ Vsc_1, Vsc_2, \ldots, Vsc_N \end{bmatrix} \tag{4}$$

Where

$$Vsa1 = \begin{bmatrix} Vsa(1) \\ Vsa(2) \\ Vsa(3) \\ . \\ . \\ . \\ Vsa(n) \end{bmatrix} \tag{5}$$

Va₁ is the sending end voltage of phase 'A' for fault case 1. N represents the total number of fault cases being used for training and n is the total number of samples for each fault case.
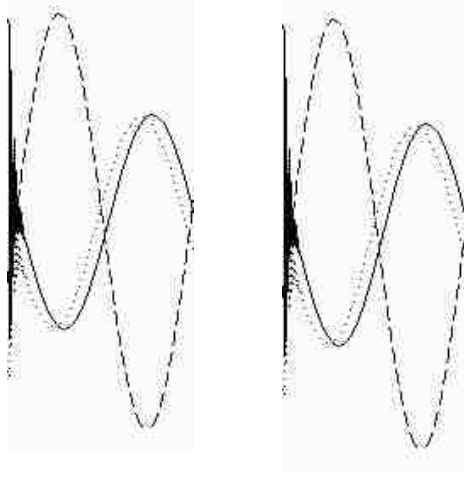

Figure 3.1: Sending end voltage waveform for grounded and non-grounded condition

The second neural network helps in learning more information about the fault case which leads to the formation of more number of clusters. All the faulted waveforms from the discrete wavelet transform block are used for training and testing this neural network. The neural network uses the sending and receiving end current waveforms as its input. The sending end and receiving end current waveforms are given as input in the matrix form according to the equation (6).

The training matrix for the second neural network is given by

$$B = \begin{bmatrix} Isa_1, Isa_2, \ldots, Isa_N \\ Isb_1, Isb_2, \ldots, Isb_N \\ Isc_1, Isc_2, \ldots, Isc_N \\ Ira_1, Ira_2, \ldots, Ira_N \\ Irb_1, Irb_2, \ldots, Irb_N \\ Irc_1, Irc_2, \ldots, Irc_N \end{bmatrix} \qquad (6)$$

Where Isa$_1$ and Ira$_1$ are the 'n' sized column vectors of the sending end and receiving end current waveforms of phase 'A' for fault case '1' and is given by

$$\text{Isa}_1 = \begin{bmatrix} Isa_1(1) \\ Isa_1(2) \\ Isa_1(3) \\ . \\ . \\ Isa_1(n) \end{bmatrix} \qquad \text{Ira}_1 = \begin{bmatrix} Ira_1(1) \\ Ira_1(2) \\ Ira_1(3) \\ . \\ . \\ Ira_1(n) \end{bmatrix} \qquad (7)$$

Where n is the total number of samples for each fault case. The *figure 3.2* shows the 1-cycle sending end voltage waveform generated by Alternate Transient Program (ATP) software for double line to ground and an aerial double line fault case.

For classifying fault distance and resistance, we need more parameters as inputs for the neural network because proper classification of faults across the boundary requires many small finely grained clusters. It was found that a single neural network with voltage and current waveforms as input takes more time for training and testing when compared to the proposed approach which uses two different neural networks. The neural network training times and accuracy of the approach is tabulated inside the results section in Chapter 4.This necessitates the usage of two neural networks for training and testing. Since this approach uses only one cycle of waveform after the occurrence of fault, it is computationally less intensive and so it takes less time for training and testing when compared to the existing approaches which is validated in Chapter 4. The outputs from the two neural networks and the fuzzy rule for fault type separation are used for classifying the fault case. This output combination gives the category which is uniquely identified for a given fault type, distance and resistance. For this thesis, two tests are performed using Matlab. In one test, a total of 60 clusters are formed with 10 different fault types, 2 fault location zones; zone1 is from 0-90% of the length of the transmission line and zone2 is from 90-100% of the length of the transmission line and 3 fault resistances (10 Ohms, 50 Ohms and 100 Ohms). The second test is performed after training 20 clusters formed with 10 different fault types and 2 fault location zones and not including the fault resistances. The results are tabulated in Chapter 4.
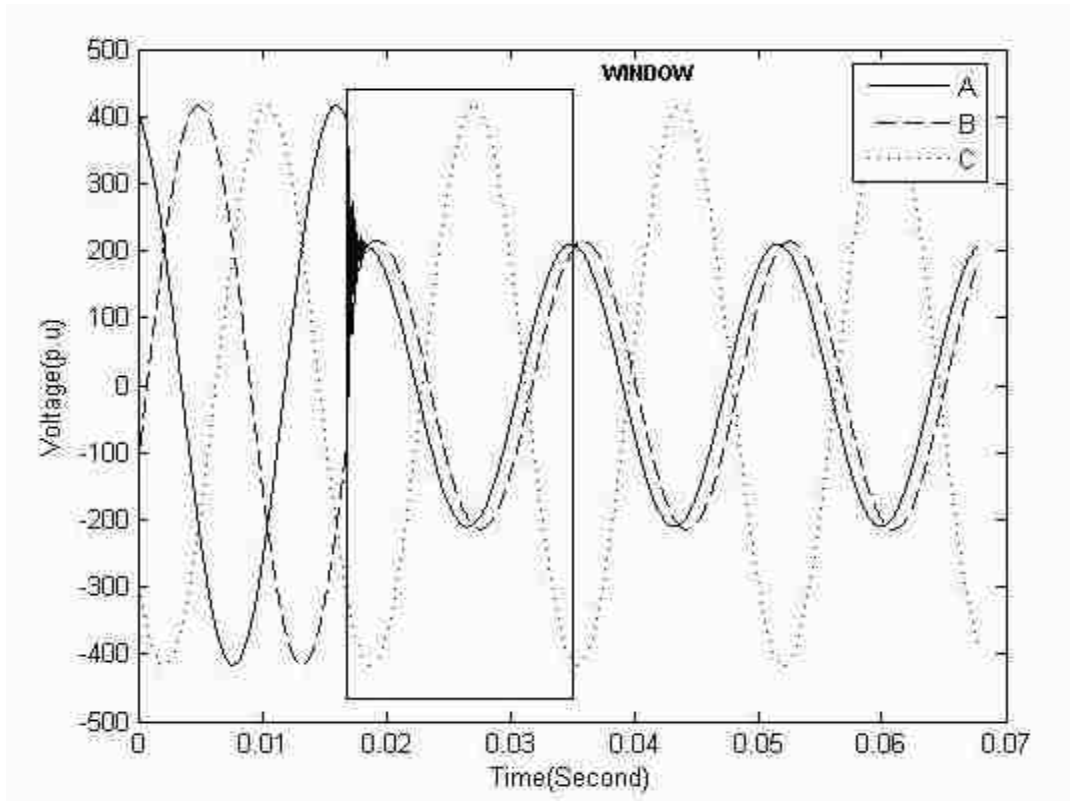
Figure 3.2: Waveform window used for training and testing

To summarize, the approach uses a discrete wavelet transform and two neural networks with different inputs for classifying faults. This optimized approach is tested with Adaptive Resonance Theory 1 (ART1) and Fuzzy Adaptive Resonance Theory (ART) Neural Networks which is explained in the later sections of this Chapter.
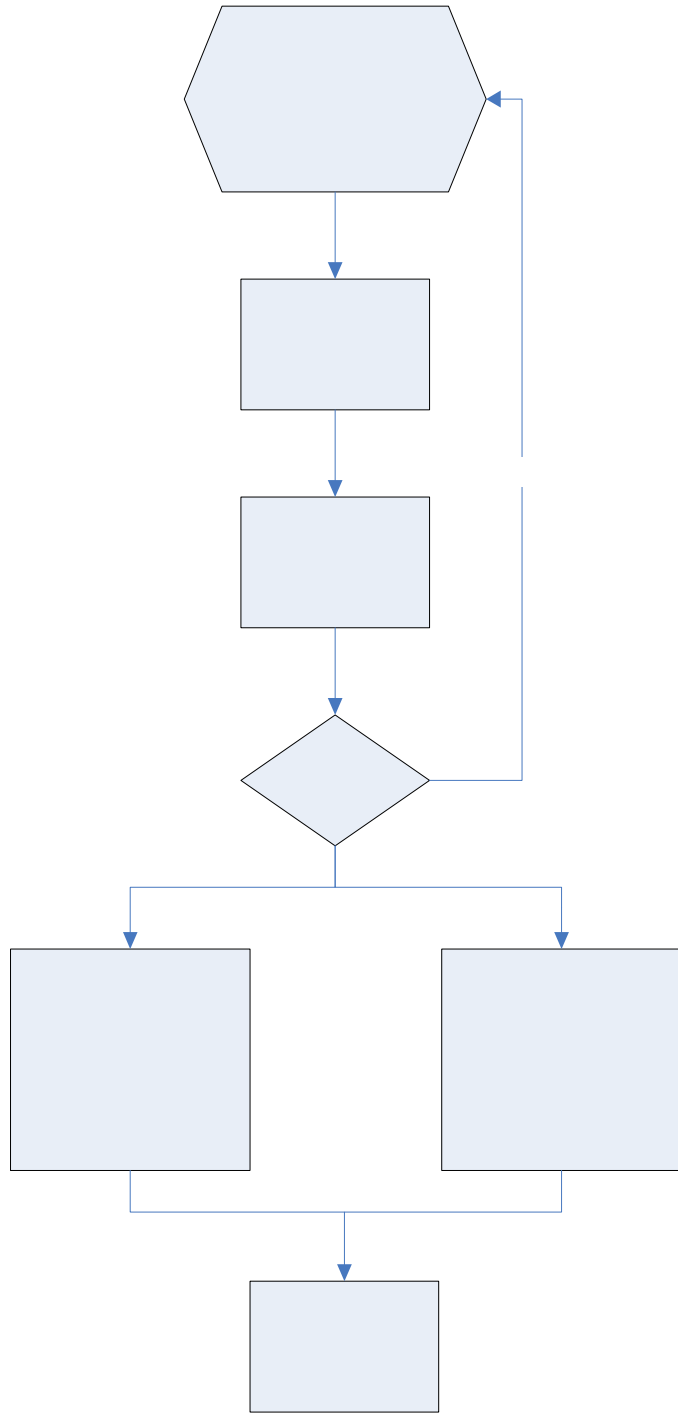
21

Figure3.3: Proposed fault detection and classification technique

3.3 Discrete wavelet transform for fault detection:

The First step in the Fault Classification is Fault detection. It is extremely important to identify those that provide the discriminative information and also identify those that may be redundant. In this project, it's the separation of faulted and non-faulted waveforms. By removing the non-faulted waveforms, the computational complexity becomes less. The faulted waveforms have to be separated from the non-faulted waveforms and the time instant at which the fault has occurred needs to be determined. The faulted wave is going to be non-stationary in nature and since we are also interested in where the spectral components exist in the signal, Fourier transform though being widely used is not suited for detecting them. Wavelet transform also overcomes the resolution problem faced when the Short Time Fourier transform (STFT) is employed [15].

Wavelet Transform is capable to giving the time and frequency information simultaneously and hence giving a time-frequency representation of the signal. The basic idea in time-frequency representations is that two parameters are needed: the frequency 'a', the position in the signal 'b' [16]. A general time-frequency waveform of a signal x will take the form:

$$x(t) \rightarrow \psi(a,b) = \int_{-\infty}^{\infty} \overline{\psi_{ab}} . x(t) \, dt \qquad (8)$$

Where ψ_{ab} is the analyzing function and $\overline{\psi_{ab}}$ is its complex conjugate.
The Discrete Wavelet transform is defined as

$$DWT(m,k) = 1/\sqrt{a_o^m} \sum_{n} x(n) \, g((k - n b_0 a_0^m)/a_0^m) \qquad (9)$$

Where $g(n)$ is the mother wavelet and $x(n)$ is the signal, the scaling and the translation parameters "a" and "b" are functions of m. the scaling gives DWT a logarithmic frequency coverage in contrast to the uniform coverage obtained in the case of a short time Fourier transform (STFT) [17]. To facilitate fault detection using wavelet transform, the mother

wavelet is first selected. This step is very important in the fault detection process because the selection of the mother wavelet is essential for accurate classification of the faulted and the non-faulted phase. Thereby the mother wavelet which would have a significant magnitude of d1 coefficient [18] and also that creates a significant difference between the faulted phase and the non-faulted phase needs to be selected. The mother wavelets considered is db4 (Daubechies), bior (biorthogonal), sym (symlets) and coif (coiflets).

The mother wavelet for this thesis was selected based on repeated iterations. Various experiments were conducted with all the four mother wavelets under different fault conditions. It was determined that the difference between the faulted and the healthy phase was substantial when db4 (Daubechies) was selected as the mother wavelet.

The *figure 3.2* shows the algorithm for detecting the transmission line faults. The discrete wavelet transform for the input waveform is calculated and sum_d1 is the absolute value of the sum of the detailed coefficients for 1-cycle [19]. This absolute sum FC is the fault criterion which is used to detect the faults and D is the duration for which the sum value should be more than the fault criterion for it to be classified as a fault. The whole process is based on a moving window approach where the 1-cycle window is moved continuously by one sample. In short, a fault is detected when the absolute sum should be more than the fault criterion FC continuously for D samples. The fault criterion FC and duration D are again calculated based on repeated experiments with different fault conditions. The value for FC and D were determined as 0.085 and 56 respectively.
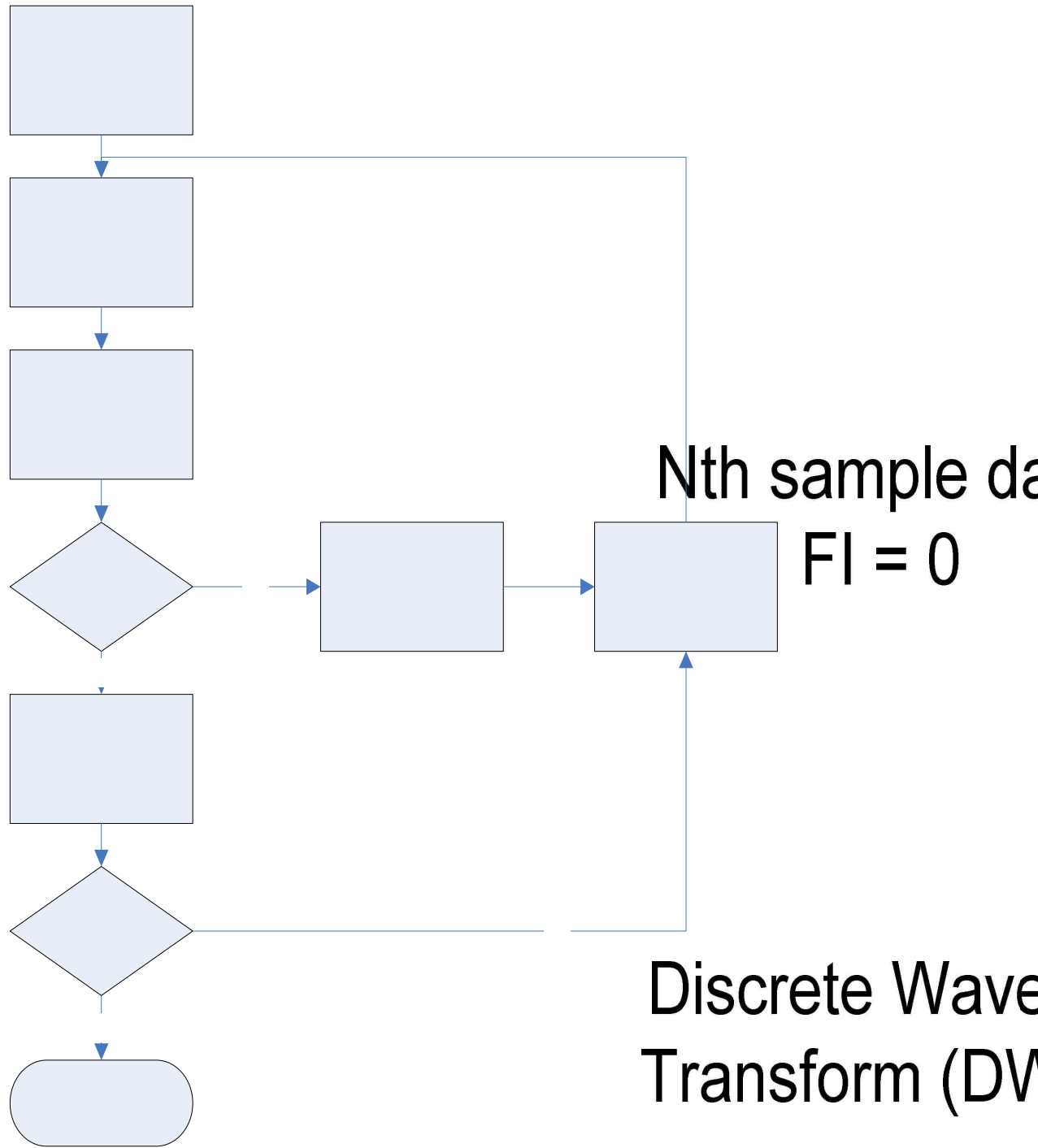
Figure 3.4: Fault detection using Discrete wavelet transform

Nth sample da
FI = 0

Discrete Wave
Transform (DW

Sum_d1 = Σ
abs(d1(k)) k =

3.4 Overview of Adaptive Resonance theory neural network:

In this Chapter, we will discuss about the Adaptive Resonance Theory neural network that is being used in this project. It comes under the category of Self-organizing networks, they learn neither by the familiar correlation rule nor by the perceptron technique. Actually they need no feedback from the outside environment during training. The network must discover for itself any relationships of interest that may exist in the input data. They will be able to translate the discovered relationships into outputs. Networks trained this way usually learn by matching certain explicit familiarity criteria. These networks can produce output that tells us how familiar it is with the present pattern. This is accomplished by comparing the present pattern with typical patterns seen in the past [20]. One important measure of similarity used for learning studied in this Chapter is the maximum value of scalar product of the weights and the input vector. Using the scalar product metric, the weight can be trained in an unsupervised mode to resemble frequent patterns, or better, pattern clusters. Another often used measure of similarity is the topological neighborhood, or distance, between the responding neurons arranged in regular geometric arrays.

Networks trained this way will not only react to values of inputs but also to their statistical parameters [21]. At the beginning of the learning the unsupervised training process, the network's response maybe implausible. The network learns gradually by itself which features need to be considered in classification without the presence of a feedback from the outside environment. The proximity measures allow the computation of important indicators of how to differentiate between more and less important features during the unsupervised training. Unsupervised training can only be implemented with redundant input data. Redundancy provides knowledge about statistical parameters of input patterns [22]. We will see in detail about the Adaptive resonance theory neural network.

3.5 Adaptive Resonance Theory (ART) neural network 1:

The Adaptive Resonance theory neural network 1(ART1) was developed by Carpenter and Grossberg. It serves the purpose of cluster discovery. The novel property of ART1 network is the controlled discovery of clusters. In addition, the ART1 neural network can accommodate new clusters without affecting the storage or recall capabilities for clusters already learned.

The network produces clusters by itself, if such clusters are identified in input data, and stores the clustering information about patterns or features without a priori information about the possible number and type of clusters. It originates the first cluster with the first input pattern received. It then creates the second cluster if the distance (Euclidean distance) of the second pattern exceeds a certain threshold; otherwise the pattern is clustered with the first cluster. This process of pattern inspection followed by either new cluster origination or acceptance of the pattern to the old cluster is the main step of ART1 Neural network production [23]. The training is summarized as follows:

1. The vigilance parameter ρ is set, and for the n-tuple input vectors and M top-layer neurons the weights are initialized. The matrices W,V are (M x n) and each is initialized with identical entries:

$$W = [1/(1+n)] \tag{10}$$
$$V = [1] \tag{11}$$
$$0 < \rho < 1$$

2. Binary unipolar input vector x is presented at input nodes, $x_i = 0,1$ for $i = 1,2,\ldots n$

3. All matching scores are computed from

$$Y_m^0 = \sum_{i=1}^{n} w_{im} x_i, \quad \text{for } m = 1, 2,\ldots m \tag{12}$$

In this step, selection of the best matching existing cluster, j is performed according to the maximum criteria as follows

$$Y_j^0 = \max \left(_{m=1,2\ldots M}(Y_m^0) \right) \tag{13}$$

4. The similarity test for the winning neuron j is performed as follows:

$$1/ \left(\|x\| \right) \sum_{i=1}^{n} v_{ij} x_i > \rho \tag{14}$$

Where ρ is the vigilance parameter and the norm $\|x\|$ is defined for the purpose of this algorithm as follows:

$$\|x\| = \sum_{i=1}^{n} |x_i| \tag{15}$$

if this test has failed, the algorithm goes to step-6 only if the top layer has more than a single active node left. Otherwise, the algorithm goes to step-5.

5. Entries of the weight matrices are updated for index j passing out of step-4. The updates are only for the entries (i,j). where $I = 1,2,\ldots M$ and are computed as follows:

$$W_{ij}(t+1) = v_{ij}(t)x_i / \left(0.5 + \sum_{i=1}^{n} v_{ij}(t)x_i \right) \tag{16}$$

$$v_{ij}(t+1) = x_i \, v_{ij}(t) \tag{17}$$

This updates the weights of the 'j'th cluster. The algorithm returns to step-2.

6. The node j is deactivated by setting $y_j = 0$. Thus this mode does not participate in the current cluster search. The algorithm goes back to step-3 and it will attempt to establish a new cluster different than j for the pattern under test.

28

The *Figure 3.4* shows a typical ART search cycle [24]. As shown the input vector "I" registers itself as a pattern X of activity across level F1. The F1 output vector S is then transmitted to the level F2. This transmission event multiplies the vector S by a matrix of adaptive weights or long term memory (LTM) traces to generate a net input vector T to level F2. The vector T is further enhanced and a compressed activity Y is generated across F2. The node that receives the maximal F1->F2 input is selected. At the end only one component of Y is non-zero after this choice is made. Activation of winner-take-all node defines the category of the input pattern I.

Figure 3.4 (a) shows the input pattern I generating the specific short term memory (STM) activity pattern X at F1. Pattern A inhibits A and generates the output signal pattern S. Signal pattern S is transformed into the input pattern T, which activates the Short term memory pattern Y across F2.

Figure 3.4 (b) indicates the pattern Y generating the top down signal pattern U, which is transformed into the prototype pattern V. If V mismatches I at F1, then a new Short term memory pattern X* is generated at F1. The reduction in total Short term memory activity occurs when X is transformed into X* causes a decrease in the total inhibition from F1 to A.

If the mismatching criterion fails to be met, A releases a non specific arousal wave to F2, which resets the STM pattern Y at F2. This is shown in *Figure 3.4 (c)*. After Y is inhibited, its top down prototype signal is eliminated, and X can be reinstated at F1. Enduring traces of the prior reset lead X to activate a different STM pattern Y* at F2. If the top-down prototype due to Y* also mismatches I at F1, then the search for the appropriate F2 code continues.
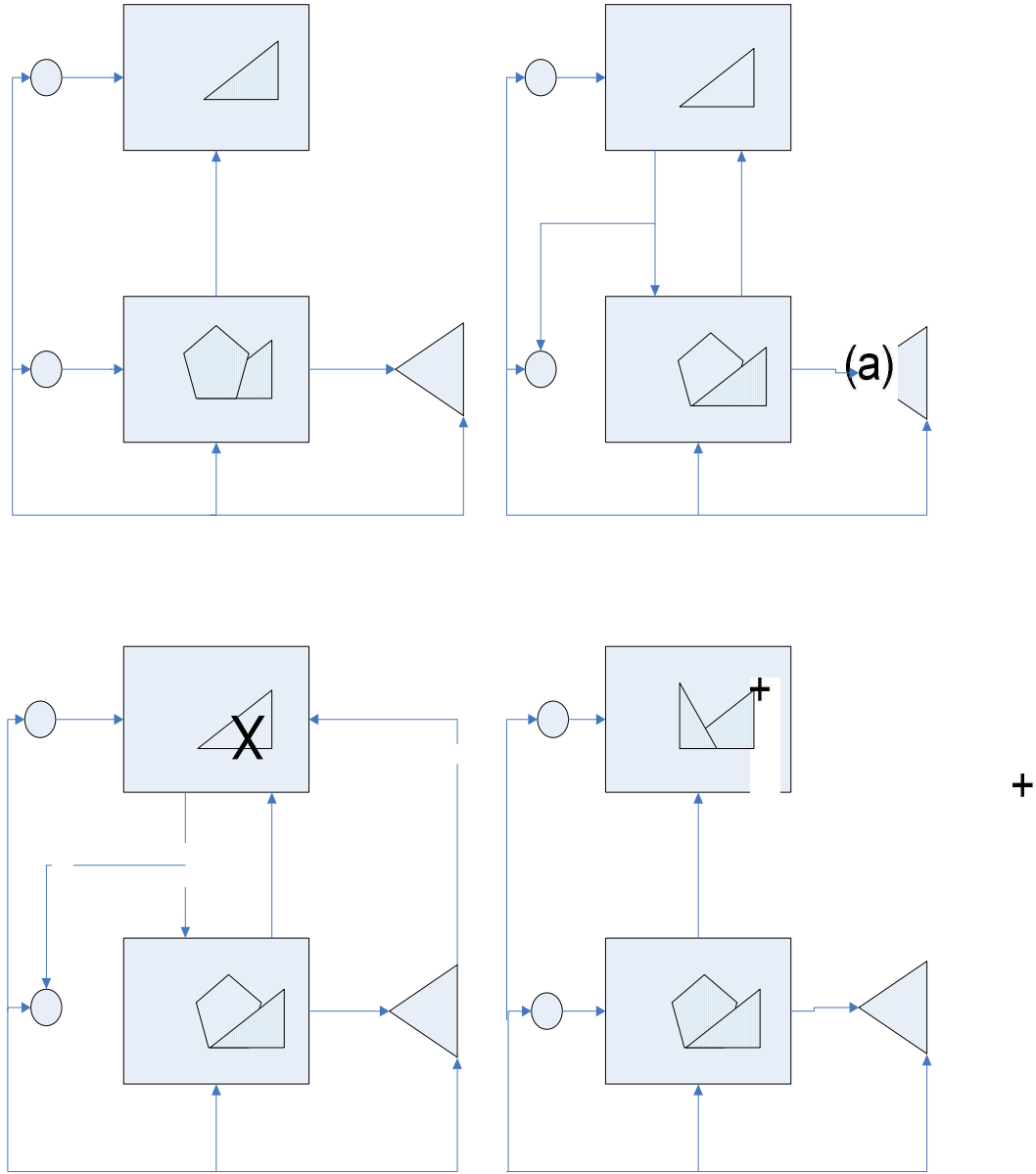
Figure 3.5: The Adaptive resonance theory 1 search cycle

3.5a Fault classification using Adaptive resonance theory Network1:

The Adaptive Resonance Theory 1 (ART 1) neural network is a vector classifier. An input vector is classified into one of a number of categories depending on similarity with the previous patterns. If the similarity criterion is not satisfied, a new category is created. Thus, the ART network solves the dilemma between plasticity and stability by enabling the learning to occur only in a resonance state [25]. In this way the network learns new information without destroying the old one. The ART1 Neural Network works with binary input patterns which are compared to prototype vectors in order of decreasing similarity until either an adequate match is found or a new category is created. The input analog waveforms from the Alternate Transients Program (ATP) software are first sampled and filtered using a low pass Finite Impulse Filter (FIR filter) and converted to bits before presenting them to the two Adaptive Resonance Theory 1 neural networks. The prototype patterns are stored in the network as connection weight vectors of the Top-Down Long Term Memory (T-D LTM) traces. The prototype vectors that have not been used until a certain time are set to 1. These T-D Long Term Memories (LTMs) give the most important property of code self-stabilization [26]. The new input patterns are also in part encoded in the Bottom-Up LTM traces (the B-U adaptive filter). The B-U LTMs play role in the first stage of determining the degree of similarity between the input pattern and the prototypes represented by the active output nodes in the upper (F2) layer. This is a coarse classification of the input in one of the active categories. The active state of an output unit indicates that the input vector belongs to the cluster represented by this unit [26].

The second stage of finding the similarity is the comparison of the feedback pattern from F2 with the input pattern. The vigilance parameter which can hold any value between 0 and 1 is used to determine whether the input vector and the output of F1 are similar enough. Consider V_j as the Top down LTM associated with the winner j in F2 and represent an already stored cluster. A vigilance parameter equal to 0.7 means that an input vector X with 10 ones will resonate with a winner vector V that has at least 7 ones of the same position as the input x. If the vigilance test is not satisfied for all categories

and the input pattern cannot be classified to neither of the existing categories, a new category is created. This process is repeated until a new cluster is formed or if the new pattern is classified in one of the existing clusters. This training process is repeated for the second neural network with current waveforms as input. While testing, the pattern or the fault case to be tested is classified as one of the existing clusters or if does not satisfy the similarity criterion, then a new cluster is formed. The outputs of the two neural networks are combined for classification.

3.6 Fuzzy Adaptive Resonance Theory (ART) neural network:

The Adaptive Resonance Theory Neural Network identifies natural groupings of data from a given input data stream and groups them into clusters i.e. similar data patterns are clustered together. The number and the size of the clusters are controlled by the tuning parameter called the threshold. If the threshold is set high, it leads to smaller number of clusters and if the threshold is set low, it creates larger number of clusters [27].

In an ART Neural Network, the data is initially processed to determine the similarity between the input patterns. The entire pattern set is processed only once. So the first pattern is formed as the first cluster and the other input patterns are categorized as the first or a new cluster is formed based on the similarity between the input patterns which is controlled by the threshold parameter. The similarity is usually measured by the Euclidean distance between the past data and the present input [24]. Since the entire data set is presented only once, this stage forms a set of unstable clusters. Hence all the unstable clusters from the previous stage are reiterated numerous times until the cluster structure becomes stable.

While testing the trained patterns, in the case of the Adaptive Resonance Theory Neural Network, the category is formed based on the neighborhood clusters [28]. But the clusters in the neighborhood are considered equally regardless of the size and the distance from the pattern. Hence there is a possibility that the pattern would be assigned to a wrong cluster. So the classification is not just a simple function of the number of categories of

32

the neighborhood clusters, but a more complex function of the size and the distance of the clusters from the pattern [29]. The fuzzy K-NN classifier calculates a set of membership values belonging to all the categories present in the K nearest clusters based on the following formula

$$\mu_c(x_i) = \sum_{k=1}^{K} \mu_c(w_k) \, 1/\|x_j - w_k\|^{2/(m-1)}/(\sum_{k=1}^{K} \rho_k / \|x_j - w_k\|^{2/(m-1)}) \qquad (18)$$

where $\mu_c(w_k)$ is given by

$$\mu_c(w_k) = \{ \rho_k, \text{ if cluster k belongs to category c}$$
$$0, \text{ otherwise} \qquad (19)$$

where $\mu_c(w_k)$ is the membership degree of cluster k belonging to category c, and is selected to be proportional to the radius of the cluster k.

Finally the pattern is assigned to the category with the highest membership degree according to the formula

$$g(x_i) = \max \{ \mu_c(x_i) \} \qquad (20)$$

The fuzzy ART neural network is capable of rapid stable learning of recognition categories in response to arbitrary sequences of analog or binary input patterns [30]. Fuzzy ART incorporates computations from fuzzy set theory into the ART neural network, which learns to categorize only binary input patterns. The generalization to learning both analog and binary input patterns is achieved by replacing appearances of the intersection operator in ART 1 by the MIN operator of fuzzy set theory. The MIN operator reduces to the intersection operator in the binary case. Category proliferation is prevented by normalizing input vectors at a preprocessing stage. A normalization procedure called complement coding leads to a symmetric theory in which the MIN operator and the MAX operator of fuzzy set theory play complementary roles.

Complement coding uses on-cells and off-cells to represent the input pattern, and preserves individual feature amplitudes while normalizing the total on-cell/off-cell vector. Learning is stable because all adaptive weights can only decrease in time. Decreasing weights correspond to increasing sizes of category boxes. Smaller vigilance values allow larger category boxes. Learning stops when the input space is covered by boxes. With fast learning and a finite input set of arbitrary size and composition, learning stabilizes after just one presentation of each input pattern. A fast-commit / slow-recode option combines initial fast learning with a forgetting rule that buffers system memory against noise. Using this option, rare events can be rapidly learned, yet previously learned memories are not rapidly erased in response to statistically unreliable input fluctuations [31].

3.6a Fault classification using Fuzzy ART neural network:

The Model of a Power System was created using the Alternate Transients Program (ATP) and the faulted waveforms for different faulted scenarios are simulated. . The training was performed by using the historical record of fault and disturbance cases obtained from the ATP. The first cycle of the waveform after the occurrence of the fault is used for training. Before, they underwent training; all the fault cases are filtered using a low-pass filter to remove all the high frequency harmonics and then sampled. The voltage and the current signals are sampled in the range [-1, 1]. Because of this technique, fault classification has just become an issue of pattern recognition instead of phasor computation and comparison. After a successful training, the system is now ready for on-line fault classification. The criterion of the acceptable match is called the Vigilance. Vigilance weighs how close the input exemplar is to the top-down prototype for resonance to occur. New patterns are classified according to their similarity to the pattern prototypes generated during training.

Both the voltage and Current Waveforms were used to training and testing. The voltage waveforms are used for training the first neural network and the current waveforms are

used for training the second neural network which separates the grounded and non-grounded type of fault. Using the unsupervised training, the space of the training patterns is transformed to the initial abstraction level containing set of clusters with corresponding prototypes, sizes and categories. These clusters are fuzzified and transformed to an intermediate level. The final level is attained when smooth boundaries are established using defuzzification. During testing, there is a possibility that new patterns which are not trained might occur. The neural network classifies them as new category (cluster) and the pattern vector is updated. Depending on the fault, the circuit breaker trips the corresponding phases until it is cleared.

Chapter Four: Experiments

4.1 Experimental setup:

The block diagram in the *Figure 4.1* shows the complete hardware and software solution, based on the Wavelet Transform and Adaptive Resonance Theory Neural Network approach.

Before fault detection and classification, it is important to consider the feature extraction and selection. Transmission line current and voltage parameters are measured with the current and voltage transformers and are converted into lower signal levels before providing them as input to the protective relay device. The frequency of the voltage and the current waveforms is 60 Hz. The waveforms are first filtered using a low pass filter with cut off frequency equal to 480 Hz; to remove the high frequency noise contents that exist in the signal. The filtered waveform is then sampled at more than twice the maximum frequency content present in the signal i.e. at a rate of 960 Hz (16 samples/cycle) to avoid aliasing. Depending on the type of the neural network used for classification, the discrete waveforms are converted to 1's and 0's. The Adaptive resonance theory neural network 1 (ART1) takes only binary inputs. So the each sample is converted to its binary equivalent. The length for each sample is taken as 8. If Fuzzy ART neural network is used for classification, then there is no necessity for converting the discrete waveforms to its binary values.

The voltage and current waveforms are now checked for the existence of the fault by analyzing the coefficients of the Discrete Wavelet Transform (DWT). The procedure for fault detection was explained in Chapter 3. If the fault is detected, the protective relay is activated and the fault has to be classified. If the fault is detected, circuit breaker trips and isolates the faulted phase. Finally the first cycle after the occurrence of the fault is used for neural network training and testing or implementation. The approach uses two different neural networks for fault classification. The first neural network takes the

sending end voltage waveforms as the input and the second neural network takes the combination of sending end and receiving end current waveforms as the input. The outputs of these two neural networks are combined to identify the fault condition. The training outputs are grouped into clusters and each cluster is uniquely identified by its fault type, fault location and fault resistance. Training is performed by accessing historical data of the fault scenario by simulation using Alternate Transients program (ATP). Testing is performed using real time data or by generating new fault scenarios using the same software. The new fault case will group itself into one of the clusters that are formed during training. Based on the cluster that the new pattern is grouped into, the information about the test pattern, fault type, fault location and fault resistance is provided.

Current and Volta
Transformers
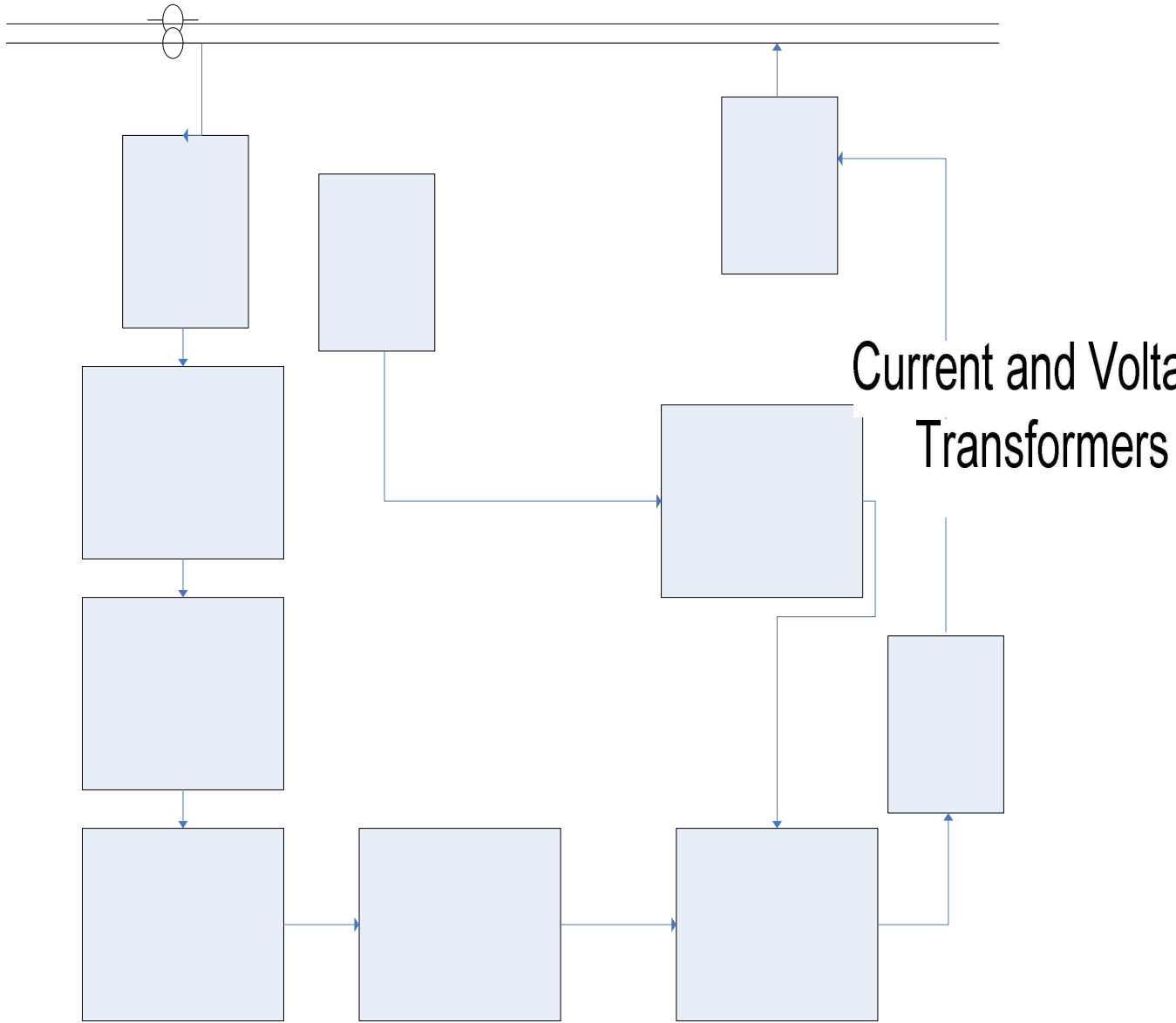
Analog Filter

Figure 4.1: Experimental setup

4.2 Simulation studies – ATP (Alternate Transients Program):

Alternate Transients Program (ATP) is the most widely used Electromagnetic Transients Program in the world. ATP is a universal program for simulating complex networks and control systems of arbitrary structures. ATP has been continuously developed through international contributions by Drs. W. Scott Meyer and Tsu-huei Liu, the co-Chairmen of the Canadian/American EMTP User Group [14]. ATP is used for detailed modeling of power network and simulation of electromagnetic transients. The procedure for network simulation scenarios and relaying algorithm is implemented in MATLAB [32], which is interfaced with the ATP.

Simulation of transmission line faults depends on three main fault parameters.

1) Fault type (10 types of fault and the non-fault type)
2) Fault distance (from 5% to 95% of the line length in increments of 5%; a total of 19 different values)
3) Fault Resistance (from 10 Ohms to 90 Ohms in increments of 5 Ohms; a total of 17 different values)

Type	AG,BG,CG,AB,BC,CA,ABG,BCG,CAG,ABC
Resistance	10,15,20,25,30,35,40,45,50,55,60,65,70,75,80,85,90 Ohms
Distance	10,20,30,40,50,60,70,80,90,100,110,120,130,140,150,160,170,180,190 miles from the sending end

Table1: Training scenarios for the neural network

The different fault scenarios for the neural network training is shown in *Table 1.*The total number of training sets presented to the neural network is 3230, which is equal to the product of 10, 19 and 17. As an example, the voltage and the current waveforms for resistance value equal to 10 Ohms and at a length of 20 km from the sending end for a Single Line to ground fault (a-g) is shown in *figures 4.2,4.3,4.4*and *4.5*. The three phase current and voltage waveforms for the other fault types are shown in Appendix.

Practical considerations such as analog filters, quantization errors due to A/D conversion are also incorporated into the simulation so that the data obtained for training the neural network is more realistic.
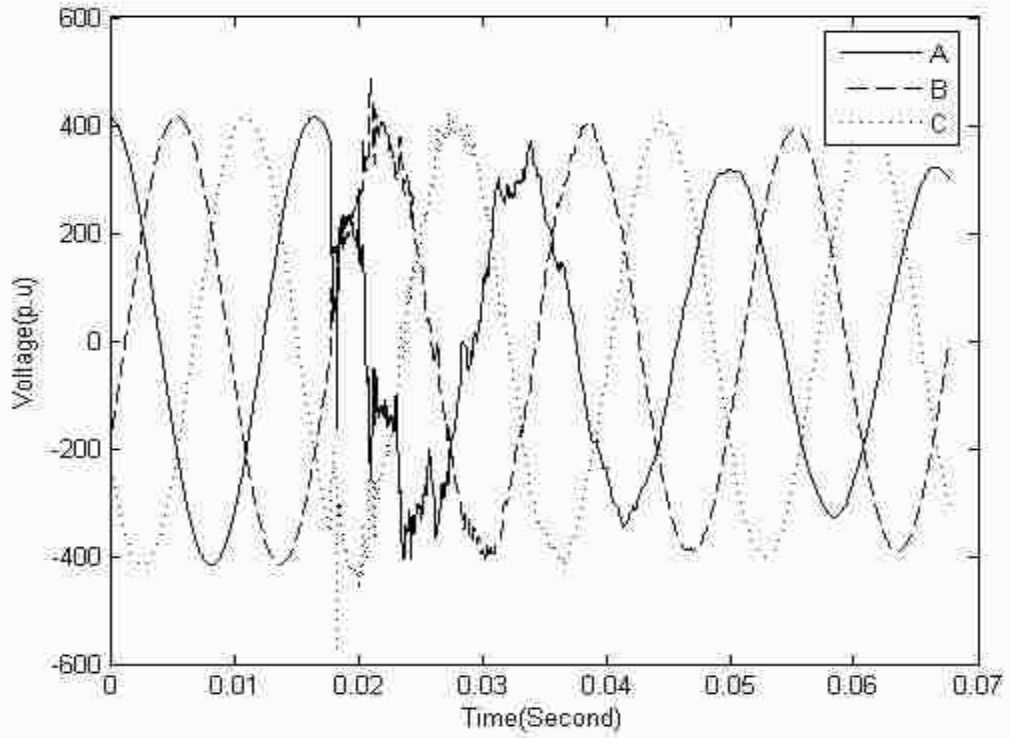
Figure 4.2: Alternate transients program output: voltage waveforms at sending end for AG fault
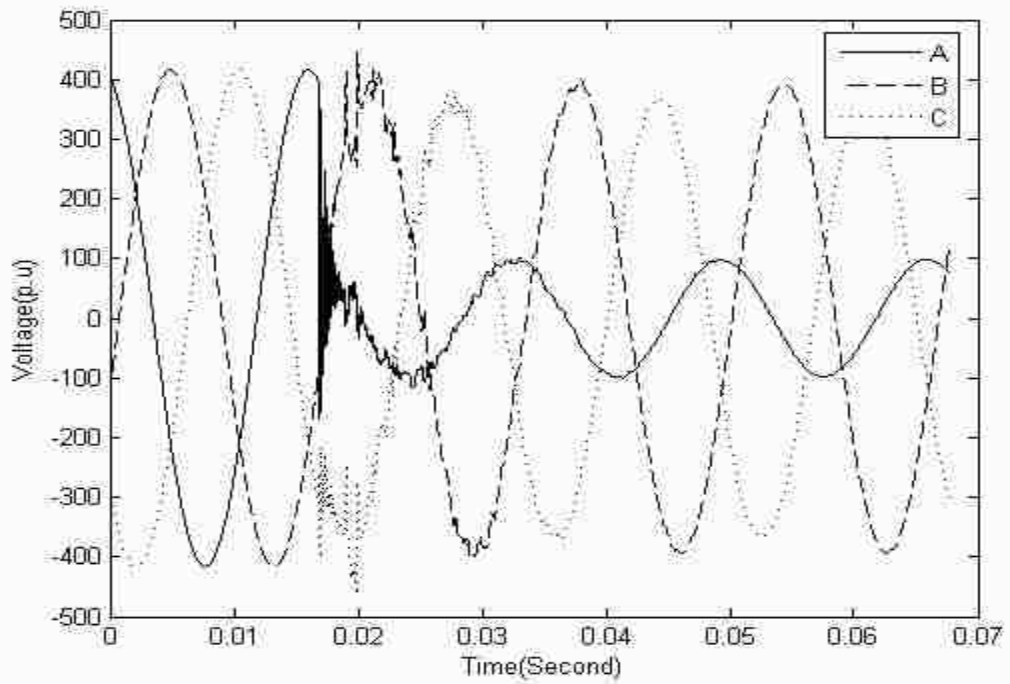
Figure 4.3: Alternate transients program output: voltage waveforms at receiving end for AG fault
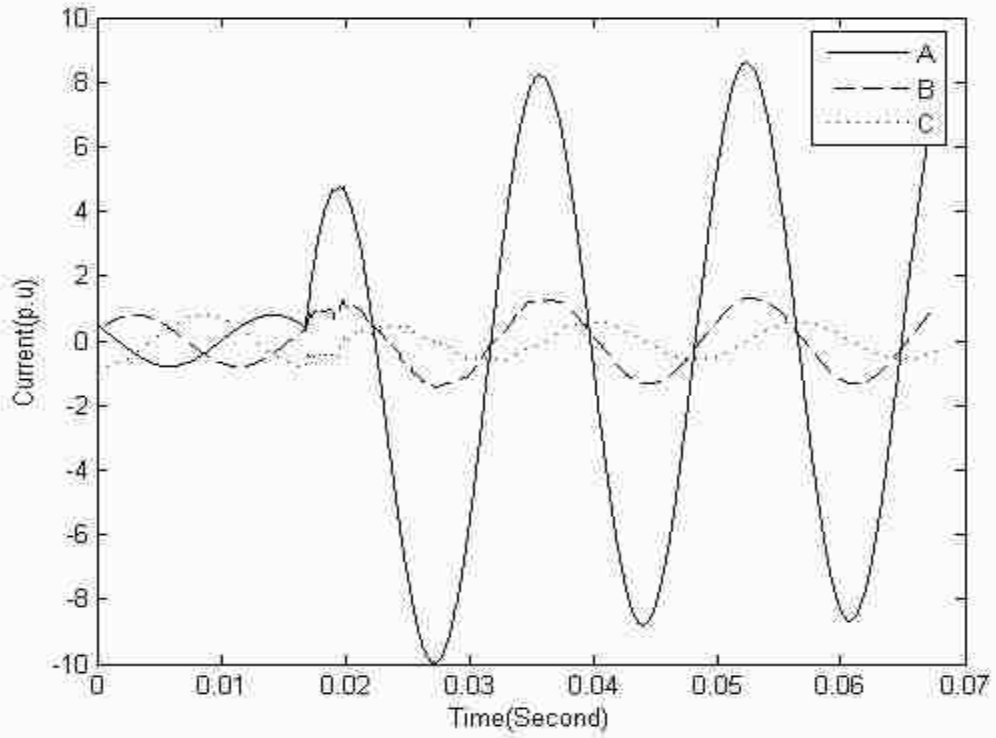
Figure 4.4: Alternate transients program output: current waveforms at sending end for AG fault
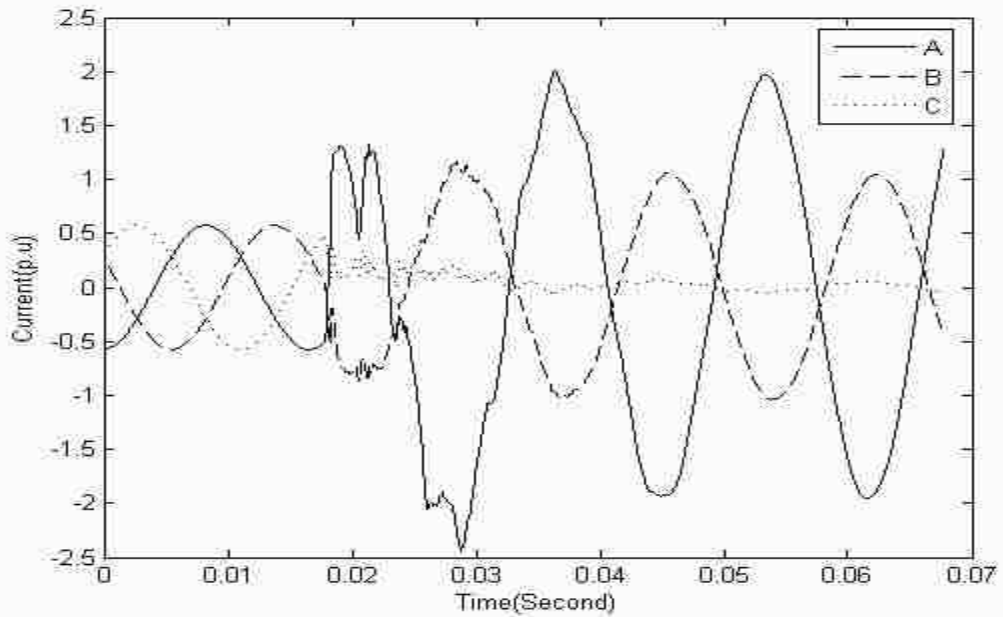
Figure 4.5: Alternate transients program output: current waveforms at receiving end for AG fault

4.3 Results:

In demonstrating the validity of our fault detection and classification system, our goal in this Chapter is to present the results at every step of the system and to compare the results with a one-step adaptive resonance theory network approach. Specifically, we will plot the outputs from all the stages and compare them with the existing techniques thereby demonstrating the superiority of the approach.

4.3a Results from the Discrete Wavelet Transform:

In the next few pages, the outputs from the Discrete wavelet transform block are plotted. Figure shows the sum_d1 vector (which was discussed under the section 3.2 from Chapter 3) for a non faulted waveform, single line to ground fault, line to line fault, line to line to ground fault and for a three phase fault. For small disturbances, the value of the total sum of coefficients value goes up the threshold but it remains for a very short interval of time and therefore is not declared as a fault.
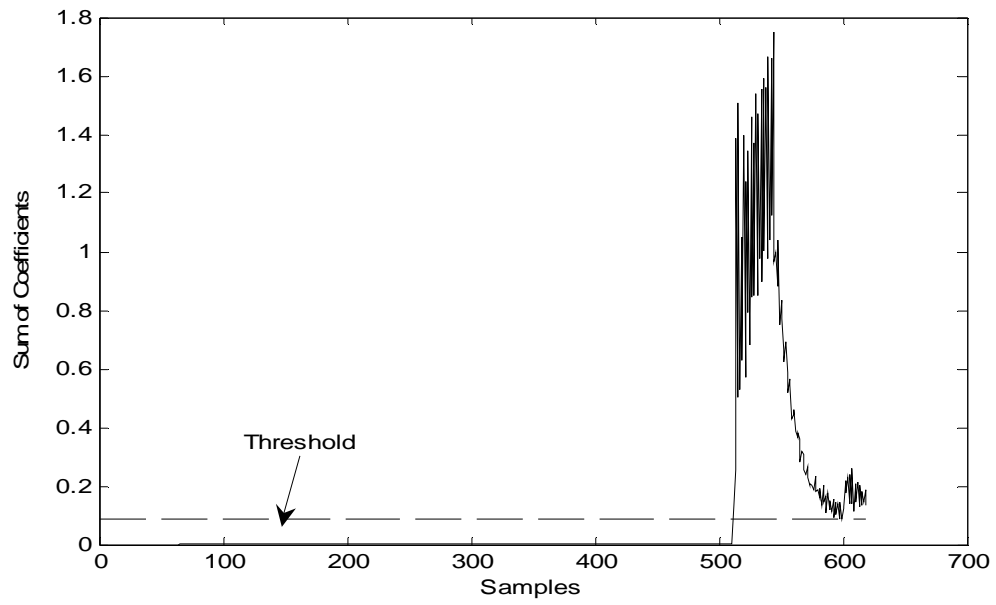
Figure 4.6: Discrete wavelet transform output for single line to ground fault
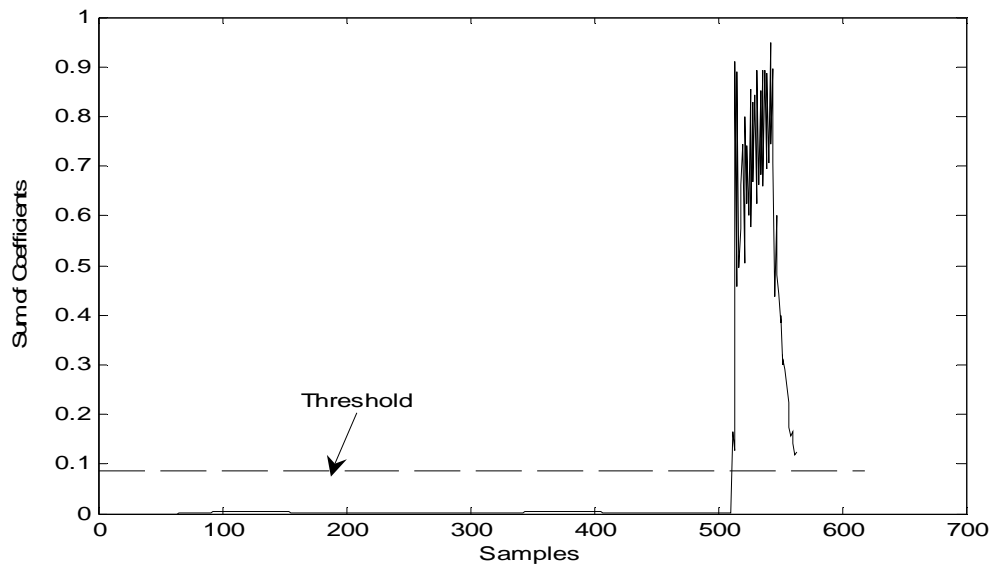
Figure 4.7: Discrete wavelet transform output for a line to line fault
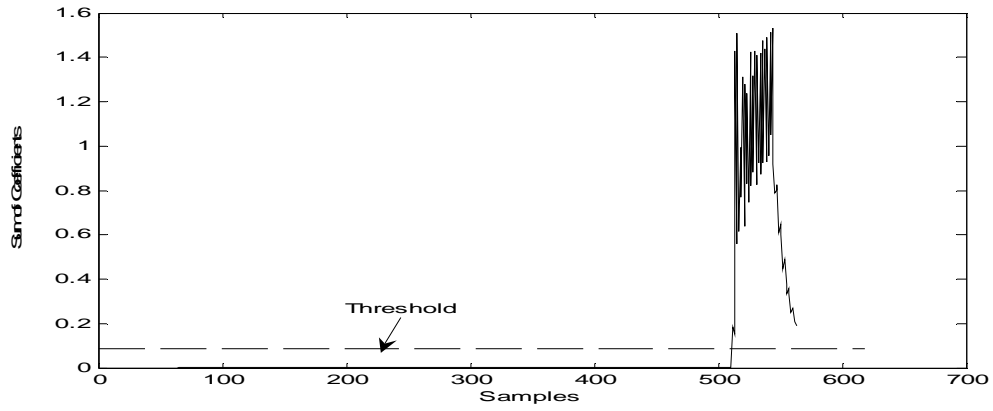
Figure 4.8: Discrete wavelet transform output for a line to line to ground fault
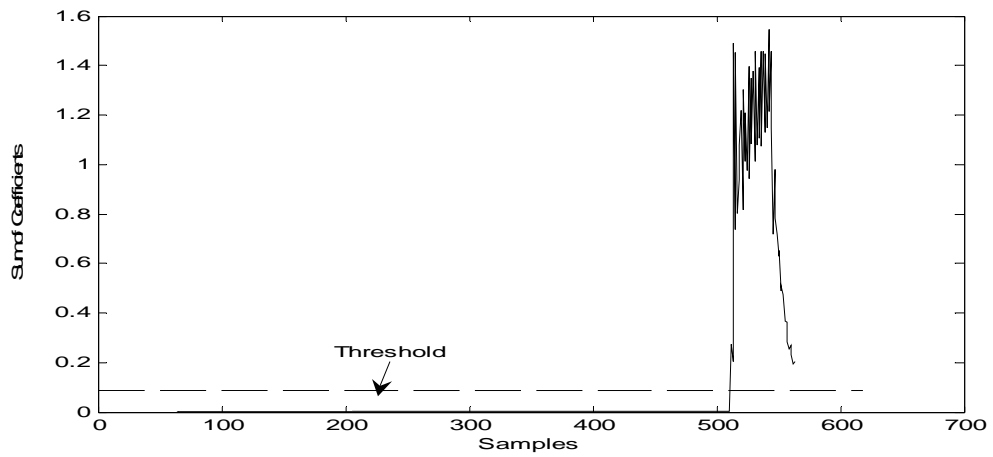
Figure 4.9: Discrete wavelet transform output for a three phase fault
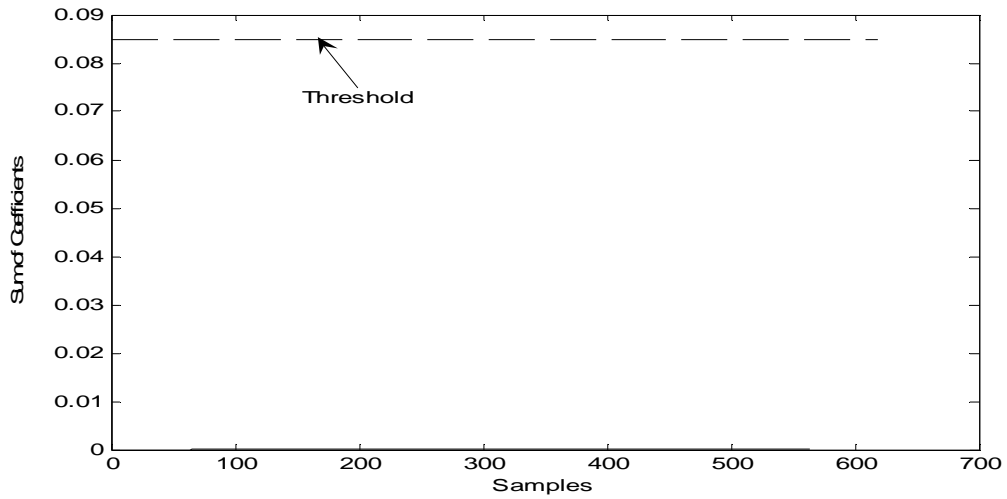
Figure 4.10: Discrete wavelet transform output for a non-faulted condition

The *Figures from 4.6- 4.9* show that once the fault occurs, the sum of the coefficients reach a maximum value (more than threshold D =0.8) and persists for more than 56 samples if it's a fault. *Figure 4.10* shows the output for a non-faulted condition.

	Sum of Coefficients			Is Sum of Coefficients greater than 0.085	Is number of samples for which Sum_d1 greater than 64	Fault or Not
	1st sample	20th sample	65th sample			
Three phase Fault Waveform	0.1858	0.8274	0.1869	YES	YES	YES
Non-faulted Waveform	6.2131e-006	4.5638e-006	1.9151e-006	NO	NO	NO

Table 2: Discrete wavelet transform: Sum of coefficients for a faulted and non-faulted condition

4.3b Results from the Neural Network:

The Neural Network outputs for the proposed approach are shown in *Table 3*. All the simulations were run on an Intel processor 1.7 GHz, 1 GB of RAM. The training time and the percentage of error for different training sets for the proposed approach is compared with the existing approach for both Adaptive Resonance Theory neural network 1 and Fuzzy Adaptive Resonance Theory neural network. The proposed approach with the help of two neural networks and a fuzzy rule classifies fault based on their type, location and resistance. The results are tabulated in Tables 3, 4 and 5 as shown. A comparison of the neural network training times for the proposed approach with the existing approach which uses a single neural network with both current and voltage waveforms as input is shown in Table 3. Both the neural networks; approach while using both Adaptive Resonance Theory neural network 1 and Fuzzy Adaptive Resonance Theory neural network which were implemented have been used for comparison. From the table, it is clear that the Fuzzy ART with analog input patterns as input takes lesser time for training when compared to the ART1 neural network which uses binary patterns as inputs for training. A comparison of the training times is made for both 1000 and 3230 different fault cases. All these fault cases are generated by the Alternate Transients Program (ATP) software. The table is plotted as a graph with the training time on the y-axis against the number of fault cases on the x-axis in *Figures 4.17 and 4.18*.

The percentage of error for fault classification is shown in Tables 4 and 5. Table 4 shows the error obtained when the test pattern is classified as one of the 60 existing clusters formed during training. The 60 clusters are obtained as a combination of 10 different types of fault, 2 fault location zones (0-90% of the length of the transmission line and 90-100% of the length of the transmission line from the sending end) and 3 fault resistances (10, 50 and 100 Ohms). The table 5 shows the percentage of error values for 20 clusters which are formed during training from 10 fault types and 2 fault location zones. A discussed in Chapter 3, there is a possibility that the voltage and current waveforms for two-phase grounded and non-grounded faults be classified in the same type and it was

proposed to use a fuzzy based rule with the absolute sum of current waveforms to check the neural network outputs. The *figures 4.11* till *4.16* clearly show the difference between the grounded and the non-grounded fault conditions when the sum of the current waveforms are used for classification. The maximum value in the case of a non-grounded fault is in the order of 10^{-7}, and for a grounded fault it typically lies between 0 and 5. Owing to these large changes in magnitude, a fuzzy rule is formed with a simple IF condition and the grounded and non-grounded faults are easily separated.

	ART1				Fuzzy ART			
	Existing Approach		Proposed Approach		Existing Approach		Proposed Approach	
Number of fault cases	1000	3230	1000	3230	1000	3230	1000	3230
Training time (seconds)	15707.72	4935.23	9877.24	28745.56	2577.45	9856.23	1377.85	4832.34

Table 3: Comparison of neural network Training times for the existing and proposed system

Number of clusters: 60 (10 fault types and 2 fault location zones(0-90% distance from sending end and 90-100% from the sending end) and 3 fault resistances (10,50 and 100 Ohms))

Test cases	3230	
Technique	ART1	Fuzzy ART
Error (%) due to fault type only	0.1548	0.0929
Error (%) due to fault location only	8.2663	6.7492
Error (%) due to fault Resistance only	46.4396	39.9071

Table 4: Percentage of Error for the proposed system (60 clusters)

Number of clusters: 20 (10 fault types and 2 fault location zones (0-90% distance from sending end and 90-100% from the sending end))

Test cases	3230	
Technique	ART1	Fuzzy ART
Error (%) due to fault type only	0.1548	0.0929
Error (%) due to fault location only	8.2663	6.7492

Table 5: Percentage of Error for the proposed system (20 clusters)
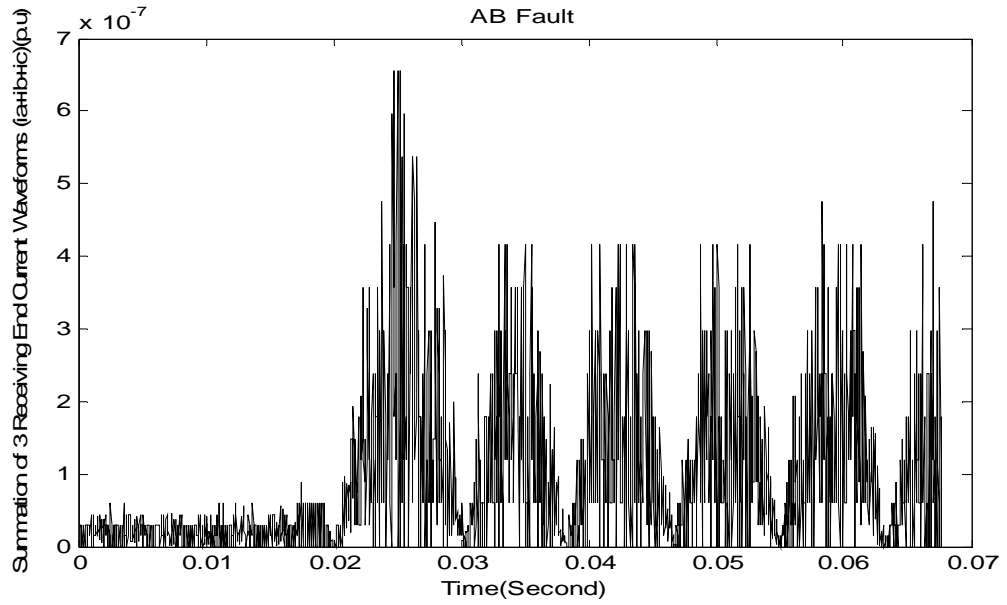
Figure 4.11: Summation of receiving end current waveforms for AB Fault
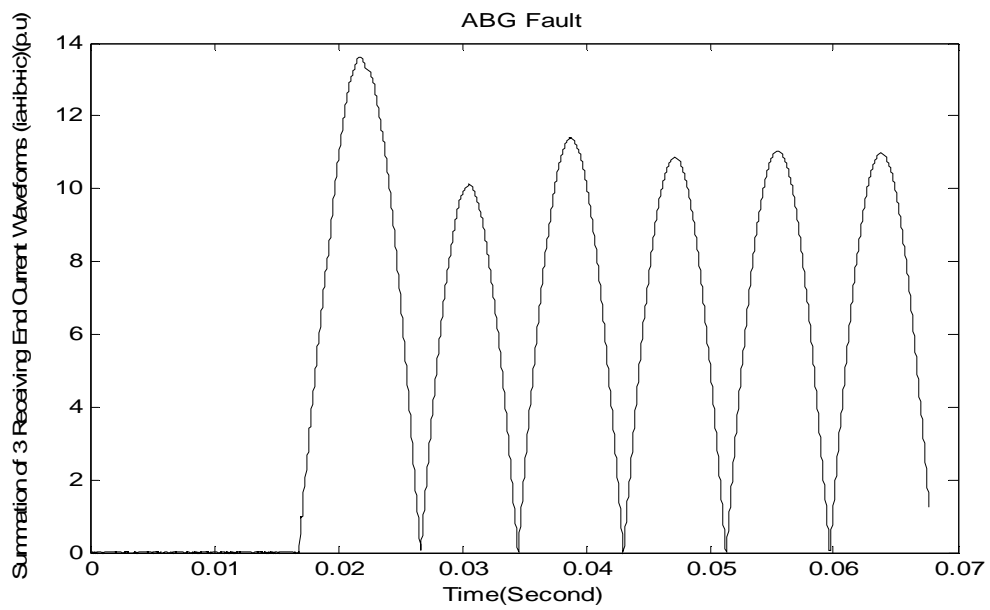
Figure 4.12: Summation of receiving end current waveforms for ABG Fault

50
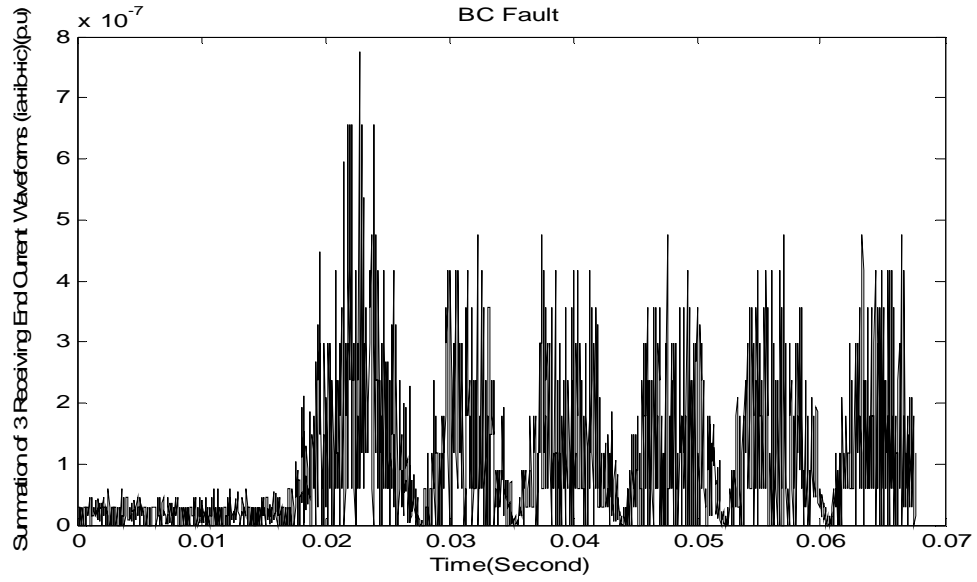
Figure 4.13: Summation of receiving end current waveforms for BC Fault
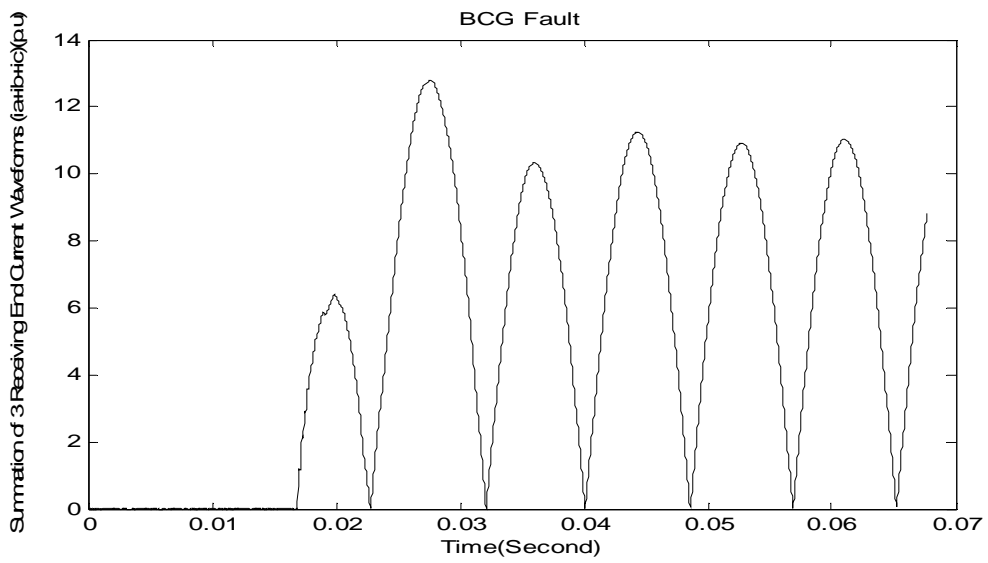
Figure 4.14: Summation of receiving end current waveforms for BCG Fault

51
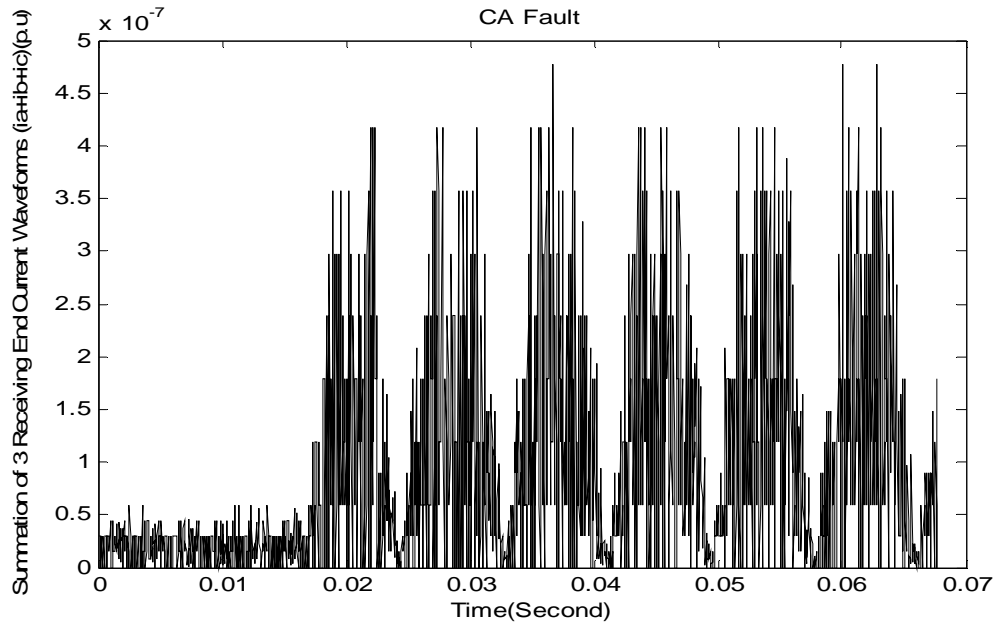
Figure 4.15: Summation of receiving end current waveforms for CA Fault
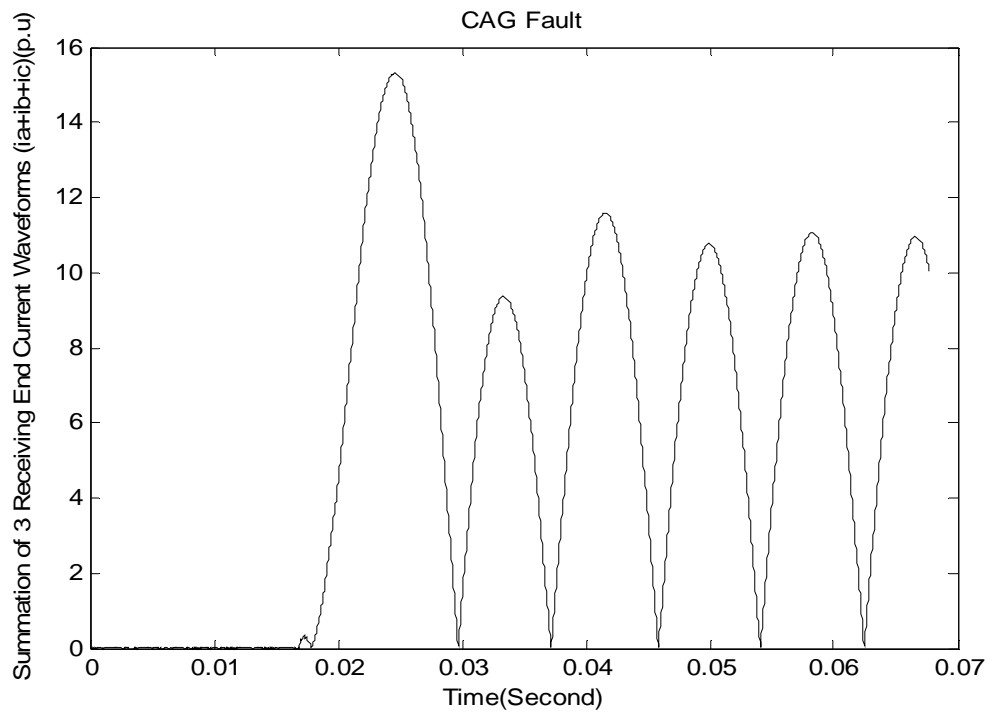
Figure 4.16: Summation of receiving end current waveforms for CAG Fault
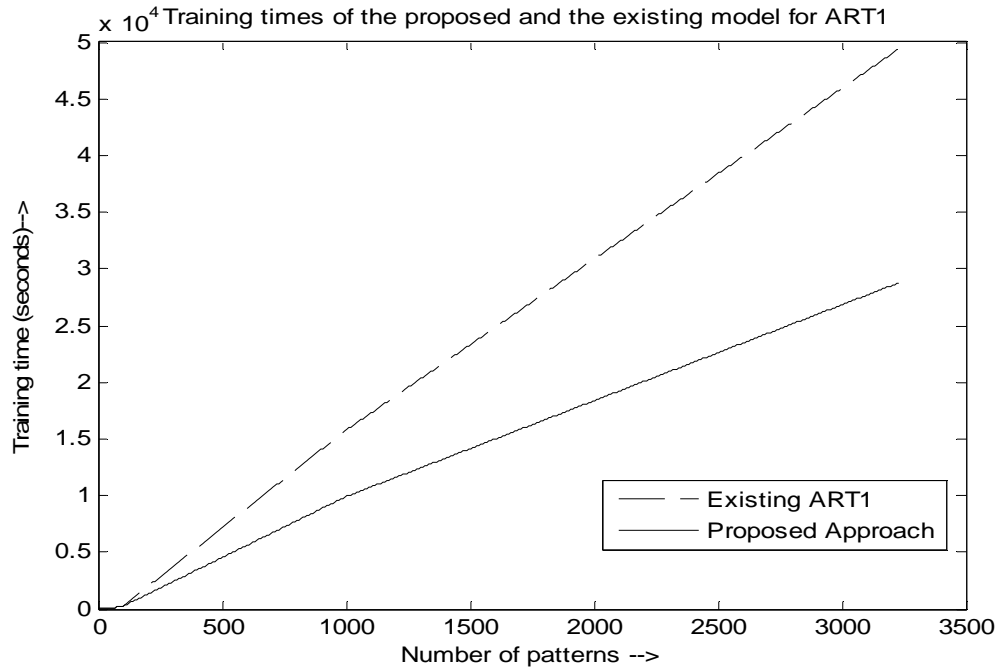
Figure 4.17: Comparison of the training times of the proposed and existing technique for ART1 neural network
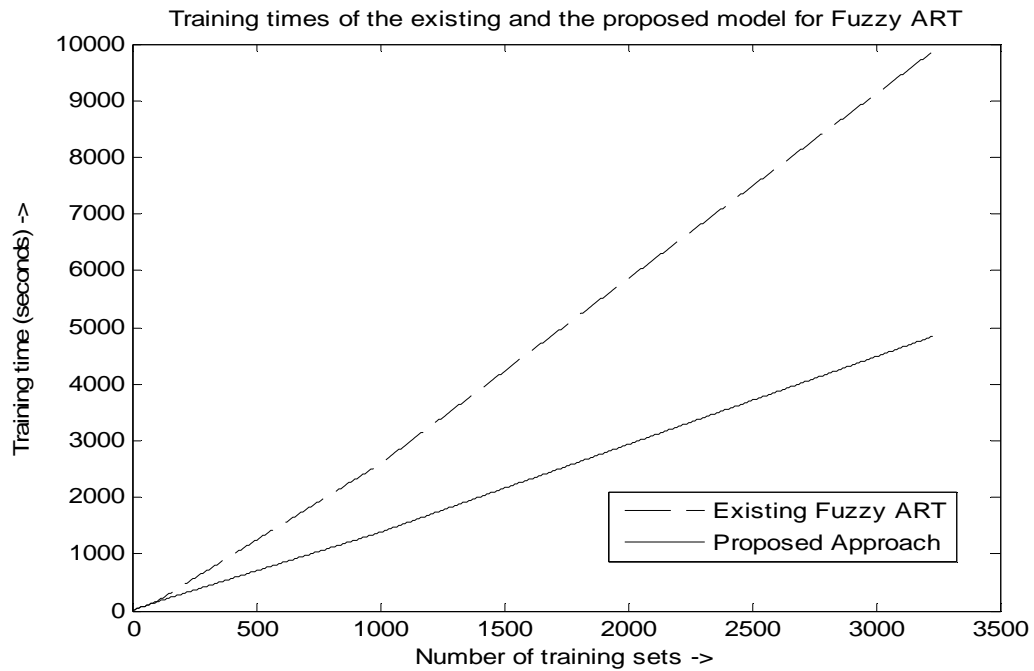
Figure 4.18: Comparison of the training times of the proposed and existing technique for Fuzzy ART neural network

Chapter Five: Summary

The results presented in the preceding Chapter shows the validity of the proposed system. The proposed approach classifies the fault based on the type, location and resistance. These three parameters are very important for clearing the fault quickly and resume normal operating conditions. The percentage of error for classifying the type of fault is minimal and is much lesser than the traditional approaches. The training time is considerably less when compared to the existing one step neural network approach. The discrete wavelet transform separates the faulted and non-faulted waveforms and a combination of the two neural networks helps in fault classification. The training and testing patterns for the first neural network are obtained only from the sending end voltage waveforms and the second neural network uses both the sending end and receiving end current waveforms for training and testing. The outputs of the two neural networks and the fuzzy rule with the absolute sum of current waveforms are used for fault classification based on type, location and resistance.

Appendix

Alternate Transients Program (ATP) Output Waveforms:
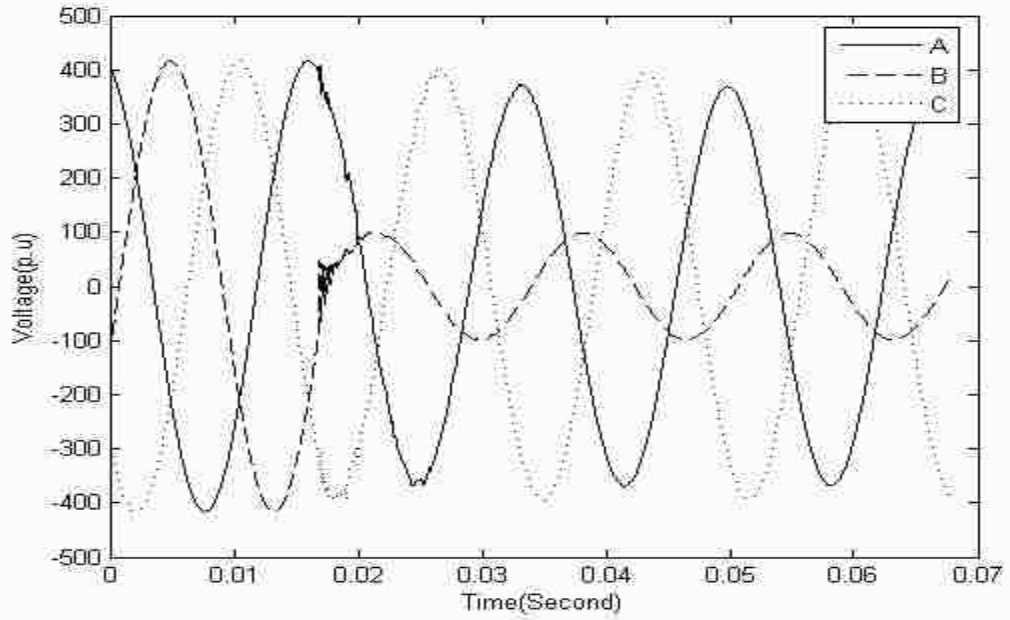
Figure A.1 (a) BG Fault: Voltage Waveforms at the Sending End and (b) Receiving End
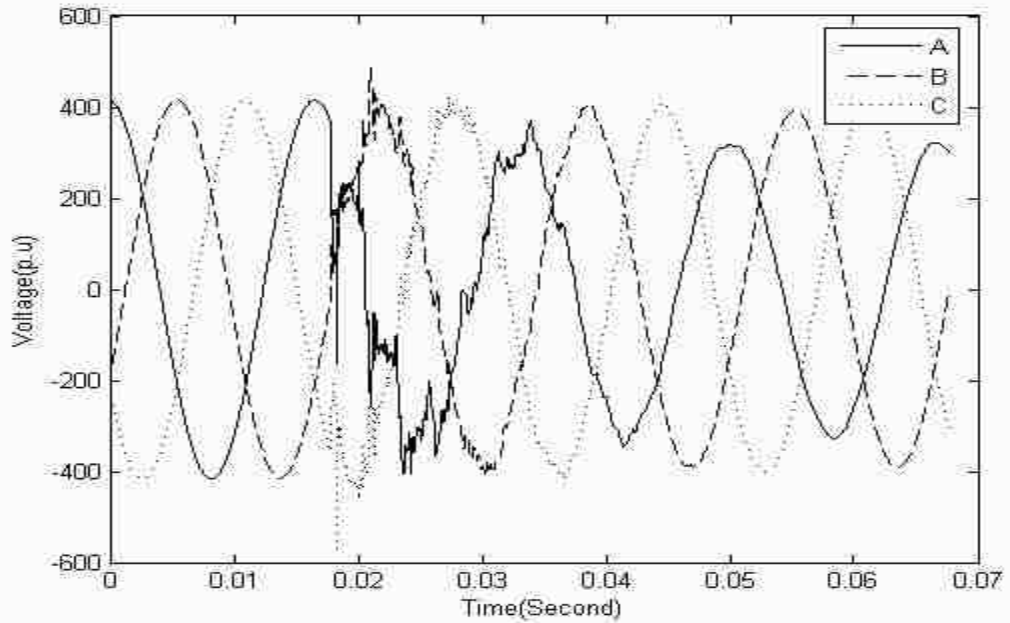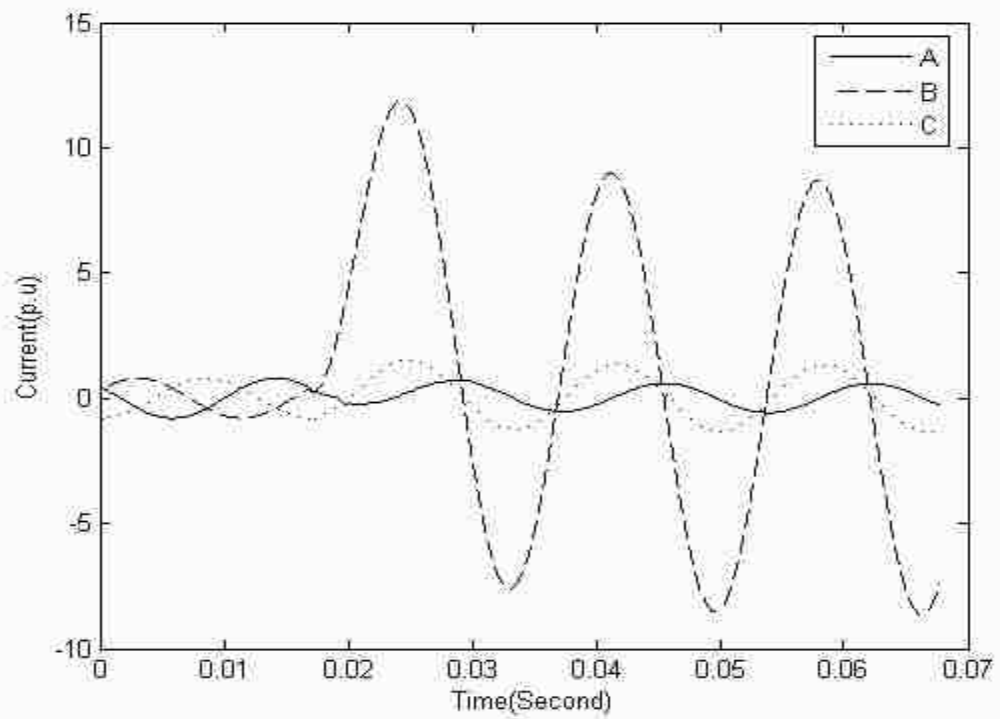
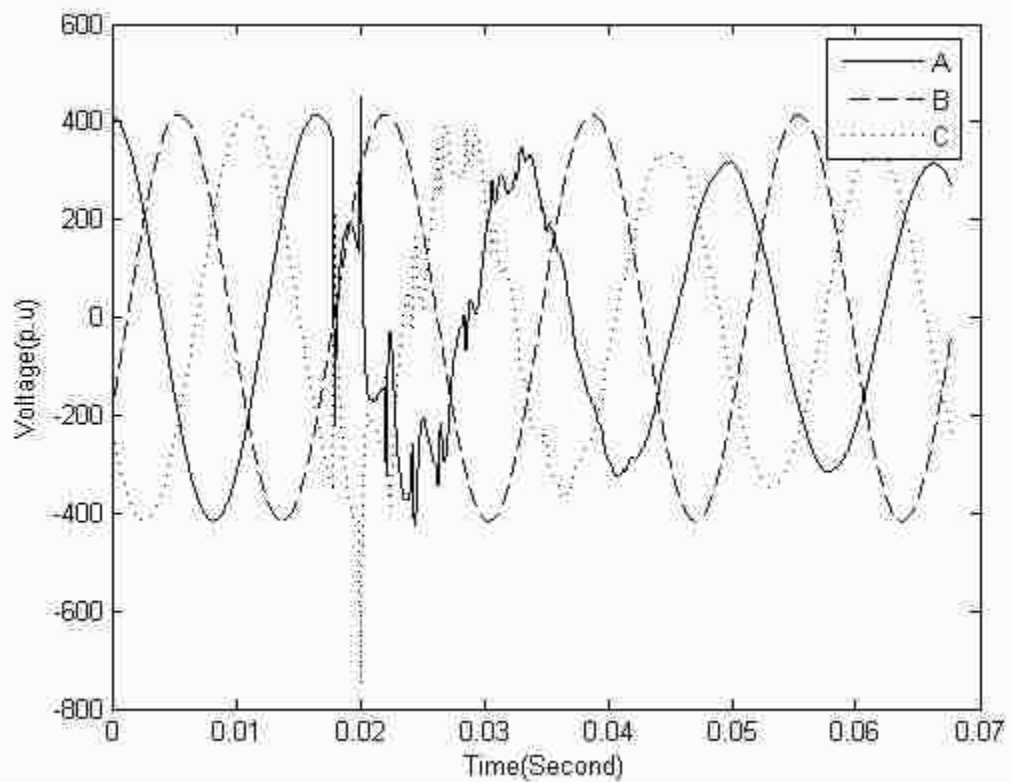*Figure A.2 (a) BG Fault: Current Waveforms at the Sending End and (b) Receiving End*

56
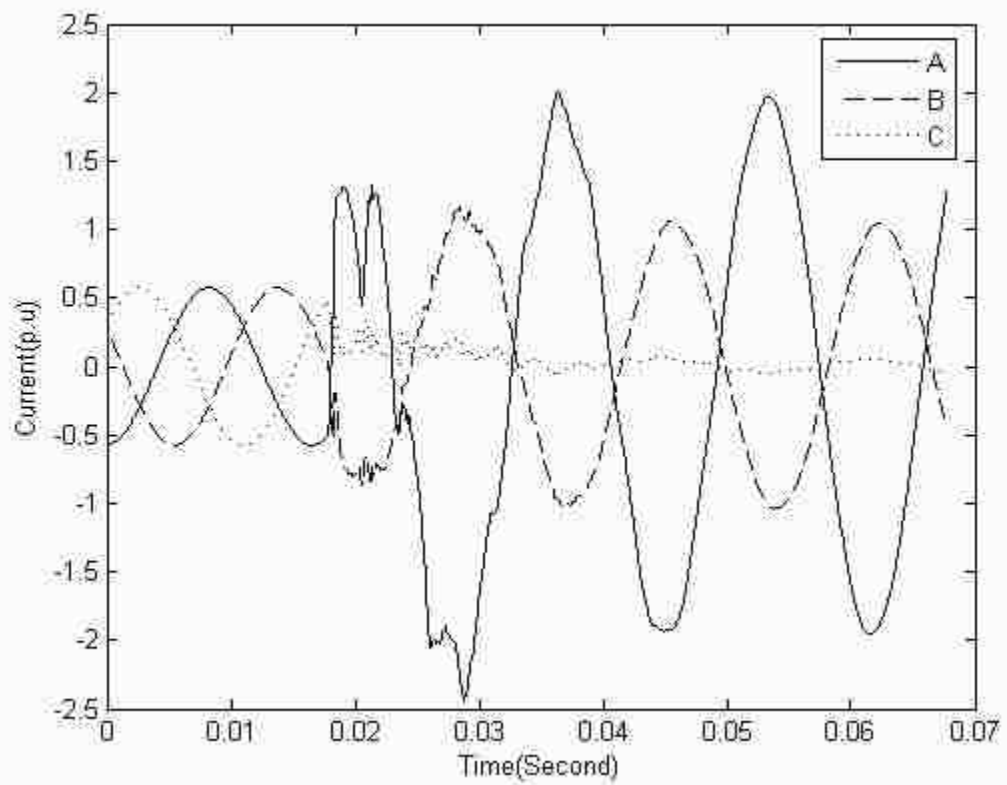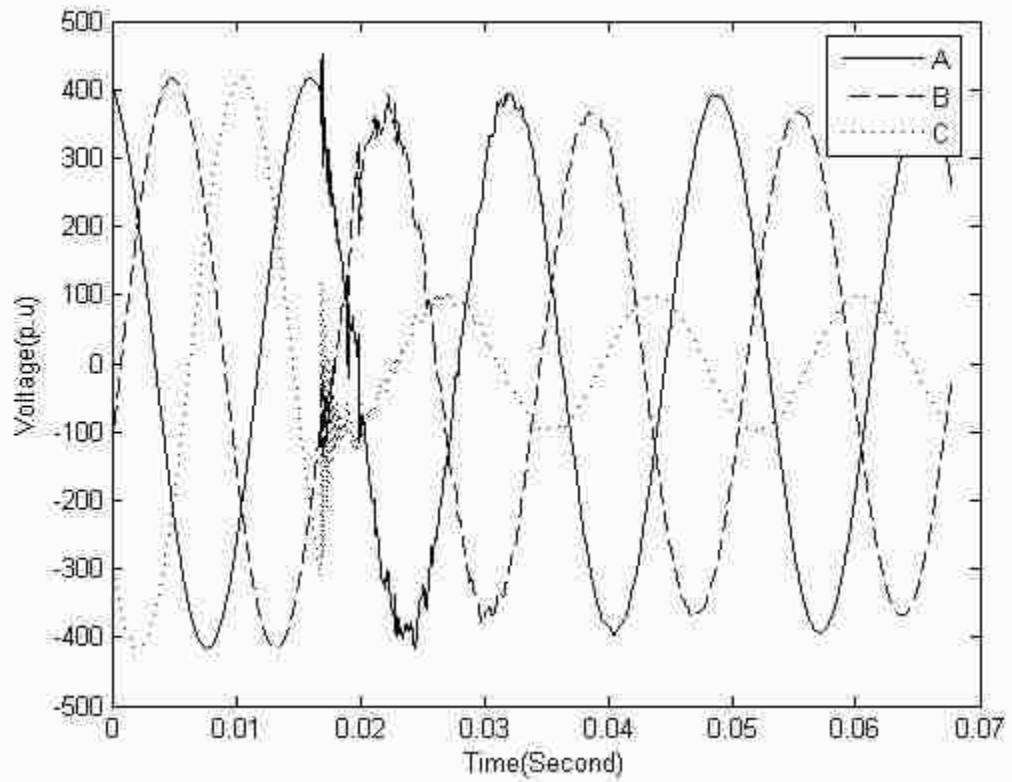
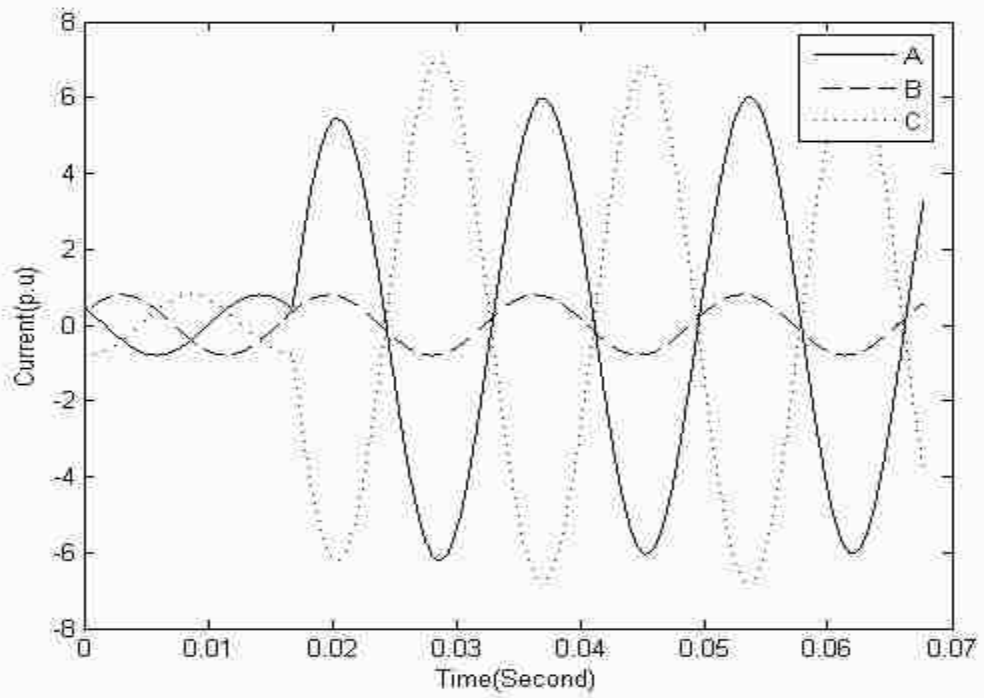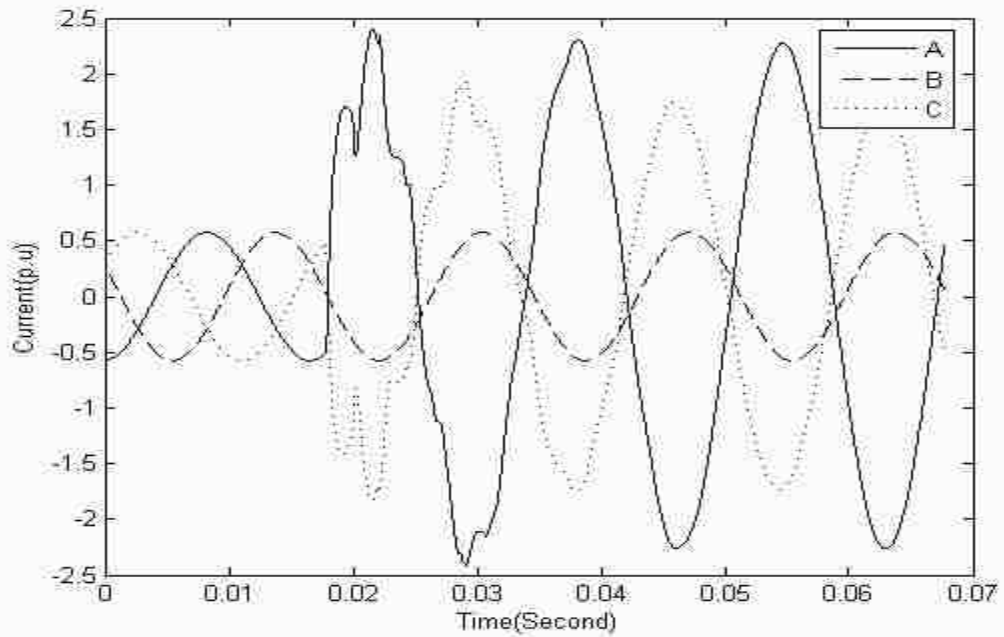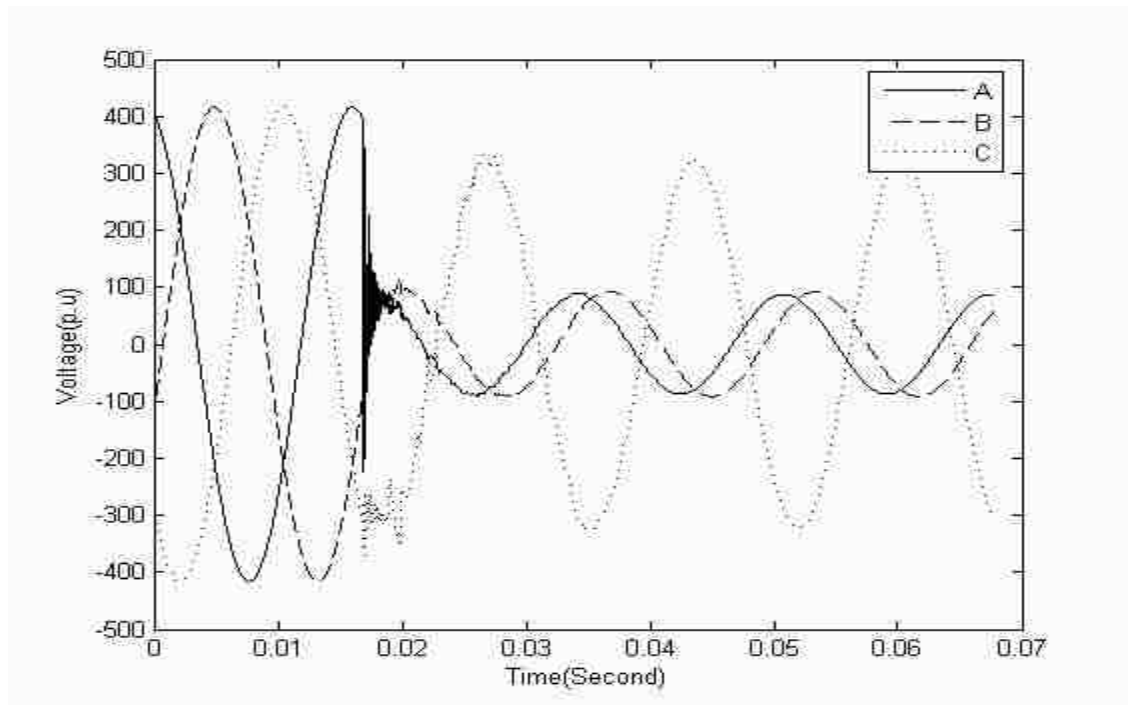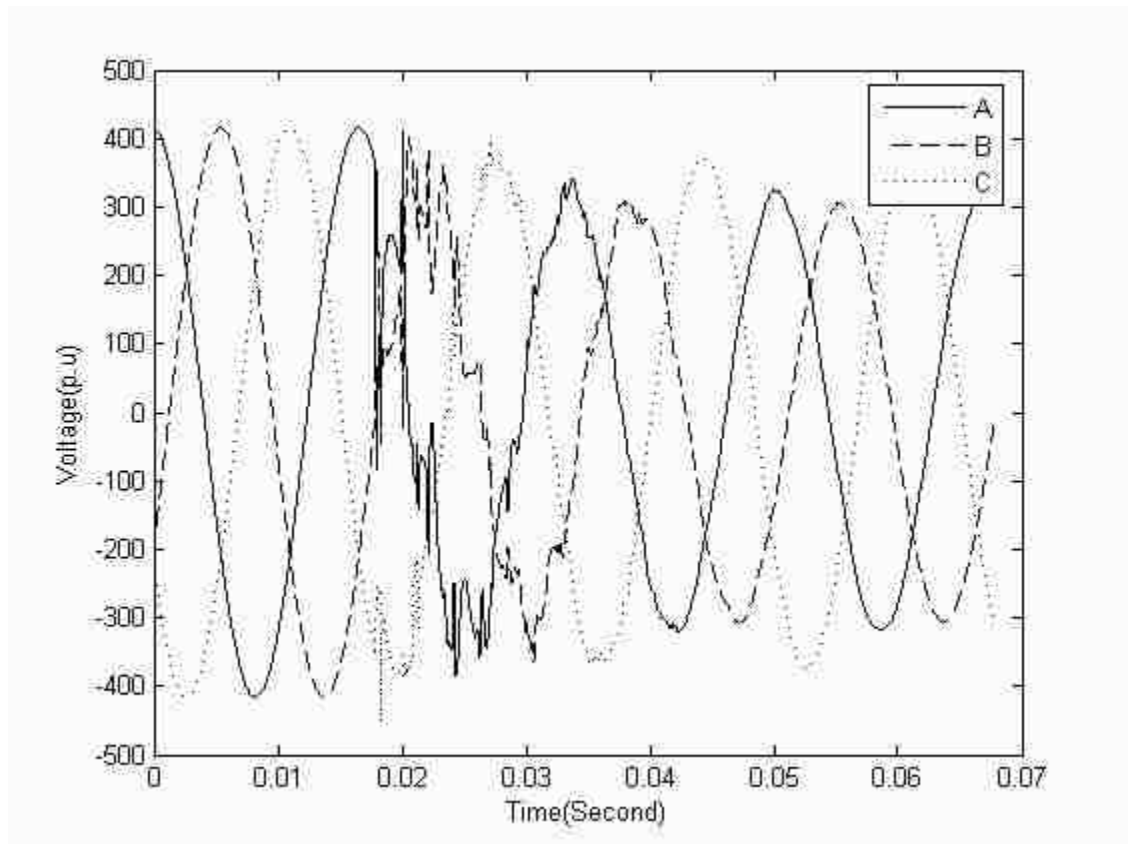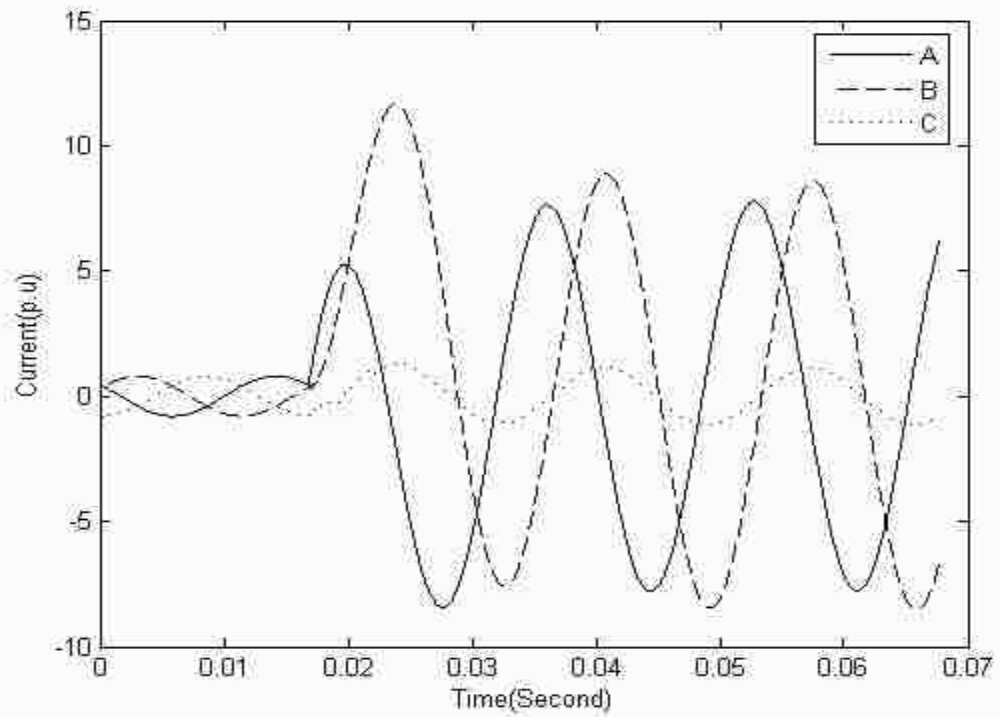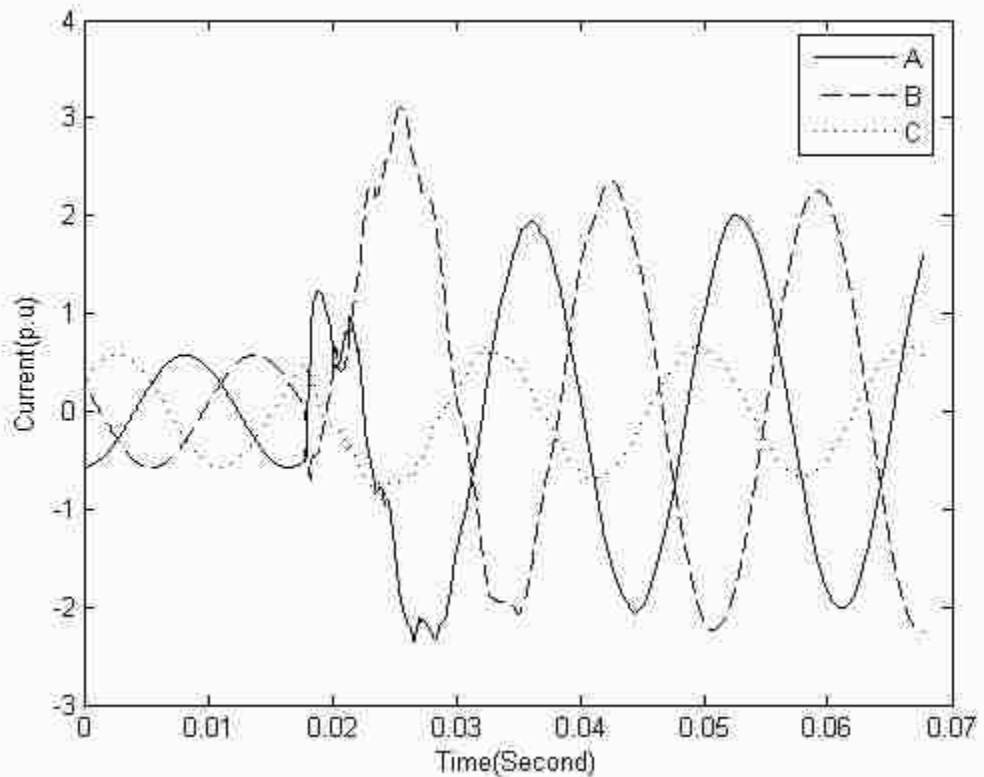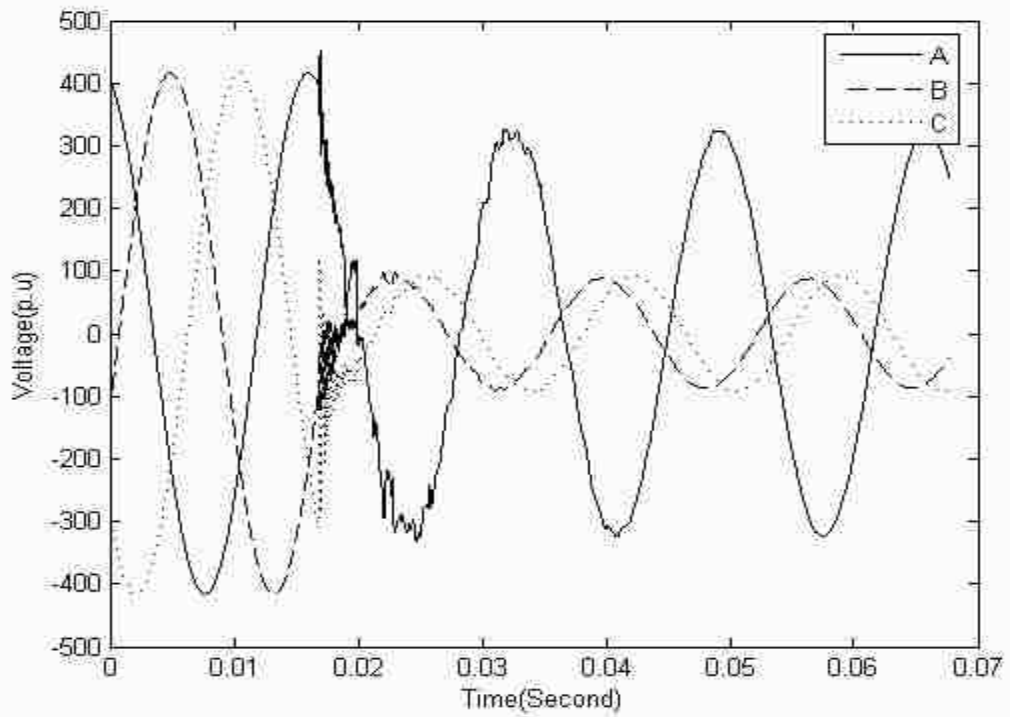*Figure A.3 (a) CG Fault: Voltage Waveforms at the Sending End and (b) Receiving End*
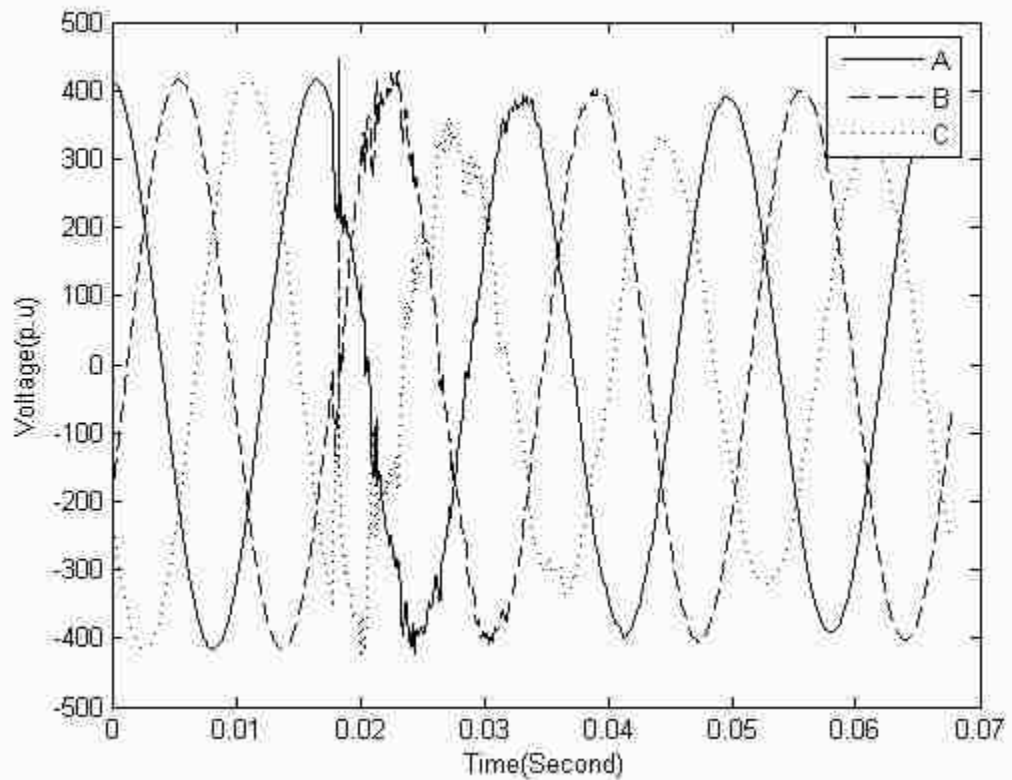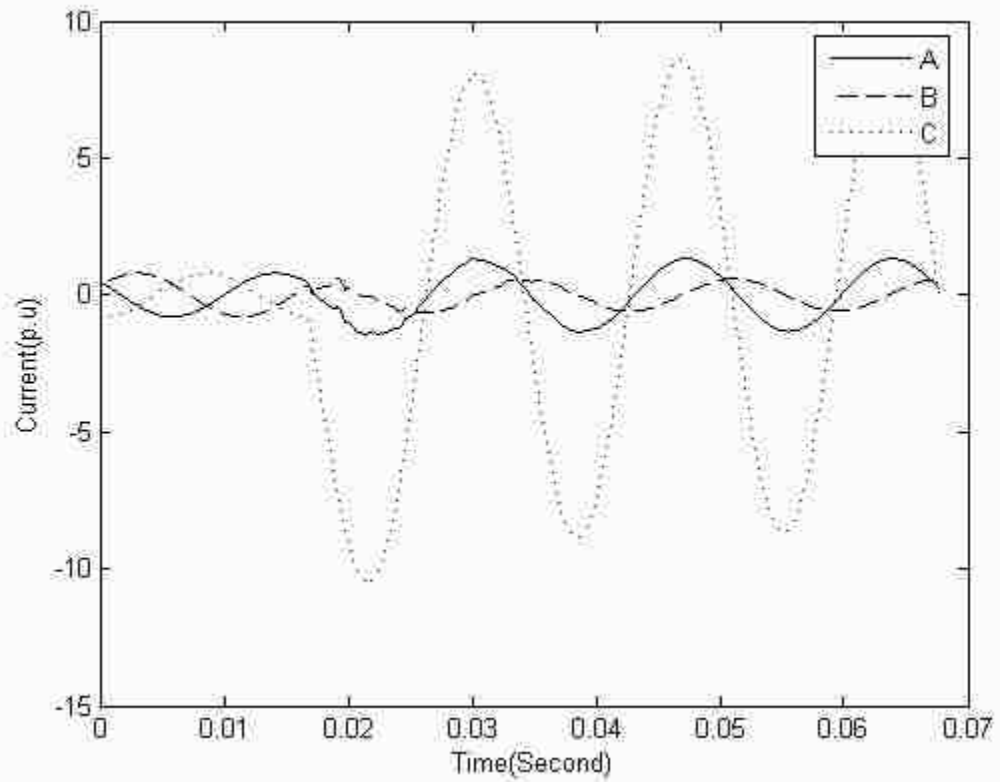
Figure A.4 (a) CG Fault: Current Waveforms at the Sending End and (b) Receiving End
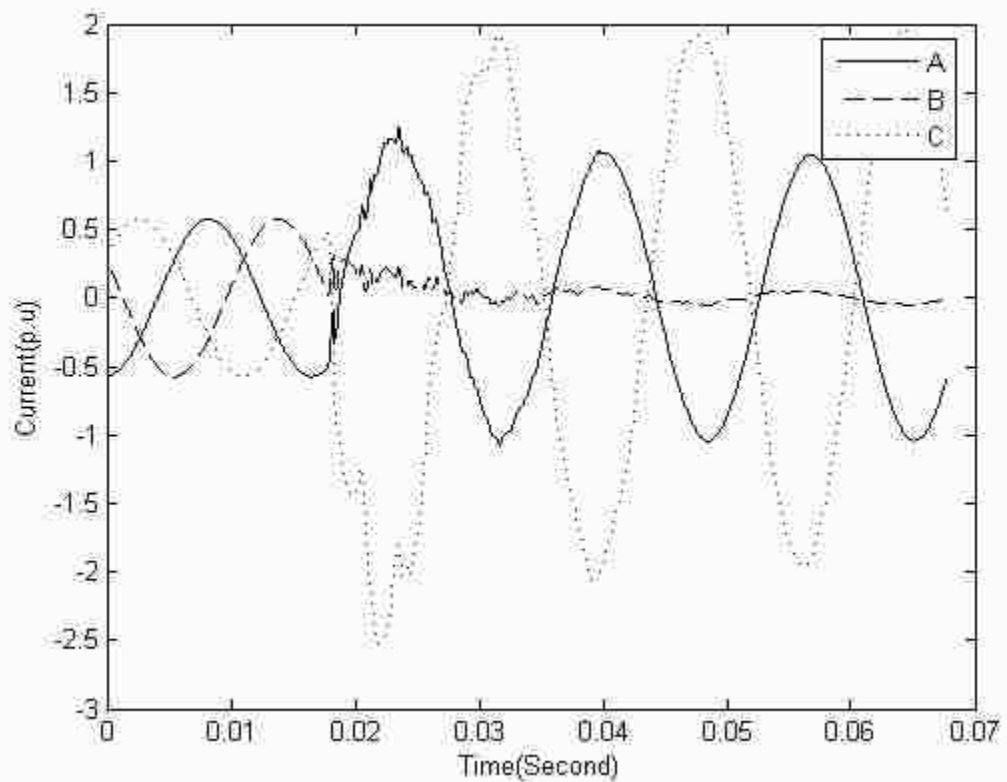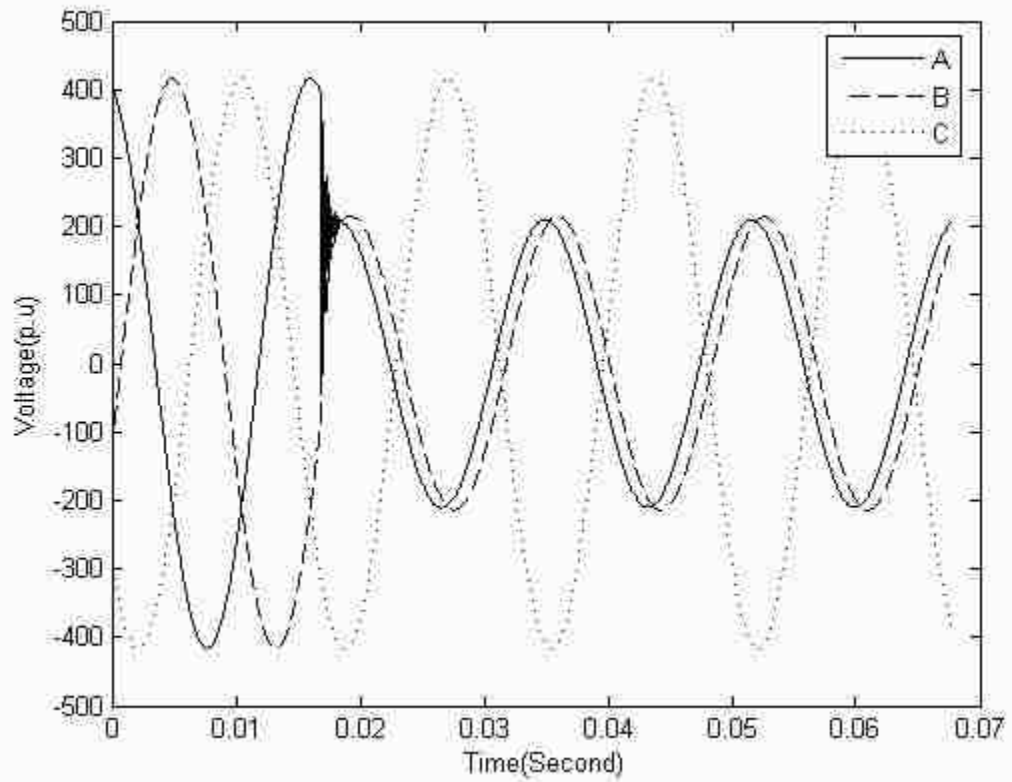
Figure A.5 (a) AB Fault: Voltage Waveforms at the Sending End and (b) Receiving End

59
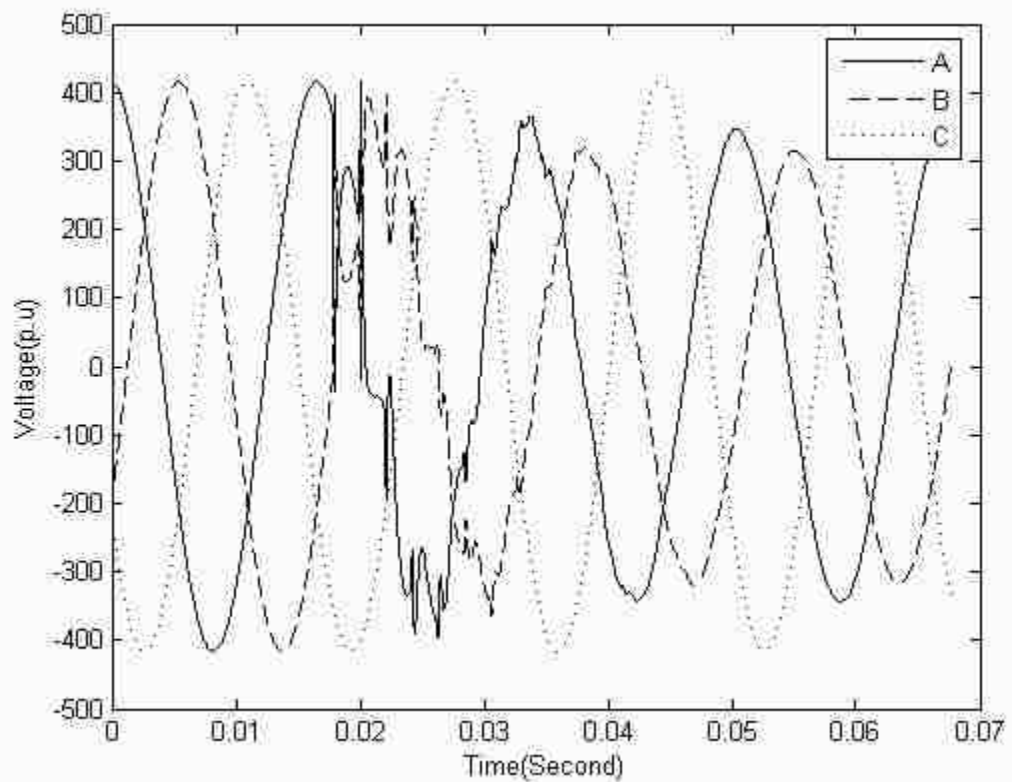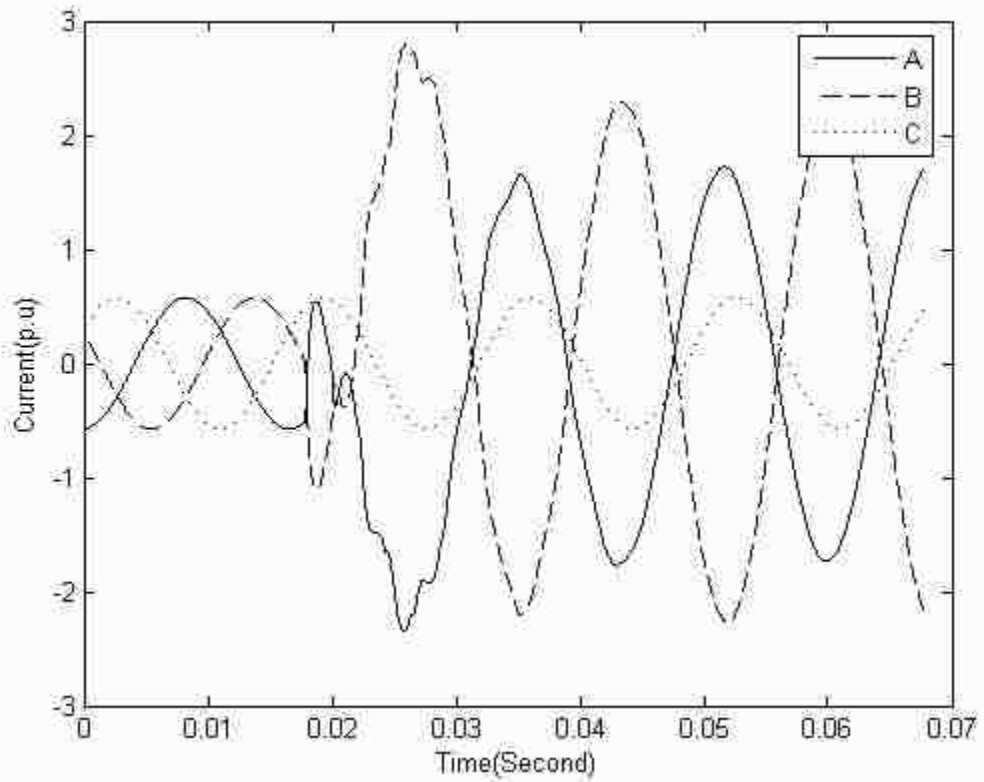
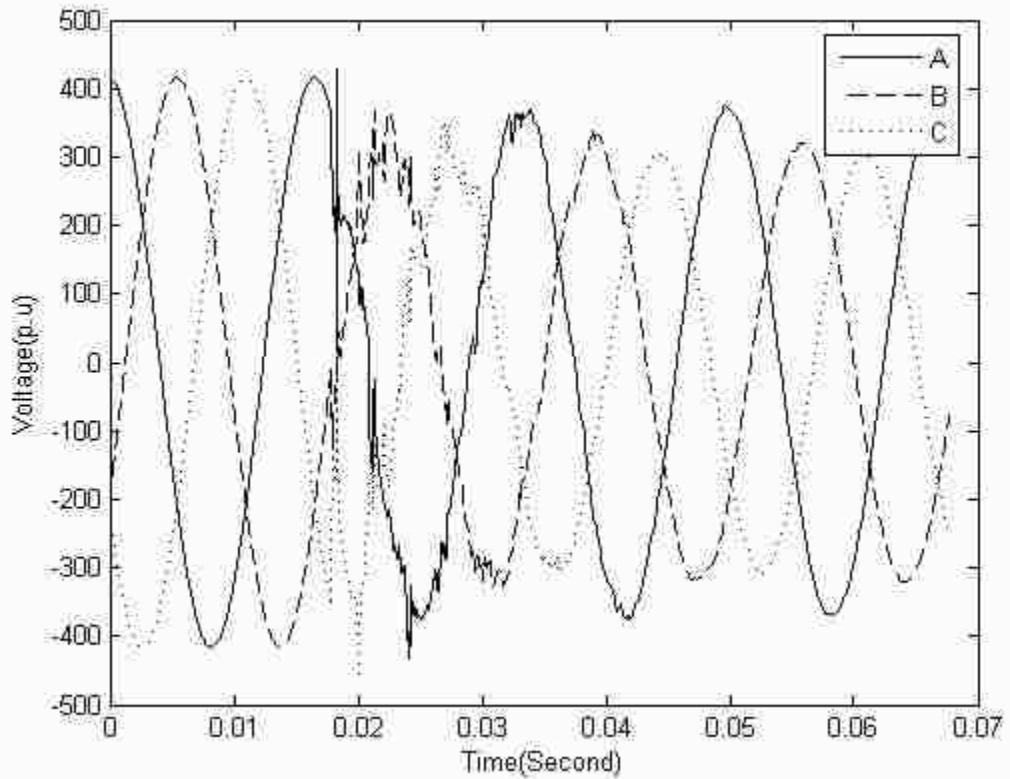*Figure A.6 (a) AB Fault: Current Waveforms at the Sending End and (b) Receiving End*
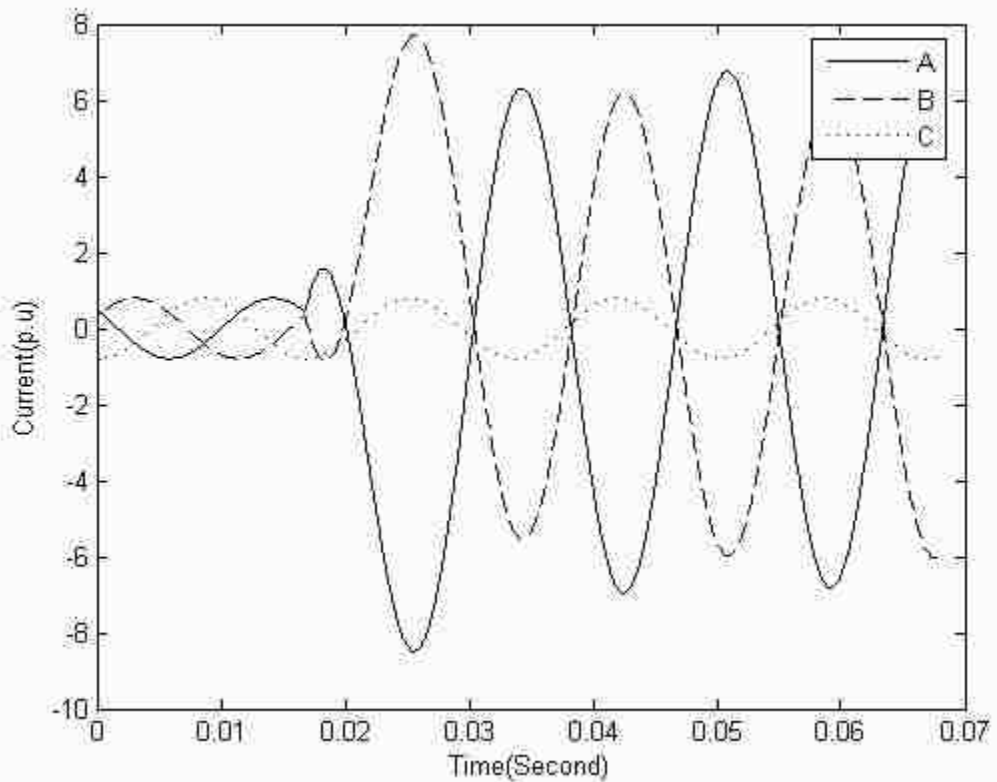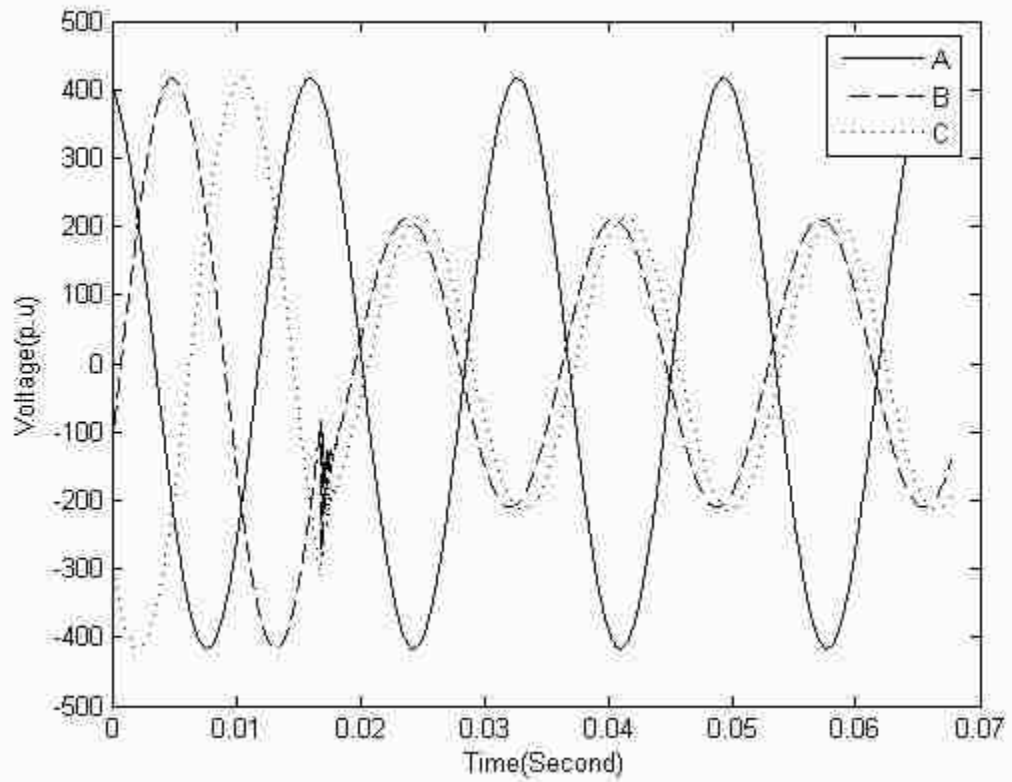
Figure A.7 (a) BC Fault Voltage Waveforms at the Sending End and (b) Receiving End
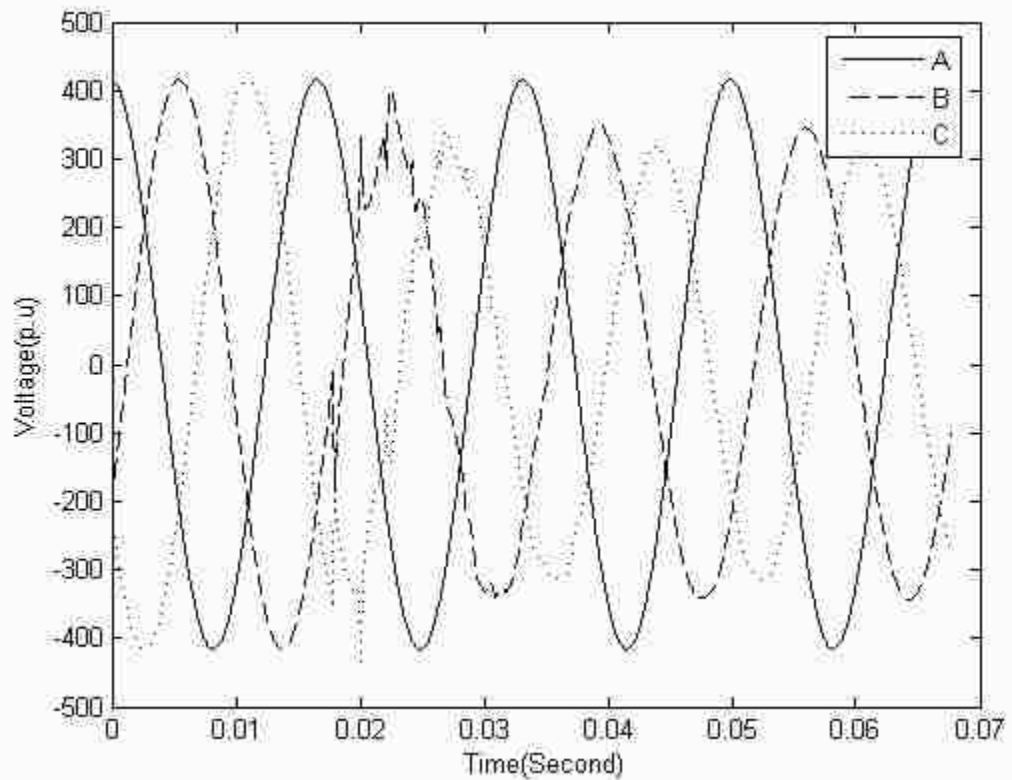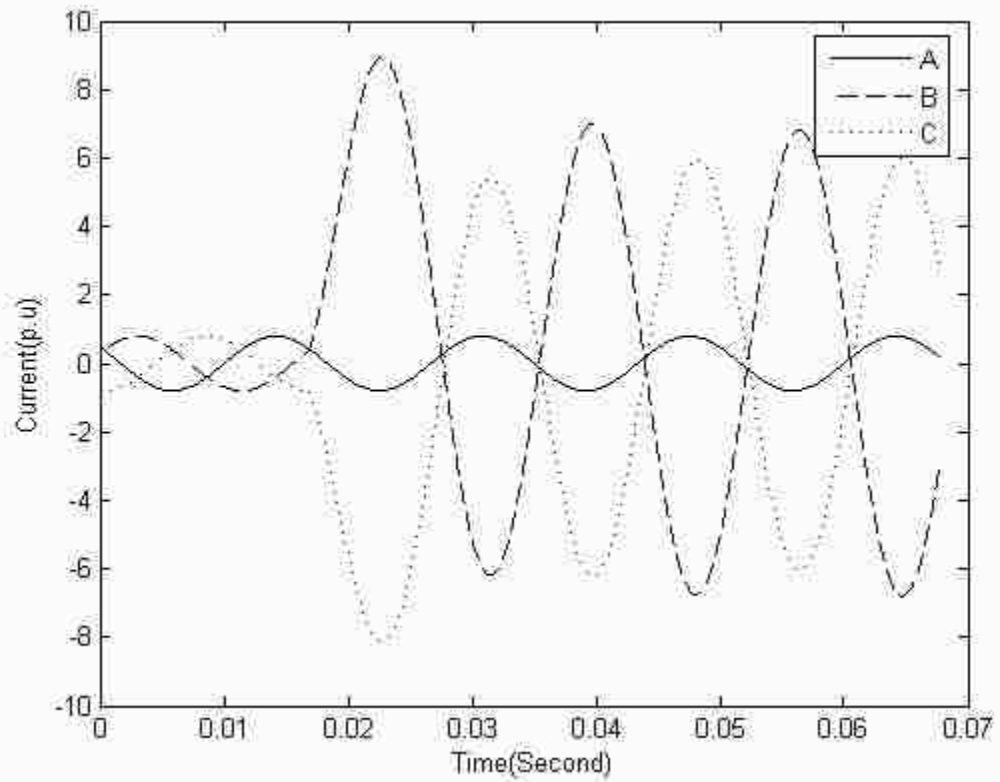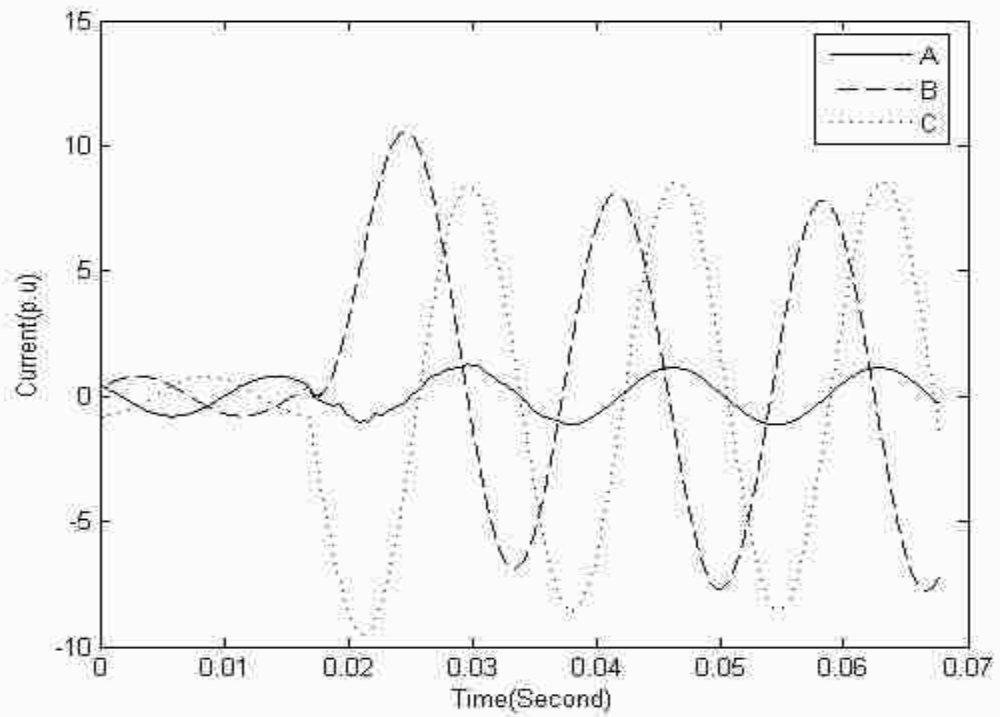
Figure A.8 (a) BC Fault: Current Waveforms at the Sending End and (b) Receiving End
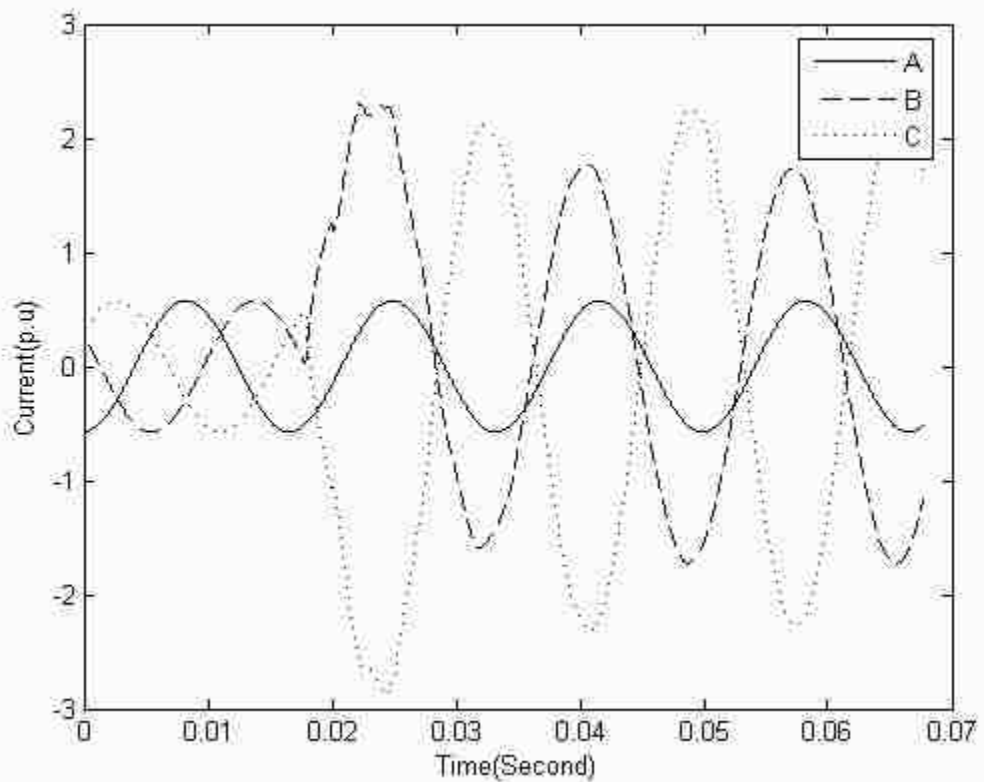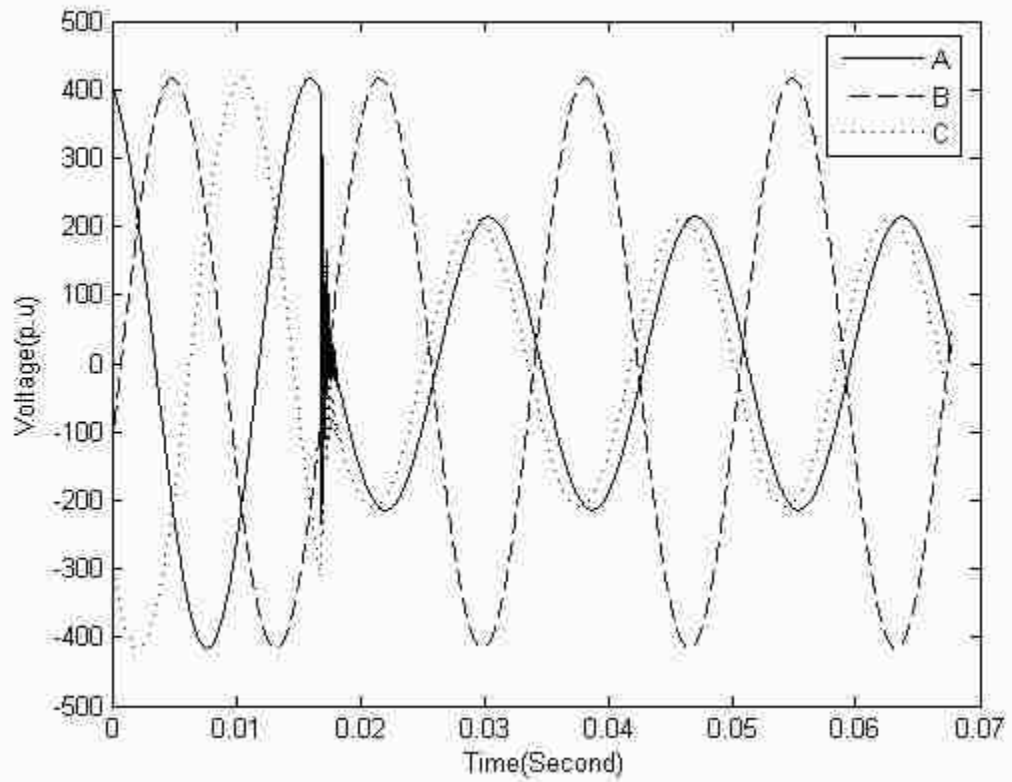
Figure A.9 (a) CA Fault: Voltage Waveforms at the Sending End and (b) Receiving End
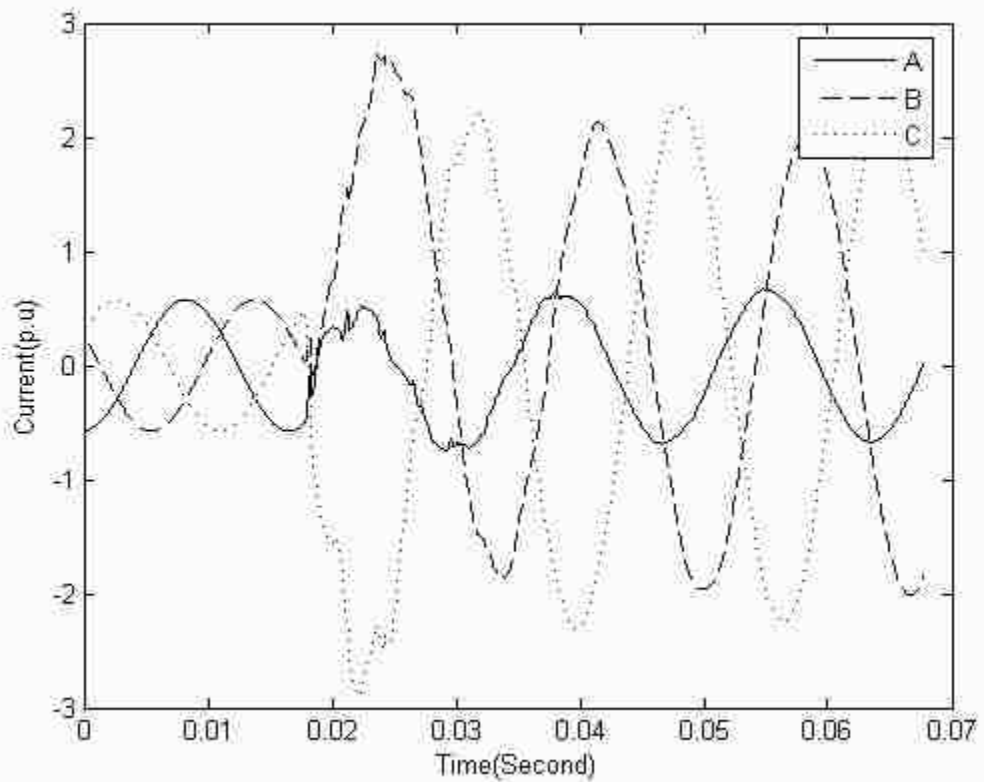
Figure A.10 (a) CA Fault: Current Waveforms at the Sending End and (b) Receiving End

Figure A.11 (a) ABG Fault: Voltage Waveforms at the Sending End and (b) Receiving End

65

Figure A.12 (a) ABG Fault: Current Waveforms at the Sending End and (b) Receiving End

Figure A.13 (a) BCG Fault: Voltage Waveforms at the Sending End and (b) Receiving End

Figure A.14 (a) BCG Fault: Current Waveforms at the Sending End and (b) Receiving End
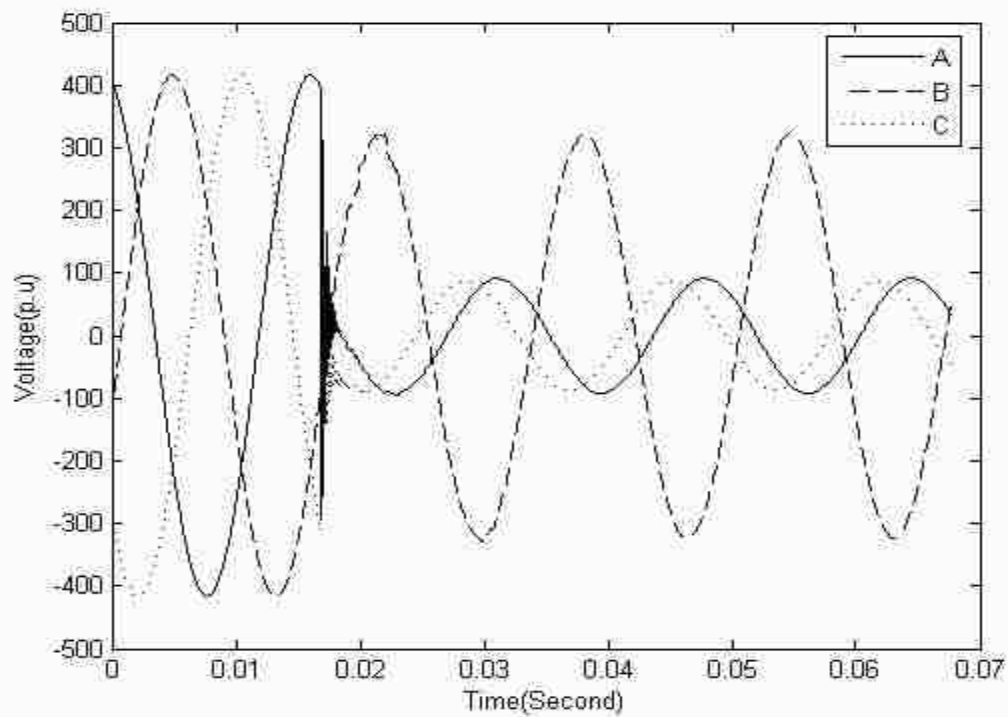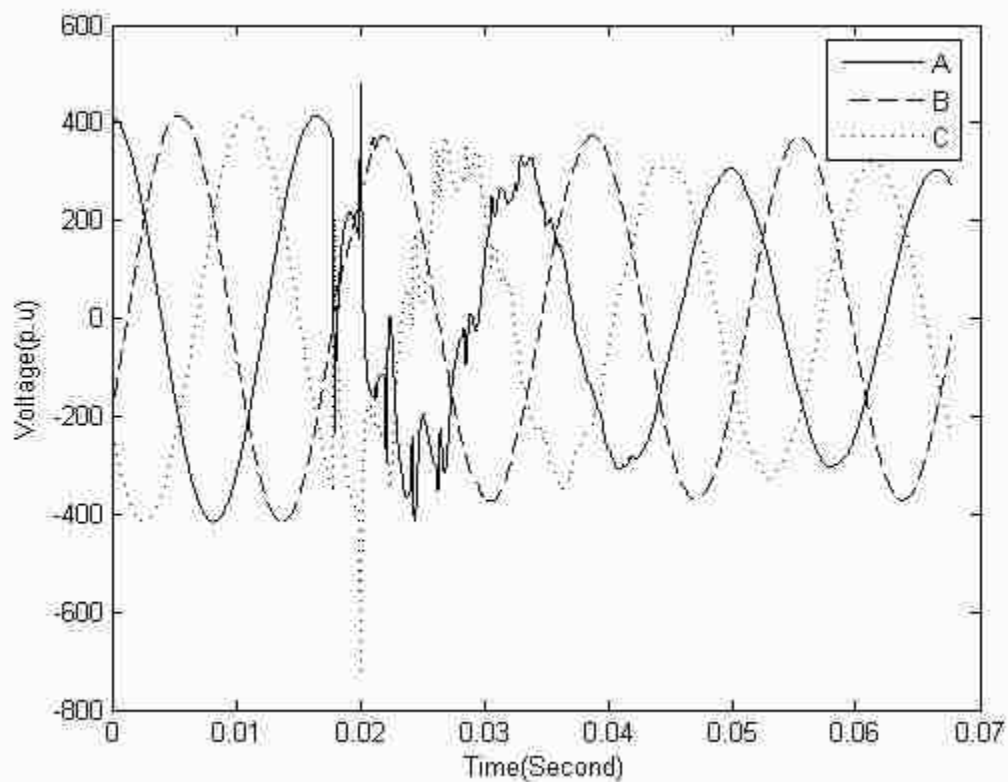
Figure A.15 (a) CAG Fault: Voltage Waveforms at the Sending End and (b) Receiving End
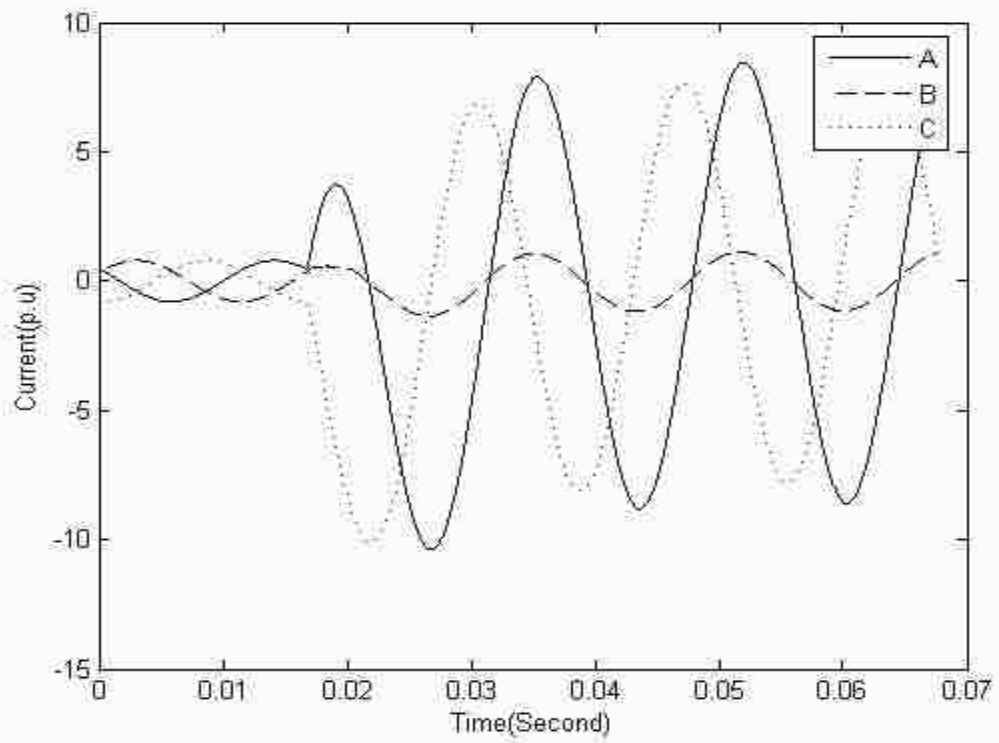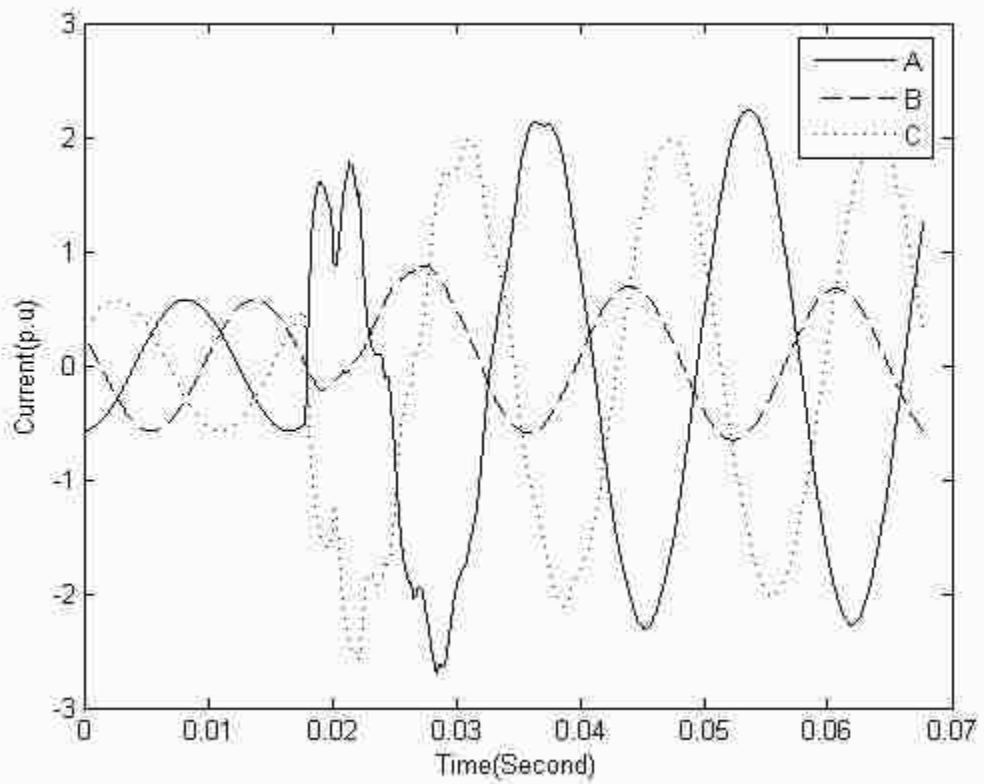
Figure A.16 (a) CAG Fault: Current Waveforms at the Sending End and (b) Receiving End
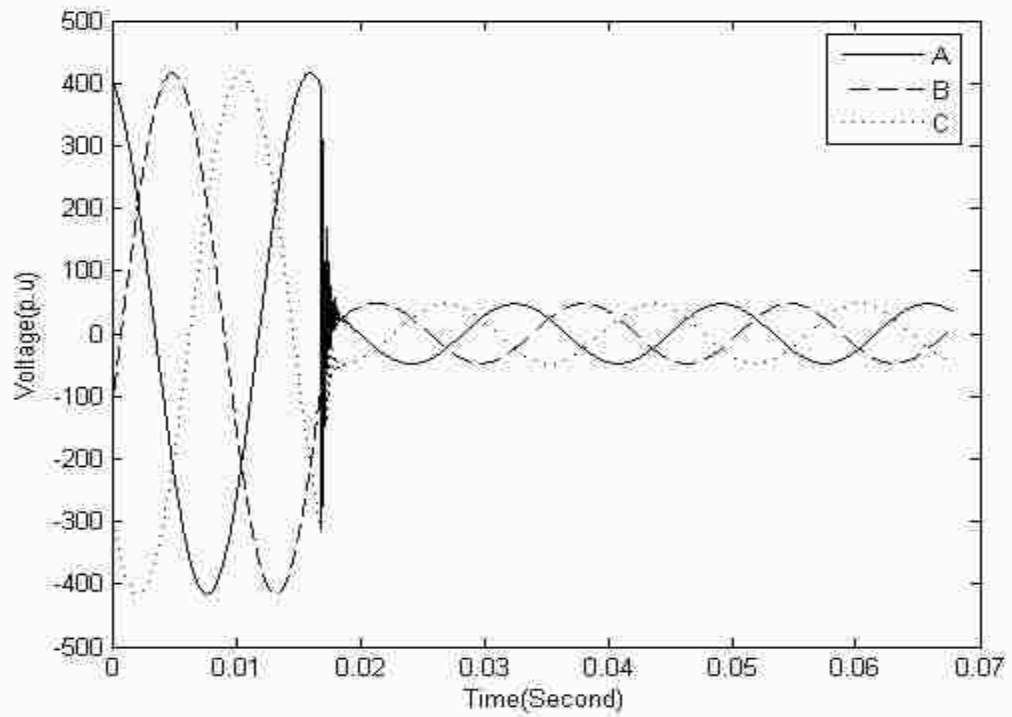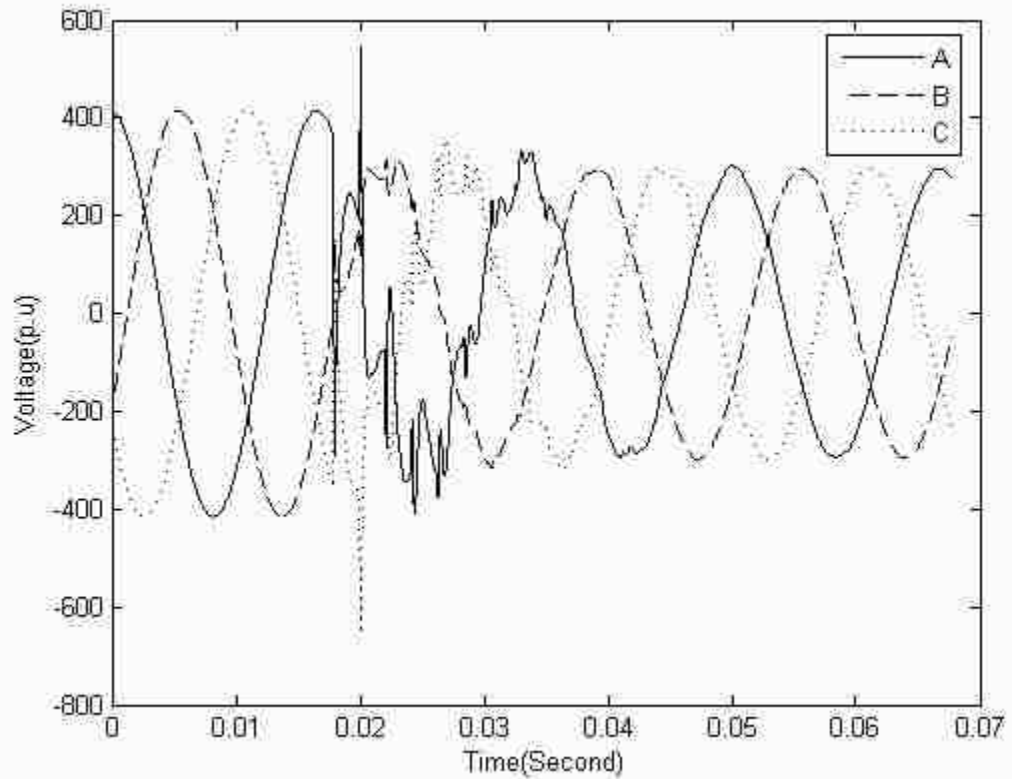
Figure A.17 (a) ABC Fault: Voltage Waveforms at the Sending End and (b) Receiving End
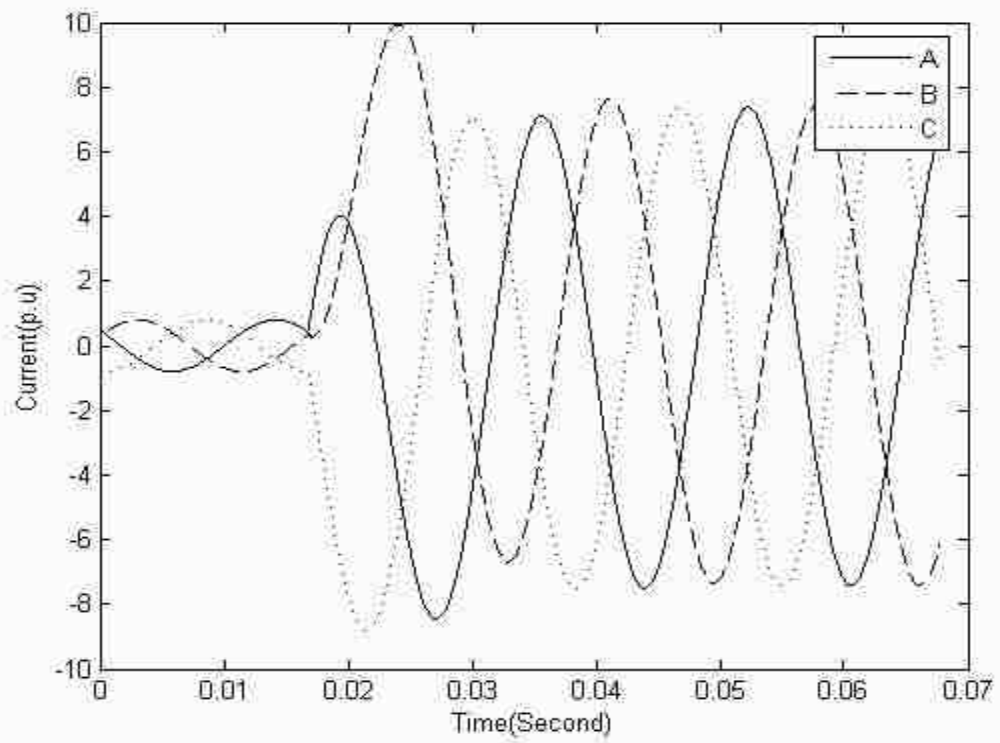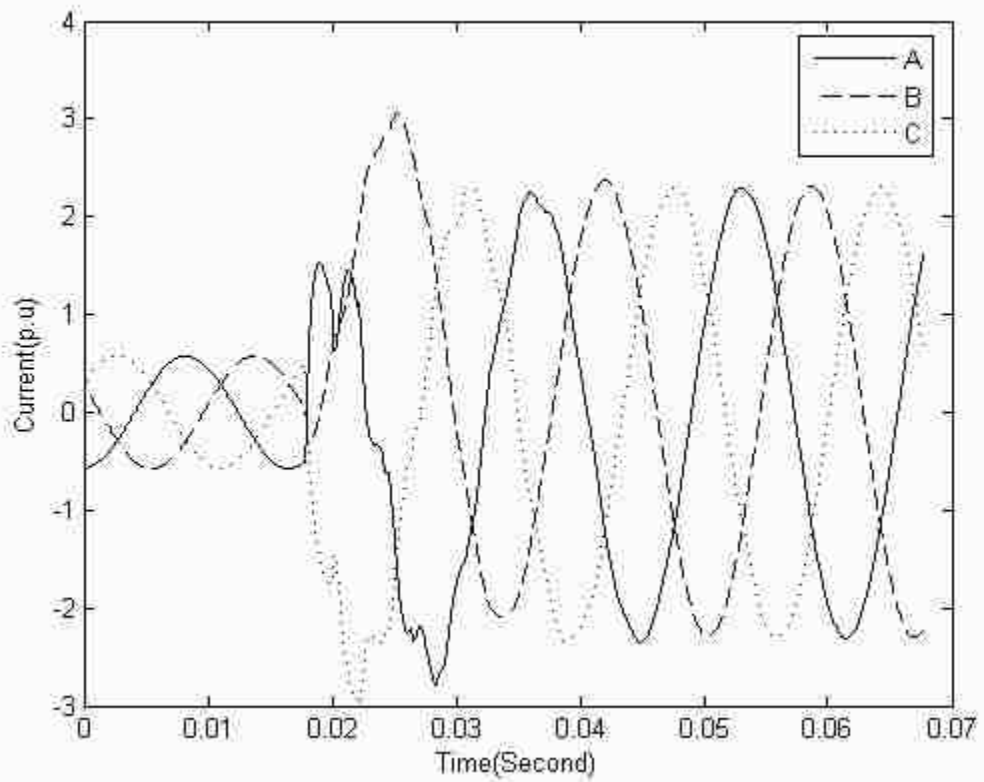
Figure A.18 (a) ABC Fault: Current Waveforms at the Sending End and (b) Receiving End

BIBLIOGRAPHY

[1] J. Lewis Blackburn, Protective Relaying Principles and Applications, Second Edition

[2] R.P. Lipmann, " Pattern Classification using Neural Networks", IEEE Communication Mag. Pp 27-62, Nov 1989.

[3] M.Tarafdar Haque and A.M. Kashtiban. "Applications of Neural Networks in Power Systems; A Review", Transactions on Engineering, Computing and Technology V6 June 2005 ISSN1305-5313.

[4] K.W. Chan, A.R. Edward, A.R. Danish, "On-Line Dynamic Security Contingency Screening Using Artificial Neural Network", IEEE Trans. Power Distribution System, pp.367-372, November 2000.

[5] Y.H. Song, Q.X. Xuan and A.T.Johns, "Comparison studies of five neural network based fault classifiers for complex transmission lines", Electric Power Systems Research, vol 43, no. 2, pp 125-132, November 1997.

[6] W.W.L. Keerthipala, C.T. Wai and W. Huisheng, "Neural Network based classifier for high speed protective relaying", IEEE Trans. Power Delivery vol. 10, no.2, pp 1002-1011, April 2005.

[7] B.Moore, "ART1 and Pattern Clustering", Proc. 1988 Connectionist Models Summer School, Moregan Kaufmann, 1998, pp.174-185.

[8] Robi Polikar, "The Engineer's ultimate guide to Wavelet analysis: The Wavelet Tutorial", http://users.rowan.edu/~polikar/WAVELETS/WTtutorial.html

[9] Thomas Dalstein, Bernd Kuliche, "Neural Network Approach to Fault Classification for High Speed Protective Relaying", IEEE Trans. on Power Delivery, Vol.10, No.2, April 1995.

[10] Philippe Crochat and Daniel Franklin, "Back-Propagation Neural Network Tutorial", Internet tutorial
"http://ieee.uow.edu.au/~daniel/software/libneural/BPN_tutorial/BPN_English/BPN_English/BPN_English.html", April 2003.

[11] R.Aggarwal, Y.Song, "Artificial neural network in power systems, I. General introduction to neural computing", Power Engineering Journal, Vol 11, No.23 pp. 129-134, June 1997.

[12] Biswarup Das and J.Vittal Reddy, "Fuzzy-Logic-Based Fault Classification Scheme for Digital Distance Protection" in IEEE Trans. on Power Delivery, Vol. 20, No.2, April 2005.

[13] Nan Zhang and Mladen Kezunovic, "Coordinating Fuzzy ART Neural Networks to improve transmission line fault detection and classification"

[14] CanAm EMTP User Group, Alternate Transients Program (ATP) Rule Book, Portland 1992.

[15] L.Daubechies, "The wavelet transform, time and frequency localization and signal analysis", IEEE Trans. Information Theory, vol.36, no.5, pp.961-1005, September 1990.

[16] P.Purkait and S.Chakravorti, "Wavelet Transform based Impulse Fault Pattern Recognition in Distribution Transformers", IEEE Transactions on Power Delivery, Vol.18, No.4, October 2003.

[17] Chul-Hwan Kim, Hyun Kim, Young-Hun Ko, Sung-Hyun Byun, Raj K. Aggarwal and Allan T.Johns, "A Novel Fault Detection Technique of High Impedence Arcing Faults in Transmission Lines using Wavelet Transform", IEEE Trans. on Power Delivery, Vol.17, No.4 October 2002.

[18] F.H. Magnago and A. Abur, "Fault Location using wavelets", IEEE Trans. Power Delivey, vol.13, pp. 1475-1480, Oct 1998

[19] Ali-Reza Sedighi, Mahmood-Reza Haghifam, O.P. Malik and Mohammad-Hasan Ghassemian, " High Impedence Fault detection based on Wavelet Transform and Statistical pattern recognition", IEEE Transactions on Power Delivery, Vol.20, No.4, October 2005.

[20] Gail A. Carpenter and Stephen Grossberg, "Adaptive Resonance Theory", The handbook of Brain Theory and Neural Networks, Second edition, Michael A.Arbib, Editor, Cambridge, Massachusetts; MIT Press.

[21] Y.H. Pao, Adaptive Pattern Recognition and Neural Networks, Reading: Wesley 1989. p.309.

[22] Michael J.Healy and Thomas P.Caudell, "A Categorical Semantic Analysis of ART Architectures

[23] M. Kezunovic, M.I. Rikalo and D.Sobajic, "High-speed fault detection and classification with neural nets". Electric Power Systems Research, vol. 34, pp. 109-116, 1995.

[24] G.A. Carpenter, Stephen Grossberg and David B.Rosen, "Fuzzy ART: Fast Stable learning and categorization of analog patterns by an Adaptive Resonance system", Neural Networks, Vol.4, pp.759-771, 1991.

[25] G.A. Carpenter and S. Grossberg, "A Massively Parallel Architecture for Self-Organizing Neural Pattern Recognition Machine," Computer Vision, Graphics and Image Processing, vol.37, 1987, pp.54-115.

[26] Gail A. Carpenter, "Distributed Learning, Recognition, and Prediction by ART and ARTMAP Neural Networks", Neural Networks, Vol.10, No.8, pp.1473-1494, 1997.

[27] S.Vasilic and M.Kezunovic, "An Improved neural network algorithm for classifying the transmission line faults", in Proc. IEEE PES Power Winter Meeting, New York, NY, Jan 2002, vo.2, pp. 918-923.

[28] Gail A. Carpenter and Stephen Grossberg, "ART2: Self-Organization of stable category recognition for analog input patterns", Applied Optics, Vol.26, No.23, December 1987.

[29]T.M. Cover and P.E.Hart, "Nearest Neighbor Pattern Classification", IEEE Transactions on Information theory, Vol. IT-13, No.1 January 1967.

[30] S. Vasilic and M. Kezunovic. "Fuzzy ART Neural Network Algorithm for Classifying the Power System Faults", IEEE Transactions on Power Delivery, pp 1-9, July 2004.

[31] Gail A. Carpenter, Stephen Grossberg and John H.Reynolds, "ARTMAP:Supervised Real-Time Learning and Classification of Nonstationary Data by a Self-Organizing Neural Network", Neural Networks, Vol.4, pp.565-588, 1991

[32] The Mathworks Inc., Using Matlab, Natick, Jul.2002.

VITA

Karthikeyan Kasinathan was born on 17th of November 1983 in Chennai, India to Kasinathan and Chitra Kasinathan. He received his higher secondary school certificate in May 2000. He then joined Sri Venkateswara College of Engineering affiliated to the University of Madras, India in August 2000 and received his Bachelor of Engineering degree in Electrical and Electronics Engineering department in May 2004. In August 2004, he enrolled at the University of Kentucky to pursue graduate studies towards the Masters degree in the Electrical and Computer Engineering.

Karthikeyan Kasinathan

(Date)