S-TRANSFORM AIDED ROUGH SET THEORY BASED IDENTIFICATION AND LOCALIZATION OF SHUNT & SERIES FAULTS ALONG THE TRANSFORMER WINDING

A THESIS SUBMITTED IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF ELECTRICAL ENGINEERING

2016

By

NILADRI RAY CHOUDHURY

EXAMINATION ROLL NO.: M4ELE1619

REGISTRATION NO.: 128912 of 2014-15

JADAVPUR UNIVERSITY

Under the Supervision of

Prof. KESAB BHATTACHARYA

DEPARTMENT OF ELECTRICAL ENGINEERING

FACULTY OF ENGINEERING AND TECHNOLOGY

JADAVPUR UNIVERSITY
KOLKATA-700032

&

Dr. SOVAN DALAI

DEPARTMENT OF ELECTRICAL ENGINEERING

FACULTY OF ENGINEERING AND TECHNOLOGY

JADAVPUR UNIVERSITY

KOLKATA-700032

JADAVPUR UNIVERSITY

Faculty of Engineering and Technology

CERTIFICATE OF RECOMMENDATION

*This is to certify that this dissertation entitled "**S-Transform aided Rough Set Theory based Identification and Localization of Shunt & Series Faults along the Transformer winding**" has been carried out by **NILADRI RAY CHOUDHURY** under our supervision and be accepted in partial fulfillment of the requirement for the degree of Master of Electrical Engineering.*

Prof. Kesab Bhattacharya
Professor
Electrical Engineering Department
Jadavpur University
Kolkata-700032

Dr. Sovan Dalai
Assistant Professor
Electrical Engineering Department
Jadavpur University
Kolkata-700032

Prof. Swapan Kumar Goswami
Head of the Department
Electrical Engineering Department
Jadavpur University
Kolkata-700032

Prof. Sivaji Bandyopadhyay
Dean,
Faculty of Engg. and Technology
Jadavpur University
Kolkata-700032

JADAVPUR UNIVERSITY

Faculty of Engineering and Technology

CERTIFICATE OF APPROVAL

The foregoing thesis is hereby approved as a credible study of an engineering subject, carried out and presented in a manner satisfactory to warrant its acceptance as pre-requisite to the degree for which it has been submitted. It is understood that by this approval the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusion drawn therein but approve the thesis only for which it has been submitted.

_____ _____

Signature of the Examiner(s) *Signature of the Supervisor*

Declaration of Originality and Compliance of Academic Ethics

I hereby declare that this thesis contains literature survey and original research work by the undersigned candidate, as part of the Master in Electrical Engineering studies.

All information in this document has been obtained and presented in accordance with academic rules and ethical conduct.

I also declare that, as required by these rules and conduct. I have fully cited and referenced all material and results that are not original to this work.

Name : **NILADRI RAY CHOUDHURY**

Examination Roll No. : **M4ELE16-19**

Registration No. : **128912 of 2014-15**

Thesis Title : ***S-transform Aided Rough Set Theory based Identification and Localization of Shunt &Series Faults Along the Transformers Winding***

.....

Signature with Date

Acknowledgements

I would like to take this opportunity to humbly express my gratitude for the innumerable gesture of help, cooperation and inspiration that I have received from my teachers, friends and well wishers during this course.

I feel honored to express my profound regard and deep sense of gratitude to my guides Prof. Kesab Bhattacharya and Dr. Sovan Dalai, Electrical Engineering Department, Jadavpur University, Kolkata, for allowing me to do my work in this exciting field. They have been the persons to instill in me a sense of commitment, dedication and optimism. I am highly obliged and grateful for their excellent guidance, endless encouragement, unique cooperation extended to me, right from the time of onset of this colossal task till its successful completion.

I would also wish to express my sincere gratitude to Prof. Abhijit Mukherjee, Prof. Sivaji Chakravorti and Dr. Biswendu Chatterjee, Dept. of Electrical Engg., Jadavpur University for their keen interest and active support in this work.

I would like to thank Prof. Swapan Kumar Goswami, Head of the Department, Electrical Engineering, and Jadavpur University for providing the necessary facilities for carrying out this thesis.

I would like to give a special thanks to Mr. Riddhi Ghosh, Mr. Nasirul Haque, Mr. Soumya Chatterjee and Mr. Suhas Deb for their spontaneous support and inspiration in carrying out this work.

I would like to thank my colleagues Arnab Baug and Amrendra Kumar for their help and endless effort to complete this work during M.E. course.

Words would be trite to express the extent of dependency, love and fortitude towards my parents, without their love and encouragement it would have been impossible to complete this work. Last but not the least; I thank almighty God for all his blessing.

Niladri Ray Choudhury

TABLE OF CONTENTS

Chapter 4: Classification

Chapter 5: Analogue Model and Circuit Parameters

Chapter 6: Results & discussions

Chapter 7: Conclusion

Chapter 1
Introduction

Chapter 1: INTRODUCTION

1.1 Transformers

A transformer is a static electromagnetic device that transfers energy between two or more circuits through the process of electromagnetic induction. Basically, a transformer consists of two windings or more that are wound around a common core to provide strong electromagnetic coupling between the windings. The law of electromagnetic induction states that "whenever there is a relative space or time variation between magnetic field and a conductor, an electromagnetic force will be induced in the respective conductors". The '*emf'* induced in the conductors of a transformer are due relative time variation between the magnetic field and the conductors wound around the core. Due to the strong electromagnetic coupling by the core, the time varying magnetic field that links on winding in the core also links every other winding in the same core with negligible loss. This process induces *'emf'* in every winding in the core, thus facilitating transfer of energy.

Since, their invention in 1855, transformers have become essential transmission, distribution and utilization of alternating current electrical energy. Continuous operation of power transformers is vital for maintaining the overall reliability of the network. Hence, transformers needs critical attention at all times and regular maintenance work need to be carried out for its proper functioning and to have critical knowledge about faults existing and developing within any part of the transformer.

1.2 Transformer testing

Different types of tests are performed on a transformer on a regular basis to assess its condition and performance. Some tests are performed by the manufacturers before delivering the transformer. Some are performed at the consumer site before commissioning and some periodically in regular or emergency basis throughout its service life. The various categories of tests that are performed on a transformer are given as follows:

1. Tests done by manufacturers: These tests are subcategorised as
    i. Type tests
    ii. Routine tests
    iii. Special tests
2. Tests done at site: These tests are again subcategorised as
    i. Pre-commissioning tests
    ii. Periodic/condition monitoring tests
    iii. Emergency tests

The type tests are performed to prove that the transformer meets customer's specifications and design expectations. Some of the type tests performed includes:

a) Transformer winding resistance measurement
b) Transformer ratio test
c) Transformer vector group test
d) Short circuit tests
e) Open circuit tests
f) Measurement of insulation resistance
g) Dielectric tests
h) Temperature rise of transformers
i) Tests on on-load tap changer
j) Vacuum tests on tank and radiators

In the context of this thesis a special type of dielectric test named as impulse test of transformer is dealt with. Impulse tests are different from general power frequency dielectric withstand tests which sometimes may not be sufficient to assess the overall integrity of the transformer. Impulse tests of transformers are subcategorised into: Lightning Impulse tests and Switching Impulse tests amongst which only lightning impulse tests is considered here.

1.3 Impulse testing of transformers

Impulse testing of transformer has been accepted as an important criteria for assessing the basic insulation level (BIL) of the transformer insulation prior to commissioning. Insulation integrity for a transformer is necessary for its proper functioning. Any weakness in the insulation may lead to catastrophic failure of the transformer at a later stage if it goes unnoticed at its inception. In such cases impulse testing of transformer

proves to be an efficient tool for identifying such incipient faults which may not be detected easily with conventional power frequency dielectric withstand test.

Transformer connection procedures, various standard specifications and safety measures to be adopted during the impulse testing of a transformer are presented in various national and international standards such as IS 2026 (Part 3):2009 [1.1] or its corresponding international standard the IEC 60076-3:2000 [1.2]. The specifications are given as follows:

i. Test criteria: The absence of significant differences between voltage and current transients recorded at reduced voltage and those recorded at full test voltage constitutes evidence that the insulation has withstood the test.
ii. Sequence of application of pulse:
iii. Test connections: The impulse test sequence is applied to each of the line terminals of the tested winding in succession. In the case of a three-phase transformer, the other line terminals of the winding shall be earthed directly or through a low impedance, not exceeding the surge impedance of the connected line. If the winding has a neutral terminal, the neutral shall be earthed directly or through a low impedance such as a current measuring shunt. The tank shall be earthed.
iv. Impulse waveform: The lightning impulse waveform applied to the terminals of the transformer should meet the standard waveshape of $1.2\mu s/50\mu s$ with a tolerance of $\pm 20\%$ in the wavefront and $\pm 50\%$ in the wavetail repectively.

1.4 Artificial Intelligence in power system

For a long time, scientists have been keen on creating a substitute that can mimic and perform the same tasks as a human. The arrival of the computer helped scientists progress at great speeds toward a feasible alternative to the human mind. Some of the most important feature that is unique to human mind is process of collecting data, reasoning with the data to acquire significant results and decision making. Academic and Engineering researchers have tried to use these techniques in a wide range of areas such as manufacturing, automation, control, problem identification and diagnostics etc. Application to power system problems is one such area.

Application of intelligent methods in the areas related to power system planning, operation, diagnosis and design has been an area of focus among researchers for quite

some time. Some of the well-established techniques has also been adopted by utilities for problem detection and mitigation. A typical power system network is huge and complex, which makes system analysis, diagnosis and regular maintenance task cumbersome and difficult, when conventional techniques are adopted. Such problems can be eliminated by the use of modern Artificial Intelligence (AI) techniques that can process a bulk amount of data easily, store important data for future references and research purposes if needed and can be suitably adapted according to designers need.

The most widely used AI techniques in power system application is mainly focused on use of various machine learning techniques to detect and classify various power system events with highest possible accuracy. Various classification tools were reported in literature such as Artificial Neural Network (ANN) [1.3] and its other variants [1.4], Fuzzy logic and Fuzzy systems [1.5], Expert Systems [1.6] and Genetic Algorithm [1.7] to name a few. For efficient working of such AI techniques it is important to extract significant features from the available system data. The technique of feature extraction using various signal processing tools is another research area that power system researchers have dwelled for a long time. The key to the task is to extract useful and compact hidden information from heaps of available data acquired through data acquisition systems that could be used for efficient training of the classification tools. Use of appropriate feature extraction method is essential for proper working of the AI tools with sufficient accuracy. Various feature extraction tools cited in literature include signal processing tools like Fourier Transform [1.8], Short-time Fourier Transform [1.9], Hilbert Transform [1.10], Wavelet transforms [1.11], Correlation techniques [1.12] etc. amongst others.

1.5 Scope of the Thesis

The aim of this thesis is to develop an intelligent tools to identify and localise internal short circuit faults from the current waveform data obtained during impulse testing of the transformer. Two Rough set based classifiers, one based on Mathematical Morphological features and another based on Stockwell transform features is developed for the purpose. To train and test the developed algorithm artificial faults are simulated in a winding model of a 3MVA, 33kV, 3-phase delta connected transformer developed at the High Tension Laboratory, Jadavpur University. The winding model was tested using a recurrent surge generator, also developed at the High Tension Laboratory, Jadavpur University and the respective fault current data are acquired using a digital oscilloscope. At last a

comparative study is done as to assess which is a better tool for analysis of the fault current waveforms and thus a better feature extraction mechanism can be developed.

1.6 Outline of the thesis

Chapter 1: This chapter provides an introduction to the thesis by introducing the importance of testing and maintenance of a power transformers in a large power network. Various categories of transformer testing methods are named and one such testing procedure namely impulse testing of transformers is discussed in details. Different IS specifications, connection rules and testing criteria are described. Further importance of application of Artificial Intelligence techniques in areas related to power systems is illustrated. Use of intelligent forms the basis of this thesis as the thesis is focused on use of new signal processing techniques for classifying internal faults and thus localizing them from the obtained data.

Chapter 2: This chapter provides in brief the details about the types of fault considered in the present work along with the specific notations adopted for easy identification of the faults and their locations.

Chapter 3: This chapter discusses in details the two different signal processing tools used in the present work to analyse the different fault current waveforms obtained through experimentations. Along with it this chapter provides a brief introduction to the general signal processing literature with importance to various time domain and time frequency domain representation of signals.

Chapter 4: This chapter gives introduction to the method of classification along with its theoretical and mathematical definitions. Literature review of various classification methodologies used in the domain of electrical systems is provided. Rough set theory is introduced next and a classifier developed based on Rough set theory is presented.

Chapter 5: This chapter gives the details about the constructional details about the analog model used in the present work for experimentation, fault simulation and data acquisition. The distributed parameter model of a transformer winding is shown which forms the basis of model construction. Determination of various lumped electrical parameters and physical parameters are presented. Brief introduction to the recurrent surge generator is given that is used to supply a low voltage impulse to the analog model. At last some of the recorded waveforms obtained during experimentation are shown.

Chapter 6: This chapter provides the results obtained during experimentation and analysis phase. A number of feature used and their variations for different fault cases are presented for both mathematical morphology and Stockwell transform based features. Various classification tables are presented from which rule bases are to be generated for identification of different fault events and their locations. At last a comparison is done over the classification accuracies of Mathematical morphology and Stockwell transform features.

Chapter 7: This chapter contains the conclusion of the thesis.

1.7 References

1.1) "Power Transformers Part - 3 Insulation Levels, Dielectric Tests and External Clearances in Air", Indian Standards, IS 2026 (Part-3): 2009, clause 13.

1.2) "Guide to lighting and switching impulse testing of power transformers and reactors", IEC Standard, Publication 60076-4, 2002.

1.3) E. Hobson, G. N. Allen, "Effectiveness of artificial neural networks for first swing stability determination of practical systems", IEEE Transactions on Power Systems, Vol. 9, 1994, pages: 1062-1068.

1.4) Zwe-Lee Gaing, "Wavelet-based neural network for power disturbance recognition and classification", IEEE Transactions on Power Delivery, Vol. 19, 2004, pages: 1560-1568.

1.5) Q. Su; C. Mi, L. L. Lai, P. Austin, "A fuzzy dissolved gas analysis method for the diagnosis of multiple incipient faults in a transformer", IEEE Transactions on Power Systems, Vol. 15, 2000, pages: 593-598.

1.6) P. Purkait, S. Chakravorti, K. Bhattacharya, "TIFDES - an expert system tool for transformer impulse fault diagnosis", High Voltage Engineering Symposium, 1999.

1.7) D. Rerkpreedapong, A. Hasanovic, A. Feliachi, "Robust load frequency control using genetic algorithms and linear matrix inequalities", IEEE Transactions on Power Systems, Vol. 18, 2003, pages: 855-861.

1.8) M. Roy, J. K. Nelson, R. K. MacCrone, L. S. Schadler, C. W. Reed, R. Keefe, "Polymer nanocomposite dielectrics-the role of the interface", IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 12, 2005, pages: 629-643.

1.9) Y. H. Gu, M. H. J. Bollen, “Time-frequency and time-scale domain analysis of voltage disturbances”, IEEE Transactions on Power Delivery, Vol. 15, 2000, pages: 1279-1284.

1.10) Stuti Shukla, S. Mishra, Bhim Singh, “Empirical-Mode Decomposition With Hilbert Transform for Power-Quality Assessment”, IEEE Transactions on Power Delivery, Vol. 24, 2009, pages: 2159-2165.

1.11) L. Satish, B. I. Gururaj, “Wavelet Analysis for Estimation of Mean-Curve of Impulse Waveforms Superimposed by Noise, Oscillations and Overshoot”, IEEE Transactions on Power Delivery, Vol. 16, 2001, pages: 116-121.

1.12) S. Dalai, B. Chatterjee, D. Dey, S. Chakvorti, K. Bhattacharya, “Rough-Set-Based Feature Selection and Classification for Power Quality Sensing Device Employing Correlation Techniques”, IEEE Sensors Journal, Vol. 13, 2013, pages: 563-573.

Chapter 2

Impulse Faults in Transformers

Chapter 2: Impulse Faults in Transformers

2.1 Introduction

Impulse testing of transformers, as introduced in chapter 1, is performed in accordance to IS 2026-part 3 to assess the basic insulation level of the transformer insulation prior to commissioning. Impulse testing confirms the insulation integrity of the transformer insulation system which is a vital part of the transformer. Any defect in the insulation is reflected as a relative change in the voltage and current waveforms at full and reduced voltages. The procedure of impulse testing is explained in chapter 1. Hence, the aim for impulse testing any transformer is detect anomalies with the insulation system which may have detrimental effects on the life and operation of the unit.

But, this technique of impulse testing may not yield desired results in all cases, particularly when the faults are at their incipient stage and exists only over a negligible portion of the winding. Examples of such cases are short circuit faults existing two winding layers or discs. Such faults may not be significant enough to cause any relative change in voltage and current waveforms at full and reduced voltages. Such fault cases if goes unnoticed may cause serious damage to the transformer insulation at a later stage. Hence, precise identification of such low level fault is of paramount importance. Also with the knowledge of exact location of the fault, time required for troubleshooting reduces drastically which may benefit the electrical supply utility in the long run. Location knowledge of the fault types and their respective approximate locations may also be useful for manufacturing companies for easy and fast troubleshooting of faults.

2.2 Impulse fault categories

Keeping the above mentioned argument in mind short circuit faults that exists in a very small section of the winding is considered for the presented work. Two different types of faults [2.1], [2.2], [2.3] are considered for identification and localization purpose.

a) Series Fault: These are short circuit faults that exists in between any two adjacent discs of the transformer winding. These faults occur when the insulation between the two adjacent discs gets damaged and creates a short circuit path. Such faults

if goes unnoticed may extend over a larger portion of the winding and cause serious damage to the transformer.

b) Shunt Faults: These are short circuit faults that occur between any disc and its adjacent grounded structure. These types of faults are created due damage in the insulation structure that lies in between the disc and mentioned grounded structure. In the present approach the grounded structure in considered as the transformer tank. This choice of ground is evident from the connection diagram adopted for impulse testing of a 3 phase delta connected transformer in which apart from the winding to be tested every other structure is shorted and connected to ground or the tank via a shunt.

2.3 Simulation of Faults:

For simulation of various internal faults along the length of a winding in a transformer only short circuit between any two adjacent disc and short between a disc and tank is considered. The faults are emulated in the analog model, description of which are presented in chapter 5, by physically shorting two adjacent discs for series faults and shorting a disc to the simulated tank for shunt faults, using a small length of wire with negligible resistance. The position or location of the faults are then varied along the length of the winding that provides different fault conditions depending on the position and nature of the fault.

2.3.1 Notations adopted for various fault cases

Suitable notation scheme has been adopted in the present work for identification different fault locations. Such as, Series1 represents series fault at location 1 i.e. fault between disc 1 and disc 2 and Shunt1 represents a shunt fault at section 1 i.e. fault between disc 1 and ground, similarly Shunt27 implies shunt fault at location 27 and so on. Altogether there are 87 series fault cases and 88 shunt fault cases. The identification procedure has been shown in Table 2.1.

Table 2.1: Fault notations

Location of Series Fault	**Corresponding Notation**	**Location of Shunt Fault**	**Corresponding Notation**
Location 1: Between discs 1 and 2	Series1	Location 1: Between disc 1 and ground	Shunt1
Location 2: Between discs 2 and 3	Series2	Location 2: Between disc 2 and ground	Shunt2
. .	. .	. .	. .
Location 87: Between discs 87 and 88	Series87	Location 88: Between disc 88 and ground	Shunt88

2.4 Reference

2.1) P. Purkait and S. Chakravorti, "Pattern Classification of Impulse Faults in Transformers by Wavelet Analysis", IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 9, 2002, pages: 555-561.

2.2) P. Purkait, A. Chatterjee, S. Chakravorti and K. Bhattacharya, "Translationally adaptive fuzzy classifier for transformer impulse fault identification", IEE Proceedings- Generation, Transmission and Distribution, Vol. 150, 2003, pages: 33-40.

2.3) D. Dey, B. Chatterjee, S. Chakravorti and S. Munshi, "Rough-granular Approach for Impulse Fault Classification of Transformers using Cross-wavelet Transform", IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 15, 2008, pages: 1297-1304.

Chapter 3

Signal Processing tools

Chapter 3: Signal Processing tools

3.1 Introduction

Data obtained from field and laboratory experimentations are in general analog (continuous) in nature. These data (signals) are converted to digital format through appropriate data acquisition systems for research and analysis purposes. Again physical data is always in time-domain i.e. the data, or the value of the physical entity, is recorded, varies or is defined with respect to time. This is specifically true for electrical signals where electrical quantities observed, generally voltage and current, are represented in time domain and popularly represented as $v(t)$ and $i(t)$ respectively.

The advantages of a time-domain signal representation diminishes when the problem at hand is to draw inferences from heaps of such similar signals which differ very minutely such as change in the signal frequencies, presence of harmonics, variation of magnitude etc. Analysing a raw signal completely in time domain becomes impossible when dealing with non-stationary signals. Under such circumstances sometimes it is worthwhile to use transformation techniques that shift the domain of the signal from one domain to another where some hidden properties of the signal becomes evident. Such features can then be utilised together with the properties observed in the original domain to analyse the signal, and thus the physical quantity, which it represents.

3.1.1 Literature Review

The first approach into this area was done by Fourier [2.1] by the introduction of Fourier Transform which converted any time domain signal to frequency domain in terms of magnitude and phase of the inherent harmonic components in the signal. Many similar transforms were then introduced later can be found in literature that have such properties and come under the categorical group of integral transforms. As the advances in the basic integral transforms came into the forefront researchers realised that these basic transforms were not as helpful in analysis of a non-stationary signal as they are with stationary signals. It is to be noted that a signal is said to be stationary if its frequency or spectral contents are not changing with respect to time, whereas a non-stationary signal is characterised by varying spectral contents with change in time. In general it is found that

apart from synthetically produced signal or some small number of situations most of the naturally occurring signal is non-stationary in nature.

In order to analyse such signals experts have come up with different approaches, one of which is the windowing technique. In this technique only a small section of the whole signal is analysed by using a short time window to enhance the local properties of the signal at a particular time instant. This special property was utilised in different time-frequency distributions (transforms) such as Short Time Fourier Transform (STFT) [2.2], Gabor Transform [2.3], Wavelet Transform [2.4], Wigner distribution function [2.5], Gabor-Winger Transform [2.6]. These new category of transforms introduced the time-frequency representation of a signal which is basically, a view of the signal in both time and frequency domain simultaneously. As opposed to Fourier Transform which gave information only about the spectral representation of the signal with no time information, these transforms gives the magnitude and phase of the frequency components of a signal along with their occurrence in time.

Apart from the various time-frequency domain representation of signals presented above there are various other mathematical techniques that has been developed for analysis of complex and large data. Some of such well-established techniques are Auto-correlation and Cross-correlation techniques [2.7]. The Hilbert Huang transform [2.8], which also involves integral operations but works on time domain only by modifying the original signal into its corresponding analytical signal.

3.2 Mathematical Morphology

3.2.1 Introduction

Morphology literally means study of shapes. Mathematical Morphology (MM) deals with the mathematical theory of describing shapes and structures using mathematical techniques such as set theory, lattice theory, topology and random functions. Developed by Serra and Matheron [2.9] [2.10], to quantify the mineral characteristics from a thin cross section using binary images, it has since been used in a wide range of applications starting from image processing (extending to grayscale) to being used by power system researches as a non-linear signal processing tool to identify and detect various power system events.

MM deals with the analysis of shape or structure and form of the data presented to it which may be either in the form of image, some signal or a random set of data. MM is different from other integral transform based methods, which are popular in the field of signal

processing, in the sense that all its operations are time domain based. Any operation using MM operators does not transform the domain of the signal, rather it modifies the shape of the data to extract some valuable information or to preserve some valuable information of the signal. MM has been proven to be very useful in de-noising images and signals. Another application of MM includes its applicability to non-periodic transient signals and being able to accurately and reliably extract the signal component without causing any distortion. Moreover the operators used are small in size and fast operating, which could be an advantage in many cases. The operations in MM does not concern with the total length of the data. Rather it deals with small chunks of data, the size of which depends on the type of operation done and the shape and size of the operator used.

The various applications reported in literature on the use of mathematical morphology in the field of power system and signal analysis are presented shortly as follows:

(a) Areas related to power quality event detection and monitoring

Lu *et al.* [2.11] used dyadic multiresolution morphological filters to detect various power disturbances. Li *et al.* [2.12] predicted the starting and ending time of various disturbances. Tingfang *et al.* [2.13] made comparison of different morphological filters for use in online monitoring of power quality signatures, etc..

(b) Transmission Line Protection

Jing *et al.* [2.14] used different units of MMFs and MMGs to filter noise and extract transient component of a signal caused by fault. Zhen Ji, Qiming Zeng, Jianghai Liao and Q. H. Wu [2.15] used morphology for locating fault in a transmission line. Khodadadi and Shahrtash [2.16] used morphological based filters for protection of three-phase transmission lines.

(c) Partial discharge de-noising

M. B. Ashtiani and S. M. Shahrtash [2.17] used morphology as a feature oriented de-noising tool of Partial Discharge signals.

(d) Transformer Inrush detection

T.V. Reddy and J. Jacob [2.18] used morphology to discriminate between transformer inrush current and internal fault currents.

(e) Current Transformer (C.T.) saturation detection

Lin *et al.* [2.19] use repeated MMG filters to identify CT saturation. Lu *et al.* [2.20] use morphological wavelets to detect and compensate CT saturation.

3.2.2. Structuring Element

In basic mathematical morphology, a structuring element (SE) is a shape, used to probe or interact with a given image, with the purpose of drawing conclusions on how this shape fits or misses the shapes in the image. It forms the basis of any morphological filter in processing an image or a signal. In the domain of signal processing a SE can be thought of as a set of data of any arbitrary length and height chosen by the user depending upon the type of signal to be analysed. In general a SE can be thought of as a set which translates over another larger set as a probe very similar to the method of convolution when one signal slides over another while operating on it. The choice of SE used to probe a given set of data is a completely random choice selected by hit and trial method. Various combinations of different SE are to be used and the one which provides the best result among others is to be chosen. One of the most basic characteristics of a SE as seen by various researchers is that with the increase in size of the data to be analysed the length of the SE also increases though not proportionately. Various popular structuring elements found in literature includes flat, linear, square, disk or ball-shaped. It may also include beeline, inclined line, curve, triangle, rotundity or polygon.

3.2.3 Basic Morphological Operators

There could be a number of operations that could be performed on a set of data in MM. The two most primitive operations to be performed are Dilation and Erosion. Every other operator in MM can be derived using these two operations. The Dilation $[d(n)]$ and Erosion $[e(n)]$ of the signal $x(n)$ by a Structuring Element $s(m)$ are denoted as $x \oplus s$ and $x \ominus s$ respectively and are given in Equations (1) and (2).

$$d(n) = \{x(n) \oplus s(m)\}(n)$$
$$= \max\{x(n-m) + s(m);\ 0 \le n-m \le n, m \ge 0\} \dots \dots (3.1)$$

$$e(n) = \{x(n) \ominus s(m)\}(n)$$
$$= \min\{x(n+m) - s(m);\ 0 \le n-m \le n, m \ge 0\} \dots \dots (3.2)$$

Other operations that could be performed are Opening and Closing denoted as $x \circ s$ and $x \cdot s$ respectively, which could be derived from dilation and erosion. Opening is defined as dilation of an eroded signal i.e. $(f \circ g)(n) = ((f \ominus g) \oplus g)(n)$ and Closing operation is defined as erosion of a dilated signal i.e. $(f \cdot g)(n) = ((f \oplus g) \ominus g)(n)$.

3.2.4 Hybrid Morphological Operations

Other similar operations are the white and black top hat transform which are defined as $f(n) - (f \circ g)(n)$ and $(f \cdot g)(n) - f(n)$ respectively and the hit or miss transform. Others include Hit or Miss Transform, Pruning transform, Morphological Skeleton, Filtering by reconstruction, Ultimate erosion and conditional bisectors, Granulometry and Geodesic distance function. Out of all the Hybrid operators mentioned are well established techniques used in the field of image processing and has no direct known application in the domain of signal processing. In the present work the author has used one of the mentioned hybrid operators namely the white and black top hat transform or in general the Top-Hat and their combination thereof. The basic top hat transform is also modified slightly in a sense that the definition of white and black top hat transform used is

$$White\ Top\ Hat\ (WTH) = f_0(n) - (f \circ g) \dots\dots (3.3)$$

$$Black\ Top\ Hat\ (BTH) = (f \cdot g)(n) - f_0(n) \dots\dots (3.4)$$

$$\&\ Top\ Hat\ (TH) = WTH - BTH \dots\dots (3.5)$$

where, $f_0(n)$ is any reference signal as selected by the user. The signal $f_0(n)$ may be the base signal that represents the healthy condition of the equipment that is analysed or in general the ideal case. This signal should be stored at the data base from which new observed data can be compared with. It should be noted that in the actual top hat transforms used in image processing applications, the function $f_0(n)$ is the signal over which the basic morphological transforms are performed.

The terminologies used for the various transforms as mentioned are derived from the corresponding image processing counterpart and has no relation to manipulation of the signals dealt with in electrical systems.

3.2.5 Some Composite Filters in Mathematical Morphology

There are some composite filters, known as Morphological Filters, which can be constructed using the aforementioned basic operations. In order to develop these composite filters one needs to use the basic and hybrid operators as mentioned in various possible combinations to arrive at the desired result. Some of these composite filters as listed in literature and different research papers are: Morphological Median Filters (MMF), which is defined as the mean of Dilation and Erosion operator. The Generalized Multi-resolution Morphological Gradient (GMMG), which is a gradient (difference) operator with Structuring Elements of different sizes and shapes operating over a same signal where

the Structuring Elements modify themselves according at hand. Open-closing Maximal and Close-opening Minimal (OCCO), the mean of opening of a closed signal and the closing of an opened signal. Multi-resolution Morphological Opening Closing (MMOC) is the simple opening and closing with different types of Structuring Elements to enhance different properties of the signal. Multiresolution Morphological Gradient (MMG), which is the same as GMMG but with a fixed number of Structuring Elements as defined by the user, etc. [2.22].

In the present work apart from the top hat transforms mentioned above, OCM2 defined as

$$OCM2 = \frac{dilation(OCM1) + erosion(OCM1)}{2} \dots\dots (3.6)$$

is used for some feature extraction.

3.3 Stockwell Transform

3.3.1 Introduction

Out of all the mentioned time-frequency transforms in section 3.1.1, the Wavelet Transform has time and again proven to be the most efficient and accurate tool for non-stationary signal analysis. Application of wavelet transform for transformer impulse fault diagnostics were reported by various researches in [2.22] [2.23]. But, the wavelet transform has certain limitations. The wavelet transform (WT) analyses the signal in terms of the translations and dilations of a basis mother wavelet function, thereby giving both time and scale information. Thus WT is effective in providing time-localized information but the spectral information is in terms of scales rather than the absolute frequency. Moreover WT in unable to retain any phase information of the original signal. To overcome the aforementioned shortcomings of the wavelet transform, a method based on S-Transform is proposed for analysing fault current data during impulse testing of transformers.

The Stockwell Transform (S-Transform) was introduced by R.G. Stockwell [2.24] to study geophysical data for analysis of non-stationary earthquake signals. Since its proposal it has been used for different areas of application ranging from biomedical applications [2.25] to mechanical gearbox vibration analysis [2.26] and has seen a wide application in areas related to electrical domain. Some of the reported applications in electrical engineering field are mentioned below:

a) Dash *et. al.* analysed power quality events using S-Transform and it modifications [2.27].
b) Dash *et. al.* used S-Transform for protection of parallel transmission lines by identifying the faulty phase out of the three phases which was in turn one of the parallel lines considered [2.28].
c) Srikanth *et. al.* used S-Transform to identify faults in a HVDC system [2.29].
d) Media *et. al.* used S-transform to detect broken bars in an Induction Motor [2.30].
e) Perveen *et. al.* used S-Transform for Fault detection in an offshore windfarm [2.31].
f) Jimenez *et. al.* used S-transform for load monitoring applications [2.32].
g) Bhende *et. al.* applied S-Transform to identify and locate switched capacitors in a power network [2.33].

3.3.2 Theoretical Development

The S-Transform can be thought of as a "phase correction" of continuous wavelet transform (CWT). The basic definition of S-Transform in terms of CWT can be defined in the following manner:

The CWT of a time series $h(t)$ is defined as

$$W(\tau, d) = \int_{-\infty}^{\infty} h(t) w(t - \tau, d) dt \dots\dots (3.7)$$

where $w(t, d)$ is the scaled version of a chosen mother wavelet. The S-Transform of the same time series $h(t)$ can be found by finding CWT with a specific mother wavelet and multiplied by a phase factor of $e^{i2\pi f\tau}$. Thus,

$$S(\tau, f) = e^{i2\pi f\tau} W(\tau, d)|_{with\ specific\ mother\ wavelet} \dots\dots (3.8)$$

The specific mother wavelet is defined as

$$w(t, f) = \frac{|f|}{\sqrt{2\pi}} e^{-\frac{t^2 f^2}{2}} e^{-i2\pi f t} \dots\dots (3.9)$$

S-transform also has close resemblance to the Short Time Fourier Transform (STFT). The approach to arrive at the definition of S-Transform from STFT point of view is presented as follows:

Given a time series $h(t)$, the S-Transform can be found by multiplying $h(t)$ with a Gaussian window $g(t)$ at any time interval $t = \tau$ as

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t) g(\tau - t) e^{-j2\pi f t} dt \ldots\ldots (3.10)$$

where τ and f denote the time instant of spectral localisation and Fourier frequency, respectively and $g(t)$ is the Gaussian window function. The Gaussian window as a function of frequency (f) and time (t) is defined as

$$g(t) = \frac{|f|}{\sqrt{2\pi}} e^{-\frac{t^2 f^2}{2}} \ldots\ldots (3.11)$$

The S-Transform can thus be defined as

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(\tau - t)^2 f^2}{2}} e^{-i2\pi f t} dt \ldots\ldots (3.12)$$

With the given definition of S-Transform it looks as a special case of Short Time Fourier Transform with a Gaussian window.

Thus, it can be seen that the S-Transform has close resemblance to, and takes the advantages of, both CWT and Fourier Transform. It retains the absolute phase information of the original time series so that retrieval of the time series data is possible using inverse transform techniques and also uses a frequency dependent scalable window for enhancing the local spectral characteristics of the signal.

Sometimes for ease of analysis and to take advantage of the FFT algorithm already available in programming languages, the S-Transform is presented as operations on the Fourier spectrum $H(f)$ of the time series $h(t)$ as,

$$S(\tau, f) = \int_{-\infty}^{\infty} H(\alpha + f) e^{-\frac{2\pi^2 \alpha^2}{f^2}} e^{i2\pi\alpha\tau} d\alpha \ ; \ f \neq 0 \ldots\ldots (3.13)$$

Proof:

A simplified approach to arrive at the above result can be derived as follows:

Considering the conventional S-Transform

$$S(\tau, f) = \int_{-\infty}^{\infty} h(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(\tau - t)^2 f^2}{2}} e^{-i2\pi f t} dt \ldots\ldots (3.14)$$

Now, consider $h(t)e^{-i2\pi f t} = x(t) \ and \ g(t) = \frac{|f|}{\sqrt{2\pi}} e^{-\frac{t^2 f^2}{2}}$

Thus we get, $S(\tau, f) = \int_{-\infty}^{\infty} x(t) g(\tau - t) dt$, which is convolution of $x(t)$ & $g(t)$.

i.e. $S(\tau, f) = x(t) * g(t)$;{ where $*$ denotes convolution}… … (3.15)

Now considering the Fourier Transforms of the individual functions, $x(t)$ & $g(t)$, we obtain

$$F\{x(t)\} = F\{h(t)e^{-i2\pi ft}\} = H(\alpha + f)\ if\ F\{h(t)\} = H(\alpha) \dots \dots (3.16).$$

Fourier Transform of the Gaussian window is derived in the next section and is found to be

$$F\{g(t)\} = F\left\{\frac{|f|}{\sqrt{2\pi}} e^{-\frac{t^2 f^2}{2}}\right\} = e^{-\frac{\sigma^2\omega^2}{2}} = e^{-\frac{2\pi^2\alpha^2}{f^2}}; \text{ where } \sigma = \frac{1}{|f|} \dots \dots (3.17)$$

Hence, we get $B(\alpha, f) = F\{x(t)\}.F\{g(t)\} = H(\alpha + f)e^{-\frac{2\pi^2\alpha^2}{f^2}}$ … … (3.18)

Here $B(\alpha, f)$ is the Fourier Transform of the $S(\tau, f)$. Hence, to obtain $S(\tau, f)$ it is required to find the inverse Fourier transform of $B(\alpha, f)$.

$$\therefore S(\tau, f) = \int_{-\infty}^{\infty} B(\alpha, f) e^{i2\pi\alpha\tau}\, d\alpha \dots \dots (3.19)$$

$$i.e.\ S(\tau, f) = \int_{-\infty}^{\infty} H(\alpha + f) e^{-\frac{2\pi^2\alpha^2}{f^2}} e^{i2\pi\alpha\tau} d\alpha \dots \dots (3.20)$$

3.3.3 Fourier Transform of the Gaussian Window

The Gaussian window in time domain is defined as:

$$g(t) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \dots \dots (3.21)$$

Where $\int_{-\infty}^{\infty} g(t)dt = 1$, i.e. the window is normalised. This condition is necessary to relate the S-Transform with the Fourier Transform.

The parameter σ defines the width of the Gaussian window. Stockwell defined σ as:

$$\sigma = \frac{1}{|f|} \dots \dots (3.22)$$

To obtain the Fourier Transform of the Gaussian window,

Differentiating both sides,

$$\frac{dg(t)}{dt} = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{t^2}{2\sigma^2}} \times \frac{d}{dt}\left(-\frac{t^2}{2\sigma^2}\right) \dots \dots (3.23)$$

$$i.e.\frac{dg(t)}{dt} = -\frac{t}{\sigma^2}g(t) \dots\dots (3.24)$$

Applying Fourier Transform to both sides yields,

$$i\omega G(\omega) = \frac{1}{i\sigma^2}\frac{dG(\omega)}{d\omega} \dots\dots (3.25)$$

$$\therefore \frac{\frac{dG(\omega)}{d\omega}}{G(\omega)} = -\omega\sigma^2 \dots\dots (3.26)$$

Now, integrating both sides

$$\int_0^{\omega} \frac{\frac{dG(\omega')}{d\omega'}}{G(\omega')} d\omega' = -\int_0^{\omega} \omega'\,\sigma^2 d\omega' \dots\dots (3.26)$$

$$i.e.\ ln\,G(\omega) - ln\,G(0) = -\frac{\sigma^2\omega^2}{2} \dots\dots (3.27)$$

As the Gaussian window is normalised, hence $G(0) = 0$. thus,

$$ln\,G(\omega) = -\frac{\sigma^2\omega^2}{2} \dots\dots (3.28)$$

Taking exponentials on both sides

$$e^{\ln G(\omega)} = e^{-\frac{\sigma^2\omega^2}{2}} \dots\dots (3.29)$$

Hence we get,

$$G(\omega) = e^{-\frac{\sigma^2\omega^2}{2}} \dots\dots (3.30)$$

This gives us the Fourier Transform expression of the Gaussian window that could be used in calculating the S-Transform using equation 3.12.

3.3.4 Summary of Stockwell Transform

The steps to obtain the S-Transform for a given time series $h(t)$ would be,

i. Find $H(\alpha) \xrightarrow{fft} h(t)$.

ii. Find $G(\alpha, f) \xrightarrow{fft} g(t, \sigma)$.

iii. Shift $H(\alpha)$ to $H(\alpha + f)$.

iv. Multiply $G(\alpha, f)$ with the shifted $H(\alpha)$.

v. Take the inverse Fourier Transform.

3.3.5 Discrete Stockwell Transform

S-Transform of a discrete time series $h[kT]$ can be defined with help of its continuous analogue equation _, by replacing $f \rightarrow \frac{n}{NT}$ & $\tau \rightarrow jT$. Thus we get,

$$S\left[jT, \frac{n}{NT}\right] = \sum_{m=0}^{N-1} H\left[\frac{m+n}{NT}\right] G(m,n) e^{\frac{i2\pi mj}{N}} \dots\dots (3.31)$$

where,

$$G(m,n) = e^{-\frac{2\pi^2 m^2}{n^2}} \dots\dots (3.32)$$

3.3.6 Generalised Stockwell Transform

The basic definition of S-Transform as laid down by Stockwell placed some restrictions on the properties (shape and size) of the Gaussian window to be used. This was considered as a disadvantage of the S-Transform in comparison to CWT which had a scalable window function. McFadden [2.26] and later Mansinha [2.34] introduced the concept of Generalised S-Transform that had an adjustable Gaussian window. They also introduced a concept that used a window other than Gaussian which too was adjustable. The generalised S-Transform expression according to Mansinha is given as:

$$S(\tau, f, \boldsymbol{p}) = \int_{-\infty}^{\infty} h(t) w(\tau - t, f, \boldsymbol{p}) e^{-i2\pi f t} dt \dots\dots (3.33)$$

In this approach the parameter $\boldsymbol{p}$ controls the shape and size of the window function w apart from the frequency f and its position τ. The only consideration that should be kept in mind while deciding for different values of the parameter $\boldsymbol{p}$ is that it should follow the normalization condition as followed by the Gaussian window proposed by Stockwell i.e.

$$\int_{-\infty}^{\infty} w(\tau - t, f, \boldsymbol{p}) dt = 1 \dots\dots (3.34)$$

As already stated earlier this condition help in maintaining a concurrent relation between S-Transform and Fourier Transform. An example of such a modified Gaussian window that is used in the presented work is:

$$G(m,n) = e^{-\frac{2\pi^2 m^2 F}{a+(b\times n^c)^2}} \dots\dots (3.35)$$

where F, a, b and c are arbitrary constants that alter the shape of the Gaussian window when set to different values.

3.4 Reference

2.1) Fourier, J. B. Joseph (1822), Théorie Analytique de la Chaleur, Paris: Chez Firmin Didot, père et fils.

2.2) E. Al-Ammar, G. G. Karady, "Transfer Function Analysis Using STFT for Improvement of the Fault Detection Sensitivity in Transformer Impulse Test", IEEE Power Engineering Society General Meeting, Vol. 2, 2005, pages: 1855-1862.

2.3) Jie Yao, P. Krolak, C. Steele, "The generalized Gabor transform", IEEE Transactions on Image Processing, Vol.4, 1995, pages: 978-988.

2.4) S. Mallat and W. L. Hwang, "Singularity Detection and Procearing with Wavelets", IEEE Transaction on Information Theory, Vol. 38, 1992, pages: 617-643.

2.5) K. H. Brenner, A. W. Lohmann, "Wigner distribution function display of complex 1D signals", Elsevier International Journal of Optics Communication, Vol.42, 1982, pages: 310-314.

2.6) N. Ali Khan, M. Noman Jaffri; S.I. Shah, " Modified Gabor Wigner Transform for Crisp Time Frequency Representation", International Conference on Signal Acquisition and Processing, 2009, pages: 119-122.

2.7) A. Kumar, Y. Bar-Shalom, "Time-domain analysis of cross correlation for time delay estimation with an autocorrelated signal", IEEE Transactions on Signal Processing, Vol.47, 1993, pages: 1664-1888.

2.8) N.E. Huang, "Hilbert-Huang transform and its applications", Interdisciplinary Mathematical Sciences, Vol. 16.

2.9) J. Serra, Image Analysis and Mathematical Morphology, New York: Academic, 1982.

2.10) G. Matheron, Random Sets and Integral Geometry, New York: Wiley, 1975.

2.11) Z. Lu, D.R. Turner, Q.H. Wu, J. Fitch and S. Mann, “Morphological Transform for Detection of Power Quality Disturbances,” International Conference on Proceedings Power System Technology, 2004 pages: 1644-1649.

2.12) G. Li, M. Zhou, Y. Luo, and Y. Ni, “Power Quality Disturbance Detection Based on Mathematical Morphology and Fractal Technique,” IEEE/PES Transmission and Distribution Conference and Exhibition: Asia and Pacific, 2005.

2.13) Y. Tingfang, L. Pei, Z. Xiangjun and K.K.Li, Application of Adaptive Generalized Morphological Filter in Disturbance Identification for Power System Signatures", International Conference on Power System Technology, 2006.

2.14) M. Jing, W. Zengping, X. Yan, and W. Jie, “A Novel Adaptive Algorithm to Identify Inrush Using Mathematical Morphology,” Conference and Exposition in Proceedings of Power Systems, 2006, pages: 1020-1028.

2.15) Z. Ji, Q. Zeng, J. Liao and Q.H. Wu, "A Novel Mathematical Morphology Filter For the Accurate Fault Location in Power Transmission Lines", TENCON, 2009.

2.16) M. Khodadadi and S. M. Shahrtash, "A New Noncommunication-Based Protection Scheme for Three-Terminal Transmission Lines Employing Mathematical Morphology-Based Filters", IEEE Transactions on Power Delivery, Vol.28, 2103, pages: 347-356.

2.17) M.B Ashtiani and S. Mohammad Shahrtash, "Feature-oriented De-noising of Partial Discharge Signals Employing Mathematical Morphology Filters", IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 19, 2102, pages: 2128-2136.

2.18) T.V. Reddy and J. Jacob, "Mathematical morphology for discrimination between internal faults and inrush currents in power transformers", Journals of Electrical Systems, Vol. 4.1, 2008, pages: 24-35.

2.19) X. Lin, L. Zou, Q. Tian, H. Weng, and P. Liu “A series multiresolution morphological gradient-based criterion to identify CT saturation,” IEEE Transactions on Power Delivery, Vol. 21, 2006, pages: 1169-1175.

2.20) Z. Lu, J. S. Smith and Q. H. Wu, “Morphological Lifting Scheme for Current Transformer Saturation Detection and Compensation,” IEEE Transactions on Circuits and Systems – to appear.

2.21) S. Gautam and S. M. Brahma, “Overview of Mathematical Morphology in Power Systems – A Tutorial Approach”, 2009.

2.22) S.K. Pandey and L. Satish, “Multiresolution Signal Decomposition: A New Tool for Fault Detection in Power Transformer Impulse tests”, IEEE Transactions on Power Delivery, Vol. 13, 1998, pages: 1994-1200.

2.23) M.R. Rao and B.P. Singh, “Detection and Localization of Interturn Fault in the HV Winding of a Power Transformer Using Wavelets”, IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 8, 2001, pages: 652-657.

2.24) R.G. Stockwell, L. Mansinha and R.P. Lowe, “Localization of the complex spectrum: the S transform”, IEEE Transactions on Signal Processing, Vol. 44, 1996, pages: 998-1001.

2.25) K.K. Poh and P. Marziliano, "Analysis of Neonatal EEG Signals using Stockwell Transform", 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2007, pages: 594-597.

2.26) P.D. McFadden, J. G. Cook AND L. M. Forster, "Decomposition of Gear Vibration Signals by the Generalised S Transform", Mechanical Systems and Signal Processing, 1999, pages: 691-707.

2.27) P. K. Dash, B. K. Panigrahi, and G. Panda, "Power Quality Analysis Using S–Transform", IEEE Transactions On Power Delivery, Vol.18, 2003, pages:406-411.

2.28) P.K. Dash, S.R. Samantaray, G. Panda and B.K. Panigrahi, "Time–frequency transform approach for protection of parallel transmission lines", IET Generation Transmission and Distribution, Vol. 1, 2007.

2.29) P. Srikanth, A. K. Chandel and K. A. Naik, "HVDC system fault identification using S-transform approach", International Conference on Power, Control and Embedded Systems (ICPCES), 2010, pages: 1-6.

2.30) M. Media, F. Martin and A. Craciunescu, "Induction Motor’s Broken Bars Detection by using Stockwell Transform", International Symposium on Power Electronics, Electrical Drives, Automation and Motion, 2012.

2.31) R. Perveen, N. Kishor and Soumya R. Mohanty, "Fault detection for offshore wind farm connected to onshore grid via voltage source converter-high voltage direct current", IET Generation, Transmission & Distribution, Research Article-2015, pages: 2544-2554.

2.32) Y. Jimenez; C. Duarte; J. Petit and G. Carrillo, "Feature extraction for nonintrusive load monitoring based on S-Transform", Power System Conference (PSC), 2014, pages: 1-5.

2.33) C. N. Bhende, "S-Transform Technique for Identifying the Location of a Switched Capacitor", 10th International Conference on Environment and Electrical Engineering (EEEIC), 2011, pages: 1-4.

2.34) C. R. Pinnegar And L. Mansinha, "The Bi-Gaussian S-Transform", Siam Journal On Scientific Computing, Vol. 24, Pages: 1678-1692.

Chapter 4
Classification

Chapter 4: Classification

4.1 Introduction

In the field of machine learning, classification refers to the approach of identifying the class to which a new observation belongs. Machine Learning by itself is one of the most widely researched area in the field of Artificial Intelligence (AI). The domain of Machine Learning can be thought of as an amalgamation of various approaches to perform some basic Artificial Intelligence tasks. Out of the various tasks that an AI system is deemed to perform the most essential ones can be narrowed down to pattern recognition and computational learning. These two features form the basis of Machine Learning. In 1959, Arthur Samuel, a pioneer of Machine Learning defined machine learning as a "Field of study that gives computers the ability to learn without being explicitly programmed". Machine Learning study deals with computational algorithms that equips any machine the ability to learn from pre-existing data and make predictions on new data.

The range of problems that different Machine Learning algorithms are designated to handle are in general broadly categorised into Supervised, Semi-supervised, Unsupervised and Reinforcement learning. The problem of classification is one of the sub-categories of the supervised learning technique which also includes regression analysis.

4.2 The method of classification

As said in section 3.1 the working principle of classification tools is to identify the class to which a new observation belongs. To understand the general classification procedure one needs to understand the theoretical aspect behind any supervised learning technique. A brief theory is presented as follows:

Supervised learning techniques needs two kinds of data. One type of data known as the training set or training data belongs to that category for which detailed information is available beforehand. Such data should be present with the algorithm developer before the beginning the tool is designed and based on which new data are to be analysed. The new set of observations or data that the classification tool is expected to identify based on the observation on the training set is called the test data.

In order to categorise the training set into different categories and keep track of the new testing set the data is divided into various subsections named class with each class pertaining to certain unique property. Any new testing data belongs to a predefined class if and only if it possess the unique characteristics that defines that class. To identify a particular class on machine language, each class is assigned a label. A model is prepared through a training process where it is required to make predictions and is corrected when those predictions are wrong. The training process continues until the model achieves a desired level of accuracy on the training data. The accuracy class of any developed classification algorithm can be defined as:

$$\%\ \text{classification accuracy} = \frac{\text{number of correctly identified new observations(test sets)}}{\text{total number of new observations(test set)}} \times 100 \ldots\ldots (4.1)$$

From the viewpoint of pattern recognition of which classification is a subdivision a classification algorithm some sort of output value to a given input value. In this case it would be the labels assigned to a class.

4.3 Mathematical definition of Classification

The working methodology of supervised learning technique can be understood more clearly when mathematical concepts are introduced. The aim is simply to map $x\ to\ y$, where $y_i \in \mathcal{Y}$ the labels of the examples are x_i. Notice that (x_i, y_i) forms the training set. If the labels are numbers, then $y = (y_i)_{i\in[n]}^{T}$ where, $[n] \Longrightarrow (1,2, \ldots, n)$, denotes the column vector of labels. Again, a standard requirement is that the pairs (x_i, y_i) are sampled independently and identically distributed from some distribution which here ranges over $\mathcal{X} \times \mathcal{Y}$. The task is well defined, since a mapping can be evaluated through its predictive performance on test examples. When $y \in \mathbb{R}$ or $y \in \mathbb{R}^d$ (or more generally, when the labels are continuous), the task is called regression [3.1].

4.4 Classification Tools

There are numerous types of classification tools that could be found in literature. The algorithms developed to implement the working principle of the theoretical mechanism is called a classifier. The list of classifier available in conventional literature are:

1) Linear Classifiers: This can be considered as a family of different sub-algorithms which are a) Fisher's linear discriminant b) Logistic Regression c) Naïve Bayes classifier and d) Perceptron. 2) Support vector machines. 3) Quadratic Classifiers. 4) Kernel estimation: k-

nearest neighbour. 5) Boosting (meta-algorithm). 6) Decision trees. 7) Neural networks. 8) Learning vector quantisation etc.

4.5 Literature Review

In this section are mentioned some of the popular classifier tools that has found wide use in the electrical system applications.

i. Strangas *et. al.* analysed fault diagnostics procedure of a permanent magnet d.c. motor using linear discrimination classifier [3.2].

ii. Chen *et. al.* studied electrical pricing variation to fundamental and strategic factors using logistic regression technique [3.3].

iii. Elnady *et. al.* used perceptron based algorithm as tool for mitigation of voltage disturbances [3.4].

iv. Koley *et al.* used SVM based classifier for impulse fault identification [3.5].

v. Gerek *et al.* used quadratic classifiers and higher order cumulants to analyse power quality events [3.6].

vi. Jarrar *et al.* used k-NN along with neural network and linear classifier to assess the hydrophobicity class of a silicone rubber material [3.7].

vii. Rovnyak *et al.* used decision tree for real time transient stability prediction [3.8].

viii. Dey *et al.* studied impulse fault diagnostics in power transformers using learning vector quantisation [3.9].

In the context of this thesis, a classification approach based on Rough Set Theory is proposed to identify fault types and location of these faults along the length of a transformer winding, on which impulse voltage is applied. The basic theoretical aspects behind the Rough Set based classifier is described in this chapter.

4.6 Rough Set Theory

4.6.1 Introduction

The foundation of Rough Set Theory (RST) were first laid down by Zdzislaw Pawlak [3.10] in the year 1982 as a new mathematical approach to tackle imperfect data. The approach is fundamentally based on set theory, i.e. use of sets to study vague data. The first approach in this aspect is dealt using the fuzzy sets that uses a fuzzy membership of its entries using membership function rather than crisp membership. By doing so, the

uncertainty about the data can be taken into account. Rough sets are a simple modification of this concept. It is one of the most important and powerful mathematical tools for feature classification where data table contains vague data, i.e. data regarding which no complete knowledge is available. Recently the use of Rough set theory has seen an increase in demand among AI researchers with applications ranging from text classification, medical applications, image classification etc. In applications to electrical systems RST has been used in identification of defective insulator from noisy Partial Discharge pulses, fraud detection in electrical power consumption, in condition monitoring of distribution feeder, data mining tool for semiconductor manufacturing and also for power generation applications [3.11].

The increase in demand of use of Rough sets stems from its significance in areas related to machine learning, knowledge acquisition, pattern recognition, knowledge recovery and decision analysis. As reported by Pawlak in the main advantages Rough set technique provides over other popular intelligent tools are that it does not need any preliminary or additional information about data, basic probability assignment or grade of membership or value. The last mentioned advantage makes rough sets somewhat superior to corresponding fuzzy sets technique. In general rough set approach:

- provides efficient algorithms for finding hidden patterns in data,
- finds minimal sets of data (data reduction),
- evaluates significance of data,
- generates sets of decision rules from data,
- is easy to understand,
- offers straightforward interpretation of obtained results,
- are particularly suited for parallel processing (most algorithms based on the rough set theory).

4.6.2 Theoretical aspect

RST as stated earlier can be considered as a modification of classical set theory to introduce the concept of vagueness to original set theory idea, which is very important to handle imprecise and superfluous data. Imprecision in RST is expressed by a boundary region of a set. The data to be classified using RST is presented in the form of tables. RST sorts the imprecise data in a boundary region through approximations i.e. lower approximation and upper approximations that makes the set rough, thereby granulating

the data. According to definition: The lower approximation of a set X with respect to R is the set of all objects, which can be for certain classified as X with respect to R (are certainly X with respect to R). The upper approximation of a set X with respect to R is the set of all objects which can be possibly classified as X with respect to R (are possibly X in view of R). And the boundary region of a set X with respect to R is the set of all objects, which can be classified neither as X nor as $notX$ with respect to R. Thus a set X is crisp or exact with respect to R if the boundary region of X is empty and is rough or inexact with respect to R if the boundary region of X is non-empty. Mathematically, the lower approximation of X is defined as

$$\underline{B}(Y) = \bigcup_{y \in U} \{B(y) : B(y) \subseteq Y\} \dots \dots (4.2)$$

The upper approximation is defined as

$$\overline{B}(Y) = \bigcup_{y \in U} \{B(y) : B(y) \cap Y \neq \emptyset\} \dots \dots (4.3)$$

The R-boundary region of X is defined as

$$Boundary(Y) = \overline{B}(Y) - \underline{B}(Y) \dots \dots (4.4)$$

The graphical representation of the mentioned approximation details is given in Figure 4.1.

Hence, approximations are expressed as granule knowledge. Each granule (subset of data) is indiscernible i.e. they contain the same information. In the process of classification, the aim would be to give each of the granulated subset a proper class or identification that serves the purpose of use i.e. classification of vague data in this case. Each granule would be precise and the union of such granulated data would be crisp. The lower approximation of a set is union of all granules which are entirely included in the set; the upper approx.-

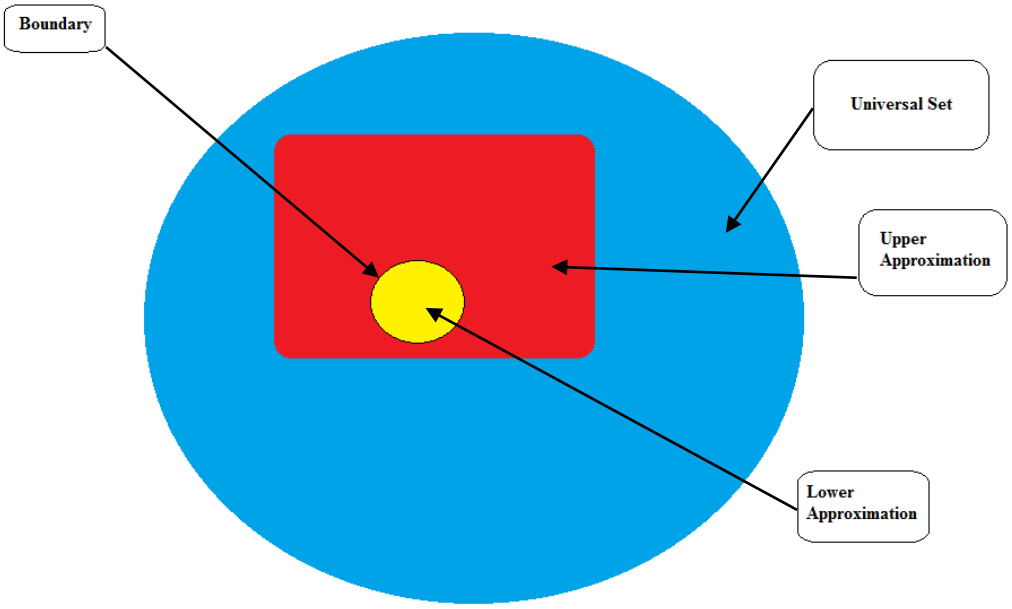


Figure 4.1: Sets representing upper and lower approximations along with boundary conditions

-imation is union of all granules which have non-empty intersection with the set; the boundary region of set is the difference between the upper and the lower approximation.

4.6.3 Rough Set based classifier

The data to be classified is arranged in form of a data table. Each row of the data table presented to RST is considered as object and each column as attribute, whereas the individual entries of the table are named as attribute values. The attributes are of two types namely, condition attributes and decision attributes. The normalised data table is then constructed for simplification of the table. In the table the objects are related to different instances or cases to be classified and the columns or attributes are special properties related to a particular object. The condition attributes are generally different features that has been extracted from the different signal processing tools as discussed in the previous chapter. As shown next, feature generation can be considered as one of the most important tool that simplifies the classifier operation through dimension redundancy reduction.

4.6.4 Importance of Pre-processing of data: Feature Extraction

One of the important feature for which signal processing tools are valuable to electrical system researchers when working in the domain of Machine learning algorithms is the property of dimensionality reduction. It can be defined as the process of mapping a large set of input data into a lower dimensional space. Witout the use of any pre-processing tool the size of data that the classifier has to handle could be huge particularly when operating with practical field data. On top of that it has to handle a number of such

different types of data with their associated variables in order to understand, draw inference and classify the data. This may require larger memory for storage, more time for processing and unnecessary large computational facility. Moreover practical field data may be contaminated with noise along with other associated redundancies. Such data when given input to the classifier may force the algorithm to perform defectively due to overfitting hence may not work properly to classify new samples.

Hence, the objective while pre-processing a large data before any application of Machine learning tools would be to probe the data for specific patterns that could summarize a large set of data with lesser number of variables. It should squeeze out information from the data such that the complete set of associated variables can be represented by a minimal set of variables. It is of paramount importance while using pre-processing tools is that the resultant outcome should be as informative as possible and should represent the parent data as accurately as possible. In the opposite scenario the total idea of dimensionality reduction could go in vain and the classifier could operate erratically.

For achieving the desired goal, some measureable and quantifiable properties are analysed out from the data. These properties are called the explanatory variables or features. The process of analysing raw data and obtaining features form observations is called the feature extraction. Typically, several features are extracted from a set of data to form the feature vector. Feature extraction reduces the variables required to represent a large set of data which facilitates easier and faster classification of that data. Hence, feature extraction pre-processes the input data to make it more explanatory and easy to manipulate.

4.6.5 Mathematical definition of Rough sets

Mathematically the decision table in RST is formulated as $T =< U, A, D, f >$ where, U is the finite set of objects $\{U_1, U_2, U_3, \dots, U_n\}$, A is the set of attributes $\{A_1, A_2, A_3, \dots, A_m\}$, $D = \bigcup_{q \in A} D_q$, where D_q is the domain of q and f symbolises decision function as, $f: U \times A \rightarrow D$.

The first step in the process is to discretise the normalised decision table. As stated earlier the aim of the classification process is to categorise each different fault condition into a separate granule which would contain the information about the fault. But the task would be too tiresome if the number of elements in the set, denoted as *card(V_q)*, is too large. Here, the term *card()* refers to the cardinality operator. Hence discretization of the data table is necessary to bring down *card(V_q)* to a lower value thus simplifying the process.

Note that this is the second step of simplification process used, the first being through the normalisation process. Through discretization some partitions in the decision table are introduced to obtain higher classification efficiency. These partitions are inducted by generating cuts in the domain of q and individual values for attributes are replaced with class-values i.e. the data within a range are given a particular class. Maximal Discernible (MD) heuristic has been employed in this present work to generate the optimum cuts and to discretize the table. The definition of heuristics according to literature is a technique developed to solve a given problem at a faster rate when conventional methods are too slow to use for any practical purposes or finding an approximate solution when classical methods fail to find any exact solution. Application of heuristics is of paramount importance to the field of Artificial Intelligence that aims in solving any conventional problem in a faster and intelligent manner. Details of MD heuristics can be found in [3.12].

The next step would be the process granulation [3.13] of the decision table. This step also helps in eliminating the superfluous features which could not be used for proper classification purpose. This is done with the help of indiscernibility relation in RST. In RST, indiscernible or indistinguishable attributes are those which are characterised by the same information. Mathematically, for $B \subseteq A$ and $(\sigma_i, \sigma_j) \in U$, where σ_i & σ_j are i^{th} & j^{th} objects, then σ_i and σ_j are indiscernible with respect to the set of attributes B, if $f(\sigma_i, q) = f(\sigma_j, q), \forall q \in B$.

The working theory behind the rough set based classifier is presented with the help of an example. Table 4.1 represents a data table arranged in the specific format as explained above. The attribute values are chosen arbitrarily here but in general usage they may represent some feature values.

Table 4.2: Example decision table

Objects	**Condition Attributes**				**Decision Attribute**
	A1	*A2*	*A3*	*A4*	
Object 1	6	8	4	5	*D1*
Object 2	6	9	4	5	*D1*
Object 3	4	7	4	8	*D2*
Object 4	6	7	1	8	*D2*
Object 5	5	7	3	5	*D3*
Object 6	4	0	4	5	*D3*

Take $B = \{A_4\}$ from Table 4.1, it can be seen that, at least four objects 1, object 2, object 5 and object 6 are indiscernible with respect to attribute A_4, as all these objects are characterised by the same information (i.e. value $= 5$). Similarly for $B = \{A_1\}$ objects $1, 2$ & 4 and again 3 & 6 are indiscernible as they possess same values (i.e. 6 & 4 respectively). These indiscernible objects form an elementary set. So, in case of $B \subseteq A$, an equivalence relation on U, named $B-$indiscernibility relation is given by I_B. Mathematically,

$$I_B = \{(\sigma_i, \sigma_j) \in \cup^2 | \forall q \in B, \quad f(\sigma_i, q) = f(\sigma_j, q)\} \dots \dots (4.5)$$

Considering again, $B = \{A_4\}$, it is seen that *B*-elementary set includes $\{1, 2, 5, 6\}$ and $\{3, 4\}$. The Information Granules are the corresponding classes of the partition introduced by the *B*-indiscernibility relation. In Table 4.1, along with the previously defined indiscernibility relation, the other relations which help in reducing the table to only indispensable data are the upper and lower approximation which are defined as $\overline{B} = \{\sigma \in B \mid I_B(\sigma) \subseteq Y\}$ and $\underline{B} = \{\sigma \in B \mid I_B(\sigma) \cap Y \neq \emptyset\}$, where Y is the rough set, and $\overline{B}, \underline{B}$ are the upper and lower approximations respectively. The objects in lower approximation $\underline{B}$ can be certainly classified as members of Y, based on knowledge in I_B. The set $Boundary_B(Y) = (\overline{B} - \underline{B})$, is called the boundary region of Y and consists of those objects which cannot be decisively classified as members of Y, based on knowledge available in I_B. Vividly, $\overline{B}$ is the set of the objects which 'possibly' belong to Y and hence the upper approximation. The set will be 'Rough' if and only if $Boundary_B(Y) \neq \phi$ i.e. the boundary is non-empty, otherwise it is 'Crisp'. Other redundant attributes are removed from the decision table with the help of this indiscernibility relation. For the assumed subset $B \subseteq A$, an attribute $q \in A$ is dispensable in B if, $I_B = I_{(B-\{q\})}$; or else it is indispensable. If none of the elements in the subset B is dispensable then B is said to be independent and vice-versa. Let $B \subseteq A$ and $P \subseteq A$ have equivalence relations in U. The B-positive region of P is denoted as, $POS_B(P) = \bigcup_{Y \in I_P} \underline{B}$. In other words, it denotes the group of elements that can be suitably classified into P-elementary sets obtained from I_P using the knowledge described by I_B. If $q \in B$ and $POS_B(P) = POS_{(B-\{q\})}(P)$ then q is P-dispensable in B, or else q is P-indispensable in B. If the set $G(G \subseteq B)$ is P-independent in B and $POS_G(P) = POS_B(P)$, then G is called as P-reduct of B or generally Reduct of B. Again looking at Table 4.1, B is taken as, $B = \{A1, A2, A3, A4\}$, and Decision Attribute$= D_i$ for i=$\{1, 2\ and\ 3\}$, then $I_B = \{1\}, \{2\}, \{3\}, \{4\}, \{5\}\ and\ \{6\}$

and $I_D = \{1,2\},\{3,4\}\ and\ \{5,6\}$. Also $POS_B(D) = \{1,2,3,4,5,6\}$. If A_2 is removed from B then, $POS_{(B-\{A_2\})}(D) = \{3,4,5,6\}$. Clearly $POS_B(D) \neq POS_{(B-\{A_2\})}(D)$, which means that the attribute set A_2 is D-indispensable in B.

But if A_3 is removed then, $POS_{(B-\{A_3\})}(D) = \{1,2,3,4,5,6\} = POS_B(D)$. Therefore A_3 is D-dispensable in P. Also we can see that the set attribute A_1 is also indispensable as $POS_{(B-\{A1\})}(D) = \{3,4,5,6\} \neq POS_B(D)$ .Also $POS_{(B-\{A_4\})}(D) = \{4,5\} \neq POS_B(D)$. Thus, the set $\{A_1, A_2, A_4\}$ is the D-reduct of P. Table 4.2 represents the reduced form of Table 4.1, which represents the decision Table without the dispensable attribute A_3. In this table, '×' represents "don't care" or dispensable condition.

Table 4.3: Reduced decision table

Objects	**Condition Attributes**				**Decision Attribute**
	A1	*A2*	*A3*	*A4*	
Object 1	6	8	×	5	*D1*
Object 2	6	9	×	5	*D1*
Object 3	4	7	×	8	*D2*
Object 4	6	7	×	8	*D2*
Object 5	5	7	×	5	*D3*
Object 6	4	0	×	5	*D3*

After reduction by elimination of all the dispensable attributes the remaining set is called a *Reduct*. *Core* is the group of relations in every *Reduct* i.e. *CORE (B) =∩REDUCT (B).* Therefore, *Core* will provide the most significant part of the decision. From the knowledge obtained from Core, decision rules are generated, which is the most important (and the final) step in the classification process. The decision rule is generally represented in an IF...THEN format.

Referring to the Table 4.2, which is the reduced form of Table 4.1, it can be deduced that the attribute condition $\{(A_1 = 6 \wedge A_2 = 8 \wedge A_4 = 5) \vee (A_1 = 6 \wedge A_2 = 9 \wedge A_4 = 5)\}$ is representative of decision class D_1. Similarly, $\{(A_1 = 4 \wedge A_2 = 7 \wedge A_4 = 8) \vee (A_1 = 6 \wedge A_2 = 7 \wedge A_4 = 8)\}$ represents the decision D_2. These are basically the Information Granules. Core for the corresponding classes are obtained as the intersected reduct of the individual decision classes (i.e. $D_1, D_2\ and\ D_3$ etc.). Table 4.3 represents the individual decision rule generated without the deduction of core. As could be seen from the table, no

Table 4.4: Individual Decisions

Decision Rule	**Generated Rule Base**	**Decision attribute**
Rule 1	$(A_1 = 6 \wedge A_2 = 8 \wedge A_4 = 5) \vee (A_1 = 6 \wedge A_2 = 9 \wedge A_4 = 5)$	D_1
Rule 2	$(A_1 = 4 \wedge A_2 = 7 \wedge A_4 = 8) \vee (A_1 = 6 \wedge A_2 = 7 \wedge A_4 = 8)$	D_2
Rule 3	$(A_1 = 5 \wedge A_2 = 7 \wedge A_4 = 5) \vee (A_1 = 4 \wedge A_2 = 0 \wedge A_4 = 5)$	D_3

CORE could be deduced for decision class D_3. This is because intersection of $\{(A_1 = 5 \wedge A_2 = 7 \wedge A_4 = 5)\ and\ (A_1 = 4 \wedge A_2 = 0 \wedge A_4 = 5)\} \Rightarrow \emptyset$. Note here that it could be said for decision class D_3 the rule should be $A_4 = 5$, but that could conflict with decision rule for D_1. The final decision with IF…THEN rule is presented in Table 4.4

Table 4.5: Decision Rule from Core and Reduct

Decision Rule number	**Statements of Rule**	
	IF	***THEN***
1	$(A_1 = 6 \wedge A_4 = 5)$	Take decision D_1
2	$(A_2 = 7 \wedge A_4 = 8)$	Take decision D_2

Chapter 7 gives details about the actual decision table used for the present work where the decision variables are different fault types and locations, and the objects are various fault current waveforms obtained during experimentation.

4.7 Reference

3.1) A. Smola and S.V.N. Vishwanathan, "Introduction to Machine Learning", Cambridge University Press.

3.2) E.G. Strangas; S. Aviyente and Syed S.H. Zaidi, "Time–Frequency Analysis for Efficient Fault Diagnosis and Failure Prognosis for Interior Permanent-Magnet AC Motors", IEEE Transactions on Industrial Electronics, Vol. 55, 2008, pages: 4191-4199.

3.3) D. Chen and D.W. Bunn, "Analysis of the Nonlinear Response of Electricity Prices to Fundamental and Strategic Factors", IEEE Transactions on Power Systems, Vol. 25, 2010, pages:595-606.

3.4) A. Elnady and M. M. A. Salama, "Mitigation of voltage disturbances using adaptive perceptron-based control algorithm", IEEE Transactions on Power Delivery, Vol. 20, 2005, pages: 309-318.

3.5) C. Koley, P. Purkait and S. Chakravorti, Wavelet-Aided SVM Tool for Impulse Fault Identification in Transformers", IEEE Transactions On Power Delivery, Vol. 21, 2006, pages: 1283-1290.

3.6) O.N. Gerek and D.G. Ece, "Power-quality event analysis using higher order cumulants and quadratic classifiers", IEEE Transactions on Power Delivery, Vol. 21, 2006, pages: 883-889.

3.7) I. Jarrar, K. Assaleh and A.H. El-hag, "Using a pattern recognition-based technique to assess the hydrophobicity class of silicone rubber materials", IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 21, 2014, pages: 2611-2618.

3.8) S. Rovnyak, S. Kretsinger, J. Thorp and D. Brown, "Decision trees for real-time transient stability prediction", IEEE Transactions on Power Systems, Vol.9, 1994, pages: 1417-1426.

3.9) A. De and N. Chatterjee, "Impulse fault diagnosis in power transformers using self-organising map and learning vector quantisation", IEE Proceedings - Generation, Transmission and Distribution, Vol. 148, 2001, pages: 397-405.

3.10) Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer, Boston, USA, 1991.

3.11) S. Biswas, D. Dey, B. Chatterjee, S. Chakravorti, "An approach based on rough set theory for identification of single and multiple partial discharge source", Electrical Power and Energy Systems, Elsevier, 2016, pages: 163-174.

3.12) Nguyen HS, “Discretization of real value attributes: a boolean reasoning approach”, PhD thesis, Dept. of Mathematics, Warsaw Univ. Poland; 1997.

3.13) Yao YY. Rough sets, neighbourhood systems and granular computing. Proceedings of IEEE conference on Electrical and Computer Engineering, Canada; 1999. Pages: 1553–1558.

Chapter 5

Analogue Model and Circuit Parameters

Chapter 5: Analogue Model and Circuit Parameters

5.1 Introduction

In this chapter a brief introduction to the constructional details of the analog model that is used for fault simulation and data acquisition is presented. The importance of developing an analog model for experimentation purpose is presented in the next section. In the subsequent sections the approach to model construction followed by tabular representation of the physical and electrical model parameters are presented. At last the fault current acquired through the use of digital oscilloscope are shown along the length of the winding. It is to be noted that the model presented in this chapter was constructed in High Tension Laboratory, Jadavpur University by Dr. Saibal Chatterjee as a part of his M.E. thesis [5.1] and the detailed calculation of the model parameters was done by Prof. Sugata Munshi for his M.E. thesis [5.2].

5.2 Need for an analog model

Impulse testing of large power transformers before commissioning is performed according to the rules pre-specified in IS 2026 (part 3), clause 13, to assess the insulation integrity of the transformer. For testing sequence of impulse voltages are applied with generally two different magnitudes. The test sequence shall consist of one impulse of reduced voltage and three subsequent impulses at full voltage. The absence of significant differences between voltage and current transients recorded at reduced voltage and those recorded at full test voltage constitutes evidence that the insulation has withstood the test. The aim of this thesis is to use intelligent methods to not only predict the existence of fault within the insulation but also to localise the faults along the length of the winding i.e. to get an idea about the section of the winding where the fault occurs with the help of only one application of full wave reduced magnitude pulse. This approach will enable the onsite experts to rectify the fault or take any appropriate precautionary step in a small interval of time thereby increasing the system reliability. For ease of accessibility, an analogue model of an actual transformer winding has been used in the present approach. Various types of faults have been emulated on the analogue model for testing the effectiveness of the proposed scheme. For this, a 33kV winding of a 3 MVA, 33/11 kV, 3-phase 50 Hz, ON, Dy 11 transformer has been chosen. The aim of the present approach

is the design of an intelligent system that would discriminate a healthy winding from a faulty one and also identify the respective locations of the faulted sections. Results from the analogue model would also help in verifying the accuracy of the proposed method. Keeping this in mind, this chapter presents the constructional details of the analogue model developed, various parameters of the analogue model, assumptions considered in the design of the model and various categories of faults that are emulated along the length the constructed model. This chapter also details the setup used for data acquisition to store the winding current details during healthy and faulted cases, which would be further used for analysis purpose.

5.3 Distributed parameter model of a transformer winding

The equivalent circuit model of a 3-phase delta (or star) connected winding of the transformer is represented as shown in Figure 5.1

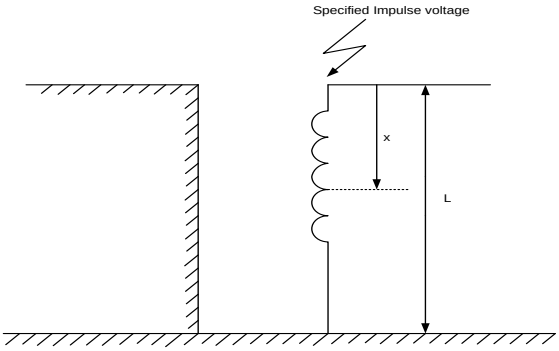


Figure 5.1: Winding representation of model

It can be seen from Figure 5.1 that the approach used for the construction of the analogue model follows the rules as specified in the IS 2026 part 3 (Clause 13), hence, every other winding other than the one which is hit by the impulse wave are short circuited to ground through tank and a high voltage co-axial shunt. The distributed parameter model for the above shown model is represented in Figure 5.2:

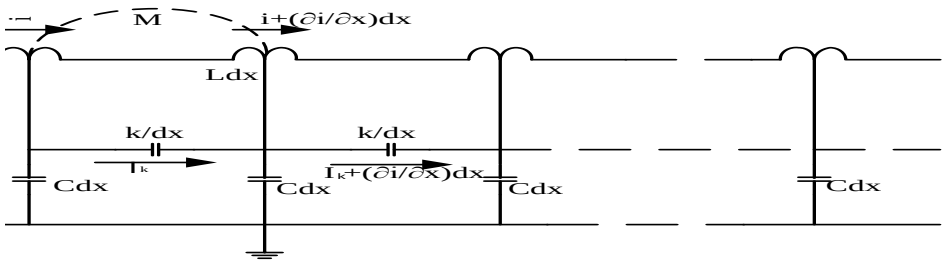


Figure 5.2: Distributed parameter model of the transformer winding

Figure 5.2 identifies the various types of distributed capacitance (series and parallel) along the length of the transformer winding. It also shows the distributed inductance values and the current distribution across the winding length. This circuit is developed by taking into consideration K.W. Wagner's theory of transient in the coil. In the given distributed parameter circuit, 'k' represents the series capacitance (i.e. capacitance between two consecutive winding turns or capacitance between two adjacent discs) and 'C' represents shunt capacitance (i.e. capacitance between the winding and the ground) of the winding, both in per unit length.

5.4 Analog Model:

Circuit for impulse testing of a delta and a star connected transformer windings is as follows:

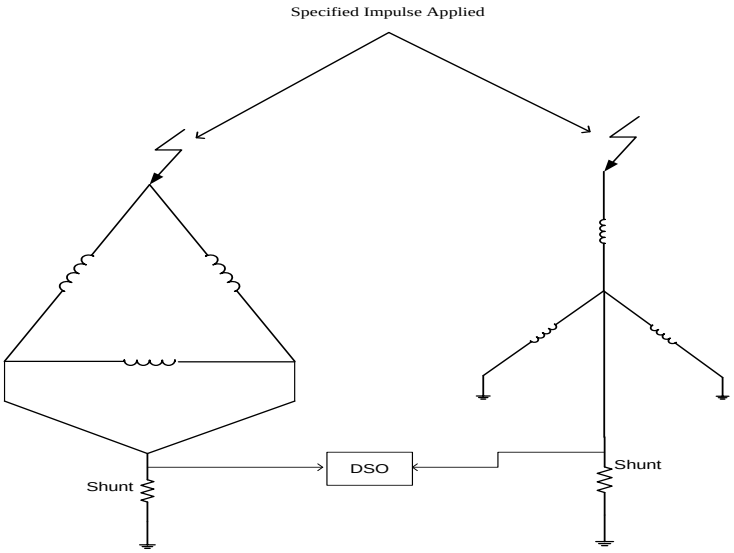


Figure 5.3: Connection Diagram for impulse testing a 3-phase delta and star transformers

As seen from Figure 5.3 in both the cases of 3-phase delta and star connected transformer, one of the terminal is subjected to lightning impulse voltage. The other two terminals are either shorted and connected to earth through a shunt (for delta connected winding) or individually grounded (for star connected winding) with shunt connected to the neutral. In both the connection methods the digital oscilloscope is connected across the shunt for data acquisition.

5.4.1 Basic assumptions for model construction:

i. The modelling is carried out by considering the winding as an isolated winding. This is because in the winding connections for impulse testing, a

three-phase delta connected winding with impulse applied at one terminal as specified by IS 2026 (Part 3) forms a circuit that results in a mutual cancellation of the main fluxes in the part of the core on which the winding to be impulse tested is wound. The cancellation of main core flux is because the flux generated in the short-circuited LV winding, which opposes the core flux. As a result the magnetic coupling by leakage flux between the impulse-hit HV winding and short-circuited LV winding is disregarded.

ii. In the model, the iron core may be replaced by an air core owing to the above mentioned cancellation of main flux. In the iron core, only the leakage flux contributes to the impulse response of the winding under test. Removal of the iron core may amount to omission of the slower electromagnetic component of transient stress distribution. This omission is insignificant in impulse studies. The magnetic dipoles in the magnetic core move at a very slower rate in comparison to the lightning phenomena. This was also reported by Miki, Hosaya and Okuyama [5.3] that the voltage response in an impulse-hit winding having an iron core is similar to the voltage response of the same winding without an iron core or simply air cored winding. Hence, the inductance of an iron core impulse hit winding is similar to an air core impulse hit winding.

5.5 Physical dimensions of the constructed model:

Physical dimensions of the analogue model are tabulated as in table 5.1:

Table 5.1: Physical Parameters

Physical Parameter	**Dimensions**
Outer Diameter	524 mm
Inner Diameter	424 mm
Winding Depth	50 mm
Axial height of a disc	6.6 mm
Number of discs	88
Number of turns per disc	19
Winding former diameter	15.5 cm
Winding former height	150 cm

The winding former is constructed using bakelite with air acting as core due to reasons described above in the assumptions. Super-enamelled copper conductor of SWG 23 is used as turns of a disc. Varnished cotton tape (0.34 mm thick) is used for disc insulation.

5.6 Lumped Circuit Parameters (Electrical)

The circuit parameters used are tabulated as given in table 5.2:

Table 5.2: Circuit Parameters

Circuit Parameters	**Values**
Disc Inductance (L_{self})	0.324 mH
Series Capacitance (C_s)	1.04 nF
Shunt Capacitance (C_g)	22.13 pF
Disc Resistance (R)	0.151 Ω

5.6.1 Determination of Self-Inductance

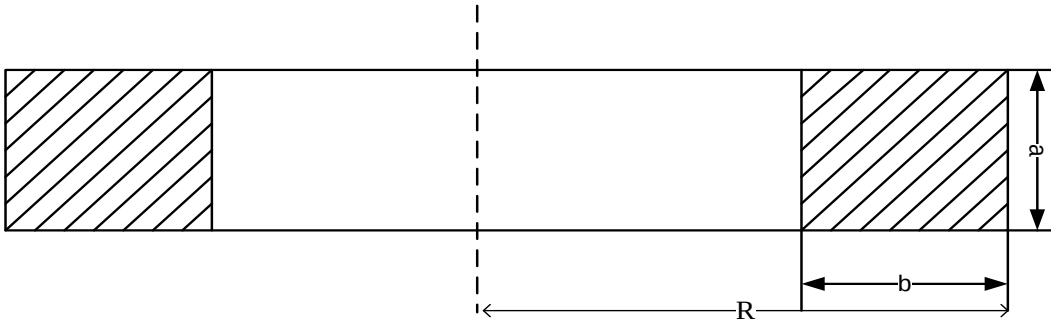
The self-inductance of a disc coil [5.2] is calculated from the following equation:

$$L_{self} = 4 \times 10^{-7} \times N_1^2 \left(ln \frac{8R}{R_1} - 2 \right) [H] \ldots\ldots (5.1)$$

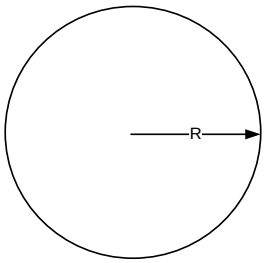
where,

$$\ln R_1 = \frac{1}{2}\ln(a^2 + b^2) - \frac{b^2}{12a^2}\ln\left(1 + \frac{a^2}{b^2}\right) - \frac{a^2}{12b^2}\ln\left(1 + \frac{a^2}{b^2}\right) + \frac{2b}{3a}tan^{-1}\left(\frac{a}{b}\right) + \frac{2a}{3b}tan^{-1}\frac{b}{a} - \frac{25}{12}$$

and N_1=number of turns and a, b and R are dimensions in meters of the coil as shown below



(a)



(b)

Figure 2.4 Coil dimensions from (a) side view & (b) top view

5.6.2 Calculation of Series (K) and Shunt (C_g) Capacitance:

The capacitance of the winding is calculated assuming the windings to be cylindrical electrodes. The details of methods of calculation of series and shunt capacitances is given in [5.2]. It is to be noted here that the shunt or ground capacitance of a transformer is calculated differently when considering the middle limb for calculation from the case when considering the side limb. The capacitance between impulse-hit limb winding and tank (i.e. C_g) or the shunt capacitance is computed by the expression for capacitance between two coaxial cylinders with certain approximations. Method of images is used to calculate the capacitance between the impulsed middle limb winding and tank.

5.7 Recurrent Surge Generator

The recurrent surge generator is an impulse generator that produces impulse waves of nominal specifications i.e. with a wave-front time of $1.2 \pm (30\%)$ μsec. and wave-tail of $50(\pm 20\%)$ μsec, as specified in the IS-2026 (part-3, Clause 13). The output from the surge generator is directly applied to the analogue model due to its low magnitude output. It has a no load output of about 136 V_{peak}. It has a repetition rate of 50 Hz i.e. at every 20 milliseconds the generator produces an output impulse wave. The waveshape of the output impulse wave may be controlled either by controlling externally the source capacitance or by the tail resistors.

Figure 2.6 Recurrent Surge Generator

5.8 Digital oscilloscope:

TEKTRONIX TBS1152 Digital Storage Oscilloscope (2 channels, 1GS/s, 150 MHz,8 bit) has been used for this data acquisition. The data from the oscilloscope can be transferred to the PC for analysis purpose through its USB port. The DSO has a recording length of 2.5k on all time bases. Input impedance of the oscilloscope is $1M\Omega$ in parallel with 20 pF. The input sensitivity range of the vertical system is 2mV/div to 5V/div and that of horizontal system is 5ns/div to 50s/div.

Figure 2.7 Digital Oscilloscope

5.9 Schematic Diagram of the model with series and shunt fault and Experimental setup

The Figure 5.8 represents the schematic diagram of the analog model. The series and shunt faults are represented which are produced by short circuiting the equivalent capacitance between two adjacent discs for series faults and short circuiting the respective disc to tank for shunt fault respectively. One of such series and shunt fault are shown.

Figure 5.9 represents the actual image of the winding model used for experimentation and Figure 5.10 shows the image of the shunt used in the experiment for data acquisition.

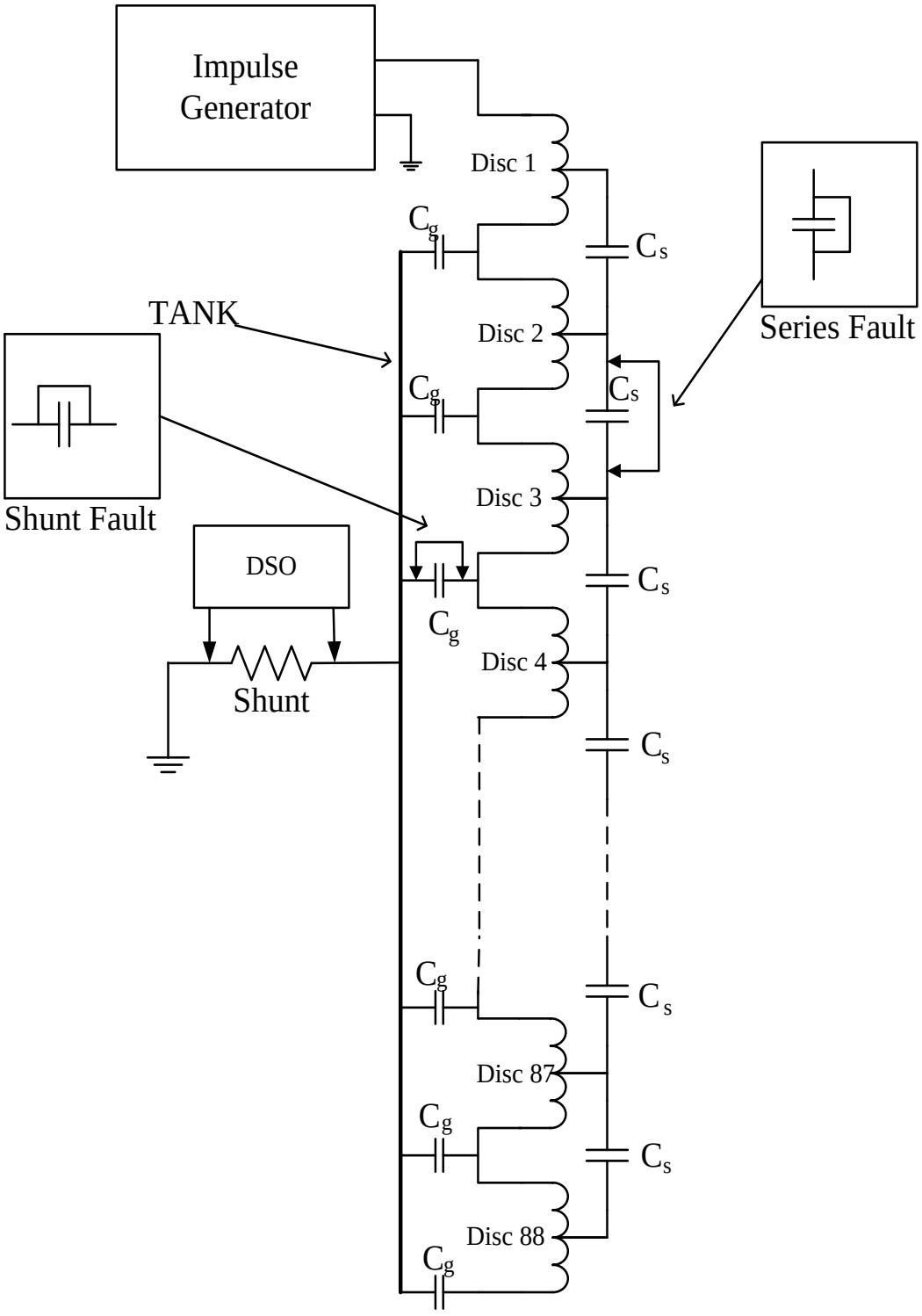


Figure 5.8 Schematic representation of analog model with emulated faults

Figure 5.9 Experimental Setup

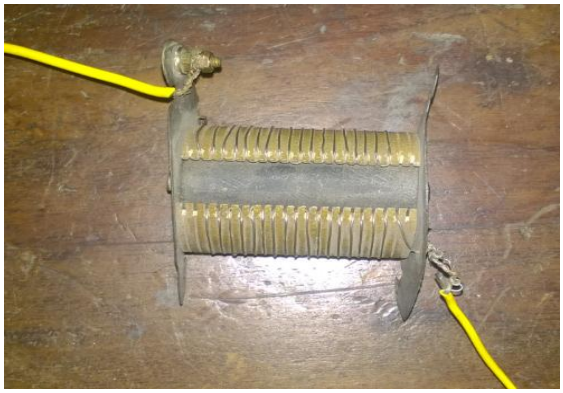
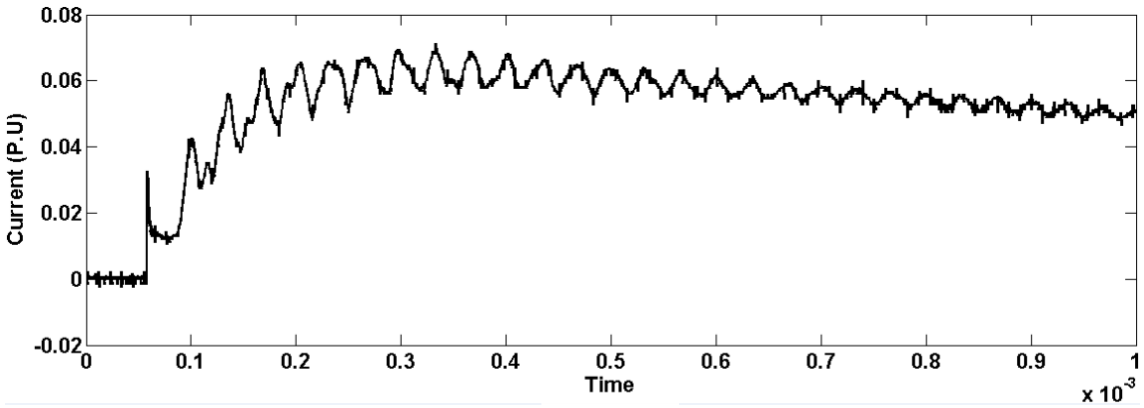


Figure 5.10 Co-Axial shunt

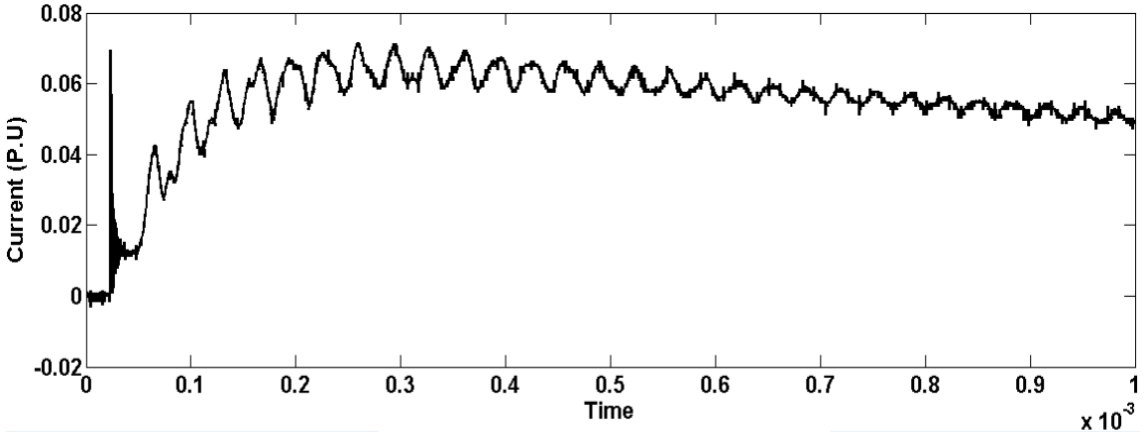
2.10 Recorded Waveforms

The figures presented below shows different waveforms obtained due fault at a different location along the winding length. Figure (a) represents the current waveform due to no fault. Figure (b), (c), and (d) represents waveforms obtained due to Series faults at upper,

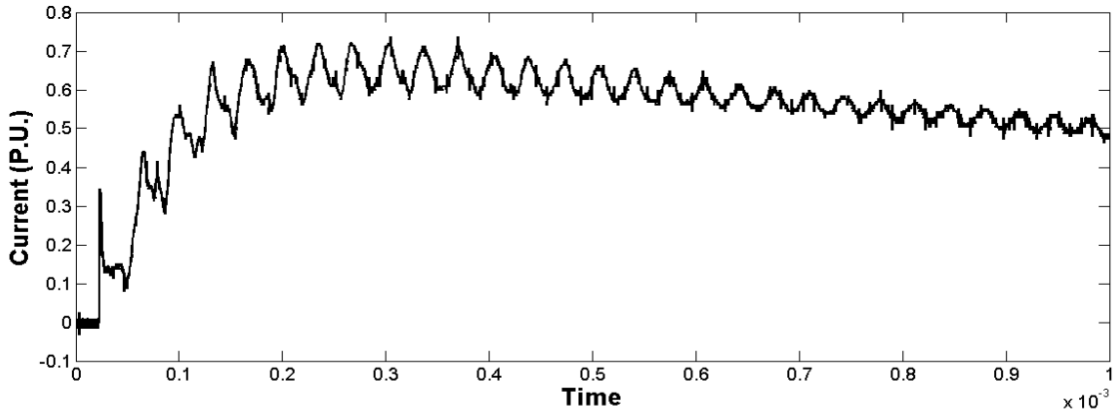
middle and lower end of the winding respectively. Similarly figure (e), (f) and (g) gives waveforms for shunt fault at upper, middle and lower end respectively.



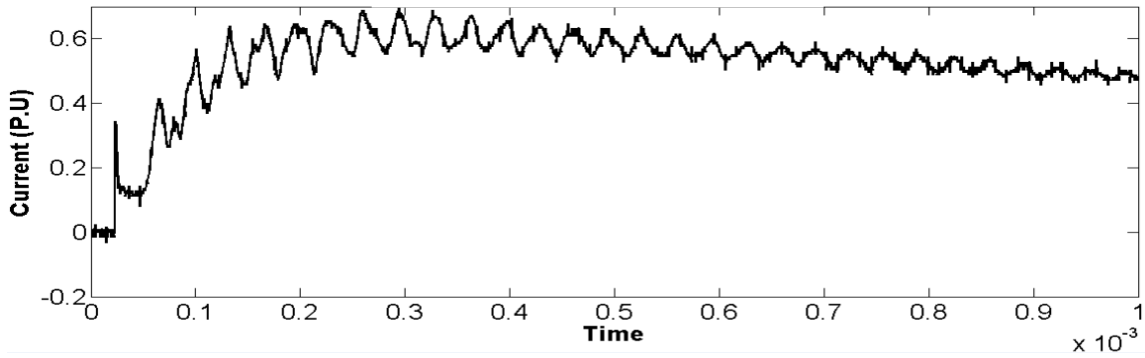
(a)



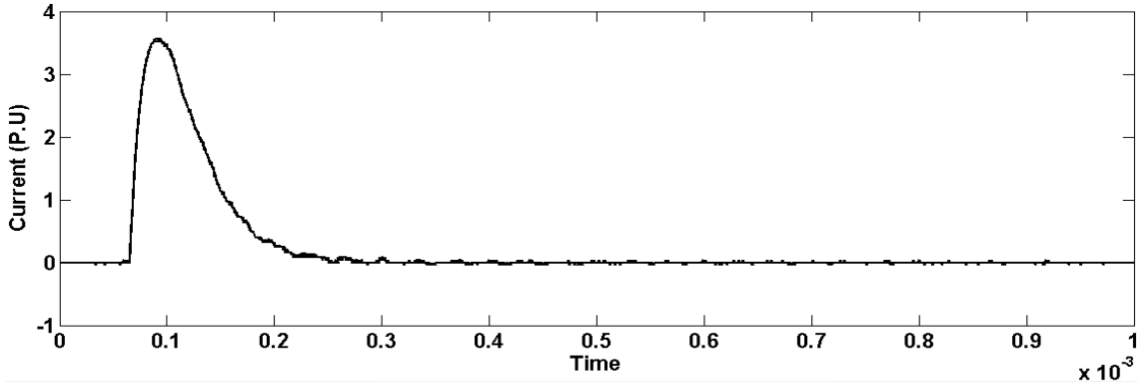
(b)



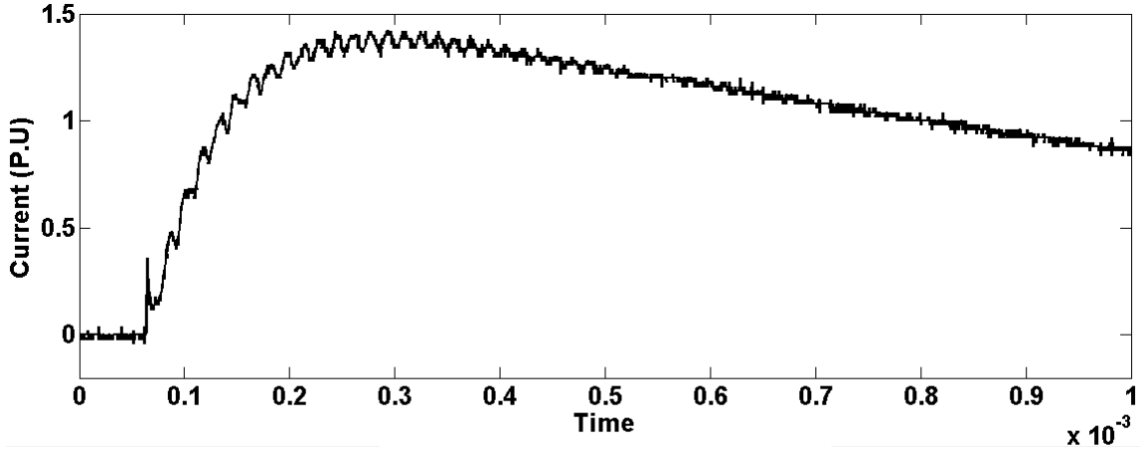
(c)



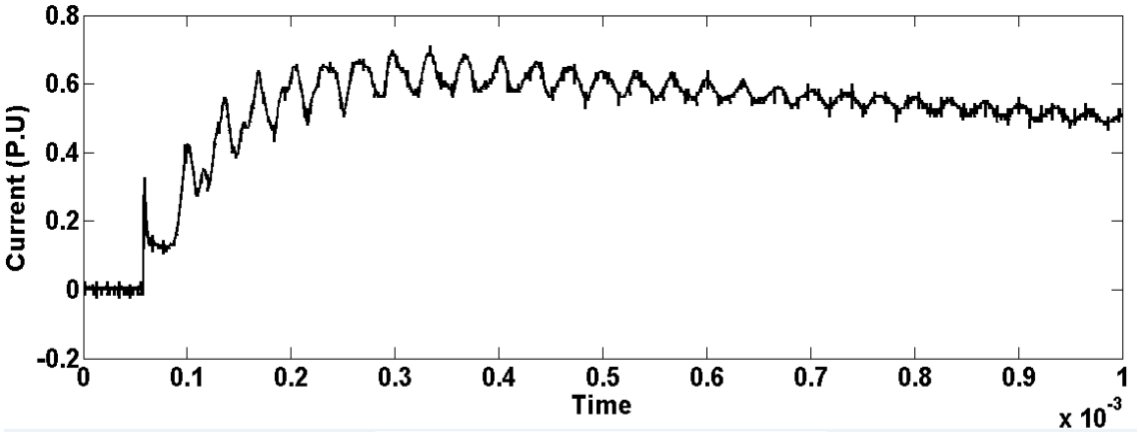
(d)



(e)



(f)



(g)

Figure 2.11: Recorded current waveforms

(a) No Fault

(b), (c) & (d) Series Faults along winding length

(e), (f) & (g) Shunt Faults along winding length

These obtained waveforms are further analysed using the signal processing tools presented in chapter 3 and at last classified using the tool presented in chapter 4 to identify fault location along the constructed analog model and also categorise the type of fault existing and depicted by the waveform. The detailed analysis is presented in the next chapter.

5.11 References

5.1) S. Chatterjee, "Analog studies on impulse performance of a H.V transformer winding with tapping coils", M.E.E thesis (Jadavpur University), 1994.

5.2) S. Munshi, "Digital Computation of Potential Distribution in a Transformer Winding due to Lightning Impulse", M.E.E thesis, Jadavpur University, 1985.

5.3) A. Miki, T. Hosoya and K. Okuyama, "A calculation method for impulse voltage distribution and transferred voltage in transformer", IEEE Transactions on Power Apparatus and Systems, Vol. PAS-97, 1978, pages: 930-938.

Chapter 6
Result and Discussion

Chapter 6: Results and Discussion

6.1 Introduction

The present chapter provides an insight into the methodology of identifying the two different types of fault considered i.e. Series and Shunt short circuits and their relative position of occurrence along the length of the winding. For classification of the different types of fault Mathematical Morphological and Stockwell transform based features and a Rough set based classifier is developed. Faults of two different types at different positions along the length of the analog model is simulated and the resultant fault currents are acquired using a Tektronix digital oscilloscope from a coaxial shunt connected across the tank and ground.

The two signal processing tools namely Mathematical Morphology and Stockwell Transform are used to extract some important features, the theoretical details of which are presented in chapter 3. For extraction of features from morphologically transformed data, only statistical features are used. Whereas, features extracted using Stockwell transform contains a mixture of statistical features along with energy and entropy of the signal.

6.2 Mathematical Morphology based feature extraction

Hybrid mathematical morphological operators have been used in the present work to transform impulse fault current data. The white and black top hat transforms are used to operate on the acquired data and statistical features are extracted based on it. A combination of two different structuring elements, a linear, with height 0.01 and length of 10, and a semi-circular, with height 0.05 and length 15 has been chosen for analysis. The choice of these structuring elements was based on a trial and error basis and these two structuring elements provided the best results. In the present scheme, three types of Hybrid Morphological operators have been used which were introduced in chapter 3 in equations 3.3, 3.4 and 3.5 respectively and are again presented below:

$$White\ Top\ Hat\ (WTH) = f_0(n) - (f \circ g) \dots\dots (6.1)$$

$$Black\ Top\ Hat\ (BTH) = (f \cdot g)(n) - f_0(n) \dots\dots (6.2)$$

$$\&\ Top\ Hat\ (TH) = WTH - BTH \dots\dots (6.3)$$

$$OCM2 = \frac{dilation(OCM1) + erosion(OCM1)}{2} \ldots\ldots (6.4)$$

Figure 6.1 presented below shows the profile of White top hat transform one at line end, one at middle end and one on the lower end of the winding for both series and shunt faults.

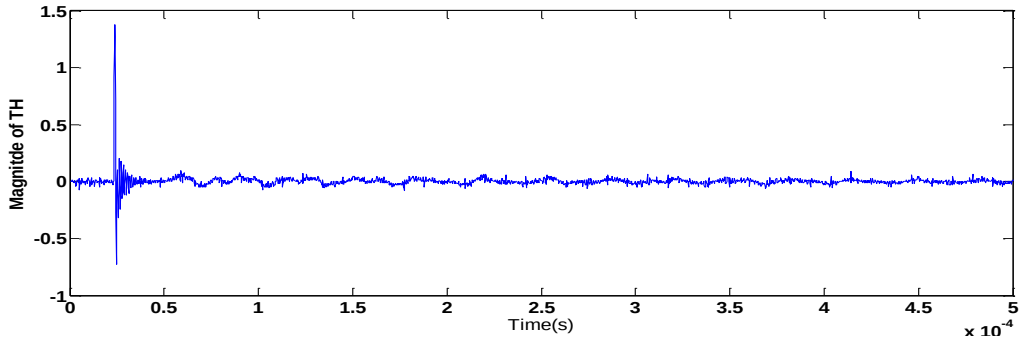


Figure 6.1: Top hat transform for Series fault at line end

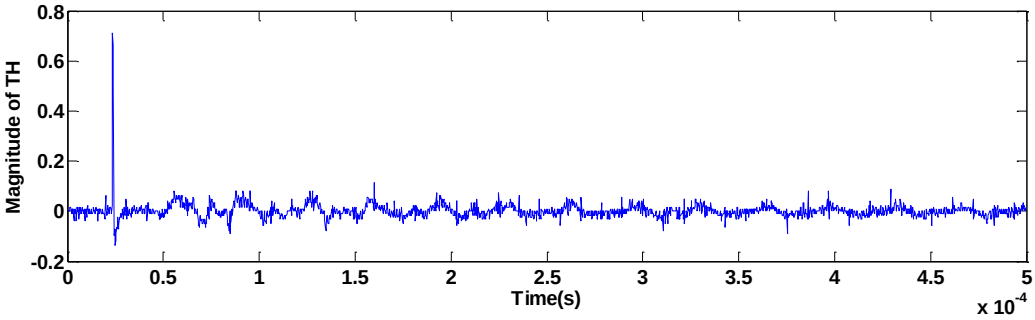


Figure 6.2: Top hat transform for Series fault at middle of the winding

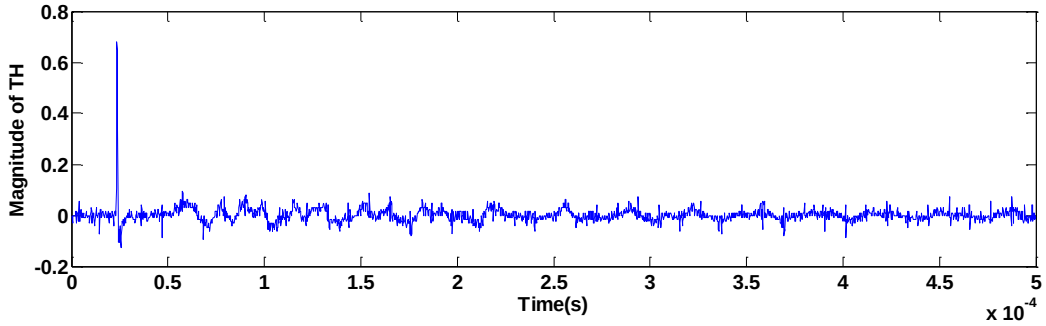


Figure 6.3: Top hat transform for Series fault at lower end of the winding

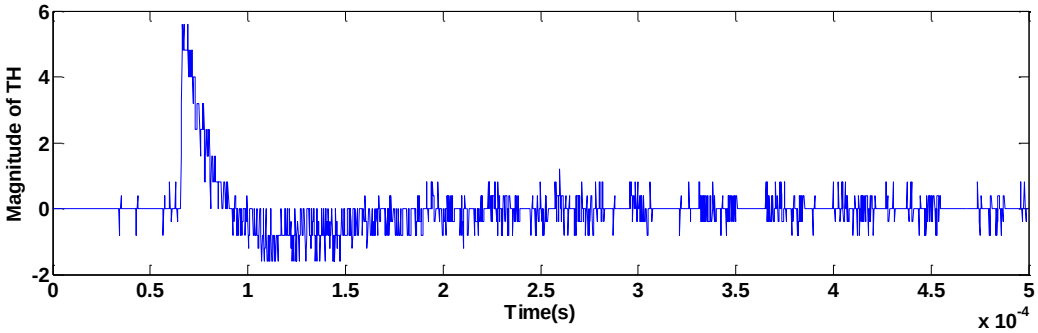


Figure 4: Top hat transform for Shunt fault at line end

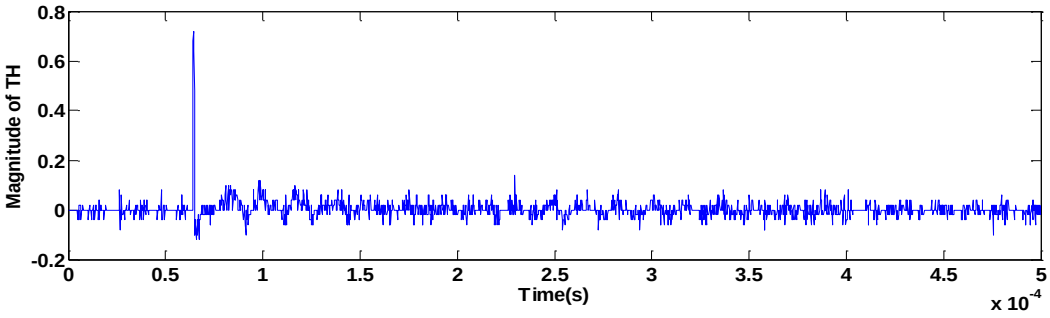


Figure 6.5: Top hat transform for Shunt fault at middle of the winding

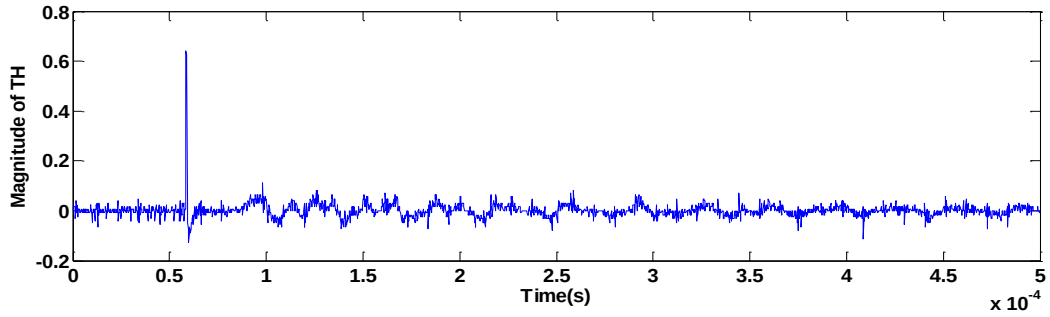


Figure 6.6: Top hat transform for shunt fault at lower end of the winding

In the present work, 13 distinct features are extracted (F_1 through F_{12}) from a mixture of WTH, BTH, TH and OCM2 as described in equations 6.1 to 6.4 respectively. These 13 statistical features are given in the Equations (6.5) to (6.17). For equations 6.5 to 6.9 the quantity $f_0(n)$ in equations 6.1 to 6.3 is the original signal $f(n)$ over which basic operations are performed. For equations 6.10 to 6.12 $f_0(n)$ in the reference signal of the healthy transformer as presented above in section 6.2.For the remaining signals i.e. equations 6.13 to 6.17 OCM2 is used to calculate statistical features. Hence, in summary the features can be partitioned as:

a) Features based on Top hat (white and black) transform with $f_0(n)$ =original signal $f(n)$ and a beeline structuring element

$$F_1 = TH_{var} = \frac{\sum_{i=1}^{N}[TH(n) - \mu]^2}{N} \dots\dots (6.5)$$

$$F_2 = WTH_{var} = \frac{\sum_{i=1}^{N}[WTH(n) - \mu]^2}{N} \dots\dots (6.6)$$

$$F_3 = TH_s = \frac{\sum_{n=0}^{N}[TH(n) - \mu]^3}{(N+1)\sigma^3} \dots\dots (6.7)$$

$$F_4 = WTH_S = \frac{\sum_{n=0}^{N}[WTH(n) - \mu]^3}{(N+1)\sigma^3} \dots\dots (6.8)$$

$$F_5 = BTH_S = \frac{\sum_{n=0}^{N}[BTH(n) - \mu]^3}{(N+1)\sigma^3} \ldots\ldots (6.9)$$

b) Features based on Top hat (white and black) transform with $f_0(n)$ =referance signal $f_r(n)$ and a beeline structuring element

$$F_6 = TH_{var} = \frac{\sum_{i=1}^{N}[TH(n) - \mu]^2}{N} \ldots\ldots (6.10)$$

$$F_7 = EW\_TH = \frac{\sum_{n=0}^{N} n * TH(n)}{TH_{max}} \ldots\ldots (6.11)$$

$$F_8 = EW\_BTH = \frac{\sum_{n=0}^{N} n * BTH(n)}{BTH_{max}} \ldots\ldots (6.12)$$

c) Features based on Opening-Closing mean (OCM2) operator with linear structuring element of height 0.01 and length of 15.

$$F_9 = OCM2_{max} \ldots\ldots (6.13)$$

$$F_{10} = OCM2_{var} = \frac{\sum_{i=1}^{N}[OCM2(n) - \mu]^2}{N} \ldots\ldots (6.14)$$

$$F_{11} = OCM2_S = \frac{\sum_{n=0}^{N}[OCM2(n) - \mu]^3}{(N+1)\sigma^3} \ldots\ldots (6.15)$$

$$F_{12} = OCM2_k = \frac{\sum_{n=0}^{N}[OCM2(n) - \mu]^4}{(N+1)\sigma^4} \ldots\ldots (6.16)$$

$$F_{13} = AVG_{OCM2} = \frac{\sum_{n=0}^{N} OCM2(n)}{(N+1)} \ldots\ldots (6.17)$$

In these 13 features, μ and σ represent the Average and Standard Deviation of the respective signals. The extracted features are described as**: F_1 = Variance of TH (TH_{var}), F_2 = Variance of BTH (BTH_{var}), F_3 = Skewness of the TH (TH_s), F_4 = Skewness of the WTH (WTH_s), F_5 = Skewness of the BTH (BTH_s), F_6 = Variance of the TH(TH_{var}), F_7 = Equivalent width of the TH (EW_TH), F_8 = Equivalent width of the BTH (EW_BTH), F_9 = Maximum value of the OCM2 ($OCM2_{max}$), F_{10} = Variance of the OCM2 ($OCM2_{var}$), F_{11} = Skewness of OCM2 ($OCM2_s$), F_{12} = Kurtosis of OCM2 ($OCM2_k$), F_{13} = Average of the OCM2 (AVG_{OCM2}).** These 13 features were found to be sufficient to classify various fault locations with an admissible accuracy. Some of the extracted features for different fault locations have been shown in Figure 6.7 to Figure 6.10.

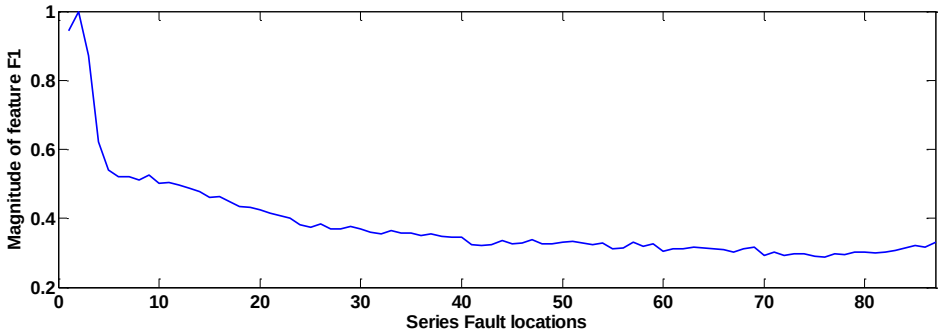


Figure 6.7: Variation of Feature F1 for Series faults only

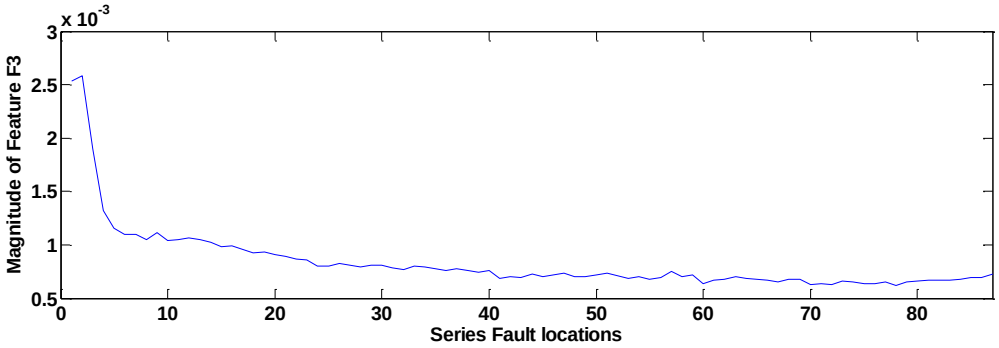


Figure 6.8: Variation of Feature F3 for series fault only

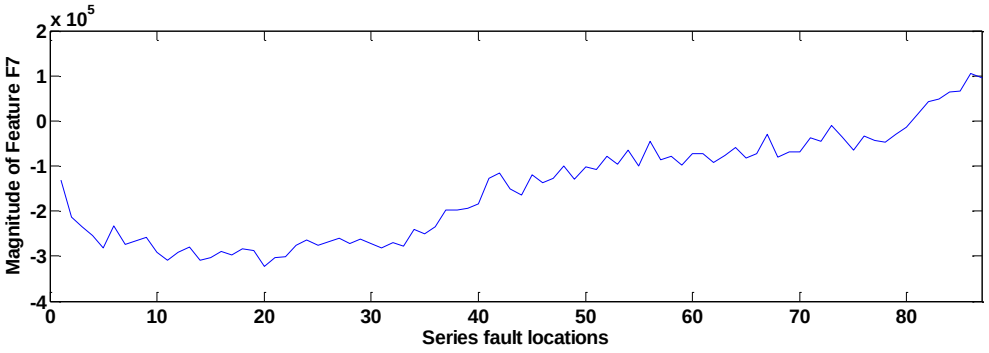


Figure 6.9: Variation of Feature F7 for series faults only

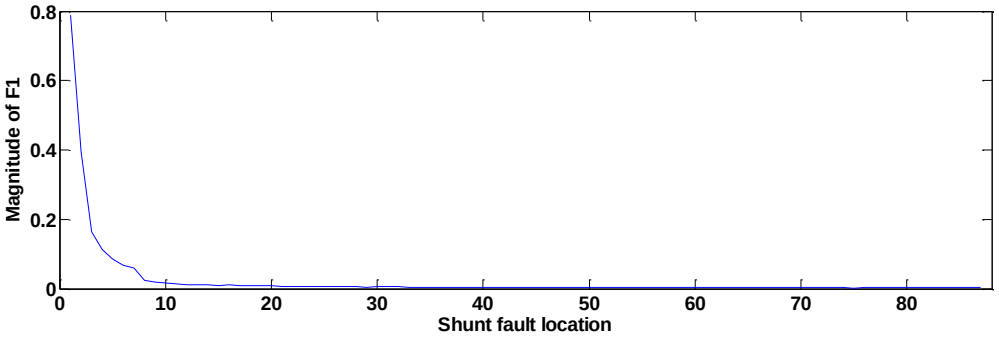


Figure 6.10: Variation of Feature F1 for Shunt faults only

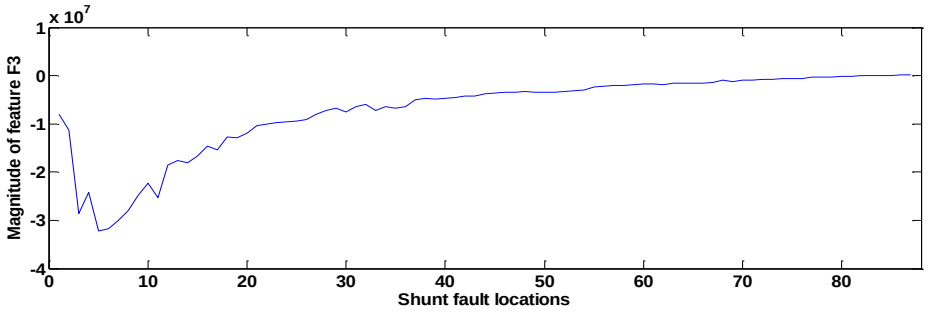


Figure 6.11: Variations of Feature F3 for shunt fault only

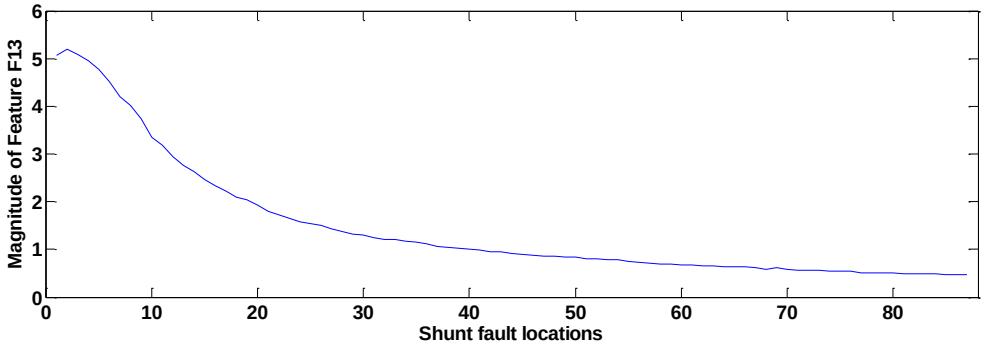


Figure 6.12: Variations of Feature F13 for shunt fault only

It can be noticed from the above figures that features for series and shunt faults are plotted separately. This is for the purpose of better classification accuracy which would be achievable with less numbers of feature to be dealt at a time. Based on the magnitude of feature F9 two types of faults i.e. series and shunt faults are identified. F9 is in the range of 0.68 to 0.758 for series faults and for shunt faults the range becomes 0.76 to 35.8. As will be discussed in section 6.5 that the classification table for series and shunt fault localization is also different.

6.3 Stockwell Transform based feature extraction

Stockwell transform with a modified Gaussian window as mentioned in section 3.3.6 of chapter 3 has been used as a tool for extracting some features from the various fault current waveforms.

As already mentioned in the chapter's introduction a mixture of statistical features with energy and entropy of the signal are used as features. By varying the parameters of equation 3.33 of chapter 3 i.e. F, a, b and c different features are extracted. For statistical features only default parameter values are only used and the values are modified only for extracting features using Energy and Entropy of the fault current waveform.

6.3.1 S-Transform contour plots for different fault cases

Various contour diagrams obtained for different identified fault cases are shown

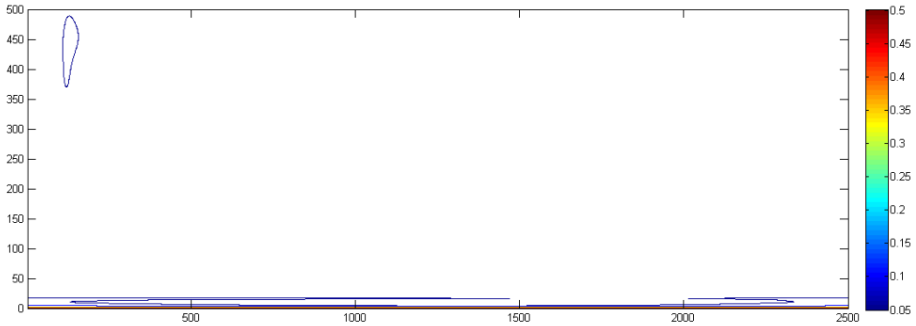


Figure 6.13: Series Fault at position 1

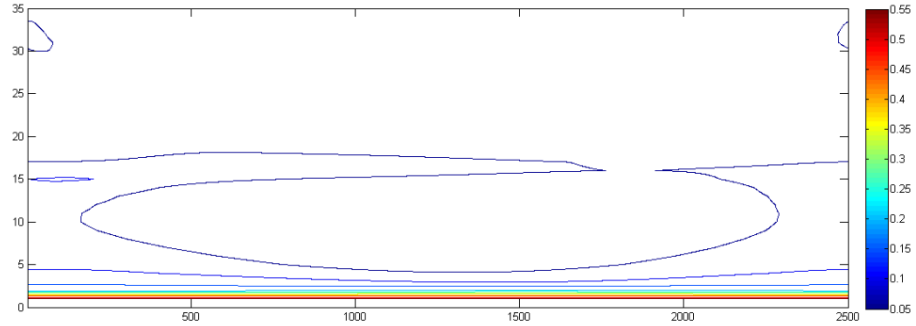


Figure 6.14: Series Fault at position 20

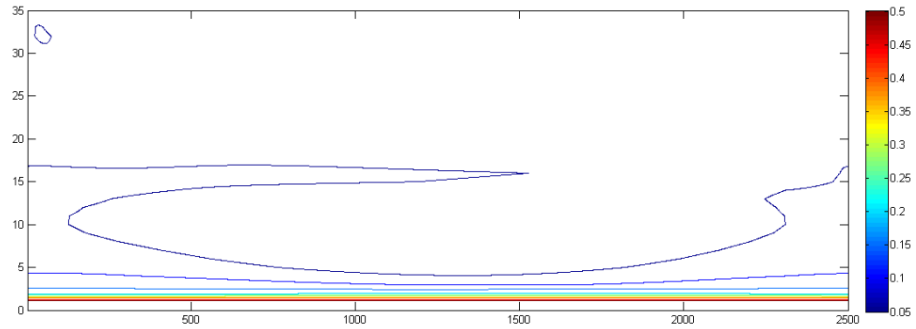


Figure 6.13: Series Fault at position 44

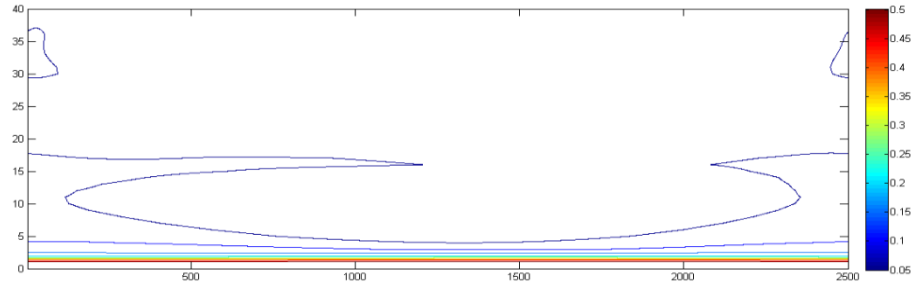


Figure 6.14: Series Fault at disc 67

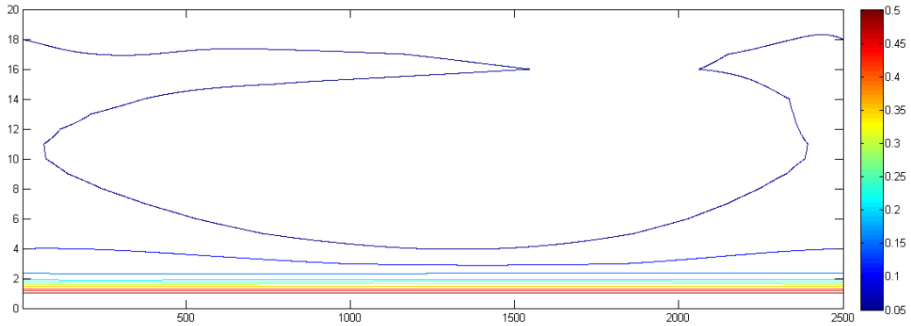


Figure 6.15: Series Fault at position 87

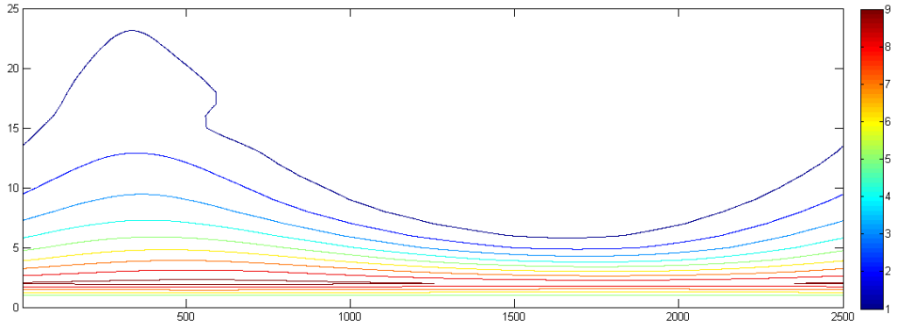


Figure 6.16: Shunt Fault at position 1

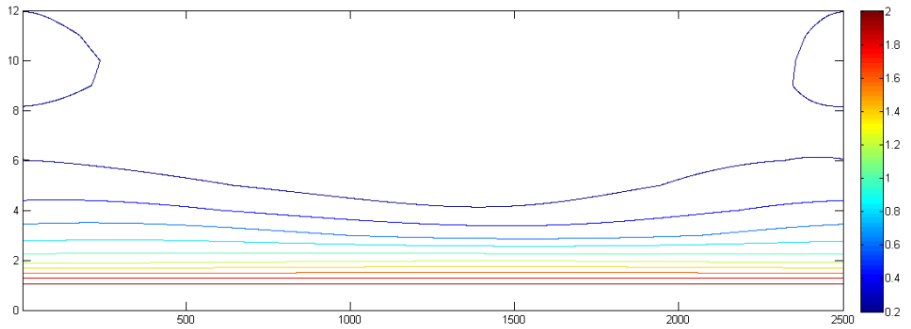


Figure 6.17: Shunt Fault at position 20

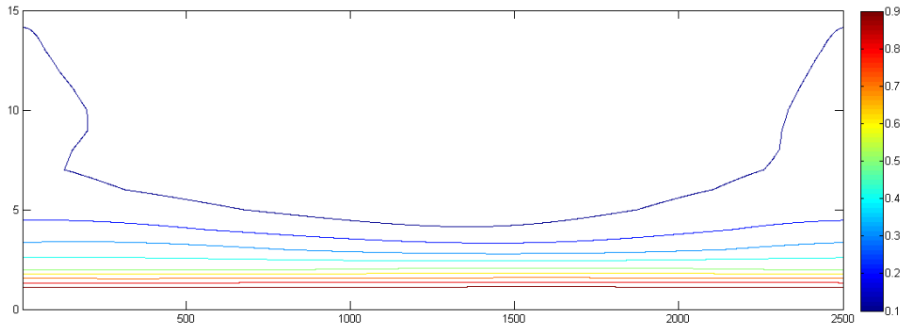


Figure 6.20: Shunt Fault at position 44

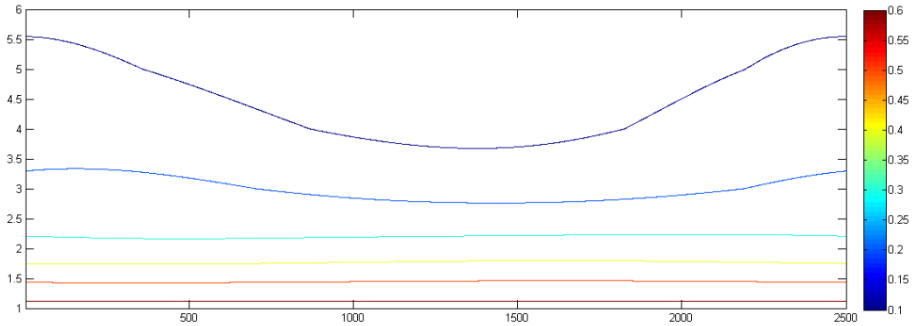


Figure 6.18: Shunt Fault at position 67

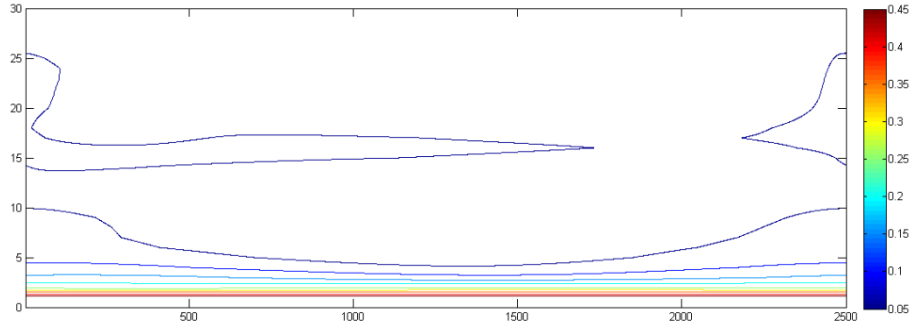


Figure 6.22: Shunt Fault at position 88

6.3.2 S-Transform based features used for fault classification

a) Energy: Energy of every signal is calculated from the Stockwell transform matrix obtained by varying the constant parameters (F, a, b and c). The formula used for calculation of energy is defined as

$$E = Real\left[\sum_{j=1}^{N}\sum_{i=1}^{M}\{S_{i,j} \times S_{i,j}^{*}\}\right] \dots\dots (6.18)$$

Where, N is the number of columns of the S-matrix and M is the number of rows in the S-matrix. $S_{i,j}^{*}$ is the conjugate of the quantity $S_{i,j}$ in the S-matrix. Three different features are extracted based on energy of the signal based on different settings of the constant parameters. The features are defined as :

$$E_{F_1} = E|_{F=1,a=0,b=1,c=1} \dots\dots (6.19)$$

$$E_{F_2} = E|_{F=1,a=0,b=1,c=1.3} \dots\dots (6.20)$$

$$E_{F_3} = E|_{F=0.33,a=1.3,b=0.33,c=1.1} \dots\dots (6.21)$$

Figure 6.22 to Figure 6.25 shows some of the features for series and shunt faults.

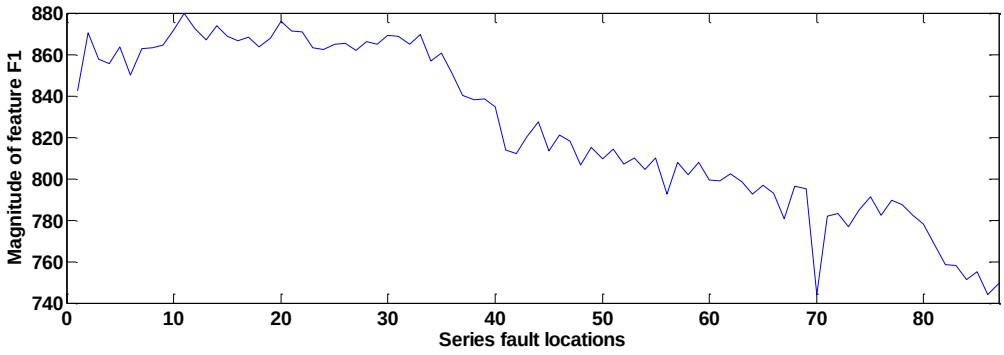


Figure 6.23: Variations of feature F1 for series fault only

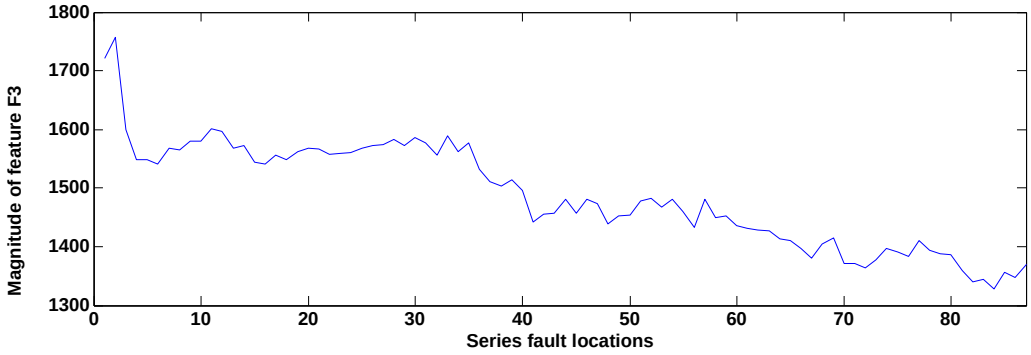


Figure 6.19: Variations of feature F3 for series faults only

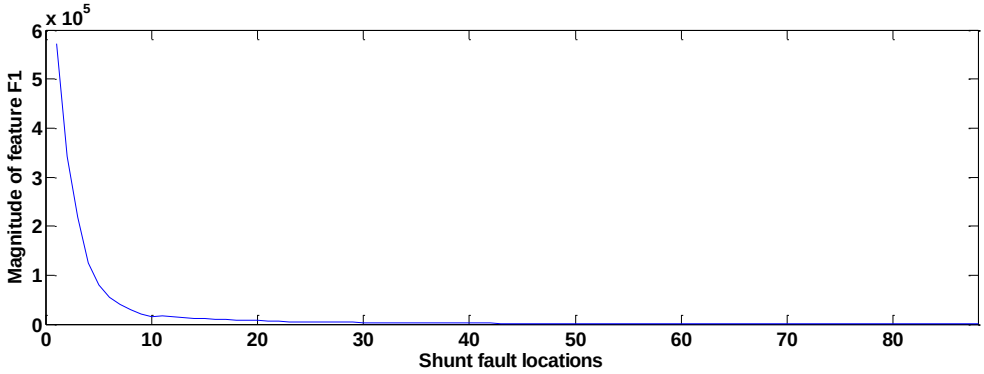


Figure 6.20: Variation of feature F1 for Shunt fault only

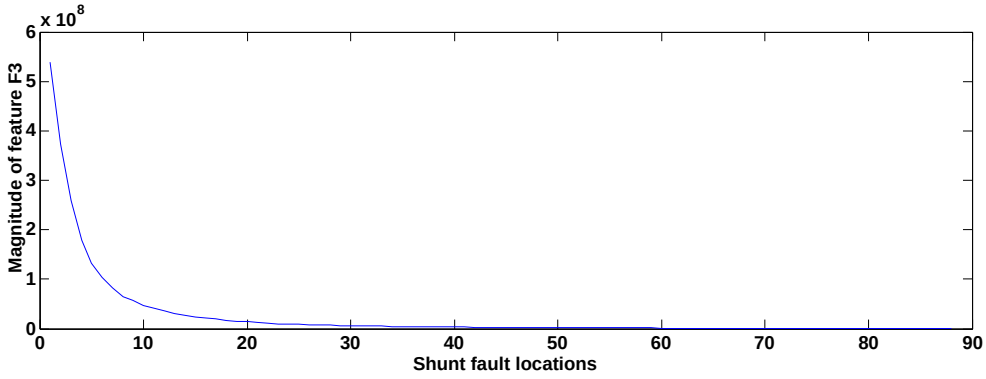


Figure 6.21: Variations of feature F3 for shunt faults only

b) Entropy: Shannon's entropy used as one of the features. Entropy of a signal is defined as the amount of information that the signal contains and is defined as follows: Suppose we have a set of possible events whose probabilities of occurrence are $p_i(i = 1, 2, \dots, n)$, the entropy of the set of probabilities is defined as

$$H = -\sum_{i=1}^{n} p_i \ln p_i \dots \dots (6.22)$$

Where, $0 < p_i < 1$ and $\sum_{i=1}^{n} p_i = 1$. For using entropy as one of the features singular value decomposition of the $(M \times N)$ S-matrix is done. Only relatively large singular values are singled out for analysis for damping the effect of noise. For selecting the singular values σ_i for $i = 1, 2, \dots, M$, the selection criteria is chosen to be

$$\frac{\sigma_j}{\sigma_1} > m \dots \dots (6.23)$$

Where, m is chosen according to the noise levels. Assuming the number of singular values selected is $'q'$ then the associated probabilities are calculated as

$$p_j = \frac{\sigma_j}{\sum_{j=1}^{q} \sigma_j} \dots \dots (6.24)$$

At last the entropy is calculated using equation 6.22 as given above.

The features based on entropy are defined as:

$$H_{F_4} = H|_{F=1,a=0,b=1,c=1} \dots \dots (6.25)$$

$$H_{F_5} = H|_{F=1,a=0,b=1,c=1.3} \dots \dots (6.26)$$

$$H_{F_6} = H|_{F=1,a=0,b=0.7,c=0.3} \dots \dots (6.27)$$

Figure 6.26 to 6.29 gives variation of calculated entropy for different values of constants (F, a, b & c) for series and shunt faults at different locations along the winding model.

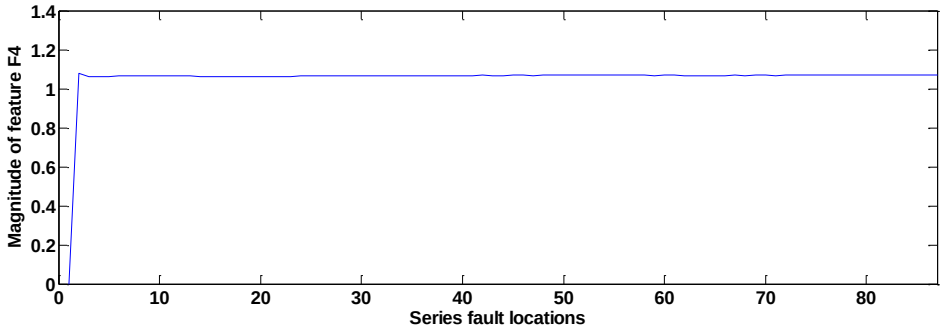


Figure 6.22: Variations of feature F4 for series faults only

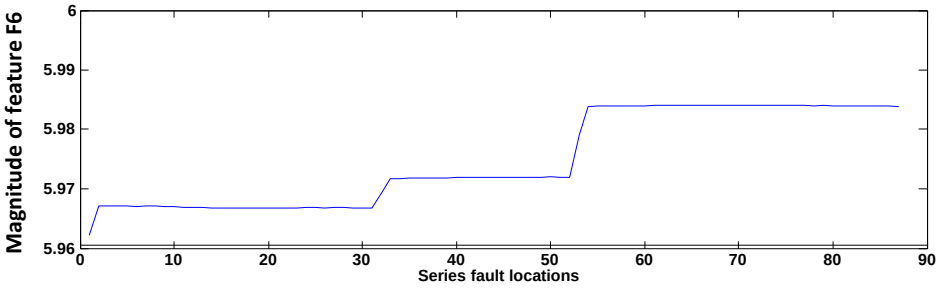


Figure 6.23: Variations of feature F6 for series faults only

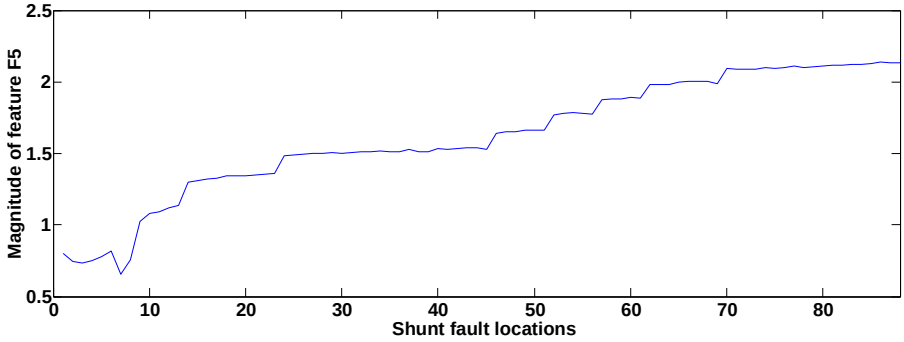


Figure 6.24: Variations of feature F5 for shunt faults only

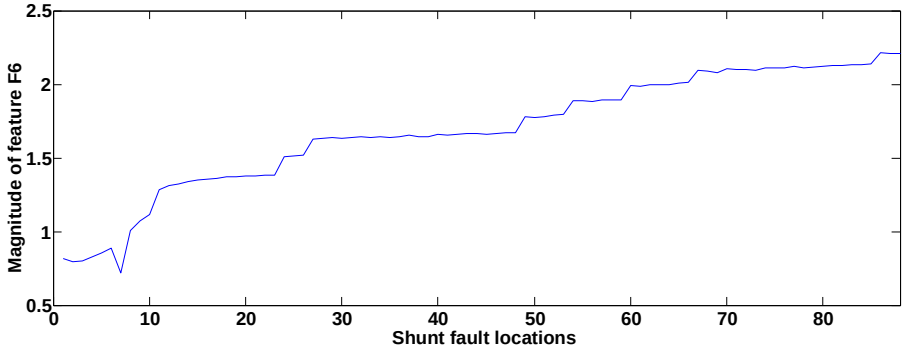


Figure 6.25: Variations of feature F6 for shunt fault only

c) Standard Deviation: One of the statistical features used for extracting features from the S-matrix is the standard deviation. The formulae used for calculating the standard deviation of the whole S-matrix is

$$STD = Real\left[\sum_{j=1}^{N}\left\{\frac{1}{M-1}\sum_{i=1}^{M}\left(S_{i,j}-\overline{S_j}\right)^2\right\}^{1/2}\right] \ldots\ldots (6.28)$$

And the feature is defined as: $S_{F_7} = STD|_{F=1,a=0,b=1,c=1} \ldots\ldots (6.29)$

Figure 6.30 and 6.31 shows the variation of feature F7 extracted based on standard deviation of the S-matrices for series and shunt faults respectively at different locations along the winding model.

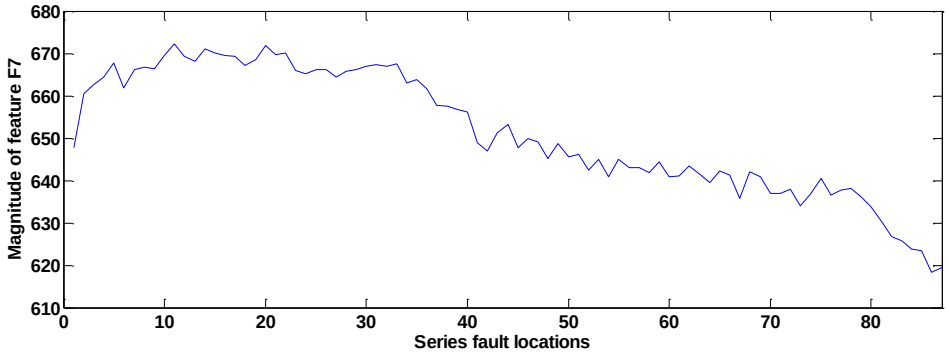


Figure 6.26 : Variations of features F7 for series faults only

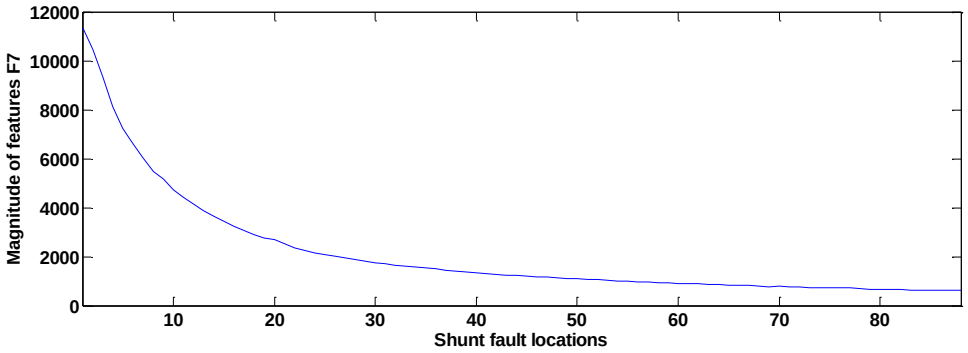


Figure 6.27: Variations of feature F7 for shunt fault only

d) Variance: The other statistical feature used for feature extraction from Stockwell transform data in the variance of each S-matrix for different fault types and locations. The formula used for calculation of variance from an S-matrix is

$$VAR = \sum_{i=1}^{N} \frac{1}{M} \sum_{j=1}^{M} \left(S_{i,j} - \overline{S_{\iota,J}}\right) \dots\dots (6.26)$$

The feature based on variance is defined as

$$VAR_{F_8} = VAR|_{F=1,a=0,b=1,c=1} \dots\dots (6.27)$$

The extracted features for series and shunt faults at different locations along the winding model are shown in figure 6.32 and6.33 respectively.

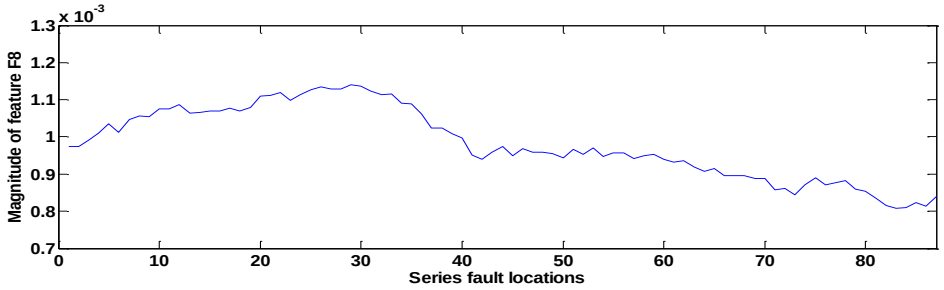


Figure 6.28: Variations of feature F8 for series faults only

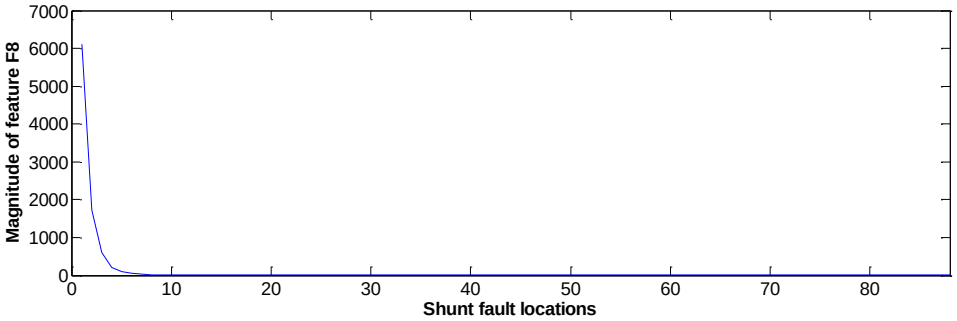


Figure 6.29: Variations of feature F8 for shunt fault only

6.4 Classification with morphological features

To classify the extracted features from morphologically transformed tank current waveforms, Rough Set Theory based classification tool has been employed. As already mentioned classification problem of series and shunt faults are dealt separately. In order to identify a fault as series and shunt, un-normalised value of the feature F_9 i.e. $OCM2_{max}$ is used. $OCM2_{max}$ is in the range of 0.68 to 0.758 for series faults and for shunt faults the range becomes 0.76 to 35.8. Hence two separate tables one for series and one for shunt faults are given in table 6.1 and 6.2 respectively in normalised form.

Table 6.6: Normalised decision table for Series faults

Objects	Condition Attributes													Decision Attributes: Location of Series Faults
	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	
1	0.94	0.91	0.97	0.23	0.96	0.50	0.45	0.38	0.93	0.92	0.76	0.95	0.96	Series1
2	1.00	1.00	1.00	0.21	0.96	0.74	0.26	0.55	0.98	0.94	0.60	0.95	0.97	Series1
3	0.87	0.88	0.73	0.09	0.98	0.86	0.20	0.69	0.96	0.95	0.54	0.96	0.98	Series2
4	0.62	0.62	0.50	0.30	1.00	0.92	0.16	0.81	0.96	0.96	0.49	0.96	0.98	Series2
5	0.54	0.60	0.51	0.31	0.77	0.93	0.09	0.83	0.95	0.95	0.48	0.97	0.99	Series3
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
84	0.31	0.32	0.26	0.82	0.21	0.78	0.90	0.13	0.90	0.85	0.73	0.95	0.92	Series28
85	0.32	0.33	0.26	0.88	0.24	0.72	0.90	0.15	0.90	0.85	0.75	0.95	0.92	Series28
86	0.31	0.32	0.26	0.86	0.15	0.58	1.00	0.04	0.90	0.84	0.89	0.94	0.91	Series29
87	0.33	0.33	0.28	0.81	0.34	0.36	0.97	0.08	0.90	0.84	0.95	0.94	0.92	Series29

Table 6.7: Normalised table for Shunt faults

Objects	Condition Attributes													Decision Attributes: Location of Shunt Faults
	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	
1	1.0000	1.0000	1.0000	1.00	0	1.00	0.79	0.27	1.00	1.00	1.00	1.00	0.96	Shunt1
2	0.2704	0.2648	0.2811	0.55	0.53	0.65	0.75	0.31	0.66	0.65	0.78	0.54	0.97	Shunt1
3	0.1358	0.1353	0.1613	0.34	0.78	0.40	0.64	0.38	0.50	0.40	0.65	0.37	1.00	Shunt2
4	0.0562	0.0566	0.0606	0.37	0.69	0.24	0.11	0.46	0.38	0.24	0.56	0.30	0.97	Shunt2
5	0.0386	0.0388	0.0435	0.27	0.81	0.15	0.24	0.55	0.31	0.15	0.48	0.27	0.95	Shunt3
. . .	. . .	. . .	. . .	. . .	. . .				. . .	. . .	. . .	. . .	. . .	. . .
84	0.0010	0.0011	0.0009	0.73	0.43	0.00009	0.99	0.29	0.02	0.0006	0.05	0.48	0.09	Shunt28
85	0.0010	0.0010	0.0009	0.72	0.47	0.00008	0.99	0.24	0.02	0.0005	0.05	0.49	0.09	Shunt29
86	0.0009	0.0011	0.0009	0.76	0.43	0.00007	0.99	0.14	0.02	0.0005	0.05	0.49	0.09	Shunt29
87	0.0009	0.0011	0.0009	0.74	0.40	0.00006	0.99	0.06	0.02	0.0005	0.04	0.50	0.09	Shunt29
88	0.0009	0.0011	0.0009	0.72	0.40	0.00005	1	0	0.02	0.0005	0.04	0.50	0.09	Shunt29

The normalised tables 6.1 and 6.2 are then discretized, to reduce $card(V_q)$ using the MD heuristics as explained in chapter 4 along with theoretical details. Table 6.3 and 6.4 presents the discretized decision table for series and shunt fault respectively.

Table 6.8: Discretized decision table for Series faults

Objects	Condition Attributes													Decision Attributes: Location of Series Faults
	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	
1	5	5	5	0	8	3	4	4	0	1	5	1	1	Series1
2	5	5	5	0	8	5	3	5	0	1	3	1	1	Series1
3	5	5	5	0	8	7	2	7	1	2	2	1	1	Series1
4	4	4	3	1	8	7	2	8	1	2	2	1	1	Series2
5	3	3	2	2	6	7	1	8	1	2	2	1	1	Series2
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
84	1	1	1	6	0	6	9	1	0	1	4	1	0	Series28
85	1	1	1	7	0	5	9	1	0	1	5	1	0	Series29
86	1	1	1	7	0	4	10	0	0	0	6	0	0	Series29
87	1	1	1	6	1	2	10	1	0	0	7	0	0	Series29

Table 6.9: Discretized decision table for Shunt faults

Objects	Condition Attributes													Decision Attributes: Location of Series Faults
	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	
1	1	1	1	8	0	2	7	2	4	2	4	4	10	Shunt1
2	1	1	1	5	4	2	6	2	4	2	4	3	10	Shunt1
3	1	1	1	2	7	2	5	3	4	2	4	2	10	Shunt2
4	1	1	1	3	6	1	0	4	3	1	4	1	10	Shunt2
5	0	0	0	2	7	1	1	5	3	1	3	1	10	Shunt3
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
84	0	0	0	6	3	0	9	2	0	0	1	3	1	Shunt28
85	0	0	0	6	3	0	9	2	0	0	1	3	1	Shunt29
86	0	0	0	7	4	0	9	1	0	0	1	3	1	Shunt29
87	0	0	0	6	3	0	9	0	0	0	0	3	1	Shunt28
88	0	0	0	6	3	0	9	0	0	0	0	3	1	Shunt29

The discretised data is subdivided into training and testing subsets. In the present approach the winding model with its 88 discs is divided into 29 sections with each section containing 3 discs except the last section which contains 4 discs. Hence, the aim would be to identify the fault within 4% of the winding as accurately as possible. Out of the three fault data obtained for any given section two are used for training the Rough set based classifier whereas the remaining one set of data is used for testing the classifier. Hence, for a series fault where there are 87 fault current waveforms available, 58 (i.e. 29×2) waveform data are used for training and 29 for testing. Similarly, for shunt fault 59 (i.e. 29×2+1) numbers of training sets are available and the number of testing sets are again 29.

Granulation as discussed in chapter 4 then helps in forming the indiscernibility relation I_B and is defined as

$$I_B = \{(\sigma_i, \sigma_j) \in \cup^2 \mid \forall q \in B, f(\sigma_i, q) = f(\sigma_j, q)\} \dots\dots (6.28)$$

This indiscernibility relation can reduce a decision table. This can be done by keeping only one element of an equivalence class and also keeping those attributes which preserve the indiscernibility relation. Thus obtained *minimal* sets of attributes are called *Reduct*. Table 6.5 and 6.6 below gives the *Reduct* tables for series and shunt faults respectively. The theoretical details of obtaining the *Reduct* is presented in chapter 4.

Table 6.10: Simplified form of decision table for Series faults

Objects	Condition Attributes													Decision Attributes: Location of Series Faults
	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	
1	5	5	5	0	8	×	×	×	0	1	×	1	1	Series1
2	5	5	5	0	8	×	×	×	0	1	×	1	1	Series1
4	×	×	×	×	×	7	×	8	1	2	2	1	1	Series2
5	×	×	×	×	×	7	×	8	1	2	2	1	1	Series2
7	3	3	2	×	×	8	1	×	1	2	×	1	1	Series3
8	3	3	2	×	×	8	1	×	1	2	×	1	1	Series3
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
82	1	1	1	×	0	6	9	2	0	1	4	1	0	Series28
83	1	1	1	×	0	6	9	2	0	1	4	1	0	Series28
85	1	1	1	7	0	×	×	×	0	×	×	×	0	Series29
86	1	1	1	7	0	×	×	×	0	×	×	×	0	Series29

It can be seen from the presented discretised decision tables that none of the features were found to be dispensable for series fault case. For the case of shunt fault localization it was found that features F_4 and F_5 were dispensable and hence are removed in the subsequent table.

Table 6.11: Simplified form of decision table for Shunt faults

Objects	Condition Attributes											Decision Attributes: Location of Series Faults
	F_1	F_2	F_3	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	F_{13}	
1	1	1	1	2	×	2	4	2	4	×	10	Shunt1
2	1	1	1	2	×	2	4	2	4	×	10	Shunt1
4	×	×	×	1	×	×	3	1	×	1	10	Shunt2
5	×	×	×	1	×	×	3	1	×	1	10	Shunt2
7	0	0	0	×	0	6	2	×	2	×	×	Shunt3
8	0	0	0	×	0	6	2	×	2	×	×	Shunt3
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.	.
82	0	0	0	0	9	2	0	0	1	3	1	Shunt28
83	0	0	0	0	9	2	0	0	1	3	1	Shunt28
85	0	0	0	0	9	×	0	0	×	3	1	Shunt29
86	0	0	0	0	9	×	0	0	×	3	1	Shunt29
87	0	0	0	0	9	×	0	0	×	3	1	Shunt29

In the next step decision rules are generated from the final table of core and *Reducts* in the form of IF…THEN rules. Tables 6.7 and 6.8 show these rules for series and shunt faults respectively.

Table 6.12: Decision Rule for localisation of Series faults

Decision Rule	Statement of Rule	
	IF	**THEN**
1	(F_1=5∧ F_2=5∧ F_3=5∧ F_4=0∧ F_5=8∧ F_9=0∧ F_{10}=1∧ F_{12}=1∧ F_{13}=1)	Fault type is Series1
2	(F_6=7∧ F_8=8∧ F_9=1∧ F_{10}=2∧ F_{11}=2∧ F_{12}=1∧ F_{13}=1)	Fault type is Series2
3	(F_1=3∧ F_2=3∧ F_3=2∧ F_6=8∧ F_7=1∧ F_9=1∧ F_{10}=2∧ F_{12}=1∧ F_{13}=1)	Fault type is Series3
. .	. .	. .
28	(F_1=1∧ F_2=1∧ F_3=1∧ F_5=0∧ F_6=6∧ F_7=9∧ F_8=2∧ F_9=0∧ F_{10}=1∧ F_{11}=4∧ F_{12}=1∧ F_{13}=0)	Fault type is Series28
29	(F_1=1∧ F_2=1∧ F_3=1∧ F_4=7∧ F_5=0∧ F_9=0∧ F_{13}=0)	Fault type is Series29

Table 6.13: Decision Rule for localisation of Shunt faults

Decision Rule	Statement of Rule	
	IF	**THEN**
1	(F_1=1∧ F_2=1∧ F_3=1∧ F_6=2∧ F_8=2∧ F_9=4∧ F_{10}=2∧ F_{11}=4∧ F_{13}=10)	Fault type is Shunt1
2	(F_6=1∧ F_9=3∧ F_{10}=1∧ F_{11}=2∧ F_{12}=1∧ F_{13}=10)	Fault type is Shunt2
3	(F_1=0∧ F_2=0∧ F_3=0∧ F_7=0∧ F_8=6∧ F_9=2∧ F_{11}=2)	Fault type is Shunt3
. .	. .	. .
28	(F_1=0∧ F_2=0∧ F_3=0∧ F_6=0∧ F_7=9∧ F_8=2∧ F_9=0∧ F_{10}=0∧ F_{11}=1∧ F_{12}=3∧ F_{13}=1)	Fault type is Shunt28
29	(F_1=0∧ F_2=0∧ F_3=0∧ F_6=0∧ F_7=9∧ F_9=0∧ F_{10}=0∧ F_{12}=3∧ F_{13}=1)	Fault type is Shunt29

Using the rules generated above the remaining data i.e. the testing set is tested for series and shunt fault localization respectively. The classification accuracy for series faults localization was found to be 86.2% (i.e. 25 out of 29 cases) and the accuracy class for shunt fault localization was about 96.6% (i.e. 28 out of 29 cases).

6.5 Classification with S-transform features

Similar to the process adopted for classification of features based on morphology, the features based on Stockwell transform are also classified. Here also the classification problem of series and shunt faults are dealt separately. In case of feature extraction and classification based on Stockwell transform features to identify a fault as series and shunt, un-normalised value of the feature F_5 i.e. magnitude of Entropy with arbitrary constants set at a value of F=1, a=0, b=1 and c=0.7 is used. Based on magnitude of feature F_5 two types of fault i.e. series and shunt faults are identified. The value of the feature F_5 is in the range of 5.9 to 6 for series faults and for shunt faults the range becomes 0.3 to 1.3.

Hence two separate tables one for series and one for shunt faults are given in table 6.9 and 6.10 respectively in normalised form.

Table 6.14: Normalised decision table for Series faults

Objects	Condition Attributes								Decision Attributes: Location of Series Faults
	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	
1	0.9673	0.9282	0.9798	0	0.9963	0.9968	0.9638	0.8543	Series1
2	0.9895	0.9594	1	1	0.9971	0.9974	0.9827	0.8530	Series1
3	0.9749	0.9676	0.9108	0.9858	0.9971	0.9944	0.9858	0.8681	Series2
4	0.9726	0.9746	0.8810	0.9865	0.9971	0.9974	0.9885	0.8862	Series2
5	0.9816	0.9853	0.8812	0.9870	0.9971	0.9973	0.9934	0.9077	Series3
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
84	0.8540	0.8610	0.7558	0.9917	0.9998	0.9999	0.9280	0.7099	Series28
85	0.8585	0.8603	0.7723	0.9923	0.9999	0.9999	0.9274	0.7210	Series28
86	0.8458	0.8468	0.7672	0.9930	0.9999	0.9999	0.9199	0.7136	Series29
87	0.8518	0.8504	0.7797	0.9932	0.9999	0.9998	0.9217	0.7372	Series29

Table 6.15: Normalised decision table for Shunt faults

Objects	Condition Attributes								Decision Attributes: Location of Shunt Faults
	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	
1	1	1	1	0.3748	0.3021	0.376	1	1	Shunt1
2	0.5958	0.6925	0.5736	0.3488	0.3396	0.3647	0.9236	0.2821	Shunt1
3	0.3806	0.4821	0.3718	0.3422	0.3813	0.3616	0.8248	0.0968	Shunt2
4	0.2172	0.3331	0.2074	0.3502	0.4264	0.3678	0.7160	0.0356	Shunt2
5	0.14	0.2468	0.1371	0.3638	0.4646	0.3773	0.6390	0.0156	Shunt3
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.
84	0.00096	0.0016	0.0012	0.9931	0.9854	0.9668	0.0559	2.45e-7	Shunt28
85	0.00091	0.0016	0.0011	0.9961	0.9910	0.9967	0.0552	2.27e-7	Shunt29
86	0.00091	0.0015	0.0011	1	1	1	0.0544	2.17e-7	Shunt29
87	0.00089	0.0015	0.0011	0.9977	0.9969	0.9982	0.0541	2.07e-7	Shunt29
88	0.00087	0.0015	0.0011	0.9974	0.9960	0.9981	0.0540	2.02e-7	Shunt29

These tables are similar in all respect to the tables presented for morphologically obtained features with one exception that the features depicted here are extracted using modified S-Transform and that here only 8 features are considered. The series fault contains 87

cases out of which 58 (i.e. 29×2) waveform data are used for training and 29 for testing. Similarly, for shunt fault 59 (i.e. 29×2+1) numbers of training sets are available and the number of testing sets are again 29.

Table 6.11 and 6.12 provides the discretised decision table for the Stockwell transform extracted features evaluated using MD heuristics with only the training sets shown.

Table 6.16: Discretised decision table for series faults

Objects	Condition Attributes						Decision Attributes: Location of Series Faults
	F_2	F_3	F_5	F_6	F_7	F_8	
1	6	8	0	0	5	6	Series1
2	7	8	4	3	6	6	Series1
4	8	5	3	3	8	6	Series2
5	8	6	3	3	9	6	Series2
7	9	6	3	3	8	7	Series3
8	9	6	3	3	8	8	Series3
. .	. .	. .	. .	. .	. .	. .	. .
82	2	0	12	11	2	1	Series28
83	2	0	12	11	2	0	Series28
85	1	0	12	9	1	1	Series29
86	0	0	12	9	1	1	Series29

As evident from the discretised table that features F_1 and F_4 were found to be dispensable and hence is removed from the table. The reduced table is shown in table 6.13.

Table 6.17: Discretised decision table for shunt fault

Objects	Condition Attributes				Decision Attributes: Location of Shunt Faults
	F_1	F_6	F_7	F_8	
1	27	1	28	27	Shunt1
2	27	1	28	27	Shunt1
4	26	1	27	26	Shunt2
5	26	1	27	26	Shunt2
7	25	1	26	25	Shunt3
8	25	1	26	25	Shunt3
. .	. .	. .	. .	. .	. .
82	2	7	1	2	Shunt28
83	1	7	1	1	Shunt28
85	1	7	1	1	Shunt29
86	1	7	1	1	Shunt29
87	1	7	1	1	Shunt29

Here, the features F_2, F_3, F_4, F_5 and F_7 are dispensable and hence are removed. The reduced decision table is shown in table 6.14.

Table 6.18: Reduced decision table for series faults

Objects	Condition Attributes						Decision Attributes: Location of Series Faults
	F_2	F_3	F_5	F_6	F_7	F_8	
1	×	8	×	×	×	6	Series1
2	×	8	×	×	×	6	Series1
4	8	×	3	3	×	6	Series2
5	8	×	3	3	×	6	Series2
7	9	6	3	3	8	×	Series3
8	9	6	3	3	8	×	Series3
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
82	2	0	12	11	2	×	Series28
83	2	0	12	11	2	×	Series28
85	×	0	12	9	1	1	Series29
86	×	0	12	9	1	1	Series29

Table 6.19: Reduced decision table for Shunt faults

Objects	Condition Attributes				Decision Attributes: Location of Shunt Faults
	F_1	F_6	F_7	F_8	
1	27	1	28	27	Shunt1
2	27	1	28	27	Shunt1
4	26	1	27	26	Shunt2
5	26	1	27	26	Shunt2
7	25	1	26	25	Shunt3
8	25	1	26	25	Shunt3
. .	. .	. .	. .	. .	. .
82	×	6	1	×	Shunt28
83	×	6	1	×	Shunt28
85	1	7	1	1	Shunt29
86	1	7	1	1	Shunt29
87	1	7	1	1	Shunt29

Tables 6.15 and 6.16 gives the rule base generated for identification of locations of series and shunt faults respectively.

Table 6.20: Rule base generated for series fault localization

Decision Rule	Statement of Rule	
	IF	**THEN**
1	(F_3=8∧ F_8=6)	Fault type is Series1
2	(F_2=8∧ F_5=3∧ F_6=3∧ F_8=6)	Fault type is Series2
3	(F_2=9∧ F_3=6∧ F_5=3∧ F_6=3∧ F_7=8)	Fault type is Series3
. .	. .	. .
28	(F_2=2∧ F_3=0∧ F_5=12∧ F_6=11∧ F_7=2)	Fault type is Series28
29	(F_3=0∧ F_5=12∧ F_6=9∧ F_7=1∧ F_8=1)	Fault type is Series29

Table 6.21: Rule base generated for shunt fault localization

Decision Rule	Statement of Rule	
	IF	**THEN**
1	(F_1=27∧ F_6=1∧ F_7=28∧ F_8=27)	Fault type is Shunt1
2	(F_1=26∧ F_6=1∧ F_7=27∧ F_8=26)	Fault type is Shunt2
3	(F_1=25∧ F_6=1∧ F_7=26∧ F_8=25)	Fault type is Shunt3
. .	. .	. .
28	(F_6=6∧F_7=1)	Fault type is Shunt28
29	(F_1=1∧ F_6=7∧ F_7=1∧ F_8=1)	Fault type is Shunt29

Using the rules generated above the remaining data i.e. the testing set is tested for series and shunt fault localization respectively. The classification accuracy for series faults localization was found to be 96.5% i.e. 28 out of 29 fault cases were identified correctly and the accuracy class for shunt fault localization was about 100% i.e. all the cases were identified accurately.

6.6 Classification Accuracy

The main aim of using classification techniques in this work is to predict the type of fault and their location along the winding model with highest possible accuracy. Table 6.17 given below provides a brief summary of the % accuracies obtained for localization of series and shunt short circuit faults based on mathematical morphology and Stockwell transform with a Rough set based classifier.

Table 6.22: Comparison table

Features based on Mathematical morphology			**Features based on Stockwell transform**		
Series Faults			***Series Faults***		
Number of individual fault events	Number of correctly identified events	% accuracy	Number of individual fault events	Number of correctly identified events	% accuracy
29	**25**	**86.2%**	**29**	**28**	**96.6%**
Shunt Faults			***Shunt Faults***		
Number of individual fault events	Number of correctly identified events	% accuracy	Number of individual fault events	Number of correctly identified events	% accuracy
29	**28**	**96.6%**	**29**	**29**	**100%**

It is clearly evident from the table that features based on Stockwell transform provides better classification accuracy as compared to features based on mathematical morphology. It is also to be noted that the presented work aims at localising faults within 3.5% of the transformer winding which is highest possible acquired so far.

6.7 Conclusion

From the above results it can be concluded that Stockwell transform based features were better capable of identifying and localizing series and shunt short circuit faults that would develop along the length of a transformer as compared to Mathematical Morphology based features. Hence, the time frequency representation based signal processing tool i.e. S-transform, was better suitable for analysing impulse fault data due to their superiority in analysing nonstationary signals as compared to time domain approach.

Chapter 7

Conclusion and Future Scope

Chapter 7: Conclusion & Future Scope

7.1 Conclusion

Impulse testing of transformers is one of the widely accepted tests for evaluation of insulation integrity of large power transformers. International and national standards such as IEC 60076-4, 2002 and IS 2026 (part-3) of 2009 provides details of the experimental procedure to conduct impulse test of a transformer and to draw inference from the experimental results. But, the established techniques only helps to identify a faulty transformer from a batch of several manufactured units, by comparison of voltage and current waveforms at full and reduced voltages. These techniques gives no physical insight into the type of fault that causes the error in observations nor any idea about the location of the fault within the insulation system. Moreover, the accuracy of identification of faults depended heavily on human observations and experience which may not always be equally reliable.

In order to take into consideration the above mentioned facts researchers have used various Machine Learning techniques developed to identify not only a faulty transformer but also the type of faults and the approximate location of the fault within the insulation. One of the most significant implication of using such automated techniques is the decrease in dependence over human intervention. With little or no human intervention the fault identification procedure can be made as accurate as possible, depending upon the algorithm developed, and also a higher degree of consistency can be maintained. In this present work also one of the machine learning technique known as classification is used to identify a faulty transformer, identify the type of fault existing within the insulation system and find their approximate position in the whole insulation system with as much high accuracy as possible. A classifier is developed based on Rough Set theory for the implication of the fault identification and localization methodology. Details into the theoretical and mathematical definition of Rough Set technique is studied.

In order to decrease the burden over the developed classifier and to make the analysis of fault events easy two different feature extraction methodology is studied and applied. Advantages of pre-processing of acquired fault current data for extraction of significant features is studied. Two signal processing techniques namely Mathematical Morphology

and Stockwell transform are studied in details and significant features are extracted from them.

Before implementing the developed algorithm for fault identification and localization it needs to trained and tested. To achieve this artificial series and shunt faults are simulated over an analog model constructed at the High Tension laboratory at Jadavpur University. The fault current waveforms are then acquired using a digital oscilloscope connected across a coaxial shunt which in turn is connected to the model.

For fault localization purpose the winding is divided into 29 sections such that each section consists of three discs. The algorithm aims at identifying the fault in these 29 sections i.e. within 3.5% of the winding. Out of the three acquired fault current data from each section two are used for training the developed classifier and one from each section is used to test. At last a comparison is made between the classifier with morphological features as one instance of feature extraction technique and another with features extracted from Stockwell transform of the fault current waveforms.

For identification of series or shunt faults in both the cases, the numerical range of one of the un-normalised feature is used. In the case of localization of faults along the analog winding model it was found that:

i. Based on Mathematical Morphological features of fault current waveforms the series faults were localized with 86.2% accuracy and the shunt faults with 96.6% accuracy.
ii. With features based on Stockwell transform of the fault current waveforms the series faults are localized with 96.6% accuracy and the shunt faults are localized with 100% accuracy.

Hence, it can be concluded that features based on Stockwell transform gives better performance in localizing various fault events as compared to features based on Mathematical Morphology.

7.2 Future Scope

The presented work aims at identifying and localizing series and shunt short circuit faults occurring along the length of a transformer winding. The developed algorithm is trained with artificial faults present at any one location at a given time. It remains to verify whether the algorithm could be modified for, trained and tested for multiple faults occurring simultaneously along the transformer winding. In such a case a rugged and

generalised algorithm is needed to be developed that could identify any number of faults along with their approximate locations with sufficient accuracy. Again, it could be said that the proposed classifier algorithm works for a particular rated transformer whose model is developed for experiments. An algorithm needs to be developed that could work with sufficient accuracy with any type of transformer with different ratings, size and makes and needs to be tested along with any type of fault occurrence conditions.