IJDACR
ISSN: 2319-4863

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 6, Issue 8, March 2018)

Detection and Classification of Transmission Lines Faults using FFT and K-Nearest Neighbour Classifier

Nilesh S.Wani[1], Dr. R. P. Singh[2], Dr. M. U. Nemade[3]
[1,3]Electrical Engineering
[1,3]Sri SatyaSai University of Technology & Medical Sciences
[1,3]Opposite OILFED, Bhopal Indore Highway, Pachama, Sehore (India) 466001
[1,3]K. J. Somaiya Institute of Engineering and Information Technology, Sion, Mumbai. 400022
nileshswani@gmail.com[1], prof.rpsingh@gmail.com[2], mnemade@somaiya.edu[3]

Abstract – **This paper presents a strategy for identifying the fault and its classification, in an electrical power distribution system. The strategy is based on a very simple technique, known as the k nearest neighbours (KNN), which simply estimates a distance between the characteristics that describe the data to be classified. When a new datum is presented to the proposed algorithm, it is classified with the same type of the example that is determined to be the closest one. For the creation of the mathematical model it is essential to have a database. The database consists of input data and output data, the input data are the detail coefficients obtained from the decomposition of the current and voltage signals using the Fourier Transform. Meanwhile, the output data are the labels assigned and with which the model can identify and classify the different types of faults. Both current signals and voltage signals are generated based on an extensive simulation of faults along the longest transmission line that has a test system.**

Keywords –**DFT, FFT, KNN, STFT, Transmission Line Faults.**

I. INTRODUCTION

The monitoring of the condition of a machine, system or process is the most efficient way to manage maintenance in many industries, since the economic savings can be exceptional in many cases, without considering the material (and human) damages that can be avoided. . Maintenance based on the condition of a machine or process that requires continuous operation, demand applications (computational algorithms) that determine or estimate the internal condition of the machine while it is in operation.

In the case of high-speed machining systems, there are two major alternatives for performing this task: vibration analysis and lubricant analysis; being the study of the vibrations the one of greater practical interest. A machining center even under normal conditions has a certain level of vibrations, when a fault occurs, these vibrations change in such a way that they can be associated with said fault [1].

Among the most important aspects of the quality of electric power are taken into account, the waveform, the continuity of the service and customer service. The location of faults is an issue that is closely related to the continuity of supply.

Distribution systems are the main means of supplying electric power to end users. A fault in these systems causes an interruption of the service to the users, and consequently economic losses due to process stoppage, loss of information, damages and detriments in machinery and supplies, among others [2]. Worldwide, several studies have been carried out in which it has been established that 80% of the total interruptions are caused by faults in the distribution systems [3].

In spite of the reliability that an energy distribution system can provide, it is inevitable that there are interruptions in the provision of the service due to line faults. The networks experience faults caused by storms, lightning, faults in the insulation and short circuits caused by plants, animals and other external agents. Therefore, the reliability of the distribution system in case of any fault that has occurred depends on its rapid isolation, repair and restoration of the service. The restoration of the service can be considerably accelerated if the location of the fault is determined with reasonable accuracy. The conventional way to find a fault is through visual inspection, which takes a long time in addition to requiring a considerable number of personnel [4] [5].

There are permanent and transient faults, and a fault locator provides information for both types of faults [6]. Generally, transient faults cause less damage and cannot be located with a simple inspection. Faced with these faults, the locators allow detecting weak points of the power system and take corrective

IJDACR International Journal Of Digital Application & Contemporary Research

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 6, Issue 8, March 2018)

actions to avoid further damage due to relapse of the fault. When a permanent fault occurs, there is a protection system with relays that typically makes a correct and quick disconnection of the fault based on measurements taken in real time.

Due to the constant growth of the energy demand, it is necessary to guarantee the continuity and quality of the energy supply, for which it is essential to have an adequate operation of the electric power systems. For this, it is important in the previous design, to consider the existence of faults in all the elements that are part of the system, especially in the transmission lines since they extend along the path of the system, so that by remaining exposed to any weather condition they become more vulnerable to possible faults. The Fourier transform is a fundamental mathematical tool when performing the analysis of certain disturbances caused by faults that can occur in the electrical system, widely used with periodic signals. However, it has limitations for the analysis of non-periodic signals, which makes it a little effective method for the analysis of transients because it detects the presence of a certain frequency but does not provide information on the evolution over time of the spectral signal characteristics [7], [8]. The analysis using the wavelet transform overcomes this limitation, performing a processing of the signal, whose result provides information in time and frequency [9]. The FFT transform will be used since it is a very efficient tool for the analysis of non-stationary signals, in this case the transients produced by faults. Under this background, the present titling project aims to determine the advantages and disadvantages, with respect to conventional methods, of using the FFT transform as a tool for fault detection.

II. Justification

Finding faults in transmission lines has been a concern for network operators, because it is known that lines are characterized by having long lengths and when a fault occurs, it is difficult to know the precise point where it is generated. This causes great uncertainty and inconvenience, since it becomes totally necessary to review section by section to locate the point causing the anomaly.

Emerging technologies such as phasor measurement units greatly improve the development of applications and tools that allow monitoring and control of the network in a more versatile, safe and reliable manner. One of these applications is the development of methods for the location of disturbances in the transmission lines, which allow the location quickly, to achieve the timely taking of protection and control actions.

Two types of fault location methods are considered in transmission lines, impedance-based methods, which are based on mathematical models of network elements and wave-based methods, which use high-frequency components, since they exploit the transient signals generated by the fault and are based on the correlation between the waves that travel forward and backward which propagate through the transmission lines [10].

If the exact point where the fault occurs is found, the repair of the faulty elements can be carried out in a timely manner, achieving a quick rehabilitation of the line, and avoiding possible fines or penalties for the delayed restoration of the service. For this reason, it is important to make a thorough review of the troubleshooting methods proposed so far, to observe and detail their best approximations and propose a new algorithm that has the same function, but with the objective that their performance and accuracy be the most optimal.

III. Characteristics Of The Distribution System That Affect The Location Of Faults

For the transmission systems excellent algorithms have been developed that allow the precise location of the point of fault [11] [12] that normally use measurements in both terminals of the line. However, such algorithms are not applicable or present a great uncertainty in the distribution networks, because they have their own characteristics that differentiate them from the transmission lines.

In recent years, different methods have been adapted and designed for application in distribution networks. Most are adaptations of existing methods used in transmission systems, therefore important characteristics of distribution systems have been ignored as described below.

A. Voltage and Current Monitoring only at the Circuit Header

Digital protection devices are generally equipped with data storage functions that are automatically activated when they detect a fault. Additionally they have the ability to record and store the waveform of the voltage and current signals in the substation. From these signals, the phasors of the pre-fault voltage and current fundamental can be obtained, and immediately after the transient caused by the fault.

The most popular methods that can be used in distribution systems are those that use the fundamental components of voltage and current in a terminal of the line, for economy, easy implementation and the limitations of the distribution system.

IJDACR
International Journal Of Digital Application & Contemporary Research

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 6, Issue 8, March 2018)

B. System Imbalance

The methods that are normally used to locate faults in distribution systems use sequence networks. The method of the sequence components generates three independent sequence impedances only in case the system is perfectly balanced. Distribution systems are usually unbalanced due to unbalanced line impedances, single-phase and biphase lateral impedances as well as unbalanced loads. Therefore work with sequencing networks in distribution systems can lead to large errors in the estimation of fault [13].

C. Unawareness of the Load Variation in Each Bar

In a distribution system the power that is delivered varies in time, therefore, the impedances are not constant. A localization method must take into account the charge variation. Because there are no measures in each bar to know each consumption in detail, it is necessary to make an approximation to take into account this phenomenon.

D. Non-Homogeneous Networks

The networks of the transmission systems are not homogeneous, since they generally have different conductor gauges in a circuit. This makes it necessary to be careful when using distance methods, since the distance to a fault is not equivalent to the proportion of the line.

E. Effect of Fault Resistance

In transmission systems it is common to ignore the effect of the fault resistance and methods are used without taking into account this phenomenon, such as that of the reactance [13]. Because the higher the load impedances are higher, in the case of distribution systems a 40Ω fault, cannot be considered very small with respect to the load. Therefore, the localization method must take into account the effects of the variation of the fault resistance and consequently the change of load flow.

IV. Fourier Analysis

It allows the decomposition of a signal in its components of sines and cosines for different frequencies and whose sum corresponds to the original signal [14]. The Fourier transform expresses a signal $f(t)$, of period T, as the sum of its harmonics. In this way a temporary signal $f(t)$ of period T, can be expressed as observed in equation (1):

$$f(t) = a_o + \sum_{k=1}^{\infty}\left(a_k \cos\frac{2\pi kt}{T} + b_k \sin\frac{2\pi kt}{T}\right) \quad (1)$$

Where the coefficients of the Fourier transform are given by equations (2), (3) and (4):

$$a_o = \frac{1}{T} + \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t)dt \quad (2)$$

$$\underset{k\geq 1}{a_k} = \frac{2}{T} + \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t)\cos\frac{2\pi kt}{T}dt \quad (3)$$

$$\underset{k\geq 1}{b_k} = \frac{2}{T} + \int_{-\frac{T}{2}}^{\frac{T}{2}} f(t)\sin\frac{2\pi kt}{T}dt \quad (4)$$

It is also possible to write the Fourier transform in integral form and complex form as shown below in equations (5) and (6) respectively:

$$f(t) = \int_{-\infty}^{\infty} F(\omega)\cdot e^{j2\pi ft}d\omega \quad (5)$$

$$F(W) = \int_{-\infty}^{\infty} f(t)\cdot e^{-j2\pi f}dt \quad (6)$$

Being $f(t)$ the representation of the signal in the time domain and $F(\omega)$ the representation of the signal in the frequency domain. In this way, thanks to the Fourier transform, we can represent a signal that is originally in the time domain $f(t)$ to another, in the frequency domain $F(\omega)$ or in the domain of the frequency $F(\omega)$ to the time domain $f(t)$ [15].

The conditions that must be fulfilled by the signals so that they can be analyzed by considering the Fourier transform are the following:

- The signals must be non-periodic.
- They must comply with Dirichlet conditions.

Below is the graph of an aperiodic signal together with its Fourier transform.

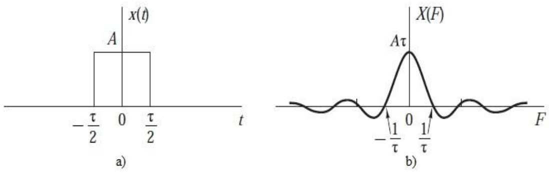

Figure 1: (a) Rectangular impulse, (b) Its Fourier transform [15]

A. Discrete Fourier Transform (DFT)

The discrete Fourier transform is a particular application of the Fourier transform. It is known that computers work only with discrete values, that is why the numerical calculation of the Fourier transform $f(t)$ requires discrete values of $f(t)$, that is, values that have the form $x[n]$ with $[n = 0,1,2,..]$ This means that it will be possible to obtain $F(\omega)$ only for discrete values of ω, that is, values will be obtained for the transformation of the form $x[k]$ with $[k = 0,1,2,..]$ [15]. Equation (7) defines the discrete Fourier transform:

$$x[k] = \frac{1}{N}\sum_{n=0}^{N-1} x(n)e^{-\frac{j2\pi kn}{N}}$$
$$n, k = 0,1,2, \ldots, (N-1) \quad (7)$$

Where:

N: Number of samples
n: n^{th} original sample
k: k^{th} term of the DFT

IJDACR
ISSN: 2319-4863
IJDACR
International Journal Of Digital Application & Contemporary Research
International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 6, Issue 8, March 2018)

In order to process a signal with the help of the DFT, it is necessary to sample it taking a single part of it. This implies that the segment being analyzed is a single period of the signal in which N number of samples are present. It turns out to be a highly inefficient algorithm due to the computational time and the use of resources that it uses to calculate the N^2 complex multiplications and $N(N-1)$ complex sums necessary to perform the calculation of a DFT of N samples [9].

B. Fourier Fast Transform (FFT)

Equation (7) supposes to realize a great number of operations since to obtain N components in frequencies of N samples, it is necessary to realize N^2 complex multiplications. To solve this problem comes the FFT that obtains the same results as the DFT but only with $N\log_2 N$ complex multiplications [15]. The FFT is an algorithm that allows to perform the DFT in a more efficient and faster way. This algorithm manages to introduce mathematical shortcuts with which the number of operations prior to the calculation of the DFT is considerably reduced.

C. Short Time Fourier Transform (STFT)

Many of the signals that describe some of the existing systems in nature are described by behaviours where the frequency changes eventually over time, this type of signals are known as non-stationary signals. The Fourier transform is not the most indicated to perform the analysis of said signals that is why mathematical tools have been developed such as the STFT (Short Time Fourier Transform).

A short-time Fourier transform can be seen as a revised and perfected version of the Fourier transform. This transformation allows to overcome the deficiencies that the Fourier transform presents and therefore the analysis of non-stationary signals, since it divides the signal into small segments in which it can be assumed that the signal is stationary and calculates the Fourier transform of each segment separately [15].

The Fourier transform is limited by the Heisenberg principle of uncertainty, where it is established that it is impossible to obtain an exact time-frequency representation of a signal, that is, it is not possible to know the value of the frequency in a moment of time, it is only possible to know the components of existing frequencies in a time interval [15].

Time-frequency localization can only be obtained with limited precision, this precision is given by the width of the time window used [15]. As the width of the window is reduced, the resolution in the time domain increases and at the same time the resolution in the frequency domain decreases, on the other hand if the window is very wide, a good frequency resolution will be achieved and a poor resolution in time [15].

V. PROPOSED METHODOLOGY

For the implementation of the method, it is only necessary to know the fundamental components of voltage and current at the time of pre-fault and fault The general algorithm is divided into sequential tasks, as shown in Figure 2.

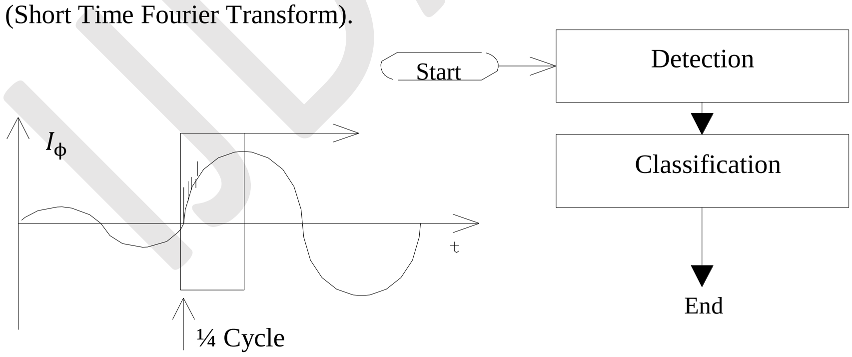

Figure 2: Task Scheduling Scheme

A. Fault Detection

This task seeks to start the algorithm, after detecting the occurrence of an event in the network. In general, the criterion commonly used in this respect is the condition that the effective value of the current in any of the system phases exceeds a preset threshold. The proposed method consists of obtaining a model, with which the type of fault can be identified and classified. According to the following procedure and with the help of MATLAB, the model is obtained:

Obtaining the Input Data Matrix: The data matrix Q is formed from the samples obtained through the fourier transform of each of the signals, both current and voltage obtained in the different simulations for each kilometre of the transmission line. The STFT of each of the signals is done to obtain the detail coefficients:

The above mentioned is summarized in equation (8):

$$STFT(\tau, f) = \int_0^t [y(t).\omega^*(t-t')].e^{j2\pi f t}dt \quad (8)$$

IJDACR
International Journal Of Digital Application & Contemporary Research

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 6, Issue 8, March 2018)

Where:
$y(t)$: Signal defined in t
$\omega^*(t - t')$: Window function delayed by a value of t'.
f: Frequency
t: Time
t': Time in which the centre of the window is indicated $(\omega(t))$

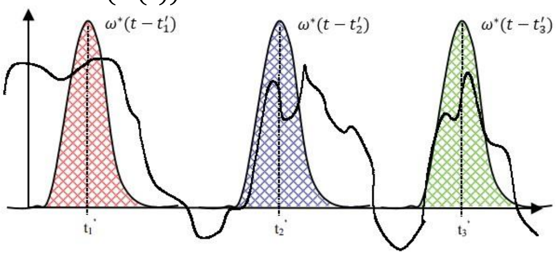

Figure 3: Graphical explanation of the STFT [15]

Figure 3 shows three windows corresponding to three Fourier transforms in three different times. The window function is Gaussian, the red function shows the window located at $t = t_1'$, the blue colour function at $t = t_2'$ and green at $t = t_3'$. For each instant t' and frequency f is calculated a new coefficient of the Fourier transform.

B. Fault Classification using KNN

In this section, the basic fundamentals of the technique based on the nearest neighbours are displayed. A practical and easily applicable way to predict or classify a new datum, based on known or past observations, is the nearest neighbour technique. As an example, the case of a physician who is trying to predict the outcome of a surgical procedure can predict that the result of the patient's surgery will be that of the most similar patient who knows, who has been subjected to the same procedure. This may be somewhat extreme, since a single similar case in which surgery failed may have an excessive influence on other cases, slightly less similar, in which the surgery was a success. For this reason the method of the nearest neighbour is generalized to use the k nearest neighbours.

This technique is based, simply, on "remembering" all the examples that were seen in the training stage. When a new datum is presented to the learning system, it is classified according to the behavior of the nearest data (Moreno, 2004).

As an example, you can see how is the classification of the nearest neighbour. We have the data belonging to the training set, as shown in Figure 4 (triangles and squares), and we want to know the label of a new datum (marked as x). Then the procedure to follow is to find the example that is nearest to this new data x, and assign its label (triangle), as shown in Figure 5.

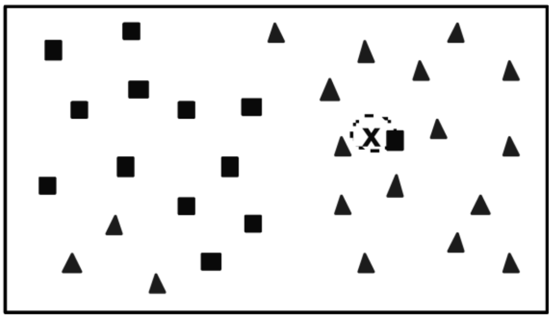

Figure 4: Location of a new data between known data

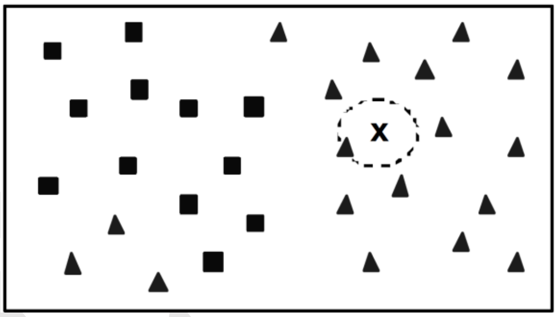

Figure 5: Prediction of the class of a new data with respect to the nearest neighbour

Now, if we consider the case where there is a square inside the data corresponding to the triangles (noise), and we want to classify the new data (x), using the nearest example as shown in Figure 6, There is a possible error.

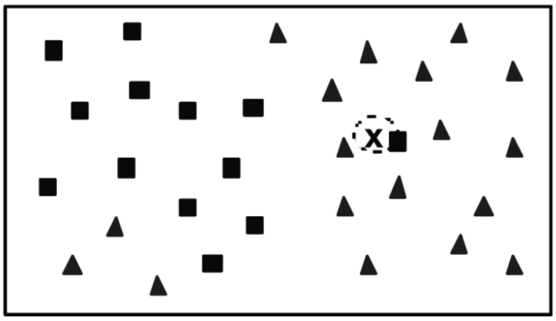

Figure 6: Predicting the class of a new data with respect to the nearest neighbor, the known data contains noise

It can be noticed due to noise, the new data is classified as square. To consider the problem of noise, you can change the classification algorithm and use a greater number of neighbors, and thus generate the label of the new data using simple majority, and not a single data. This generalization of the method is called k-nearest neighbors (Moreno, 2004). In this case, k = 5 is made and it can be seen that the new data belongs to the class triangles, as graphically shown in Figure 7.

IJDACR
International Journal Of Digital Application & Contemporary Research

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 6, Issue 8, March 2018)

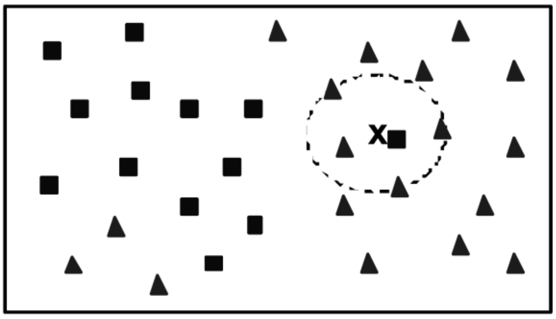

Figure 7: Prediction of the class of a new data with respect to the five nearer neighbors, the known data contain noise

With this new approach, the problem of noise is solved. The larger k is, the more robust the noise classification. However, the value of k has a limit, if any new data is maximized, it will always have the class label with the most data in the training set (Aha et al., 1991). For example, for the case presented in Figure 4, if k = 31 is assigned, the new data will always be classified as triangle because there are 18 triangles and 13 squares.

VI. Simulation Results

A. Simulation Parameters

The cases considered for the evaluation of the algorithm, for the particular system previously described, were the following:

1. Type of fault: The eleven types of possible faults (AG, BG, CG, ABC, BCG, CAG, AB, BC, CA, ABC and ABCG) were considered.
2. Fault impedance (resistive only - Rf): 100Ω, 300Ω, 500Ω.
3. Fault timing: 0.23 sec.
4. Fault location: 5 to 195 km.

B. Results

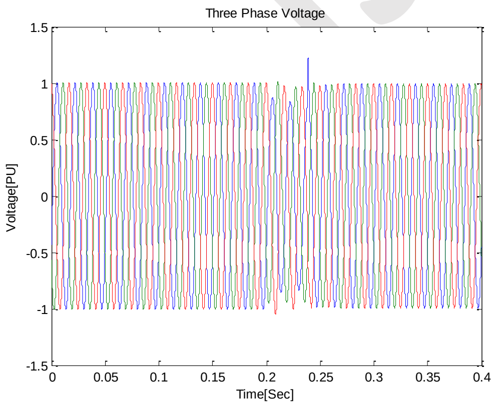

Figure 8: Voltage response of three phase circuit

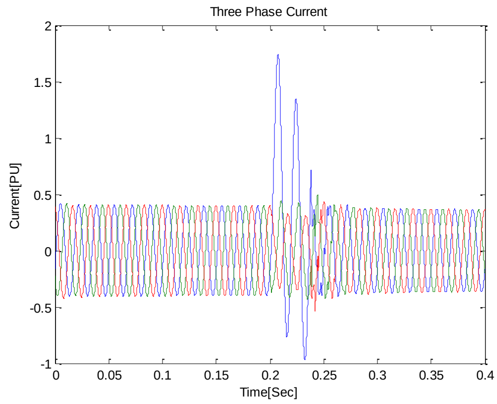

Figure 9: ABC fault in three phase circuit

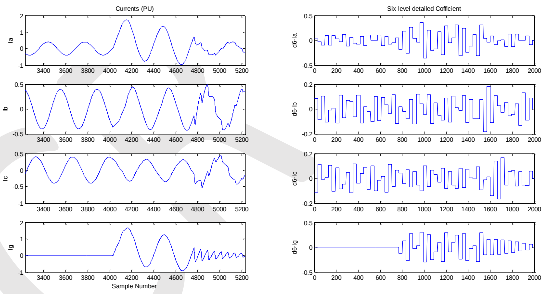

Figure 10: Detailed coefficient for various faults

VII. Conclusions

This paper describes the application of Fourier analysis and k nearest neighbours for the protection of energy transport lines. Based on the analysis of Fourier theory and the simulation of line protection, the potentialities and advantages of the Short Time Fourier Transform to obtain and analyze the transient components of voltage and current signals, highlighting in these methodologies the relative low computational effort, allowing them to be used as part of a high protection system speed. In the different tasks studied, we observed the effectiveness of each of the algorithms and criteria proposed, highlighting the case of classification and location of faults.

Reference

[1] Koley, E., Verma, K. and Ghosh, S., 2015. An improved fault detection classification and location scheme based on wavelet transform and artificial neural network for six phase transmission line using single end data only. *SpringerPlus*, *4*(1), p.551.
[2] Chui, C.K., 2016. *An introduction to wavelets*. Elsevier.
[3] Eldin, A.H. and Refaey, M.A., 2011. A novel algorithm for discrimination between inrush current and internal faults in power transformer differential protection based on discrete wavelet transform. *Electric Power Systems Research*, *81*(1), pp.19-24.
[4] Cheong, W.J. and Aggarwal, R.K., 2002, October. Accurate fault location in high voltage transmission systems comprising an improved thyristor controlled series capacitor model using wavelet transforms and neural network. In *Transmission and Distribution Conference*

International Journal of Digital Application & Contemporary Research
Website: www.ijdacr.com (Volume 6, Issue 8, March 2018)

and Exhibition 2002: Asia Pacific. IEEE/PES (Vol. 2, pp. 840-845). IEEE.

[5] Marti, J.R., 1982. Accuarte modelling of frequency-dependent transmission lines in electromagnetic transient simulations. *IEEE transactions on power apparatus and systems*, (1), pp.147-157.

[6] Chiesa, N., Mork, B.A. and Høidalen, H.K., 2010. Transformer model for inrush current calculations: Simulations, measurements and sensitivity analysis. *IEEE Transactions on Power Delivery*, *25*(4), pp.2599-2608.

[7] Schettler, F., Huang, H. and Christl, N., 2000. HVDC transmission systems using voltage sourced converters design and applications. In *Power Engineering Society Summer Meeting, 2000. IEEE* (Vol. 2, pp. 715-720). IEEE.

[8] Grainger, J.J. and Stevenson, W.D., 1994. *Power system analysis* (Vol. 621). New York: McGraw-Hill.

[9] Agarwal, S., Swetapadma, A., Panigrahi, C. and Dasgupta, A., 2017, April. Fault detection in direct current transmission lines using discrete fourier transform from single terminal current signals. In *Electronics, Materials Engineering and Nano-Technology (IEMENTech), 2017 1st International Conference on* (pp. 1-5). IEEE.

[10] Dasgupta, A., Nath, S. and Das, A., 2012. Transmission line fault classification and location using wavelet entropy and neural network. *Electric Power Components and Systems*, *40*(15), pp.1676-1689.

[11] Takagi, T., Yamakoshi, Y., Baba, J., Uemura, K. and Sakaguchi, T., 1981. A New Alogorithm of an Accurate Fault Location for EHV/UHV Transmission Lines: Part I-Fourier Transformation Method. *IEEE Transactions on Power Apparatus and Systems*, (3), pp.1316-1323.

[12] Horowitz, S.H. and Phadke, A.G., 2008. *Power system relaying* (Vol. 22). John Wiley & Sons.

[13] Choi, M.S., Lee, S.J., Lee, D.S. and Jin, B.G., 2004. A new fault location algorithm using direct circuit analysis for distribution systems. *IEEE transactions on power delivery*, *19*(1), pp.35-41.

[14] C. Sidney Burrus, R. A. Gonipath, H. Guo. 1998. Introduction to Wavelets and Wavelet Transforms. *Prentice Hall*.

[15] Parlitz, U., Chua, L.O., Kocarev, L., Halle, K.S. and Shang, A., 1992. Transmission of digital signals by chaotic synchronization. *International Journal of Bifurcation and Chaos*, *2*(04), pp.973-977.