JAVA Journal of Electrical and Electronics Engineering
Volume 12, Number 2, October 2014

Sensitive Detection for High Impedance Fault on Transmission Line using Wavelet and Naive Bayes

Dimas Anton Asfani, I Made Yulistya Negara, and Muhammad Rajavalens
Electrical Engineering Department
Institut Teknologi Sepuluh Nopember (ITS)
Surabaya, Indonesia
anton@ee.its.ac.id, yulistya@ee.its.ac.id, va_lens@yahoo.com

***Abstract*— This paper presents the combination method of wavelet transform and naive bayes classifier to detect and classify the high impedance fault of transmision line. The fault's current signal is transformed using wavelet. The transformed signal produces coefficients with certain pattern according to the type of fault that occurs. Then, coefficients of transformed signal are variated become 7 variabels, based on the algorithm of classification. Those variabels are classified using naive bayes classifier to detect and classify the fault of transmission line. Three types of mother wavelet used in this study are Daubechies-5 (Db5), Daubechies-8 (Db8), and Coiflet-5 (Coif5). Every mother wavelet produces different coefficients. However, they have similar pattern to the algorithm of classification. The highest accuracy of classification was obtained using coeffisients of Daubechies-5 (Db5) at 5th level. The classification accuracy is 97.09% using normal distribution, and 99.78% using kernel distribution.**

Keywords*— *Fault detection, Fault Classification, Transmision line, Wavelet Transform, Naive Bayes Classifier

I. INTRODUCTION

Transmission line is a very important component in electric power systems. Transmission line must be able to ensure the continuous availability of electrical energy in every load that is connected to the system. Most of fault in electric power system occur on transmission line [1]. Fault of the transmission line could hinder the continuity of electrical energy distribution. Therefore, faults need to be detected, classified, and repaired as soon as possible.In a transmission line protection, detection and classification of faults are two important things that need to be dealt with reliably and accurately. One of the most common faults is short circuit. The existence of this fault can not be eliminated. It required treatment as soon as possible to avoid further consequence of the fault. To be able to handle it well, then this type of fault must be known accurately.

In most of power system relaying algorithm, the first step always involves fault detection and the next step involves the classification of fault. Such information is necessary in the determination of distance algorithm. Most of the algorithms for detection and classification of fault on transmission line using wavelet multi resolution analysis (MRA) are based on the measurement and comparison of sharp variations in the value of the current for the three phases in the first stage MRA detail signals extracted from the original signal [2]. The Sum of the coefficients are evaluated during fault condition for detection and classification of fault. In [3], describes the algorithm for the detection and classification using mother wavelet Daubechies-4 (Db4) at third level of decompositon is good enough. This study uses mother wavelet decomposition up to 5 level for detection and classification algorithms.Different simulations also conducted in this study, where the fault resistance will be set to a large value (11 kΩ) so that the increase of fault current caused by short circuit will be very small. This very small increase would basically look like signal in normal operation.This study uses combination method of wavelet transform and Naive Bayes classifier to detect and classify the type of faulton transmission line. The fault's current signal is transformed using wavelet. The transformed signal produces coefficients with certain pattern according to the type of fault that occurs. Then, coefficients of transformed signal are variated become some variabels, based on the algorithm of classification. Those variabels are classified using naive bayes classifier to detect and classify the fault of transmission lines.

II. RESEARCH METHOD

Basic theories used in the execution of this study include the theory of short circuit and their types, the theory of wavelet transform and what filter will be used, and Naive Bayes classifier theory.

A. *Short Circuit Fault*

In the operation of power systems, faults may occur that result interference of the power supply. Faults of the transmission line occur about 85% s/d 87% of the entire fault on the power system [1].Faults are obstaclefor system which operating or power distribution systems state that deviate from normal conditions [4]. Most of the common faults are short circuit. Short circuit is a fault that occurs because of an error between voltage parts.

Faults can be classified into several types, which aresingle phase to ground fault (LG), a two-phase fault (LL), two-phase to ground fault (LLG), three-phase fault (LLL), and three-phase to ground fault (LLLG). Three-phase fault and three-phase to ground fault is types of symmetrical fault. While single phase to ground fault, two-phasefault, two-phaseto ground fault is a type of unsymmetrical fault

B. *Wavelet Transform*

Wavelet transform is divided into two types, continuous wavelet transform and discrete wavelet transform [5]. By providing a wave function f (t), continuous wavelet transform (CWT) produces too many wavelet coefficients (wavelet transform Coefficient / WTC) . This causes the resulting data becomes redundant (redundancy) [6]. And data redundancy problem can be solved by the use of discrete wavelet transform (DWT). In DWT, only a few samples are taken from WTC. Which means, DWT reduces excess WTC from CWT.

Figure 1 shows a diagram of DWT decomposition which based on Mallat algorithm. The input signal is divided into two sub-signals with low frequency l(n) and high frequency h(n) section. Sub-section low-frequency signals divided again into two sub with different frequencies. This process happens over and over again in accordance with the number of decomposition levels of wavelet transform used

Figure 1. Decomposition diagram of DWT

Daubechies is one of the most famous mother wavelet in wavelet research. Daubechies' filter (db) have the order of N (N positive integer numbers). Have a support width length of 2N-1. The filter length is 2N.Coiflet is designed to generate value trend resembles the original signal value. All types of wavelet Coiflet defined in the same way. Coiflet's filter (Coif) has Order N (N = 1, 2, 3, 4, 5). Its support width is 6N -1. The length of the filter is 6N.

C. *Naïve Bayes Classifier [7]*

Bayes classifications are statistical classification which can predict the probability class member. For simple Bayes classification known as Naive Bayes classification, it can be assumed that the effect of an attribute value of a given class is independent of other attributes. This assumption is called class conditional independence which made to facilitate the calculations. In Naive Bayes, it is predicted that attributes are assumed to not depend on class or were not affected by other attributes.Given set of data with many attributes, it will be very difficult in the computational calculations. To alleviate this problem, it's common to make the assumption that each of the two coordinates of the document is a vector, when viewed as a random variable, statistically independent of each other, this is not dependent assumptions that can be expressed by the following equation:

$$P(\mathrm{d}|\mathrm{c}_j) = \prod_{k=1}^{m} P(w_k|c_j) \tag{7}$$

Where m is the number of data,w in the collection of data. Classification probabilities using this assumption is called Naive Bayes classifier and most of the calculations using the probability approach to classify the data.

Naive Bayes classifier algorithm is to first calculate the probability of each term which occurs in the training data, then combine the probabilities associated with the data found in the test data to classify and predict that the possibility of this document include different classes . Eventually will provide test data to mark the class with highest probability.

III. RESULTS AND ANALYSIS

A. *Transmission Line Model*

This study uses a model of 500kV transmission line with a line length of 250 km. The line parameterused for modeling is the parameters of the line that connects the Surabaya Barat bus (East Block) and Ungaran bus (west block) in the Jamali transmission system.

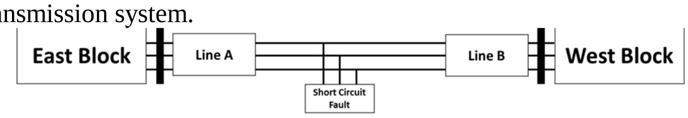

Figure 2. Transmission line's model of Ungaran – Surabaya Barat

The east block is power source, the parameters used are Vrms 489000 volts,phase angel A11.868, Frequency 50 Hz, Yg internal connections,3-phase short circuit level 38623.6 MVA, base voltage Vrms500000 ph – ph, X/R ratio 28, and the bus-type is swing. While for the west block, the parameters used are Vrms 467000 volts, phase angle A 6.108 , a frequency of 50 Hz, internal connections clincher, 3-phase short circuit level 18878.8MVA, the base voltage Vrms 500000 ph - ph , the X / R ratio 12 , and the type of bus is swing.

The line parameters used were a frequency of 50 Hz, the resistance of positive sequence and zero sequence[r1, r0] 0.0293 and 0.3864 Ω/km, positive and zero sequence inductance [l1,l0] 0,896x10-3and 4,1264 x10-3H/km, positive and zero sequence capacitance [c1,c0] 12,74 x10-9dan 7,751 x10-9F/km, and the line length is 250 km .

B. *Short Cirtuit Simulation*

Short circuit simulation on a transmission line's model is to simulate some of the conditions on the transmission line, which is the state when the normal condition and also in the event of a short circuit fault. For normal conditions, a total of 110 conditions are simulated with currents between 450A to 500A. Meanwhile, for the condition of a short circuit, it's includes several simulated conditions and different parameters. The type of fault that simulated are single-phase short circuit to ground (a-g), short circuit between phases (a-b), two- phase short circuit to ground (a-b-g), three- phase short circuit to ground (a-b-c-g). The performed simulations also have parameters such as distance fault (250 kilometer intervals along 25 kilometers), Fault Inception Angle (0^o - 180^o with 20^o intervals), fault resistance (0.001Ω , 35Ω , 65 Ω , 95 Ω , 8000 Ω , 9000 Ω , Ω 10000 , 11000 Ω).

Fault Inception Angel (FIA) is the initial angle of the fault. The greater the FIA, the fault start time will be more backward. For the current shape differences based on Fault Inception Angle (FIA) on single phase to ground fault, fault resistance 95 Ω, and the distance of 250km , can be seen in Figure 3.

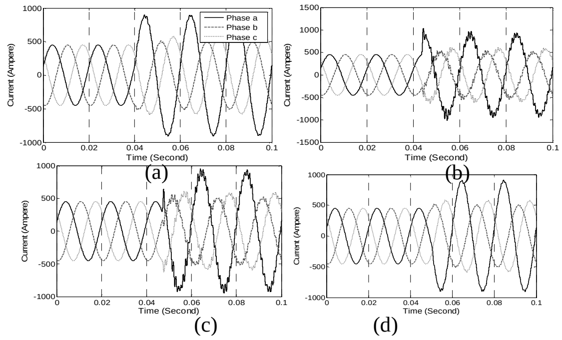

Figure 3.Current shape differences based on*Fault Inception Angle*
(a) *Fault Inception Angel* 0º (b) *Fault Inception Angel* 60º
(c) *Fault Inception Angel* 120º (d) *Fault Inception Angel* 180º

Fault resistance (Rf) is a very influential on the fault current surge. The larger the value, the smaller Rf fault current surge will be. The smaller the fault current, it will be more difficult to detect as well.In Figure4(e), 4(f), 4(g), 4(h), are a short circuit fault simulation with high resistance. This causes a fault current surge is very small so it looks like a normal operation signal. The current value differences based on fault resistance (Rf) for single phase to ground short circuit, FIA 0o, and the distance of 250 km can be seen in Figure 6. And the comparison is shown in Table 1.

From the short circuit simulation results, current samples taken with a sample frequency of 20 kHz along 0.1 seconds. So we get 5current cycle with 2000 samples at each simulated signal. The sample then flows in the filter with 3 kind of different mother wavelet.

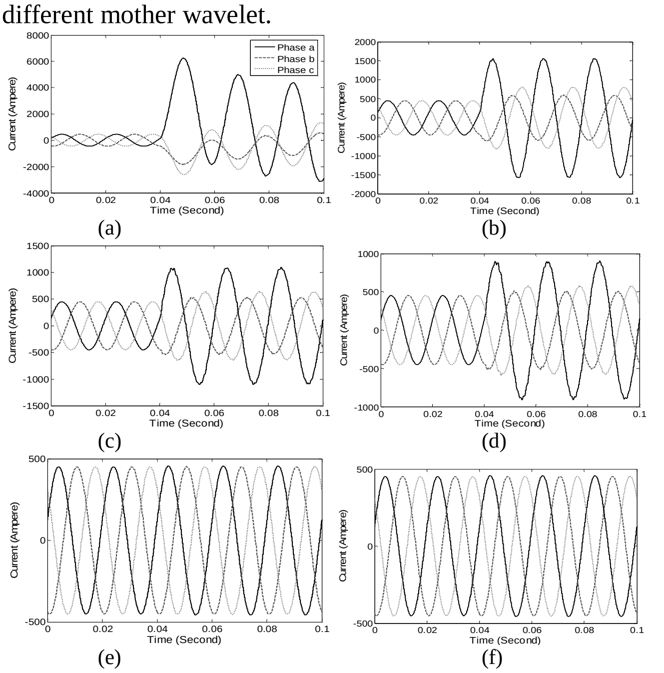

Figure 4.Current Value differences based on fault resistance (Rf)
(a) Rf = 0.001 Ω (b) Rf = 35 Ω (c) Rf = 65 Ω (d) Rf = 95 Ω(e) Rf = 8000 Ω (f)Rf = 9000 Ω (g) Rf = 10000 Ω(h) Rf = 11000Ω

Table 1. Differences of fault current value based on fault resistance (*Rf*)

Rf	**Normal Current (A)**	**Lowest fault current**		**Highest fault current**	
		(A)	**surge**	**(A)**	**surge**
0	450	6240	1286.00%	115400	25544.00%
35	450	1560	246.00%	11070	2360.00%
65	450	1080	140.00%	6200	1278.00%
95	450	900	100.00%	4420	882.00%
8000	450	457	1.56%	507	12.60%
9000	450	457	1.56%	501	11.33%
10000	450	456	1.33%	496	10.22%
11000	450	456	1.33%	492	9.33%

C. *Wavelet Transform Simulation*

Wavelet transform simulation use three kinds of mother wavelet which are Daubechies-5 (DB5), Daubechies-8 (db8), and Coiflets-5 (Coif5). Phase current signals obtained from the simulation process is transformed into a wavelet shape using 3 different filters as much as five levels of decomposition to get wavelet signal detail, d1for level 1, d2for level 2, d3 for level 3, d4 for level 4, and d5 forlevel 5. Signal decomposition details of each level can be seen in Figure 7.

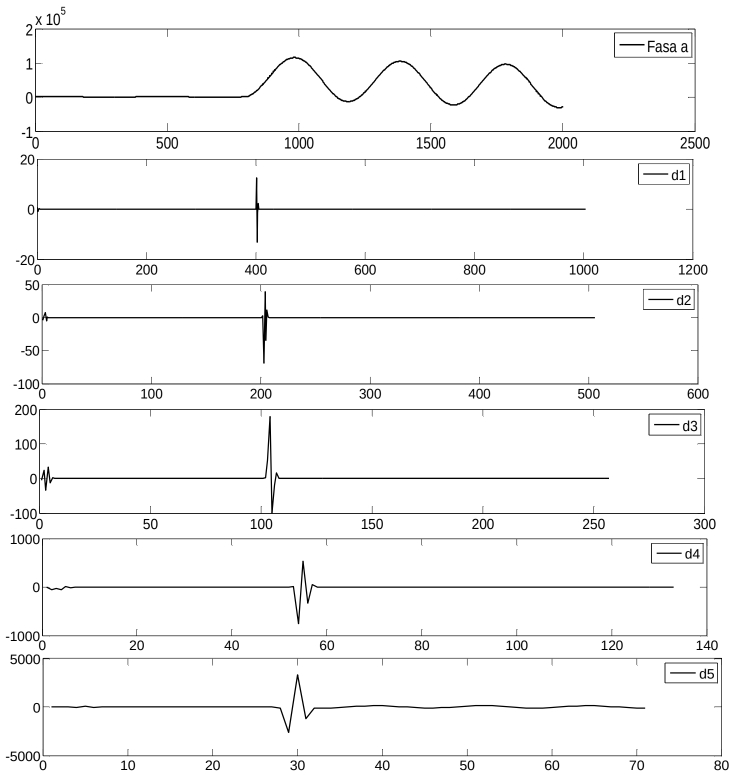

Figure 5. Detail signal of five level decomposition using*Db5*filter

From Figure 5, it shows that when the sample into two fifths of the total sample wavelet signals detail have been decomposed, there was a sharp spike coefficient. This is because the type of Db5mother wavelet which detects differences in fault current of normal current. Any differences that occur outside the current value of the signal characteristic properties, then the current differences that will result in the value of the wavelet detail coefficients are high, while the normal current value will have a coefficient of 0.

Of each level, which has been transformed and summed signal equal to 3 cycles, then the sum of the coefficients called coefficient S, so that will be a 3 coefficients are obtained, namely Sa for the phase A, Sb for phase b, and Sc for phase c. Equations used in the determination of the coefficient S at each level are:

$$S_n = \sum_{i=1}^{\frac{1200}{2^n}} x_{(i)} \qquad (8)$$

Where n is the level of decomposition of the signal and i is an integer. 1200 is the number of samples of the original 3signal cycle .

Obtained coefficient then varied into several variables to look at the characteristics of a variable pattern of each type of fault. Variables that are formed are SumSaSbSc (Sa, Sb, and Sc Addition), SumSaSb (Sa and Sb Addition), SumSaSc (Sa and Sc Addition), SumSbSc (Sb and Sc Addition), |Sa|-|Sb| (Sa and Sb absolute value reduction), |Sa|-|Sc| (Sa and Sc Reduction in absolute value), and |Sb|-|Sc| (Sb and Sc absolute value reduction) . Those variables will be used as features in classification using Naive Bayes.

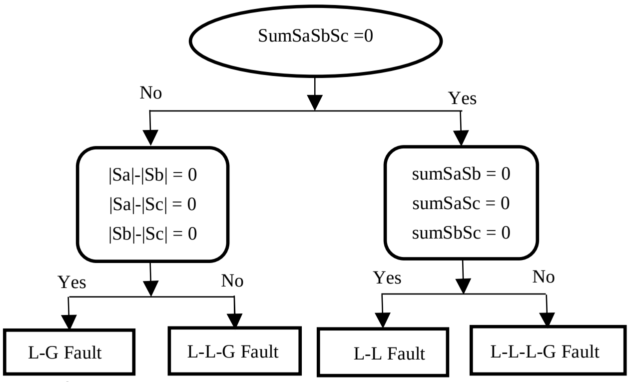

Figure 6.Classification algorithm using wavelet MRA

The algorithm in Figure 6 shows the variation patterns are formed from each of the disorder based on the results of simulation and analysis [3] .

Figure 7 shows the influence of Fault Inception Angle of the Sa, Sb, Sc, and SumSaSbSc voltage absolute coefficients. Data sample of single phase to groundfault, taken from a distance of 0 km, fault resistance 0 Ω and on the 5th level.

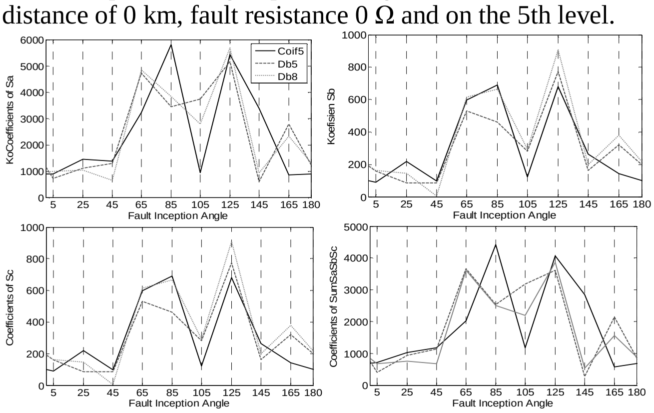

Figure 7.The influence ofFault inception Angletoabsolute value of coefficientS

D. Classification Using Naïve Bayes

Naive Bayesian classifier is used to classify the types of fault based on the features that have been described previously. In this classification, two types of distributions areused, which are normal distribution and the kernel distribution.Each level of the wavelet decomposition with different mother wavelet are taken 90% of the amount of data to be the training data, and 10 % are used to test data. All data are the result of the wavelet filter DB5, Db8, Coif5 on decomposition level 1-5. 3630 total data for each level. 3267 total data used for training data and 363 data used for test data.

After trained a training system for each type of mother wavelet at each level, the training system is used to test the accuracy of the test data. Before used to test data, it was examined on the training system itself (validation). Simulation conditioned on five conditions, the first condition is when the normal condition (class 0), the second is when the single phase to ground fault (class 1), the third is a two-phase short circuit fault (class 2), the fourth is a two-phase to groundfault (class 3) and the fifth is a three-phase to ground fault (class 4).

Table 3. Sample data

Condition	**Training Data**		**Test data**	
	Total data	***Filter***	**Total data**	***Filter***
Normal	99	*Db5, Db8, Coif5* (level 1-5)	11	*Db5, Db8, Coif5* (level 1-5)
1 phase to ground fault	792		88	
2 phase fault	792		88	
2 phase to ground fault	792		88	
3 phaseto grond fault	792		88	
Total	3267	15	363	15

Validation test results show thatfor normal distribution, the best results obtained by using the coefficients of the mother wavelet Db5 on fifth level. If using the mother wavelet coefficients of the other two at fifth level, the accuracy rate shown does not vary much. For the kernel distribution, the best results are also obtained using the coefficients of the mother wavelet Db5 on fifth level. This is because coefficients of mother wavelet Db5 on fifth level generates a sizeable and more densely distributed for each variable, so that the classification accuracy rate is higher.

Table 4. Accuracy of validationusing normal Distribution

Filter	**Level Accuracy (%)**				
	1	**2**	**3**	**4**	**5**
Db5	71.75	86.41	85.22	94.09	97.09
Db8	71.69	92.41	84.96	96.11	96.97
Coif5	57.64	86.23	86.13	93.69	96.48

Table 5. Accuracy of validationusing kernelDistribution

Filter	**Level Accuracy (%)**				
	1	**2**	**3**	**4**	**5**
Db5	98.75	97.49	98.29	98.32	99.78
Db8	98.93	97.46	98.65	98.56	99.36
Coif5	92.29	96.27	98.62	98.75	99.39

Table 6. Test data classification accuracy using normal distribution

Filter	Level Accuracy (%)				
	1	**2**	**3**	**4**	**5**
Db5	73.55	87.88	83.20	94.21	100
Db8	73.00	93.11	83.20	95.87	100
Coif5	55.10	84.85	87.88	93.98	98.90

Table 7. Test data classification accuracy using kernel distribution

Filter	Level Accuracy (%)				
	1	**2**	**3**	**4**	**5**
Db5	97.25	93.39	99.17	98.35	100
Db8	96.14	98.90	98.90	99.45	100
Coif5	82.09	98.35	98.35	99.17	100

After performing the validation for the training data, a number of data test are tested to determine the results of training that has been done. Classification results using test data can be seen in Table 6 and 7.The results from the table above shows the better accuracy of the validation, the better accuracy of the generated test data. But for the distribution of the kernel is not so, due to the distribution of the kernel if there is a test that is worth of data on the outer limits of the data training, then the data can't be classified. That's because the kernel generates Probability Density distribution Function (PDF) is very close to the original shape of the data distribution.

Classification of the whole simulation has been carried out, using the classification of the mother wavelet coefficients DB5on fifth level and distribution type kernels showed the best accuracy rate. Validation of classification reaches 99.78% accuracy rate. Means in the 3267 classification training data, only 7 failed data correctly classified. It happened on 5 types of data with a single phase to ground fault, where the data for the variables of the valued sumSaSbSc is very small (<10-2). The variable pattern is closer to the fault pattern of two phase faultsvariable, so the data is incorrectly classified. Errors also occur in one type of data with a two phase to ground fault, which for the variable |Sb|-|Sc|data-value is very small (<10-1), so the pattern is closer to the fault pattern of single phase to ground. Also at sumSaSbSc of 1 data with a very small (<10-2)variable data-value, so the pattern is closer to the fault pattern of two phases fault. Details of classification validation results using the type of mother wavelet DB5 on fifth level with kernel distribution can be seen in Table 8. Class 0 for normal conditions, class 1 for LG fault, class 2 for LL fault, class 3 for LLG fault, and class 4 for LLLG fault.

Table 8. Positive Predictive Value of validation using Db5 5^{th} level with kernel distribution

		Output					
	Kelas	**0**	**1**	**2**	**3**	**4**	**Akurasi kelas**
Input	**0**	99	0	0	0	0	100.00%
	1	0	787	5	0	0	99.37%
	2	0	0	792	0	0	100.00%
	3	0	1	1	790	0	99.75%
	4	0	0	0	0	792	100.00%

Based on the results obtained from simulation and analysis in this study, several conclusions can be drawn which are conditions that can be detected and classified using Wavelet Transform and Bayesian Classifier is a normal condition and short circuit fault conditions . This method is able to detect and classify the type of short circuit on a transmission line with a high fault resistance. The use of kernels type distribution estimation are better than the normal type , because the kernel type is able to estimate the Probability Distribution Function (PDF) of the data variability is closer to the original form of the data distribution. The highest accuracy of classification was obtained using coeffisients of Daubechies-5 (Db5) at 5th level. The classification accuracy is 97.09% using normal distribution, and 99.78% using kernel distribution.. This is because the type of mother wavelet DB5 at 5thlevel is able to detect transient fault current conditions and generate greater detail signals coefficient than others.

REFERENCES

[1] M. Mirzaei, M.Z. A Ab Kadir, E. Moazami, H. Hizam "Review of Fault Location Methods for Distribution Power System". Australian Journal of Basic and Applied Sciences , 3(3): 2670-2676, 2009 ISSN 1991-8178

[2] Liang J, Elangovan S, Devotta JBX. "A wavelet multiresolution analysis approach to fault detection and classification in transmission lines". Electr Power Energ Sys 1998;20 (5):327–32.

[3] M. Jayabharata Reddy, D.K. Mohanta "A wavelet-fuzzy combined approach for classification and location of transmission line faults" Electrical Power and Energy Systems 29 (2007) 669–678

[4] Daman Suswanto, 2009 "Sistem Distribusi Tenaga Listrik"

[5] James S. Walker "A Primer on WAVELETS and Their Scientific Applications" second edition 2008 pp41

[6] Dimas Anton A, Adi Soeprijanto, Mauridhi Heri P, "Klasifikasi Gangguan Hubung Singkat pada Saluran Transmisi yang Dikompensasi Seri Menggunakan Kombinasi Wavelet dan ANFIS". Seminar Nasional Efisiensi dan Konservasi Energi FISERGI, Semarang, 12 Desember 2005, hal. B158-B165

[7] Aritz Pérez *, Pedro Larrañaga, Iñaki Inza "Bayesian classifiers based on kernel density estimation: Flexible classifiers" International Journal of Approximate Reasoning 50 (2009) 341–362