Rough-granular Approach for Impulse Fault Classification of Transformers using Cross-wavelet Transform

D. Dey, B. Chatterjee, S. Chakravorti and **S. Munshi**
Jadavpur University
Department of Electrical Engineering
Kolkata 700032, India

ABSTRACT

A novel approach based on information granulation using Rough sets for impulse fault identification of transformers has been proposed. It is found that the location and type of fault within a transformer winding can be classified efficiently by the features extracted from cross-wavelet spectra of current waveforms, obtained from impulse test. Results show that the proposed methodology can localize the fault within 5% of the winding length with a high degree of accuracy. The basic concepts of feature extraction using cross-wavelet transform and the method of classification of those features by rough-granular method are also explained.

Index Terms — **Cross-wavelet transform, cross-wavelet spectrum, impulse fault identification, information granulation, rough set.**

1 INTRODUCTION

IN case of failure during impulse testing of transformers the oscillographic traces of applied voltage and corresponding current waveforms are investigated for judging the condition of the transformers, which requires significant human expertise and knowledge. Impulse test is an acceptance test for power transformers. Guidelines for performing these tests are described in standards [1]. Different computer-aided methodologies, such as, Artificial Neural Network (ANN) [2], Fuzzy systems [3] and wavelet based analysis [4] have already been reported for the classification and localization of fault within the transformer winding. In the present work an information granulation based approach using Rough set analysis is proposed. In the proposed scheme transformer tank currents due to applied lightning impulse are processed by cross-wavelet transform and some significant features are extracted from the cross-wavelet spectra. A Rough-granular classifier is used as a new paradigm to determine the fault location within the transformer winding and also to identify the fault type.

It is worth mentioning here that during conventional impulse testing, current waveforms for both standard full test voltage and reduced voltage are compared. In the present approach the nature of time variation of the current waveforms for different fault conditions are investigated in time-frequency domain. These waveforms are normalized before processing. So, absolute magnitude of the applied voltage as well as that of the corresponding current waveform is not an important parameter in this case.

Manuscript received on 21 January 2008, in final form 18 March 2008.

Noise is an important issue for all types of measurements. As the impulse testing of transformer is usually performed within the shielded laboratory, severe noise contamination has not been experienced by the authors. However, in real-life experimentation it is commonly observed that the waveforms are superposed with very low level random noise. Cross-wavelet transform usually takes care of this type of random un-correlated noise. As the cross-wavelet transform gives correlation between two waves in time-frequency domain, low level random noise has minimal effect on it. So, in this present work no denoising scheme is applied for prior processing of the waveforms. However, if the noise level is found to be high then any suitable denoising technique may be applied before processing the recorded waveforms. Authors have already reported different denoising schemes employing hybrid-filtering in real-time [5] as well as wavelet based decomposition [4].

Rough set based analysis is suitable where knowledge is imprecise or the information system contains superfluous information. The data about a system can be reduced using Rough sets keeping all the information or features of the system intact [6]. Performance shows that the proposed method can identify and localize the fault accurately within 5% of the winding length with a high degree of accuracy.

2 SIMULATION OF SERIES AND SHUNT FAULT IN ANALOG AND DIGITAL MODELS

To analyze the performance of the proposed scheme, impulse faults are simulated on the analog model of a 3-MVA, 33/11 kV, three phase, 50 Hz, ON (i.e. oil natural cooling), Dy11 (i.e. vector group delta-star-11) transformer. Results of

analysis on Electromagnetic Transient Program (EMTP) based digital model are also shown in the paper.

Details of the analog model of the transformer, used in this work can be found in [4] and [7]. The model is having 89 discs. In this work the entire winding has been divided into 22 sections, namely L1, L2,…, L22 each involving approximately 5% of the total length of the winding. Each section consists of 4 discs sequentially. The terminal section L22 is having the last five discs. In the present case of study, two types of faults, series and shunt, are simulated. Series-fault (i.e. insulation failure between the discs or between turns) is simulated by short-circuiting the corresponding disc. On the other hand, for shunt-fault (i.e. failure of insulation between winding and earthed components) the disc is connected to ground. The photograph of analog model is shown in Figure 1. Corresponding EMTP based digital model is also shown in Figure 2. Implementation of the series and shunt faults are also shown in the exploded views of the Figure 2.

The types of faults in transformer windings are topologically of two types, namely, series and shunt faults. These two types mostly cover the significant faults that occur in transformers. These may characteristically be classified as static or complex. The method of simulation and analysis of partial or complex faults are different altogether from static faults. So, in the present work static series and shunt faults are analyzed. The simulation and analysis of complex faults lie within the scope of future work.

Even if the transformer winding is symmetric in geometry, while applying the impulse voltage to the line end of the winding, the other end is earthed. The impulse voltage distribution along the transformer winding is non-linear. Therefore, from earthed end or from line end of the winding no two points are symmetric from electric circuit point of view under the application of impulse voltage.

Example of recorded currents waveforms obtained from the analog model, by tank-current method for no-fault condition and for series fault at a certain disc position is shown in Figure 3. These waveforms have been normalized before processing.

3 FEATURE EXTRACTION FROM CROSS-WAVELET SPECTRUM

Cross-wavelet transform may be considered as an extension of wavelet based analysis, which has not been explored extensively in fault analysis in the field of power engineering. A very brief overview of cross-wavelet transform is given in this section. However, the detailed mathematical background can be found in [8-9].

The cross-wavelet transform of two signals, x(t) and y(t) is defined as:

$$W^{xy}(s,\tau) = \frac{1}{k_\psi}\int_{-\infty-\infty}^{+\infty+\infty} W^x(a,b)W^{y^*}(\frac{a}{s},\frac{b-\tau}{s})\frac{dadb}{a^2} \quad .$$

Here, $W^x(s,\tau)$ and $W^y(s,\tau)$ are the wavelet transform of x(t) and y(t) respectively with respect to a mother wavelet ψ(t). The choice of mother wavelet depends on the nature of the waveform being processed. In this case Morlet mother wavelet is used. Comparing the results of other mother wavelets like Paul and Gaussian derivative functions for the present problem, it is found that the performance of the scheme is better with Morlet wavelet function. k_ψ is a constant, defined as, $k_\psi = \int_{-\infty}^{+\infty}\frac{|\Psi(\omega)|^2}{|\omega|}d\omega < \infty$. The cross-wavelet spectrum gives a measure of correlation between two waveforms in time-frequency domain. So, cross-wavelet spectrum shows regions in time-frequency space where two waveforms are having high common power. The Cross-wavelet spectrum the magnitude of W^{xy} and the phase angle, $\phi = \tan^{-1}\frac{\Im\{W^{xy}\}}{\Re\{W^{xy}\}}$ are plotted.

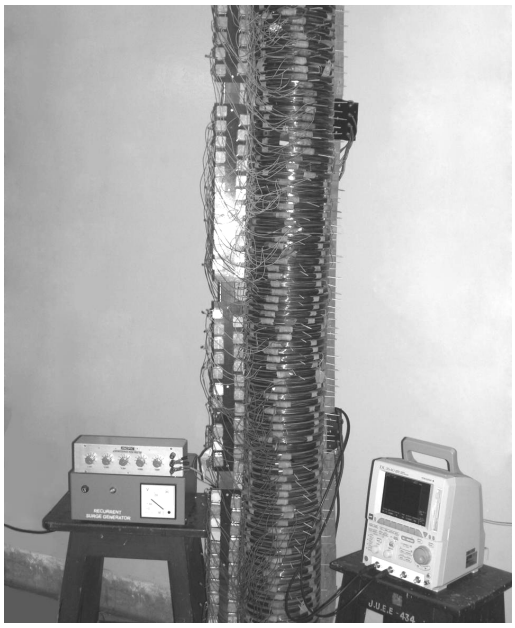

Figure 1. Photograph of the analog model of the transformer.

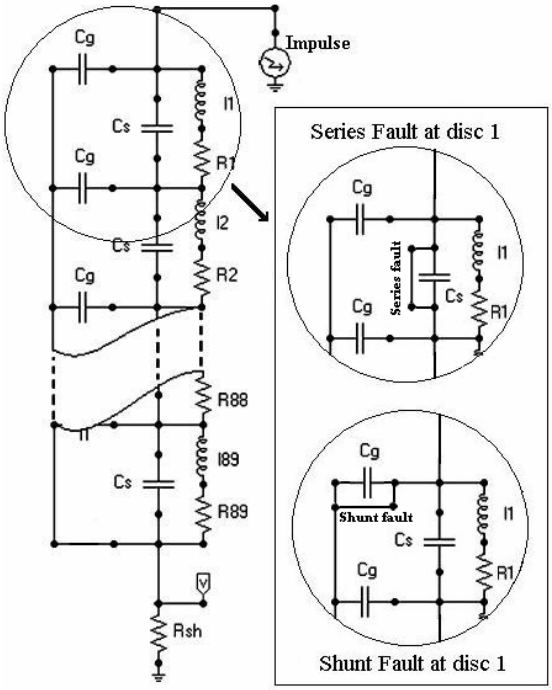

Figure 2. EMTP model of the transformer along with implementation of series and shunt fault.

The cross-wavelet spectra of different fault current waveforms with the impulse current at no-fault condition are analyzed. Typical no-fault current waveform along with a fault-current waveform is shown in Figure 3. The corresponding cross-wavelet spectrum is given in Figure 4. The 'U' shaped and black colored line shows the "cone of influence" (COI). COI indicates the region where edge effects due to zero padding are significant. Similar to spectral analysis, errors will occur at the edges (i.e. beginning or end) of the waveform in the case of cross-wavelet because of the finite length of time series. Padding with zeros introduces discontinuities and the amplitude of the coefficients near the edge decreases at larger scales as more zeroes are added. The cone of influence is the region of the cross-wavelet spectrum in which these effects become important. This is defined as the e-folding time for the autocorrelation of wavelet power at each scale. That means the power for a discontinuity drops by a factor of e^{-2} (where, e= 2.7182) and the edge effects are negligible beyond this point.

In Figure 4, $|W^{xy}|$ values at different 'time' and 'scale' are plotted. The x-axis is considered as 'time' axis and y-axis shows the 'scale', which is related to the inverse of frequency. The color of the figure at a point shows the value of $|W^{xy}|$ at that time-frequency space. The color-bar given in the right side of Figure 4 indicates the value corresponding to a color. Higher the value higher the common power at that time-frequency point.

Black arrows show the phase angle. Arrows shown in Figure 4, pointing towards right indicate "in-phase" (i.e. phase difference in zero) and arrows pointing left indicate "anti-phase" (i.e. phase difference is 180 degrees) conditions. For classification and identification of fault some features are extracted from $|W^{xy}|$. The features are described below:

1. $F_1=\dfrac{\sum_s\sum_\tau s\tau|W^{xy}(s,\tau)|}{\sum_s\sum_\tau|W^{xy}(s,\tau)|}$; 2. $F_2=\sqrt{\dfrac{\sum_s\sum_\tau s^2\tau^2|W^{xy}(s,\tau)|}{\sum_s\sum_\tau|W^{xy}(s,\tau)|}}$

3. $F_3=\dfrac{\sum_s\sum_\tau|W^{xy}(s,\tau)|}{|W^{xy}(s,\tau)|_{peak}}$; 4. $F_4=\dfrac{\sum_s\sum_\tau|W^{xy}(s,\tau)|}{(s_{max}-s_{min})(\tau_{max}-\tau_{min})}$

5. $F_5=\sqrt{\dfrac{\sum_s\sum_\tau(F_4-|W^{xy}(s,\tau)|)^2}{(s_{max}-s_{min})(\tau_{max}-\tau_{min})}}$

6. F_6= "s" at peak of $|W^{xy}(s,\tau)|$ i.e. $|W^{xy}(s,\tau)|_{peak}$

7. F_7= "τ" at peak of $|W^{xy}(s,\tau)|$ i.e. $|W^{xy}(s,\tau)|_{peak}$

Five more features, similar to F_1 to F_5, are also extracted from the phase angle $\phi(s,\tau)$ data. They are:

8. $F_8=\dfrac{\sum_s\sum_\tau s\tau|\phi(s,\tau)|}{\sum_s\sum_\tau|\phi(s,\tau)|}$; 9. $F_9=\sqrt{\dfrac{\sum_s\sum_\tau s^2\tau^2|\phi(s,\tau)|}{\sum_s\sum_\tau|\phi(s,\tau)|}}$

10. $F_{10}=\dfrac{\sum_s\sum_\tau|\phi(s,\tau)|}{|\phi(s,\tau)|_{peak}}$; 11. $F_{11}=\dfrac{\sum_s\sum_\tau|\phi(s,\tau)|}{(s_{max}-s_{min})(\tau_{max}-\tau_{min})}$

12. $F_{12}=\sqrt{\dfrac{\sum_s\sum_\tau(F_{11}-|\phi(s,\tau)|)^2}{(s_{max}-s_{min})(\tau_{max}-\tau_{min})}}$

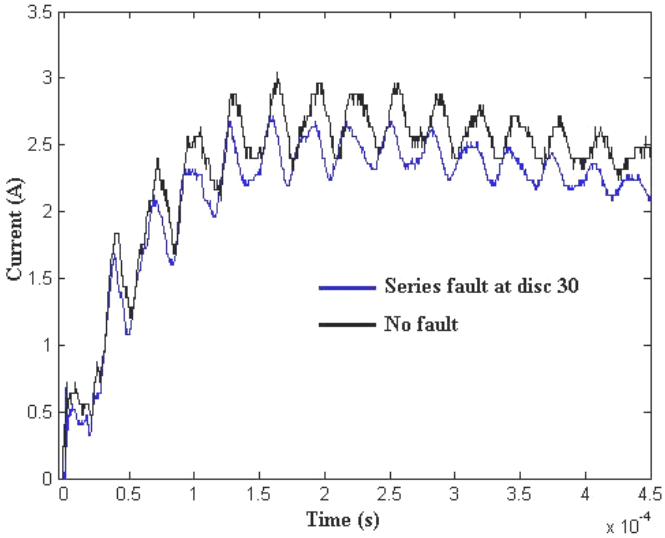

Figure 3. Tank current waveforms for no-fault condition and series fault at disc no. 30.

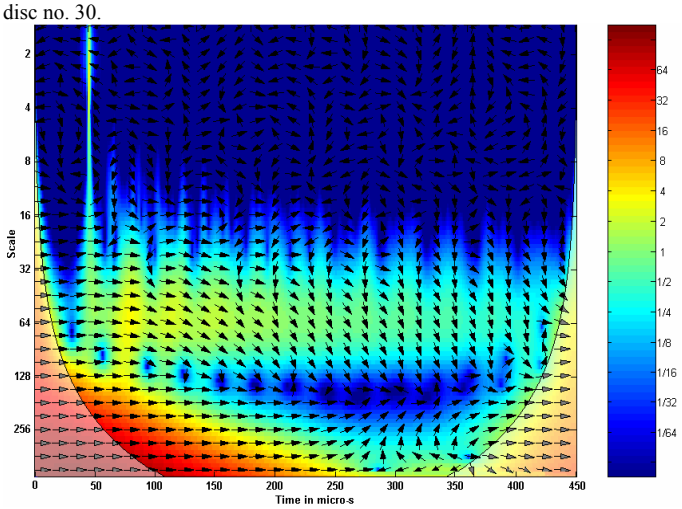

Figure 4. Cross-wavelet spectrum of no-fault current and series fault current for the fault at disc no. 30.

So, the total number of features taken from the cross-wavelet spectrum for the classification of fault is 12. These features are chosen because of the fact that they represent the salient features of the cross-wavelet spectrum. However, one may choose some other features like, location of local peaks of $|W^{xy}|$ and $\phi(s,\tau)$ surfaces, if any, or some more features from $|W^{xy}|$ and $\phi(s,\tau)$ depending upon the nature of the problem. In the present case the above-mentioned twelve features are found to be sufficient, because using these features the localization and identification of the faults can be done with reasonable accuracy. To improve the accuracy or for another problem some other features, as stated earlier, may be used.

4 INFORMATON GRANULATION USING ROUGH SETS

To classify the extracted features a Rough-Granular approach is used. It is evident that there is no apriori knowledge regarding which features of the cross-wavelet spectrum will be suitable and should be taken for classification of impulse faults. So, the data table obtained after the extraction of proposed features may contain imprecise or superfluous information. A Rough-granular classifier is used here as Rough sets are well-suited for in this kind of problems. For example, Rough Set Theory (RST) has successfully been used for condition monitoring of distribution feeder [10], for fraud detection in electrical energy consumers [11], in data-mining for semiconductor manufacturing [12] and also in case generation [13].

Information granulation is basically representation of information in the form of some aggregates, called granules that hold a number of individual entities. Granulation implies that some subsets of the universe can be described approximately. In RST these approximations are termed as lower and upper approximations. Rough sets granulate the information by partition [13-14]. Details of the Rough set theory can be found in [6, 10-15].

In RST, data is presented in a decision table in which each row represents an object (e.g. impulse fault information at different disc position) and each column represents an attribute. For example, in the present problem decision table contains extracted features (F_1-F_{12}) as the 12 condition attributes and the fault location and type as the decision attribute. For decision attribute the following notation is followed in the present problem. 'SE1' indicates the series fault in the section L1. Similarly 'SH2' implies shunt fault in section L2 and so on. The normalized decision table is shown in Table 1. Mathematically, information system $T = \langle U, Q, V, f \rangle$. Here, U is the finite set of objects and Q is the set of attributes. $V = \bigcup_{q \in Q} V_q$, where V_q is the domain of the values of q and f denotes decision function as, $f : U \times Q \to V$. If the table is having a large number of attribute values i.e. *card*(V_q) is very high for some $q \in Q$, then there is a very low chance of classifying a new object by the rules generated directly from the table. Here, *card()* means cardinality operator, which means "number of elements of a set". Therefore, discretization of the decision table is required for large real-valued decision table. Discretization of a data table indicates some partitioning of the attribute values. In the present problem Maximal Discernible (MD) heuristic is followed which is discussed in details in [15]. The discretized decision table is shown in Table 2.

In RST, for different attributes, objects are called indiscernible, i.e. similar, if they are characterized by the same information. If $P \subseteq Q$ and $x_i, x_j \in U$, then x_i and x_j are indiscernible wrt the set of attributes P, if $f(x_i, q) = f(x_j, q), \forall q \in P$. For example, let P= {F_1}. The part of the discretized decision table given in Table 2 shows that, at least objects 1, 3, 4 are indiscernible with attribute F_1, as all these objects are having same value (i.e. value=6) for feature F_1. Similarly objects 2,…, 84 and 87 are also indiscernible using F_1 as they possesses same values (i.e. value =3) for these objects. An elementary set is the set of all indiscernible objects. So, for $P \subseteq Q$, an equivalence relation on U, called P-indiscernibility relation is given by I_P. Considering the previous example of P= {F_1}, it can be found that the P-elementary set includes elementary sets like {1,3,4,…}, {2,…,84,87}, {5,6,…,86,89} and {…,85,88} etc., because all the elements of each of these sets are having same attribute values for attribute F_1. The equivalence classes of the partition induced by the P-indiscernibility relation are called *information granules*. This is explained later.

Table 1. Decision table considered for Rule generation

Object	Condition Attributes												Decision Attribute
	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	
1	0.95	0.92	0.44	0.84	0.30	0.49	0.83	0.17	0.28	0.92	0.41	0.32	SE1
2	0.43	0.93	0.20	0.52	0.18	0.59	0.76	0.97	0.46	0.84	0.30	0.96	SE1
3	0.90	0.97	0.25	0.20	0.19	0.62	0.77	0.27	0.06	0.90	1.00	0.72	SH1
4	0.95	0.40	0.22	0.67	0.68	0.44	0.70	0.25	0.98	0.87	0.01	0.41	SH1
5	0.80	0.13	0.97	0.33	0.10	0.51	0.54	0.87	0.25	0.88	0.56	0.35	SE2
6	0.76	0.91	0.92	0.01	0.54	0.66	0.44	0.73	0.42	0.98	0.97	0.26	SE2
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
84	0.45	0.41	0.01	0.68	0.15	0.34	0.59	0.13	0.51	0.46	0.99	0.43	SH21
85	0.01	0.89	0.74	0.37	0.69	0.28	0.62	0.01	0.33	1.00	0.48	0.93	SE22
86	0.82	0.05	0.74	0.83	1.00	0.34	0.79	0.35	0.43	0.92	0.43	0.68	SE22
87	0.45	0.05	0.93	0.50	0.86	0.53	0.85	0.19	0.32	0.90	0.49	0.21	SE22
88	0.01	0.81	0.46	0.70	0.65	0.92	0.92	0.49	0.57	0.05	0.21	0.83	SH22
89	0.79	0.02	0.41	0.42	0.59	0.30	0.98	0.66	0.76	0.05	0.64	0.62	SH22

Table 2. Discretized Decision table for Rule generation

Object	Condition Attributes												Decision Attribute
	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	
1	6	5	1	3	2	1	2	0	0	1	1	1	SE1
2	3	5	0	2	1	1	2	2	1	1	1	4	SE1
3	6	5	0	0	1	1	2	0	0	1	3	3	SH1
4	6	2	0	2	5	0	2	0	2	1	0	2	SH1
5	5	0	4	1	0	1	1	2	0	1	1	1	SE2
6	5	5	4	0	3	1	1	2	1	1	3	0	SE2
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
84	3	2	0	2	1	0	1	0	1	0	3	2	SH21
85	0	4	3	1	5	0	2	0	1	1	1	4	SE22
86	5	0	3	3	7	0	2	0	1	1	1	3	SE22
87	3	0	4	2	6	1	2	0	1	1	1	0	SE22
88	0	4	1	2	4	2	3	1	2	0	0	3	SH22
89	5	0	1	1	4	1	3	1	2	0	2	3	SH22

For any rough set *Y*, $\underline{PY}$ and $\overline{PY}$ are called *P*-lower and *P*-upper approximation of Y and defined as, $\underline{PY} = \{x \in Y \mid I_P(x) \subseteq Y\}$ and $\overline{PY} = \{x \in Y \mid I_P(x) \cap Y \neq \phi\}$ respectively. 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. In other words, keeping all the information intact and removing the superfluous attributes. Thus obtained *minimal* sets of attributes are called *Reduct*. The CORE is the set of relations occurring in every *Reduct*, i.e. $CORE(P) = \bigcap RED(P)$. Therefore, CORE represents the *most important* part of the knowledge. From the CORE and Reducts one can generate the decision rules. Usually these rules are considered in "IF…THEN" formats. This is illustrated in the following paragraphs. For better understanding of the method a part of the decision table is considered as shown in grey color in Table 2. The truncated decision table is shown in Table 3.

For a given subset $P \subseteq Q$, an attribute $q \in P$ is dispensable in *P* if and only if, $I_P = I_{(P-\{q\})}$; otherwise *q* is indispensable. If every element in *P* is indispensable then *P* is called *independent* otherwise *dependent*. Let $P \subseteq Q$ and $D \subseteq Q$ have equivalence relations in *U*. The *P-positive* region of *D* is indicated as, $POS_P(D) = \bigcup_{Y \in I_D} \underline{PY}$. In other words, it denotes the set of elements that can correctly be classified into D-elementary sets obtained from I_D using the knowledge described by I_P. If $q \in P$ and $POS_P(D) = POS_{(P-\{q\})}(D)$ then *q* is *D*-dispensable in *P*, otherwise *q* is *D*-indispensable in *P*. If the set of attributes *G* ($G \subseteq P$) is a *D*-independent in *P* and $POS_G(D) = POS_P(D)$, then *G* is called *D*-reduct of *P* or in general *Reduct* of *P*.

All these definitions can be explained using Table 3. For example, if *P* is taken as, *P* = {F_1, F_2, F_3}, and *D* = {*'Location and type of fault'*} (i.e. decision attribute), then I_P= {1}, {2}, {3}, {4}, {5} and {6}; I_D = {1,2}, {3,4} and {5,6}. Also, $POS_P(D) = \{1,2,3,4,5,6\}$. If the attribute F_1 is removed from *P* then, $POS_{(P-\{F_1\})}(D) = \{1,4,5,6\}$. Clearly, $POS_{(P-\{F_1\})}(D) \neq POS_P(D)$. Therefore the attribute F_1 is *D*-indispensable in *P*. Similarly, removing attribute F_2 gives, $POS_{(P-\{F_2\})}(D) = \{1,2,3,4,5,6\} = POS_P(D)$. Therefore attribute F_2 is *D*-dispensable in *P*. Again, for attribute F_3 it is easy to observe that, $POS_{(P-\{F_3\})}(D) = \{2,4,5,6\} \neq POS_P(D)$. Similarly, F_3 is also *D*-indispensable in *P*. Thus, the set {F_1, F_3} is the *D*-reduct of *P*. Therefore, the simplified or reduced form of Table 3 is given in Table 4. '-' indicates "don't care" (i.e. dispensable) condition. It can be said that, attribute values, (F_1=6 $\wedge$ F_3=1) $\vee$ (F_1=3$\wedge$ F_3=0) are the characteristic for decision class 'SE1'. '$\wedge$' and '$\vee$' are logical "AND" and "OR" operators respectively. Similarly, (F_1=6 $\wedge$ F_3=0) is the characteristic of decision class 'SH1' and (F_1=5 $\wedge$ F_3=4) is the characteristic of 'SE2'. These are called *information granules*. Intersections of these reduct values for each of the decision class (i.e. SE1, SH1 and SE2) will give the CORE for the respective class. For the decision class 'SE1' no such CORE value is obtained from the Table 4, as the intersection of (F_1=6 $\wedge$ F_3=1) and (F_1=3 $\wedge$ F_3=0) is null. Similarly, for decision class 'SH1' the intersection of (F_1= 6 $\wedge$ F_3= 0) and (F_1=6 $\wedge$ F_3=0) gives CORE values F_1=6 and F_3=0. Again, for 'SE2' the

Table 3. Truncated Decision table for illustration

Object	Condition Attributes			Decision Attribute
	F_1	F_2	F_3	
1	6	5	1	SE1
2	3	5	0	SE1
3	6	5	0	SH1
4	6	2	0	SH1
5	5	0	4	SE2
6	5	5	4	SE2

CORE values are F_1=5 and F_3=4. Furthermore, this reduced Table 4 can be used to generate decision rules. The decision rules obtained from the Reduct and CORE values are given in Table 5.

Table 4. Reduced form of Table 3

Object	Condition Attributes			Decision Attribute
	F_1	F_2	F_3	
1	6	-	1	SE1
2	3	-	0	SE1
3	6	-	0	SH1
4	6	-	0	SH1
5	5	-	4	SE2
6	5	-	4	SE2

Table 5. Decision Rules obtained from CORE and Reducts

Decision Rule No.	Statement of the Rule	
	IF	THEN
1	$(F_1=6 \wedge F_3=1) \vee (F_1=3 \wedge F_3=0)$	Location and type of fault is 'SE1'
2	$(F_1=6 \wedge F_3=0)$	Location and type of fault is 'SH1'
3	$(F_1=5 \wedge F_3=4)$	Location and type of fault is 'SE2'

It is worth mentioning here, that this derivation of decision rules is a demonstrative example of the whole process. While applying these procedures to the complete decision table the rules may be different from these rules.

5 RESULTS AND DISCUSSIONS

As mentioned earlier, series and shunt faults are simulated at every disc position and corresponding tank-current waveforms due to applied impulse are recorded. For each of these waveforms a cross-wavelet spectrum is obtained from cross-wavelet transform of the fault current data with the current waveform at no-fault condition. Thereafter the feature vector for each of the fault condition is extracted. These feature vectors along with the corresponding fault condition are tabulated in a table and forms the decision table. The complete data set is having 178 cases of different fault conditions, i.e. series and shunt faults at each of the 89 disc locations. 50% of the data, randomly chosen, are used to develop the decision rules and the remaining 50% data are used for testing purpose. Therefore, the data table used for rule generation is having 89 rows (objects) and 13 columns (excluding the column of object index), as evident from Table 1. The first twelve columns consist of the features from F_1-F_{12}, which are called condition attributes and the last column is for the decision attribute, i.e. fault location and type, as stated earlier. The values of the condition attributes are obtained from the equations for F_1-F_{12}. To obtain the decision rules for classification of faults following steps are followed,

Step 1: *The data table is discretized.* This means that each condition attribute values are divided into ranges. In the present problem Maximal Discernible (MD) heuristic is followed which is discussed in details in [15], as stated in the previous section. The discretized table is shown in Table 2.

Step 2: *Identical Attributes and cases are eliminated.* As in the present case no identical attribute or case is observed, the discretized table remains unchanged after this step for the present problem.

Step 3: *Dispensable attributes are removed*. The method described in the previous section is followed to obtain dispensable attributes and it was found that attributes (features) F_4, F_6, F_9 and F_{10} are dispensable and the remaining attributes are indispensable. So the reduced decision table is having 8 columns for condition attributes.

Step 4: *Reducts and CORE are obtained.* The sample computation of Reduct and CORE is described earlier. Following the same procedure the final form of the decision table is obtained and is shown in Table 6.

Step 5: *Decision rules are generated from the final table of CORE and Reducts.* The decision rules are constructed from the *granulated* knowledge and some of the rules are shown in Table 7. Thus, 44 IF…THEN rules are obtained. The test dataset is discretized and tested by the generated decision rules to judge the validity of the rules.

Table 6. Simplified form of the complete Decision table

Object	Condition Attributes								Decision Attribute
	F_1	F_2	F_3	F_5	F_7	F_8	F_{11}	F_{12}	
1	-	5	-	-	2	-	1	-	SE1
2	-	5	-	-	2	-	1	-	SE1
3	6	-	0	-	2	0	-	-	SH1
4	6	-	0	-	2	0	-	-	SH1
5	5	-	4	-	1	2	-	-	SE2
6	5	-	4	-	1	2	-	-	SE2
•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•	•	•
84	3	-	-	1	-	0	-	2	SH21
85	-	-	-	-	2	0	1	-	SE22
86	-	-	-	-	2	0	1	-	SE22
87	-	-	-	-	2	0	1	-	SE22
88	-	-	1	4	3	1	-	3	SH22
89	-	-	1	4	3	1	-	3	SH22

Among the 89 test cases the Rule set obtained from proposed information granulation technique has correctly classified the fault 81 times. So, the percentage of success is almost 90%, which is reasonable for localization of the fault within 5% of the winding length (i.e. within 4 discs out of total 89 discs in the winding under investigation). It is to be mentioned here that earlier researchers have already reported localization of fault within 20% [3] and 10% [4] of the winding length.

The results shown above are obtained from the simulation of faults on the analog model. The performance of the proposed scheme is also verified with an EMTP based digital model as shown in Figure 2. Fault simulation, feature extraction and decision rule generation are performed similarly as described for the case of analog model. The final rule set is shown in Table 8. Among 89 test cases, developed rules can successfully classify the fault type and location for 83 times. So the percentage of success is 93%.

Table 7. Decision Rules obtained from the analysis of analog model

Decision Rule No.	Statement of the Rule	
	IF	THEN
1	(F_2=5∧ F_7=1∧ F_{11}=1)	Location and type of fault is 'SE1'
2	(F_1=6 ∧ F_3=0 ∧ F_7=2 ∧ F_8=0)	Location and type of fault is 'SH1'
3	(F_1=5 ∧ F_3=4 ∧ F_7=1 ∧ F_8=2)	Location and type of fault is 'SE2'
•	•	•
•	•	•
•	•	•
42	(F_1=3 ∧ F_5=1 ∧ F_8=0 ∧ F_{12}=2)	Location and type of fault is 'SH21'
43	(F_7=2 ∧ F_8=0 ∧ F_{11}=1)	Location and type of fault is 'SE22'
44	(F_3=1 ∧ F_5=4 ∧ F_7= 3 ∧ F_8=1 ∧ F_{12}=3)	Location and type of fault is 'SH22'

Table 8. Decision Rules obtained from the analysis of EMTP model

Decision Rule No.	Statement of the Rule	
	IF	THEN
1	(F_2=4∧ F_7=1∧ F_{11}=1)	Location and type of fault is 'SE1'
2	(F_1=5 ∧ F_4=1∧ F_7=2 ∧ F_8=1)	Location and type of fault is 'SH1'
3	(F_1=4 ∧ F_4=4 ∧ F_7=1 ∧ F_8=1)	Location and type of fault is 'SE2'
•	•	•
•	•	•
•	•	•
42	(F_1=4 ∧ F_4=2 ∧ F_8=0 ∧ F_{12}=2)	Location and type of fault is 'SH21'
43	(F_7=2 ∧ F_4=3 ∧ F_8=0 ∧ F_{11}=1)	Location and type of fault is 'SE22'
44	(F_3=1 ∧ F_2=2 ∧ F_7= 3 ∧ F_8=2)	Location and type of fault is 'SH22'

6 CONCLUSIONS

A novel approach based on information granulation using Rough sets for impulse fault identification of transformers is proposed and investigated. The performance of the scheme is shown for both analog and EMTP based digital models of transformers. Results show that cross-wavelet transform may be considered as an efficient tool for feature extraction and the Rough-granulation based classification approach is suitable for transformer impulse fault classification and localization.

REFERENCES

[1] *Guide to the Lightning and Switching Impulse Testing of Power Transformers and Reactors*, IEC Std., Publ. 722, 1982.

[2] A. De and N. Chatterjee, "Recognition of impulse fault patterns in transformers using Kohonen's self-organizing feature map", IEEE Trans. Power Del., Vol. 17, pp. 489–494, 2002.

[3] A. De and N. Chartterjee, "A fuzzy ARTMAP fault classifier fo impulse testing of power transformer", IEEE Trans. Dielectr. Electr. Insul., Vol. 11, pp. 1026–1036, 2004.

[4] C. Koley, P. Purkait and S. Chakravorti, "Wavelet-Aided SVM Tool fo Impulse Fault Identification in Transformers", IEEE Trans. Power Del., Vol. 21, pp. 1283-1290, 2006.

[5] D.Dey, B. Chatterjee, S. Chakravorti and S.Munshi, "Hybrid Filtering Scheme for Proper Denoising of Real-Time Data in Dielectric Spectroscopy", IEEE Trans. Dielectr. Electr. Insul., Vol. 14, pp. 1323-1331, 2007.

[6] Z. Pawlak, *Rough Sets: Theoretical Aspects of Reasoning about Data*, Kluwer, Boston, USA, 1991.

[7] P. Purkait and S. Chakravorti, "Time and frequency domain analyzer based expert system for impulse fault diagnosis in transformers", IEEE Trans. Dielectr. Elect. Insul., Vol. 9, pp. 433-445, 2002.

[8] B. G. Ruessink, G. Coco, R. Ranasinghe, and I. L. Turner, "A cross-wavelet study of alongshore non-uniform near shore sandbar behavior", Proc. Joint Conference on Neural Networks, Vancouver, Canada, pp. 4310-4317, 2006.

[9] C. Torrence and G. P. Compo, "A Practical Guide to Wavelet Analysis", J. American Meteorological Soc., Vol. 79, pp. 61-79, 1998.

[10] J.T. Peng, C.F. Chien and T.L.B. Tseng, "Rough set theory for data mining for fault diagnosis on distribution feeder", IEE Proc.-Gener. Transm. Distrib., Vol. 151, pp. 689-697 , 2004.

[11] J.E. Cabral, J.P. Pinto, E.M Gontijo and J.R .Filho, "Fraud Detection in Electrical Energy Consumers Using Rough Sets", Proc. IEEE Conf. Systems, Man and Cybernetics, pp. 3625-3629, 2004.

[12] A. Kusiak, "Rough Set Theory: A Data Mining Tool for Semiconductor Manufacturing", IEEE Trans. Electronics Packaging Manufacturing, Vol. 24, pp. 44-50, 2001.

[13] S. K. Pal and P. Mitra, "Case Generation Using Rough Sets with Fuzzy Representation", IEEE Trans. on Knowledge and Data Engg., Vol. 16, pp. 292-300, 2004.

[14] Y.Y. Yao, "Rough Sets, Neighborhood systems and Granular computing", Proc. IEEE Conf. on Electr. and Comp. Engg., pp. 1553-1558, Canada, May 1999.

[15] H. S. Nguyen, *Discretization of Real Value Attributes: A Boolean Reasoning Approach*, PhD Thesis, Dept. of Mathematics, Warsaw Univ. Poland, 1997.

Debangshu Dey was born in Kolkata, West Bengal in April 1980. He got his B.E.E and M.E.E. degrees from Jadavpur University, Kolkata, India in 2003 and 2005 respectively. Presently he is working as Senior Research Fellow in High Tension Laboratory, Electrical Engineering Department, Jadavpur University, Kolkata, India. He has published ten research papers. His areas of interest are sensor linearization, intelligent instrumentation, signal conditioning and application of optimization and computational intelligence in electrical measurements.

Biswendu Chatterjee was born in Kolkata, West Bengal, India, in February 1973. He obtained the M.E.E. degree from Jadavpur University, Kolkata, India in 2004. Presently he is working as a Research Associate in High Tension Laboratory, Electrical Engineering Department, Jadavpur University, Kolkata, India. His areas of interest are partial discharge (PD) measurements and instrumentation, EMI reduction techniques and condition monitoring of large electrical equipment.

Sivaji Chakravorti (M'90-SM'00) obtained his B.E.E., M.E.E. and Ph.D. degrees from Jadavpur University, Kolkata, India in 1983, 1985 and 1993 respectively. Since 1985 he has been a full-time faculty member of the Electrical Engineering Department of Jadavpur University, where he is currently Professor in Electrical Engineering. In 1984 he worked at the Indian Institute of Science, Bangalore as Indian National Science Academy Visiting Fellow. He worked at the Technical University Munich as Humboldt Research Fellow in 1995-96, 1999 and 2007. He served as Development Engineer in Siemens AG in Berlin in 1998. He has also worked as Humboldt Research Fellow in ABB Corporate Research at Ladenburg, Germany in 2002. In 2003 he worked as US-NSF guest scientist at the Virginia Tech, USA. He has published about 100 research papers and has authored a book. He is the recipient of Technology Day Award of AICTE for best project work in 2003. His current fields of interest are numerical field computation, computer aided design and optimization of insulation system, condition monitoring of large electrical equipment and signal conditioning in high voltage systems.

Sugata Munshi obtained the B.E.E and M.E.E. degrees from Jadavpur University, Kolkata, India in 1980 and 1985, respectively. He worked as an Engineer in the Plasma Physics Division of Saha Institute of Nuclear Physics, India from 1985 to 1990. In 1986, he had a training on 'Tokamak' machine in the 'Heavy Engineering Works' of Toshiba Corporation in Japan. In 1990 he joined the Electrical Engineering Department of Jadavpur University as a faculty member. At present, he is a Reader in this department. He has published about 30 research papers in refereed journals. He was the joint recipient of The President of India's Prize (English) in 1989-90, The Pandit Madan Mohan Malaviya Memorial Prize in 1989-90, The Sir Thomas Ward Memorial Prize in 1994-95, The Tata Rao Medal, awarded by the Institution of Engineers (India) in 1996-97 and Certificate of Merit from IE (India) in 1996-97.His current fields of interest are signal processing, surge phenomena in power equipment and sensor systems.