BUSBAR PROTECTION USING RANDOM FOREST TECHNIQUE

A DISSERTATION

Submitted in partial fulfilment of the

requirements for the award of the degree

of

MASTER OF TECHNOLOGY

in

ELECTRICAL ENGINEERING

(With specialization in Instrumentation and Signal Processing)

By

NIDHI NARAYAN

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY ROORKEE
ROORKEE - 247 667 (INDIA)
MAY, 2016

INDIAN INSTITUTE OF TECHNOLOGY ROORKEE

CANDIDATE'S DECLARATION

I hereby declare that this thesis report entitled BUSBAR PROTECTION USING RANDOM FOREST TECHNIQUE, submitted to the Department of Electrical Engineering, Indian Institute of Technology, Roorkee, India, in partial fulfilment of the requirements for the award of the Degree of Master of Technology in Electrical Engineering with specialization in Instrumentation and Signal Processing is an authentic record of the work carried out by me during the period June 2015 through May 2016, under the supervision of Dr. B. Bhalja and Dr. M. Tripathy, Department of Electrical Engineering, Indian Institute of Technology, Roorkee. The matter presented in this thesis report has not been submitted by me for the award of any other degree of this institute or any other institutes.

Date: Nidhi Narayan

Place: Roorkee

CERTIFICATE

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Dr. BHAVESH BHALJA
Associate Professor
Department of Electrical Engineering
Indian Institute of Technology
Roorkee- 247667, India

Dr. MANOJ TRIPATHY
Associate Professor
Department of Electrical Engineering
Indian Institute of Technology
Roorkee- 247667, India

ACKNOWLEDGEMENT

I owe my profound gratitude to Dr. Bhavesh Bhalja**,** Associate Professor, Department of Electrical Engineering, Indian Institute of Technology, Roorkee, for his constant guidance, valuable feedbacks and suggestions at various stages of this work and for providing relevant resources without which this thesis would not have been completed so well.

I would like to extend my sincere thanks to Dr. Manoj Tripathy**,** Associate Professor, Department of Electrical Engineering, Indian Institute of Technology, Roorkee, and Dr. R P Maheshwari, Professor, Department of Electrical Engineering, Indian Institute of Technology, Roorkee for showing keen interest and for being helpful.

I also want to thank my parents and my brother for their constant love, support and motivation.

NIDHI NARAYAN

ABSTRACT

The busbar, where all the lines converge, is very critical element of a power system. Faults on busbar can lead to disturbances and major shut down of power system. On other hand, the unwanted isolation of busbar leads to a major power loss in the system. Therefore, in order to prevent hazard and major power loss to the power system, effective and reliable protection schemes for busbars needs to be developed.

This thesis presents two schemes for fault zone identification for busbar. The first scheme is based on the Random Forest Classifier while the second is based on *dq* transform. The schemes depend on extracting features from the measured data of current signals of all the bays (lines) connected to a busbar. Validation of the scheme has been carried out by modelling an existing 400 kV Indian busbar system using PSCAD/EMTDC software package. A wide variety of simulation cases (33,600) consisting of internal and external faults have been generated by varying the system and fault parameters. By giving post fault data of one cycle duration of all the bay currents as input the algorithm based on RF classifier, and using 30% data for training and 70% for testing, an accuracy of more than 98% has been obtained.

Furthermore, a scheme based on *dq* transform has been developed. The i_d and i_q components of all the line currents are calculated and summed up. The post fault data samples of one cycle are added to find out $i_{d_{SOS}}$ (Sum of Samples of i_d components) and $i_{q_{SOS}}$ (Sum of Samples of i_q components). If both these values are less than the threshold, it is a case of external fault and if it is more than a threshold, it is a case of internal fault. The overall accuracy obtained with this scheme is of the order of 98%.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER-1
INTRODUCTION

1.1 General

The junction of the electrical power system where a number of transmission line and different elements of power system are connected is called busbar. The protection of power system bus is one of the most critical relaying applications. When the fault occurs within the busbar areas of a power system, fault current may be dangerously elevated. Other than this, some of the CTs of the circuits connected to the busbar may have insufficient ratings. This may lead to a risk of saturation of CTs which would disturb the busbar protection scheme.

The protection scheme for busbar must remain stable during any external or through fault as instability would lead to false tripping of the relay. A false tripping of distribution bus may disrupt several feeders, cause problem to many customers, as well as disconnect sub transmission lines. A false trip of busbar may alter the topology of the system tremendously and the security of the power system may be affected.

Conversely, busbar faults generate enormous fault currents which if not cleared quickly can harm the whole substation because of build-up of dynamic forces and thermal effects.

Prime requirement of the busbar protection scheme is to be able to identify the fault associated with a particular region and disconnect the breakers concerned with that region so that the other healthy regions of the power system are not affected. Thus, unit protection is preferred to time graded scheme.

Damage to the equipment depends on the fault type, fault duration, fault level, and the withstanding capability of switchgear. The isolation of a busbar disrupts all the circuits coupled to the busbar. The busbar protection system, therefore, must be cautiously examined to avoid accidental operations. Hence, if a system is simple to design and easy to apply, it is likely to provide reliable protection.

Accordingly, a bus protection system should have the following features:

- High speed protection of busbar system during in-zone faults to reduce damage and increase stability.
- It should remain stable for all external faults to avoid unnecessary interruption of supply.
- It should provide proper discrimination between two zones, i.e. it should be tripping minimum number of relays.
- It should give reliable operation to avoid widespread destruction to the equipment, harm to human resources, and disturbance of service.

Bus faults' frequency of occurrence is very low but when such faults arise, the risk of damage is very high because huge currents are involved. It has the capacity to destroy an entire power system. Bus faults have been observed very rarely, around 6-7% of all faults compared with the line faults which are over 60%. Out of the total busbar faults, LG faults are 67%, LLLG faults are 19% and LLG faults are 15%. Causes of the bus faults are:

- Breakdown of insulation as the material deteriorates.
- Continued and unwarranted overvoltages leading to flashover.
- Failure of switchgear and circuit breakers.
- Human errors during operation and maintenance of switchgear.
- Contacts by animals

1.2 Literature Review

Different busbar protection schemes based on sequence impedance and directional comparison have been presented by researchers. [2] - [5]. But, the main limitation of the above mentioned schemes is that they maloperate is case of an external fault. Thereafter, a scheme for protection of busbar based on direction comparison that uses the method of superimposition of components of currents was proposed by Zadeh *et al.* [6]. Although this scheme incorporates IEC61850 substation protocol, one has to measure the voltages on the line side as well as the line currents which make

the scheme costly and more complex. Afterwards, many digital protection schemes for busbar have been presented by the researchers using Wavelet transform [7] - [10]. However, the high frequency noise signals penetrate when the current signals are decomposed due to which the accuracy of these schemes has been reduced. Thereafter, researchers have proposed [11] - [12] traveling wave based busbar protection scheme. However, the above schemes provide unsatisfactory results if the signal available during transient fault condition is weak.

Afterwards, Chothani *et al.* [13] - [14] presented a busbar protection scheme using Support Vector Machine (SVM) which provides better accuracy for identification of the zone of fault. But the kernel function of SVM needs to fulfil Mercer's condition which makes the predicted results non probabilistic. Furthermore, the number of support vectors (SVs) and the regularization parameter (C) for least error bound (ε) has to be calculated in this method. Hence, owing to the complex computations involved in SVM, it is difficult to use it in the practical field. Thereafter, Song *et al.* [15] proposed a busbar protection scheme using polarity comparison of superimposed current. Nevertheless, improper discrimination between high resistance internal fault and heavy through fault with CT saturation are several limitations of the above scheme

CHAPTER-2
SYSTEM MODELLING

The single line diagram of the Indian power system network with three sources represented by their Thevenin's equivalent has been presented in Fig.2.1. Three lines—L1, L2 and L3 connect the sources to the bus which has been considered. Here, line L1 is compensated transmission line whereas L2 and L3 are uncompensated transmission lines. The line and system parameters are given in Appendix. 'Base Source Impedance Values (BSIV)' stands for the source impedance values given in the Appendix.

The source impedance (SI) values (S_{z1}, S_{z2}, S_{z3}) have been varied and taken into account. This is required to check effectiveness of the presented scheme when the system & fault conditions are varied. Here, 120% of BSIV, 80% of BSIV and 100% of BSIV are considered for generating a large number of data sets. Seven combinations of these source impedance values, as shown in Table 2.1, are utilized for generation of simulation cases.

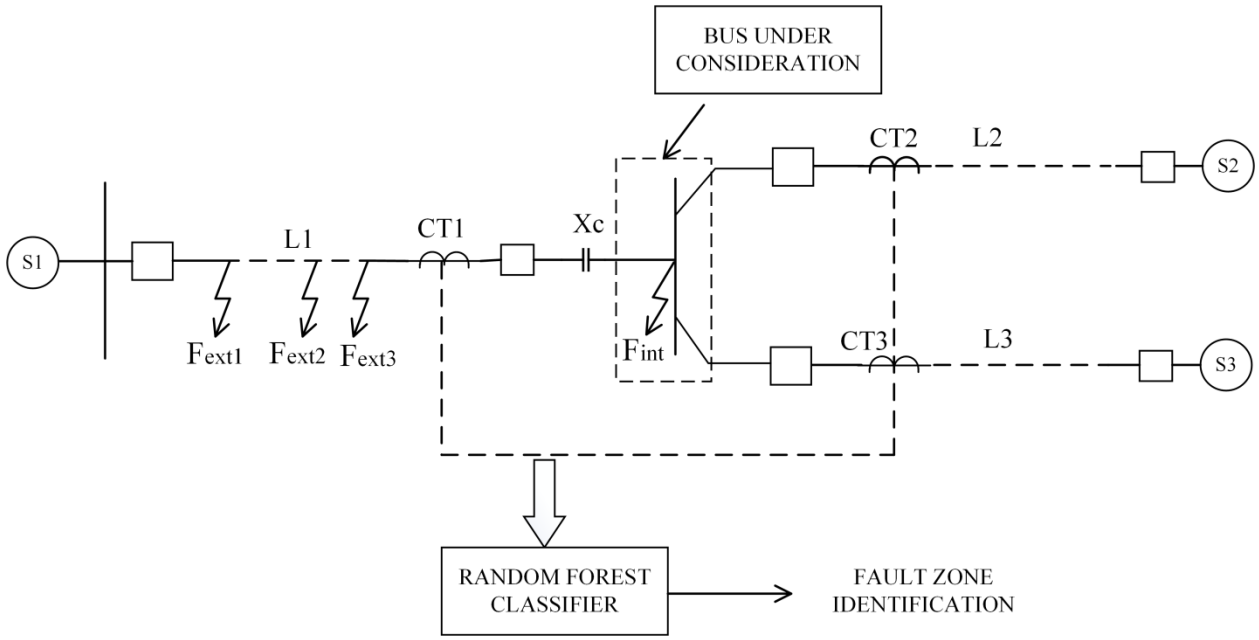


Fig. 2.1 Single line diagram of the power system model

Table 2.1 Source Impedances combinations considered during generation of simulation cases.

Cases	S_{Z1}%	S_{Z2}%	S_{Z3}%
SI 1	100	100	100
SI 2	100	80	100
SI 3	80	100	100
SI 4	100	100	80
SI 5	100	100	120
SI 6	100	120	100
SI 7	120	100	100

Furthermore, the load angle (δ), types of fault (Ftype), fault resistance (R_f), fault inception angle (FIA) and fault locations (FL) on line L1 (Fex1, Fex2, Fex3) are varied for data generation. The value of δ and R_f is varied from 0 degree to 30 degrees and 0 Ω to 50 Ω in steps of 10^0 and 10 Ω, respectively. Subsequently, the values of 0^0, 45^0, 90^0,115^0 and 150^0 are considered for FIA. All the ten types of faults (LG, LL, LLG and LLL/LLLG) are considered for generation of simulation cases for internal and external faults. External faults at three different locations i.e. 10%, 30% and 50% on line L1 have been simulated.

Thus, for each combination of source impedance as shown in Table 2.1, a total of 4 (δ) × 6 (R_f) × 5 (FIA) × 10 (Ftype) = 1200 cases of internal faults have been generated whereas a total of 4 (δ) × 6 (R_f) × 5 (FIA) × 10 (Ftype) × 3 (FL) = 3600 cases of external fault have been generated. In this manner, 8400 (1200 × 7 (SI)) cases of internal fault and 25,200 (3600 × 7 (SI)) cases of external fault have been investigated in this work. Hence, for the purpose of performance evaluation of the presented scheme a total of 33,600 (8400+ 25,200) cases have been considered. An automatic model for generation of fault data has been developed in the PSCAD/EMTDC software package [16], to avoid manual error and to make the simulation process simpler. This is achieved by using the 'Multirun' blocks available in PSCAD. The 'Bergeron' line model has been used for representing the transmission line.

For achieving accurate and effective classification, proper training of the RF classifier is necessary. Out of the total 33,600 data cases, 30% of the data i.e. 10,080 cases with wide variety of parameters have been used for training the RF classifier. Conversely, the remaining 23,520 data cases (70%

of the total data sets) have been utilized for the purpose of testing the RF based algorithm. The fault discrimination accuracy (η) of the scheme is given by:

$$\eta = \frac{Number of\ correct\ fault classification\ cases}{23{,}520} \times 100 \tag{2.1}$$

CHAPTER-3

IDENTIFICATION OF ZONE OF FAULT OF BUSBAR USING RANDOM FOREST CLASSIFIER

3.1 Random Forest Classifier

Random forest is a classifier based on an ensemble of decision trees. Rapid training process, models with strong predictive results and removal of the problem of overfitting are some of the advantages associated with this classifier. In this classifier, a random subset of the data samples available for training purpose, is used for growing a tree. Also, the attributes are chosen in such a way that every node splitter is determined randomly. These are the two ways in which randomness is introduced in the decision trees. The accuracy of classification and make choice or vote for the most popular class is considerably enhanced when the ensemble trees are extended. The random vectors not only dictate the growth of each and every tree but also nourish all the trees which have been formed. A random vector θ_n, independent of the preceding random vectors θ_1, ..., θn_{-1} from the same distribution is used for the formation of n^{th} tree. The training set, θn, helps to develop a tree and thus results in formation of a classifier $h(x, \theta n)$ with x as the input vector. In this way a large number of trees are formed. Thereafter, the most popular class is chosen by all the trees for a given input vector and the process is called Random Forests [17] - [18].

The correlation and strength of all the trees which have been grown to infinity depth determine the forest error rate. As the strength of the trees rises, the forest error rate falls. On the other hand, as the correlation between trees increases, the forest error rate increases. The correlation as well the accuracy can be decreased by decreasing the number of possible predictors. It is possible to achieve this by finding out the Out-of-Box (OOB) error. The Out-of-Bag (OOB) error helps to ascertain the optimal importance of correlation and strength. Out of the total original data, two third data is utilized for the purpose of training and the OOB error is calculated using the remaining data in the process of application of OOB data.

Given (X_i, Y_i) pairs, ($i=1, 2, ..., M$) where, X_i stands for the feature predictor variable which has been input and $Y_i \epsilon \{-1, 1\}$ is the response for classification of the negative and positive class. The function estimator for each pair can be given by,

$g(.)=h_M((X_1, Y_1),...,(X_M, Y_M))(.)$ (3.1)

Where, $h(.)$ is the function representing the estimator as a function of the data. The training samples $(X_1^*, Y_1^*),...,(X_M^*, Y_M^*)$ are constructed by randomly withdrawing samples from the original data $(X_1, Y_1),...,(X_M, Y_M)$ with replacement and the estimator function is used to compute its estimator which is given by,

$g(.)=h_M((X_1^*, Y_1^*),...,(X_M^*, Y_M^*))(.)$ (3.2)

Theoretically, the process continues for an infinite number of times or practically, for the number of times selected by the user, after which the resultant probabilities are averaged. Once the training is complete, predictions for unseen feature vector x' is done by pushing the particular feature vector x' through all the M trees constructed until it reaches corresponding leaves. The average of predictions ($\hat{w}$) from all the individual regression trees on x' gives the output class label of the feature vector:

$$\hat{w} = \frac{1}{B}\sum_{b=1}^{B} \hat{w}_M(x') \quad (3.3)$$

The output class which had obtained votes in majority in each case is the class which was OOB after we finish one run. The OOB error estimate is given as the number of times the response is not equal to the actual response class Y_i divided by the total number of cases. The RF classifier can deal with uneven data sets having missing variables very well, and still train faster than all the other methods of machine learning.

3.2 The Scheme

The schematic block diagram of the RF classifier based scheme is shown in Fig. 3.1. Firstly, by using the data acquisition system, post fault samples of current signals of all CTs for 3 bays are acquired. To make sure that all the disturbances occurring on the connected bays and the busbar is captured, CT signals of all the substation bays are utilized. Generally, off-line configuration of the relay is performed based on the RF classifier using the 10,080 simulation cases available as training dataset. The fundamental frequency is 50 Hz and 1 kHz of sampling frequency (i.e. total 20 samples in each cycle) is used. Samples of one cycle post fault current signals of all three phases from all CTs (CT1, CT2 & CT3) are acquired (20 samples for one cycle × 3 phases × 3 signals from CT1, CT2 & CT3 = 180 samples for each simulation case). For each simulation case, a feature

vector with 180 samples in a row is given to the RF classifier as input. In this way a $10{,}080_{row}$ (train cases) × 180_{column} dimension vector is formulated for 10,080 training data set.

After the RF classifier has been configured, discrimination between fault condition and normal condition is carried out by using the fault detection algorithm [19]. The first stage of work of any conventional digital/numerical relay performs is fault detection. A relay uses two parallel data acquisition system which are independent of each other. The first one is used for fault detection (Fault detector unit) whereas the second is used for identification of fault zone.

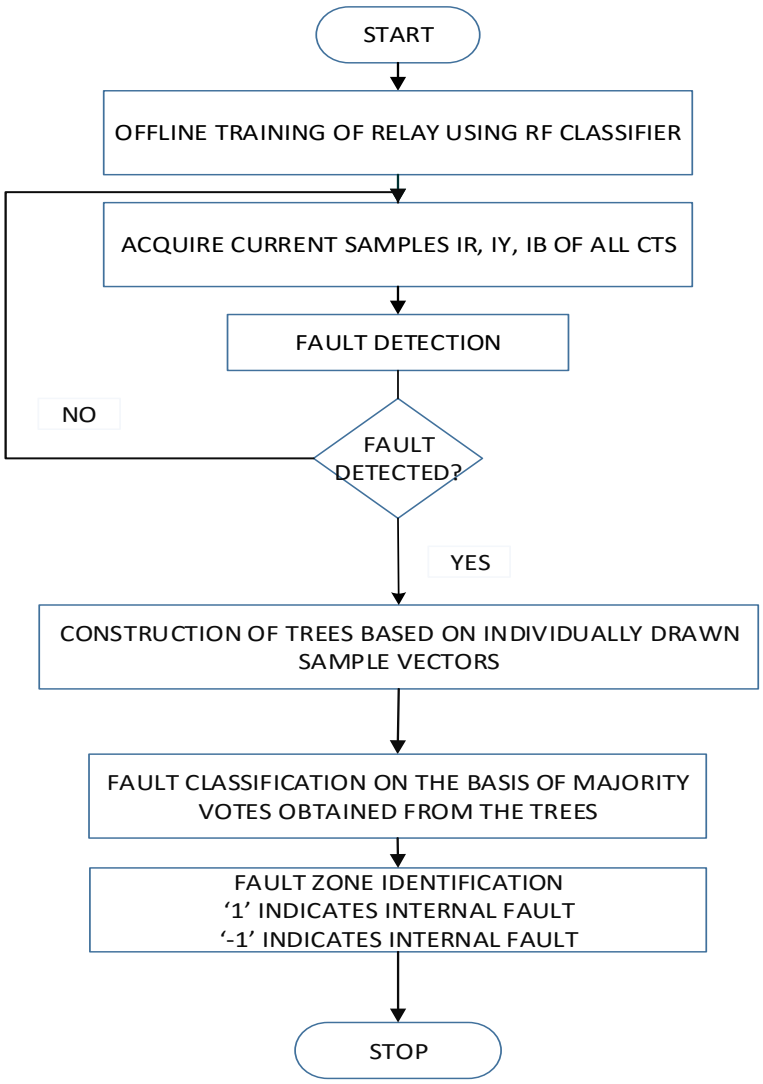


Fig. 3.1 Schematic block diagram of the proposed scheme

Whenever a fault is detected by fault detection algorithm, different feature vectors are formed by combining post fault current (20 samples/phase) samples of one cycle of all the CTs. For every

fault case scenario, a feature vector with 180 samples is formed using 3 signals from the 3 CTs (3 signals × 3-phases × 20 samples per cycle). Thus, using the testing data set for the cases already mentioned in the previous section, a simulation data base of *post fault samples* × test *data length* i.e. 180 × 23520 is generated. Soon after, all the feature vectors are utilized for the purpose of testing the algorithm based on RF classifier. The classifier output (obtained by majority voting of all the trees) '1' denotes an internal fault whereas '-1' denotes an external fault.

3.3. Selection of Parameters for RF Classifier

Once the training cases have been chosen, it is necessary to choose the optimal parameters for the training and testing of the RF classifier. For this purpose, it is required to run the algorithm for a number of iterations by varying the number of trees grown. The numbers of trees are varied from 100 to 750 and the accuracy has been observed. Fig. 3.2 shows variation of accuracy of RF classifier against the number of trees. It has been observed from Fig. 3.2 that the highest accuracy of the order of 98.49% is obtained with 139 trees. Highest In-zone fault accuracy of 97.0238% and Out-of-zone fault accuracy of 98.9796% is obtained with the method as seen from Fig 3.3.

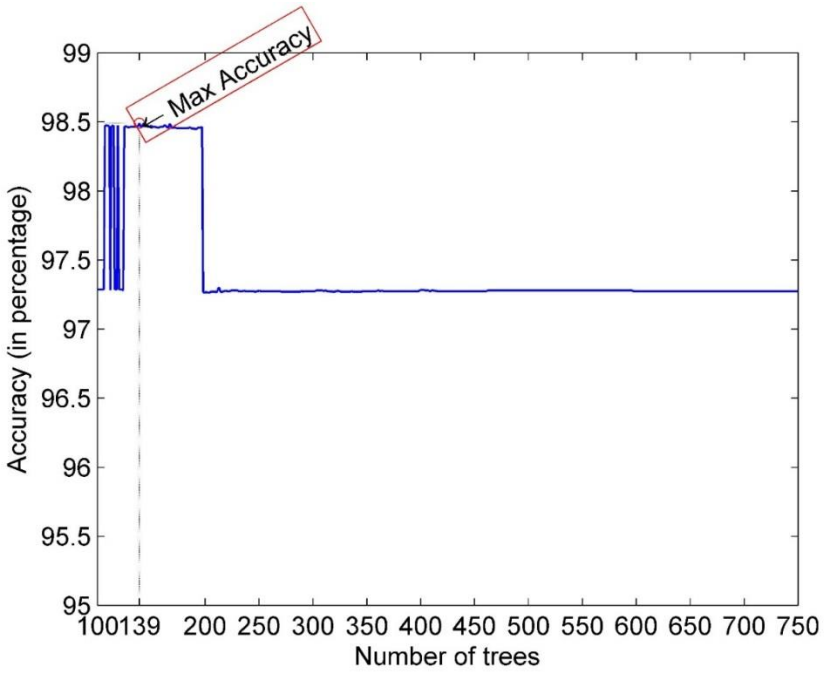


Fig. 3.2 Variation of accuracy against number of trees

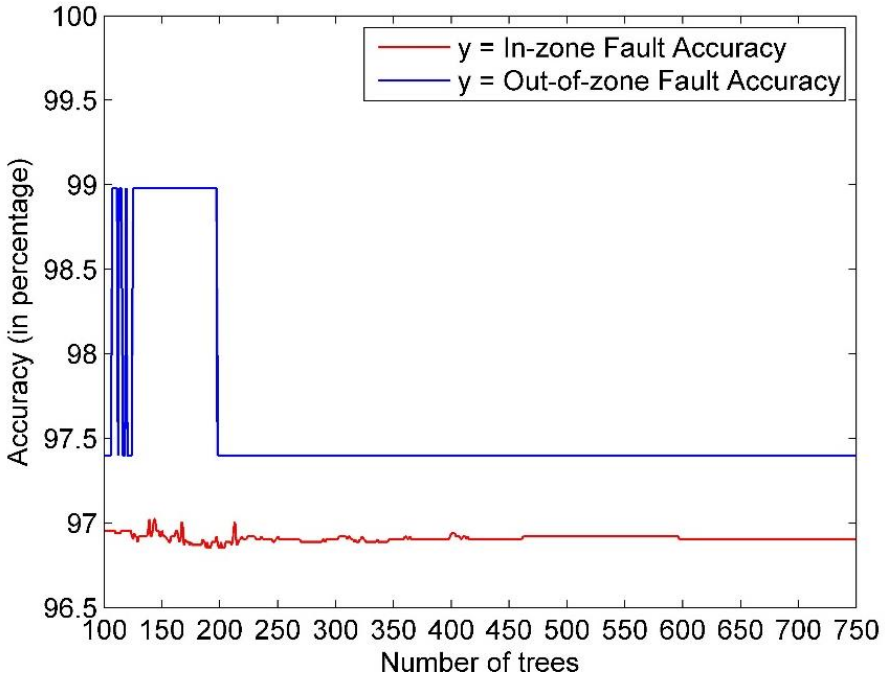


Fig 3.3 Variation of In-zone fault accuracy and Out-of-zone fault accuracy with number of trees

3.4 Performance Measurement

The main objective of building the classification model is to find out correctness of the model. This is achieved by calculating accuracy of the model which provides sum of true positives and true negatives with reference to the total cases. However, accuracy is not always a valid criterion for measuring the performance of the model. This is due to the fact that a system with zero true positives and all true negatives or all true positives and zero true negatives can also give 100% accuracy. Hence, such a system would be of no relevance as far as classification is concerned. Thus, many other performance evaluation indices are required to authenticate the classification model. Based on the confusion/contingency matrix, as shown in Table 3.1, various performance indices have been derived and also described below.

Table 3.1 Confusion Matrix

Term	Positive Predictions	Negative Predictions
Positive Class	True Positive (TP)	False Negative (FN)
Negative Class	False Positive (FP)	True Negative (TN)

(i) **True Positive Rate (*TPR*):** Likelihood a test being actually positive when it is predicted a positive. It is also called Recall or Sensitivity and given by equation (3.4). Recall can be thought of as a measure of completeness of classifiers. A low recall indicates many False Negatives.

$$TPR = \frac{TP}{TP+FN} \tag{3.4}$$

(ii) **False Positive Rate (*FPR*)**: Likelihood a test being actually negative when it is predicted a positive. It is the proportion of cases that are incorrectly classified as positive and are actually a negative. A low *FPR* is an indication of good classification and given by equation (3.5).

$$FPR = \frac{FP}{FP+TN} \tag{3.5}$$

Fig. 3.4 shows the plots of *TPR* and *FPR* against the number of trees grown. It has been observed from Fig. 3.4 that the best true positive rate of 0.9694 is obtained with 139 trees. Moreover, it is to be noted that the true positive rate and the false positive rate remains close to 1 and 0, respectively. Receiver Operating Characteristic (ROC) which is a plot between *FPR* and *TPR* would essentially be a point near (0,1) representing 100% sensitivity (all true positive) and 100% specificity (no false positive).

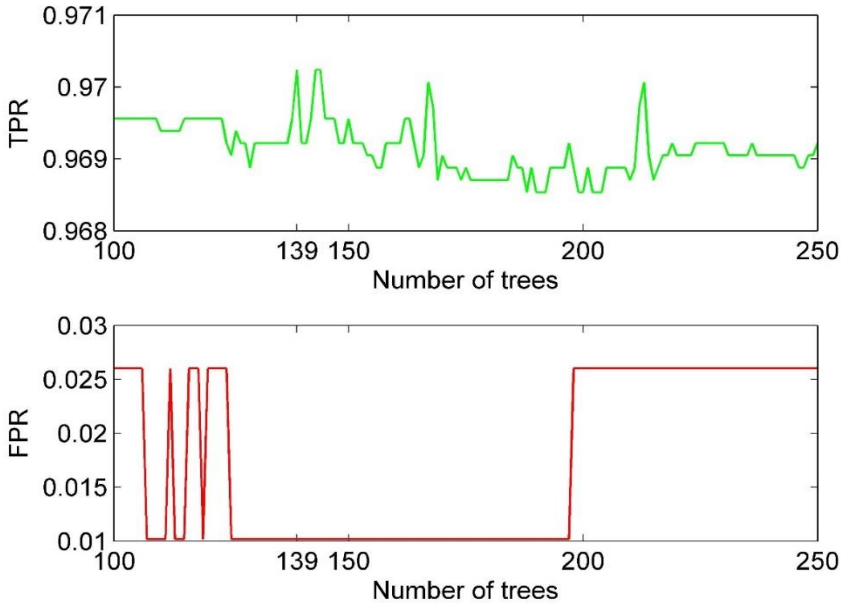


Fig. 3.4 *TPR* and *FPR* as a function of number of trees grown

(iii) **Precision**: It is the measure of how often the algorithm is correct when it predicts a correct value. It is the ratio of the number of positive predictions obtained to the total number of positive class values predicted. Precision is the measure of exactness of a classifier and given by equation (3.6). A low precision indicates a large number of False Positives.

$$Precision = \frac{TP}{TP+FP} \tag{3.6}$$

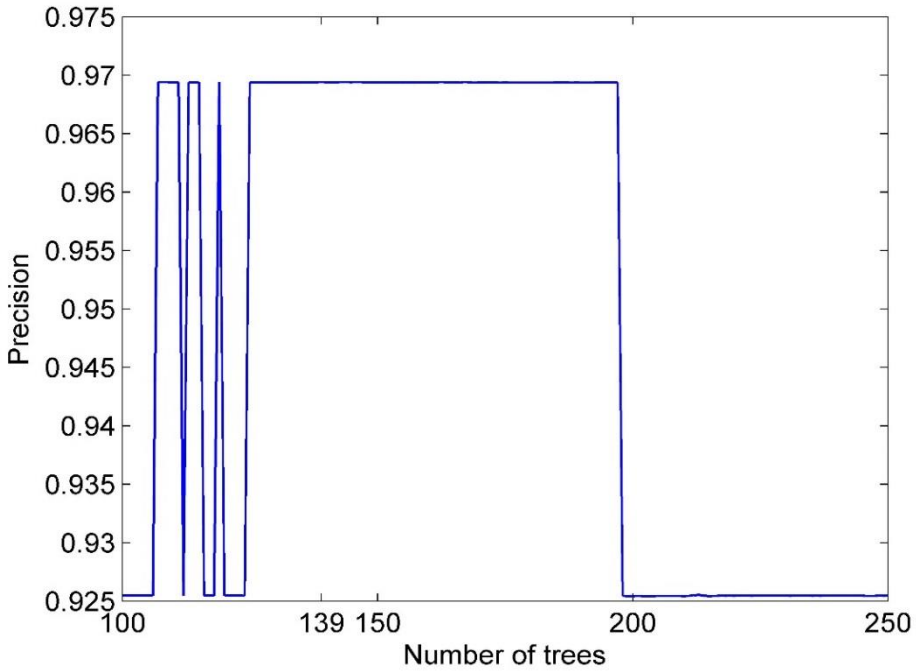


Fig. 3.5 Variations of *Precision* against number of trees grown

(iv) ***F-score*:** The advantage of having the two numbers, Precision and Recall is that one is often more important than the other in many circumstances and there is usually a trade-off between the two. *F-score* gives balance between the precision and the recall. It is the weighted average of Precision & Recall (TPR) and given by equation (3.7).

$$F-Score = \frac{2\times \mathrm{Pr}ecision\times \mathrm{Re}call}{\mathrm{Pr}ecision+\mathrm{Re}call} \tag{3.7}$$

(v) **Out of Bag (OOB) error:** In order to get unbiased test results, Random Forest has its own mechanism of cross validation. Each tree is grown on a different bootstrap sample drawn

from the original data. Each tree is constructed from about 63% of the original training data (due to the bootstrap sampling process) and rest of 37% data is available for testing any single tree. Based on this, average error rate is computed and it is known as OOB error. Fig. 3.6 shows the graph of OOB error and *F-score* against the number of trees grown.

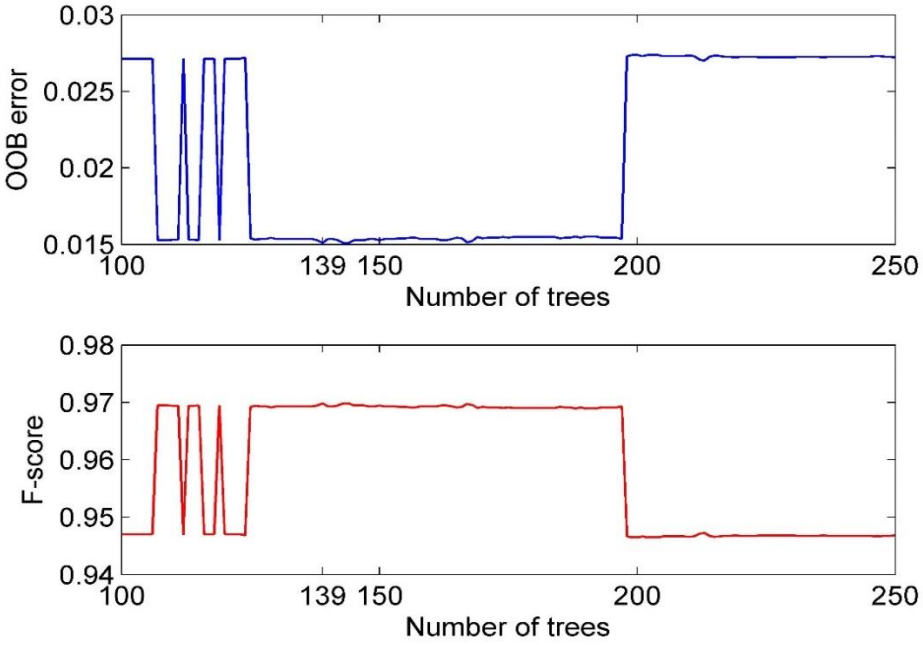
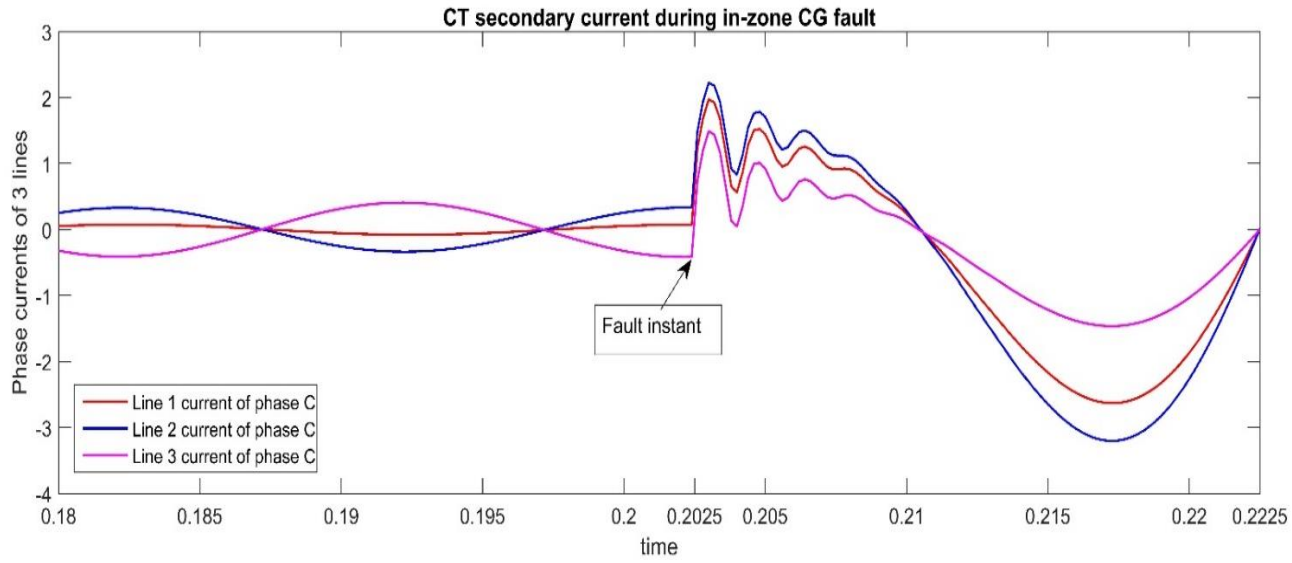


Fig. 3.6 Variations of OOB error and *F-score* against number of trees grown

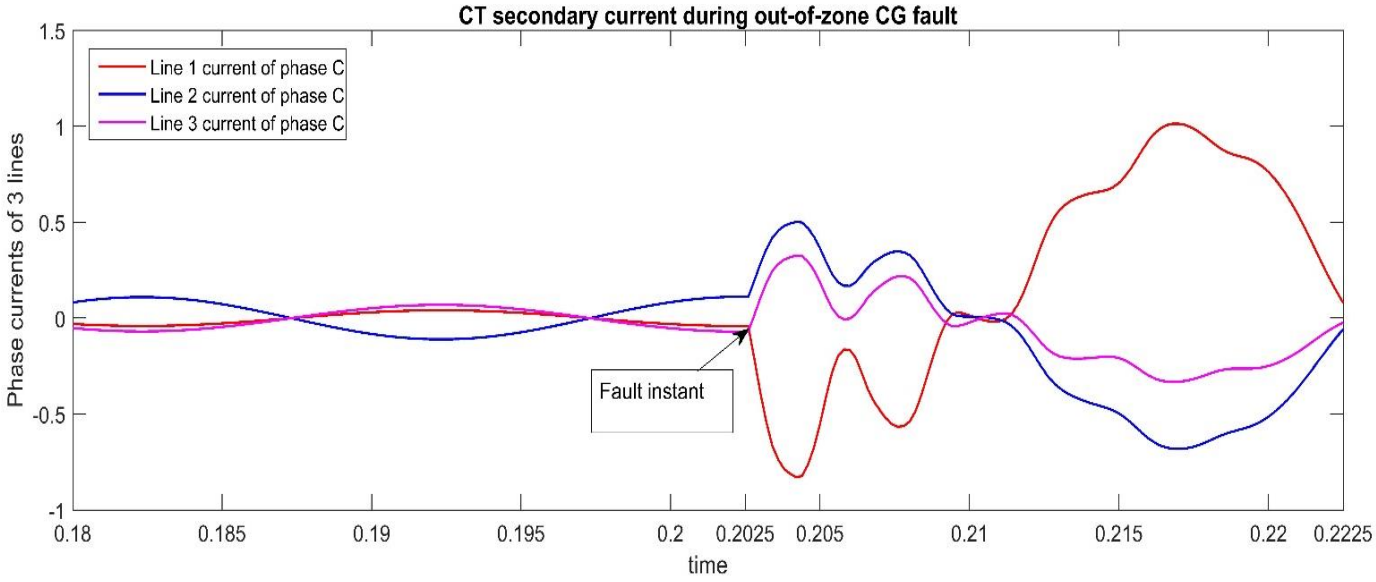
A classifier with *F-score* close to 1 has better information retrieval capability. It has been observed from Fig. 3.6 that the graph of *F-score* always stays from 0.94 to 1.0 and maximum score is obtained as 0.9698 with 139 trees. Subsequently, OOB error also reaches a minimum of value of 0.015 with 139 trees.

3.5 Results and Discussion

In order to evaluate the performance of the presented scheme, simulation of 23,520 test cases have been carried out by considering variations in system and fault parameters. Since the best performance measures are obtained using 139 trees, the same value has been selected for the purpose of validating the presented scheme. Fig. 3.7 shows the patterns of CT secondary currents during in-zone and out of zone single line to ground (C-G) fault. These patterns are given as an input to the presented RF based classifier.



(a)



(b)

Fig. 3.7 Current waveforms of faulted phase of all the bays during (a) in-zone and (b) out-of-zone single line to ground fault

It is to be noted from Fig. 3.7 that the signature of currents during out of zone and in-zone faults are entirely different. Therefore, the presented scheme discriminates internal faults from external faults accurately.

3.5.1 Effect of varying the fault resistance

Table 3.2 shows the effect of change in fault resistance values on the accuracy of the scheme. A total of 3920 cases have been utilized for testing the algorithm in which 980 cases are for internal faults whereas 2940 cases are for external faults. It has been seen from Table 3.2 that the overall accuracy of the order of more than 97% is obtained irrespective of fault resistance values.

Table 3.2 Fault classification accuracy (η) with variable fault resistance (R_f)

R_f (Ω)	Internal Faults				External Faults				Total			
	Total	TP	FN	η (%)	Total	TP	FN	η (%)	Total	TP	FN	η (%)
0	980	917	63	93.57	2940	2907	33	98.88	3920	3824	96	97.55
10	980	936	44	95.51	2940	2934	6	99.79	3920	3870	50	98.72
20	980	976	4	99.59	2940	2935	5	99.83	3920	3911	9	99.77
30	980	977	3	99.69	2940	2929	11	99.63	3920	3906	14	99.64
40	980	977	3	99.69	2940	2925	15	99.49	3920	3902	18	99.54
50	980	878	102	89.59	2940	2934	6	99.79	3920	3812	108	97.25

3.5.2 Effect of change in source impedance values

As shown in Table 2.1, total of 7 source impedance values have been used for data generation in three different combinations. The robustness of the scheme has been investigated separately for each source impedance combinations. Here, each source impedance combinations contain 3360 cases out of which 840 cases are for internal faults whereas 2520 cases are for external faults. All the seven source impedance combinations which include 23,520 (3360×7) data cases have been considered to carry out evaluation of performance of the scheme and the results are presented in

Table 3.3. It can be seen from Table 3.3 and Fig 3.8 that an overall accuracy of more than 94% is obtained from the RF classifier based scheme for all source impedance combinations.

Table 3.3 Fault classification accuracy (η) with source resistance combinations

SOURCE CASES	INZONE FAULTS				OUT OF ZONE FAULTS				TOTAL			
	TOTAL	TP	FN	η (%)	TOTAL	TP	FN	η (%)	CASES	TP	FN	η (%)
S1	840	763	78	90.8333	2520	2475	45	98.2142	3360	3238	123	96.3690
S2	840	826	14	98.3333	2520	2478	42	98.3333	3360	3304	56	98.3333
S3	840	826	14	98.3333	2520	2520	0	100	3360	3346	14	99.5833
S4	840	826	14	98.3333	2520	2360	160	93.6507	3360	3186	174	94.8214
S5	840	826	14	98.3333	2520	2478	42	98.3333	3360	3304	56	98.3333
S6	840	826	14	98.3333	2520	2478	42	98.3333	3360	3304	56	98.3333
S7	840	826	14	98.3333	2520	2478	42	98.3333	3360	3304	56	98.3333

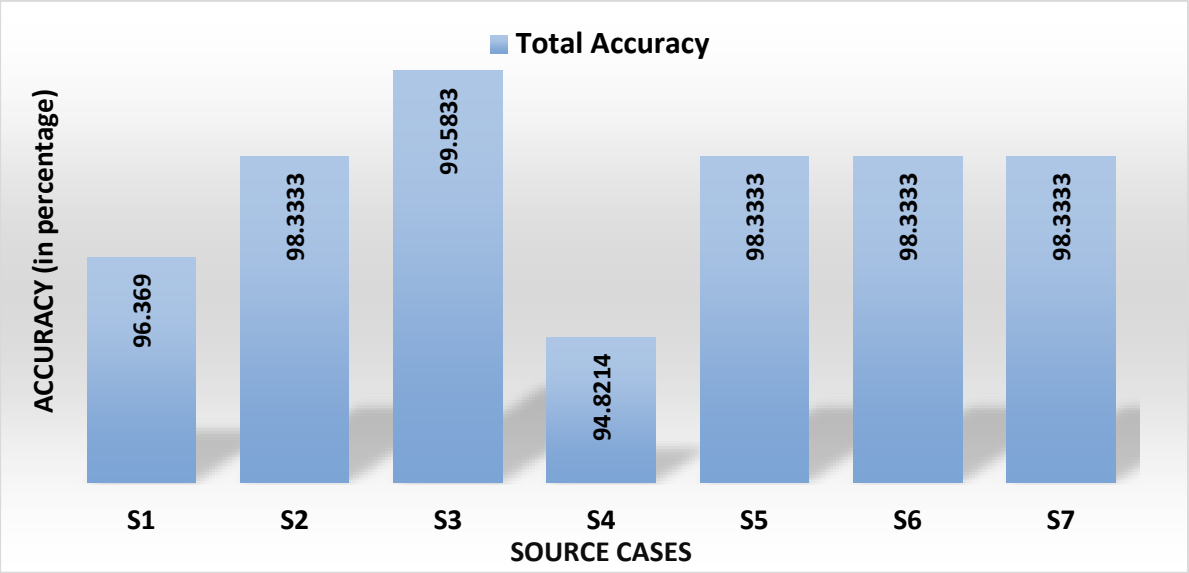


Fig. 3.8 Percentage accuracy given by the algorithm for different source impedance combinations

3.5.3 Effect of change in fault inception angle

Table 3.4 shows the percentage accuracy given by the scheme in case of change in fault inception angles with a total of 4704 cases. It is to be noted from Table 3.4 that the scheme maintains an accuracy of more than 90% for different values of fault inception angles.

Table 3.4 Fault classification accuracy (η) with fault inception angle (FIA)

FIA	INZONE FAULTS				OUT OF ZONE FAULTS				TOTAL			
	TOTAL	TP	FP	η (%)	TOTAL	TN	FN	η (%)	CASES	TP	FP	η (%)
0	1176	1157	19	98.3843	3528	3294	234	93.3673	4704	4451	253	94.6216
45	1176	1157	19	98.3843	3528	3471	57	98.3843	4704	4628	76	98.3844
90	1176	1157	19	98.3843	3528	3079	449	87.2732	4704	4236	468	90.051
115	1176	1157	19	98.3843	3528	3293	235	93.339	4704	4450	524	94.6003
150	1176	1161	15	98.7244	3528	3422	106	96.9954	4704	4583	121	97.4277

3.5.4 Effect of change in types of fault

Various types of faults have been considered to carry out performance evaluation of the scheme and the results are presented in Table 3.5. It can be observed from Table 3.5 that the overall accuracy offered by the presented scheme is more than 94% irrespective of type of faults. On the other hand, it gives an accuracy of more than 98% for single line-to-ground faults, occurrence of which is maximum compare to other types of faults on busbar [20].

Table 3.5 Fault classification accuracy for various types of faults

Fault Type	Internal faults				External faults				Overall accuracy			
	Total Cases	TP	FN	η (%)	Total Cases	TP	FN	η (%)	Total Cases	TP	FN	η (%)
LG	1764	1764	0	100	5292	5211	81	98.47	7056	6975	81	98.85
LL	1764	1659	105	94.05	5292	4990	302	94.29	7056	6649	407	94.23
LLG	1764	1764	0	100	5292	4906	386	92.71	7056	6670	408	94.53
LLLG	588	575	13	97.79	1764	1764	0	100	2352	2339	13	99.45

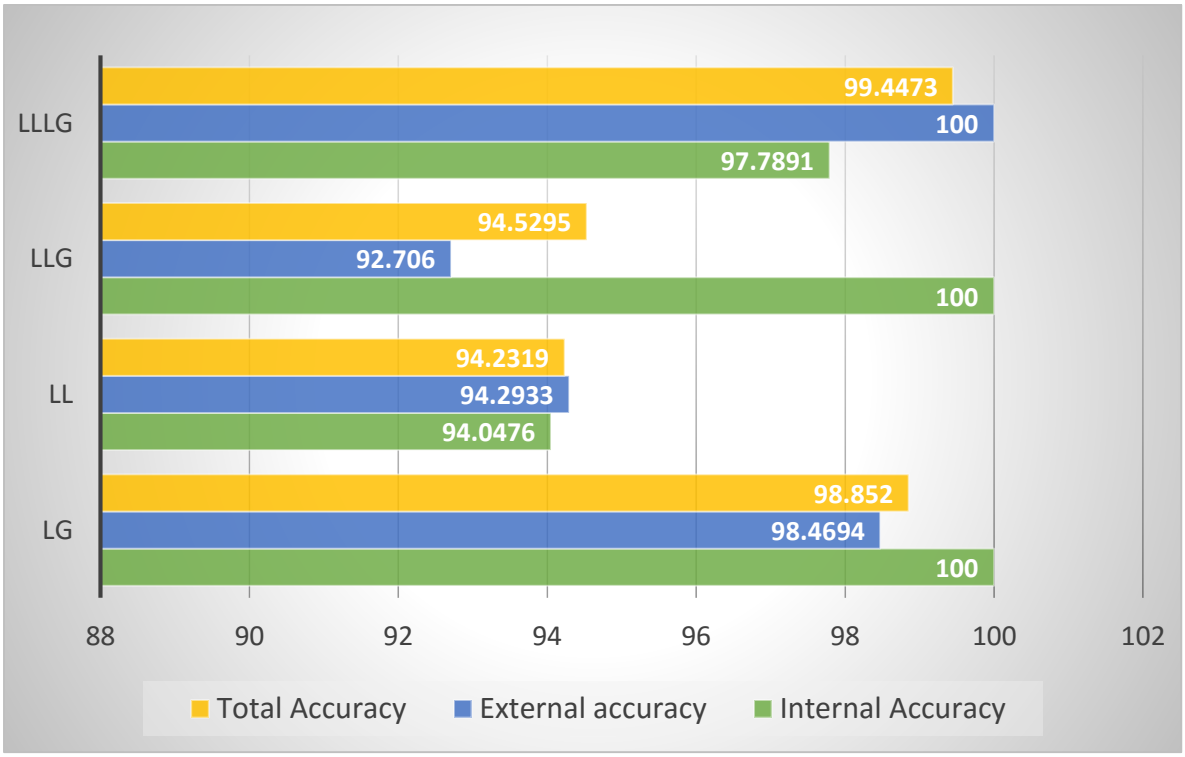


Figure 3.9 Bar chart representation of accuracy with different types of fault.

3.5.5 Effect of CT Saturation

It is crucial to study the effect of Current Transformer (CT) saturation on a busbar protection scheme in case a heavy through fault occurs. It is important for the protection scheme to maintain stability during external fault and offer high sensitivity during low current internal faults.

For validating the authenticity of the scheme when Current Transformer (CT) saturation occurs, different through faults have been simulated on line L1 to L3. All these faults are simulated at very close distance of the order of 50 to 100 m from the busbar with varying fault parameters. Furthermore, severe most close-in fault has been simulated at a point on a connector connecting the CVT to the transmission line towards outer gantry side. Fig. 3.10 shows CT secondary current waveform during heavy through fault with saturation of CT. It has been observed from Fig. 3.10 that the pattern of the waveform of current during CT saturation condition is completely different than the waveform of currents during in-zone fault. The results of simulations are depicted in Table 3.6. As seen from Table 3.6, the RF based scheme remains inoperative when severe CT saturation condition occurs, as accuracy obtained by the presented scheme is very close to 100%.

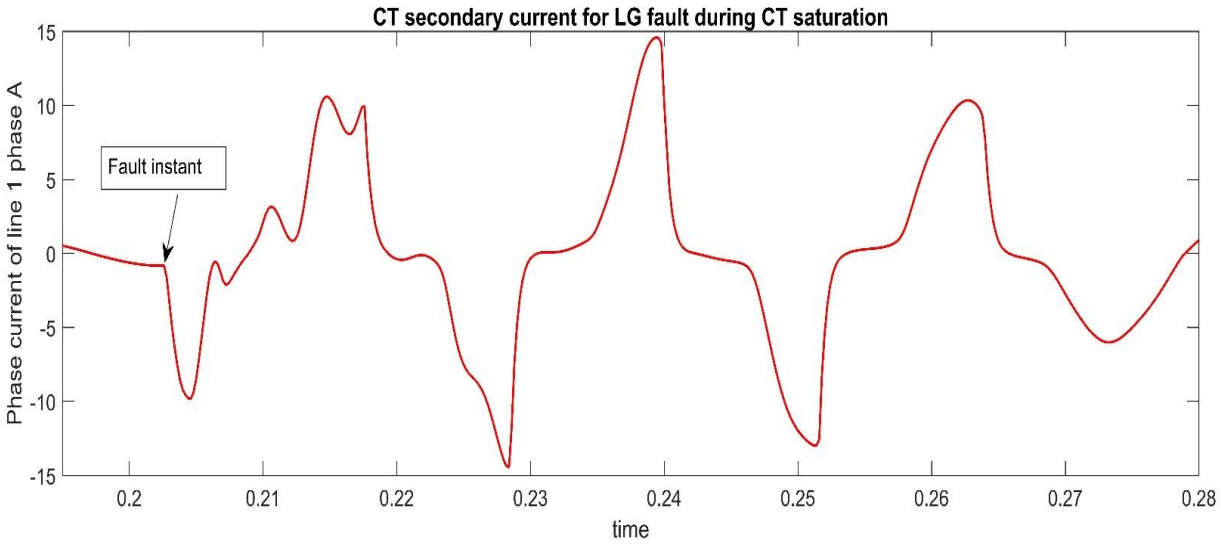


Fig. 3.10 CT secondary currents during heavy through faults (considering saturation of CT)

Table 3.6 Performance of the presented scheme during CT saturation phenomena

Faults Cases	**Total test cases**	**Total Classified Data**		**Accuracy (η) %**
		TP	**TN**	
Faults on lines L1-L3 (at 50 to 100 m with CT saturation)	480	478	02	99.58

3.5.6 Comparison between existing scheme based on SVM and RF classifier based scheme

It is worthwhile to check the sensitivity of the presented scheme during in-zone ground fault with high fault resistance and the stability of the same scheme during cross country fault. Hence, to validate the authenticity of the presented scheme, different types of internal faults (LG and LLG) and cross country fault with $R_f = 50\ \Omega$ have been simulated on line L1. The performance of an existing scheme based on SVM has also been evaluated on the above mentioned situations. Table 3.7 shows comparison between the presented scheme and the scheme based on SVM [21] for cross country external faults and internal fault with high value of resistance. It can be concluded from Table 3.7 that the presented scheme is able to distinguish cross country faults and high resistance in-zone faults with an accuracy of more than 99% while the SVM based scheme provides 90.81% and 89.51% accuracy, respectively, for the said two conditions. Therefore, the RF classifier based scheme is superior to the scheme based on SVM. This is because of the fact that the presented scheme effectively discriminates between the above two situations as the signatures of the waveform are entirely different whereas they are almost identical in case of scheme based on SVM.

Table 3.7 Comparison between the RF classifier based scheme and SVM based scheme.

Total Cases	**Test cases**	**Proposed RF based Scheme**		**Existing scheme based on SVM**	
		Total Classified Data	**Accuracy (η) %**	**Total Classified Data**	**Accuracy (η) %**
Internal faults with considerable resistance ($R_f = 50\Omega$)	1764	1753	99.37	1579	89.51
Cross country faults on line L1	882	877	99.43	801	90.81

3.6 Advantages of the presented scheme over existing schemes

1. The algorithm utilizes only 30% of the total data generated (10,080 out of 33,600) for training. Conversely, all the existing schemes require much larger data sets for training purpose. Thus, the scheme is superior to almost all the existing schemes.
2. The *F-score* obtained for this scheme is as high as 0.9698. Thus, the information retrieval capability of the used scheme is very high. Moreover, there is a high degree of balance between Precision and Recall of the scheme.
3. The *TPR* stays constantly above 0.97 and the *FPR* stays below 0.26 for the scheme. The ROC would therefore be only a point close to (0,1) which is the case of an excellent classification indicating 100% specificity and sensitivity.
4. The scheme gives 97.02 % accuracy with 139 trees for internal faults which should be correctly identified and cleared as fast as possible for minimum damage to the busbar.
5. The scheme gives 98.98 % accuracy for external faults during which the relay must be stable and should not operate. A faulty operation of the relay during out-of-zone faults will result into undesirable black out of the complete substation.

6. It has been observed that the existing scheme (using Support Vector Machine) gives poor accuracy of the order of 90% in case of cross country external faults. Cross country external faults have been simulated by considering a double circuit line in Fig. 2.1 between the bus section and source S3. The performance of the scheme based on RF classifier has been evaluated for the above faults and found that it gives an accuracy of more than 95% which is higher than the accuracy given by the existing scheme.

CHAPTER-4

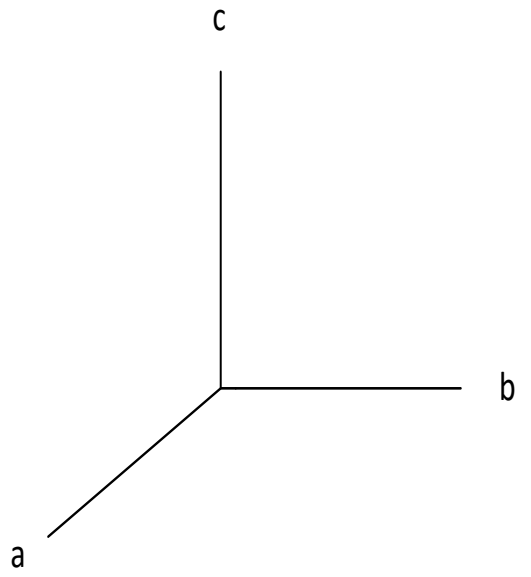
DQ TRANSFORM BASED FAULT ZONE IDENTIFICATION SCHEME FOR BUSBAR

4.1 The Synchronous Reference Frame Theory or the *dq* transform

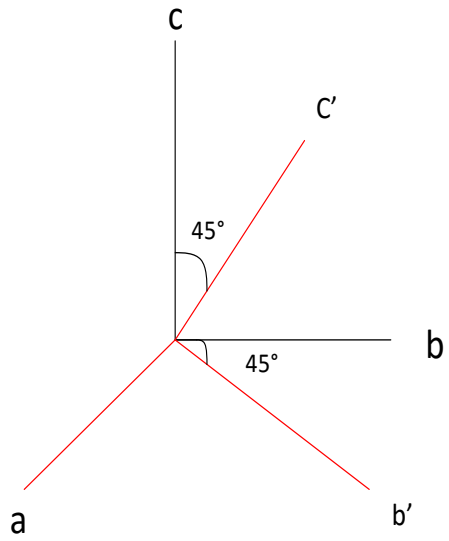
The Synchronous Reference Frame Theory or *dq* transformation is a mathematical approach to rotate the reference frame of a three-phase system so that the analysis of 3 phase circuits can be simplified.

This transformation is a two-step process:

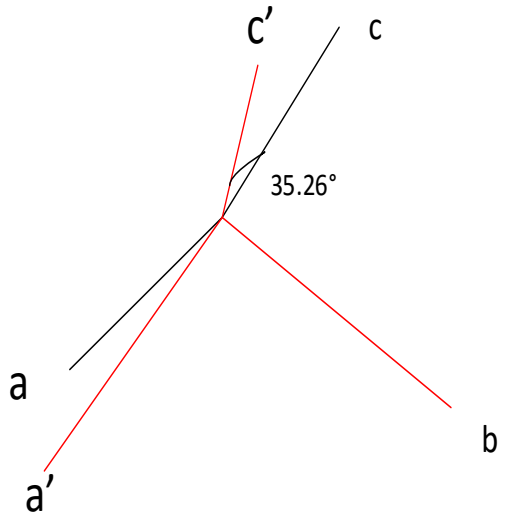
1) Firstly, the three-phase stationary coordinate system is transformed into the two-phase, stationary coordinate system or the $\alpha\beta$ coordinate system. After this rotation, α, β, and z are the new rotated a, b and c axes with z axis equidistant from all the three original a, b, and c axes. For a balanced system, the values on the 3 axes would be balanced thus making the value on the z-axis zero. This reduction of the number of relevant frames is one of the fundamental properties of the dq0 transformation.



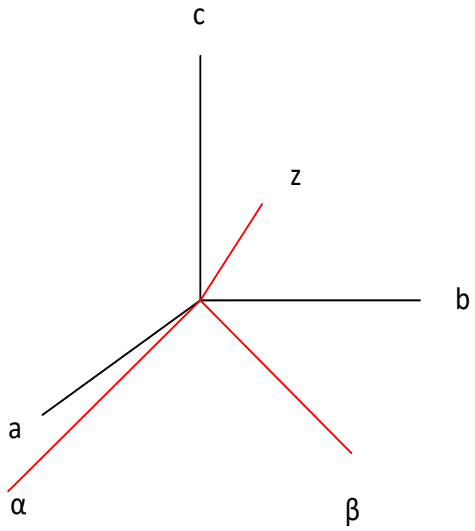
(a) Three phase coordinate system



(b) Rotating a axis by -45°



(c) Rotating new b' axis by 35.26°



(d) The α*β* stationary coordinates

Fig. 4.1 Transformation of three phase coordinate system into α*β* coordinate system

2) Then, transformation of the α*β* coordinate system into the *dq* coordinate system takes place. The former is a stationary coordinate system while the latter is a rotating coordinate system. This transformation is carried out by rotating the α*β* frame at fundamental frequency.

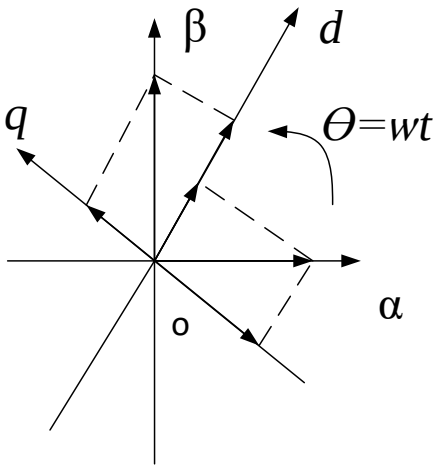


Fig. 4.2 Transformation of α*β* coordinate system into *dq* coordinate system

The three phase currents are given as:

$$i_a = \sqrt{2}\, I\sin(\omega_s t) \tag{4.1}$$

$$i_b = \sqrt{2}\, I\sin(\omega_s t + \tfrac{2\pi}{3}) \tag{4.2}$$

$$i_c = \sqrt{2}\, I\sin(\omega_s t - \tfrac{2\pi}{3}) \tag{4.3}$$

Clark's transformation modifies the 3 phase system into the following 2 phase orthogonal system:

$$i_\alpha = \tfrac{2}{3}\, i_a - \tfrac{1}{3}(i_b - i_c) \tag{4.4}$$

$$i_\beta = \tfrac{2}{\sqrt{3}}(i_b - i_c) \tag{4.5}$$

$$i_z = \tfrac{2}{3}(i_a + i_b + i_c) \tag{4.6}$$

Transformation of $\alpha\beta$ coordinate system (stationary) to the dq coordinate system (rotating) is performed by the following formula:

$$i_d = i_\alpha.\cos(\omega t) + i_\beta.\sin(\omega t) \tag{4.7}$$

$$i_q = -i_\alpha.\sin(\omega t) + i_\beta.\cos(\omega t) \tag{4.8}$$

Thus, the transformation of a three phase system into a two phase synchronously rotating orthogonal system can be performed in a single step by following equation:

$$\begin{bmatrix} i_d \\ i_q \\ i_0 \end{bmatrix} = \frac{2}{3}\begin{bmatrix} \cos(\omega t) & \cos(\omega t - \frac{2\pi}{3}) & \cos(\omega t + \frac{2\pi}{3}) \\ -\sin(\omega t) & -\sin(\omega t - \frac{2\pi}{3}) & -\sin(\omega t + \frac{2\pi}{3}) \\ \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \end{bmatrix}.\begin{bmatrix} i_a \\ i_b \\ i_c \end{bmatrix} \tag{4.9}$$

The major benefit of this theory is the mapping of AC signals at fundamental frequency into DC values. Calculations can be greatly simplified using these DC values. Thereafter, inverse dq0 transform can be applied to get back the original three-phase AC results. Furthermore, it is known to reduce the effects of harmonic distortion.

4.2 The Scheme

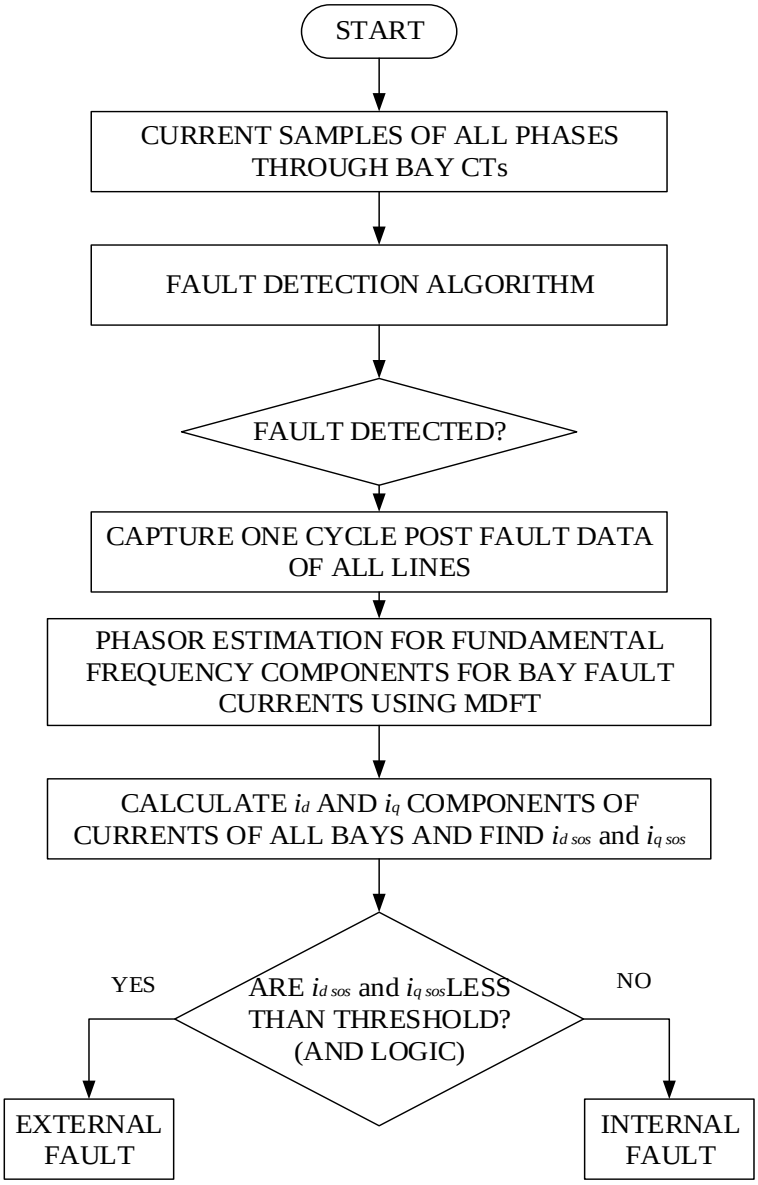


Fig. 4.3 Algorithm for identification of fault zone based on *dq* transform

The algorithm of the scheme is shown in the above Fig. 4.3. In this algorithm, the sampling frequency is 1 kHz whereas the operating fundamental frequency is 50Hz. Primarily, the data acquisition system acquires samples of current from all bay CTs (*n*-lines). Then discrimination between normal condition and fault condition is carried out using fault detection algorithm.

Whenever the fault detection algorithm discovers a fault, 3 phase post fault current samples of one cycle (20 × 3 samples) for all the connected bay CTs are sent to the modified discrete Fourier transform (MDFT) block [22], which computes the fundamental frequency components of currents. Thereafter, i_d and i_q components of the currents of all the lines are calculated using the fundamental frequency signals with the help of equation (4.9).

The i_d and i_q components of all the lines are DC values which are summed over using the following formula:

$$i_{d_{net}} = \sum_{l=1}^{L} i_d(l) \tag{4.10}$$

$$i_{q_{net}} = \sum_{l=1}^{L} i_q(l) \tag{4.11}$$

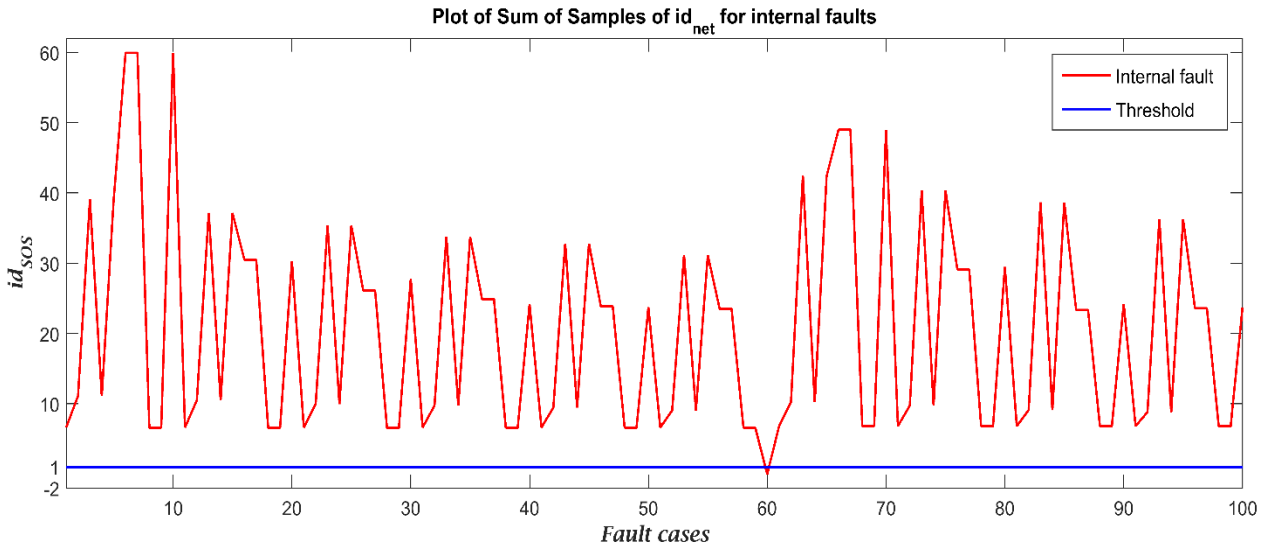
where, L is the number of transmission lines connected to the busbar. It can be observed from Fig 4.4 that the $i_{d_{net}}$ and $i_{q_{net}}$ for out-of-zone faults remain close to zero. The $i_{d_{net}}$ and $i_{q_{net}}$ are summed over for one post fault cycle to obtain their Sum of Samples (SOS) given as $i_{d_{SOS}}$ and $i_{q_{SOS}}$:

$$i_{d_{SOS}} = \left|\sum_{n=1}^{N} i_{d_{net}}\right| \tag{4.12}$$

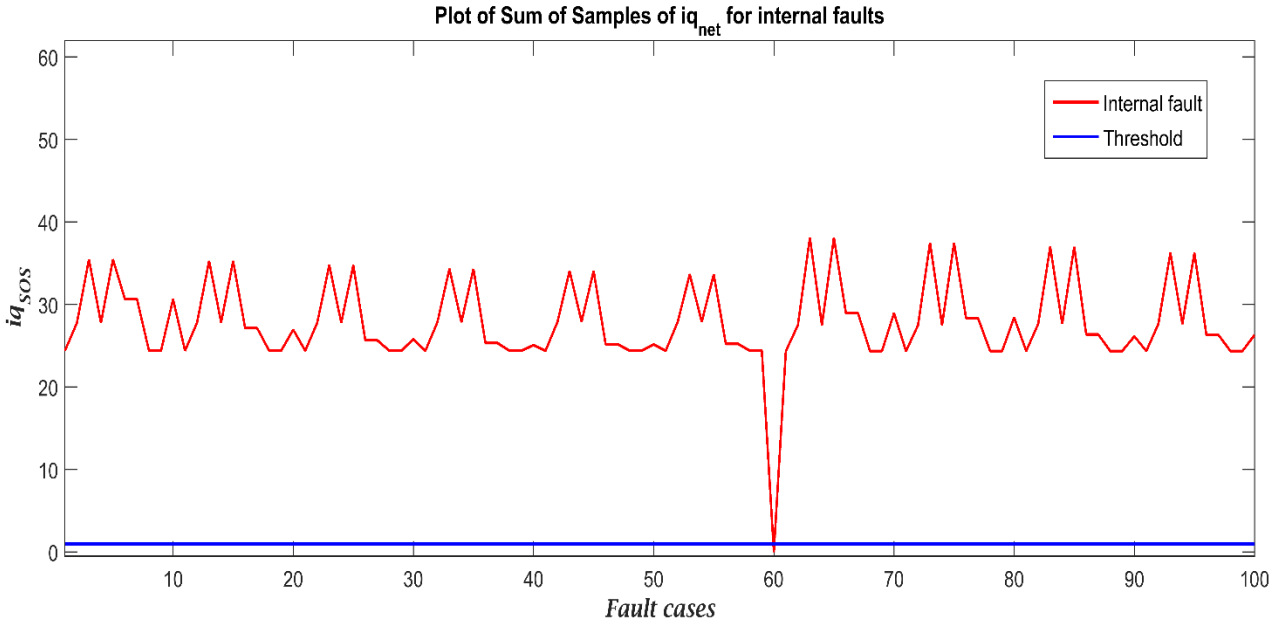
$$i_{q_{SOS}} = \left|\sum_{n=1}^{N} i_{q_{net}}\right| \tag{4.13}$$

Where, N is the number of samples in one post fault cycle. A suitable threshold is selected and $i_{d_{SOS}}$ as well $i_{q_{SOS}}$ are given as input to a comparator unit. If $i_{d_{SOS}}$ and $i_{q_{SOS}}$ are less than the threshold, the comparator's output is '-1' indicating external fault whereas, if both are greater than the threshold, the output of the comparator is '+1' indicating internal fault. The threshold selected for this purpose is 1 kA. Finally, to discriminate between internal and external faults, the output of both comparator is combined using AND logic. The MDFT algorithm has been utilized because it removes the decaying DC component in addition to the integer harmonics. High speed performance and easy implementation can be obtained employing the hardware that are presently being used in the digital relays with the help of the presented scheme as it involves less mathematical computations. The presented scheme has been developed on the supposition that at

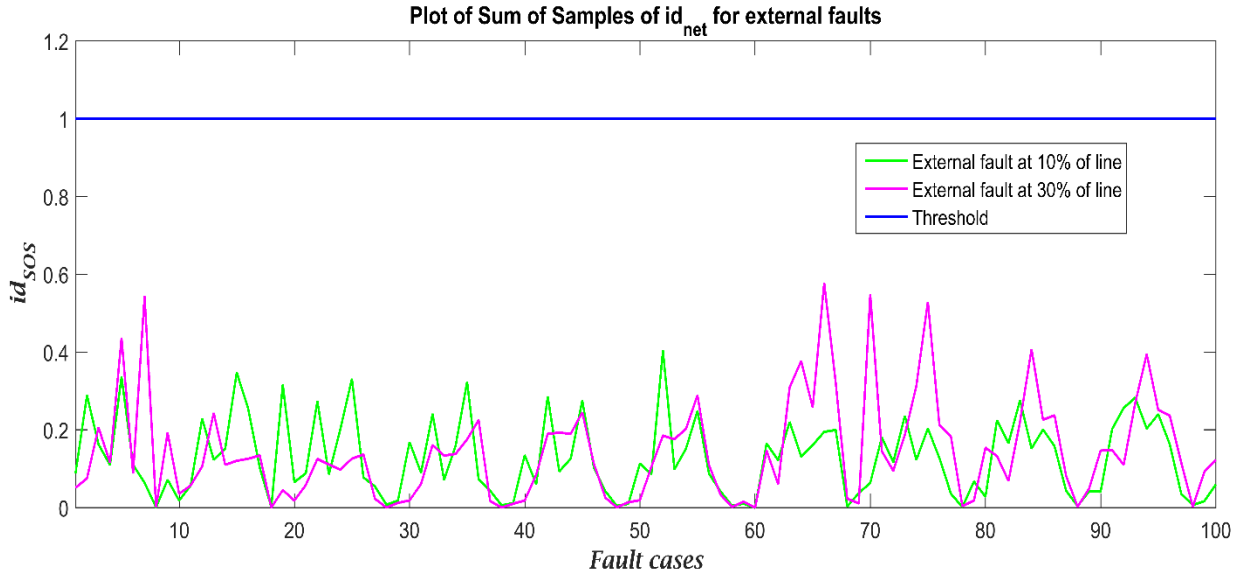
least two bays are connected to the bus which is considered. When that is not the case, the comparator unit is not initiated and the protection philosophy of back-up/feeder should be used.



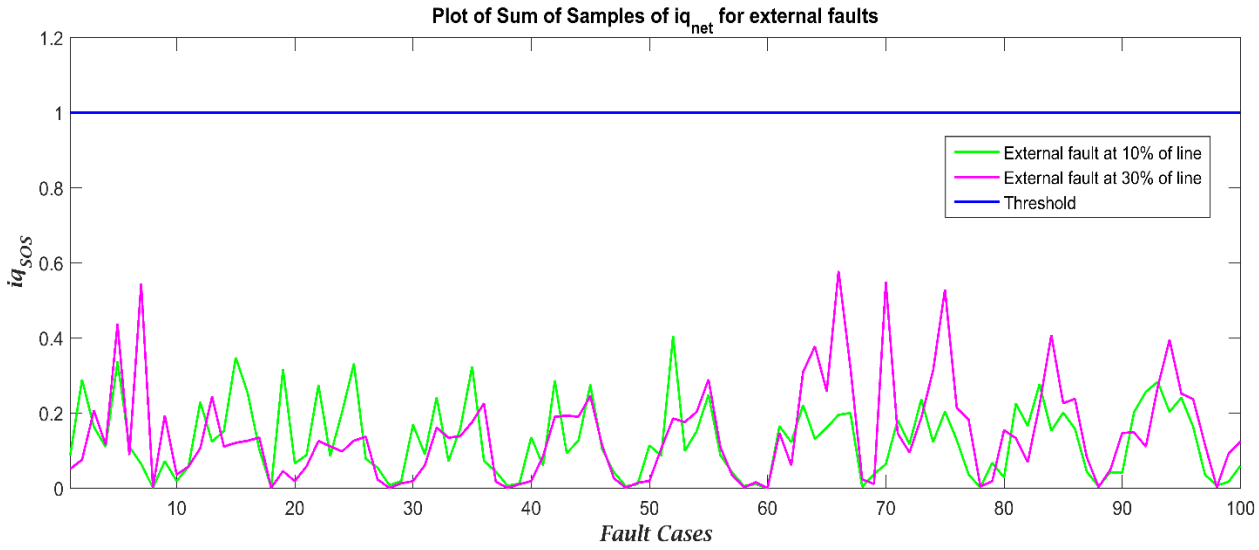
(a)



(b)



(c)



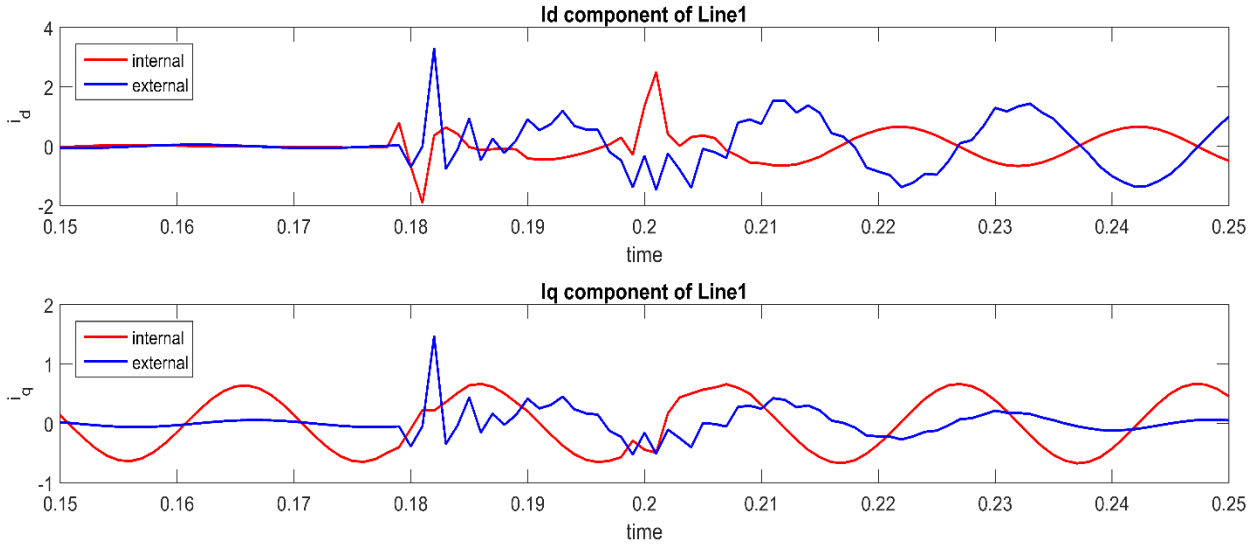
(d)

Fig. 4.4 $i_{d_{SOS}}$ and $i_{q_{SOS}}$ in kA for 100 fault cases for (a)-(b) internal faults (c)-(d) external faults

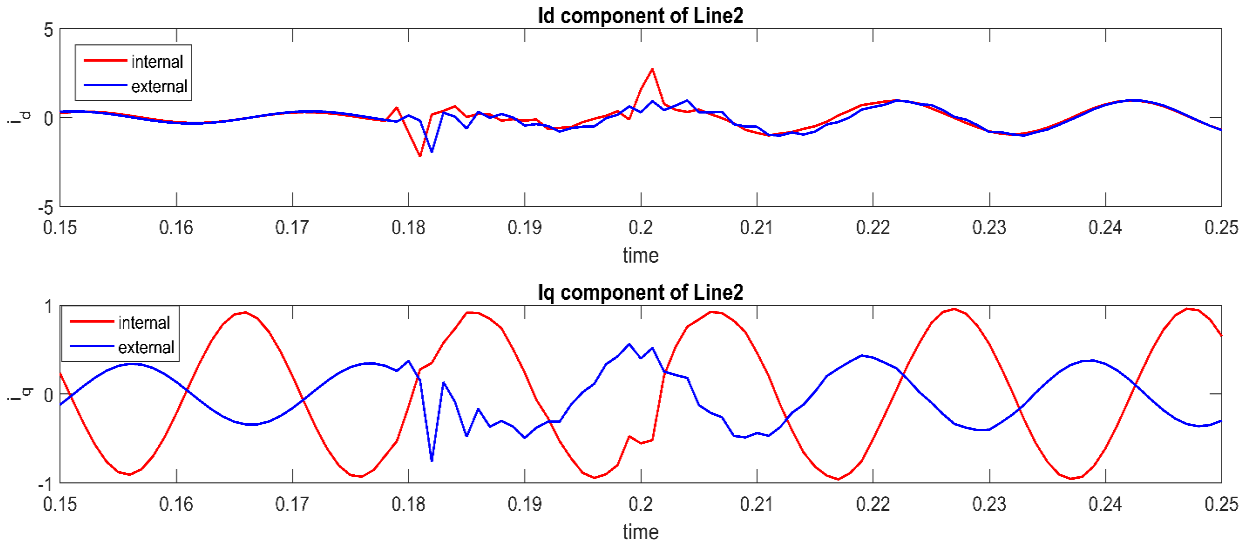
It can be seen from Fig 4.4 that both $i_{d_{SOS}}$ as well $i_{q_{SOS}}$ are greater than the threshold for internal faults and less than the threshold for external faults.

4.3 Results of Simulations and their Discussion

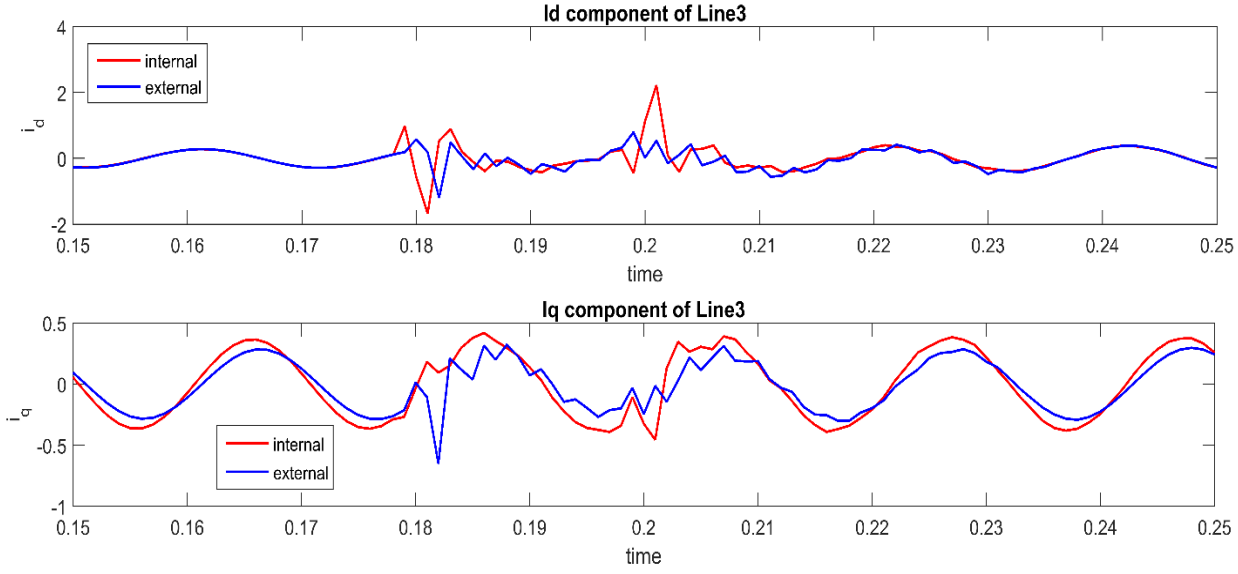
The cases which have been validated correctly are represented as true positive (TP) and those validated incorrectly are represented as and true negative (TN). For the purpose of validation of the presented scheme a total of 3600 cases have been considered which include 1200 in-zone fault cases (1 (SI) x 4 (δ) × 6 (R_f) × 5 (FIA) × 10 (Ftype)) and 2400 out-of-zone fault cases (1 (SI) × 4 (δ) × 6 (R_f) × 5 (FIA) × 10 (Ftype) × 2(FL)) and the results are shown in Table 4.1. It can be observed from Table 4.1 that this scheme effectively discriminates internal faults from external faults. The overall accuracy obtained with the scheme for in-zone as well as out-of-zone faults is 99.41% thereby establishing that the presented scheme is stable in case of occurrence of through fault and thus is effective in practical scenario. Besides, this scheme is very reliable as it provides accuracy as high as 98.47% for internal faults, which should be identified accurately for minimum damage.



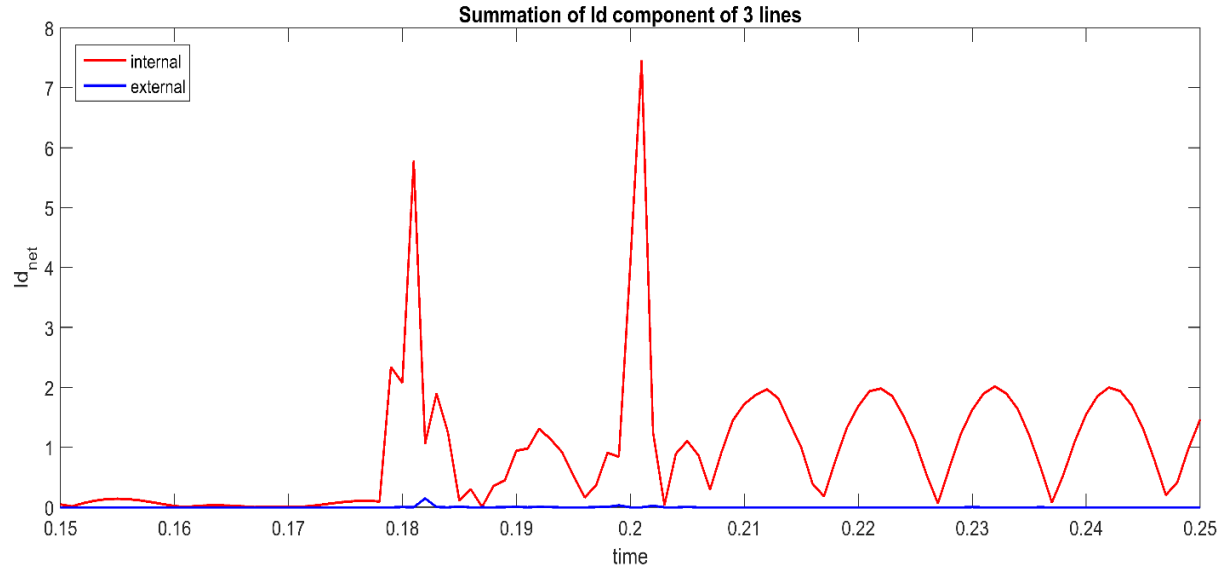
(a)


Id component of Line2
internal
external
i_d
time
Iq component of Line2
internal
external
i_q
time

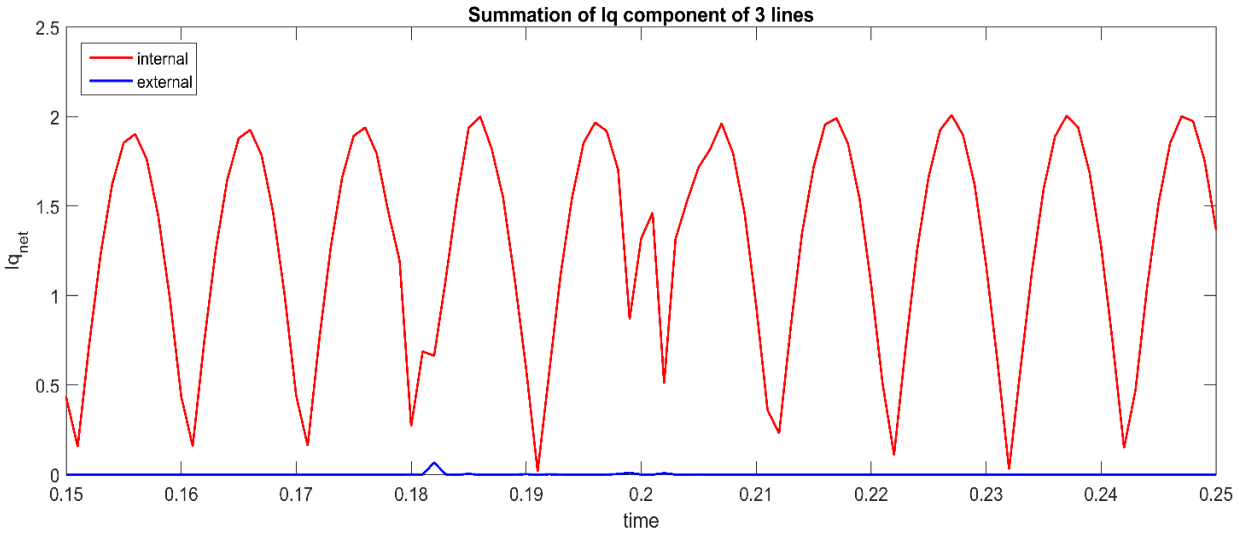
(b)


Id component of Line3
internal
external
i_d
time
Iq component of Line3
internal
external
i_q
time

(c)



(d)



(e)

Fig. 4.5 i_d and i_q components of currents in kA of the three lines during internal and external faults(a)-(c) Summation of the components (d) $i_{d_{net}}$ in kA and (e) $i_{q_{net}}$ in kA

Figures 4.5(a)-(e) show the waveforms of i_d and i_q components of all currents of all bays during internal (BC) fault and an external fault on bay L1 and the summations of i_d and i_q components, $i_{d_{net}}$ and $i_{q_{net}}$ respectively. The fault instant is 0.2 sec on time scale with a fault resistance (R_f) of 10 ohms for both the cases. It can be observed from Figure 4.5(d)-(e) that both summations

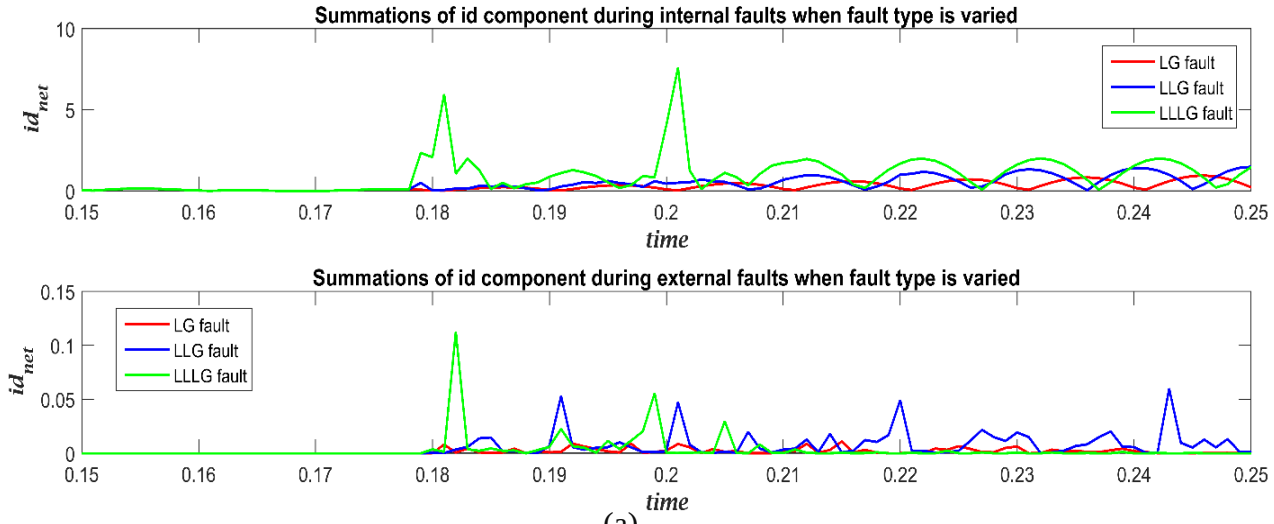
$i_{d_{net}}$ and $i_{q_{net}}$ for out-of-zone fault remains less than 1 kA. On the other hand, as shown in Figure 4.5(e), summations $i_{d_{net}}$ and $i_{q_{net}}$ for in-zone fault remain greater than 1 kA. While generating these two fault cases (in-zone and out-of-zone) the effects of decaying DC components in fault current have been taken into account. The algorithm can provide fast decomposition of currents into its fundamental frequency components and suppress the decaying DC component effectively.

Table 4.1 Accuracy of fault classification during in-zone and out-of-zone faults

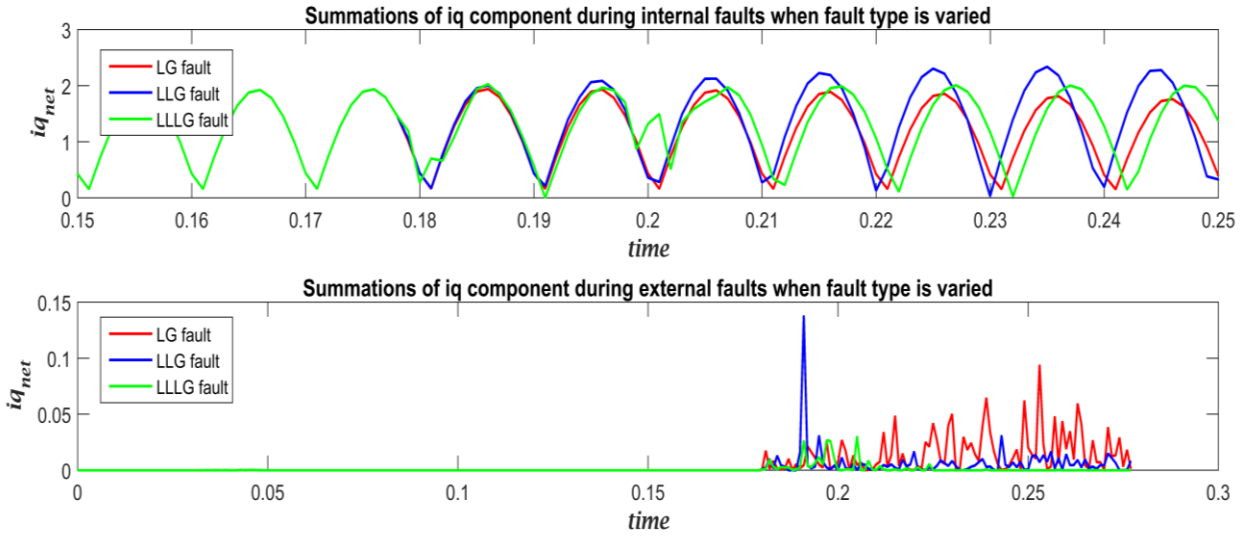
Fault cases	Number of test cases	TP	FP	Accuracy (%)
Internal	1200	1181	19	98.41
External	2400	2400	0	100
Total	3600	3581	19	99.47

4.3.1 Variation of fault type

Fig. 4.6 shows plots of $i_{d_{net}}$ and $i_{q_{net}}$ when the type of fault is varied. The fault resistance is maintained at 10 ohms and the fault instant is 0.2 secs. It can be observed from the figure that both $i_{d_{net}}$ and $i_{q_{net}}$ remain less than the threshold for external fault and greater than the threshold for internal fault case.



(a)



(b)

Fig. 4.6 $i_{d_{net}}$ and $i_{q_{net}}$ in kA as a function of time for LG, LLG, and LLLG type of faults.

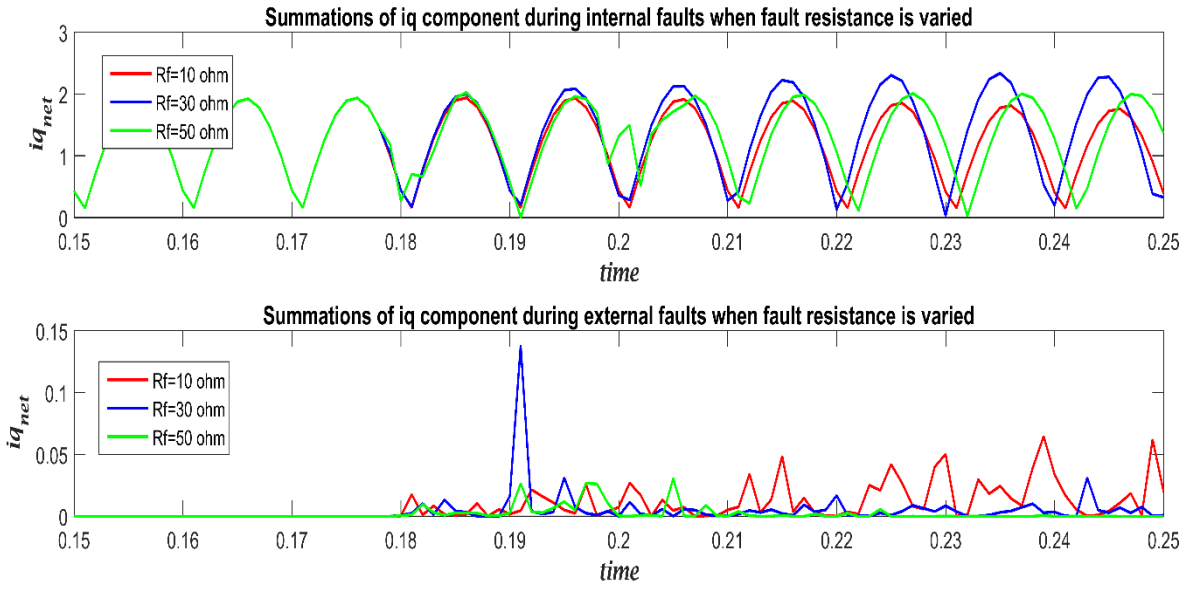
Table 4.2 $i_{d_{SOS}}$ and $i_{q_{SOS}}$ in kA for internal and external faults when the fault type is varied

Fault type	Internal faults		External faults			
			Fault at 10% of L1		Fault at 10% of L1	
	$i_{d_{SOS}}$	$i_{q_{SOS}}$	$i_{d_{SOS}}$	$i_{q_{SOS}}$	$i_{d_{SOS}}$	$i_{q_{SOS}}$
AG	6.6358	24.4020	0.0875	0.0875	0.0512	0.0512
BG	11.1384	27.7799	0.2891	0.2891	0.0754	0.0754
CG	39.0887	35.4849	0.1629	0.1629	0.2063	0.2063
ABG	11.1217	27.7823	0.1098	0.1098	0.1156	0.1156
ACG	39.0750	35.4788	0.3369	0.3369	0.4373	0.4373
BCG	59.9418	30.6393	0.1104	0.1104	0.0879	0.0879
ABCG	59.9418	30.6393	0.0639	0.0639	0.5449	0.5449
AB	6.6090	24.4094	0.0008	0.0008	0.0016	0.0016
AC	6.5934	24.4133	0.0719	0.0719	0.1937	0.1937
BC	59.9418	30.6393	0.0185	0.0185	0.0358	0.0358

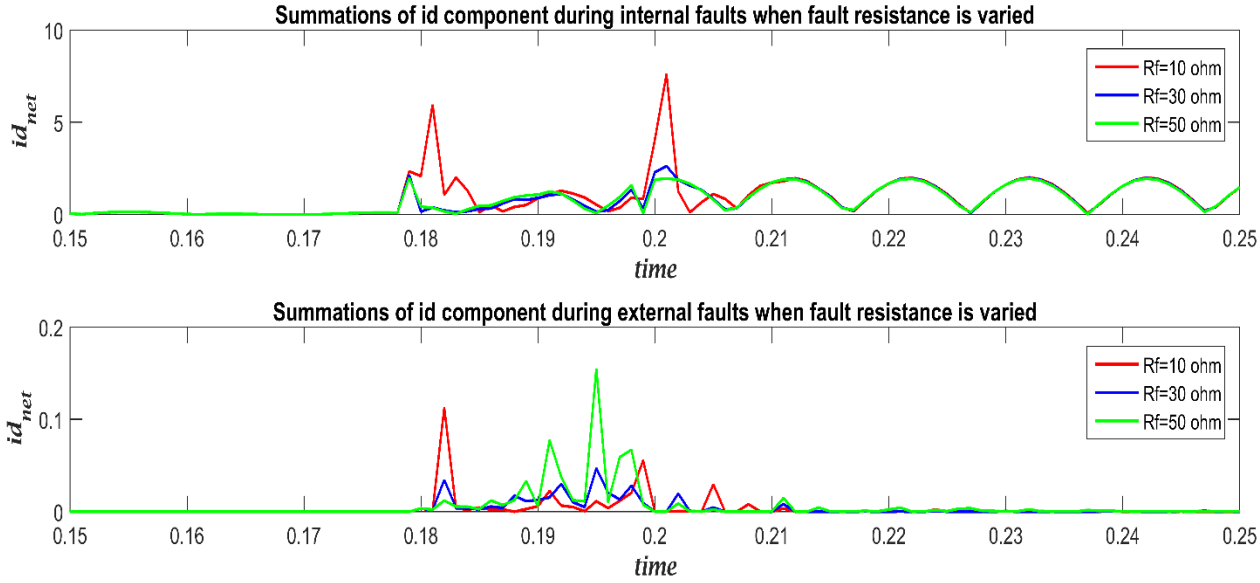
Table 4.2 shows the values of $i_{d_{SOS}}$ and $i_{q_{SOS}}$ for different types of faults. The fault instant in this case in 0.2 secs and the fault resistance is zero ohms. As can be seen from the above table both the values are always less than the threshold.

4.3.2 Variation of fault resistance

Fig. 4.7 shows plots of $i_{d_{net}}$ and $i_{q_{net}}$ when the fault resistance is varied. The time of fault is 0.2 secs and a three phase to ground fault is considered for investigation. Here also, both the parameters $i_{d_{net}}$ and $i_{q_{net}}$ remain less than the threshold for an out-of-zone fault and therefore, the algorithm is successful in discrimination of in-zone fault from out-of-zone fault irrespective of change is fault resistance.



(a)



(b)

Fig. 4.7 $i_{q_{net}}$ and $i_{d_{net}}$ in kA as a function of time when the fault resistance is changed.

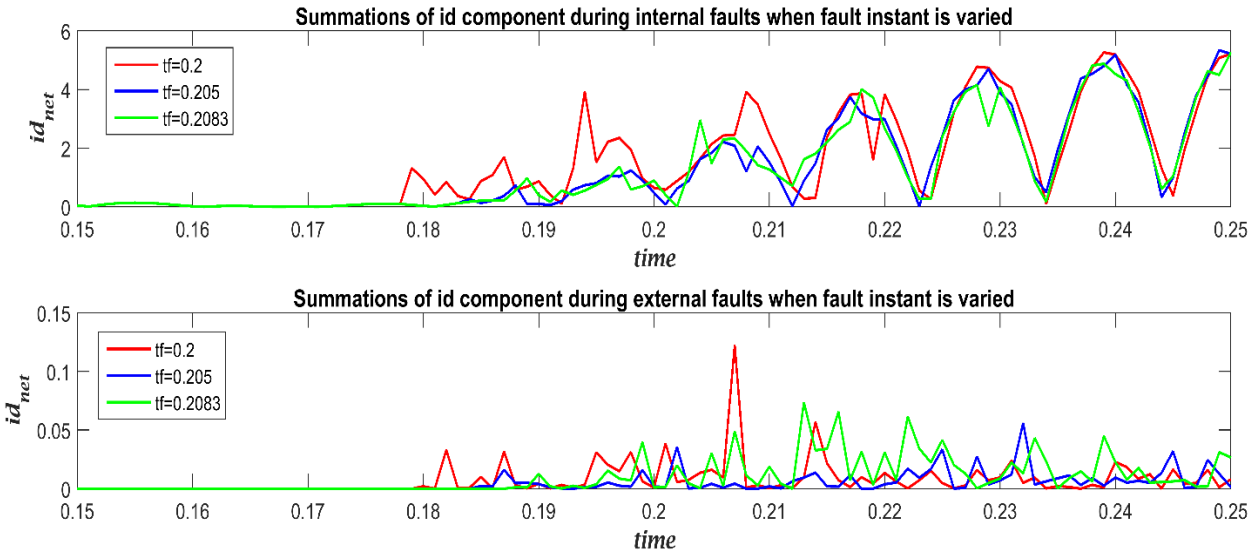
Table 4.3 $i_{d_{SOS}}$ and $i_{q_{SOS}}$ in kA for internal and external faults when the fault resistance is varied

R_f (Ω)	Internal faults		External faults			
			Fault at 10% of L1		Fault at 10% of L1	
	$i_{d_{SOS}}$	$i_{q_{SOS}}$	$i_{d_{SOS}}$	$i_{q_{SOS}}$	$i_{d_{SOS}}$	$i_{q_{SOS}}$
0	11.1217	27.7823	0.1098	0.1098	0.1156	0.1156
10	10.5014	27.8325	0.1523	0.1523	0.1105	0.1105
20	9.9728	27.7944	0.2026	0.2026	0.0976	0.0976
30	9.7511	27.8667	0.1588	0.1588	0.1383	0.1383
40	9.4559	27.9231	0.1265	0.1265	0.1903	0.1903
50	9.0505	27.9078	0.1517	0.1517	0.2029	0.2029

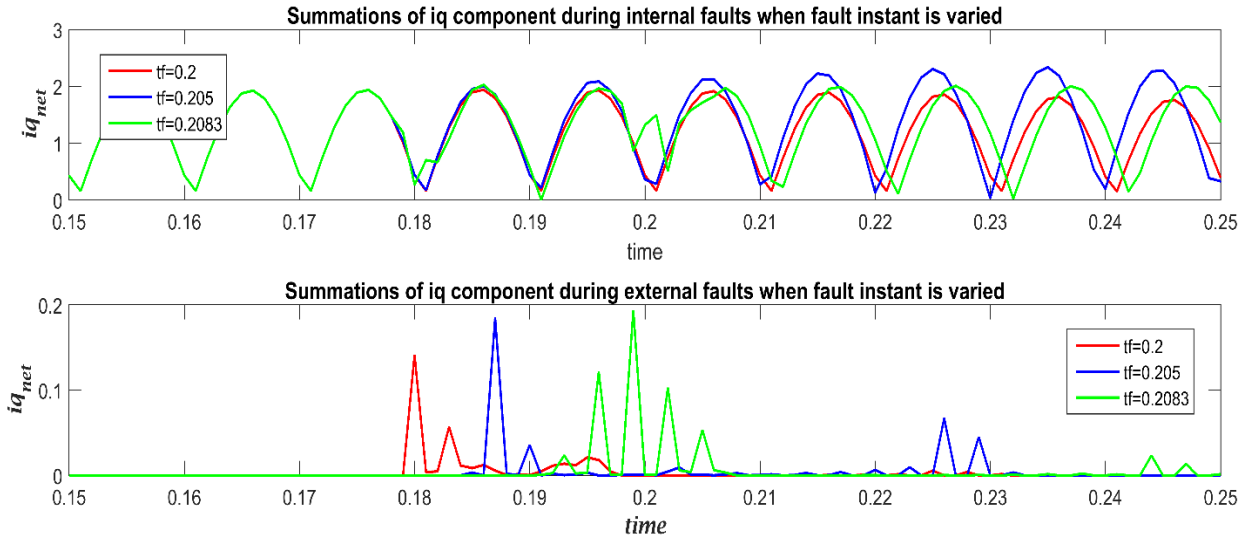
Table 4.3 shows the values of $i_{d_{SOS}}$ and $i_{q_{SOS}}$ during LLG (ABG) fault for different values of fault resistances. The fault instant in this case in 0.2 secs. As can be observed from the above table both the values are always less than the threshold for external faults.

4.3.3 Variation of fault instant

Plots of $i_{d_{net}}$ and $i_{q_{net}}$ are shown in Fig. 4.8 when the fault instant or the fault inception angle is varied. The fault resistance here is 0 ohms and an LG fault is considered for the purpose of validation of the scheme. The scheme effectively discriminates between internal and external faults owing to the fact that $i_{d_{net}}$ and $i_{q_{net}}$ are greater than the threshold for internal fault and less than threshold for external fault. Table 4.4 presents the values of $i_{d_{SOS}}$ and $i_{q_{SOS}}$ during LLG (ABG) for different values of fault inception angles.



(a)



(b)

Fig. 4.8 $i_{d_{net}}$ and $i_{q_{net}}$ in kA as a function of time when the fault instant is changed.

Table 4.4 $i_{d_{SOS}}$ and $i_{q_{SOS}}$ in kA for internal and external faults when the fault inception angle/ fault instant is varied

Fault instant (secs)	Internal faults		External faults			
			Fault at 10% of L1		Fault at 10% of L1	
	$i_{d_{SOS}}$	$i_{q_{SOS}}$	$i_{d_{SOS}}$	$i_{q_{SOS}}$	$i_{d_{SOS}}$	$i_{q_{SOS}}$
0.2	59.9418	30.6393	0.0639	0.0639	0.5449	0.5449
0.2025	49.0002	28.9679	0.1999	0.1999	0.3216	0.3216
0.2050	27.9578	25.2103	0.0557	0.0557	0.1545	0.1545
0.2064	31.2624	24.7584	0.1001	0.1001	0.0627	0.0627
0.2083	49.8857	27.3814	0.2068	0.2068	0.0534	0.0534

4.3.4 High resistance in-zone fault

Table 4.5 shows the performance of the technique during high resistance in-zone faults. It can be seen from the table that despite change in fault type and fault inception angle both $i_{d_{SOS}}$ and $i_{q_{SOS}}$ remain above the threshold.

Table 4.5 $i_{d_{SOS}}$ and $i_{q_{SOS}}$ values in kA for high resistance in-zone faults

Fault cases		Rf = 40 Ω		Rf = 50 Ω	
		$i_{d_{SOS}}$	$i_{q_{SOS}}$	$i_{d_{SOS}}$	$i_{q_{SOS}}$
AG	tf = 0.2 secs	6.6358	24.4020	6.6358	24.4020
ABG	tf = 0.2 secs	9.4559	27.9231	9.0505	27.9078
AG	tf = 0.205 secs	6.8383	24.3619	6.8383	24.3619
ABG	tf = 0.205 secs	8.5433	27.6437	8.2560	27.6341

4.5 Advantages of the scheme

1. The scheme gives 98.47 % accuracy for internal faults which must be correctly identified and cleared as fast as possible for minimum damage to the busbar.
2. The scheme gives 100 % accuracy for external faults during which the relay must be stable and should not operate. A faulty operation of the relay during out-of-zone faults will result into undesirable black out of the complete substation.
3. The main drawback directional protection schemes for busbar based on voltage and current signals is that the voltage remains near to zero during close-in faults. This scheme on the other hand uses only current signals.
4. This scheme can work effectively during internal faults with high-resistance. The reliability of the scheme in case of high-resistance internal faults has been validated by simulating 200 cases of in-zone faults.
5. CTs and CVTs are required at each and every bay in the schemes suggested in [3, 6]. Furthermore, if the communication between the IEDs and the CPU fails, these schemes

would no longer be reliable. No such additional hardware is required in this scheme based on *dq* transform

6. The MDFT algorithm has been used in this algorithm to perform all mathematical calculations. This guarantees its reliability.

CHAPTER-5
CONCLUSION

A fault zone identification scheme for protection of busbar using RF classifier is presented in the thesis. It is based on the process of feature extraction from the measured current signals of all the bays (lines) that are connected to a busbar. To validate the authenticity of the RF classifier based scheme, PSCAD/EMTDC software package has been used to model an existing Indian busbar system of 400kV. As many as 33,600 cases have been simulated including both in-zone and out-of-zone faults by varying the system and fault parameters. By providing post-fault current samples of one cycle of all the bays as input to RF classifier, the overall accuracy of fault zone identification obtained is more than 98%. Furthermore, the scheme is unaffected by variation in type of fault, fault location, fault inception angle and source impedance values. In addition, the presented scheme maintains stability when severe CT saturation phenomena occur as well as during heavy through fault. At the end, comparison between the RF classifier based scheme and the schemes which already exist clearly indicates that this scheme is superior in terms of sensitivity during internal faults and stability during external faults & disturbances.

Further, a new algorithm based on *dq* transform of secondary currents of all bay currents is presented here for bus zone identification. Post fault samples of currents of all bay CTs are given to the MDFT block for estimating its fundamental frequency components. Synchronous Reference Frame Theory or *dq* transformation is used to find out$i_{d_{SOS}}$ as well as $i_{q_{SOS}}$. Tripping of relays are carried out if both $i_{d_{SOS}}$ as well $i_{q_{SOS}}$ are greater than the threshold else it is blocked. To check whether this technique is feasible or not, it has been tested on an existing 400 kV Indian busbar system which has been modelled in PSCAD/EMTDC software package. 3,600 fault cases have been simulated by varying the fault and system parameters to validate the authenticity of this scheme. The presented scheme gives accuracy more than 98% for all in-zone faults including faults with high value of resistance. Furthermore, this scheme remains stable during different types of out-of-zone faults including through faults with heavy current.

CHAPTER-6
FUTURE WORK

The following objectives need to be incorporated in the future work to expand the study and improve the accuracy:

- Line compensation of line L1 is fixed to 50%. Performance of both the algorithms is to be checked when the compensation of the line is varied.
- The algorithm based on *dq* transform has to be applied to cases of CT saturation. The scheme needs to be stable in case of severe CT saturation conditions.
- The scheme based on *dq* transform has to be validated for cross-country faults. Cross country faults are type of two phases to earth faults. These are simultaneous earth faults which occur on different line sections and on different phases of same circuit. However, the probability of occurrence of cross-country fault is quite low.
- Comparative study of the scheme with the already existing schemes needs to carried out.

REFERENCES

[1] Working Group K14 of the IEEE PES Power System Relaying Committee, "Exploring the IEEE C37.234 Guide for Protective Relay Application to Power System Buses," *IEEE Transactions on Power Delivery*, vol. 26, no. 2, pp. 936-943, April 2011.

[2] M. S. Sachdev, T. S. Sidhu, and H. S. Gill, "A busbar protection technique and its performance during CT saturation and CT ratio-mismatch," *IEEE Transactions on Power Delivery*, vol. 15, no. 3, pp. 895–901, July 2000.

[3] A. P. Apostolov, "High speed Peer-to-Peer communications based bus protection," *In Proc. IEEE Power Engineering Society*, Winter Meet, vol. 2, pp. 693–698, February 2001.

[4] R. M. Rifaat, "Considerations in Applying Power Bus Protection Schemes to Industrial and IPP Systems," *IEEE Transactions on Industry Applications*, vol. 40, no. 6, pp. 1705-1711, November 2004.

[5] N. G. Chothani and Bhavesh Bhalja, "Development of a New Bus Zone Identification Algorithm Based on Phase Angle Comparison Using Sequence Components of Currents," *Electric Power Components and Systems*, vol. 42, no. 2, pp. 215-226, January 2014.

[6] M. R. D. Zadeh, T. S. Sidhu, and A. Klimek, "Implementation and Testing of Directional Comparison Bus Protection Based on IEC61850 Process Bus", *IEEE Transactions on Power Delivery*, vol. 26, no. 3, pp. 1530- 1537, July 2011.

[7] M. E. Mohammed, "High speed differential busbar protection using wavelet packet transform", *Proceeding IET Generation, Transmission & Distribution*, vol. 152, no. 6, pp. 927-933, November 2005.

[8] S. P. Valsan and K. S. Swarup, "Computationally efficient wavelet-transform based digital directional protection for busbars," *IEEE Transactions on Power Delivery*, vol. 22, no. 3, pp.

1342-1350, July 2007.

[9] S. A. Gafoor and P. V. Ramana Rao "A transient current based busbar protection scheme using Wavelet Transforms", *Electrical Power and Energy Systems,* vol. 33, no. 4, pp. 1049–1053, May 2011.

[10] M. M. Eissa, "A new digital busbar protection technique based on frequency information during CT saturation," *Electrical Power and Energy Systems,* vol. 45, no. 1, 2013, pp. 42–49, February 2013.

[11] Guibin Zou and Houlei Gao, "A Traveling-Wave-Based Amplitude Integral Busbar Protection Technique," *IEEE Transactions on Power Delivery*, vol. 27, no. 2, pp. 602–609, April 2012.

[12] Zhenwei Guo, Jiangang Yao and Zhewen Tan, "Hilbert-Huang transform based transient busbar protection algorithm," *IET Generation, Transmission & Distribution,* vol. 9, no. 14, pp. 2032-2039, November 2015.

[13] Nilesh Chothani, Bhavesh Bhalja and Urmil Parikh, "A New Fault Zone Identification Scheme for Busbar using Support Vector Machine," *IET Generation, Transmission & Distribution*, vol. 5, no. 10, pp. 1073-1079, October 2011.

[14] N. G. Chothani, B. R. Bhalja and Urmil Parikh, "Development of a new bus zone identification algorithm using support vector machine," *IET Generation Transmission and Distribution,* vol. 6, no. 7, 2012, pp. 710–718, July 2012.

[15] S. Song and G. Zou, "A Novel Busbar Protection Method Based On Polarity Comparison of Superimposed Current," *IEEE Transactions on Power Delivery,* vol. 30, no. 4, pp.1914 - 1922, August 2015.

[16] "PSCAD/EMTDC Power Systems Simulation Manual," Winnipeg, MB, Canada, 2001.

[17] C. Chen, A. Liaw and L. Breiman, "Using Random Forest to Learn Imbalanced Data," Department of Statistics, UC Berkeley, 2004.

[18] L. Andy, and M. Wiener, "Classification and regression by Random Forest,'' *R news* 2.3: 18-22, 2002.

[19] S. R Mohanty, A. K. Pradhan and A. Routray, "A Cumulative Sum-Based Fault Detector for Power System Relaying Applications," *IEEE Transaction on Power Delivery*, vol. 23, no. 1, pp. 79-86, January 2008

[20] P. M. Anderson, "Power System Protection," McGraw-Hill: IEEE Press, New York, pp. 645-646, 1999.

[21] N. G. Chothani and B. R. Bhalja, "A New Algorithm for Busbar Fault Zone Identification Using Relevance Vector Machine," *Electric Power Components and Systems*, vol. 44, no. 2, pp. 193-205, January 2016.

[22] S. L. Yu and J. C. Gu, "Removal of decaying DC in current and voltage signals using a modified Fourier filter algorithm,'' *IEEE Transaction on Power Delivery,* vol. 16, no. 3, pp. 372-379, July 2001.

[23] B. Crowhurst, E. F. El-Saadany, L. El Chaar, and L. A. Lamont, "Single phase grid-tie inverter control using DQ transform for active and reactive load power compensation," in *Proc. IEEE International Conference Power Energy*, Kuala Lumpur, Malaysia, pp. 489–494, Nov./Dec. 2010.

APPENDIX

Source data

Positive-sequence impedance of all sources	: 0.871 + j9.96 Ω
Zero-sequence impedance of all sources	: 1.743 + j19.92 Ω
Frequency	: 50 Hz

Change in source impedance

1. 100% of SI	: Z1=0.871 + j9.96 V, Z0 = 1.743 + j19.92
2. 80% of SI	: Z1=0.697 + j7.96 V, Z0 = 1.395 + j15.93 Ω
3. 120% of SI	: Z1=1.04 + j11.95 V, Z0 = 2.092 + j23.90 Ω

Transmission-line data

Length	: 100 km
Voltage	: 400 kV
Positive-sequence impedance	: 0.0297 + j0.332 V/km
Zero-sequence impedance	: 0.162 + j1.24 V/km
Positive-sequence capacitance	: 9.23 nF/km
Zero-sequence capacitance	: 6.72 nF/km