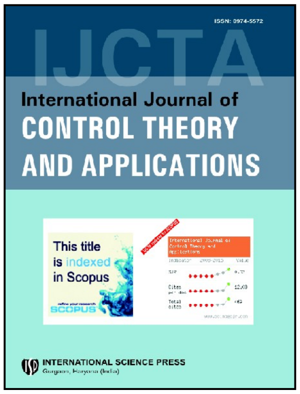

International Journal of Control Theory and Applications

ISSN : 0974-5572

Volume 10 • Number 6 • 2017

An Innovative Approach for Multiple Faults Detection in Induction Motor Using Statistical Time Measures and Random Forest Classifier

Deepak M. Sonje[1], P. Kundu[2] and A. Chowdhury[3]

[1] *Electrical Engineering Department, RHSCOEMS& R, Nasik, India, E-mail: deepaksonje123@gmail.com*
[2,3] *Electrical Engineering Department, SVNIT, Surat, India, E-mails: pk@eed.svnit.ac.in, [3]ac@eed.svnit.ac.in*

Abstract: Fault diagnosis and detection is the vital area in health monitoring of electrical machines. It is always necessary to detect the incipient fault in the induction machine at an early stage to avoid the further consequences. This paper proposes an innovative approach for fault detection in induction machine based on statistical features and classified using random forest (RF) algorithm. Initially, stator currents are acquired from the induction machine under various conditions. After preprocessing the currents, the Park's Vector Modulus (PVM) is obtained from three phase currents using d-q transformation. Next, Fourteen statistical time measures of PVM are estimated. These parameters are considered as inputs to the classifier. The stator, rotor and simultaneous faults (rotor and stator faults) are detected and classified using the proposed method. The obtained performance measures are compared with direct current analysis method. The results of proposed method show the much better performance and more accurate than direct current. For demonstration of planned fault diagnosis algorithm, experimentally obtained results are considered.

Keywords: Fault Detection, Park's Vector Modulus (PVM), Random Forest Classifier (RF), Statistical Time Measures.

I. INTRODUCTION

Induction machines are playing most important part of the process by providing the uninterrupted continuation and production in many industries. They are mainly subjected to mechanical, electrical, and thermal stresses during running condition. If any of these stresses become severe enough then various faults may initiate in the induction machine. The faults in the machine can be segregated into mainly stator, rotor, bearing and eccentricity related faults. If the faults are not sensed at an initial stage, results in premature damage of the machine and costly downtime of the plant. Numerous methods and scheme are designed for fault identification.

Many existing methods are applicable for bigger size induction machines, and very few are applicable in small size machines due to restrictions related to sensor size and cost of data acquisition. The major techniques for monitoring the fault are based on vibration, current, temperature, air gap torque, magnetic flux, and partial discharge measurement [1-4]. The methods which are non-invasive and non-intrusive are mostly considered for fault diagnosis of induction machines which monitors the motor's condition using only electrical parameters. Motor current signature analysis (MCSA) is traditional non-invasive technique which utilizes the spectral analysis of stator current for fault analysis [5]. But the limitation of MCSA is that the magnitudes of characteristic frequency components are relying on the load variation which becomes difficult for fault analysis in the motor. The sensitivity of MCSA is enhanced by combining conventional MCSA method with wavelet transform, short time Fourier transform and expert system [6].

The condition based monitoring and fault analysis of induction machines have stimulated in recent years from conventional methods to artificial intelligence (AI) methods. These methods don't require knowledge of machine parameters or any modeling of the system required. It should be noted that plenty of neural based methods are available in literature [7]–[10]. However, in the neural network (NN) based methods, architecture of a NN are not known in advance; and are obtained after a trial-and-error procedure. Some of AI methods use expert systems [11] and support vector machine (SVM) [12]. These classifiers for fault analysis are affected when the class is more in number. It is also found that the response of the classifiers becomes slow. Many detection schemes are very costly and also applicable to large size machines. By analyzing the amplitudes of current signals in the time domain, experimental results with bearing, stator, and rotor faults are tested using different pattern classification methods under varied power supply and mechanical loading conditions in [13].

The main intention of the paper is to build up a novel method to detect and classify simultaneous and individual fault in induction machine. It has been observed that as numbers of inputs to the classifier are reduced then time to build model will also reduce. In this scheme three phase stator currents are captured, preprocessed and Park's Vector Modulus (PVM) is obtained. Fourteen simple statistical parameters are estimated from PVM and these features are considered as the inputs to the classifier to detect and classify four conditions of motor. In classifier based fault-detection schemes, along with accuracy other performance measures are also important parameter to judge classifier performance. Accordingly, accuracy and other performance parameters such as true positive (TP) rate, false positive (FP) rate, precision, recall and F-measure are evaluated. For demonstration of planned fault analysis techniques, experimental results are generated to build the method more realistic. The rest of the paper is organized as follows. The section II covers the system description and data acquisition. Section III discusses about the statistical time feature extraction. The background of proposed classifier is included in section IV. The results and discussions are provided in section V. Paper is concluded in section VI with brief remarks and inferences.

II. DATA ACQUISITION AND PREPROCESSING

In this paper three main faults which are normally occur in induction motor are investigated. The experimentation and data acquisition is performed on 2.2 kW, 415 V, 4- pole, 50 Hz custom designed induction machine. The machine is coupled with mechanical arrangement for loading i.e. adjusting weight on spring and belt. The schematic diagram of the scheme is given in Figure 1 and actual experimental set-up in shown in Figure 2. The diagnostic instrumentation system used is National Instruments data acquisition card model NI-6212. The sampling frequency of 16.896 kHz is adjusted. The instrumentation system is supported with Lab VIEW 2015 and Matlab R2014a is used for the processing and analysis. In the stator current, undesirable frequency components may present due to the load and supply conditions under normal running operation of motor. In order to eliminate these harmonic components that do not provide useful failure information, selective low pass filter is designed. The statistical tool box of Matlab is used to estimate the statistical time features of PVM which is obtained from filtered currents.

2.1. Stator Faults

To simulate inter turn fault, the star connected stator winding has been customized by addition of a number of tappings to the coils. The winding of the machine is having 180 turns per path per phase and has two parallel paths. The tappings are brought outside and connected to an external panel from which turns for creating inter turn faults can be accessed. Minor inter turn fault involving 1.39% (5 turns) of total turns in each phase of the winding is possible with the setup. Machine is tested under different conditions of minor inter-turn faults are furnished in Table I.

2.2. Rotor Faults

For demonstrating the rotor fault, three identical rotors having 28 bars are considered. Out of these rotors, one is considered as healthy and in the other two rotors damaged artificially by drilling the holes on the rotor bars. Three bar broken and five bar broken are considered for the rotor fault analysis. The actual pictures of faulty as well as healthy rotor are given in Figure 3.

2.3. Simultaneous Faults

For the analysis of simultaneous faults, rotor and stator faults are combined. The various mixed combinations are considered for the analysis.

The motor is run at no load and various loading conditions with the healthy faulty cases. The four classes have been considered as healthy (H), stator fault (SF), rotor fault (RF) and simultaneous fault (MF) for classification problem. The recorded observations have been arranged systematically for the classification purpose. The data size for each class is 200 x 10000 samples chosen.

Table I
Stator inter- turn fault cases

Sr. No.	*Inter-turn fault type*	*No. of turns shorted*	*% of stator winding turns per phase shorted*
1	Healthy	0	0
2	Fault1	5	1.39
3	Fault2	10	2.77
4	Fault3	15	4.16
5	Fault4	20	5.55
6	Fault5	25	6.94

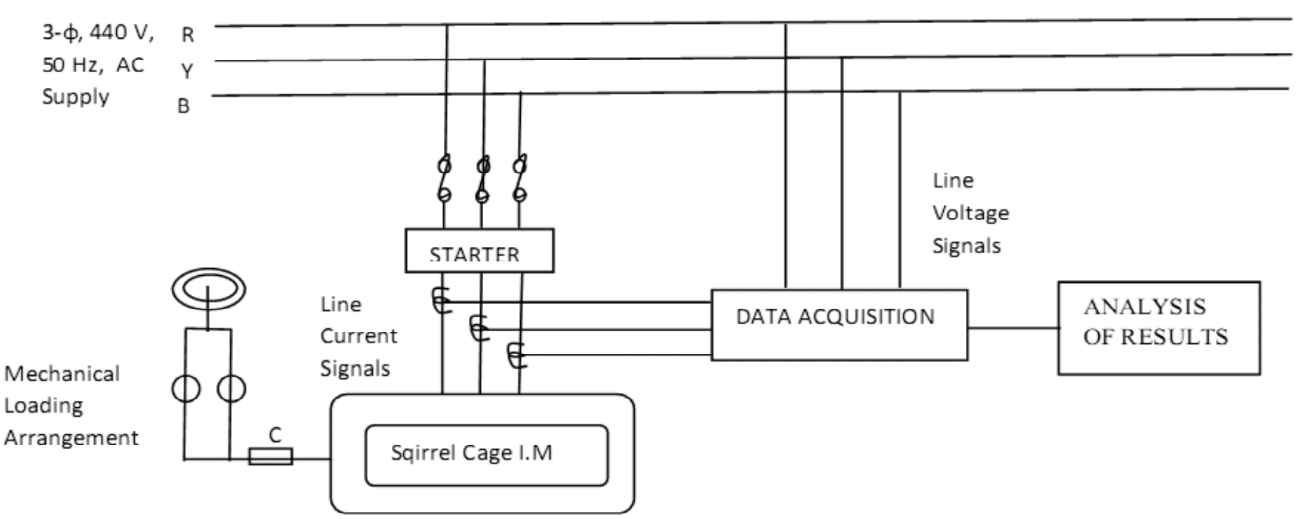

Figure 1: Block diagram of proposed scheme

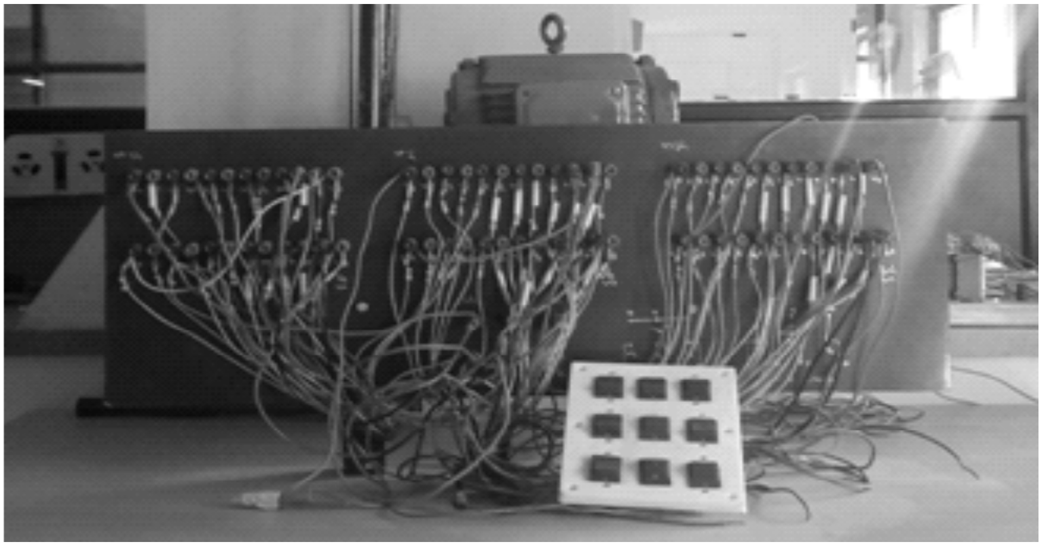

Figure 2: Actual experimental setup

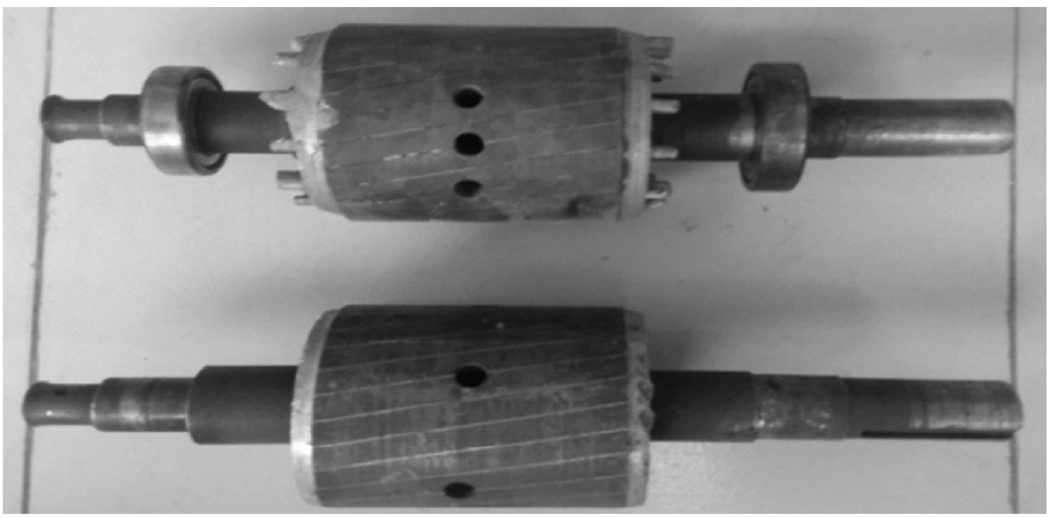

Figure 3: Healthy and faulty rotor

III. FEATURE EXTRACTION

In conventional fault diagnostic methods, electrical parameters such as voltage, currents and mechanical parameter such as vibration data are considered as features. During experimentation, no noticeable changes are observed in current waveforms of stator under healthy and various faulty condition as given in Fig. 4. There is necessity to extract some unique features from these currents. Accordingly, commonly used D-Q representation based on the Park's transformation is used to describe three-phase induction motor performance. The orthogonal Park's vector components (,) are determined using three phase line currents from (1) and (2), respectively [14-16].

$$i_d = \sqrt{\frac{2}{3}}\left[i_r \cos\theta + i_y \cos\left(\theta - \frac{2\pi}{3}\right) + i_b \cos\left(\theta + \frac{2\pi}{3}\right)\right] \tag{1}$$

$$i_q = \sqrt{\frac{2}{3}}\left[i_r \sin\theta + i_y \sin\left(\theta - \frac{2\pi}{3}\right) + i_b \sin\left(\theta + \frac{2\pi}{3}\right)\right] \tag{2}$$

Under ideal or healthy condition of motor and supply system, these components are given by,

$$i_d = \frac{\sqrt{6}}{2} i_f \sin(2\pi f_e t) \tag{3}$$

$$i_q = \frac{\sqrt{6}}{2} i_f \sin\left(2\pi f_e t - \frac{\pi}{2}\right) \tag{4}$$

Where, i_f = maximum value of the motor current

Further, these two current components has been reduce to single component known as Park's Vector Modulus (PVM) which is given by (5),

$$PVM = \sqrt{i_d^2 + i_q^2} \tag{5}$$

Further, set of common and simple statistical- time features are estimated from PVM. Fourteen features are suggested to each condition which is estimated from (6)-(19). The least number of time features to be analyzed covers the maximum, minimum, mean, median values, sum and standard deviation. The mean and median are estimated on individual dimension basis. The mean or variance, are utilized to describe the probability density function of a time-varying signal. The symmetry of distribution is measured by skewness. The higher order statistical time features such as kurtosis, which gives warning of the proportion of samples which diverge from the mean by a minute value compared with those which diverge by a big value. These features have the important property that they are not sensitive to Gaussian distributed noise.

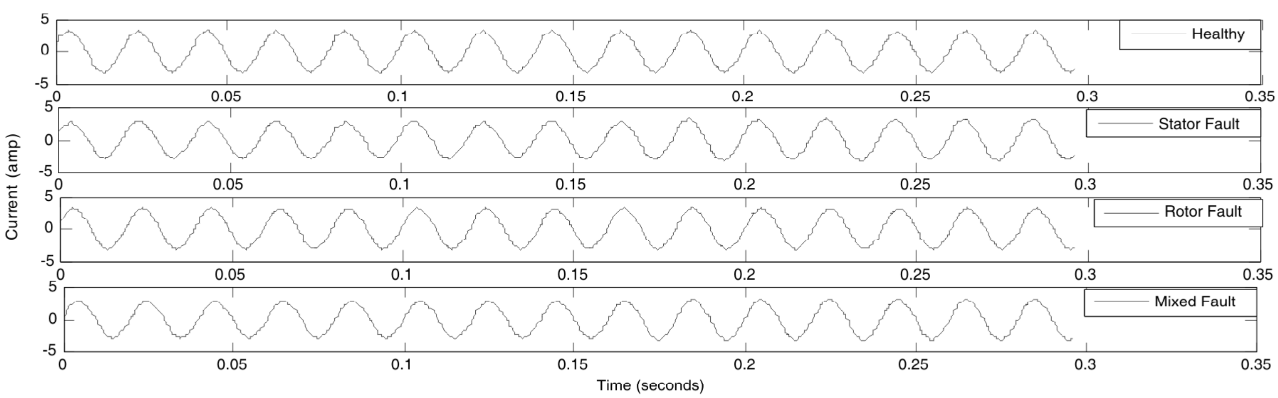

Figure 4: Current waveforms acquired at constant load condition: Healthy, stator fault, rotor fault and mixed fault

Minimum Value $\quad X_{min} = \min(x_i) \quad$ (6)

Maximum Value $\quad X_{max} = \max(x_i) \quad$ (7)

Mean $\quad \mu = \frac{1}{N}\sum_{i=1}^{N} x_i$

N= the total number of samples

Median
$$median = \left(\frac{(N+1)}{2}\right)^{th} value \tag{9}$$

Standard Deviation
$$S = \sqrt{\frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N-1}} \tag{10}$$

Variance
$$S^2 = \frac{\sum_{i=1}^{N}(x_i - \mu)^2}{N-1} \tag{11}$$

Sum
$$Sum = \sum_{i=1}^{N} x_i \tag{12}$$

Skewness
$$Skewness = \frac{1}{N}\frac{\sum_{i=1}^{N}(x_i - \mu)^3}{\sigma^3} \tag{13}$$

Kurtosis
$$Kurtosis = \frac{1}{N}\frac{\sum_{i=1}^{N}(x_i - \mu)^4}{\sigma^4} \tag{14}$$

Energy
$$Energy = \sum_{i=1}^{N} x_i^2 \tag{15}$$

R.M.S. value
$$x_{rms} = \frac{\sqrt{\sum_{i=1}^{N} x_i^2}}{N} \tag{16}$$

Absolute value of sum
$$Abs(Sum) = \left|\sum_{i=1}^{N} x_i\right| \tag{17}$$

Shape Factor
$$S.F. = \frac{x_{rms}}{Abs(Sum)} \tag{18}$$

Peak Factor
$$C.F. = \frac{x_{max}}{x_{rms}} \tag{19}$$

IV. RANDOM FOREST ALGORITHM [17]

RF algorithm is based on the grouping of trees for regression and classification. It is a combination of tree predictors such that each tree depends on the values of a random vector sampled independently and with the same distribution for all trees in the forest. The distinctive training set is completed by using bagging. The procedure bagging means extracts a fixed quantity from a training set randomly improve classification, and regression models according to stability and classification accuracy. This process decreases variance and avoids over-fitting. The absolute results makes by adding the scores of component predictors on every class and then choosing the successor class in terms of the number of scores to it. The error for forests converges to an optimized value as the number of trees in the forest becomes large. The error of classifiers also depends on the strength of the individual trees and the correlation between them.

V. RESULT AND DISCUSSION

As mentioned in earlier, fourteen statistical features are calculated for PVM. Then, these features are given as an input to the classifier. For testing and training the algorithm, Weka software (Ver. 3.8.0) is used as a tool [18]. The use of this data mining and machine learning software is very common in both academic researches and industrial studies [17-18]. The designed parameters used for algorithm are the same as default values determined by the software. The algorithm is tested on data sets of 200 numbers of observations to check the performance of classifier.

Initially small data is presented for training and large data is kept for testing. Gradually, training data size is increased and classifier performance is observed. In the next stage, the classifier is tested on cross validated (CV) data of 10 fold. The results of the proposed method are compared with conventional method. As the numbers of inputs to the classifier are reduced, the time to build the model will also reduce. It is observed from Table II that time to build the model using proposed classifier is reduced from 0.18 sec to 0.10 sec for cross validated data and 0.13 sec to 0.09 seconds for the standard training and testing ratio of 0.66. The accuracy is also enhanced from 97% to 98.5 %. The desired values of TP, FP, precision, recall and F-measure are also obtained as furnished Table III. The confusion matrix gives the information about the false and correct observations

Table II
Comparison of Performance of Direct Current Signatures and Proposed Method

	Performance using current signatures		*Performance using PVM*	
	Cross Validation Fold (10)	*Training and Testing Ratio (0.66)*	*Cross Validation Fold (10)*	*Training and Testing Ratio (0.66)*
Time taken to build model (seconds)	0.18	0.13	0.1	0.09
% Overall accuracy	97.5	97.0588	97	98.5294

Table III
Detailed Accuracy by class

Classified Output	*Performance using current signatures*								*Performance using PVM*							
	Cross Validation Fold (10)				*Training and Testing Ratio (0.66)*				*Cross Validation Fold (10)*				*Training and Testing Ratio (0.66)*			
	H	SF	RF	MF	H	SF	RF	MF	H	SF	RF	MF	H	SF	RF	MF
H	50	0	0	0	23	0	0	0	50	0	0	0	23	0	0	0
SF	0	49	0	1	0	10	0	1	2	48	0	0	0	11	0	0
RF	0	0	47	3	0	0	17	0	0	0	50	0	0	0	17	0
MF	0	0	1	49	0	0	1	16	1	0	3	46	0	0	1	16

represented in Table IV. Out of 200 observations false prediction rate is reduced to zero in case of proposed method as given in table IV. The results obtained from the proposed method gives more accurate results as compared to conventional method. Hence, the proposed method has been found more suitable for the fault detection and classification in induction machine.

Table IV
Confusion matrix

Parameter	*Performance using current signatures*								*Performance using PVM*							
	Cross Validation Fold (10)				*Training and Testing Ratio (0.66)*				*Cross Validation Fold (10)*				*Training and Testing Ratio (0.66)*			
	H	*SF*	*RF*	*MF*	*H*	*SF*	*RF*	*MF*	*H*	*SF*	*RF*	*MF*	*H*	*SF*	*RF*	*MF*
TP Rate	1.000	0.980	0.940	0.980	1.000	0.909	1.000	0.941	1.000	0.960	1.000	0.920	1.000	1.000	1.000	0.941
FP Rate	0.000	0.000	0.007	0.027	0.000	0.000	0.020	0.020	0.020	0.000	0.020	0.000	0.000	0.000	0.020	0.000
Precision	1.000	1.000	0.979	0.925	1.000	1.000	0.944	0.941	0.943	1.000	0.943	1.000	1.000	1.000	0.944	1.000
Recall	1.000	0.980	0.940	0.980	1.000	0.909	1.000	0.941	1.000	0.960	1.000	0.920	1.000	1.000	0.941	0.970
F-Measure	1.000	0.990	0.951	0.951	1.000	0.945	0.971	0.922	0.971	0.980	0.971	0.958	1.000	1.000	0.971	0.970

VI. CONCLUSION

This paper has presented innovative and accurate approach for individual and simultaneous fault in induction machine. After preprocessing the currents, using standard Park's transformation PVM is obtained. In the next stage, fourteen statistical measures of PVM are estimated using statistical tool box of MATLAB software. These features are fed as input to the random forest algorithm. From obtained results, it has been observed that as numbers of inputs to the classifier are reduced then time to build model will also reduce. The time to build the model is reduced from 0.18 seconds to 0.10 seconds for cross validated data and 0.13 seconds to 0.09 seconds for standard training and testing ratio of 0.66. The fault detection accuracy rate is also increased up to 98.50% hence various errors are drastically reduced to very small value almost 0% using proposed algorithm. The false prediction of the classifier is also reach to zero in the 200 observations of data set. The classifier accuracy is almost constant for cross validated data and testing and training data sets. The results obtained using proposed method is promising therefore; proposed method can be better choice for individual and simultaneous fault analysis and classification in induction machine.

ACKNOWLEDGMENT

The authors thankfully acknowledge the support provided by S.V.N.I.T., Surat for conducting the research work. Deepak M. Sonje also thanks KKWIIER, Nashik and RHSCOEMS&R, Nashik for their help in assisting while conducting the experiments. The author also thankful to SPPU,Pune, Maharashtra, India for their financial support toward the research project.

REFERENCES

[1] P. Zhang, Y. Du, T. G. Habetler, B. Lu, "A survey of condition monitoring and protection methods for medium-voltage induction motors," IEEE Trans. on Industry Applications, vol.47, no.1, pp. 34-46, 2011.

[2] A. Siddique, G. S. Yadava, B. Singh., "A review of stator fault monitoring techniques of induction motors ", IEEE Trans. on Energy Conversion, vol. 20, pp. 106-114, 2005.

[3] Deepak M. Sonje, A. Chowdhury, P. Kundu, " Park's Vector Approach for Fault Diagnosis of Induction Motor," Proc. of IEEE Conference on Advances in Electrical Engineering, ICAEE, pp.1-4, 2014

[4] A. Bellini, F. Filippetti, C. Tassoni, and G. A. Capolino, "Advances in diagnostic techniques for induction machines," IEEE Trans. on Industrial Electronics, vol. 55, pp. 4109–4126, Dec. 2008.

[5] W. T. Thomson and M. Fenger, “Current signature analysis to detect induction motor faults,” IEEE Industrial Application Magazine, vol. 7, pp. 26–34, Aug. 2001.

[6] M. E. H. Benbouzid, “A review of induction motors signature analysis as a medium for faults detection,” IEEE Trans. on Industrial Electronics, vol.47, 5, pp. 984–993, 2000.

[7] V. N. Ghate, S.V. Dadul, “Cascade neural-network-based fault classifier for three-phase induction motor”, IEEE Trans. on Industrial Electronics, vol. 58, pp. 1555–1563, Dec. 2011.

[8] C. T. Kowalski and T. Orlowska-Kowalska, “Neural networks application for induction motor faults diagnosis,” Math. Comput. Simul. vol. 63, no. 3–5, pp. 435–448, Nov. 2003.

[9] M. B. K. Bouzid, G. Champenois, N. M. Bellaaj, L. Signac, and K. Jelassi, “An effective neural approach for the automatic location of stator inter turn faults in induction motor,” IEEE Trans. on Industrial Electronics, vol. 55, no. 12, pp. 4277–4289, Dec. 2008.

[10] F. Filippetti, G. Franceschini, C. Tassoni, and P. Vas, “Recent developments of induction motor drives fault diagnosis using AI techniques,” IEEE Trans. on Industrial Electronics, vol. 47, no. 5, pp. 994–1004, Oct. 2000.

[11] B. Ayhan, M. Y. Chow, and M. H. Song, “Multiple discriminant analysis and neural-network-based monolith and partition fault-detection schemes for broken rotor bar in induction motors,” IEEE Trans. on Industrial Electronics, vol. 53, no. 4, pp. 1298–1308, Jun. 2006.

[12][13] V. N. Ghate, S.V. Dadul, “Induction machine fault detection using support vector machine based classifier”, WSEAS Trans. on Systems, vol. 8, pp. 591-603, May 2009.R. H. Cunha Palácios, I. N. da Silva, A. Goedtel, W. Godoy, “A Comprehensive Evaluation of Intelligent Classifiers for Fault Identification in Three-Phase Induction Motors”, Electric Power Systems Research, vol. 127, 249–258, 2015.

[14] S. Cruz, A. Cardoso. Stator winding fault diagnosis in three-phase synchronous and asynchronous motors by the Extended Park’s Vector Approach, IEEE Transactions on Industrial Applications, vol. 37, 1227-1233, 2001.

[15] Mansour Ojaghi1, Mahdi Sabouri, “Faiz, Diagnosis methods for stator winding faults in three-phase squirrel-cage induction motors”, International Transactions on Electrical Energy Systems, vol. 24, 891-912, 2014.

[16] D. M. Sonje, A. Chowdhury, P. Kundu, “Fault diagnosis of induction motor using Park’s Vector approach” , IEEE Conference Proceedings on Advances in Electrical Engineering, pp.1-4, 2014.

[17] Raj Kumar Patel, V.K. Giri, “Feature selection and classification of mechanical fault of an induction motor using random forest classifier”, Elsevier Engineering and Material Sciences, vol.8, 334-337, 2016.

[18] *http://www.cs.waikato.ac.nz/me/weka*