

# Scalable fault models for diagnosis in a synchronous generator using feature mapping and transformation techniques

R. Gopinath<sup>1</sup>, C. Santhosh Kumar<sup>2</sup>, and K. I. Ramachandran<sup>3</sup>

<sup>1,2,3</sup> *Machine Intelligence Research Lab., Department of Electronics and Communication Engineering,  
Amrita School of Engineering, Coimbatore,  
Amrita Vishwa Vidyapeetham, Amrita University, India-641112*

*rgopinath.gct@gmail.com*

*cs\_kumar@cb.amrita.edu*

*ki\_ram@cb.amrita.edu*

## ABSTRACT

Condition based maintenance (CBM) needs data acquired during healthy and faulty conditions to develop intelligent system for fault diagnosis. However, fault injection is not allowed/possible in a highly expensive components of complex/critical systems to collect fault condition data. Therefore, proto-type/small working models are used to conduct experiments for abnormal/fault conditions, to obtain and scale the intelligence of the system for effective health monitoring of complex system. This methodology is referred as scalable fault models. For proof of concept, in this work, we considered two different capacity synchronous generators with rating of 3 kVA and 5 kVA to emulate the behavior of proto-type/small working model and complex system respectively, for scalable fault models. We explored feature mapping and transformation techniques to achieve effective scalability.

From the preliminary experiments, it is observed that the baseline system performance deteriorated due to the changes in the system (capacity) and its characteristics with load changes. We therefore, expressed the input features in terms of load and system independent manner, to make the features less dependent on load and system variations. We explored locality constrained linear coding (LLC) to express the features load/system independently. It is observed that experimenting LLC with the backend support vector machine (SVM) classifier gave the best fault classification performance for linear kernel, suggesting that the faults are linearly separable in the new feature space.

Since the LLC mapped feature space is linearly separable, we then explored linear feature transformation technique, nuisance attribute projection (NAP) on the LLC mapped feature

space to further minimize the load/system specific variations. We observed that LLC-NAP improved the overall accuracy and sensitivity of the classifier significantly. We also noted that the performance of NAP was limited in the original feature space since the feature space (NAP without LLC) is non-linear with load/system variations.

## 1. INTRODUCTION

Over the past few years, condition based maintenance (CBM) strategy is preferred over preventive maintenance approach in most of the industries due to its reduced down time and maintenance cost, and increased reliability of the machines (Jardine, Lin, & Banjevic, 2006). Condition monitoring of a line replaceable unit (LRU) in aerospace applications is part of the integrated vehicle health management (IVHM) system (Felke, Hadden, Miller, & Mylaraswamy, 2010). In modern aircrafts, brushless synchronous machines are used mostly to generate electrical power for operation of aircraft subsystems. Maintenance of the flight critical components (eg., power generator) is performed based on time or usage. However, in this practice, the components are removed prematurely irrespective of its remaining useful life (RUL). Therefore, effective maintenance strategy needs to be carried out for reduced maintenance cost and high reliability of the systems (Batzel & Swanson, 2009). CBM's of various aerospace subsystems are reported towards effective implementation of IVHM (Tolani, Yasar, M. Shin, & Ray, 2005; Keller & Ray, 2001; Kirkland, Pombo, Nelson, & Ferghout, 2004).

To perform a data driven CBM of any engineering system, it requires the data during normal and fault conditions from the system to be monitored (Tavner, 2008; Nandi, Toliyat, & Li, 2005). However, in highly expensive complex systems, we cannot inject fault and collect data to learn the intelligence about the system. Further, deriving an accurate mathematical model of many of the complex systems will be extremely

Gopinath, R. et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

difficult if not impossible (Narasimhan, Roychoudhury, Balaban, & Saxena, 2010). To overcome this problem, prototype models were used, where fault injection is possible to learn the system intelligence/knowledge and scale the intelligence, to health monitoring of the complex (actual) system without conducting experiments for fault conditions (Oh et al., 2014).

In condition monitoring of wind turbines, a wind turbine simulator tool can be used to validate and verify the algorithms before implementing them on the actual wind turbine due to the practical limitations. Oh et al. proposed a wind turbine simulator with a small capacity of 20 kW that has similarities to 3 MW wind turbine for studying the characteristics of 3 MW wind turbine using a 20 kW wind turbine simulator (Oh et al., 2014). The experimental results reported that the acquired vibration and strain values from the simulator had similarities to the actual wind turbine. Oh et al. suggested that the data from the simulator could be used for developing an intelligent fault diagnosis system to monitor the actual wind turbine (Oh et al., 2014). However, no such study has been conducted from the simulator model to monitor the condition of actual wind turbine.

Godoy et al. experimented with two different specification of induction motors at various loading conditions to monitor the condition of machines (Godoy, da Silva, Goedel, & Palcios, 2015). The performance of the fault diagnosis system is compared using different classifiers, multilayer perceptron network, fuzzy adaptive resonance theory network and support vector machine (SVM). However, the fault models are generated in a machine dependent manner. Zhang et al. (Zhang, Ma, Lin, Ma, & Jia, 2015) developed a rotor dynamic model to simulate the various operating conditions of the aero-engines and acquired the data at various fault conditions, unbalance, misalignment and rub impact from the simulator. An improved empirical mode decomposition (EMD) and statistical parameters are used to extract the time-frequency features. Faults are diagnosed using an optimized fuzzy SVM based on genetic algorithm (Zhang et al., 2015). However, the proposed algorithms are tested for the simulated data rather than real aero-engine data.

Duque-Perez et al. proposed additive models to diagnose the faults in the rotor bars of induction motors when it was fed by different types of frequency inverters and control (Duque-Perez, Garcia-Escudero, Morinigo-Sotelo, Gardel, & Perez-Alonso, 2015). The study reported that same fault condition causes different fault signatures due to variations in power supply, motor and load conditions (Duque-Perez et al., 2015). However, the proposed approach has been experimented when the data from the different motors were combined together.

In our recent work, we used two different capacity three phase generators to consider as small working model/prototype and complex system, for proof of concept of scalable fault models (Gopinath, Kumar, Upendranath, & Kiran, 2016). Nuisance

attribute projection algorithm (NAP) was used to reduce system dependent attributes/features from input features. Decision tree classifier was experimented for fault classification. Though, NAP improved the system performance, it was noted that NAP requires linear space to be effective in reducing system attributes that are affecting the system performance. Therefore, the scope of the NAP was limited due to use of non-linear statistical features. It was suggested that expressing the input features in terms of load independent manner using locality constrained linear coding (LLC) helps improve the fault classification performance of load independent system dependent fault diagnosis (Gopinath, Kumar, Vishnuprasad, & Ramachandran, 2015).

In this paper, we therefore explore LLC to minimize the load and system specific variations in the input features from the prototype and complex system to make the features robust across the systems. We then explore NAP in the LLC mapped feature space to further minimize the load and system specific variations effectively. We also check the performance of the NAP in the original feature space (without LLC) to illustrate the effectiveness of experimenting NAP in the LLC mapped feature space. We use SVM as a backend classifier for fault diagnosis. Experimental setup and feature extraction process are described in Section 2. Methodologies and experiments are discussed in Section 3 and 4 respectively. Conclusions are made in Section 5.

## 2. EXPERIMENTAL SETUP AND FEATURE EXTRACTION

In this work, synchronous generators with capacity of 3 kVA and 5 kVA are customized to experiment stator winding short circuit faults by taking the leads out from the coils to the front panel as illustrated in Figure 1 and 2 (Gopinath et al., 2013). Figure 3 represents experimental facility. The taps are made in each coils at different points of the stator winding coils. In our experiments, we shorted only 6 turns (3.57%), 8 turns (4.76%) and 14 turns (8.33%) of 168 turns in each phase respectively. Each phase of the stator winding has 18 taps. The specifications of the 3 kVA and 5 kVA synchronous generator is listed in Appendix A.

Current signatures are acquired from synchronous generators during healthy and inter-turn fault conditions for various loading conditions from 0.5A to 3.5A with an increment of 0.5A load. Three phase resistive load bank is used for loading the generator at different loads. Current sensors are interfaced to the computer with the help of NI-PXI 6221<sup>1</sup>, DAQ. The block diagram of the experimental setup is shown in Figure 4. Short circuit fault conditions (inter-turn) are induced in the three phases of winding (stator) separately. The current signal acquired during the inter-turn short circuit condition for the 3 kVA and 5 kVA synchronous generator is shown in Figure 5. Current signals are collected for ten seconds of duration at

<sup>1</sup><http://sine.ni.com/ds/app/doc/p/id/ds-15/lang/en>

1kHz sampling rate for every experiments. A window size of 512 points is used to divide the signals into multiple frames. Further, statistical features are extracted in frequency domain from the input current signals (Lei, He, & Zi, 2008). Features used in this work are listed in Table 1.

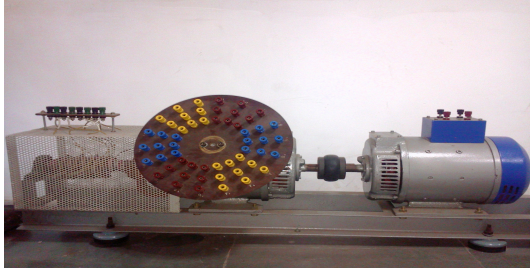


Figure 1. 3 kVA synchronous generator

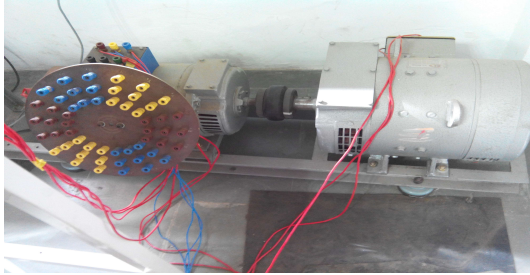


Figure 2. 5 kVA synchronous generator



Figure 3. Experimental facility

### 3. METHODOLOGY

#### 3.1. Locality constrained linear coding (LLC)

LLC is a feature mapping technique used to represent the non-linear features in a higher dimensional space where it can be linearly separable. Originally, LLC was introduced in image classification application for improving the performance of the classifier with the use of linear kernel (Wang et al., 2010). In this work, we explore LLC for minimizing load/system specific variations in the input features that affects fault classification performance.

LLC is an unsupervised learning method which represents the input feature vectors as a weighted sum of  $k$  nearest code-

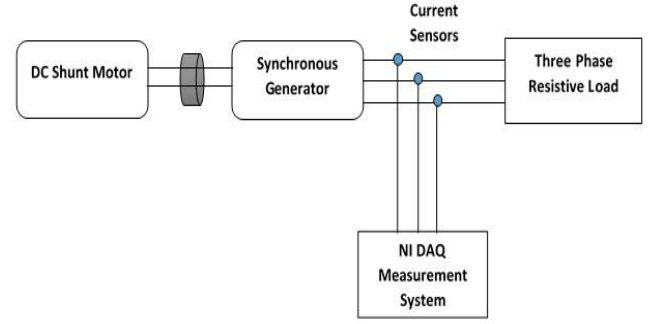


Figure 4. Block diagram of the experimental setup

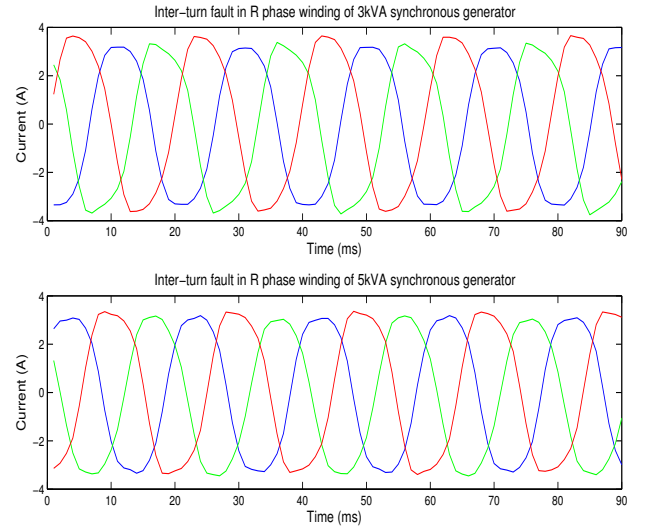


Figure 5. Current signals acquired during inter-turn fault in the 3 kVA and 5 kVA synchronous generators

books/basis vectors (Wang et al., 2010). The input features are represented using few elements (sparse) in a higher dimensional space. Codebooks are computed with the help of  $k$  means clustering algorithm from the training data set. Let  $X = [x_1, x_2, \dots, x_N] \in \mathbb{R}^{D \times N}$  be the input feature vector and given the codebook  $B = [b_1, b_2, \dots, b_M] \in \mathbb{R}^{D \times M}$ , every feature vector is mapped into a  $M$ - dimensional code. Basis vectors  $[b_1, b_2, \dots, b_M]$  identifies the inherent structures in the feature vectors. LLC can be represented as:

$$\min_C \sum_{i=1}^N \|X_i - \tilde{c}_i B_i\|^2 \quad (1)$$

$$s.t. 1^T \tilde{c}_i = 1, \forall i$$

| Description                      | Feature   |
|----------------------------------|---|
| Mean                             | $Feat_1 = \frac{\sum_{k=1}^K s(k)}{K}$  |
| Variance                         | $Feat_2 = \frac{\sum_{k=1}^K (s(k) - Feat_1)^2}{K-1}$                                       |
| Skewness                         | $Feat_3 = \frac{\sum_{k=1}^K (s(k) - Feat_1)^3}{K(\sqrt{Feat_2})^3}$                        |
| Kurtosis                         | $Feat_4 = \frac{\sum_{k=1}^K (s(k) - Feat_1)^4}{K(Feat_2)^2}$                               |
| Frequency centre                 | $Feat_5 = \frac{\sum_{k=1}^K f_k s(k)}{\sum_{k=1}^K s(k)}$                                  |
| Standard deviation               | $Feat_6 = \sqrt{\frac{\sum_{k=1}^K (f_k - Feat_5)^2 s(k)}{K}}$                              |
| Root Mean Square frequency       | $Feat_7 = \sqrt{\frac{\sum_{k=1}^K f_k^2 s(k)}{\sum_{k=1}^K s(k)}}$                         |
| Spectrum power convergence       | $Feat_8 = \sqrt{\frac{\sum_{k=1}^K f_k^4 s(k)}{\sum_{k=1}^K f_k^2 s(k)}}$                   |
| Stability factor                 | $Feat_9 = \frac{\sum_{k=1}^K f_k^2 s(k)}{\sqrt{\sum_{k=1}^K s(k) \sum_{k=1}^K f_k^4 s(k)}}$ |
| Coefficient variability          | $Feat_{10} = \frac{Feat_6}{Feat_5}$   |
| Skewness frequency               | $Feat_{11} = \frac{\sum_{k=1}^K (f_k - Feat_5)^3 s(k)}{K(Feat_6)^3}$                        |
| Kurtosis frequency frequency     | $Feat_{12} = \frac{\sum_{k=1}^K (f_k - Feat_5)^4 s(k)}{K(Feat_6)^4}$                        |
| Spectrum power positional factor | $Feat_{13} = \frac{\sum_{k=1}^K (f_k - Feat_5)^{\frac{1}{2}} s(k)}{K\sqrt{Feat_6}}$         |

where  $s(k)$  is the spectrum for  $k = 1, 2, \dots, K$ ,  $K$  is the number of spectrum lines  $f_k$  is the frequency value of the  $k^{th}$  spectrum.

Table 1. Frequency Domain Features

where  $\tilde{C} = [\tilde{c}_1, \tilde{c}_2, \tilde{c}_3, \dots, \tilde{c}_N]$  is the set of coding coefficients for input feature vector  $X$ . The shift invariant requirement of LLC is obtained using the constraint  $1^T \tilde{c}_i = 1$ . The pictorial representation of LLC is shown in Figure 6. It illustrates the selection of nearest basis vectors for representing the input feature vector. From Figure 6, it is also noted that the basis vectors are shared for the similar input feature vectors. Thus, LLC ensures similar input vectors will have similar codes. Since LLC prefers locality, coding coefficients will have smaller value for the basis vectors far away from the local input feature vectors.

In this work, we experimented LLC to improve the fault classification performance of scalable fault models. To make the fault model effective, the feature vectors should be robust to load/system specific variations. For this, we derived the basis vectors in a load/system independent manner and expressed the feature vectors using these basis vectors. The load/system independent basis vectors or codebooks are computed using  $k$  means clustering algorithm from the input training feature vectors. The computed codebooks are used to express the input training and test feature vectors load/system independently. LLC selects the  $k$  nearest load/system independent basis vectors to represent the input features. The new training and test features are used for generating and test the model

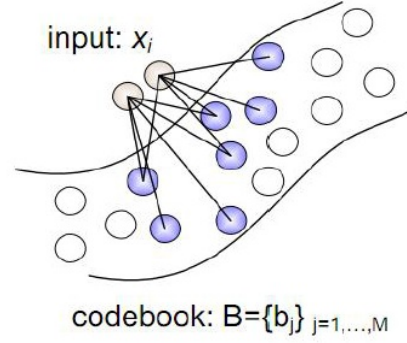


Figure 6. LLC

using SVM classifier.

### 3.2. Nuisance attribute projection (NAP)

Originally, NAP was proposed for reducing the channel effects to improve the robustness in speaker recognition system (Solomonoff, Campbell, & Quillen, 2007). The same person speaking through different channels (carbon button handset and mobile phone) could be identified as different speakers. The effect of the channels are considered as the nuisance which affects the classification performance of the system (Solomonoff et al., 2007). NAP uses eigenvalue analysis to remove the nuisance attributes.

In this work, we consider that the fault is expressing itself through proto-type (3 kVA generator) and actual system (5 kVA generator). Though the nature of fault and its effect on the generators are same, it is noted that there is a significant difference in the signatures from generators. We attribute this difference in the signal characteristics to the differences in generator characteristics (proto-type/actual system) through which the fault is expressing itself. In this paper, capacity of the system is considered to be a attribute that affects the fault diagnosis system performance. We experiment NAP to reduce system attributes from feature space (input features) to develop system to be robust in diagnosing faults.

Though, there were various approaches to remove the nuisance attribute from the input feature space, projection approach is most preferred due to its easiness. Projection is a feature space transformation that helps to remove the vector components in the specified subspace. NAP algorithm is illustrated pictorially in Figure 7. From this figure, we interpret that the input feature vector with system information (capacity)  $m^{(s,c)}$  is projected on to the subspace which contains the system information  $v$  to obtain the system independent feature vector  $m^{(s)}$ . Here, we learn the dimensions that corresponds to system capacity and project out those dimensions from the input feature vector to make the feature vectors less dependent on system capacity.

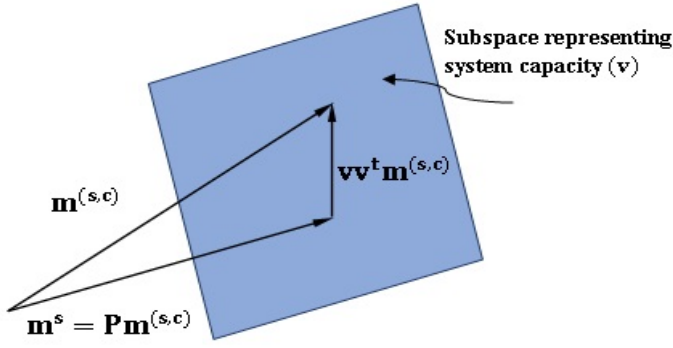


Figure 7. Nuisance attribute projection (NAP)

Projection  $P$  can be written as:

$$P = I - vv^t \quad (2)$$

where  $v$  contains the information about the system attributes and  $I$  denotes the identity matrix. The input feature vectors of training data set  $\{x_i\}_{i=0}^n$  are labeled with the nuisance attribute. A symmetric weight matrix is generated for training feature vectors by applying the criteria:

$$W_{ij} = \begin{cases} 1; & \text{system}(x_i) \neq \text{system}(x_j) \\ 0; & \text{system}(x_i) = \text{system}(x_j) \end{cases} \quad (3)$$

Here, we move together closer for the data from the same capacity generator and move away the data from different capacity generator. The figure of merit (FOM)  $\delta$  is minimized over  $P$  to improve robustness of feature vectors between the different capacity generators. FOM is expressed as (Solomonoff et al., 2007):

$$\delta = \sum_{ij} W_{ij} \|P(x_i - x_j)\|^2 \quad (4)$$

Substituting,  $P = I - vv^t$ , in Eq. (4):

$$\delta = \sum_{ij} W_{ij} \|(I - vv^t)(x_i - x_j)\|^2 \quad (5)$$

Eq. (5), is reduced by performing algebra operations:

$$\delta = \sum_{ij} W_{ij} \left[ \|(x_i - x_j)\|^2 - (v^t(x_i - x_j))^2 \right] \quad (6)$$

Since, the first term in Eq. (6), does not have system information  $v$ , it can be ignored:

$$\delta = - \sum_{ij} W_{ij} (v^t(x_i - x_j))^2 \quad (7)$$

Eq. (7), can be simplified and expressed as:

$$\delta = -2 \sum_i (\sum_j W_{ij}) v^t x_i x_i^t v + 2 \sum_{ij} W_{ij} v^t x_i x_j^t v \quad (8)$$

Further, Eq. (8), can be expressed in terms of  $A = [x_1, x_2, \dots, x_n]$ :

$$\delta' = -2v^t A \text{diag}(W.1) A^t v + 2v^t A W A^t v \quad (9)$$

$$= 2v^t A (W - \text{diag}(W.1)) A^t v \quad (10)$$

The FOM  $\delta'$  should be minimized, subject to  $\|v\| = 1$ , which is equal to identifying the smallest eigenvalue of the symmetric eigenvalue analysis.

$$A(W - \text{diag}(W.1)) A^t v = \lambda v \quad (11)$$

Eq. (11), can be further simplified as:

$$Az(W) A^t v = \lambda v \quad (12)$$

Eq. (12), is called as NAP equation. Finally, the objective function  $\delta(P)'$  can be minimized by keeping the columns of  $v$  to be  $k$  most principal eigenvectors of the eigenvalue analysis, that has the information about the system. Therefore, the system dependent features can be removed by projecting it out from the input space (features).

#### 4. EXPERIMENTS AND RESULTS

In this work, we developed scalable fault models for fault diagnosis of the synchronous generator using two different capacity, 3 kVA and 5 kVA synchronous generators. We model the faults using the 3 kVA generator (no-fault and fault data) and no-fault data from the 5 kVA generator, and scale this model to diagnose faults in the 5 kVA generator. The 5 kVA fault condition data is used for testing purpose only and not while building the model. In this work, first we develop the baseline system for scalable fault models using SVM classifier. Then, we express the feature vectors in terms of load and system independent manner to make the feature vectors less dependent on load/system variations for effective scalability. Dataset used in the experiments is listed in Table 2.

We measure the performance of the system by calculating the overall accuracy, sensitivity (alarm accuracy) and specificity

(no-alarm accuracy). Sensitivity and specificity are used to analyze the performance of a binary classifier. Sensitivity calculates the proportion of actual faults (positives) which are correctly classified as fault, whereas specificity calculates the proportion of actual no-faults (negatives) which are correctly classified as no-fault. Overall accuracy measures the ability to classify the no-fault and fault conditions correctly (Zhu, Zeng, & Wang, 2010).

In this work, we consider systems used in mission critical applications that require high sensitivity for fault detection. In such systems, every possibility of an alarm (fault) needs to be detected, and the alarm detection accuracy has to be as high as possible. Short circuit faults (inter-turn) are experimented in all the phases of the stator winding separately and acquired the output current signals. In this work, experiments are performed as a two class classification problem. i.e., R phase fault vs no-fault, Y phase fault vs no-fault, B phase fault vs no-fault.

| Data     | 3 kVA Generator |       | 5 kVA Generator |       |
|----------|-----------------|-------|-----------------|-------|
|          | Healthy         | Fault | Healthy         | Fault |
| Training | 13300           | 10640 | 13300           | -     |
| Testing  | -               | -     | 5852            | 4683  |

Note: The numbers in table are the number of data points used for training and testing.

Table 2. Data set

#### 4.1. Scalable fault models using SVM: Baseline system

For the baseline system, we first combined the data acquired from the different loads under normal and fault operating conditions. Three phase current signals are collected at 0.5 A, 1 A, 1.5 A, 2 A, 2.5 A, 3 A and 3.5 A resistive loads. Inter-turn faults are injected in the R, Y and B phases separately and acquired the current signals respectively. Subsequently, the acquired time domain signals are transformed into frequency domain using Fast Fourier Transform (FFT). We then extracted the frequency domain statistical features (13 x 3 phase= 39 features) from the current signal acquired in the 3 kVA and 5 kVA generators.

The features from 3kVA generator (no-fault and fault condition) and 5kVA generator (no-fault condition) are used for training the model using SVM classifier. The features from 5 kVA generator (no-fault and fault conditions) are used as a test data set (See Table 2) for fault classification using SVM. No-fault and fault conditions are labeled as two separate classes. The block diagram for scalable fault models baseline system is shown in Figure 8. The experiments with R, Y and B phase inter-turn faults are treated as a independent two class classification problems. It may be noted that a  $N$  class problem may be realized as  $N$  two class problems. For its simplicity, we chose to experiment with the two class problem, no-fault or fault, with the R or Y or B phase of the gener-

ator. For each experiment, overall accuracy, sensitivity and specificity of the baseline system are calculated from SVM testing. Linear kernel is used for fault classification. The performance of the baseline system for scalable fault models is listed in Table 3. The baseline system has the overall accuracy of 53.79%, 52.09% and 52.91% for the R, Y and B phase faults respectively. From experiments and results, it is observed that the baseline system performance deteriorated due to the changes in the system (capacity) and its characteristics with load changes. Therefore, the effective fault discrimination has not been achieved due to the load and system specific variations in the input feature vectors. For effective scalability, the input features should be robust across the 3 kVA and 5 kVA generators. We therefore express the feature vectors from the 3 kVA and 5 kVA generators in terms of load/system independent basis vectors to make the features less dependent on load/system variations.

| Fault | Overall Accuracy (%) | Sensitivity (%) | Specificity (%) |
|-------|----------------------|-----------------|-----------------|
| R     | 53.79                | 31.86           | 71.34           |
| Y     | 52.09                | 63.63           | 42.86           |
| B     | 52.91                | 56.46           | 50.09           |

Table 3. Baseline system performance of scalable fault models using linear SVM

#### 4.2. Scalable fault models using LLC

From the baseline system performance of scalable fault models, it is observed that the fault classification accuracy deteriorated due to the changes in the system (capacity) and its characteristics with load changes. We therefore experimented with LLC to minimize the load/system specific variations in the feature vectors for effective scalability. The block diagram of LLC-SVM system is shown in Figure 9.

In this experiment, input features from the training data (39 features) are used to compute load/system independent basis vectors (codebooks). Subsequently, the computed codebook is used for the representation of features from the training and test datasets. The optimum size of the codebook is obtained empirically for the best classification performance. The experiments are carried out for different  $k$  nearest neighbors/basis vectors to represent the feature vectors effectively in a load and system independent manner. The new feature vectors are used for generating and test the fault model. The block diagram of LLC for the fault diagnosis of synchronous generator is shown in Figure 10.

The performance of LLC for scalable fault models is listed in Table 4. The best classification accuracy is obtained for 1024 codebook size with an overall accuracy of 71.60 %, 78.27% and 77.65% for the R, Y and B phase faults. From experiments and results, we observed that experimenting LLC with backend SVM classifier gave the best fault classification per-

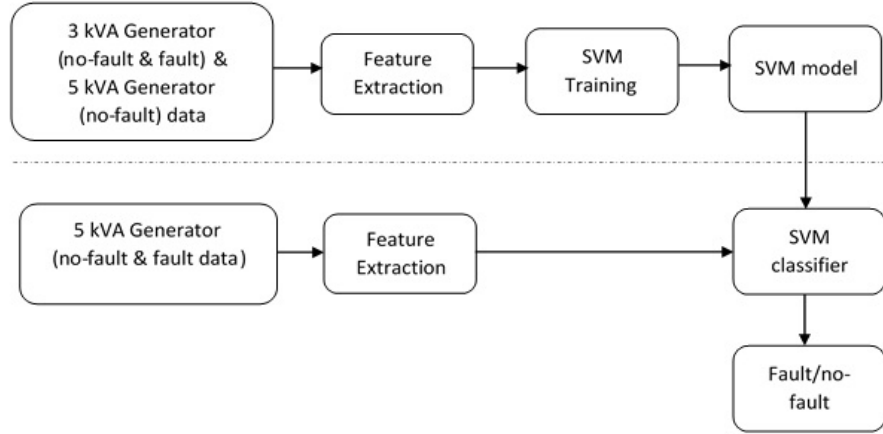


Figure 8. Baseline system for scalable fault models of synchronous generator



Figure 9. Scalable fault models using LLC-SVM

formance for the linear kernel, suggesting that the faults are linearly separable in the new feature space. We note that LLC improves the overall classification accuracy and specificity significantly when compared to the baseline system. However, better sensitivity is desired in mission critical applications.

| Fault | $k$ -NN | Overall (%)<br>Accuracy | Sensitivity<br>(%) | Specificity<br>(%) |
|-------|---------|-------------------------|--------------------|--------------------|
| R     | 1       | 71.60                   | 36.58              | 99.62              |
| Y     | 2       | 78.27                   | 52.65              | 98.77              |
| B     | 1       | 77.65                   | 50.11              | 99.69              |

Table 4. Performance of scalable fault models using LLC (1024 Codebooks) for linear kernel

From the experiments, we observed that LLC helps minimize the load/system specific variations significantly by expressing the feature vectors in a load/system independent manner. However, sensitivity of the classifier needs to be improved further for effective scalability. Since the LLC mapped feature space is linearly separable, we then explored linear feature transformation technique, NAP on the LLC mapped feature space to further minimize the load/system specific variations. The experiments with NAP are discussed in the following section.

### 4.3. Scalable fault models using LLC-NAP

In this experiment, we explored supervised feature transformation technique, NAP to identify and remove the load/system specific variations in the LLC mapped feature space effec-

tively. The block diagram of LLC-NAP system is shown in Figure 11. The detailed block diagram of NAP is shown in Figure 12. We first computed the load/system independent basis vectors (codebook) from the input training feature vectors to express the features in terms of load/system independent basis vectors. The codebook and  $k$  nearest neighbors are chosen empirically by experimenting with different sizes of codebooks and its nearest neighbors. It may be noted that LLC makes the features less dependent on load/system variations. From the previous experiments with LLC (Section 4.2), we observed that LLC gave best performance for the linear kernel, suggesting that the faults are linearly separable in the new feature space. Since the LLC mapped feature space is linearly separable, we then explored NAP to further remove the load/system attributes in the LLC mapped feature space.

NAP (Solomonoff et al., 2007) has been used widely in speaker recognition systems to remove the effects of the nuisance attributes. NAP filters out the effect of the nuisance attributes using eigenvalue analysis. Similarly, the performance of the fault diagnosis system also depends on the system dependent factors such as, power supply, motor and loads (Duque-Perez et al., 2015). In this work, we explored NAP in minimizing system dependent factors in the feature vectors in linear space for effective scalability.

In NAP, initially, the input LLC mapped training feature vectors from the two different capacity synchronous generators (3 kVA & 5 kVA) are labeled as two different classes. Then,  $W_{ij}$  (weight matrix) is computed using Eq. 3 to develop the model in a system independent manner. Now we minimize figure of merit (FOM)  $\delta'(P)$  (Eq. 4) to remove the system dependent factors and make the features robust across the two different capacity generators. This objective can be obtained by selecting  $k$  most principal eigenvectors of the eigenvalue of  $v$ , where  $v$  is a matrix with orthonormal columns representing the system information. The system attributes ( $k$

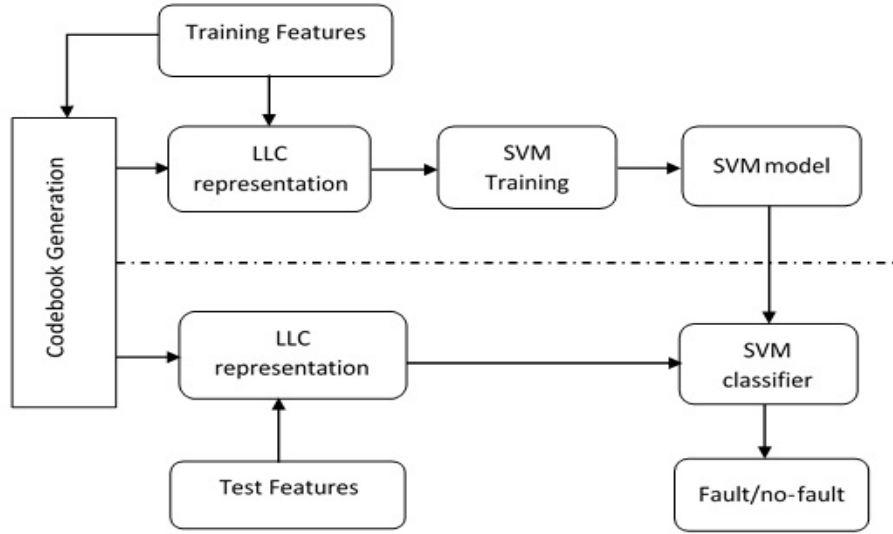


Figure 10. LLC-SVM methodology

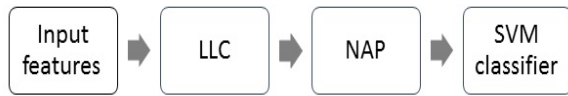


Figure 11. Scalable fault models using LLC-NAP

most principal eigenvectors) are identified and removed from the LLC feature mapped training feature vectors using Eq. 2. The optimum principal eigenvectors are obtained empirically for the best fault classification performance. Finally, projection matrix is multiplied with LLC mapped training and test feature vectors to make the features robust to the load/system specific variations. The NAP transformed training and test feature vectors are used for generating and test the fault model. SVM backend classifier is used for fault classification.

The fault classification performance of LLC-NAP for the scalable fault models is listed in Table 5. The best performance is achieved for the codebook size of 512 using LLC. The optimum principal eigenvectors are selected empirically using NAP. It is noted that LLC-NAP improves the overall accuracy and sensitivity by 25.94%, 26.09% and 29.85%, and 40.31%, 8.31%, and 14.05% for R, Y and B phase faults respectively when compared to baseline system. LLC-NAP has the overall accuracy of 79.73%, 78.18% and 82.76% for R, Y and B phase faults respectively.

It is observed that NAP minimizes the effect of load/system variations significantly in the linear space. We also compare the performance of NAP with LLC (LLC-NAP-SVM), and NAP without LLC (NAP-SVM). In the following section, we discuss the fault classification performance of NAP without

LLC.

| Fault | $k$ -NN | Eigen Vector | Overall Accuracy | Sensitivity | Specificity |
|-------|---------|--------------|------------------|-------------|-------------|
| R     | 1       | 1:285        | 79.73            | 72.17       | 85.78       |
| Y     | 1       | 1:238        | 78.18            | 71.94       | 83.18       |
| B     | 1       | 1:268        | 82.76            | 70.51       | 92.56       |

Table 5. Performance of scalable fault models using LLC-NAP for linear kernel (512 Codebooks)

#### 4.3.1. Scalable fault models using NAP without LLC (NAP-SVM)

In this experiment, we explore NAP in the input (original) feature space to check its effectiveness in minimizing the load and system specific variations. The input frequency domain feature vectors are used for fault modeling. The block diagram of NAP-SVM system is shown in Figure 13.

As discussed in previous section, we compute weight matrix  $W_{ij}$ , to model the system in a system independent feature space using NAP. The figure of merit (FOM) over projection,  $\delta'(P)$  (Eq. 4) is minimized to make the features robust across the 3 kVA and 5 kVA generator. The system dependent information is identified and projected out by finding the  $k$  most principal eigenvectors of eigenvalue analysis. We project out these system-dependent features to make the features system-independent. The projected vectors are multiplied with training and test data sets to make the training and test feature vectors, system-independent. The new transformed training and test data sets are used to generate and test the fault model. We used SVM backend classifier for fault classification. The performance of NAP-SVM for the scalable fault model is listed in Table 6.



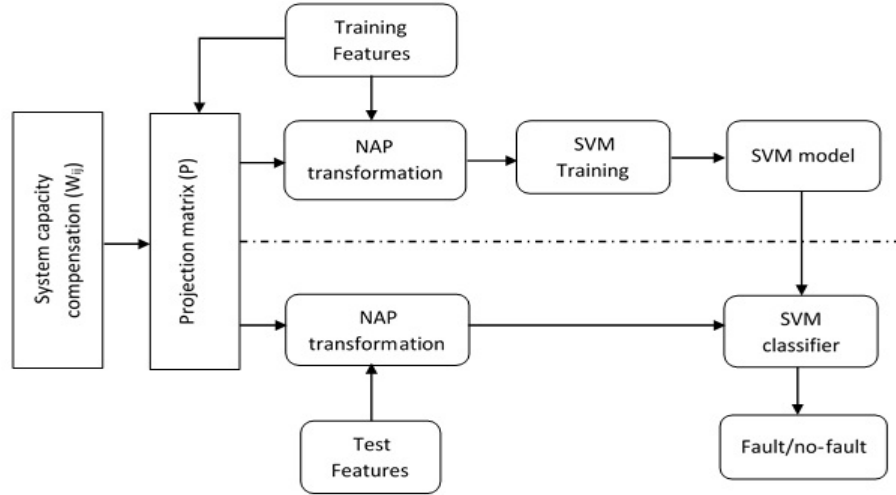


Figure 12. NAP-SVM methodology



Figure 13. Scalable fault models using NAP-SVM

| Fault | Eigen vector | Overall Accuracy | Sensitivity | Specificity |
|-------|--------------|------------------|-------------|-------------|
| R     | 1            | 57.99            | 18.55       | 89.55       |
| Y     | 1            | 75.03            | 44.78       | 99.23       |
| B     | 1            | 70.61            | 43.18       | 92.57       |

Table 6. Fault classification performance of NAP without LLC (NAP-SVM) for linear kernel

From experiments and results, we observed that NAP did not improve the fault classification performance significantly when compared to the LLC-NAP system (Table 5). It is noted that the performance of NAP is limited in the input/original feature space (without LLC) since the feature space is non-linear with load/system variations. The performance of the baseline SVM, LLC-SVM, LLC-NAP and NAP-SVM (NAP without LLC) systems are compared using receiver operating characteristics (ROC) curve. The ROC curves are shown in Figure 14-16. The area under the curve (AUC) for LLC-NAP was improved by 0.3, 0.25, and 0.29 for fault conditions (R, Y, and B phase) when compared to baseline system. From the experiments, we observed that NAP helps to improve the performance of scalable fault models effectively in the LLC mapped feature space.

## 5. CONCLUSIONS

In this work, we explored the concept of scalable fault models in a laboratory setup. We used synchronous generators with two different capacities as a case study. From the preliminary

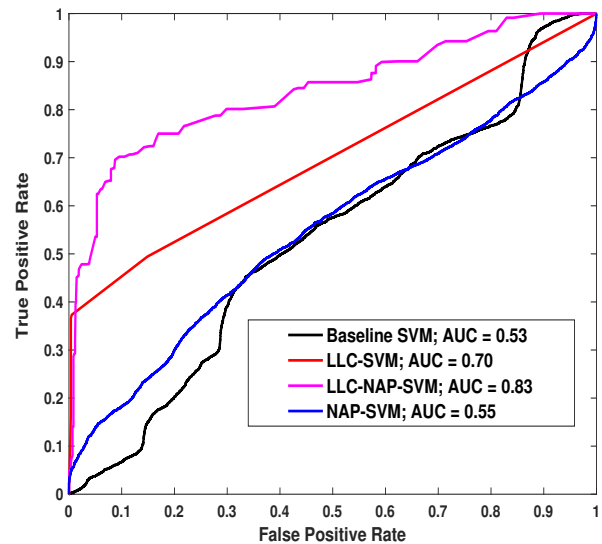


Figure 14. Performance comparison of scalable fault models for R phase fault

experiment, we noted that the baseline system performance deteriorated due to the changes in the system (capacity) and its characteristics with load changes. We explored feature mapping and feature transformation techniques to improve the fault classification performance of synchronous generators. We experimented with LLC to express the input features in terms of load and system independent manner. It was seen that the LLC with the backend SVM classifier gave the best fault classification performance for the linear kernel, suggesting that the faults are linearly separable in the new feature space. As the LLC mapped feature space is linearly separable, we then explored linear feature transforma-

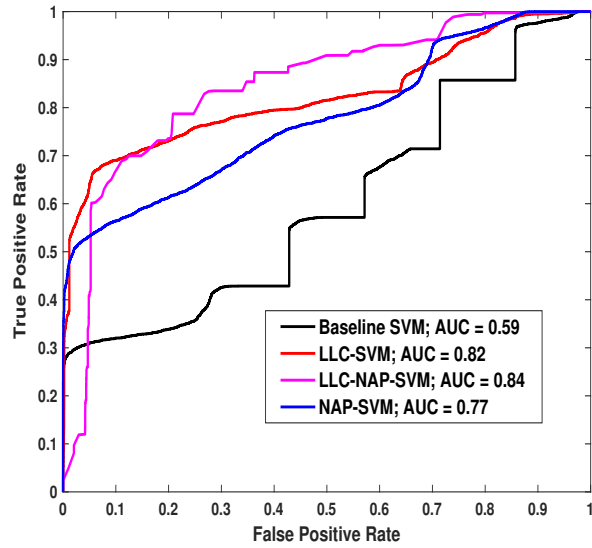


Figure 15. Performance comparison of scalable fault models for Y phase fault

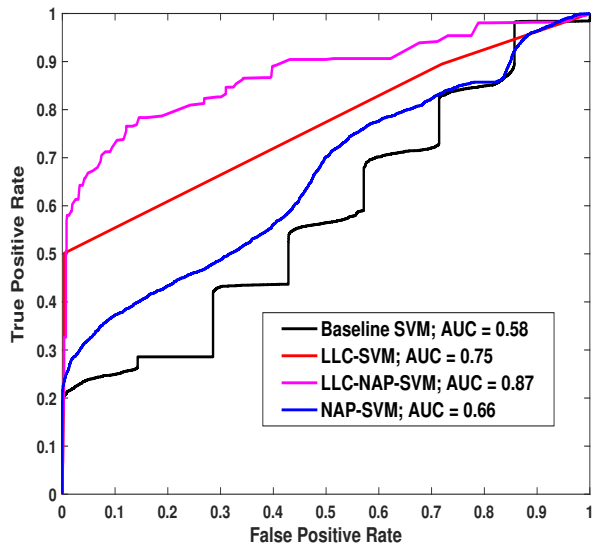


Figure 16. Performance comparison of scalable fault models for B phase fault

tion technique, nuisance attribute projection (NAP) on the LLC mapped feature space to further minimize the load and system specific variations. We observed that LLC-NAP improved the fault classification performance significantly. Further, we also illustrated the effectiveness of NAP in the original feature space. We noted that the performance of the NAP was limited in the original feature space since the feature space (NAP without LLC) is nonlinear with load/system variations. The performance of the classifier was evaluated using receiver operating characteristics (ROC) curve. We ob-

served that LLC-NAP outperforms the other systems for scalable fault models. The authors also suggest that the proposed LLC-NAP approach could be used in various applications, where the classification performance of the system gets affected due to the presence of nuisance variations in the data.

#### ACKNOWLEDGMENTS

The authors are thankful to Aeronautical Development Agency (ADA), India for providing financial support (Research Grant No. PARC #3) for the experimental facility. M/S. Bharath Electricals, Coimbatore, India helped us in manufacturing generators with our specifications at low cost.

#### APPENDIX A

| Parameter        | Generator 1 | Generator 2 |
|------------------|-------------|-------------|
| Rated power      | 3 kVA       | 5 kVA       |
| Rated voltage    | 415 V       | 415 V       |
| Rated frequency  | 50 Hz       | 50 Hz       |
| Connection type  | star        | star        |
| Number of poles  | 4           | 4           |
| Number of phases | 3           | 3           |
| Speed            | 1500 rpm    | 1500 rpm    |
| Current          | 4.2 A       | 6.5 A       |
| Power factor     | 0.8         | 0.8         |

Synchronous generator specifications

#### REFERENCES

- Batzel, T. D., & Swanson, D. C. (2009). Prognostic health management of aircraft power generators. *IEEE Transactions on Aerospace and Electronic Systems*, 45, 473-483.
- Duque-Perez, O., Garcia-Escudero, L. A., Morinigo-Sotelo, D., Gardel, P. E., & Perez-Alonso, M. (2015). Analysis of fault signatures for the diagnosis of induction motors fed by voltage source inverters using anova and additive models. *Electric Power Systems Research*, 121, 1-13.
- Felke, T., Hadden, G., Miller, D., & Mylaraswamy, D. (2010). Architectures for integrated vehicle health management. In *Aiaa infotech@ aerospace 2010* (p. p.3433).
- Godoy, W. F., da Silva, I. N., Goedel, A., & Palcios, R. H. C. (2015). Evaluation of stator winding faults severity in inverter-fed induction motors. *Applied Soft Computing*, 32, 420-431.
- Gopinath, R., Kumar, C. S., Vishnuprasad, K., & Ramachandran, K. I. (2015). Feature mapping techniques for improving the performance of fault diagnosis of synchronous generator. *International Journal of Prognostics and Health Management*, 6, p. 12.
- Gopinath, R., Kumar, K. I., C. S. Ramachandran, Upen-dranath, V., & Kiran, P. S. (2016). Intelligent fault

- diagnosis of synchronous generators. *Expert Systems with Applications*, 45, 142-149.
- Gopinath, R., Nambiar, T. N. P., Abhishek, S., Pramodh, S., Pushparajan, M., Ramachandran, K. I., ... Thirugnanam, R. (2013). Fault injection capable synchronous generator for condition based maintenance. In *7th international conference on intelligent systems and control (isco)* (p. 60-64).
- Jardine, A. K., Lin, D., & Banjevic, D. (2006). A review on machinery diagnostics and prognostics implementing condition based maintenance. *Mechanical systems and signal processing*, 1483-1510.
- Keller, E., & Ray, A. (2001). Real-time nondestructive evaluation of airframe structures for health monitoring and residual life prediction. In *Proceedings of the 20th digital avionics systems conference* (Vol. 1: 3D5/13D5/1, p. 2).
- Kirkland, L. V., Pombo, T., Nelson, K., & Ferghout, F. (2004). Avionics health management: searching for the prognostics grail. In *Proceedings of the IEEE aerospace conference* (Vol. 5, p. 3448-3454).
- Lei, Y., He, Z., & Zi, Y. (2008). A new approach to intelligent fault diagnosis of rotating machinery. *Expert Systems with Applications*, 1593-1600.
- Nandi, S., Toliyat, H. A., & Li, X. (2005). Condition monitoring and fault diagnosis of electric motors- a review. *IEEE Transactions on Energy Conversion*, 719-729.
- Narasimhan, S., Roychoudhury, I., Balaban, E., & Saxena, A. (2010). Combining model-based and feature-driven diagnosis approaches a case study on electromechanical actuators. *21st International Workshop on Principles of Diagnosis*, 1-8.
- Oh, K. Y., Lee, J. K., Bang, H. J., Park, J. Y., Lee, J. S., & Epureanu, B. I. (2014). Development of a 20 kw wind turbine simulator with similarities to a 3 mw wind turbine. *Renewable Energy*, 62, 379-387.
- Solomonoff, A., Campbell, W. M., & Quillen, C. (2007). Nuisance attribute projection. *Speech Communication*, 1-73.
- Tavner, P. J. (2008). Review of condition monitoring of rotating electrical machines. *IET, Electric Power Applications*, 215-247.
- Tolani, D., Yasar, M. Shin, C., & Ray, A. (2005). Anomaly detection for health management of aircraft gas turbine engines. In *Proceedings of the 2005 american control conference* (Vol. 1, p. 459-464).
- Wang, J., Yang, J., Yu, K., Lv, F., Huang, T., & Gong, Y. (2010). Locality-constrained linear coding for image classification. *2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3360-3367.
- Zhang, J., Ma, W., Lin, J., Ma, L., & Jia, X. (2015). Fault diagnosis approach for rotating machinery based on dynamic model and computational intelligence. *Measurement*, 59, 73-87.
- Zhu, W., Zeng, N., & Wang, N. (2010). Sensitivity, specificity, accuracy, associated confidence interval and roc analysis with practical sas implementations. *NESUG proceedings: health care and life sciences, Baltimore, Maryland, 19*, 67.