Bearing Condition Monitoring Using Multiple Sensors and Integrated Data Fusion Techniques

S.L. Chen[1], M. Craig[1], R.J.K. Wood[1], L. Wang[1], R. Callan[2], H.E.G. Powrie[2]
[1]Surface Engineering and Tribology Group, University of Southampton, Southampton, SO17 1BJ, UK
[2]GE Aviation, Digital Systems, Chandlers Ford, SO 53, 4YG, UK

ABSTRACT

Recently, at the University of Southampton, a series of taper roller bearing tests have been conducted to evaluate the effectiveness of using on-line sensing technologies to detect incipient faults. The test rig instrumentation included vibration and electrostatic sensors as well as off-line processes such as debris analysis and tribological assessment of bearing condition, which are used for corroborative purposes. The test results indicate that the combination of these techniques are capable of detecting bearing deterioration shortly before complete failure, but the expected precursor events related to fatigue failure initiation are indistinguishable from the conventional univariate signal plots. Therefore, more advanced data fusion techniques have been developed to extract further information and enhance the detection of abnormal trends. This paper describes work on intelligent processes of both training and testing data and demonstrates how the prognostic window is significantly improved relative to original processed features. The approach also identifies the main variables which drive the anomaly detection so as to provide diagnostic information.

Keywords: Rolling Element Bearing, Condition Monitoring, Vibration Monitoring, Electrostatic Monitoring, Data Fusion.

1. INTRODUCTION

Vibration analysis has been used in bearing condition monitoring for over 50 years, and has achieved various degree of success. Electrostatic charge detection was originally developed by GE Aviation (formerly Smiths Aerospace) for the identification of debris in the gas path of jet engines [1]. Further research at the University of Southampton has shown its potential to detect precursor processes that indicate contact distress and wear in lubricated environments [2]. Recently, a series of taper roller bearing tests have been conducted to evaluate the effectiveness of using on-line sensing technologies to detect incipient faults [3]. The test rig instrumentation included vibration accelerometer and electrostatic sensors as well as off-line processes such as debris analysis and tribological assessment of bearing condition, which are used for corroborative purposes. Although the test results indicate that the combination of these techniques is capable of detecting bearing deterioration shortly before complete failure, expected precursor events related to initiation of bearing fatigue failure are indistinguishable from the conventional univariate signal plots. Moreover, these techniques require a lot of manpower to analyze the results in order to achieve detection and diagnostic information. Thus, developing a systematic approach that can automatically extract bearing abnormal events and diagnostic information is crucial.

So far, various approaches have been tried, including artificial neural networks (ANNs) [4], fuzzy logic [5], evolutionary algorithms [6] etc. to develop automatic fault detection and diagnosis systems. Although the performances of these systems are consistently improved with the adaptation of the algorithms or implementation of the data pre-processing [7] and feature selection [8] techniques, there is still a major concern over this development strategy: it is not easy to obtain multi-class training data covering all aspects of bearing symptoms that can appear very different and confusing. Nevertheless, data representing the normal condition of the bearings is the most easy-to-obtain dataset, and the systems can be built to detect whether the bearing condition deviates from the 'normal' region rather than mapping collected data into several established categories of the bearing conditions. This idea has been used as the central development strategy by novelty detection [9, 10] and statistical process control (SPC) [11, 12] techniques in image processing and chemical plant applications.

In this paper, a Gaussian mixture model (GMM) is proposed to summarize the probabilistic density of the training data (normal condition) and to build the normal model. However, the true condition of the training data is unknown and it might contain a number of anomalous data due to the sensor malfunction or environmental effects etc., so the GMM could not be expected to respond to these anomalies [13]. Therefore, the model is adapted with a novel approach to suppress regions with anomalies. Further, Hotelling's T-squared statistic is applied to extract abnormal information by measuring the distance between the test sample and the origin of the adapted normal model. Finally, the contribution values to the calculated T-squared statistic are calculated for each extracted feature, which are related to bearing components as well as to the sensor locations.

This paper is organized as follows: a brief introduction of the proposed approach is presented in section 2. Section 3 describes the experiments, extracted signals in both time and frequency domains and applied datasets. Section 4 provides the training process of the normal model, and the anomaly detection and diagnosis results of the test data is described in the section 5.

2. METHODOLOGY

Fig. 1 shows the flow chart of the proposed scheme. At the beginning, multiple features from time domain, i.e. RMS value and frequency domain, i.e. energies at bearing defect frequencies are processed for both the training and testing data. Next, the GMM approach is applied to describe the training data. Thirdly, the innovative methods are adopted to locate and trim Gaussian components in the existing GMM associated with anomalies. The scheme then utilizes principle component analysis (PCA) in each of the remaining Gaussian components, so that the multiple PCA model can be achieved. Finally, the multivariate testing data is fed into the multiple PCA models, so the Hotelling's T-squared statistic and variable contribution values can be calculated to obtain the anomaly detection and diagnosis information respectively.

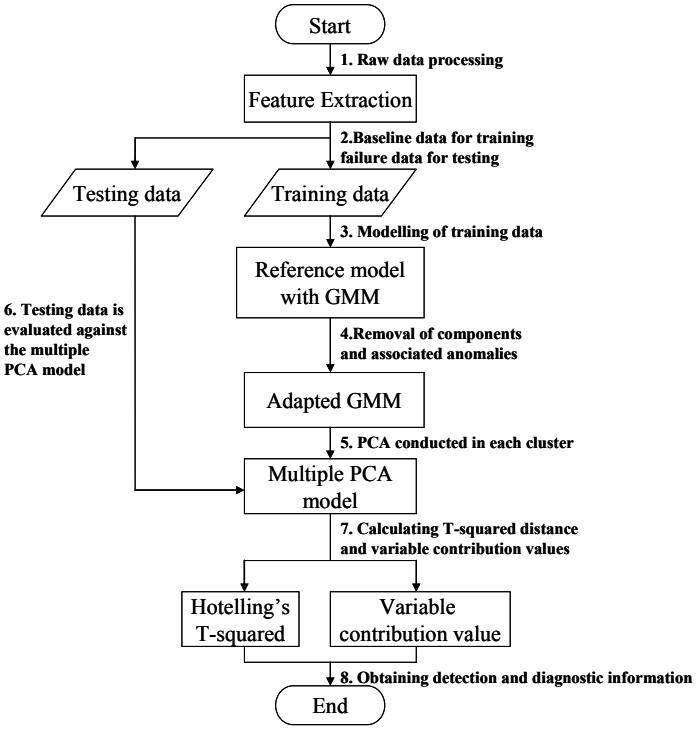

Fig. 1 Methodology flowchart of the proposed intelligent scheme

2.1. Gaussian Mixture Model

Often it is assumed that normal data may be described by a single Gaussian distribution. However, real world data can be distributed in a more complicated way. Data from different time intervals might occupy their own space of normality (e.g. vibration levels varies at different times due to a different load regime) or it may contain unknown anomalies. Clearly data with these elements will not follow a single Gaussian distribution. In this case, the data may be characterized by a combination of Gaussian distributions, or a Gaussian Mixture Model. Details of the GMM and its associated Expectation-Maximization (EM) algorithm used for clustering the training data can be found in [14]. It should be noted that the Bayesian Information Criterion (BIC) [15] was utilised in the current study to indicate the optimized number of clusters.

2.2. Adaptation of GMM to remove anomalies in the training data

As indicated in the introduction section, the true condition of the training data is unknown and it

might contain a number of outlying data which could affect the fault detection results. In this paper, a novel approach is introduced to identify areas of the cluster space that might be associated with anomalies.

To define the proposed approach, two assumptions need to be made according to the general distribution of the anomalies: 1) The anomalies are occurring at occasional time intervals in the time series; 2) The clusters associated with anomalies are distant from the clusters with normal data.

- **Entropy based method**

To find out anomalies under the first assumption, the Entropy statistic from the information theory is applied. For n partitioned categories (equal time intervals) in the time series, and for a training vector x with all samples having a cluster ID v, the Entropy E is defined as [16]:

$$E(x=v) = \sum_{i=1}^{n} -p_i \log_2 p_i \qquad (1)$$

Where p_i is the probability of training vector x coming from cluster v of category i occurring.

Fig. 2 shows the concept and steps of using the Entropy statistic for finding the anomalies embedded in the GMM subspace. Hence, the main idea of the entropy based anomaly detection method is to explore whether the support cases in a particular cluster are from occasional time intervals or multiple time intervals. If data in a cluster is just from one time interval, the probability of the data points from this time interval in the current cluster will be as high as 100%, and the probabilities of the data points from other time intervals are 0%, so that the entropy score for the current cluster will be as low as zero. Regarding the first assumption, the current cluster could be recognized as a region associated with anomalies.

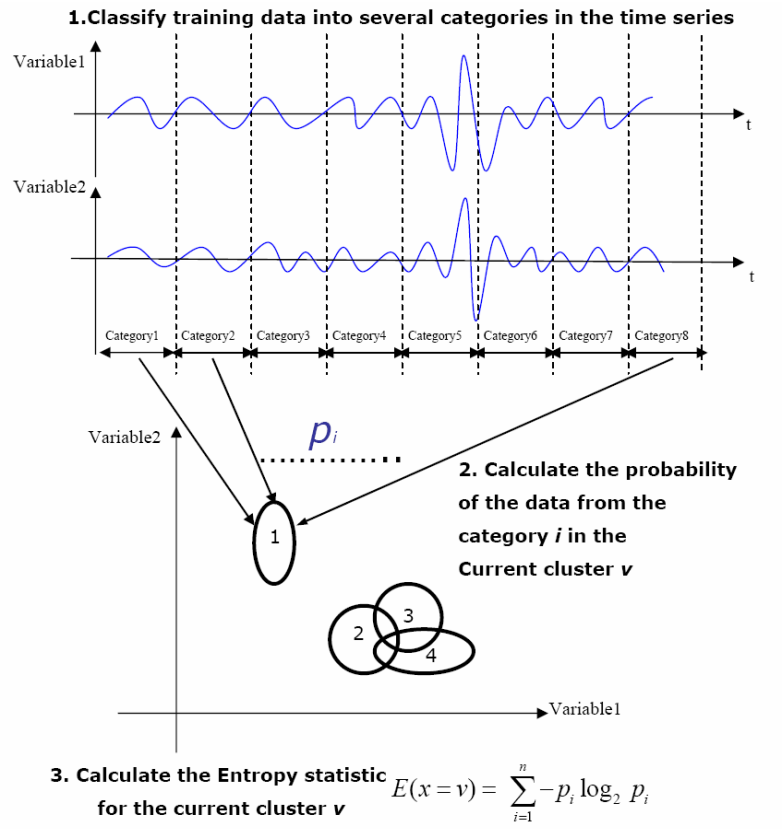

Fig.2 Steps of the entropy based approach to remove anomalies In the training data

- **Distance based method**

Although anomalies are found occurring occasionally in the time series of the training data in most situations, however anomalies are also appearing frequently throughout the test, and the entropy based method will be insensitive to these anomalies. Hence, a distance based method is developed to measure the data dispersion-clusters with distant points from other regions being eliminated from the calculation. In this study, the Hotelling's T-squared distance is calculated between two clusters [17], using equations (2) and (3):

$$T^2 = \frac{n_A n_B}{n_A + n_B}(\mu_A - \mu_B)^T \hat{\Sigma}_{AB}^{-1} (\mu_A - \mu_B) \qquad (2)$$

$$\hat{\Sigma}_{AB} = \frac{1}{n_A + n_B - 2}((n_A - 1)\hat{\Sigma}_A + (n_B - 1)\hat{\Sigma}_B) \qquad (3)$$

Where n_A and n_B are the number of support cases in clusters A and B; μ_A and μ_B are sample means of

clusters A and B, respectively; $\hat{\Sigma}_A$ and $\hat{\Sigma}_B$ are estimates of covariance matrices of the clusters A and B, respectively.

2.3. Anomaly detection and diagnosis

After characterization of the training data and adaptation of the GMM, the PCA is applied in each of the remaining Gaussian component, so the multiple PCA models can be generated. For detailed introduction of PCA, please refer to [18]. And then, for the given testing vector y_i, the sum of normalized squared scores, known as Hotelling's T-squared statistic is calculated to measure the variation of testing vector y_i within the PCA model. The T-squared statistic is defined as:

$$T_i^2 = t_i \lambda^{-1} t_i^T = y_i P_c \lambda^{-1} P_c^T y_i^T \tag{4}$$

where t_i refers to the i^{th} row of Tc , the m by c matrix of c scores vectors from the PCA model, and λ is a diagonal matrix containing the eigenvalues (λ1 through λc) corresponding to c eigenvectors (principal components) retained in the model. On the other hand, the anomaly detection threshold is defined by maximum Hotelling's T-squared distance of the training data.

It should be noted that Hotelling's T-squared statistic is calculated based on a single component within the multiple PCA models for each testing vector. The component is chosen with the minimum Mahalanobis distance, d_k, from its centre to the testing vector [9], given by equation (5):

$$d_k = \sqrt{(y_i - \mu_k)^T \sum_K^{-1} (y_i - \mu_k)} \tag{5}$$

Where μ_k and Σ_k are mean vector and covariance matrix of k^{th} Gaussian component respectively.

Furthermore, the T-squared contribution values describe how individual variables contribute to the Hotelling's T-squared value for a given testing sample.

$$t_{con,i} = t_i \lambda^{-1/2} P_c^T = y_i P_c \lambda^{-1/2} P_c^T \tag{6}$$

Thus, diagnostic information is achieved by calculating the variables contribution value, to identify which variables are the main driven factors for the detected anomalies.

3. EXPERIMENT AND DATASETS

3.1. Experiment

The bearing rig tests discussed in this paper form part of a series of tests undertaken at the University of Southampton. Details of the rig may be found in other references (see [3] for example), so only a brief overview is provided here.

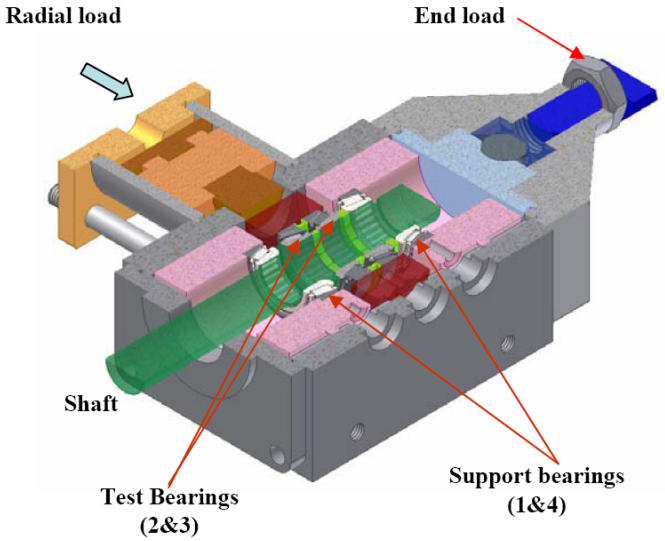

Fig.3 Bearing rig chamber showing main test components

The bearing rig comprises four taper roller bearings housed in a chamber, as shown in Fig. 3. The bearings are mounted on a shaft driven by an electric motor. The outer two bearings (#1 and #4) are support bearings, whilst the inner bearings (#2 and #3) are test bearings. The two test bearings have a radial load applied to accelerate their failure. In addition, #2 was pre-indented on the inner race. The baseline test used non-indented bearings. The bearing types were LM67010 (cup) and LM67048 (cone), which are all steel bearings. The rig was run at a

constant speed of 2500 rpm under ambient conditions and fully flooded with the Shell Vitrea 32 lubricant at the oil flow rate of 4 litres min^{-1}.

The rig was instrumented with a number of condition monitoring sensors. In the test chamber there was one vibration sensor (mounted externally) and three electrostatic wear-site sensors (WSS) - one each on bearings #1 and #4 and one looking at bearings #2 & #3. The oil re-circulation pipework included an electrostatic oil-line debris sensor. Additional oil-line debris sensors included an inductive debris sensor, which provided wear particle counts for > 100 microns.

The testing also included a number of complimentary off-line analyses. Oil samples were taken during various stages of the testing and these were analyzed for sub-100 micron debris content. Tribological analyses of the bearing condition pre and post-test were conducted and included photographic evidence and mass loss calculations. Where applicable, the oil-line sensors and post-test / off-line analyses were used to help interpret the responses of the vibration and WSS sensors during the tests.

3.2. Feature extraction

The post-processed signal features from the three types of sensors (vibration, WSSs and OLS) were extracted to construct both training and testing data set for the developed intelligent scheme.

To form the vibration and WSSs training set, six dimensional features were selected for each of the two sensing technologies. In theses features, the first and second dimensions were chosen as the RMS value of the time domain and the energy at the rotational frequency respectively. The third to sixth features are based on the energies at the bearing defect frequencies, i.e. races, rollers and cage. These features are used to reveal the information on the bearing components. Each energy was externally clocked using once per revolution signal from the shaft tachometer. Data sets of RMS value and energies at bearing defect frequencies were produced at two points per minute. The bearing defect frequencies are listed as the equations (7) ~ (10) given by [19],

-Cage frequency

$$\omega_c = \frac{\omega_s}{2}\left(1-\frac{d}{D}\cos\alpha\right) \quad (7)$$

-Roller frequency

$$\omega_{re} = 2\omega_b = \frac{D\omega_s}{d}\left(1-\frac{d^2}{D^2}\cos\alpha^2\right) \quad (8)$$

-Outer race frequency

$$\omega_{od} = \frac{Z\omega_s}{2d}\left(1-\frac{d}{D}\cos\alpha\right) \quad (9)$$

-Inner race frequency

$$\omega_{id} = z\left(\omega_s-\omega_c\right) = \frac{Z\omega_s}{2}\left(1+\frac{d}{D}\cos\alpha\right) \quad (10)$$

Where ω_s is the shaft rotational speed in rad /s, d is the diameter of the roller, D is the pitch diameter, Z is the number of rolling elements and α is the contact angle.

Vibration	**WSS1**	**WSS2**	**WSS3**	**OLS**
RMS	RMS	RMS	RMS	RMS1
Tacho Energy	Tacho Energy	Tacho Energy	Tacho Energy	RMS2
Cage Energy	Cage Energy	Cage Energy	Cage Energy	Indicator
Roller Energy	Roller Energy	Roller Energy	Roller Energy	
OuterRace Energy	OuterRace Energy	OuterRace Energy	OuterRace Energy	
InnerRace Energy	InnerRace Energy	InnerRace Energy	InnerRace Energy	
Front housing	**Vicinity of Bearing1**	**Vicinity of Bearings2&3**	**Vicinity of Bearing4**	**Oil-line**

Table1 Extracted features for the multivariate data analysis

For the OLS analysis, three time domain parameters including RMS values of the two oil line sensors and the cross-correlated value that couples two OLSs to identify charge moving at the oil flow rate, were utilised to generate data set. This training and testing sets were produced one point every two seconds. Table 1 tabulates the selected features for the bearing wear detection.

3.3. Bearing test datasets

In this paper, 2 sets of data are utilized to validate the proposed anomaly detection and diagnostic scheme. Table 2 summarizes the datasets and their test conditions and functions. From the test condition perspective, there are two modes of test, the first one is the baseline test, in which bearings are free from defects; the second one is the defect test, in which the inner race of one of the test bearings was pre-indented to accelerate the fatigue failure process. According to the investigation of the proposed condition monitoring scheme, the central idea of utilization of these data is: the baseline datasets are collected to build the reference model representing the normal operating condition, while the defect test datasets are evaluated against the established reference model to detect anomalies as well as diagnosing detected anomalies.

Test No.	Test load/kN	Test duration/hours	Defect bearing	Speed/rpm	Application
1	20	80	none	2500	Training data
2	20	62	#2	2500	Testing data

Table2 Bearing test data used in this paper

4. TRAINING OF THE NORMAL MODEL

4.1. Construction of the classic mixture model

The first step is to build the suitable GMM to correctly describe the training data. Fig. 4(a) shows the Bayesian Information Criterion (BIC) for model-based method applied to test 1 data. It was found that the local minimum for the mixture model occurs with eleven clusters, hence, this was the number of clusters chosen and the trained normal model is illustrated in Fig. 4 (b).

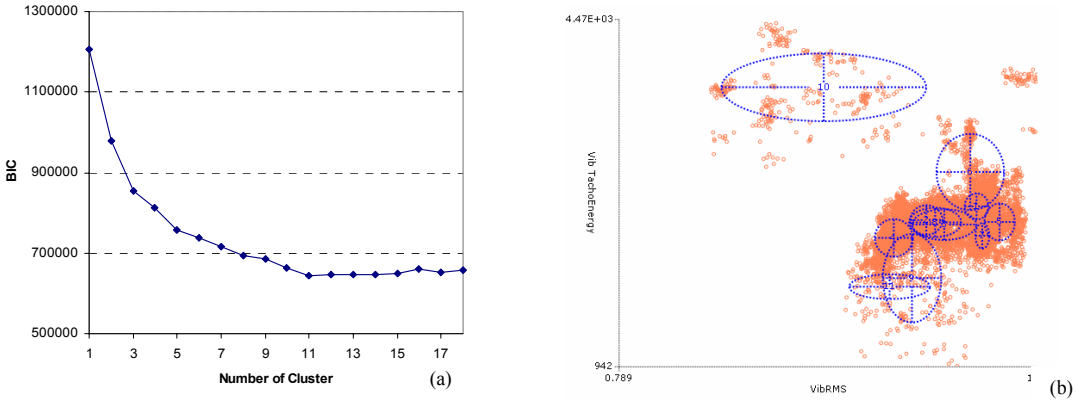

Fig.4 BIC score for GMM applied to test 1 data (a) Trained GMM for the test 1 data (b)

4.2. Adaptation of GMM

It was indicated in section 2.2 that the training data (test 1 in this case) might contain a number of anomalies, and need to be eliminated by the developed entropy and distance based approaches.

At the beginning of the approach, the training data in the time series needs to be partitioned into several categories. The difficult task here is to determine the number

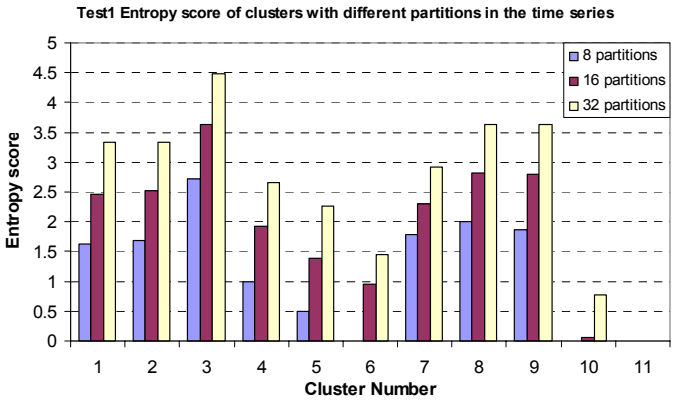

Fig.5 Entropy score of the clusters with three partition numbers

of partitions. Mostly the partition needs to be assigned properly, as it directly affects the anomaly detection result. Thus the parameter of number of partitions has to be set explicitly. In this paper, the number of partitions is varied, and cross validation helps to find an appropriate parameter.

For the test 1 data, three partition numbers 8, 16 and 32 were assigned accordingly. Fig. 5 shows the calculated entropy score for each of the 11 trained clusters with 3 different partitions in the time series. It can be seen that the entropy score for each cluster increases lineally with increasing number of partitions. From the Fig.5, the entropy scores of clusters 10 and 11 are seen as significantly lower than the other scores of clusters, no matter how many partitions are assigned. Thus, these two clusters are highly suspected to be the regions associated with anomalies, and are suggested to be removed from the classical GMM.

Fig. 6 shows vibration RMS value before and after the removal of clusters 10 and 11 and their support cases. It was found a distinctive running-in region existed in the original plot (Figure 6(a)), if these running-in data points are used to build the GMM, they would lead to a coarser view of 'normality' which may not prove so effective when trying to detect real faults. After the removal of clusters 10 and 11 and their support cases as the Figure 6(b) shows, the running-in region data points have been successfully trimmed, and a new GMM can be generated for the anomaly detection.

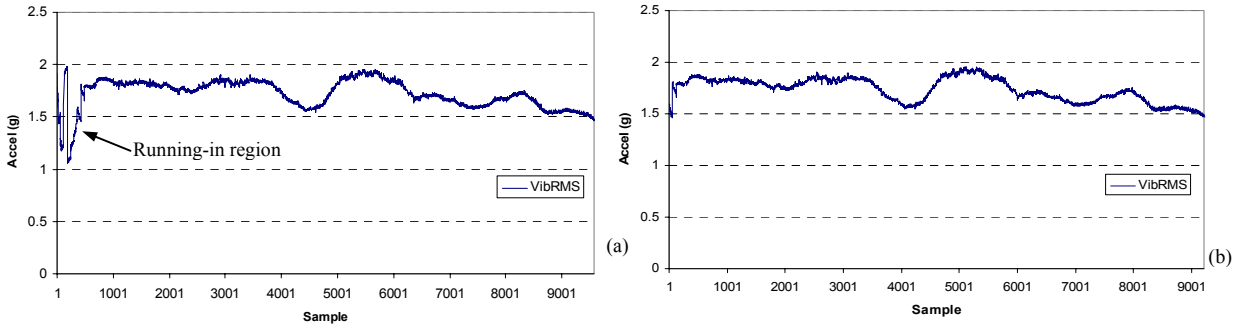

Fig.6 Vibration RMS plot before (a) and after (b) the removal of clusters 10, 11 and their support cases

However, the anomalies might be occurring frequently or distributed throughout the dataset, and the entropy based method would be ineffective in detecting these anomalies. This situation also occurred with the test 1 data. Drilling down to the extracted features in the frequency domains, as Fig. 7(a) illustrates, a significant number of anomalies still exist throughout the test, although clusters 10 and 11 and their associated support cases were eliminated.

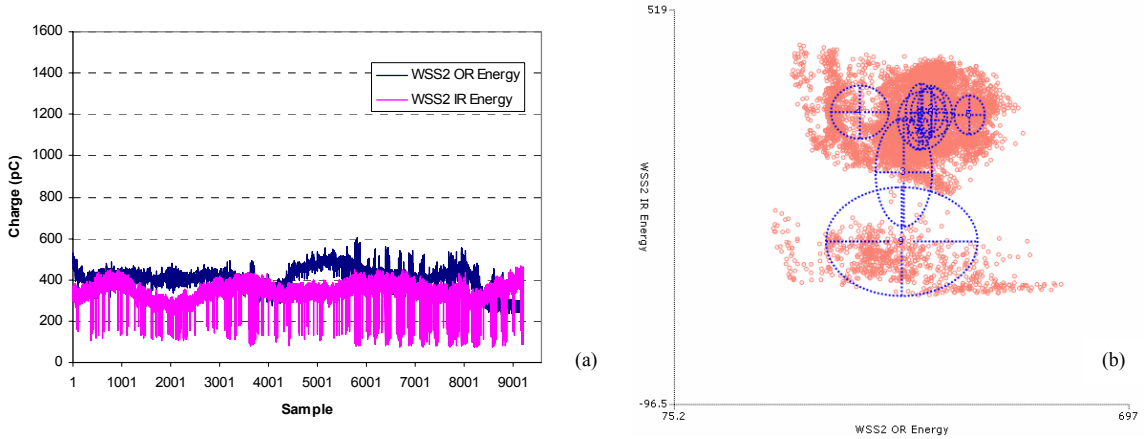

Fig.7 WSS2 OR and IR plot after the removal of cluster 10, 11 (a) GMM in the space of WSS2 OR and IR energy (b)

If the clustering space is plotted with the dimensions of WSS2 OR and WSS2 IR in Fig. 7(b). The cluster 9 might be the region with the anomalies discovered in Fig. 7(a). Going back to the entropy

score plot (Fig. 5), cluster 9 is the region with high entropy score with 3 types of partitions. Hence, this cluster should not be removed.

Under this circumstance, the second strategy to remove anomalies in the training data is triggered. With the second anomaly detection strategy, Hotelling's T-squared distance is calculated between each of the Gaussian components. Fig.8 (a) illustrates the calculated result. It can be seen that clusters 3 and 9 are recognized as distant from the other clusters, so the support cases associated with these two clusters are identified as anomalies. Fig.8 (b) shows the WSS2 OR and WSS2 IR data after the removal of clusters 3 and 9 and their associated data points. Most of the defined anomalies have been filtered out from the original data plot, and shown a reasonable flat trend for the usage of the normal training data.

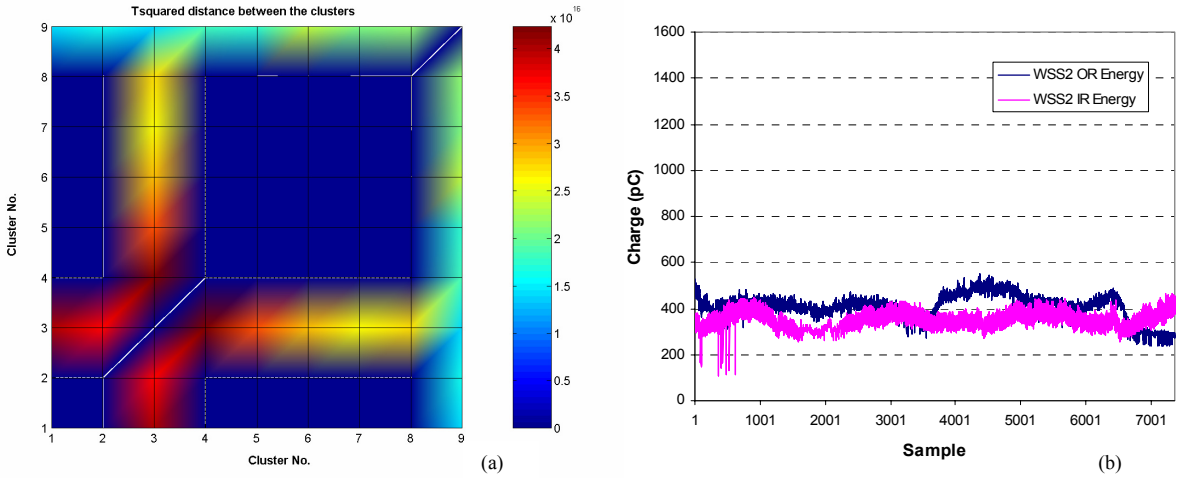

Fig.8 T-squared distance between the trained clusters (a)
WSS2 OR and IR energy plot after the removal of clusters 3 and 9 (b)

4.3. Principle Component Analysis in the remained Gaussian components

After the construction and adaptation of the normal reference model, the Principle Component Analysis (PCA) is applied in each of the remaining Gaussian component, so that the multiple PCA subspace can be built. Table 3 shows the number of the Principle Components (PCs) and their captured variance for each of the remained Gaussian components of the test 1 data.

Remained clusters	Number of PCs	Cumulative Variance (%)
1	14	82.61
2	11	82.04
4	14	81.18
5	13	80.42
6	13	82.49
7	12	80.73
8	13	80.40

Table 3 Summary of number of extracted PCs in the remained clusters

5. ANOMALY DETECTION AND DIAGNOSIS RESULTS

5.1.Anomaly detection

The effectiveness of using equipped multiple sensing technologies to detect impending failure of the bearings, and the multivariate dataset, marked as test 2 in this paper are reported in [3]. There are two aims behind the further analysis of the data. First to determine whether an earlier (prognostic) indication of the failure can be achieved, i.e. can the initiation of fatigue failure be detected in advance of the finial few hours? Second, the capability for determining what and where the fault is i.e. is there clear diagnostic information available? In the field, both of these are important if goals such as optimized maintenance planning, minimum logistic footprint and maximum equipment are to be

achieved.

Fig. 9(a) shows the Hotelling's T-squared statistic with the test 2 data, which has been run through the normal reference model trained with test 1 data. From this figure there is evidence of abnormal behaviour at the start of the test which, is due to the bearing running in. The run out to failure from about 54 hours, dominates. However, it is encouraging to find evidence of change in advance of 54 hours: an overall increase at around 36 or so hours, with significant peaks between 43 and 45 hours. It is also noted that these detected prognostic activities could not be seen clearly in the original extracted features, which proves the advantage of the proposed anomaly detection approach.

So far, it is not clear the cause of the period of activity between 43 and 45 hours in Fig. 9(a), they may result from the generated wear debris by delamination from the pre-indent on the inner race of bearing #2 or surface debounding by subsurface cracks. It is also found in Fig.10, that the off-line debris counter could detect increases in debris production at 43 hours. This correlates with the detected activities by the T-squared statistic over the similar period of time. Hence, these abnormal events are invaluable, as they could be used to detect debris related abnormal conditions that result from the initiation of fatigue failure.

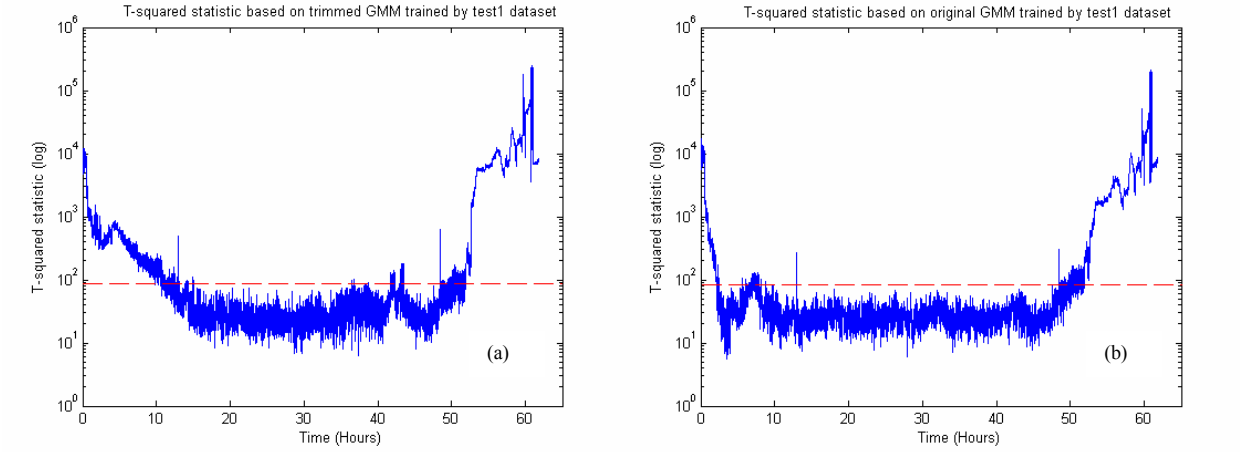

Fig.9 T-squared statistic with the test 2 data based on trimmed GMM (a)
T-squared statistic with the test 2 data based on original GMM (b)

In order to verify the importance of the model adaptation technique, the Hotelling's T-squared statistic is also calculated against the un-trimmed original GMM as shown in Fig. 9(b). The prognostic activities between 43 and 45 hours can not be visualized, this is due to a number of anomalies contained in the training data (test1) which is used to build the model of normality, and the abnormal events in the testing data which are similar to the anomalies in the training data become undetectable.

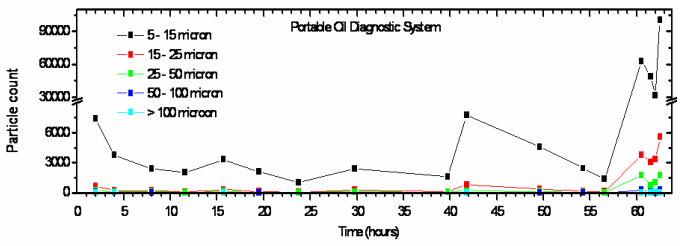

Fig.10 Off-line debris counter

5.2. Anomalies diagnosis

The anomaly detection results indicate two periods of interest as shown in Fig. 9(a). In the field, it might be insufficient to conduct appropriate maintenance plan with only the knowledge of detected anomalies, as it is meaningful to understand where these anomalies occur. As stated before, the applied features could reveal the condition of different bearings and components within the bearing. Hence, there is a need to investigate which variables are the main driving factors for the detected

anomalies to achieve a diagnostic function. In this study, contribution values for each of the variables are calculated to identify the main influencing factors that drive the detected anomalies of interest.

Fig. 11(a) shows the absolute contribution values of the rolling element energy for the four applied sensors. Over the first period of interest (between 37 and 43 hours, as Fig. 11(b) shows), the WSS3 rolling element energy dominates the trend with additional peaks, while the other rolling element energies show insignificant contributions to the detected abnormal trend. On the other hand, it is clear to see that the vibration and WSS3 rolling element energies are the main driving factors to the increasing trend during the second period of interest (54-63 hours). For the contribution value of outer race energies (see Fig.11 (c) and (d)), vibration and WSS3 elements dominate both periods of interest, but do not seem as significant as the rolling element energies. For the contribution value of inner race energies (not shown here), the difference from the previous two bearing elements is that contribution values of inner race energies can not be distinguished to see which variables are the major contribution resource for both two areas of abnormal, and have the relatively weak impact. Apart from the contribution values of the bearing element energies, the oil-line variables are also examined, and found that the OLS variables were seen to be significant in the second period interest.

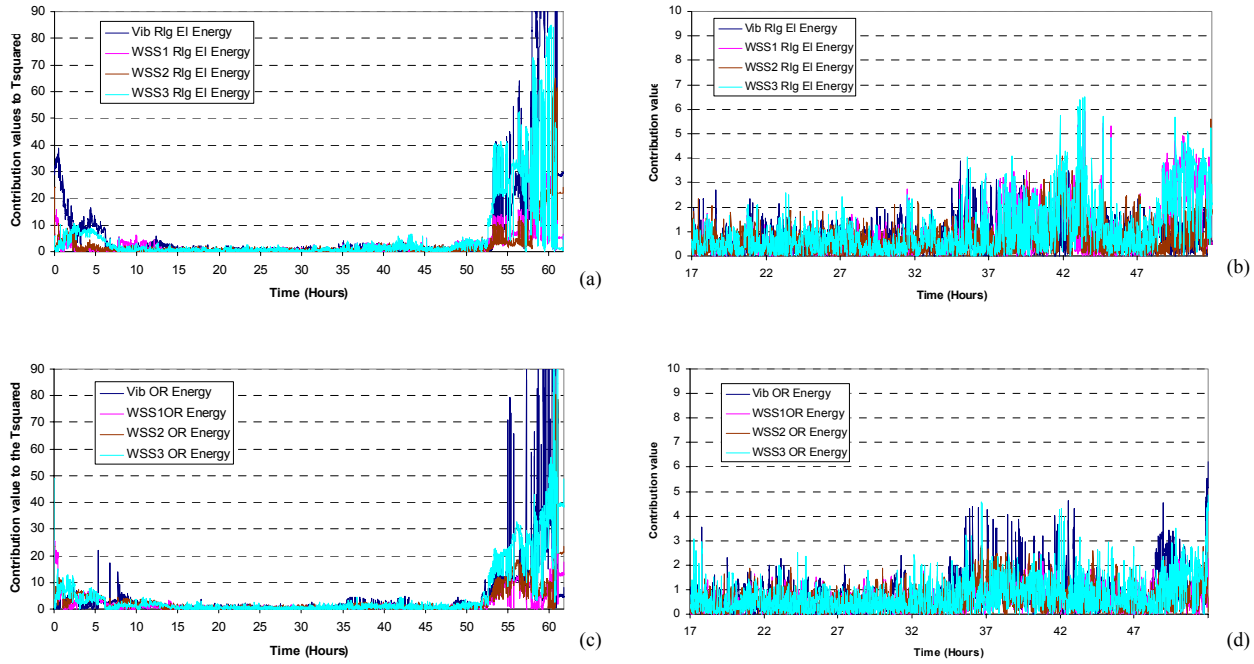

Fig.11 Contribution charts of the applied variables for the bearing diagnosis

Therefore, it can be seen that the WSS 3 rolling element and vibration outer race energies obtain the highest impact on the detected prognostic events (between 37 and 43 hours), from which two diagnostic information is generated: 1) abnormal conditions occurred within the bearing #4, since it is monitored by WSS3. 2) As the WSSs are designed to detect charge generated between the contacting surfaces and wear debris, it can be inferred that the delamination or initial spallation occurring between the contacting surfaces of rolling elements and outer race of bearing #4, and amount of debris generated which is detected by off-line debris counter.

On the other hand, nearly all the variables have the impact on the abnormal condition at the end stage (54-63 hours), but the rolling element and outer race energies of vibration and WSS3 are the strongest. This could be explained as spallation occurred within the bearing #4 due to rolling contact fatigue, hence large amounts of debris and hugely modified contact surfaces generate high charge and strong energies which are reflected by the WSS and vibration sensors, respectively. Furthermore, OLS

variables also show a significant trend during this period. This is because the OLSs are installed in the oil-line, and the generated debris was brought by the oil flow passing through the OLSs, causing significant charges.

5.3.Bearing inspection

In order to relate the physical condition of the bearing elements to the diagnostic results, both these faults are corroborated by post-test inspections of the bearings. It is clear to see in the Fig. 12 that there is significant damage to the outer race of bearing #4 and several rolling elements of the same bearing suffered material loss. This confirms that, for the two periods of interest, the fatigue initiation and spallation occurred between the contacting surfaces of the rolling elements and the outer race of bearing #4.

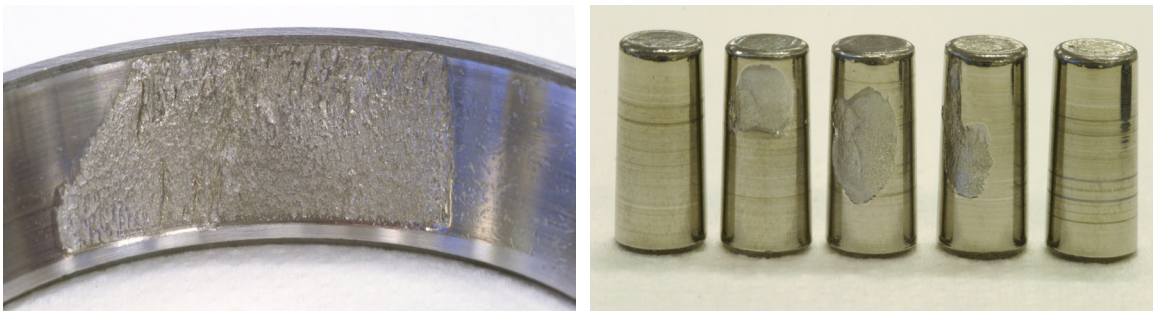

Fig.12 Post-analysis of the failed bearing #4 of outer race and rolling elements

6. CONCLUSIONS AND FUTURE WORK

A new intelligent scheme has been developed to detect early signals of bearing distress. From the work conducted to date the following conclusions can be made:

1) A novel approach based on entropy statistic and T-squared distance has been developed to effectively remove the anomalies in the training data, and the demonstration results show the importance of adaptation of the normal model.

2) The T-squared distance has been utilized for extracting abnormal events from the multivariate data. Prognostic information has been extracted approximately 10 hours before the severe wear occurred.

3) The contribution values of the extracted features can assist operators locate faulty bearings as well as indicate that the vibration features have the strongest impact on the T-squared statistic at the second period of interest while the electrostatic features are more sensitive to the precursors. Furthermore, off-line debris analysis and final bearing surface inspection also provide strong physical evidence for the precursors and predicted faulty bearing locations.

4) The vibration features were examined by the developed scheme for the first time to obtain the physical understanding of the vibration sensor. From the testing results, the vibration sensor could be used to detect and locate prognostic events that are related to fatigue initiation, this capability was complementary with the electrostatic wear-site sensors which also provide contribution to detect prognostic events. On the other hand, vibration sensor is found to be more sensitive for monitoring severe wear of the bearing (e.g. large amounts of wear debris were entrained between the raceways and rolling elements due to the spallation giving rise to significant shocks).

Although the test results have shown the effectiveness of the scheme, there are still several concerns that need to be addressed by future work:

1) In the approach of reference model adaptation, number of the partitions assigned in the time series is set explicitly. A mechanism is required to set this parameter by investigating number of the data points and dimensionality in the training data.

2) The threshold value for the anomaly detection needs to be optimized.

3) To achieve the automated reasoning, the system needs to be assigned with the knowledge that represents different conditions of the bearings. The current diagnostic method using contribution values is not an automatic approach. Hence, it is important to develop diagnostic training data representing different bearing symptoms (bearing fault mechanism, and fault location) by discovering knowledge from the run-to-failure test and recording characteristic features from such datasets.

Acknowledgement
The authors would like to thank GE Aviation for funding this project and Dr. Harvey from Southampton University for providing the test data.

References

[1] Powrie H, McNicholas K. Gas path condition monitoring during accelerated mission testing of a demonstrator engine, AIAA 97, 1997, Paper 97-2904.
[2] Morris S, Wood RJK, Harvey TJ, Powrie H. Use of electrostatic charge monitoring for early detection of adhesive wear in oil lubricated contacts, ASME Journal of Tribology, 2002,124: 288-96.
[3] Harvey TJ, Wood, RJK, Powrie H. Electrostatic wear monitoring of rolling element bearings, Wear, 2007, 263: 1492-1501.
[4] Wang L, Hope AD. Bearing fault diagnosis using multi-layer neural networks. Insight: Non-Destructive Testing and Condition Monitoring 2004; 46(8): 451-55.
[5] Liu TI, Singonahalli JH, Iyer NR. Detection of roller bearing defects using expert system and fuzzy logic. Mechanical Systems and Signal Processing 1996; 10(5): 595-614.
[6] Zhang L, Jack LB, Nandi AK. Fault detection using genetic programming. Mechanical Systems and Signal processing 2005; 19: 271-89.
[7] Yang J, Zhang Y, Zhu Y. Intelligent fault diagnosis of rolling element bearing based on SVMs and fractal dimensions. Mechanical Systems and Signal Processing 2007; 21 (5): 2012-24.
[8] Sun W, Chen J, Li J. Decision tree and PCA-based fault diagnosis of rotating machinery. Mechanical Systems and Signal Processing 2007; 21 (3): 1300-17.
[9] Roberts SJ, Novelty detection using extreme value statistics, IEE Proc. on Vision, Image and Signal Processing. 1999; 146(3): 124-29.
[10] Banister PR, Tarassenko L. Learning jet engine vibration response for novelty detection. The 2nd World Congress on Engineering Assessment Management and the 4th International Conference on Condition Monitoring. 2007; 229-38.
[11] Kresta JV, MacGregor JF, Marlin TE. Multivariate statistical monitoring of process operating performance. Can. J. Chem. Eng. 1991; 69: 35–47.
[12] Kourti T, MacGregor JF, Process analysis, monitoring and diagnosis, using multivariate projection methods. Chemometrics Intell. Lab. Syst. 1995; 28: 3–21.
[13] Callan R, Larder B, Standiford J. An integrated approach to the development of an intelligent prognostic health management system. IEEE Aerospace Conference, Montana, USA, 2006.
[14] Dempster AP, Laird NM, Rubin DB, Maximum likelihood from incomplete data via the EM algorithm, Journal of the Royal Statistical Society 1977; B 39:1-38.
[15] Fraley C, Raftery AE. How many clusters? Which clustering method? Answer via model-based cluster analysis. The Computer Journal 1998; 41(8): 578-88.
[16] Callan R, Artificial Intelligence, Palgrave, UK, ISBN 0-333-80136-9, 2003.
[17] Anderson T. Am Introduction to Multivariate Statistical Analysis. New York: John Wiley and Sons, 2nd edn, 1984.
[18] Bishop CM, Neural Networks for Pattern Recognition. 1995, Oxford, UK, ISBN 0 19 853864 2 (pbk).
[19] Tandon N, Choudhury A. A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings. Tribol. Int. 1999; 32: 469-80.