Performance Improvement of Classifier in Fault Diagnosis of Rotating Machines Using Sensor Fusion Techniques

S. Meenakshi Sundaram, M. Saimurugan

Abstract: The shaft, rotor, bearing and gear are the important elements of the rotating machines. Most of the problems in rotating machines are caused due to bearings and shaft. The failure of rotating machine causes production downtime and economic & safety issues. Vibration signal analysis is highly accepted technique in fault diagnosis of rotating machine. For automation of fault diagnosis, machine learning approach has been followed. Machine learning classifies fault based on variation in signatures pattern of the machine. But its effectiveness gets reduced when it is used for multi fault class problem. So in the present work, sound signals are also used along with vibration signals for applying sensor fusion techniques. In sensor fusion, signals from various sensors are fused in three levels such as data fusion, feature fusion and decision level fusion and the fused data sets are used for fault classification using machine learning algorithm. The performance of each technique is studied in detail and compared using classification accuracy. A new method is proposed by combination of fusion techniques to enhance the performance.

I. INTRODUCTION

The rotating machines are largely used in various industries such as textile, paper, automobile, aerospace etc. Automotive and aircraft engines, gearboxes, compressors, pumps, turbines, electric motors, etc., belong to the category of rotating machines. The fault occurring in any of the component in these rotating machines lead to unexpected machine breakdown or downtime, productivity loss, reduction in quality and economic loss. Machine condition monitoring and fault diagnosis was rapidly growing due to its ability to avoid machinery breakdown and prevent catastrophic failure which gives early warning signals and detect the fault before it leads to machinery failure. Fault diagnosis of rotating machinery has been studied in detail during the last sixty years. In fault diagnosis, parameters of the machine are acquired and monitored for identification of the fault (Mechefske, 2005). So it is necessary to detect the fault early by suitable techniques for maintaining the machinery in good condition. Vibration signal analysis is the most important technique used for fault diagnosis approach. Two important parameters which are used extensively for detection and diagnosis are amplitude and frequency. Amplitude will be varying if any faulty component runs inside the machine. With the help of frequency domain analysis various faults can be diagnosed such as bearing defect, worn out gear, bent shaft, unbalanced rotor etc. Initially, vibration analysis was performed by using time domain signal. Vibration monitoring is widely used in identification of crack.

(Doebbling et al., 1998), rotor unbalance (Parkinson, 1991; Foiles et al., 1998), coupling misalignment (Gibbons, 1976; Sekhar and Prabhu, 1995). The statistical parameters such as probability density and kurtosis can be effectively used for identification of bearing fault (Dyer and Stewart, 1978). Later, research on sound signals are used. Cost of acquiring sound signals became more affordable than vibration signals because the transducer for acquiring sound data is very less when compared to cost of vibration transducer. Gustafsson and Tallian (1962) proved that kurtosis is ineffective for finding the fault at an initial stage. Mitchell(2000) used statistical parameters for machinery fault diagnosis. The concept of machine learning was introduced before the 1990s. Li & Wu(1989) introduced the pattern recognition system. Pattern recognition system works on the principle of classification of objects based on their subject. W.J Wang(1989) developed an automated fault diagnosis of ball bearings using pattern recognition system. After 1990s, Artificial neural networks are developed. The machine learning has 3 main steps feature extraction, feature selection and feature classification. Saimurugan et al.,(2011) reported the effectiveness of decision tree algorithm and support vector machine in multi component fault diagnosis. In 2005, Lia et al used hidden Markov model for fault diagnosis of rotating machine. W.J.Wang(2012) integrated wavelet transform and spectral analysis along with enveloping and used them in machine learning approach. Acquired signals are represented in the form of statistical features, histogram and wavelet features. Samanta and Al-Baulshi(2003) used decision tree for selection of prominent features. Saravanathan(2009) used decision tree algorithm for selecting the prominent features and performed the classification process for identification of fault in spur bevel gearbox. In 2004 I.Kuncheva stated that when number of fault classes gets increases classification accuracy gets reduced. In sensor fusion, H.F Darren(1988) white described three models of multisensor fusion and their effectiveness in the application of sensors used in robots. B.V Dasarathath defined various types of data fusion techniques. G.T Mckee defined three types of sensor fusion (direct, indirect and fusion of them both). Dasarathath(1997) proved that the effectiveness of sensor fusion algorithm increases with addition of sensors. R. R. Brooks,S. S. Iyengar(1998) gave the facts on complementary data fusion about its easiness and reduced complexity . In 1969 Fraser and potter derived an equation which uses that Kalman filters in both front and backward direction which increases the weight of the variance estimates. Garry A.Finicke(2012) mentioned smoother equation can be used for processing of sound signals instead of using normal filters.

Revised Manuscript Received on August 20, 2019

Dr. Saimurugan M Associate Professor at Department of Mechanical Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore.

Meenakshi Sundaram Bachelor’s degree in Mechanical Engineering and a Master’s degree in Engineering design.
L. Kuncheva, J.C. Bezdek (2001) defined classifier fusion consists of two stages, classifier selection and classifier combination. Many classifiers can be used in selection but only one classifier can be used for decision making. R. Duin and D. Tax (1999) defined various combination methods for weak classifiers and proved its effectiveness. L. Xu suggested a method by grouping the classifier outputs in their journal. Chao Wang (2018) applied Majority voting in financial markets. Further, Gavin brown classified majority voting into 3 parts individual accuracy, good and bad diversity.

Kearn and Valiant initially raised a question about a performance of weak learning algorithm par with strong learning algorithm. Schapire developed the polynomial-time boosting algorithm in 1989. Drucker, Schapire and Simard conducted experiments on boosting algorithms which was developed by Freund (1990). The Adaboost algorithm, introduced in 1995 by Freund and Schapire, solved many of the practical difficulties of the earlier boosting algorithms. Most of the research work done in this area considered one or two components with small number of fault classes. The critical components of the rotating machine are shaft, bearing, rotor and gear. These four components with twenty-four fault combinations are extensively studied in this work. Vibration and sound signals are extracted for these 24 fault classes. Data fusion techniques are performed using decision tree.

II. EXPERIMENTAL STUDIES

2.1 Experimental setup: Fault diagnosis is carried out using Machinery fault simulator with optimal conditions. It is the variable speed drive machine with rpm range from 220 to 1440. It consists of shaft with rotor where conditions like balancing and unbalancing of shaft can be done, rolling element bearings where the fault is simulated, belt drive, and simple spur bevel gearbox. Control knob is used to vary the speed of the motor. The primary components are attached to the aluminium working table for reduction of vibration. In the control panel, digital displays are being provided to display the speed, bearing temperature and motor current.

![Machinery故障模拟器](image)

**Fig 1-Machinery Fault simulator**

Data acquisition system is defined as the process of acquiring physical parameters such as temperature, acceleration, voltage that are being measured in digital signals. The components in the DAQ are

- Sensors
- DAQ device
- Computer with DAQ software

The acceleration is one of the characteristics of vibration are acquired using the piezoelectric accelerometer. A piezoelectric accelerometer is fixed on top of the bearing housing are used to collect acceleration values. The accelerometer sends the values to the computer with the help of data acquisition system. Similarly, a microphone is used for collection of sound signals from bearing. The vibration and sound signals are stored in the computer and the .wav file of sound signals are converted into digital data with the help of MATLAB.

2.2 Experimental procedure:

Two shaft conditions (good and bent), 2 rotor conditions (good and unbalanced), 3 bearing conditions (good, Inner race fault (IRF) and outer race fault (ORF)) and 2 gear conditions (good and broken) are considered in this work and combination of these gives 24 fault classes. SKF6206 ball bearing is used in this study. Bearing faults are created using Electrical discharge machining (EDM). Cut is made with the depth of 0.45mm in outer race and depth of .35mm for inner race fault. Before starting the procedure connections between the sensor and the DAQ device and the computer should be checked. The values obtained from the DAQ device are default in time domain format. Experiments are conducted for 24 fault classes and three different speeds. The speed is adjusted by the knob which is provided on the control panel. Once the speed is set, the model is created in LabVIEW software to acquire signals.

The statistical analysis of acquired vibration signals for various fault conditions gives statistical parameters. These parameters bring out the information from the time domain signals of various faults. These parameters are called statistical features which are used for detection of faults. The statistical features yield better classification accuracy in fault diagnosis of rolling element bearings (Kankar et al., 2011). The statistical features used in this study are mean, standard deviation, median, standard deviation, sample variance, kurtosis, skewness, range, minimum, maximum and sum. These statistical features are explained below.

- Mean: It is defined as the average of the data.
- Variance: It is defined as the measure of each value from the mean value and variation of each point.
- Standard deviation: It is used to measure the number of variations between data points and is calculated by a square root of the variance.
- Mode: It is the measure of value in a dataset with high no of occurrences and repetitions.
- Median: It is defined as the measure of middle value in the given data.
- Range: It is the difference between the largest value and the smallest value in data.
- Kurtosis: It is the measure of the heaviness of given distribution. Whether they are heavily tailed or lightly tailed relative to the normal distribution.
- Skewness: It is the measure of the lack of symmetry or unbalances in the given signal.
- Maximum: It is the measure of the maximum value in the signal.
• Minimum: It is the measure of the minimum value in the signal.

These ten features were extracted for both the signals. You can send paper in the given email address of the journal. There are two email addresses. It is compulsory to send paper in both email addresses.

III. DECISION TREE ALGORITHM

The acquired huge vibration and sound data cannot be given as an input to machine learning algorithm. They are reduced into features and given as input to machine learning algorithm. So the prominent features are selected using decision tree algorithm. Decision tree uses tree like model to represent an information with leaves representing the various fault classes. Here J48 algorithm was used which has roots, branches, nodes and leaves. A chain of nodes start from roots to leaves are known as branches and each node are associated with attribute. The class variables are expressed as leaf nodes. Entropy reduction and information gain are used for selection of optimum features. The C4.5 Quinlan (1993) is the most commonly used algorithm for decision tree. The prominent features are the one that increases the classification accuracy. They are shown in top of the decision tree.

The features are arranged in the descending order of importance in the decision tree. The best features are selected from the first layer.

There are two ways to select the number of dominant features for classification study. They are

• Choose the number of features which maximizes classification efficiency
• Choose the number which ensures higher enough classification efficiency and satisfies the consequence of dimensionality reduction.

Fig 2- Decision tree for 500rpm

The features are arranged in the descending order of importance in the decision tree. The best features are selected from the first layer.

There are two ways to select the number of dominant features for classification study. They are

• Choose the number of features which maximizes classification efficiency
• Choose the number which ensures higher enough classification efficiency and satisfies the consequence of dimensionality reduction.

IV. DATA FUSION

If you are Data fusion is defined as the process of combination of multiple data sources to produce more consistent, accurate, and useful information than that provided by any individual data source. It is mainly done to produce more accurate and informative data for machine learning approach. There are various types of data fusion techniques available.

4.1 Feature fusion

Feature fusion is defined as the combination of features from various signals to form new set of features that is more informative. Set of features can be fully combined or partially combined. For example, statistical features from the time domain data are combined with wavelet features or cepstrum features in order to produce an improved model for the estimation of fault. Feature fusion gets the discriminatory information in the fusion process and eliminates unnecessary information which causes misclassification. Feature fusion is classified into serial fusion and parallel fusion.

Serial fusion is done by integrating the entire set of features from various vibration analysis techniques into a single set of data and then using it for classification purpose. Here features from various techniques are joined together and not reduced. T.Praveenkumar et al., (2019) detailed the effectiveness of wavelet features fusion in gearbox fault diagnosis. Feature fusion is classified into serial fusion and parallel fusion.

Parallel fusion is defined as the combination of a set of selective features from various vibration analysis techniques. Features are selected based on an accuracy of the individual analysis is given by the classifier.

4.2 Sensor fusion

It is defined as the process of combining information from multiple sensors to enhance the performance of the system. K. E. Foote and D. J. Rueben (1995) specified that sensor fusion is mainly used in order to overcome an unpredictability of the single sensor. N.S.V Rao (2001) compared the effectiveness of measurements of fused sensor and single sensor and concluded that system complexity gets reduced in sensor fusion. The combination of multiple sensors increases the strength of one type can compensate for the weakness of other types of sensor. Examples of sensor types are radar, thermal, acoustic, laser, optical sensors, accelerometers, seismic sensors, sonar, magnetic sensors, and chemical detection devices etc. There are various types of sensor fusion. Complementary fusion is the most widely used technique. Complementary level fusion and Fraser potter smoother interval equation are used in this work.

4.2.1 Complementary fusion

A. Hoover and B. D. Olsen (1999) performed complementary level fusion on multiple cameras used to capture images. It is one of the widely used raw level data fusion techniques. This method can be used when more than two different sensors are used. It is the summation of data from various signals to form more accurate raw level data. The signal from sensor one
and the signal from the sensor two are added together to obtain the combined signal. It is mainly used to overcome the uncertainty of any of signals obtained from the individual sensor and to produce more accurate data. Complementary fusion method can only be applied to sensors where the sampling rate of both the sensors are the same.

4.2.2 Fraser potter equation

Fraser potter equation is also called as smoother interval equation. In 1969 Fraser and potter derived an equation which uses that Kalman filters in both front and backward direction which increases the weight of the variance estimates. Garry A.Finicke(2012) used smoother used for processing of sound signals instead of normal filters. It is the modified method of complementary level fusion.

\[ x_3 = \sigma_3^2 (x_1^2 + x_2^2) \]
\[ \sigma_3^2 = 1 (\sigma_1^2 + \sigma_2^2) \]

- \( x_1 \) = measurement from sensor 1.
- \( x_2 \) = measurement from sensor 2.
- \( x_3 \) = combined signal
- \( \sigma_1^2 \) = Variance of \( x_1 \)
- \( \sigma_2^2 \) = Variance of \( x_2 \)
- \( \sigma_3^2 \) = Variance of combined estimate \( \sigma_2^2 \) & \( \sigma_3^2 \).

Unlike complementary level fusion it combines data by using variances of signals from individual sensors. This method can be used to combined signals only with same sampling rate. Then features are obtained from the new signal obtained from this equation and is further sent for classification into the machine learning algorithm.

4.3 Classifier fusion

4.3.1 Majority voting

Majority voting is one of the most commonly used fusion technique. While selecting the classifier fusion two important criteria considered are a selection of classifiers and a combination of classifiers. Majority voting is used in the area of financial markets as its effectiveness is very high. Multiple classifiers are used to create various sub models in majority voting. Generally used classifiers in majority voting are decision tree, Support vector machine, Naïve Bayes, Random forest etc. Various classifiers are trained with the given data. The classifier with the most number of predicted classes is predicted as output by majority voting. The classifier outputs can be combined with other combination methods like an average of probabilities, min, maximum etc. Decision tree and support vector machine is used here. Majority voting is done individually for both vibration and sound signal.

4.3.2 Adaboost algorithm

Unlike majority voting which uses multiple classifiers, Adaboost uses the only single classifier. The performance of Adaboost algorithm has been tested by many researchers. Freund performed Adaboost on UCI benchmark datasets with decision tree algorithm. It uses method called boosting, which uses multiple models and arrange them in a series where the prediction errors made out by the previous model gets corrected to a maximum extent. Most commonly used classifier in Adaboost is a decision tree. Initially, it creates a normal model on the given dataset. Each model is weighted based on its misclassification and accuracy. The weights get updated based on the accuracy of the model on every iteration. The model with high misclassification is given high weight when compared to others. The preset models of weak learners are created on a certain iteration. Finally, the model which further can’t reduce the error rate is concluded as classifier output.

V. RESULTS AND DISCUSSIONS

The fused data is further classified using decision tree which gives confusion matrix as an output. The classification accuracy for various data fusion techniques like feature fusion techniques such as parallel fusion and serial fusion, sensor fusion techniques such as complementary level fusion and Fraser potter equation and classifier fusion techniques such as majority voting and Adaboost algorithm are shown in table 1 and 2.

Table 1- mean classification accuracy for individual signals and Feature fusion

<table>
<thead>
<tr>
<th>Speed (rpm)</th>
<th>Classification accuracy in %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Individual Signals</td>
</tr>
<tr>
<td>Vibration</td>
<td>Sound</td>
</tr>
<tr>
<td>500</td>
<td>83</td>
</tr>
<tr>
<td>700</td>
<td>84</td>
</tr>
<tr>
<td>900</td>
<td>84.40</td>
</tr>
<tr>
<td>Mean</td>
<td>83.38</td>
</tr>
</tbody>
</table>

The mean classification accuracies for individual vibration signals perform better than sound signals. The sound signal result indicates a worst performance of less than 50%. Data fusion techniques improve the classification accuracy compared to individual signals. In feature fusion, serial fusion yielding higher classification accuracy than parallel fusion.

Table 2- mean classification accuracy for sensor fusion and classifier fusion

<table>
<thead>
<tr>
<th>Speed (rpm)</th>
<th>Classification accuracy in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensor Fusion</td>
<td>Classifier fusion</td>
</tr>
<tr>
<td>CLF</td>
<td>Fraser Eqn</td>
</tr>
<tr>
<td>500</td>
<td>88.80</td>
</tr>
<tr>
<td>700</td>
<td>86.70</td>
</tr>
<tr>
<td>900</td>
<td>87</td>
</tr>
<tr>
<td>Mean</td>
<td>87.50</td>
</tr>
</tbody>
</table>

For sensor fusion, the mean classification accuracy for complementary level fusion is at 87.5% and for Fraser potter equation is at 80%. Complementary level fusion performs better than Fraser potter equation in sensor fusion. In classifier fusion, majority voting and Adaboost

Retrieval Number F8345088619/2019/BEIESP
DOI: 10.35940/ijeat.F8345.088619
Published By: Blue Eyes Intelligence Engineering & Sciences Publication
algorithm are performed individually for both vibration and sound signals. In majority voting the classification accuracy for vibration signal is 84.1% and for sound signal is 44%. And for Adaboost algorithm of vibration signal is 86.6% and for sound signal is 46.7%. Adaboost can be used for better classification but limited only for vibration signals. Fused data shows better result than the individual signals and it is efficient in fault diagnosis approach.

In the various fusion techniques, the performance of serial fusion and the complementary fusion is satisfactory. For further improvement of classification accuracy, the feature sets of serial fusion and complementary fusion feature sets are given as an input to adaboost algorithm. Table 2 shows the comparison of classification accuracy of serial fusion and complementary fusion with adaboost algorithm. Serial fusion performed with Adaboost has mean classification accuracy of 91% which is nearly 10% more than individual vibration signal. Complementary fusion performed using adaboost algorithm is having mean classification accuracy of 90% which is nearly 7% higher than individual vibration signal. The serial fusion with a adaboost algorithm is best suitable combination for automated fault diagnosis of rotating machines. In the various fusion techniques, the performance of serial fusion and the complementary fusion is satisfactory. For further improvement of classification accuracy, the feature sets of serial fusion and complementary fusion feature sets are given as an input to adaboost algorithm.

### Table 3-Classification accuracy for serial fusion and complementary fusion performed with Adaboost algorithm

<table>
<thead>
<tr>
<th>Speed(rpm)</th>
<th>Vibration signals</th>
<th>Serial fusion with adaboost</th>
<th>Complementary fusion with adaboost</th>
</tr>
</thead>
<tbody>
<tr>
<td>500</td>
<td>83%</td>
<td>94.458%</td>
<td>91%</td>
</tr>
<tr>
<td>750</td>
<td>84%</td>
<td>89.6%</td>
<td>88%</td>
</tr>
<tr>
<td>1000</td>
<td>84.4%</td>
<td>90%</td>
<td>90.5%</td>
</tr>
<tr>
<td>Mean</td>
<td>83.38%</td>
<td>91%</td>
<td>90%</td>
</tr>
</tbody>
</table>

Table 3 shows the comparison of classification accuracy of serial fusion and complementary fusion with adaboost algorithm. Serial fusion performed with Adaboost has mean classification accuracy of 91% which is nearly 10% more than individual vibration signal. Complementary fusion performed using adaboost algorithm is having mean classification accuracy of 90% which is nearly 7% higher than individual vibration signal. The serial fusion with a adaboost algorithm is best suitable combination for automated fault diagnosis of rotating machines.

### VI. CONCLUSION

Experiments are conducted for 24 fault classes including bearing faults, shaft faults, rotor faults and gear faults. The statistical features are obtained from both vibration and sound signals. The obtained statistical features are sent to decision tree algorithm for fault classification. The individual performance of both vibration signals and sound signals are compared. To improve the classification accuracy, data fusion techniques are performed for both the signals. In feature fusion, serial fusion yields better classification accuracy than parallel fusion technique. In sensor fusion, classification accuracy of complementary level fusion is better than Fraser potter equation. And in classifier fusion, adaboost algorithm which fuses the classification models of single classifier performs better than majority voting which the various classifier models. For further improvement of classification accuracy, combination of data fusion techniques is done. The best fusion techniques, serial fusion and complementary fusion feature sets are classified using adaboost algorithm for better improvement of classification accuracy. Both the fusion feature sets yield a classification accuracy of more than 90%. It shows the advantages of adaboost algorithm. Serial fusion with adaboost algorithm is the best method for automation of rotating machine fault diagnosis.

### REFERENCES

Performance Improvement of Classifier in Fault Diagnosis of Rotating Machines Using Sensor Fusion Techniques

34. Quinlan, J. R., C4. 5: programs for machine learning(Vol. 1). Morgan Kaufmann, 1993

AUTHORS PROFILE

Dr. Saimurugan M currently serves as Associate Professor at Department of Mechanical Engineering, Amrita School of Engineering, Amrita Vishwa Vidyapeetham, Coimbatore. He completed his Ph.D. in the field of Machine Condition Monitoring and published around 25 research papers. His areas of research include Vibration Analysis, Machine Learning and Machine Condition Monitoring.

Meenakshi Sundaram S holds a Bachelor’s degree in Mechanical Engineering and a Master’s degree in Engineering design. He is an active researcher in the area of machine learning. His research includes use of various data fusion techniques to improve the performance of individual machine learning algorithms.