

Research Article

Uncertainty Reduced Novelty Detection Approach Applied to Rotating Machinery for Condition Monitoring

S. Ma,^{1,2} S. M. Li,¹ and Y. P. Xiong²

¹College of Energy and Power Engineering, Nanjing University of Aeronautics and Astronautics, Nan Jing 210016, China

²Faculty of Engineering and the Environment, University of Southampton, Southampton SO16 7QF, UK

Correspondence should be addressed to S. Ma; sm13g14@soton.ac.uk

Received 7 May 2015; Revised 25 July 2015; Accepted 27 July 2015

Academic Editor: Mickaël Lallart

Copyright © 2015 S. Ma et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Novelty detection has been developed into a state-of-the-art technique to detect abnormal behavior and trigger alarm for in-field machine maintenance. With built-up models of normality, it has been widely applied to several situations with normal supervising dataset such as shaft rotating speed and component temperature available meanwhile in the absence of fault information. However, the research about vibration transmission based novelty detection remains unnoticed until recently. In this paper, vibration transmission measurement on rotor is performed; based on extreme value distributions, thresholds for novelty detection are calculated. In order to further decrease the false alarm rate, both measurement and segmentation uncertainty are considered, as they may affect threshold value and detection correctness heavily. Feasible reduction strategies are proposed and discussed. It is found that the associated multifractal coefficient and Kullback-Leibler Divergence operate well in the uncertainty reduction process. As shown by in situ applications to abnormal rotor with pedestal looseness, it is demonstrated that the abnormal states are detected. The higher specificity value proves the effectiveness of proposed uncertainty reduction method. This paper shows novel achievements of uncertainty reduced novelty detection applied to vibration signal in dynamical system and also sheds lights on its utilization in the field of health monitoring of rotating machinery.

1. Introduction

Novel behaviors of a dynamic system illustrate transition in its running states and overall performance, but to detect the novelty and alarm instantly when fault happens might be a tricky task. In traditional fault diagnosis or health monitoring research, it is assumed that the supervising datasets are “balanced,” which means not only normal data can be acquired, but also the data generated from fault or malfunction in dynamic system. Based on this assumption, soft computing tools such as artificial neural network (ANN) and Support Vector Machine (SVM) are well trained with both normal and abnormal data to separate out features representing abnormality in sample set. Ever since, this series of diagnosis methods has become a powerful method for offline application and proved its value for machine maintenance [1]. However, in many industrial situations, either high risk of abnormal running states or rather limited period of fault makes the abnormal data unavailable and impossible to obtain well

trained model for diagnosis. Under this circumstance, novelty detection method is developed to avoid the deficiencies of “unbalanced” dataset in addressing abnormality detection problems. Novelty, or in some research called outlier, is defined as novel behaviour in feature space which is extracted from supervising dataset. Based on this definition, two types of models are widely used: static and dynamic model. For a static model, time domain supervising data like rotating speed or temperature will be transformed into another feature domain and novelty is detected whenever it gets across the preset threshold; this causes a delay in online monitoring time and may lead to a biased detection rate. On the contrary, with dynamic models, supervising data is compared with a time-evolving threshold in real time and alarm could be triggered without time delay [2].

It is well known that soft computing techniques could be modified into detectors in both static and dynamic applications and their performance mainly depends on the quality of training datasets. To be specific, as one of the pioneer

algorithms, ANN has been proved useful and reported in [1–9]. Bishop first presented the topic of ANN based novelty detection [1]. In his research, the “relationship between the novelty of input data and validity of network outputs” is discussed according to probabilistic model. Based on this, those events with lower probability are defined as novelty. Addison et al. made an argument in their article [2] that “a synthesis of various neural networks with linear regression provides comparable performance,” and the comparison between test results shows that this feature reduction process ensures best use of the training samples. Hwang and Cho analysed the output characteristics of trained auto-associative MLP (multilayer perceptron) in [3] and concluded that this algorithm is a reliable solution for novelty detection. However, in their research, ambiguity in the input layer is found to be inevitable for AOMLP based novelty detection. Another important form of ANN algorithm called radial basis function network is discussed by Albrecht et al. [4]. A Bayesian classifier is developed by Bayesian learning to fulfil the task of novelty detection. With this learning process, quality of classifier is promoted and a better performance is acquired. Brotherton and Johnson [5] also combine the distance based metric with RBF NN classifier to perform novelty detection. In review article [6] and recent research reported by Barreto and Frola [7], almost all neural network paradigms based novelty detection methods are evaluated. Furthermore, the sensitivity of ANN training parameters is assessed in both supervised and unsupervised cases. More interestingly, a handling strategy as labelling novelty for further classification purposes is proposed in [8]. Above all, although ANN with its derivative algorithms has been demonstrated to be valuable and frequently reported in novelty detection applications, due to the deficiency of training data and time consumption requirements, it may give unstable results in real time tasks and this makes usage of this methodology discounted.

The development of soft computing based novelty detection approaches arises from the need of intuitive explanation for detection results. Support Vector Machine (SVM) algorithm is among the best for addressing the problem of ambiguous interpretation. Support Vector Data Description (SVDD) as well as series of homogenous algorithms is proposed for novelty detection successively [10–13]. In Tax and Duin's paper [10], SVDD is found to be capable of dealing with “sparse and complex datasets” and according to the claim of Schölkopf et al. [11], this algorithm is proved to be well behaved in his research. Lian made a combination of SVM and PCA to extract features in positive dataset [12], which is an enlightening insight for others to further apply this technique in more specific areas. Guo et al. [13] proposed a bound updating method so that the overfitting caused false negative rate could be decreased when processing real datasets. For other static novelty detection applications such as nearest neighbour-based and clustering-based methods, one can refer to [14–18]. But just as Wang et al. mentioned [19]: adaptability of a novelty detector is an important ability which offers the dynamic model based method a better real time performance during condition monitoring. With a dynamic model, the novelty detected in monitoring process could be adopted as reference for further training. Compared to

the static model, by reporting the occurrence of novel events or data to bring negative samples into this process, it is actually the transition from unsupervised learning to supervised learning hierarchy. As stated earlier in this paper, it is our expectation to develop a dynamical novelty detection which could be applied to raw monitoring data (accelerations) simultaneously overcoming the defect of time delay and fixed novelty threshold.

Since the use of these aforementioned soft computing techniques applied to dynamic novelty detection remains a challenge for their incompetence to deal with the dimension curse, computation complexity or rather the unbalanced dataset, statistical methods are believed to be competent to give a reliable solution to the time evolving novelty detection meanwhile circumventing this drawback. Statistical novelty detection or, more specifically, the extreme value statistical novelty detection will be detailedly discussed in this paper to provide a foundation for this significant task. In previously reported research [20], Ntalampiras et al. performed acoustic novelty detection with selected statistical features and statistical modelling techniques. Filippone and Sanguinetti used KL Divergence as the metric for novelty decision to reduce false alarm rate in small sample cases [21]. Breaban and Luchian proposed a feature extraction method based on projection pursuit algorithm to detect outlier with the novel metric given by kurtosis [22]. Clifton et al. proposed an extreme value statistics based novelty detection framework [23–29] which is applied to gas-turbine engine monitoring, biomedical signal processing, and so on. For its ability to overcome the flip-side of aforementioned methods, this method will be used here to detect rotor faults with time series of raw vibration data incorporating measurement and segmentation uncertainty during statistical modelling period.

This paper is organised as follows. A vibration transmission measurement for supervising data collection is presented in Section 2. Section 3 discusses the extreme value theory and corresponding novelty threshold setting rules. The measurement and segmentation uncertainty reduction strategy with their applications and discussions about detection rate are presented in Sections 4 and 5, respectively. Finally, we present our conclusions in Section 6.

2. Vibration Transmission Measurement

Surface vibration transmission measurement is usually used to define and calculate the transfer function to further obtain a modal analysis result in frequency domain. However, more importantly, it is also a simplified but operative way to implement energy flow analysis to monitoring or diagnosis tasks especially for rotating systems. For energy flow analysis framework, the basic governing equations are given in [30–32]. In most of previous research, information about the vicinity of the damage or fault is assumed to be known in advance, which is not actual for many cases. To avoid this drawback, model analysis based approaches have been proposed, but most of them are applied to finite element analysis simulation with regular structures [33]. For rotating systems, one point should be noticed is that, with torque transmitted interiorly, it is unrealistic to measure surface radial velocity

by sensors attached to shaft or pedestal. This makes the power flow calculation infeasible. Hence, in our research, a relative surface vibration transmission is measured; it is worth noting that it is not the vibration transmissibility defined as ratio of two measurement points. The relative surface vibration transmission is given by the following simple equation:

$$A_{\text{rel}} = A_{\text{abs}} - A_{\text{ref}}, \quad (1)$$

where “rel,” “abs,” and “ref” stand for relative value, absolute value, and reference acceleration value, respectively.

In our novelty detection experiment, reference measurement point is located to vibration source, which can be seen in Figure 4 near the electric motor. The absolute measurement point could be selected arbitrarily, not required to be near the fault position shown in Figure 4. The main idea of this test is energy loss effect derived from structure fault. As structural parameters of systems under vibration could be used as dictator of its running states, similarly, for a rotor system with inner fault, the physical quantity such as transmitted acceleration will be transformed into another form (thermal or acoustic), thus causing vibrational energy loss compared to a normal rotor.

3. Extreme Value Theory Based Novelty Threshold Setting

Extreme value theory has been widely used in many areas such as risk management and reliability estimation. The novel idea of this statistical modelling method is about using the maximum and minimum value sets extracted from original time series to obtain more accurate density models which is tricky for traditional statistical models. This surely makes difference in prediction results and it has been shown that, for novelty detection applications, due to its high dependence on the threshold value as well as its corresponding setting strategy, a more accurate model ensures a satisfying detection rate especially when in situ measurement uncertainty during the whole process is unavoidable. Also, for vibration data of rotating machines, it is found that extreme value distribution is applicable and more appropriate for novelty detection as the threshold value will rise, thus decreasing the false alarm rate compared to threshold setting strategy based on Gaussian distribution.

It has been discovered by Fisher and Tippett [23, 34] that the limit forms of the probability of observed extremum are considered to be one of the three following distributions:

$$\begin{aligned} \text{Gumbel: } f_G(y) &= \exp(-\exp(-y)), \\ \text{Fréchet: } f_F(y) &= \begin{cases} \exp(-y^{-\alpha}) & \text{if } y > 0 \\ 0 & \text{if } y \leq 0, \end{cases} \\ \text{Weibull: } f_W(y) &= \begin{cases} 1 & \text{if } y > 0 \\ \exp(-(-y)^{-\alpha}) & \text{if } y \leq 0, \end{cases} \end{aligned} \quad (2)$$

where y is named as reduced variate, which is

$$y = \frac{x - c_m}{d_m}. \quad (3)$$

In (3), the so-called location and scale parameters are given in

$$\begin{aligned} c_m &= \sqrt{2 \ln m} - \frac{\ln \ln m + \ln 4\pi}{2\sqrt{2 \ln m}}, \\ d_m &= \frac{1}{\sqrt{2 \ln m}}, \end{aligned} \quad (4)$$

where m stands for the number of extremum data; $\alpha \in \mathbb{R}^+$.

These three distributions may also be regarded as special cases of the generalized extreme-value (GEV) distribution as is shown by (3) succinctly:

$$f_{\text{GE}}(y) = \begin{cases} \exp(-(1 + \xi y)^{-1/\xi}) & \text{if } \xi \neq 0 \\ \exp(-\exp(-y)) & \text{if } \xi = 0, \end{cases} \quad (5)$$

where ξ is a shape parameter. The cases $\xi < 0$, $\xi = 0$, and $\xi > 0$ give the Gumbel, Fréchet, and Weibull distributions.

In our research, the extreme value distribution of raw acceleration data is fitted. According to the formulation of EVD given by (2) and the definition equation of y , the uncertainty of modelling result is mainly generated from measurement and segmentation process. Namely, measurement uncertainty or, say, noise becomes influential, because extreme values extracted from original data will be biased severely and this impairs the corresponding statistical distribution accuracy. On the other hand, segmentation uncertainty, as shown by its name, comes from the segmenting process for separating acquired time series into pieces before extracting extreme values from every piece of the datasets. As it is admittedly difficult to avoid these two types of uncertainty and meanwhile their negative influence on statistical distributions further hampers the calculation of a desired novelty threshold, effective uncertainty reduction strategy will be proposed in Section 4; in other words, it is to answer two questions:

- (1) How many metadata in the original time series should be clustered into a segment for following extreme value extraction?
- (2) Which of all the segments should be used for statistical modelling? This substantially determines which extreme values could be used to fit the probabilistic density in (2).

As it is fairly important to set a threshold for novelty detection, three main setting strategies are taken into account before the next step; they are (1) cumulative function, (2) inequality [35], and (3) quantile value based methods. Comparative result shows the difference between those three could be neglected when parameters adjusted properly so that the first one is chosen in this research because it seems that more applications are accomplished and give convincing detection rate with the norm of “cumulative value exceeds 0.99.” Based on this, the determination of segmentation number is becoming even more important in some real cases, when the cumulative threshold could be easily exceeded with the segmentation number increases [36]. To avoid this shortcoming, a Kullback-Leibler Divergence metric is resorted to establish stable extreme value statistical distribution.

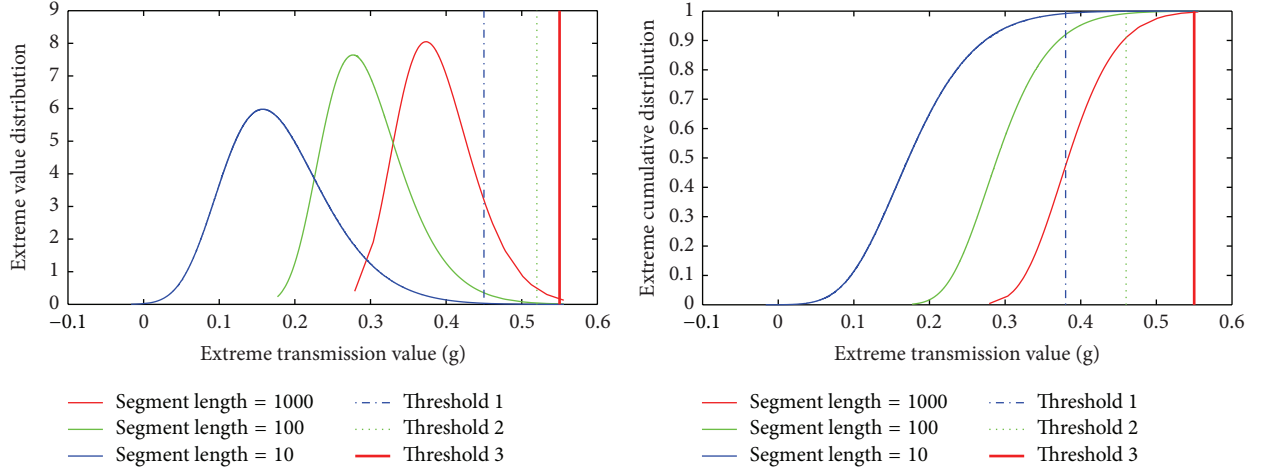


FIGURE 1: Extreme value distribution with relative threshold affected by segment length variation.

4. Segmentation and Measurement Uncertainty Reduction

Several forms of measurement uncertainty have been observed by engineers, such as poor sampling or noisy measurement [37, 38]. It should be emphasized that because extreme value sets are inclined to be affected by measurement uncertainty, this entire distribution fitting procedure highly relies on the quality of the raw acceleration data. As modelling process starts in extracting extreme value (maximum and minimum value) in every segment, here, firstly, segmentation uncertainty is discussed in the following part as well as its reduction strategy.

4.1. Segmentation Uncertainty Reduction. In our extreme value analysis based novelty detection procedure, for an acceleration time serial, the first step is to segment it into pieces. The maximum and minimum value of each segment will be extracted and then fitted into extreme value distribution function for the following threshold setting. Obviously the segmentation strategy, or to be more specific the segment length, is a key factor. In Figure 1, one of the datasets collected from rotor in good condition is processed for exemplification; three strategies are adopted to segment it into pieces with different length (1000, 100, and 10 points). Three extreme value distribution curves are fitted and the corresponding thresholds are calculated (threshold and corresponding distribution are in the same colour as legend shows). As the figure shows, with the segment length increases from 10 to 1000, the novelty threshold moves from 0.45 to 0.55. This leads us to a dilemma for if a new data sample falls into the zone between 0.45 and 0.55, it could be considered as novelty according to threshold 1, and, however, it could also be considered normal according to threshold 3. This could cause an increase on false alarm rate when we choose to use threshold 1 generated from segmentation strategy 1 (segment length = 10).

It is illustrated that the segment length affects the extreme value distribution, thus causing severe disruption to threshold setting. And because of this disruption, the norm

“cumulative probability exceeds 0.99” becomes meaningless for novelty detection as the false alarm rate becomes so high and unacceptable.

A multifractal stationarity based segmentation strategy has been proposed in [39]. Multifractal stationarity is supposed to be dependable to give logical segmentation results based on the self-similarity characteristic; it is a metric to evaluate complexity of local variation, for an acceleration segment with multifractal coefficient value almost equals the data entirety; it is considered as element of the original dataset. And the data entirety is composed of these segments. Accordingly, the extracted local maximum and minimum are naturally representative for extreme distribution fitting.

The multifractal coefficient is defined as (6) shows:

$$f(\alpha(q)) = q\alpha(q) - D_q(q-1), \quad (6)$$

where parameter q is called moment of order q , α is the Lipchitz-Holder exponent, and D_q is the generalized dimensions [40].

The corresponding segmentation algorithm is given as follows:

```

Fix the global average multi-fractal value  $m_0$ ;
Initialise Segment Length =  $L_i$  ( $i = 1, \dots, N$ );
Function (length)
Begin
Original time serial is segmented into pieces with
length  $L_i$ ;
Calculate multi-fractal value of each piece;
Calculate the average multi-fractal value  $m_i$ ;
Set a small value  $\varepsilon > 0$ ;
If  $|m_i - m_0| \leq \varepsilon$  then:
Final segment length is  $L_i$ 
Else
Segment length =  $L_{i+1}$ ;

```


Do Function (length);
End

To run this code, the global average multifractal value m_0 is calculated previously; it is 80.61. The final result shows “segment length = 50” and corresponding m_i is 80.58. In this part, original acceleration data is segmented into pieces with every 50 points; the result data pieces obtained from this periodic clustering method will be handled for extremum extraction in the following part. At the same time, an imperative measurement uncertainty reduction is incorporated to assure that the distribution used in the analysis is stable.

4.2. Measurement Uncertainty Reduction. As an unavoidable phenomenon, measurement uncertainty affects the extreme distribution through biased maximum and minimum value extracted from every piece of acceleration data. When measured data suffers from random noise or absolute drifting, it is quite necessary to determine the number of extremum for distribution fitting, as, with metadata number increasing, the distribution becomes almost invariant and the uncertainty is reduced. During the application of this approach, impulse noise is frequent in measurement; it is most likely generated from sensor failure and will heavily affect the extreme distribution as maximum and minimum value will always be the impulse amplitude. However, this is not expected for the distribution fitted for novelty detection. As the amplitude of impulse is much bigger than that of the normal data, calculated thresholds will be exceptionally deviated that all sample data are claimed normal in this situation. Based on this observation, impulse noise should be cleaned out of the original dataset.

The definition of Kullback-Leibler Divergence, or Relative Entropy, is given in

$$D(P \parallel Q) = \sum_{x \in X} P(x) \log \frac{P(x)}{Q(x)}, \quad (7)$$

where $P(x)$ and $Q(x)$ are candidate probability distribution. In our research, they are extreme distribution $f_i(x)$ and $f_{i+1}(x)$, $i \in [100, 200, \dots, N]$.

An effective method is utilized in this research to prevent impulse noise from ruining threshold: Kullback-Leibler Divergence is calculated between a dataset and the updated one. Here probability distribution P stands for the already obtained one fitted from extreme data in last step. The probability distribution Q stands for the updated one fitted with extreme data from both last step and new sample. In this process, appending new sample to existing dataset could lead to discrepancy between the two distributions. For mitigating influence of measurement uncertainty, two rules are set for obtaining a reliable extreme distribution for novelty detection:

- (1) When Kullback-Leibler Divergence value exceeds upper threshold, impulse noise is considered detected and appended dataset will be discarded. After deleting segment including impulse noise, the iteration will move on.

- (2) When Kullback-Leibler Divergence value exceeds lower threshold, it is believed the reliable distribution has been obtained under the law of large numbers.

In order to clearly show every step of this iteration, the pseudocode is given as follows:

```

Fix the Kullback Leibler Divergence threshold value  $D_0$ ;
Initialise Length of extremum data =  $L_i$  ( $i = 1, \dots, N$ );
Function (length)
Begin
Original extremum serial is modelled into distribution  $P_i$ ;
Update Length of extremum data =  $L_{i+1} = L_i + \Delta L$ ;
Model the lengthened extremum data into distribution  $P_{i+1}$ ;
Calculate the Kullback Leibler Divergence value  $D_i$  of  $P_i$  and  $P_{i+1}$ ;
Set small values  $\varepsilon_1$  (lower threshold) and  $\varepsilon_2$  (upper threshold)  $> 0$  ( $\varepsilon_2 > \varepsilon_1$ );
If  $|D_i - D_0| \geq \varepsilon_2$  then:
Discard this new dataset  $\Delta L$  ( $L_{i+1} = L_i$ )
Else
If  $|D_i - D_0| \leq \varepsilon_1$  then:
Final length for modelling is  $L_i$ 
Else
Length of extremum data =  $L_{i+1}$ ;
Do Function (length);
End

```

In Figures 2 and 3, distribution curves of both maximum and minimum are plotted. Red line and green line represent the first two distributions with lengths $L_1 = 100$ and $L_2 = 200$, respectively. It can be clearly observed that, with data amount increases, distribution functions intend to converge to a stable one as shown by those blue curves.

After both segmentation and measurement uncertainty are effectively controlled, optimal segment length is guaranteed. These two adopted strategies dispose of the shortcomings of classical extreme value distribution based novelty detection method. More importantly, this preparatory work is finished for only one purpose: to give an optimal novelty threshold. The title “optimal” is actually entitled for a low False Positive rate and a high True Negative rate when detecting novel events. However, lower False Positive rate may lead to a higher false negative rate in most of applications for there are mutual datasets of normal and abnormal states.

5. Application to Faulty Rotating Machine

For novelty detection, normality based model will be established before its applications. After that, input data carrying novel information, which generally indicates abnormal

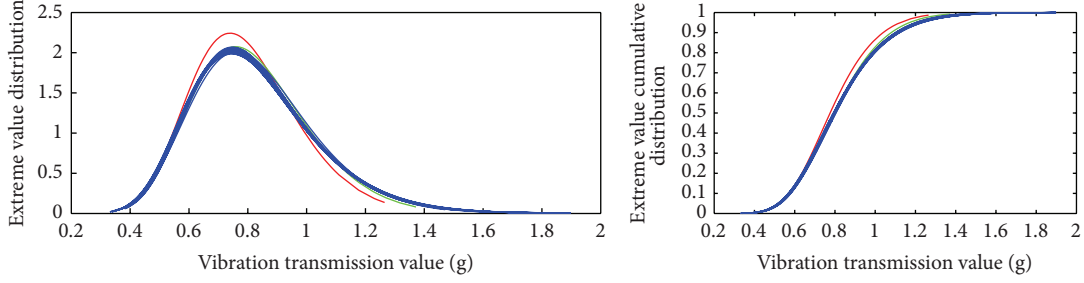


FIGURE 2: Maximum distribution curves with increasing data amount.

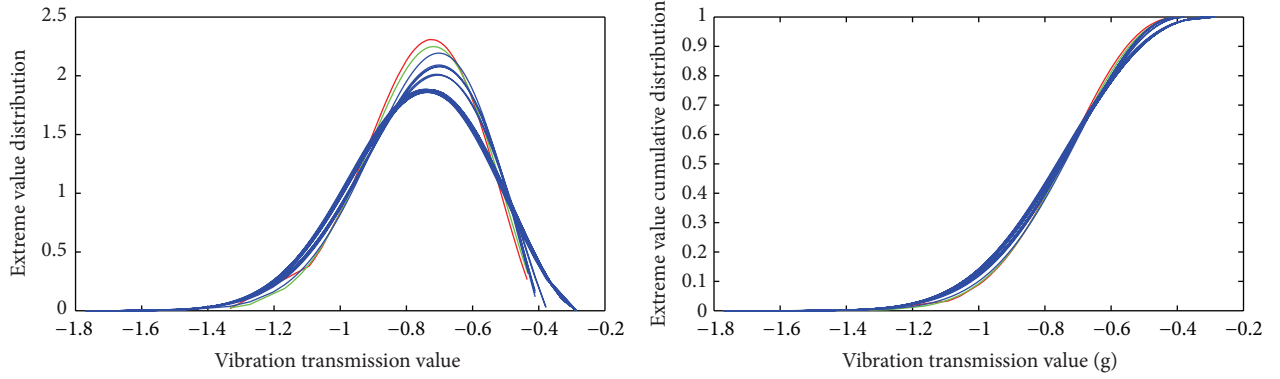


FIGURE 3: Minimum distribution curves with increasing data amount.

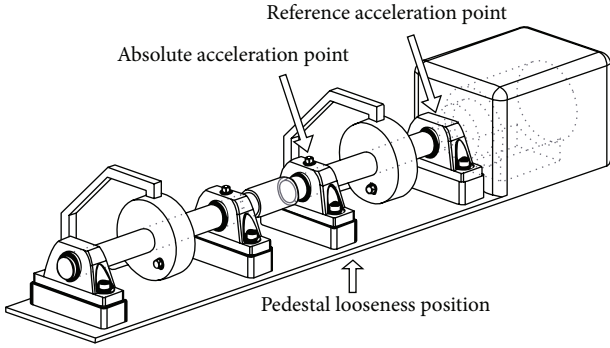


FIGURE 4: Cascade rotor system for vibration transmission measurement with pedestal looseness.

behavior of the rotor system, will be detected and warned. In order to test the effectiveness of the modified method, pedestal looseness on supporting component and minor rub-impact fault are set to the cascade rotor system deliberately. Several groups of vibration transmission data are collected by two accelerometers between absolute and reference points. The modified method is applied to detect this abnormal state.

The experiment rig for vibration transmission measurement is shown in Figures 4 and 6. Two balanced disks dimensioned 75 mm (OD) \times 23 mm (thickness) are set on two shafts, respectively. This system is comprised of two single disk rotors symmetrically arranged with shafts connected by a rigid coupling unit. Four oil lubricated journal bearings and supporters are mounted on stiff pedestals. The total length

of two shafts is 0.6 m and the diameter is 10 mm. One shaft is connected to a motor with an acoustic shield installed to reduce the motor noise. Rotary speed control for acceleration, deceleration, and steady states is implemented by PC. With a first critical speed of 2000 RPM, experimental data are all collected when speed stabilizes at 1900 RPM; thus speed fluctuation would not compel cascade system to pass the first critical speed. Four accelerometers are installed to this equipment. The sampling frequency is set at 4096 Hz for series of data acquisition.

5.1. Threshold Setting with Normal Data. The fitted distributions of normal rotor are illustrated in Figures 2 and 3 and the thresholds will be shown in Figure 5. Two criteria to claim novelty are stated as (1) maximum exceeds upper threshold of maximum accumulative distribution and (2) minimum exceeds lower threshold of minimum accumulative distribution. Those events which accord with these two criteria will be declared novelty.

To the best of our knowledge of normality training, most attention is paid to the threshold and detection rate of given methods. Here the implication of detection rate is depicted by both sensitivity and specificity. They are simply defined as follows:

$$\text{Sensitivity} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}}, \quad (8)$$

$$\text{Specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Postive}}.$$

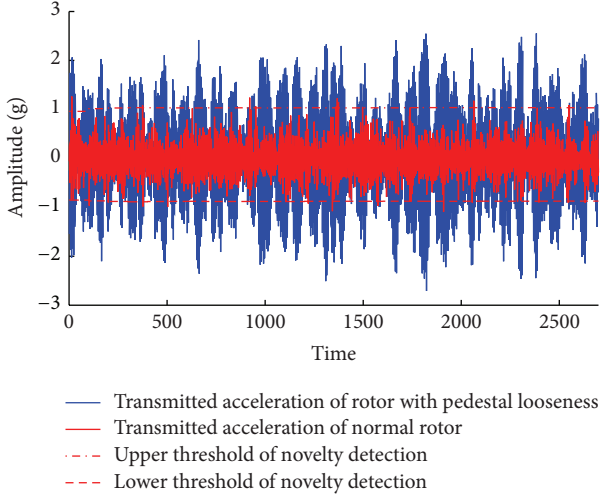


FIGURE 5: Upper/lower thresholds and acceleration time series of normal/pedestal looseness rotor.

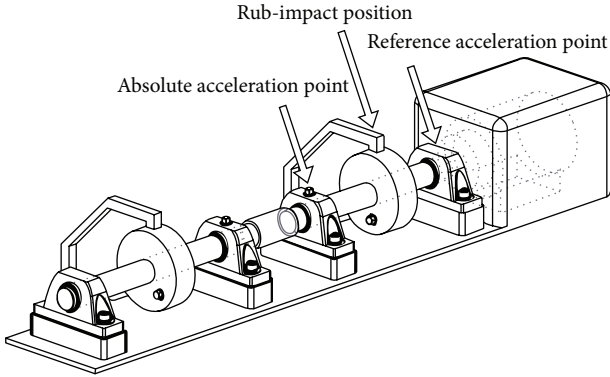


FIGURE 6: Cascade rotor system for vibration transmission measurement of rub-impact rotor.

For an ideal novelty detection result, the values of sensitivity and specificity are all supposed closed to 1. However, for some specific situations when these two cannot be considered simultaneously, it is a priority to set sensitivity in the first place because controlling risk and preventing disasters should be crucial. The results of upper and lower thresholds are given in Figure 5. As can be seen, the normal data are surrounded by upper threshold and lower threshold tightly while the pedestal looseness abnormal data exceeds the thresholds in almost all time zones.

5.2. Novelty Detection Rates. In this section, detection rates calculated in five different situations are given in Tables 1 and 2. As can be seen in the table, for a dataset, detection rate of novelty will be influenced by segmentation strategy. The optimal segmentation strategy is supposed to give sensitivity and specificity; both approximate to 1.

For both pedestal looseness and rub-impact fault, multifractal analysis based segmentation strategy gives the best detection rates. According to our calculation results shown in two tables, the value of sensitivity is 1, which means

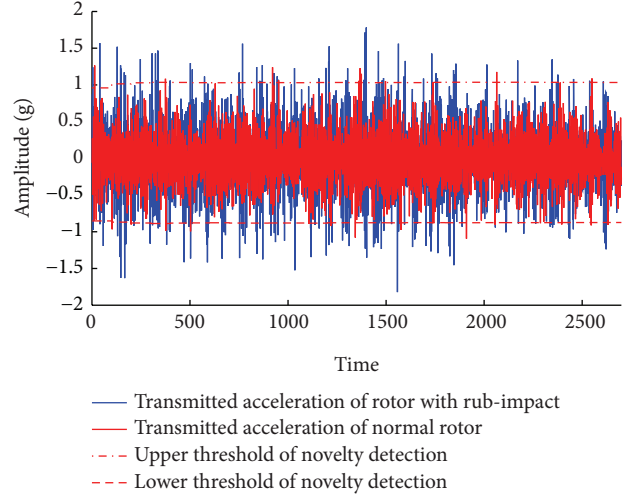


FIGURE 7: Upper/lower thresholds and acceleration time series of normal/rub-impact rotor.

TABLE 1: Pedestal looseness detection rate of different segmentation strategies.

Segmentation	10	20	50*	100	200	500
Sensitivity	1	1	1	1	1	1
Specificity	0.87	0.9	0.98	0.91	0.87	0.7

* Segmented dataset based on multifractal analysis.

TABLE 2: Rub-impact detection rate of different segmentation strategies.

Segmentation	10	20	50*	100	200	500
Sensitivity	1	1	1	1	1	1
Specificity	0.68	0.72	0.74	0.69	0.74	0.7

* Segmented dataset based on multifractal analysis.

the modified method produces no false negative, meanwhile the value of specificity is 0.98, meaning the false negative is still acceptable, although some false alarms are still observed. Similarly, shown by Figure 7, sensitivity of rub-impact fault is 1 and specificity is 0.74. Compared to sensitivity and specificity values of faulty rotor with pedestal looseness, an evident drop of specificity can be observed; it may be explained as follows: (1) rub-impact influence could be minor that the thresholds we set are conservative to avoid the false negative; (2) the position of vibration transmission measurement point is crucial because the energy level could be discrepant between each other. Considering that the purpose of this research is finding a way to lower the false alarm rate, it is indeed obtained by adopting this uncertainty reduced novelty detection approach.

6. Conclusions

Uncertainty reduced extreme value distribution of raw acceleration transmission data is established to develop a novelty detection method for abnormal dynamic rotor system.

The effects of segmentation and measurement uncertainty on stability of distribution and thresholds are investigated. Multifractal coefficient and Kullback-Leibler Divergence metrics to control uncertainty effects are examined. An abnormality detection system is constructed upon upper threshold of maximum distribution and lower threshold of minimum distribution calculated with normal data only. The pedestal looseness and minor rub-impact fault of the experimental rig are studied experimentally to demonstrate this approach. With measurement and segmentation uncertainty reduced, lower false alarm rate (higher specificity value) is obtained. Based on these investigations, the following conclusions about extreme value distribution based abnormality detection are drawn:

- (1) Acceleration transmission measurement is an easily performed supervisory instrumentality, with which supervising data could be applied to evaluate the running state of targets as a representative of energy loss. With normal data only, novelty detection criterion could be established based on extreme value distribution theory. Upper threshold and lower threshold of extremum accumulative distribution are found to be competent to give a satisfactory detection rate.
- (2) Measurement and segmentation may introduce uncertainty which impacts detection rate significantly, especially for novelty detection tasks highly relying on stable thresholds. Therefore, multifractal coefficient is adopted to reduce segmentation uncertainty by segmenting raw data into pieces with an average coefficient value equal to precalculated original one. Kullback-Leibler Divergence is also utilized to control the effect of measurement uncertainty as a criterion in iteration process of extremum distribution fitting. These proposed two metrics are validated in testing process of abnormal rotor given that the lower false alarm rate is obtained.
- (3) For rotor abnormality detection, the utility of extremum probabilistic method has spread far and wide, but here raw acceleration data has been employed. As a hundred-percent time domain detection methodology, a better time reality is obtained through this framework although it has to be admitted that this is still a static novelty detection hierarchy incompetent to deal with time-evolving threshold.

As found in this research, based on an extremum probability understanding of normal rotor, those behaviors of abnormal rotor could be detected by rule and line. Through usage of impactful strategies, uncertainty in measurement and segmentation can be controlled to ensure a more feasible approach with lower false alarm rate. For future research, the modified method is planned for us to achieve a well acceptable detection rate on composite faulty rotor with a lower false alarm rate.

Highlights

- (i) An uncertainty reduced novelty detection method based on vibration transmission measurement is proposed; uncertainty reduced extreme distributions of normal rotor system are calculated for setting novelty thresholds and to trigger alarm for abnormalities.
- (ii) Measurement and segmentation uncertainty are considered in this research. Multifractal coefficient and Kullback-Leibler Divergence have been adopted as effective metrics for corresponding uncertainty reduction and eventually guarantee the better detection accuracy.
- (iii) An abnormality detection procedure for rotating machinery is performed and yields a satisfying experimental result.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The research described in this paper is fully supported by the Aviation Science Fund of China (Grant no. 2012ZD52054) and the open project of State Key Laboratory for Strength and Vibration of Mechanical Structures (SV2015-KF-01). It is also supported by Funding of Jiangsu Innovation Program for Graduate Education (Grant no. CXZZ13_0180) and the Fundamental Research Funds for the Central Universities (no. NZ2015103). The authors would also like to thank Ponprot Boonpratpai and Dr. Yeping Xiong of Fluid Structure Interaction Group, University of Southampton, UK, for their contributions to this study. The authors also would like to thank the China Scholarship Council for supporting the first author of this paper while studying in University of Southampton, UK.

References

- [1] C. M. Bishop, "Novelty detection and neural network validation," *IEEE Proceedings—Vision, Image and Signal Processing*, vol. 141, no. 4, pp. 217–222, 1994.
- [2] J. F. D. Addison, S. Wermter, and J. MacIntyre, "Effectiveness of feature extraction in neural network architectures for novelty detection," in *Proceedings of the 9th International Conference on Artificial Neural Networks (ICANN '99)*, vol. 2, pp. 976–981, IEEE, Edinburgh, UK, 1999.
- [3] B. Hwang and S. Cho, "Characteristics of auto-associative MLP as a novelty detector," in *Proceedings of the International Joint Conference on Neural Networks (IJCNN '99)*, vol. 5, pp. 3086–3091, Washington, DC, July 1999.
- [4] S. Albrecht, J. Busch, M. Kloppenburg, F. Metzke, and P. Tavan, "Generalized radial basis function networks for classification and novelty detection: self-organization of optimal Bayesian decision," *Neural Networks*, vol. 13, no. 10, pp. 1075–1093, 2000.
- [5] T. Brotherton and T. Johnson, "Anomaly detection for advanced military aircraft using neural networks," in *Proceedings of the IEEE Aerospace Conference*, pp. 63113–63123, March 2001.

- [6] M. Markou and S. Singh, "Novelty detection: a review-part 2: neural network based approaches," *Signal Processing*, vol. 83, no. 12, pp. 2499–2521, 2003.
- [7] G. A. Barreto and R. A. Frola, "A unifying methodology for the evaluation of neural network models on novelty detection tasks," *Pattern Analysis and Applications*, vol. 16, no. 1, pp. 83–97, 2013.
- [8] P. Crook and G. Hayes, "A robot implementation of a biologically inspired method for novelty detection," in *Proceedings of the Towards Intelligent Mobile Robots Conference*, Manchester, UK, 2001.
- [9] M. F. Augusteijn and B. A. Folkert, "Neural network classification and novelty detection," *International Journal of Remote Sensing*, vol. 23, no. 14, pp. 2891–2902, 2002.
- [10] D. M. J. Tax and R. P. W. Duin, "Support vector data description," *Machine Learning*, vol. 54, no. 1, pp. 45–66, 2004.
- [11] B. Schölkopf, R. Williamson, A. Smola, J. Shawe-Taylor, and J. Piatt, "Support vector method for novelty detection," in *Proceedings of the 13th Annual Neural Information Processing Systems Conference (NIPS '99)*, pp. 582–588, December 1999.
- [12] H. Lian, "On feature selection with principal component analysis for one-class SVM," *Pattern Recognition Letters*, vol. 33, no. 9, pp. 1027–1031, 2012.
- [13] S. M. Guo, L. C. Chen, and J. S. H. Tsai, "A boundary method for outlier detection based on support vector domain description," *Pattern Recognition*, vol. 42, no. 1, pp. 77–83, 2009.
- [14] G. G. Cabral, A. L. I. Oliveira, and C. B. G. Cahú, "Combining nearest neighbor data description and structural risk minimization for one-class classification," *Neural Computing and Applications*, vol. 18, no. 2, pp. 175–183, 2009.
- [15] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding," *Science*, vol. 290, no. 5500, pp. 2323–2326, 2000.
- [16] I. Tziakos, A. Cavallaro, and L.-Q. Xu, "Event monitoring via local motion abnormality detection in non-linear subspace," *Neurocomputing*, vol. 73, no. 10–12, pp. 1881–1891, 2010.
- [17] M. Masud, J. Gao, L. Khan, J. Han, and B. M. Thuraisingham, "Classification and novel class detection in concept-drifting data streams under time constraints," *IEEE Transactions on Knowledge and Data Engineering*, vol. 23, no. 6, pp. 859–874, 2011.
- [18] M. Moshtaghi, T. C. Havens, J. C. Bezdek et al., "Clustering ellipses for anomaly detection," *Pattern Recognition*, vol. 44, no. 1, pp. 55–69, 2011.
- [19] S. Wang, J. Yu, E. Lapiac, and J. Lee, "A modified support vector data description based novelty detection approach for machinery components," *Applied Soft Computing*, vol. 13, no. 2, pp. 1193–1205, 2013.
- [20] S. Ntalampiras, I. Potamitis, and N. Fakotakis, "Probabilistic novelty detection for acoustic surveillance under real-world conditions," *IEEE Transactions on Multimedia*, vol. 13, no. 4, pp. 713–719, 2011.
- [21] M. Filippone and G. Sanguinetti, "Information theoretic novelty detection," *Pattern Recognition*, vol. 43, no. 3, pp. 805–814, 2010.
- [22] M. Breaban and H. Luchian, "Outlier detection with nonlinear projection pursuit," *International Journal of Computers, Communications and Control*, vol. 8, no. 1, pp. 30–36, 2013.
- [23] D. A. Clifton, S. Huguency, and L. Tarassenko, "Novelty detection with multivariate extreme value statistics," *Journal of Signal Processing Systems*, vol. 65, no. 3, pp. 371–389, 2011.
- [24] D. A. Clifton, L. Tarassenko, N. McGrogan, D. King, S. King, and P. Anuzis, "Bayesian extreme value statistics for novelty detection in gas-turbine engines," in *Proceedings of the IEEE Aerospace Conference*, pp. 1–11, IEEE, Big Sky, Mont, USA, March 2008.
- [25] D. A. Clifton, L. Clifton, S. Huguency, D. Wong, and L. Tarassenko, "An extreme function theory for novelty detection," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 1, pp. 28–37, 2013.
- [26] D. A. Clifton, L. Clifton, S. Huguency, and L. Tarassenko, "Extending the generalised pareto distribution for novelty detection in high-dimensional spaces," *Journal of Signal Processing Systems*, vol. 74, no. 3, pp. 323–339, 2014.
- [27] S. J. Roberts, "Extreme value statistics for novelty detection in biomedical data processing," *IEEE Proceedings: Science, Measurement and Technology*, vol. 147, no. 6, pp. 363–367, 2000.
- [28] S. J. Roberts, "Novelty detection using extreme value statistics," *IEEE Proceedings: Vision, Image and Signal Processing*, vol. 146, no. 3, pp. 124–129, 1999.
- [29] H. Sohn, D. W. Allen, K. Worden, and C. R. Farrar, "Structural damage classification using extreme value statistics," *Journal of Dynamic Systems, Measurement and Control*, vol. 127, no. 1, pp. 125–132, 2005.
- [30] H. G. D. Goyder and R. G. White, "Vibrational power flow from machines into built-up structures, part I: introduction and approximate analyses of beam and plate-like foundations," *Journal of Sound and Vibration*, vol. 68, no. 1, pp. 59–75, 1980.
- [31] P. Boonprapai and Y. P. Xiong, "Investigation of vibrational power flow patterns in damaged plate structures for damage localisation," in *Analysis and Design of Marine Structures V*, pp. 115–123, CRC Press, 2015.
- [32] Y. P. Xiong, X. P. Wang, J. T. Xing, and W. G. Price, "Hybrid active and passive control of vibratory power flow in flexible isolation system," *Shock and Vibration*, vol. 7, no. 3, pp. 139–148, 2000.
- [33] X. Q. Wang, W. O. Wong, and L. Cheng, "Modal power flow with application to damage detection," *International Journal of Engineering Science*, vol. 47, no. 4, pp. 512–523, 2009.
- [34] R. A. Fisher and L. H. C. Tippett, "Limiting forms of the frequency distribution of the largest or smallest member of a sample," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 24, no. 2, pp. 180–190, 1928.
- [35] L. Hu, N. Hu, X. Zhang, F. Gu, and M. Gao, "Novelty detection methods for online health monitoring and post data analysis of turbopumps," *Journal of Mechanical Science and Technology*, vol. 27, no. 7, pp. 1933–1942, 2013.
- [36] S. Huguency, D. A. Clifton, and L. Tarassenko, "Novelty detection with multivariate extreme value theory, part II: an analytical approach to unimodal estimation," in *Proceedings of the IEEE International Workshop on Machine Learning for Signal Processing (MLSP '09)*, pp. 1–6, IEEE, Grenoble, France, September 2009.
- [37] F. Attivissimo, A. Cataldo, L. Fabbiano, and N. Giaquinto, "Systematic errors and measurement uncertainty: an experimental approach," *Measurement*, vol. 44, no. 9, pp. 1781–1789, 2011.
- [38] M. Eidukeviciute and V. Volkovas, "Measurement uncertainty in vibromonitoring systems and diagnostics reliability evaluation," *Journal of Sound and Vibration*, vol. 308, no. 3–5, pp. 625–631, 2007.

- [39] K. Daoudi and J. L. Véhel, “Signal representation and segmentation based on multifractal stationarity,” *Signal Processing*, vol. 82, no. 12, pp. 2015–2024, 2002.
- [40] R. Lopes and N. Betrouni, “Fractal and multifractal analysis: a review,” *Medical Image Analysis*, vol. 13, no. 4, pp. 634–649, 2009.

