DOI: 10.3901/CJME.2011.03.481, available online at www.cjmenet.com; www.cjmenet.com.cn

Turbopump Condition Monitoring Using Incremental Clustering and One-class Support Vector Machine

HU Lei[1, *], HU Niaoqing[1], QIN Guojun[1], and GU Fengshou[2]

1 *College of Mechatronics Engineering and Automation, National University of Defense Technology, Changsha 410073, China*
2 *University of Huddersfield, Huddersfield, QueensgateHD13DH, UK*

Received June 23, 2010; revised December 15, 2010; accepted January 25, 2011; published electronically January 26, 2011

Abstract: Turbopump condition monitoring is a significant approach to ensure the safety of liquid rocket engine (LRE). Because of lack of fault samples, a monitoring system cannot be trained on all possible condition patterns. Thus it is important to differentiate abnormal or unknown patterns from normal pattern with novelty detection methods. One-class support vector machine (OCSVM) that has been commonly used for novelty detection cannot deal well with large scale samples. In order to model the normal pattern of the turbopump with OCSVM and so as to monitor the condition of the turbopump, a monitoring method that integrates OCSVM with incremental clustering is presented. In this method, the incremental clustering is used for sample reduction by extracting representative vectors from a large training set. The representative vectors are supposed to distribute uniformly in the object region and fulfill the region. And training OCSVM on these representative vectors yields a novelty detector. By applying this method to the analysis of the turbopump's historical test data, it shows that the incremental clustering algorithm can extract 91 representative points from more than 36 000 training vectors, and the OCSVM detector trained on these 91 representative points can recognize spikes in vibration signals caused by different abnormal events such as vane shedding, rub-impact and sensor faults. This monitoring method does not need fault samples during training as classical recognition methods. The method resolves the learning problem of large samples and is an alternative method for condition monitoring of the LRE turbopump.

Key words: novelty detection, condition monitoring, incremental clustering, one-class support vector machine, turbopump

1 Introduction

Turbopump is the propulsion machinery of a large-scale liquid rocket engine (LRE). Once faults of the turbopump occur, they will threaten the safety of the engine. Thus, condition monitoring is important for turbopump to insure reliable operation throughout the course of turbopump's service and ground test[1]. Vibration is a main cause of turbopump destructions and it directly images the condition of the turbopump. Thus vibration monitoring is one of significant ways for turbopump condition monitoring.

For the turbopump of this LRE, large amounts of vibration signals have been collected through a number of accelerometers mounted on the turbopump outside. Most of these signals are from normal turbopump. Abnormal signals can hardly be sampled due to the reasons as follows. (1) The turbopump is manufactured with high reliability, thus faults may rarely occur and less historic cases can be referred. (2) Once a novel event is suspected, the turbopump will be turned off immediately to avoid any further damages, which makes it less possible to have chance to collect this sample signal. (3) Taking into account of loss that may be caused by faults, faults embedding experiment has never been performed on the turbopump. And it is very expensive to construct an experimental facility that is fairly similar to the real system. Thus it is difficult to collect abnormal samples of most fault modes by fault simulation.

In the case of lacking fault samples, vibration monitoring relies on comparative assessments of the status under testing with normal status. This strategy is named novelty detection, one-class classification or outlier detection.

Novelty detection is the identification of new or unknown data or signal that a machine learning system is not aware of during training[2]. For machine condition monitoring, the philosophy of such approach is to establish a description of normality using the features that represent the normal condition, and then test for deviation from the normality when new data becomes available[3]. It shows that research efforts have been made in novelty detection and kinds of approaches have been exploited as reviewed in Refs. [2, 4–5].

* Corresponding author. E-mail: lake_hl@hotmail.com
This project is supported by National Natural Science Foundation of China (Grant No. 50675219) and Hu'nan Provincial Science Committee Excellent Youth Foundation of China (grant No. 08JJ1008)

As for the turbopump, previous work by the authors has concentrated on online detection algorithms and a real-time fault detection system[1, 6]. The length of training set for online methods is restricted. As incoming feature vectors are tested and then added to the training set, elder vectors are removed. Thus the detectors yielded by online framework are adaptive but not complete.

In this paper, a novelty detection method in classical framework is presented to establish a complete description of normality for the turbopump. In order to solve the learning problem of large samples, one-class support vector machine (OCSVM) is integrated with incremental clustering algorithm. Incremental clustering is used to extract a small quantity of representative vectors from large amounts of available samples. Representative vectors extracted are supposed to distribute uniformly in the object region and fulfill the region. Then a detector can be yielded by training OCSVM on these representative vectors. The layout of this paper is as follows: Section 2 describes the novelty detection method integrating OCSVM with incremental clustering. Section 3 contains the novelty detection results of the LRE turbopump. The paper concludes with some discussion in section 4.

2 Novelty Detection Method Integrating OCSVM with Incremental Clustering

2.1 Boundary description using OCSVM

OCSVM has two basic descriptions, Tax's hypersphere description[7] and Schölkopf's hyperplane description[8]. The former named support vector data description (SVDD) tries to find a hypersphere that encompasses most feature points in the training set with the minimum radius. Points located outside the sphere will be rejected as outliers. While the latter named υ-support vector classifier (υ-SVC) tries to find a hyperplane separates the dataset from the origin with maximal margin. When the data is preprocessed to have unit norm, υ-SVC is equivalent to the SVDD approach[7].

Here the SVDD method will be briefly introduced. Suppose $\boldsymbol{X} = \{\boldsymbol{x}_i,\ i = 1,\ 2,\ \cdots,\ l\}$ to be the training set, in which $\boldsymbol{x}_i$ is one of feature vectors and l is the size of the training set. A hypersphere can be obtained by solving the quadratic programming optimization problem

$$\min_{\alpha_i,\alpha_j} \quad P(\alpha) = \sum_{i,j=1}^{l} \alpha_i \alpha_j (\boldsymbol{x}_i \cdot \boldsymbol{x}_j)\,, \tag{1}$$

$$\text{s.t.} \quad 0 \leqslant \alpha_i \leqslant (\upsilon l)^{-1},\ i = 1,2\cdots,\ l, \tag{2a}$$

$$\sum_{i=1}^{l} \alpha_i = 1\,, \tag{2b}$$

where α_i and α_j are Lagrange multipliers. The trade-off parameter $\upsilon \in (0,\ 1)$ is the upper bound on the fraction of outliers over all training samples. And υ can be used to control the trade-off between the volume of the sphere and the number of outliers. Solving such a quadratic problem means finding a set $\{\alpha_i\}$ that minimize $P(\alpha)$ with subject to the constraints of Eq. (2).

The training vectors with α=0, $0 < \alpha < (\upsilon l)^{-1}$, and α= $(\upsilon l)^{-1}$ are respectively called non-support vectors (NSVs), boundary support vectors (BSVs), and non-boundary support vectors (NBSVs). And they are respectively located in, on and outside the sphere as illustrated in Fig. 1.

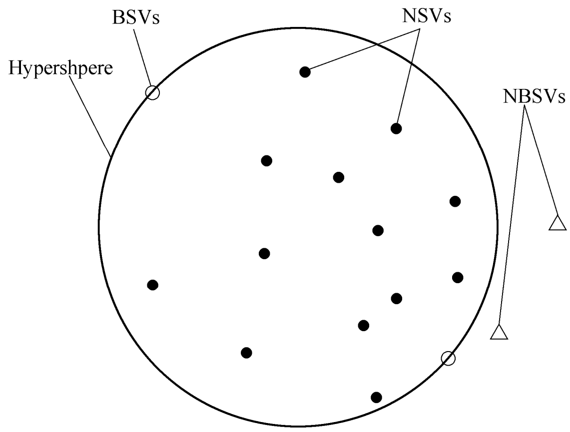


Fig. 1. Illustration of hypersphere, non-support vectors (NSVs), boundary support vectors (BSVs) and non-boundary support vectors (NBSVs) in SVDD

The decision function or test function of SVDD is

$$f(\boldsymbol{z}) = \sum_k \alpha_k (\boldsymbol{x}_k \cdot \boldsymbol{z}) - b\,, \tag{3}$$

where $\boldsymbol{z}$ is a test vector, $\boldsymbol{x}_k$ is one of BSVs or NBSVs and α_k is its corresponding Lagrange multiplier, b is the threshold or offset. If $f(\boldsymbol{z}) \geqslant 0$, $\boldsymbol{z}$ will be accepted as a normal sample, otherwise $\boldsymbol{z}$ will be excluded as a novel sample. As BSVs lie on the sphere, the threshold can be yielded by inputting one of BSVs to the decision function

$$b = \delta \sum_k \alpha_k (\boldsymbol{x}_k \cdot \boldsymbol{x}_{\mathrm{BSVs}})\,, \tag{4}$$

where δ is a scaling parameter used to decrease the threshold to reduce false alarms. δ can be set smaller than and close to 1.

A flexible description can be yielded by replacing the inner product $(\boldsymbol{x}_i \cdot \boldsymbol{x}_j)$ in Eqs. (1), (3), and (4) with a kernel function $K(\boldsymbol{x}_i,\ \boldsymbol{x}_j)$. And OCSVM mostly uses Gaussian kernel

$$K(\boldsymbol{x}_i, \boldsymbol{x}_j) = \exp\left(-\frac{\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|^2}{\sigma^2}\right), \tag{5}$$

where σ is the width parameter of the Gaussian kernel.

2.2 Density based Incremental clustering

Density based clustering algorithms are serious candidates for processing large samples as they require only a few passes on the training set and have the merits of high computational efficiency[9]. Here a simple density based incremental clustering algorithm is proposed to extract representative points from large historical test data of the turbopump. When new training samples are available, the cluster described by representative points can be updated incrementally without taking past training set into account.

Define $V_\varepsilon(\boldsymbol{x}_i)$ to be a ε-neighborhood of $\boldsymbol{x}_i$, which is a hypersphere centered at $\boldsymbol{x}_i$ with radius ε defined by user. For any $\boldsymbol{x} \in V_\varepsilon(\boldsymbol{x}_i)$, $\|\boldsymbol{x} - \boldsymbol{x}_i\| \le \varepsilon$. The "density" of the neighborhood of $\boldsymbol{x}_i$ is the number of points of $\boldsymbol{X}$ lying in $V_\varepsilon(\boldsymbol{x}_i)$, denoted by $N_\varepsilon(\boldsymbol{x}_i, \boldsymbol{X})$. Then two definitions are made as follows.

Definition 1. (dense point and sparse point) A point $\boldsymbol{x}_i$ is a dense point of $\boldsymbol{X}$ if $N_\varepsilon(\boldsymbol{x}_i, \boldsymbol{X}) > q$, $q \in N$, otherwise it is a sparse point. Here q is the maximum density defined by user.

Definition 2. (the set of representative points) $\boldsymbol{P}$ is a set of $\boldsymbol{X}$'s representative points, if

(i) $\boldsymbol{P} \subseteq \boldsymbol{X}$.

(ii) $\forall\, \boldsymbol{p} \in \boldsymbol{P}$, $N_\varepsilon(\boldsymbol{p}, \boldsymbol{X}) \le q$.

(iii) $\forall\, \boldsymbol{x} \in \boldsymbol{X}$, $\exists\, \boldsymbol{p} \in \boldsymbol{P}$, satisfy $\boldsymbol{x} \in V_\varepsilon(\boldsymbol{p})$.

For current novelty detection scheme, there is only one cluster in the algorithm. The set of representative points $\boldsymbol{P}$ can be extracted from training set $\boldsymbol{X}$ using the algorithm as follow.

Density based incremental clustering algorithm.

Let $\boldsymbol{X}^{\mathrm{un}}$ be the set of points in $\boldsymbol{X}$ that have not been considered yet.

Set $\boldsymbol{X}^{\mathrm{un}} = \boldsymbol{X}$, $\boldsymbol{P} = \Phi$
While $\boldsymbol{X}^{\mathrm{un}} \neq \Phi$, do
  Arbitrarily select a $\boldsymbol{x} \in \boldsymbol{X}^{\mathrm{un}}$
  If $N_\varepsilon(\boldsymbol{x}, \boldsymbol{P}) \leqslant q$, set
    $\boldsymbol{P} = \boldsymbol{P} + \{\boldsymbol{x}\}$
    $\boldsymbol{X}^{\mathrm{un}} = \boldsymbol{X}^{\mathrm{un}} - \{\boldsymbol{x}\}$
  End {if}
End {while}

When a new training set $\boldsymbol{X}^{\mathrm{new}}$ is available, $\boldsymbol{P}$ can be updated by letting $\boldsymbol{X}^{\mathrm{un}} = \boldsymbol{X}^{\mathrm{new}}$, and performing the incremental clustering algorithm. Elder training set $\boldsymbol{X}$ needs not to be considered any more.

2.3 Integrating OCSVM with incremental clustering

Support vector machine (SVM) can be used to solve problems in various subjects such as novelty detection, classification and regression. It becomes popular in machine condition monitoring and fault diagnosis due to its global solution and excellent generalization [4, 6, 10–13].

A major limitation of SVM or OCSVM is the high computational burden required especially during training phase[10, 14]. A quadratic programming optimization problem needs to be solved when training SVM, and solving such a problem sets high demands on computer memory requirements for large training sets. Though decomposition methods enable SVM to go through large training sets[14-17], they are still time consuming as training vectors extracted from turbopump vibration have mounted up to thousands upon thousands.

In order to process the large data sets in vibration monitoring of the turbopump, density based incremental clustering is used firstly to reduce the size of the training set. In this way, a more representative sample is obtained and presented to OCSVM for more efficient classification.

2.4 Parameter tuning

There are four parameters in this novelty detection method:

υ —Trade-off parameter of OCSVM,

σ—Width parameter of Gaussian kernel,

q —Maximum density of $V_\varepsilon(\boldsymbol{x})$,

ε —Radius of $V_\varepsilon(\boldsymbol{x})$.

Determination of parameters υ and σ has been deeply analyzed in Ref. [7]. Generally, a smaller υ should be chosen if few training vector is allowed to be rejected. And σ can be set around $d_{\max}/2$, where $d_{\max}$ is the largest distance between training points. A bigger σ is suggested in fault detection to reduce false alarms on normal data.

For a fixed training set, the choice of q and ε lies on how many representative vectors are supposed to be obtained. Let n be the number of representative vectors extracted. It's not hard to see that the bigger ε, the bigger the region represented by each representative vector, and the smaller the number of vectors needed to represent the complete object region. It's easier to see that the smaller q, the smaller the n for a fixed ε. When $q = 1$, there is only one representative vector in each ε-neighborhood, and representative vectors yielded distribute uniformly in the object region and none of them is redundant. Thus $q = 1$ is suggested first. Then for an acceptable interval $[n_{\min}, n_{\max}]$, the value of ε can be modified according to the number of representative vectors n. Reduce ε to increase n when $n < n_{\min}$ and increase ε to reduce n when $n > n_{\max}$.

Fig. 2 displays a simulation case in which the initial training set consists of 500 points. 39 representative vectors are extracted by performing incremental clustering ($q = 1$, $\varepsilon = 4$) on the initial training set. A hypersphere is obtained by training OCSVM ($\upsilon = 0.001$, $\sigma = 11$, $\delta = 0.9$) on these 39 representative vectors. Fig. 3 shows the variation of n with ε and q. It can be seen that n increases with the decrease of ε or the increase of q.

3 Turbopump Condition Monitoring

3.1 Feature selection

A mass of vibration signals including a small quantity of fault samples have been recorded from a kind of LRE turbopump in a series of ground tests. Restricted by space and structure, sensors used to record vibrations are mounted on the outside of the turbopump. Vibration signal of the turbopump is a superposition of components and it is hard to recognize faults of the turbopump according to frequency-domain analysis. While faults occur, abrupt

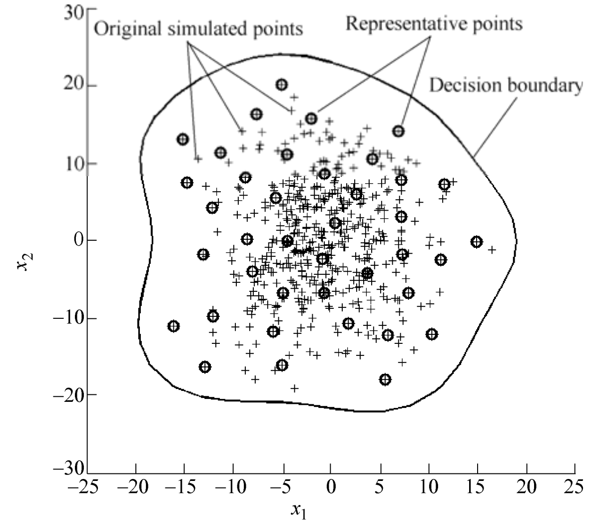


Fig. 2. Illustration of result yielded by performing density based incremental clustering and OCSVM on simulated data

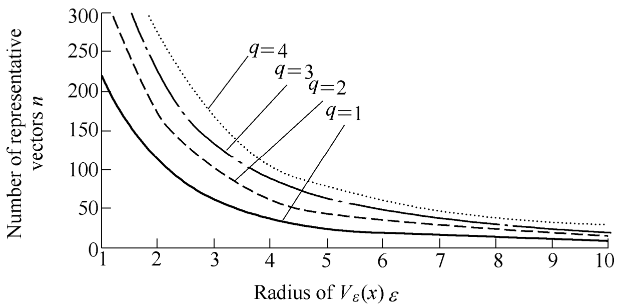
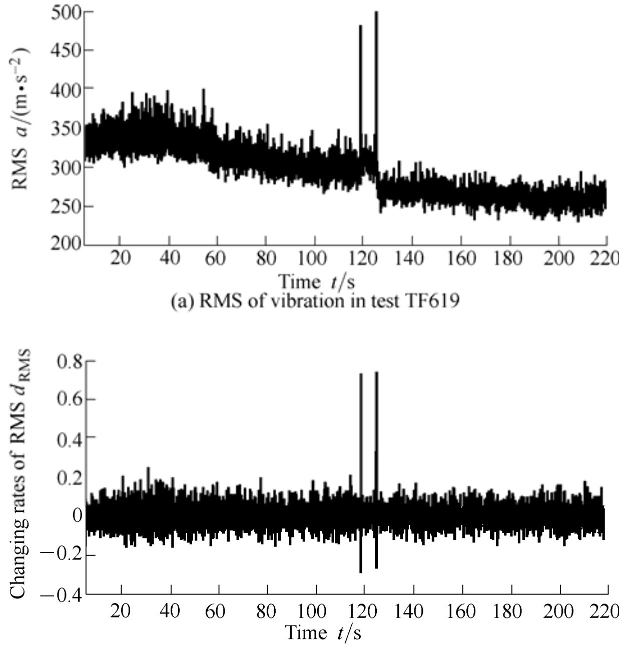


Fig. 3. Variation of n with ε and q for the case in Fig. 2



(a) RMS of vibration in test TF619

(b) Changing rates of RMS

Fig. 4. Comparison of RMS and its changing rates d_{RMS}

shock will arise along with thrust generated by the turbopump. Correspondingly, statistic characteristics of vibration signals will change in the form of energy or waveform. These changes can be described with time-domain features, such as mean, stand deviation, root mean square (RMS), kurtosis factor (KF), clearance factor (CF) and one-step autocorrelation coefficient. Time-domain features have the merit of low cost of computation and they are widely used for monitoring the turbopump of rocket engine in practical engineering. The dependency and fault sensitivity of these time-domain features have been analyzed with a large amount of test data. And RMS, KF and CF are selected finally[18].

Vibration of turbopump is also sensitive to kinds of random factors and its statistic characteristics may vary with tests carried out at different time. In order to enable these detection features to have better consistency and uniform data scaling, the changing rates of RMS, KF and CF instead of themselves are used as final detection features. For a time-domain feature x, its changing rate is defined as

$$d_x = \frac{x(i+1) - x(i)}{x(i)}. \tag{6}$$

Fig. 4 shows the RMS of vibration in test TF619 and its changing rates d_{RMS}.

3.2 Fault detection

Vibration signals of the turbopump were sampled at 50 kHz. Features were calculated every 0.05 s. 91 representative points were extracted by performing density based incremental clustering on 36 800 feature vectors. And these 36 800 feature vectors were extracted from 10 historical tests, which are identified by the field engineer to cover various normal operating conditions. The maximum density of $V_\varepsilon(\boldsymbol{x})$ was q = 1 and the radius of $V_\varepsilon(\boldsymbol{x})$ was ε = 0.12. A novelty detector was yielded by training OCSVM on these 91 representative points. Parameters of the OCSVM are υ = 0.01, σ = 0.8, δ= 0.98. Then the novelty detector was used to detect abnormality in historical data records.

By taking $f(\boldsymbol{x})$ in Eq. (3) as novelty index, Fig. 5 displays the features of test TF619. During test TF619, turbo vanes shed at 120.83 s and 127.18 s, which caused short shocks in the vibration waveform. But these faults had not been given enough attention. It was fortunate that nasty accident had not been caused and the test continued to the end. Fig. 6 shows the novelty indexes of the vibration signal in test TF619.

Fig. 7 displays the features of test TF627 during which serious rub-impact occurred from 11.73 s. Emergency measure was taken to shut down the turbopump. Fig. 8 shows the novelty indexes of the vibration signal in test TF627.

Fig. 9 displays the features of test TN618. During test TN618, the turbopump itself was normal, but sensor faults occurred. The vibration sensor disabled during time intervals [91.52 s, 93.11 s] and [103.66 s, 104.98 s]. Fig. 10 shows the novelty indexes of the vibration signal in test TN618.

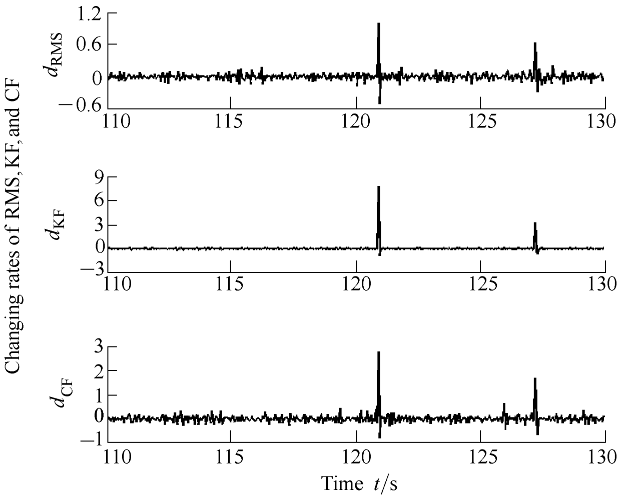


Fig. 5. Detection features of test TF619

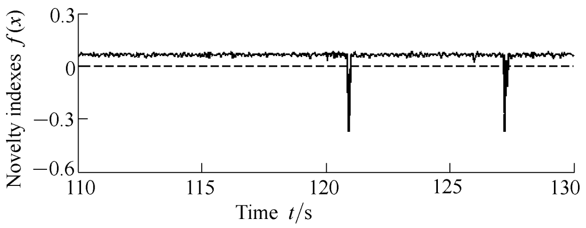


Fig. 6. Novelty detection results of test TF619

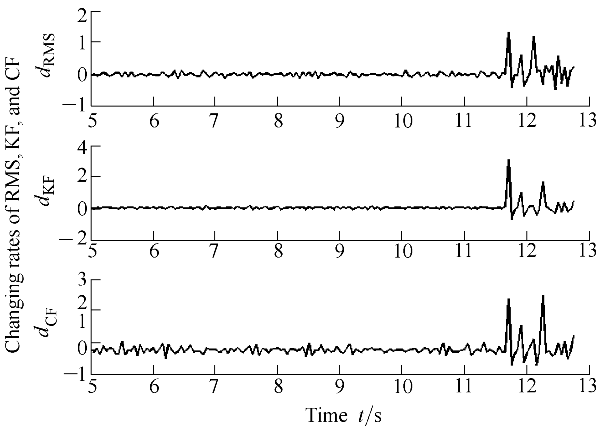


Fig. 7. Detection features of test TF627

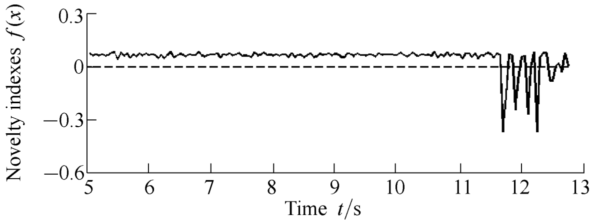


Fig. 8. Novelty detection results of test TF627

It can be seen from Fig. 6, Fig. 8 and Fig. 10 that the novelty detector derived from the combination of incremental clustering and OCSVM can detect these kinds of faults effectively. Detection results of all the normal tests that we have (15 tests including those 10 tests used for training the detector) showed no false alarm.

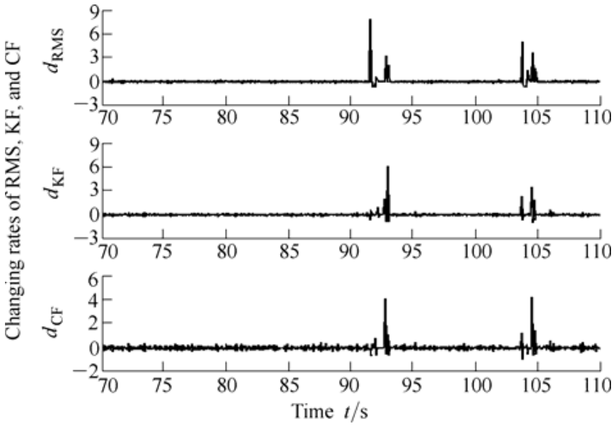


Fig. 9. Detection features of test TN618

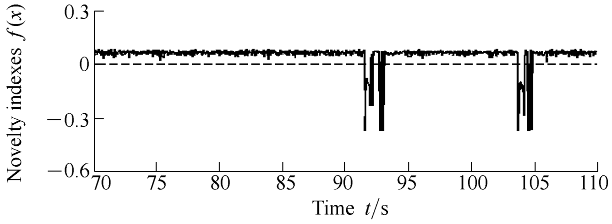


Fig. 10. Novelty detection results of test TF618

4 Conclusions

(1) A novelty detection method is developed for turbopump vibration monitoring.

(2) This method extracts a more representative data set from large historical data sets with density based incremental clustering algorithm. By this pre-processing, the data points in the representative set distribute uniformly and cover the entire object region.

(3) When new tests are available, the set of representative points can be updated easily leaving elder training set out of account. Moreover, the size of the representative set is under control.

(4) A novelty detector used to describe the boundary of the object region is yielded by training OCSVM on the representative set.

(5) Applying this method to the condition monitoring of a LRE turbopump has proved that it can identify different spikes in vibration signals caused by abnormal events such as vane shedding, rub-impact and sensor faults.

References

[1] HU Niaoqing, QIN Guojun, HU Lei, et al. Real-time fault detection system for turbopump of liquid rocket engine based on vibration measurement signals[C]//Proceeding of the 6th International Workshop on Structural Health Monitoring, Stanford, 2007: 590–597.

[2] MARKOU M, SINGH S. Novelty detection: a review, Part 1: statistical approaches[J]. *Signal Processing*, 2003, 83(12): 2 481–2 497.

[3] WORDEN K, MANSON G, ALLMAN D. Experimental validation of a structural health monitoring methodology. Part I. Novelty detection on a laboratory structure[J]. *Journal of Sound and Vibration*, 2003, 259(2): 323–343.

[4] MARKOU M, SINGH S. Novelty detection: a review, Part 2: Neural network based approaches[J]. *Signal Processing*, 2003, 83(12): 2 499–2 521.
[5] HODGE V J, AUSTIN J. A survey of outlier detection methodologies[J]. *Artificial Intelligence Review*, 2004, 22(2): 85–126.
[6] HU Lei, HU Niaoqing, QIN Guojun. Online fault detection algorithm based on double-threshold OCSVM and its application[J]. *Journal of Mechanical Engineering*, 2009, 45(3): 169–173. (in Chinese)
[7] TAX D, DUIN R. Support vector data description[J]. *Machine Learning*, 2004, 54(1): 45–66.
[8] SCHÖLKOPF B, WILLIANMSON R, SMOLA A, et al. Support vector method for novelty detection[C]//*Advances in neural information processing systems 12: proceedings of the 1999 conference*, MIT Press, 2000: 582–588.
[9] THEODORIDIS S, KOUTROUMBS K. *Pattern recognition*[M]. 3rd edition. Beijing: China Machine Press, 2006.
[10] NELLO C, JOHN S. *An introduction to support vector machines and other kernel-based learning methods*[M]. Cambridge: Cambridge University Press, 2000.
[11] WIDODO A, YANG B S. Support vector machine in machine condition monitoring and fault diagnosis[J]. *Mechanical System and Signal Processing*, 2007, 21(6): 2 560–2 574.
[12] HAYTON P, UTETE S, KING D, et al. Static and dynamic novelty detection methods for jet engine health monitoring[J]. *Philosophical Transactions of the Royal Society A*, 2007, 365: 493–514.
[13] DAVY M, DESOBRY F, GRETTON, A, et al. An online support vector machine for abnormal events detection[J]. *Signal Processing*, 2006, 86(8): 2 009–2 025.
[14] PLATT J. *Sequential minimal optimization: a fast algorithm for training support vector machines*[R]. Technical Report MSR-TR-98-14, Microsoft Research, 1998.
[15] CHANG Chih-chuang, HSU Chih-wei, LIN Chihjen. The analysis of decomposition methods for support vector machine[J]. *IEEE Transactions on Neural Networks*, 2000, 11(4): 1 003–1 008.
[16] MATTERA D, PALMIERI E, HAYKIN S. An explicit algorithm for training support vector machines[J]. *IEEE Signal Processing Letters*, 1999, 6(9): 243–246.
[17] HU Lei, HU Niaoqing, QIN Guojun. Support vector machines detect method for turbopump test data analysis[J]. *Journal of Propulsion Technology*. 2008, 29(2): 244–248. (in Chinese).
[18] XIE Guangjun. *Research on real-time fault detection technology and system for liquid rocket engine turbopump*[D]. Changsha: National University of Defense Technology of China, 2006. (in Chinese).

Biographical notes

HU Lei, born in 1981, is currently in *National University of Defense Technology*, *China*. His research interests focus on condition monitoring, diagnostics and prognostics.
Tel: +86-371-4574971; E-mail: lake_hl@hotmail.com

HU Niaoqing, born in 1967, is currently a professor in *National University of Defense Technology, China*. He received his PhD degree from *National University of Defense Technology*, *China*, in 2002. His research interests include condition monitoring, diagnostics, prognostics and health management.
E-mail: hnq@nudt.edu.cn

QIN Guojun, born in 1970, is currently an associate professor in *National University of Defense Technology*, *China*. He received his PhD degree from National *University of Defense Technology*, *China*, in 1998. His research interests include condition monitoring, diagnostics, and wireless sensor network.
E-mail: qgj@nudt.edu.cn

GU Fengshou is currently a senior research fellow of machine diagnosis engineering at *University of Huddersfield*, *UK*. He is one of the experts in the field of machine diagnosis and vibro-acoustics analysis, with over 20 years of research experience.
E-mail: f.gu@hud.ac.uk