Applied Soft Computing 13 (2013) 1193–1205

Contents lists available at SciVerse ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

A modified support vector data description based novelty detection approach for machinery components

Shijin Wang[a,*], Jianbo Yu[b], Edzel Lapira[c], Jay Lee[c]

[a] Department of Management Science and Engineering, School of Economics & Management, Tongji University, Shanghai 200092, PR China
[b] College of Mechatronics Engineering and Automation, Shanghai University, Shanghai 200072, PR China
[c] Center for Intelligent Maintenance Systems, Department of Mechanical Engineering, University of Cincinnati, Cincinnati, OH 45221, United States

ARTICLE INFO

Article history:
Received 30 October 2011
Received in revised form
28 September 2012
Accepted 8 November 2012
Available online 19 November 2012

Keywords:
Novelty detection
Support vector data description
One-class classification
Parameter selection
Bearing

ABSTRACT

Novelty detection is an important issue for practical industrial application, in which there is only normal operating data available in most cases. This paper proposes a systematic approach for novelty detection of mechanical components, using support vector data description (SVDD), a kernel approach for modeling the support of a distribution. To reduce the false alarm rate and increase the detection accuracy, a parameter optimization estimation scheme is proposed based on a grid search method that relies on the performance trade-off between the minimum fraction of support vectors and the maximum dual problem objective value. An evaluation value (*E*-value) chart based on the kernel distance for detection result is also designed to facilitate the decision visualization. To illustrate the effectiveness of the proposed method, novelty detection was applied to a particular kind of tapered roller bearing used in an industrial robot, which is investigated as a case study. The experimental results, in comparison to other methods, demonstrate that the proposed SVDD can conduct novelty detection of the monitored mechanical component effectively with higher accuracy.

1. Introduction

Realizing that new input data is novel can be extremely important since it can indicate that certain change or novel event has occurred within the system. Novelty detection can be used in this case to highlight such data. Novelty detection is the process of detecting abnormal behavior in a system by learning the normal behavior [1]. It is different from conventional fault diagnosis, which focuses on detecting and recognizing situations or behaviors which have previously been known to occur or can be anticipated [2]. Therefore, novelty detection has been of great interest to researchers in different domains, especially in areas where it is difficult or expensive to obtain examples of abnormal or novel behavior [1], such as in machine fault detection [3], network intrusion detection [4], and image analysis [5,6].

In the field of mechanical components condition monitoring, with the steady growth of complexity and sophistication of machinery components, it becomes increasingly difficult and costly to perform induced-failure tests on the designed system at all possible operating conditions. This leads to sparse and rare of various failure data (i.e., outlier or novelty data) ready for a priori training [7]. Novelty detection offers a solution to such kind of complications, since it is one-class classification problem [8] and has the advantage of not requiring a priori knowledge of various failure data. Still, there are challenges that need to be addressed in novelty detection such as increasing the detection accuracy as well as increasing the sensitivity to incipient novelty or new events.

Due to its significance in theory and its frequent occurrence in practical industrial applications, a number of different approaches for novelty detection have been developed. There are three main categories of approaches to novelty detection, namely, statistical approach, neural network and support vector based approaches. For a thorough overview of novelty detection in machine learning before 2004, readers can refer to Refs. [9–11].

From a statistical point of view, novelty detection is based on a sequential likelihood-ratio test between normal training statistical situation and the statistical situation for new coming data [12]. The popular statistic method is to estimate the underlying probability density function (PDF) for the distribution of training data with the assumption of distribution (e.g., Gaussian distribution). The most commonly used models for PDF estimation are the Gaussian mixture model (GMM), *k*-nearest neighbor method and Parzen windows [9]. When using a statistical approach, one needs to specify or make assumptions on the nature of training data. However, the actual distribution is usually unknown and is also not easy to estimate accurately [13]. To illustrate, a supposed normally

* Corresponding author. Tel.: +86 21 65982806; fax: +86 21 65982806.
E-mail address: shijinwang0223@yahoo.com.cn (S. Wang).

1568-4946/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.asoc.2012.11.005

distributed data needs the assumption that a sufficiently large collection of samples will satisfy the central limit theorem, on the other hand, non-normal distributions are commonly encountered in real data. Density estimation is more difficult if the underlying probability measure does not possess an absolutely continuous distribution function.

As for neural networks-based novelty detection, multi-layer perceptron (MLP) [14], radial basis function network (RBF) [15], and self organizing maps (SOMs) [16] are popular methods. Compared with other kinds of neural networks, SOM is more widely used. For example, based on a new statistical features extraction technique for vibration signal, Wong et al. [3] proposed a modified SOM-based novelty detector with one-class training data. The experiments on bearing data obtained from a test rig and data from a US Navy helicopter show the high accuracy and robustness of their proposed method.

The reasons for the wide use of SOM in novelty detection could be summarized as follows: (1) SOM is unsupervised and a priori information on class label is not necessary. Only the data collected from the healthy machine is needed to construct an SOM. (2) This method places emphasis on domain description compared to probability density estimation. (3) It offers different intuitive visualizations, which can be used for decision making, such as *U*-matrix, histogram and operating point and trajectory. Despite these numerous features that contribute to its appeal, the selection of an appropriate threshold has been an issue that has not been properly addressed up to now [11].

Recently, inspired by the support vector machine theory (a kernel based machine learning method), Tax and Duin [17,18] derived the support vector data description (SVDD) method. SVDD is a one-class classification method that estimates the distributional support of a data set. A flexible enclosing boundary is used to separate target data on the insides from outliers on the outside [7]. Due to its several generic benefits including ability to handle data sparsity, good generalization and use of kernels [5], SVDD has been widely used into novelty detection of mechanical components. Tax et al. [19] applied the SVDD into pump failure detection. Based on the vibration measurements from three accelerometers, features such as power spectrum, envelope spectrum, autoregressive modeling, RMS, kurtosis are extracted. In terms of both the ratio of target acceptance and outlier rejection, the performance of the SVDD is evaluated through different combinations of feature sets and width parameter σ. The results showed that the SVDD can distinguish normal working station of a pump from erroneous behavior. Considering that SVDD is good at describing the boundary for the non-Gaussian distributed variables, Xie and Kruger [20] employed SVDD to monitor abnormal plant operation based on the particle swarm optimization and independent component analysis (PSO-ICA) feature extraction technique. Case study based on Tennessee Eastman process simulator showed the effectiveness of their method. Pan et al. [21] employed the SVDD method for bearing performance degradation assessment based on the normal data's feature vectors composed of node energies of improved wavelet packet decomposition (IWPD). A health index (HI) is then designed based on the distance to the sphere boundary, which can be used to represent either the normal condition or degradation conditions (larger HI, worse degradation) according to the HI value. The experimental results in a bearing accelerated life test show the feasibility and effectiveness of the proposed method of IWPD and SVDD. Liu et al. [22] proposed a fast SVDD (F-SVDD) to reduce the testing time complexity of SVDD in novelty detection. For demonstration of the effectiveness of the F-SVDD, a practical industry example regarding liquid crystal display micro-defect inspection is used to compare the applicability of SVDD and the F-SVDD when faced with mass data input. The proposed F-SVDD is further used for defect detection of real surface images of LCD panels in thin film transistor liquid crystal display (TFT-LCD) array process [23]. Chou et al. [24] proposed a genetic algorithm (GA)-SVM based virtual metrology (VM) system for wafer quality prediction system, in which SVDD is chosen as the VM model and it is used for novelty detection automatically. Based on the real-time data of 10 sensors from chemical vapor deposition (CVD) equipment, the kernel principal components analysis (KPCA) is used to extract significant features, which in turn are input in the SVDD module for novelty detection. Experiments demonstrate that the SVDD can achieve lower classification errors compared with the non-SVDD methods in both training and testing data.

All work mentioned above demonstrated the promising outlook on SVDD in novelty detection in practical industrial application for different mechanical components. However, it deserves further exploration. It still needs a generic methodology that is applicable across applications without a priori knowledge of known data distributions. Moreover, there are still two issues for the SVDD-based novelty detection that need to be addressed.

(1) One is the selection of the free hyper-parameter σ and C. The performance of SVDD is very dependent on the choice of these parameters. The only way to set σ and C is through a heuristic approach and if not enough data is present for validation purpose this can become very difficult. Tax and Duin [18] mentioned that σ can be estimated based on the expectation of errors using leave-one-out (LOO) cross validation of training data. But LOO requires high computational expense. The ensemble of SVDD models with different σ might lessen the impact of σ, but how to determine the appropriate number of base SVDDs in the ensemble is another problem.
(2) Another is the visualization of decision. SVDD is weaker in decision visualization compared with SOM. The support boundary of SVDD can only be shown for the training data with two or three dimension (2D or 3D) features.

In this context, this paper proposes an SVDD-based systematic novelty detection approach for mechanical components by solving two issues mentioned above. The relatively optimal σ and C are obtained based on a parameter optimization estimation scheme, in which a grid search method is optimized based on a trade-off between the minimum fraction of support vectors and the maximum dual problem objective value. A straightforward evaluation value (*E*-value) chart based on kernel distance will be defined to visualize the testing decision.

The remainder of the paper is organized as follows. Section 2 describes the systematic scheme for SVDD-based novelty detection. Section 3 introduces the theoretical background of SVDD. The modified algorithm with parameter optimization and visualization chart is illustrated in detail in Section 4. Novelty detection for a tapered roller bearing of industrial robot is presented in Section 5. Finally, conclusions are drawn and future work is described in Section 6.

2. Systematic approach for novelty detection

There are several important issues related to novelty detection. A systematic novelty detector should possess the following characteristics [1,9]:

(1) Novelty detection and degradation evaluation: a novelty detector can detect novelty system behavior deviating from the normal condition, and can find out the gradual performance degradation of the monitored components and systems.
(2) Parameter optimization and parameter number minimization: a novelty detector should aim to minimize the number of