



ELSEVIER

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/jnlabr/ymssp

Application of relevance vector machine and logistic regression for machine degradation assessment

Wahyu Caesarendra^a, Achmad Widodo^b, Bo-Suk Yang^{a,*}

^a School of Mechanical Engineering, Pukyong National University, San 100, Yongdang-dong, Nam-gu, Busan 608-739, South Korea

^b Mechanical Engineering Department, Diponegoro University, Tembalang, Semarang 50275, Indonesia

ARTICLE INFO

Article history:

Received 20 August 2009

Received in revised form

14 October 2009

Accepted 15 October 2009

Available online 27 October 2009

Keywords:

Relevance vector machine

Logistic regression

Failure degradation

Prognostics

ABSTRACT

Degradation parameter or deviation parameter from normal to failure condition of machine part or system is needed as an object of prediction in prognostics method. This study proposes the combination between relevance vector machine (RVM) and logistic regression (LR) in order to assess the failure degradation and prediction from incipient failure until final failure occurred. LR is used to estimate failure degradation of bearing based on run-to-failure datasets and the results are then regarded as target vectors of failure probability. RVM is selected as intelligent system then trained by using run-to-failure bearing data and target vectors of failure probability estimated by LR. After the training process, RVM is employed to predict failure probability of individual units of machine component. The performance of the proposed method is validated by applying the system to predict failure time of individual bearing based on simulation and experimental data. The result shows the plausibility and effectiveness of the proposed method, which can be considered as the machine degradation assessment model.

© 2009 Elsevier Ltd. All rights reserved.

1. Introduction

Starting from a new condition, a machine part that was properly manufactured and installed into a system without any errors will operate at a particular level of performance, which is ideally operated at its design requirement. As its operating life progresses, degradation occurs. Regardless of the reasons for degradation, the item can no longer meet its original service requirements and its level of performance falls. Here, failure degradation assessment becomes a key point. Failure degradation assessment can be used to monitor the level of performance falls of a part or system. By detecting the loss-in-condition of the part, it can be studied when the degradation was started. Moreover, the change in performance level is useful to forecast the impending failure. In maintenance activity, failure degradation assessment can be used as an important feature or information that indicates the performance of a part or a machine at particular or the range of time, usually through condition monitoring and fault diagnostics system. In order to make the appropriate decision in system health management and predict the impending failure, the condition monitoring and fault diagnostics of machine performance have to be complemented with the prognostics system. Now, the term prognostics is introduced. Prognostics hold the promise for considerable cost savings (for example by avoiding unscheduled maintenance and by increasing equipment usage) and operational safety improvements. Therefore, it is not surprising that finding solutions to the prognostics problem is a very active area. Many authors have defined and classified the term machine prognostic based on the utilization of data input and method. Jardine et al. [1] explained that the approaches to prognosis fall into three main

* Corresponding author. Tel.: +82 51 629 6152; fax: +82 51 629 6150.
E-mail address: bsyang@pknu.ac.kr (B.-S. Yang).

categories: *statistical approaches*, *artificial intelligence (AI) approaches* and *model based approaches*. Heng et al. [2] broke down the methods for predicting rotating machinery failure into two models: *physics-based prognostics models* and *data-driven prognostics models*. Each prognostics method has merits and demerits. Currently, the AI approach or the data driven approach is still a popular method that can be defined as a method to build models based on learning process or training process of condition monitoring (CM) data to produce the prediction output in terms of one-step-ahead or k -step-ahead prediction. In this category, artificial neural network (ANN) is the most commonly found in the prognostics literature [3–7]. Other methods were reported using support vector machine [8], regression tree and neuro-fuzzy [9,10], and the Dempster–Shafer regression [11].

One approach that is rarely implemented due to limitation of maintenance records including both normal data and failure data is the probability prognostic approach. Probability prognostic is a more appropriate method since the behavior of failure is unpredictable and uncertain. This study proposes the combination of probability approach and data-driven approach based on logistic regression (LR) and relevance vector machine (RVM). The utilization of either LR or RVM for machine health prognostics is still in its infancy. LR is frequently used in statistic science and biomedical research, e.g. [12,13]. In mechanical engineering domain, Yan et al. [14] combined the LR and ARMA models for machine performance assessment and estimate remaining useful life (RUL). Yan et al. [15] also presented a prognostic method for on-line performance degradation assessment and root cause classification using multiple logistic regressions. Comprehensive discussion of LR can be found in [16].

RVM, which was first introduced by Tipping, [17] is often used for classification and pattern recognition [18–20]. Previous work concerning classification has been devoted to fault diagnosis of low speed bearing using a multi-class RVM [21]. This study uses RVM in the application of prognostics in order to address the failure degradation assessment and prediction of incipient failure until final failure occurred. Feature named kurtosis is calculated initially. Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution [22]. Kurtosis is also one of the statistical time domain feature that indicates the performance degradation of bearing and provides potential damage detections at an earlier stage. The reason is when the defects impact the rolling element bearing, it produces a signal that has probability density sharper than a normal bearing. In this case, kurtosis is the best measure that indicates the peaked distribution related to rolling element bearing fault. Literatures [23–25] have studied the use of kurtosis in assessing the bearing condition and have shown the effectiveness of kurtosis in bearing defect detection. After kurtosis is calculated, target vectors of failure probability are determined using a LR method. Simulation and experimental bearing run-to-failure data are used for validating the proposed method. RVM is used to train the kurtosis bearing data and target vectors of failure probability estimated by LR. After the training process, RVM is employed to predict the failure probability of an individual unit of machine component.

2. Theoretical background

2.1. Logistic regression (LR)

Logistic regression is a variation of ordinary regression method that is used when the dependent variable is a dichotomous variable (which is usually represented by the occurrence or non-occurrence of some output event, usually coded as 0 and 1). The goal of logistic regression is to find the best fitting model to describe the relationship between the dichotomous characteristic of the dependent variable and a set of independent variables [14].

In the logistic regression method, the dependent variable is the probability that an event will occur; hence, the output value has a discrete number of responses that is constrained between 0 and 1. The logistic function is

$$\text{Prob(event)} = P(\vec{x}) = \frac{1}{1 + e^{-g(\vec{x})}} = \frac{e^{g(\vec{x})}}{1 + e^{g(\vec{x})}} \quad (1)$$

where $P(\vec{x})$ is the probability of some output event, $\vec{x}(x_1, x_2, \dots, x_k)$ is an input vector corresponding to the independent variables (predictors), and $g(\vec{x})$ is the logit model. The logit model of multiple logistic regressions can be written as

$$g(\vec{x}) = \log\left(\frac{P(\vec{x})}{1 - P(\vec{x})}\right) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k \quad (2)$$

where $g(\vec{x})$ is a linear combination of the independent variables x_1, x_2, \dots, x_k and $\alpha, \beta_1, \beta_2, \dots, \beta_k$ are known as the regression coefficient. Logistic regression applies maximum likelihood estimation after transforming the dependent into a logit variable in order to calculate the parameter $\alpha, \beta_1, \beta_2, \dots, \beta_k$.

In this paper, logistic regression is used to estimate the failure probability of run-to-failure bearing data. This failure probability also represents the failure propagation from incipient failure coded as 0 to final failure conditions coded as 1. Furthermore, the failure probability result will be used as a target vector input of RVM.

2.2. Relevance vector machine (RVM)

RVM is special case of a sparse kernel model, which represents a Bayesian treatment of a generalized linear model of identical functional form to support vector machine (SVM). It differs with SVM in the case of solution, which provides

متن کامل مقاله

دریافت فوری ←

ISIArticles

مرجع مقالات تخصصی ایران

- ✓ امکان دانلود نسخه تمام متن مقالات انگلیسی
- ✓ امکان دانلود نسخه ترجمه شده مقالات
- ✓ پذیرش سفارش ترجمه تخصصی
- ✓ امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
- ✓ امکان دانلود رایگان ۲ صفحه اول هر مقاله
- ✓ امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
- ✓ دانلود فوری مقاله پس از پرداخت آنلاین
- ✓ پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات