Degradation Assessment and Fault Modes Classification Using Logistic Regression

Jihong Yan
e-mail: yanjh@uwm.edu

Jay Lee

NSF I/UCRC Center for Intelligent Maintenance Systems,
University of Wisconsin-Milwaukee,
Milwaukee, WI 53211

Real-time health monitoring of industrial components and systems that can detect, classify and predict impending faults is critical to reducing operating and maintenance cost. This paper presents a logistic regression based prognostic method for on-line performance degradation assessment and failure modes classification. System condition is evaluated by processing the information gathered from controllers or sensors mounted at different points in the system, and maintenance is performed only when the failure/malfunction prognosis indicates instead of periodic maintenance inspections. The wavelet packet decomposition technique is used to extract features from non-stationary signals (such as current, vibrations), wavelet package energies are used as features and Fisher's criteria is used to select critical features. Selected features are input into logistic regression (LR) models to assess machine performance and identify possible failure modes. The maximum likelihood method is used to determine parameters of LR models. The effectiveness and feasibility of this methodology have been illustrated by applying the method to a real elevator door system. [DOI: 10.1115/1.1962019]

Keywords: Prognostics, Logistic Regression, Degradation Assessment, Fault Modes Classification

1 Introduction

Considerable efforts have been made to develop methods and tools to diagnose failures. However, limited results have been given on prognostics that can detect, analyze and correct equipment problems well before failures actually manifest, also provide system operators with a sufficient time window to schedule maintenance without disrupting the operations [1–3].

This paper presents a prognostic method for on-line performance degradation assessment and root cause classification using multiple logistic regression (LR). The paper is organized as follows. Section 2 provides state of the art of the prognostic methodology along with related mathematics. Section 3 illustrates testbed setup of an elevator door system and assessment/classification results obtained from an application of the proposed schemes on real data. Section 4 concludes the paper with a summary and future research directions.

2 Methodology

The prognostic scheme is based on monitored data which contain equipment/machinery incipient failure signatures; intelligent mathematical techniques which can be incorporated to detect, evaluate the risk of failure over some protracted period of time and also classify which particular type of failure may occur.

2.1 Procedures of the Methodology. There are three major steps of the methodology. Step 1: signal pre-processing that involves data format conversion, sampling, normalization, denoising etc. Step 2: feature extraction and selection that extracts and determines appropriate features for degradation assessment and root cause classification as well as to reduce the search space for fast computation. Step 3: performance assessment and root cause classification by means of the logistic regression method.

2.2 Feature Extraction Using Wavelet Packet Decomposition. The wavelet packet transform (WPT) is a generalization of wavelet analysis to enhance the decomposition procedures. A wavelet packet function $\Psi^n_{j,k}(t)$, has three indices, where integers n, j and k are the modulation, the scale and the translation parameters, respectively

$$\Psi^n_{j,k}(t) = 2^{j/2}\Psi^n(2^j t - k) \quad n = 1,2,\ldots \tag{1}$$

The decomposed wavelet packet component signal $f^n_j(t)$ can be expressed by a linear combination of wavelet packet functions as follows:

$$f^n_j(t) = \sum_{k=-\infty}^{\infty} c^n_{j,k}\Psi^n_{j,k}(t). \tag{2}$$

The wavelet packet coefficients $c^n_{j,k}$ can be obtained from

$$c^n_{j,k} = \int_{-\infty}^{\infty} f(t)\Psi^n_{j,k}(t)dt. \tag{3}$$

Usually, direct assessment from all wavelet packet coefficients often leads to inaccurate decisions [4]. In our study, the wavelet packet node energy is defined as

$$e_{j,n} = \sum_k (c^n_{j,k})^2, \tag{4}$$

which measures the signal energy contained in some specific frequency band indexed by parameters j and k.

2.3 Feature Selection Using Fisher's Criterion. Feature selection in feature measurement space is to select the feature components that contain discriminant information and discard those feature components that provide little information. The feature subset can be selected from the available features that have larger criterion function values by using Fisher's criterion [5]. The feature components $\{f_l | \ell = 1,2,\ldots,n\}$ can be ranked as

$$J(f_1) \geqslant J(f_2) \geqslant \cdots \geqslant J(f_{n-1}) \geqslant J(f_n), \tag{5}$$

where $J(\bullet)$ is a criterion function for measuring the discriminant power of a specific feature component. Fisher's criterion was used as criterion function and is defined as

$$J_{f_l}(i,m) = \frac{|\mu_{i,f_l} - \mu_{m,f_l}|^2}{\sigma^2_{i,f_l} + \sigma^2_{m,f_l}}, \tag{6}$$

where μ_{i,f_l}, μ_{m,f_l} are the mean values of the lth feature, f_l, for class i and m, and σ^2_{i,f_l}, σ^2_{m,f_l} are the variance of the ℓth feature, f_l, for class i and m correspondingly.

The feature subset can be selected from the available features that have larger criterion function values. This will significantly simplify the design of the logistic regression classifier and enhance the generalization capability of performance assessment process.

2.4 Logistic Regression Method. The machine condition description from daily maintenance records/logs is a dichotomous problem (either normal or failed) which can be represented using a logistic regression function [6]. The goal of logistic regression is

Contributed by the Manufacturing Engineering Division for publication in the ASME JOURNAL OF MANUFACTURING SCIENCE AND ENGINEERING. Manuscript received September 18, 2003; final revision received July 22, 2004. Associate Editor: K. Danai.

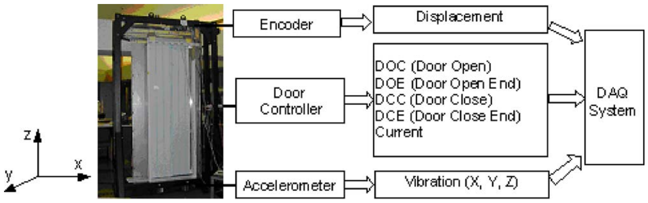

Fig. 1 Elevator door data acquisition system

to find the best fitting model to describe the relationship between the categorical characteristic of dependent variable (the probability of an event, constrained between 0 and 1) and a set of independent variables. The logistic function is

$$\text{Prob(event)} = P(\mathbf{x}) = \frac{1}{1+e^{-g(\mathbf{x})}} = \frac{e^{g(\mathbf{x})}}{1+e^{g(\mathbf{x})}}. \tag{7}$$

The logistic or logit model is

$$\text{Logit} = g(\mathbf{x}) = \log\frac{P(\mathbf{x})}{1-P(\mathbf{x})} = \alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_K x_K, \tag{8}$$

where $g(\mathbf{x})$ is a linear combination of the independent variables $x_1, x_2, \ldots x_K$.

The pre-condition for figuring out $P(\mathbf{x})$ is determining parameters α and $\beta_1, \ldots, \beta_K$ in advance. Due to the fact that the dichotomous dependent variable makes estimation using ordinary least squares inappropriate, rather than choosing parameters that minimize the sum of squared errors, estimation in logistic regression chooses parameters of α and $\beta_1, \ldots, \beta_K$ using the maximum likelihood method [6]. Then, the probability of failure for each input vector $\mathbf{x}$ can be calculated according to Eq. (7).

3 Application

The methodology has been implemented in an elevator door system (see, Fig. 1) to evaluate door performance dynamically. Also fault mode analysis was performed in order to find out the possible root cause.

3.1 Data Acquisition System Description. An encoder was mounted to measure door displacement. Four command switch signals and a current signal are acquired from the door control board, as well as three vibration signals acquired from an installed accelerometer, with a sampling rate as 500 Hz.

3.2 Feature Extraction and Selection. First of all, displacement and the corresponding command switch signals are used to calculate speed profile and cycle time of each open-close cycle. Second, six-level wavelet packet decomposition using Daubechies wavelet (DB4) was adopted for each vibration/current signal, then package energies were used as features. A subset of feature components was determined using Fisher's criterion. In this case, after feature selection, a 19-dimension feature vector (including maximum speed, cycle time, two packet energies from current, two packet energies from vibration-X, 11 packet energies from vibration-Y, two packet energies from vibration Z) is finally selected as feature vector for performance assessment and fault mode classification.

3.3 LR Models Training.

- LR model trained for performance assessment.
  80 open-close cycles were used as training data, including 40 normal cycles ($P(\mathbf{x})=0$) **versus** 40 fault cycles ($P(\mathbf{x})=1$). The parameters α, $\beta_1 \ldots \beta_{19}$ were estimated using the maximum likelihood method to eventually obtain the model for performance assessment as LR1.

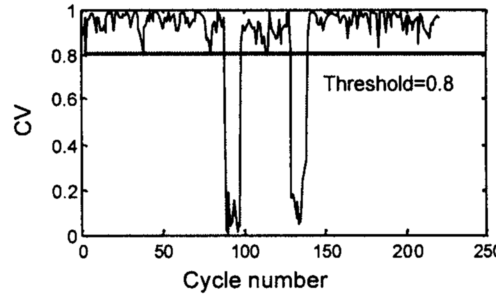

Fig. 2 Assessment result of the 220 cycles

- LR models trained for root cause classification.
  Fault mode 1 (friction): 20 friction cycles ($P(\mathbf{x})=1$) **versus** 60 non-friction cycles ($P(\mathbf{x})=0$) were used to train the classification model (LR2) for fault mode 1 using the maximum likelihood method;
  Fault mode 2 (vibration): 20 vibration cycles ($P(\mathbf{x})=1$) **versus** 60 non-vibration cycles ($P(\mathbf{x})=0$) (40 normal cycles and 20 friction cycles) were used to train the classification model (LR3) for fault mode 2 using the maximum likelihood method.

3.4 Validation. Two hundred twenty continuous door open-close cycles were used for validation. Note that the confidence value (CV) is calculated based on the probability of failure. Define $\text{CV}=1-P(\mathbf{x})$. When the machine is running normally, CV is close to 1, if the machine is going to fail, CV is approaching 0 correspondingly. If the confidence value is less than a prefixed threshold, for example, 0.8, the root cause classification module will be triggered, and features are input into fault modes classification models to calculate the probability of each fault. In the 220 open-close cycles friction was imposed on the door from cycles 89 to 98, also vibration was interfered from cycles 129 to 138. Figure 2 shows the overall performance assessment of the 220 cycles using model LR1, the probability of different fault modes conducted from LR2 and LR3 is shown in Fig. 3.

In Fig. 2, both friction and vibration problems can be detected from the two CV drops, but it is hard to clarify what the difference is between the two drops and what caused the drops. In this methodology, the root cause classification module is triggered as long as the confidence value is below a predetermined threshold (0.8) by inputting the corresponding features into the trained models (LR2 and LR3) to calculate the probability of fault modes. From cycles 89 to 98, the probability of fault 1 (friction) is very high, see the solid line in Fig. 3. The probability of fault 2 (vibration) is rather high from cycles 129–138 shown as the dashed line in Fig. 3. Also there are other points with CV slightly below threshold which might be caused by process noise or disturbances, consequently minor probability of failure of these points (judged as vibration fault) is indicated by circles in Fig. 3.

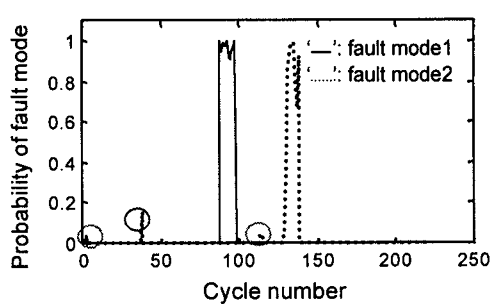

Fig. 3 Probability of failure modes 1 and 2

4 Conclusions

A logistic regression based approach for machine degradation evaluation and root cause classification has been performed in this paper. LR combined with maximum likelihood technique is an effective and efficient tool for performance assessment and root cause classification dynamically. WPT combined with Fisher's criterion is a suitable feature extraction step where appropriate features can be obtained from non-stationary signals. The method is generic and shows promising results for the analysis of both stationary and non-stationary signals, which could be applied to other industrial manufacturing systems.

ARMA model based time to failure (TTF) prediction has been accomplished using simulation data [1]. The prediction module implementation in real life application is in process. Also when the process is not time shifted, the coefficients of WPT can be used as features directly instead of using packet energy, which will be investigated further in future research and application.

References

[1] Yan, J., Lee, J., and Koc, M., 2002, "Predictive Algorithm for Machine Degradation Detection Using Logistic Regression," *Fifth International Conference on Managing Innovations in Manufacturing (MIM)*, Milwaukee, WI, 9–11 September, pp.172–178.

[2] Kacprzynski, G. J., and Roemer, M. J., 2000, "Health Management Strategies for 21st Century Condition-Based Maintenance Systems," *International COMADEM Congress*, Houston, TX, December.

[3] Roemer, M. J., and Kacprzynski, G. J., 2000, "Advanced Diagnostics and Prognostics for Gas Turbine Engine Risk Assessment," *Aerospace Conference Proceedings*, IEEE, Vol. 6, pp. 345–353.

[4] Yen, G. G., and Lin, K. C., 2000, "Wavelet Packet Feature Extraction for Vibration Monitoring," IEEE Trans. Ind. Electron., **47**(3), pp. 650–667.

[5] Fukunaga, K., 1992, *Introduction to Statistical Pattern Recognition*, Academic, NY.

[6] Houston, W. M., and Woodruff, D. J., 1997, "Empirical Bayes Estimates of Parameters from the Logistic Regression Model,"ACT Research Report Series, June.