An Approach for Remaining Useful Life Estimation Based on Combination of Multidimensional Information

Hongju Yue, Yanling Qian, Long Wang, Zezhong Li

National University of Defense Technology, Changsha, China

290934339@qq.com

Keywords: health management; remaining useful life; performance assessment; prognostics

Abstract. Previously, the traditional approaches for remaining useful life (RUL) prediction only use the one-dimensional information relating to the malfunction. In actually, it is very difficult to collect the information, and the accuracy is low. The RUL prediction is limited by previous reasons. Except for the traditional information relating to the malfunction, there are many kinds of information, such as operations, performance monitoring data and maintenance record, which can reflect the performance state and degenerative process. Obtaining and analyzing the multidimensional information in operating condition should be effective for modeling degeneration and the estimation of remaining useful life. This paper describes an approach for remaining useful life estimation based on combination of multidimensional information.

1 Introduction

RUL estimation is the most common task in Prognostics and Health management. In recent years, more and more engineering technical personnel and statistical workers measure the degradation data of certain performance to estimate the RUL of the product. This method is better than the reliability model based on failure, because it gets rid of the leakage of the experimental information and solves the problems of small samples, long life and correlation which the reliable model cannot achieve. Nelson [1] provided the degradation model and the review of related degradation model .For the linear regression model, Lu, etc, [2] put forward a kind of the linear degradation path model with random intercept and slope. However, the degradation path is a nonlinear function of time invariably. Lu, etc, [3] come up with a nonlinear mixed effect model and the two-step method to estimate model parameters.

The traditional approaches for remaining useful life (RUL) prediction only use the one-dimensional information relating to the malfunction. While the system will expose a lot of information in running state. We can get that the information contains the relationship with the real performance degradation. Using the measured sensitive data to confirm the health indicator (HI) of system and finding out the relationship with the real degradation of health indicator is a good way to create the real degradation model of system. Combining with multidimensional varying conditions in practice, this paper presents an approach for remaining useful life estimation based on combination of multidimensional information. We can take the fusion of working conditions and the characterization data as the HI of system, then structure the degradation model and estimate the RUL of system.

2 Methodology

The approach consists of two essential procedures: training phase and testing phase. Before the performance assessment, the feature extraction is necessary. It contains the classification of working conditions, the analysis of the change rule of every sensor, the selection of sensors with obvious degradation.
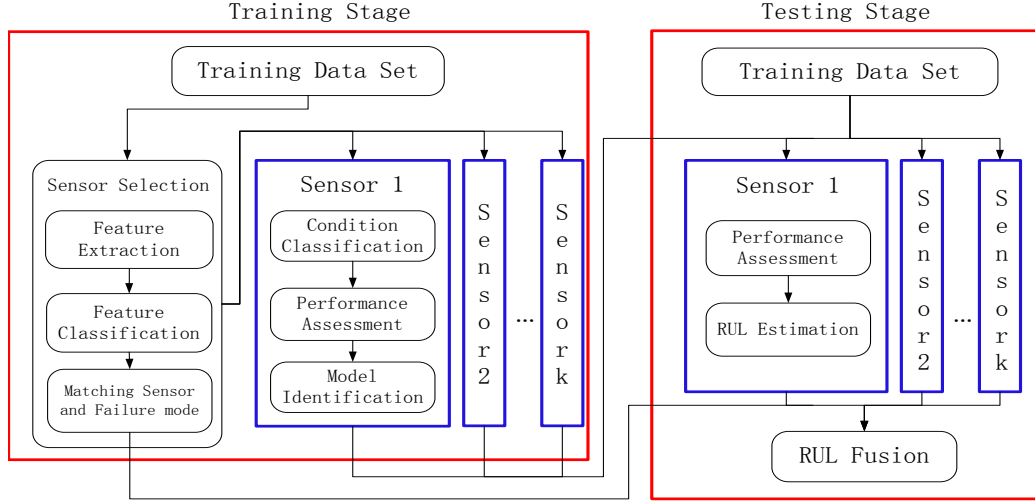
Figure 1. Flow chart of RUL estimation

A. training phase

The training phase consists of four steps:

1) Sensor selection

There may be many feature modes for one system. We should identify and simplify the modes first.It is called feature extraction.However,each sensor has different trends, but there are some rules.The classification for the trends is necessary. For example, the rise and fall of the sensor has obvious degradation trend, which can be used as the raw data of the transformation of HIs. Rise and fall can be understood as the corresponding health indicators of different faults, which can form a competitive failure model. The process is named as matching sensor and failure mode, and it provides the basis for the estimation.

2) Classification of working conditions

In practice, the working conditions have effects on the output of sensor and the degradation speed, so the process for each case should be analyzed separately. The classification of working conditions becomes necessary for the follow-up work. For the multidimensional, the dimension reduction processing is carried out.

3) Performance evaluation

In practice, the working conditions have effects on the output of sensor. The conditions are varying and multidimensional, so that the translations of sensors are different. The logistic regression method can translate multiple features into HI [6]. Beyond that, ARMA model can be used to predict machine performance. The history of each unit ends when it reaches one failure condition. Linear regression ignores this, so it cannot be conducted. In order to simplify the problem, we take the single sensor with obvious trend as the HI. Linear translation or index translation can be used to translate the sensor data into performance indicator (Linear translation in this paper).

$$y_i = f_k(x_i) + \varepsilon \qquad i = 1,2 \cdots m , \ k = 1,2 \cdots n. \tag{1}$$

4) Model identification

By analyzing the degradation mechanism, we can choose preliminary degradation model. In this case,index (nonlinear) regression model is used to describe the relationship between the cycle and health indicators:

$$y = a\left(e^{bt} - 1\right) + \varepsilon. \tag{2}$$

where are the model parameters. is the noise and it obeys normal distribution of the mean to 0. y is the HI. t is the cycle.

Using the training data set, the parameters can be estimated. Because of the difference between the manufacturing process and system performance of each unit,there is difference between each set of parameters.

B. testing phase

Testing phase consists of three steps

1) Performance assessment

Using the transfer function of health indicators of training data, the original sensor data of sensor from testing set can turn into the time sequence of HIs. Due to observation error or other factors, the transformation of HIs may be inaccurate. To estimate the end performance, Lsqcurvefit function can be used. After Getting the parameter set, the end HI can be obtained. Figure 2 shows the assessment of the HI.
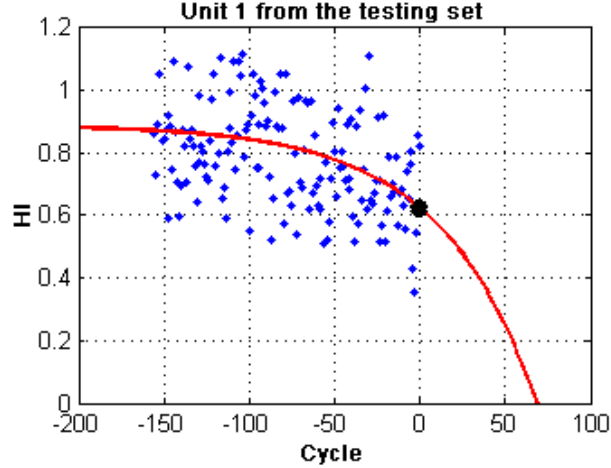

Figure 2. Assessment of the HI

2) RUL estimation

After putting the end HIs of the estimations into the curves of training set, the forecast results are showed out. When the values are histogram statistics. The maximum likelihood can be used to estimate the parameters. In order to normal distribution as an example,the probability density function(PDF)is

$$p(RUL) = \frac{1}{RUL\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(\ln RUL-\mu)^2}{2\sigma^2}} \tag{3}$$

And then to solve the following optimization problem

$$\max_{(\mu,\sigma)} L(\mu, \sigma) = \prod_{i=1}^{N} p(\mu, \sigma). \tag{4}$$

3）RUL Fusion

For single failure mode system, the above estimation is the final result. For the multiple, there are different failure modes, which are playing roles on the system.The final RUL of the test unit can be estimated through weighted sum of the obtained RULs:

$$RUL = \sum_{j}\left(\alpha_j \cdot RUL_j\right), \sum_{j} \alpha_j = 1. \tag{5}$$

Or the method of taking the minimum :

$$RUL = \min\left(RUL_j\right). \tag{6}$$

3 Application instance

The data set, provided by the 2008 PHM Data Challenge Competition[5], consists of multivariate time series that are collected from multiple units of an unspecified component. A library of degradation patterns can be created from these units with complete run-to-failure data (called training units). A unit whose remaining life

will be predicted (called a test unit) also has its historical data recorded continuously. Each time series is from a different instance of the same complex engineered system. The sensor data are contaminated with noise. As soon as one HI achieves the threshold value, the unit fails and stops working. Multiple failure modes are in series connection relationship.

The approach was chosen based on the following assumptions:

1) Record the run-to-failure historical data from multiple units of a common system/component;

2) Neglect the influence of different working conditions for the system performance degradation speed;

3) System has a finite number of independent failure modes, and any failure is possible in the experiment;

For the rise and fall, the most representative ones with the biggest change range are the sensor 25 and sensor 9. In summary, the final selection of sensor 9 and sensor 25 to analyze and research.
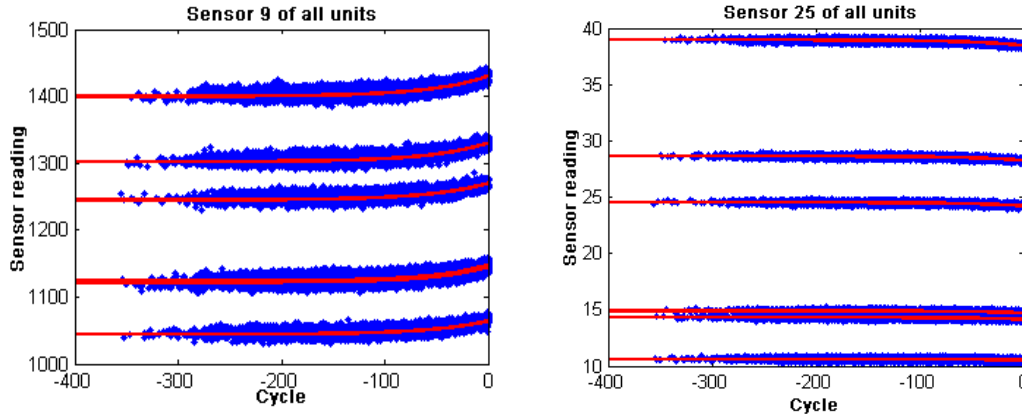

Figure 3. Data statistics and fitting curves of sensors 9 and 25

In order to simplify the problem, we believe that there are two different failure modes with sensor 9 and 25, which are playing a role on the system. To estimate the end HI, Lsqcurvefit function can be used. Figure 4 shows the predicted distribution of unit 1.
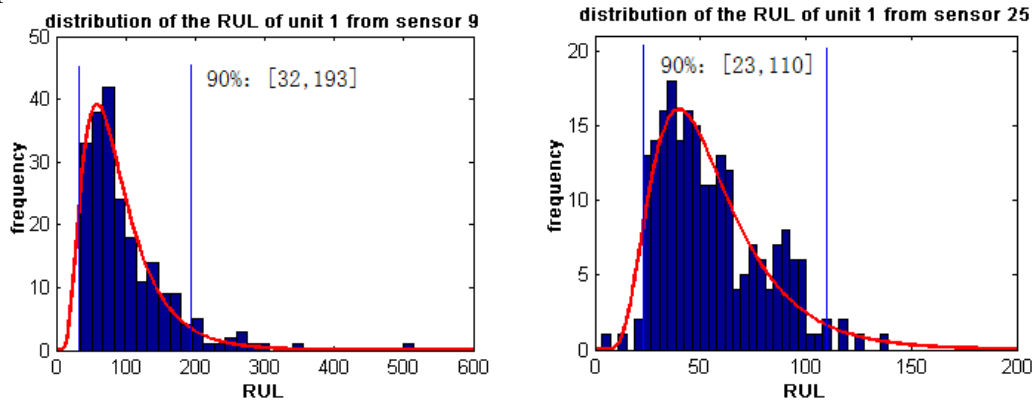

Figure 4. Distribution of the RUL of unit 1.(a) the RUL of sensor 9;(b) the RUL of sensor 25

The minimum function is the way to get the final prediction of the remaining life of the unit. The following is the formula:

$$RUL = \min\left(RUL_9, RUL_{25}\right)$$

（7）

4 Results and Discussions

The adequate training data sets provides the basis and methods for the degradation model transformations of all kinds of failure modes. Using the method that we get from the training data sets, we can assess the HIs of units at the end of the test. The HIs will be brought into the curves of the training data sets, then we can get the RUL statistics for one failure mode. For single failure mode system, the RUL distribution is the final result.

For multiple failure modes system, the RULs of each failure mode should be combined, and the fusion is the result. The estimation is compared with the truth that Phm08 provided. For the 10 units, there are nine in the 90% RUL distribution range. We can say that the approach is effective. The method makes full use of historical data .It filters and translates the multidimensional information. Due to the manufacturing process or environmental stress, the initial values of HI of each unit and individual attributes are both different. The method is suitable for both single failure mode system and multiple failure modes system. In the actual production, it is of great significance.

References

[1] Nelson W. Accelerated testing: Statistical methods, test plans and data analysis[M].New York,1990.

[2] Lu J C, Park J, Yang Q. Statistical inference of a time to failure distribution derived from linear degradation data[J].Technometrics,1997,39(4):391-400.

[3] Lu J C, Meeker W Q. Using degradation measures to estimate a time to failure distribution[J]. Technometrics, 1993, 35(2):161-174.

[4] Jay Lee SMITH T A.USAF condition-based maintenance plus(CBM+) initiative [R]. Alabama, USA: Air Force Logistics Management Agency,2003:3-19.

[5] 2008 PHM data challenge competition, [Online] http://www.phmconf.org/ OCS/ index. php/phm/2008/challenge.

[6] J. Yan, M. Koc, and J. Lee, . "A prognostic algorithm for machine performance assessment and its application,." Production Planning & Control, Vol. 15, No. 8, 2004, pp. 796.‐801.

[7] Meeker, W. Q. and Escobar, L. A.(1998).Statistical Methods for Reliability Data, New York: John Wiley and Sons.

[8] Gebraeel, N., Lawley, M., Liu, R., Parmeshwaran, V Residual life predictions from vibration-based degradation signals: a neural network approach[J]. IEEE Transactions on Industrial Electronics, 2004, 51(3): 694-700.

[9] A similarity-based prognostics approach for Remaining Useful Life estimation of engineered systems. 10.1109/PHM.2008.4711421.

[10] Si Xiaosheng, Wang Wenbin, Hu Changhua, et al. Remaining Useful Life Estimation－ A Review on the Statistical Data Driven Approaches. European Journal of Operational Research，2011，213(1) : 1－14.

[11] Sauter F C. A system Approach to Condition Based Maintenance［C］//American Helicopter Society Specialises'Meeting on Conditon Based Maintenance , February 12－13 2008.