Data-driven Machinery Prognostics Approach using in a Predictive Maintenance Model

Wenzhu Liao*
Department of Industrial Engineering, Chongqing University, Chongqing, P.R. China
Email: liaowz@cqu.edu.cn
*The corresponding author

Ying Wang
Department of Industrial Engineering and Logistics Engineering, Shanghai Jiao Tong University, Shanghai, P.R. China
Email: ying.w18@163.com

Abstract—Nowadays, more and more manufacturers realize the importance of adopting new maintenance technologies to enable systems to achieve near-zero downtime, so machinery prognostics that enables this paradigm shift from traditional fail-and-fix maintenance to a predict-and-prevent paradigm has arose interests from researchers. Machinery prognostics which could estimate machine condition and degradation strongly support predictive maintenance policy. This paper develops a novel data-driven machine prognostics approach to predict machine’s health condition and describe machine degradation. Based on machine’s prognostics information, a predictive maintenance model is well constructed to decide machine’s optimal maintenance threshold and maintenance cycles. Through a case study, this predictive maintenance model is verified, and the computational results show that this proposed model is efficient and practical.

Index Terms—prognostics, predictive maintenance, cost, optimization

I. INTRODUCTION

Many manufacturing systems suffer increasing wear with usage and age as a deterioration process, which causes low reliability and huge losses [1-2]. Nowadays, maintenance management becomes an important part in manufacturing systems and has been widely used to keep machines in good operation to decrease failures [3].

Since 1960s, the analysis and modeling of maintenance operations have aroused the interests from researchers [4]. Barlow and Hunter first proposed a simple replacement model with minimal repair, in which minimal repair is performed immediately after machine failure to restore the machine to its prior state before failure [5]. Based on this time-based maintenance model, a lot of scheduled maintenance models in which fixed time intervals is predetermined to perform maintenance operations (i.e. characteristics of the unavailability periods are known in advance) were developed. For example, Khandelwal, Sharma and Ray researched an application of periodic maintenance model for a machine [6]. Yak, Dillon and Forward used a periodic maintenance model to achieve the reliability requirements for a fault-tolerant computer system [7]. Comparing to the failure-based maintenance models, scheduled maintenance models show they are more positive and efficient [8]. However, how to decide maintenance interval is a crucial work [9-10]. If the maintenance interval is too long, although it can decrease maintenance operations so as to reduce maintenance cost, machine reliability will be low and more failures will occur, which leads to higher breakdown cost. If the maintenance interval is too short, although the machine remains in good operation condition, maintenance cost will be much higher. Hence, neither too long nor too short maintenance interval is suitable for maintenance model due to economic loss [11]. If the maintenance operations could be well planned according to machine’s condition, much more resources would be saved. Machine’s condition should thus be known ahead to help arrange flexible maintenance intervals for appropriate maintenance operations. Therefore, based on this scheme, preventive maintenance is studied for the optimization of those cases in which the maintenance operation is controllable [12-13]. However, nowadays there is some research on preventive maintenance models, most of them involve the use of common function distributions to describe machine degradation, which seems unpractical [14].

Therefore, it is essential to develop good maintenance planning based on machine’s real deterioration process. With the development of complex manufacturing process, Condition-based maintenance which is one kind of maintenance programs that recommends maintenance operations based on machine's condition is being implemented [15]. Condition-based maintenance attempts to avoid unnecessary maintenance tasks by performing operations only when there is evidence of abnormal behaviours of the machine, which requires machine’s condition information. In recent years, with the applications of embedded agent techniques and tether-free techniques [16], it is possible to monitor machine’s condition information continuously. However, most of available condition-based maintenance research only set several machine's failed cases, then identify current machine's condition whether belonging to those cases.
Obviously, this kind of condition-based maintenance ignores machine degradation.

Because traditional maintenance models usually ignore machine reliability, it becomes important to improve maintenance models if machine degradation can be well known [17]. This conducts predictive maintenance with machinery prognostics information nowadays. Predictive maintenance is one positive and useful maintenance methodology [18-19]. It attempts to avoid unnecessary maintenance tasks according to machine’s prognostics information [20]. Based on this scheme, prognostics that are the capability to provide early detection and isolation of the precursor and/or incipient fault condition to a machine failure condition has been applied to the field of predictive maintenance [21]. Prognostics is to know before, to predict the future as a result of rational study and analysis of available pertinent data [22]. In manufacturing processes, machine degradation obtained by machinery prognostics provides the important data for predictive maintenance [23]. Hence, it becomes vital to develop a good approach to estimate machine’s health condition and predict machine degradation to construct a suitable predictive maintenance model.

This paper gives a data-driven machinery prognostics approach to describe machine degradation. Then, based on this machine prognostics information, a predictive maintenance ($H^*_s, N^*$) model is proposed with the aim to minimize the long-term average cost. When machine’s health index reaches the maintenance threshold $H^*_s$, a predictive maintenance operation is performed to restore this machine. Once it is the $N^{th}$ time for the machine to reach the maintenance threshold $H^*_s$, the machine should be replaced.

II. PROBLEM DESCRIPTION

This paper studies a single machine maintenance arrangement problem in which the machine is not continuously available due to machine’s deterioration process. Hence, predictive maintenance operations with flexible time intervals during the planning horizon are performed. The maintenance operations which can reduce the increasing risk of machine failures are supposed to be able to restore the machine to a “as good as new” status (i.e., the machine is renewed). In general, machine’s natural life is a time period between machine installation and replacement. As maintenance operation can help the machine shift back to its operating status before failures, machine’s natural life may contain several maintenance cycles. Fig. 1 presents a machine’s deterioration process briefly. In this figure, a factor of machine’s health index is introduced to describe machine’s condition that ranges from 0 to 1. $H_o (0 \leq H_o \leq 1)$ denotes machine’s health index at the beginning. $H_{new}$ denotes machine’s health index when machine’s condition is “as good as new”, $H_{fail}$ denotes machine’s health index when the machine fails and $H_s$ denotes machine’s maintenance threshold (i.e., the machine should be maintained at the time). It can be viewed that there are several maintenance cycles during machine’s natural life. Along with time, machine’s health index degrades. As machine breakdowns will result in process interruption and huge losses, a safety maintenance threshold of $H_s$ is set before the degradation reaches $H_{fail}$ which will trigger a predictive maintenance operation, preventing the breakdown from occurring before maintenance can be conducted. The time period from $H_o$ to $H_s$ is denoted to be one predictive maintenance cycle which equals to $T_{H_s} - T_{H_o}$. Thus, predictive maintenance operations are performed based on this predicted time period, subject to machine’s deterioration process during the planning time horizon (seen in Fig.1).
III. THE MACHINERY PROGNOSTICS APPROACH

In this section, a data-driven machinery prognostics approach is developed to estimate and predict machine’s health condition. There are two steps:
1) Estimate machine’s health condition;
2) Predict machine’s future condition and describe machine degradation.

A. The Estimation of Machine’S Health Condition

First, because the collected monitoring data usually contains multidimensional data sets, feature extraction is applied to obtain dominant information. This paper uses PCA (PCA, principle component analysis) method for feature extraction and dimension reduction. Then, the dominant feature is clustered by considering SPR (SPR, statistical pattern recognition) approach. The main idea is: by designing a decision boundary, a given sample can be clustered into one certain pattern with the decision function. Given a set of patterns of machine’s condition \{\omega_1, \omega_2, \ldots, \omega_c\}, \(c\) is the number of patterns. If \(\mathbf{X} = (x_1, x_2, \ldots, x_j)^T\) is the feature vector for one certain machine’s condition, for two patterns of machine’s condition, shown as \(\omega_1\) and \(\omega_2\), \(\mathbf{X}\) is clustered into \(\omega_1\) (e.g. \(\mathbf{X} \in \omega_1\)) with the following decision criterion.

\[
l_i(\mathbf{X}) = \frac{p(\mathbf{X} | \omega_1)}{p(\mathbf{X} | \omega_2)} > p(\omega_2)
\]

where \(l_i(\mathbf{X})\) is the likelihood function. It illustrates the similarity between these two patterns.

Finally, based on the recognized patterns of machine’s condition, Chi-square test is used to estimate machine’s health index. Given a Gaussian distribution of multi-variables \(\bar{X} \sim MVN(\mu, \Sigma)\), through applying feature extraction and dimension reduction, it transforms to \(\bar{X} \sim MVN(0, I_p)\), thus, machine’s health index can be obtained by

\[
H(\bar{X}) = 1 - F_{X_p}(s \sum_{i=1}^{p} \hat{x}_i^2)
\]

where \(s\) is sensitivity factor (usually to be 0.25).

B. The prediction of machine degradation

With the above machine’s health condition information, a general machine performance prediction model can be constructed to predict machine’s health index, given as

\[
H_t - \sum_{i=1}^{p} \hat{\phi}_i \cdot H_{t-i} = \epsilon_t - \sum_{j=1}^{q} \hat{\theta}_j \cdot \epsilon_{t-j}
\]

where \{\(\epsilon_t\)\} is white noise, \(\hat{\phi}_1, \hat{\phi}_2, \ldots, \hat{\phi}_p\) is auto-regressive coefficient and \(\hat{\theta}_1, \hat{\theta}_2, \ldots, \hat{\theta}_q\) is moving average coefficient. \(p\) and \(q\) present their order, respectively. Equation (3) means machine’s future health index is influenced by previous health index and the turbulence occurred during this interval.

The mathematical statements of machine degradation prediction model are presented as followed. Given a set of samples of machine’s health index \{\(H_1, H_2, \cdots, H_n\)\}, where \(n\) is the length.

First, a regression factor \(\hat{\alpha}\) and residual variance \(\hat{\sigma}^2\) are estimated by

\[
\hat{\alpha}(n) = \left(\hat{\alpha}_1, \hat{\alpha}_2, \cdots, \hat{\alpha}_p\right)^T = \left[\hat{y}(j-i)\right]_{p \times (n-p)}^T \left[\hat{y}(i)\right]_{p \times 1}
\]

and

\[
\hat{\sigma}_p^2 = \hat{\gamma}(0) - \left[\hat{y}(i)\right]_{p \times 1}^T \left[\hat{y}(j-i)\right]_{p \times (n-p)} \left[\hat{y}(i)\right]_{p \times 1}
\]

where \(\hat{\gamma}(K) = \frac{1}{n} \sum_{t=1}^{n-K} H_t H_{t+K}\) and \(P_n\) is a multiple of \(\lg n\).

Then, residual \(\hat{\epsilon}_t\) is obtained to be

\[
\hat{\epsilon}_t = H_t - \sum_{i=1}^{p} \hat{\alpha}_i H_{t-i}
\]

Later, prediction model’s parameters \(\hat{\beta}\) and \(\hat{\sigma}^2_{pq}\) can be estimated as below

\[
\hat{\beta} = \left(\begin{array}{c} \hat{\phi} \\ \hat{\theta} \end{array}\right) = \left(\begin{array}{c} \hat{\gamma}_H (j-i)_{p \times p} \\hat{\gamma}_{H\epsilon} (j-i)_{p \times q} \\ \hat{\gamma}_{\epsilon H} (j-i)_{q \times p} \\hat{\gamma}_\epsilon (K)_{q \times q} \end{array}\right) \left(\begin{array}{c} \hat{\gamma}_H (K)_{p \times p} \\ \hat{\gamma}_{H\epsilon} (K)_{p \times q} \\ \hat{\gamma}_{\epsilon H} (K)_{q \times p} \\hat{\gamma}_\epsilon (K)_{q \times q} \end{array}\right)^{-1} \left(\begin{array}{c} \phi \\ \theta \end{array}\right)
\]

(7)

and

\[
\hat{\sigma}^2_{pq} = \hat{\gamma}(0) - \left(\begin{array}{c} \phi \\ \theta \end{array}\right)^T \left(\begin{array}{c} \hat{\gamma}_H (K)_{p \times p} \\ \hat{\gamma}_{H\epsilon} (K)_{p \times q} \\ \hat{\gamma}_{\epsilon H} (K)_{q \times p} \\hat{\gamma}_\epsilon (K)_{q \times q} \end{array}\right)^{-1} \left(\begin{array}{c} \phi \\ \theta \end{array}\right)
\]

(8)

where

\[
\hat{\gamma}_H^r (K) = \frac{1}{n-K} \sum_{t=1}^{n-K} H_t H_{t+K}, \quad K = 1, 2, \cdots, p
\]

(9)

\[
\hat{\gamma}_\epsilon^r (K) = \frac{1}{n-P_n-K} \sum_{t=1}^{n-K} \epsilon_t \epsilon_{t+K}, \quad K = 1, 2, \cdots, q
\]

(10)

\[
\hat{\gamma}_{H\epsilon}^r (K) = \frac{1}{n-P_n-K} \sum_{t=1}^{n-K} H_t \epsilon_{t+K}
\]

(11)
\[
\hat{\gamma}^H(K) = \hat{\gamma}^H(-K)
\]  

(12)

Finally, the order \( p \) and \( q \) are determined by SBC criterion [24], shown as

\[
SBC = \log \sigma_p^2 + \frac{(p+q+1) \log(n-p)}{(n-p)}
\]  

(13)

After the order \( p \) and \( q \) are determined, machine degradation prediction model can be well constructed and used to predict machine’s future health index, which can determine machine’s predictive maintenance cycle time.

Therefore, with the predicted health index information, the time when machine’s health index reaches to \( H_s \) can be predicted. Based on this prediction information, the predictive maintenance cycle can be determined (e.g. the first predictive maintenance cycle time is \( T_{H_s} - T_{H_k} \)). It is obvious that through applying this developed machinery prognostics approach, machine’s health index can be estimated and predicted and machine’s predictive maintenance cycle time can be determined, which could strongly support the construction of machine’s predictive maintenance model.

IV. THE CONSTRUCTION OF PREDICTIVE MAINTENANCE MODEL

This section establishes the mathematical framework for predictive maintenance model to prove the structural characteristics of the optimum subject to assumptions made as follows:

1) A single machine is studied;
2) The machine is subject to a deterioration process;
3) The machine begins a new degradation process after predictive maintenance operations;
4) The maintenance operation can be performed only once in a predictive maintenance cycle.

As mentioned above, with the advancements in sensor and intelligent prognostic technologies, machine’s health index can be obtained and machine degradation can be estimated and. In this model, the time period between each two adjacent maintenance operations is called one maintenance cycle. At the beginning of one cycle, the remaining time for conducting next maintenance will be determined first by machine’s deterioration information. Then with the obtained machine degradation information, the optimal planning for maintenance operations will be generated subject to the pre-determined objective.

Suppose there are the maintenance cost \( C_{pm} \) and the replacement cost \( C_o \), and both of them are constant. The operational cost \( C_o \) is variant that changes according to \( t \) and \( t \). The operational cost is constructed by three parts: \( C_{oc} \), \( C_{or} \) and \( C_{ov} \cdot t \), represents the fixed cost for operating; \( C_{or} \) represents the relative variant cost rate according to the maintenance cycles; and \( C_{ov} \) represents the relative variant cost rate according to time. Here, \( C_{or} \) and \( C_{ov} \) could be deduced from the history maintenance data of the system. And operational cost is written as:

\[
C_o = C_{oc} + C_{or} \cdot i + C_{ov} \cdot t
\]  

(14)

Then in order to minimize the long-term average cost, the cost function must be well constructed. During each predictive maintenance cycle, there should be: a probable predictive maintenance cost, a probable replacement cost and an operational cost. Here, \( \int_0^T C_o(i,t)dt \) represents the operational cost for each maintenance cycle, where \( T_i \) is the time interval for the \( i^{th} \) predictive maintenance cycle. Thus, machine’s expected long-term average cost \( ETC \) for each predictive maintenance cycle is inferred as:

\[
ETC_i = \frac{\int_0^T C_o(i,t)dt + C_{pm}}{T_i}, \quad 0 < i < N
\]  

(15)

This proposed predictive maintenance policy assumes that once it is the \( N^{th} \) time to reach the maintenance threshold \( H_s \), the machine should be replaced. Hence, the expected long-term average cost \( ETC \) for the \( N^{th} \) predictive maintenance cycle should be such that:

\[
ETC_N = \frac{\int_0^T C_o(i,t)dt + C_r}{T_i}, \quad i = N
\]  

(16)

Therefore, based on (15) and (16), from the machine installation to replacement, the expected long-term average cost \( ETC \) can be obtained. That is

\[
ETC = \frac{\sum_{i=1}^{N-1} ETC_i \cdot T_i + ETC_N \cdot T_N}{\sum_{i=1}^{N-1} T_i}, \quad 0 < i \leq N
\]  

(17)

With a given maintenance threshold in real situations, the entire optimization is implemented as a two-variables search, where the variable \( N \) is incremented. When the machine is available within a permitted operating region, the minimal \( ETC \) could be obtained by comparing all the local optimal results corresponding to different maintenance threshold \( H_s \). The procedures of the search algorithm are outlined as follows:

1) Fix the upper bound of predictive maintenance cycle \( N_{up} \) according to the related maintenance data.
2) Initialize \( C_{pm} \) beyond the \( N_{up}^{th} \) predictive maintenance as a very larger number, say \( 10^5 \).
3) Initialize \( ETC^* \) as a very larger number, say \( 10^7 \) (\( ETC^* \) is used to store the minimal \( ETC \)).
4) For the given machine’s health index...
maintenance threshold region \([H_1, H_2] \) where 
\( H_1 < H_2 \), let \( H_s = H_1 \).

5) Search \( N \) from one in step of one until the value of \( ETC \) cannot be further reduced. For a given value of \( N \):

(5.1) Calculate \( \{T_1, \ldots, T_N\} \) by (3) \( \sim \) (13).

(5.2) Calculate \( ETC \) by (17).

(5.3) If the calculated \( ETC \) is smaller than the current \( ETC^* \), replace the current \( ETC^* \) by the calculated \( ETC \) (i.e. \( ETC^* = ETC \)). And the current value of \( N \) and \( \{T_1, \ldots, T_N\} \) are stored as the local optimal result, \( N^* = N \).

6) Let \( H_s = H_s + 0.01 \), if \( H_s \leq H_2 \), return to step (5). Otherwise, stop.

Note that the traversal of machine's health index belonging to maintenance threshold region is spaced with unit 0.01, as such kind of precision of health index can satisfy real maintenance processes. Also, it can simplify computation effectively. At the end of the entire search, the minimal \( ETC^* \) (i.e. the optimal \( ETC^* \)) can be identified, thus the corresponding predictive maintenance policy \((H_s^*, N^*)\) is determined. Moreover, the obtained maintenance cycle time \( \{T_1, \ldots, T_N\} \) could strongly help manufacturers prepare those maintenance operations.

V. A CASE STUDY

In this section, a predictive maintenance model is constructed for a drilling machine. As machine’s spindle load signal can well reflect machine’s health condition according to maintenance engineers’ experience, it is collected to be used for data analysis. Because machine’s spindle load signal is acquired to estimate and predict machine’s health index, some sample data is collected. First, 50 sets of samples of machine’s perfect condition (seen in Fig.2) and 30 sets of samples of machine’s failure condition (seen in Fig.3) are collected. For these data, its sampling interval is 135 seconds. By training these samples, the patterns of “good” and “bad” machine conditions are recognized.

By using this prediction model to predict machine’s health index, the comparison between predicted results and real results is provided. Fig. 4 shows the comparison between the actual health index and the predicted health index for these collected 500 sets of samples. It can be seen that the 85% of variances belong to \([-2\sigma, +2\sigma]\) range, which means the result satisfies T-test or F-test. Thus, it can be proved that this performance prediction model is reasonable and effective.

Finally, assume machine’s health index maintenance threshold is \( H_s = 0.30 \) for performing maintenance operations. By applying this machinery prognostics approach, it can be obtained that machine’s residual life is 17.68 hours. Through monitoring machine’s real operation process, it is found that after 17.50 hours, machine’s health index reaches 0.30. The computation error is less 0.5%. Then, by comparing the predicted results with the actual results, it can be seen that the error of the comparison between machine’s predicted health index and the actual health index is less than 0.07 (seen in Fig.5). For the error percentage of the comparison between machine’s predicted health index and the actual health index, it is less than 1.2% (seen in Fig. 6). These two computational results can satisfy T-test or F-test,
which means the computational prediction model is good and effective. Thus, the computational results shows this developed data-driven machinery prognostics approach can well describe machine degradation.

![Figure 4. Comparison between actual data and estimated data of the 500 samples.](image)

Figure 4. Comparison between actual data and estimated data of the 500 samples.

![Figure 5. The error of the comparison between machine’s predicted health index and the actual health index.](image)

Figure 5. The error of the comparison between machine’s predicted health index and the actual health index.

Later, in order to construct the predictive maintenance model, assume the related cost factors in this maintenance model are: $C_{pm} = 500$ and $C_r = 5000$. And for the parameters of the variant operational cost, they are: $c_{oo} = 40$, $c_{oi} = 10$, $c_{vi} = 3$. Generally, the cost factors are required to be pre-designed ahead by technicians. In real manufacturing processes, maintenance engineers usually should be responsible for the design of these cost factors. In addition, the time of maintenance operations is negligible.

![Figure 6. The error percentage of the comparison between machine’s predicted health index and the actual health index.](image)

Figure 6. The error percentage of the comparison between machine’s predicted health index and the actual health index.

![Figure 7. Relationship between $ETC$ and $H_s$.](image)

Figure 7. Relationship between $ETC$ and $H_s$.

According to the procedures of the search algorithm presented above, the optimal result for this predictive maintenance model could be solved. The computational results with the aim to minimize the long-term average cost are presented in Fig.7. It indicates that among the whole maintenance threshold region, there is a minimal long-term average cost. The minimal long-term average cost is $ETC = 148.47$. Based on this minimal result, it can be obtained that its corresponding optimal predictive maintenance policy is $(0.31, 6)$. This $(0.31, 6)$ predictive maintenance policy indicates that in order to minimize the long-term average cost, the maintenance threshold of machine’s health index should be set to be 0.31. Once machine’s health index degrades to the maintenance threshold 0.31, one predictive maintenance operation should be performed. And for the whole maintenance process, the machine should run 6 predictive maintenance cycles. When it is the 6th time for the machine to reach the maintenance threshold of machine’s health index, the machine should be replaced.
IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING, VOL. 7, NO. 2, APRIL 2010

VI. CONCLUSIONS

Because the traditional maintenance policies usually ignore machine degradation, this paper is devoted to propose a predictive maintenance model considering machine’s deterioration process to minimize the long-term average cost. As an increasing importance of machinery prognostics for modern manufacturing process nowadays, machinery prognostics approaches that can describe machine degradation should be researched to support predictive maintenance planning. Thus, this paper develops a novel data-driven machinery prognostics approach to assess and predict machine’s health index. Through the case study about a drilling tool, the computational results show that this developed machinery prognostics approach can be well verified. And based on this, a predictive maintenance model is provided. The computational results can well match machine’s real degradation, which proves this predictive maintenance model is efficient and practical. With this proposed predictive maintenance model, the cost can be reduced and the time intervals for maintenance cycles can be decreased.

REFERENCES


Wenzhu Liao is a researcher of the Department of Industrial Engineering, Chongqing University, P. R. China. She received her Doctor’s Degree from Shanghai Jiao Tong University, P.R. China. Her 20 papers have been published. Her main research interests include maintenance planning and modeling, quality management and control technologies, reliability management, integrated manufacturing process and project management.

Ying Wang is a graduate student of the Department of Industrial Engineering and Logistics Management, Shanghai Jiao Tong University, P.R. China. She has published 5 papers. Her research interests include maintenance management, quality management and control technologies, statistical process control, automatic process control and cost management.