

Research Article

A Hidden Semi-Markov Model with Duration-Dependent State Transition Probabilities for Prognostics

Ning Wang,¹ Shu-dong Sun,² Zhi-qiang Cai,² Shuai Zhang,² and Can Saygin³

¹Department of Automobile, Chang'an University, Xi'an 710064, China

²Department of Industrial Engineering, Northwestern Polytechnical University, Xi'an 710072, China

³Mechanical Engineering Department, University of Texas San Antonio, San Antonio, TX 78249, USA

Correspondence should be addressed to Ning Wang; ningwang@chd.edu.cn

Received 5 January 2014; Accepted 7 March 2014; Published 14 April 2014

Academic Editor: Manyu Xiao

Copyright © 2014 Ning Wang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Realistic prognostic tools are essential for effective condition-based maintenance systems. In this paper, a Duration-Dependent Hidden Semi-Markov Model (DD-HSMM) is proposed, which overcomes the shortcomings of traditional Hidden Markov Models (HMM), including the Hidden Semi-Markov Model (HSMM): (1) it allows explicit modeling of state transition probabilities between the states; (2) it relaxes observations' independence assumption by accommodating a connection between consecutive observations; and (3) it does not follow the unrealistic Markov chain's memoryless assumption and therefore it provides a more powerful modeling and analysis capability for real world problems. To facilitate the computation of the proposed DD-HSMM methodology, new forward-backward algorithm is developed. The demonstration and evaluation of the proposed methodology is carried out through a case study. The experimental results show that the DD-HSMM methodology is effective for equipment health monitoring and management.

1. Introduction

Fault is a change from the normal operating condition of a system to an abnormal condition, which occurs as a result of system performance degradation over time [1]. Diagnostics indicates the occurrence of a fault and its root cause. Prognostics is fault prediction method; it involves detection of a pending fault before it occurs, identifying its root cause and estimating the remaining useful life (RUL), which is also known as time-to-failure [2]. Condition-based Maintenance (CBM) is a maintenance program that recommends maintenance actions based on the information collected through condition monitoring.

A CBM program can be used to do diagnostics or prognostics; however, regardless of the application, it follows three steps [3–5]. First, data relevant to events and system health are collected through data acquisition techniques. Data acquisition in CBM includes event-type data (i.e., information of what happened) and condition monitoring data, which are the measurements related to system health.

Second, event and condition monitoring data are interpreted for better understanding in the data processing step. Finally, maintenance decisions are made based on the interpretation and analysis of data. In particular, to identify the weakest components and states and improve the efficiency of CBM, the integrated importance measure of multistate system was introduced by Si et al. [6, 7]. An extensive survey on machine diagnostics and prognostics implementing condition-based monitoring can be found in Jardine et al. [8] and Heng et al. [9].

Data analysis for event data only is reliability analysis, which maps the event data over a time axis to determine the probability of events and uses the probability distribution to predict failures. On the other hand, data acquisition in CBM provides event and condition monitoring data. Therefore, it is more effective to combine events and conditions in a model in order to do diagnostics or prognostics. Hidden Markov model (HMM) is a technique for modeling and analyzing event and condition monitoring data together. It consists of two stochastic processes: (1) a Markov chain with

finite number of states describing an underlying mechanism and (2) an observation process depending on the hidden state [10–12]. An HMM contains finite states connected by transitions. Each state is characterized by a transition probability and an observation probability [13].

Researchers have proposed a number of techniques to address these limitations. Continuous variable duration HMM is adopted in the speech recognition. Compared to standard HMM, results show that the absence of a correct duration model increases the error rate by 50% [14–16]. Another example is in handwritten word recognition area; due to the inherent ambiguity related to the segmentation process in handwritten words, it is a practical idea to use the variable duration model for the states in a HMM-based handwritten word recognition system [17, 18]. Recently, some researchers apply HMM in the area of diagnostics and prognostics in machining process [19, 20]. However, these studies use only ordinary HMM technique. The inherent limitation of HMM as mentioned above still exists in these models.

Prognostic methods used in CBM are often a combination of statistical inference and machine learning methods [4, 21–24]. Model-based methods assume that measured information is stochastically correlated with the actual machine condition. HMM identifies the actual machine conditions from observable monitored data through a statistical approach. HMM has been very effective in various applications ranging from speech recognition [10, 14, 25–27] to tool wear monitoring and machining [3, 20, 28, 29].

The primary advantage of HMM is its robust mathematical foundation that can allow for many practical applications and different areas of use. An added benefit of employing HMMs is the ease of model interpretation in comparison with pure “black-box” modeling methods such as artificial neural networks that are often employed in advanced diagnostic models [28]. However, an inherent limitation of HMM approach is that its state duration follows an exponential distribution. In other words, HMM does not provide adequate representation of temporal structure.

To overcome the limitations of HMM in prognosis, Dong and He [30, 31] propose a Hidden Semi-Markov Model-based (HSMM) methodology by adding an explicit temporary structure into HMM to predict RUL of equipment. In this model, the states of HSMM are used to represent the health status of equipment. The trained HSMM can be used to diagnose the health state of equipment. Through parameter estimation of the health-state duration probability distribution and the proposed backward recursive equations, the RUL of the equipment can be predicted [32]. Although the results from HSMM are promising, the deterioration in the same state of the system is not taken into consideration in this model. It assumed that the state transition probabilities stay the same in the same state, which assumes all observations are independent, which typically does not hold in real world applications.

This paper presents a new approach that expands the HSMM methodology [30, 31, 33] with duration-dependent state transition probabilities. Different from HSMM, the proposed Duration-Dependent Hidden Semi-Markov Model

(DD-HSMM) does not follow the unrealistic Markov chain’s memoryless assumption and therefore provides a more realistic and powerful modeling and analysis capability. The major contribution of the DD-HSMM methodology is that it allows explicit modeling of the transition probabilities, which (1) do not only depend on the state but also (2) vary with the duration of each state, and (3) it provides the capability to relax observations’ independence assumption by accommodating a link between consecutive observations, which makes it more realistic in real world applications.

2. Theoretical Background

2.1. Description of General HSMM. A Hidden Semi-Markov Model (HSMM) is an extension of HMM by allowing the underlying process to be a semi-Markov chain with a variable duration or sojourn time for each state. The HSMM model is an ideal mathematical model for estimating the unobservable health states with observable sensor signal. For example, a small change in a bearing alignment could cause a small nick in the bearing, which could cause scratches in the bearing race and additional nicks, leading to complete bearing failure. This process can be well described by the HSMM. Let s_t be the hidden state at time t and O the observation sequence; a HSMM is characterized by its parameters. The parameters of a HSMM are as follows: the initial state distribution (denoted by π), the transition model (denoted by A), the observation matrix (denoted by C), and the state duration distribution (denoted by D). Thus, a HSMM can be written as $\lambda = (\pi, A, C, D)$. For a given state S , C is the probability matrix of observation being o_t at time t and o_{t+1} at time $t + 1$.

2.2. State Transitions. In HSMM, there are N states, and the transitions between the states are according to the transition matrix A ; that is, $P(i \rightarrow j) = a_{ij}$. Similar to standard HMM, the state s_0 at time $t = 0$ is a special state “START.”

Although the distinct health-state transition $s_{q_{t-1}} \rightarrow s_{q_t}$ is Markov,

$$P[s_{q_t} = j | s_{q_{t-1}} = i] = a_{ij}, \quad (1)$$

the state transition $s_{t-1} \rightarrow s_t$ is usually not Markov. It is the reason why the model is called “semi-Markov” [32], which means in the HSMM case, the conditional independence between the past and the future is only ensured when the process moves from one health state to another distinct health state.

2.3. Inference Procedures. Similar to HMM, HSMM also has basic problems to deal with, that is, evaluation, recognition, and training problems:

- (1) evaluation (also called classification): given the observation sequence $O = o_1 o_2 \cdots o_T$ and a HSMM λ , what is the probability of the observation sequence given the model, that is, $P(O | \lambda)$;
- (2) decoding (also called recognition): given the observation sequence $O = o_1 o_2 \cdots o_T$ and a HSMM λ , what

sequence of hidden states $S = s_1 s_2 \cdots s_T$ most probably generates the given sequence of observations;

- (3) learning (also called training): how do we adjust the model parameters $\lambda = (\pi, A, D, B)$ to maximize $P(O | \lambda)$.

Different algorithms have been developed for above three problems. The most straightforward way of solving the evaluation problem is enumerating every possible state sequence of length T (the number of observations). However, the computation burden for this exhaustive enumeration is prohibitively high. Fortunately, there is a more efficient algorithm that is based on dynamic programming, called forward-backward procedure. The goal for decoding problem is to find the optimal state sequence associated with the given observation sequence. The most widely used optimality criterion is to find the single best state sequence (path), that is, to maximize $P(S | O, \lambda)$ that is equivalent to maximizing $P(S, O | \lambda)$. Viterbi algorithm is used to find this single best state sequence, which is based on dynamic programming methods. For learning problem, there is no known way to obtain analytical solution. However, we can adjust the model parameter $\lambda = (\pi, A, D, B)$ such that $P(O | \lambda)$ is locally maximized using an iterative procedure, such as the Baum-Welch method (or equivalently the Expectation-Maximization algorithm).

3. Inference and Learning Mechanisms of DD-HSMM

3.1. Model Structure. Although HSMM has explicit state duration probability distribution $P_i(d)$, the state transition probabilities a_{ij} are duration invariant. In this paper, we replace duration-invariant state transition probabilities with duration-dependent state transition probabilities. The parameters for a DD-HSMM are as follows: the initial state distribution denoted by $\pi = \{\pi_i\}$ $1 \leq i \leq N$ ($\pi_i = P[s_1 = i]$ $1 \leq i \leq N$); the transition model denoted by $A = \{a_{ij}(d)\}$ ($1 \leq i, j \leq N, 1 \leq d \leq D$); the observation matrix denoted by $B = \{b_i(k)\}$ ($b_i(k) = P[v_k | s_t = i]$, where $1 \leq i \leq N, 1 \leq k \leq M$, M is the observation number in i , and $v_i = \{v_1, v_2, \dots, v_M\}$ is the observation symbols in state i); and the state duration distribution denoted by $D(P_i(d))$ ($i = 1, 2, \dots, N$). Thus, a DD-HSMM can be written as $\lambda = (\pi, A, D, B)$. Here, for the duration in given state S is $d_t(i) = d$, the state transition probability $a_{ij}(d) = P(q_{t+1} = j | q_t = i, d_t(i) = d)$, $1 \leq i, j \leq N, 1 \leq d \leq D_i$ (N is the state number, D_i is the max staying time in state i). And the state transition probabilities satisfy the constraint: $\sum_{j=1}^N a_{ij}(d) = 1, 1 \leq d \leq D_i, 1 \leq i \leq N$.

3.2. Duration-Dependent State Transition Probability. In DD-HSMM, the state transition probability distribution $A = \{a_{ij}(d)\}$, $1 \leq i, j \leq N, 1 \leq d \leq D$. We define duration-dependents state transition probabilities as follows:

$$a_{ij}(d) = P(q_{t+1} = j | q_t = i, d_t(i) = d), \quad (2)$$

where N and D are the number of states and the maximum duration in any states, respectively. Equation (2) represents the transition from state i to state j , given that the duration in state i at time t is $d_t(i) = d$. It indicated that, in the DD-HSMM case, the state transition probability is not only state dependent, but also duration variant.

3.3. Inference Procedures. Similar to HSMM, DD-HSMM also has basic problems to deal with, that is, evaluation, recognition, and training problems. To facilitate the computation in the proposed DD-HSMM-based health prediction model, in the following, new forward-backward variables are defined and modified forward-backward algorithm is developed.

A dynamic programming scheme is employed for the efficient computation of the inference procedures. To implement the inference procedures, a forward variable $\alpha_t(i, d)$ is defined as the probability of generating o_1, o_2, \dots, o_t and ending in state i and the duration $d_t(i) = d$:

$$\alpha_t(i, d) = P(o_1, o_2, \dots, o_t, q_t = i, d_t(i) = d | \lambda). \quad (3)$$

The initial conditions are established at time $t = 1$ as follows:

$$\begin{aligned} \alpha_1(1, d_1) &= 1, \\ \alpha_1(j, 1) &= \alpha_1(1, d_1) a_{1j}(d_1) b_j(o_1), \end{aligned} \quad (4)$$

$$\alpha_1(N, d_N) = \sum_{i=2}^N \alpha_1(i, 1) P_i(1) a_{iN}(1).$$

All unspecified α values are zero. For time $t = 2, \dots, T$,

$$\begin{aligned} \alpha_t(j, 1) &= \sum_{\substack{i=2 \\ i \neq j}}^{N-1} \sum_{d=1}^{D_i} \alpha_{t-1}(i, d) P_i(d) a_{ij}(d) b_j(o_t), \\ \alpha_t(j, d) &= \alpha_{t-1}(j, d-1) b_j(o_t), \end{aligned} \quad (5)$$

$$\alpha_t(N, d_N) = \sum_{i=2}^{N-1} \sum_{d=1}^{D_i} \alpha_t(i, d) P_i(d) a_{iN}(d),$$

where $a_{ij}(d)$ is the state transition probability from state i to state j , given that the duration in state i at time t is $d_t(i) = d$. $b_j(o_t)$ is the output probability of observation vector o_t from state j and $p_i(d)$ is the state duration probability of state i . N is the number of states in DD-HSMM and D_i is the maximum duration in state i .

Similar to the forward variable, the backward variable can be written as

$$\beta_t(i, d) = P(o_t, \dots, o_T | q_t = i, d_t(i) = d, \lambda). \quad (6)$$

For the backward probability, the initial conditions are set at time $t = T$ as follows:

$$\beta_T(N, d_N) = 1,$$

$$\beta_T(i, d) = P_i(d) a_{iN}(d), \quad (7)$$

$$\beta_T(1, d_1) = \sum_{j=2}^{N-1} \alpha_{1j}(d_1) b_j(o_T) \beta_T(j, 1).$$

For time $0 < t < T$,

$$\beta_t(i, d) = P_i(d) \sum_{\substack{j=2 \\ j \neq i}}^{N-1} a_{ij}(d) b_j(o_{t+1}) \beta_{t+1}(j, 1) + b_j(o_{t+1}) \beta_{t+1}(i, d+1) \quad 1 < i < N, \quad (8)$$

$$\beta_t(1, d_1) = \sum_{j=2}^{N-1} a_{1j}(d_1) b_j(o_t) \beta_t(j, 1). \quad (9)$$

Then the total probability can be computed by

$$P_r = P(O | \lambda) = \sum_{i=1}^N \sum_{d=1}^{D_i} \alpha_t(i, d) \beta_t(i, d). \quad (10)$$

3.4. Modified Forward-Backward Algorithm for DD-HSMM.

In order to give reestimation formulas for all variable of the DD-HSMM, one DD-HSMM-featured forward-backward variable is defined:

$$\xi_t(i, j, d) = P(t = q_n, s_t = i, t' = q_{n+1}, s_{t'} = j, d_t(i) = d | O, \lambda'). \quad (11)$$

In this equation, $\xi_t(i, j, d)$ is the probability of state transition from state i to state j at time $t + 1$ after being in state i for a duration of $d_t(i) = d$, given the model λ and observation O . From the definition of the forward-backward variables, we can derive $\xi_t(i, j, d)$ as follows:

$$\begin{aligned} & \xi_{t,t'}(i, j, d) \\ &= P(t = q_n, s_t = i, t' = q_{n+1}, s_{t'} = j, \\ & \quad d_t(i) = d | O_0^T, \lambda) \\ &= P(t = q_n, s_t = i, t' = q_{n+1}, s_{t'} = j, \\ & \quad d_t(i) = d, O_0^T | \lambda) \\ & \quad \times (P(O_0^T | \lambda))^{-1} \\ &= P(t = q_n, s_t = i, t' = q_{n+1}, s_{t'} = j, \\ & \quad d_t(i) = d, O_0^t, O_{t+1}^t, O_{t+1}^t | \lambda) \\ & \quad \times (P(O_0^T | \lambda))^{-1} \\ &= P(t = q_n, s_t = i, d_t(i) = d, O_0^t | \lambda) \\ & \quad \times P(t' = q_{n+1}, s_{t'} = j, d_t(i) = d, O_{t+1}^t | t \\ & \quad = q_n, s_t = i, d_t(i) = d, \lambda) \\ & \quad \times (P(O_0^T | \lambda))^{-1} \\ &= \alpha_t(i, d) P(t' = q_{n+1}, s_{t'} = j, O_{t+1}^t, O_{t+1}^t | t \\ & \quad = q_n, s_t = i, d_t(i) = d, o_t, \lambda) \\ & \quad \times (P(O_0^T | \lambda))^{-1} \\ &= \alpha_t(i, d) P(t' = q_{n+1}, s_{t'} = j | t = q_n, \\ & \quad s_t = i, d_t(i) = d, \lambda) \\ & \quad \times P(O_{t+1}^t, O_{t+1}^t | t = q_n, s_t = i, \\ & \quad t' = q_{n+1}, s_{t'} = j, d_t(i) = d, o_t, \lambda) \\ & \quad \times (P(O_0^T | \lambda))^{-1} \\ &= \alpha_t(i, d) a_{ij}(d) P(O_{t+1}^t, O_{t+1}^t | t = q_n, s_t = i, \\ & \quad t' = q_{n+1}, s_{t'} = j, d_t(i) = d, o_t, \lambda) \\ & \quad \times (P(O_0^T | \lambda))^{-1} \\ &= \begin{cases} \frac{\alpha_t(1, d_1) a_{1j}(d_1) b_j(o_t) \beta_t(j, 1)}{P(O_0^T | \lambda)}, & i = 1 \\ \alpha_t(i, d) p_i(d) a_{ij}(d) \\ \quad \times b_j(o_{t+1}) \beta_{t+1}(j, 1) \\ \quad \times (P(O_0^T | \lambda))^{-1}, & 2 \leq i, j < N \\ \alpha_t(i, d) p_i(d) \\ \quad \times a_{iN}(d) \beta_t(i, d) \\ \quad \times (P(O_0^T | \lambda))^{-1}, & 2 \leq i < N, j = N. \end{cases} \end{aligned} \quad (12)$$

Then, we have

$$\begin{aligned} & \xi_t(i, j, d) \\ &= P(t = q_n, s_t = i, t' = q_{n+1}, s_{t'} \\ & \quad = j, d_t(i) = d | O, \lambda') \end{aligned}$$

$$= \begin{cases} \frac{1}{P_r} \alpha_t(1, d_1) a_{1j}(d_1) b_j(o_t) \beta_t(j, 1), & i = 1 \\ \frac{1}{P_r} \alpha_t(i, d) p_i(d) a_{ij}(d) \\ \quad \times b_j(o_{t+1}) \beta_{t+1}(j, 1), & 2 \leq i, j < N \\ \frac{1}{P_r} \alpha_t(i, d) p_i(d) & 2 \leq i < N, j = N. \\ \quad \times a_{iN}(d) \beta_t(i, d), & \end{cases} \quad (13)$$

and the probability in state i at time t with duration of d is defined as $\gamma_t(i, d)$, and, from the definition of the forward-backward variables, we can easily derive $\gamma_t(i, d)$ as follows:

$$\gamma_t(i, d) = p(q_t = i, d_t = d | O, \lambda) = \frac{1}{P_r} \alpha_t(i, d) \beta_t(i, d). \quad (14)$$

The forward-Backward algorithm computes the following probabilities.

Forward Pass. The forward pass of the algorithm computes $\alpha_t(i, d)$.

Step 1 (initialization ($t = 1$)). The forward variable is shown as follows:

$$\alpha_1(i, d) = \begin{cases} 1 & j = 1, d = d_1 \\ \alpha_1(1, d_1) a_{1j}(d_1) b_j(o_1) & 1 < j < N, d = 1 \\ \sum_{i=2}^N \alpha_1(i, 1) P_i(1) a_{iN}(1) & j = N, d = d_N. \end{cases} \quad (15)$$

Step 2 (forward recursion ($t > 1$)). For $t = 2, \dots, T$,

$$\alpha_t(j, d) = \begin{cases} \sum_{i \neq j}^{N-1} \sum_{d=1}^{D_i} \alpha_{t-1}(i, d) P_i(d) a_{ij}(d) b_j(o_t) & d = 1 \\ \alpha_{t-1}(j, d-1) b_j(o_t) \\ \sum_{i=2}^{N-1} \sum_{d=1}^{D_i} \alpha_t(i, d) P_i(d) a_{iN}(d) & j = N. \end{cases} \quad (16)$$

Backward Pass. The backward pass computes $\beta_t(i, d)$

Step 1 (initialization ($t = T$)). The backward variable is shown as follows:

$$\beta_T(i, d) = \begin{cases} 1 & i = N, d = d_N \\ P_i(d) a_{iN}(d) & 1 < i < N, d_1 < d < d_N \\ \sum_{j=2}^{N-1} \alpha_{1j}(d_1) b_j(o_T) \beta_T(j, 1) & i = 1, d = d_1. \end{cases} \quad (17)$$

Step 2 (backward recursion ($t < T$)). For $t = 2, \dots, T$,

$$\beta_t(i, d) = \begin{cases} P_i(d) \sum_{\substack{j=2 \\ j \neq i}}^{N-1} a_{ij}(d) b_j(o_{t+1}) \beta_{t+1}(j, 1) \\ + b_j(o_{t+1}) \beta_{t+1}(i, d+1) & 1 < i < N \\ \sum_{j=2}^{N-1} a_{1j}(d_1) b_j(o_t) \beta_t(j, 1) & i = 1, d = d_1. \end{cases} \quad (18)$$

3.5. Parameter Reestimation for DD-HSMM. The reestimation formula for initial state distribution is the probability that state i was the first state, given O :

$$\bar{\pi}_i = \frac{\pi_i \left[\sum_{d=1}^D \beta(i, d) P(d | i) (c_{o_0o_1}^i c_{o_0o_1}^i \cdots c_{o_{d-1}o_d}^i) \right]}{P(O | \lambda)}. \quad (19)$$

The reestimation formula of state transition probabilities is the ratio of expected number of transition from state i to state j , to the expected number of transitions from state i :

$$\bar{a}_{ij}(d) = \frac{\sum_{t=1}^T \xi_t(i, j, d)}{\sum_{t=1}^T \gamma_t(i, d)}. \quad (20)$$

3.6. Training of State Duration Models Using Parametric Probability Distributions. In this paper, state duration densities are modeled by single Gaussian distribution estimated from training data. The existing state duration estimation method is through the simultaneous training DD-HSMM and their duration densities. However, these techniques are inefficient because the training process requires huge storage and computational load. Therefore, a new approach is adopted for training state duration models. In this approach, state duration probabilities are estimated on the lattice (or trellis) of observations and states, which is obtained in the DD-HSMM training stage.

The mean μ_i and variance $\sigma(i)$ of duration probability of state i are determined by

$$\mu(i) = \frac{\sum_{t=1}^T \sum_{d=1}^{D_i} \chi_t(i, d) d}{\sum_{t=1}^T \sum_{d=1}^{D_i} \chi_t(i, d)}, \quad (21)$$

$$\bar{\sigma}(i) = \frac{\sum_{t=1}^T \sum_{d=1}^{D_i} \chi_t(i, d) d^2}{\sum_{t=1}^T \sum_{d=1}^{D_i} \chi_t(i, d)} - [\bar{\mu}(i)]^2,$$

In these equations, $\chi_t(i, d)$ is the probability of state i at time t with the duration of $d_t(i) = d$ and $\chi_t(i, d)$ can present as

$$\chi_t(i, d) = \frac{1}{P_r} \alpha_t(i, d) p_i(d) \left[\sum_{\substack{j=2 \\ j \neq i}}^{N-1} a_{ij}(d) b_j(o_{t+1}) \beta_{t+1}(j, 1) + a_{iN}(d) \beta_t(N, d_N) \right]. \quad (22)$$

TABLE 1: Prognostics results.

System state	Normal 0	Degradation level 1	Degradation level 2	Failure 3	Recognition accuracy
Normal 0	29	1	0	0	96.7%
Degradation 1	2	27	1	0	90%
Degradation 2	0	1	28	1	93.3%
Failure 3	0	0	1	29	96.7%
Total accuracy					94.2%

TABLE 2: State transition probability ($d_i(1) = 1$).

System state	Normal 0	Degradation level 1	Degradation level 2	Failure 3
Normal 0	0.8813	0.0454	0.0732	0.0001
Degradation 1	0.0000	0.6231	0.3473	0.0296
Degradation 2	0.0000	0.0000	0.9263	0.0737
Failure 3	0.0000	0.0000	0.0000	1.0000

4. DD-HSMM Based Health Prognostic

Many applications in the actuarial, econometric, engineering, and medical literature involve the use of the hazard rate function [33]. The mathematical properties of HR function can reveal a variety of features in the data.

Let T denote the time to failure of an item under consideration, with lifetime distribution function $F(t)$ and reliability function $R(t)$, where $F(t) + R(t) = 1$ and $F(0) = 0$. Assume that $F(0) = 0$ and density function $f(t) = F'(t)$ exist; then the HR function can be defined as:

$$\lambda(t) = \lim_{\substack{N \rightarrow \infty \\ \Delta t \rightarrow 0}} \frac{\Delta m(t)}{[N - m(t)] \Delta t} = \frac{dm(t)}{[N - m(t)] dt} \quad (23)$$

$$= \frac{dm(t)/M}{1 - F(t)} = \frac{f(t)}{R(t)}.$$

In which, M is the total number of sample items, $m(t)$ is the number of items that fail before time t , and $\Delta m(t)$ is the number of items that fail during the time interval $(t, t + \Delta t)$. The ERL function $\mu(t)$ is the expected time remaining to failure, given that the system has survived to time t ; then $\mu(t) = E(T - t \mid T > t) = (1/R(t)) \int_t^\infty R(x) dx$, for t such that $R(t) > 0$. Therefore, $\lambda(t)$ can be approximated as the conditional probability of failure during the time interval $(t, t + \Delta t)$ given survival to time t .

Suppose that a machine will go through health states i , ($i = 1, 2, \dots, N - 1$) before entering failure state N . Let $D(i)$ denote the expected duration of the machine staying at health state i ; based on the parameters estimated above, we can get $D(i)$ as follows:

$$D(i) = \mu(i) + \rho \delta^2(i). \quad (24)$$

And ρ can be denoted by

$$\rho = \frac{(T - \sum_{l=0}^{L-1} \mu(i))}{\sum_{l=0}^{L-1} \delta^2(i)}. \quad (25)$$

Then, once the machine has entered the health state i , its expected residual life equals the summation of the expected

residual duration of the machine staying at health state i and the total remaining staying in the future health states before failure. Denote $\widehat{D}(i^d)$ as the expected residual duration of the machine staying in the health state i for d . When the equipment entered state i at time t_i , the conditional probability of failure during $(t_i + d, t_i + (d + \Delta t))$ can be defined as the probability that the machine will transit to any other state during the coming Δt and the probability that the machine still stay at state i . It can be seen from (9) and (10) that $\widehat{\lambda}(t + d)\Delta t$ can be denoted as follows:

$$\widehat{\lambda}(t + d)\Delta t = \frac{\xi_t(i, j, d)}{\gamma_t(i, d)}. \quad (26)$$

Then

$$\begin{aligned} \widehat{D}(i^d) &= D(i) (1 - \widehat{\lambda}(t + d)\Delta t) \\ &= D(i) \left(1 - \frac{\xi_t(i, j, d)}{\gamma_t(i, d)} \right). \end{aligned} \quad (27)$$

The DD-HSMM equipment health prediction procedure is given as follows.

Step 1. From the DD-HSMM training procedure (i.e., parameter estimation), the state transition probability for the DD-HSMM can be obtained.

Step 2. Through the DD-HSMM parameter estimation, the duration probability density function for each health-state can be obtained. Therefore, the duration mean and variance can be calculated.

Step 3. By classification, identify the current health status of the equipment.

Step 4. The remaining useful life (RUL) of equipment can be predicted by the following formula (suppose that the equipment currently stays at health state i with duration of d):

$$RUL_i^{(d)} = \widehat{D}(i^d) + \sum_{j=i+1}^{N-1} D(j). \quad (28)$$

TABLE 3: Mean and variance of duration in each state ($d_i(1) = 1$).

System state	Normal 0	Degradation level 1	Degradation level 2	Failure 3
Mean	7.9577	7.2105	7.1435	3.3435
Variance	1.3953	0.7429	0.7924	0.5452

TABLE 4: State transition probability ($d_i(1) = 4$).

System state	Normal 0	Degradation level 1	Degradation level 2	Failure 3
Normal 0	0.8813	0.0454	0.0732	0.0001
Degradation 1	0.0000	0.4728	0.4154	0.1118
Degradation 2	0.0000	0.0000	0.9263	0.0737
Failure 3	0.0000	0.0000	0.0000	1.0000

5. Case Study

In this case study, long-term wear experiments on rolling element bearings were conducted [1]. In order to collect adequate amount of data sets for the validation of the proposed scheme, three experiments with normal operating conditions, three experiments with cage defect fault, and three experiments each of inner and outer race defect faults were performed until the bearing reached a complete failure state and stopped operating. Bearing characteristic frequencies in the frequency domain are extracted from the vibration signals corresponding to different degrees of the health states of the bearing acquired during experiments.

During the test running, under each condition, vibration signals were collected. These signals were extracted using a Mahalanobis-Taguchi System (MTS) based model in the original paper [1] and used for the proposed DD-HSMM methodology in this paper. The expert judgment is made of four integer numbers ranging from 0 to 3, representing 4 system states, as follows:

- 0 → the bearing is operating normally;
- 1 → the bearing is operating and shows signs of deterioration; it is advisable to take some preventive action at the next planned maintenance;
- 2 → the bearing is operating but requires immediate attention;
- 3 → the bearing has failed.

5.1. Operation State Identification. In order to identify the accuracy of the operation state identification method proposed in this paper, experimental data with normal operating condition were obtained. The experimental data set included 50 samples for each state (denoted by 0, 1, 2, and 3). Of these data points, 20 of them were used to train the model, and the remaining 30 samples were used to validate the model.

In the DD-HSMM, mixture Gaussian distribution and the single Gaussian distribution were used to model the output probability distribution and the state duration densities separately, in which the number of states is 4. The maximum number of iterations in training process is set to 100 and the convergence error to 0.000001.

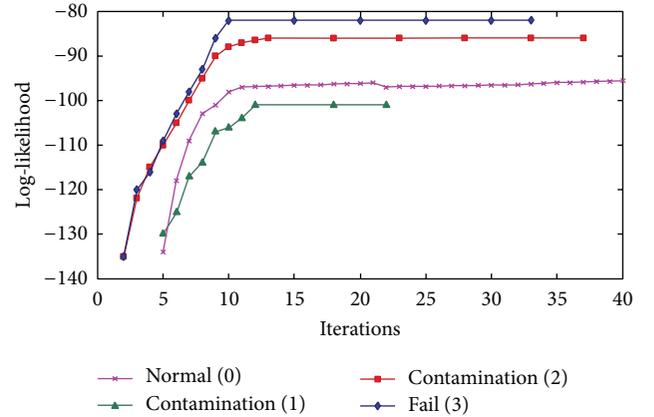


FIGURE 1: Training curve of the DD-HSMM model.

The DD-HSMM-based training model is shown as Figure 1. The x -axis shows the training steps and the y -axis represents the likelihood probability of different states. As can be seen from Figure 1, the progression of the four states reaches the set error in less than 40 steps. This demonstrates the potential of the model to have a strong real-time signal processing capability.

The classification results obtained on the remaining 30 data samples are shown in Table 1. As indicated in the results, the accuracy of the DD-HSMM method is 94.2%.

5.2. Health Prediction for RUL. As described before, a four-state DD-HSMM prediction model is constructed. In the training process, even if the device is in the same running condition, the dwell time is different, transition probabilities between states and the mean or variance of duration in each state are not the same. Tables 2 and 3 show the state transition probability, the mean, and variance of duration in each state when $dt(1) = 1$, representing the bearing in state 1 with duration of 1. Tables 4 and 5 show the state transition probability, the mean, and variance of duration in each state when $dt(1) = 4$, representing the bearing in state 1 with duration of 4.

First, the state i of the current operating state based on the recognition results is determined; then the residence

TABLE 5: Mean and variance of duration in each state ($d_i(1) = 4$).

System state	Normal 0	Degradation level 1	Degradation level 2	Failure 3
Mean	7.9577	5.3105	7.1435	3.3435
Variance	1.3953	1.1328	0.7924	0.5452

TABLE 6: Comparison of DD-HSMM versus HSMM.

Actual RUL	DD-HSMM model		HSMM model	
	Predicted RUL	Error (%)	Predicted RUL	Error (%)
27.0000	26.4027	2.212		0.0548
24.0000	23.6596	1.418	26.9852	12.438
22.0000	22.0856	0.389		22.66
17.0000	17.6142	3.613		0.437
15.0000	15.7246	4.831	17.0734	13.829
12.0000	11.6945	2.546		12.334
11.0000	10.0602	8.544	10.5199	4.365
9.0000	9.1944	2.16		16.888
5.0000	5.1253	2.506		27.792
3.0000	3.0122	0.407	3.6104	20.347

time $\sum_{j=i+1}^{N-1} D(j)$ is calculated according to the duration parameters of the operating state in training process. Then, the remaining effective life in the current operational state is calculated using (25). Finally, the RUL of the bearing can be calculated using (26). Suppose that the bearing is now at state 1 with a duration of 1; then the following can be obtained: $D(2) + D(3) = 10.9426$, $\widehat{D}(1^1) = 6.0875$ by (25), and $RUL_1^{(1)} = 17.0211$ by (26).

5.3. Prediction Comparison. In order to compare the prognostic method based on the DD-HSMM with the prognostic method based on the HSMM, (29) is used to evaluate the life error. In (29), RUL_{actual} represents the actual life of the component, and $RUL_{\text{forecasted}}$ represents the expected life predicted by DD-HSMM or HSMM:

$$\text{Error} = \frac{100\% \times |RUL_{\text{actual}} - RUL_{\text{forecasted}}|}{RUL_{\text{actual}}}. \quad (29)$$

Table 6 shows the prediction comparison of DD-HSMM versus HSMM. Failure prediction of the HSMM method is only state dependent, while the DD-HSMM method uses both state dependency and duration dependency. The DD-HSMM method has a self-updating capability, in which the historical data on states are used in the calculation of state transition probability matrix. As indicated in the results, the DD-HSMM method is more accurate than the HSMM method.

6. Conclusion

This paper presents a Duration-Dependent Hidden Semi-Markov Model (DD-HSMM) for prognostics. As opposed to the Hidden Semi-Markov Model (HSMM), failure prediction capability of the DD-HSMM method uses state dependency

and duration dependency. The two important aspects of equipment health monitoring, which are the stages and the rate of aging, are taken into consideration in an integrated manner in the proposed DD-HSMM model. The duration-dependent state transition probability in the Hidden Semi-Markov model makes the decision-making more relevant to real world applications.

In order to facilitate the computational procedure, a new forward-backward algorithm and reestimation approach are developed. By using autoregression, the interdependency between observations is established in the model. By incorporating an explicitly defined temporal structure into the model, the DD-HSMM is capable of predicting the remaining useful life of equipment more accurately.

The demonstration of the proposed model is carried out using experimental data on rolling element bearings. The proposed model provides a powerful state recognition capability and very accurate results in terms of remaining useful life prediction. In order to draw general conclusion on the capabilities of the proposed DD-HSMM, more experimental data in various prognostics areas are needed.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors gratefully acknowledge the financial support for this research from the National High Technology Research and Development Program of China (no. 2012AA040914), the National Natural Science Foundation of China (Grant no. 71101116), and the Basic Research Foundation of NPU (Grant no. JC20120228).

References

- [1] A. Soylemezoglu, S. Jagannathan, and C. Saygin, "Mahalanobis taguchi system (MTS) as a prognostics tool for rolling element bearing failures," *Journal of Manufacturing Science and Engineering*, vol. 132, no. 5, Article ID 051014, 2010.
- [2] A. Soylemezoglu, S. Jagannathan, and C. Saygin, "Mahalanobis-Taguchi system as a multi-sensor based decision making prognostics tool for centrifugal pump failures," *IEEE Transactions on Reliability*, vol. 60, no. 4, pp. 864–878, 2011.
- [3] C. Bunks, D. McCarthy, and T. Al-Ani, "Condition-based maintenance of machines using hidden Markov models," *Mechanical Systems and Signal Processing*, vol. 14, no. 4, pp. 597–612, 2000.
- [4] P. Orth, S. Yacout, and L. Adjengue, "Accuracy and robustness of decision making techniques in condition based maintenance," *Journal of Intelligent Manufacturing*, vol. 23, no. 2, pp. 255–264, 2012.
- [5] S. Ambani, L. Li, and J. Ni, "Condition-based maintenance decision-making for multiple machine systems," *Journal of Manufacturing Science and Engineering*, vol. 131, no. 3, pp. 0310091–0310099, 2009.
- [6] S. Si, H. Dui, Z. Cai, S. Sun, and Y. Zhang, "Joint integrated importance measure for multi-state transition systems," *Communications in Statistics. Theory and Methods*, vol. 41, no. 21, pp. 3846–3862, 2012.
- [7] S. Shubin, G. Levitin, D. Hongyan, and S. Shudong, "Component state-based integrated importance measure for multi-state systems," *Reliability Engineering and System Safety*, vol. 116, pp. 75–83, 2013.
- [8] A. K. S. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mechanical Systems and Signal Processing*, vol. 20, no. 7, pp. 1483–1510, 2006.
- [9] A. Heng, S. Zhang, A. C. C. Tan, and J. Mathew, "Rotating machinery prognostics: State of the art, challenges and opportunities," *Mechanical Systems and Signal Processing*, vol. 23, no. 3, pp. 724–739, 2009.
- [10] L. R. Rabiner, "Tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 257–286, 1989.
- [11] R. J. Elliott, L. Aggoun, and J. B. Moore, *Hidden Markov Models: Estimation and Control*, vol. 29, Springer, New York, NY, USA, 1995.
- [12] M. D. Le and C. M. Tan, "Optimal maintenance strategy of deteriorating system under imperfect maintenance and inspection using mixed inspection scheduling," *Reliability Engineering & System Safety*, vol. 113, pp. 21–29, 2013.
- [13] Y. Xu and M. Ge, "Hidden Markov model-based process monitoring system," *Journal of Intelligent Manufacturing*, vol. 15, no. 3, pp. 337–350, 2004.
- [14] M. Ostendorf and S. Roukos, "Stochastic segment model for phoneme-based continuous speech recognition," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 12, pp. 1857–1869, 1989.
- [15] A. Ljolje and S. E. Levinson, "Development of an acoustic-phonetic hidden Markov model for continuous speech recognition," *IEEE Transactions on Signal Processing*, vol. 39, no. 1, pp. 29–39, 1991.
- [16] A. Kannan and M. Ostendorf, "Comparison of trajectory and mixture modeling in segment-based word recognition," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 327–330, April 1993.
- [17] M. Y. Chen, A. Kundu, and J. Zhou, "Off-line handwritten work recognition using a hidden Markov model type stochastic network," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 5, pp. 481–496, 1994.
- [18] M. Y. Chen, A. Kundu, and S. N. Srihari, "Variable duration hidden Markov model and morphological segmentation for handwritten word recognition," *IEEE Transactions on Image Processing*, vol. 4, no. 12, pp. 1675–1688, 1995.
- [19] L. Atlas, M. Ostendorf, and G. D. Bernard, "Hidden Markov models for monitoring machining tool-wear," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 3887–3890, June 2000.
- [20] L. Wang, M. G. Mehrabi, and E. Kannatey-Asibu Jr., "Hidden Markov model-based tool wear monitoring in turning," *Journal of Manufacturing Science and Engineering*, vol. 124, no. 3, pp. 651–658, 2002.
- [21] S. Lee, L. Li, and J. Ni, "Online degradation assessment and adaptive fault detection using modified hidden markov model," *Journal of Manufacturing Science and Engineering*, vol. 132, no. 2, pp. 0210101–0210101, 2010.
- [22] S. Si, H. Dui, Z. Cai, and S. Sun, "The Integrated Importance Measure of Multi-State Coherent Systems for Maintenance Processes," *IEEE Transactions on Reliability*, vol. 61, no. 2, pp. 266–273, 2012.
- [23] E. Zio and M. Compare, "Evaluating maintenance policies by quantitative modeling and analysis," *Reliability Engineering & System Safety*, vol. 109, no. 203, pp. 53–65, 2013.
- [24] Z. Cai, S. Sun, S. Si, and B. Yannou, "Identifying product failure rate based on a conditional Bayesian network classifier," *Expert Systems with Applications*, vol. 38, no. 5, pp. 5036–5043, 2011.
- [25] K. Tokuda, H. Zen, and A. W. Black, "An HMM-based speech synthesis system applied to English," in *Proceedings of the IEEE Workshop on Speech Synthesis*, pp. 227–230, 2002.
- [26] H. Zen, K. Tokuda, T. Masuko, T. Kobayasih, and T. Kitamura, "A hidden semi-Markov model-based speech synthesis system," *IEICE Transactions on Information and Systems*, vol. 90, no. 5, pp. 825–834, 2007.
- [27] K. Hashimoto, Y. Nankaku, and K. Tokuda, "A Bayesian approach to hidden semi-Markov model based speech synthesis," in *Proceedings of the 10th Annual Conference of the International Speech Communication Association*, pp. 1751–1754, September 2009.
- [28] P. Baruah and R. B. Chinnam, "HMMs for diagnostics and prognostics in machining processes," *International Journal of Production Research*, vol. 43, no. 6, pp. 1275–1293, 2005.
- [29] T. Boutros and M. Liang, "Detection and diagnosis of bearing and cutting tool faults using hidden Markov models," *Mechanical Systems and Signal Processing*, vol. 25, no. 6, pp. 2102–2124, 2011.
- [30] M. Dong and D. He, "A segmental hidden semi-Markov model (HSMM)-based diagnostics and prognostics framework and methodology," *Mechanical Systems and Signal Processing*, vol. 21, no. 5, pp. 2248–2266, 2007.
- [31] M. Dong and D. He, "Hidden semi-Markov model-based methodology for multi-sensor equipment health diagnosis and prognosis," *European Journal of Operational Research*, vol. 178, no. 3, pp. 858–878, 2007.
- [32] S. Yu, "Hidden semi-Markov models," *Artificial Intelligence*, vol. 174, no. 2, pp. 215–243, 2010.

- [33] M. Dong, "A tutorial on nonlinear time-series data mining in engineering asset health and reliability prediction: concepts, models, and algorithms," *Mathematical Problems in Engineering*, vol. 2010, Article ID 175936, 22 pages, 2010.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

