Robotics and Computer-Integrated Manufacturing 27 (2011) 581–590

Contents lists available at ScienceDirect

Robotics and Computer-Integrated Manufacturing

journal homepage: www.elsevier.com/locate/rcim

# Equipment PHM using non-stationary segmental hidden semi-Markov model

Ming Dong [a,*], Ying Peng [b]

[a] College of Economics and Management, Shanghai Jiao Tong University, 535 Fahua Zhen Road, Shanghai 200052, PR China
[b] School of Mechanical Engineering, Shanghai Jiao Tong University, 800 Dongchuan Road, Shanghai 200240, PR China

ARTICLE INFO

*Article history:*
Received 31 March 2009
Received in revised form
1 October 2010
Accepted 4 October 2010
Available online 18 October 2010

*Keywords:*
PHM
Non-stationary segmental hidden semi-Markov model
Hazard rate
Aging factor
Remaining useful life

ABSTRACT

Health monitoring and prognostics of equipment is a basic requirement for condition-based maintenance (CBM) in many application domains where safety, reliability, and availability of the systems are considered mission critical. As a key complement to CBM, prognostics and health management (PHM) is an approach to system life-cycle support that seeks to reduce/eliminate inspections and time-based maintenance through accurate monitoring, incipient faults. Conducting successful prognosis, however, is more difficult than conducting fault diagnosis. A much broader range of asset health related data, especially those related to the failures, shall be collected. The asset health progression can then be possibly extracted from the congregated data, which has proved to be very challenging. This paper presents a non-stationary segmental hidden semi-Markov model (NSHSMM) based prognosis method to predict equipment health. Unlike previous HSMMs, the proposed NSHSMM no longer assumes that the state-dependent transition probabilities keep the same value all the time. That is, the probability of transiting to a less healthy state does not increase with the age. "Non-stationary" means the transition probabilities will change with time. In the proposed method, in order to characterize a deteriorating equipment, three kinds of aging factor that discount the probabilities of staying at current state while increasing the probabilities of transitions to less healthy states are introduced. The performances of these aging factors are compared by using historical data colleted from three hydraulic pumps. The hazard function (h.f.) has been introduced to analyze the distribution of lifetime with a combination of historical failure data and on-line condition monitoring data. Using h.f., PHM is based on a failure rate that is a function of both the equipment age and the equipment conditions. The state values of the equipment condition considered in PHM, however, are limited to those stochastically increasing over time and those having non-decreasing effect on the hazard rate. The estimated state duration probability distributions can be used to predict the remaining useful life of the systems. With the equipment PHM, the behavior of the equipment condition can be predicted.



## 1. Introduction and literture review

A prerequisite to the deployment of condition-based maintenance (CBM) is effective diagnosis and prognosis. Diagnosis is an assessment about the current and past health, which deals with fault detection, isolation and identification based on observed symptoms of a system when abnormity occurs. Prognosis is an assessment of the future health, which deal with fault and degradation prediction before their occurring. CBM is a decision making strategy to enable real-time diagnosis of impending failures and prognosis of future equipment health, where the decision to perform maintenance is reached by observing the "condition" of the system and its components. The condition of a system can be quantified by getting data from various sensors in the system periodically or even continuously. CBM increases system efficiency and availability through elimination of unnecessary maintenance. Equipment health diagnosis and prognosis for implementing condition-based maintenance becomes a basic and desirable requirement in many application domains where safety, reliability, and availability of the systems are considered mission critical. The economic ramifications of CBM are many folds since it affects labor requirements, replacement part costs, and the logistics of scheduling routine maintenance. Prognostics and health management (PHM) is an approach to system life-cycle support that seeks to reduce/eliminate inspections and time-based maintenance through accurate monitoring, incipient faults [1]. PHM is a key complement to CBM that also mitigates the variability in the maintenance process inherent in CBM driven by automated fault detection or periodic inspection. CBM and PHM have evident synergies. Both require in-depth knowledge of failure modes and effects, with detail understanding of failure probability as a function of usage and state—at the individual component level. PHM adds a valuable option to the menu of available CBM tasks

* Corresponding author.
*E-mail address:* mdong@sjtu.edu.cn (M. Dong).

0736-5845/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.rcim.2010.10.005

by mitigating the impact of failures on "safety, environment, operations and economics" [2].

Conducting successful prognosis, however, is more difficult than conducting fault diagnosis. Being able to perform precise and reliable prognostics is the key of CBM for engineered systems, and it is also critical for improving safety, planning missions, scheduling maintenance, reducing maintenance costs and down time [3]. The objective of prognosis is predicting the progression of a fault condition to component failure and estimating remaining useful life (RUL) of the component. The literature on prognostic methods is limited but the concept has been gaining importance in recent years. Unlike numerous methods available for diagnostics, prognosis is still in its infancy, and literature is yet to present a working model for effective prognosis [4,5]. Many traditional methodologies used successfully in other areas or new methodologies have been introduced into the prognostics field. Many researchers have combined two or more techniques and methods together to improve the performance of prognostic models. Samanta and Nataraj presented a system for monitoring and prognostics of machine conditions using soft computing (SC) techniques, namely adaptive neuro-fuzzy inference system (ANFIS). Comparison with a machine learning method, namely support vector regression (SVR), is also presented [20]. Usually there are three categories of prognostic model: physical model, knowledge based model, and data-driven model. Physical model based approaches usually employ mathematical models that are directly tied to the physical processes that have direct or indirect effects on the health of the components. It is usually a tough task to accurately build a mathematical model for a physical system with prior principles in real world applications. So the uses of physical model based methodology are limited. For such a reason, knowledge-based methodology such as expert systems and fuzzy logic requiring no physical model is proposed by some researchers. The data-driven prognostic model is based upon statistical and learning techniques, most of which originated from the theory of pattern recognition. Data-driven models are usually developed from collected input/output data. These models can process a wide variety of data types and exploit the nuances in the data that cannot be discovered by physical systems. HMM and HSMM belong to the data-driven model. Baruah and Chinnam first pointed out that standard hidden Markov model could be applied in the area of prognostics in machining processes [6]. An integrated fault diagnostic and prognostic approach for bearing health monitoring and CBM was introduced [7]. The proposed scheme consists of principle component analysis (PCA), HMM, and an adaptive stochastic fault prediction model. Camci proposed an integrated diagnostic and prognostic architecture that employed support vector machine (SVM) and HMM [8]. But HMMs have some inherent limitations. One is the assumption that successive system behavior observations are independent. The other is the Markov assumption itself that the probability in a given state at time $t$ only depends on the state at time $t-1$ is sometimes untenable in practical applications. In order to cope with the inaccurate durational modeling of HMM, some authors have proposed Hidden semi-Markov model (HSMM) to model explicitly the state duration [9,10]. A common idea is to replace the duration probability density function with some well-chosen probability functions close to the durational distribution of real-life applications. Dong and He presented an integrated framework based on HSMM for multi-sensor equipment diagnosis and prognosis [11,12]. In this framework, they used states of HSMM to represent the health status of a component. The trained HSMM can be used to diagnose the health status of a component. Through parameter estimation of the health state duration probability distribution and the proposed backward recursive equations, the RUL of the component can be predicted [13].

The objective of this work is to develop a new prognostic methodology for an NSHSMM model. Three kinds of aging factors that discounts the probabilities of staying at current state while increasing the probabilities of transitions to less healthy states will be introduced into the model. A statistical algorithm with aging factor will be designed to compute the transition probability between two health states when the equipment deterioration happened. With the historical data and the real-time collected data, the remaining useful life of the equipment at time $t$ (RUL($t$)) can be calculated.

## 2. NSHSMM based modeling framework for prognostic

### 2.1. Transition matrix considering equipment age

HMMs characterize doubly embedded stochastic process with an underlying hidden stochastic process that can be observed through some probabilistic behavior, this is where its name "hidden" comes from. HMM is a parametric model, its parameters can be estimated by the vast experimental data using statistical techniques. HMMs have some distinct characteristics that are not possessed by some traditional methods. They could not only reflect the randomicity of machine behaviors but also reveal their hidden states, changing processes. Furthermore, HMMs have a well constructed theoretical basis and easy to realize in software. The principle of HMM based prognostics is as follows: first, build and train $N$ HMMs for all component health states. Between $N$ trained HMMs, it is usually assumed that the estimated vectors of state transition times follow some multivariate distribution. Once the distribution is assessed, the conditional probability distribution of a distinct state transition given the previous state transition points can be estimated. The coordinates of the points of intersection of the log-likelihood trajectories for different HMMs along the life/usage axis represent the estimated 'state transition time instants'. It is these state transition points that would allow one to extend the using of HMMs for prognostics. However, only standard HMM based approaches are proposed in such model.

HSMM is constructed by adding a temporal component into the well-defined HMM structures. HMMs with such an explicitly added state durational probability functions are called HSMM, because the transition properties are no longer governed by a Markov process. It is like a HMM except each state can emit a sequence of observations. HSMM models the observations during the stay in state $H_i$ as a whole. Hidden semi-Markov chains possess both the flexibility of hidden Markov chains for approximating complex probability distributions and the flexibility of semi-Markov chains for representing temporal structures.

For a machine, it usually evolves through several distinct health states prior to reaching a failure. Here failure means the machine breakdown. Suppose now the health state has been classified into $n$ discrete states 1, 2, 3, …, $n-1$, and $F$. The classification can be done by experience. Since the health state can be tested at each sampling time point by a trained HSMM diagnostic model [12], it can be viewed as a stochastic process $H=\{H_t\colon t\geq 0\}$. If $H_t=i$, the equipment is said to be in state $H_i$ at time $t$. We assume here that when the process is in state $H_i$, there is a fixed probability $P_{i,j}$ that the health state will be in state $j$ at the next time point. For a hidden Markov chain, the conditional distribution of any future state $H_t+1$ given the earlier states $H_1$, …, $H_t$ becomes [14]

$$P(H_{t+1}=j|H_t=i,H_{t-1}=i_{t-1},\ldots,H_2=i_2,H_1=i_1)$$
$$=P(H_{t+1}=j|H_t=i)=P_{ij}.$$