Mechanical Systems and Signal Processing ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

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

Ascertainment-adjusted parameter estimation approach to improve robustness against misspecification of health monitoring methods

P. Juesas [a], E. Ramasso [a,b,*]

[a] Department of Applied Mechanics, France
[b] Department of Automatic Control and Micro-Mechatronic Systems, FEMTO-ST institute, UMR CNRS 6174-UFC/ENSMM/UTBM, 26 chemin de l'Epitaphe, 25000 Besançon, France

A R T I C L E   I N F O

Article history:
Received 15 October 2015
Received in revised form
23 March 2016
Accepted 24 March 2016

Keywords:
Time-series
Noisy and uncertain labels
Weighted distribution theory
Health monitoring
CMAPSS datasets

A B S T R A C T

Condition monitoring aims at ensuring system safety which is a fundamental requirement for industrial applications and that has become an inescapable social demand. This objective is attained by instrumenting the system and developing data analytics methods such as statistical models able to turn data into relevant knowledge. One difficulty is to be able to correctly estimate the parameters of those methods based on time-series data. This paper suggests the use of the Weighted Distribution Theory together with the Expectation–Maximization algorithm to improve parameter estimation in statistical models with latent variables with an application to health monotonic under uncertainty. The improvement of estimates is made possible by incorporating uncertain and possibly noisy prior knowledge on latent variables in a sound manner. The latent variables are exploited to build a degradation model of dynamical system represented as a sequence of discrete states. Examples on Gaussian Mixture Models, Hidden Markov Models (HMM) with discrete and continuous outputs are presented on both simulated data and benchmarks using the turbofan engine datasets. A focus on the application of a discrete HMM to health monitoring under uncertainty allows to emphasize the interest of the proposed approach in presence of different operating conditions and fault modes. It is shown that the proposed model depicts high robustness in presence of noisy and uncertain prior.

1. Introduction

The statistical representation of multi-dimensional time-series originating from a dynamical system consists in finding a concise and meaningful mathematical model that can be easily interpreted and used to understand the behavior of the system. It is an important problem encountered in a wide range of real applications such as localization and mapping for mobile robot exploring an unknown environment [44,43], structural health monitoring under different loading conditions [1,49,40], forecasting and prognostics of various systems [53,47,19,24] or human motion analysis [39,20]. The context of real-world applications generally involves temporal processes subject to uncertainty which is generally managed either by

* Corresponding author at: Department of Automatic Control and Micro-Mechatronic Systems, FEMTO-ST institute, UMR CNRS 6174-UFC/ENSMM/UTBM, 26 chemin de l'Epitaphe, 25000 Besançon, France.
E-mail address: emmanuel.ramasso@femto-st.fr (E. Ramasso).

http://dx.doi.org/10.1016/j.ymssp.2016.03.022
0888-3270/© 2016 Published by Elsevier Ltd.

the theory of belief functions, set-membership approaches or probability theory [23,48,4,5].

This paper is focused on the management of uncertainty in statistical models of time-series originating from in-situ monitoring of the health state of industrial equipments. The monitoring is ensured by sensors that continuously record data (observations) based on which the current and future degradation levels of the system has to be inferred for detection (or diagnostics) and prognostics purposes [52].

The literature on data-driven Prognostics and Health Monitoring (PHM) is mainly focused on supervised or unsupervised models [51,28,24,41]: For training, the degradation level (or state) is either known precisely or hidden and, during testing, the current and future levels are inferred from sensor data. In situations where the degradation level can be represented as a sequence of discrete states, statistical models with discrete latent variables have been widely used. Hidden Markov Model (HMM) [34,45] represents one of those models. HMMs have indeed been widely exploited for PHM [26,9,15,30,24,35] where it is assumed that the degradation of the equipment follows a doubly stochastic process: one for the dynamics of the hidden states and one to account for the distribution of the observations.

Generally, the *generative* form of HMM is used: one model is built for each possible degradation level, and, during monitoring, a similarity-based approach is applied to find the likeliest model, that is then used to infer the current state (detection and diagnostics) and the future trends (forecasting and prognostics). We propose an alternative in using HMM for PHM by considering *discriminative* learning where the parameters are estimated with the aim to improve the classification into degradation levels.

Compared to classical generative approaches for HMM-based PHM, and more generally to latent model-based PHM, the proposed approach is based on the idea that sensor data related to different degradation levels have to be put together in a training dataset. Each level is then assigned a *soft label*: it allows one to use uncertain or noisy labels [10] according to the quantity and the quality of the prior knowledge about the degradation level. By this way, some of the latent variables may now represent one degradation level and multiple levels can share common feature subspaces through the use of uncertain labels.

The solution for incorporating soft labels in models with latent variables is based on a modification of the conditional expectation of the log-likelihood in the Expectation–Maximization algorithm by using the Weighted Distribution Theory [33]. An application of this criterion is shown for Gaussian Mixture Models, continuous and discrete HMM. A particular attention is paid to Maximum Likelihood Estimation (MLE) of parameters in HMM with discrete-valued observations (DHMM) in presence of soft labels.

Standard DHMMs have been used in the past in many applications [31,57,42,8,3,26] and in particular for noisy speech and character recognition [7,50,17,12]. In the context of PHM, DHMM has also been widely used, for instance in [25] for predictive modeling dedicated to intelligent maintenance in complex semiconductor manufacturing processes, in [29] for incipient fault detection and diagnosis in turbine engines, in [2] for failure isolation for cognitive robots and in [16] for anomaly detection in electronic systems.

While the inference phase is very similar between HMM with continuous (CHMM) and discrete observations (DHMM), the learning phase presents a fundamental difference since the observations in the latter are discrete. The learning phase for huge datasets is thus generally faster for the DHMM by using matrix encoding. Moreover, it has been shown that the DHMM may be better in presence of noise with unknown (nonGaussian) characteristics [50].

The contribution of this paper is two-fold:

- We propose a framework for PHM based on latent variables. It is based on the Expectation–Maximization algorithm and the Weighted Distribution Theory. This work is inspired from a previous work proposed by Denoeux [14] based on Dempster–Shafer's theory of belief functions and plausibility weights. The main difference with the present work is the consideration of almost unrestricted weights (satisfying only positiveness) based on the work of Patil [33].
- This framework is used to improve the performance of DHMM for PHM. An application to turbofan health monitoring is presented where we evaluate the sensitivity of DHMM to vector quantization with respect to the quality and quantity of prior.

The problem of incorporating partial knowledge about latent variables is discussed in Section 2. Section 3 presents the application of this method to DHMM. Section 4 is dedicated to the analysis of the proposed model on several datasets.

2. Incorporating prior knowledge on latent variables in EM

The Expectation–Maximization algorithm (EM) [13] allows to estimate the parameters of a statistical model with latent variables. It has been adapted in [14] in order to take *uncertain* prior information into account. The prior is supposed to be represented by a collection of T Basic Belief Assignments (BBA) denoted as $\mathbf{m} = \{m_1^{\Omega_\mathbf{y}}, \ldots m_T^{\Omega_\mathbf{y}}\}$ defined on the set of discrete states $\Omega_\mathbf{y}$ with $\sum_{A \subseteq \Omega_\mathbf{y}} m_t^{\Omega_\mathbf{y}}(A) = 1$, and such that $m_t^{\Omega_\mathbf{y}}(\emptyset) = 0$ (normalized BBA). The adaptation of EM concerns the auxiliary function (conditional expectation of the log-likelihood) to be maximized at each iteration q:

Please cite this article as: P. Juesas, E. Ramasso, Ascertainment-adjusted parameter estimation approach to improve robustness against misspecification of health monitoring methods, Mech. Syst. Signal Process. (2016), http://dx.doi.org/10.1016/j.ymssp.2016.03.022