Control Engineering Practice 74 (2018) 84–94

Contents lists available at ScienceDirect

Control Engineering Practice

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

A hidden-Gamma model-based filtering and prediction approach for monotonic health factors in manufacturing

Gian Antonio Susto [a,*], Andrea Schirru [b], Simone Pampuri [b], Alessandro Beghi [a], Giuseppe De Nicolao [c]

[a] Department of Information Engineering, University of Padova, via G. Gradenigo 6/B, 35131 Padova, Italy
[b] Statwolf LTD, 38-39 Baggot Street Lower, Dublin, Ireland
[c] Department of Computer Engineering and Systems Science, University of Pavia, via Ferrata 1, 27100 Pavia, Italy

ARTICLE INFO

Keywords:
Decision support system
Gamma distribution
Industry 4.0
Monte Carlo methods
Particle filters
Predictive maintenance
Prognostic and health management
Semiconductor manufacturing

ABSTRACT

In the context of Smart Monitoring and Fault Detection and Isolation in industrial systems, the aim of Predictive Maintenance technologies is to predict the happening of process or equipment faults. In order for a Predictive Maintenance technology to be effective, its predictions have to be both accurate and timely for taking strategic decisions on maintenance scheduling, in a cost-minimization perspective. A number of Predictive Maintenance technologies are based on the use of "health factors", quantitative indicators associated with the equipment wear that exhibit a monotone evolution. In real industrial environment, such indicators are usually affected by measurement noise and non-uniform sampling time. In this work we present a methodology, formulated as a stochastic filtering problem, to optimally predict the evolution of the aforementioned health factors based on noisy and irregularly sampled observations. In particular, a hidden Gamma process model is proposed to capture the nonnegativity and nonnegativity of the derivative of the health factor. As such filtering problem is not amenable to a closed form solution, a numerical Monte Carlo approach based on particle filtering is here employed. An adaptive parameter identification procedure is proposed to achieve the best trade-off between promptness and low noise sensitivity. Furthermore, a methodology to identify the risk function associated to the observed equipment based on previous maintenance data is proposed. The present study is motivated and tested on a real industrial Predictive Maintenance problem in semiconductor manufacturing, with reference to a dry etching equipment.

1. Introduction

Advanced monitoring is a fundamental activity in the Industry 4.0 scenario to implement control, maintenance, quality, reliability, and safety policies (Arinton, Caraman, & Korbicz, 2012; Chioua, Bauer, Chen, Schlake, Sand, Schmidt, et al., 2016; Ma, Dong, Peng, & Zhang, 2017). In particular, Fault Detection and Isolation (FDI) (Ma et al., 2017) and Predictive Maintenance (PdM) (Nguyen, Do, & Grall, 2015) technologies have proliferated in the past recent years for diagnosis and prognosis of process/tool failures (Sikorska, Hodkiewicz, & Ma, 2011). While the aim of such technologies is similar and partly overlapped, PdM technologies are more focused on prognosis. Prognosis can be defined as the capability to provide early detection of the precursor and/or incipient fault condition of a component, and to design tools for managing and predicting the progression of such fault condition to component failure (Engel, Gilmartin, Bongort, & Hess, 2000). Given their goal, PdM technologies are typically applied to failures that are associated with wear and usage of the system/process (Susto, Schirru, Pampuri, McLoone, & Beghi, 2015), or, more generally, to failures that can be predicted in advance (Lewin, 1995; Susto, McLoone, Pagano, Schirru, Pampuri, & Beghi, 2013). Examples of such type of faults are the breaking of the source in ion-implantation processes in semiconductor manufacturing (Susto et al., 2015), the flute wear in cutting tool equipment (Benkedjouh, Medjaher, Zerhouni, & Rechak, 2015), and the lifespan of lithium-ion batteries (Liao & Köttig, 2016).

In this work we focus on the so-called 'Health Factors' (HFs), an important concept in prognostic.[1] HFs are quantitative indexes used to define the current status of a tool/process and to assess the future

[1] Health Factors are also indicated as 'Component Health' (Sikorska et al., 2011), 'State of Health'/'Health State' (Si, Wang, Hu, & Zhou, 2011; Zhou, Stein, & Ersal, 2017) or as 'Health Indicators' (Benkedjouh et al., 2015; Wang, Yu, Siegel, & Lee, 2008) by different authors and they are closely in relation with the concept of 'degradation data' (Chen, Lio, Ng, & Tsai, 2017).

* Corresponding author.
E-mail addresses: gianantonio.susto@dei.unipd.it (G.A. Susto), andrea.schirru@statwolf.com (A. Schirru), simone.pampuri@statwolf.com (S. Pampuri), beghi@dei.unipd.it (A. Beghi), giuseppe.denicolao@unipv.it (G. De Nicolao).

https://doi.org/10.1016/j.conengprac.2018.02.011
Received 9 July 2017; Received in revised form 3 January 2018; Accepted 23 February 2018

statuses of the system under exam (or of one of its components/subsystems), and its Remaining Useful Life (RUL) (Bressel, Hilairet, Hissel, & Bouamama, 2016; Butler & Ringwood, 2010; Wang et al., 2008), so that strategic decisions regarding maintenance scheduling and dynamic sampling plans can be taken (Nguyen et al., 2015). Being in direct relationship with wear, usage or stress of an equipment/component or system, HFs generally have a monotone evolution. A HF can be of very different nature: in its simplest form, HFs can be observable parameters that, thanks to specific domain expertise, can be associated with equipment/process health status. Example of health factors as quantities that are directly related to system health, such as the thermal index of a polymeric material (Xie, Jin, Hong, & Van Mullekom, 2017), the scar width in sliding metal wear (Hu, Li, & Hu, 2017), and the temperature difference in semiconductor manufacturing epitaxy processes (Susto, Beghi, & Luca, 2012). HFs can also be the output of Soft Sensor modules (Souza & Araujo, 2014; Wang, Liu, & Srinivasan, 2010), where the status health is impossible/costly to be monitored. Moreover, HFs can be the residual of first principle FDI models (Zhang & Canova, 2015). In fact, in many practical examples (Arinton et al., 2012; Butler & Ringwood, 2010; Hast, Findeisen, & Streif, 2015; Zhang & Canova, 2015), residuals have a monotonic behavior and threshold-based policies to maintenance management are implemented on such quantities. HFs are therefore relevant quantities in both model-based (Dey, Biron, Tatipamula, Das, Mohon, Ayalew, et al., 2016; Hast et al., 2015; Xu, Lee, Zhou, & Yang, 2015) and model-free (Arinton et al., 2012; Bakdi, Kouadri, & Bensmail, 2017; Ge, Song, & Gao, 2013; Ma et al., 2017) prognostic approaches.

In the present paper, the problem of designing a HF for Predictive Maintenance (PdM) purposes is considered (Ding, Yin, Peng, Hao, & Shen, 2013; Filev, Chinnam, Tseng, & Baruah, 2010; Susto et al., 2015). In particular, the issue of assessing the probability distribution of the HF future values given its past measurements is addressed, under the following assumptions: (i) the HF is monotonically increasing; (ii) its measurements are subject to random noise that may conceal its monotonic nature; (iii) measurements are non-uniformly sampled over time. The aforementioned features are typical traits of HF signals (Butler & Ringwood, 2010; Gorinevsky, 2004; Saha, Goebel, & Christophersen, 2009; Susto et al., 2012; You, Li, Meng, & Ni, 2010), but they are generally not simultaneously accounted for in the related literature. Non-stochastic models (see Si et al., 2011 for a broad review on RUL estimation) for HFs have been presented in literature, as well as inspection and intervention approaches for increasing maintenance actions effectiveness and decreasing the associated costs. However, such methodologies are well suited for noise-free scenarios and, given the aforementioned assumptions on the HF signals, it is here proposed to adopt a stochastic filtering paradigm (Wang, Hussin, & Jefferis, 2012). With the proposed approach, the HF is treated as a stochastic process, with the possibility to combine prior knowledge on the HF with statistical information regarding the observed noisy data. A simple approach to deal with the problem at hand is provided by the Wiener and Kalman predictors (Abdennadher, Venet, Rojat, Rétif, & Rosset, 2010; Lu, Tu, & Lu, 2007; Susto et al., 2012; Yang & Liu, 1999), which are statistically optimal for linear Gaussian models. However, such classical approaches may be considered suboptimal for signals with the characteristics given in assumptions (i)–(iii). As a matter of fact, far from being Gaussian, the HF derivative is in this work considered to be a nonnegative random variable.

Given such premises, a framework for HF filtering and prediction based on the Gamma distribution is here proposed. PdM applications employing Gamma distributions has been developed since the 1970s (Abdel-Hameed, 1975), especially in mechanical and civil engineering applications (Cinlar, Osman, & Bazoant, 1977; Lawless & Crowder, 2004; Lu, Pandey, & Xie, 2013) and, recently, in industrial environments (LeSon, Fouladirad, & Barros, 2016). Indeed, if the HF is modeled as the sum of Gamma distributed random variables, such sum is still Gamma distributed, with the advantage that convenient estimation and prediction algorithms can be derived. Given that in real-world industrial applications HFs are usually observed with noise, the approach proposed in this work considers the HF as a monotonic Gamma process (with time-varying shape parameter) corrupted by Gaussian noise (***hidden-Gamma*** model). Such assumptions lead to the lack of closed-form solutions for the estimation of model parameters in the proposed approach. However, it will be shown that the prediction problem can be efficiently solved by resorting to particle filtering methods (Alrowaie, Gopaluni, & Kwok, 2012; Doucet, 1998; Doucet, DeFreitas, & Gordon, 2001), employing Monte Carlo (MC) simulations to derive the target posterior distributions. Finally, a recursive procedure to estimate the time-varying shape parameter is proposed. Such procedure allows to optimize a trade-off between the need for promptness and noise insensitivity/outlier rejection.

The paper is organized as follows. In Section 2 the hidden-Gamma model is presented. In Section 3.1 the principles of Particle Filtering (PF) are briefly summarized and adapted to Gamma processes. In Section 4 an adaptive recursive scheme for estimation of monotone HFs is presented. Section 5 is dedicated to the definition and estimation of an appropriate Risk Function for the proposed model. In Section 6 some experimental results on synthetic datasets are reported, whereas in Section 7 a real PdM semiconductor manufacturing problem related to dry etching is tackled.[2]

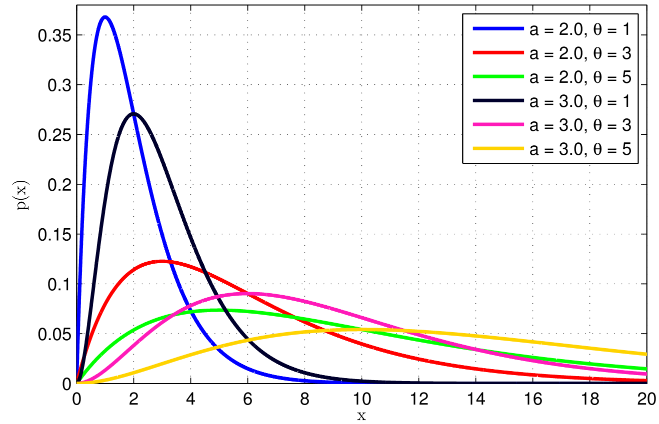

Fig. 1. Gamma probability distributions for different values of a and θ.

2. The hidden Gamma process

2.1. Gamma probability distribution

The most notable property of Gamma distributions is their nonnegative support. We consider a random variable x with Gamma distribution $\Gamma(a, \theta)$, where a is the shape parameter and θ is the scale factor. The first two statistical moments of x are $E[x] = a\theta$ and $Var[x] = a\theta^2$ and the probability density function (PDF) is $p(x) = \frac{x^{a-1}e^{-\frac{x}{\theta}}}{\Gamma(a)\theta^a}$. Gamma distributed random variables enjoy the following property:

Property 1 (*Infinite Divisibility*). *If $x_1 \sim \Gamma(a_1, \theta)$ and $x_2 \sim \Gamma(a_2, \theta)$, then the sum $x = x_1 + x_2$ has a Gamma distribution with shape $a_1 + a_2$ and scale factor θ.*

The shape of the Gamma probability distribution for different values of a and θ is shown in Fig. 1.

[2] The present work is an extended version of Schirru, Pampuri, and DeNicolao (2010). Additional material concerns implementation details, the derivation of a risk function associated with the maintenance operation, and the use of synthetic data to better assess performance of the algorithms.