Computers in Industry 86 (2017) 1–14

Contents lists available at ScienceDirect

Computers in Industry

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

Comparative case study of life usage and data-driven prognostics techniques using aircraft fault messages

Marcia Baptista [a,*], Ivo P. de Medeiros [b], Joao P. Malere [b], Cairo Nascimento Jr. [c], Helmut Prendinger [d], Elsa M.P. Henriques [a,*]

[a] Instituto Superior Tecnico, Universidade de Lisboa, Lisbon 1049-001, Portugal
[b] Technol. Dev. Dept., Embraer SA, Sao Jose dos Campos, Brazil
[c] Instituto Tecnologico de Aeronautica (ITA), 12228-900 Sao Jose dos Campos, SP, Brazil
[d] National Institute of Informatics, 2-1-2 Hitotsubashi, Chiyoda-ku, Tokyo 101-8430, Japan

ARTICLE INFO

Article history:
Received 15 June 2016
Received in revised form 28 October 2016
Accepted 31 December 2016
Available online

MSC:
00-01
99-00

Keywords:
Case study
Aircraft prognostics
Data-driven techniques
Life usage modeling
Fault messages

ABSTRACT

Prognostics are a key activity in repair and maintenance operations. A recent approach to condition-based maintenance is the data-driven approach. This approach has been mostly based on past failure time measures, and sensed measurements of component degradation to derive estimates of the remaining useful life of equipment. An alternative source of data, rarely used in these models, is the stream of automatic messages derived from diagnostics systems, which consist of fault codes indicating abnormal events or deviations from optimal operation. Despite the richness and concise nature of these messages, their difficult interpretation poses significant challenges to its use in prognostics. This paper aims to show that data-driven prognostics based on this type of messages can be better suited to maintenance than time-based approaches. We illustrate this comparison with an industrial case study involving the removal times of a bleed valve from the aircraft air management system. Our experimental results reveal a significant accuracy improvement over the contrasting time-based models. We also establish the contribution to this improvement of the data-driven methods and message-related predictors.

1. Introduction

Engineering prognostics plays a central role in failure prevention and mitigation. The ability to forecast and estimate when a component or system can develop faults and irregularities based on the knowledge of its past usage, current state, and future conditions, is a competitive advantage in almost any industry. The discipline of prognostics is even more important in sectors where the criticality of the systems may compromise safety such as the aeronautics, automotive and nuclear industries. Despite significant research conducted over recent years, prognostics, in practice, continues to be a field predominantly intuitive that relies extensively on hard time preventive measures and human decision making. This is not surprising, as research has been mostly theoretical in nature, restricted to a small number of models and failure modes [1]. The number of complex prognostics models applied in real systems in the industrial field is also limited [1].

Originally from the fields of artificial intelligence and machine learning, the data-driven approach to prognostics consists in the extraction of statistical patterns from streams of data. As an empirical approach, it attempts to capture the connection between the system state variables – input, and output variables – using analytical methods such as linear models or neural nets, which rely on an implicit representation of the analyzed data instead of on the explicit knowledge of the physical behavior of the system.

Despite the efforts to disseminate the use of data-driven methods to predict potential failure and maintenance needs, many industries, namely the airline or the nuclear industries, are still reluctant to apply these kind of models to their most critical operational systems. Given the "black-box" nature of data-driven models, their interpretation is often difficult and their validity is dependent on the availability of field data.

In this paper we aim to explore how a range of sophisticated data-driven methods, combined with fault messages, can help improve prognostics and estimates of equipment lifetime. Fault messages are a kind of data which are seldom used in prognostics

* Corresponding authors.
E-mail addresses: marcia.baptista@ist.utl.pt (M. Baptista), elsa.h@tecnico.ulisboa.pt (Elsa M.P. Henriques).

http://dx.doi.org/10.1016/j.compind.2016.12.008
0166-3615/

[2]. They are generated by diagnostics systems when the value of an equipment sensory signal, such as vibration, or temperature, falls outside of its programmed parameters. They act as operational alerts, allowing maintenance teams to diagnose and predict future problems.

The proposed work involves an industrial case study, focusing on a two-valve subsystem of the aircraft air management system, which is subject to a specific range of operating conditions. In particular, we aim to show that experience-based maintenance (determined from Weibull analysis of past failure times) can be enhanced both by (1) artificial intelligence and machine learning techniques and (2) predictors related to fault messages. The novelty of our approach is twofold: a comprehensive analysis and comparison of a set of machine learning techniques and the investigation of fault messages as a new kind of health monitoring indicator for real industrial problems.

The remaining paper is organized as follows. Section 2 discusses related work in predictive maintenance and data-driven prognostics. Section 3 provides an overview of the case study. Section 4 describes the modeling approaches and the evaluation methods. Results are presented in Section 5. Finally, the paper is concluded with a summary of the results and an outline of future work in Section 6.

2. Theory and related work

2.1. Time-based maintenance and condition-based maintenance

The concept and practice of predictive or condition-based maintenance (CBM) originated back in the 1960s [3]. By the late 1950s the cost of maintenance activities was becoming unsustainable in many industries and pushed companies to rethink their preventive repair and maintenance operations. Back then, the concept of predictive maintenance based on performance and/or structural health monitoring, was novel and appealing but difficult to grasp and implement. Until then, companies had mostly relied on preventive or time-based maintenance (TBM), scheduling its operations according to fixed intervals defined by the manufacturer of the system/equipment with base on laboratory tests and past experience. For example, a common practice was to replace or renew bearings after a specified number of operating hours, assuming that the bearing failure rate increased with time in service.

The central generalized assumption of TBM was that the reliability of an equipment was directly related to its operating age [4]. Despite the apparent straightforwardness of this notion, in reality, many types of failures were difficult to prevent with base on statistical correlations of likelihood of failure and operational age alone [3]. Immediate solutions were the introduction of redundancy mechanisms together with strict TBM maintenance. Essential functions were supported by redundancy systems to ensure that, in the event of a failure, safety and standard operation were not compromised. It was also often preferred to exchange or repair a component, even without strong evidence of failure.

Condition-based maintenance took a different approach to the definition of a maintenance program by not focusing solely on the TBM strategy and on the notion of operational age. Usage concerns were integrated with the temporal evolution of the health condition of the equipment to develop enhanced prognostics. The goal here was for companies to be able to determine the most effective maintenance to each of its physical assets from a technical and economic standpoint [5].

2.2. Prognostics in time-based maintenance

Experience-based, failure rate or *life usage* models are considered the most widely used and accepted form of prognostics in aviation [6]. These models are used to drive preventive time based maintenance (TPM) plans that define the equipment 'hard time' intervals. They are often considered the most suitable alternative when the criticality or the failure risk of the equipment is low, when an efficient health monitoring sensor network to assess its condition is too costly or not possible to deploy, or when the equipment exhibits a linear or constant failure rate [7].

Life usage models rely exclusively on historical data of past failure times, either of the equipment or a sample of legacy systems. The underlying assumption is that the equipment failure times (or usage) follows a parametric distribution that is fitted from a sample of past data [8]. Here, a number of statistical distributions have been proposed [4], such as Exponential, Weibull, Log-normal, and Gamma distributions, with the Exponential and Weibull being the most common [9]. The former is simple and easy to apply while the later has the ability to adjust to different reliability stages namely, infant, mature and wear out phases [9]. A variety of parameter estimation methods, such as least squares, moments and maximum likelihood can be used to fit the time-to-failure data to these distributions.

Life usage modeling has various advantages – data analysis is relatively straightforward and can be performed by reliability personnel without specialized expertise. These models show some resistance to overfit and tend to have better performance with small training data sets. Life usage analysis is also not constrained to a single failure mode or a single-unit system, and can, in theory, be suited for reasonably complex systems with multiple failure mechanisms and/or several constituting components prone to failure [1].

The simplicity of usage-based modeling is an advantage but also a limitation. Due to its particular statistical treatment of the data, life usage models aim to capture average behavior. For this reason, these estimates are usually conservative, relying on cautious measures. For instance, in life usage models based on the Weibull distribution, failure is predicted in reference to the characteristic life. Given their focus on the capture of standard behavior these models have some difficulty to capture all the variability in the data thus losing predictive accuracy.

2.3. Model-based prognostics in condition-based maintenance

Prognostic methodologies for CBM have only recently been introduced [10]. According to the type of data inference, prognostics models for CBM can be classified as *model-based* or *data-driven* [11]. The model-based approach attempts to hand-code the real and specific interplay of factors determining the degradation and failure of a given component [1]. These models usually consist of a set of dynamic ordinary or partial differential equations [12] that can be solved with Lagrangian or Hamiltonian dynamics, approximation methods, and distributed models among other techniques [13]. They can also be described using state-space methods, that is, state variables describe the system by a set of first-order differential or difference equations, rather than by one or more nth-order differential or difference equations [14].

The most sophisticated model-based methods include rule-based expert systems such as SHINE [15] or Gensym G2 [16]. Other examples of model-based AI techniques are finite-state machines, as in [17,18] and qualitative reasoning, as in [19]. It should be noted that these model-based approaches, in contrast with life usage models, tend to derive indirect estimations of the equipment remaining useful life. They usually derive expected condition measurements that can be compared against the actual sensory signals of the equipment. A fault or failure is detected when a residual, that is, the difference between the observed and estimated (nominal) values of a measurement, is determined to be statistically significant [20]. A comprehensive, yet not extensive, review of physical models is provided in [21].