ARTICLE IN PRESS

Measurement xxx (2017) xxx–xxx

Contents lists available at ScienceDirect

Measurement

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

A data-driven prognostic approach based on statistical similarity: An application to industrial circuit breakers

Giacomo Leone [a,*], Loredana Cristaldi [a], Simone Turrin [b]

[a] Dipartimento di Elettronica, Informazione e Bioingegneria (DEIB), Politecnico di Milano, Piazza Leonardo da Vinci, 32, 20133 Milano, Italy
[b] ABB AG, Corporate Research Center, Wallstadter Str. 59, 68526 Ladenburg, Germany

ARTICLE INFO

Article history:
Received 25 October 2016
Received in revised form 18 January 2017
Accepted 9 February 2017
Available online xxxx

Keywords:
Data-driven
Industrial circuit breakers
Prognostics
Remaining useful life
Sub-fleet
Statistical test

ABSTRACT

In this paper, a data-driven prognostic algorithm for the estimation of the Remaining Useful Life (RUL) of a product is proposed. It is based on the acquisition and exploitation of run-to-failure data of homogeneous products, in the followings referred as fleet of products. The algorithm is able to detect the set of products (sub-fleet of products) showing highest degradation pattern similarity with the one under study and exploits the related monitoring data for a reliable prediction of the RUL. In particular, a novel methodology for the sub-fleet identification is presented and compared with other solution found in literature. The results obtained for a real application case as Medium and High Voltage Circuit Breaker, have shown a high prognostic power for the algorithm, which therefore represents a potential tool for an effective Predictive Maintenance (PdM) strategy.

1. Introduction

The Remaining Useful Life (RUL) of a system is defined as the useful life left at a particular time instant, that is the remaining time interval in which it will be able to meet its operating requirements. RUL estimation represents the core of the Prognostics and Health Management (PHM) programs which aim to a reduction of maintenance and life-cycle management costs, an increase of the systems availability and the adoption of Predictive Maintenance (PdM) strategies [1,2].

In the literature, the prognostic algorithms for the RUL estimation are usually classified in three different categories. The first class is related to the model-based approaches that refer to physical models describing the behavior of the systems under study. Such models can be very accurate but often require a strong and detailed knowledge of the inherent physics-of-failure. It follows that they are often very specific to the case study and their implementation is not always possible.

On the other hand, data-driven algorithms, the second main category found in the literature, are mainly based on the exploitation of the collected run-to-failure data and usually do not require particular knowledge about the inherent failure mechanisms. They provide a good trade-off between model complexity and results accuracy.

Finally, hybrid approaches attempt to leverage the advantages of combining the prognostics models in the aforementioned different categories for RUL prediction.

In the last years, data-driven approaches have experienced a wide diffusion. One reason is their suitability for applications related to complex engineered systems for which the definition of analytical models is a complex and resources demanding task. Another factor is the increasing availability of cheap monitoring systems that allow the collection of condition monitoring data in substantial quantities. A complete and detailed review about them is given in [3].

The same authors of this paper already presented two data-driven prognostic algorithms, one based on the statistical extraction and exploitation through Monte Carlo (MC) simulations of reliability and maintenance knowledge [4] and one based on a Machine Learning solution, in particular a Neural Network (NN) architecture [5]. A key and novel differentiator of the proposed approaches was the concept of fleet of products. A fleet of products is a set of homogenous products, with respect to the function for which they are intended, clustered together following different possible criteria such as belonging to the same customer, being installed in the same region or same industrial application and so on. The advantage of this practice is the possibility to extract fleet-specific usage and degradation profiles that can be exploited for the RUL prediction of a specific element (i.e. a specific product) of the selected fleet. In particular, the contribution of the acquisition of knowledge at fleet level on the improvement of the prognostic ability is quite

* Corresponding author.
E-mail addresses: giacomo.leone@polimi.it (G. Leone), loredana.cristaldi@polimi.it (L. Cristaldi), simone.turrin@de.abb.com (S. Turrin).

http://dx.doi.org/10.1016/j.measurement.2017.02.017
0263-2241/

relevant when the estimation of the product RUL at its early stage is of interest. Predicting the future behavior, in fact, is tied to the ability to learn from the past [1], and this is quite limited if the product is not in the mature stage of its life.

The application cases considered in [4,5], as well as in this paper, are Medium Voltage (MV) and High Voltage (HV) Circuit Breakers (CBs). MV and HV CBs are crucial elements in power transmission and distribution systems. They are usually characterized by a long useful life and high reliability, but at the same time, even a single failure of them may cause severe damages, from technical, economical and safety point of view. It follows that reliable prognostic models are necessary for such kind of engineered systems. Defining physical models for such devices, however, is a time consuming and laborious task, since several different failure mechanisms depending on many parameters (number of completed opening operations, contacts degradation, short circuit current) may occur and also interfere with each other. All these aspects motivate the choice of applying data-driven algorithms for the estimation of their RUL.

As already said, the approaches proposed in [4,5] are based on the concept of fleet of products. The application case proposed, however, represents one of the many classes of products for which collecting a relevant number of run-to-failure curves is difficult. A striking example for this are the Vacuum Circuit Breakers (VCBs). They are based on a relatively recent technology and producers claim for them a Mean Time To Failure (MTTF) of over 30 years. It follows that acquisition of Condition Monitoring (CM) data spanning all the lifetime for many VCBs that can be considered to work in similar conditions and industrial applications is often not possible.

These considerations have driven the authors to propose in [6] a new approach for the selection of a sub-fleet of products showing highest degradation pattern similarity with the product under study and for which the RUL estimation is required. Then, the related monitoring data are exploited for a reliable prediction of the RUL, offering a potential tool for an effective Predictive Maintenance strategy. The main resulting advantage is the lower constraint required for the definition of the product fleet (the reference library can now be constituted by run-to-failure data of products belonging to different customers, different geographical regions and industrial application), as the proposed methodology automatically selects the products with the most affine degradation profiles, discarding the ones that would weight in negative way in the RUL estimation.

In this paper, an alternative strategy for the selection of the products sub-fleet is proposed. In particular, it is based on the concept of degradation rate, which is explained further, and the application of a suitable statistical test for discarding products showing a degradation profile statistically not homogenous with the one of the product under analysis.

This article is structured in the following way: in Section 2 the proposed methodology for the selection of an appropriate sub-fleet is described and compared with the one presented by the authors in [6] and other works found in the literature. In Section 3, the prognostic stage exploiting the CM data of the sub-fleet products for the RUL prediction is illustrated. Finally, in Section 4, the results obtained for the application case are reported and compared with the ones obtained with the distance-based sub-fleet identification of [6]. The paper is ended with the conclusions.

2. Sub-fleet selection

The definition of a sub-fleet of products consists in the selection among a given set of products of a subset of them that show higher similarity, in terms of observed degradation (OD) in time, with respect to the item for which the estimation of the RUL is required. In the literature, some methods for a sub-fleet definition already exist. In particular, they are based on a similarity-based approach that consists in the evaluation of the similarity between the test trajectory pattern (monitored degradation pattern for the item for which the RUL has to be predicted) and the reference trajectory patterns stored in the database and use the RULs of these latter to estimate the RUL of the former, accounting for how similar they are [7]. In [8], the authors propose the definition of a similarity coefficient based on the sum of the squared errors between the monitored test pattern and the reference trajectories. In particular, given a test product x and a reference item j, the similarity coefficient sc_{xj} is calculated as:

$$sc_{xj} = \sum_{i=1}^{I} (OD_{ji} - OD_{xi})^2 \tag{1}$$

where OD_{xi} (OD_{ji}) is the observed degradation for the test product x (library specimen j) at cycle i and I is the number of observed cycles. In the estimation of the RUL, in order to give more weight to the library specimens with larger similarity coefficients, a weighting coefficient for a product j is defined as:

$$w_{xj} = \exp\left(-\frac{sc_{xj}}{\beta}\right) \tag{2}$$

where β is selected according to the desired selectivity (i.e. if β is small, few specimens are influential). Finally, the *RUL* for the product x is computed according to (3):

$$RUL_x = \frac{\sum_{j=1}^{J} w_{xj} RUL_j}{\sum_{j=1}^{J} w_{xj}} \tag{3}$$

This approach has been used also in [9], whereas in [10] a slight modification is proposed. The modifications are made in the RUL estimation (i.e. Eq. (3)), in which the most similar P percent number of the library samples are utilized rather than using whole dataset.

A different approach is suggested in [11]. The authors, in fact, propose a definition of a deterministic model M_i for each i-th training unit of the library, so that, for a given time t, an estimated value y of the Health Indicator (HI) variable that describes the degradation pattern of the item is provided. At this point, if a sequence $Y = y_1, y_2, \ldots, y_r$ of values of the HI for a test unit is available, a distance metric between the model M_i and Y is defined as the sum of the squared errors between the monitored test pattern and estimations provided by the model, divided by the prediction variance of the model itself. Then the RUL estimation for the test product is equal to a weighted sum of the RUL of the reference products. The weights can be assigned according to different principles. One of them is to apply the k-nearest neighbor method that is to select the products with the k smallest distance values and apply a weight $1/k$ to their RULs.

In this paper, an alternative methodology for the sub-fleet identification is proposed. In particular, it is based on the application of a statistical test for the identification of those products presenting a statistical distribution of the degradation rate similar to the target product. This approach deals with the identification of homogenous products through a different point of view, with respect to the methodology presented in [6]. The rest of the Section is structured as follows: first, a short overview (Section 2.1) on the condition monitoring data considered for the study is provided. Then, in order to highlight the novelty of the proposal, in Section 2.2 the sub-fleet identification presented in [6] is briefly recalled and discussed. Finally, last Section 2.3 presents the new identification approach.