Fleet Prognostics and Health Management – A General Process Model for Data-Driven Fleet Analytics

Carolin Wagner, MScIS

Westfälische Wilhelms-Universität Münster, 48149 Münster, Germany
wagner@ercis.uni-muenster.de

ABSTRACT

Current prognostics and health management approaches heavily rely on the availability of data, which in most cases is only accessible to a limited extend. To improve existing techniques, knowledge that is captured in the fleet of machines could be included. Even though machines are usually not identical, algorithms can learn from similar situations. In many cases, fleet knowledge is exclusively used to build a general prediction model disregarding the different characteristics and operating conditions. This is mainly attributed to the limited knowledge about the reasonable handling of fleet knowledge, especially considering the unique characteristics of the fleet. As a result, the introduction of an appropriate prognostics and health management approach is often impeded or even hindered. The thesis therefore strives to develop guiding principles as part of a general process model in order to support practitioners during the implementation of a prognostics and health management system using fleet knowledge.

1. PROBLEM STATEMENT

The reliability of machines is a critical factor in today's highly optimized business environment. Unplanned machine breakdown leads to high maintenance and downtime costs. Traditional maintenance strategies, including corrective and time-based preventive maintenance, however are no longer able to address this requirement in a satisfactory way. For this purpose, prognostics and health management (PHM) gains in importance both for practice and research. Through the increased availability of sensors and machine data, the condition of the machine can be monitored and analyzed continuously. This data provides information to identify and forecast device failures, resulting in more precise maintenance and spare parts requirements. The success of a PHM strategy highly depends on the implementation of the two main tasks, namely fault diagnostics and prognostics. Fault diagnostics focuses on the detection, isolation and identification of the failure type. On the other hand, prognostics attempt to estimate the remaining useful life (time to failure) of a machine. Approaches for both tasks can generally be categorized into model-based and data-driven. Model-based approaches use physics models to describe the machine behavior. Contrary, data-driven approaches aim to extract relevant information in order to build a model that represents best the available historical data. While the former requires substantial knowledge about the system and its behavior, in the latter case information are exclusively obtained through data analysis.

Current data-driven PHM approaches are often not able to meet expectations due to their limited ability to accurately detect abnormal machine conditions, identify failures and estimate the remaining useful life. This is in many cases attributed to the lack of real data and insufficient knowledge about the component or machine under consideration. Instead, experimental data is often used for algorithm training, which is not able to reflect the complexity of real-world systems. To improve PHM approaches condition data from fleets of machines rather than single units can be taken into consideration. Even though machines are usually not identical with regard to their characteristics and operating conditions, algorithms can learn from similar situations. Therefor machine conditions are assessed against situations encountered by machines in the same fleet and knowledge is transferred to allow algorithms to intelligently learn and improve their capabilities. This is referred to in the following as fleet prognostics and health management.

Several approaches for fault diagnostics and prognostics have recently been presented in the literature which make use of fleet knowledge. These approaches are usually designed for specific fleet compositions and characteristics. The potential for fleet-based analytics has not been exhausted to its full extend. In most cases, approaches using a set of similar machines either apply a general prediction model disregarding the different working conditions, environmental influences and machine specific characteristics or built individual models for each machine overlooking the great potential of learning from the fleet. This can, on the one side, be attributed to missing knowledge about the reasonable handling of fleet knowledge, especially considering the unique characteristics of the fleet under consideration. On the other side, fleet-based analytics lacks a general process model to support practitioners to select a suitable approach including required preparation tasks and evaluation steps to implement and/ or improve their data-driven prognostics and health management strategy relying on the available knowledge of their fleets. In order to incorporate the usage of fleet knowledge into existing PHM strategies, promising techniques and guiding principles have to be identified, which support practitioners applying fleet-based analytics dealing with the specific characteristics.

2. EXPECTED CONTRIBUTIONS

This doctoral thesis aims to develop a general process model for data-driven fleet fault diagnostics and prognostics. It will present guiding principles for the implementation of fleet-based analytics in the context of PHM with particular regard to different fleet characteristics and requirements. Given the existing fleet-based approaches, a process model will be designed to capture the required characteristics of the fleet and to identify the most promising approaches for fleet knowledge integration. It will include several general steps, which need to be performed independently of the specific characteristics as well as steps only necessary for subsets of fleets. Figure 1 provides a preliminary outline of a procedural design including the main phases to be investigated during this thesis. This steps are identified based on existing approaches for fleet-based analytics and generalized for the proposed process model.

Figure 1. Course structure of main tasks

The first phase *Fleet Scenario* focuses on the identification of different fleet characteristics and resulting requirements for fleet-based approaches. Depending on these characteristics it is possible to determine whether fleet-based approaches are applicable in general to the specific case. Based on these results, crucial *data processing* tasks are presented and evaluated. Among general data processing tasks, such as feature selection and handling imbalanced data sets, additional tasks specifically addressing fleet data are outlined and depicted in a summarized form. This is followed by the optional phase *sub-fleet definition.* For some specific fleet characteristics (e.g. highly heterogeneous fleets) or techniques (e.g. case-based reasoning), it is necessary to group machines which exhibit similar technical features or behavior in order to focus the analysis on the data providing relevant information. Relevant sub-fleets are defined to best fit the specific fleet characteristics as well as the analysis target. The fourth phase *Methodical Approach* depicts the central element of the process model. The requirements of the specific fleet are matched against the properties of existing fleet-based approaches both for fault diagnostics and prognostics. A set of possible techniques are presented to the practitioner showing corresponding benefits and drawbacks. *Performance evaluation* of the selected approach is subsequently performed with regard to several criteria. Existing performance measurements are taken and extended for fleet-based analytics.

Summarizing, the main objective is the development of guiding principles for the implementation of fleet-based analytics. In order to solve this objective, the following research questions will be answered in the thesis:

(RQ1) What are relevant dimensions for fleet classification in the context of groups of machines and what are the possible characteristics?

(RQ2) Which fleet-based approaches exist for prognostics and health management and what are their distinguishing properties?

(RQ3) How can these results be integrated in a general process model to facilitate the implementation of fleet prognostics and health management?

3. RESEARCH PLAN

The thesis will be conducted using a step by step approach of the main tasks and objectives, considering especially the identified research questions. Each of the research question is answered using structured literature research and complemented by expert interviews to additionally include the practitioner's perspective and problems. Additional findings are continuously integrated and corresponding adaptations are performed. The structured model is validated using available data sets. Table 1 depicts the current research plan in respect of the main thesis objectives and the planned time horizon (periods of six months).

Table 1: Research Plan

Objectives /Term	2017 T1	2017 T2	2018 T1	2018 T2	2019 T1	2019 T2
Fleet Scenario	■	■				
Fleet-based Approaches		■	■			
Process Model		■	■	■		
Evaluation			■	■	■	
Writing/ Adjustment				■	■	■

3.1 Work Performed

Work performed has mainly targeted the first two objectives of the proposed research plan. With regard to the first objective, the thesis extends the existing fleet classification by Al-Dahidi, Di Maio, Baraldi, & Zio, 2016. In their work, they present three groups of fleets, namely identical, homogeneous and heterogeneous fleets. This frequently used distinction makes reference to the structural dimension of fleets. For fleet-based approaches, however additional dimensions should be taken into account. These include among others the operating condition in the fleet (e.g. identical, different, or dynamically changing) and the type of available data (e.g. sensor reading, context data, textual description). The previously developed data classification is taken as foundation for the fleet data dimension (Wagner, Saalmann, & Hellingrath, 2016a). Interim results are shown in the publication which will be submitted to the PHM Annual Conference 2017. This is followed by a deeper analysis of fleet characteristics as well as clustering them to form recurrent fleet scenarios to be used afterwards to fleet-based technique selection.

Besides the definition of fleet characteristics, a deeper analysis of benefits and drawbacks for the integration of fleet knowledge has already been conducted for the second objective also presenting techniques for sub-fleet evaluation and fleet-based fault diagnostics and prognostics (Wagner & Hellingrath, 2017). These interim results have recently been extended using a structured literature research approach based on different notions of the term fleet to yield an extensive overview of existing fleet-based approaches as well as their unique properties.

3.2 Remaining Work

Existing results are to be continuously improved incorporating new findings as well as using expert interviews to gain additional insights from practitioners in order to complete the first two objectives. Besides existing techniques, a general neural-network based fleet prognostics approach is currently developed, which is built on existing neural network approaches and designed for homogeneous as well as heterogeneous fleets. This approach will also be taken into consideration for the process model.

With regard to the last phase, both remaining steps, data processing and performance evaluation, will be primary based on existing knowledge. For that, a structured literature review is planned to extract data processing techniques as well as performance metrics from prognostics and health management in general. Additionally, the work previously performed on imbalanced data sets will be integrated (Wagner, Saalmann, & Hellingrath, 2016b). The results are subsequently to be extended and adapted to fit the requirements of fleet-based analytics.

In the last part of the thesis, the results are evaluated using different available data sets. To present, the usage of the PHM data set 2008 as well as an online available hard drive data set ("Hard Drive Test Data - Determining Failure Rates and More") is envisioned. While former presents different benchmark options, the latter enables testing the process model on a complete new data set exhibiting its own properties and difficulties.

4. CONCLUSION

The thesis aims to develop a process model for fleet PHM. It will support practitioners to develop or improve their PHM strategy provided that data from groups of machines is available. Based on different fleet characteristics, relevant techniques are selected which fit best the specific fleet requirements. Results are subsequently embedded into a process model to assist the implementation of fleet prognostics and health management using guiding principles.

The final results of the thesis are based on existing works on fleet characteristics and fleet-based approaches. These interim results are to be extended following the presented procedural design and subsequently validated on different data sets.

REFERENCES

Al-Dahidi, S., Di Maio, F., Baraldi, P., & Zio, E. (2016). Remaining useful life estimation in heterogeneous fleets working under variable operating conditions. *Reliability Engineering & System Safety*, *156*, 109–124. https://doi.org/10.1016/j.ress.2016.07.019

Hard Drive Test Data - Determining Failure Rates and More. Retrieved from https://www.backblaze.com/b2/hard-drive-test-data.html

Wagner, C., & Hellingrath, B. (2017). Fleet-based Diagnostics and Prognostics – Investigating the Usage of Fleet Knowledge for Condition-Based Maintenance. *Logistikmanagement 2017 Stuttgart*.

Wagner, C., Saalmann, P., & Hellingrath, B. (2016a). An Overview of Useful Data and Analyzing Techniques for Improved Multivariate Diagnostics and Prognostics in Condition-Based Maintenance. In *Annual Conference of the Prognostics and Health Management Society 2016, PHM Conference 2016.*

Wagner, C., Saalmann, P., & Hellingrath, B. (2016b). Machine Condition Monitoring and Fault Diagnostics with Imbalanced Data Sets based on the KDD Process. *IFAC-PapersOnLine*, *49*(30), 296–301. https://doi.org/10.1016/j.ifacol.2016.11.151